

Anhang: Die Funktionsbeziehungen des Flussdiagramms

Funktionsbeziehung I: Sie beschreibt die Fahrleistungsreduktion des PKW-Freizeitverkehrs bei steigendem Kraftstoffpreis ($\epsilon = -0,17$ bzw. $\epsilon = 0,13$).

Funktionsbeziehung II: Es wird die Wirkung der Kraftstoffkosten auf die Verkehrsmittelwahl nachgebildet ($\epsilon = -0,05$ bzw. $\epsilon = -0,08$).

Funktionsbeziehung III: Die Beziehung zeigt den Zusammenhang zwischen der Verkehrsmittelwahl und den Fahrleistungen im Freizeitverkehr auf ($\epsilon = -0,05 + (-0,17)$ bzw. $\epsilon = 0,08 + (-0,13)$).

Funktionsbeziehung IV: Sie zeigt, welche Auswirkungen die Einkommenshöhe auf die Verkehrsmittelwahl hervorruft.

Funktionsbeziehung V: Hier wird der Zusammenhang zwischen der Einkommenshöhe und den PKW-Fahrleistungen aufgezeigt.

Funktionsbeziehung VI: Diese Funktion zeichnet die Wirkungsweise der Kraftstoffkosten auf die Kfz-Technik nach.

Funktionsbeziehung VII: Hier wird die Wirkung der Kfz-Technik auf die PKW-Fahrleistungen beschrieben.

Funktionsbeziehung VIII: Sie zeigt auf, wie sich eine verbrauchsmindernde Kfz-Technik auf die streckenbezogenen Kraftstoffkosten auswirkt.

Modellierung von Mobilitätsdaten mit Methoden der Künstlichen Intelligenz

VON JOACHIM HUGO, BERLIN

Im Rahmen dieser Arbeit wird der herkömmliche Ansatz zur Modellierung des Modal-Splits (Logit-Modelle) verglichen mit Ansätzen der Künstlichen Intelligenz. Dies sind neuronale Netzwerke, Fuzzy-Systeme sowie eine Kombination beider Methoden, Neuro-Fuzzy-Systeme. Es wird exemplarisch die Anwendbarkeit der Ansätze zur Modellierung von Veränderungen des Modal-Splits zugunsten öffentlicher Verkehrsmittel untersucht.

1. Einleitung

Mobilitätsbefragungen wie die KONTIV 1989 oder neuerdings das Haushaltspanel [ZUM98] bilden eine wichtige Datengrundlage zur Erstellung von Verflechtungsmatrizen des Personenverkehrs sowohl für den gegenwärtigen Zustand als auch darauf aufbauend für Prognosen des Personenverkehrs. Ein Bestandteil des den Prognosen zugrundeliegenden Vierstufenmodells sind Verkehrsmittelwahlmodelle, sogenannte Modal-Split-Modelle.

Im Rahmen dieser Arbeit wird die Anwendbarkeit von herkömmlichen Modal-Split-Modellen wie Logit-Modellen (als ein Beispiel für ein diskretes Entscheidungsmodell) verglichen mit Ansätzen, die in der Künstlichen Intelligenz (KI) entwickelt wurden und die bislang kaum auf verkehrswissenschaftliche Fragestellungen angewandt wurden. Dies sind neuronale Netzwerke, Fuzzy-Systeme sowie eine Kombination beider Ansätze, sogenannte Neuro-Fuzzy-Systeme.

Es wird geklärt, wie die Methoden der Künstlichen Intelligenz, deren Anwendungsbereich bislang fast ausschließlich auf technische Fragestellungen beschränkt ist, sich auf sozialwissenschaftliche Aufgaben anwenden lassen. Das Hauptunterscheidungsmerkmal von sozialwissenschaftlichen gegenüber technischen Aufgaben besteht in der Einbeziehung von irrationalen Elementen, die notwendigerweise mit menschlichen Entscheidungsprozessen verknüpft sind. Schwerpunkte der vergleichenden Bewertung sind die Abbildung nichtlinearer Effekte wie Verstärkungen, Synergien und Sättigungseffekte.

Die zugrundeliegende Datenbasis [BRÜ95], die 485 Personen umfaßt, ist Teil der „Münchener Umweltstudie 1993“, die am Institut für Soziologie der Ludwig-Maximilians-Universität

Anschrift des Verfassers:
Dr. Joachim Hugo
Rheinsberger Straße 48
10435 Berlin

München erhoben wurden. Erfragt wurde u.a. der Verkehrszweck zwischen Wohnort und Arbeitsplatz in Form von Haushaltsbefragungen.

2. Logit-Modelle, neuronale Netzwerke, Fuzzy-Systeme

2.1 Logit-Modelle

Die Aufgabe von Entscheidungsmodellen besteht darin, das Entscheidungsverhalten von Individuen (Agenten) unter Berücksichtigung der vorhandenen Alternativen zu beschreiben. In Abhängigkeit der Mischbarkeit der Alternativen lassen sich zwei extreme Ansätze unterscheiden: stetige und diskrete Entscheidungsmodelle. Während stetige Entscheidungsmodelle benutzt werden, wenn die Alternativen in beliebigen Mischungsverhältnissen gewählt werden können, liegt der Anwendungsbereich diskreter Entscheidungsmodelle bei Entscheidungssituationen, wo unter einer endlichen Anzahl klar unterscheidbarer Möglichkeiten gewählt werden kann und die Alternativen nicht in beliebigen Anteilen gemischt werden können [MAI90].

Eine entscheidende Rolle bei der Entwicklung diskreter Entscheidungsmodelle spielten Untersuchungen der Verkehrsökonomie und der Verkehrsplanung. Ein Schwerpunkt waren Modal-Split-Modelle [LAV69, BEN77]. Hierauf aufbauend wurden diskrete Entscheidungsmodelle angewandt für Routenwahl- und Zielwahlmodelle [DOM75] sowie für allgemeinere Untersuchungen in der Stadtplanung. Dies bewirkte die Kombination mehrerer Modellierungsebenen der Verkehrsmittel- und der Routenwahl, teilweise unter Umständen unter Hinzunahme weiterer Komponenten [BEN77]. Dies können die von der Stadt- und Regionalökonomie untersuchten Fragen der Standortentscheidung sein (also der Entscheidung von Personen und Haushalten über ihren Wohnstandort bzw. die Wohnungsnachfrage im allgemeinen) [AUF86].

Zur Bewertung der diskreten Alternativen wird für jede Alternative eine Nutzenfunktion $U(a_{in})$ definiert. Für sie gilt:

$$U(a_{in}) = U(\mathbf{C}_{in}, \mathbf{S}_n)$$

Hierbei werden die verschiedenen diskreten Alternativen a_{in} (zusammengefaßt durch die Alternativenmenge $A_n = \{a_{1n}, \dots, a_{In}\}$) durch einen Vektor \mathbf{C}_{in} , der die Charakteristika der Alternativen i , und durch einen Vektor \mathbf{S}_n , der die sozioökonomischen Charakteristika des Entscheidungsträgers n beschreibt, dargestellt.

Kennt das Individuum seine Nutzenfunktion und die Charakteristika aller Alternativen, so wird es diejenige Alternative wählen, die den höchsten Nutzen verspricht.

$$a_{opt,n} = \{a_{in} | U(a_{in}) \geq U(a_{jn}), \quad j = 1, \dots, m, a_{in} \in A_n\}$$

Während im stetigen Fall die Entscheidung durch den Marginalkalkül bestimmt wird (die optimale Alternative ist dann erreicht, wenn durch eine kleine Änderung in der Zusammensetzung der Alternative keine verfügbare Alternative mit höherem Nutzen erreicht werden kann), kann bei diskreten Entscheidungsmodellen die Reaktion auf marginale Änderungen sehr unterschiedlich sein. Sie reicht von der Beibehaltung der bisherigen Nachfragestruktur bis zu einer sprunghaften Entscheidungsänderung. Einen Ausweg aus dem Dilemma bietet die Zufallsnutzentheorie („random utility theory“) [MAN73, MAN77].

Sie geht davon aus, daß die oben beschriebene Optimierung zwar vom Entscheider durchgeführt wird, vom beobachteten Analytiker jedoch nicht in allen Einzelheiten dargestellt werden kann. Für ihn stellt der Nutzen daher eine Zufallsvariable dar. Der unvollständige Einblick kann folgende Ursachen haben:

- der Analytiker kennt nicht alle relevanten Charakteristika der Alternativen
- er kennt nicht alle relevanten sozioökonomischen Charakteristika des Individuums
- er kann die Werte der Charakteristika, die er als relevant erkennt, nicht exakt messen
- manche der relevanten Charakteristika sind nicht direkt, sondern nur indirekt über Instrumentvariable zu messen

Um eine formal und numerisch praktikable Untersuchung der Zufallsnutzentheorie zu ermöglichen, separiert man die Nutzenfunktion in einen deterministischen Anteil V und in einen stochastischen Anteil $\tilde{\epsilon}_{in}$:

$$\tilde{U}_{in} = V(\mathbf{C}_{in}^*, \mathbf{S}_n^*) + \tilde{\epsilon}_{in}$$

Die Wahrscheinlichkeit dafür, daß ein Individuum n die Alternative i wählt, ist danach gleich der Wahrscheinlichkeit dafür, daß die Alternative i den höchsten (Zufalls-) Nutzen aufweist:

$$P(a_{in} = a_{opt,n}) = \text{Prob}(U_{in} \geq U_{jn}, \quad j = 1, \dots, m, a_{in} \in A_n) \text{ bzw.}$$

$$\begin{aligned} P_{in} &= \text{Prob}(V_{in} + \epsilon_{in} - V_{jn} \geq \epsilon_{jn}, \quad j = 1, \dots, I_n) \\ &= \text{Prob}(V_{in} - V_{jn} \geq \epsilon_{jn} - \epsilon_{in}, \quad j = 1, \dots, I_n) \end{aligned}$$

Im allgemeinen wird als deterministische Nutzenfunktion V eine in den Parametern β_k lineare Funktion verwendet:

$$V_{in} = \sum_k \beta_k X_{kin} \quad \text{mit} \quad X_{kin} = g_k(\mathbf{C}_{in}, \mathbf{S}_n)$$

Die Funktionen g dürfen keine zu schätzenden Parameter enthalten. Der lineare Ansatz ist beliebt, da für die zu schätzenden Parameter gilt:

$$\partial V_{in} / \partial \beta_k = X_{kin}$$

Diskrete Entscheidungsmodelle werden danach eingeteilt, welche Verteilungsfunktion die stochastische Nutzenkomponente $\tilde{\epsilon}_{in}$ besitzt.

Ist die stochastische Nutzenkomponente $\tilde{\epsilon}_{in}$ unabhängig identisch Gumbel-verteilt, so erhält man das Logit-Modell [JOH70, DOM75, BEN85]. Es ist das am häufigsten angewandte multinomiale (d.h. mehr als zwei Entscheidungsalternativen) diskrete Entscheidungsmodell, dem auch die Berechnungen dieser Arbeit zugrunde liegen. Die Gumbel-Verteilung ist durch folgende Verteilungsfunktion charakterisiert:

$$\text{Prob}(X \leq c) = F(c) = \exp[-e^{-\mu(c-\eta)}] \quad \mu > 0$$

Damit gilt für die Dichtefunktion

$$f(c) = \mu e^{-\mu(c-\eta)} \exp[-e^{-\mu(c-\eta)}] \quad \mu > 0$$

Aus der Verteilung der Gumbel-Verteilung ergibt sich für die Auswahlwahrscheinlichkeit:

$$P_{in} = \frac{e^{\mu V_{in}}}{\sum_{j \in A_n} e^{\mu V_{jn}}}$$

Das Maximum mehrerer Gumbel-verteilter Zufallsvariablen mit identischem Streuungsparameter ist wiederum Gumbel-verteilt. Bei der Maximierung des Nutzens ist dies eine wichtige Eigenschaft. Die Verteilung ist rechtsschief. Die Ursache hierfür liegt darin, daß sie die Verteilung des Maximums von Zufallsvariablen darstellt und bereits ein hoher Wert für einen hohen Wert des Maximums ausreicht.

Eine besondere Eigenschaft des Logit-Modells ist dessen Unabhängigkeit des Verhältnisses der Auswahlwahrscheinlichkeiten zweier Alternativen von der Verfügbarkeit und den Charakteristika anderer (ähnlicher) Alternativen. Es gilt:

$$\frac{P_{in|A}}{P_{jn|A}} = \frac{P_{in|\tilde{A}}}{P_{jn|\tilde{A}}} \quad \text{mit } A \subseteq \tilde{A}, \text{ da}$$

$$\frac{P_{in}}{P_{jn}} = \frac{\exp(\mu V_{in})}{\sum_{k \in A_n} \exp(\mu V_{kn})} / \frac{\exp(\mu V_{jn})}{\sum_{k \in A_n} \exp(\mu V_{kn})} = e^{\mu(V_{in}-V_{jn})} = \text{konst}$$

Die Eigenschaft ist gleichzeitig Stärke und Schwäche des Logit-Modells. Es ist zwar leicht möglich, die Auswirkung von Veränderungen der Anzahl der Alternativen ohne besondere Schwierigkeiten zu simulieren.

Problematisch ist jedoch die Einführung einer weiteren, sehr ähnlichen Alternative („red bus-blue bus-Paradoxon“). Als Beispiel diene folgende Situation: es stehen bei einer Verkehrsmittelwahlentscheidung ein Auto und ein (roter) Bus zur Auswahl. Der Einfachheit halber gelte: $P(\text{Auto}) = P(\text{Bus}) = 0,50$. Bei einer weiteren Alternative eines (blauen) Busses wird erwartet, daß die zusätzliche Alternative zu Lasten des (roten) Busses geht und man $P(\text{Auto}) = 0,50$ sowie $P(\text{blauer Bus}) = P(\text{roter Bus}) = 0,25$ erhält. Nach dem Logit-Modell hingegen gilt: $P(\text{blauer Bus}) = P(\text{roter Bus}) = P(\text{Auto}) = 0,33$.

Ursache des Problems ist die Annahme unabhängig verteilter Zufallsterme in der Nutzenfunktion der Wirtschaftssubjekte. Sie verbietet es, Ähnlichkeiten zwischen den Alternativen zu modellieren, wie dies mit einer expliziten Kovarianzstruktur der einzelnen Zufallsterme möglich wäre. Eine Möglichkeiten zur Überwindung der Schwächen des Logit-Modells wäre das Probit-Modell, dessen numerische Behandlung jedoch wesentlich aufwendiger ist.

2.2 Neuronale Netzwerke

Neuronale Netzwerke sind Systeme, die einige der im menschlichen Gehirn erkannten oder vermuteten Organisationsprinzipien ausnutzen [ECC90, ECC94]. Sie bestehen aus einer Vielzahl einfacher Prozessoren, den Neuronen, die über gewichtete Verbindungen miteinander Nachrichten austauschen.

Neuronale Netzwerke stehen für ein neues Paradigma in der Künstlichen Intelligenz (KI). Bisherige Systeme bildeten Wissen regelbasiert auf der Grundlage herkömmlicher Computer ab. Diese Art der Wissensrepräsentation wird als symbolisch bezeichnet. Im Gegensatz dazu wird Wissen in neuronalen Netzwerken über das ganze Netzwerk in Gewichten verteilt (subsymbolisch) gespeichert. Die Speicherung basiert auf der Modellierung assoziativer und intuitiver Vorgänge mit Hilfe selbstorganisierender Prozesse. Wissen muß in weiten Teilen nicht mehr formalisiert werden, sondern kann sich durch den Lernprozeß selbst heranbilden. Formale Neuronen können als ein Addierer mit Schwellwert betrachtet werden, mit dem sich prinzipiell jede beliebige logische Funktion aufbauen läßt. Im einfachsten Fall berechnet das Modellneuron die resultierende Aktivität z formal als Skalarprodukt:

$$z(\mathbf{w}, \mathbf{x}) = \sum_j w_j x_j = \mathbf{w}^T \mathbf{x} - s$$

$$\text{mit } \mathbf{x} \rightarrow \mathbf{x} = (x_1, \dots, x_n)^T \text{ und } \mathbf{w} \rightarrow \mathbf{w} = (w_1, \dots, w_n)^T.$$

Die Aktivität y am Neuronenausgang wird bestimmt durch die Ausgabefunktion („activation function“) $y = S(z)$. Eine typische Funktionen $S(z)$ ist die sigmoide Ausgabefunktion (Fermi-

$$\text{Funktion) } S(z) = \frac{1}{1 + \exp(-kz)}.$$

Neuronen können zu einem neuronalen Netzwerk verschaltet werden. Sie verarbeiten ihre Eingaben vollständig parallel und unabhängig voneinander. Ein Hauptunterscheidungsmerkmal hinsichtlich ihrer Architektur sind Netze mit und ohne Rückkopplungen. Bei einem Netz ohne Rückkopplungen (einem feed-forward Netz), wie sie hier benutzt werden, unterscheidet man zwischen verschiedenen Schichten des Netzes. Verbindungen existieren nur zwischen zwei aufeinanderfolgenden Schichten. Die Verbindungen sind gerichtet und laufen von einer unteren zu einer darüberliegenden Schicht. Derartige Netzwerke werden als Bottom-Up-Modelle bezeichnet.

Lernalgorithmen ermöglichen die adaptive Veränderung der Gewichtematrix W während der Lernphase. Sie stellen die interessanteste Komponente neuronaler Netze dar. Das Ziel des Lernvorgangs ist es, W so zu bestimmen, daß es auf eine bestimmte Eingabe mit einer gewünschten Ausgabe reagiert und auf diese Weise auch auf neue (unbekannte) Eingaben mit einer geeigneten Reaktion antwortet. Diese Fähigkeit wird als Generalisierung bezeichnet. Dazu muß das Netz während der Lernphase die signifikanten Merkmale der Eingabemuster extrahieren. Sie gestattet es neuronalen Netzwerken, auf gestörte, verrauschte oder unvollständige Muster in etwa so zu reagieren, als läge eine korrekte Eingabe vor. Dies gelingt jedoch nur, wenn sich das unbekannte Muster einem der gelernten Klassen zuordnen läßt. Zur Bestimmung von W wird die erhaltene mit der erwünschten Ausgabe verglichen. Dies erfordert die Einführung eines Ähnlichkeits- oder Fehlermaßes, das von der Netzwerkarbeit und vom Lernvorgang abhängt.

Die hier benutzten Backpropagation-Netzwerke besitzen die in Abbildung 1 dargestellte Architektur. Im konkreten Fall werden dem Netzwerk über die Neuronen der Eingangsschicht die Antworten der befragten Personen zu den Fragen des im Anhang dargestellten Fragebogens zugespielt. Die Neuronen der Ausgangsschicht beinhalten die verschiedenen wählbaren Verkehrsmittel. In der Trainingsphase werden dem Netzwerk sämtliche Daten einschließlich der Verkehrsmittelwahl präsentiert. In der nachfolgenden Testphase wird überprüft, inwieweit das Netzwerk in der Lage ist, unbekannte Datensätze einem Verkehrsmittel zuzuordnen.

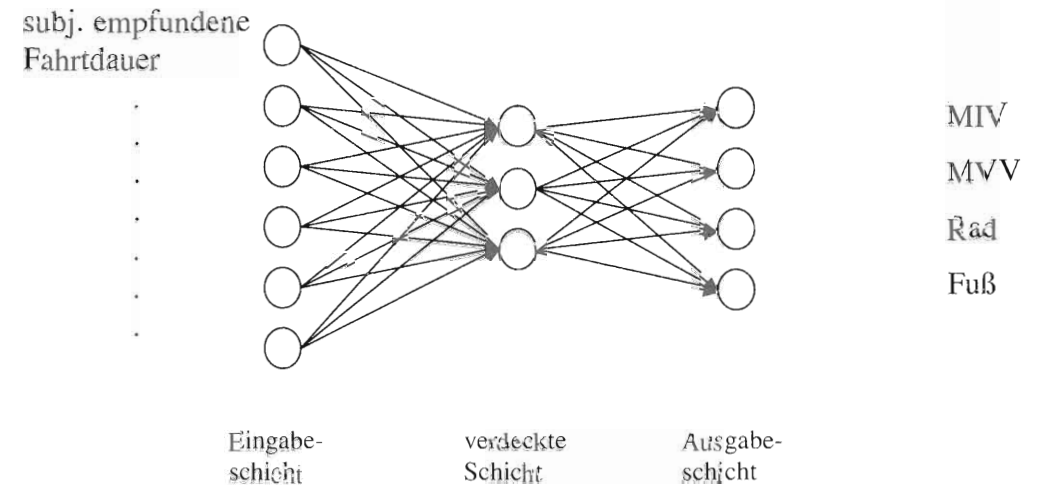


Abbildung 1: Schema des Backpropagation-Netzwerks, sämtliche Eingangsparameter des Netzwerks sind im Anhang aufgelistet

Der zu dem Netzwerk gehörende Lernalgorithmus wird als Backpropagation-Algorithmus bezeichnet [WER74]. Man versteht darunter die rückwärtige Ausbreitung eines Fehlersignals durch das Netzwerk. Der Grundgedanke der Regel besteht darin, nach der Ausbreitung des Musters durch das Netzwerk das erhaltene Ausgabemuster mit der erwünschten Ausgabe zu vergleichen. Das ermittelte Fehlersignal breitet sich rückwärts durch das Netz aus. Dies erlaubt nicht nur die Berechnung der Fehler jeder Ausgabeeinheit (und damit die Ermittlung der Fehler der Gewichte, die zu den Ausgabeeinheiten führen), sondern auch die Ermittlung der Fehler der zu den inneren Neuronen führenden Gewichte. Auf der Ermittlung der Fehlerbestimmung der inneren Neuronen beruht die besondere Leistung des Lernverfahrens.

2.3 Fuzzy-Systeme

Fuzzy-Systeme als ein Beispiel für Expertensysteme bestehen aus einer problemunabhängigen Komponente (dem Inferenzmechanismus) und einer problemabhängigen Komponente (der Wissensbasis). Häufig sind in beiden Komponenten große Mengen an Wissen zu integrieren. Eine zu große Menge an Wissen ist jedoch kritisch für den Erfolg eines Systems. Eine Möglichkeit der Bewältigung des Wissens besteht in einer Zunahme der Systemorganisation. Eine weitere Alternative ist die bewußte Einbeziehung der Reduktion des Präzisionsgrads der einfließenden Information unter Benutzung von Vagheits- und Unsicherheitsaspekten. Dies geschieht in Systemen mit Fuzzy-Logik [ZAD65, ZAD73]. Analog den herkömmlichen Expertensystemen werden auch in Fuzzy-Systemen Methoden der Wissensrepräsentation und -propagation benötigt.

Eine weitere Motivation von Fuzzy-Systemen besteht darin, daß komplizierte Steuerungsaufgaben häufig ohne die Kenntnis eines physikalisch-mathematischen Modells lösbar sind. So sind menschliche Bediener in der Lage, Fahrrad zu fahren, ohne die genauen mechanischen Gesetze zu kennen. Es bietet sich daher an, den Prozeß so zu modellieren und zu simulieren, wie ihn ein menschlicher Experte beschreibt. Dies wird als kognitive Analyse bezeichnet. Die Formulierung der Regeln geschieht durch linguistische Variablen („groß“, „schnell“).

Fuzzy-Systeme wurden ursprünglich im Bereich regelungstechnischer Anwendungen als Fuzzy-Control angewandt. Die Lösung einer regelungstechnischen Aufgabe besteht in der Angabe einer Kontrollfunktion $\varphi: X_1 \times \dots \times X_n \rightarrow Y$, die zu jedem Tupel von Meßwerten $(x_1, \dots, x_n) \in X_1 \times \dots \times X_n$ einen geeigneten Stellwert $y = \varphi(x_1, \dots, x_n)$ bestimmt. Ein Fuzzy-Regler muß demnach eine Interpolationsaufgabe lösen. Er ist dazu in der Lage, da er prinzipiell jede stetig differenzierbare Funktion approximieren kann. Fuzzy-Systeme können daher auch wie im Rahmen dieser Arbeit im Bereich der Mustererkennung eingesetzt werden.

Im Gegensatz zur klassischen Mengenlehre, wo Elemente entweder zu einer Menge gehören oder nicht, weisen die meisten Begriffe der Umgangssprache (linguistische Variable) keine derartigen Sprungstellen auf, sondern eher gleitende Übergänge zwischen einzelnen Extrema („jung-alt“). Mathematisch lassen sich gleitende Übergänge durch die Einführung gradueller Zugehörigkeitsgrade aus dem kompakten Intervall $[0,1]$ beschreiben. Derartige verallgemeinerte charakteristische Funktionen sind in der Lage, nicht nur scharfe klassische Mengen, sondern auch unscharfe Mengen (Fuzzy-Mengen) zur Darstellung linguistischer Variablen zu beschreiben. Eine Fuzzy-Menge μ läßt sich durch folgende Funktion beschreiben:

$$\mu: X \rightarrow [0,1]$$

Wissenspropagation in Fuzzy-Systemen geschieht durch die Inferenz (das Zusammenwirken) mehrerer Regeln (Implikationen), die zu einer Regelbasis zusammengefaßt werden. Die Regeln R_i besitzen folgende Form:

$$R_i: \text{if } X_1 \text{ is } \mu_{\alpha 1} \text{ and...and } X_n \text{ is } \mu_{\alpha n} \text{ then } Y \text{ is } \mu_{\beta}$$

Im Gegensatz zu herkömmlichen regelungstechnischen Methoden, bei denen der Prozeß modelliert wird, existiert die Möglichkeit, das Verhalten eines Menschen zu modellieren und zu simulieren, der diesen Prozeß regeln kann. Das Aufstellen eines Modells für das Verhalten eines „menschlichen Regelungsexperten“ wird als kognitive Analyse bezeichnet. Ein solches Modell setzt sich aus folgenden Komponenten zusammen:

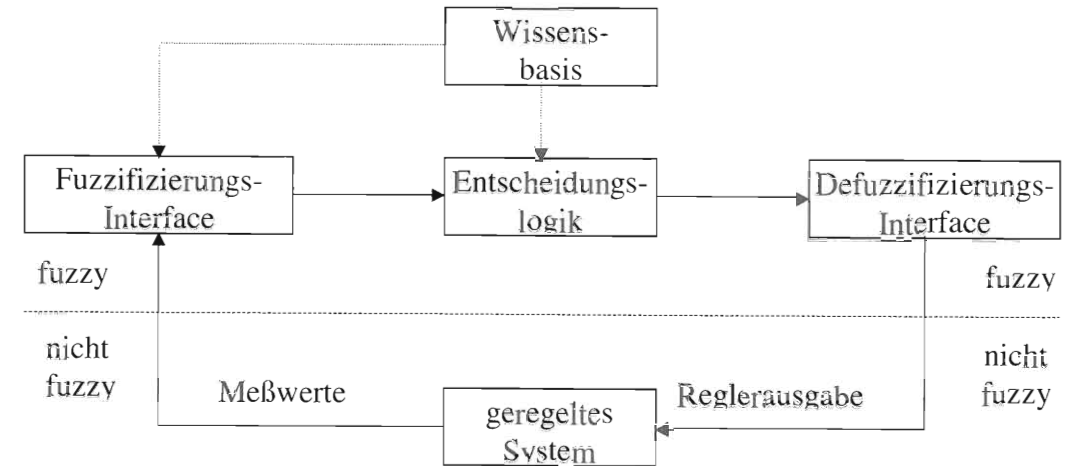


Abbildung 2: Architektur eines Fuzzy-Reglers

- das Fuzzifizierungs-Interface nimmt den aktuellen Meßwert auf und führt gegebenenfalls eine Transformation in einen geeigneten Wertebereich durch (u.a. in einen linguistischen Term oder in eine Fuzzy-Menge)
- die Wissensbasis setzt sich aus einer Daten- und aus einer Regelbasis zusammen. Die Datenbasis beinhaltet die Informationen über die Wertebereiche der Meß- und Stellgrößen, eventuelle Normierungen sowie die zu den Termen assoziierten Fuzzy-Mengen. Die Regelbasis beinhaltet die linguistischen Kontrollregeln
- die Entscheidungslogik stellt das Rechenwerk des Fuzzy-Reglers dar. Mit ihm werden aus den Meßgrößen mit Hilfe der Wissensbasis Informationen über die Stellgröße gewonnen
- das Defuzzifizierungs-Interface ermittelt aus der von der Entscheidungslogik gelieferten Information über die Stellgröße einen scharfen Stellwert

2.4 Neuro-Fuzzy-Systeme

Neuronale Netzwerke und Fuzzy-Systeme sind zwei Teilbereiche des Forschungsgebiets „Künstliche Intelligenz“ (KI). Sie lassen sich unter dem Begriff „Soft-Computing“ zusammenfassen. Man versteht darunter wissensbasierte Techniken, die in der Lage sind, mit unvollständigen, gestörten oder ungenauen Eingaben dennoch eine akzeptable Ausgabe zu liefern.

Beide Verfahren sind im Gegensatz zu herkömmlichen Systemen der KI in der Lage, ihnen unbekannte Daten aufzunehmen und damit neue Fakten zu definieren. Sie besitzen die Fähigkeit, approximative Folgerungen ziehen zu können, die unabhängig von den mathematischen Modellen der Datenanalyse sind. Werden allerdings neuronalen Netzwerken bzw. Fuzzy-Systemen Daten präsentiert, deren Sachverhalt während des Lernvorgangs nicht angemessen berücksichtigt wurde bzw. nicht durch die Regelbasis definiert wurde, so kann dies zu einem Versagen beider Systeme führen. Die Flexibilität beschränkt sich daher bei beiden Systemen auf die Bildung von Analogieschlüssen.

Die Fähigkeit zur Bildung von Analogieschlüssen beider Alternativen kommt bei der Modellierung von Veränderungspotentialen des Modal-Splits zum Tragen. Es existiert bislang keine Formel, die alle Einflußfaktoren des Modal-Splits explizit darstellt. Statistische Methoden bieten einen Ausweg; problematisch wird dieser Weg jedoch bei zu vielen Einflußfaktoren. Es besteht dann die Gefahr, daß die Beziehungen nur sehr vage geschätzt werden können und die Statistik wegen zu vieler ungenauer Variablen nichtssagend wird [SPI93].

Die Stärken neuronaler Netzwerke liegen vor allem in deren Lernfähigkeit, ohne im herkömmlichen Sinne programmiert sein zu müssen. Daneben erlaubt ihre verteilte Struktur, auch hochgradig parallele Implementierungen und Hardware-Realisierungen zuzulassen. Sie liefern akzeptable Lösungen, auch wenn ein Teil ihrer Eingaben unvollständig, gestört oder ungenau ist.

Jedoch ist die Problemlösung bei neuronalen Netzwerken implizit in den Gewichten kodiert und nur in den seltensten Fällen in Form von Regeln sinnvoll interpretierbar. Man spricht daher von einem Black-Box-Verhalten des Netzwerks. Es erfüllt zwar eine gewünschte Funktion, läßt aber nicht erkennen, wie es diese ermittelt. Daraus folgt, daß sich weder bereits vorhandenes Wissen zur Initialisierung in ein Netzwerk integrieren noch daß sich (Teil-) Wissen extrahieren läßt.

Vorteilhaft an Fuzzy-Systemen ist deren Fähigkeit, Wissensdarstellung und Wissenspropagation durch die Verwendung einer Regelbasis transparent darstellen zu können. Durch die Trennung in eine problemunabhängige Komponente (den Inferenzmechanismus) und in eine problemabhängige Komponente (die Wissensbasis) läßt sich jederzeit Wissen integrieren bzw. extrahieren.

Fuzzy-Logik kann die Daten jedoch nur dann analysieren, wenn die Wissens- und die Regelbasis vollständig und ohne Widersprüche sind. Ist dies nicht der Fall, ist eine Nachbearbeitung der Regeln und/oder der linguistischen Werte der Zugehörigkeitsfunktionen („Fine-Tuning“) unumgänglich. Hierfür muß heuristisch vorgegangen werden, da keine formalen Methoden existieren. Bei komplizierteren Fuzzy-Systemen ist eine solche Anpassung bzw. Optimierung der Zugehörigkeitsgrade nahezu unmöglich. Werden neben dem Hinzufügen, Entfernen oder Ändern einzelner Regeln der Regelbasis weitere Verfahren wie zum Beispiel

die Gewichtung von Regeln eingesetzt, so wird die Semantik des Fuzzy-Systems aufgegeben [ALT92, KOS92].

Neuronale Netzwerke	Fuzzy-Logik
Vorteile	
<ul style="list-style-type: none"> - kein mathematisches Prozeßmodell notwendig - kein Regelwissen notwendig - verschiedene Lernalgorithmen 	<ul style="list-style-type: none"> - kein mathematisches Prozeßmodell notwendig - a-priori (Regel-) Wissen nutzbar - einfache Interpretation und Implementation
Nachteile	
<ul style="list-style-type: none"> - Black-Box-Verhalten - kein Regelwissen extrahierbar - heuristische Wahl der Netzparameter - Anpassung an veränderte Parameter ist eventuell schwierig und kann eine Wiederholung des Lernvorgangs erfordern - kein a-priori Wissen verwendbar - der Lernvorgang konvergiert nicht garantiert - unter Umständen lange Trainingsdauer 	<ul style="list-style-type: none"> - Regelwissen muß verfügbar sein - nicht lernfähig - keine formalen Methoden für „tuning“ - semantische Probleme bei der Interpretation „getunter“ Regler - Anpassung an veränderte Parameter eventuell schwierig - ein „Tuning“-Versuch kann erfolglos bleiben

Tabelle 1: Gegenüberstellung der Vor- und Nachteile von neuronalen Netzwerken und Fuzzy-Systemen [NAU96]

Da die artspezifischen Vorteile der neuronalen Netze gerade den Nachteilen von Fuzzy-Systemen entsprechen und umgekehrt, erscheint es sinnvoll, beide Systeme miteinander zu kombinieren. Es kann die Lernfähigkeit neuronaler Netze dazu genutzt werden, die Parameter eines Fuzzy-Systems in einem Adaptionsprozeß an die gegebenen Anforderungen anzupassen. Derartige Systeme werden als neuronale Fuzzy-Systeme oder Neuro-Fuzzy-Systeme bezeichnet. Sie ermöglichen die Integration von Lernverfahren in Fuzzy-Systemen. Erlernen bedeutet hierbei die vollständige Erzeugung der Regelbasis bzw. der Zugehörigkeitsfunktionen.

Neuronale Netzwerke profitieren davon, da sich durch Fuzzy-Logik a-priori Wissen in das System integrieren läßt. Bereits bekannte Regeln und Zugehörigkeitsfunktionen können für eine Initialisierung des Systems genutzt werden, was eine erhebliche Verkürzung des Lernvorgangs bewirkt. Das Black-Box-Verhalten des neuronalen Netzwerks läßt sich vermeiden, da der Lernprozeß mit einer Anpassung der Zugehörigkeitsfunktionen bzw. der Regelbasis

endet, die weiterhin als Standard-Fuzzy-System interpretierbar sind [NAU92]. Auf diese Weise läßt sich dem System sogar neues Wissen entnehmen. Der Gewinn für Fuzzy-Systeme besteht darin, daß die Systeme lernfähig sind.

Die Architektur wird meist durch die Regeln und die unscharfen Mengen festgelegt [BER92, KRU92, NAU93]. Dadurch entfällt die heuristische Bestimmung von Netzparametern (wie der Anzahl der inneren Einheiten des neuronalen Netzwerks).

Jedoch läßt sich auch durch die Kombination beider Methoden nach wie vor kein Lernerfolg bzw. analog keine Verbesserung des Regelverhaltens durch Tuning garantieren.

3. Ergebnisse

Von der Befragtengruppe nutzten 41,0 % das Auto, 42,9 % den Münchner Verkehrsverbund (MVG), 9,3 % das Fahrrad als Verkehrsmittel, 6,8 % gingen zu Fuß zur Arbeit. Teilweise wurde eine Kombination verschiedener Verkehrsmittel genutzt. In diesem Fall wurde von den Befragten das anteilmäßig wichtigste Verkehrsmittel genannt. Abhängig von diversen Bestimmungsgrößen lassen sich zum Teil erhebliche Schwankungen der Verkehrsmittelwahl (des Modal-Split) feststellen. Derartige Parameter sind die Verfügbarkeit einer problemlosen Parkmöglichkeit am Arbeitsplatz, die Entfernung Wohnort-Arbeitsplatz, die Umsteigehäufigkeit mit öffentlichen Verkehrsmitteln, das monatliche Nettoeinkommen sowie die subjektive Einschätzung der Zeit- (einschließlich der Umsteigehäufigkeit), Bequemlichkeits- und Kostenparameter der einzelnen Verkehrsmittel.

3.1 Darstellung der Veränderungspotentiale des Modal-Splits mit Logit-Modellen

Die Verkehrsmittelwahl läßt sich als ein zweistufiger Entscheidungsprozeß auffassen, der u.a. durch ein Nested Logit-Modell [MAI90] beschrieben werden kann. Auf der ersten Stufe entscheidet man zwischen einem motorisierten und einem unmotorisierten Verkehrsmittel. Hat man sich für ein motorisiertes Verkehrsmittel entschieden, so hat man auf der zweiten Stufe die Wahl zwischen dem Auto und dem Münchner Verkehrsverbund (MVG). Die Modellierung ist jedoch recht kompliziert.

Da die erste Entscheidung die im wesentlichen eingeschlossene Benutzergruppe von Rad und Fuß betrifft und für verkehrspolitische Belange die erste Entscheidungsstufe weniger von Bedeutung ist, interessiert vor allem die Frage, wie und in welchem Umfang Änderungen des Verkehrsmittelwahlverhaltens zwischen Auto und MVG erhalten werden können. Beim Vorhandensein dieser Entscheidungsstruktur können die Parameter der zweiten Stufe auch ohne Beachtung der ersten Stufe unverzerrt geschätzt werden. Man benutzt als statistisches Analyseverfahren das Logit-Modell.

Zur Schätzung der Elastizitäten wird das Programmpaket LIMPED [GRE92] benutzt. Der deterministische Teil der Nutzenfunktion ist wie üblich linear, so daß die Elastizitäten Konstanten sind.

Als unabhängige Variablen werden die drei von jedem Individuum erhobenen subjektiven Kostenattribute sowie die Personenmerkmale (Geschlecht, Alter, Bildung, in Ausbildung befindlich, Einkommen, Umweltbewußtsein) berücksichtigt. Als zusätzliche Kontrollvariable dient die Frage, ob eine Person ein Auto besitzt.

Die Elastizitäten der Kostenattribute geben die Gewichtung an, mit denen die Einflußfaktoren in die Nutzenfunktion eingehen. Da ihre absolute Größen (wie generell üblich bei Elastizitäten) von der Skalierung abhängig sind, erlauben sie nur relative Aussagen über das Verhältnis zweier Größen. Der Bezugspunkt der ermittelten Effekte ist die Wahrscheinlichkeit der Wahl des Autos als Verkehrsmittel.

Variable	Modell 1	Modell 2
Zeitkosten		-0,50 (3,83)
Unbequemlichkeit		-0,75 (4,42)
Geldkosten		-0,42 (3,35)
Konstante	-2,72 (2,47)	-5,89 (3,25)
Frau	-0,22 (0,75)	-0,67 (1,45)
Alter in Jahren / 10	0,05 (0,38)	0,16 (0,82)
Mittlere Reife	0,25 (0,63)	0,66 (1,14)
(Fach-) Abitur	0,06 (0,16)	0,10 (0,19)
In Ausbildung	-1,54 (2,66)	-0,98 (1,18)
Einkommen in DM / 1000	0,09 (1,12)	0,08 (0,67)
Umweltbewußtsein	0,02 (0,36)	0,08 (1,11)
Person besitzt Auto	2,56 (5,45)	2,75 (4,05)
Pseudo R^2	19,7 %	57,8 %
Fallzahl	268	268

Tabelle 2: Konditionale Logit-Modelle der Verkehrsmittelwahl [BRÜ95]

Das Modell 1 enthält zur Erklärung des Verkehrsmittelwahlverhaltens nur Personenmerkmale. Es entspricht damit dem gewöhnlichen Logit-Regressionsmodell, wie es in sozialwissenschaftlichen Untersuchungen häufig benutzt wird. Allerdings ist der Erklärungsbeitrag der Personenmerkmale relativ gering, wie der Pseudo R^2 Wert von 19,7 % zeigt.

Werden hingegen wie in Modell 2 zusätzlich die subjektiven Kostenattribute (quantifiziert durch Schulnoten) aufgenommen, so steigt die Erklärungskraft des Modells erheblich. Die subjektiven Einstellungen zu den Verkehrsmitteln spielen demnach eine entscheidende Rolle.

Die negativen Vorzeichen der Kostenattribute zeigen, daß eine Kostenerhöhung den Nutzen einer Alternative senkt und damit auch deren Wahrscheinlichkeit. Die stärkste Abhängigkeit besteht beim Bequemlichkeitsparameter, die geringste beim Preis. Es kann daher erwartet werden, daß eine Senkung des Tarifs des MVV oder eine Erhöhung des Benzinpreises nur eine geringe Veränderung auf das Verkehrsmittelwahlverhalten ausüben werden. Um den Modal-Split zugunsten des MVV zu verschieben, haben Maßnahmen, die die Bequemlichkeit des MVV erhöhen, einen ungleich größeren Einfluß.

Von Interesse ist weiter, wie leicht oder schwer sich die getroffene Verkehrsmittelwahl verändern läßt. Das Zufallsnutzenmodell ermöglicht es, für jede Person den erwarteten Nutzen jeder Alternative zu errechnen. Je nach Höhe der Nutzendifferenz zwischen der Auto- und der MVV-Alternative erhält man Auskunft, wie festgefahren eine Entscheidung für ein Verkehrsmittel ist. Es zeigt sich, daß i.allg. die Nutzendifferenz der MVV-Nutzer näher am Nullpunkt liegt als die der Auto-Nutzer. Allerdings treten bei einem Teil der MVV-Nutzer sehr hohe Werte auf. Dies dürften Personen sein, die ein Auto nie benutzen (Personen unter 18 Jahre, Personen ohne Führerschein,...). Da die Mehrzahl der MVV-Nutzer dennoch näher am Nullpunkt liegt als die Auto-Nutzer, kann davon ausgegangen werden, daß bei einer Fahrpreiserhöhung eher eine Wanderung zum Auto hin stattfindet als umgekehrt bei einer Fahrpreiserhöhung eine Wanderung zum MVV.

3.2 Darstellung der Veränderungspotentiale des Modal-Splits mit neuronalen Netzwerken

Zur Modellierung des neuronalen Netzwerks wird das Programm NNU der Firma IBM benutzt [HUG98]. Wie üblich bei neuronalen Netzwerken, wird der insgesamt 485 Muster umfassende Datensatz in einen 400 Muster umfassenden Trainingsdatensatz und in einen 85 Muster umfassenden Testdatensatz aufgeteilt. Der Trainingsdatensatz klassifiziert 99 % der Muster richtig. Dies entspricht vier falsch zugeordneten Mustern. Beim Testdatensatz werden 62,4 % aller Muster richtig klassifiziert (dies entspricht 53 richtig klassifizierten Mustern). Die restlichen Muster wurden zu gleichen Anteilen entweder falsch klassifiziert oder sie konnten keinem der Verkehrsmittel zugeordnet werden.

Der Anteil von 62,4 % richtig klassifizierten Mustern erscheint sehr gering. Jedoch ist zu beachten, daß Verkehrsmittelwahlmodelle eine sozialwissenschaftliche Untersuchung sind und somit immer das irrationale menschliche Element berücksichtigt werden muß. Die meisten bisherigen Anwendungen neuronaler Netzwerke betrafen technische Systeme, die sich vollkommen rationell verhalten. Eine Folge davon ist, daß auch die Präsentation unbekannter Muster bestimmten nachvollziehbaren Situationen entspricht, die einen höheren Anteil richtig klassifizierter Testmuster bewirken.

Präsentiert man dem fertig trainierten Netz als erste Gütekontrolle den gesamten, 485 Muster umfassenden Datensatz, so wird durch das Netz folgender Modal-Split berechnet:

	Auto	MVV	Rad	Fuß
Ausgangsdatensatz	41,0 %	42,9 %	9,3 %	6,8 %
durch das trainierte Netz klassifiziert	42,8 %	43,0 %	7,6 %	6,6 %

Tabelle 3: erste Gütekontrolle des trainierten Netzes durch Präsentation des Ausgangsdatensatzes

Die Anteile der Verkehrsträger MVV und des zu Fußgehens werden nahezu identisch wiedergegeben (0,1 bzw. 0,2 % Abweichung). Hingegen liegt die Abweichung bei den Verkehrsträgern Auto bzw. Rad bei fast 2 %. Die Ursache hierfür sind die falsch bzw. gar nicht klassifizierten Muster des Testdatensatzes.

Zur Berechnung der Veränderungspotentiale des Modal-Splits werden dem fertig trainierten Netz, dessen Gewichtswerte sich nicht mehr ändern, modifizierte Formen des ursprünglichen (485 Muster umfassenden) Datensatzes präsentiert. Die Modifizierung kann sowohl die Veränderung einzelner Parameter (u.a. der subjektiv empfundenen Reisezeit des MVV) betreffen als auch das Zusammenwirken mehrerer Veränderungsmaßnahmen (Veränderung der subjektiv empfundenen Parameter Reisezeit und -bequemlichkeit des MVV). Je nach Stärke und Kombination der Abweichungen werden nichtlineare Effekte wie Verstärkungen, Sättigungen und Synergieeffekte erwartet.

Exemplarisch sei im folgenden (neben der Änderung eines Parameters, wie sie in der diesem Artikel zugrundeliegenden Dissertation durchgeführt wurde,) die Modifizierung des Modal-Splits bei gleichzeitiger Veränderung der Einschätzung von Fahrdauer und -bequemlichkeit des MVV sowie der Umsteigehäufigkeit dargestellt. Ein zusätzlicher bzw. eingesparter Umsteigevorgang entspricht einer Veränderung der Reisezeit um ± 5 min. Die Veränderungspotentiale werden derart kombiniert, daß eine um eine Note verbesserte Bewertung von Fahrdauer und -bequemlichkeit des MVV gemeinsam mit einem eingesparten Umsteigevorgang untersucht werden. Die Simulation wird unter der Bezeichnung „ein Punkt zugunsten MVV“ zusammengefaßt. Analog werden die übrigen Veränderungspotentiale zugunsten und zuungunsten des MVV dargestellt.

Es zeigt sich, daß mehrere Einflußfaktoren so kombiniert werden können, daß sie sich in ihrer Wirkung verstärken. Sie bewirken größere Verschiebungspotentiale als die separate Behandlung einzelner Einflußfaktoren. Bei einer Verbesserung um eine bzw. zwei Noten treten Verstärkungseffekte im Vergleich zur Simulation mit ausschließlich veränderter Fahrdauer und -bequemlichkeit auf. Allerdings treten wiederum Sättigungseffekte bei einer Veränderung um ± 3 Bewertungseinheiten auf. Dies deutet darauf hin, daß ein Anteil zwischen 65 und 70 % des MVV am Modal-Split des Berufsverkehrs ohne Beeinflussung des

Individualverkehrs einen gewissen Grenzwert darstellt. Vergleichbare Anteile des öffentlichen Verkehrs am Modal-Split werden derzeit in Zentren von Ballungsräumen beobachtet.

	Auto	MVV	Rad	Fuß
drei Punkte zugunsten MVV	15,5 %	69,8 %	9,7 %	5,0 %
zwei Punkte zugunsten MVV	22,2 %	64,6 %	7,8 %	5,3 %
ein Punkt zugunsten MVV	31,1 %	54,8 %	7,8 %	6,3 %
Ausgangsverteilung	42,8 %	43,0 %	7,6 %	6,6 %
ein Punkt zuungunsten MVV	54,0 %	31,3 %	5,7 %	9,0 %
zwei Punkte zuungunsten MVV	61,4 %	24,5 %	4,5 %	9,6 %
drei Punkte zuungunsten MVV	64,5 %	21,2 %	4,4 %	9,9 %

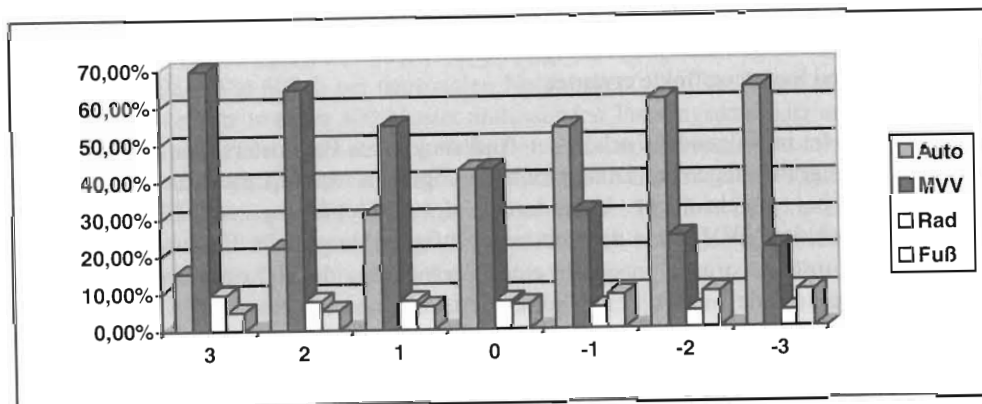


Abbildung 3: Abhängigkeit des Modal-Splits aus einer Kombination von subjektiv empfundener Fahrdauer und -bequemlichkeit sowie der Umsteigehäufigkeit

Ein Unsicherheitsfaktor neuronaler Netzwerke ist die Anzahl der nichtklassifizierbaren und die der falsch klassifizierten Muster. Der Anteil der falsch klassifizierten Muster des Trainings- und des Testdatensatzes beträgt zusammen 20 Muster. Dies entspricht 4,1 % aller Muster. Es kann davon ausgegangen werden, daß ihre Anzahl auch bei veränderten Eingangsmustern konstant bleibt, da sie auf den oben genannten irrationalen menschlichen

Verhaltensweisen beruhen. Sie stellen damit einen unvermeidlichen, aber konstanten Fehler dar. Da die Ausgangsmuster trotzdem weitgehend befriedigend klassifiziert wurden, wird diese mögliche Fehlerquelle als weitgehend vernachlässigbar angesehen und daher nicht mehr weiter untersucht.

Etwas anders ist die Lage bei den nicht klassifizierbaren Mustern. Während das Netz bei der Präsentation des ursprünglichen Datensatzes 25 Muster nicht klassifizieren kann (dies entspricht 5,1 % aller Muster), steigt der Anteil je nach Stärke der Veränderung der Präsentationsmuster auf bis zu 94 nichtklassifizierten Mustern an (dies entspricht 19,4 % aller Muster). Es kann hier also nicht angenommen werden, daß die nicht klassifizierbaren Muster eine konstante Fehlerquelle darstellen. Wengleich nicht davon ausgegangen werden kann, daß alle nichtklassifizierbaren Muster nur von einer Verkehrsart stammen, so relativiert sich die Aussagekraft der Simulationsergebnisse eher auf Abschätzungen denn auf Vorhersagen im Dezimalbereich. Diese Schwierigkeit wird allerdings dadurch relativiert, daß bei gleichzeitiger Änderung mehrerer Parameter um mehrere Größenordnungen derart große Veränderungen des Modal-Splits prognostiziert werden, wie sie derzeit ohnehin nicht denkbar sind.

Es hat sich darüber hinaus gezeigt, daß das neuronale Netzwerk nicht eines externen (menschlichen) Experten entbehrt, der überprüft, ob alle als wichtig vermuteten Einflußparameter in etwa gleichberechtigt vertreten sind. Ansonsten besteht die Gefahr, daß das Netzwerk einen Einfluß fälschlicherweise als unbedeutend erkennt.

Gegenüber Logit-Modellen, die i.allg. lineare Regressionsverfahren benutzen, erlauben neuronale Netzwerke deutlich weitergehende Analysen. Der Hauptvorteil neuronaler Netzwerke besteht darin, daß in Abhängigkeit der Stärke von Veränderungsszenarien Nichtlinearitäten dargestellt werden können. Es besteht die Möglichkeit, die Stärke der Veränderung des Modal-Splits in Abhängigkeit auch mehrerer Einflußfaktoren gleichzeitig nahezu stetig darzustellen. Statistische Modelle hingegen erlauben unter Verwendung der linearen Regression nur die Darstellung linearer Abhängigkeiten eines einzelnen Einflußfaktors. Es besteht die Möglichkeit, mittels nichtlinearer Regression gleichzeitig mehrere Parameter zu verändern. Im Vergleich zu neuronalen Netzwerken ist dieses Verfahren jedoch wesentlich umständlicher und es liefert Resultate, die weniger signifikant sind.

Neuronale Netzwerke erlauben i.allg. zuverlässige Vorhersagen, obwohl die Genauigkeitsanforderungen wesentlich höher sind als die an statistische Modelle. Während bei statistischen Modellen die Alternative als gewählt angenommen wird, deren Attraktivität den höchsten Wert aufweist, unabhängig von deren Betrag, wird bei neuronalen Netzwerken unter Verwendung eines Klassifikationsobjekts nur die Alternative als gewählt angesehen, deren Attraktivitätswert über 0,9 liegt (mögliche Attraktivitätswerte liegen zwischen 0 und 1).

Problematisch ist jedoch nach wie vor der mathematische Kontext neuronaler Netzwerke, der sich bislang durch deren Black-Box-Verhalten noch nicht allgemeingültig darstellen ließ. Besondere Anstrengungen wurden in jüngster Zeit unternommen, um zu klären, inwieweit neuronale Netze mit bekannten Methoden aus der Statistik korrespondieren und wie sich beide Verfahren ergänzen können [CHE93].

Eine Möglichkeit der Kombination besteht darin, mehrschichtige Backpropagation-Netzwerke als nichtlineare, nichtparametrische Regressionsverfahren aufzufassen [WHI91, GEM92]. Neuronale Netzwerke sind zwar nominell parametrische Modelle. Aufgrund ihrer Lernfähigkeit können sie jedoch tatsächlich als nichtparametrische Modelle angesehen werden. Um die Transparenz zu erhöhen, bietet es sich in diesem Zusammenhang an, vor der eigentlichen Informationsverarbeitung durch das Netzwerk eine lineare Regression des Datensatzes durchzuführen. Dem Netz bliebe nur noch die nichtlineare Regression.

3.3 Darstellung der Veränderungspotentiale des Modal-Splits mit Fuzzy-Systemen und Neuro-Fuzzy-Systemen

Fuzzy-Systeme lassen sich dann benutzen, wenn sich die Aufgabe durch wenige Regeln darstellen läßt und die Partitionierung der unscharfen Mengen eindeutig ist. In der konkreten Anwendung können die Partitionierungen der unscharfen Mengen jedoch nicht definiert werden (wann verdient der öffentliche Verkehr die Schulnote „gut“)? Aus diesem Grund können Fuzzy-Systeme nicht dazu herangezogen werden, Veränderungspotentiale des Modal-Splits zu modellieren.

Neuro-Fuzzy-Systeme können sowohl eine Regelbasis automatisch erzeugen als auch die Partitionierungen von Fuzzy-Mengen optimieren.

Bei der Präsentation des Datensatzes ermittelt das Programm NEFCLASS-PC (eine Beschreibung des Programms findet sich u.a. in [NAU96]) 125 Regeln zur Beschreibung des Verkehrsmittelwahlverhaltens. Aus Gründen der Übersichtlichkeit wird die Anzahl der Regeln auf 20 beschränkt, wobei diejenigen ausgewählt werden, die die meisten Entscheidungen erklären. Nach Beendigung des Trainings wurden 192 Muster falsch klassifiziert. Dies entspricht einem Anteil an richtig klassifizierten Mustern von 59,1 %. Dies entspricht etwa der Klassifikationsleistung des Testdatensatzes des neuronalen Netzwerks; damit besitzt das Neuro-Fuzzy-System eine deutlich schlechtere Klassifikationsleistung gegenüber dem neuronalen Netzwerk.

Ursächlich hierfür ist die Tatsache, daß Neuro-Fuzzy-Systeme (und damit auch Fuzzy-Systeme) dann eingesetzt werden können, wenn die Regeln möglichst separate Klassen (Cluster) darstellen. Im konkreten Fall hingegen findet eine vielfache unscharfe Überlagerung der einzelnen Klassen statt, da im Extremfall jede Person ihre eigene Regelbasis besitzt. Diese kann für sich genommen schon irrational oder widersprüchlich sein. Darüber

hinaus können die Regelbasen der verschiedenen Personen irrational oder widersprüchlich sein.

Es bleibt jedoch die Möglichkeit, durch die ermittelte Regelbasis die durch die neuronalen Netzwerke erhaltenen Ergebnisse zu veranschaulichen. So wurden trotz der deutlich schlechteren Klassifikationsleistung des Neuro-Fuzzy-Systems plausible Regelkonklusionen gebildet.

3.4 Vergleichende Bewertung der Simulationswerkzeuge

In den vorangegangenen Abschnitten konnte gezeigt werden, daß die Simulationsverfahren diskrete Entscheidungsmodelle, neuronale Netzwerke und Neuro-Fuzzy-Systeme unterschiedliche Systemeigenschaften besitzen und sich somit unterschiedliche Anwendungsschwerpunkte anbieten.

Fuzzy-Systeme und damit auch Neuro-Fuzzy-Systeme können dann benutzt werden, wenn sich die Aufgabe durch die Angabe weniger Regeln darstellen läßt. Um eine möglichst hohe Anzahl korrekt präsentierter Muster zu gewährleisten, sollten sich die einzelnen Regeln dabei möglichst wenig überlappen. Es kann daher davon ausgegangen werden, daß sich Neuro-Fuzzy-Systeme nicht zur Modellierung von sozialwissenschaftlichen Fragestellungen eignen, sondern i. allg. zur Modellierung (nichtlinearer) technischer Systeme.

Neuronale Netzwerke können auch bei hochdimensional nichtlinearen sozialwissenschaftlichen Fragestellungen eingesetzt werden. Da sozialwissenschaftliche Fragestellungen notwendigerweise irrationale Bewertungselemente enthalten, erreichen neuronale Netzwerke zwar nicht die hohe Generalisierungsfähigkeit (bewertet durch die Anzahl richtig klassifizierter Testmuster) wie bei technischen Anwendungen. Sofern die Art der Fragestellung eines Datensatzes jedoch alle möglichen Fälle ausreichend berücksichtigt, erfüllen neuronale Netzwerke die an sie gestellten Erwartungen als Analyse- und Prognoseinstrument. Insofern wiegt der Nachteil der mangelnden Transparenz weniger schwer, da die Regelbasis eines Neuro-Fuzzy-Systems, sofern sie überhaupt befriedigende Ergebnisse liefern würde, im Falle einer sozialwissenschaftlichen Fragestellung aufgrund der zu erwartenden hohen Regelanzahl unübersichtlich werden würde.

Neuronale Netzwerke erfüllen demnach die an sie gerichteten Erwartungen als Analyse- und Prognoseinstrument hochdimensionaler nichtlinearer Systeme. Sie sind in der Lage, Analogieschlüsse zu ziehen und nichtlineare Effekte wie Verstärkungen, Sättigungen und Synergien abzubilden. Im konkreten Fall eines Datensatzes mit 27 Eingabevariablen und vier Ausgabevariablen erfüllten sie die Erwartungen ohne zusätzliche Unterstützung. Häufig sind die Dimensionen der zu untersuchenden Systeme jedoch wesentlich höher. Damit geht ein starker Anstieg der unbestimmten Parameter (v.a. der Gewichte) einher. Die Anzahl der Präsentationsmuster steigt jedoch nicht dem Umfang. Zum Einsatz kommen in solchen Fällen Pruning-Verfahren [KUD96].

Simulationswerkzeug	Anwendungsbereich
Fuzzy-Systeme / Neuro-Fuzzy-Systeme	(nichtlineare) technisch-naturwissenschaftliche Systeme, deren Verhalten durch wenige Regeln beschreibbar ist
Neuronale Netze	(nichtlineare) technische und sozialwissenschaftliche Systeme
Neuronale Netze + Pruning-Verfahren	(nichtlineare) technische und sozialwissenschaftliche Systeme, deren Dimension jedoch so hoch ist, daß Pruning-Verfahren den Lernprozeß unterstützen

Tabelle 4: Vergleich der Einsatzgebiete von (Neuro-) Fuzzy-Systemen und von neuronalen Netzen

3.5 Weitere Fragestellungen

Künftige Aufgabenstellungen von neuronalen Netzwerken im Bereich der Verkehrswissenschaft sind vor allem in der Integration weiterer Einflußfaktoren zur Berechnung des Modal-Splits zu sehen. Dies ist zum einen die Einbeziehung stadt- und regionalplanerischer Überlegungen, zum anderen die Kombination von Verkehrsmittelwahlmodellen mit Routen- und Zielwahlmodellen. Schließlich kann die Dimensionierung der räumlichen Ebene geändert werden. Anstatt Fragen zur Stadt- oder zur Regionalplanung zu analysieren, kann der Untersuchungsgegenstand die gesamtstaatliche Planung sein. Wegen der deutlich erhöhten Dimension solcher Netze wird angenommen, daß Pruning-Verfahren unterstützend zur Trainingsleistung zum Einsatz kommen werden.

Hinsichtlich ihrer Prognosesicherheit müssen sich neuronale Netzwerke mit herkömmlichen statistischen Verfahren vergleichen lassen. Wegen des Black-Box-Verhaltens neuronaler Netzwerke ist dies bislang allerdings nur eingeschränkt möglich. Es gilt daher, Kombinationsverfahren neuronaler Netze mit statistischen Methoden zu entwickeln. Eine Möglichkeit besteht darin, den linearen Teil der Prognose durch lineare Regression zu ermitteln, um darauf aufbauend die nichtlinearen Korrekturen durch das neuronale Netzwerk abzubilden.

Analog der automatischen Optimierung der Regelbasis und der Partitionierung eines Fuzzy-Systems durch neuronale Netzwerke besteht die Möglichkeit der automatischen Optimierung der Architektur neuronaler Netzwerke durch genetische Algorithmen und Evolutionsstrategien [SCH96]. Der zentrale Gedanke des Ansatzes besteht darin, den genetischen Algorithmus bzw. die Evolutionsstrategie nach einer optimalen Netzwerktopologie des Problems suchen zu lassen. Dadurch läßt sich eine der zentralen Schwierigkeiten neuronaler Netzwerke vermeiden, der heuristischen Bestimmung der Netzwerkparameter.

Abstract

In this paper four models to describe the modal split in the urban area are tested and evaluated. These models are: statistical models (logit models) and three models of the artificial intelligence (neural networks, fuzzy systems and a combination of neural networks and fuzzy systems (so called neuro fuzzy systems)). So far, models of the artificial intelligence are mainly used to solve problems in nature or in engineering sciences. In this paper, the models are applied to solve a problem of social sciences: the modal split in an urban area. The main difference between nature and social sciences is the occurrence of irrational decisions from the human beings in social sciences. It was shown that neural networks can be used in nature as well as in social sciences. Over and above that they provide more detailed information than statistical methods. Fuzzy systems can only be used in technical questions. Neuro fuzzy systems can be used to complete the information from neural networks.

Anhang: Die Fragen des Fragebogens

1. Bewertung des zeitlichen Aufwands mit Schulnoten von Auto, MVV und Rad
2. Bewertung der Bequemlichkeit mit Schulnoten von Auto, MVV und Rad
3. Bewertung der Kosten mit Schulnoten von Auto, MVV und Rad
4. Geschlecht der befragten Person
5. höchster allgemeiner Schulabschluß der befragten Person
6. Alter der befragten Person
7. Art der Erwerbstätigkeit der befragten Person (vollzeit, teilzeit, nebenher, Hausfrau/-mann, Rentner, arbeitslos, Schüler(in), Student(in), Wehr-/Zivildienstleistender, sonstiges)
8. berufliche Stellung der befragten Person (Angestellte(r), Arbeiter(in), Beamte(r), Selbständige(r), Auszubildende(r), mithelfende(r) Familienangehörige(r), sonstiges)
9. Höhe des monatlichen Nettoeinkommens des gesamten Haushalts
10. Familienstand der befragten Person
11. sechs Fragen zur Bewertung der Umweltsituation
12. Autoverfügbarkeit bei befragter Person
13. Entfernung Wohnort-Arbeitsplatz der befragten Person
14. Parkmöglichkeit an der Arbeitsstätte bei befragter Person
15. Umsteigehäufigkeit der öffentlichen Verkehrsmittel beim Weg Wohnort-Arbeitsplatz der befragten Person
16. Fahrtdauer der öffentlichen Verkehrsmittel beim Weg Wohnort-Arbeitsplatz der befragten Person überwiegend getroffene Verkehrsmittelwahl der befragten Person im letzten Monat

Literatur

- [ALT92] C. van Alrock, B. Krause, H. Zimmermann (1992). Advanced Fuzzy Logic Control Technologies in Automotive Applications. In: Proc. IEEE Conf. on Fuzzy Systems 1992, 835-842, San Diego.
- [AUF86] E. Aufhauser, M.M. Fischer, H. Schönhofer (1986). A Disaggregated Probabilistic Approach to a Regulated Housing Market with Emphasis on the Demand Side: The Vienna Case. Papers of the Regional Science Association, Vol. 60, 133-153.
- [BEN77] M. Ben-Akiva, T. Atherton (1977). Methodology for Short-Range Travel Demand Predictions. Journal of Transportation and Economic Policy, Vol. 26, 321-341.
- [BEN85] M. Ben-Akiva, S. Lerman (1985). Discrete Choice Analysis: Theory and Application to Travel Demand. MIT-Press, Cambridge.
- [BER92] H.R. Berenji (1992). A Reinforcement Learning-Based Architecture for Fuzzy Logic Control. Int. J. Approximate Reasoning, 6: 267-292.
- [BRÜ95] J. Brüderl, P. Preisendörfer (1995). Der Weg zum Arbeitsplatz: Eine empirische Untersuchung zur Verkehrsmittelwahl, in: A. Diekmann, A. Franzen (Hrsg.): Kooperatives Umwelthandeln - Modelle, Erfahrungen, Maßnahmen. Rüegger, Zürich.
- [CHE93] B. Cheng, D.M. Titterton (1993). Neural Networks: A Review from a Statistical Perspective. Statistical Science 9(1), 2-54.
- [DOM75] T.A. Domencich, D. McFadden (1975). Urban Travel Demand - A Behavioral Analysis. North Holland, Amsterdam.
- [ECC90] J.C. Eccles (1990). Das Gehirn des Menschen. Piper, München.
- [ECC94] J.C. Eccles (1994). Die Evolution des Gehirns - die Erschaffung des Selbst. Piper, München.
- [GEM92] S. Geman, E. Bienenstock, R. Doursat (1992). Neural Networks and the Bias/Variance Dilemma. Neural Computation, 4, 1-58.
- [GRE92] W. Greene (1992). LIMPED, User's Manual and Reference Guide, Version 6.0. Econometric Software, New York.
- [HUG98] J. Hugo (1998). Untersuchung von neuronalen Netzwerken und von Neuro-Fuzzy-Systemen in den Sozialwissenschaften. Dissertation, Universität Stuttgart, Fakultät Physik.
- [JOH70] N.L. Johnson, S. Kotz (1970). Continuous Univariate Distributions. Boston: Houghton Mifflin.
- [KOS92] B. Kosko (1992). Neural Networks and Fuzzy Systems. Prentice Hall, Englewood Cliffs NJ.
- [KRU92] R. Kruse (1992). On the Semantic Foundations of Fuzzy Probability Theory and Fuzzy Statistics. In: H. Bandemer (Hrsg.), Modelling Uncertain Data, 131-135. Akademie Verlag, Berlin.
- [KUD96] F. Kudermann (1996). Optimierung eines neuronalen Netzwerks mittels Pruningverfahren. Diplomarbeit, Universität Stuttgart, Fakultät Physik.
- [LAV69] C.A. Lave (1969). A Behavioral Approach to Modal Split Forecasting. Transportation Research, Vol. 3, 463-480.
- [MAI90] G. Maier, P. Weiss (1990). Modelle diskreter Entscheidungen. Springer, Wien/New York.
- [MAN73] Manski, C.F. (1973). The Analysis of Qualitative Choice. Ph.D. Dissertation, Department of Economics, MIT, Cambridge, Mass.
- [MAN77] Manski, C.F. (1977). The Structure of Random Utility Models, Theory and Decision, Vol. 8, 229-254.
- [NAU92] D. Nauck, F. Klawonn, R. Kruse (1992). Fuzzy Sets, Fuzzy Controllers and Neural Networks. Wissenschaftliche Zeitschrift der Humboldt-Universität zu Berlin, R.Medizin, 41(4): 99-120
- [NAU93] D. Nauck, F. Klawonn, R. Kruse (1993). Combining Neural Networks and Fuzzy-Controllers. In: E.P. Klement, W. Slany (Hrsg.): Fuzzy Logic in Artificial Intelligence (FLAI93), 35-46. Springer, Berlin.
- [NAU96] D. Nauck, F. Klawonn, R. Kruse (1996). Neuronale Netze und Fuzzy-Systeme. Vieweg, Braunschweig/Wiesbaden.
- [SCH96] E. Schöneburg, F. Heinzmann, S. Feddersen (1996). Genetische Algorithmen und Evolutionsstrategien. Addison-Wesley, Bonn.
- [SPI93] M. Spies (1993). Unsicheres Wissen. Spektrum Akademischer Verlag, Heidelberg.
- [WER74] P.J. Werbos (1974). Beyond regression: new tools for prediction and analysis in the behavioral sciences. Ph.D. thesis, Harvard University, Cambridge MA.
- [WHI91] H. White (1991). Connectionist Nonparametric Regression: Multilayer Feed-forward can learn Arbitrary Mappings. Neural Networks 3, 535-550.
- [ZAD65] L.A. Zadeh (1965). Fuzzy Sets. Information and Control, 8: 338-353.
- [ZAD73] L.A. Zadeh (1973). Outline of a New Approach to the Analysis of Complex Systems and Decision Processes. IEEE Trans. Systems, Man & Cybernetics, 3: 28-44.
- [ZUM98] D. Zumkeller (1998). Haushaltspanel 1997/98 – Auswertung. Schlußbericht zum Forschungsauftrag BMV FE 96524/97.