Intelligent
Embedded Systems

6

Sebastian Breker

Klassifikation von Niederspannungsnetzen mit Support Vector Machines

Bewertung des Aufnahmevermögens für Dezentrale Erzeugungsanlagen

kassel
university
press

Intelligent
Embedded Systems Band 6

Herausgegeben von
Prof. Dr. Bernhard Sick, Universität Kassel

Sebastian Breker

Klassifikation von Niederspannungsnetzen mit Support Vector Machines

Bewertung des Aufnahmevermögens
für Dezentrale Erzeugungsanlagen

kassel
university
press

Die vorliegende Arbeit wurde vom Fachbereich Elektrotechnik / Informatik als Dissertation zur Erlangung des akademischen Grades eines Doktors der Ingenieurwissenschaften (Dr.-Ing.) angenommen.

Erster Gutachter: Prof. Dr. rer. nat. Bernhard Sick
Zweiter Gutachter: Prof. Dr.-Ing. Albert Claudi

Tag der mündlichen Prüfung: 24. Juni 2015

Bibliografische Information der Deutschen Nationalbibliothek
Die Deutsche Nationalbibliothek verzeichnet diese Publikation in der Deutschen Nationalbibliografie; detaillierte bibliografische Daten sind im Internet über http://dnb.d-nb.de abrufbar

Zugl.: Kassel, Univ., Diss. 2015
ISBN 978-3-7376-0014-9 (print)
ISBN 978-3-7376-0015-06(e-book)
URN: http://nbn-resolving.de/urn:nbn:de:0002-400152

Printed in Germany

ABSTRACT

The thesis adresses various issues of low-voltage grid classification regarding their hosting capacity for distributed generators. Against this backdrop, some efficient approaches to assess low-voltage grids are presented. The relation to the application domain and the capability of the elaborated methods is demonstrated by a plenty of practically relevant experiments with data from real low-voltage grids. Hence, the application of the proposed methods is enabled and will not only help Distribution System Operators to face the challenges in future grid planning as well as to focus their work on further enhancement of weak grid structures, but it will also be valuable in choosing relevant grids for detailed surveys, e.g., by choosing various grids from each ordinal class for the survey. Hereby, the amount of „weak“ and „strong“ grids which will be investigated can be controlled to get representative results with regard to all classes.

In a first step, discriminating features which characterize low-voltage grids are identified and packed in two different data sets to face the problem of gathering appropriate base data. To assess the hosting capacity of low-voltage grids for distributed generators with ordinal classes (or labels), two new grid specific approaches are proposed. The first approach is based on knowledge elicitation from human experts while the second one uses stochastic simulations to classify a grid. Although both approaches have a different focus, obviously, the empirical investigations show that the results of both approaches do not contradict each other and, on top of this, are statistically concordant.

Because of the different focus of the grid specific approaches and the fact that the statement of a single expert is for many reasons more or less uncertain, partially different classification results are expected for a single low-voltage grid. Thus, this thesis also considers the problem of combining several, potentially wrong labels to get a combined label which is afflicted with lower uncertainty than the single labels. Furthermore, the combination of human expert statements with automatically gathered simulation results leads to more comprehensive labels with regard to the assessment of low-voltage grids. A novel combination rule, the Extended Imprecise Dirichlet Model Rule which is based on a k-nearest-neighbor approach and Dirichlet distributions, i.e., second-order distributions for multinomial distributions, is proposed. It can be shown that, in comparison to some known combination rules, the use of our new combination rule leads to better results where ordinal class labels are used and only a few, uncertain expert statements are available.

To assess additional low-voltage grids more efficiently, without using the grid specific classification approaches, different classification concepts are empirically evaluated using Support Vector Machines (SVM). Using SVM-classifiers, various kinds of uncertainty can occur, e.g., regarding the allocation of samples to classes, the exact feature values, or the parameterization with a limited amount of data. In this thesis, samples are not distinctly allocated to one class (e.g., various perhaps different expert statements for one low-voltage grid). The uncertain information is processed by using a combination rule prior to the classification. The subsequent modeling of a SVM is reduced to the prediction of the combined label. In addition to the investigation of different classification concepts, the influence of the feature set, with which the classifier is trained, and the influence of the applied combination rule on the classification accuracy is analyzed. It is demonstrated that the Extended Imprecise Dirichlet Model Rule outperforms the other combination rules concerning the prediction of the combined labels considerably. Additionally, it is shown that the SVM classifier can be further improved considering ordinal information from the labels to solve the multiclass classification problem.

Altogether, the classification results can be rated as very satisfying. The amount of missclassifications which is reached with the best classification concept is 15%. Particularly, this missclassifications mainly differ one class from the desired output of the classifier.

ZUSAMMENFASSUNG

In dieser Arbeit werden verschiedene Fragestellungen zur Klassifikation von Niederspannungs-(NS)-Netzen hinsichtlich ihrer Aufnahmekapazität für dezentrale Erzeugungsanlagen (DEA) adressiert. Vor diesem Hintergrund werden effiziente Ansätze zur Bewertung von NS-Netzen vorgestellt. Die Anwendungsnähe und Einsatztauglichkeit der erarbeiteten Methoden wird konsequent durch praxisnahe Experimente an Daten einer Vielzahl realer NS-Netze unterstrichen. Ein Einsatz in der Praxis ist daher direkt möglich und kann z. B. zur Erhöhung der Planungssicherheit bei der Steuerung von Investitionsmitteln auf der NS-Ebene dienen. Weiterhin wird durch die Methoden eine strukturierte Möglichkeit zur Auswahl von relevanten NS-Netzstrukturen für detaillierte Untersuchungen geschaffen, indem z. B. von jeder Klasse eine bestimmte Anzahl an NS-Netzen für die Untersuchungen gewählt wird. Auf diese Weise kann der Anteil an „schwachen“ und „starken“ Netzen, die untersucht werden sollen, gesteuert werden, um repräsentative Ergebnisse über alle Klassen zu erhalten.

Um das Grundproblem der Bildung einer geeigneten Datengrundlage zur Charakterisierung von NS-Netzen anzugehen, werden zunächst diskriminierende Merkmale identifiziert und in zwei unterschiedlichen Datensätzen gebündelt. Zur Bewertung der NS-Netze hinsichtlich ihres Aufnahmevermögens für DEA anhand von ordinalen Klassen (bzw. Labeln) werden zwei neue netzspezifische Ansätze erarbeitet. Neben einem expertenbasierten Konzept zur Erhebung von ordinalen Klassen mit Hilfe menschlicher Experten wird ein zweiter Ansatz vorgestellt, der auf stochastischen Simulationen basiert. Obwohl beide Ansätze offensichtlich unterschiedliche Schwerpunkte in der Bewertung setzen, zeigt sich in durchgeführten empirischen Analysen, dass sich die Klassifikationsergebnisse nicht widersprechen und sogar mit statistischer Signifikanz konkordant sind.

Auf Grund der unterschiedlichen Schwerpunkte der erarbeiteten netzspezifischen Klassifikationsansätze und der Tatsache, dass die Aussage eines einzelnen Experten aus vielen Gründen mehr oder weniger unsicherheitsbehaftet ist, sind dennoch zum Teil unterschiedliche Ergebnisse zu erwarten. Diese Arbeit widmet sich daher auch dem Problem der Kombination von mehreren, potenziell falschen Klassifikationseinschätzungen, um eine im Vergleich zu den einzelnen Klassifikationseinschätzungen weniger unsicherheitsbehaftete Gesamtaussage zu erhalten. Weiterhin führt die Kombination von Einschätzungen menschlicher Experten mit automatisch erhobenen Simulationsergebnissen zu einer umfassenderen Bewertung von NS-Netzen. In diesem Zuge wird ei-

ne neue Kombinationsregel, die Erweiterte Imprecise-Dirichlet-Model-Regel (EIDMR), vorgestellt, die auf einem k-nächste-Nachbarn-($k$nn)-Ansatz und Dirichlet-Verteilungen (z. B. konjugierte Verteilungen für Multinomialverteilungen) basiert. Die Ergebnisse der durchgeführten empirischen Experimente zeigen, dass die neue Kombinationsregel den bekannten überlegen ist, wenn zum einen ordinale Klassen verarbeitet werden müssen, und zum anderen, wenn die vorliegende Anzahl an Klassifikationseinschätzungen für ein einzelnes Objekt gering ist.

Um für die Bewertung weiterer NS-Netze auf die Anwendung der netzspezifischen Klassifikationsansätze verzichten zu können und damit eine Effizienzsteigerung bei der Klassifikation zu generieren, werden in einem weiteren Schritt verschiedene Support-Vector-Machine-(SVM)-basierte Klassifikationskonzepte empirisch untersucht. Bei der Anwendung von SVM kann Unsicherheit prinzipiell an vielen Stellen auftreten, z. B. hinsichtlich der Zuordnung von Datenpunkten zu einer Klasse, der exakten Featurewerte oder der Parametrisierung der SVM anhand einer begrenzten Datenmenge. In dieser Arbeit liegt der Fall vor, in dem Datenpunkte nicht eindeutig einer Klasse bzw. gleichzeitig mehreren Klassen zugeordnet sind (z. B. mehrere evtl. verschiedene Expertenaussagen für ein einzelnes NS-Netz). Der unsicheren Information wird durch vorherige Anwendung einer Kombinationsregel begegnet. Die nachgelagerte Anwendung der SVM reduziert sich damit auf die Prognose der kombinierten Label. In diesem Zuge wird neben unterschiedlichen Klassifikationskonzepten auch der Einfluss der verwendeten Features und der angewendeten Kombinationsregel untersucht. Dabei zeigt sich hinsichtlich des Klassifikationsverhaltens der SVM ein deutlicher Vorteil bei Verwendung der neuen Kombinationsregel EIDMR. Weiterhin wird dargelegt, dass der SVM-Klassifikator durch Berücksichtigung der in den Labeln beinhalteten ordinalen Information bei der Lösung des Mehrklassenklassifikationsproblems weiter verbessert werden kann.

Die erreichten Klassifikationsgüten sind aus Anwendungssicht als sehr zufriedenstellend zu bewerten. Die Anzahl der Fehlklassifikationen mit dem besten Klassifikationskonzept liegen bei 15%, wobei diese Fehlklassifikationen hauptsächlich nur um ein Klasse von der gewünschten Ausgabe des Klassifikators abweichen.

DANKSAGUNG

Die vorliegende Arbeit wurde neben meiner beruflichen Tätigkeit im Asset Management der EnergieNetz Mitte GmbH an der Universität Kassel angefertigt. Für die Unterstützung der EnergieNetz Mitte und die Möglichkeit zur Verwendung realer Netzdaten in dieser Arbeit möchte ich mich an dieser Stelle bedanken.

Die Begutachtung an der Universität Kassel wurde durch Herrn Prof. Dr. rer. nat. Bernhard Sick und Herrn Prof. Dr.-Ing. Albert Claudi übernommen. Für Ihre kritischen und hilfreichen Ratschläge bin ich Ihnen sehr dankbar. Bernhard Sick sei an dieser Stelle für die zahlreichen fachlichen Diskussionen ein besonders herzlicher Dank ausgesprochen. Mein Dank gilt auch den fünf Experten der EnergieNetz Mitte GmbH Herrn Dipl.-Ing. Kai Boldt, Herrn Dipl.-Ing. M.Sc. Marvin Reiting, Herrn Dipl.-Ing. M.Sc. Timm Eberwein, Herrn M.Eng. Johannes Rudolph und Herrn Wolfgang Adams für ihre Teilnahme an der Expertenbefragung sowie den wissenschaftlichen Mitarbeitern des Fachgebietes „Intelligent Embedded Systems“ des Fachbereichs 16 der Universität Kassel für die geleistete Unterstützung.

Vor allen anderen gilt der Dank meiner Familie, die diesem Vorhaben stets Interesse geschenkt, mir viele Freiräume geschaffen und mich jederzeit persönlich unterstützt haben.

Kassel, im März 2015 Sebastian Breker

INHALTSVERZEICHNIS

AKRONYME

ANN	Artificial Neural Network
CART	Classification and Regression Trees
CI	Computational Intelligence
DSR	Dempster-Shafer-Regel
DST	Dempster-Shafer-Theorie
DEA	Dezentrale Erzeugungsanlage
EE	Erneuerbaren Energien
EIDMR	Erweiterte Imprecise-Dirichlet-Model-Regel
GMM	Gauß'sches Mischmodell
GA	Genetische Algorithmen
HS	Hochspannung
HöS	Höchstspannung
KVS	Kabelverteilerschrank
KKT	Karush-Kuhn-Tucker
IDM	Imprecise Dirichlet Model
IDMR	Imprecise-Dirichlet-Model-Regel
$k$nn	k-nächste-Nachbarn
MS	Mittelspannung
MLP	Multilayer Perceptron
MR	Murphy's Regel
NS	Niederspannung
PV	Photovoltaik
SMO	Sequential Minimal Optimization
SVM	Support Vector Machine
VDE	Verband der Elektrotechnik, Elektronik, Informationstechnik e.V.
VNB	Verteilnetzbetreiber

1

EINLEITUNG

1.1 HERAUSFORDERUNGEN IN DER NIEDERSPANNUNGSNETZPLANUNG

1.1.1 *Wandel im elektrischen Verteilnetz*

Die Anforderungen an die Planung und den Betrieb eines elektrischen Verteilnetzes haben sich innerhalb der letzten Jahre stark gewandelt. Durch viele äußere Einflüsse und Treiber wie der Liberalisierung der Energiemärkte, der Regulierung des Netzbetriebes, dem Trend zur Rekommunalisierung und dem Beschluss zur Energiewende findet sich ein regionaler Verteilnetzbetreiber (VNB) heute in einem Spannungsfeld aus Renditeerwartungen, kommunalpolitischen Ansprüchen, regulatorischen Auflagen, gesellschaftlicher Verantwortung sowie der Aufrechterhaltung einer effizienten und zuverlässigen Energieversorgung wieder. Insbesondere durch die angestrebten energiepolitischen Ziele werden in Deutschland massive Investitionen notwendig [1], um eine gleichbleibend hohe Versorgungsqualität und -zuverlässigkeit zu gewährleisten, denn die massive Steigerung der installierten Leistung zur Stromerzeugung aus Erneuerbaren Energien (EE) hat vielfältige Auswirkungen auf die verschiedenen Spannungsebenen.

Vor diesem Hintergrund werden zukünftig optimierte und vor allem starke Infrastrukturen zur Stromversorgung erforderlich, deren Aufbau von einem hohen Bedarf an finanziellen Mitteln begleitet werden wird. Weiterhin kommt dem Zustand des Anlagevermögens eines Netzbetreibers neben der regulatorischen und betriebswirtschaftlichen eine zusätzliche Bedeutung für die Versorgungssicherheit und -qualität zu. Um unter diesen Voraussetzungen zukünftig auch weiterhin wirtschaftlich effizient handeln zu können, ist für Netzbetreiber eine sorgfältige und vorausschauende Investitionsplanung unabdingbar. In der Vergangenheit wurden zur strategischen Netzplanung in der Mittelspannung (MS) und Niederspannung (NS) häufig Zielnetzplanungen eingesetzt. Die Entwicklung von Zielnetzstrukturen bedingt allerdings ein gewisses Maß an Sicherheit bei der Antizipation zukünftiger Umweltbedingungen und Investitionsrückflüsse. Die

heutigen Anforderungen an die Planung von Verteilnetzen unterscheiden sich jedoch signifikant von denen der Vergangenheit, in der Netzplanungen nahezu ausschließlich anhand von Lastprognosen und innerhalb eines vergleichsweise statischen Unternehmensumfeldes durchgeführt wurden. Im Vergleich zur Lastzunahme der Vergangenheit ist z. B. der heute angestrebte lokale Ausbau der EE – insbesondere auf den unteren Verteilnetzebenen – nicht nur von örtlichen Gegebenheiten und der regionalen Kaufkraft, sondern auch von der Akzeptanz innerhalb der Bevölkerung und von Subventionsleistungen durch den Staat abhängig. Es lässt sich folglich festhalten, dass die Anschlussmöglichkeit einer dezentralen Erzeugungsanlage (DEA) an einem Knoten im Verteilnetz auch von vielen nichttechnischen Faktoren abhängig ist.

1.1.2 *Auswirkungen auf die Niederspannungsnetzplanung*

Die NS-Ebene ist die Netzebene, auf der die meisten Endkunden – z. B. Haushaltskunden – an das elektrische Energieversorgungssystem angeschlossen sind. Zu Beginn der Elektrizitätsversorgung ist dieses System sukzessive hierarchisch geplant, erweitert und optimiert worden. Dies bedeutet, dass der Transport der elektrischen Energie hauptsächlich von zentral gelegenen Großkraftwerken aus zu den Endkunden vorgenommen worden ist. Mit den wechselnden politischen Rahmenbedingungen der letzten Jahre, aufstrebenden neuen Markttrends, der wachsenden Belastung der Umwelt und der sinkenden Verfügbarkeit von fossilen Primärenergieträgern, findet derzeit ein Paradigmenwechsel im elektrischen Versorgungssystem statt [2, 3, 4, 5]. Insbesondere in ländlichen und vorstädtischen Bereichen wurde, hervorgerufen durch die dort auftretenden hohen Potenziale für eine Energieumwandlung unter der Nutzung von EE, der Ausbau und die Installation von DEA – insbesondere der Photovoltaik (PV) – innerhalb des letzten Jahrzehnts forciert.

Treten in den oberen Netzebenen hauptsächlich Stabilitätsprobleme auf (hauptsächlich in der Höchstspannung (HöS), selten in der Hochspannung (HS)), so spielen diese Erscheinungen in den MS- und NS-Netzen auf Grund der vergleichsweise geringen Entfernungen im Leitungsnetz so gut wie keine Rolle. In den unteren Verteilnetzebenen ergibt sich die Problematik daher hauptsächlich aus der Stromrichtungsumkehr zu Zeitpunkten mit hohen dezentralen Erzeugungsleistungen. Durch wachsende Einheitsgrößen und der zum Teil ungünstigen Verteilung von DEA im Netz (z. B. bei kleinen Onshore-Windparks), wird die Einhaltung des in den Normen und Richtlinien (vgl. Abschnitt 1.2) festgelegten Spannungsbandes auf Basis der heutigen Netzstrukturen zukünftig nicht mehr jederzeit gewährleistet werden können. Ohne geeignete Gegenmaßnahmen wird eine stark steigende installierte Leistung auf der NS-Ebene folglich Betriebsmittelüberlastungen und Spannungsbandverletzungen zur Folge haben. Das Auftreten von technischen Engpässen im NS-Netz ist dabei sehr stark von der Struktur des Netzes und der darin befindlichen Anordnung der DEA abhängig. Eine Überlas-

tung von Betriebsmitteln tritt in der Vielzahl der Fälle erst nach Verletzung des Spannungsbandes bei sehr hohen Einspeiseleistungen auf. Häufig ist der Netztransformator thermisch das begrenzende Element [6]. Gelegentlich können an Netzknoten mit sehr geringen Kurzschlussleistungen weitere Beeinträchtigungen der Versorgungsqualität z. B. in Form von Flickererscheinungen, starken Unsymmetrien oder einer unzulässig hohen Oberschwingungsbelastung auftreten [7].

Bezüglich der genannten technischen Probleme ist es Aufgabe des verantwortlichen VNB, netzspezifische Verstärkungsmaßnahmen und langfristige Entwicklungen in seinen NS-Netzen durchzuführen. Aus dem in Abschnitt 1.1.1 beschriebenen Spannungsfeld resultieren neben technischen Problemen durch den regulierten Markt ebenfalls begrenzte finanzielle Spielräume (z. B. Anreizregulierung), wodurch die Entscheidung immer mehr an Bedeutung gewinnt, in welches Netz aus einer Vielzahl von NS-Netzen Investitionsmittel platziert werden. Die Entscheidung wird dabei durch die Tatsache erschwert, das vorhandene Netzstrukturen nicht etwa in ihrer Gesamtheit geplant wurden, sondern historisch gewachsene Strukturen aufweisen, deren Entwicklung stark durch lokale und geographische Abhängigkeiten (z. B. Flüsse, Bodenbeschaffenheit, Bebauungsplanung) beeinflusst wurde. Zusammenfassend lassen sich die Arbeitsschritte bis zu einer Verstärkung eines NS-Netzes in folgende Schritte unterteilen:

1. Auswahl einer Menge von in der Zukunft potenziell problematischen, stark durch DEA penetrierten NS-Netzen,

2. Bewertung/Priorisierung/Klassifikation der ausgewählten NS-Netze hinsichtlich ihres Aufnahmevermögens für DEA,

3. Planung von adäquaten Netzausbaumaßnahmen in den identifizierten schwachen NS-Netzen und

4. Projektierung und operative Umsetzung der Maßnahmen.

Insbesondere dem zweiten Arbeitsschritt kommt eine hohe Bedeutung zu, um die Entwicklung der Betriebsmittelbelastungen und Spannungsniveaus bei fortschreitendem DEA-Zubau nur für eine ausgewählte Menge an NS-Netzen beobachten zu müssen. Auf dieser Basis können anschließend gezielt Netzverstärkungsmaßnahmen geplant werden. Da bei der Planung von Netzverstärkungmaßnahmen immer auch die zukünftige Entwicklung des DEA-Zubaus antizipiert wird, kann die Projektierung und operative Umsetzung dieser Maßnahmen entsprechend der Geschwindigkeit und der Eintrittswahrscheinlichkeit der in der Planung unterstellten Rahmenbedingungen (z. B. Anschlusspunkte und Leistungen zukünftiger DEA) auch zu einem späteren Zeitpunkt veranlasst werden. Vor diesem Hintergrund ist die Planung dann gegebenenfalls nochmal an zwischenzeitlich geänderte Rahmenbedingungen anzupassen.

1.2 RELEVANTE NORMEN UND RICHTLINIEN

Beim Betrieb und der Planung seiner Verteilnetze ist ein VNB an eine Vielzahl technischer Normen und Richtlinien gebunden. Daher wird im Folgenden eine Auswahl der für die NS-Ebene relevanten Normen und Richtlinien zusammengestellt. Im Abschnitt 1.1.2 sind bereits Auswirkungen durch den zunehmenden Anschluss von DEA in den Verteilnetzen auf den Netzbetrieb vorgestellt und insbesondere auf die Betriebsmittelauslastungs- und Spannungsproblematik im NS-Netz hingewiesen worden. Die Definition und Beschreibung von Merkmalen, die eine Versorgungsspannung in MS- und NS-Netzen an der Übergabestelle zum Netznutzer aufzuweisen hat, erfolgt in der Norm DIN EN 50160 [8]. Festgelegt werden dort im Wesentlichen Merkmale bezüglich der Frequenz, Höhe, Kurvenform und Symmetrie der Leiterspannungen. Zur Einhaltung der Versorgungsqualität in NS-Netzen bildet diese Norm daher die wichtigste Grundlage. Bezüglich der Spannungshöhe und -änderung werden in der Norm langsame und schnellen Spannungsänderungen unterschieden. Langsame Spannungsänderungen bezeichnen die durch unterschiedliche Belastungszustände des Netzes hervorgerufenen Abweichungen in der Höhe der Versorgungsspannung. Der Begriff schnelle Spannungsänderung bezeichnet hingegen die Differenz der Effektivwerte zweier aufeinanderfolgender Spannungswerte von nicht näher festgelegter Dauer [8]. Für die Höhe der Versorgungsspannung im NS-Netz gelten gemäß [8] die Anforderungen:

LANGSAME SPANNUNGSÄNDERUNGEN: Langsame Änderungen von $U_N = 400\text{V}$ sollten $\pm 10\%$ nicht überschreiten, wobei Fehler- und Unterbrechungssituationen von dieser Vorgabe ausgeschlossen werden. Hinsichtlich der Überprüfung dieser Anforderungen an den Übergabestellen zum Netznutzer muss gewährleistet werden, dass sich „95% der 10-Minuten-Mittelwerte des Effektivwertes der Versorgungsspannung jedes Wochenintervalls“ (siehe [8]) innerhalb dieses Bereichs befinden und dass „alle 10-Minuten-Mittelwerte des Effektivwertes der Versorgungsspannung innerhalb des Bereichs U_N +10%/-15% liegen“ (siehe [8]).

SCHNELLE SPANNUNGSÄNDERUNGEN: Die Anforderung an die Versorgungsspannung hinsichtlich der schnellen Spannungsänderungen ist in der Norm etwas weniger präzise formuliert. Diese dürfen „in der Regel 5% der Nennspannung U_N nicht überschreiten“ (siehe [8]), allerdings erlaubt die Norm zusätzlich unter nicht näher beschriebenen Umständen das Auftreten von „schnellen Spannungsänderungen bis zu 10% U_N mit kurzer Dauer mehrmals am Tag“ (siehe [8]).

Aus den genannten Anforderungen an langsame Spannungsänderungen wird deutlich, dass auf Grund der hohen Toleranz von $U_C \pm 10\%$ in der MS die volle Ausnutzung dieses Spannungsbereichs die Einhaltung der Vorgaben in der NS stark erschweren würde. Das im Netzbetrieb genutzte MS-Band wird deshalb vom VNB so eingeschränkt, dass die Einhaltung der Spannungsgrenzen in der NS noch für alle Betriebsfälle gewähr-

leistet werden kann. Dies bedeutet, dass die Planung im MS-Netz häufig auch einen Einfluss auf die unterlagerten NS-Netze mit sich bringen wird. Die Einhaltung der Grenzwerte für z. B. Oberschwingungsbelastungen, Flickererscheinungen und Unsymmetriezustände wird in der NS beim Anschluss von Endkunden und DEA mittels in weiteren technischen Richtlinien vorgegebenen Grenzwerten und vereinfachten Rechenverfahren überprüft. Für den Anschluss von DEA ist dabei die vom Verband der Elektrotechnik, Elektronik, Informationstechnik e.V. (VDE) veröffentlichte Arbeitsrichtlinie VDE AR 4105 [9] anzuführen, die z. B. auch den in dieser Arbeit durchgeführten Simulationen zu Grunde gelegt wird. Neben der DIN EN 50160 basiert diese Richtlinie auch wesentlich auf den technischen Normen DIN VDE 0838 [10] und DIN VDE 0839 [11]. Eine inhaltliche Zusammenfassung der Inhalte dieser Normen findet man z. B. in [7].

1.3 MOTIVATION, ZIELSETZUNG UND AUFBAU DER ARBEIT

Der Beurteilung der Aufnahmekapazität von Verteilnetzen für DEA und der effizienten Steuerung der im regulierten Netzgeschäft verknappten finanziellen Mittel kommt vor dem Hintergrund der aktuellen energiepolitischen Ziele in Deutschland eine immer größere Bedeutung zu. Insbesondere auf der NS-Ebene wird die Netzbeurteilung und die anschließende Steuerung von Investitionsmitteln auf Grund der Vielzahl an Netzen, deren Daten darüber hinaus häufig nicht in digitaler Form zu Verfügung stehen, stark erschwert. Vor diesem Hintergrund hat diese Arbeit folgende Zielsetzung:

- Identifikation und Erhebung relevanter Merkmale zur Charakterisierung von NS-Netzen,
- Entwicklung netzspezifischer Ansätze zur Klassifikation von NS-Netzen hinsichtlich ihres Aufnahmevermögens für DEA auf Basis dieser Mermale,
- Kombination unsicherer Klassifikationseinschätzungen (z. B. von menschlichen Experten), um Trainingsdaten bzw. Vergleichsdaten für Klassifikatoren zu generieren und
- Effizienzsteigerung bei der Klassifikation von NS-Netzen durch Anwendung von Support Vector Machines (SVM), um eine optimale Performanz bei der Klassifikation zu erhalten.

Insbesondere wird erwartet, dass die Ergebnisse langfristig zu einer Erhöhung der Planungssicherheit bei der Steuerung von Investitionsmitteln auf der NS-Ebene führen könnten. Dazu wird mit den erarbeiteten Methoden direkt eine Grundlage zur Verbesserung des zweiten der in Abschnitt 1.1.2 genannten Schritte bis zur Verstärkung eines NS-Netzes geschaffen. Als Folge daraus könnte eine erhöhte Planungssicherheit

einen VNB dazu ermutigen, den Ausbau seiner NS-Netze vorausschauender zu planen und in der Konsequenz technische Effizienzen in der Verteilung elektrischer Energie zu heben. Zusätzlich kann auf Basis der Ergebnisse die Auswahl von relevanten NS-Netzstrukturen für detaillierte Untersuchungen strukturierter gestaltet werden, indem z. B. von jeder Klasse eine bestimmte Anzahl an NS-Netzen für die Untersuchungen gewählt wird.

Der Aufbau der Arbeit ist eng an die formulierte Zielstellung angelehnt. Nachdem in diesem Kapitel eine Einleitung in die Thematik erfolgt, schließt sich in Kapitel 2 eine Übersicht zu verwandten Arbeiten an, deren Inhalt thematisch an den Einsatz von Methoden der Computational Intelligence (CI) in Netzplanung und -betrieb angelehnt ist. Darüber hinaus erfolgt auf dieser Grundlage eine Einordnung der Inhalte dieser Arbeit in den wissenschaftlichen Kontext. Der Stand des Wissens zu in dieser Arbeit angewandten Methoden wird einzeln jeweils im Rahmen von Vorbemerkungen dargelegt. Zur Klassifikation von Objekten werden diskriminierende Merkmale benötigt, die einer Beschreibung der Objekte in nachfolgenden Verarbeitungsschritten dienen. Idealerweise ist dies eine Menge von wenigen relevanten, starken Merkmalen, um die Informationsflut zu reduzieren und gleichzeitig einen hinreichend hohen Informationsgehalt in den Daten zu wahren [12]. Kapitel 3 widmet sich daher der Identifikation und Erhebung geeigneter Datengrundlagen zur Charakterisierung von NS-Netzen. In Kapitel 4 werden zwei netzspezifische Ansätze zur Klassifikation von NS-Netzen hinsichtlich ihres Aufnahmevermögens für DEA erarbeitet. Diese Ansätze bilden die Grundlage zur Bewertung von NS-Netzen und einer späteren Anwendung von SVM. Auf Grund der unterschiedlichen Schwerpunkte, die in den netzspezifischen Klassifikationsansätzen Berücksichtigung finden und der Tatsache, dass die Aussagen eines einzelnen Experten aus vielen Gründen mehr oder weniger unsicherheitsbehaftet sind, widmet sich Kapitel 5 der Kombination von mehreren, potenziell falschen Klassifikationseinschätzungen. Um für die Bewertung weiterer NS-Netze auf die Anwendung der netzspezifischen Klassifikationsansätze verzichten zu können und vor diesem Hintergrund eine Effizienzsteigerung bei der Klassifikation zu generieren, werden in Kapitel 6 drei verschiedene SVM-basierte Klassifikationskonzepte vorgestellt und empirisch untersucht. Einer Reflexion der Arbeit mit einem anschließenden Fazit widmet sich Kapitel 7. Abschließend erfolgt ein Ausblick zu weiterem Forschungsbedarf und eine Diskussion offengebliebener Fragestellungen.

Zu den in Kapitel 4 beschriebenen netzspezifischen Klassifikationsansätzen sind bereits erste Ergebnisse in den „IEEE Transactions on Power Systems“ veröffentlicht worden [13]. In dem Artikel wird ein simulationsbasierter Klassifikationsansatz für NS-Netze vorgestellt, der an den normativen Stand der Technik zur Bewertung der Aufnahmekapazität von NS-Netzen für DEA angelehnt ist. Der Ansatz wird anhand der Daten von 300 realen, ländlichen und vorstädtischen NS-Netzen statistisch gegen Klassifikationseinschätzungen von Experten aus der Verteilnetzplanung evaluiert.

2

EINSATZ VON COMPUTATIONAL INTELLIGENCE IN NETZPLANUNG UND -BETRIEB

2.1 ÜBERSICHT ZU VERWANDTEN ARBEITEN

Innerhalb der letzten Jahre wurde bereits in einigen Arbeiten der Einsatz von Paradigmen der CI in Planung und Betrieb von Verteilnetzen untersucht. Die Anwendung konzentrierte sich bisher im Wesentlichen auf die Bereiche Lastvorhersage, Netzbetrieb und -regelung, Beurteilung der Systemzuverlässigkeit und -sicherheit sowie die Klassifikation von Betriebszuständen und Systemfehlern im Netz [14]. Im Folgenden wird ein Überblick zu publizierten Forschungsergebnissen geschaffen, um auf dieser Grundlage in Abschnitt 2.2 eine Einordnung dieser Arbeit in den wissenschaftlichen Kontext vorzunehmen. Die Darstellung der publizierten Forschungsergebnisse wird entsprechend der im Rahmen dieser Arbeit durchgeführten Untersuchungen weitestgehend auf die NS-Ebene fokussiert.

Der Einsatz von Methoden der CI wird häufig im Zusammenhang mit der Entwicklung intelligenter Stromnetze unter dem Schlagwort Smart Grids gesehen [15, 16, 17]. Innerhalb der letzten Jahre wurden einige Arbeiten von initialem Charakter zur Analyse der Potentiale zum Einsatz von CI in Smart Grids veröffentlicht. In [15] werden Rahmenbedingungen für eine zukünftige Einbindung von Methoden der CI zur Entwicklung von Smart Grids gegeben. Das Hauptaugenmerk liegt dabei vor dem Hintergrund des wachsenden Anteils von EE und einem zukünftigen Lastzuwachs durch Elektrofahrzeuge auf der Verbesserung der Datenverarbeitung sowie der Zuverlässigkeit und Sicherheit im Netz. Die Machbarkeit wird anhand einiger Anwendungsbeispiele demonstriert. Einen ähnlichen Ausblick zu Potenzialen für die Nutzung von CI in Smart Grids findet man in [16]. Der Schwerpunkt des Ausblicks liegt dabei ebenfalls auf der Bewältigung der durch weitere Integration von EE und einem erwarteten Anschluss von Elektrofahrzeugen hervorgerufenen Komplexitätszunahme in Netzplanung, -betrieb und regelung. Dazu werden einige Methoden der CI vorgestellt und deren Potenziale anhand von Simulationsergebnissen demonstriert. Zusätzlich wird in [17] eine Übersicht zur An-

wendung von Methoden der CI im Stromnetz der Zukunft vorgeschlagen. Als spezielle Anwendungsfälle werden beispielhaft Einsatzmöglichkeiten einer situationsbedingten Weitbereichsbeobachtung sowie einer auf adaptiver dynamischer Programmierung basierenden intelligenten Regelung diskutiert. Eine ähnliche Vision zum Stromnetz der Zukunft und dem damit verbundenen Einsatz von Methoden der CI, die im Rahmen des „IEEE Computer Society Smart Grid Vision Projektes“ entwickelt wurde, findet man in [18]. Darüber hinaus werden erste Ansätze zur dynamischen Preissetzung im Smart Grid mit CI-Methoden in [19] untersucht.

Eine wichtige Funktion im zukünftigen Netzbetrieb kommt der Klassifikation von Betriebszuständen und Systemfehlern im Netz zu, damit Fehler sicher erkannt und Gegenmaßnahmen effizient eingeleitet werden können. Zhang et al. entwickeln auf der Basis eines SVM-Klassifikators einen Modellierungsansatz, mit der im Schutzrelais anhand von Messdaten zwischen normalen und fehlerhaften Betriebsbedingungen im Netz diskriminiert werden kann. Die zur Klassifikation der Betriebszustände benötigten Features können dabei online aufgenommen werden. Die Anpassung an verschiedene Systemzustände wird durch ein Onlineupdate des SVM-Klassifikators gewährleistet. Durch Simulationsergebnisse wird gezeigt, dass SVM eine brauchbare Methodik ist, um die Schutzrelaistechnik intelligenter und adaptiver zu gestalten [20]. In [21] stellen Pisica und Eremia eine neue, auf überwachtem Lernen basierende Methodik zur Zustandsklassifikation vor. Es wird insbesondere darauf hingewiesen, dass die Methodik durch Integration in vorhandene Messsysteme geeignet sei, um den Entwurf und die Entwicklung von Smart Grids, aber auch die autonome Behebung von Systemfehlern zu unterstützen. Die Klassifikation des Systemzustandes wird sowohl anhand eines Meta-Algorithmus (Adaptive Boosting), als auch anhand von Classification and Regression Trees (CART) vorgenommen. Die Klassifikationsgüte der entwickelten Methodik wird anhand des „IEEE 39-Knoten-Referenznetzes“ demonstriert. Zwei weitere neuartige Ansätze zur Fehlerdetektion und -klassifikation auf selektiven Leitungsstrecken werden in [22] entwickelt. Die erste Methode nutzt dazu zeitliche und attributive Korrelationen von an der Leitungsstrecke während der transienten Phase aufgenommenen Messdaten. Der zweite Ansatz basiert auf einer Ein-Klassen-SVM-Formulierung und nutzt die Korrelation zwischen Attributen nur zur automatischen Fehlerklassifikation. Darüber hinaus beschäftigen sich Haruna et al. mit dem Einsatz von CI in Lastabwurfkonzepten [23].

Zur effizienten Versorgung von Endkunden kommt zukünftig der Prognose von Kundengewohnheiten und Lastverhältnissen im Netz eine immer bedeutendere Rolle zu. Li et al. zeigen, dass Algorithmen der CI zur Prognose von Kundengewohnheiten und -vorzügen in der Nutzung von vier Haushaltsgeräten genutzt werden können [24]. In den Untersuchungen wird die Vorhersagegüte von SVM, Multilayer Perceptren (MLP) und Naive-Bayse-Klassifikatoren auf der Basis von synthetischen Daten zum Nutzerverhalten und -gewohnheiten verglichen, die anhand eines Fragebogens von 32 Nutzern erhoben wurden. Yuanzhe et al. nutzen SVM zur Prognose der aggregierten Last im

Netz [25]. Zur Suche der Parameter für die SVM werden Genetische Algorithmen (GA) eingesetzt. Die Vorhersagegüte des resultierenden Modells wird für ein reales Stadtnetz in der Provinz Hebei mit einer herkömmlichen Methode verglichen, die eine Lastprognose auf Basis ähnlicher Tage vornimmt. Einen Schritt weiter gehen Chandler und Hughes. Sie stellen einen experimentellen, intelligenten Regler vor, der in ein vorhandenes Regelungssystem integriert wird und Echtzeitdaten, historische Daten und Vorhersagedaten nutzt, um den Energiebedarf im Netz für die nächsten 72 Stunden zu prognostizieren [26]. Auf der Basis der Prognose wird versucht, eine nahezu optimale Versorgungsplanung für die nächsten 24 Stunden abzuleiten.

Auch die Planung und Bewertung von Netzstrukturen erfährt durch den Einsatz von CI ein erhebliches Verbesserungspotenzial. Stark und Krost entwickeln z. B. ein Entwurfstool für Micro-Grids, das eine systematische Planung auf Basis von Methoden der CI erlaubt [27]. Die präsentierte Entwurfsmethodik basiert auf dem Einsatz von Artificial Neural Netzworks (ANN), eines Experten- sowie eines Fuzzy-Systems. Gross et al. präsentieren einen Ansatz zur Bewertung der Ausfallwahrscheinlichkeit von Versorgungsleitungen in einem Stadtnetz. Die Leitungen werden dabei anhand von aggregierten physischen Daten und Zeitreihen mittels SVM in eine Rangfolge gebracht. Die Evaluation des Ansatzes erfolgt anhand von realen Daten eines Stadtnetzes in New York [28]. Rudin et al. entwickeln diese Egebnisse zu einem allgemeinen Prozess weiter, der einem VNB die Fehlervorhersage und präventive Wartung im Netzbetrieb erleichtern soll [14].

2.2 KONTEXTUELLE EINORDNUNG UND INNOVATIONSGEHALT

Diese Arbeit lässt sich thematisch in den Bereich des Einsatzes von Methoden der CI zur Planung und Bewertung von Netzstrukturen einordnen. Die bereits in diesem Bereich publizierten Ergebnisse von Stark und Krost [27] beziehen sich auf den Entwurf von Micro-Grids, wobei ANN und ein Fuzzy-System Anwendung finden. Die eingesetzten Paradigmen der CI unterscheiden sich folglich von den in dieser Arbeit verwendeten und werden nicht zur Bewertung, sondern zur verbesserten Planung von autarken NS-Netzen genutzt. Gross et al. [28] und Rudin et al. [14] setzen ähnlich zu dieser Arbeit SVM zur Bewertung von Netzstrukturen ein. Die Anwendung konzentriert sich dabei auf die Erzeugung von Rangfolgen bezüglich des Ausfallverhaltens von Versorgungsleitungen und Betriebsmitteln. Zum Training der SVM werden sowohl physikalische Betriebsmitteldaten, als auch historische Zeitreihen zur Belastung der Betriebsmittel analysiert und vorverarbeitet. Die Datennutzung und die generierten Rangfolgen beziehen sich allerdings jeweils nur auf einen speziellen Betriebsmitteltyp. Eine Bewertung von betriebsmittelübergreifenden Netztopologien wird nicht vorgenommen. Hervorzuheben ist, dass die entwickelte Methodik anhand eines großen realen Datenbestandes evaluiert wird.

Die Ergebnisse dieser Arbeit zielen, ähnlich zu den in [14, 28] entwickelten Methoden, auf die Bewertung von realen Netzen ab. Der Beitrag zum in Abschnitt 2.1 abgegrenzten wissenschaftlichen Kontext liegt dabei in der erstmaligen Erarbeitung effizienter Ansätze zur Klassifikation von NS-Netzen. Die Ansätze sind dabei auf eine effiziente Bewertung der gesamten Netztopologie ausgerichtet. Vor diesem Hintergrund ist auch der Einsatz von SVM zur Bewertung von NS-Netzen als neuartig anzusehen.

Im Detail werden erstmalig zwei netzspezifische Ansätze zur Bewertung des Aufnahmevermögens von NS-Netzen für DEA anhand von ordinalen Klassen vorgestellt. Zusätzlich wird eine neue Regel, die Erweiterte Imprecise-Dirichlet-Model-Regel (EIDMR), zur Kombination von unsicherheitsbehafteten Klassifikationsergebnissen erarbeitet und mit einigen bekannten Kombinationsregeln verglichen. Dabei zeigt sich, dass die Verwendung der EIDMR vorteilhaft ist, wenn zum einen ordinale Klassen verarbeitet werden müssen und zum anderen die vorliegende Anzahl an Klassifikationseinschätzungen für ein einzelnes Objekt gering ist. Darüber hinaus werden drei unterschiedliche Klassifikationskonzepte zum Einsatz von SVM bei unsicheren Labeln (mehrere evtl. verschiedene Label für einen Datenpunkt) untersucht. Im Gegensatz zu bekannten Ansätzen zur Begegnung dieser Unsicherheit reduziert sich durch den Einsatz einer Kombinationsregel die Anwendung der SVM auf die Prognose der kombinierten Label. Hier zeigt sich in den empirischen Experimenten eine deutlich höhere Klassifikationsgüte bei Verwendung der EIDMR gegenüber den verwendeten anderen Kombinationsregeln. Weiterhin wird dargelegt, dass der SVM-Klassifikator durch Berücksichtigung von ordinalen Informationen bei der Lösung des Mehrklassenklassifikationsproblems weiter verbessert werden kann. Um den Anwendungsbezug der in dieser Arbeit untersuchten Methoden zu untermauern, werden diese jeweils anhand von 300 realen, ländlichen und vorstädtischen NS-Netzen in empirischen Experimenten evaluiert. Über die hier erfolgte zusammenfassende Darlegung hinaus, wird der methodische Innovationsgehalt im weiteren Verlauf dieser Arbeit jeweils im Rahmen von Vorbemerkungen diskutiert.

3

MERKMALE ZUR CHARAKTERISIERUNG VON NIEDERSPANNUNGSNETZEN

Zur Klassifikation von Objekten werden diskriminierende Merkmale (bzw. Features) benötigt, anhand derer die zu klassifizierenden Objekte für nachfolgende Verarbeitungsschritte in geeigneter Weise repräsentiert werden können. Idealerweise ist dies eine Menge von wenigen relevanten, starken Merkmalen, um die Informationsflut zu reduzieren und gleichzeitig einen hinreichend hohen Informationsgehalt zu wahren. Das Grundproblem besteht also darin, eine geeignete Datengrundlage zur Charakterisierung der zu klassifizierenden Objekte zu identifizieren [12]. Obwohl Daten ein abstraktes Konzept zur Beschreibung von elektrischen Netzen darstellen und diese entsprechend dem Untersuchungsgegenstand interpretiert und ausgelegt werden müssen, bieten sie im Gegensatz zur reinen visuellen Darstellung von Netzen (z. B. in einem Netzplan) ein geeignetes Konzept zur Diskriminierung. Der Grund dafür liegt in der Möglichkeit, unmittelbar numerische Vergleiche zwischen einer großen Anzahl verschiedener Netze durchzuführen. Weiterhin ermöglicht eine Charakterisierung anhand von Daten im Gegensatz zur visuellen Beschreibung eine einfache und effiziente Weiterverarbeitung in nachfolgenden Verarbeitungsschritten [29].

Inhaltlich ist dieses Kapitel wie folgt aufgebaut. In Abschnitt 3.1 werden zehn netzspezifische Merkmale identifiziert, die einen starken Bezug zur Verteilnetzplanung aufweisen und deren Erhebung folglich ein gewisses Maß an Anwendungswissen voraussetzt. Zusätzlich wird die Relevanz der Merkmale hinsichtlich einer Bewertung des Aufnahmevermögens von NS-Netzen für DEA erläutert. Zur Veranschaulichung werden anschließend die Daten zweier realer Musternetze miteinander verglichen. In Abschnitt 3.2 werden weitere 22 topologische Merkmale aus der Graphentheorie (siehe z. B. [30]) identifiziert und definiert, die zur Charakterisierung von NS-Netzen geeignet sind. Diese graphspezifischen Merkmale werden anschließend ebenfalls anhand einiger ausgewählter Merkmale für zwei Musternetze miteinander verglichen.

3.1 NETZSPEZIFISCHE MERKMALE

3.1.1 *Identifikation netzspezifischer Merkmale*

Die Identifikation und Erhebung von netzspezifischen Merkmalen, die einen Bezug zur Verteilnetzplanung haben, benötigt Anwendungswissen und kann, je nachdem welche Merkmale erhoben werden, aufwendig sein. Oft liegen in der Praxis Informationen zur NS-Ebene noch nicht vollständig in digitaler Form vor. Allerdings ist netzspezifischen Merkmalen eine Nähe zum Untersuchungsgegenstand immanent. Die Nutzung von netzspezifischen Merkmalen kann daher vorteilhaft sein, wenn NS-Netze z. B. von menschlichen Experten bewertet werden sollen. Zusammen mit fünf Experten aus der Verteilnetzplanung wurden folgende zehn netzspezifische Merkmale identifiziert:

1. Anzahl der Hausanschlüsse,
2. Anzahl der Transformatorstationen (ohne kundeneigene Stationen),
3. Summe der Transformatornennleistungen in kVA,
4. Anzahl der Kabelverteilerschränke (KVS),
5. Summe der Leitungslängen in m,
6. Summe der vermaschten Leitungslängen in m,
7. Vermaschungsgrad in %,
8. Anteil der Leitungen des Typs NAYY 4x150mm^2 in %,
9. maximaler Stationsradius in m und
10. durchschnittlicher Stationsradius in m.

Die ersten fünf dieser Merkmale verkörpern in der Praxis übliche, eindeutige Begrifflichkeiten zur Beschreibung von Netzstrukturen. Die übrigen Merkmale sind bisher nicht eindeutig definiert. Im Folgenden wird daher eine Definition der Merkmale sowie eine Motivation zu deren Auswahl vorgenommen:

1. Anzahl der Hausanschlüsse:
Die Hausanschlüsse eines NS-Netzes verkörpern nicht nur dessen Verbrauchsstellen, sondern im Zuge des Ausbaus von EE häufig auch dessen Einspeisepunkte (z. B. bei

PV-Anlagen). Ihre Anzahl kann somit als Maß für die Lastsituation und das Ausbaupotenzial für DEA im NS-Netz betrachtet werden.

2. Anzahl der Transformatorstationen:
Transformatorstationen stellen im NS-Netz die Einspeisestellen aus der übergelagerten MS-Ebene dar. Das einem Transformator nachgelagerte Leitungsnetz der NS-Ebene wird bis auf wenige Ausnahmen als separater Stationsbereich, ohne betriebsmäßig geschlossene galvanische Verbindung zu einem anderen Transformator, geschaltet. Im Allgemeinen gilt, dass in großen NS-Netzen auch eine höhere Anzahl an Transformatorstationen für die Versorgung bereitgestellt werden muss. Die sich ergebende Anzahl an Transformatorstationen enthält daher Informationen über die Größe und Ausdehnung eines NS-Netzes.

3. Summe der Transformatornennleistungen in kVA:
Die Nennleistung eines Transformators beschreibt dessen Übertragungsvermögen für eine elektrische Scheinleistung. Sie wird in der Praxis gemäß der Last- bzw. Erzeugungsverhältnisse des Stationsbereiches dimensioniert und ist ein begrenzender Faktor für den Leistungsaustausch mit dem übergelagerten MS-Netz. Die Summe aller Transformatornennleistungen gibt damit Auskunft über die maximal mögliche Bezugs- bzw. Rückspeiseleistung des NS-Netzes.

4. Anzahl der KVS:
KVS dienen im NS-Netz als selektive Punkte. In ihnen werden ankommende Leitungen entweder – je nach Schutzkonzept des Netzes – über Sicherungen bzw. Trennmesser verbunden oder aber zur Einrichtung einer Netztrennstelle offen geschaltet. Sie bieten einem VNB damit Möglichkeiten, den Betrieb der Netze an die Last- bzw. Erzeugungssituation eines Stationsbereichs anzupassen und Reserveleistung zwischen verschiedenen Stationsbereichen durch Schalthandlungen bereitzustellen. Ein Zusammenschalten von Leitungen in einem KVS kann z. B. zur Erhöhung der Vermaschung bzw. Stärkung der Netzstruktur und folglich unmittelbar zur Reduktion der Netzimpedanz führen. Die gesamte Anzahl der KVS eines NS-Netzes stellt damit ein Maß für die Stärke der vorhandenen Netzstruktur bereit.

5. Summe der Leitungslängen in m:
Dieses Merkmal beinhaltet – unbeachtet des Leitungstyps – die Summe der Längen aller in einem NS-Netz verlegten Leitungsabschnitte. Bei gegebener räumlicher Ausdehnung führt eine höhere Summe der Leitungslängen zu einer stärkeren Netzstruktur, da bei gleicher Anzahl von Hausanschlüssen im Netz mehr Leitungen zu deren Versorgung mit elektrischer Energie eingesetzt werden.

6. Summe der vermaschten Leitungslängen in m:
Die Summe der vermaschten Leitungslänge umfasst die Summe aller Leitungslängen eines NS-Netzes, die betriebsmäßig nicht als Stichleitung betrieben werden. In der ver-

maschten Leitungslänge werden somit auch nicht die häufig in NS-Netzen vorzufindenden offenen Ringtopologien berücksichtigt, sondern lediglich die Topologien, die auf Grund der Verschaltung Ihrer Leitungen zu einer Impedanzverringerung – und folglich zur Stärkung der Strukturen – im Netz führen.

7. Vermaschungsgrad in %:
Auf Basis der letztgenannten Merkmale wird der Vermaschungsgrad als das Verhältnis zwischen der Summe der vermaschten Leitungslänge und der Summe aller Leitungslängen in einem NS-Netz definiert. Da eine Vermaschung der Netzstruktur unmittelbar eine Verringerung der resultierenden Netzimpedanz auf den betroffenen Leitungen zur Folge hat, beinhaltet dieses Merkmal eine von der räumlichen Ausdehnung des NS-Netzes unabhängige Information zur Stärke der Netzstruktur.

8. Anteil der Leitungen des Typs NAYY 4x150mm^2 in %:
Um den Wert für den Anteil der Leitungen des Typs NAYY 4x150 mm^2 zu ermitteln, werden zunächst alle Leitungslängen der Kabel vom Typ NAYY 4x150 mm^2 aufsummiert, das in der Praxis bei einer Vielzahl von VNB aktuell als Standardkabel genutzt wird. Die so ermittelte Länge wird schließlich – in gleicher Weise wie beim Vermaschungsgrad – ins Verhältnis zur Summe der Leitungslängen des betrachteten NS-Netzes gesetzt.

9. Maximaler Stationsradius in m:
Bei der Einhaltung der Spannungsbandgrenzen aus [8] bzw. der Grenzwerte für die Netzverträglichkeit aus [9] ist die räumliche Ausdehnung der Stationsbereiche eines NS-Netzes ein wichtiger Einflussfaktor. Je weiter ein Stationsbereich räumlich ausgedehnt ist, desto höher wird in der Regel auch die am entferntesten Punkt vorliegende Netzimpedanz ausfallen. Eine hohe Netzimpedanz kann beim Energietransport hohe Spannungsfälle bzw. -hübe verursachen und ist eine wichtige Größe bei der Abschätzung von Netzrückwirkungen durch DEA (siehe z. B. [7]). Um die Ausdehnung zu erfassen, wird der Stationsradius als Luftlinienentfernung zwischen dem Transformator und dem am weitesten entlegenen Knoten im Stationsbereich festgelegt. Der maximale Stationsradius gibt das Maximum über alle Stationsbereiche eines NS-Netzes an.

10. Durchschnittlicher Stationsradius in m:
Entsprechend der Definition des Stationsradius ergibt sich der durchschnittliche Stationsradius als Mittelwert über alle Stationsbereiche im NS-Netz.

3.1.2 *Vergleich der netzspezifischen Merkmale zweier Musternetze*

Um die Relevanz der identifizierten netzspezifischen Merkmale hinsichtlich der Charakterisierung von NS-Netzen zu veranschaulichen, erfolgt nun beispielhaft eine Interpre-

tation dieser Merkmale für zwei reale, ländliche NS-Netze. Die Struktur dieser Musternetze ist in Abbildung 1 gezeigt. Darüber hinaus sind in Tabelle 1 die Daten dieser Musternetze ausgewiesen.

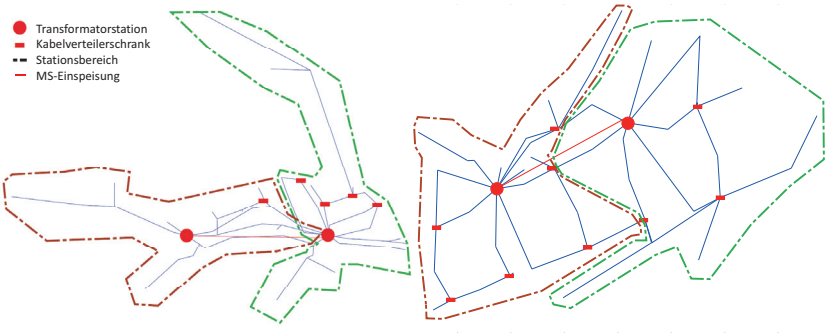


Abbildung 1: Struktur zweier realer, ländlicher Niederspannungsnetze, Netz I (links) und Netz II (rechts).

MERKMAL	NETZ I	NETZ II
1 Anzahl der Hausanschlüsse	92	100
2 Anzahl der Transformatorstationen	2	2
3 Summe der Transformatornennleistungen [kVA]	880	1260
4 Anzahl der Kabelverteilerschränke	5	9
5 Summe der Leitungslängen [m]	7475	5375
6 Summe der vermaschten Leitungslängen [m]	1764	3105
7 Vermaschungsgrad [%]	23,6	57,8
8 Anteil der Leitungen des Typs NAYY 4x150 mm^2 [%]	50,3	62,8
9 Maximaler Stationsradius in [m]	1082	354
10 Durchschnittlicher Stationsradius in [m]	840	344

Tabelle 1: Netzspezifische Merkmale der beiden Musternetze.

Es ist ersichtlich, dass sich beide Netze offensichtlich nicht nur stark in ihren Strukturen, sondern auch in den Werten ihrer Merkmale unterscheiden. Wie Merkmal 2 zeigt, bestehen zwar beide Netze aus jeweils zwei Stationsbereichen und versorgen mit jeweils ungefähr 100 Hausanschlüssen eine ähnliche Anzahl an Endkunden (Merkmal 1), aber ein Vergleich der Merkmale 9 und 10 beider Netze lässt erkennen, dass die durch Netz I versorgte Fläche eine ungefähr um den Faktor 2,5-mal größere räumliche Ausdehnung aufweist als die von Netz II. Die galvanische Auftrennung der beiden Netze in jeweils zwei Stationsbereiche erfolgt in der Praxis in den in Abbildung 1 gezeigten KVS, die sich in der Nähe der parallel verlaufenden Kanten der Polygone um die Stationsbereiche

befinden. Hinsichtlich der unterschiedlichen räumlichen Dimensionierung der versorgten Flächen tritt – bei gleicher Anzahl von Stationsbereichen – im Netz II ein höherer Vermaschungsgrad als in Netz I auf. Dies wird direkt durch die beiden Ausprägungen des Merkmals 7 deutlich. Die Begründung für die unterschiedliche Vermaschung liegt in der ökonomischen Effizienz bei der Versorgung der Endkunden. Ein weiterer Indikator für diesen Fakt ist die stark unterschiedliche Anzahl von installierten KVS, den selektiven Knoten in einem NS-Netz, die neun in Netz II und lediglich fünf in Netz I beträgt. Im Gegensatz zu Netz II besteht das Netz I nur aus einem vermaschten Bereich in dessen Zentrum, wohingegen die in den Außenbereichen liegenden Endkunden durch strahlenförmige Leitungszüge versorgt werden. Die Interpretation der netzspezifischen Merkmale der beiden Musternetze zeigt, dass Netz II eine wesentlich robustere Struktur zur Integration von DEA aufweist als Netz I.

3.2 GRAPHSPEZIFISCHE MERKMALE

3.2.1 *Identifikation und Definition graphspezifischer Merkmale*

Im Folgenden werden die Strukturen eines NS-Netzes abstrakt als Graph interpretiert. Vor diesem Hintergrund werden 22 graphspezifische Merkmale aus der Graphentheorie identifiziert, die zur Charakterisierung von NS-Netzen geeignet sind. Diese Herangehensweise ermöglicht eine effiziente, automatisierte Erhebung einer abstrakten Datengrundlage für eine große Anzahl von Netzen, die kein Anwendungswissen erfordert.

Graphen werden häufig zur Abstraktion von Problemstellungen eingesetzt [31]. Im Laufe der letzten Jahre wurden im Wesentlichen Methoden zur Klassifikation von Graphen entwickelt, die auf der Anwendung von SVM basieren. Hierbei wird eine Kernelmatrix genutzt, um Ähnlichkeiten zwischen Graphenpaaren anhand einzelner, komplex zu berechnender graphspezifischer Gemeinsamkeiten (z. B. anhand gemeinsamer Untergraphen) zu bewerten. Die in der Kernelmatrix enthaltenen Ähnlichkeiten ermöglichen anschließend eine Klassifikation mit SVM. Trotz nachgewiesener Effektivität dieser Graph-Kernel-Methoden verursachen sie einen sehr hohen Berechnungsaufwand bei der Ermittlung der über die jeweils ausgewählte graphspezifische Gemeinsamkeit gemessenen Ähnlichkeiten zwischen den Graphen [29]. Häufig werden in der Kernelmatrix Gemeinsamkeiten zwischen den Graphen bezüglich zufälliger Wege [32, 33, 34], kürzester Wege [35], zyklischer Muster [36], Baumstrukturen [37, 38, 39, 40] und Subgraphen [40, 41, 42] bestimmt. Im Gegensatz zu den genannten Graph-Kernel-Methoden untersuchen Li et al. einen einfacheren Ansatz. Sie sammeln viele verschiedene, weniger aufwendig zu berechnende graphspezifische Merkmale von Graphen und berechnen auf dieser Basis die Ähnlichkeiten zwischen Graphen. Ihr Vorgehen beruht auf der Annahme, dass Graphen einer gleichen Klasse auch viele ähnliche graphspezifische Merkmale

haben sollten und diese Merkmale somit auch direkt zur Charakterisierung von Graphen eingesetzt werden können [29]. Eine rechnerisch komplexer Abgleich der Unterstrukturen von Graphen wie bei Graph-Kernel-Methoden ist damit nicht erforderlich. Die graphspezifischen Merkmale können direkt als Features interpretiert und zum Training einer SVM genutzt werden. Die Untersuchungsergebnisse von Li et al. auf realen Benchmark-Datensätzen zeigen, dass bei mindestens gleich guten Klassifikationsresultaten im Vergleich zu einigen bekannten Graph-Kernel-Methoden wesentlich weniger Rechenaufwand erforderlich ist. Diese grundlegende Idee der Nutzung vieler graphspezifischer Merkmale zur Charakterisierung von Graphen wird hier adaptiert. Viele der hier genutzten graphspezifischen Merkmale werden bereits in [29] zur Klassifikation von Graphen vorgeschlagen. Um kein Anwendungswissen bei der Datenerhebung vorauszusetzen und diese so effizient wie möglich zu gestalten, werden einige der in [29] genutzten graphspezifischen Merkmale, die gewichtete Graphen (anwendungsspezifische Zusatzinformationen sind annotiert) betreffen, vernachlässigt und durch andere, zusätzliche Merkmale ungewichteter Graphen ersetzt.

Vor der Beschreibung der hier verwendeten graphspezifischen Merkmale, wird zunächst der Begriff des ungerichteten, ungewichteten Graphen definiert [31]:

Definition 3.1 (Ungerichteter, ungewichteter Graph) *Ein ungerichteter, ungewichteter Graph G werde aus einem Tupel (V, E) gebildet. V bezeichne eine endliche und nichtleere Menge von Knoten der Mächtigkeit n. Die Menge E bestehe aus einer Teilmenge der zweielementigen Teilmengen von V, formal also $E \subseteq \binom{V}{2} := \{(u_i, u_j) | u_i, u_j \in V, u_i \neq u_j\}$. Ihre Elemente werden als Kanten des Graphen bezeichnet.*

Auf der Basis dieser Definition bieten Graphen ein abstraktes Konzept zur Beschreibung der Struktur eines NS-Netzes, das nicht nur die Anzahl der im Netz vorhandenen Knoten widergibt, sondern auch deren Vernetzung durch die im NS-Netz vorhandenen elektrisch leitenden Verbindungen (Leitungen, Stationstransformator) veranschaulicht. Zur kompakten, formalen Darstellung der in Graphen enthaltenen Strukturinformationen wird in der Graphentheorie die Adjazenzmatrix $\mathbf{A}$ genutzt. Sie gibt an, welche Knoten in einem Graphen adjazent sind und wird wie folgt gebildet [31, 43, 44, 45]:

Definition 3.2 (Adjazenzmatrix) *Die Adjazenzmatrix $\mathbf{A}$ eines ungewichteten, ungerichteten Graphen sei die reelle, symmetrische $n \times n$-Matrix mit Zeilenindex i, Spaltenindex j sowie Elementen a_{ij} der Form*

$$a_{ij} = \begin{cases} 1 & \textit{für } (u_i, u_j) \in E, \\ 0 & \textit{sonst.} \end{cases}$$

Auf dieser Basis wurden folgende graphspezifische Merkmale ausgewählt, die Strukturinformationen zu NS-Netzen bereitstellen:

1. Gesamtanzahl der Knoten,
2. Gesamtanzahl der Kanten,
3. durchschnittlicher Knotengrad,
4. Anteil der Knotenzahl des größten verbundenen Untergraphen,
5. prozentualer Anteil von Endknoten,
6. durchschnittlicher Clusterkoeffizient,
7. durchschnittliche effektive Exzentrizität,
8. minimale effektive Exzentrizität,
9. maximale effektive Exzentrizität,
10. prozentualer Anteil zentraler Knoten,
11. durchschnittliche Nähezentralität,
12. durchschnittliche Intermediationszentralität,
13. Anzahl der Zusammenhangskomponenten,
14. durchschnittliche Knotenzahl der Zusammenhangskomponenten,
15. minimale Knotenzahl der Zusammenhangskomponenten,
16. durchschnittliche Kantenzahl der Zusammenhangskomponenten,
17. minimale Kantenzahl der Zusammenhangskomponenten,
18. spektraler Radius,
19. zweitgrößter Eigenwert,
20. Anzahl der verschiedenen Eigenwerte,
21. Energie und

22. prozentualer Anteil der Artikulationspunkte.

Es ist ersichtlich, dass die identifizierten Merkmale im Wesentlichen aus Basismerkmalen, pfadbasierten Merkmalen sowie spektralen Merkmalen der Graphen bestehen. Die spektralen Merkmale werden anhand der Eigenwerte der Adjazenzmatrix, dem Spektrum eines Graphen, bestimmt. Aus dem Spektrum des Graphen lassen sich auf der Grundlage der spektralen Graphentheorie Rückschlüsse auf dessen Struktur ziehen [43, 44]. In Analogie zu Abschnitt 3.1.1 werden nun die in dieser Arbeit verwendeten graphspezifischen Merkmale definiert und erläutert:

1. Gesamtanzahl der Knoten:
Dieses Merkmal gibt die Gesamtzahl n der Knoten bzw. die Mächtigkeit der Knotenmenge V des Graphen wider.

2. Gesamtanzahl der Kanten:
Dieses Merkmal gibt die Gesamtzahl e bzw. die Mächtigkeit der Kantenmenge E im Graphen wider.

3. Durchschnittlicher Knotengrad:
Der Grad $d(u_i)$ eines Knotens u_i wird aus der Anzahl der mit diesem Knoten direkt verbundenen (benachbarten) Kanten bestimmt. Der durchschnittliche Knotengrad eines Graphen ergibt sich dann durch Mittelwertbildung über alle n Knoten im Graphen: $\overline{d}(G) = \frac{1}{n}\sum_{i=1}^{n} d(u_i)$ [29].

4. Anteil der Knotenzahl des größten verbundenen Untergraphen:
Der Anteil der Knotenzahl des größten verbundenen Untergraphen am gesamten Graphen wird anhand des verbundenen Untergraphen ermittelt, der die höchste Anzahl an Knoten im Graphen aufweist. Diese Knotenzahl wird ins Verhältnis zur gesamten Knotenanzahl gesetzt [29].

5. Prozentualer Anteil von Endknoten:
Der prozentuale Anteil von Endknoten im Graphen wird durch die Knotenanzahl vom Grad eins bestimmt. Diese Knotenanzahl wird ins Verhältnis zur gesamten Knotenanzahl n im Graphen gesetzt.

6. Durchschnittlicher Clusterkoeffizient:
Der Clusterkoeffizient $c(u_i)$ eines Knotens u_i ist als Verhältnis zwischen den tatsächlich verlaufenden Kanten und der theoretisch möglichen Kantenzahl (bei z_i Nachbarn: $\frac{1}{2}z_i(z_i-1)$) zwischen seinen Nachbarknoten definiert. Der durchschnittliche Clusterkoeffizient resultiert aus der Mittelwertbildung über alle n Knoten im Graphen: $C(G) = \frac{1}{n}\sum_{i=1}^{n} c(u_i)$ [29].

7. Durchschnittliche effektive Exzentrizität:
Die Exzentrizität $\mathrm{ex}(u_i) = \max\{d(u_i, u_j)|u_j \in V\}$ eines Knotens u_i gibt das Maximum der kürzesten Wege $d(u_i, u_j)$ zu allen anderen Knoten im Graphen an. Bei ungewichteten Graphen hat ein Weg zwischen zwei benachbarten Knoten u_i, u_j die Länge eins. Die effektive Exzentrizität $\mathrm{ex}_{eff.}(u_i) = \max\{d(u_i, u_j)|u_j \in V\}$ wird als die maximale Länge der kürzesten Wege von u_i definiert, mit der mindestens 90% der Knoten im Graphen erreichbar sind. Die durchschnittliche effektive Exzentrizität eines Graphen wird durch Mittelwertbildung über alle n Knoten im Graphen bestimmt: $\mathrm{Ex}_{eff.}(G) = \frac{1}{n}\sum_{i=1}^{n} \mathrm{ex}_{eff.}(u_i)$ [29].

8. Minimale effektive Exzentrizität:
Die minimale effektive Exzentrizität wird über das Minimum der effektiven Exzentrizität im Graphen über alle n Knoten ermittelt und entspricht somit dem effektiven Radius des Graphen: $\mathrm{rad}(G) = \min\{\mathrm{ex}_{eff.}(u_i)|u_i \in V\}$ [29].

9. Maximale effektive Exzentrizität:
Die maximale effektive Exzentrizität gibt den maximalen Wert der effektiven Exzentrizität über alle n Knoten im Graphen wider und entspricht dem effektiven Durchmesser eines Graphen: $\mathrm{diam}(G) = \max\{\mathrm{ex}_{eff.}(u_i)|u_i \in V\}$ [29].

10. Prozentualer Anteil zentraler Knoten:
Ein Knoten u_i ist ein zentraler Knoten im Graph, wenn seine Exzentrizität gleich der minimalen effektiven Exzentrizität des Graphen ist. Der prozentuale Anteil zentraler Knoten eines Graphen wird durch das Verhältnis der zentralen Knoten zur Gesamtzahl der Knoten im Graphen gebildet.

11. Durchschnittliche Nähezentralität:
Die Nähezentralität eines Knotens u_i ist durch den Kehrwert der durchschnittlichen Pfadlänge zwischen u_i und allen anderen Knoten im Graphen definiert. Die durchschnittliche Nähezentralität $\mathrm{Close}(G)$ wird durch Mittelwertbildung über alle n Knoten im Graphen ermittelt, wobei die Nähezentralität wie folgt beschrieben werden kann: $\mathrm{close}(u_i) = \frac{n-1}{\sum_{u_j \in V, u_j \neq u_i} d(u_i, u_j)}$ [29, 46].

12. Durchschnittliche Intermediationszentralität:
Die Intermediationszentralität misst die Wichtigkeit eines Knotens in einem Graphen anhand der kürzesten Wege, auf denen dieser Knoten liegt. Die Anzahl der kürzesten Wege zwischen zwei Knoten u_i und u_j sei durch $\sigma_{ij} = \sigma_{ji}$ gegeben. Weiterhin sei $\sigma_{ij}(u_k) = \sigma_{ji}(u_k)$ als die Anzahl der kürzesten Wege zwischen den Knoten u_i und u_j definiert, auf denen der Knoten u_k als Zwischenknoten liegt. Die durchschnittliche Intermediationszentralität $\mathrm{Betw}(G)$ wird wiederum durch Mittelwertbildung über alle n Knoten im Graphen ermittelt, wobei die Intermediationszentralität eines Knotens u_k formal wie folgt beschrieben werden kann: $\mathrm{betw}(u_k) = \sum_{i \neq j \neq k} \frac{\sigma_{ij}(u_k)}{\sigma_{ij}}$ [46].

13. Anzahl der Zusammenhangskomponenten:
Diese Eigenschaft gibt die Anzahl der zusammenhängenden Komponenten im Graph (entspricht der Anzahl der Transformatorstationen bzw. den Stationsbereichen eines Niederspannung (NS)-Netzes) an.

14. Durchschnittliche Knotenzahl der Zusammenhangskomponenten:
Es wird zunächst für jede Zusammenhangskomponente im Graphen die Knotenanzahl bestimmt und anschließend der Mittelwert über alle Zusammenhangskomponenten gebildet.

15. Minimale Knotenzahl der Zusammenhangskomponenten:
Es wird zunächst für jede Zusammenhangskomponente im Graphen die Knotenanzahl bestimmt und anschließend das Minimum über alle Zusammenhangskomponenten ermittelt.

16. Durchschnittliche Kantenzahl der Zusammenhangskomponenten:
Es wird zunächst für jede Zusammenhangskomponente im Graphen die Kantenanzahl bestimmt und anschließend der Mittelwert über alle Zusammenhangskomponenten bestimmt.

17. Minimale Kantenzahl der Zusammenhangskomponenten:
Es wird zunächst für jede Zusammenhangskomponente im Graphen die Kantenanzahl bestimmt und anschließend das Minimum über alle Zusammenhangskomponenten ermittelt.

18. Spektraler Radius:
Der spektrale Radius eines Graphen ist definiert als betragsmäßig größter Eigenwert der Adjazenzmatrix $\mathbf{A}$ des Graphen. Mit s betragsmäßig verschiedenen Eigenwerten $|\lambda_1| > |\lambda_2| > \cdots > |\lambda_s|$ der Adjazenzmatrix ist der spektrale Radius $\rho(G)$ wie folgt berechenbar: $\rho(G) = |\lambda_1|$ [29, 43, 44].

19. Zweitgrößter Eigenwert:
Entsprechend der Definition des spektralen Radius eines Graphen gibt dieses Merkmal den betragsmäßig zweitgrößten Eigenwert $|\lambda_2|$ der Adjazenzmatrix $\mathbf{A}$ des Graphen an [29, 43, 44].

20. Anzahl der verschiedenen Eigenwerte:
Dieses Merkmal gibt die Anzahl der verschiedenen Eigenwerte (ohne Berücksichtigung der Betragsfunktion) $s \leq n$ der Adjazenzmatrix $\mathbf{A}$ des Graphen an [29, 43, 44].

21. Energie:
Die Energie eines Graphen ist definiert als quadratische Summe der Eigenwerte der Adjazenzmatrix $\mathbf{A}$ des Graphen: $E(G) = \sum_{i=1}^{n} \lambda_i^2$ [29, 43, 44].

22. Prozentualer Anteil der Artikulationspunkte:
Ein Knoten u_i in einem Graphen heißt Artikulationspunkt, wenn durch sein Entfernen der Graph zerfällt, d. h. sich die Zahl der zusammenhängenden Komponenten erhöht. Mit diesem Merkmal wird das Verhältnis der Artikulationspunkte zur Gesamtzahl an Knoten im Graphen gemessen [30].

3.2.2 *Vergleich der graphspezifischen Merkmale zweier Musternetze*

In Analogie zu Abschnitt 3.1.2 erfolgt nun eine Interpretation der für die beiden Musternetze aus Abbildung 1 erhobenen Merkmalwerte, um die Relevanz der identifizierten graphspezifischen Merkmale zur Charakterisierung von NS-Netzen zu zeigen. Auf Grund der großen Zahl an graphspezifischen Merkmalen wird die Erläuterung auf einige ausgewählte Merkmale eingeschränkt.

MERKMAL	NETZ I	NETZ II
3 Durchschnittlicher Knotengrad	2	2,18
7 Durchschnittliche effektive Exzentrizität	6,94	4,96
8 Minimale effektive Exzentrizität	3	3
9 Maximale effektive Exzentrizität	11	8
10 Prozentualer Anteil zentraler Knoten	0,15	0,18
12 Durchschnittliche Intermediationszentralität	123,0	73,5
13 Anzahl der Zusammenhangskomponenten	2	2
22 Prozentualer Anteil der Artikulationspunkte	25,0	15,8

Tabelle 2: Graphspezifische Merkmale der beiden Musternetze.

Bereits in Abschnitt 3.1.2 wurde festgestellt, dass beide Netze jeweils zwei Stationsbereiche und ungefähr die gleiche Anzahl an Hausanschlüssen versorgen, sich allerdings stark in ihren Strukturen unterscheiden. Die Anzahl der Stationsbereiche wird bei den graphspezifischen Merkmalen durch die Anzahl der Zusammenhangskomponenten (Merkmal 13) erfasst, der für beide Musternetze entsprechend einen Wert von zwei annimmt (Tabelle 2). Trotz gleicher Anzahl an Zusammenhangskomponenten unterscheiden sich die Musternetze in ihrer räumlichen Ausdehnung. Die von Netz I versorgte Fläche übersteigt die von Netz II ungefähr um den Faktor 2,5. Da die effektive Exzentrizität Auskunft über die maximale Länge der kürzesten Wege eine Knotens im Netz gibt, spiegeln sich diese Unterschiede in der räumlichen Ausdehnung der Netze in den Werten der effektiven Exzentrizität der Graphen wider (Merkmale 7, 8 und 9). Stimmen die Werte für die minimale effektive Exzentrizität, die dem effektiven Radius des Graphen entspricht, noch überein, so übersteigt der effektive Durchmesser von Netz I , der durch die maximale effektive Exzentrizität beschrieben wird, den Wert von Netz II ungefähr um das 1,4-fache. Auch die durchschnittlichen Werte der effekti-

ven Exzentrizität stehen im gleichen Verhältnis zueinander. Da die graphspezifischen Merkmale anhand von ungewichteten Graphen erhoben wurden, tritt der Unterschied in der räumlichen Ausdehnung beider Netze allerdings nicht direkt im realen Verhältnis (wie bei den netzspezifischen Merkmalen) auf. Die unterschiedliche Vermaschung der Netze kann den Werten des durchschnittlichen Knotengrads (Merkmal 3) und der Intermediationszentralität (Merkmal 12) entnommen werden. Die relativ geringen Werte des durchschnittlichen Knotengrads von 2,0 bzw. 2,18 beider Netze ist typisch für ländliche und vorstädtische NS-Netze. Beispielsweise liegen die Werte aller 300 untersuchten Netze zwischen 1,5 und 2,5. Die Intermediationszentralität gilt als Maß für die Wichtigkeit eines Knotens im Graphen. Je geringer der auftretende Wert, desto höher ist die Vermaschung der Netze. Die für beide Musternetze auftretenden Werte von 123,0 und 73,5 geben den unterschiedlichen Vermaschungsgrad der Netze deutlicher wider, als die Werte des durchschnittlichen Knotengrads. Der prozentuale Anteil der Artikulationspunkte gibt Auskunft über die Kompaktheit eines Graphen. Da durch das Entfernen eines Artikulationspunktes der Graph zerfällt, treten diese Knoten vor allem in strahlenförmigen NS-Netzen auf. Dabei gilt offensichtlich, je höher der Anteil der Artikulationspunkte, desto „strahlenförmiger“ ist das Netz. Diese Interpretation wird durch die in den beiden Musternetzen auftretenden Werte von 25,0% und 15,8% gestützt. Analog zu Abschnitt 3.1.2 zeigt auch die Interpretation einiger graphspezifischer Merkmale beider Musternetze, dass Netz II eine wesentlich robustere Struktur zur Integration von DEA aufweist als Netz I.

3.3 EINORDNUNG DER ERGEBNISSE

In diesem Kapitel wurden diskriminierende Merkmale identifiziert und in zwei unterschiedlichen Datensätzen, den netzspezifischen und den graphspezifischen Merkmalen, gebündelt. Damit wurde eine Grundlage geschaffen, um im weiteren Verlauf dieser Arbeit NS-Netze zu charakterisieren. Die netzspezifischen Merkmale finden dabei zunächst in Abschnitt 4.1 Verwendung, um die Strukturen von NS-Netzen menschlichen Experten im Rahmen eines expertenbasierten Klassifikationsansatzes zu beschreiben. Darüber hinaus werden die Merkmale beider Datensätze in Abschnitt 6.3.3 als Features zur Implementierung von SVM-Klassifikatoren genutzt.

Eine Bewertung der Relevanz der Merkmale zur Bewertung von NS-Netzen wurde bis zu dieser Stelle anhand von Interpretationen der Merkmalwerte zweier Musternetze vorgenommen. Die Frage, welchen Merkmalen zur Klassifikation von NS-Netzen eine hohe Bedeutung zukommt, wird in Abschnitt 6.3.2 nochmal aufgegriffen und unter Berücksichtigung einer großen Anzahl von NS-Netzen mit zwei Verfahren der Feature-Selection untersucht.

4

NETZSPEZIFISCHE KLASSIFIKATIONSANSÄTZE

In technischen Anwendungen ist die Erhebung von Daten zur Charakterisierung der zu untersuchenden Objekte häufig mit wesentlich weniger Aufwand verbunden, als die Einordnung dieser Objekte in vordefinierte Klassen. Dies ist beispielsweise der Fall, wenn die Datenerhebung automatisiert erfolgt, für die Klassifikation allerdings menschliche Experten erforderlich sind. Die Bearbeitung einer Klassifikationsaufgabe erfordert in der Vielzahl der Anwendungen nicht unerhebliche Ressourcen an Zeit und finanziellen Mitteln. Dies gilt prinzipiell auch für die Klassifikation von NS-Netzen. Vor diesem Hintergrund werden in diesem Kapitel zwei Ansätze zur Klassifikation von NS-Netzen hinsichtlich ihres Aufnahmevermögens für DEA erarbeitet. Abschnitt 4.1 beschreibt einen expertenbasierten Ansatz zur Erhebung von ordinalen Klassen mit Hilfe menschlicher Experten. Ergänzend dazu wird in Abschnitt 4.2 ein weiterer Ansatz zur Klassifikation von NS-Netzen anhand von stochastischen Simulationen vorgestellt. Beide Ansätze werden für die bereits in den Abschnitten 3.1.2 und 3.2.2 genutzten Musternetze veranschaulicht. Die Erarbeitung von zwei verschiedenen Klassifikationsansätzen liegt darin begründet, dass eine „wahre“ Bewertung der Aufnahmekapazität von NS-Netzen im Sinne einer „ground truth“ nur durch einen tatsächlichen, realen DEA-Ausbau erfolgen kann, was zu Testzwecken offensichtlich nicht möglich ist. Zum Abschluss des Kapitels werden in Abschnitt 4.3 die Ergebnisse beider Klassfikationsansätze empirisch miteinander verglichen. Dabei wird anhand eines einseitigen Hypothesentests statistisch nachgewiesen, dass sich die empirischen Ergebnisse beider Ansätze zum einen nicht grundlegend widersprechen und darüber hinaus sogar eine signifikante Übereinstimmung zwischen diesen vorliegt.

4.1 EXPERTENBASIERTE ORDINALE KLASSIFIKATION

Im Folgenden wird ein Klassifikationsansatz zur Bewertung von NS-Netzen mit ordinalen Klassen durch menschliche Experten vorgestellt. Die Beschreibung erfolgt dabei zunächst an vielen Stellen sehr generisch und abstrakt. Dies ermöglicht durch Anpas-

sung des Klassifikationsansatzes (z. B. Auswahl entsprechender Merkmale) auch dessen Einsatz zur Bewertung anderer netzwerkbasierter Infrastrukturen (z. B. Straßennetze). Die Anwendung des Klassifikationsansatzes wird im Anschluss an NS-Netzen illustriert. Die Bereitstellung von Informationen für die Experten während der Klassifikation erfolgt auf Basis eines festen Klassifikationsschemas. Dessen konkrete Ausgestaltung und die Durchführung der Klassifikationsaufgabe erfolgt anhand eines der Aufgabenstellung angepassten Fragebogens. In dem Fragebogen ist zusätzlich die Erhebung weiterer Informationen von den Experten vorgesehen, die ex post eine Analyse der Qualität der Expertenaussagen ermöglichen.

Zunächst wird in Abschnitt 4.1.1 einführend eine Übersicht zur Erhebung und Verarbeitung von Expertenwissen in technischen Anwendungen dargestellt. Weiterhin werden subjektive Einflüsse auf die Klassifikationseinschätzungen von Experten diskutiert. Anschließend wird in Abschnitt 4.1.2 das übergeordnete Klassifikationsschema beschrieben. Das Design des Fragebogens wird in Abschnitt 4.1.3 erläutert. Daraufhin zeigt Abschnitt 4.1.4 Ansätze und Methoden zur Ex-post-Bewertung der Messqualität auf. Eine experimentelle Anwendung und Evaluation des expertenbasierten Klassifikationsansatzes anhand von NS-Netzen wird in Abschnitt 4.1.5 durchgeführt.

4.1.1 *Vorbemerkungen*

Die Erhebung und Verarbeitung von Expertenwissen ist in technischen Anwendungen ein weitverbreiteter Ansatz. Es wurde bereits viel Forschungsaufwand betrieben, um diesen Ansatz bestmöglich umzusetzen. Grundsätzlich sind viele Lösungsansätze denkbar. Dabei erfordert die Erhebung und Bewertung eines menschlichen Urteils oftmals auch ein hohes Maß an Interdisziplinarität. Häufig finden daher Methoden verschiedener Disziplinen Anwendung, wie z. B. der Psychologie. In [47] wird z. B. die Conjoint-Analyse in einem Erhebungsinterface eingesetzt, mit dem Expertenmeinungen gesammelt und in Modellparameter überführt werden können. Zur Erhebung von Expertenwissen ist aber zunächst der Einsatz eines problemspezifischen Fragebogens naheliegend. Bei der Konzeption des Fragebogens werden verschiedene Fragetypen wie z. B. erfahrungsbasierte, verhaltensorientierte und quantitative Fragen eingesetzt [48]. Keeney und Winterfeld beschreiben die Erstellung des Fragebogens als ein kooperatives Schachspiel mit den Experten, das eine hohe Bedeutung hinsichtlich der Wissensextraktion einnimmt und bei gutem Ergebnis zur Erleichterung der Eingewöhnung der Experten an die Befragung beiträgt [49]. Die Gestaltung der Struktur des Fragebogens hat dabei einen wichtigen Einfluss auf die Qualität der Wissenserhebung. Moser vertritt den Standpunkt, das Layout auf dem Trichterprinzip aufzubauen, wobei die allgemeinen und übergreifenden Fragen bzw. Aufgaben zu Beginn des Fragebogens gestellt werden sollten. Die konkreten und detaillierten Fragen bzw. Aufgaben folgen danach [48]. Eine Schwierigkeit liegt

dabei darin, den Fragebogen zum einen optimal an die Problemstellung anzupassen und dabei zum anderen die Einschätzungen der Experten vergleichbar zu halten [50].

Im Allgemeinen können Expertenurteile entweder quantitativ oder qualitativ erhoben werden. Zur Erhebung numerischer Werte können z. B. Skalen verwendet werden [51]. Allerdings werden hier auch innovativere Ansätze verfolgt. In [52] wird z. B. eine Methodik zum expertengeführten Clustering von Daten vorgestellt, die Experteninteraktionen während des Klassifikationsprozesses über ein visuelles Interface ermöglicht. Oftmals fühlen sich Experten jedoch hinsichtlich der Abgabe eines quantitativen Urteils unbehaglich, weil sie befürchten, dass die Angabe eines konkreten Wertes den (falschen) Eindruck erwecken könne, sie seien sich mit dieser Einschätzung sicherer, als sie es wirklich sind. Auch bei hohem Anwendungsbezug ist es für Experten oftmals schwierig, hochdimensionale und komplexe Aufgabenstellungen zu bearbeiten [53]. Im Gegensatz dazu kann der reine Gebrauch von Sprache im Sinne einer qualitativen Einschätzung unpräzise und missverständlich ausfallen [49]. Vor diesem Hintergrund stellt eine Experteneinschätzung auf Basis vordefinierter ordinaler Klassen einen begründeten Kompromiss zwischen der Erhebung von rein quantitativem und ausschließlich qualitativem Expertenwissen dar. Um eine Auskunft zur Unsicherheit einer Expertenaussage zu erhalten, kann z. B. für jedes Objekt eine zusätzliche Einschätzung der Schwierigkeit von den Experten erhoben werden. Die Entscheidung eines Experten für eine ordinale Klasse wird aber auch durch subjektive Einflüsse beeinflusst:

1. Experten haben unterschiedliche Erfahrungslevel,
2. ihre Tagesform kann sich unterscheiden,
3. sie besitzen verschiedene Auffassungen von „Strenge“ und
4. Experten haben eine individuelle Neigung, keine extremen Beurteilungen vorzunehmen.

Auf einer abstrakten Ebene, können die Erfahrungslevel und Tagesformen von Experten als Wahrscheinlichkeit interpretiert werden, dass ihnen bei ihrer Einschätzung ein Fehler unterläuft. Vor diesem Hintergrund beeinflussen die Erfahrung und die Tagesform eines Experten folglich die Zuverlässigkeit seiner Aussagen. So kann z. B. der Anteil an Aufgaben in einem Fragebogen, der für einen Experten mit sehr großem Erfahrungsschatz als leicht eingestuft wird, im Vergleich zu einem anderen Experten mit einer vergleichsweise geringen Erfahrung wesentlich höher ausfallen. Ebenso kann dieses Verhältnis für einen einzelnen Experten auf Grund verschiedener Tagesformen variieren. Die Strenge mit der ein Experte seine Beurteilung durchführt, bildet einen zusätzlichen subjektiven Einfluss auf die Urteilsfindung. Sie äußert sich modellhaft darin, dass ein Experte mit gewisser Wahrscheinlichkeit eine zur wahren Klasse niedrigere (oder höhere) Klassifikationeinschätzung vornimmt. Die individuelle Neigung eines Ex-

perten, keine extremen Urteile abzugeben, äußert sich in einer Tendenz, ein Objekt dessen wahre Klasse sich nahe einer Randklasse befindet (z. B. der niedrigsten oder höchsten Klasse der Klassenordnung) eher in eine der mittleren Klassen einzuordnen. Diese Neigung (die bei einem erfahrenen Experten weniger ausgeprägt sein dürfte) führt also dazu, dass die während der Beurteilung vom Experten verwendete Bandbreite der Klassen gering ist.

Zusammengefasst ist aus den genannten Gründen die Entscheidung eines Experten für eine ordinale Klasse mehr oder weniger unsicherheitsbehaftet. Daher sollte eine finale Klassifikationsentscheidung in einer Anwendung nicht anhand eines einzelnen Expertenurteils getroffen werden. Die Urteile mehrerer Experten können dabei nicht übereinstimmen. Die Kombination von mehreren unsicherheitsbehafteten Expertenaussagen zu einer Gesamtaussage unter Anwendung von geeigneten Kombinationsregeln wird in Kapitel 5 adressiert.

4.1.2 *Klassifikationsschema*

Um eine fundierte Klassifikationseinschätzung von Experten zu erhalten, muss ein Ansatz zwei grundlegende Funktionen erfüllen. Zum einen müssen den Experten charakteristische Informationen zu den betrachteten Objekten zur Verfügung gestellt werden und zum anderen ist eine präzise Instruktion der Experten zur Durchführung der Befragung notwendig. Im Allgemeinen können diese beiden Funktionen anhand eines problemangepassten, in Eigenständigkeit zu bearbeitenden Fragebogens oder durch einen instruierten Interviewer eingenommen werden [48]. Um den Einfluss Dritter auf die Klassifikationsentscheidung eines Experten auf ein Mindestmaß zu reduzieren und um den oftmals bei Expertenbefragungen entstehenden zeitlichen und finanziellen Aufwand in einem überschaubaren Rahmen zu halten, wird die Wissenserhebung hier anhand eines eigenständig zu bearbeitenden Fragebogens vorgenommen. Dies bringt insbesondere Vorteile mit sich, wenn Urteile von einer Vielzahl an Experten zu erheben sind, die räumlich weit voneinander getrennt sind. Allerdings resultieren aus der Anforderung, dass Experten die Beantwortung der Fragen in Eigenständigkeit durchführen können, wiederum Herausforderungen bei der Informationsbereitstellung und Instruktion der Experten, in der nunmehr nur in geringem Maße Mehrdeutigkeiten und Interpretationsspielräume zugelassen werden können. Aus diesem Grund wird hier ein Klassifikationsschema eingeführt, das den Prozess der Klassifikation sowie die Bereitstellung von Informationen darstellt und als Grundlage für das Design des Fragebogens dient.

Zusammenfassend besteht die Aufgabe eines Experten während der Klassifikation darin, eine Beziehung zwischen zur Verfügung gestellten Informationen und vordefinierten ordinalen Klassen herzustellen. Das vom Experten intern zur Klassifikationsentscheidung durchgeführte Kalkül bleibt dabei im Verborgenen und kann abstrakt mit der

Messung einer latenten Variable verglichen werden [54]. In Analogie zur Messung einer physikalischen Größe, ist es daher wichtig, die Qualität dieser abstrakten Form der Messung zu bewerten. Als Konsequenz daraus müssen die den Experten zur Klassifikation bereitgestellten Informationen und Instruktionen neben der bereits erwähnten Eindeutigkeit zusätzlich daran orientiert sein, die Qualität der Expertenaussagen ex post bewerten zu können, ohne dabei den Freiheitsgrad der Experten bei ihrer Klassifikationsentscheidung zu stark einzuschränken.

Das als Basis zur Erhebung der Expertenaussagen dienende Klassifikationsschema ist in Abbildung 2 dargestellt. Die Grundlage für die Klassifikationsentscheidungen im Rahmen des Klassifikationsprozesses bildet die Bereitstellung von Basis- und Zusatzinformationen über die zu untersuchenden Objekte. Die Basisinformationen bestehen dabei aus jeweils einem Datenpunkt pro untersuchtem Objekt, in dem die Ausprägungen netzspezifischer Merkmale (z. B. Anzahl der Knoten in einem elektrischen Netz oder einem Straßennetz, siehe z. B. Abschnitt 3.1) enthalten sind. Auf Grund der Tatsache, dass eine adäquate Charakterisierung von Objekten in vielen Anwendungen nur anhand einer großen Menge von Merkmalen möglich ist, kann die Bereitstellung von Basisinformationen in einem komplexen Merkmalraum enden, der es Experten stark erschwert die Klassifikationsaufgabe konsistent zu lösen. Daher wird im Klassifikationsschema die Basisinformation anhand weniger relevanter Merkmale vorgenommen und zusätzlich durch eine Visualisierung des zu untersuchenden Objektes gestützt (z. B. ein Plan eines elektrischen Netzes oder eine Straßenkarte).

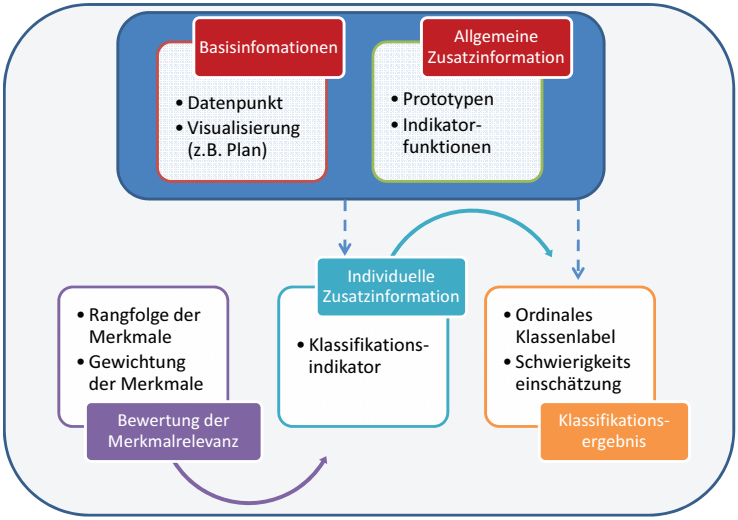


Abbildung 2: Übergeordnetes Klassifikationsschema

Bis zu diesem Punkt ist die Bereitstellung der Informationen naheliegend. Es wird allerdings hier der Ansatz verfolgt, eine Erhöhung der Reliabilität und Konsistenz der Expertenaussagen durch Bereitstellen von Zusatzinformationen herbeizuführen. Die Nutzung dieser Zusatzinformationen sollte den Experten allerdings ausdrücklich freigestellt bleiben, da die Intention der Zusatzinformationen nicht darin liegt, den Freiheitsgrad der Experten einzuschränken. Die erste Zusatzinformation besteht in der Zuordnung von je einem prototypischen Objekt zu jeder Klasse. Die Prototypen sollten aus der zu untersuchenden Menge an Objekten ausgewählt werden und den Experten analog zur Basisinformation in Form eines Datenpunktes im Raum, der durch die gewählten Merkmale aufgespannt wird, und dessen Visualisierung zu Beginn der Befragung zugänglich gemacht werden. Jeder Experte erhält damit die Möglichkeit seine Klassifikationsentscheidungen anhand eines Vergleichs mit den Prototypen zu stützen.

Um den Experten weitere Zusatzinformationen zu geben, wird der in einer Anwendung des Klassifikationsansatzes auftretende Wertebereich jedes Merkmals in Intervalle unterteilt, deren Anzahl der Klassenanzahl enstpricht. Jedes dieser Intervalle repräsentiert dann eine der vordefinierten Klassen. Für jeden Wert eines Merkmals lässt sich unter Berücksichtigung dieser Intervalle eine Klasse auswählen. Daraus resultiert formal für jedes Merkmal eine Indikatorfunktion, die einem Merkmalwert eine Klasse zuordnet. Die daraus entstehende Information kann als Klassifikationsempfehlung interpretiert werden, falls die Entscheidung nur in Abhängigkeit des zu Grunde liegenden Merkmals zu treffen wäre. Wird der Klassifikationsansatz für eine Klassifikationsaufgabe erstmalig angewendet, stehen keine Klasseninformationen bereit, anhand derer geschlussfolgert werden kann, dass die Klassen durch die gewählten Intervalle bestmöglich repräsentiert werden. In diesem Fall ist die Wahl der jeweiligen Intervallgrenzen folglich erschwert. Um bei fehlenden Klasseninformationen trotzdem eine repräsentative Wahl der Intervallgrenzen treffen zu können, ist eine Erhebung weiterer Hinweise im Vorfeld der Klassifikation sinnvoll. Hier bietet sich z. B. eine Vorabstimmung zur Wahl der Intervallgrenzen mit den an der Befragung beteiligten Experten und eine anschließende nochmalige Plausibilisierung der gewählten Werte durch die Experten an. Eine Methodik zur Durchführung dieser Vorabstimmung stellt z. B. die Delphi-Methode bereit [55]. In manchen Fällen können auch Informationen aus Ergebnissen anderer Klassifikationsansätze (z. B. simulationsbasierte Klassifikation) zur Wahl der Intervallgrenzen genutzt werden. Formal können die Indikatorfunktion wie folgt beschrieben werden:

$$f_i : M_i \rightarrow \{1, \ldots, C\}, \quad i = 1, \ldots, N_M, \tag{1}$$

mit

f_i:	i-te Indikatorfunktion,
M_i:	i-tes Merkmal,
N_M:	Anzahl der Merkmale,
C:	Anzahl der ordinalen Klassen.

Die Festlegung der Berechnungsvorschriften der Indikatorfunktionen sollte jeweils anwendungsspezifisch beim Fragebogendesign getroffen werden.

Nun wird eine Möglichkeit erarbeitet, den Experten individuelle Zusatzinformationen zur Klassifikation zur Verfügung zu stellen. Als Grundlage dazu wird von den Experten zu Beginn der Befragung ihre persönliche Präferenz hinsichtlich der Relevanz jedes Merkmals in ihrem internen Klassifikationskalkül abgefragt. Das Ziel liegt darin, eine Abschätzung des Merkmaleinflusses auf die Klassifikationseinschätzungen in Form einer Rangfolge zu erlangen. Zusätzlich zur Angabe der individuellen Rangfolge soll durch die Experten außerdem eine Gewichtung der Merkmale erfolgen. Durch die Gewichtung besteht die Möglichkeit, Merkmalen mit unterschiedlichem Rang eine gleiche Relevanz zuzuordnen. Auf dieser Grundlage kann den Experten während der Befragung für jedes zu klassifizierende Objekt ein persönlicher Klassifikationsindikator zur Verfügung gestellt werden. In dem Klassifikationsindikator soll zum einen die allgemeine Zusatzinformation und zum anderen die persönlichen Präferenzen des Experten bezüglich des Merkmaleinflusses berücksichtigt werden. Formal wird der Indikator daher auf Basis der Indikatorfunktionen und der gewählten Merkmalgewichtungen wie folgt gebildet:

$$ci_j \;=\; w_{j1} \cdot f_1 + w_{j2} \cdot f_2 + \cdots + w_{jN_M} \cdot f_{N_M} \tag{2}$$

mit

ci_j:	Klassifikationsindikator des Experten j,
f_i:	i-te Indikatorfunktion,
w_{ji}:	Gewichtung des i-ten Merkmals durch Experte j,
N_M:	Anzahl der Merkmale.

Die Bereitstellung des Klassifikationsindikators wird durch die Untersuchungsergebnisse von Mullin getragen. Ein Ansatz zur Erhebung von Expertenwissen sollte sich nach Mullin nämlich auch auf die Identifikation von Schlüsselmerkmalen durch die Experten zur anschließenden Verwendung in einfachen algebraischen Modellen konzentrieren [56]. Die Begründung erfolgt durch Mullin anhand von einigen weiteren Untersuchungen [57, 58], die belegen, dass dieser Ansatz zu guten Resultaten führt, auch wenn die Modellierung von der durch die Experten intern vorgenommenen Integration der Informationen abweichen kann.

Das Ausmaß, in dem einem Experten allgemeine und individuelle Zusatzinformationen zur Verfügung gestellt werden, sollte von der Schwierigkeit der Klassifikationsaufgabe und der Erfahrung des Experten hinsichtlich der Bearbeitung dieser Aufgabe abhängig gemacht werden. Dies bedeutet insbesondere, dass einem Experten nicht für jedes zu klassifizierende Objekt Zusatzinformationen bereitgestellt werden müssen. Vor diesem Hintergrund sind Zusatzinformationen zum einen gut geeignet, um einen unerfahrenen Experten durch intensive Bereitstellung zu Beginn der Befragung an eine Klassifikationsaufgabe heranzuführen, zum anderen kann aber auch ein gelegentliches Offenlegen

einen erfahrenen Experten im Laufe der Befragung in seinen Entscheidungen leiten. Darüber hinaus könnte sogar ganz auf die Bereitstellung von Zusatzinformationen verzichtet werden, wenn eine Klassifikation durch einen Experten regelmäßig durchgeführt wird.

Um eine Indikation zur Unsicherheit der Expertenaussagen zu erhalten, ist im Klassifikationsschema weiterhin für jedes Objekt eine Experteneinschätzung zur Schwierigkeit der Klassifikationsaufgabe vorgesehen. In Analogie zur Klassifikationseinschätzung wird die Beurteilung des Schwierigkeitsgrades auf einer ordinalen Skala (z. B. „einfach“, „mittel“ und „schwer“) erhoben. Zur Beurteilung der Messqualität wird eine gedoppelte Erhebung der Expertenurteile für eine bestimmte Anzahl von Objekten vorgesehen. Die Entscheidung für wie viele Objekte gedoppelte Expertenurteile zu erheben sind, ist anwendungsspezifisch und folglich während der Konzeption des Fragebogens zu treffen.

4.1.3 *Konzeption des Fragebogens*

Anhand des übergeordneten Klassifikationsschemas wird nun die Konzeption des Fragebogens vorgenommen. Bei der ordinalen Klassifikation von Objekten handelt es sich allgemein um eine erfahrungsbasierte und quantitative Aufgabenstellung. Dies bedeutet, dass zur adäquaten Erfüllung der Aufgaben zum einen gewisse anwendungsspezifische Erfahrungen der Experten vorausgesetzt werden und die Aufgabenstellung zum anderen zur Erhebung quantitativer Daten in Form von ordinalen Klassifikationseinschätzungen angelegt ist. Vor diesem Hintergrund sind im Wesentlichen folgende Arbeitsschritte im Fragebogen zu berücksichtigen:

- Zuordnung eines Rangs r_i zu jedem Merkmal M_i,
- Festlegung der Gewichtung w_i der Merkmale,
- Treffen der Klassifikationsentscheidungen und
- Angabe der Schwierigkeitseinschätzungen.

Die Erstellung der Merkmalrangfolge mit Gewichtung ist zu Beginn des Fragebogens vorzusehen, da diese während der Befragung zur Bereitstellung von Zusatzinformationen (z. B. Klassifikationsindikatoren) genutzt wird. Es handelt sich um eine übergeordnete Aufgabe, die auch nach Moser [48] an den Anfang des Fragebogens zu stellen ist. Im Gegensatz dazu sind die einzelnen Klassifikationsentscheidungen und Schwierigkeitseinschätzungen objektbezogen. Sie haben damit im Fragebogen der Angabe einer Merkmalrangfolge nachzufolgen [48]. Damit ist die grundlegende Struktur des Fragebogens festgelegt und konform zu den Grundsätzen aus [48].

M_i / c	1		2		...		C	
1	l_{11}	l_{12}	l_{12}	l_{13}	...	...	l_{1C}	l_{1C+1}
2	l_{21}	l_{22}	l_{22}	l_{23}	...	...	l_{2C}	l_{2C+1}
...	...	...	...	...	...	...	...	...
N_M	l_{N_M1}	l_{N_M2}	l_{N_M2}	l_{N_M3}	...	...	l_{N_MC}	l_{N_MC+1}

Tabelle 3: Berechnungsvorschriften der Indikatorfunktionen in tabellarischer Form.

Die Herausforderung in der Konzeption besteht nun vielmehr darin, die Informationen für die Experten übersichtlich im Fragebogen anzuordnen[1]. Dies ist insbesondere zur Einhaltung einer hinreichend hohen Messqualität während der Befragung essentiell. Um die Experten vor der Befragung über die Berechnungsvorschriften der Indikatorfunktionen zu informieren, werden diese wie in Tabelle 3 anhand der Grenzen l_{ic} ($i = 1, \ldots, N_M$, $c = 1, \ldots, C$) gezeigt, vor der Erhebung der Merkmalrangfolge und -gewichtung auf dem Fragebogen angeordnet.

NR.	NAME DES MERKMALS	RANG $r(M_i)$
1	*Merkmal* M_1	
2	*Merkmal* M_2	
...	...	
i	*Merkmal* M_i	
...	...	
N	*Merkmal* F_{N_M}	

Tabelle 4: Tabelle zur Angabe der Merkmalrangfolge.

$r(M_i)$	NAME DES MERKMALS	GEWICHT w_i DES MERKMALS
1	*Merkmal mit Rang* 1	
2	*Merkmal mit Rang* 2	
...	...	
i	*Merkmal mit Rang* i	
...	...	
N_M	*Merkmal mit Rang* N_M	

Tabelle 5: Gewichtungstabelle mit entsprechend der Rangfolge angeordneten Merkmalen.

Die Erhebung der Merkmalrelevanz erfolgt anschließend in einer weiteren Tabelle (Tabelle 4). Um den Experten die Einschätzung der Gewichtung zu erleichtern, werden die

1 Einige Expertenaussagen werden im weiteren Verlauf der Befragung genutzt und sind daher direkt im Anschluss an die Experteneinschätzung weiterzuverarbeiten (z. B. Berechnung der Klassifikationsindikatoren anhand der Einschätzungen zur Merkmalrelevanz). Die Durchführung der Befragung sollte daher computergestützt vorgenommen werden.

Merkmale entsprechend der Rangfolge in einer weiteren Tabelle neu angeordnet. Zur Gewichtung wird ein aggregiertes Gewicht W vorgegeben, das durch die Experten auf die Merkmale M_i zu verteilen ist (Tabelle 5).

OBJEKT	M_1, $r(M_1)$		...		M_N, $r(M_N)$		ci_j
1	$M_1(1)$	$f_1(M_1(1))$	...	...	$M_N(1)$	$f_N(M_N(1))$	$ci_j(1)$
2	$M_1(2)$	$f_1(M_1(2))$	...	...	$M_N(2)$	$f_N(M_N(2))$	$ci_j(2)$
...	...	...	...	...	...	...	...
n	$M_1(n)$	$f_1(M_1(n))$	...	...	$M_N(n)$	$f_N(M_N(n))$	$ci_j(n)$

Tabelle 6: Tabellarische Anordnung der Informationen im Fragebogen.

Auf die Bewertung der Merkmale folgen – wie oben erwähnt – die objektspezifischen Klassifikationsaufgaben. Um den Experten bei jeder Entscheidung eine Berücksichtigung aller Informationen zu ermöglichen, werden die gewählte Merkmalrangfolge, die Werte der Indikatorfunktionen sowie die Klassifikationsindikatoren zusätzlich zu den Datenpunkten der Objekte auf dem Fragebogen dargestellt. Die Visualisierung der Objekte sowie die Bereitstellung der Prototypen erfolgt separat, um den Experten Vergleiche zwischen verschiedenen Objekten zu erleichtern. Eine Möglichkeit zur übersichtlichen Anordnung aller Informationen im Fragebogen ist in Tabelle 6 gezeigt. Die Werte der Indikatorfunktionen werden dort direkt neben dem jeweiligen Wert des Merkmals angeordnet. Zusätzlich wird der Klassifikationsindikator neben den Datenpunkten des Objektes angezeigt. Die Ergebnisse der einzelnen Indikatorfunktionen können z. B. vereinfachend durch C verschiedene Symbole (z. B. Pfeile) oder Einfärbungen der Merkmalswerte visualisiert werden. Diese Art der Darstellung reduziert die Anzahl numerischer Werte im Fragebogen. Die Häufigkeit und Regelmäßigkeit, mit der Zusatzinformationen im Laufe des Fragebogens bereitgestellt werden, ist an dieser Stelle in Abhängigkeit der Expertenerfahrungen bezüglich der Klassifikationsaufgabe festzulegen.

OBJEKT/KLASSE	1	2	...	C
1				
2				
...				
n				

Tabelle 7: Tabelle zur Dokumentation der Klassifikationsergebnisse.

Zur Dokumentation der Klassifikationsentscheidungen und der Schwierigkeitseinschätzungen der Experten wird neben Tabelle 6 jeweils eine ordinale Skala vorgegeben. Bei dieser Anordnung können die Expertenaussagen in Form von Kreuzen in den betreffenden Feldern der Skalen festgehalten werden (Tabellen 7 und 8).

OBJEKT/SCHWIERIGKEIT	1	2	...	D
1				
2				
...				
n				

Tabelle 8: Tabelle zur Dokumentation der Schwierigkeitseinschätzungen.

4.1.4 *Bewertung der Messqualität*

Der Beurteilung der Messqualität kommt bei der Durchführung einer Expertenbefragung eine besondere Rolle zu. Wie bereits zu Beginn dieses Kapitels am Beispiel von NS-Netzen erwähnt wurde, steht häufig keine Vergleichsgröße im Sinne einer „ground truth“ zur Verfügung. Weiterhin wurde angeführt, dass eine Klassifikationseinschätzung eines Experten abstrakt als Messung einer latenten Variable interpretiert werden kann. Verhaltenswissenschaftler werden in ihrer Disziplin häufig mit dieser abstrakten Form des Messens konfrontiert. Aus diesem Grund wurde der Begriff des Messens von Stevens wie folgt erweitert: „Messen ist die Zuweisung von Zahlen zu Objekten oder Ereignissen anhand von Regeln“ [59]. Analog zur Messung einer physikalischen Größe hat eine Messung den Qualitätskriterien Objektivität, Reliabilität und Validität zu genügen [54]. In diesem Zusammenhang werden den genannten Qualitätskriterien folgende Bedeutungen zugeordnet [54]:

OBJEKTIVITÄT: Jede Messung ist so weit wie möglich unabhängig von möglichen Einflüssen durch einen Interviewer. Dies bedeutet insbesondere, dass zwei unterschiedliche Interviewer auch zu gleichen Messergebnissen kommen sollten.

RELIABILITÄT: Jede Messung ist zu einem späteren Zeitpunkt unter sonst gleichen Messbedingungen reproduzierbar. Zur Einhaltung der Reliabilität ist die Objektivität der Messung Voraussetzung, d. h. ein notwendiges aber nicht hinreichendes Kriterium.

VALIDITÄT: Die Validität entspricht dem Grad zu dem eine Messung die zu messende Eigenschaft eines Objektes reflektiert.

Bezieht man diese Qualitätskriterien auf den Klassifikationsansatz, so wird die Einhaltung der Objektivität durch die Entwicklung und Anwendung eines problemspezifischen Fragebogens gewährleistet, der durch die Experten ohne Anwesenheit eines Interviewers bearbeitet werden kann. Die Wahrscheinlichkeit einer Beeinflussung der Klassifikationsentscheidungen durch einen Interviewer ist somit nicht vorhanden.

Die Reliabilität der Klassifikationsergebnisse wird auf Basis der erhobenen Expertenaussagen ex post analysiert. Dazu wird die Intraraterkonkordanz jedes einzelnen Experten bewertet. Die Intraraterkonkordanz wird in der Statistik zur Messung der Übereinstimmung von wiederholten kategorischen oder ordinalen Beurteilungen durch einen Experten genutzt. In diesem Zusammenhang kann auf ähnliche Weise auch die Übereinstimmung zwischen den Beurteilungen verschiedener Experten untersucht werden, welche als Interraterkonkordanz bezeichnet wird.

Insbesondere die Beurteilung der Validität der Messergebnisse gestaltet sich schwierig, da häufig keine „wahre“ Vergleichsgröße im Sinne einer „ground truth“ verfügbar ist. Um dennoch eine Indikation zur Validität der Messergebnisse zu erhalten, kann die Interraterkonkordanz zwischen einzelnen Experten analysiert werden. Dieser Ansatz basiert auf der Annahme, dass sich die Validität einer Messung mit zunehmender Anzahl von Experten erhöht, deren Urteile dann zur Verifikation untereinander genutzt werden können. Zusätzlich ist es sinnvoll, zur Beurteilung der Validität, falls möglich, die Interraterkonkordanz zwischen den Expertenaussagen und den Ergebnissen einer weiteren unabhängigen Bewertung vorzunehmen (z. B. Simulationsergebnisse).

		Urteil 1				
		1	2	3	4	5
Urteil 2	1	1	0.75	0.5	0.25	0
	2	0.75	1	0.75	0.5	0.25
	3	0.5	0.75	1	0.75	0.5
	4	0.25	0.5	0.75	1	0.75
	5	0	0.25	0.5	0.75	1

Tabelle 9: Beispiel für eine symmetrische lineare Gewichtung.

Eine Standardmethode zur statistischen Beurteilung der Intra- und Interraterkonkordanz von ordinalen Urteilen [60] und somit zur Bewertung der Reliabilität und Validität der Klassifikationseinschätzungen im Sinne der hier getroffenen Annahmen, ist Cohen's gewichtetes Kappa κ_w. Die κ_w-Statistik erlaubt die Nutzung von Gewichten, um den Abstand zwischen verschiedenen Klassen zu berücksichtigen, auf denen eine Ordnung definiert ist. Auf diese Weise kann das Ausmaß der Unstimmigkeit zwischen den einzelnen erhobenen Urteilen in der Bestimmung der Konkordanz berücksichtigt werden [61]. Um κ_w zu berechnen, werden die n Klassifikationsergebnisse für die C definierten ordinalen Klassen in einer $C \times C$-Konfusionsmatrix gegenübergestellt. Ein Eintrag o_{rs} ($r = 1, \ldots, C$ und $s = 1, \ldots, C$) repräsentiert dann die absolute Anzahl der Objekte, für die das erste Urteil zur Klasse s und das zweite Urteil zu Klasse r geführt hat. Damit beinhaltet die Diagonale der Konfusionsmatrix die absolute Anzahl der Objekte, für die beide Urteile exakt übereinstimmen. In einem weiteren Schritt wird für jede Zelle rs der Konfusionsmatrix ein Gewicht w_{rs} ($w_{rs} \in [0, 1]$) festgelegt, das den Grad der Übereinstimmung zwischen den Urteilen entsprechend des Abstandes der Klassen widergibt.

In dieser Arbeit wird eine symmetrische lineare Gewichtung verwendet, welche neben der Anwendung von symmetrischen quadratischen Gewichten die im Zusammenhang mit κ_w am häufigsten genutzte Art der Gewichtung darstellt [62]. Ein Beispiel für die lineare symmetrische Gewichtung zeigt Tabelle 9 für den Fall $C = 5$.

Im nächsten Schritt kann zur Berechnung von κ_w die gewichtete relative Häufigkeit A_0 unter Verwendung von Gleichung (3) bestimmt werden, die ein Indikator für die insgesamt beobachtete Übereinstimmung verkörpert. Der korrespondierende Indikator A_ϵ für die relative Häufigkeit der Übereinstimmung, die unter Unabhängigkeit der Urteile zu erwarten ist, kann mit Gleichung (4) berechnet werden [54]:

$$A_0 = \sum_{r=1}^{C} \sum_{s=1}^{C} w_{rs} \cdot o_{rs}/n, \tag{3}$$

$$A_\epsilon = \sum_{r=1}^{C} \sum_{s=1}^{C} w_{rs} \cdot \epsilon_{rs}/n, \tag{4}$$

mit

$$\begin{aligned} \epsilon_{rs} &= o_{r.} \cdot o_{.s}/n, \\ o_{r.} &= \sum_{s=1}^{C} o_{rs}, \\ o_{.s} &= \sum_{r=1}^{C} o_{rs}. \end{aligned}$$

Wenn A_0 und A_e bekannt sind, ergibt sich κ_w zu:

$$\kappa_w = \frac{A_0 - A_\epsilon}{1 - A_\epsilon}. \tag{5}$$

Intuitiv kann das gewichtete Kappa aus Gleichung (5) als zufallsbereinigter Übereinstimmungsindex interpretiert werden, da im Zähler die Differenz zwischen der beobachteten Häufigkeit der Übereinstimmung A_0 und der Übereinstimmung A_ϵ, die per Zufall zu erwarten ist, gebildet wird [54]. Aus Gleichung (5) lässt sich weiterhin ableiten, dass κ_w bei perfekter Übereinstimmung einen Wert von eins annimmt. Wenn κ_w sich zu null ergibt, wurde eine Übereinstimmung vom Ausmaß der per Zufall zu erwartenden Übereinstimmung beobachtet. Ein negativer Wert indiziert demnach also eine Konkordanz, die kleiner als eine zufällig zustande gekommene ist [63]. Letztendlich ist in einer Anwendung aber die Interpretation des berechneten Wertes von κ_w grundlegend freigestellt. Eine feste Skala zur Interpretation des Wertes für κ_w existiert in der Literatur

nicht. Dies ist maßgeblich durch die Abhängigkeit der Werte von κ_w von der gewählten Gewichtung begründet. Zur Analyse der statistischen Signifikanz einer beobachteten Übereinstimmung, bietet sich der Einsatz eines Hypothesentests an [54].

Um die Reliabilität des entwickelten expertenbasierten Klassifikationsansatzes zu bewerten, können die Objekte genutzt werden, die gemäß des Klassifikationsschemas von den Experten doppelt klassifiziert wurden. Die gedoppelten Klassifikationsergebnisse ermöglichen nämlich in Zusammenhang mit der κ_w-Statistik die Erhebung einer Indikation zur Intraraterkonkordanz. Je höher die durch κ_w errechnete Übereinstimmung zwischen den gedoppelten Klassifikationsergebnissen ist, desto höher ist auch die Reliabilität des Klassifikationsansatzes zu bewerten. Liegt eine insgesamt sehr niedrige Reliabilität des Klassifikationsansatzes vor, könnte zum einen die spezifizierte Klassifikationsaufgabe zu schwer sein, oder zum anderen der Klassifikationsansatz generell keine reliablen Resultate zurückliefern und somit kein verlässlicher Ansatz zur Messung der gewünschten Eigenschaft sein. Die Anzahl der Übereinstimmungen für die doppelt zu klassifizierenden Objekte hängt aber auch davon ab, inwieweit diese von den Experten als doppelt erkannt werden. Aus diesem Grund sollten zum einen die Visualisierungen (z. B. Pläne) der gedoppelten Objekte verfremdet werden (z. B. Rotation der Pläne) und zum anderen die doppelten Objekte in einer hinreichend großen Stichprobe entsprechend gut verteilt sein. Dies bedeutet z. B., dass ein Objekt nicht zweimal kurz hintereinander in der Befragung mit einer gleichen Visualisierung auftreten sollte.

Die Intraraterkonkordanz, die bei den einzelnen Experten ermittelt wird, kann zusätzlich Auskunft darüber geben, wie erfahren diese hinsichtlich der Bearbeitung der gestellten Aufgaben sind oder, ob sie die Bearbeitung grundsätzlich mit der notwendigen Ernsthaftigkeit oder Konzentration durchgeführt haben. Diese Information ist von hoher Bedeutung, da das Bewertungsverhalten einzelner Experten die gesamte Reliabilität der Klassifikationsresultate beeinflusst. Dabei nimmt der Einfluss eines einzelnen Experten mit steigender Anzahl von Experten ab. Analog zur Bewertung der Reliabilität wird κ_w ebenfalls dazu genutzt, um eine Indikation zur Validität des Klassifikationsansatzes zu erhalten. Wie oben bereits argumentiert, wird dazu die Interraterkonkordanz zwischen einer Anzahl von Experten paarweise ausgewertet. Hierzu können im Gegensatz zur Berechnung der Intraraterkonkordanz alle in der Untersuchung beinhalteten Objekte genutzt werden.

4.1.5 *Experimentelle Untersuchung und Evaluation*

In diesem Abschnitt wird der Klassifikationsansatz zur Bewertung der Aufnahmekapazität von NS-Netzen angewandt, um empirisch nachzuweisen, dass dieser aussagekräftige Ergebnisse zur Bewertung von NS-Netzen liefert. Nach der bis zu diesem Punkt bewusst generisch und abstrakt gehaltenen Erläuterung des expertenbasierten Klassifikations-

ansatzes wird zunächst die empirische Ausgestaltung der Befragung kurz erläutert. Zur Veranschaulichung der Klassifikationsergebnisse werden die Resultate für die beiden in Abschnitt 3.1.2 vorgestellten Musternetze diskutiert. Im Anschluss daran wird auf Basis der Befragungsergebnisse die Nutzung der Zusatzinformationen durch die Experten ex post analysiert und deren Mehrwert für die Klassifikation bewertet. Abschließend erfolgt auf Basis der Überlegungen aus Abschnitt 4.1.4 eine empirische Evaluation der Messqualität anhand der Befragungsergebnisse.

Empirische Ausgestaltung der Befragung:
Die Befragung wurde mit fünf Netzplanungsexperten eines regionalen VNB durchgeführt, die mit in der Praxis auftretenden netzplanerischen Aufgabenstellungen vertraut sind (z. B. Anschlussbeurteilung für DEA, Dimensionierung von Betriebsmitteln). Die Anzahl von fünf Experten stellt einen Kompromiss zwischen der Erreichung einer möglichst hohen Messqualität und der Reduktion der Kosten für die Befragung dar. Jeder Experte bewertete im Rahmen der Befragung die Struktur von 300 realen, ländlichen NS-Netzen anhand der folgenden $C = 5$ ordinalen Klassen:

1. sehr schwach,
2. schwach,
3. durchschnittlich,
4. stark und
5. sehr stark.

Die gewählte Anzahl der Klassen dient als Kompromiss zwischen einer feinen Einteilung der betrachteten NS-Netze und einem moderaten Schwierigkeitsgrad der Klassifikationsaufgabe. Die neben den Klassifikationsergebnissen zu erhebenden Schwierigkeitseinschätzungen der Experten wurden anhand von drei Schwierigkeitsgraden („leicht“, „mittel“ und „schwer“) ermittelt.

Zur Charakteristierung der NS-Netze auf dem Fragebogen wurden die netzspezifischen Merkmale verwendet (Abschnitt 3.1.1). Die Basisinformationen bestehen folglich aus insgesamt zehn Merkmalen. Der unmittelbare Anwendungsbezug der verwendeten Merkmale stellt sicher, dass diese den Experten zur Bearbeitung des Fragebogens bekannt waren. Auf diese Weise wurde den Experten ermöglicht, ihre Erfahrungen aus der Verteilnetzplanung zur Klassifikation der Netze einzusetzen. Dies ist bei Verwendung von abstrakten (z. B. graphspezifischen) Merkmalen nicht ohne Weiteres möglich. Um die Bewertung der Stärke der Netzstrukturen durch die Experten unabhängig von der räumlichen Ausdehnung der Netze zu gestalten, wurden die Merkmale „Anzahl der Hausanschlüsse“, „Summe der Transformatornennleistungen“, „Anzahl der KVS“, „Summe

der Leitungslängen“ und „Summe der vermaschten Leitungslängen“ in der Befragung als durchschnittliche Werte je Stationsbereich angegeben. Die bereits installierte DEA-Leistung in den Netzen wurde nicht als Merkmal verwendet, da sie nicht in der Menge der zehn netzspezifischen Merkmale enthalten ist, und damit bei der Klassifikation vernachlässigt. Diese Vernachlässigung ist sinnvoll, um eine gleiche Ausgangssituation für jedes Netz und folglich vergleichbare Klassifikationsergebnisse zu erhalten. Zusätzlich zu diesen Basisdaten wurde den Experten in einem Katalog für jedes Netz ein Plan zur Verfügung gestellt, der zur Visualisierung der Objekte während der Befragung diente (vgl. z. B. Abbildung 1).

	1 (sehr schwach)		2 (schwach)		3 (durchschn.)		4 (stark)		5 (sehr stark)	
1	178	95	95	77	77	59	59	37	37	0
2	1	2	2	3	3	4	4	5	5	19
3	100	276	276	400	400	441	441	516	516	630
4	0	2,1	2,1	3,1	3,1	4,1	4,1	5,1	5,1	12
5	0	1827	1827	2585	2585	3063	3063	3720	3720	8607
6	0	633	633	1131	1131	1571	1571	2125	2125	6425
7	0	0,308	0,308	0,440	0,440	0,537	0,537	0,622	0,622	1
8	0	0,686	0,686	0,773	0,773	0,835	0,835	0,900	0,900	1
9	2613	975	975	739	739	592	592	454	454	0
10	2156	639	639	533	533	432	432	361	361	0

Tabelle 10: Berechnungsvorschriften der Indikatorfunktionen für die netzspezifischen Merkmale M_i aus Tabelle 1.

Zur Festlegung der Berechnungsvorschriften für die Indikatorfunktionen (entsprechend der allgemeinen Darstellung in Tabelle 3) der zehn verwendeten Merkmale wurde vor der Befragung in Einzelgesprächen eine Vorabstimmung mit den an der Befragung beteiligten Experten durchgeführt. Die auf Basis der Gespräche gewählte Festlegung wurde anschließend nochmal jedem der beteiligten Experten zur Plausibilisierung vorgelegt. Als Ergebnis des beschriebenen Abstimmungsprozesses mit den Experten wurden die Berechnungsvorschriften der Indikatorfunktionen so gewählt, dass diese jeweils 20% der Netze einer Klasse zuordnen. Diese Zuordnung wurde hauptsächlich gewählt, um die Bandbreite der Klassen optimal durch die Indikatorfunktionen zu repräsentieren und deren Nutzung nicht bereits durch Festlegungen im Vorfeld der Befragung einzuschränken. Die auf diese Weise für die Befragung festgelegten Berechnungsvorschriften sind in Tabelle 10 gezeigt. Die Nummerierung der Merkmale M_i erfolgt in der Tabelle analog zu Tabelle 1. Am Beispiel des Merkmals M_4, das die durchschnittliche Anzahl der KVS in einem Netz angibt, lässt sich die Indikatorfunktion wie folgt erläutern: nimmt M_4 für ein Netz einen Wert an, der größer oder gleich 2,1 und geringer als 3,1 ist, so gibt die Indikatorfunktion von M_4 im Ergebnis Klasse 2 aus. Als weiteres Beispiel gibt die Indikatorfunktion für das Merkmal M_9, dem maximalen Stationsradius, im Ergebnis

Klasse 5 aus, wenn der Wert des Merkmals kleiner oder gleich 360,5m und größer als 0m ist. Die Ergebnisse der Indikatorfunktionen wurden den Experten bei der Hälfte der Netze zur Verfügung gestellt. Die Anordnung im Fragebogen erfolgte dabei in Blöcken zu je fünf Netzen. Auf einen Block von Netzen mit Ergebnissen der Indikatorfunktionen folgte jeweils ein Block ohne ausgewiesene Ergebnisse der Indikatorfunktionen. Diese Anordnung wurde gewählt, da die Experten bisher keine Erfahrungen mit dieser oder ähnlichen Befragungen besaßen und um diese daher über die gesamte Befragung in der Bearbeitung der Klassifikationsaufgaben zu leiten. Zur Illustration der Anordnung von Informationen im Fragebogen gemäß Tabelle 6 ist die Ausgestaltung des Fragebogens für den hier diskutierten Anwendungsfall beispielhaft für die ersten 15 NS-Netze der Befragung von Experte 5 in Abbildung 3 gezeigt. Die Bereitstellung der Ergebnisse der Indikatorfunktionen in Blöcken zu je fünf Netzen ist darin deutlich erkennbar. Zur Visualisierung der Ergebnisse wurden fünf farbige Pfeilsymbole genutzt.

Die Prototypen für die Netze wurden im Vorfeld der Befragung auf Basis der mit den Experten abgestimmten Berechnungsvorschriften der Indikatorfunktionen ausgewählt. Im Detail kommt bei dieser Art der Auswahl ein Netz als Prototyp für eine Klasse in Frage, wenn die Ergebnisse der Indikatorfunktionen überwiegend den Wert dieser Klasse ausgeben und ihr Durchschnitt den Wert dieser Klasse näherungsweise widerspiegelt. Ausgewählt wurde aus dieser Menge an potenziellen Prototypen jeweils das Netz, das die jeweilige Klasse am besten repräsentiert (d. h. der Durchschnitt lag möglichst nahe am Wert der Klasse und möglichst viele Indikatorfunktionen haben den Wert der Klasse ausgegeben). Der für die Klasse 1 genutzte Prototyp weist beispielsweise folgende Wertepaare $(M_i,f_i(M_i))$ für die zehn netzspezifischen Merkmale auf: „durschnittliche Anzahl der Hausanschlüsse“ ($M_1 = 25$, $f_1(M_1) = 5$), „Anzahl der Transformatorstationen“ ($M_2 = 1$, $f_2(M_2) = 1$), „durchschnittliche Summe der Transformatornennleistungen“ ($M_3 = 250$, $f_3(M_3) = 1$), „durchschnittliche Anzahl der Kabelverteilerschränke“ ($M_4 = 1$, $f_4(M_4) = 1$), „durchschnittliche Summe der Leitungslängen“ ($M_5 = 1624$, $f_5(M_5) = 1$), „durchschnittliche Summe der vermaschten Leitungslängen“ ($M_6 = 0$, $f_6(M_6) = 1$), „Vermaschungsgrad“ ($M_7 = 0$, $f_7(M_7) = 0$), „Anteil der Leitungen des Typs NAYY 4x150mm^2“ ($M_8 = 0{,}82$, $f_8(M_8) = 3$), „maximaler Stationsradius“ ($M_9 = 729$, $f_9(M_9) = 3$) und „durchschnittlicher Stationsradius“ ($M_{10} = 729$, $f_{10}(M_{10}) = 3$). Der Durchschnitt der Indikatorfunktionen liegt damit unter Berücksichtigung aller Merkmale bei 1,8. Unter Vernachlässigung des Ausreißers der Indikatorfunktion für die Anzahl der Hausanschlüsse reduziert sich der Durchschnitt auf 1,4. Die auf diese Weise ausgewählten fünf Prototypen wurden den Experten in der Befragung in einem separaten Katalog zur Verfügung gestellt, um einen schnellen Vergleich mit einem zu beurteilenden Netz zu ermöglichen.

Zur Gewichtung der Merkmale wurde das zu verteilende aggregierte Gewicht auf einen Wert von $W = 20$ festgelegt. Die Wahl des aggregierten Gewichts ist letztendlich willkürlich. Allerdings erleichtert ein hoher Wert für W den Experten eine differenzierte Gewichtung der Merkmale. Um die Verteilung des aggregierten Gewichts auf die ein-

Merkmalrangfolge	8	10	3	9	5	6	4	7	2	1	Klass.-indikator entsprechend gewählter Rangfolge und Gewichtung
Nr.	HA [Anz./ST]	ST [Anz.]	P Trafo [kVA/ST]	KVS [Anz./ST]	Leitungs-länge [m/ST]	Vermaschte Leitungs-länge [m/ST]	Vermaschungs-grad [%]	Anteil NAYY 150 [%]	Max. Entfernung von ST [m]	Durchschnittliche max. Entfernung von ST [m]	
1	66	2	325	5.0	3550	1525	42.9%	71.9%	578	505	3.0
2	8	1	160	0.0	928	0	0.0%	76.9%	256	256	3.0
3	73	2	273	4.0	2998	1155	38.5%	96.0%	1082	784	1.5
4	69	3	400	2.7	2716	1529	56.3%	89.2%	487	321	4.0
5	95	1	400	3.0	3759	1594	42.4%	55.4%	529	529	3.4
6	14	1	250	0.0	915	0	0.0%	80.3%	612	612	
7	58	3	360	3.3	2356	1129	47.9%	81.4%	693	429	
8	107	2	515	4.5	3987	2412	60.5%	89.8%	428	407	
9	83	2	365	5.0	2299	928	40.4%	100.0%	650	338	
10	79	3	437	4.0	3014	1814	60.2%	70.4%	714	558	
11	50	4	363	3.3	3155	1878	59.5%	96.3%	570	406	3.6
12	112	7	531	4.6	3527	1829	51.8%	79.2%	668	446	3.6
13	79	3	427	4.7	3488	1927	55.2%	73.4%	1118	576	2.6
14	14	4	153	1.4	1030	308	29.9%	74.3%	632	355	2.6
15	45	6	355	2.7	2099	975	46.5%	83.7%	772	547	2.2

Abbildung 3: Beispiel zur Ausgestaltung des Fragebogens von Experte 5 für die ersten 15 Netze.

zelnen Merkmale weiter zu vereinfachen wurde diese den Experten nur für die ersten fünf Merkmale ihrer persönlichen Rangfolge abverlangt. Das Ziel dieser Einschränkung ist eine Fokussierung der Experten bei der Gewichtung auf die wichtigsten Merkmale zu erreichen. Entsprechend sind bei der Bildung der persönlichen Klassifikationsindikatoren gemäß Gleichung (2) für jeden Experten nur diese fünf Merkmale berücksichtigt worden. Ein weiterer Vorteil der geringen Anzahl gewichteter Merkmale liegt damit in für die Experten leichter nachvollziehbaren Ergebnissen der Klassifikationsindikatoren. Analog zu den Ergebnissen der Indikatorfunktionen wurde der Klassifikationsindikator für die Hälfte der zu klassifizierenden Netze in Blöcken zu je fünf Netzen im Fragebogen bereitgestellt. Zur Illustration ist die vom Experten 5 angegebene Rangfolge der Merkmale in der ersten Zeile des Fragebogenausschnitts in Abbildung 3 zu erkennen. Das aggregierte Gewicht W wurde von Experte 5 in der Befragung wie folgt auf die ersten fünf Merkmale der Rangfolge verteilt: „durchschnittlicher Stationsradius“ $w_{10} = 6$, „maximaler Stationsradius“ $w_9 = 4$, „durchschnittliche Summe der Transformatornennleistungen“ $w_3 = 4$, „Vermaschungsgrad“ $w_7 = 3$ und „durchschnittliche Summe der Leitungslängen“ $w_5 = 3$. Anhand der Rangfolge und der Gewichtung lassen sich die Klassifikationsindikatoren von Experte 5 mittels Gleichung (2) nachvollziehen. Für das erste Netz des Fragebogens berechnet sich der im Fragebogenausschnitt angegebene Klassifikationsindikator wie folgt: $ci_5(1) = 1/20 \cdot (6 \cdot 3 + 4 \cdot 4 + 4 \cdot 2 + 3 \cdot 2 + 3 \cdot 4) = 3{,}0$.

Die Anzahl der doppelt zu klassifizierenden NS-Netze wurde auf zehn festgelegt und entspricht damit einem Anteil von ca. 3% an der Gesamtheit der zu klassifizierenden NS-Netze. Diese Wahl stellt einen Kompromiss zwischen der Generierung einer ausreichend hohen Zahl von gedoppelten Klassifikationseinschätzungen zur Beurteilung der Reliabilität der Messung (Abschnitt 4.1.4) und der Vermeidung des Falles dar, dass einem Experten die doppelte Erhebung der Klassifikationseinschätzung während der Befragung auffällt. Die doppelten NS-Netze wurden gleichmäßig über die Befragung verteilt, um ein direktes Auftreten der Netze kurz nacheinander zu vermeiden. Zusätzlich wurden beim zweiten Auftreten der gedoppelten Netze deren Pläne jeweils um 180 Grad gedreht, um die Dopplung für die Experten schwerer erkennbar werden zu lassen und den Experten eine „neue“ Klassifikationsaufgabe zu suggerieren.

Nutzung der Zusatzinformationen durch die Experten:
Anhand der Klassifikationsergebnisse wird nun die Nutzung der während der Befragung bereitgestellten Zusatzinformationen untersucht. Dazu werden im Folgenden einige statistische Zusammenhänge untersucht. Kann ein statistischer Zusammenhang entdeckt werden, so ist formal nicht direkt auch ein kausaler Zusammenhang, wie z. B. „Experte X hat Zusatzinformationen Y häufig genutzt“, ableitbar. Allerdings lassen sich auf Basis statistischer Zusammenhänge Vermutungen hinsichtlich kausaler Zusammenhänge ableiten. Diese Vermutungen können dann z. B. in Gesprächen mit den Experten diskutiert werden.

Eine Analyse, inwiefern die Prototypen durch die Experten in ihre Klassifikationsentscheidungen einbezogen wurden, kann anhand der Klassifikationsergebnisse der verwendeten Prototypen untersucht werden. Das linke Netzdiagramm in Abbildung 4 zeigt für jeden der fünf Experten die Klassifikationsergebnisse der Prototypen. Die Einschätzungen der Experten sind darin den Klassen gegenübergestellt, die durch die Prototypen repräsentiert werden. Zum Vergleich ist zusätzlich im rechten Netzdiagramm eine perfekte Nutzung der Prototypen dargestellt. Die Netzdiagramme zeigen, dass eine perfekte Übereinstimmung nur bei Experte 1 auftritt. Aus diesem statistischen Zusammenhang lässt sich eine intensive Nutzung der Prototypen durch Experte 1 im Laufe der Befragung vermuten. Mit nur einer einzelnen Abweichung, die für den Prototypen der Klasse 2 auftritt, liegt bei Experte 5 ebenfalls eine hohe Übereinstimmung zwischen den Klassifikationseinschätzungen und den Klassen vor, die durch die Prototypen repräsentiert werden. Diese Übereinstimmung lässt für Experte 5 also ebenfalls eine deutliche Tendenz zur Nutzung der Prototypen erwarten. Bei Experte 3 treten in der Stichprobe insgesamt drei Übereinstimmungen auf. In diesem Fall lässt sich damit nur eine schwache Tendenz zur Nutzung der Prototypen vermuten. Die Klassifikationseinschätzungen der Prototypen stimmen bei den Experten 2 und 4 lediglich in zwei Fällen mit den repräsentierten Klassen überein. Dies lässt erwarten, dass diese Experten während der Befragung tendenziell keinen Abgleich der Netze mit den prototypischen Netzen vorgenommen haben. Insbesondere fällt auf, dass die Experten 2 und 4 im Vergleich zu den durch die Prototypen repräsentierten Klassen zum Teil auch die Ordnung der Prototypen getauscht haben. So wurde z. B. der Prototyp der Klasse 2 durch Experte 2 in Klasse 3 eingeordnet. Diese Klassifikationsergebnisse könnten ein Indiz für eine geringere Expertise der Experten 2 und 4 im Vergleich zu den anderen drei Experten sein. Diese Vermutung wird allerdings in den folgenden Untersuchungsergebnissen zur Reliabilität nicht bestätigt .

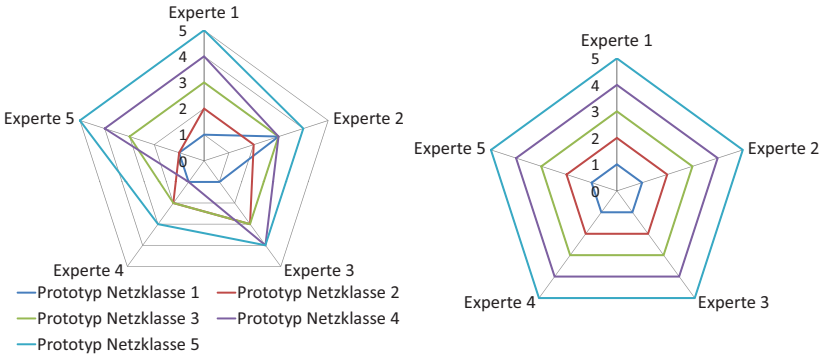


Abbildung 4: Netzdiagramme zur realen (links) und perfekten (rechts) Nutzung der Prototypen durch die Experten.

Vermutungen, inwieweit die Klassifikationsindikatoren durch die einzelnen Experten genutzt wurden, können auf Basis der Interraterkonkordanz zwischen den gerundeten Klassifikationsindikatoren und den Experteneinschätzungen getroffen werden. Die zu dieser statistischen Untersuchung nutzbaren Stichproben besitzen – entsprechend der Häufigkeit der Angabe der Klassifikationsindikatoren während der Befragung (Abbildung 3) – einen Umfang von 150 Netzen. Beginnend mit Experte 1 und endend mit Experte 5 führten die Berechnungen zu folgenden Konkordanzen: $\kappa_{w1} = 0{,}386$, $\kappa_{w2} = 0{,}734$, $\kappa_{w3} = 0{,}533$, $\kappa_{w4} = 0{,}004$, and $\kappa_{w5} = 0{,}428$. Mit einem Wert nahe null ist bei Experte 4 damit lediglich eine Übereinstimmung zu beobachten, die einer per Zufall enstandenen Übereinstimmung entspricht. Ein bewusste Nutzung des Klassifikationsindikators durch Experte 4 lässt sich auf Basis dieses Ergebnisses nicht vermuten. Der höchste Grad der Übereinstimmung zwischen den gerundeten Klassifikationsindikatoren und den in der Befragung vergebenen Klassen tritt bei Experte 2 auf. Dieses Ergebnis führt zu der Vermutung, dass die Klassifikationsindikatoren durch Experte 2 stark genutzt wurden. Die bei den Experten 1, 3 und 5 berechneten Werte für κ_w sind deutlich größer als null. Es liegt somit nicht nur eine zufällige Übereinstimmung vor. Dies lässt bei diesen Experten ebenfalls eine häufige Nutzung der Klassifikationsindikatoren erwarten.

Experte	1	2	3	4	5
Protoypen	ja	nein	nein	nein	ja
Indikatorfunktionen	häufig	häufig	häufig	selten	häufig
Klassifikationsindikator	mittel	sehr hoch	hoch	keine	hoch

Tabelle 11: Übersicht zur Nutzung der Zusatzinformationen durch die Experten.

Inwieweit die Ergebnisse der einzelnen Indikatorfunktionen, die als Grundlage zur Berechnung der Klassifikationsindikatoren für die Experten genutzt wurden, lässt sich anhand der erhobenen Informationen kaum nachvollziehen. Daher wurden die Experten direkt nach Abschluss der Klassifikation befragt, in welchem Ausmaß sie die im Fragebogen bereitgestellten Ergebnisse der Indikatorfunktionen bei der Klassifikation berücksichtigt haben. Zusätzlich wurden in den Gesprächen mit den Experten auch die aus den statistischen Zusammenhängen abgeleiteten Vermutungen zur individuellen Nutzung der Zusatzinformationen diskutiert. In den Gesprächen wurden die vermuteten kausalen Zusammenhängen weitestgehend bestätigt. Einzig die vermutete Nutzung der Prototypen durch Experte 3 hat sich im Nachhinein nicht bestätigt. Die drei beobachteten Übereinstimmungen aus Abbildung 4 sind lediglich per Zufall entstanden. Die Nutzung der Zusatzinformationen durch die einzelnen Experten ist in Tabelle 11 nochmal in übersichtlicher Form zusammengefasst. Ein Abgleich der Netze mit den Prototypen wurde dabei in der Befragung von zwei der Experten durchgeführt. Bei vier von fünf Experten liegt eine gute statistische Übereinstimmung zwischen den Klassifikationsindikatoren und ihren Klassifikationseinschätzungen vor. Es lässt sich festhalten, dass die Zusatzinformationen insgesamt rege durch die Experten im Laufe der Befra-

gung genutzt wurden. Die Bereitstellung der Zusatzinformationen kann auf Basis dieser empirischen Analysen als grundsätzlich sinnvoll erachtet werden.

Klassifikationsergebnisse zweier Musternetze:
Nachdem in Abschnitt 3.1.2 bereits die Strukturen der beiden Musternetze aus Abbildung 1 anhand einiger netzspezifischer Merkmale verglichen wurden, folgt nun eine Gegenüberstellung der Klassifikationseinschätzungen der Experten zu beiden Netzen. Zunächst werden allerdings beispielhaft die Ergebnisse der Indikatorfunktionen für die Merkmale 4, 7, 9 und 10 aus Tabelle 1 diskutiert, um die getroffene Festlegung der Berechnungsvorschriften zu evaluieren. Diese Merkmale wurden in Abschnitt 3.1.2 bereits für beide Musternetze diskutiert. Bei Merkmal 4 (Anzahl der KVS) führt die Berechnungsvorschrift (Tabelle 10) für Netz I zu Klasse 2 und für Netz II zu Klasse 4. Betrachtet man die Indikatorfunktion des Merkmals 7 (Vermaschungsgrad), so ergibt sich für Netz I eine Indikation zu Klasse 1. Für Netz II wiederholt sich die Indikation zu Klasse 4. Die Berechnungsvorschriften der Merkmale 9 (maximaler Stationsradius) und 10 (durchschnittlicher Stationsradius) indizieren beide jeweils eine Zugehörigkeit von Netz I zu Klasse 1 und für Netz II zu Klasse 5. Ein Vergleich dieser Analyse mit der Interpretation der Netze in Abschnitt 3.1.2 zeigt, dass die Berechnungsvorschriften der Indikatorfunktionen in Abstimmung mit den an der Befragung beteiligten Experten sinnvoll festgelegt wurden.

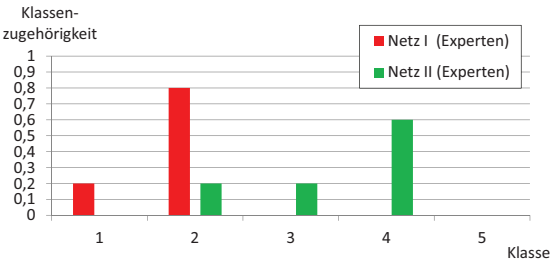


Abbildung 5: Verteilung der Experteneinschätzungen für die Netze I und II.

In Abbildung 5 sind die relativen Häufigkeiten der Experteneinschätzungen für beide Musternetze dargestellt. Sie resultieren aus den einzelnen Klassifikationseinschätzungen und werden hier mit Schwierigkeitseinschätzung zur Klassifikation durch die Experten angegeben. Netz I: (Klasse 2, „leicht“), (Klasse 1, „mittel“), (Klasse 2, „mittel“), (Klasse 2, „mittel“), (Klasse 2, „leicht“). Netz II: (Klasse 4, „leicht“), (Klasse 4, „mittel“), (Klasse 3, „leicht“), (Klasse 2, „mittel“), (Klasse 4, „leicht“). Es ist ersichtlich, dass Netz I mit einer hohen Klassenzugehörigkeit von 0,8 von den Experten zur Klasse 2 zugeordnet wird. Zusätzlich tritt eine geringe Klassenzugehörigkeit von 0,2 zu Klasse 1 auf. Im Gegensatz dazu tritt für Netz II eine relativ hohe Klassenzugehörigkeit von 0,6 zu Klasse 4 und eine jeweils geringe Klassenzugehörigkeit von 0,2 zu den Klassen 2

und 3 auf. Trifft man für beide Netze vereinfacht eine Entscheidung auf Basis der maximal auftretenden Klassenzugehörigkeit (Bayes'sches Risikominierungsprinzip), so wird Netz I der Klasse 2 und Netz II der Klasse 4 zugeordnet. Auch hier zeigt der Vergleich mit der Interpretation der Netze in Abschnitt 3.1.2, dass die Klassifikationsentscheidungen der Experten sinnvoll getroffen wurden.

Evaluation der Messqualität:
Es erfolgt nun auf Basis der Überlegungen aus Abschnitt 4.1.4 ex post eine empirische Evaluation der Messqualität. Die Objektivität der Befragung wird – wie in Abschnitt 4.1.4 erläutert – durch die eigenständige Bearbeitung des Fragebogens gewährleistet.

EXPERTE	1	2	3	4	5	$\mu(\kappa_w)$
k_w	0,31	0,34	0,38	0,58	0,66	0,45

Tabelle 12: Intraraterkonkordanz der Experten für die doppelt erhobenen Klassifikationsresultate.

Die Reliabilität der Befragung wird anhand der Intraraterkonkordanz der gedoppelten Klassifikationsergebnisse bewertet und mit der κ_w-Statistik gemessen. Tabelle 12 zeigt die resultierenden Werte. Der Mittelwert von $\mu(\kappa_w) = 0{,}45$ über alle Experten spricht für eine insgesamt gute Intraraterkonkordanz in der Befragung, woraus sich eine reliable Messung schlussfolgern lässt. Die Einzelwerte von κ_w weisen die höchste Intraraterkonkordanz bei Experte 5 auf. Experte 5 hat – wie oben untersucht – die Zusatzinformationen im Vergleich aller Experten am häufigsten genutzt. Bemerkenswert ist allerdings auch, dass bei Experte 4, der die Zusatzinformationen im Vergleich zu den anderen Experten am geringsten genutzt hat, ebenfalls eine hohe Intraraterkonkordanz mit einem Wert von 0,58 vorliegt. Im Gegensatz dazu hat Experte 1 auch vergleichsweise häufig die Zusatzinformationen genutzt, z. B. die Prototypen (Abbildung 4). Allerdings weisen die Klassifikationsergebnisse von Experte 1 mit einem Wert von 0,31 die im Vergleich geringste Intraraterkonkordanz auf. Es lässt sich damit kein direkter Zusammenhang zwischen der Reliabilität der Expertenaussagen und der Nutzung der Zusatzinformationen vermuten. Vielmehr lassen die Ergebnisse den Schluss zu, dass die Nutzung der Zusatzinformationen eher vom Erfahrungsniveau der Experten und deren individuellen Neigungen zur Bearbeitung der Klassifikationsaufgabe abhängig ist. Dieses Ergebnis unterstreicht das in Abschnitt 4.1.3 vorgeschlagene Vorgehen, dass, sofern Wissen zur Expertise der Experten vorhanden ist, die Häufigkeit der Bereitstellung von Zusatzinformationen bei der Erstellung des Fragebogens individuell an die Erfahrung der Experten angepasst werden sollte (z. B. in einer nächsten Befragung). Weiterhin lässt sich aus der empirisch nachgewiesenen Reliabilität der Klassifikationsergebnisse schließen, dass die Experten die Befragung offensichtlich mit der notwendigen Ernsthaftigkeit und Konzentration durchgeführt haben und keine Expertenaussagen aus diesem Grund ausgeschlossen werden müssen.

κ_w	Exp. 1	Exp. 2	Exp. 3	Exp. 4	Exp. 5
Exp. 1	—	0,099	0,48	-0,08	0,47
Exp. 2	—	—	0,14	0,06	0,19
Exp. 3	—	—	—	-0,02	0,48
Exp. 4	—	—	—	—	-0,08
Exp. 5	—	—	—	—	—

Tabelle 13: Paarweise Interraterkonkordanz zwischen den fünf Experten.

Die Bewertung der Validität gestaltet sich auf Grund einer fehlenden Vergleichsbasis schwierig. Hier wird die Validität anhand der paarweisen Konkordanz zwischen den Aussagen der Experten durchgeführt. Dies bedeutet, dass die Expertenaussagen gegenseitig mit Hilfe der Interraterkonkordanz verifiziert werden. Die aus den Befragungsergebnissen resultierenden Werte für κ_w sind in Tabelle 13 ausgewiesen. Es zeigt sich, dass die Klassifikationsaussagen der Experten 1, 2, 3 und 5 mit Werten von $\kappa_w > 0$ übereinstimmmen. Dabei weist nur die Konkordanz zwischen den Experten 2 und 3 einen geringeren Wert als 0,1 auf. Die Übereinstimmung zwischen Experte 4 und den übrigen Experten ist jeweils mit Werten von $\kappa_w < 0$ geringer als per Zufall zu erwarten. Bei genauerer Betrachtung der nicht übereinstimmenden Ergebnisse in den Konfusionsmatrizen lässt sich feststellen, dass die Klassifikationseinschätzungen von Experte 4 tendenziell niedriger sind, als die der übrigen Experten. Es lässt sich daher vermuten, dass in der Befragung von Experte 4 eine im Vergleich zu den anderen Experten strengere Bewertung vorgenommen wurde. Die höchsten Interraterkonkordanzen sind zwischen den Experten 1, 3 und 5 zu beobachten. Das Bewertungsverhalten dieser Experten führt damit zu guten Übereinstimmungen in den Klassifikationsergebnissen. Der Mittelwert $\mu(\kappa_w)$ bezüglich der paarweisen Interraterkonkordanz ergibt sich zu 0,175 und liegt deutlich über null. Insgesamt zeigen die ermittelten paarweisen Interraterkonkordanzen damit auf, dass aus der Befragung valide Klassifikationsergebnisse hervorgehen. Diese Aussage wird darüber hinaus durch den nachfolgend in Abschnitt 4.3 durchgeführten empirischen Vergleich mit den Ergebnissen des simulationsbasierten Klassifikationsansatzes gestützt, dessen Ergebnisse zwar auch keine „ground truth“ darstellen, allerdings im Gegensatz zu Expertenaussagen keinen subjektiven Einflüssen unterliegen.

Zusätzlich zur Analyse der Interraterkonkordanz sind in Abbildung 6 nochmal beispielhaft die absoluten Häufigkeiten der Klassen für die Experten 1, 4 und 5 ausgewiesen. Im Vergleich der drei aggregierten Klassenzugehörigkeiten fällt auf, dass sich deren Gestalt voneinander unterscheidet. Das Maximum tritt z. B. jeweils für eine unterschiedliche mittlere Klasse auf: Klasse 4 bei Experte 1, Klasse 2 bei Experte 4 und Klasse 3 bei Experte 5. Im Gegensatz dazu befindet sich das Minimum jeweils bei Klasse 5. Die bereits erwähnte tendenziell strengere Bewertung durch Experte 4 im Vergleich zu den anderen Experten spiegelt sich ebenfalls im Vergleich der absoluten Häufigkeiten wider. So wurden die Klassen 4 und 5 durch den Experten 4 nur selten vergeben. Bei Exper-

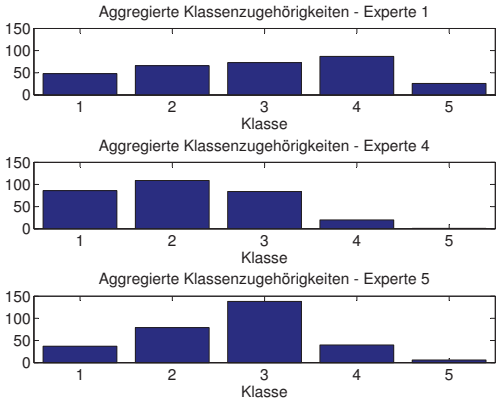


Abbildung 6: Gegenüberstellung der aggregierten Klassenzugehörigkeiten für die Experten 1, 4 und 5.

te 5 ist auffällig, dass dieser die Randklassen 1 und 5 nur selten genutzt hat. Dafür ist bei dem Experten 5 das Maximum bei Klasse 3 vergleichsweise stark ausgeprägt. Dies lässt vermuten, dass bei Experte 5 eine deutliche Tendenz aufgetreten ist, Netze nicht in eine der Randklassen einzuordnen. Bei Experte 1 wird die Bandbreite der Klassen im Vergleich am besten genutzt. Allerdings ist auch bei Experte 1 das Minimum für Klasse 5 deutlich ausgeprägt. Aus den absoluten Häufigkeiten lässt sich ableiten, dass ein Großteil der bewerteten Netze eine grundsätzlich robuste Struktur zur Integration von DEA aufweist. Die geringen absoluten Häufigkeiten von Klasse 5 zeigen allerdings auch, dass reale NS-Netze innerhalb der letzten Jahrzehnte zur ökomischen Versorgung von Kunden optimiert worden sind und zur Integration einer hohen DEA-Leistung verstärkt werden müssten.

Zusammenfassend zeigt die empirische Analyse der Messqualität, dass die Anwendung des entwickelten Klassifikationsansatzes zu reliablen und validen Klassifikationsresultaten mit einer hinreichend hohen Qualität führt.

4.2 SIMULATIONSBASIERTE STOCHASTISCHE KLASSIFIKATION

In diesem Abschnitt wird ein weiterer Ansatz zur Klassifikation von NS-Netzen vorgestellt, der auf einer dreischrittigen Klassifikationsstrategie basiert und bereits in [13] veröffentlicht wurde. Die erarbeitete Methodik ist an den normativen Stand der Technik zur Bewertung der Aufnahmekapazität von NS-Netzen für DEA angelehnt, der

in der Praxis zur Integration von DEA in der NS-Ebene angewendet wird. Da eine starke Abhängigkeit zwischen der Aufnahmekapazität eines NS-Netzes für DEA und der DEA-Konfiguration im Netz (Position des Netzanschlusses, Nennleistung der Anlage) exisitiert, erlaubt der Klassifikationsansatz die Berücksichtigung einer Vielzahl von zufallsgenerierten DEA-Konfigurationen im Netz. Dies führt zu einer differenzierten Bewertung der DEA-Kapazität eines Netzes bei der Klassifikation. Das Vorgehen bei der Klassifikation der Netze mit Hilfe des Ansatzes kann in folgende Schritte zusammengefasst werden:

1. Stochastische Simulation bezüglich der Aufnahmekapazität für DEA,

2. parametrische stochastische Modellierung der Aufnahmekapazität mit Weibullverteilungen und

3. Klassifkation auf Basis des ermittelten Weibullmodells.

In Abschnitt 4.2.1 wird zunächst ein Überblick zum Stand des Wissens zur simulationsbasierten Bewertung des Aufnahmevermögens von NS-Netzen für DEA dargelegt und der Klassifikationsansatz wird in den wissenschaftlichen Kontext eingeordnet. Daraufhin wird in Abschnitt 4.2.2 die maximale DEA-Kapazität von NS-Netzen durch Ableitung von zwei Kriterien definiert. Auf dieser Grundlage wird ein stochastischer Simulationsalgorithmus zur Bewertung der maximalen DEA-Kapazität von NS-Netzen erarbeitet. Anschließend werden in Abschnitt 4.2.3 Ansätze zur stochastischen Modellierung der Simulationsergebnisse mit Weibullverteilungen und deren statistischer Validierung vorgestellt. Die Klassifikation von NS-Netzen auf Basis der Weibullverteilungen wird in Abschnitt 4.2.4 beschrieben. Zum Abschluss wird in Abschnitt 4.2.5 eine empirische Evaluation des Klassifikationsansatzes durchgeführt.

4.2.1 *Vorbemerkungen*

Zur Untersuchung der Einflüsse einer steigenden Anzahl von DEA in der NS-Ebene und zur Ableitung von Gegenmaßnahmen wurde bereits viel Forschungsarbeit geleistet. Papaioannou et. al. entwickeln ein detailliertes Modell einer PV-Einspeisung im NS-Netz [64]. Die Ergebnisse einer anhand dieses Modells durchgeführten Simulationsstudie zeigen, dass auch bei hoher PV-Einspeisung im NS-Netz keine Verletzungen der Oberschwingungsgrenzwerte zu erwarten sind, jedoch zur Einhaltung der Spannungsgrenzwerte aus [8] Gegenmaßnahmen erforderlich werden. Ergänzend dazu untersuchen Inam et al. den Einfluss der Größe und Position einer PV-Anlage auf die DEA-Kapazität eines NS-Netzes [65]. Die Analysen zeigen, dass das Aufnahmevermögen beim Anschluss vieler verteilter Kleinanlagen höher ausfällt, als beim Anschluss einzelner großer PV-Anlagen. Dieses Ergebnis wird von einer Untersuchung der Aus-

wirkungen von PV-Anlagen auf ein typisches NS-Netz in Malaysia gestützt [66]. Auf Grund der relativ kurzen Leitungslängen und dem hohen Verkabelungsgrad des Netzes resultiert eine hohe PV-Leistung, die vom NS-Netz aufgenommen werden kann. Zudem zeigt sich, dass der Einfluss einer PV-Anlage auf das NS-Netz sehr stark netzstruktur- und ortsabhängig ist. Ansätze zur Erhöhung des DEA-Aufnahmevermögens durch Regelungsstrategien für PV-Wechselrichter und regelbare Ortsnetztransformatoren findet man z. B. in [67].

Bei der Erarbeitung des simulationsbasierten Klassifikationsansatzes bestand das Ziel darin, eine Methodik mit hoher praktischer Relevanz bezüglich einer Anwendung dieser Methodik an einer großen Anzahl an realen NS-Netzen bereitzustellen. Daher ist der entwickelte Ansatz stark an die Untersuchungen von Kerber und Witzmann in [6, 68] angelehnt und stellt eine Verfeinerung der darin vorgestellten Ansätze dar. In [68] wird ein Ansatz zur Modellierung verschiedener netzspezifischer Merkmale mit Weibullverteilungen für drei Netzklassen vorgestellt (z. B. Verteilung der Transformatornennleistung). Die Einteilung der untersuchten NS-Netze in die Klassen „ländlich“, „dörflich“ und „vorstädtisch“ wird anhand der bebauten Fläche vor den Analysen festgelegt und nicht anhand von Simulationen bestimmt. Folglich werden im Gegensatz zum hier vorgestellten Klassifikationsansatz die Weibullverteilungen nicht zur Modellierung des DEA-Aufnahmevermögens, sondern zur Beschreibung der Verteilung einzelner Merkmale verwendet. Auf der Basis der erhobenen stochastischen Modelle wird zusätzlich ein Ansatz zur Konzeption von synthetischen Referenznetzen für die drei vordefinierten Netzklassen vorgestellt. Dazu wird vorgeschlagen, typische Merkmalwerte anhand der Erwartungswerte der Verteilungsfunktionen zu bestimmen und die Referenznetze so zu erzeugen, dass in diesen die ermittelten typischen Merkmalwerte in Kombination auftreten. Darüber hinaus untersuchen sie in [6] die DEA-Kapazität von NS-Netzen anhand von Lastflussberechnungen an einigen realen Netzstrukturen. Die DEA-Leistung wird dabei homogen auf die Hausanschlüsse in den Netzen verteilt und bis zu einer Verletzung der durch [8] und [69] definierten Grenzwerte oder einer Betriebsmittelüberlastung gleichmäßig hochskaliert. Die genutzten Grenzwerte sind zwar von hoher praktischer Relevanz, allerdings erscheint die vorgenommene homogene Verteilung der DEA-Leistung auf die Hausanschlüsse vor diesem Hintergrund nicht angemessen zu sein. Die über Jahre real auftretende Erhöhung der DEA-Leistung in einem NS-Netz wird nämlich von vielen nichttechnischen Faktoren beeinflusst. Wichtige nichttechnische Einflussfaktoren sind die lokalen Potenziale für die Installation von DEA (z. B. die Dachfläche bei PV-Anlagen), die Bereitstellung von energieträgerabhängigen Subventionen durch den Staat, die Entwicklung der Marktpreise und die regionale Kaufkraft.

Im Gegensatz zu den vorgestellten Arbeiten wird in dem hier vorgestellten Ansatz die Klassifikation von NS-Netzen nicht bereits als gegeben angesehen (z. B. auf Basis der vorhandenen Bebauung [68]), sondern hinsichtlich der DEA-Kapazität des NS-Netzes durchgeführt. Weiterhin wird durch die stochastische Modellierung ein differenzierter Ansatz zur Bewertung des Aufnahmevermögens für DEA bereitgestellt, in dem eine

Vielzahl realistischer Szenarien bezüglich der Erhöhung der DEA-Leistung berücksichtigt werden. Einen ähnlichen Simulationsansatz allerdings ohne anschließende stochastische Modellierung und Klassifikation findet man in [1]. Die hier verwendete Systematik stellt folglich eine direkte Erweiterung der simulationsbasierten Bewertungen um eine stochastische Modellierung und modellbasierte Klassifikation dar. Die Resultate können entweder als probabilistische Klassenzugehörigkeiten interpretiert oder in scharfe Klassenzugehörigkeiten überführt werden. Die Anwendbarkeit des Klassifikationsansatzes wird im Gegensatz zu vielen anderen Studien nicht nur an synthetischen Modellnetzen (wie z. B. in [1]) oder einer geringen Zahl realer Netze, sondern an einer Vielzahl realer NS-Netze demonstriert (300 reale ländliche und vorstädtische Netze). Um die Vorteile zu verdeutlichen, wird für diese Netze ergänzend der Simulationsansatz aus [6] angewendet. Der empirische Vergleich zeigt, dass die Ergebnisse des Ansatzes aus [6] eine gute Näherung des Erwartungswertes der Weibullverteilungen liefern, die mit der hier vorgestellten dreischrittigen Klassifikationsstrategie bestimmt wurden.

4.2.2 *Stochastische Simulation*

Im Folgenden werden zunächst zwei formale Kriterien zur Abschätzung der Aufnahmekapazität von NS-Netzen für DEA auf Basis der relevanten Normen und Richtlinien (Abschnitt 1.2) festgelegt. Anschließend wird darauf aufbauend ein stochastischer Simulationsalgorithmus erarbeitet, mit Hilfe dessen eine Datengrundlage zur stochastischen Modellierung gebildet wird.

Kriterien zur Beurteilung der DEA-Aufnahmekapazität von NS-Netzen:
Wie bereits im Abschnitt 1.2 beschrieben, sind in der Spannungsqualitätsnorm [8] Grenzen für die Versorgungsspannung im Verteilnetz definiert worden. Um deutsche MS- und NS-Netze hinsichtlich des Anschlusses von DEA entkoppelt voneinander betrachten zu können, ohne die in [8] definierten Spannungsbandgrenzen zu verletzen, wird in [9] für die NS ein 3%-Kriterium zur Bewertung der relativen Spannungsänderung bereitgestellt, die durch alle DEA im Netz hervorgerufen wird. Dieses Kriterium kann genutzt werden, ohne bei der Anschlussbeurteilung das überlagerte MS-Netz mit in Betracht zu ziehen. Das Kriterium beinhaltet die Empfehlung, dass die durch alle DEA im Netz hervorgerufene relative Spannungsänderung (bezogen auf den analogen Lastfall ohne DEA) ein Sicherheitslimit von 3% nicht überschreiten darf. Ein weiterer Vorteil des Kriteriums ist, dass es durch den Bezug auf den analogen Lastfall nahezu unabhängig von gewählten Lastannahmen wird. Aus diesen Gründen wird das 3%-Kriterium hier als erstes Kriterium (Kriterium 1) zur Beurteilung der DEA-Aufnahmekapazität von NS-Netzen gewählt. Formal kann das Kriterium 1 wie folgt formuliert werden:

$$\Delta u = \left(\max_{\gamma=1,\ldots,N_N} \frac{|\underline{U}_{DEA,\gamma} - \underline{U}_{0,\gamma}|}{|\underline{U}_{0,\gamma}|} - 1 \right) \leq 0,03, \tag{6}$$

mit

Δu:	Relative Spannungsänderung,
γ:	Knotenindex,
N_N:	Anzahl der Knoten,
$\underline{U}_{DEA,\gamma}$:	Komplexe Spannung im Knoten γ mit Erzeugung,
$\underline{U}_{0,\gamma}$:	Komplexe Spannung im Knoten γ ohne Erzeugung.

Für ein NS-Netz mit nur einer einzigen DEA, kann Kriterium 1 auch vereinfacht aus dem Verhältnis der Kurzschlussleistung der DEA und der Kurschlussleistung des Netzes im Netzverknüpfungspunkt der DEA bestimmt werden [69]. Dieser vereinfachte Ansatz für eine einzige DEA kann wie folgt interpretiert werden: Wenn das Verhältnis einen Wert von 3% nicht überschreitet, so ist die Netzstruktur stark genug, um die DEA in das Netz zu integrieren. De facto werden auch viele weitere Netzrückwirkungsarten anhand des Verhältnisses aus Kurzschlussleistung der DEA und der Netzkurzschlussleistung bestimmt (z. B. Emission von Harmonischen). Aus dieser Betrachtung heraus resultiert, dass – unter gleichzeitiger Berücksichtigung der Betriebsmittelbelastungen im Netz – die Anwendung von Kriterium 1 mittelbar auch zur Begrenzung von vielen weiteren Netzrückwirkungen durch DEA im Netz dient. Diese Annahme wird z. B. auch durch die Untersuchungsergebnisse in [64] gestützt. Die Anwendung des Kriteriums ist also hinsichtlich einer Abschätzung der DEA-Aufnahmekapazität als adäquat zu beurteilen. Bezüglich der fluktuierenden Einspeisung von DEA findet bei Anwendung dieses Kriteriums eine Betrachtung des kritischen Lastflusses statt, in dem nahezu keine Abnahme, aber die maximale dezentrale Einspeisung im Netz auftritt.

Zusätzlich zur Einhaltung von Kriterium 1 darf die Netzintegration von DEA zu keinem Zeitpunkt zu einer unzulässig hohen Betriebsmittelbelastung im Netz führen. Diese Bedingung kann wie folgt formalisiert werden und dient als Kriterium 2 (Gleichung (7)). Der Einfluss der Temperatur auf die Betriebsmittelbelastung wird vereinfachend nicht berücksichtigt:

$$I_{max} = \max_{\beta=1,\ldots,N_\beta} \frac{\left|\underline{I}_{L,\beta}\right|}{I_{N,\beta}} \leq 1, \tag{7}$$

mit

I_{max}:	Maximalbelastung eines Betriebsmittels im Netz,
β:	Betriebsmittelindex,
N_β:	Anzahl der elektrischen Betriebsmittel,
$\underline{I}_{L,\beta}$:	Komplexer Strom in Equipment β
$I_{N,\beta}$:	Nennstrom von Betriebsmittel β.

Die erfolgte Ableitung der beiden Kriterien zur Definition des maximalen Aufnahmevermögens eines NS-Netzes für DEA aus europäischen Normen stellt sicher, dass der normative Stand der Technik Berücksichtigung findet, der in der Praxis zu einer Netz-

verstärkung (und -investition) führt. Daher können die unter Anwendung dieser Kriterien erhaltenen Ergebnisse als hochrelevant bezüglich der Netzintegration von DEA in europäischen NS-Netzen betrachtet werden.

Stochastischer Simulationsalgorithmus:
Bezüglich der Aufnahmekapazität eines NS-Netzes für DEA haben die Nennleistungen und die Anordnung der DEA im Netz einen sehr starken Einfluss. Die DEA-Aufnahmekapazität eines Netzes ist z. B. bei Anschluss von vielen, allerdings kleinen PV-Anlagen höher, als im Vergleich zum Anschluss weniger Großanlagen [65, 66]. Dies bedeutet, dass – um zuverlässige und robuste Abschätzungen der DEA-Kapazität eines Netzes zu erhalten – nicht nur eine spezielle Anordnung von DEA betrachtet werden kann (z. B. Hochskalieren der Leistung von homogen im Netz verteilten DEA). Auch die Betrachtung einer einzelnen aus einer Dachflächenanalyse resultierenden speziellen Anordnung von DEA im Netz ist nicht ausreichend, um zuverlässige Abschätzungen zu erhalten, da in der Realität weitere Einflüsse wie z. B. mikrofundierte Kaufkräfte auf den Ausbau von DEA wirken. Vielmehr ist die DEA-Aufnahmekapazität als Zufallsvariable zu behandeln. Dieser Tatsache wird im Folgenden durch die Entwicklung eines auf den definierten Kriterien basierenden, stochastischen Simulationsalgorithmus Rechnung getragen, der eine anschließende Modellierung der Verteilungsfunktion dieser Zufallsvariable ermöglicht.

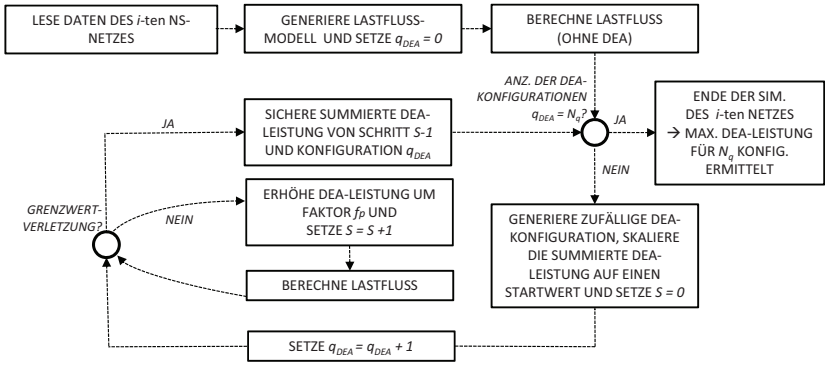


Abbildung 7: Flussdiagramm der stochastischen Simulation.

Vor dem Hintergrund der vielfältigen nichttechnischen Einflüsse auf die Installation von DEA wird die Wahrscheinlichkeit, dass eine bestimmte Leistung in einem Netzknoten auftritt, anhand einer Gleichverteilung modelliert. Dies bedeutet, dass jedem möglichen Leistungswert einer DEA die gleiche Wahrscheinlichkeit zugeordnet wird ($p(P_{DEA}) = \text{konst.}$, d. h. Wirkleistungswert der dezentralen Erzeugungsleistung P_{DEA} unterliegt einer Gleichverteilung). Diese Annahme ist auch insofern sinnvoll, als dass die Modelle realer NS-Netze oftmals nicht im Detail jeden Hausanschluss enthalten

(z. B. die in dieser Arbeit verwendeten Netzmodelle). In diesem Fall beinhalten die Modelle lediglich Knoten an den Stoßstellen zweier Leitungsabschnitte (z. B. Muffen oder Kabelverteilerschränke). In der Realität werden in einem NS-Netz Knoten auftreten, in denen auch in Zukunft keine DEA angeschlossen sein werden. Dementgegen steht allerdings die Ungewissheit über den finalen Durchdringungsgrad von DEA in einem Netz, der stark abhängig von Änderungen der politischen Rahmenbedingungen ist. Weiterhin wird Nebenwissen über die simulierten Netze (z. B. Dachflächenpotenziale) nicht berücksichtigt, um einen Simulationsalgorithmus bereitzustellen, welcher auf eine große Anzahl von NS-Netzen angewendet werden kann. Der erarbeitete Simulationsalgorithmus wird in Abbildung 7 allgemein für das i-te zu untersuchende Netz illustriert. Das Ziel der Simulationen ist die Generierung einer empirischen Verteilungsfunktion unter Berücksichtigung einer großen Zahl N_q verschiedener DEA-Konfigurationen im Netz. Der Wert von N_q gibt damit die Anzahl der berücksichtigten, unterschiedlichen DEA-Konfigurationen im Netz an.

Zu Beginn der Simulation werden zunächst die Daten des i-ten Netzes geladen und darauf basierend die nichtlinearen Gleichungen zur Lastflussberechnung im Netz vorbereitet. Das resultierende Gleichungssystem fußt auf der physikalischen Begebenheit, dass sich die Knotenspannungen im Netz gemäß des Gleichgewichtszustandes für die Wirk- und Blindleistungen in jedem Knoten einstellen:

$$\mathrm{Re}\{3\underline{\boldsymbol{U}}_K\underline{\boldsymbol{Y}}^*_{KK}\underline{\boldsymbol{u}}^*_K\} = \Delta\boldsymbol{p} = 0, \tag{8}$$

und

$$\mathrm{Im}\{3\underline{\boldsymbol{U}}_K\underline{\boldsymbol{Y}}^*_{KK}\underline{\boldsymbol{u}}^*_K\} = \Delta\boldsymbol{q} = 0, \tag{9}$$

mit

$\underline{\boldsymbol{U}}_K$:	Diagonalmatrix der komplexen Spannungen,
$\underline{\boldsymbol{Y}}_{KK}$:	Komplexe Knotenadmittanzmatrix,
$\underline{\boldsymbol{u}}_K$:	Vektor der komplexen Spannungen,
$\Delta\boldsymbol{p}$:	Vektor der Wirkleistungsdifferenzen,
$\Delta\boldsymbol{q}$:	Vektor der Blindleistungsdifferenzen,
*:	Operator für komplexe Konjugation.

Die bereits installierten DEA sollten vernachlässigt bleiben, um eine gleiche Ausgangssituation für jedes Netz und folglich vergleichbare Simulationsergebnisse zu erhalten. Diese Vernachlässigung liegt auch darin begründet, dass eine bereits installierte, hohe Zahl an DEA (z. B. wenn die durch die Kriterien 1 und 2 definierte Leistung fast erreicht ist) eine fundierte Schätzung der empirischen Verteilungsfunktion nicht zulässt. Ein Vergleich der Robustheit verschiedener Netzstrukturen wäre in der Folge kaum möglich. Im nächsten Schritt erfolgt eine erste Lastflussberechnung, in der keine dezentrale Erzeugung im Netz berücksichtigt wird. Die aus dieser Berechnung resultierenden Knotenspannungen dienen später als Referenzwerte zur Auswertung des 3%-Kriteriums

(Kriterium 1). Während der Berechnung wird das nichtlineare Gleichungssystem bestehend aus den Gleichungen (8) und (9) zunächst durch Taylorreihenentwicklung linearisiert und anschließend unter Anwendung des Newton-Raphson Algorithmus [70] gelöst. Nach erfolgter Berechnung des Referenzlastflusses wird am ersten Knoten des Flussdiagrammes die Anzahl der bereits berücksichtigten DEA-Anordnungen im betrachteten Netz überprüft. Wird die gewählte Grenze der N_q zu berücksichtigenden Konfigurationen erreicht, so ist die Simulation für das i-te Netz vollständig durchlaufen. Ist N_q noch nicht erreicht, so wird für jeden Knoten des Netzes ein Zufallswert für die dort installierte Erzeugungsleistung bestimmt ($P_{DEA,\gamma} \in [0,1]$, $\gamma = 1,\ldots,N_N$, N_N: Anzahl der Knoten, Erzeugung der Zufallswerte basiert auf einer Gleichverteilung). Nachdem jedem Knoten ein zufälliger Wert für die installierte Erzeugungsleistung zugeordnet worden ist, wird die Summe der Erzeugungsleistungen über alle Knoten auf einen festen Startwert skaliert, z. B. wie hier auf 1kW, um die Simulation mit einer möglichst geringen summierten Erzeugungsleistung zu beginnen. Nun wird abwechselnd – in Form einer weiteren Schleife – eine Lastflussberechnung mit erhöhter Erzeugungsleistung durchgeführt und die Verletzung der definierten Kriterien überprüft (3%-Kriterium, Betriebsmittelbelastungen). Die Hochskalierung in jedem Durchlauf dieser Schleife wird anhand eines Faktors f_p vorgenommen, der die Erzeugungsleistungen in jedem Knoten des Netzes gleichmäßig erhöht. Dies bedeutet, dass die berücksichtigte Erzeugungsleistung in einem Knoten mit f_p multipliziert wird. Der Abbruch der Schleife erfolgt bei derjenigen Summe der Erzeugungsleistungen, bei der entweder Kriterium 1 oder Kriterium 2 verletzt wird. Die Summe der in der vorletzten Lastflussberechnung aufgetretenden Erzeugungsleistungen in den Knoten ($P_{DEA} = \sum_{\gamma=1}^{N_N} P_{DEA,\gamma}$) entspricht dann dem abgeschätzten Aufnahmevermögen des Netzes für die getroffene zufällige Anordnung der DEA. Damit resultieren nach Durchlauf des Simulationsalgorithmus unter Zufallseinfluss N_q Ausprägungen des DEA-Aufnahmevermögens, die zur Erstellung einer empirischen Verteilungsfunktion und einer anschließenden stochastischen Modellierung nutzbar sind. Die Simulationsergebnisse berücksichtigen dabei keinen Blindleistungsaustausch zwischen den DEA und dem NS-Netz. Die Berücksichtigung könnte allerdings anhand einer weiteren Simulation mit veränderten Phasenwinkeln der DEA erfolgen.

4.2.3 *Parametrische stochastische Modellbildung*

Die aus dem in Abschnitt 4.2.2 vorgestellten Simulationsalgorithmus resultierenden Ergebnisse werden nun zur parametrischen stochastischen Modellierung der DEA-Aufnahmekapazität eines NS-Netzes genutzt. Die Grenzwertverletzungen bezüglich der Kriterien 1 und 2 können abstrakt mit dem Ausfall eines Systems bei fortschreitender Lebensdauer verglichen werden. Die Knoten und Betriebsmittel des Netzes entsprechen in diesem Vergleich den einzelnen Komponenten des Systems, die selber ein belastungsabhängiges Ausfallverhalten hätten (Knoten → Kriterium 1, Betriebsmittel → Kriterium 2). Der Ausfall des Systems tritt dann auf, wenn die Lebensdauer eines

der Bauteile erreicht ist. Die wachsende dezentrale Erzeugungsleistung entspricht der fortschreitenden Lebensdauer der Komponenten bzw. des Systems. Die aufgezeigten Analogien sind als starkes Signal dafür zu werten, dass eine stochastische Modellierung der DEA-Aufnahmekapazität mittels Weibullverteilungen ein adäquater parametrischer Ansatz ist. Dies liegt darin begründet, dass Weibullverteilungen in vielen technischen Anwendungen zur Lebensdaueranalyse von Komponenten eingesetzt werden. Für die hier vorgesehene Anwendung ist die Nutzung der dreiparametrigen Form der Weibullverteilung sinnvoll, da sie zur Berücksichtigung einer minimalen Erzeugungsleistung, die in jedem Fall ohne Grenzwertverletzungen in das Netz integriert werden kann, einen Verschiebungsparameter bereitstellt. Allgemein kann die Dreiparameterform der Weibulldichtefunktion wie folgt dargestellt werden [71]:

$$f(x) \;=\; ab^{-1}\left(\frac{x-v}{b}\right)^{a-1}\exp\left(-\left(\frac{x-v}{b}\right)^{a}\right), \qquad (10)$$
$$\text{mit } x > v \quad \text{und} \quad a, b > 0.$$

Der Parameter a wird als Form- und b als Skalierungsparameter bezeichnet. Der Parameter v ist der bereits erwähnte Verschiebungsparameter, der die minimale Erzeugungsleistung verkörpern wird, die normkonform in ein Netz integriert werden kann. Die Zweiparameterform der Weibullverteilung erhält man aus Gleichung (10), indem v zu null gesetzt wird. Weiterhin liefert die Integration von Gleichung (10) die allgemeine Form der dreiparametrigen Weibullverteilungsfunktion:

$$F(x) \;=\; 1-\exp\left(-\left(\frac{x-v}{b}\right)^{a}\right), \qquad (11)$$
$$\text{mit } x > v \quad \text{und} \quad a, b > 0.$$

In den in dieser Arbeit durchgeführten Untersuchungen bildet die DEA-Aufnahmekapazität eines Netzes P_{DEA} die Zufallsvariable x des parametrischen Modells in den Gleichungen (10) und (11).

Um nun zunächst aus den Simulationsergebnissen eine empirische Weibullfunktion für das DEA-Aufnahmevermögen zu erhalten, werden die N_q Ausprägungen der Zufallsvariable P_{DEA} zunächst in aufsteigender Reihenfolge sortiert. Weiterhin werden die Werte von P_{DEA} noch durch die Anzahl der Transformatorstationen im Netz geteilt, so dass sie eine durchschnittliche Erzeugungsleistung je Stationsbereich repräsentieren. Die Normierung auf die Anzahl der Stationsbereiche eines Netzes wird vorgenommen, um die Ergebnisse von Netzen mit unterschiedlicher räumlicher Ausdehnung vergleichbar zu gestalten. Das Sortieren der Simulationsergebnisse in aufsteigender Reihenfolge kann als Rangfolge interpretiert werden, die eine Zuordnung von kumulierten Wahrscheinlichkeiten $\widetilde{F}(\overline{k}_r)$ zu jeder innerhalb der Rangfolge auftretenden Ausprägung $\overline{k}_r$ der Zufallsvariable ermöglicht. Die Zuordnung der Wahrscheinlichkeiten kann – wie in dieser Arbeit – z. B. mittels Bernard's Approximation des Medianranges erfolgen [72]:

$$\widetilde{F}(\overline{k}_r) \;\approx\; \frac{r-0.3}{\widehat{n}+0.4}, \qquad (12)$$

mit

$\overline{k}_r$:	Sortierte Ausprägung der Zufallsvariable,
r:	Rang,
$\widehat{n}$:	Anzahl der Ausprägungen der Zufallsvariable.

Die sortierten Ausprägungen entsprechen dann den $\overline{k}_r$ in Gleichung (12). Mit Hilfe dieser Vorarbeit kann die Verteilungsfunktion des dreiparametrigen Weibullmodells nun an die Simulationsergebnisse angepasst werden. Die Schätzung der Modellparameter für die dreiparametrige Weibullfunktion wird – verglichen mit dem Zweiparameterfall – durch das Auftreten des Verschiebungsparameters v erschwert. Einige hilfreiche Schätzungsmethoden werden z. B. in [73, 74, 75] vorgestellt. Da der Berechnungsaufwand im Vergleich zum in Abschnitt 4.2.2 vorgestellten Simulationsalgorithmus verschwindend gering ist, kann pragmatisch eine Nutzung der Methode der kleinsten Quadrate [76] oder der Maximum-Likelihood Methode [77] zur Bestimmung der Zweiparameterform der Weibullverteilung erfolgen. Um eine Schätzung für den Verschiebungsparameter zu erhalten, werden dabei die empirischen Daten durch verschiedene Verschiebungsparameter angepasst und anschließend derjenige gewählt, für den sich die beste Modellanpassung ergibt. Die fehlenden Schätzungen der Form- und Skalierungsparameter ergeben sich dann aus dem korrespondierenden Zweiparametermodell. In den im Rahmen dieser Arbeit erfolgten empirischen Untersuchungen wurde zur Parametrisierung des Zweiparametermodells die Methode der kleinsten Quadrate verwendet.

Um zu testen, wie gut ein parametrisches Modell an eine empirische Datenreihe angepasst wurde, können statistische Anpassungstests durchgeführt werden. Diese Anpassungstests werden in der Literatur als Goodness-of-fit-Tests bezeichnet (siehe z. B. [71]) und entsprechen im Wesentlichen Hypothesentests, die auf vom zu überprüfenden Modell abhängigen Verteilungen basieren. Dabei wird die Diskrepanz zwischen den empirischen Daten und den dazu korrespondierenden, aus dem Modell resultierenden Werten gemessen. Zur Evaluation der Anpassung der dreiparametrigen Weibullverteilung an die Simulationsresultate, wurden in dieser Arbeit die Korrelationsstatistiken vom Shapiro-Wilk Typ genutzt, die in [71] veröffentlicht wurden. Dort werden die Teststatistiken R_W^2, für R_W aus Gleichung (13), definiert und für Anpassungstests an das Dreiparametermodell genutzt. Anschaulich stellt R_W dabei den Korrelationskoeffizienten einer Quantil-Quantil-Grafik dar [71].

$$R_W = \frac{\sum_{r=1}^{\widehat{n}} (\overline{k}_r - \overline{k}) m s_{w,r}}{\left(\sum_{r=1}^{\widehat{n}} (\overline{k}_r - \overline{k})^2 \sum_{r=1}^{n} (m s_{w,r} - \overline{ms})^2 \right)^{0,5}} \tag{13}$$

und

$$ms_{w,r} = \left(-\ln\left(1 - \frac{r - 0,3175}{\widehat{n} + 0,365}\right)\right)^{1/\hat{a}}, \tag{14}$$

mit

$\widehat{n}$:	Anzahl der Ausprägungen der Zufallsvariable,
$\overline{k}_r$:	Sortierte Ausprägung der Zufallsvariable,
$\overline{k}$:	Durchschnitt der Ausprägungen der Zufallsvariable,
$ms_{w,r}$:	Median Score,
$\overline{ms}$:	Durchschnittlicher Median Score,
$\hat{a}$:	Geschätzter Formparameter.

Der Median Score stellt eine Näherungsformel zur Bewertung des statistischen Zusammenhangs zwischen empirischen Daten und der dreiparametrigen Weibullfunktion dar und hängt zusätzlich von der Schätzung $\hat{a}$ des Formparameters a ab, der hier wie beschrieben mittels der Methode der kleinsten Quadrate bestimmt wurde. Für die bezüglich eines Signifikanzniveaus geltenden kritischen Werte werden Näherungsformeln angegeben, die ebenfalls eine Abhängigkeit der kritischen Werte von der Parameterschätzung $\hat{a}$ berücksichtigen [71]. Zur Berechnung der kritischen Werte für das 99%-Signifikanzniveau ($\alpha = 0,01$) ergibt sich folgende Abhängigkeit [71]:

$$\begin{aligned} \mathrm{RWS}_{\alpha=0,01} = {} & 0,99910494 - \frac{2,66292}{\widehat{n}} \\ & + \frac{12,86169089}{\widehat{n}^2} - 0,0403293 \\ & + 0,02112679\,\hat{a} - 0,00250893\,\hat{a}^2, \end{aligned} \tag{15}$$

mit

$\widehat{n}$:	Anzahl der Ausprägungen der Zufallsvariable
$\hat{a}$:	Geschätzter Formparameter

Auf dieser Basis kann eine Bewertung der Modellgüte vorgenommen werden. Dazu wird nochmals die bereits zur Ermittlung der empirischen Verteilungsfunktion ermittelte Rangfolge der N_q Ausprägungen des Aufnahmevermögen für DEA benötigt. Die sortierten Ausprägungen entsprechen den $\overline{k}_r$ in Gleichung (13). Der Wert $\widehat{n}$ wird entsprechend gleich der in den Simulationen berücksichtigten DEA-Konfigurationen N_q gewählt. Der Wert des geschätzten Formparameters $\hat{a}$ resultiert direkt aus der oben beschriebenen Parameterschätung. Auf dieser Basis können die weiteren zur Bewertung des Modells benötigten Variablen der Gleichungen (13) bis (15) berechnet werden. Zur Bewertung der Modellgüte wurde in dieser Arbeit anschließend ein Hypothesentest auf dem 99%-Signifikanzniveau durchgeführt. Dies bedeutet im Konkreten, die Hypothese, dass die dreiparametrige Weibullverteilungsfunktion mit gewählter Signifikanz an die empirischen Daten angepasst werden kann, wird zurückgewiesen, wenn R^2_W kleiner oder

gleich dem ermittelten kritischen Wert $RWS_{\alpha=0.01}$ ist. Der α-Fehler (oder Typ I Fehler) beschreibt die Wahrscheinlichkeit, dass die Hypothese zurückgewiesen wird, obwohl sie in Wirklichkeit wahr ist und beträgt auf dem 99%-Signifikanzniveau 1%. Im Gegensatz zum α-Fehler kann über den β-Fehler (oder Typ II Fehler), der die Wahrscheinlichkeit beschreibt, dass eine Hypothese nicht zurückgewiesen wird, obwohl sie nicht wahr ist, keine numerische Aussage getroffen werden.

4.2.4 *Simulationsbasierte Klassifikation*

Auf Basis der parametrischen, stochastischen Modellierung in Abschnitt 4.2.3 wird nun ein neuer Ansatz zur Klassifikation von NS-Netzen abgeleitet. Dazu werden $C = 5$ verschiedene ordinale Klassen genutzt, die in aufsteigender Rangfolge die Stärke der Netzstrukturen hinsichtlich des Aufnahmevermögens für DEA repräsentieren:

1. sehr schwach,
2. schwach,
3. durchschnittlich,
4. stark und
5. sehr stark.

Um ein Netz in eine dieser Klassen einzuordnen, werden zunächst Klassengrenzen für die durchschnittliche Erzeugungsleistung je Stationsbereich P_{DEA} bzw. allgemein die Zufallsvariable im Modell definiert. Zur Klassifikation wird für jede Klasse die Klassenzugehörigkeit anhand der beiden Klassengrenzen aus der Weibullverteilung bestimmt. Die Klassenzugehörigkeit gibt folglich an, mit welcher Wahrscheinlichkeit die durchschnittliche Erzeugungsleistung je Stationsbereich zwischen den Klassengrenzen einer Klasse in den Simulationen auftritt:

$$\begin{aligned} p(c) &= p(l_{u,c} \leq P_{DEA} \leq l_{o,c}) \\ &= F(l_{o,c}) - F(l_{u,c}), \end{aligned} \tag{16}$$

und

$$c = 1, \ldots, C = 5,$$

mit

P_{DEA}: Zufallsvariable,

c: Klassenindex,
$l_{u,c}$: Untere Klassengrenze der Klasse c,
$l_{o,c}$: Obere Klassengrenze der Klasse c.

Die Anwendung dieser Methodik führt direkt zu probabilistischen Klassenzugehörigkeiten und gibt somit einen sehr differenzierten Eindruck zur Stärke von Netzstrukturen. Falls scharfe Klassenzugehörigkeiten benötigt werden, kann für ein Netz die Klasse mit der höchsten Klassenzugehörigkeit gewählt werden. An dieser Stelle werden keine speziellen Klassengrenzen vorgegeben, da bei Anwendung der Methodik ein gewisser Freiheitsgrad bestehen sollte, um die Klassengrenzen an die betrachtete Grundgesamtheit der Netze anzupassen. Dies ist insofern sinnvoll, als dass z. B. jeder VNB eine eigene Planungs- und Betriebsphilosophie verfolgt (Selektivität, Netzstruktur, Dimensionierung von Betriebsmitteln usw.)

4.2.5 *Experimentelle Untersuchung und Evaluation*

Empirische Anwendung des Klassifikationsansatzes:
Für die Simulationen wurden digitale Netzmodelle erstellt, die eine Anwendung des Simulationsalgorithmus aus Abbildung 7 ermöglichen. Um eine effiziente Simulation vieler Lastflussszenarien durchzuführen, wurde eine selbstimplementierte Software auf Basis des Newton-Raphson Algorithmus genutzt. Dabei wurden die Resultate des Programmes anhand einiger synthetischer Musternetze gegen die Ergebnisse einer kommerziellen Software (PSS Sincal) validiert. Weiterhin wurde zur Bildung der komplexen Knotenadmittanzmatrix in den Gleichungen (8) und (9) für die empirischen Untersuchungen jeweils in allen Betriebsmitteln eine Leitertemperatur von 20°C unterstellt. Die Daten der realen Netze wurden von einem regionalen VNB in MS-Access bereitgestellt (Datenformat von PSS Sincal) und in die Software importiert. An dieser Stelle sei darauf hingewiesen, dass in der Simulation der Netze nur der aktuelle Schaltzustand sowie die vorhandene Netzstruktur bewertet werden. Manchmal kann allerdings eine Veränderung des Schaltzustandes die Aufnahmekapazität eines Netzes für DEA erhöhen (oder verringern). Weiterhin besteht die Möglichkeit, dass auch durch geringfügige Veränderungen der Netzstruktur – z. B. durch Hinzufügen oder Ersetzen eines kurzen Leitungsabschnittes – die aufnehmbare dezentrale Erzeugungsleistung erhöht wird. Eine Veränderung des Schaltzustandes oder der Netzstruktur kann allerdings nur behutsam unter Aufrechterhaltung der Selektivität im Netz, der Berücksichtigung der Struktur des vorgelagerten MS-Netzes sowie weiterer ökonomischer Randbedingungen erfolgen, und stellt damit eine planerische Aufgabe dar, die im Rahmen von Simulationen nur mit sehr hohem Aufwand berücksichtigt werden kann. Aus diesen Gründen bilden die simulationsbasierten Klassifikationsresultate immer nur eine Aussage über die Stärke des aktuellen Netzzustandes. Mögliche Verstärkungsmaßnahmen werden aus den oben genannte Gründen nicht berücksichtigt.

Klassifikationsresultate zweier Musternetze:
Vor der Evaluation aller 300 simulationsbasierten Klassifikationsresultate werden die Ergebnisse nun zunächst für die in Abschnitt 3.1.2 vorgestellten zwei Musternetze (Abbildung 1) diskutiert. Die für $N_q = 100$ resultierenden empirischen Verteilungsfunktionen und die parametrisierten Weibullmodelle sind in Abbildung 8 gezeigt.

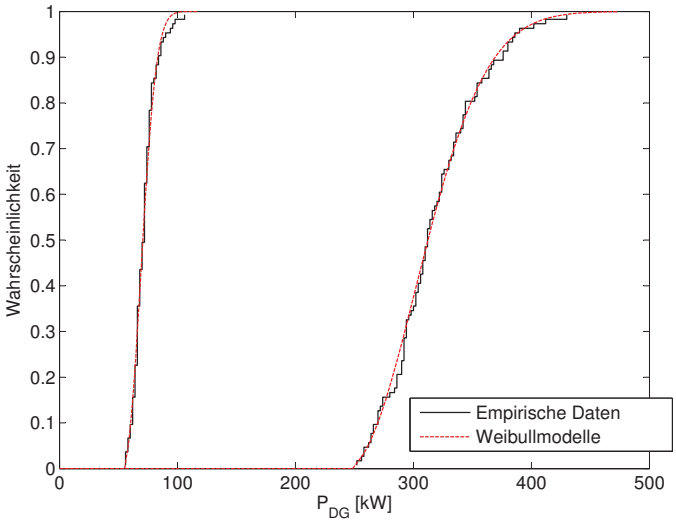


Abbildung 8: Simulationsergebnisse und probabilistische Modelle der DEA-Kapazitäten von Netz I (linke Kurven) und II (rechte Kurven).

Beide Modelle erfüllen den Anpassungstest auf dem 99%-Signifikanzniveau. Es ist direkt ersichtlich, dass sich die unterschiedlichen Strukturen und Merkmale (Abschnitt 3.1.2) der beiden Netze in den Simulationsergebnissen widerspiegeln. Die minimale Erzeugungsleistung, die normkonform in das Netz II integriert werden kann, ist ungefähr 4,3-mal höher als in Netz I. Auch die Gestalt der Weibullverteilungen ist offensichtlich verschieden. Während der Graph von Netz I nach dem Schwellwert stark ansteigt, weist der von Netz II eine wesentlich flachere Steigung auf. Die Ursache für die unterschiedliche Steigung ist in der unterschiedlichen Vermaschung und räumlichen Ausdehnung der beiden Netze zu finden. Das Netz II ist stark vermascht und seine Stationsbereiche weisen eine geringere räumliche Ausdehnung auf als die in Netz I. Das Netz I hat dagegen nur eine vermaschte Struktur im Zentrum seines Versorgungsbereiches. Außenliegende Endkunden werden wegen der höheren räumlichen Ausdehnung aus ökonomischen Gründen über lange Ausläuferleitungen versorgt. Weiterhin führen in beiden Netzen alle $N_q = 100$ zufälligen Anordnungen von DEA zu einer Verletzung von Kriterium 1, was zunächst auf eine ausreichende Dimensionierung der im Netz vorhandenen Betriebsmittel schließen lässt. Andersherum reicht – hinsichtlich des Spannungsniveaus

bei dezentraler Erzeugung – die Struktur beider Netze nicht aus, um die durchschnittliche installierte Nennleistung der Transformatoren im Netz zu erreichen. Unter diesem Aspekt sind Verstärkungsmaßnahmen durch den Verteilnetzbetreiber sinnvoll, um die Transformatoren im Netz auch im Erzeugungsfall auszulasten. Bezüglich der unterschiedlichen Gestalt beider Weibullverteilungen kann geschlussfolgert werden, dass die vermaschte und dichte Struktur von Netz II von der Anordnung der DEA im Netz stärker abhängig ist als die Struktur von Netz I. Diese Aussage wird durch eine genauere Betrachtung hinsichtlich der Grenzwertverletzungen gestützt. Das Kriterium 1 wird in Netz I häufig in den Netzbereichen mit langen Ausläuferleitungen und damit in denselben Netzbereichen zuerst verletzt. Im Gegensatz dazu führt die homogenere, vermaschte Struktur in Netz II in Abhängigkeit der Anordnung der DEA zu räumlich verteilteren Verletzungen von Kriterium 1.

Die Anwendung des Simulationsansatzes aus [6], in dem eine homogen auf die Knoten des Netzes verteilte Erzeugungsleistung bis zu einer Grenzwertverletzung hochskaliert wird, liefert nur einen Einzelwert für die DEA-Kapazität zurück. Dieser Wert beträgt 70kW für Netz I und 326kW für Netz II. Im Vergleich dazu resultieren aus den ermittelten Verteilungsfunktionen Erwartungswerte von 71kW für Netz I und 322kW für Netz II. Offensichtlich liefert also die Methodik aus [6] für die beiden Musternetze eine gute Näherung des Erwartungswertes der beiden probabibilistischen Modelle des DEA-Aufnahmevermögens. Dieses interessante Phänomen wird im Zuge der Evaluierung aller Klassifikationsergebnisse näher analysiert. Zuvor wird allerdings noch anhand der beiden Beispielnetze der Einfluss der in der stochastischen Simulation berücksichtigten Anzahl an DEA-Konfigurationen N_q auf die Parameter der Verteilungsfunktionen in einer Sensitivitätsanalyse bewertet. Dazu werden die Modellparameter für verschiedene Werte von N_q geschätzt. Um verlässliche Werte zu erhalten, sollte die Simulation für jedes N_q einige Male wiederholt werden.

In den Abbildungen 9 bis 11 sind die sich bei sukzessiver Erhöhung der berücksichtigten DEA-Konfigurationen einstellenden Verläufe der empirischen Erwartungswerte für die drei Modellparameter beider Musternetze gezeigt. Für jeden Wert von N_q wurden dabei die Simulationen 30-mal wiederholt. Um die Veränderung der Modellgenauigkeiten bewerten zu können, sind zusätzlich zu den Erwartungswerten die $\mu \pm \sigma$-Intervalle in den Diagrammen ausgewiesen. Darauf aufbauend werden Rückschlüsse auf die Modellgenauigkeit bei zunehmendem N_q aus den auftretenden Trends in den $\mu \pm \sigma$-Intervallen abgeleitet.

Abbildung 9 zeigt den Verlauf des Skalierungsparameters. In Lebensdaueranalysen beschreibt der Skalierungsparameter der Weibullverteilung die charakteristische Lebensdauer. Vor dem Hintergrund der hier durchgeführten Analysen kann der Skalierungsparameter daher als charakteristische durchschnittliche DEA-Kapazität des Netzes interpretiert werden. Gemäß dieser Interpretation ist es schlüssig, dass der Skalierungsparameter auf Grund der zur Integration von DEA robusteren Struktur für Netz II

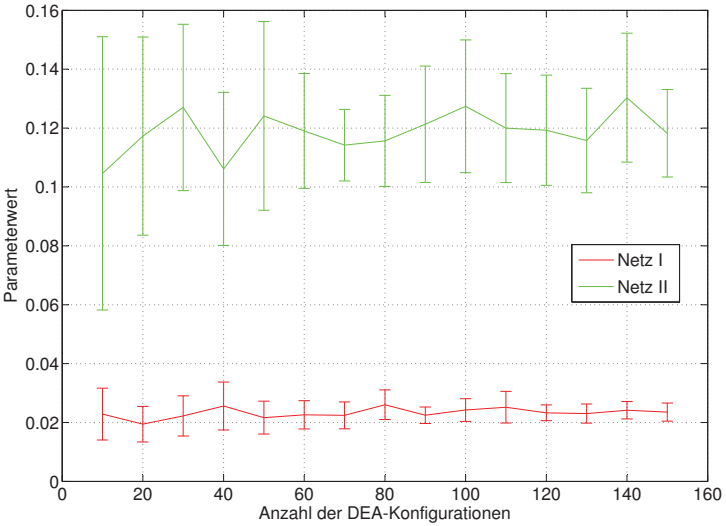


Abbildung 9: Sensitivität des Skalierungsparameters b für beide Musternetze.

den ungefähr 5- bis 6-fachen Wert im Vergleich zu Netz I annimmt. Weiterhin zeigen die absoluten Werte des Skalierungsparameters, dass der Schätzfehler für Netz I im Vergleich zu Netz II deutlich geringer ausfällt. Dies liegt darin begründet, dass in den Simulationen von Netz I eine Verletzung des 3%-Kriteriums häufig in den gleichen, die langen Ausläuerleitungen beinhaltenden Netzteilen auftritt. Im Gegensatz dazu sind diese Grenzwertverletzungen in den Simulationen von Netz II homogen im Netz verteilt, woraus in der Folge höhere Werte für σ resultieren. Die Trends in den $\mu \pm \sigma$-Intervallen zeigen für den Erwartungswert des Skalierungsparameters von Netz II eine deutlich abnehmende Tendenz mit Zunahme von N_q auf. Die Modellierungsgenauigkeit steigt bei Erhöhung von N_q auf einen Wert von 100 auf das ungefähr 2,2-fache an. Werden in den Simulationen stattdessen 150 DEA-Konfigurationen berücksichtigt, so verbessert sich die Modellierungsgenauigkeit lediglich noch um das 1,6-fache. Bei Erhöhung der DEA-Konfigurationen von 10 auf 100 in den Simulationen von Netz I sinkt das $\mu \pm \sigma$-Intervall auf das ungefähr 0,5-fache ab. Diese Abnahme entspricht ähnlich wie bei Netz II einer Verdopplung der Modellierungsgenauigkeit. Bei weiterer Erhöhung von N_q auf einen Wert von 150 resultiert eine zusätzliche Verbesserung um das 1,3-fache. Die relativen Abnahmen in den $\mu \pm \sigma$-Intervallen verhalten sich folglich in beiden Netzen ähnlich. Bei Erhöhung von N_q auf einen Wert von 100 stellt sich zunächst eine starke Verbesserung der Modellierungsgenauigkeit ein. Die zusätzliche Verringerung der $\mu \pm \sigma$-Intervalle bei weiterer Zunahme von N_q ist nicht mehr ähnlich stark ausgeprägt wie die zuvor beobachtete.

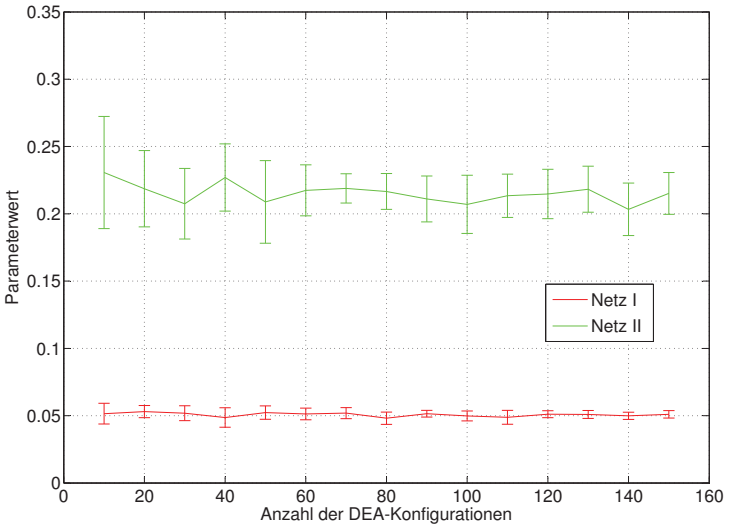


Abbildung 10: Sensitivität des Verschiebungsparameters v für beide Musternetze.

Die Sensitivität des Verschiebungsparameters ist in Abbildung 10 gezeigt. Dieser Parameter, der die minimal in ein Netz integrierbare DEA-Leistung beschreibt, fällt für Netz II ungefähr 4,3-mal höher aus als für Netz I. Analog zum Verhalten der Skalierungsparameter fällt der Schätzfehler beim Verschiebungsparameter im Vergleich zu Netz II für Netz I gering aus. Außerdem tritt bei einer Erhöhung auf $N_q = 100$ wiederum ungefähr eine Verdopplung der Modellierungsgenauigkeit in beiden Netzen auf, wohingegen aus einer weiteren Erhöhung von N_q auf einen Wert von 150 nicht mehr so signifikante Abnahmen der $\mu \pm \sigma$-Intervalle resultieren.

Die Formparameter der Weibullverteilungen liegen bei beiden Netzen in einer ähnlichen quantitativen Größenordnung (Abbildung 11). Dies ist symptomatisch für die Grenzwertverletzungen in dem stochastischen Simulationsalgorithmus, die im Ergebnis nicht zu grundsätzlich unterschiedlichen Formen der Weibullverteilungen führen (die hier durchgeführten Untersuchungen führten unter Berücksichtigung von $N_q = 100$ DEA-Konfigurationen über alle Netze z. B. auf durchschnittliche Werte von $\mu = 2{,}9$ und $\sigma = 1{,}2$). In der Sensitivitätsanalyse ist bei einer Erhöhung von N_q auf einen Wert von 100 für Netz I eine deutliche Erhöhung der Modellierungsgenauigkeit zu beobachten (Faktor 2,1). Im Gegensatz dazu tritt für Netz II lediglich eine geringfügige Verbesserung auf. Allerdings kann für beide Netze keine weitere deutliche Verbesserung bei zusätzlicher Erhöhung von N_q erreicht werden.

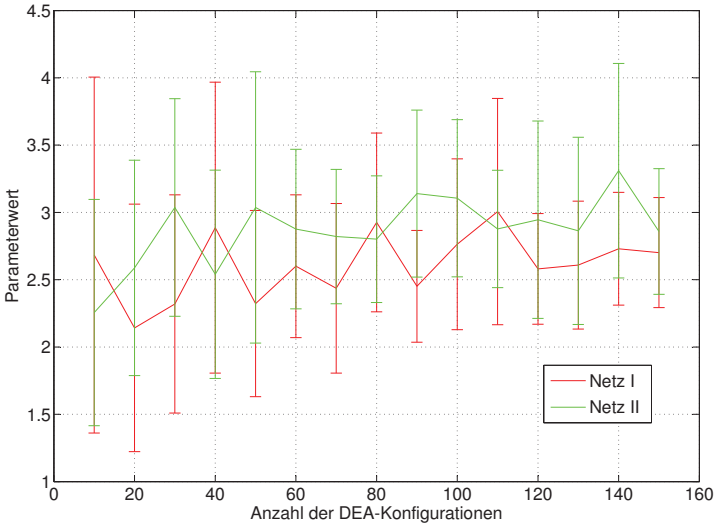


Abbildung 11: Sensitivität des Formparameters a für beide Musternetze.

Die Ergebnisse der Sensitivitätsanalysen für die beiden Musternetze lassen den Rückschluss zu, dass die Wahl von $N_q = 100$ zu einem guten Kompromiss zwischen Rechenaufwand und Modellierungsgenauigkeit führt. Bei weiterer Anwendung des Klassifikationsansatzes kann dieser Wert selbstverständlich an die jeweils spezifischen Untersuchungsziele angepasst werden.

KLASSENINDEX	$l_{lo,c}$ [KW]	$l_{up,c}$ [KW]
1	0	90
2	90	130
3	130	190
4	190	280
5	>280	

Tabelle 14: Empirisches Beispiel zur Festlegung der Klassengrenzen.

Nachdem die stochastischen Simulationsergebnisse für zwei Musternetze diskutiert wurden, können nun die Klassenzugehörigkeiten dieser Netze bestimmt werden. Die dazu gewählten Klassengrenzen sind in Tabelle 14 ausgewiesen (durchschnittliche dezentrale Erzeugungsleistung je Stationsbereich). Die Klassengrenzen wurden anhand der empirischen Ergebnisse des expertenbasierten Klassifikationsansatzes (Abschnitt 4.1.5) abgeleitet, so dass die aggregierten Verteilungen der Klassenzugehörigkeiten über die einzelnen Klassen für beide Klassifikationsansätze annähernd übereinstimmen. Dass es sich

bei den Experten um Mitarbeiter des VNB handelt, der die Netzdaten zur Verfügung gestellt hat, führt implizit zu einer Berücksichtigung der spezifischen betrieblichen und netzplanerischen Philosophie des VNB bei der Ableitung der Klassengrenzen.

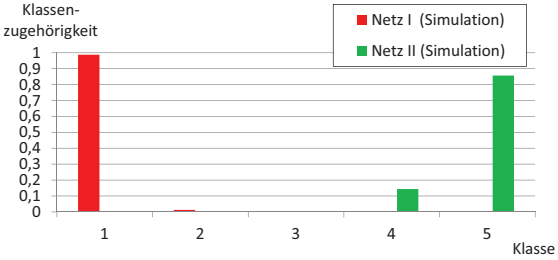


Abbildung 12: Verteilung der Klassenzugehörigkeiten für die Netze I und II.

Die aus den Simulationsergebnissen berechneten Klassifikationsergebnisse für die beiden Musternetze zeigt Abbildung 12. Netz I weist eine hohe Klassenzugehörigkeit nahe eins zur Klasse 1 auf, wodurch eine sehr schwache Netzstruktur hinsichtlich der Integration von DEA signalisiert wird. Die Wahrscheinlichkeit des Auftretens einer durchschnittlichen dezentralen Erzeugungsleistung in den Simulationen, die zwischen den Klassengrenzen von Klasse 2 liegt, ist nur vernachlässigbar höher als null. Netz II wird mit einer Klassenzugehörigkeit, die einen Wert von 0,8 übersteigt, der Klasse 5 zugeordnet, welches ein klares Indiz für die Stärke der vorhandenen Netzstruktur darstellt. Gleichzeitig tritt eine geringe Zugehörigkeit zu Klasse 4 auf. Leitet man nun aus diesen Ergebnissen scharfe Klassenzughörigkeiten durch Auswahl der Klasse mit der jeweils höchsten Klassenzugehörigkeit ab (gemäß dem Bayes'schen Risikominimierungsprinzip), erhält man für das Netz I eine Zuordnung zur Klasse 1 und für Netz II eine Klassifikation in Klasse 5. Damit korrespondieren auch die simulationsbasierten Klassifikationsergebnisse der Musternetze (Abbildung 1) gut mit den in Kapitel 3 erfolgten Interpretationen der Merkmale dieser Netze (Tabelle 1). Die diskutierten Klassifikationsergebnisse der Musternetze sind vor diesem Hintergrund als sinnvoll zu bewerten und motivieren eine empirische Evaluation der Resultate aller 300 klassifizierten Netze.

Evaluation der Ergebnisse:
Im nächsten Schritt wird nun eine Evaluation der Ergebnisse aller 300 untersuchten NS-Netze durchgeführt. Gemäß den Resultaten der exemplarischen Sensitivitätsanalyse wurde die Anzahl der im Rahmen der Simulationen berücksichtigten DEA-Konfigurationen für die gesamte empirische Untersuchung zu $N_q = 100$ gewählt. Die Klassengrenzen wurden entsprechend den Werten aus Tabelle 14 festgelegt. Abbildung 13 zeigt die sich auf dieser Basis ergebende aggregierte Klassenverteilung. In der Aggregation tragen die meisten der untersuchten NS-Netze zur Klassenzugehörigkeit der Klassen 3 bis 5 bei. Dieses Ergebnis zeigt in Analogie zu den Ergebnissen

der einzelnen Experten aus Abbildung 6, dass ein Großteil der bewerteten Netze eine grundsätzlich robuste Struktur zur Integration von DEA aufweist. Das Maximum der aggregierten Klassenverteilung tritt für Klasse 3 auf, wohingegen sich das Minimum bei Klasse 5 befindet. Insbesondere die geringe aggregierte Zugehörigkeit zu Klasse 5 unterstreicht, dass reale NS-Netze historisch gewachsene, von lokalen und geographischen Abhängigkeiten geprägte Strukturen aufweisen und zur ökomischen Versorgung von Kunden innerhalb der letzten Jahrzehnte weiterentwickelt worden sind. Folglich sind bei einem zukünftig hohem Penetrationsgrad der Netze durch DEA in vielen Fällen Netzverstärkungen zu erwarten.

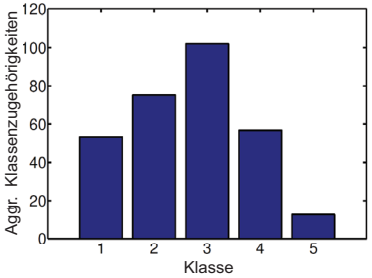


Abbildung 13: Aggregierte Verteilung der probabilistischen Klassenzugehörigkeiten.

Die Geeignetheit der Weibullverteilungen wird hinsichtlich einer 99%-Signifikanz bereits während der Klassifikation anhand des Goodness-of-fit-Tests überprüft. Bezüglich der statistischen Richtigkeit der Modelle sei an dieser Stelle darauf hingewiesen, dass 92% der Weibullverteilungen mit mindestens 99%-iger Signifikanz die zu Grunde liegenden empirischen Daten repräsentieren (für $N_q = 100$). Während der Analyse der beiden Musternetze wurde bereits beobachtet, dass der Simulationsansatz aus [6] jeweils eine gute Näherung der Erwartungswerte der estimierten Weibullmodelle liefert. Im weiterführenden Vergleich der empirischen Ergebnisse aller 300 Netze zeigt sich, dass sich bei rund 88% der Netze die mit dem Simulationsansatz von Kerber und Witzmann bestimmten Werte für die DEA-Kapazität in den $\mu \pm \sigma/4$-Intervallen der Weibullverteilungen befinden. Vergrößert man die Intervalle auf einen Wert von $\mu \pm \sigma/3$, so trifft diese Aussage sogar für 96% der Netze zu. Aus diesen Ergebnissen lässt sich schließen, dass der hier entwickelte Klassifikationsansatz eine direkte Erweiterung des von Kerber und Witzmann in [6] vorgestellten Simulationsansatzes darstellt. Eine weitere empirische Evaluation der Klassifikationsergebnisse erfolgt nun durch direkten Vergleich der Ergebnisse mit den Expertenaussagen.

4.3 VERGLEICH DER NETZSPEZIFISCHEN KLASSIFIKATIONSANSÄTZE

Zusätzlich zu den bisher erfolgten empirischen Bewertungen der Ergebnisse beider Klassifikationsansätze erfolgt nun eine direkte Gegenüberstellung der empirischen Daten beider Konzepte. Die Intention besteht darin, dass beide Klassifikationsansätze zu ähnlichen Ergebnissen führen sollten. Vor allem dürfen sich die Ergebnisse nicht grundsätzlich widersprechen. Weiterhin dient der Vergleich als weitere, die bisherigen Analysen ergänzende Evaluation der Messvalidität des expertenbasierten Klassifikationsansatzes anhand der Simulationsergebnisse, die zwar auch keine „ground truth“ darstellen, allerdings im Gegensatz zu Expertenaussagen keinen subjektiven Einflüssen unterliegen. Unterschiede in den Klassifkationsergebnissen können z. B. durch Unsicherheiten in den Expertenaussagen oder durch methodisch unterschiedliche Schwerpunkte (wie z. B. der Tatsache, dass in den Simulationen lediglich der aktuelle Netzzustand berücksichtigt wird oder mögliche unaufwändige Netzverstärkungsmaßnahmen bei der Klassifikation durch menschliche Experten berücksichtigt werden können) hervorgerufen sein.

		Simulationsansatz					
		1	2	3	4	5	$f_{r.}$
Experten	1	**17**	12	13	7	1	50
	2	9	**20**	26	17	2	74
	3	20	37	**38**	26	5	126
	4	5	9	13	**13**	6	46
	5	1	1	1	1	**0**	4
	$f_{.s}$	52	79	91	64	14	**300**

Tabelle 15: Gegenüberstellung der Klassifikationsergebnisse in einer Konfusionsmatrix

Zum Vergleich der Ergebnisse beider Klassifikationsansätze wird auch hier die κ_w-Statistik (Abschnitt 4.1.4) durchgeführt. Um mit beiden Ansätzen scharfe Klassenzugehörigkeiten zu erhalten wird gemäß dem Bayes'schen Risikominimierungsprinzip hier zunächst (der Einfachheit halber) jeweils ein Mehrheitsentscheid durchgeführt. Tabelle 15 zeigt eine Gegenüberstellung der Klassifikationsergebnisse in einer Konfusionsmatrix [13]. Die Ergebnisse stimmen demnach für 88 Netze überein (Summe der Einträge auf der Hauptdiagonalen) und unterscheiden sich für 240 Netze um maximal ± 1 Klasse (Summe der Einträge auf Haupt- und Nebendiagonalen). Die Anwendung der Gleichungen (3), (4) und (5) liefert unter Verwendung der linearen Gewichtung w_{rs} aus Tabelle 9 einen Wert von $\kappa_w = 0{,}108$ für den Vergleich der Expertenaussagen mit den Simulationsergebnissen. Die Klassifkationsergebnisse stimmen damit nicht nur rein zufällig überein. Um ergänzend zu untersuchen, ob diese Konkordanz auch von statistischer

Signifikanz ist, wird zunächst die Varianz von κ_w benötigt. Dies liegt darin begründet, dass

$$u = \frac{\kappa_w}{\sqrt{\mathrm{V}(\kappa_w)}} \tag{17}$$

näherungsweise normalverteilt ist [54]. Die Berechnung der Varianz von κ_w kann – unter Berücksichtigung der Gleichung (4) zur Bestimmung von $o_{r.}$ und $o_{.s}$ sowie der linearen Gewichtung w_{rs} aus Tabelle 9 – wie folgt bestimmt werden [54]:

$$\begin{aligned} \mathrm{V}(\kappa_w) = & \frac{1}{n \cdot (1-p_\epsilon)^2} \cdot \left(\sum_{r=1}^{C} \sum_{s=1}^{C} p_{r.} \cdot p_{.s} \right. \\ & \left. \cdot \left(w_{rs} - (\overline{w}_{r.} + \overline{w}_{.s})\right)^2 - p_\epsilon^2 \right), \end{aligned} \tag{18}$$

mit

$$\begin{aligned} p_{r.} &= \frac{o_{r.}}{n}, \\ p_{.s} &= \frac{o_{.s}}{n}, \\ p_\epsilon &= \sum_{r=1}^{C} p_{r.} \cdot p_{.r}, \\ \overline{w}_{r.} &= \sum_{s=1}^{C} w_{rs} \cdot p_{.s}, \\ \overline{w}_{.s} &= \sum_{r=1}^{C} w_{rs} \cdot p_{r.}. \end{aligned}$$

Zum Nachweis der statistischen Signifikanz der Konkordanz bei linearer Gewichtung w_{rs} wird ein einseitiger Hypothesentest mit folgender Formulierung der Nullhypothese H_0 für das 99%-Signifikanzniveau durchgeführt:

H_0: Die Resultate beider Klassifikationsansätze stimmen mit einer Signifikanz von 99% überein.

Der Berechnung der Varianz liefert zunächst $\mathrm{V}(\kappa_w) = 0{,}00168$. Darüber hinaus folgt aus Gleichung (17) $u = 2{,}99$. Dieser Wert muss nun mit dem kritischen Wert der Standardnormalverteilung $u_{krit.} = 2{,}33$ für das 99%-Signifikanzniveau verglichen werden. Der Vergleich führt zu folgendem Testergebnis (E) und der damit verbundenen Interpretation (I):

E: Der sich im Test ergebende Wert $u = 2{,}48$ ($> u_{krit.} = 2{,}33$) ist signifikant hinsichtlich des 99%-Niveaus.

I: Die Nullhypothese H_0 wird im Test bestätigt. Die Ergebnisse beider Klassifikationsergebnisse stimmen überein.

4.4 EINORDNUNG DER ERGEBNISSE

Im Ergebnis wurden in diesem Kapitel zwei netzspezifische Ansätze zur Klassifikation von NS-Netzen erarbeitet. Die unter Anwendung dieser Ansätze erhobenen empirischen Ergebnisse zum Aufnahmevermögen der untersuchten Netze für dezentrale Erzeugungsanlagen sind auf Grund des Stichprobenumfangs von 300 Netzen auf reale ländliche und vorstädtische Netze verallgemeinerbar. Streng genommen ist den Ergebnissen allerdings auch eine geringe Abhängigkeit von der Betriebsphilosophie (z. B. Schaltzustände der Netze, Experten sind bei gleichem VNB angestellt) immanent, der sowohl die Netzdaten, als auch die Experten für die Untersuchungen zur Verfügung gestellt hat. Die erarbeiteten Klassifikationsansätze werden in dieser Arbeit genutzt, um neben einer Bewertung von NS-Netzen auch Beispielergebnisse zum Training von SVM-Klassifikatoren zu erzeugen. Diese Beispieldaten legen in Kombination mit den Merkmalen aus Kapitel 3 die Grundlage für eine Implementierung von SVM-Klassifikatoren in Abschnitt 6.3.3. In diesem Zuge wird ebenfalls analysiert, mit welchen Merkmalen aus Kapitel 3 die besten Klassifikationsergebnisse erreicht werden. Da beide Klassifikationsansätze unterschiedliche Schwerpunkte in der Bewertung von NS-Netzen haben und die Nutzung von Expertenwissen aus vielen Gründen mehr oder weniger unsicherheitsbehaftet ist, werden in Kapitel 5 Kombinationsregeln untersucht, mit deren Hilfe die einzelnen Klassifikationseinschätzungen zu einer Gesamtaussage fusioniert werden können.

5

KOMBINATION UNSICHERER KLASSIFIKATIONSERGEBNISSE

Die Nutzung von Expertenwissen ist aus vielen Gründen (siehe Abschnitt 4.1.1) mehr oder weniger unsicherheitsbehaftet: Expertenwissen kann z. B. unpräzise, imperfekt oder fehlerhaft sein. Die Klassifikation von Objekten durch Experten führt daher häufig zu unterschiedlichen Ergebnissen. Daher sollte eine finale Klassifikationsentscheidung in einer Anwendung nicht anhand einer einzelnen Expertenaussage getroffen werden. Liegen mehrere unsicherheitsbehaftete Klassifikationseinschätzungen vor, können diese in geeigneter Weise kombiniert werden. Eine adäquate Fusion kann z. B. anhand einer Kombinationsregel vorgenommen werden. Das Ziel dieser Regeln liegt in der Prognose einer kombinierten Klasse, die der wahren (und in den meisten Anwendungen unbekannten) Klasse mit höherer Treffsicherheit entspricht, als die zu Grunde gelegten individuellen Klassifikationseinschätzungen.

Der Kombination von mehreren, potenziell falschen Klassifikationseinschätzungen für ein einzelnes Objekt mit einer Kombinationsregel ist dieses Kapitel gewidmet. Vor dem Hintergrund der Klassifikationsansätze aus Kapitel 4 wird dabei unterstellt, dass es sich um eine ordinale Klassenstruktur und nur wenige verfügbare Klassifikationseinschätzungen handelt. In diesem Zuge wird eine neue Kombinationsregel, die EIDMR, vorgestellt und gezeigt, dass diese einigen exisitierenden Kombinationsregeln überlegen ist, wenn zum einen die Klassen untereinander eine Ordnung aufweisen und zum anderen die Anzahl an zu kombinierenden Klassifikationseinschätzungen (z. B. Anzahl der Expertenaussagen) für ein einzelnes Objekt gering ist. Die EIDMR basiert auf einem k-nächste-Nachbarn-($k$nn)-Ansatz und Dirichlet-Verteilungen (z. B. Verteilungen zweiter Ordnung für eine Bayes'sche Schätzung der Parameter von Multinomialverteilungen). Sie erweitert eine existierende Imprecise-Dirichlet-Model-Regel (IDMR) durch die Berücksichtigung der Klassenordnung und die Verarbeitung von zusätzlichen, ebenfalls unsicheren Klassifikationseinschätzungen, die für ähnliche Objekte vorliegen (gemessen in einem adäquaten Featureraum mit einem geeigneten Ähnlichkeitsmaß). Die mit der Kombination verbundene Unsicherheit betrifft die Entscheidung für eine Klasse bei Anwendung einer Kombinationsregel. Diese kann bei der EIDMR mit Hilfe von zwei

Wahrscheinlichkeitsgrenzen ermittelt werden. Die Nutzung der daraus resultierenden Unsicherheitswerte ist optional. Sie können z. B. genutzt werden, um einen Vergleich dieser Unsicherheiten für verschiedene Objekte durchzuführen oder um die Kombinationsergebnisse weiteren Resultaten, z. B. von einer zweiten Expertengruppe, gegenüberzustellen. Wurden zusätzlich zu den zu kombinierenden Klassifkationseinschätzungen Schwierigkeitseinschätzungen erhoben (z. B. von Experten im Zuge der Klassifikation abgegeben, siehe Abschnitt 4.1), können diese ebenfalls optional bei Anwendung der EIDMR berücksichtigt werden. Dabei wird unterstellt, dass die Schwierigkeitseinschätzungen die Unsicherheit der Experten bei der Abgabe ihrer Klassifikationseinschätzung in Form einer Selbstbewertung widerspiegeln. Die Ergebnisse der EIDMR werden empirisch mit Ergebnissen dreier bereits existierender Kombinationsregeln verglichen, der IDMR, der Dempster-Shafer-Regel (DSR) und Murphy's Regel (MR).

Dieses Kapitel ist wie folgt untergliedert: In Abschnitt 5.1 wird zunächst der Stand des Wissens zu Kombinationsregeln überblicksartig dargestellt. Daraufhin gibt Abschnitt 5.2 eine grundlegende Einführung zum Imprecise Dirchlet Model (IDM) und stellt die Kombinationsregeln IDMR und EIDMR vor. Auf dieser Basis erfolgt in Abschnitt 5.3 ein experimenteller Vergleich dieser Kombinationsregeln mit der DSR und MR hinsichtlich ihrer Güte zu Detektion der wahren Klassen auf künstlichen Daten, deren Charakteristik und wahre Klassen bekannt sind. Zum Abschluss des Kapitels werden in Abschnitt 5.4 die Kombinationsregeln auf Daten von 300 NS-Netzen angewandt und zur Kombination der experten- und simulationsbasierten Klassifikationsergebnisse aus Kapitel 4 genutzt. In diesem Zuge werden auch die Ergebnisse der einzelnen Kombinationsregeln gegenübergestellt.

5.1 VORBEMERKUNGEN

Zur Behandlung von Unsicherheiten und Widersprüchen bei der Verarbeitung von Experteneinschätzungen wurden bereits verschiedenste Ansätze verfolgt. In [50] werden die Expertenaussagen z. B. als Menge von Wahrscheinlichkeitsverteilungen modelliert und über einen Aggregationsansatz zu einer einzigen Verteilungsfunktion fusioniert. Im Gegensatz dazu wird in [78] ein Modellierungsansatz zur Überführung von Expertenmeinungen in Fuzzymengen anhand eines Ähnlichkeitsmaßes vorgestellt. Martin und Osswald hingegen behandeln Einschätzungen von verschiedenen menschlichen Experten mit zwei aus der Evidenztheorie [79] abgeleiteten Ansätzen [80]. Zusätzlich zur Evidenztheorie werden auch Rough-Set-basierte Modelle zur Begegnung der Unsicherheit in Expertenaussagen eingesetzt [81]. Diese Arbeit beschränkt sich auf den Einsatz von Kombinationsregeln. Diesbezüglich werden in [82, 83] vier Anforderungen formuliert, die durch Andrade et al. [84] um eine fünfte Anforderung erweitert werden. Demnach muss eine Kombinationsregel folgende fünf grundlegende Eigenschaften aufweisen [84]:

1. Unabhängigkeit von der Reihenfolge der Kombination (Kommutativität und Assoziativität),

2. Verringerung der Unsicherheit des Kombinationsergebnisses mit steigender Anzahl von Informationsquellen,

3. konkordante Aussagen führen zu einer Erhöhung des Vertrauens in die Aussage,

4. widersprüchliche Aussage führen zu einer Verringerung des Vertrauens in die Aussage und

5. persistente Widersprüche werden im Kombinationsergebnis widergespiegelt.

Es exisitieren mittlerweile viele Kombinationsregeln zur Wissensfusion, die sehr vielfältig eingesetzt werden können. Die Fusion von Klassifikationseinschätzungen stellt nur eine von sehr vielen möglichen Anwendungen dar. Viele Kombinationsregeln finden ihren Ursprung in der Dempster-Shafer-Theorie (DST), z. B. die DSR [79] und MR [82]. Eine Studie zu DST-basierten Kombinationsregeln wurde von Guo Hua-wei durchgeführt [85]. Die meisten der DST-basierten Alternativen zur DSR berücksichtigen eine Umverteilung des Konfliktgrades. Das unintuitive Verhalten bei paradoxen Problemen [86], bei denen z. B. viele widersprüchliche und nur einige wenige übereinstimmende Expertenaussagen vorliegen, ist selten berücksichtigt [87]. Neben den DST-basierten Regeln existieren weitere Regeln, die z. B. aus der Dezert-Smarandache Theorie [88] abgeleitet sind und einem nicht-Bayes'schen Schlussfolgerungsansatz gleichkommen, wie z. B. die PCR5 [89]. Zusätzlich stellen Smarandache et al. einige Gegenbeispiele vor, in denen die DSR keine schlüssigen (oder gar keine) Ergebnisse liefert [90].

Zusammenfassend lässt sich feststellen, dass alle bekannten DST-basierten Regeln nicht die zusätzlich von Andrade et al. formulierte Anforderung an Kombinationsregeln erfüllen. Aus diesem Grund stellen Andrade et al. die IDMR [91, 92, 93] vor und zeigen, dass diese Regel alle dieser Anforderungen erfüllt. Dies gilt auch für die hier erarbeitete EIDMR, da es sich um eine Erweiterung der IDMR handelt. Dass die EIDMR ebenfalls die fünfte Anforderung erfüllt, wird im Folgenden nochmal explizit gezeigt (Abschnitt 5.2.2). Hinsichtlich der hier vorgenommenen Anwendung zur Kombination von Klassifikationseinschätzungen sei an dieser Stelle erwähnt, dass im Gegensatz zur EIDMR keine der genannten Regeln die Ordnung der Klassen sowie zusätzliche Klassifikationseinschätzungen zu ähnlichen Objekten berücksichtigt.

5.2 METHODISCHE GRUNDLAGEN ZU KOMBINATIONSREGELN

5.2.1 *Grundlagen des Imprecise Dirichlet Model*

Ausgangspunkt zur Motivation und Herleitung des IDM seien C gegenseitig exklusive Ereignisse, in diesem Fall C Klassen $c = 1, \ldots, C$. Die einzelnen Wahrscheinlichkeiten θ_c für die Wahl einer Klasse c bei einer Klassifikation seien in einem Vektor $\boldsymbol{\theta}$ zusammengefasst. Die numerische Ausprägung des Vektors $\boldsymbol{\theta}$ ist im Regelfall bei einer Klassifikation unbekannt. Geht man von einer N_E-maligen Klassifikation (z. B. N_E Expertenaussagen) eines Objektes aus, so lässt sich das Klassifikationsergebnis durch die Häufigkeiten des Auftretens der einzelnen Klassen ebenfalls in einem Vektor $\boldsymbol{n} = (n_1, \ldots, n_C)^{\mathrm{T}}$ mit $\sum_{c=1}^{C} n_c = N_E$ beschreiben. Die Wahrscheinlichkeit eines Klassifikationsergebnisses $\boldsymbol{n}$ ist dann durch die folgende Likelihood-Funktion bestimmbar [93]:

$$p(\boldsymbol{n}|\boldsymbol{\theta}) \;=\; \prod_{c=1}^{C} (\theta_c)^{n_c} \,. \tag{19}$$

Da der Vektor $\boldsymbol{\theta}$ im Regelfall unbekannt ist, wird in einem weiteren Schritt eine Verteilung erforderlich, mit der ein bestimmter Vektor $\boldsymbol{\theta}$ auftritt, der selbst als Zufallsvariable angenommen wird. Hierzu kann die Dirichlet-Verteilung genutzt werden, die in einer Bayes'schen Parameterschätzung als Prior-Verteilung dient. Sie lautet unter Berücksichtigung zweier Hyperparameter h und $\boldsymbol{t} = (t_1, \ldots, t_C)^{\mathrm{T}}$ [92]:

$$p(\boldsymbol{\theta}) \;=\; \frac{\Gamma(h)}{\prod\limits_{c=1}^{C} \Gamma(ht_c)} \cdot \prod_{c=1}^{C} (\theta_c)^{ht_c - 1} \,. \tag{20}$$

Es gilt $h > 0$, $0 < t_c < 1$ für $c = 1, \ldots, C$ sowie $\sum_{c=1}^{C} t_c = 1$. Γ stellt die Gammafunktion dar [94]. Der Vektor $\boldsymbol{t}$ spiegelt das Prior-Wissen über $\boldsymbol{\theta}$ wieder. Durch Kombination von Gleichung (19) und (20) gemäß $p(\boldsymbol{\theta}|\boldsymbol{n}) \;\propto\; p(\boldsymbol{n}|\boldsymbol{\theta}) \cdot p(\boldsymbol{\theta})$ erhält man folgende Verteilung zur Bestimmung der Posteriori-Wahrscheinlichkeit [92]:

$$p(\boldsymbol{\theta}|\boldsymbol{n}) \;=\; \frac{\Gamma\left(\sum\limits_{c=1}^{C} (n_c + ht_c)\right)}{\prod\limits_{c=1}^{C} \Gamma(n_c + ht_c)} \cdot \prod_{c=1}^{C} (\theta_c)^{n_c + ht_c - 1} \,. \tag{21}$$

Da die Dirichlet-Verteilung aus Gleichung (20) die konjugierte Verteilung der in Gleichung (19) verwendeten multinomialen Verteilung ist, resultiert zur Bestimmung der Posteriori-Verteilung wieder eine Dirichlet-Verteilung. Mit dem Hyperparameter h wird der Einfluss des Vektors $\boldsymbol{t}$ auf die Posterior-Wahrscheinlichkeit bestimmt [93]. Anhand der Erweiterung

$$n_c + ht_c - 1 \;=\; (N_E + h)\frac{n_c + ht_c}{N_E + h} - 1$$

kann gezeigt werden, dass die Hyperparameter der Posteriori-Verteilung, die im Folgenden als h^* und t_c^* bezeichnet werden, anhand der ursprünglichen Parameter aus Gleichung (20) bestimmt werden können [91]:

$$h^* = N_E + h, \tag{22}$$

$$t_c^* = \frac{n_c}{N_E + h} + \frac{h}{N_E + h} t_c. \tag{23}$$

In einer realen Klassifikation ist häufig kein Vorwissen über die Parameter h und $\boldsymbol{t}$ verfügbar. Mittels der Verteilung aus Gleichung (21) wird im IDM nun anstatt einer einzelnen Dirichlet-Verteilung eine Menge von Dirichlet-Verteilungen bei festem Hyperparameter h betrachtet [93]. Die Menge der Dirichlet-Verteilungen $(h, \boldsymbol{t})$ wird dabei so gewählt, dass weiterhin $\sum_{c=1}^{C} t_c = 1$ gilt [91]. Durch Maximierung bzw. Minimierung über t_c ergeben sich unter diesen Bedingungen die folgende Ober- bzw. Untergrenze für die Wahrscheinlichkeit, dass ein zu klassifizierendes Objekt einer Klasse c zugeordnet wird [92]. Dies bedeutet, dass für jeden Vektor $\boldsymbol{t}$ die korrespondierende Dirichletverteilung bestimmt wird [93]. Für den Fall das kein Vorwissen über $\boldsymbol{\theta}$ verfügbar ist, sind die Wahrscheinlichkeitsgrenzen auf Grund der Linearität der Gleichungen (22) und (23) durch die einseitigen Grenzübergänge $t_c \to 0$ (für die Untergrenze) und $t_c \to 1$ (für die Obergrenze) bestimmbar [91]:

$$\underline{p}(c|n_c, N_E) = \frac{n_c}{N_E + h}, \tag{24}$$

und

$$\overline{p}(c|n_c, N_E) = \frac{n_c + h}{N_E + h}. \tag{25}$$

In diesen Gleichungen für die Wahrscheinlichkeitsgrenzen wird c als Zufallsvariable für das Kombinationsergebnis betrachtet. Die Wahrscheinlichkeitsgrenzen können ähnlich zu den Vertrauens- und Plausibilitätsfunktionen der DST interpretiert werden. Ihre Differenz hängt nur von der Anzahl der Klassifikationseinschätzungen N_E ab und kann wiederum ähnlich zur DST als Unsicherheit des Kombinationsergebnisses gedeutet werden. Der Hyperparameter h hat einen konstanten Wert und sollte in einer Anwendung entweder auf 1 oder 2 gesetzt werden [91]. Er bestimmt, wie schnell die obere und untere Wahrscheinlichkeitsgrenze mit steigender Anzahl N_E an verfügbaren Klassifikationsaussagen konvergieren. Aus diesem Grund wurde h von Walley allgemein als Anzahl der Beobachtungen bezeichnet, die benötigt werden, um die Unsicherheit zu halbieren [91]. Bei Nutzung der Wahrscheinlichkeitsgrenzen muss in einer Anwendung der Vektor $\boldsymbol{t}$ nicht näher spezifiziert werden.

5.2.2 *Auf dem Imprecise Dirichlet Model basierende Kombinationsregeln*

Imprecise-Dirichlet-Model-Regel:
Gegeben seien N_E Aussagen c_j (c_j ist die j-te dieser Klassifikationseinschätzungen). In den Gleichungen (24) und (25) ist n_c die Anzahl der Klassifikationseinschätzungen, die ein gegebenes Objekt der Klasse c zuordnen. Für N_E Klassifikationseinschätzungen gilt $0 \leq n_c \leq N_E$. Dieser Ansatz berücksichtigt bisher keine mit der Klassifikation verbundene Angabe zur Sicherheit einer Klassifikationseinschätzung, z. B. basierend auf Schwierigkeitseinschätzungen von Experten. Um diese in der Kombination zu berücksichtigen, wird für jede Klassifikationseinschätzung ein Sicherheitsgewicht w_j ($w_j \geq 1$, $j = 1, \ldots, N_E$) eingeführt. Weiterhin wird eine Indikatorfunktion $I_{j,c}$ genutzt, die 1 ausgibt, wenn die Klassifikationseinschätzung j das zu klassifizierende Objekt der Klasse c zuordnet, und sonst den Wert 0 annimmt. Die Anzahl mit der ein Objekt der Klasse c zugeordnet wurde, wird auf dieser Grundlage in den Gleichungen (24) und (25) durch

$$n_c = \sum_{j=1}^{N_E} w_j \cdot I_{j,c} \tag{26}$$

ersetzt. Je höher das Gewicht w_j einer Klassifikationseinschätzung c_j gewählt wird, desto mehr Einfluss hat diese Aussage bei der Kombination. Mit Hilfe dieser Wahl von n_c können nun die Wahrscheinlichkeitsgrenzen aus den Gleichungen (24) and (25) ausgewertet werden.

Bei Anwendung der IDMR können die Wahrscheinlichkeitsgrenzen einer Klasse c auf verschiedene Arten genutzt werden. Wenn (wie in dieser Arbeit) eine scharfe Entscheidung für eine Klasse c getroffen werden soll, kann man vorsichtig vorgehen und die unteren Grenzen $\underline{p}(c|n_c, N_E)$ zur Entscheidung nutzen. Alternativ kann die Klasse c' mit

$$c' = \arg\max_c n_c, \tag{27}$$

gewählt werden. Die Klasse c', deren Anzahl an gewichteten Klassifikationseinschätzungen im Vergleich am höchsten ist, wird folglich als kombiniertes Ergebnis gewählt. Die Differenz der Wahrscheinlichkeitsgrenzen $\overline{p}(c|n_c, N_E) - \underline{p}(c|n_c, N_E)$ kann zur Beschreibung der Unsicherheit dieser Entscheidung ermittelt werden. Diese Werte können z. B. genutzt werden, um einen Vergleich der Unsicherheiten für verschiedene Objekte durchzuführen oder um die Kombinationsergebnisse weiteren Resultaten, z. B. von einer zweiten Expertengruppe, gegenüberzustellen.

Erweiterte Imprecise-Dirichlet-Model-Regel:
Es wird nun eine neue Kombinationsregel, die EIDMR, vorgestellt, die auf der gleichen Idee wie die IDMR basiert. Allerdings stellt die EIDMR eine Erweiterung der IDMR dar. Im Vergleich zur IDMR werden zusätzliche Informationen verarbeitet, die

zu ähnlichen Objekten vorhanden sind, um die Unsicherheit des Kombinationsergebnisses hinsichtlich der Entscheidung für eine Klasse zu reduzieren. „Ähnlich“ bedeutet in diesem Zusammenhang, dass die Datenpunkte der Objekte bezüglich zuvor ausgewählter Features ähnlich zueinander sind, wenn man diese mit einem adäquaten Ähnlichkeitsmaß (z. B. basierend auf einer Metrik) vergleicht. Zur Auswahl der zusätzlichen Objekte wird die IDMR um einen $k$nn-Ansatz erweitert, der im Featureraum Anwendung findet. Auf dieser Grundlage werden die Klassifikationseinschätzungen des zu klassifizierenden Objektes und die seiner k nächsten Nachbarn bei der Kombination berücksichtigt. Die Anwendung des $k$nn-Ansatzes stellt sicher, dass nur Objekte in der gewichteten Kombination berücksichtigt werden, die eine ähnliche Charakteristik zum zu klassifizierenden Objekt haben. Unter der Annahme, dass N_E Klassifikationseinschätzungen für das zu klassifizierende Objekt und seine k Nachbarn vorliegen, werden bei Anwendung der EIDMR $(k+1) \cdot N_E$ Klassifikationseinschätzungen berücksichtigt, um ein spezifisches Objekt zu bewerten. Bei der Kombination mit der EIDMR werden die einzelnen Aussagen nicht nur auf Basis der Schwierigkeitseinschätzungen wie oben beschrieben gewichtet, sondern auch in Abhängigkeit der Differenz zwischen den für das zu klassifizierende Objekt vorliegenden Klassifikationseinschätzungen und denen seiner k Nachbarn. Vor diesem Hintergrund wird mit der EIDMR ebenfalls die Ordnung der Klassen berücksichtigt: Größere Differenzen zwischen den Klassen führen zu einem geringeren Einfluss eines Objektes bei der Komination.

Um den Ansatz formal zu beschreiben, wird zunächst ein Index i für mehrere Variablen eingeführt, da nun nicht mehr nur ein Objekt bei der Kombination, sondern zusätzlich dessen k Nachbarn berücksichtigt werden. Dem zu klassifizierenden Objekt wird der Index 0, seinen Nachbarn die Indizes $i = 1, \ldots, k$ zugewiesen. Für ein Objekt i $(i = 0, \ldots, k)$ liegen dann N_E Klassifikationseinschätzungen zusammengefasst in einem Vektor $\boldsymbol{c}_i = (c_{i,1}, \ldots, c_{i,N_E})^{\mathrm{T}}$ mit dazu korrespondierenden Sicherheitsgewichten $\boldsymbol{w} = (w_{i,1}, \ldots, w_{i,N_E})^{\mathrm{T}}$ vor. Da die Klassen untereinander eine Ordnung aufweisen, kann die Differenz zwischen einem Vektor $\boldsymbol{c}_i$ und dem Vektor $\boldsymbol{c}_0$ anhand einer Standardmetrik, wie z. B. hier mit dem Euklidischen Abstand $\|\boldsymbol{c}_i - \boldsymbol{c}_0\|$, bestimmt werden. Auf dieser Grundlage kann für die Objekte eine Ähnlichkeitsgewichtung g_i (mit $i = 0, \ldots, k$) definiert werden, die diese Differenzen wie folgt berücksichtigt:

$$g_i \;=\; \frac{1}{k} \cdot \left(1 - \frac{\|\boldsymbol{c}_i - \boldsymbol{c}_0\|}{\sum\limits_{l=1}^{k} \|\boldsymbol{c}_l - \boldsymbol{c}_0\|} \right). \tag{28}$$

Um eine Division durch Null zu vermeiden, ist der sehr spezielle Fall, in dem alle Vektoren $\boldsymbol{c}_i$ gleich sind (in diesem Fall wird $g_0 = 1$ und $g_i = 0$ für $i = 1, \ldots, k$ gewählt) auszuschließen. Die so definierte Ähnlichkeitsgewichtung hat die Eigenschaft $\sum_{i=0}^{k} g_i = 1$.

Nach dieser Vorarbeit werden die Wahrscheinlichkeitsgrenzen der IDMR wie folgt erweitert:

$$\underline{p}(c|n_{c,0},\ldots,n_{c,k},N_E) \;=\; \frac{\sum\limits_{i=0}^{k} g_i \cdot n_{c,i}}{\sum\limits_{c=1}^{C}\sum\limits_{i=0}^{k} g_i \cdot n_{c,i} + h}, \tag{29}$$

$$\overline{p}(c|n_{c,0},\ldots,n_{c,k},N_E) \;=\; \frac{\sum\limits_{i=0}^{k} g_i \cdot n_{c,i} + h}{\sum\limits_{c=1}^{C}\sum\limits_{i=0}^{k} g_i \cdot n_{c,i} + h} \tag{30}$$

mit

$$n_{c,i} \;=\; \sum_{j=1}^{N_E} w_{i,j} \cdot I_{i,j,c}. \tag{31}$$

$I_{i,j,c}$ gibt nun an, ob ein Objekt i durch die Klassifikationseinschätzung j der Klasse c zugeordnet wurde (vgl. Gleichung (26) oben).

Analog zur Anwendung der IDMR können die Wahrscheinlichkeiten direkt genutzt werden oder es kann eine scharfe Entscheidung für eine Klasse c'mit

$$c' \;=\; \arg\max_{c} \sum_{i=0}^{k} g_i \cdot n_{c,i} \tag{32}$$

getroffen werden. Dies bedeutet, dass die Klasse c mit der höchsten Anzahl an gewichteten Klassifikationseinschätzungen gewählt wird. Die Unsicherheit, von der diese Entscheidung begleitet wird, kann mit $\overline{p}(c|n_{c,0},\ldots,n_{c,k},N_E) - \underline{p}(c|n_{c,0},\ldots,n_{c,k},N_E)$ bestimmt werden. Auch hier unterscheiden sich die Unsicherheiten bei der Entscheidung für eine Klasse eines einzelnen Objektes nicht. Allerdings ist zu erwarten, dass sich die Unsicherheiten durch den Einfluss der k zusätzlich in der Kombination berücksichtigten Objekte sowie der Ähnlichkeitsgewichte g_i für verschiedene Objekte im Vergleich zur IDMR häufiger unterscheiden.

Dass die EIDMR persistente Widersprüche im Kombinationsergebnis widerspiegelt und so die in Abschnitt 5.1 formulierte fünfte Eigenschaft erfüllt, kann in einer zu [84] analogen Vorgehensweise gezeigt werden. Damit persistente Widersprüche bei der Kombination erhalten bleiben muss folgende Bedingung erfüllt sein:

$$\lim_{\alpha\to\infty} \underline{p}(c|n_{c,0} = \alpha\cdot\widetilde{n}_{c,0},\ldots,n_{c,k} = \alpha\cdot\widetilde{n}_{c,k},\alpha\cdot N_E) \;=$$

$$\lim_{\alpha\to\infty} \overline{p}(c|n_{c,0} = \alpha\cdot\widetilde{n}_{c,0},\ldots,n_{c,k} = \alpha\cdot\widetilde{n}_{c,k},\alpha\cdot N_E) \;=\; \frac{\sum\limits_{i=0}^{k} g_i \cdot \widetilde{n}_{c,i}}{\sum\limits_{c=1}^{C}\sum\limits_{i=0}^{k} g_i \cdot \widetilde{n}_{c,i}}.$$

Durch entsprechendes Einsetzen in Gleichungen (29) und (30) sowie Kürzen beider Brüche mit α lässt sich leicht feststellen, dass diese Bedingung auch für die EIDMR erfüllt ist. Zur Verdeutlichung ist die Anwendung der EIDMR nochmal in Algorithmus 1 zusammengefasst.

Algorithmus 1 : EIDMR-Algorithmus für ein Objekt

Input : Menge der Objekte mit Featurevektoren, Klassifikationseinschätzungen zu den Objekten, Schwierigkeitseinschätzungen zu den Klassifikationseinschätzungen; Objekt 0 wird klassifiziert

1 Suche die k nächsten Nachbarn des Objektes 0 im Featureraum mit dem Euklidischen Abstand;
2 **for** $i = 0$ *to* k **do**
3 | Berechne die Ähnlichkeitsgewichte g_i;
4 **end**
5 **for** $c = 1$ *to* C **do**
6 | Berechne die Wahrscheinlichkeitsgrenze $\underline{p}(c|n_{c,0}, \ldots, n_{c,k}, N_E)$;
7 | Berechne die Wahrscheinlichkeitsgrenze $\overline{p}(c|n_{c,0}, \ldots, n_{c,k}, N_E)$;
8 **end**
9 Wähle Klasse c' mit $c' = \arg\max_c \sum_{i=0}^{k} g_i \cdot n_{c,i}$;

Output : Klasse c' und obere und untere Wahrscheinlichkeitsgrenze $\overline{p}$ und $\underline{p}$

5.2.3 *Illustratives Beispiel zur Anwendung der Kombinationsregeln*

Anwendung und Unterschiede der vorgestellten Kombinationsregeln werden nun anhand eines einfachen Beispiels verdeutlicht. Zu diesem Zweck ist in Abbildung 14 eine beispielhafte Anordnung von drei verschiedenen Datenpunkten in einem Featureraum dargestellt, die jeweils eines von drei Objekten charakterisieren. Dem gezeigten Beispiel liegt für jedes Objekt eine fiktive Klassifikationseinschätzung von zwei Experten auf der Basis von drei angeordneten Klassen (1, 2 und 3) zu Grunde. Zusätzlich wird angenommen, dass zu jeder Klassifikationseinschätzung eine fiktive Schwierigkeitseinschätzung basierend auf drei Schwierigkeitsgraden vorliegt (entweder „leicht“ mit Gewicht $w_{i,j} =$ 1,5, „mittel“ mit Gewicht $w_{i,j} = 1{,}25$ oder „schwer“ mit Gewicht $w_{i,j} = 1{,}0$). In Abbildung 14 sind die Klassifikationseinschätzungen mit den korrespondierenden Sicherheitsgewichten bereits in den Vektoren $\boldsymbol{c}_i$ und $\boldsymbol{w}_i$ $(i = 0, \ldots, 2)$ zusammengefasst. Es ist ersichtlich, dass die Sicherheitsgewichte der Klassifkationseinschätzungen für das zu untersuchende Objekt (rot) geringer sind als für dessen zwei nächste Nachbarn (grün). Diese Anordnung simuliert den Fall, in dem ein Objekt schwer zu klassifzieren ist, jedoch zwei ähnliche Objekte exisitieren, für die eine Bewertung leichter erscheint.

Anhand dieses Beispiels werden nun die IDM-basierten Kombinationsregeln auf das Beispiel angewandt. Für die IDMR resultieren $n_1 = 1$, $n_2 = 1$ und $n_3 = 0$. Demnach

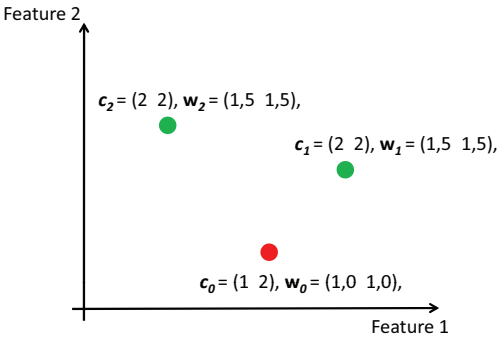


Abbildung 14: Beispielhafte Anordnung eines Datenpunktes (rot) und seiner $k = 2$ nächsten Nachbarn (grün) im Featureraum.

kann mit der IDMR keine eindeutige Entscheidung für eine Klasse getroffen werden. Die Werte für die untere und obere Wahrscheinlichkeitsgrenze (mit h=1) resultieren zu $\underline{p}(1) = \underline{p}(2) = 1/3$, $\overline{p}(1) = \overline{p}(2) = 2/3$ und $\underline{p}(3) = 0$, $\overline{p}(3) = 1/3$. Die Unsicherheit bezüglich der Entscheidung für eine Klasse beträgt mit der IDMR in diesem Beispiel folglich jeweils $1/3$.

Zur Anwendung der EIDMR werden zunächst die drei Distanzen bestimmt. Zu diesem Zweck wird hier der Euklidische Abstand verwendet. Dies führt zu $\|\boldsymbol{c}_0 - \boldsymbol{c}_0\| = 0$, $\|\boldsymbol{c}_1 - \boldsymbol{c}_0\| = 1$, und $\|\boldsymbol{c}_2 - \boldsymbol{c}_0\| = 1$. Auf dieser Basis kann im nächsten Schritt die Ähnlichkeitsgewichtung berechnet werden:

$$\begin{aligned} g_0 &= \frac{1}{2}(1-0) = \frac{1}{2}, \\ g_1 &= \frac{1}{2}(1-\frac{1}{2}) = \frac{1}{4}, \\ g_2 &= \frac{1}{2}(1-\frac{1}{2}) = \frac{1}{4}. \end{aligned}$$

Mit diesen Ergebnissen werden dann die $n_{c,i}$ mit Gleichung (31) bestimmt:

$$\begin{aligned} n_{1,0} &= 1, \\ n_{2,0} &= 1, \\ n_{2,1} &= 3, \\ n_{3,2} &= 3. \end{aligned} \tag{33}$$

Alle weiteren $n_{c,i}$ ergeben sich im Beispiel zu null. Diese Werte führen zur folgenden Klassenentscheidung:

$$\arg\max_c \{\frac{1}{2}, 2, 0\} = 2. \tag{34}$$

Folglich führt die in den zwei nächsten Nachbarn enthaltene Zusatzinformation zu der eindeutigen Entscheidung, dass Klasse 2 als kombiniertes Klassifikationsergebnis gewählt wird. Die jeweiligen Wahrscheinlichkeitsgrenzen ergeben sich für die EIDMR zu $\underline{p}(1) = 1/7$, $\overline{p}(1) = 3/7$, $\underline{p}(2) = 4/7$, $\overline{p}(2) = 6/7$ und $\underline{p}(3) = 0$, $\overline{p}(3) = 2/7$. Die Unsicherheit der Entscheidung für eine Klasse beträgt jeweils $2/7$.

Zusammenfassend führt die Anwendung der EIDMR in diesem einfachen Beispiel zu einer eindeutigeren Entscheidung mit reduzierter Unsicherheit.

5.3 EXPERIMENTELLER VERGLEICH DER KOMBINATIONSREGELN AUF KÜNSTLICHEN DATEN

Nachdem zunächst ein anschauliches Beispiel analysiert worden ist, folgt in diesem Abschnitt ein empirischer Vergleich der in Abschnitt 5.2.2 vorgestellten Kombinationsregeln anhand von künstlichen Daten. Da die Daten künstlich erzeugt wurden, sind die wahren Klassen für alle Datenpunkte bekannt. Dies erlaubt für jede Kombinationsregel eine direkte empirische Analyse der Kombinationsgüte. Zunächst wird jedoch in Abschnitt 5.3.1 die Vorgehensweise zur Erzeugung der künstlichen Daten erläutert. Im Anschluss daran wird in Abschnitt 5.3.2 das Vorgehen zur Erzeugung künstlicher Expertenaussagen für die Datenpunkte dargelegt. Zum Abschluss erfolgt in Abschnitt 5.3.3 eine empirische Untersuchung und Gegenüberstellung der Güte der Kombinationsregeln auf den künstlichen Daten.

5.3.1 *Generierung künstlicher Daten*

Zur Analyse der vier Kombinationsregeln wurden drei künstliche Datensätze mit jeweils 1000 Datenpunkten und fünf ordinalen Klassen erzeugt. Der Featureraum wurde dabei bewusst auf zwei Dimensionen eingeschränkt, um die Verteilung der Datenpunkte und Klassen darstellen zu können. Die Verteilung der Daten im Featureraum wurde jeweils anhand eines Gauß'schen Mischmodells (GMM) mit fünf Komponenten erzeugt. Eine Komponente im Modell repräsentiert jeweils eine der fünf wahren Klassen. Die Mischkoeffizienten des GMM wurden jeweils auf einen Wert von 0,2 gesetzt, um eine näherungsweise gleiche Verteilung der Klassen in den Datensätzen zu erhalten. Jede der Klassen tritt damit ca. 200-mal in einem Datensatz auf. Die zur Erzeugung der Datensätze verwendeten GMM unterscheiden sich lediglich in der Lage der Erwartungswerte ihrer Komponenten im Featureraum. Dabei wurde die Distanz zwischen den Erwartungswerten (Mittelwerten) sukzessive verringert. Im ersten Datensatz (Datensatz I) treten damit die größten Abstände und im dritten Datensatz (Datensatz III) die ge-

ringsten Abstände zwischen den Erwartungswerten auf. Als Konsequenz daraus ist die Separierung der Klassen im Datensatz I im Vergleich zu den anderen beiden Datensätzen am besten möglich. Im Gegensatz dazu überlappen im Datensatz III die fünf Komponenten des GMM, die den verschiedenen Klassen zugeordnet werden, am stärksten. Durch die Erzeugung von Datensätzen mit einer jeweils unterschiedlich hohen Überlappung der Komponenten des GMM soll der Einfluss einer zunehmend imperfekten Charakterisierung der Objekte durch die gewählten Features berücksichtigt werden. Der Grund dafür ist, dass in einer realen Anwendung die Separierbarkeit der Klassen von den zur Beschreibung der Objekte genutzten Features abhängig ist. Um die Ergebnisse der Datenerzeugung zu verdeutlichen, sind in Abbildung 15 die Datensätze I (geringste Überlappung) und III (höchste Überlappung) gegenübergestellt. Um die Separierbarkeit der Klassen für verschiedene Datensätze zu vergleichen, kann z. B. der Davies-Bouldin-Index DB genutzt werden. Je besser die Separierbarkeit der Klassen im Datensatz ist, desto geringere Werte nimmt dieser Index an. Datensatz I weist einen Davies-Bouldin-Index von $\mathrm{DB}_I = 0{,}43$ auf. Im Gegensatz dazu führt die Datenverteilung im Datensatz III zu $\mathrm{DB}_{III} = 0{,}94$ und liegt damit deutlich über dem Wert von Datensatz I. Für Datensatz II ergibt sich im Vergleich zu den anderen beiden Datensätzen entsprechend ein mittlerer Wert von $\mathrm{DB}_{II} = 0{,}63$.

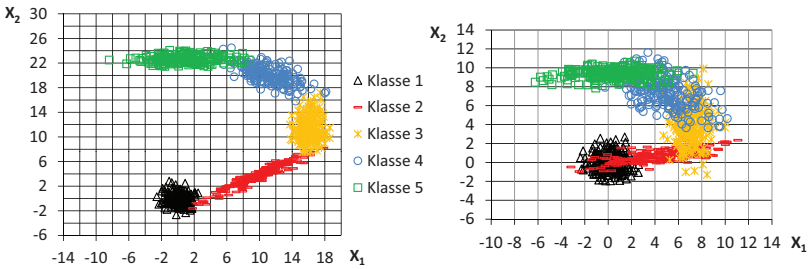


Abbildung 15: Darstellung der mit unterschiedlichen GMM erzeugten Datensätze, Datensatz I (links) und Datensatz III (rechts).

5.3.2 *Generierung künstlicher Expertenaussagen*

Nach Erzeugung der Datensätze wurden 30 künstliche und (mehr oder weniger) unsichere Expertenaussagen für jeden Datenpunkt erzeugt. Die Bildung der Klassifikationseinschätzungen erfolgte durch Modifikation der wahren Klassen gemäß der in Abschnitt 4.1.1 identifizierten Einflüsse. Jeder dieser Einflüsse wurde separat modelliert, so dass seine Ausprägung für jeden künstlichen Experten zufällig in unterschiedlicher Ausprägung berücksichtigt werden konnte. Im Einzelnen wurde jedem Experten zunächst eine individuelle Erfahrung (die Tagesform wurde nicht separat modelliert) zugeord-

net. Auf dieser Basis wurden – zum Beispiel aus Abschnitt 5.2.3 analoge – Schwierigkeitseinschätzungen der Experten für jeden Datenpunkt anhand von drei angeordneten Schwierigkeitsgraden erzeugt (entweder „leicht“, „mittel“ oder „schwer“). Die einem Experten zugeordnete Erfahrung beeinflusste dabei die einzelnen Wahrscheinlichkeiten für die Wahl einer dieser Schwierigkeitsgrade für einen Datenpunkt. Dies führt in der Simulation beispielsweise dazu, dass ein Experte mit hoher Erfahrung mehr (weniger) Datenpunkten den Schwierigkeitsgrad „leicht“ („schwer“) zuordnet, als ein Experte mit geringerer Erfahrung. Zusätzlich zur Erfahrung wurde den Experten ebenfalls eine individuelle Auffassung von Strenge sowie eine individuelle Neigung, keine extremen Beurteilungen vorzunehmen, zugewiesen. Für jeden Experten wurden dabei drei unterschiedlich starke Ausprägungen dieser subjektiven Einflüsse modelliert, um bei der Generierung der Klassifikationseinschätzungen die den Datenpunkten von diesem Experten zugeordneten Schwierigkeitsgrade berücksichtigen zu können. Bei der Generierung einer Expertenaussage für einen Datenpunkt mit dem Schwierigkeitsgrad „leicht“ („schwer“) wurde z. B. entsprechend die jeweils geringste (stärkste) Ausprägung verwendet. Die auf diese Weise modellierten Einflüsse führen in Summe zu einer bestimmten Wahrscheinlichkeit, dass ein künstlicher Experte einen Datenpunkt in eine im Vergleich zur wahren Klasse verschiedene Klasse einordnet. Ein Experte mit einer ausgeprägten Strenge wird z. B. eine hohe Wahrscheinlichkeit aufweisen, eine Klasse zu wählen, die niedriger als die wahre Klasse ist.

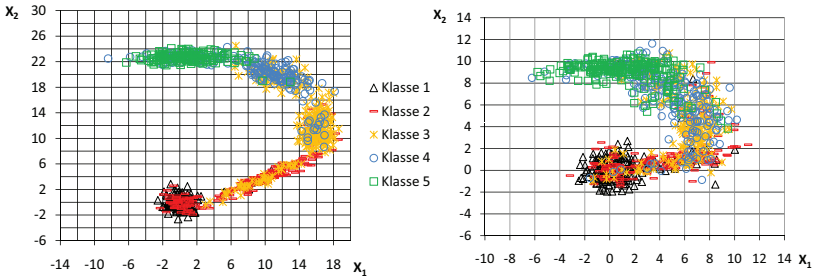


Abbildung 16: Darstellung künstlich generierter Klassifikationseinschätzungen für Datensatz I (links) und III (rechts).

Die Generierung der Klassifikationseinschätzungen berücksichtigt keinen Einfluss der Anordnung der Datenpunkte im Featureraum auf die Entscheidung eines Experten. Die wahren Klassen werden hier nur entsprechend der modellierten subjektiven Einflüsse auf eine Expertenentscheidung (Abschnitt 4.1.1) modifiziert. Diese Einschränkung basiert auf der Tatsache, dass es selbst für Anwendungsexperten schwierig ist, hochdimensionale und komplexe Aufgabenstellungen zu bearbeiten (siehe z. B. [53]). Vor diesem Hintergrund wird hier angenommen, dass ein Experte die Anordnung aller Datenpunkte im Featureraum nicht direkt berücksichtigen wird, während er die Datenpunkte nacheinander klassifiziert. In der Realität wird allerdings ein (mehr oder weniger

stark ausgeprägter) impliziter Einfluss von der Datenverteilung im Featureraum auf die Entscheidung eines Experten auftreten.

Um das Ergebnis der Modifikation der wahren Klassen zu veranschaulichen, sind in Abbildung 16 künstlich generierte Klassifikationseinschätzungen für die Datensätze I und III gezeigt. Die Modifikationen sind gegenüber den ursprünglichen Datensätzen (Abbildung 15) deutlich erkennbar. Obwohl die Datenstruktur noch erkennbar bleibt, treten z. B. bei den Datenpunkten, die einer Klasse zugeordnet sind, deutlich mehr Ausreißer auf. Um das Ausmaß der Veränderung zum ursprünglichen Datensatz I aufzuzeigen, kann ebenfalls der Davies-Bouldin-Index genutzt werden. Der ursprüngliche Wert von $\mathrm{DB}_{III} = 0{,}94$ erhöht sich durch das simulierte Klassifikationsverhalten des künstlichen Experten z. B. für den Datensatz III auf einen Wert von 1,51.

5.3.3 *Eigenschaften der Kombinationsregeln*

Die vier Kombinationsregeln EIDMR, IDMR, DSR und MR wurden auf die künstlichen Daten angewendet, um die generierten Expertenaussagen zu kombinieren. Vor Anwendung der EIDMR wurden die Featurewerte z-transformiert. Dies ist wichtig, um eine Dominanz von Features mit großen Wertebereichen gegenüber Features mit kleinen Werten in der Berechnung der Distanzen im Featureraum zu vermeiden [95]. Die Featuretransformation wird nur bei Anwendung der EIDMR benötigt, da die übrigen drei Kombinationsregeln keine Klassifikationseinschätzungen zu benachbarten Datenpunkten bei der Kombination berücksichtigen. Zur Analyse des Einflusses der Expertenanzahl auf die Güte der Kombinationsregeln, wurde die Anzahl der bei der Kombination berücksichtigten Expertenaussagen Schritt für Schritt auf 30 erhöht. Damit die Ergebnisse vor dem Hintergrund der künstlichen Generierung von Expertenaussagen statistisch fundiert ausgewertet werden können, wurde die Modifikation der wahren Klassen und die Kombination der Expertenaussagen für jeden Datensatz zehn Mal wiederholt. Dass die wahren Klassen der Datensätze bekannt sind, ermöglicht für jede Kombinationsregel eine direkte Untersuchung der Kombinationsgüte hinsichtlich der Detektion dieser Klassen. Die Bewertung der Kombinationsgüte der einzelnen Regeln erfolgt anhand der Interraterkonkordanz zwischen den Kombinationsergebnissen und den wahren Klassen unter Verwendung von Cohen's κ_w (Abschnitt 4.1.4).

Zur Anwendung der IDMR und EIDMR sind die generierten Schwierigkeitseinschätzungen der Experten in eine Sicherheitsgewichtung $w_{i,j}$ ($w_{i,j} > 0$) überführt worden, um auf Basis seiner Schwierigkeitseinschätzung die Sicherheit eines Experten bei der Zuordnung einer Klasse zu einem Datenpunkt bei der Kombination zu berücksichtigen. Dazu wurden in Abstimmung mit den in dieser Arbeit befragten Experten die numerischen Werte 1,00, 1,25 und 1,50 genutzt (siehe auch Abschnitt 5.2.3). Die Klassifikationseinschätzung eines Experten, der für einen Datenpunkt die Schwierigkeit der Klassifikation

als „mittel“ einschätzt, wird z. B. mit $w_{i,j} = 1{,}25$ bei der Kombination gewichtet. Die Anwendung der DSR und MR wurde unter Verwendung von Simple Support Functions (vgl. [84]) vorgenommen. Die Gewichtung musste dazu durch Multiplikation mit z. B. 0,6 in ein geeignetes Intervall überführt werden (z. B. 0,60, 0,75 und 0,90).

In einem ersten Experiment wurde der Einfluss der Datenverteilung im Featureraum auf die Kombinationsgüte anhand der drei vorbereiteten Datensätze untersucht. Zusätzlich wurde bei Anwendung der EIDMR der Wert von k variiert, um den Einfluss der berücksichtigten, benachbarten Datenpunkte zu analysieren.

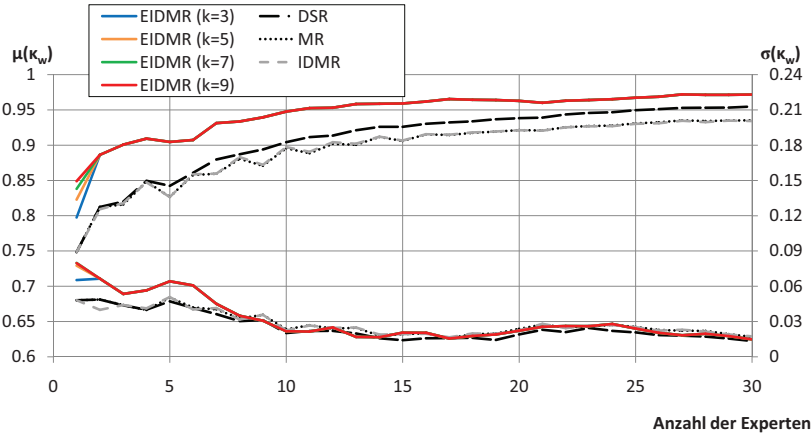


Abbildung 17: Verlauf der Kombinationsgüten für den Datensatz I.

In Abbildung 17 sind die sich in der Simulation ergebenden Verläufe der Mittelwerte $\mu(\kappa_w)$ und Standardabweichungen $\sigma(\kappa_w)$ (zehn Wiederholungen, siehe oben) für den Datensatz I dargestellt. Dieser Datensatz weist die in der Untersuchung geringste Überlappung der fünf Komponenten des GMM im Featureraum auf, die den verschiedenen Klassen zugeordnet sind. Es zeigt sich, dass die Verwendung der EIDMR für jede hier untersuchte Anzahl an Experten und jeden untersuchten Wert von k vorteilhaft ist. Dies ist durch die geringe Überlappung der Klassenverteilungen begründet, wodurch die Berücksichtigung benachbarter Datenpunkte eine deutliche Erhöhung der Kombinationsgüten verursacht. Die Entscheidung auf Basis der Klassifikationseinschätzungen der $k + 1$ Datenpunkte führt häufiger zum richtigen Kombinationsergebnis als bei Anwendung der übrigen Kombinationsregeln. Durch die geringe Überlappung der Komponenten des GMM beeinflusst der Wert von k die Kombinationsgüte ab einer Anzahl von zwei Experten nicht mehr. Allerdings treten gegenüber den anderen Kombinationsregeln höhere Standardabweichungen auf, was auf den unterschiedlichen Einfluss der Klassifikationseinschätzungen der k benachbarten Datenpunkte zurückzuführen ist.

Weiterhin zeigen die Ergebnisse einen sehr ähnlichen Verlauf der Kombinationsgüten zwischen der MR und IDMR. Abweichungen zwischen beiden Verläufen treten nur in geringem Maße auf. Dies ist durch das Vorgehen bei der MR begründet, bei der zunächst eine Durchschnitts-Belief-Funktion gebildet und auf diese anschließend $(N_E - 1)$-mal die DSR angewendet wird, woraus sehr ähnliche Ergebnisse verglichen mit der IDMR resultieren.

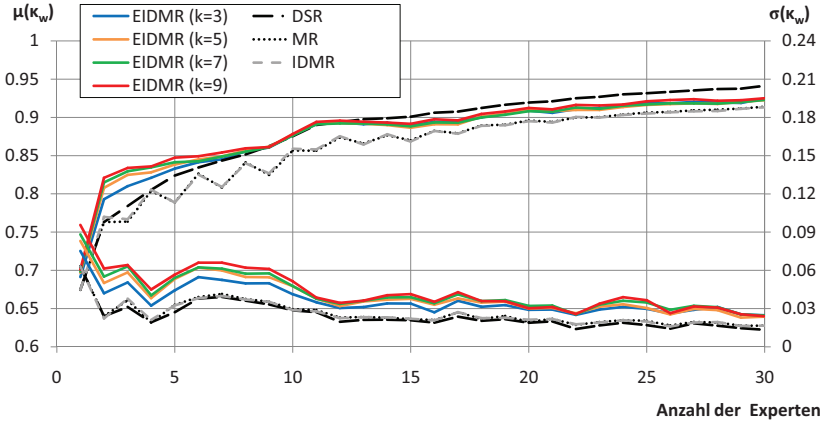


Abbildung 18: Verlauf der Kombinationsgüten für den Datensatz II.

Die Simulationsergebnisse für den Datensatz II sind in Abbildung 18 zu sehen. Im Vergleich zu den Datensätzen I und III weist dieser Datensatz eine mittlere Überlappung der Komponenten des GMM auf. Vergleicht man die durchschnittlichen Kombinationsgüten der DSR, MR und IDMR, so sind geringe Unterschiede in den Verläufen erkennbar, falls die Anzahl der Experten kleiner als fünf ist. Analog zum Datensatz I nehmen die Kombinationsgüten der MR und IDMR sehr ähnliche Verläufe an. Die höchste Kombinationsgüte wird bei Anwendung der EIDMR für $k = 9$ erzielt. Auf Grund der zusätzlichen Information, die durch den $k$nn-Ansatz verarbeitet wird, liefert die EIDMR deutlich höhere Kombinationsgüten als alle übrigen Kombinationsregeln, wenn die Anzahl der Experten geringer als neun ist. Allerdings sind die Standardabweichungen ebenfalls höher. Mit darüber hinaus steigender Expertenzahl (Anzahl an Experten > 12) übertrifft die DSR alle anderen Kombinationsregeln in der Detektion der wahren Klassen.

Abbildung 19 zeigt die Simulationsergebnisse für den Datensatz III mit der höchsten Überlappung der Komponenten des GMM. In den Fällen mit einer geringen Anzahl von Expertenaussagen (z. B. Anzahl der Experten ≤ 3) liefert die EIDMR im Vergleich zu allen anderen Regeln die durchschnittlich höchsten Kombinationsgüten hinsichtlich der Werte von $\mu(\kappa_w)$. Im Vergleich zu den anderen Datensätzen hat die Variation

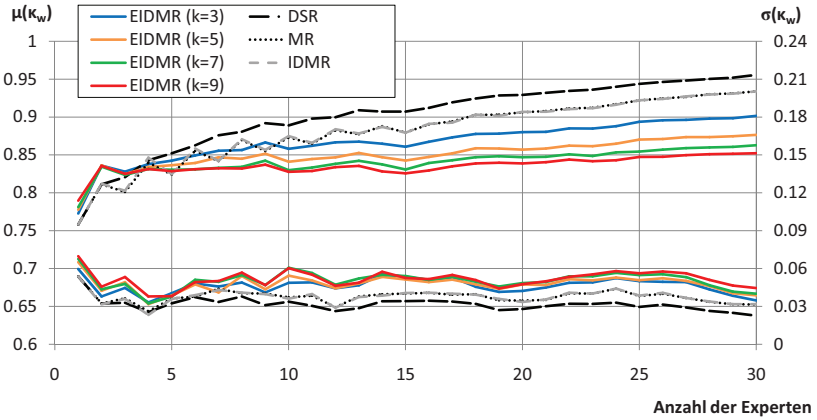


Abbildung 19: Verlauf der Kombinationsgüten für den Datensatz III.

von k den größten Einfluss auf die Kombinationsgüte. Die höchste Kombinationsgüte wird für $k = 3$ erreicht. Bei zunehmender Anzahl von Experten (Anzahl der Experten > 3) nimmt der Mehrwert durch die Berücksichtigung von benachbarten Datenpunkten ab. Der Einfluss führt zu deutlich geringeren Kombinationsgüten verglichen mit den Ergebnissen der übrigen Regeln. Beide beschriebenen Effekte sind durch die hohe Überlappung der Komponenten des GMM begründet, die bei einer hohen Anzahl unterschiedlicher Klassifikationseinschätzungen für einen Datenpunkt häufiger zur Entscheidung für eine falsche Klasse führt. Die Berücksichtigung benachbarter Datenpunkte führt bei verhältnismäßig schlechter Separierbarkeit der wahren Klassen im Featureraum also nicht zu einer Verbesserung der Kombinationsgüten, wenn die Anzahl der verfügbaren Experten hinreichend hoch ist. In diesen Fällen (Anzahl der Experten > 3) liefert in den Simulationen die Anwendung der DSR die besten Ergebnisse.

Nach erfolgter Gegenüberstellung der Kombinationsgüten der verschiedenen Kombinationsregeln für alle drei Datensätze werden nun weitere Eigenschaften der EIDMR anhand zweier zusätzlicher Experimente untersucht. Im ersten dieser weiteren Experimente wird die Anzahl der Datenpunkte im Datensatz III variiert, um den Einfluss der Datendichte auf die mit der EIDMR erreichte Kombinationsgüte analysieren zu können. Abbildung 20 zeigt die Kombinationsergebnisse für 250, 750 und alle 1000 in Datensatz III enthaltenen Datenpunkte. Der Mittelwert (zehn Wiederholungen, siehe oben) der Kombinationsgüten erhöht sich mit zunehmender Anzahl an Datenpunkten deutlich. Diese Erhöhung ist hauptsächlich auf die zunehmende Datendichte zurückzuführen, die zu einer Auswahl von – bezüglich der Features – ähnlicheren k benachbarten Datenpunkten durch den $k$nn-Ansatz führt. Der beschriebene Effekt ist für $k = 9$ deutlich stärker ausgeprägt als für $k = 3$.

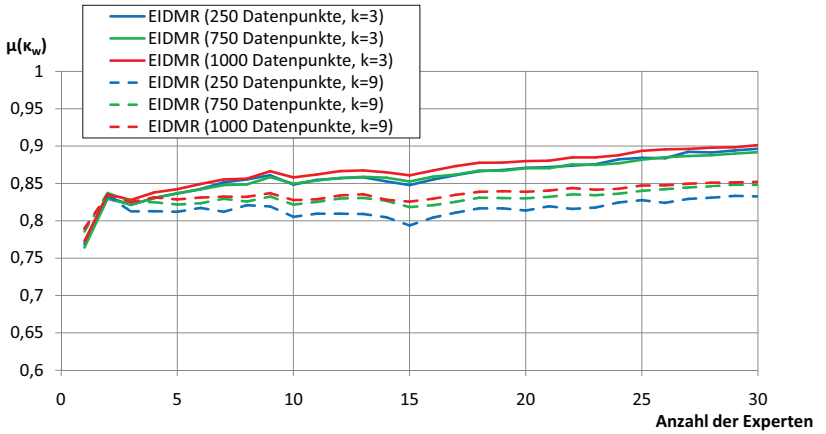


Abbildung 20: Kombinationsgüten der EIDMR bei variierender Anzahl von Datenpunkten im Datensatz I.

Das zweite zusätzliche Experiment dient zur Untersuchung des Einflusses der Sicherheitsgewichtung $w_{i,j}$ und der Berücksichtigung der ordinalen Information bei Anwendung der EIDMR. Zur Durchführung des Experiments wurde jedes Sicherheitsgewicht $w_{i,j}$ auf einen Wert gesetzt, der einer Vernachlässigung der durch die Experten abgegebenen Schwierigkeitseinschätzungen für jeden Datenpunkt äquivalent ist. Um die ordinale Information der Klassen zu vernachlässigen, kann z. B. jeder Wert des Ähnlichkeitsgewichts g_i zu $1/(k+1)$ gewählt werden. In dem Experiment wurden beide Fälle sowie ihre Kombination untersucht. Der letztgenannte Fall führt zu einem einfachen Mehrheitsentscheid bei Anwendung der EIDMR. Die Ergebnisse in Abbildung 21 zeigen, dass die Sicherheitsgewichtung und die Berücksichtigung der ordinalen Information einen deutlich positiven Effekt auf die Kombinationsgüte haben.

Unter Berücksichtigung der diskutierten Experimente auf den drei Datensätze kann festgestellt werden, dass die DSR eine sehr gute Kombinationsregel zur Detektion der wahren Klassen darstellt. Allerdings kann die Kombinationsgüte in einigen Fällen durch Anwendung der EIDMR weiter erhöht werden, vor allem, wenn die Anzahl der verfügbaren Expertenaussagen eher gering ist. Der Mehrwert der EIDMR hängt vom verwendeten Datensatz ab. Die Ergebnisse können wie folgt zusammengefasst werden:

- Bei geringer Überlappung der Komponenten des GMM liefert die EIDMR für jede untersuchte Anzahl von Experten durch die Berücksichtigung von Informationen benachbarter Datenpunkte bessere Kombinationsergebnisse als die bekannten Regeln. Das Kombinationsergebnis ist durch die gute Separierbarkeit der Klassen nahezu unabhängig von den in den Experimenten berücksichtigten Werten für k.

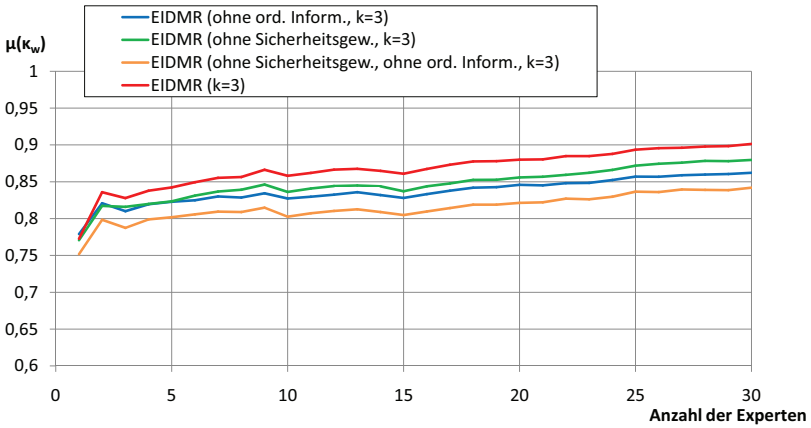


Abbildung 21: Einfluss der Sicherheitsgewichtung und der ordinalen Information auf die Kombinationsgüte für Datensatz I.

- Mit zunehmend schlechterer Separierbarkeit der Klassen im Featureraum (wie z. B. im Datensatz II) verliert die in benachbarten Datenpunkten enthaltene Information bei der Kombination einer hohen Anzahl von Expertenaussagen (> 9) mit der EIDMR an Bedeutung. Der Einfluss von k bei der Kombination mit EIDMR steigt im Vergleich zu Datensatz I. Die höchste Kombinationsgüte stellt sich in den Experimenten für $k = 9$ ein. Die EIDMR liefert ab einer Expertenzahl von 12 schlechtere Ergebnisse als die DSR.

- Bei hoher Überlappung der Komponenten des GMM und einer hinreichend hohen Anzahl verfügbarer Expertenaussagen stellt die DSR die beste Alternative zur Kombination der Expertenaussagen dar. Durch die schlechtere Separierbarkeit der Klassen im Featureraum treten bei Anwendung der EIDMR die höchsten Kombinationsgüten für $k = 3$ auf. Ist allerdings nur eine geringe Expertenzahl (< 4) verfügbar, z. B. durch hohe Kosten für die Erhebung der Expertenaussagen, so liefert die Anwendung der EIDMR bessere Ergebnisse als die DSR.

- Neben der Überlappung der Komponenten des GMM und der Wahl von k hängt die Kombinationsgüte der EIDMR von der Dichte der Daten ab. In den Experimenten wirkte sich eine höhere Anzahl von Datenpunkten positiv auf das Kombinationsergebnis auf.

- Die Berücksichtigung der Sicherheitsgewichtung und der ordinalen Information haben jeweils einen deutlich positiven Effekt auf die Kombinationsgüte. Durch Kombination beider Fälle wird die Kombinationsgüte weiter erhöht.

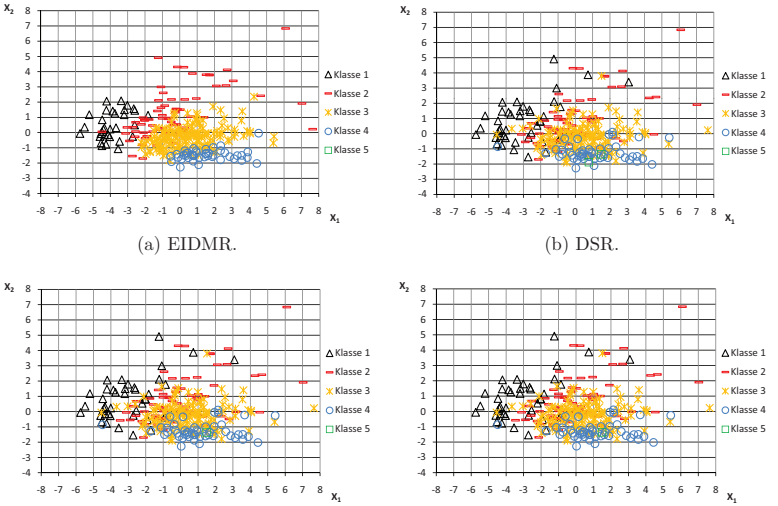
5.4 ANWENDUNG DER KOMBINATIONSREGELN AUF NIEDERSPANNUNGSNETZE

Die im voranstehenden Abschnitt anhand von künstlichen Daten untersuchten Kombinationsregeln werden nun zur Kombination der in Kapitel 4 erhobenen Klassifikationseinschätzungen (fünf Expertenaussagen und ein simulationsbasiertes Klassifikationsergebnis) genutzt. Im Gegensatz zum in Abschnitt 4.3 durchgeführten statistischen Vergleich zwischen den Expertenaussagen und den Simulationsergebnissen auf Grundlage eines Mehrheitsentscheides wird daher im Folgenden nicht nur eine komplexere Methodik zur Ermittlung einer Gesamtaussage angewendet, sondern auch die Ergebnisse beider Klassifikationsergebnisse zu einem resultierenden Ergebnis zusammengeführt. Durch die Kombination des Expertenwissens mit den automatisch erhobenen Simulationsergebnissen wird eine umfassendere Bewertung der untersuchten NS-Netze erwartet. In diesem Zusammenhang wird im Folgenden zum einen der Einfluss der Simulationsergebnisse auf die Übereinstimmung zu den einzelnen Klassifikationseinschätzungen und zum anderen der Einfluss auf die Verteilung der Kombinationsergebnisse im Featureraum untersucht. Die Analyse dieser Einflüsse hat das Ziel, im Rahmen dieser Arbeit einen guten Kompromiss zwischen einer angemessenen Berücksichtigung der Simulationsergebnisse in der Gesamtaussage (um eine umfassende Bewertung der Netze zu erhalten) und einer möglichst guten Separierbarkeit der Kombinationsergebnisse im Featureraum (zur Anwendung eines SVM-Klassifikators, Kapitel 6) zu erhalten.

Zur Anwendung der Kombinationsregeln wurden – analog zum Vorgehen für die künstlich generierten Expertenaussagen in Abschnitt 5.3 – für jede Expertenaussage auf der Grundlage der im Rahmen der Expertenbefragung erhobenen Schwierigkeitseinschätzungen („leicht“,“ mittel“ und „schwer“, siehe Abschnitt 4.1) jeweils ein Sicherheitsgewicht (bei IDMR und EIDMR: 1,00, 1,25 und 1,50, bei DSR und MR: 0,60, 0,75 und 0,90) verwendet. Bei den simulationsbasierten Klassifikationsergebnissen wurde für jedes Netz die Klasse mit der höchsten probabilistischen Klassenzugehörigkeit $p(c)$ ($c = 1, \ldots, 5$) ausgewählt. Der Wert $p(c)$ der auf diese Weise ausgewählten Klasse c kann direkt als Sicherheitsgewicht der Entscheidung für diese Klasse interpretiert werden. Zur Anwendung der IDMR und EIDMR musste die Gewichtung entsprechend durch Multiplikation mit 1,5 auf ein zu der Gewichtung der Expertenaussagen analoges Intervall überführt werden. Als Featureraum zur Auswahl der benachbarten Datenpunkte wurden die zuvor z-transformierten, netzspezifischen Merkmale aus Abschnitt 3.1 verwendet. Entsprechend der Ergebnisse aus Abschnitt 5.3.3 wurde die Anzahl der berücksichtigten, benachbarten Datenpunkte bei Anwendung der EIDMR eher niedrig zu $k = 5$ gewählt.

Die Güte der Kombinationsergebnisse kann für die in dieser Arbeit verwendeten realen Daten von NS-Netzen nicht wie in Abschnitt 5.3 evaluiert werden. Dies liegt darin begründet, dass die wahren Klassen der NS-Netze nicht bekannt sind. Um die Kombi-

nationsregeln auf den realen Daten miteinander zu vergleichen, werden in dieser Arbeit in Kapitel 6 die kombinierten Label genutzt, um SVM-Klassifikatoren zu modellieren und hinsichtlich ihrer Klassifikationsgüte miteinander zu vergleichen. Dieser Untersuchung liegt die Annahme zu Grunde, dass die angewandte Kombinationsregel nicht nur eine hohe Güte zur Detektion der wahren Klassen aufweisen, sondern auch zu einer guten Separierbarkeit der kombinierten Label im Featureraum beitragen sollte, um im Ergebnis einen guten Klassifikator zu erhalten.



(a) EIDMR. (b) DSR.

(c) MR. (d) IDMR.

Abbildung 22: Darstellung der Kombinationsergebnisse auf den zwei Hauptachsen.

Abbildung 22 zeigt die Ergebnisse der Kombinationsregeln anhand der ersten beiden Hauptachsen, die anhand einer Hauptachsentransformation der netzspezifischen Merkmale ermittelt wurden. In der Kombination wurden die simulationsbasierten Klassifikationseinschätzungen wie eine zu den fünf Experten zusätzliche, gleichgewichtige Experteneinschätzung behandelt, so dass im Ergebnis insgesamt sechs Klassifikationseinschätzungen kombiniert worden sind. Auf Basis der duch die Hauptachsentransformation vorgenommenen Dimensionsreduktion lässt sich ein (wenn auch auf Grund des mit der Dimensionsreduktion einhergehenden Informationsverlustes nicht ganz vollständiger) visueller Eindruck über die Verteilung der Daten im Featureraum gewinnen. Insgesamt unterscheiden sich die in Abbildung 22 gezeigten Ergebnisse der DSR, MR und IDMR visuell nur geringfügig. Dieser Eindruck wird von den in Tabelle 16 ausgewiesenen, paarweise ermittelten Interraterkonkordanzen κ_w gestützt. Die Ergebnisse der drei bekannten Kombinationsregeln sind mit sehr hohen Werten für κ_w konkordant, die zwi-

schen 0,92 und 0,97 liegen. Im Gegensatz dazu weisen die Resultate der EIDMR zwar immer noch eine gute, allerdings deutlich geringere Konkordanz zu den Ergebnissen der drei anderen Kombinationsregeln auf. Die Ergebnisse der EIDMR unterscheiden sich damit deutlich von denen der anderen Regeln. Auffällig ist für alle gezeigten Datensätze, dass die Kombination der einzelnen Klassifikationseinschätzungen zu einer sehr geringen Anzahl von Datenpunkten der Klasse 5 führt. Bei Anwendung der DSR werden nur zwei Netze, bei Anwendung der MR und IDMR wird jeweils nur ein Netz der Klasse 5 zugeordnet. In den Ergebnissen der EIDMR ist die Klasse 5 sogar gar nicht mehr vorhanden. Die Ursache für die geringe Anzahl an Netzen, die der Klasse 5 zugeordnet werden, liegt in der bereits geringen Anzahl von Netzen der Klasse 5 in den einzelnen Klassifikationseinschätzungen (vgl. Abbildungen 6 und 13). Vor diesem Hintergrund führt die Berücksichtigung der $k = 5$ benachbarten Datenpunkte bei Anwendung der EIDMR dazu, dass Klasse 5 im Ergebnis keinem der untersuchten Netze zugeordnet wird.

κ_w	EIDMR	DSR	MR	IDMR
EIDMR	—	0,653	0,648	0,645
DSR	—	—	0,964	0,927
MR	—	—	—	0,924
IDMR	—	—	—	—

Tabelle 16: Paarweise Interraterkonkordanz zwischen den Kombinationsergebnissen.

Die Berücksichtigung von benachbarten Datenpunkten bei der Kombination führt allerdings im Vergleich zu den anderen drei Kombinationsregeln zu einer deutlich besseren Separierbarkeit der Klassen. Ein numerischer Vergleich zwischen den Kombinationsergebnissen lässt sich wieder anhand der Davies-Bouldin-Indizes durchführen (alle zehn z-transformierten Merkmale wurden berücksichtigt). Für die Anordnung der unter Anwendung der EIDMR bestimmten Kombinationsergebnisse im Featureraum ergibt sich ein Wert von $DB_{EIDMR} = 2{,}28$. Entsprechend des anhand von Abbildung 22 erhaltenen visuellen Eindrucks resultieren für die übrigen Kombinationsregeln deutlich höhere Werte von $DB_{DSR} = 3{,}19$, $DB_{MR} = 3{,}03$ und $DB_{IDMR} = 3{,}03$. Der im Vergleich deutlich niedrigere Wert des Davies-Bouldin-Index bei Anwendung der EIDMR lässt vermuten, dass die Kombinationsergebnisse auf Basis der verwendeten Merkmale besser anhand eines SVM-Klassifikators prognostizierbar sein werden. Diese Erwartung liegt darin begründet, dass die Klassifikationsgüte einer SVM maßgeblich von der Verteilung der Daten im Featureraum abhängig ist und wird in Kapitel 6 nochmal aufgegriffen und in Experimenten mit SVM-Klassifikatoren näher analysiert.

Bisher wurden die simulationsbasierten Klassifikationseinschätzungen in der Kombination nur als ein zusätzlicher Experte berücksichtigt. Dies entspricht (je nachdem wie die einzelnen Schwierigkeitseinschätzungen eines Datenpunktes ausfallen) einer Gewichtung von ca. $1/6$. Um den Ergebnissen der stochastischen Simulationen mehr Gewicht

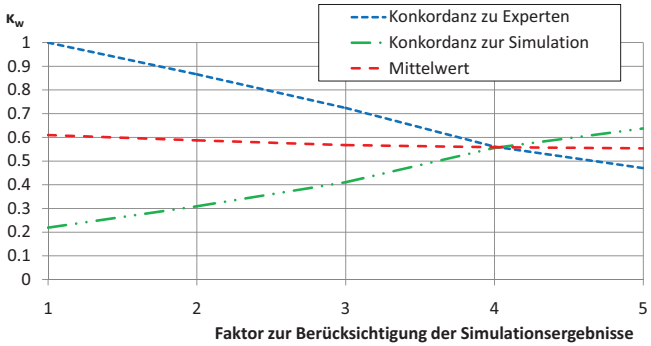


Abbildung 23: Übereinstimmung der Kombinationsergebnisse zu den Experten und zur Simulation.

bei der Kombination zu verleihen, können diese mehrfach, d. h. wie jeweils zwei oder mehr Experten gewichtet werden. In einem letzten Schritt wird nun der Einfluss der Gewichtung der Simulationsergebnisse auf die Kombinationsergebnisse untersucht. Dazu wurde die EIDMR mehrfach angewendet und darin die Gewichtung der Simulationsergebnisse von der bisher erfolgten einfachen bis auf eine fünffache Gewichtung erhöht. Abbildung 23 zeigt die Interraterkonkordanz κ_w zwischen zum einen den auf diese Weise bestimmten Kombinationsergebnissen und zum anderen den (unter Verwendung der EIDMR ohne Berücksichtigung der Simulationsergebnisse, $k = 5$) kombinierten Expertenaussagen sowie den simulationsbasierten Klassifikationseinschätzungen. Entsprechend nimmt die Konkordanz zwischen den Kombinationsergebnissen und den kombinierten Expertenaussagen mit steigender Gewichtung der Simulationsergebnisse sukzessive ab und dazu gegensätzlich die Konkordanz zwischen den Kombinationsergebnissen und den einzelnen Simulationsergebnissen sukzessive zu. Bei fünffacher Gewichtung der Simulationsergebnisse und damit einer ungefähr gleichen Gewichtung zwischen Expertenaussagen und Simulationsergebnissen, weisen die Kombinationsergebnisse eine etwas höhere Konkordanz zu den Simulationsergebnissen als zu den kombinierten Expertenaussagen auf. Es stellt sich die Frage, welche Klassifikationseinschätzungen, die expertenbasierten oder die simulationsbasierten, bei der Kombination höher zu gewichten sind. Das Ergebnis dieser Fragestellung wird immer eine Abwägung zwischen gewünschter Gewichtung und erzielter Klassifikationsgüte bei Anwendung einer SVM erfordern. Zusätzlich zu den bereits disktutieren Konkordanzverläufen ist in Abbildung 23 ebenfalls jeweils der Mittelwert zwischen den beiden Konkordanzen gebildet worden, um eine Gesamtindikation zur Übereinstimmung zwischen dem Kombinationsergebnis und den kombinierten Expertenaussagen sowie den Simulationsergebnissen zu erhalten. Es zeigt sich, dass der Mittelwert bei einfacher Berücksichtigung der Simulationsergebnisse am höchsten ist und mit zunehmend höherer Gewichtung leicht abnimmt. Die Abnahme im Vergleich zum Maximum der Mittelwerte ist allerdings nicht stark ausgeprägt. Um

zusätzlich die Auswirkungen einer zunehmenden Gewichtung der Simulationsergebnisse auf die Datenverteilung im Featureraum zu bewerten, ist für jeden der fünf Datensätze der Davies-Bouldin-Index bestimmt worden. Im Vergleich zum bereits für die einfache Berücksichtigung bestimmten Wert von $DB_{EIDMR} = 2{,}28$ (ohne Berücksichtigung der Simulationsergebnisse, d. h. nur Expertenaussagen: $DB_{EIDMR} = 2{,}32$) erhöht sich der Davies-Bouldin-Index bei zunehmender Gewichtung der Simulationsergebnisse sukzessive auf einen Wert von 3,99 bei fünffacher Gewichtung. Eine Erhöhung der Gewichtung beeinflusst folglich die Datenverteilung im Featureraum deutlich negativ, woraus sich eine Verschlechterung der Klassifikationsgüte bei Verwendung einer SVM zur Prognose der Kombinationsergebnisse auf Basis der verwendeten Merkmale erwarten lässt. Auch dieses Ergebnis wird in Kapitel 6 in Experimenten mit SVM-Klassifikatoren näher analysiert.

Zusammenfassend bleibt an dieser Stelle festzuhalten, dass die EIDMR im Vergleich zu den anderen verwendeten Kombinationsregeln zu einer – bereits visuell auf Basis der zwei Hauptachsen bemerkbar – besseren Separierbarkeit der fünf Klassen führt. Eine zunehmende Gewichtung der Simulationsergebnisse bei der Kombination mit der EIDMR beeinflusst die Datenverteilung im Featureraum deutlich negativ. Vor dem Hintergrund dieser Ergebnisse erscheint eine maximal zweifache Gewichtung der Simulationsergebnisse bei der Kombination sinnvoll, um zum einen die Simulationsergebnisse im Gesamtergebnis nicht zu überschätzen und gleichsam eine gute Separierbarkeit im Featureraum zur Anwendung von SVM-Klassifkatoren zu erhalten. Die Analyse der Kombinationsergebnisse sind als Indikationen zu bewerten und werden in Kapitel 6 im Rahmen von Experimenten mit SVM-Klassifikatoren weiter evaluiert.

5.5 EINORDNUNG DER ERGEBNISSE

In diesem Kapitel wurden vier Kombinationsregeln zur Fusion unsicherer Klassifikationseinschätzungen zu einem Gesamtergebnis untersucht. Auf Basis eines Vergleichs der Kombinationsgüten dieser Regeln auf künstlichen Daten wurden die Kombinationsregeln anschließend genutzt, um die mit Hilfe der in Kapitel 4 vorgestellten Ansätze bestimmten Klassifikationseinschätzungen zu einer Gesamtaussage zu vereinen. Im Ergebnis wurden damit für die weiteren Untersuchungen dieser Arbeit zusätzliche Beispieldaten (die Kombinationsergebnisse) erzeugt, die zum Training von SVM-Klassifikatoren in Abschnitt 6.3.3 genutzt werden können. Die Frage, mit welcher der Kombinationsregeln sich bei der Modellierung eines SVM-Klassifikators die höchste Klassifikationsgüte erreichen lässt, wird ebenfalls in Abschnitt 6.3.3 aufgegriffen und empirisch untersucht. Der Untersuchung liegt die Annahme zu Grunde, dass die angewandte Kombinationsregel nicht nur eine hohe Güte zur Detektion der wahren Klassen aufweisen, sondern auch zu einer guten Verteilung der Kombinationsergebnisse im Featureraum beitragen sollte, um im Ergebnis einen guten Klassifikator zu erhalten.

6

KLASSIFIKATION MIT SUPPORT VECTOR MACHINES

Auf der Grundlage der bisher im Verlauf dieser Arbeit erhobenen Beispieldaten kann durch Anwendung eines SVM-Klassifikators eine Effizienzsteigerung für die Bewertung weiterer, in dieser Arbeit nicht untersuchter, NS-Netze bewirkt werden. Eine Anwendung der Klassifikationsansätze aus Kapitel 4 sowie einer Kombinationsregel aus Kapitel 5 ist nach erfolgter Modellierung des Klassifikators dann nämlich nicht mehr erforderlich. Der Klassifikator prognostiziert die Klasse eines nicht in den Beispieldaten enthaltenen Netzes nur anhand von dessen Features.

In diesem Kapitel werden drei SVM-Klassifikationskonzepte zur Prognose der Kombinationsergebnisse (kombinierte Label) vorgestellt. Zunächst erfolgen in Abschnitt 6.1 einige grundlegende Vorbemerkungen zu maschinellen Lernverfahren und darüber hinaus eine Erläuterung der mathematischen Grundlagen von SVM-Klassifikatoren. Auf dieser Basis schließt sich in Abschnitt 6.2 eine detaillierte Erläuterung der drei Klassifikationskonzepte zur Prognose der kombinierten Label mit SVM an. Daraufhin werden diese Konzepte in Abschnitt 6.3 hinsichtlich ihrer Klassifikationsgüte in empirischen Experimenten gegenübergestellt. In diesem Zuge erfolgt darüber hinaus eine Analyse des Einflusses der zur Modellierung verwendeten Features. Entsprechend der Untersuchungen in Abschnitt 5.4 wird auch der Einfluss der Simulationsergebnisse auf die Klassifikationsgüte des Klassifikators analysiert (mehrfache Gewichtung der Simulationsergebnisse bei Anwendung der EIDMR, d. h. wie jeweils zwei oder mehr Experten). Weiterhin wird dargelegt, dass der SVM-Klassifikator durch Berücksichtigung von ordinalen Informationen bei der Lösung des Mehrklassenklassifikationsproblems weiter verbessert werden kann.

6.1 GRUNDLAGEN

6.1.1 *Vorbemerkungen*

Allgemein können mit Klassifikatoren Objekte automatisiert anhand ihrer Features in vorgegebene Klassen eingeordnet werden. Dazu muss der Klassifikator vorher anhand von Beispieldaten trainiert und getestet werden. Die Prognosegüte für einen unbekannten Datenpunkt hängt stark vom Klassifikatortyp sowie der Qualität und Struktur der genutzten Daten ab. Die zum Training des Klassifikators verwendete Menge von Beispieldaten wird als Trainingsdatenmenge T bezeichnet. Die Ausprägungen der Features eines Datenpunktes können allgemein durch einen Featurevektor $\boldsymbol{x}$ beschrieben werden. Diesem Vektor ist im Ausgaberaum ein Wert y zugeordnet (z. B. eine Klasse bzw. ein Label c). In der Computational Intelligence (CI) werden zum Training von Klassifikatoren Lernverfahren der folgenden übergeordneten Kategorien unterschieden. Dabei wird zur Modellierung eines SVM-Klassifikators am häufigsten das überwachte Lernen angewendet:

ÜBERWACHTES LERNEN: Zu jedem Datenpunkt beschrieben durch einen Featurevektor $\boldsymbol{x}$ der Trainingsdatenmenge T im Featureraum ist ebenfalls ein zugeordneter Wert y im Ausgaberaum bekannt. In der Trainingsphase wird die tatsächliche Ausgabe $\hat{y}$ des Klassifikators mit dem korrekten (gewünschten) Ausgabewert y verglichen.

UNÜBERWACHTES LERNEN: Für alle Featurevektoren $\boldsymbol{x}$ einer Trainingsdatenmenge T im Featureraum ist kein zugeordneter Wert y im Ausgaberaum bekannt. Nach der Trainingsphase sollten typischerweise ähnliche Eingaben (nahe zueinanderliegende Featurevektoren) auch zu ähnlichen Ausgabewerten des Klassifikators führen.

HALBÜBERWACHTES LERNEN: Es liegt nur für einen (häufig geringen) Teil der Featurevektoren $\boldsymbol{x}$ im Featureraum ein zugeordneter Wert y im Ausgaberaum vor. In der Trainingsphase kann somit nur vereinzelt ein Vergleich zwischen der tatsächlichen Ausgabe $\hat{y}$ des Klassifikators und dem korrekten Ausgabewert y durchgeführt werden. Beim halbüberwachten Lernen kann z. B. ein Ziel sein, einen Klassfikator durch Kombination von überwachtem und unüberwachtem Lernen und der damit verbundenen Vergrößerung der zum Training verfügbaren Beispieldaten gegenüber dem reinen überwachten Lernen zu verbessern.

AKTIVES LERNEN: Das aktive Lernen bildet einen Spezialfall des halbüberwachten Lernens. Zunächst liegt wie im Fall des unüberwachten Lernens in der Trainingsdatenmenge T zu keinem Featurevektor $\boldsymbol{x}$ ein zugeordneter Ausgabewert y vor.

Daraufhin wird in der Trainingsphase interaktiv für ausgewählte Featurevektoren $\boldsymbol{x}$ der Ausgabewert y von einem „Orakel“ (z. B. menschlicher Experte) abgefragt.

Die SVM wurde zuerst von Vapnik und Cordes vorgestellt [96]. Der Vorteil der SVM gegenüber vielen anderen Methoden liegt in einem guten Kompromiss zwischen der Modellkomplexität und der Lernfähigkeit bei gleichzeitig hoher Vorhersagegüte [25]. Die gute Fähigkeit zur Generalisierung von SVM wurde bereits relativ früh in vielen Anwendungen demonstriert [97, 98]. Bei der Anwendung von SVM kann Unsicherheit prinzipiell an vielen Stellen auftreten, z. B. hinsichtlich der Zuordnung der Datenpunkte zu einer Klasse, der exakten Featurewerte oder der Parametrisierung der SVM anhand einer begrenzten Datenmenge. Vor diesem Hintergrund bezieht sich die Verwendung von SVM bei unsicheren Informationen hier auf den Fall, in dem Datenpunkte nicht eindeutig einer Klasse bzw. gleichzeitig mehreren Klassen zugeordnet sind (wie z. B. bei den in dieser Arbeit erhobenen einzelnen experten- und simulationsbasierten Klassifikationseinschätzungen). Um diese Problemstellung direkt in einer SVM zu adressieren, können z. B. die Fuzzy-Theorie [99] oder die Rough-Set-Theorie [100, 101] genutzt werden. Im Gegensatz dazu wird die unsichere Informationen hier durch vorherige Verwendung einer Kombinationsregel behandelt. Die nachgelagerte Anwendung der SVM reduziert sich damit auf die Prognose der kombinierten Label.

Mehrklassenklassifikationsprobleme werden bei SVM über mehrere Zweiklassenklassifikationsprobleme gelöst. Die Klassen werden dabei als nominalskalierte Werte verarbeitet, d. h. eine eventuell vorhandene Ordnung der Klassen wird vernachlässigt [102]. Die Berücksichtigung ordinaler Klassen kann z. B. durch Anwendung des allgemeinen Ansatzes von Frank und Hall für Klassifikatoren zur Lösung von ordinalen Mehrklassenklassifikationsproblemen erfolgen [103]. Ein weiterer Ansatz zur direkten Verarbeitung von ordinalen Klassen in einer SVM wird in [104] vorgestellt. In dieser Arbeit wird die Ordnung der Klassen zunächst durch Optimierung der gewichteten Interraterkonkordanz κ_w (Abschnitt 4.1.4) zwischen den kombinierten Labeln und der Klassifikatorausgabe berücksichtigt, die als Gütemaß zur Bewertung der SVM-Klassifikationskonzepte genutzt wird. Darauf aufbauend wird in den hier durchgeführten Experimenten der Ansatz von Frank und Hall zur ordinalen Klassifikation mit SVM verwendet und gezeigt, dass sich mit diesem Ansatz die Klassifikationsgüte hinsichtlich einer Optimierung der gewichteten Interraterkonkordanz κ_w zwischen den kombinierten Labeln und der Klassifikatorausgabe noch verbessern lässt [103].

6.1.2 *Grundlagen zu Support Vector Machines*

Im einfachsten Fall wird beim Einsatz von SVM für ein Zweiklassenproblem unter Berücksichtigung der Trainingsdatenmenge eine Hyperebene zur linearen Trennung der Datenpunkte im Eingaberaum bestimmt. Die Lage dieser Hyperbene wird während

der Trainingsphase dahingehend optimiert, dass der Abstand der Datenpunkte, die der Trennebene am nächsten liegen, zur Trennebene maximiert wird. Es wird so anschaulich durch Variation der Lage der Trennebene ein optimaler Rand um die Entscheidungsgrenze gebildet, um möglichst sichere Klassifikatorausgaben zu erhalten. Der Rand beschreibt dabei den Abstand zwischen der Trennebene und dem nächstliegenden Datenpunkt, einem Stützvektor. Diese Vektoren führten zur entsprechenden Benennung des Verfahrens. Oft ist jedoch eine exakte lineare Trennung der Datenpunkte im Eingaberaum nicht möglich. Man spricht dann von einem nicht linear trennbaren Klassifikationsproblem (Abbildung 24). Bei SVM kann diesem Problem durch eine nichtlineare Abbildung in einen im Vergleich zum Eingaberaum im Allgemeinen höherdimensionalen Featureraum begegnet werden. Dabei wird mit Hilfe einer Kernelfunktion die nichtlineare Abbildung und gleichzeitig die Bildung des Skalarprodukts für zwei Datenpunkte im Featureraum realisiert. Die Entwicklung einer nichtlinearen Kurve im Eingaberaum zur Trennung der Datenpunkte ist somit nicht erforderlich.

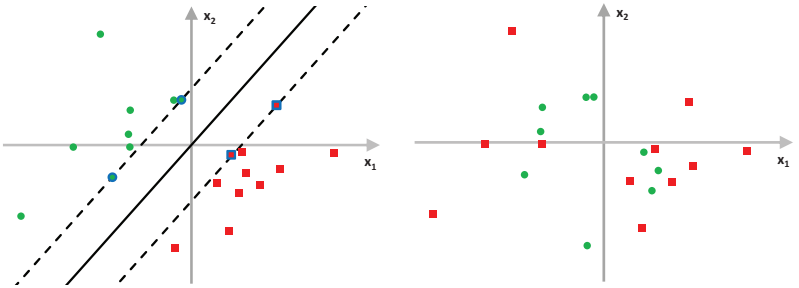


Abbildung 24: Schematische Gegenüberstellung eines linear trennbaren (links) und eines nicht linear trennbaren (rechts) Klassifikationsproblems.

Nach Beschreibung der Funktionsweise von SVM folgt nun deren exakte mathematische Beschreibung. Dazu wird zunächst eine allgemeine Formulierung des linearen SVM-Klassifikators für ein Zweiklassenklassifikationsproblem hergeleitet. Jedem Featurevektor $\boldsymbol{x}$ sei dazu ein Label y (mit $y \in \{-1, 1\}$) zugeordnet. Allgemein kann ein linearer Klassifikator mit Hilfe eines Gewichtsvektors $\boldsymbol{w}$ und eines zusätzlichen Parameters b formal wie folgt beschrieben werden [97, 102, 105]:

$$\widehat{k}(\boldsymbol{x}) \;=\; \operatorname{sign}\left(\boldsymbol{w}^T\,\boldsymbol{x} + b\right). \tag{35}$$

Anschaulich erhält man mit Gleichung (35) die zur linearen Trennung der Datenpunkte erforderliche Hyperebene in ihrer Normalenform, wenn man $\boldsymbol{w}^T\,\boldsymbol{x} + b = 0$ setzt. Der Gewichtsvektor $\boldsymbol{w}$ verkörpert geometrisch also den Normalenvektor der Hyperebene. Damit liegt ein Datenpunkt, für den $\widehat{k}(\boldsymbol{x}) = 0$ erfüllt ist, direkt auf der Entscheidungsgrenze des linearen Klassifikators. Es stellt sich nun die Frage, wie die Hyperebene im Eingaberaum zu wählen ist, um der oben genannten Anforderung nachzukommen,

dass der Abstand der Datenpunkte, die der Trennebene am nächsten liegen, maximiert wird. Anschaulich entspricht diese Anforderung einer Maximierung des Randes um die Trennebene, in dem keine Datenpunkte zu finden sind. Für eine gegebene linear trennbare Trainingsdatenmenge $T = \{(\boldsymbol{x}_i, y_i)\}$ $(i = 1, \ldots, n)$ wird zunächst ein Funktional zur Berechnung des Randes $\hat{r}$ zur Trennebene eingeführt [97, 102, 105]:

$$\hat{r}_i \;=\; y_i \left(\boldsymbol{w}^T \boldsymbol{x}_i + b \right). \tag{36}$$

Ein positiver Wert für $\hat{r}_i$ tritt also immer dann auf, wenn die Lage und Position der Trennebene zur korrekten Ausgabe für den Trainingsvektor $\boldsymbol{x}_i$ führt. Weiterhin nimmt mit steigendem Wert für $(\boldsymbol{w}^T \boldsymbol{x}_i + b)$ auch die Entfernung von $\boldsymbol{x}_i$ zur Trennebene und damit auch die Sicherheit der anhand von $\widehat{k}(\boldsymbol{x}_i)$ ermittelten Ausgabe zu. Der funktionale Rand eines linearen Klassifikators bezüglich einer Trainingsdatenmenge T ist daher wie folgt definiert [97, 102, 105]:

$$\hat{r} \;=\; \min_{i=1,\ldots,n} \hat{r}_i. \tag{37}$$

Im Gegensatz zum geometrischen Abstand ist $\hat{r}$ skalierbar. Es sei an dieser Stelle erwähnt, dass man den geometrischen Rand des Klassifikators durch Normierung der rechten Seite von Gleichung (36) auf die L_2-Norm von $\boldsymbol{w}$ erhält. Die Ermittlung der Lage der Trennebene kann unter Ausnutzung der Skalierbarkeit von $\hat{r}$ als ein quadratisches, konvexes Optimimierungsproblem mit linearen Nebenbedingungen formuliert werden, das mit konventioneller quadratischer Programmierung effizient gelöst werden könnte (zum Training von SVM existieren allerdings spezielle, effizientere Verfahren, siehe unten). Dazu wird zur Maximierung des geometrischen Randes $\hat{r}/\|\boldsymbol{w}\|$ der funktionale Rand auf einen Wert von $\hat{r} = 1$ skaliert. Die daraus resultierende Maximierung von $1/\|\boldsymbol{w}\|$ kann dann gleichwertig durch eine Minimierung von $\|\boldsymbol{w}\|^2$ ersetzt werden. Unter Berücksichtigung der Nebenbedingungen ergibt sich das folgende Optimierungsproblem zur Maximierung des Randes um die Trennebene [97, 102, 105]:

$$\min_{\boldsymbol{w}} \; \frac{1}{2}\|\boldsymbol{w}\|^2 \tag{38}$$

mit

$$y_i \left(\boldsymbol{w}^T \boldsymbol{x}_i + b \right) \geq 1, \quad \text{für } i = 1, \ldots, n. \tag{39}$$

Das durch die Gleichungen (38) und (39) beschriebene Optimierungsproblem lässt sich durch Umstellen der Nebenbedingungen in eine Form bringen, die eine Lösung durch Anwendung des Lagrange-Ansatzes ermöglicht [97, 102, 105]. Über eine primäre Form des Optimierungsproblems wird dann, unter Berücksichtigung der so genannten Karush-Kuhn-Tucker-(KKT)-Bedingungen, das zu dieser Lagrangefunktion korrespondierende duale Optimierungsproblem gefunden [97, 102, 105]:

$$\max_{\boldsymbol{\alpha}} \quad W(\boldsymbol{\alpha}) \;=\; \sum_{i=1}^{n} \alpha_i - \frac{1}{2} \sum_{i=1}^{n} \sum_{j=1}^{n} y_i y_j \alpha_i \alpha_j \boldsymbol{x}_i^T \boldsymbol{x}_j \tag{40}$$

mit

$$\boldsymbol{\alpha} = (\alpha_1, \ldots, \alpha_n)^{\mathrm{T}}, \tag{41}$$

$$\alpha_i \geq 0 \quad i = 1, \ldots, n, \tag{42}$$

$$\sum_{i=1}^{n} \alpha_i y_i = 0. \tag{43}$$

Im Anschluss an die Maximierungsaufgabe bezüglich $\boldsymbol{\alpha}$ können die optimalen Werte $\boldsymbol{w}^*$ mittels $\boldsymbol{w} = \sum_{i=1}^{n} \alpha_i y_i \boldsymbol{x}_i$ berechnet werden. Interessant ist, dass nur diejenigen α_i Werte ungleich null annehmen werden, die direkt am Rand um die Trennebene liegen. Dies sind – wie bereits oben erwähnt – nur einige Vektoren der Trainingsdatenmenge, die Stützvektoren. Die Verschiebung b^* wird in der Praxis durch Mittelwertbildung bestimmt [97, 102, 105], z. B.:

$$b^* = -\frac{1}{2} \cdot \left(\max_{i:y_i=-1} \boldsymbol{w}^{*T} \boldsymbol{x}_i + \min_{i:y_i=1} \boldsymbol{w}^{*T} \boldsymbol{x}_i \right). \tag{44}$$

Um die Parameter der linearen SVM an die Trainingsdatenmenge T anzupassen, benötigt die duale Form in Gleichung (40) nur noch das innere Produkt zwischen den Datenpunkten $\boldsymbol{x}_i$ und $\boldsymbol{x}_j$. Zur Erweiterung der linearen SVM wird nun eine Abbildungsvorschrift $\boldsymbol{\phi}(\boldsymbol{x})$ eingeführt, die Datenpunkte aus dem Eingaberaum in den Featureraum transformiert. Dazu muss die Abbildungsvorschrift $\boldsymbol{\phi}(\boldsymbol{x})$ nicht explizit bekannt sein, sondern lediglich deren inneres Produkt $\boldsymbol{\phi}(\boldsymbol{x}_i)^T \boldsymbol{\phi}(\boldsymbol{x}_j)$. Das innere Produkt der Abbildungsvorschrift wird als Kernelfunktion $K(\boldsymbol{x}_i, \boldsymbol{x}_j)$ bezeichnet und misst die Ähnlichkeit der Featurevektoren im Feature- und ursprünglichen Eingaberaum. Die lineare SVM entspricht damit dem Spezialfall der SVM, in dem als Kernelfunktion $K(\boldsymbol{x}_i, \boldsymbol{x}_j)$ das Skalarprodukt der Datenpunkte $K(\boldsymbol{x}_i, \boldsymbol{x}_j) = \boldsymbol{x}_i^T \boldsymbol{x}_j$ genutzt wird. Der Eingaberaum und der Featureraum sind in diesem Fall identisch. Durch Verwendung einer Kernelfunktion können damit auf dem Skalarprodukt basierende Algorithmen effizient auf hochdimensionale (sogar unendlich-dimensionale) Featureräume verallgemeinert werden.

Für eine gegebene Trainingsdatenmenge lässt sich anhand einer Kernelfunktion auch die zugehörige $n \times n$-Kernelmatrix $\boldsymbol{K}$ mit $K_{ij} = K(\boldsymbol{x}_i, \boldsymbol{x}_j)$ berechnen. Damit ein gültiger Kernel vorliegt, müssen bestimmte Bedingungen erfüllt sein [106]. Ein häufig in Zusammenhang mit SVM genutzter Kernel, der diese Bedingungen erfüllt, ist der Gauß'sche Kernel. Die Abbildungsvorschrift dieses Kernels lautet [97, 102, 105]:

$$K(\boldsymbol{x}_i, \boldsymbol{x}_j) = \exp\left(-\frac{\| \boldsymbol{x}_i - \boldsymbol{x}_j \|^2}{2\sigma^2} \right) \tag{45}$$

und entspricht einer Abbildung in einen unendlich dimensionalen Featureraum. Für ähnliche – nahe beieinander liegende – $\boldsymbol{x}_i, \boldsymbol{x}_j$ ist der Wert des Kernels nahe bei 1, wohingegen bei Unähnlichkeit ein Wert nahe 0 angenommen wird.

Die durch $\boldsymbol{\phi}$ erreichte hochdimensionale Abbildung erhöht die Wahrscheinlichkeit einer Trennung der Datenpunkte mit der SVM. Es gibt allerdings keine Garantie, dass dieser Fall eintritt. Um auch mit Datensätzen umgehen zu können, die auch unter Verwendung einer Kernelfunktion nicht separierbar sind, muss ein gewisses Maß an Fehlklassifikationen geduldet werden. Dies verringert gleichzeitig auch die Empfindlichkeit des Algorithmus gegenüber Ausreißern. Dazu wird das Optimierungsproblem nochmal um einen Parameter C erweitert und die Nebenbedingungen durch Einführung von Schlupfvariablen θ_i wie folgt gelockert [97, 102, 105]:

$$\min_{\gamma, \boldsymbol{w}} \quad \frac{1}{2} \parallel \boldsymbol{w} \parallel^2 + C \sum_{i=1}^{n} \theta_i \tag{46}$$

mit

$$y_i(\boldsymbol{w}^T \boldsymbol{x}_i + b) \geq 1 - \theta_i, \quad i = 1, \ldots, n \tag{47}$$

$$\theta_i \geq 0, \quad i = 1, \ldots, n. \tag{48}$$

Durch Einführung der Schlupfvariablen wird weiterhin ein Rand bevorzugt, der die Trainingsdaten korrekt klassifiziert. Allerdings erfolgt eine Abschwächung der Nebenbedingungen. Für Fehlklassifikation und für Datenpunkte, die zwar korrekt klassifiziert werden, aber innerhalb des Randes liegen, können Bestrafungen erfolgen. Mit dem Parameter C wird die Gewichtung zwischen den Optimierungszielen gesteuert, einen möglichst breiten Rand um die Trennebene zu erhalten und sicherzustellen, dass möglichst viele Datenpunkte einen funktionalen Rand haben, der größer als 1 ist. Lagrangeansatz und Umformung mit Hilfe der KKT-Bedingungen führen dann wieder zu einem dualen Optimierungsproblem in der Form [97, 102, 105]:

$$\max_{\boldsymbol{\alpha}} \quad W(\boldsymbol{\alpha}) = \sum_{i=1}^{n} \alpha_i - \frac{1}{2} \sum_{i=1}^{n} \sum_{j=1}^{n} y_i y_j \alpha_i \alpha_j K(\boldsymbol{x}_i, \boldsymbol{x}_j) \tag{49}$$

mit

$$0 \leq \alpha_i \leq C, \quad i = 1, \ldots, n$$

$$\sum_{i=1}^{n} \alpha_i y_i = 0. \tag{50}$$

Im dualen Problem verschwinden die Schlupfvariablen und die Konstante C ergibt eine Beschränkung der Lagrange-Multiplikatoren α_i [97, 102, 105]. Die Berechnung von b^* muss allerdings noch angepasst werden. Gleichung (44) ist nicht mehr anwendbar [107]. Die Entscheidungsfunktion resultiert zu

$$\widehat{k}(\boldsymbol{x}) = \operatorname{sign}\left(\boldsymbol{w}^T \phi(\boldsymbol{x}) + b\right) = \operatorname{sign}\left(\sum_{i=1}^{n} y_i \alpha_i K(\boldsymbol{x}_i, \boldsymbol{x}) + b\right). \tag{51}$$

Eine effiziente Lösung der allgemeinen Form der SVM, die deutlich performanter gegenüber der Anwendung einer quadratischen Programmierung ist, ermöglicht der von Platt entwickelte Algorithmus Sequential Minimal Optimization (SMO) [107].

Die beschriebene allgemeine Form der SVM kann bisher lediglich zur Lösung von Zweiklassenklassifikationsproblemen genutzt werden. Eine Verallgemeinerung auf Mehrklassenklassifikationsprobleme erfolgt in der Literatur anhand der Ansätze „One-against-all SVM“, „Pairwise SVM“, “Error-correcting Output Code SVM“ und „All-at-once SVM [102]. In dieser Arbeit wird zunächst der Ansatz „Pairwise SVM“ genutzt. Darin wird das Mehrklassenklassifikationsproblem in mehrere Zweiklassenklassifikationsprobleme (für C Klassen in $C(C-1)/2$ Zweiklassenklassifikationsprobleme) überführt. Diese Zweiklassenklassifikationsprobleme werden zunächst, wie oben beschrieben gelöst. Auf Basis der Ausgaben $\widehat{k}(\boldsymbol{x})$ der paarweisen Klassifikatoren wird die Ausgabe $\hat{y}$ der Mehrklassen-SVM bestimmt [102, 108]. Die genannten Ansätze zur Realisierung von Mehrklassen-SVM berücksichtigen nicht eine eventuelle Ordnung der Klassen. Da in dieser Arbeit mit einer SVM ordinale Klassen prognostiziert werden sollen, wird nebem dem Ansatz „Pairwise SVM“ auch der von Frank und Hall vorgeschlagene allgemeine Ansatz zur Lösung von ordinalen Mehrklassenklassifikationsproblemen verwendet [103]. Auch bei diesem Ansatz wird das Mehrklassenklassifikationsproblem in mehrere Zweiklassenklassifikationsprobleme (für C Klassen in $C-1$ Zweiklassenklassifikationsprobleme) zerlegt. Bei der Formulierung der einzelnen Zweiklassenklassifikationsprobleme wird allerdings die Ordnung der Klassen berücksichtigt: (Klasse 1 gegen $2, 3, 4, 5, \ldots, C$), (Klassen $1, 2$ gegen $3, 4, 5, \ldots, C$), $\ldots$, (Klassen $1, 2, 3, 4, 5, \ldots, C-1$ gegen C). Die Ausgaben der einzelnen zur Lösung dieser Zweiklassenklassifikationsprobleme modellierten Klassifikatoren werden dann nach Frank und Hall auf Basis ihrer Wahrscheinlichkeitsausgaben zur Schätzung der Wahrscheinlichkeiten der ursprünglichen ordinalen Klassen kombiniert [103]:

$$\begin{aligned} p(1) &= 1 - p(y > 1), \\ p(i) &= p(y > c-1) - p(y > c), \quad 1 < c < C \\ p(C) &= p(y > C-1). \end{aligned} \tag{52}$$

Die Klasse mit der höchsten Wahrscheinlichkeit bildet die Ausgabe $\hat{y}$ des Klassifikators und wird dem zu klassifizierenden Datenpunkt zugeordnet. Zur Bestimmung von Wahrscheinlichkeitsausgaben für eine SVM, die zur Kombination der einzelnen Klassifikatoren dienen, wird in dieser Arbeit die Implementierung in [108] genutzt. Darin werden die paarweisen Klassenwahrscheinlichkeiten auf Basis der Ausgaben $\widehat{k}(\boldsymbol{x})$ der paarweisen Klassifikatoren mit Hilfe einer Sigmoid-Funktion geschätzt [108].

6.2 MODELLIERUNG UND PROGNOSE DER KOMBINIERTEN LABEL

6.2.1 *Training einer Support Vector Machine mit kombinierten Labeln*

Dieser Abschnitt erläutert den wohl naheliegendsten Ansatz zur Prognose der kombinierten Label anhand eines SVM-Klassifikators. In diesem werden zur Modellierung des Klassifikators neben den zur Verfügung stehenden Datenpunkten ausschließlich die kombinierten Label verwendet (Konzept 1). Eine schematische Übersicht zur Bildung des Klassifikators anhand der kombinierten Label ist in Abbildung 25 dargestellt. Die N_E Einzellabel werden demnach bereits auf der Datenebene durch Anwendung einer Kombinationsregel vorverarbeitet (Abschnitt 5.4). Nach erfolgter Fusion der Einzellabel liegt die Trainingsdatenmenge T analog zu Abschnitt 6.1.2 direkt in der Form $T = \{(\boldsymbol{x}_i, y_i)\}$ $(i = 1, \ldots, n)$ vor. Die anschließende Modellierung des Klassifikators reduziert sich damit auf die Formulierung und Lösung des Klassifikationsproblems aus Abschnitt 6.1.2.

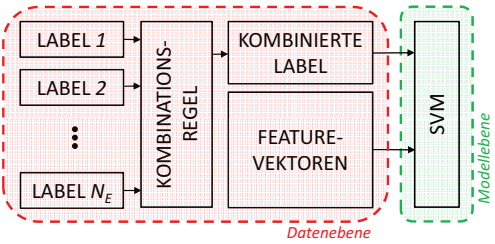


Abbildung 25: Schematische Darstellung der Modellierung des SVM-Klassifikators mit kombinierten Labeln (Konzept 1).

6.2.2 *Training einer Support Vector Machine mit Einzellabeln*

Im voranstehenden Abschnitt wurde die Modellierung des SVM-Klassifikators auf Basis der kombinierten Label beschrieben. Dazu gegensätzlich wird nun ein Ansatz vorgestellt, der zur Modellierung direkt auf die N_E unsicheren Einzellabel zurückgreift. Der Unterschied zum Training der SVM mit kombinierten Labeln liegt darin, dass für ein Objekt die auf Basis einer Kombinationsregel vorgenommene dichotome Zuordnung zu einer Klasse aufgegeben und stattdessen alle zur Verfügung stehenden Einzellabel zum Training der SVM genutzt werden. Die Trainingsdatenmenge T des Klassifikators ist entsprechend zu erweitern.

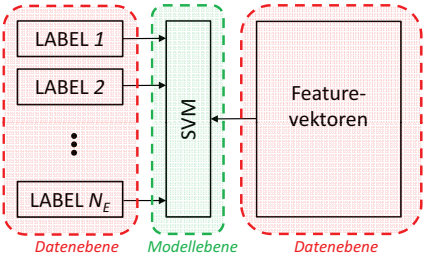


Abbildung 26: Schematische Darstellung der Modellierung des SVM-Klassifikators mit Einzellabeln (Konzept 2).

Abbildung 26 stellt die Umsetzung dieses Modellierungsansatzes schematisch dar (Konzept 2). Zur Umsetzung wird beim Training der SVM jeder Featurevektor $\boldsymbol{x}_i$ genau N_E-mal genutzt. Dabei wird jedes Einzellabel genau einmal mit $\boldsymbol{x}_i$ kombiniert. Die Trainingsdatenmenge hat dann die Form $T = \{(\boldsymbol{x}_i, y_i)\}$ $(i = 1, \ldots, n \cdot N_E)$. Durch dieses Vorgehen wird sich die Lage der Hyperebene im Featureraum stark an den Einzellabeln orientieren, die jeweils am häufigsten für die duplizierten Datenpunkte in den Trainingsdaten enthalten sind. Die Modellierung des Klassifikators reduziert sich wieder auf die Formulierung und Lösung des Klassifikationsproblems aus Abschnitt 6.1.2. Es ist zu erwarten, dass die Umsetzung dieses Ansatzes zu ähnlichen Ergebnissen führt, wie man bei graduellen Label erwarten dürfte, da alle Datenpunkte gleich oft dupliziert werden.

6.2.3 *Bildung eines Ensembles mit Support Vector Machines*

In den beiden bisher verfolgten Ansätzen zur Modellierung des Klassifikators wurde eine einzige SVM verwendet. Ergänzend dazu wird nun ein dritter Ansatz entwickelt, in dem die Zusammenführung der Information erst auf der Modellebene an der Ausgabe mehrerer Einzelklassifikatoren erfolgt. In der Literatur sind diese Ansätze als Klassifikator-Ensembles bekannt. Dies betrifft zum einen die Zusammenführung von Einzelklassifikatoren, die verschiedene oder unterschiedlich gewichtete Eingangsdaten nutzen (z. B. bagging, boosting), aber auch den Fall der Zusammenführung von Klassifikatoren, die auf der Basis verschiedener Ausgabewerte modelliert wurden. Die Ensembletechniken unterscheiden sich dabei unter anderem in der Art der Gewichtung der Einzelklassifikatoren zur Bestimmung ihrer Ausgabe. Eine schematische Übersicht zur Bildung des Ensembles zeigt Abbildung 27.

Anhand der für jedes Objekt zur Verfügung stehenden N_E Einzellabel und der Featurevektoren $\boldsymbol{x}_i$ $(i = 1, \ldots, n)$ werden zunächst N_E Trainingsdatenmengen T_j der Form

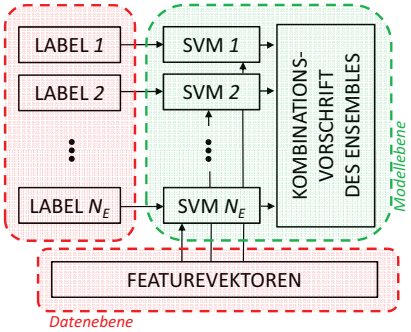


Abbildung 27: Schematische Darstellung der Modellierung des SVM-Ensembles.

$T_j = \{(\boldsymbol{x}_i, y_{ij})\}$ $(i = 1, \ldots, n)$ gebildet. Auf Basis dieser Trainingsdaten erfolgt die Modellierung von N_E einzelnen SVM (Formulierung und Lösung des Klassifikationsproblems wieder analog zu Abschnitt 6.1.2). Die Modellebene besteht damit, wie beschrieben, nicht mehr nur aus einer einzelnen SVM. Die Ausgaben der SVM sind so zu gewichten, dass die kombinierten Label bestmöglich prognostiziert werden können. Einen standardmäßigen Ansatz liefert die Bestimmung der Gewichtung anhand der (kreuzvalidierten) Güten auf den Trainingsdaten der einzelnen SVM (Konzept 3*a*). In diesem Fall finden die kombinierten Label allerdings in der Berechnung keine Berücksichtigung. Es wird lediglich dem Klassifikator, der auf den Trainingsdaten die höchste Klassifikationsgüte aufweist, auch anteilig der größte Einfluss auf die Gesamtausgabe des Ensembles zugestanden. Eine Überanpassung der Klassifikatoren auf die Trainingsdaten überträgt sich folglich direkt auf die Ausgabequalität des Ensembles. Im Gegensatz zum standardmäßigen Ansatz wird daher hier folgende Idee verfolgt: Bei der Ermittlung der Gewichte der SVM im Ensemble, soll jeweils die Ähnlichkeit zwischen den N_E Einzellabeln der zugeordneten Trainingsdatenmenge T_j und den kombinierten Labeln berücksichtigt werden. Dies bedeutet, dass der Einfluss der Ausgabe einer SVM auf die Ausgabe des Ensembles mit zunehmender Ähnlichkeit der SVM-Ausgabe zu den kombinierten Labeln größer werden soll. Da in dieser Arbeit ausschließlich Label von ordinalem Charakter betrachtet werden, kann die Zusatzinformation, die in der Ordnung der Klassen enthalten ist, zur Bestimmung der Ähnlichkeit genutzt werden (Konzept 3*b*). Ein in dieser Arbeit bereits genutztes Kalkül zur Berechnung der Ähnlichkeit zwischen ordinalen Labelvektoren ist Cohen's gewichtetes Kappa κ_w (Abschnitt 4.1.4). Auf Grundlage dieser Überlegungen lassen sich die Ähnlichkeiten $\kappa_{w,j}$ $(j = 1, \ldots, N_E)$ zwischen den Ausgaben der einzelnen SVM-Klassifikatoren des Ensembles und den kombinierten Labeln bestimmen. Die Ermittlung der Gewichtungen

w_j $(j = 1, \ldots N_E)$ im Ensemble resultieren dann aus der Normierung der $\kappa_{w,j}$ mit $\sum_{j=1}^{N_E} \kappa_{w,j}$:

$$w_j = \frac{\kappa_{w,j}}{\sum\limits_{j=1}^{N_E} \kappa_{w,j}}. \tag{53}$$

Die Ausgabe des Ensembles wird auf Basis der Werte w_j mit dem gerundeten, gewichteten Mittel der Einzelausgaben $\hat{y}_j$ $(j = 1, \ldots, N_E)$ im Ensemble bestimmt:

$$\hat{y}_{Ens.} = \lfloor 0,5 + \sum_{j=1}^{N_E} w_j \cdot \hat{y}_j \rfloor. \tag{54}$$

6.3 EMPIRISCHE EVALUATION DER MODELLIERUNGSANSÄTZE

Nachdem in den voranstehenden Abschnitten die mathematischen Grundlagen von SVM erläutert und darauf aufbauend drei Ansätze zur Modellierung von Klassifikatoren vorgestellt wurden, erfolgt nun eine empirische Untersuchung dieser Ansätze anhand der Daten von 300 realen NS-Netzen. Abschnitt 6.3.1 gibt dazu zunächst eine Übersicht zu den verfügbaren Beispieldaten. Im Anschluss daran erfolgt in Abschnitt 6.3.2 eine Analyse der Featurerelevanz zur Klassifikation. In Abschnitt 6.3.3 wird darauf aufbauend ein empirischer Vergleich der Klassifikationsgüten der drei Modellierungsansätze durchgeführt. In diesem Zuge erfolgt darüber hinaus eine Analyse des Einflusses der zur Modellierung verwendeten Features sowie der Gewichtung der Simulationsergebnisse bei Anwendung der EIDMR auf die Klassifikationsgüte des Klassifikators. Weiterhin wird untersucht, ob der SVM-Klassifikator durch Berücksichtigung von ordinalen Informationen bei der Lösung des Mehrklassenklassifikationsproblems weiter verbessert werden kann.

6.3.1 *Übersicht zu den empirischen Daten*

Als Features zur Modellierung der Klassifikatoren können die netzspezifischen, die graphspezifischen oder eine Menge aus beiden dieser Features verwendet werden (Kapitel 3). Zusätzlich ist auch der Einsatz weiterer simulationsspezifischer Features, die bei der simulationsbasierten Klassifikation erhoben wurden (z. B. Parameter der Weibullverteilungen), möglich. Für jedes Netz stehen darüber hinaus fünf Einzellabel von Experten (Abschnitt 4.1) und ein simulationsbasiertes Einzellabel (Abschnitt 4.2) zur Verfügung.

Um den Einfluss der vier Kombinationsregeln aus Kapitel 5 auf die Prognosegüte des SVM-Klassifikators zu untersuchen, wurden in Abschnitt 5.4 alle vier Kombinationsregeln (DSR, MR, IDMR und EIDMR) angewendet, um aus den oben genannten sechs Einzellabeln für jedes Netz ein kombiniertes Label zu generieren. Aus der Anwendung der Kombinationsregeln resultieren damit für jeden Datenpunkt zusätzlich zu den sechs Einzellabeln vier kombinierte Label. Die Label, die für die NS-Netze anhand der stochastischen Simulationen bestimmt wurden, sind bei Anwendung der Kombinationsregeln wie ein einzelner Experte berücksichtigt worden (bei gleicher Gewichtung wie ein einzelner Experte). Bei der Anwendung der EIDMR zur Kombination der Einzellabel wurden die netzspezifischen Features verwendet (Abschnitt 5.4).

6.3.2 *Relevanz der Features bei der Klassifikation*

In Kapitel 3 wurden mit den netzspezifischen und den graphspezifischen Features zwei Datengrundlagen zur Charakterisierung von NS-Netzen gebildet. Die erstgenannten Features wurden in Abschnitt 4.1 zur expertenbasierten Klassifikation von NS-Netzen genutzt. Es ist daher zu erwarten, dass die Label der Experten direkt von den Ausprägungen der netzspezifischen Features abhängen. Dabei bildeten diese eine wichtige Basisinformation für die Klassifikationsentscheidung der Experten. Weiterhin wurden den Experten auf der Grundlage der netzspezifischen Features Indikatorfunktionen und Klassifikationsindikatoren bereitgestellt. Die Relevanz der einzelnen Features hinsichtlich ihrer Entscheidungen wurde im Rahmen der Befragung sogar von den Experten in Form einer gewichteten Rangfolge abgefragt. Inwieweit die Featureauswahl dazu beiträgt, dass die gelabelten Datenpunkte im durch diese Features aufgespannten Raum so gut wie möglich separiert werden können, lässt sich anhand von Verfahren der Feature Selection untersuchen.

Vorgehen zur Bewertung der Relevanz der Features:
Um die Relevanz von Features hinsichtlich einer erfolgten Klassifikationsentscheidung ex post zu untersuchen, können verschiedene Methoden der Feature Selection angewendet werden. In dieser Arbeit werden dazu Filterverfahren angewendet. Im Gegensatz zu Wrapper-Verfahren, die jede potentielle Featuremenge auf den Klassifikator anwenden, wird keine Abhängigkeit zum später verwendeten Klassifikator berücksichtigt [109]. Der Hauptvorteil bei Anwendung von Filterverfahren liegt darin, dass ihre Ergebnisse mit wenig Aufwand berechenbar sind. Allerdings müssen die Ergebnisse bei der Modellierung eines Klassifikators hinsichtlich der erreichten Klassifikationsgüte noch evaluiert werden. Weiterhin werden Ergebnisse verschiedener Filterverfahren wegen der ihnen zu Grunde liegenden, unterschiedlichen Bewertungskriterien und Vorgehensweisen mehr oder weniger stark differieren. Aus diesem Grund ist für eine Bewertung und Auswahl von Features die Nutzung von zwei (oder mehreren) verschiedenen Filterverfahren empfehlenswert. Auf Basis der angestellten Überlegungen wird im Folgenden eine Bewer-

tung der Datengrundlage hinsichtlich der Prognose der kombinierten Label (EIDMR) unter Anwendung des Relief-F-Algorithmus [110] und der χ^2-Methode [111] vorgenommen. Zur Bewertung der Feature-Relevanz mit den beiden geannten Filterverfahren werden die Punkte

- Differenzierung der netzspezifischen und graphspezifischen Features untereinander,
- Bewertung der Relevanz von zusätzlichen simulationsspezifischen Features (z. B. Parameter der Weibullverteilungen) für die expertenbasierte Klassifikation sowie
- Bewertung der Kombination von netz- und graphspezifischen Features zu einer Featuremenge

untersucht. Diese Untersuchung ist von Bedeutung, um die Wichtigkeit einzelner Features zur Beurteilung des DEA-Aufnahmevermögens von NS-Netzen (Implikationen für die Netzentwicklung) und Hinweise zur Verwendung einzelner Features bei der Modellierung von SVM-Klassifikatoren zu erhalten.

Anhand der im ersten Punkt genannten Unterscheidung der Relevanz der Features untereinander können Rückschlüsse auf die Datenerhebung gezogen werden. Kommt einem Feature ex post nur eine sehr geringe Bedeutung bei der Klassifikation zu, kann es z. B. bei der Datenerhebung für weitere Netze vernachlässigt werden. Weiterhin liefert insbesondere die Bewertung der netzspezifischen Features untereinander interessante Hinweise zur zukünftigen Entwicklung und Stärkung von NS-Netzstrukturen. Simulationsspezifische Features können, bei hinreichend hoher Relevanz, z. B. zur Prognose der expertenbasierten Klassifikationseinschätzungen eingesetzt werden. Eine Bewertung der Kombination von netz- und graphspezifischen Features zu einer Featuremenge ist zur Modellierung eines SVM-Klassifikators hilfreich. Auf diese Weise erhält man bereits im Vorfeld der Modellierung Indikationen, welche Features dieser Featuremenge einen positiven Beitrag auf die Klassifikationsgüte haben könnten.

Die Analyse wird auf die mit der EIDMR kombinierten Label eingeschränkt. Diese Einschränkung kann getroffen werden, da alle vier kombinierten Label eines Datenpunktes jeweils aus denselben sechs Einzellabeln (fünf Expertenaussagen und ein Simulationsergebnis bei gleicher Gewichtung, Abschnitt 5.4) hervorgehen und folglich nur geringfügige Unterschiede in der Feature-Relevanz bei Verwendung der weiteren kombinierten Label zu erwarten sind. Vor Analyse der Features wurde jeweils eine z-Transformation der Features durchgeführt.

Bewertung der Relevanz der netzspezifischen Features:
Tabelle 17 zeigt die Ergebnisse der Bewertung der netzspezifischen Features. Die ausgewiesenen Ergebnisse des Relief-F-Algorithmus sowie die der χ^2-Methode sind auf

das Intervall $[0, 1]$ normiert. Die Einzelergebnisse der Ex-Post-Analyse verdeutlichen, dass beide Filterverfahren die Relevanz der Features untereinander ähnlich bewerten. Die durch den Relief-F-Algorithmus bestimmten Bewertungen nehmen im Vergleich zu denen der χ^2-Methode deutlich stärker ab. Zusätzlich wurde eine weitere Bewertung vorgenommen, indem die von den Experten im Rahmen der Befragung vergebenen Gewichtungen der Features (wurde von jedem Experten für die ersten fünf Features seiner Rangfolge angegeben, Abschnitt 4.1.5) summiert und anschließend ebenfalls auf das Intervall $[0, 1]$ normiert worden sind. Dies ermöglicht einen direkten Vergleich der Voreinschätzung der Experten zum Einfluss der Features mit den Ergebnissen der ex post bewerteten Feature-Relevanz.

FEATURE	RELIEF-F	χ^2	EXPERTEN
Vermaschungsgrad	1,00	1,00	0,55
Summe der Transformatornennleistungen	0,68	0,53	0,40
Anzahl der Hausanschlüsse	0,32	0,57	0,80
Summe der vermaschten Leitungslängen	0,32	0,51	0,60
Summe der Leitungslängen	0,12	0,38	0,55
Anzahl der Kabelverteilerschränke	0,10	0,40	0,40
Anzahl der Transformatorstationen	0,17	0,10	0,15
Maximaler Stationsradius	0,06	0,25	0,55
Durchschnittlicher Stationsradius	0,06	0,20	1,00
Anteil Leitungstyp NAYY 4x150 mm^2	0,00	0,00	0,00

Tabelle 17: Übersicht zur Relevanz der netzspezifischen Features untereinander.

Der Vermaschungsgrad eines NS-Netzes wird sowohl durch den Relief-F-Algorithmus, als auch durch die χ^2-Methode als Feature mit höchster Relevanz bewertet. Auch in der Bewertung der Experten kommt diesem Feature eine überdurchschnittlich hohe Bewertung zu, so dass es insgesamt als Feature mit hoher Relevanz zu bewerten ist. Weitere Features, denen sowohl durch beide Filterverfahren, als auch durch die Expertenbewertung eine hohe Relevanz zugeordnet wird, sind die Summe der Transformatornennleistungen, die Anzahl der Hausanschlüsse und die Summe der vermaschten Leitungslängen. Den Features „Summe der Leitungslängen“ und „Anzahl der Kabelverteilerschränke“ kommt in allen drei Bewertungen jeweils eine mittlere Relevanz zu. Für das Feature „durchschnittlicher Stationsradius“ liegen sehr konträre Einschätzungen aus Ex-post-Analyse und Expertenbewertung vor. Auf Grund dieser starken Abweichung werden die Ränge und Gewichtungen, die diesem Feature durch die einzelnen Experten zugeordnet wurden, nochmal in Tabelle 18 gegenübergestellt. Es zeigt sich, dass vier Experten dem durchschnittlichen Stationsradius sowohl einen hohen Rang, als auch eine hohe Gewichtung zugeordnet haben. Nurch durch Experte 1 wird dieses Feature als weniger relevant bewertet. Die aus den Gewichtungen der Experten abgeleitete hohe Bewertung des durchschnittlichen Stationsradius in Tabelle 17 wird damit durch die Einzelbetrachtung der Expertenbewertungen gestützt.

EXPERTE	RANG	GEWICHTUNG
Experte 1	9	0
Experte 2	3	4
Experte 3	1	5
Experte 4	3	4
Experte 5	1	6

Tabelle 18: Übersicht zur Expertenbewertung des durchschnittlichen Stationsradius.

Die insgesamt niedrigste Relevanz weist der Anteil des Leitungstyps NAYY 4x150 mm^2 auf. Dieses Ergebnis kann als Indiz für eine mögliche Vernachlässigung dieses Features bei der Klassifikation gewertet werden. Den übrigen, nicht explizit erwähnten Features kommt aus der Bewertung eine mittlere bis niedrige Relevanz zu. Es lässt sich festhalten, dass bei der Weiterentwicklung von ländlichen und vorstädtischen NS-Netzen zur Erhöhung der DEA-Leistung der Vermaschung und der im Netz verlegten Leitungslänge eine besondere Rolle zukommt. Dies bedeutet, dass bei Verlegung weiterer Leitungen diese immer auch zur Erhöhung der Vermaschung des Netzes beitragen sollten. Aus physikalischer Sicht ist dieses Ergebnis insofern einleuchtend, als dass mehr Leitermaterial und eine höhere Vermaschung im Netz zu einer Verringerung der Netzimpedanz und folglich zu einer Entlastung der Betriebsmittel sowie zu einer Reduzierung der Spannungsniveaus im Falle starker Erzeugung beitragen.

Bewertung der Relevanz der graphspezifischen Features:
Analog zu den diskutierten Ergebnissen der netzspezifischen Features sind in Tabelle 19 die Ergebnisse des Relief-F-Algorithmus und der χ^2-Methode für die graphspezifischen Features ausgewiesen. Auch hier ist zu beobachten, dass sich die Ergebnisse beider Methoden nur geringfügig unterscheiden bzw. nicht widersprechen. Eine Einschätzung der Experten liegt für diese Features auf Grund ihres abstrakten Charakters nicht vor, so dass ein Vergleich zwischen Vorabeinschätzung und Ex-Post-Bewertung an dieser Stelle nicht möglich ist. Die Darstellung der Ergebnisse ist zur besseren Übersicht auf die zehn Features beschränkt worden, die durch beide Filterverfahren überdurchschnittlich hoch bewertet werden. Dem Feature „Betrag des zweitgrößten Eigenwerts“ wird dabei jeweils eine sehr hohe Relevanz zugeordnet. Die beiden Features „spektraler Radius“ und „durchschnittliche Pfadlänge“ sind im Ergebnis ebenfalls von hoher Relevanz. Der betragsmäßig zweitgrößte Eigenwert und der spektrale Radius der Graphen sind beides Features, die aus der Spektralen Graphentheorie [43, 44] abgeleitet worden sind. Features die durch beide Filterverfahren sehr niedrig bewertet werden und folglich bei der Klassifikation eher vernachlässigt werden könnten, sind z. B. die „minimale Knotenzahl der verbundenen Komponenten“, „die minimale effektive Exzentrität“ und die „prozentuale Anzahl der Artikulationspunkte“.

FEATURE	RELIEF-F	χ^2
Betrag zweitgrößter Eigenwert	0,61	1,00
Spektraler Radius	0,56	0,99
Durchschnittliche Pfadlänge	0,52	0,85
Knotenzahl größter verb. Untergraphen	1,00	0,40
Durchschnittlicher Knotengrad	0,32	0,81
Gesamte Kantenanzahl	0,27	0,62
Energie	0,27	0,62
Durchschn. Kantenzahl verb. Komponenten	0,22	0,71
Anzahl verschiedenartiger Eigenwerte	0,28	0,55
Gesamte Knotenanzahl	0,28	0,53

Tabelle 19: Übersicht zur Relevanz der graphspezifischen Features untereinander.

Bewertung der Relevanz von zusätzlichen simulationsspezifischen Features für die expertenbasierte Klassifikation:
Zusätzlich zu den netzspezifischen und den graphspezifischen Features werden nun aus der stochastischen Simulation abgeleitete Features bewertet. Diese könnten z. B. eingesetzt werden, um einen möglichen Mehrwert zur Prognose der expertenbasierten Klassifikationseinschätzungen mit einem SVM-Klassifikator zu generieren. Eine Verwendung von simulationsspezifischen Features zur Prognose von kombinierten Labeln, zu deren Berechung die Simulationsergebnisse verwendet wurden, ist sowohl aus methodischer, als auch aus Sicht der Anwendung nicht sinnvoll, da zur Bestimmung der simulationsspezifischen Features eine stochastische Simulation für jedes NS-Netz erforderlich wäre. Im Detail wurden im Rahmen der Simulationen die folgenden Features erhoben:

- Formparameter a der Weibullverteilung,
- Skalierungsparameter b der Weibullverteilung,
- Verschiebungsparameter v der Weibullverteilung,
- Anteil der Verletzungen von Kriterium 1 (3%-Kriterium),
- Anteil der Verletzungen von Kriterium 2 (Betriebsmittelbelastung) und
- geschätzter Wert des Korrelationskoeffizienten R_w einer Quantil-Quantil-Grafik.

Zur Bewertung dieser Features wurde ihre Relevanz hinsichtlich der fusionierten fünf Einzellabel der Experten ($N_E = 5$ bei gleicher Gewichtung, Anwendung der EIDMR, Abschnitt 5.4) ebenfalls mit dem Relief-F-Algorithmus und der χ^2-Methode untersucht. Dabei wurden die Features jeweils einmal mit den netzspezifischen und einmal mit den

graphspezifischen Features zu einer gemeinsamen Featuremenge zusammengeführt, um den möglichen Mehrwert der simulationsspezifischen Features gegenüber den anderen Features bewerten zu können. In den Ergebnissen dieser Analyse zeigte sich, dass die simulationsspezifischen Features nur eine sehr geringe Relevanz im Vergleich zu den netzspezischen Features haben. Bei Anwendung beider Filterverfahren belegten die simulationsspezifischen Features jeweils die letzten sechs Ränge. Ein ähnliches Bild ergab der Vergleich mit den graphspezifischen Features. In den Ergebnissen beider Filterverfahren befand sich kein simulationsspezifisches Feature unter den ersten 15 Rängen. Zusammenfassend ist damit kein Mehrwert durch die Verwendung der simulationsspezifischen Features bei der Modellierung eines SVM-Klassifikators zu erwarten.

Bewertung der Kombination von netz- und graphspezifischen Features:
Es wird nun eine Bewertung der Kombination von netz- und graphspezifischen Features zu einer gemeinsamen Featuremenge vorgenommen. Diese Bewertung ist für eine anschließende Modellierung eines SVM-Klassifikators hilfreich, da sie Indikationen zur Selektion bzw. Kombination bestimmter Features liefert. Die Ergebnisse des Gesamtvergleichs sind in Tabelle 20 in bereits bekannter Form gezeigt.

FEATURE	RELIEF-F	χ^2
Betragsmäßig zweitgrößter Eigenwert	0,63	1,00
Spektraler Radius	0,58	0,99
Vermaschungsgrad	0,66	0,81
Durchschnittliche Pfadlänge	0,52	0,86
Knotenzahl größter verb. Untergraphen	1,00	0,44
Summe der Transformatornennleistungen	0,53	0,58
Anzahl der Hausanschlüsse	0,42	0,66
Durchschnittlicher Knotengrad	0,33	0,82
Summe der vermaschten Leitungslängen	0,38	0,70
Gesamte Kantenanzahl	0,36	0,64
Energie	0,36	0,64
Anzahl verschiedenartiger Eigenwerte	0,37	0,57
Gesamte Knotenanzahl	0,37	0,56
Durchschn. Kantenzahl verb. Komponenten	0,26	0,73
Durchschn. Knotenzahl verb. Komponenten	0,25	0,65

Tabelle 20: Übersicht zur Relevanz der Gesamtheit der Features untereinander.

Die Darstellung ist dabei auf die ersten 15 Features beschränkt, denen durch beide Filterverfahren jeweils eine überdurchschnittlich hohe Bewertung zukommt. Bei Betrachtung der Ergebnisse fällt auf, dass sich in dem gezeigten Ausschnitt lediglich vier der zehn netzspezifischen Features befinden. Die restlichen 11 Features sind graphspezifische Features. Zwei Features mit sehr hoher Relevanz sind analog zu Tabelle 19 der betragsmäßig zweitgrößte Eigenwert und der spektrale Radius der Graphen. Die durch-

schnittliche Pfadlänge kommt in der Betrachtung aller Features ebenfalls eine hohe Relevanz zu. Die Einordnung der graphspezifischen Features in der Gesamtbewertung wird damit durch das Ergebnis der Bewertung der graphspezifischen Features untereinander gestützt. Ein weiteres Feature mit hoher Relevanz in der Gesamtbewertung ist der Vermaschungsgrad, der aus der Menge der netzspezifischen Features stammt. Auch dieses Ergebnis korrespondiert mit der Bewertung der Features untereinander, da der Vermaschungsgrad in Tabelle 17 die höchste Relevanz aufweist. Damit stehen die Ergebnisse der Bewertung der Gesamtheit der Features jeweils nicht im Widerspruch zur Bewertung Features untereinander und liefern konsistente Hinweise zur Kombination und Vernachlässigungen von Features bei der Modellierung eines SVM-Klassifikators.

Die hier gezogenen Vergleiche hinsichtlich einer Selektion und Kombination von Features zur Klassifikation enthalten erste Implikationen für die Modellierung eines SVM-Klassifikators. Mit Hilfe welcher Features sich die beste Klassifikationsgüte erreichen lässt, wird im Folgenden experimentell näher untersucht.

6.3.3 *Vergleich der Klassifikationsgüten*

In diesem Abschnitt werden Untersuchungsergebnisse zur Qualität der Modellierung von SVM-basierten Klassifikatoren vorgestellt. Das primäre Ziel liegt darin, eine möglichst hohe Klassifikationsgüte hinsichtlich der kombinierten Label (Abschnitt 5.4) zu erreichen. Dieser Zielsetzung liegt die Annahme zu Grunde, dass eine zur Fusion der Einzellabel angewandte Kombinationsregel nicht nur eine hohe Güte zur Detektion der (in diesem Fall unbekannten) wahren Klassen aufweisen, sondern auch zu einer guten Separierbarkeit der kombinierten Label im Featureraum beitragen sollte, um im Ergebnis einen guten Klassifikator zu erhalten. In den empirischen Experimenten werden zur Modellierung eines SVM-Klassifikators folgende Aspekte untersucht:

- Training und Evaluation eines SVM-Klassifikators unter Verwendung aller Einzellabel,
- Training und Evaluation je eines SVM-Klassifikators für jedes Einzellabel,
- Einfluss der vorgestellten Klassifikationskonzepte auf die Klassifikationsgüte,
- Einfluss durch Selektion und Kombination von Features auf die Klassifikationsgüte,
- Einfluss der Kombinationsregeln auf die Klassifikationsgüte,
- Einfluss der Simulationsergebnisse auf die Klassifikationsgüte und

- mögliche Optimierung durch Anwendung des Ansatzes von Frank und Hall zur Lösung von ordinalen Mehrklassenklassifikationsproblemen [103].

Bevor die Ergebnisse dieser Untersuchungen im Detail diskutiert werden, erfolgt zunächst eine Erläuterung des Vorgehens zur Durchführung der Experimente.

Vorgehen zur Durchführung der Experimente:
Zur Implementierung der SVM wurde die LIBSVM-Standardbibliothek mit einem RBF-Kernel (standard Gaußkernel) verwendet [108]. Vor der Anwendung der SVM wurden die Features jeweils mit einer z-Tranformation skaliert, um während der Optimierung eine Dominanz von Features mit großen Wertebereichen gegenüber Features mit kleinen Werten und daraus resultierende Überanpassungen des Modells zu vermeiden [95]. Zur Abschätzung guter Parameter für die SVM, und um zusätzlich eine Überanpassung auf bestimmte Datenpunkte bzw. Klassen zu vermeiden, wurde in den Experimenten zunächst jeweils eine stratifizierte dreifach Kreuzvalidierung durchgeführt. Die Arbeit mit kreuzvalidierten Parameterwerten ermöglicht einen fairen Vergleich zwischen verschiedenen Klassifikationskonzepten, da so jeweils optimierte Parameter verwendet werden und die Gefahr einer Überschätzung des Potenzials eines Konzeptes minimiert wird. Eine Stratifizierung der Mengen ist bei der Kreuzvalidierung insofern vorzusehen, als dass bei empirischen Daten die Label selten gleich verteilt auftreten. Die zufällige Anordnung der Datenpunkte in die einzelnen Teilmengen ist daher so vorzunehmen, dass diese Teilmengen die Verteilung der Label bestmöglich repräsentieren. Um eine Vielzahl an Parameterkombinationen in strukturierter Weise zu berücksichtigen, wurde während der Kreuzvalidierung eine Gittersuche durchgeführt. Bei der Gittersuche dürfen allerdings die Testdaten der Kreuzvalidierung jeweils nicht genutzt werden. Daher wurde für die Gittersuche eine zusätzliche unterlagerte Kreuzvalidierung durchgeführt, um eine robuste Parametersuche zu erhalten, die lediglich die jeweiligen Trainingsdaten der übergeordneten Kreuzvalidierung nutzt. Dieses Vorgehen führt zu einer geschachtelten Anwendung von zwei Kreuzvalidierungen. Die innere Kreuzvalidierung dient zur robusten Parametersuche auf den Trainingsdaten und die äußere jeweils zur Beurteilung der Klassifikationsgüte des Modells. Eine Besonderheit bei der Beurteilung der Klassifikationsgüte liegt darin, dass die Label einen ordinalen Charakter aufweisen, in einer SVM jedoch als nominale Klassen interpretiert und verarbeitet werden. Daher wurde die Optimierung der Parameter in der Gittersuche nicht anhand des Klassifikationsfehlers, sondern anhand der Übereinstimmung zwischen den kombinierten Labeln der Testdaten und den Ausgaben des Klassifkators unter Verwendung von Cohen's κ_w durchgeführt (Abschnitt 4.1.4). Die auf diese Weise bestimmten Parameterkombinationen wiesen insgesamt nur marginale Unterschiede auf. Daher konnte für die Experimente die Parameterkombination $C = 10$ und $\gamma = 0,03$ ausgewählt und für die Modellierung aller SVM genutzt werden. Für den Parameter σ der Kernelfunktion gemäß Gleichung (45) wird in der Standardinitialisierung der LIBSVM-Standardbibliothek γ verwendet ($\gamma = 1/(2 \cdot \sigma^2)$). Um statistisch aussagekräftige Ergebnisse zu erhalten, wurde mit diesen Parametern in allen Experimenten jeweils zehnmal eine drei-

fach Kreuzvalidierung durchgeführt. Die zehn Wiederholungen unterschieden sich in der Zusammensetzung der jeweils per Zufall gezogenen, stratifizierten Mengen, die zur Kreuzvalidierung genutzt wurden.

Benchmark SVM-Klassifikatoren zum Vergleich der Klassifikationsgüten: Um eine Ausgangsbasis für die im Rahmen der Experimente untersuchten, komplexeren Klassifikationskonzepte zu erhalten, werden zunächst SVM-Klassifikatoren ausschließlich anhand der Einzellabel als Benchmarks trainiert und evaluiert. Diese entspricht einer Betrachtung, in dem weder eine Kombinationsregel, noch ein komplexeres Klassifikationskonzept zur Verarbeitung der Daten Anwendung finden. Dazu werden im Folgenden zwei Experimente durchgeführt. Im ersten Experiment wird jeder Featurevektor sowohl in der Trainings-, als auch in der Testmenge N_E-mal mit jeweils einem zugeordneten Einzellabel genutzt, um eine SVM zu trainieren (Benchmark 1). Dieser Fall entspricht nahezu dem in Abschnitt 6.2.2 vorgestellten Klassifikationskonzept. Der Unterschied liegt darin, dass nicht ein anhand der Einzellabel bestimmtes, kombiniertes Label, sondern die Einzellabel selbst als Vergleichslabel zur Bestimmung der Testgüte des SVM-Klassifikators genutzt werden. Im zweiten Experiment wird je eine SVM für jedes Einzellabel trainiert (Benchmark 2). Dieses Vorgehen entspricht dem in Abschnitt 6.2.3 vorgestellten Klassifikatorkonzept. Im Gegensatz zum Ensemble wird allerdings keine Kombinationsvorschrift angewendet, um die Ausgaben der N_E SVM-Klassifikatoren zu einer Gesamtausgabe zu vereinen. Zur Bestimmung der Güte der einzelnen SVM-Klassifikatoren werden jeweils die Einzellabel verwendet.

	Bewertungsmaß			
	#	$\kappa_{w,train.}$	$\kappa_{w,test}$	$e_{train.}$ [%]	e_{test} [%]
Ergebnisse	1	0,343	0,276	50,7	56,6
	2	0,343	0,264	50,9	57,3
	3	0,339	0,253	51,0	58,7
	4	0,336	0,284	51,3	55,8
	5	0,337	0,284	51,0	55,9
	6	0,332	0,284	51,6	56,2
	7	0,340	0,264	51,3	57,0
	8	0,343	0,275	51,4	56,4
	9	0,334	0,307	51,6	54,9
	10	0,336	0,265	51,4	56,8
	$\boldsymbol{\mu}$	**0,338**	**0,276**	**51,2**	**56,5**
	$\boldsymbol{\sigma}$	**0,004**	**0,015**	**0,3**	**1,0**

Tabelle 21: Klassifikationsgüte κ_w und -fehler e der nur anhand der Einzellabel in Kombination mit allen netzspezifischen Features trainierten und getesteten SVM (Benchmark 1).

Die im erstgenannten Experiment unter Verwendung aller netzspezifischen Features resultierenden Trainings- und Testergebnisse sind in Tabelle 21 gezeigt. Neben den

Klassifikationsgüten $\kappa_{w,train.}$ und $\kappa_{w,test}$ auf den Trainings- und Testdaten sind zusätzlich auch die entsprechenden Klassifikationsfehler $e_{train.}$ und e_{test} für jede der zehn Kreuzvalidierungen gezeigt ($e \neq 1 - \kappa_w$). Der Klassifikationsfehler entspricht dem Anteil der jeweils falsch klassifizierten Netze. Die Tabelle dient zur Veranschaulichung der im Rahmen der Experimente vorgenommenen Bewertung der Klassifikationskonzepte. Im weiteren Verlauf dieses Abschnitts werden die Ergebnisse der trainierten SVM-Klassifikatoren jeweils in graphischer Form anhand der in den unteren beiden Zeilen der ersten zwei Spalten berechneten Mittelwerte $\mu(\kappa_w)$ und Standardabweichungen $\sigma(\kappa_w)$ gegenübergestellt. Eine zusätzliche graphische Gegenüberstellung der gemittelten Klassifikationsfehler (mittlerer Anteil falsch klassifizierter Netze) $\mu(e_{train.})$ und $\mu(e_{test})$ sowie der entsprechenden Standardabweichungen $\sigma(e_{train.})$ und $\sigma(e_{test})$ sind jeweils dem Anhang beigefügt.

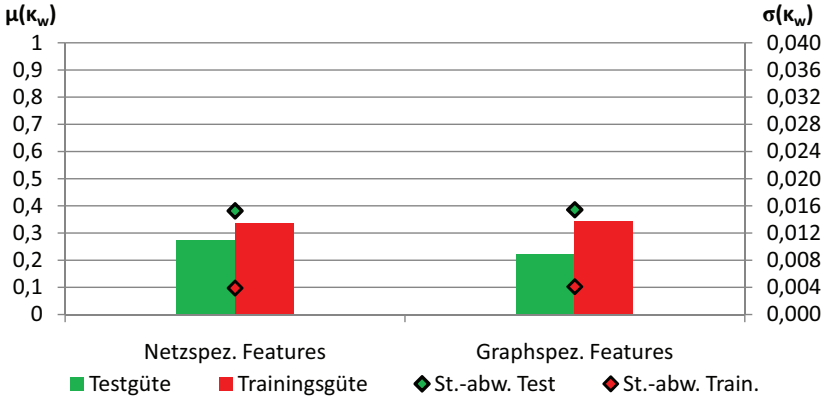


Abbildung 28: Klassifikationsgüte einer nur anhand der Einzellabel trainierten und getesteten SVM (Benchmark 1).

In Abbildung 28 sind die Klassifikationsgüten des SVM-Klassifikators nochmal bei Verwendung aller netzspezifischen und aller graphspezifischen Features gegenübergestellt. Die resultierenden Trainingsgüten der beiden Featuresätze stimmen mit Werten von $\mu(\kappa_{w,train.}) = 0,34$ nahezu überein. Auch die sich ergebenden Standardabweichungen auf den Trainings- und den Testdaten, $\sigma(\kappa_{w,train.}) = 0,004$ und $\sigma(\kappa_{w,test}) = 0,015$, nehmen jeweils ähnliche Werte an. Im Gegensatz dazu wird mit den netzspezifischen Features eine Testgüte von $\mu(\kappa_{w,test}) = 0,28$ erreicht, die den unter Verwendung der graphspezifischen Features erreichten Wert von $\mu(\kappa_{w,test}) = 0,22$ deutlich übersteigt. Damit weist die Verwendung der netzspezifischen Features zur Modellierung der ersten Benchmark-SVM Vorteile gegenüber der Verwendung der graphspezifischen Features auf. Allerdings liegen die erreichten Klassifikationsgüten auf einem insgesamt sehr niedrigen Niveau. Diese niedrigen Werte lassen sich durch die jeweils N_E-malige Verwendung der Featurevektoren mit den Einzellabeln in den Trainings- und Testdaten

begründen. Wie in Abschnitt 6.2.2 beschrieben, wird sich durch dieses Vorgehen die Lage der Hyperebene im Bildraum stark an den Labeln orientieren, die jeweils für einen Datenpunkt am häufigsten in den Trainingsdaten enthalten sind. Für die Featurevektoren, denen ein davon abweichendes Einzellabel zugeordnet ist, führt die Ausgabe der SVM daher zu Fehlklassifikationen. Zusammenfassend führt die Verwendung aller Einzellabel in den Trainingsdaten zu einer (gemäß der Ausführungen in Abschnitt 6.2.2) sinnvollen Generierung einer Gesamtausgabe der SVM. Wie oben beschrieben mangelt es bei dieser Benchmark-SVM allerdings an einem entsprechenden sinnvollen Vergleichslabel, wie es z. B. in den komplexeren Klassifikationskonzepten aus Abschnitt 6.2 durch Verwendung der kombinierten Label vorgesehen ist.

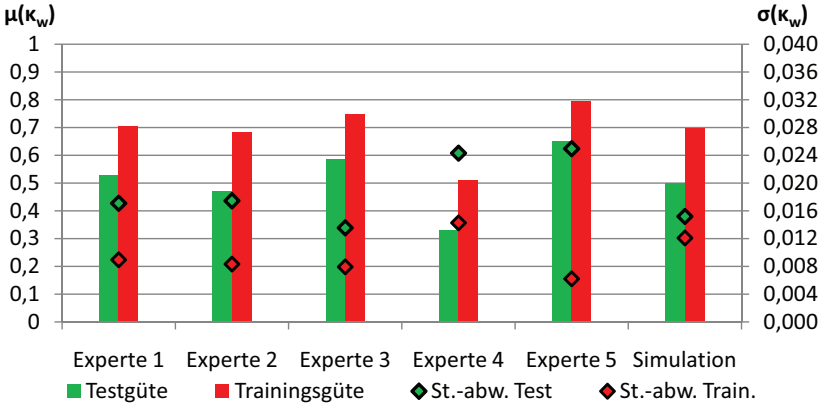


Abbildung 29: Klassifikationsgüte für jeweils anhand von Einzellabeln in Kombination mit netzspezifischen Features trainierte SVM (Benchmark 2).

Die sich im zweiten Experiment (Benchmark 2) einstellenden Klassifikationsgüten der einzelnen SVM sind in Abbildung 29 bei Verwendung aller netzspezifischen Features dargestellt. Die Klassifikationsergebnisse sind im Vergleich zur SVM, die mit allen Einzellabeln trainiert wurde, deutlich verbessert (vgl. Abbildung 28). Diese Verbesserung liegt hauptsächlich darin begründet, dass in den verwendeten Trainings- und Testdaten nicht für einen Datenpunkt mehrere (möglicherweise verschiedene) Label vorliegen. Allerdings treten bei den einzelnen SVM teilweise höhere Standardabweichungen auf (z. B. bei Verwendung der Label von Experte 4 oder Experte 5). Dies ist auf die geringere Anzahl an Datenpunkten zum Training und Test der SVM im Vergleich zum Benchmark 1 zurückzuführen (300 im Vergleich zu $5 \cdot 300 = 1500$). Die höchste Klassifikationsgüte tritt für die anhand der Label von Experte 5 trainierte SVM mit einem Wert von $\mu(\kappa_{w,test}) = 0,65$ auf. Der schlechteste Klassifikator wird beim Training einer SVM mit den Labeln von Experte 4 realisiert ($\mu(\kappa_{w,test}) = 0,28$). Die unter Verwendung der simulationsbasierten Label trainierte SVM weist im Vergleich zu den expertenbasierten SVM mit einem Wert von $\mu(\kappa_{w,test}) = 0,49$ eine mittlere Klassifikationsgüte auf.

Einfluss der Klassifikationskonzepte auf die Klassifikationsgüte:
Nachdem mit Hilfe der Benchmark-SVM eine Vergleichsbasis für die Klassifikationsgüte komplexerer Klassifikationskonzepte gelegt wurde, werden in einem nächsten Schritt die mit den drei in Abschnitt 6.2 vorgestellten Konzepten erzielten Klassifikationsgüten zur Prognose der mittels EIDMR kombinierten Label gegenübergestellt und mit den Resultaten der Benchmark-SVM verglichen. Wie in Abschnitt 6.3.1 beschrieben, wurden die Label, die für die NS-Netze anhand der stochastischen Simulationen bestimmt wurden, in den Kombinationsregeln wie ein einzelner Experte berücksichtigt. Dieser Vergleich wird hier unter Verwendung aller netzspezifischen Features zum Training der SVM durchgeführt. Ergänzende Ergebnisse für diesen Vergleich bei Verwendung aller graphspezifischen Features befinden sich im Anhang. Zusätzlich wird der standardmäßige Gewichtungsansatz zur Bestimmung der Ausgabe eines Ensembles aus Konzept 3*a*, in dem der Einfluss der Einzelklassifikatoren auf die Ensemble-Ausgabe anhand der jeweils kreuzvalidierten Güte auf den Trainingsdaten der SVM gebildet wird, mit in den Vergleich einbezogen. Dies ermöglicht zusätzlich die Bewertung eines möglichen Mehrwerts durch das in Abschnitt 6.2.3 vorgestellte Konzept 3*b*, bei dem die Gewichtung anhand der mit Cohen's κ_w gemessenen Ähnlichkeit zwischen den Einzellabeln der jeweiligen Trainingsdatenmenge T_j und den zu prognostizierenden kombinierten Labeln gebildet wird. Eine übersichtliche Zusammenfassung der Klassifikationskonzepte ist nochmal in Tabelle 22 dargestellt.

KLASSIFIKATORKONZEPT	BESCHREIBUNG	REFERENZ
Konzept 1	Training mit kombinierten Labeln	Abschnitt 6.2.1
Konzept 2	Training mit Einzellabeln	Abschnitt 6.2.2
Konzept 3*a*	Ensemble mit stand. Gewichtung	Abschnitt 6.2.3
Konzept 3*b*	Ensemble mit neuer Gewichtung	Abschnitt 6.2.3

Tabelle 22: Klassifikationskonzepte zur Prognose der kombinierten Label.

Abbildung 30 zeigt die Ergebnisse des Vergleichs der Klassifikationskonzepte. Die Trainingsgüte im Konzept 2 entspricht mit einem Wert von $\mu(\kappa_{w,train.}) = 0{,}34$ erwartungsgemäß exakt der Trainingsgüte der SVM in Benchmark 1. Der Unterschied zum Benchmark 1 liegt in der Verwendung der kombinierten Label als Vergleichslabel zur Bestimmung der Testgüte. Aus diesem Grund steigt die Testgüte im Vergleich zum Benchmark 1 mit einem Wert von $\mu(\kappa_{w,test}) = 0{,}74$ auf mehr als das Zweifache an. Die Ausgabe der SVM in Konzept 2 liefert damit bereits eine gute Übereinstimmung zu den kombinierten Labeln. Auch gegenüber den Ergebnissen der einzelnen SVM in Benchmark 2 zeigt sich eine deutliche Verbesserung der Klassifikationsgüte. Die Klassifikationsgüte des Ensembles aus Konzept 3b liegt mit einem Wert von $\mu(\kappa_{w,test}) = \mu(\kappa_{w,train.}) = 0{,}74$ auf dem gleichen guten Niveau. Allerdings weisen die Standardabweichungen sowohl beim Training, als auch beim Test den Zweifachen bzw. den 1,5-fachen Wert im Vergleich zum Training einer SVM mit Einzellabeln auf. Gegenüber dem Ensemble mit standardmäßiger Gewichtung aus Konzept 3a zeigt sich für das Konzept 3b ei-

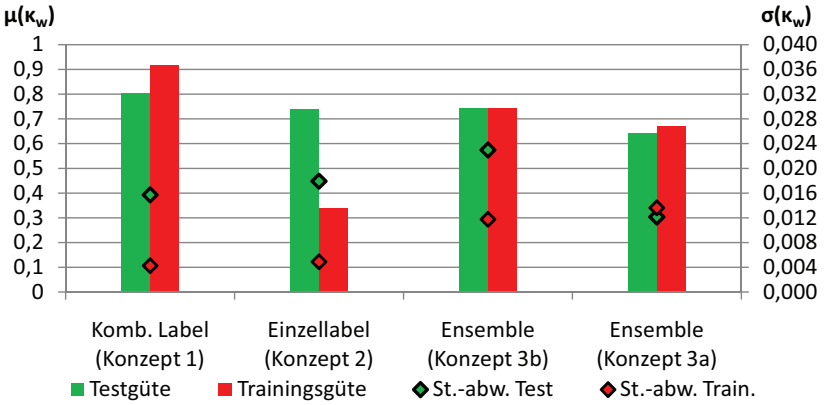


Abbildung 30: Gegenüberstellung der Klassifikationsgüten verschiedener Klassifikationskonzepte zur Prognose der kombinierten Label (EIDMR) mit allen netzspezifischen Features.

ne deutliche Verbesserung der Klassifikationsgüte. Die neue Gewichtung anhand der mit Cohen's κ_w gemessenen Ähnlichkeit zwischen den Einzellabeln und den kombinierten Labeln bringt folglich einen deutlichen Mehrwert gegenüber der Verwendung der Standard-Gewichte. Die höchste Klassifikationsgüte wird allerdings mit Konzept 1 erzielt, in dem die Modellierung des Klassifikators direkt auf Basis der kombinierten Label erfolgt. Die Ausgaben der SVM weisen eine insgesamt sehr gute Übereinstimmung zu den kombinierten Labeln auf. Mit Werten von $\mu(\kappa_{w,train.}) = 0{,}91$ und $\mu(\kappa_{w,test}) = 0{,}80$ übertreffen sowohl die Trainingsgüte, als auch die Testgüte von Konzept 1 die Werte der anderen Konzepte deutlich. Die auftetenden Standardabweichungen für die Trainings- und Testgüte sind im Vergleich ebenfalls gering. Einzig die bei Anwendung von Konzept 3a auftretende Standardabweichung der Testgüte weist einen geringfügig niedrigeren Wert auf. Es lässt sich an dieser Stelle festhalten, dass alle Klassifikationskonzepte eine höhere Klassifikationsgüte als die Benchmark-SVM erzielen. Die Modellierung des Klassifikators mit Konzept 1 führt im Vergleich der Klassifikationskonzepte zur höchsten Klassifikationsgüte.

Einfluss der Features auf die Klassifikationsgüte:
Auf Basis von Konzept 1 soll nun eine fundierte Aussage getroffen werden, welche Features zur besten Prognose der mit der EIDMR kombinierten Label führen. Im Benchmark 1 haben die Ergebnisse aus Tabelle 21 und Abbildung 28 gezeigt, dass die netzspezifischen Features zu einer höheren Klassifikationsgüte führen als die graphspezifischen Features. Um den Einfluss der Features näher zu untersuchen, ist in Abbildung 31 die Klassifikationsgüte bei Modellierung des Klassifikators mit allen netzspezifischen und allen graphspezifischen Features gegenübergestellt. Zusätzlich wird die

Klassifikationsgüte zur Prognose der mit der EIDMR kombinierten Expertenaussagen (ohne Simulationsergebnisse) mit allen simulationsspezifischen Features in den Vergleich einbezogen, die sich unter sonst gleichen Bedingungen einstellt. Das Ergebnis des Vergleichs bestätigt, dass die Verwendung der netzspezifischen Features zu einer höheren Klassifikationsgüte führt. Die bereits diskutierte hohe Klassifikationsgüte bei Nutzung der netzspezifischen Features von $\mu(\kappa_{w,test}) = 0{,}80$ übersteigt den sich unter Verwendung der graphspezifischen Features einstellenden Wert von $\mu(\kappa_{w,test}) = 0{,}75$ deutlich. Auch die Standardabweichung nimmt für die netzspezifischen Features deutlich niedriger Werte auf den Testdaten an, als unter Verwendung der graphspezifischen Features. Die in Abschnitt 6.3.2 ermittelte niedrige Relevanz der simulationsspezifischen Features spiegelt sich ebenfalls in einer sehr niedrigen Klassifikationsgüte von $\mu(\kappa_{w,test}) = 0{,}12$ wider.

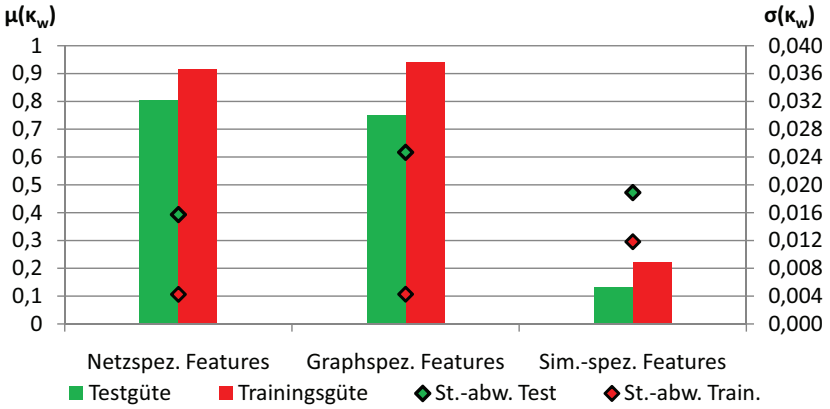


Abbildung 31: Gegenüberstellung der Klassifikationsgüte für verschiedene Featuresätze zur Prognose der kombinierten Label (EIDMR).

Nach erfolgter Gegenüberstellung der Klassifikationsgüte der drei gesamten Featuresätze, stellt sich nun die Frage, wie sich eine Selektion einiger für die Klassifikation höherrelevanter Features eines Datensatzes sowie eine Kombination höherrelevanter Features verschiedener Datensätze auf die Klassifikationsgüte auswirkt. Den Ausgangspunkt zur Untersuchung dieser Fragestellung bildet die Analyse der Featurerelevanz in Abschnitt 6.3.2. Zunächst wurde die Menge der netzspezifischen Features anhand der Ergebnisse in Tabelle 17 auf die ersten sechs Features dieser Tabelle reduziert, da diese sowohl von beiden Filterverfahren, als auch von den Experten vergleichsweise hoch bewertet worden sind. Zusätzlich wurde dieser Menge das Feature „durchschnittlicher Stationsradius“ hinzugefügt, das von den Experten am höchsten bewertet worden ist. Bei den graphspezifischen Features wurden von allen 22 Features die zehn in Tabelle 19 dargestellten Features ausgewählt, da diese von beiden Filterverfahren überdurchschnittlich hoch bewertet worden sind. Die Ergebnisse zur Selektion der Features sind

in Abbildung 32 dargestellt. Durch die Reduktion der Featuremenge verschlechtert sich die Klassifikationsgüte der mit netzspezifischen Features trainierten SVM deutlich auf einen Wert von $\mu(\kappa_{w,test}) = 0{,}77$. Bei den graphspezifischen Features hingegen bewirkt die Reduktion der Featuremenge eine leichte Verbesserung der Klassifikationsgüte auf einen Wert von $\mu(\kappa_{w,test}) = 0{,}76$.

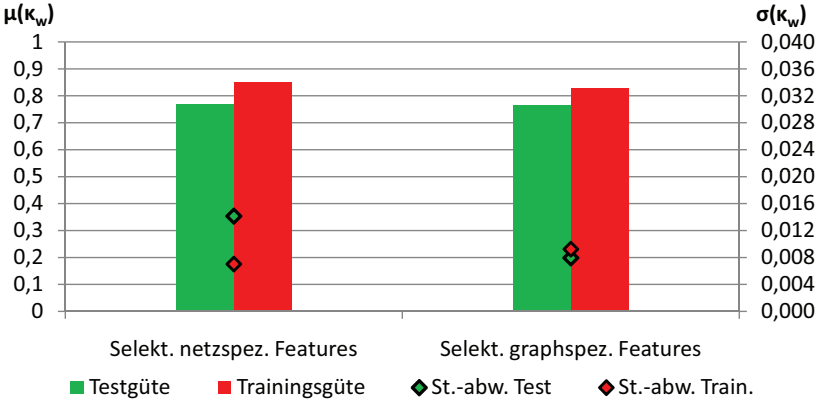


Abbildung 32: Gegenüberstellung der Klassifikationsgüte bei Selektion von Features zur Prognose der kombinierten Label (EIDMR, bei simulationsspezifischen Features nur expertenbasierte Label).

Die Ergebnisse zur Kombination von Features verschiedener Datensätze zeigt Abbildung 33. Neben der sich unter Verwendung aller Features einstellenden Klassifikationsgüte, sind diese für die 15 Features aus Tabelle 20, denen eine überdurchschnittlich hohe Relevanz durch beide Filterverfahren zugeordnet wurde, sowie für eine um weitere fünf Features reduzierte Featuremenge (nur noch die ersten zehn Features aus Tabelle 20 wurden verwendet) gezeigt. Die niedrigste Klassifikationsgüte auf den Testdaten stellt sich mit einem Wert von $\mu(\kappa_{w,test}) = 0{,}72$ bei Verwendung aller Features ein. Für die relevantesten 15 Features nimmt die Klassifikationsgüte $\mu(\kappa_{w,test})$ einen Wert von 0,77 an. Bei weiterer Reduktion der Featuremenge auf die zehn relevantesten Features sinkt die Klassifikationsgüte leicht auf $\mu(\kappa_{w,test}) = 0{,}76$. Aus der mit sinkender Anzahl berücksichtigter Features abnehmenden Trainingsgüte ist ersichtlich, dass die Featurezahl bei der relativ geringen Anzahl von 300 Datenpunkten nicht zu groß gewählt werden darf, um eine Überanpassung des Klassifikators auf die Trainingsdaten zu vermeiden (für alle Features sogar: $\mu(\kappa_{w,train}) \approx 1{,}00$). Die höchste Testgüte bei niedrigster Standardabweichung tritt mit einem Wert von $\mu(\kappa_{w,test}) = 0{,}77$ bei Verwendung der 15 relevantesten Features auf. Damit liegt die erreichte Klassifikationsgüte auf dem gleichen Wert, wie bei Selektion der relevantesten graphspezifischen Features erreicht wurde. Allerdings liegt dieser Wert unter der mit allen netzspezifischen Features erreichten Testgüte von $\mu(\kappa_{w,test}) = 0{,}80$.

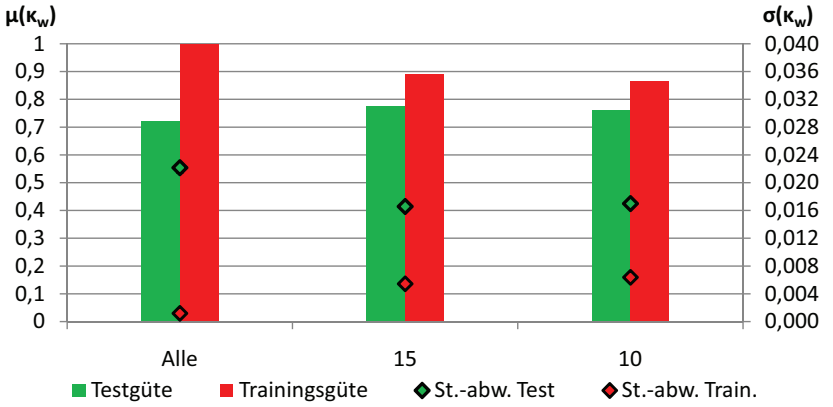


Abbildung 33: Ergebnisse zur Kombination von Features zur Prognose der kombinierten Label (EIDMR).

Aus den Ergebnissen lässt sich festhalten, dass die in Abschnitt 6.3.2 verwendeten Filterverfahren – wie dort bereits angeführt – zur Bewertung der Featurerelevanz in den hier durchgeführten empirischen Untersuchungen zwar erste Indikationen zur Verwendung der Features geben, diese allerdings noch während der Modellierung des Klassifikators zu verifizieren sind, da bei der Featurebewertung keine Abhängigkeit zum später angewendeten Klassifikator berücksichtigt wird. So hat sich z. B. die bewertete niedrige Relevanz der simulationsspezifischen Features zur Prognose der kombinierten Expertenaussagen auch in einer niedrigen Klassifikationsgüte widergespiegelt. Hinsichtlich der auf den Testdaten erreichten Klassifikationsgüte werden die relevantesten Features von den zehn netzspezifischen Features gebildet.

Einfluss der Kombinationsregeln auf die Klassifikationsgüte:
In den bisher diskutierten Experimenten hat sich gezeigt, dass die höchste Klassifikationsgüte zur Prognose der unter Anwendung der EIDMR kombinierten Label mit allen netzspezifischen Features und Konzept 1 erzielt wird. Für diese erfolgversprechendste Kombination aus Featureauswahl und Klassifikationskonzept wird nun der Einfluss der verwendeten Kombinationsregel auf die Klassifikationsgüte untersucht. Es stellt sich die Frage, ob durch Verwendung einer der drei weiteren Kombinationsregeln aus Abschnitt 5 (der DSR, MR oder IDMR) die zur Prognose der kombinierten Label erreichte Testgüte von $\mu(\kappa_{w,test}) = 0{,}80$ noch übertroffen werden kann. Das Ergebnis dieses Vergleichs ist in Abbildung 34 dargestellt.

Es ist ersichtlich, dass unter Verwendung der weiteren drei Kombinationsregeln auf den Testdaten mit Werten von $\mu(\kappa_{w,test})$ zwischen 0,55 und 0,57 nur eine mäßige Klassifikationsgüte erzielbar ist. Die Verwendung der EIDMR ist folglich hinsichtlich der

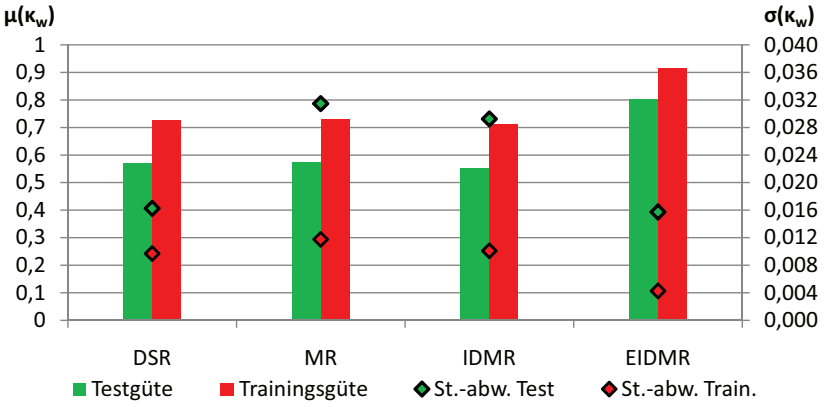


Abbildung 34: Einfluss der verwendeten Kombinationsregel auf die Klassifikationsgüte bei Verwendung aller netzspezifischen Features und Konzept 1.

Modellierung eines Klassifikators deutlich vorteilhaft. Neben den deutlich höheren Klassifikationsgüten treten bei Verwendung der EIDMR auch im Vergleich der Standardabweichungen die niedrigsten Werte auf. Dies bedeutet, dass die EIDMR nicht nur, wie in Abschnitt 5.3 auf künstlichen Daten analysiert wurde, eine hohe Güte zur Detektion der wahren Klassen aufweist, sondern auch zu einer hohen Klassifikationsgüte bei Anwendung einer SVM führt. Diese Tendenz zeigte sich bereits in Abschnitt 5.4 in einer im Vergleich zu den übrigen Kombinationsregeln (anhand des Davies-Bouldin-Index bewerteten) besseren Separierbarkeit der kombinierten Label im Featureraum[1]. Über die Ergebnisse aus Kapitel 5 hinaus, in dem die Güte auf allen Beispieldaten untersucht wurde, zeigt sich bei der hier simulierten Generalisierung auf „unbekannte" Datenpunkte (NS-Netze), dass bei Anwendung der EIDMR in Kombination mit einer SVM im Vergleich zu den anderen drei Kombinationsregeln die besten Klassifikationsergebnisse erreicht werden.

Einfluss der Simulationsergebnisse auf die Klassifikationsgüte:
Bis zu diesem Punkt wurden die Label, die für die NS-Netze anhand der stochastischen Simulationen bestimmt wurden, in den Kombinationsregeln wie ein einzelner Experte gewichtet. Um den Ergebnissen der stochastischen Simulationen mehr Gewicht bei der Bestimmung der kombinierten Label zu geben, können diese – wie in Abschnitt 5.4 beschrieben – mehrfach, d. h. wie jeweils zwei oder mehr Experten, bei der Kombination gewichtet werden. Es wird nun in einem letzten Schritt der Einfluss der Gewichtung der simulationsbasierten Label auf die Klassifikationsgüte untersucht. Bei einer Anwendung

1 Neben der Bewertung wurde in Abschnitt 5.4 auch eine Visualisierung in dem Raum vorgenommen, der durch die ersten zwei Hauptkomponenten aller netzspezifischen Features aufgespannt wird.

in der Praxis sollte zwischen gewünschter Gewichtung und erzielter Klassifikationsgüte abgewägt werden.

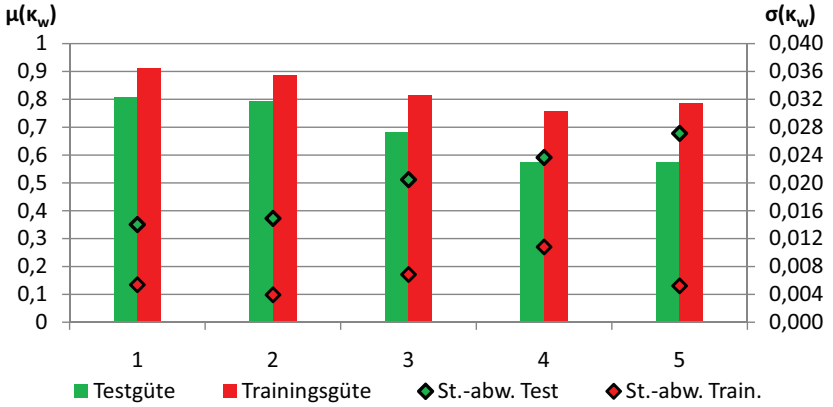


Abbildung 35: Einfluss der Gewichtung der Simulationsergebnisse auf die Klassifikationsgüte für netzspezifischen Features und Verwendung der EIDMR.

Abbildung 35 zeigt die Klassifikationsgüte bei Anwendung des Konzeptes 1 mit kombinierten Labeln (EIDMR) für die ein- bis fünffache Gewichtung der simulationsbasierten Label bei der Kombination der Einzellabel (analog zu Abschnitt 5.4: Kombination mit EIDMR, $k = 5$ berücksichtigte Nachbarn). Zur Modellierung des Klassifikators wurden alle netzspezifischen Features verwendet. Die Klassifikationsgüte nimmt im Vergleich zur einfachen Gewichtung bei zweifacher Gewichtung nur leicht ab. Ab der dreifachen Gewichtung der simulationsbasierten Label ist eine deutliche Abnahme der Klassifikationsgüte $\mu(\kappa_{w,test})$ auf Werte von bis zu 0,57 zu beobachten. Zusammenfassend kann also eine zweifache Gewichtung der simulationsbasierten Label bei der Kombination ohne große Abnahme der Prognosequalität vorgenommen werden. Eine darüber hinausgehende mehrfache Gewichtung ist hinsichtlich der deutlichen Abnahme der Klassifikationsgüte allerdings nicht empfehlenswert. Die Ergebnisse decken sich mit der in Abschnitt 5.4 durchgeführten Analyse der Kombinationsergebnisse. Insgesamt führen damit sowohl die Analyse der Kombinationsergebnisse, als auch die der Klassifikationsgüte zu der Erkenntnis, dass die simulationsbasierten Label nicht höher als zweifach bei der Kombination gewichtet werden sollten.

Anwendung des Ansatzes von Frank und Hall zur Lösung von ordinalen Mehrklassenklassifikationsproblemen:
In einem letzten Experiment wird nun überprüft, ob sich durch Anwendung des allgemeinen Ansatzes zur Lösung von ordinalen Mehrklassenklassifikationsproblemen nach Frank und Hall eine weitere Erhöhung der Klassifikationsgüte erzielen lässt [103] (Abschnitt 6.1.2). Prognostiziert werden sollen die mit der EIDMR kombinierten Label bei

einfacher Gewichtung der simulationsbasierten Label. In den vorhergehenden Experimenten wurde die Ordinalität der Label durch Optimierung von Cohen's κ_w zwischen SVM-Ausgabe und den kombinierten Labeln bei der Parametersuche für die SVM berücksichtigt (für alle SVM wurden die Parameter $C = 10{,}0$ und $\gamma = 0{,}03$ verwendet). Aus diesem Grund wird zunächst für das Konzept 1 in Kombination mit allen netzspezifischen Features nochmal eine Optimierung der SVM-Parameter C und γ hinsichtlich Cohen's κ_w vorgenommen. Mit Werten von $C = 9{,}8$ und $\gamma = 0{,}04$ weichen die neuen Parameter nur gering von den zuvor in allen Experimenten verwendeten Parametern $C = 10{,}0$ und $\gamma = 0{,}03$ ab. Für die einzelnen SVM zur Lösung der Zweiklassenklassifikationsprobleme im Ansatz von Frank und Hall wurde ebenfalls jeweils eine Optimierung der Parameter durchgeführt. In aufsteigender Reihenfolge wurden für die vier SVM folgende Parameterkombinationen genutzt: $C_1 = 0{,}8$ und $\gamma_1 = 0{,}20$, $C_2 = 0{,}7$ und $\gamma_2 = 0,14$, $C_3 = 3{,}2$ und $\gamma_3 = 0{,}27$ sowie $C_4 = 10{,}0$ und $\gamma_4 = 0{,}10$.

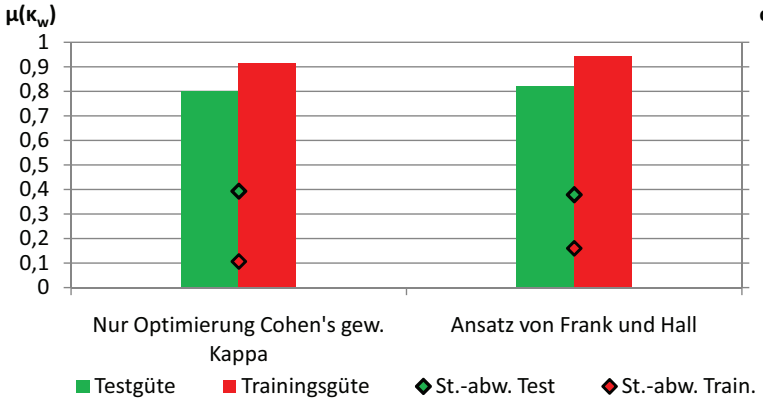


Abbildung 36: Verbesserung der Klassifikationsgüte zur Prognose der kombinierten Label (EIDMR) mit dem Ansatz von Frank und Hall.

Die sich unter diesen Voraussetzungen einstellenden Klassifikationsgüten beider Ansätze sind in Abbildung 36 gegenübergestellt. Es zeigt sich, dass durch die Parameteroptimierung für das Konzept 1 keine weitere nennenswerte Verbesserung der Klassifikationsgüte mehr erreicht wird. Wie in den Experimenten zuvor ergibt sich ein Wert von $\mu(\kappa_{w,test}) = 0,80$. Im Vergleich dazu wird unter Anwendung des Ansatzes von Frank und Hall nochmals eine Verbesserung der Klassifikationsgüte auf einen Wert von $\mu(\kappa_{w,test}) = 0,82$ erzielt. Die Standardabweichungen beider Konzepte sind mit Werten von ca. $\sigma(\kappa_{w,test}) = 0,0016$ ähnlich. Die Ausnutzung der ordinalen Information in den Labeln schafft damit eine leichten Mehrwert bei der Modellierung des SVM-Klassifikators und führt in Kombination mit allen netzspezifischen Features zu den besten Klassifikationsergebnissen in den hier durchgeführten Experimenten zur Prognose der kombinierten Label (EIDMR). Eine nähere Untersuchung der Ergebnisse

des Klassifikators nach Frank und Hall wird nun anhand der drei Konfusionsmatrizen (Abschnitt 4.1.4) für die Testdaten der zehnten Kreuzvalidierung durchgeführt:

$$\begin{pmatrix} 9 & 0 & 0 & 0 & 0 \\ 2 & 15 & 2 & 0 & 0 \\ 0 & 8 & 48 & 1 & 0 \\ 0 & 0 & 2 & 13 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix} \quad \begin{pmatrix} 10 & 0 & 0 & 0 & 0 \\ 1 & 21 & 3 & 0 & 0 \\ 0 & 2 & 48 & 4 & 0 \\ 0 & 0 & 1 & 10 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix} \quad \begin{pmatrix} 10 & 1 & 0 & 0 & 0 \\ 0 & 15 & 0 & 1 & 0 \\ 0 & 7 & 48 & 3 & 0 \\ 0 & 0 & 4 & 11 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix} \tag{55}$$

Entsprechend der durchgeführten dreifach Kreuzvalidierung ergibt die Summe der Einträge in jeder Konfusionsmatrix einen Wert von 100 (NS-Netzen). Wie bereits in Abschnitt 5.4 festgestellt wurde, treten bei Anwendung der EIDMR auf alle sechs Einzellabel bei entsprechend gleicher Gewichtung keine NS-Netze der Klasse 5 auf[2]. Für die drei Konfusionsmatrizen ergeben sich folgende Werte für die Interraterkonkordanz κ_w: 0,83, 0,87 und 0,80. Die Anteile der falsch klassifizierten Netze ergeben sich entsprechend zu: 15%, 11% und 16%. Weiterhin lässt sich aus den Konfusionsmatrizen ablesen, dass bei einer fehlerhaften Klassifikation die Ausgabe des Klassifikators hauptsächlich nur um eine Klasse vom korrekten Ergebnis abweicht, was auf die Ausnutzung der ordinalen Information der Klassen nach Frank und Hall zurückzuführen ist. Abweichungen nach unten oder nach oben liegen in ungefähr gleichem Maße vor. Nur in der dritten Konfusionsmatrix tritt ein einzelner Eintrag mit einem geringen Wert von eins (entspricht durchschnittlich 0,33% der NS-Netze über alle drei Konfusionsmatrizen) außerhalb der Haupt- und Nebendiagonalen auf. Absolut gesehen treten die meisten Fehlklassifikationen für die mittleren Klassen auf, was zum einen durch die Verteilung der Klassen (mittlere Klassen treten häufiger auf als Randklassen) und zum anderen durch die Kombination der Ausgaben von zwei SVM für die mittleren Klassen im Ansatz von Frank und Hall bedingt ist (Abschnitt 6.1.2).

6.4 EINORDNUNG DER ERGEBNISSE

Nach einer einführenden Erläuterung der mathematischen Grundlagen von SVM wurden drei Konzepte zur Modellierung von SVM-Klassifikatoren vorgestellt und empirisch untersucht. Ausgangsbasis bildeten dabei Benchmark-SVM, die ohne Anwendung einer Kombinationsregel und eines komplexeren Klassifikationskonzeptes trainiert wurden. In den Experimenten zeigte sich durch Vergleich mit den Ergebnissen dieser Benchmark-SVM deutlich, dass bei unsicheren Labeln eine Erhebung mehrerer Einzellabel sowie die Anwendung der Kombinationsregel EIDMR zu wesentlich zuverlässigeren SVM-Klassifikatoren als in den Benchmarks führen. Die höchste Klassifikationsgüte zur Prognose der kombinierten Label wurden mit dem Konzept 1 (Training der SVM mit kombinierten Labeln) erreicht. Die Anwendung der EIDMR zur Kombination der Einzellabel

2 Der vierte Einzelklassifikator im Konzept von Frank und Hall wird daher nur bei Anwendung einer der drei anderen Kombinationsregeln benötigt (Ergebnisse im Anhang).

führte dabei zu deutlich besseren SVM-Klassifikatoren verglichen mit den anderen drei Kombinationsregeln DSR, MR und IDMR. Hinsichtlich der Verwendung von Features resultierte bei Nutzung aller netzspezifischen Features die höchste Klassifikationsgüte. Es zeigte sich allerdings auch, dass die Nutzung von bestimmten graphspezifischen Features nur zu einer geringfügig schlechteren Klassifikationsgüte führt. Durch eine Kombination von netzspezifischen und graphspezifischen Features konnte in den Experimenten keine Erhöhung der Klassifikationsgüte herbeigeführt werden. Zudem zeigte sich in den Experimenten, dass durch Anwendung des Ansatzes von Frank und Hall zur Lösung von ordinalen Mehrklassenklassifikationsproblemen die Klassifikationsgüte des Konzeptes 1 weiter verbessert werden konnte. Die Nutzung der ordinalen Informationen bei der Modellierung des Klassifikators hat in dieser Arbeit somit einen Mehrwert bei der Klassifikation mit SVM geschaffen. Die erreichte Klassifikationsgüte liegt bei zehn durchgeführten dreifach Kreuzvalidierungen bei $\mu(\kappa_{w,test}) = 0,82$ und ist aus Anwendungssicht als sehr zufriedenstellend zu bewerten. Die Anzahl der Fehlklassifikationen beträgt dabei durchschnittlich 15%, wobei die Analysen der einzelnen Konfusionsmatrizen der zehnten Kreuzvalidierung zeigt, dass durch die Ausnutzung der ordinalen Informationen in den Labeln zur Lösung des Mehrklassenklassifikationsproblems, diese Fehlklassifikationen zu einem sehr großen Teil nur um ein Klasse von dem gewünschten Ergebnis abweichen.

7

REFLEXION, FAZIT UND AUSBLICK

7.1 REFLEXION UND FAZIT

In dieser Arbeit wurden verschiedene Fragestellungen zur effizienten Klassifikation von NS-Netzen hinsichtlich ihrer Aufnahmekapazität für DEA adressiert. In Abschnitt 1.3 wurden vor diesem Hintergrund konkrete Zielsetzungen formuliert, die im Folgenden aufgegriffen und hinsichtlich ihrer Erfüllung diskutiert werden.

Das erste Ziel dieser Arbeit bestand in der Identifikation und Erhebung relevanter Merkmale zur Charakterisierung von NS-Netzen. Um diesem Grundproblem der Bildung einer geeigneten Datengrundlage zu begegnen, wurden zunächst diskriminierende Merkmale identifiziert und in zwei Datensätzen gebündelt. Der erste Datensatz setzt sich aus zehn netzspezifischen Merkmalen zusammen, die einen starken Bezug zur Verteilnetzplanung aufweisen und deren Erhebung ein gewisses Maß an Anwendungswissen voraussetzt. Ergänzend dazu beinhaltet der zweite Datensatz 22 abstraktere, graphspezifische Merkmale. Verglichen mit den netzspezifischen Merkmalen ist die Erhebung dieser graphspezifischen Merkmale leicht automatisierbar und setzt kein Anwendungswissen voraus. Die Relevanz dieser Merkmale hinsichtlich der Klassifikation von NS-Netzen wurde darauf aufbauend anhand von zwei Filterverfahren, Expertenbewertungen und bei der Modellierung von SVM-Klassifikatoren bewertet. Aus den Analysen geht hervor, dass alle netzspezifischen sowie ein Großteil der graphspezifischen Merkmale (z. B. „Betrag des zweitgrößten Eigenwerts“, „spektraler Radius“) sehr gut zur Klassifikation geeignet sind. Die Kombination von Merkmalen beider Datengrundlagen führte in den Experimenten zu keinem Mehrwert bei der Klassifikation. Aus der Bewertung der netzspezifischen Merkmale untereinander konnte aus Anwendungssicht folgende Implikation für die Entwicklung von NS-Netzen abgeleitet werden: Bei der Weiterentwicklung von ländlichen und vorstädtischen NS-Netzen zur Erhöhung der DEA-Leistung kommt der Vermaschung und der im Netz verlegten Leitungslänge eine besondere Rolle zu. Dies bedeutet, dass bei Verlegung weiterer Leitungen diese immer auch zur Erhöhung der Vermaschung des Netzes beitragen sollten.

Insgesamt ist in technischen Anwendungen die Erhebung von Daten zur Charakterisierung der Objekte häufig mit wesentlich weniger Aufwand verbunden, als die Einordnung dieser Objekte in vordefinierte Klassen. Die Bearbeitung der Klassifikationsaufgabe erfordert in der Vielzahl der Anwendungen nicht unerhebliche Ressourcen an Zeit und finanziellen Mitteln. Dies gilt prinzipiell auch für die Klassifikation von NS-Netzen. Dementsprechend bestand eine weitere Zielstellung dieser Arbeit in der Erarbeitung neuer netzspezifischer Ansätze zur Klassifikation von NS-Netzen hinsichtlich ihres Aufnahmevermögens für DEA. Neben einem expertenbasierten Konzept zur Erhebung von ordinalen Klassen mit Hilfe menschlicher Experten, wurde ein weiterer Ansatz auf der Grundlage stochastischer Simulationen vorgestellt. Aus methodischer Sicht handelt es sich grundsätzlich um zwei verschiedene Ansätze. In beiden Ansätzen wurden unterschiedliche Schwerpunkte zur Bewertung von NS-Netzen gesetzt. Der Schwerpunkt des expertenbasierten Ansatzes liegt in der Generierung einer hohen Messqualität hinsichtlich zuvor identifizierter subjektiver Einflüsse auf das Klassifikationsurteil eines Experten. Im Gegensatz dazu wurde der Schwerpunkt der Simulation auf die Bewertung einer Vielzahl von DEA-Konfigurationen in einem NS-Netz gelegt. Es zeigte sich in den hier durchgeführten empirischen Analysen allerdings, dass sich die Ergebnisse der Ansätze zum einen nicht widersprechen und zum anderen sogar mit statistischer Signifikanz konkordant sind. Aus Sicht der Anwendung sind die erhobenen empirischen Ergebnisse zum Aufnahmevermögen der untersuchten Netze für DEA auf Grund des großen Stichprobenumfangs von 300 Netzen auf reale ländliche und vorstädtische Netze verallgemeinerbar. Streng genommen ist den Ergebnissen allerdings auch eine geringe Abhängigkeit von der Betriebsphilosophie (z. B. Schaltzustände der Netze, Experten sind bei gleichem VNB angestellt) immanent, der sowohl die Netzdaten, als auch die Experten für die Untersuchungen zur Verfügung gestellt hat.

Auf Grund der unterschiedlichen Schwerpunkte der erarbeiteten netzspezifischen Klassifikationsansätze und der Tatsache, dass die Nutzung von Expertenwissen aus vielen Gründen mehr oder weniger unsicherheitsbehaftet ist, sind dennoch zum Teil unterschiedliche Ergebnisse für einzelne NS-Netze zu erwarten. Aus diesem Grund bestand ein weiteres Ziel dieser Arbeit in einer adäquaten Kombination von mehreren, potenziell falschen Klassifikationseinschätzungen für ein einzelnes Netz. In diesem Zuge wurde eine neue Kombinationsregel, die EIDMR, erarbeitet, die auf einem $k$nn-Ansatz und Dirichlet-Verteilungen basiert. Neben dieser Regel wurden auch drei weitere, bereits bekannte Kombinationsregeln vorgestellt und hinsichtlich ihrer Kombinationsgüte auf künstlichen Daten verglichen, deren Charakteristik und wahre Klassen bekannt sind. Weiterhin wurden anhand der künstlich generierten Daten verschiedene Eigenschaften der EIDMR untersucht. Aus methodischer Sicht zeigte sich in den Experimenten, dass von den drei weiteren untersuchten Kombinationsregeln die DSR zu den besten Kombinationsergebnissen führt. Allerdings konnte die Kombinationsgüte in einigen Fällen durch Anwendung der EIDMR weiter erhöht werden. Der Mehrwert der EIDMR hing dabei vom verwendeten Datensatz ab. Eine Verbesserung gegenüber der DSR konnte vor allem beobachetet werden, wenn die Anzahl der verfügbaren Expertenaussagen oder die

Überlappung der Komponenten des die Datenpunkte generierenden GMM eher gering ist. Neben der Überlappung der Komponenten des GMM hing die Kombinationsgüte auch von der Anzahl k der Nachbarn und der Dichte der Daten ab. In den Experimenten wirkte sich eine höhere Anzahl von Datenpunkten positiv auf das Kombinationsergebnis aus. Die in der EIDMR vorgesehene Berücksichtigung einer Sicherheitsgewichtung und der ordinalen Information hatten jeweils einen deutlich positiven Effekt auf die Kombinationsgüte. Durch Kombination beider Fälle wurde die Kombinationsgüte weiter erhöht. Ein Einsatz der EIDMR zur Kombination von Klassifkationsergebnissen ist über die in dieser Arbeit vorgestellte Anwendung auf dem Gebiet der Klassifkation von NS-Netzen auch in anderen Anwendungen (z. B. Kombination von Expertenaussagen hinsichtlich einer Fragestellung in einem anderen Themengebiet) möglich.

Das letzte Ziel dieser Arbeit lag in einer Effizienzsteigerung bei der Klassifikation von NS-Netzen durch Anwendung von SVM, um eine optimale Performanz bei der Klassifikation zu erhalten. Dazu wurden verschiedene SVM-basierte Klassifikationskonzepte empirisch untersucht. Aus methodischer Sicht bestand die Schwierigkeit darin, dass die Label zum einen geordnet und zum anderen auf Grund ihrer Erhebung unsicherheitsbehaftet waren. Ausgangsbasis der Untersuchung bildeten Benchmark-SVM, die ohne Anwendung einer Kombinationsregel und eines komplexeren Klassifikationskonzeptes modelliert wurden. In den Experimenten zeigte sich durch Vergleich mit den Ergebnissen dieser Benchmark-SVM deutlich, dass eine Erhebung mehrerer Klassifikationseinschätzungen sowie die Anwendung der Kombinationsregel EIDMR bei unsicheren Labeln zu wesentlich zuverlässigeren SVM-Klassifikatoren als im Benchmark führten. Die höchste Klassifikationsgüte zur Prognose der kombinierten Label wurde beim Training der SVM mit kombinierten Labeln (Konzept 1) erreicht. Die Anwendung der EIDMR zur Kombination der unsicheren Einzellabel führte dabei zu deutlich besseren SVM-Klassifikatoren verglichen mit den anderen drei Kombinationsregeln DSR, MR und IDMR. Durch Nutzung der ordinalen Informationen der Label bei der Modellierung des Klassifikators (Ansatz von Frank und Hall zur Lösung von ordinalen Mehrklassenklassifikationsproblemen) konnte der SVM-Klassifikator weiter verbessert werden. Aus Anwendungssicht liegen anhand der durchgeführten Untersuchungen nun Ergebnisse zur effizienten Klassifikation einer Vielzahl von weiteren NS-Netzen in einer praktischen Anwendung vor. Die Nutzung von graphspezifischen Merkmalen bringt Vorteile, da ihre Erhebung keines Anwendungswissens bedarf und zudem leicht automatisiert werden kann. Darüber hinaus haben Untersuchungen zur unterschiedlichen Gewichtung der Simulationsergebnisse bei Anwendung der EIDMR gezeigt, dass ein guter Kompromiss zwischen einer im Vergleich zu den Expertenaussagen adäquaten Berücksichtigung der Simulationsergebnisse bei der Kombination und einer hohen Klassifikationsgüte des SVM-Klassifikators vorliegt, wenn die Simulationsergebnisse nicht höher als zweifach bei der Kombination gewichtet werden. Die erreichte durchschnittliche Klassifikationsgüte liegt bei zehn durchgeführten dreifach Kreuzvalidierungen bei $\mu(\kappa_{w,test}) = 0{,}82$ und ist als sehr zufriedenstellend zu bewerten. Die Anzahl der Fehlklassifikationen beträgt dabei durchschnittlich 15%, wobei durch die Ausnutzung der

ordinalen Informationen in den Labeln diese Fehlklassifikationen zu einem sehr großen Teil nur um ein Klasse von dem gewünschten Ergebnis abweichen.

Vor dem Hintergrund der Erfüllung aller einzelnen Zielstellungen dieser Arbeit kann auch das übergeordnete Ziel, effiziente Ansätze zur Bewertung der Aufnahmefähigkeit von NS-Netzen für DEA zu erarbeiten, als erreicht angesehen werden. Insbesondere wurden auch die Anwendungsnähe und die Einsatztauglichkeit der erarbeiteten Methoden konsequent durch praxisnahe Experimente an Daten einer Vielzahl realer NS-Netze unterstrichen. Der Einsatz der Methoden in der Praxis ist somit direkt möglich und kann zur Erhöhung der Planungssicherheit bei der Steuerung von Investitionsmitteln auf der NS-Ebene dienen. Eine erhöhte Planungssicherheit sollte einen VNB dazu ermutigen, den Ausbau seiner NS-Netze vorausschauender zu planen und in der Konsequenz technische Effizienzen in der Verteilung elektrischer Energie zu heben. Weiterhin wird durch die erarbeiteten Methoden eine Möglichkeit zur Auswahl von relevanten NS-Netzstrukturen für detaillierte Untersuchungen geschaffen, indem z. B. von jeder Klasse eine bestimmte Anzahl an NS-Netzen für die Untersuchungen ausgesucht wird. Auf diese Weise kann der Anteil an „schwachen“ und „starken“ Netzen, die untersucht werden sollen, gesteuert werden, um repräsentative Ergebnisse über alle Klassen zu erhalten.

Unter Betrachtung der wissenschaftlichen und methodischen Aspekte dieser Arbeit ist festzustellen, dass durchaus (wie in jeder wissenschaftlichen Arbeit) noch weiterer Untersuchungsbedarf zu offengebliebenen Fragestellungen vorliegt. Eine Anregung zu weiteren Forschungsaktivitäten sowie eine Diskussion einiger offengebliebener Fragestellungen findet sich im anschließenden Abschnitt.

7.2 AUSBLICK

Abschließend werden offene Punkte zu eingesetzten Methoden diskutiert und Anregungen für weitere Forschungsarbeiten auf den Themenfeldern abgeleitet. Hinsichtlich der genutzten Datenbasis sollte diese idealerweise aus wenigen relevanten, starken Merkmalen bestehen, um die Informationsflut zu reduzieren und gleichzeitig einen hinreichend hohen Informationsgehalt zu erhalten. Es stellt sich an dieser Stelle also die Frage, welche weiteren Merkmale zur Charakterisierung von NS-Netzen genutzt werden können. Sollte ein hier nicht betrachtetes Merkmal einen hohen Informationsgehalt zur Diskriminierung von NS-Netzen besitzen, so ist zumindest eine Erweiterung der Datenbasis oder aber gar die Substitution gegen ein bisher genutztes Feature mit geringerer Relevanz zu betrachten. Auch der Einsatz weiterer kompletter, hier nicht untersuchter Datensätze erscheint sinnvoll (z. B. städtischer NS-Netze).

Im expertenbasierten Klassifikationsansatz ist bisher eine direkte Einordnung der Objekte in eine Menge vordefinerter ordinaler Klassen vorgesehen. Um die Experten bezüglich der Relation zu anderen in der Befragung zu klassifizierenden Objekten zu unterstützen, wurden Zusatzinformationen wie z. B. ein individueller Klassifikationsindikator eingeführt. Da insbesondere zur Klassifikation von NS-Netzen bisher keine ähnliche Methodik bekannt ist, stellt der expertenbasierte Klassifikationsansatz einen ersten Ansatz zur strukturierten Bewertung von Netzstrukturen durch menschliche Experten dar. Das bedeutet im Umkehrschluss, dass hinsichtlich der Methodik noch viel Untersuchungsbedarf vorliegt. Denkbar wäre z. B. der Einsatz weiterer Zusatzinformationen für die Experten. Die direkte Einordnung der Objekte in eine ordinale Klasse könnte weiterhin z. B. durch einen direkten Vergleich zweier Objekte, in dem ein Experte sich für das stärkere der beiden Netze entscheidet ersetzt werden. Bei einer hinreichend hohen Anzahl an Vergleichen könnte dann anhand der Vergleichsergebnisse anschließend mit geeigneten mathematischen Methoden eine Rangfolge der bewerteten Objekte berechnet werden.

Auch in dem vorgestellten simulationsbasierten Klassifikationsansatz lassen sich direkt einige Verbesserungspotenziale identifizieren. So werden zur Klassifikation der NS-Netze lediglich der aktuelle Netz- und Schaltzustand berücksichtigt. Mögliche sinnvolle Änderungen der Netzstruktur werden damit nicht betrachtet. Außerdem werden auch keine Regelungsmöglichkeiten der installierten DEA (z. B. durch Blindleistungsaustausch mit dem Netz) in die Klassifikation einbezogen. Die Untersuchung von Einflüssen auf die Zuverlässigkeit eines Netzes sowie die Veränderung der Netzverluste durch den Zubau von DEA erscheint ebenfalls interessant. Insbesondere zur Berücksichtigung des Einflusses von DEA auf die Netzverluste wurden bereits wertvolle Grundlagen geschaffen [112, 113, 114].

Die Kombination unsicherer Klassifikationseinschätzungen beschränkt sich in dieser Arbeit auf den Einsatz von Kombinationsregeln. Dabei wurden je zwei DST- und zwei IDM-basierte Kombinationsregeln untersucht. Mittlerweile exisitiert allerdings bereits eine Vielzahl an Kombinationsregeln. Über den Einsatz von Kombinationsregeln hinaus werden zur Behandlung von Unsicherheiten und Widersprüchen in Experteneinschätzungen verschiedenste Ansätze verfolgt. Zu nennen sind hier beispielhaft die Modellierung von Expertenaussagen als Menge von Wahrscheinlichkeitsverteilungen und deren Fusion zu einer einzigen Verteilungsfunktion. Auch Rough-Set-basierte Modelle werden häufig zur Begegnung der Unsicherheit in Expertenaussagen eingesetzt. In dieser Arbeit wurden die mit der EIDMR berechneten Unsicherheitswerte bei der Entscheidung für eine Klasse nicht weiter verwendet. Auch hier liegt ein Potenzial für weiterführende Untersuchungen zur Verwendung dieser Unsicherheitswerte.

Bei der Anwendung von SVM stellte sich die höchste Klassifikationsgüte beim Training mit kombinierten Labeln (Konzept 1) unter Anwendung des Ansatzes von Frank und Hall zur Lösung von ordinalen Mehrklassenklassifikationsproblemen ein. Um die

Selektion und Kombination von Features über die hier durchgeführten Untersuchungen hinaus näher zu analysieren, könnten z. B. auch Wrapper-Verfahren eingesetzt werden, die bei der Suche einer guten Featuremenge diese jeweils auf den Klassifikator anwenden. Vor diesem Hintergrund könnten in zukünftigen Anwendungen weitere komplexere Methoden angewendet werden, um die Klassifikationsgüte weiter zu erhöhen. Hier stellt sich die Frage, ob durch andere Techniken zur Berücksichtigung ordinaler Label in der SVM verbesserte Ergebnisse erreicht werden können. Zudem beschränkte sich hier die Anwendung lediglich auf SVM. Es ist allerdings durchaus denkbar, dass in weiteren Anwendungen unter dem Einsatz anderer Methoden höhere Klassifikationsgüten zu erreichen sind.

ANHANG

A

WEITERE ERGEBNISSE

Im Folgenden werden ergänzende Ergebnisse zu den in Abschnitt 6.3.3 durchgeführten Experimenten dargelegt. Die Ergebnisse umfassen dabei zu den diskutierten Experimenten analoge Untersuchungen, in denen unterschiedlicher Features (z. B. graphentheoretische Features) oder unterschiedliche Kombinationsregeln angewendet worden sind (z. B. DSR).

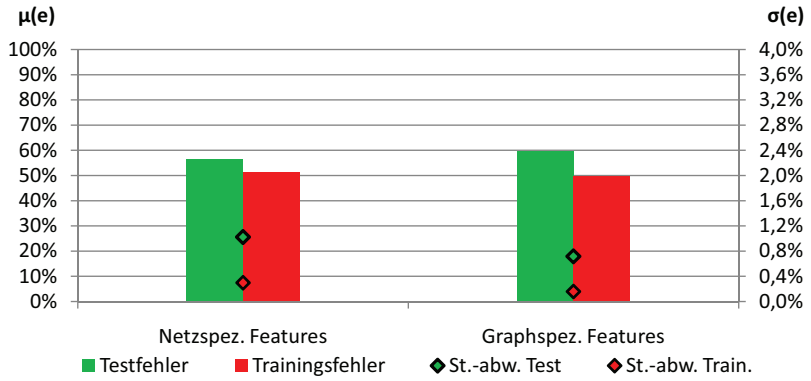


Abbildung 37: Klassifikationsfehler einer nur anhand der Einzellabel trainierten und getesteten SVM (Benchmark 1).

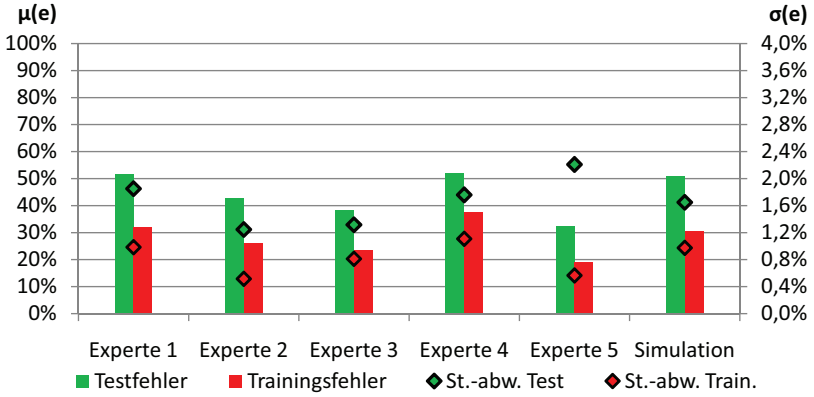


Abbildung 38: Klassifikationsfehler für jeweils anhand von Einzellabeln in Kombination mit netzspezifischen Features trainierte SVM (Benchmark 2).

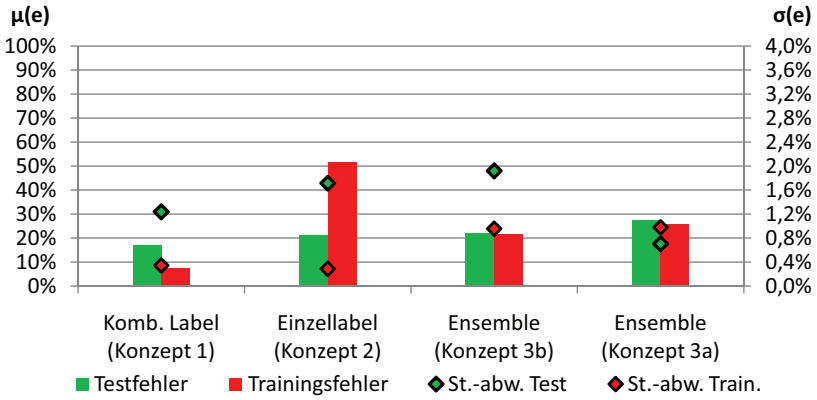


Abbildung 39: Gegenüberstellung der Klassifikationsfehler verschiedener Klassifikationskonzepte für kombinierte Label (EIDMR) und alle netzspezifischen Features.

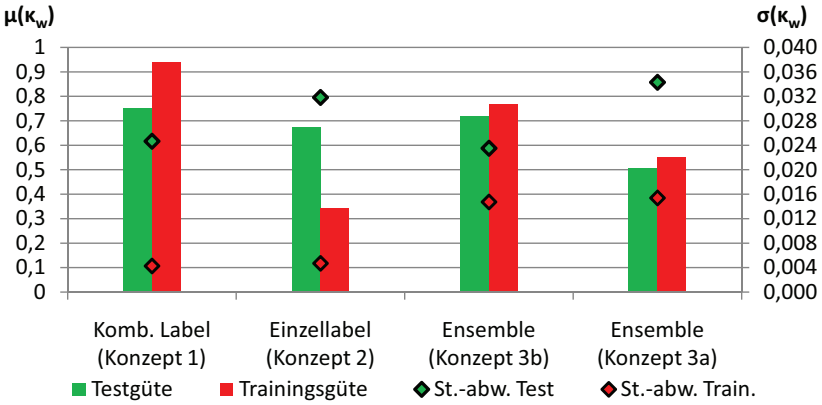


Abbildung 40: Gegenüberstellung der Klassifikationsgüten zur Prognose der kombinierten Label (EIDMR) für verschiedene Klassifikationskonzepte und alle graphspezifischen Features.

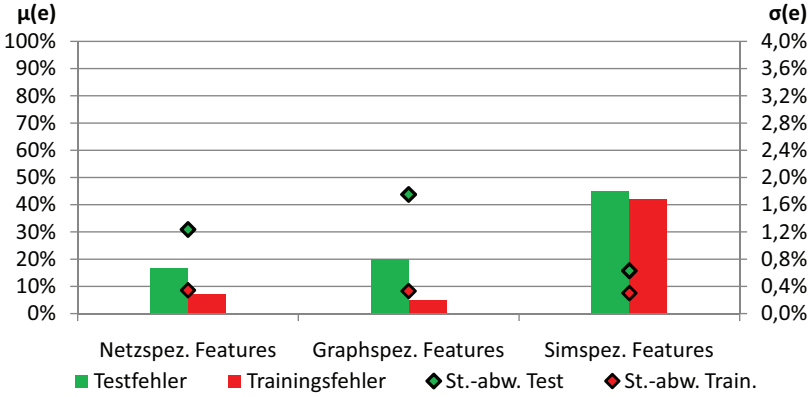


Abbildung 41: Gegenüberstellung der Klassifikationsfehler zur Prognose der kombinierten Label (EIDMR) für verschiedene Featuresätze.

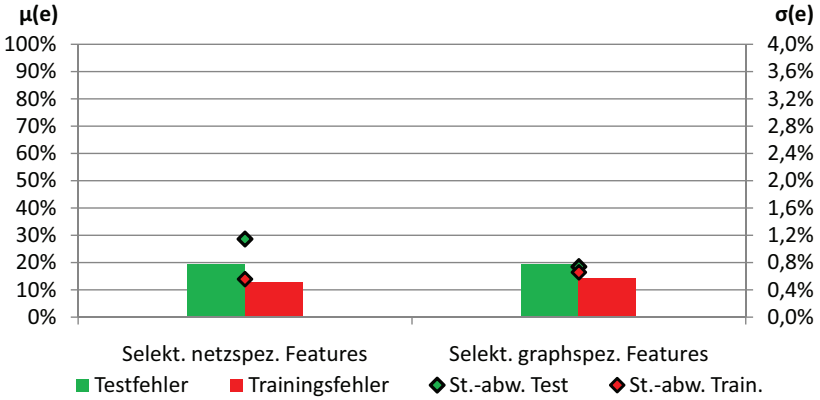


Abbildung 42: Gegenüberstellung der Klassifikationsfehler bei Selektion von Features mit höherer Relevanz zur Prognose der kombinierten Label (EIDMR).

Abbildung 43: Klassifikationsfehler bei Kombination von Features zur Prognose der kombinierten Label (EIDMR).

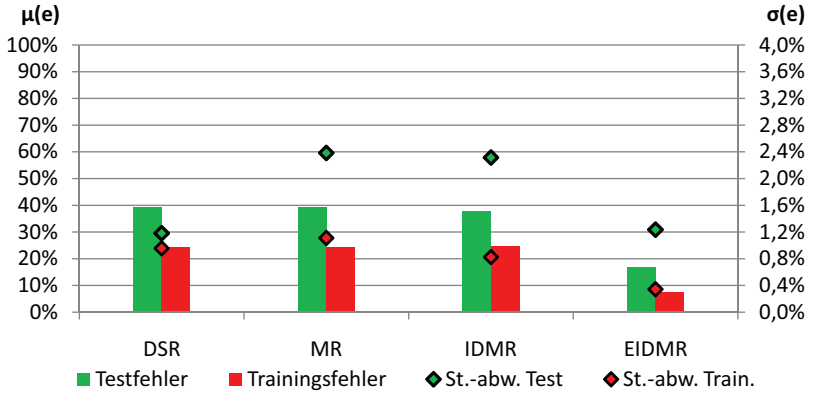


Abbildung 44: Einfluss der verwendeten Kombinationsregel auf den Klassifikationsfehler für alle netzspezifischen Features.

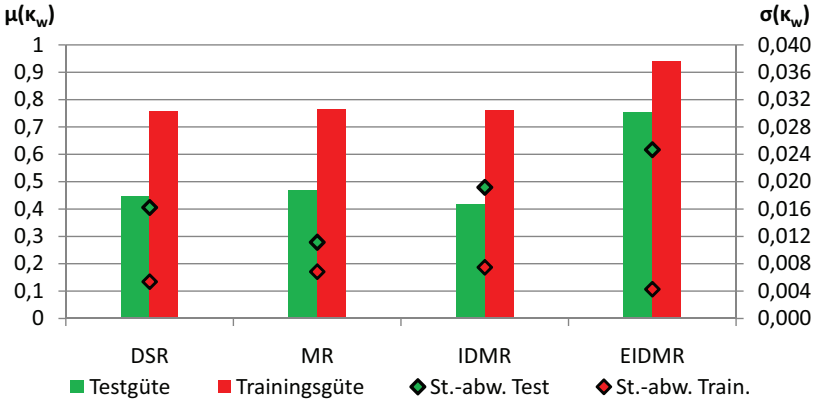


Abbildung 45: Einfluss der verwendeten Kombinationsregel auf die Klassifikationsgüte für alle graphspezifischen Features.

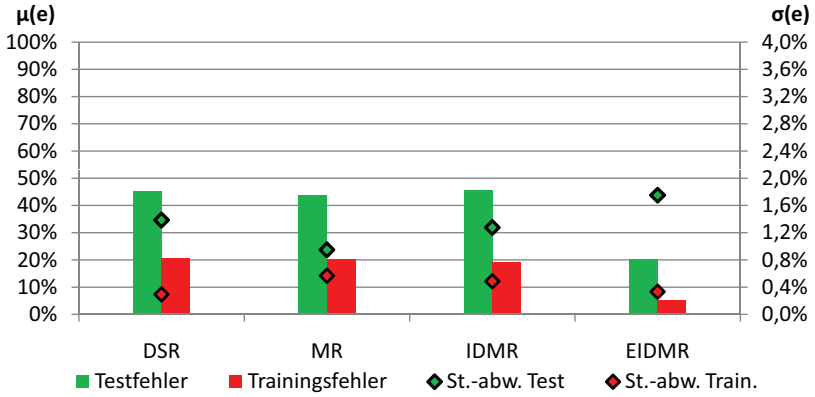


Abbildung 46: Einfluss der verwendeten Kombinationsregel auf den Klassifikationsfehler für alle graphspezifischen Features.

Abbildung 47: Einfluss der Gewichtung der Simulationsergebnisse auf den Klassifikationsfehler für alle netzspezifischen Features.

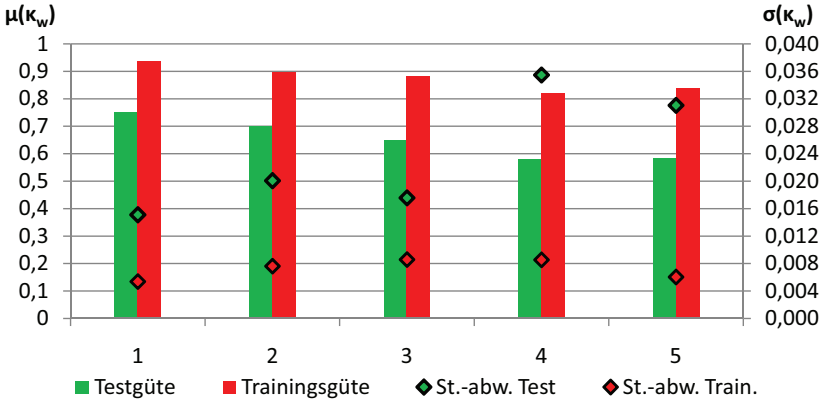


Abbildung 48: Einfluss der Gewichtung der Simulationsergebnisse auf die Klassifikationsgüte zur Prognose der kombinierten Label (EIDMR) für alle graphspezifischen Features.

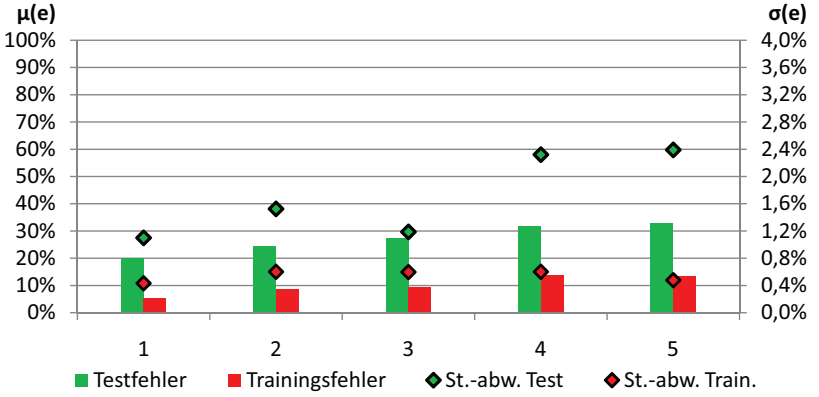


Abbildung 49: Einfluss der Gewichtung der Simulationsergebnisse auf den Klassifikationsfehler zur Prognose der kombinierten Label (EIDMR) für alle graphspezifischen Features.

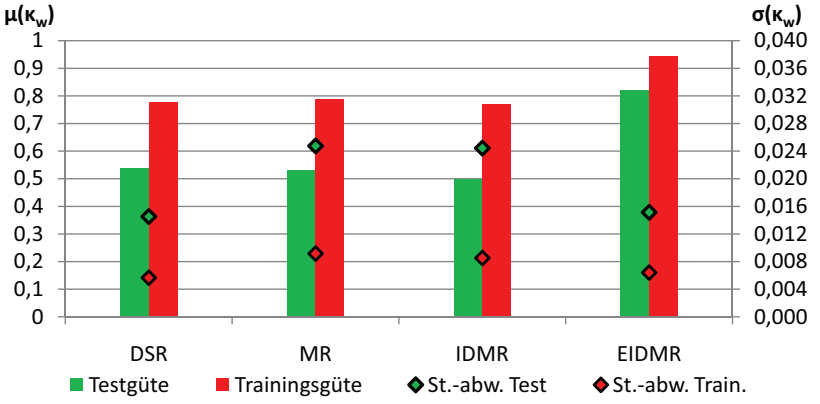


Abbildung 50: Klassifikationsgüte zur Prognose der kombinierten Label mit dem Ansatz von Frank und Hall für alle netzspezifischen Features.

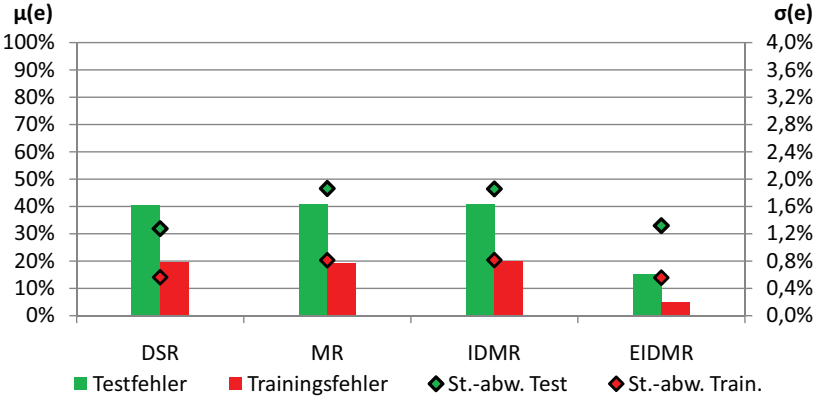


Abbildung 51: Klassifikationsfehler zur Prognose der kombinierten Label mit dem Ansatz von Frank und Hall für alle netzspezifischen Features.

LITERATURVERZEICHNIS

[1] E-Bridge, Institut für Elektrische Anlagen und Energiewirtschaft, Offis: Abschlussbericht zum Forschungsprojekt Nr. 44/12: Moderne Verteilernetze für Deutschland (Verteilernetzstudie). In: *Studie im Auftrag des Bundesministeriums für Wirtschaft und Energie (BMWi)* (2014)

[2] Ault, G. W. ; McDonald, J. R. ; Burt, G ..: Strategic Analysis Framework for Evaluating Distributed Generation and Utility Strategies. In: *IEEE Proceedings on Generation, Transmission and Distribution* 150 (2003), Nr. 4, S. 475–481

[3] Puttgen, H. B. ; MacGregor, P. R. ; Lambert, F. C.: Distributed Generation: Semantic Hype or the Dawn of a new era? In: *IEEE Power and Energy Magazine* 1 (2003), Nr. 1, S. 1–5

[4] Chiradeja, P.: Benefit of Distributed Generation: A Line Loss Reduction Analysis. In: *Transmission and Distribution Conference and Exhibition: Asia and Pacific, Yokohama* (2005), S. 1–5

[5] Driesen, J. ; Belmans, R.: Distributed Generation: Challenges and Possible Solutions. In: *Power Engineering Society General Meeting, Montreal* (2006), S. 1–5

[6] Kerber, G. ; Witzmann, R.: Aufnahmefähigkeit der Verteilnetze für Strom aus Photovoltaik. In: *Energiewirtschaft* 106 (2007), Nr. 4, S. 50–54

[7] Schulz, D.: *Netzrückwirkungen-Theorie, Simulation, Messung und Bewertung.* VDE-Verlag GmbH-Berlin,Offenbach, 2004

[8] CENELEC: European Comitee for Electrotechnical Standardization: *Voltage Characteristics of Electricity Supplied by Public Distribution Networks.* German version EN 50160:2010 + Cor., 2010

[9] Verband der Elektrotechnik Elektronik Informationstechnik e.V. (VDE): *Technische Mindestanforderungen für Anschluss und Parallelbetrieb von Erzeugungsanlagen am Niederspannungsnetz.* VDE-AR-N 4105 Erzeugungsanlagen am Niederspannungsnetz, 2011

[10] VDE-Verband der Elektrotechnik Elektronik Informationstechnik e.V.: *DIN VDE 0838.* Technische Norm, 2011

[11] VDE-Verband der Elektrotechnik Elektronik Informationstechnik e.V.: *DIN VDE 0839*. Technische Norm, 2011

[12] Schölkopf, B. ; Müller, K.-R. ; Smola, A. J.: Support-Vektor-Methoden zur Analyse hochdimensionaler Daten. In: *Informatik Forschung und Entwicklung* 14 (1999), Nr. 1, S. 154–163

[13] Breker, S. ; Claudi, A. ; Sick, B.: Capacity of Low-Voltage Grids for Distributed Generation: Classification by Means of Stochastic Simulations. In: *IEEE Transactions on Power Systems* 30 (2015), Nr. 2, S. 689–700

[14] Rudin, C. ; Waltz, D. ; Anderson, R. N. ; Boulanger, A.: Machine Learning for the New York City Power Grid. In: *IEEE Transactions on Pattern Analysis and Machine Intelligence* 34 (2012), Nr. 2, S. 328–345

[15] Ma, Y. ; Long, Z. ; Tse, N. ; Osman, A.: An Initial Study on Computational Intelligence for Smart Grid. In: *International Conference on Machine Learning and Cybernetics, Baoding* (2009), S. 3425–3429

[16] Venayagamoorthy, G. K.: Potentials and Promises of Computational Intelligence for Smart Grids. In: *Power & Energy Society General Meeting, Calgary* (2009), S. 1–6

[17] Haibo, H.: Toward a Smart Grid: Integration of Computational Intelligence Into Power Grid. In: *The 2010 International Joint Conference on Neural Networks (IJCNN), Barcelona* (2010), S. 1–6

[18] Cartes, D. ; Chow, J. H. ; McCaugherty, D. ; Widergren, S.: The IEEE Computer Society Smart Grid Vision Project Opens Opportunites for Computational Intelligence. In: *2013 IEEE Conference on Evolving and Adaptive Intelligent Systems (EAIS), Singapore* (2013), S. 144–150

[19] Byung-Gook, K. ; Yu, Z. ; Schaar, M. van d. ; Lee, L. J.: Dynamic Pricing for Smart Grid with Reinforcement Learning. In: *IEEE Conference on Computer Communications Workshops, Toronto* (2014), S. 640–645

[20] Zhang, Y. ; Ilic, M. D. ; Tonguz, O.: Application of Support Vector Machine Classification to Enhanced Protection Relay Logic in Electric Power Grids. In: *Large Engineering Systems Conference on Power Engineering, Montreal* (2007), S. 31–38

[21] Pisica, I. ; Eremia, M.: Making Smart Grids Smarter by Using Machine Learning. In: *Proceedings of 2011 International Universities' Power Engineering Conference (UPEC), Soest* (2011), S. 1–5

[22] Shahid, N. ; Aleem, S. A. ; Naqvi, I. H. ; Zaffar, N.: Support Vector Machine Based Fault Detection & Classification in Smart Grids. In: *2012 IEEE Globecom Workshops, Anaheim* (2012), S. 1526–1531

[23] Haruna, Y .. ; Bakare, G. A. ; Aliyu, U. O.: Adaptive Static Load Shedding Schemes for Nigerian Grid System Based on Computational Intelligence Techniques. In: *Power Systems Conference, Clemson* (2014), S. 1–6

[24] Li, B. ; Gangadhar, S. ; Cheng, S. ; Verma, P. K.: Predicting User Comfort Level Using Machine Learning for Smart Grid Environments. In: *IEEE PES Innovative Smart Grid Technologies (ISGT), Anaheim* (2011), S. 1–6

[25] Yuanzhe, C. ; Qing, X. ; Chengqiang, W. ; Fangcheng, L.: Short-Term Load Forecasting for City Holidays Based on Genetic Support Vector Machines. In: *2011 International Conferene on Electrical and Control Engineering (ICECE), Yichang* (2011), S. 3144–3147

[26] Chandler, S. A. ; Hughes, J. G.: Smart Grid Distribution Prediction and Control Using Computational Intelligence. In: *2013 IEEE Conference on Technologies for Sustainability (SusTech), Portland* (2013), S. 86–89

[27] Stark, M. ; Krost, G.: Design Tool for Isolated Micro-Grids Based on Methods of Computational Intelligence. In: *International Conference on Intelligent System Application to Power Systems (ISAP), Hersonissos* (2011), S. 1–6

[28] Gross, P. ; Salleb-Aouissi, A. ; Dutta, H. ; Boulanger, A.: Ranking Electrical Feeders of the New York Power Grid. In: *International Conference on Machine Learning and Applications, Miami Beach* (2009), S. 359–365

[29] Li, G. ; Semerci, M. ; Yener, B. ; Zaki, M. J.: Graph Classification via Topological and Label Attributes. In: *9th Workshop on Mining and Learning with Graphs, San Diego* (2011), S. 1–9

[30] Newman, M. E. J.: *Networks: An Introduction*. Oxford University Press, Oxford, 2010

[31] Steger, A.: *Diskrete Strukturen*. Springer, Berlin, 2007

[32] Gärtner, T. ; Flach, P. ; Wrobel, S.: On Graph Kernels: Hardness Results and Efficient Alternatives. In: *16th Annual Conference on Computational Learning Theory, Washington DC* (2003), S. 129–143

[33] Kashima, H. ; Tsuda, K. ; Inokuchi, A.: Marginalized Kernels Between Labeled Graphs. In: *International Conference on Machine Learning, Los Angeles* (2003), S. 321–328

[34] Borgwardt, K. M. ; Ong, C. S. ; Schönauer, S. ; Vishwanathan, S. V. N.: Protein Function Prediction via Graph Kernels. In: *International Conference on Intelligent Systems for Molecular Biology, Detroit* (2005), S. 47–56

[35] Borgwardt, K. M. ; Kriegel, H.-P.: Shortest Path Kernels on Graphs. In: *5th IEEE International Conference on Data Mining, Washington DC* (2005), S. 74–81

[36] Horvath, T. ; Gärtner, T. ; Wrobel, S.: Cyclic Pattern Kernels for Predictive Graph Mining. In: *10th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, Seattle* (2004), S. 158–167

[37] Mahe, P. ; Vert, J.-P.: Graph Kernels Based on Tree Patterns for Molecules. In: *Machine Learning* 75 (2009), Nr. 1, S. 3–35

[38] Ralaivola, L. ; Swamidass, S. J. ; Saigo, H. ; Baldi, P.: Graph Kernels for Chemical Informatics. In: *Neural Networks* 18 (2005), Nr. 8, S. 1093–1110

[39] Ramon, J. ; Gärtner, T.: Expressivity Versus Efficiency of Graph Kernels. In: *1st International Workshop on Mining Graphs, Trees and Sequences, Osaka* (2003)

[40] Shervashidze, N. ; Borgwardt, K. M.: Fast Subtree Kernels on Graphs. In: *23rd Annual Conference on Neural Information Processing Systems, Curran* (2009), S. 1660–1668

[41] Kondor, I. R. ; Shervashidze, N. ; Borgwardt, K. M.: The Graphlet Spectrum. In: *International Conference on Machine Learning, Montreal* 382 (2009), S. 67

[42] Shervashidze, N. ; Vishwanathan, S. V. ; Petri, T. H. ; Mehlhorn, K. ; Borgwardt, K. M.: Efficient Graphlet Kernels for Large Graph Comparison. In: *12th International Conference on Artificial Intelligence and Statistics, Clearwater Beach* (2009), S. 488–495

[43] Butler, S.: *Eigenvalues and Structures of Graphs.* Dissertation, University of California, San Diego, 2011

[44] Chung, F.: *Spectral Graph Theory.* Conference Board of the Mathematical Sciences, Fresno, 1997

[45] Chung, F. ; Butler, S.: *Handbook of Linear Algebra: Spectral Graph Theory.* CRC Press, Boca Raton, 2013

[46] Freeman, L.: A set of measures of centrality based on betweenness. In: *Sociometry* 40 (1977), Nr. 1, S. 35–41

[47] Walsh, S. ; Dalton, A. ; Whitney, P. ; White, A.: Parameterizing Bayesian Network Representations of Social-Behavioral Models by Expert Elicitation. In: *International Conference on Intelligence and Security Informatics (ISI), Vancouver* (2010), S. 227–232

[48] Moser, S.: *Konstruktivistisch forschen – Methodologie, Methoden, Beispiele*. Verlag für Sozialwissenschaften, Wiesbaden, 2011

[49] Stevens, S. S.: On the use of Expert Judgement on Complex Technical Problems. In: *IEEE Transactions on Engineering Management* 36 (1989), Nr. 2, S. 83–86

[50] Chytka, T. M. ; Conwell, B. A. ; Unal, R.: An Expert Judgment Approach for Addressing Uncertainty in High Technology System Design. In: *Technology Management for the Global Future, Istanbul* (2006), S. 444–449

[51] Siordia, O. S. ; Diego, I. M. ; Conde, C. ; Reyes, G.: Driving Risk Classification Based on Experts Evaluation. In: *IEEE Intelligent Vehicles Symposium (IV), San Diego* (2008), S. 1098–1103

[52] MacInnes, J. ; Santosa, S. ; Wright, W.: Visual Classification: Expert Knowledge Guides Machine Learning. In: *IEEE Computer Graphics and Applications* 30 (2010), Nr. 1, S. 8–14

[53] Sandri, S. A. ; Dubois, D. ; Kalfsbeek, H. W.: Elicitation, Assessment, and Pooling of Expert Judgments Using Possibility Theory. In: *IEEE Transactions on Fuzzy Systems* 3 (1995), Nr. 3, S. 313–335

[54] Bortz, J. ; Lienert, G. A. ; Boehnke, K.: *Verteilungsfreie Methoden in der Biostatistik*. Springer, Heidelberg, 2008

[55] Häder, M.: *Delphi-Befragungen. Ein Arbeitsbuch*. Westdeutscher Verlag, Wiesbaden, 2002

[56] Mullin, T. M.: Experts' Estimation of Uncertain Quantities and its Implications for Knowledge Acquisition. In: *IEEE Transactions on Systems, Man, and Cybernetics* 19 (1989), Nr. 3, S. 616–625

[57] Dawes, R. M.: The Robust Beauty of Improper Linear Models in Decision Making. In: *American Psychology* 34 (1979), Nr. 7, S. 571–582

[58] Dawes, R. M.: Man Versus Model of Man: A Rationale Plus Some Evidence for a Method of Improving Clinical Inference. In: *Psychatry Bulletin* 73 (1970), Nr. 6, S. 422–432

[59] Stevens, S. S.: *Handbook of Experimental Psychology: Mathematics, Measurement, and Psychophysics*. Wiley, New York, 1951

[60] Warrens, M. J.: Equivalences of Weighted Kappas for Multiple Raters. In: *Statistical Methodology* 9 (2012), Nr. 3, S. 407–422

[61] Maclure, M. ; Willett, W. C.: Misinterpretation and Misuse of the Kappa Statistic. In: *American Journal of Epidemiology* 126 (1987), Nr. 2, S. 161–169

[62] Mielke, P. W. ; Berry, K. J.: A Note on Cohen's Weighted Kappa Coefficient Of Agreement With Linear Weights. In: *Statistical Methodology* 6 (2009), Nr. 5, S. 439–446

[63] Warrens, M. J.: Cohen's Quadratically Weighted Kappa is Higher Than Linearly Weighted Kappa for Tridiagonal Agreement Tables. In: *Statistical Methodology* 9 (2012), Nr. 3, S. 440–444

[64] Papaioannou, I. T. ; Alexiadis, M. C. ; Demoulias, C. S. ; Dokopoulos, P. S.: Modeling and Field Measurements of Photovoltaic Units Connected to LV Grid. Study of Penetration Scenarios. In: *IEEE Transactions on Power Delivery* 26 (2011), Nr. 2, S. 979–987

[65] Jenicek, J. ; Inam, W. ; Ilic, M.: Locational Dependence of Maximum Installable PV-Capacity in LV-Networks while Maintaining Voltage Limits. In: *North American Power Symposium (NAPS), Boston* (2011), S. 1–5

[66] Tie, C. H. ; Gan, C. K.: Impact of Grid-Connected Residential PV Systems on the Malaysian Low-Voltage Distribution Network. In: *7th International Power Engineering and Optimization Conference (PEOCO), Langkawi* (2013), S. 670–675

[67] Stetz, T. ; Marten, F. ; Braun, M.: Improved Low Voltage Grid-Integration of Photovoltaic Systems in Germany. In: *IEEE Transactions on Sustainable Energy* 4 (2013), Nr. 2, S. 534–542

[68] Kerber, G. ; Witzmann, R.: Statistische Analyse von NS-Verteilungsnetzen und Modellierung von Referenznetzen. In: *Energiewirtschaft* 107 (2008), Nr. 6, S. 22–26

[69] Verband der Elektrizitätswirtschaft e. V. (VDEW): *Richtlinie Eigenerzeugungsanlagen am Niederspannungsnetz*. VDEW, 2001

[70] Ben-Israel, A.: A Newton-Raphson Method for the Solution of Systems of Equations. In: *Journal of Mathematical Analysis and Applications* 15 (1966), Nr. 2, S. 243–253

[71] Evans, J. W. ; Johnson, R. A. ; Green, D. W.: Two- and Three-Parameter Weibull Goodness-of-Fit Test. In: *United States, Department of Agriculture, Madison, Wisconsin* PL-RP-493 (1989), S. 1–27

[72] Stapelberg, R. F.: *Handbook of Reliability, Availability, Maintainability and Safety in Engineering Design*. Springer, London, 2009

[73] Ng, H. K. T. ; Luo, L. ; Hu, Y. ; Duan, F.: Parameter Estimation of Three-Parameter Weibull Distribution Based on Progressively Type-II Censored Samples. In: *Journal of Statistical Computation and Simulation* 82 (2012), Nr. 11, S. 1661–1678

[74] Nagatsukaa, H. ; Kamakuraba, T. ; Balakrishnanc, N.: A Consistent Method of Estimation for the Three-Parameter Weibull Distribution. In: *Computational Statistics and Data Analysis* 58 (2013), Nr. 3, S. 210–226

[75] Moeinia, A. ; Kouroush, J. ; Mohsen, M. ; Mehdi, F.: Fitting the Three-Parameter Weibull Distribution With Cross Entropy. In: *Applied Mathematical Modelling* 37 (2013), Nr. 9, S. 6354–6363

[76] Myers, R. H.: *Classical and Modern Regression With Applications*. Duxbury Classic Series, Pacific Grove, 1990

[77] Brandt, S.: *Data Analysis*. Springer, New York, 1999

[78] Li, J. D. ; Zhang, X. J. ; Chen, Y. S.: Applying Expert Experience to Interpretable Fuzzy Classification System Using Genetic Algorithms. In: *Fourth International Conference on Fuzzy Systems and Knowledge Discovery, Haikou* (2007), S. 129–133

[79] Shafer, G.: *A Mathematical Theory of Evidence*. Princeton University Press, Princeton, 1976

[80] Martin, A. ; Osswald, C.: Experts Fusion and Multilayer Perceptron Based on Belief Learning for Sonar Image Classification . In: *3rd International Conference on Information and Communication Technologies (ICTTA): From Theory to Applications, Damascus* (2008), S. 1–6

[81] Panchal, V. K. ; Gupta, S. ; Gupta, N. ; Monga, M.: Eliciting Conflict In Expert's Decision For Land Use Classification. In: *International Conference on Environmental Engineering and Applications (ICEEA), Singapore* (2010), S. 30–33

[82] Murphy, C.: Combining Belief Functions When Evidence Conflicts. In: *Decision Support Systems* 29 (2000), Nr. 1, S. 1–9

[83] Sentz, K. ; Ferson, S.: Combination of Evidence in Dempster-Shafer-Theory. In: *Sandia National Laboratories, Technical Report* (2002)

[84] Andrade, D. ; Horeis, T. ; Sick, B.: Knowledge Fusion Using Dempster-Shafer Theory and the Imprecise Dirichlet Model. In: *IEEE Conference on Soft Computing in Industrial Applications, Muroran* (2008), S. 142–148

[85] Hua-wei, G. ; Wen-kang, S. ; Yong, D.: Evidential Conflict and its 3D Strategy: Discard, Discover and Disassemble. In: *Systems Engineering and Electronics* 29 (2007), Nr. 6, S. 890–898

[86] Zadeh, L. A.: Review of Books: A Mathematical Theory of Evidence. In: *The AI Magazine* 5 (1984), Nr. 3, S. 81–83

[87] Yi-jie, D. ; She-liang, W. ; Xin-dong, Z. ; Miao, L. ; Xi-qin, Y.: A Modified Dempster-Shafer Combination Rule Based on Evidence Ullage. In: *International Conference on Artificial Intelligence and Computational Intelligence, Shanghai* (2009), S. 387–391

[88] Smarandache, F. ; Dezert, J.: *Advances and Applications of DSmT for Information Fusion: Collected Works*. American Research Press, Rehoboth, 2009. – 108–113 S.

[89] Tchamova, A. ; Dezert, J.: On the Behavior of Dempster's Rule of Combination and the Foundations of Dempster-Shafer Theory. In: *IEEE International Conference on Intelligent Systems, Sofia* (2012), S. 108–113

[90] Smarandache, F. ; Dezert, J. ; Kroumov, V.: Examples Where the Conjunctive and Dempster's Rules are Insensitive. In: *Proceedings of the 2013 International Conference on Advanced Mechatronic Systems, Luoyang* (2013), S. 7–9

[91] Walley, P.: Inferences From Multinomial Data: Learning About a bag of Marbles. In: *Journal of Royal Statistical Society, Series B (Methodological)* 58 (1996), Nr. 1, S. 3–57

[92] Utkin, L. V.: Extensions of Belief Functions and Possibility Distributions by Using the Imprecise Dirichlet Model. In: *Fuzzy Sets and Systems* 154 (2005), Nr. 3, S. 413–431

[93] Troffaes, M. C. M. ; Coolen, F.: On the Use of the Imprecise Dirichlet Model With Fault Trees. In: *Technical Report, Durham University.* (2007), S. 1–8

[94] Whitaker, E. T. ; Watson, G. N.: *A Course Of Modern Analysis*. Cambridge University Press, Cambridge, 1996

[95] Chang, C.-C. ; Lin, C.-J.: A Practical Guide to Support Vector Classification. In: *Technical Report, National Taiwan University.* (2010), S. 1–16

[96] Cortes, C. ; Vapnik, V.: Support-Vector Networks. In: *Machine Learning* 20 (1995), Nr. 3, S. 273–297

[97] Christianini, N. ; Shawe-Taylor, J.: *An Introduction to Support Vector Machines and Other Kernel-based Learning Methods*. Cambridge University Press, Cambridge, 2000

[98] Sanchez, V. D.: Advanced Support Vector Machines and Kernel Methods. In: *Neuro Computing* 55 (2003), Nr. 1, S. 5–20

[99] Lin, C. F. ; Wang, S. D.: Fuzzy Support Vector Machines. In: *IEEE Transactions on Neural Networks* 13 (2002), Nr. 2, S. 464–471

[100] Chen, D. G. ; He, Q. ; Wang, X. Z.: FRSVMs: Fuzzy Rough Set Based Support Vector Machine. In: *Fuzzy Sets and Systems* 161 (2010), Nr. 4, S. 596–607

[101] Zhang, J. H. ; Wang, Y. Y.: A Rough Margin Based Support Vector Machine. In: *Information Sciences* 178 (2008), Nr. 9, S. 2204–2214

[102] Abe, S.: *Support Vector Machines for Pattern Classification*. Advances in Pattern Recognition, Springer, London, 2010

[103] Frank, E. ; Hall, M.: A Simple Approach to Ordinal Classification. In: *Proceedings of the European Conference on Machine Learning, Freiburg* (2001), S. 145–165

[104] Chu, W. ; Keerthi, S.: Support Vector Ordinal Regression. In: *Neural Computing* 19 (2007), Nr. 3, S. 792–815

[105] Steinwart, I. ; Christmann, A.: *Support Vector Machines*. Information Science and Statistics, Springer, London, 2008

[106] Mercer, J.: Functions of Positive and Negative Type and their Connection with the Theory of Integral Equations. In: *Philosophical Transactions of the Royal Society* 1 (1909), Nr. 209, S. 415–446

[107] Platt, J. C.: *Sequential Minimal Optimization: A Fast Algorithm for Training Support Vector Machines*. Advances in Kernel Methods - Support Vector Learning, MIT Press, Cambridge, 1998. – `http://research.microsoft.com/apps/pubs/default.aspx?id=68391` (13.01.2015)

[108] Chang, C.-C. ; Lin, C.-J.: LIBSVM: A Library for Support Vector Machines. In: *ACM Transactions on Intelligent Systems and Technology* 2 (2011), Nr. 3, S. 1–27. – Software available at `http://www.csie.ntu.edu.tw/~cjlin/libsvm` (13.01.2015)

[109] Guyon, I. ; Elisseeff, A.: An Introduction to Variable and Feature Selection. In: *Journal of Machine Learning Research* 3 (2003), Nr. 1, S. 1157–1182

[110] Liu, H. ; Motoda, H.: *Computational Methods of Feature Selection*. Chapman & Hall, Boca Raton, 2008

[111] Liu, H. ; Setiono, R.: Chi^2: Feature Selection and Discretization of Numeric Attributes. In: *Proceedings on 7th IEEE International Conference on Tools with Artificial Intelligence, Herndon* (1995), S. 338–391

[112] Chiradeja, P. ; Ramakumar, R.: An Approach to Quantify the Technical Benefits of Distributed Generation. In: *IEEE Transactions on Energy Conversion* 19 (2004), Nr. 4, S. 764–773

[113] Atwa, Y. M. ; El-Saadany, E. F. ; Salama, M. M. A. ; Seethapathy, R.: Optimal Renewable Resources Mix for Distribution System Energy Loss Minimization. In: *IEEE Transactions on Power Systems* 25 (2010), Nr. 1, S. 360–370

[114] Borges, C. L. T. ; Falcao, D. M.: Optimal Distributed Generation Allocation for Reliability, Losses and Voltage Improvement. In: *Electrical Power and Energy Systems* 28 (2006), Nr. 6, S. 413–420

ISBN 978-3-7376-0014-9