

Künstliche Neuronale Netze als Verfahren der betrieblichen Entscheidungsunterstützung

Die Leistungsfähigkeit des menschlichen Gehirns in der Verarbeitung komplexer Signale, z.B. in der Mustererkennung von Geräuschen oder Gesichtern, übertrifft noch immer die konventioneller von-Neumann-Computer. Aus dieser Motivation heraus wurden Algorithmen entwickelt, die sich an der parallelen und verteilten Informationsverarbeitung in natürlichen neuronalen Netzen orientieren; vgl. Zell (1997). Im Kontext betriebswirtschaftlicher Entscheidungsprobleme repräsentieren **künstliche neuronale Netze (KNN)** eine Klasse mathematisch-statistischer Verfahren für die computergestützte Planung.

KNN als Untersuchungsgegenstand der Wirtschaftsinformatik

Die Erforschung von KNN reicht bis in die Anfänge computergestützter Informationsverarbeitung zurück. Nach grundlegenden neurophysiologischen Arbeiten der 40er Jahre wurde bereits 1958 das **Perzeptron** zur Erkennung handgeschriebener Zeichen entwickelt. Doch erst 1985 gelang der Durchbruch dank der Entwicklung eines mathematischen Verfahrens zur Anpassung mehrschichtiger KNN, so genannter **Multilayer Perzeptrons**, an komplexe Problemstellungen allein aus Beobachtungen. KNN zeichnen sich aufgrund ihrer spezifischen Informationsverarbeitung zur Mustererkennung insbesondere durch

- hohe Freiheitsgrade in der Modellierung,
- Lernfähigkeit und Adaptionfähigkeit,
- Modellierung ohne Strukturwissen über das zu lösende Problem,
- Fehlerrobustheit und Fehlertoleranz sowie
- Nichtlinearität aus; vgl. Alex (1998).

Problemstellungen der **Mustererkennung** finden sich auch in der betrieblichen Realität, z.B. in der Erkennung von Regelmäßigkeiten in Absatzzeitreihen, dem Kaufverhalten von Kunden oder der Rückzahlung von Krediten. Realwirtschaftlich-soziale Systeme wie Betriebe, Branchen oder Märkte zeichnen sich dabei häufig durch hohe Komplexität, Dynamik und Intransparenz aus. Bezogen auf die Strukturdeterminanten dieser Systeme bedeutet dies eine

- hohe Anzahl relevanter Variablen,
- hohe Vernetzung der Variablen mit statischen und dynamischen Interdependenzen,
- erschwerte Ableitung kausaler Ursache-Wirkungs-Beziehungen des realen Systems und
- nicht-lineare Beziehungen der Variablen.

Aufgrund der übereinstimmenden Merkmale bieten sich KNN als Verfahrensklasse zur Lösung betriebswirtschaftlicher Entscheidungsprobleme, insbesondere der **Prognose und Klassifikation**, an. Zur computergestützten Entscheidungsfindung im Management werden KNN in so genannten Decision-Support-Systemen eingesetzt, was ihre Betrachtung als Untersuchungsobjekt der Wirtschaftsinformatik verdeutlicht; vgl. Voß/Gutenschwager (2001). KNN konkurrieren dabei mit exakten, heuristischen oder meta-heuristischen Verfahren der multivariaten Statistik und des Operations Research. Daher stellt sich die Frage, inwieweit KNN in der Lage sind, betriebswirtschaftliche Problemstellungen effizienter zu lösen als herkömmliche Verfahren. Nach einer Einführung in die Informationsverarbeitung in KNN werden ihre Modellierungsvorzüge exemplarisch an einem Entscheidungsproblem der Klassifikation zur Kreditvergabe und einer Umsatzprognose dargestellt

Grundlagen der Informationsverarbeitung in KNN

Biologische Nervensysteme bestehen aus vielen einfachen Nervenzellen, **Neuronen** genannt, die den Grad ihrer Reizung über Verbindungen untereinander austauschen. Natürliche neuronale Netze nehmen Eingangsreize auf, berechnen den Grad der Reizung (die **Aktivierung**) im Nervensystem und erzeugen daraus eine Handlung. Die Informationsverarbeitung in den computergestützten Modellen der KNN entsteht in Analogie

zur biologischen Signalverarbeitung. Die Verarbeitung von Eingangssignalen x_i in Ausgabesignale y_j erfolgt durch eine simple mathematische Berechnung (tangens hyperbolicus der summierten Eingabewerte) einer positiven Ausgabe, sofern die Summe der Eingangssignale einen Schwellenwert θ_j übersteigt. Diese Berechnung erfolgt **parallel** in vielen künstlichen Neuronen, die in funktionsgleiche **Schichten** aufgeteilt werden:

- einer Eingabeschicht, die die Werte der Außenwelt annimmt,
- von außen versteckte Schichten, die die Reizberechnung umsetzen, und
- einer Ausgabeschicht, die eine Ausgabe als Reaktion auf die Eingabe ausgibt.

Die Neuronen sind über **gewichtete Verbindungen** w_{ij} miteinander verbunden, wobei das Gewicht einer Verbindung die Stärke der Beziehungen zwischen den Neuronen bestimmt. Die Verbindungen leiten die mathematischen Ausgabewerte o_j von der Eingabeschicht zur Ausgabeschicht weiter; vgl. Bishop (1995). Eine mögliche KNN-Architektur zeigt Abb.1:

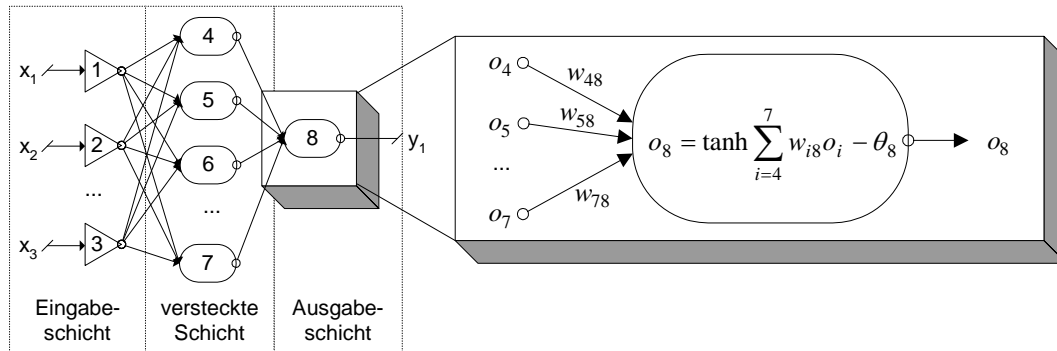


Abb. 1: Architektur eines KNN und die Informationsverarbeitung im Neuron

Hohe Freiheitsgrade in der Modellierung von KNN

Betriebswirtschaftliche Entscheidungen erfordern die Berücksichtigung einer Vielzahl unterschiedlich skaliert, interdependent verknüpfter Variablen. Für die Beurteilung der Bewilligung oder Ablehnung eines Kredites können unterschiedliche Eingangsvariablen wie z.B. *Einkommen*, *Alter*, *Geschlecht* etc. wichtig sein. Für eine Umsatzprognose können z.B. vergangene *Umsatzwerte*, *Werbeaktivitäten*, *Feiertageinflüsse* etc. relevant sein. KNN können eine **hohe Anzahl unabhängiger Eingangsvariablen** auf eine oder mehrere abhängige Ausgangsvariablen abbilden. Zusätzliche Variablen werden einfach als weitere Neuronen der Eingabeschicht integriert. Die vernetzte Verarbeitung aller Eingabe- zu Ausgabevariablen erlaubt die Abbildung von **statischen Interdependenzen** zwischen den Variablen. Dabei können simultan Realisationen der Variablen zu früheren Zeitpunkten in das Modell einfließen und somit **dynamische Interdependenzen** der Variablen, z.B. als zeitlich verzögerte Absatzwirkung einer Werbeaktion in nachfolgenden Perioden, modelliert werden. Eine KNN-Architektur kann somit sowohl zur Klassifikation einer Kreditbewilligung als auch zur Umsatzprognose angewendet werden, lediglich durch die Variation der verwendeten Eingangs- und Ausgabewerte sowie der Verbindungsgewichte, wie in Abb. 2 ersichtlich:

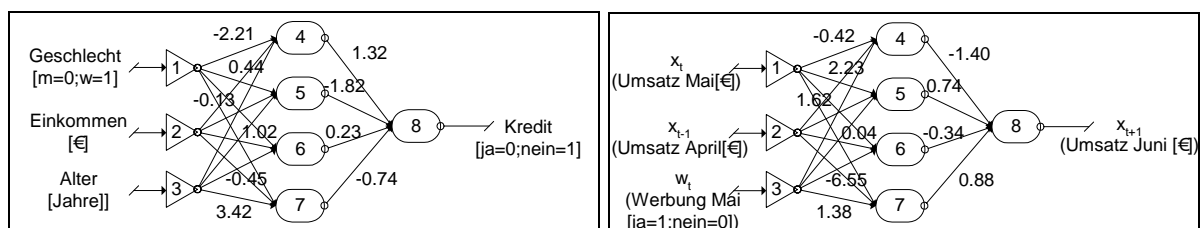


Abb. 2: Architektur eines KNN zur Kreditvergabe oder zur Umsatzprognose

KNN erlauben durch die Verarbeitung von Input- und Outputdaten **beliebiger Skalierung** die Modellierung unterschiedlicher statistischer Verfahren. Ein KNN kann die funktionalen Beziehungen des Erklärungsmodells zwischen metrisch skalierten (z.B. *Einkommen*: $[0, \dots, \infty]$, *Alter*) sowie nominal skalierten (z.B. *Geschlecht*: [männlich; weiblich]) Eingabe- und

Ausgabevariablen lernen. Somit simulieren KNN nichtlineare Varianten unterschiedlicher, statistischer Verfahren:

- Regressionsanalyse (metrische unabhängige → metrische abhängige Variablen),
- Diskriminanzanalyse (metrische unabhängige → nominale abhängige Variablen),
- Varianzanalyse (nominale unabhängige → metrische abhängige Variablen) und
- Kontingenzanalyse (nominale unabhängige → nominale abhängige Variablen).

Ein KNN mit fester Architektur ersetzt somit allein durch unterschiedlich skalierte Input- und Outputvariablen verschiedene statistische Verfahren zur Prognose und Klassifikation. Ferner können metrisch und nominal skalierte Variablen gleichzeitig als Eingabe- oder Ausgabevariablen dienen oder mehrere abhängige Variablen simultan ausgegeben werden; vgl. Reed (1998).

Eine Kreditentscheidung kann somit durch ein KNN mit identischer Architektur in Eingabe-, versteckter und Ausgabeschicht und bei identischen Eingangsdaten als nominale [Ja;Nein]-Entscheidung, als metrische Prognose der erwarteten Höhe der Kreditrückzahlung oder als Wahrscheinlichkeit der Kreditrückzahlung im Intervall [0;1] modelliert werden. Metrische Umsatzprognosen können den erwarteten Umsatz in [€] vorhersagen oder als nominale Prognose eine [steigt;fällt]-Aussage tätigen. Ein KNN beliebiger Architektur bietet somit als Verfahren **höhere Freiheitsgrade** in der Modellierung von Entscheidungsproblemen als ein einzelnes statistisches Verfahren; vgl. Backhaus et al. (1998).

Lernfähigkeit und Adaptionfähigkeit in KNN

Natürliche neuronale Netze lernen aus Beispielen. Bei einer fehlerhaften Handlung auf ein Eingangssignal, z.B. beim Greifen eines Gegenstandes, wird der Versuch solange wiederholt, bis die Aktion korrekt gelernt wurde. Das **Lernen** findet dabei über eine Veränderung der neuronalen Verbindungen statt. Im Fall einer Kreditvergabe lernt der Sachbearbeiter die relevanten Einflussfaktoren für eine erfolgreiche Rückzahlung des Kredites ebenfalls aus Beispielen anhand fehlerhafter und korrekter Klassifikation. Das gebildete Erfahrungswissen repräsentiert die Verbindung der unabhängigen Variablen aus *Alter, Einkommen, Geschlecht* etc. auf die Vorrausage der *Kreditgewährung*; vgl. Reed (1998).

Entsprechend lernen KNN problembezogenes Wissen aus empirischen Beobachtungen. Der Lernprozess eines KNN für die Kreditvergabe erfolgt durch **Training** aus Beispielen vergangener erfolgreicher und fälschlicherweise vergebener Kredite, indem Eingabewerte jedes Kreditantrages an die Eingabeschicht des KNN angelegt werden und daraus ein Ausgabewert der Kreditbewilligung oder Ablehnung berechnet wird.

- Wird der Kredit vom KNN korrekt bewilligt, entsteht kein Fehler.
- Wird der Kredit vom KNN abgelehnt, obwohl er in Wirklichkeit zurückgezahlt wurde, ist der Ausgabewert falsch und es entsteht ein Fehler.
- Wird der Kredit bewilligt, obwohl er nicht zurückgezahlt wurde, entsteht ein Fehler.

Durch einen mathematischen Lernalgorithmus werden die fehlerverursachenden Verbindungen bestimmt und das Gewicht der Verbindung so verändert, dass der Fehler sinkt. Alle vorliegenden Beispiele werden dem KNN so lange präsentiert, bis der Fehler gegen Null sinkt und das KNN den zu erlernenden Zusammenhang zwischen *Alter, Einkommen, Geschlecht* und einer korrekten *Kreditbewilligung* erlernt hat. Dabei kann das KNN auch auf neue, nicht erlernte Beispiele angewendet werden, was als **Generalisierungsfähigkeit** bezeichnet wird. Die **Veränderung der Verbindungsgewichte** im Lernen entspricht der Parametrisierung statistischer Verfahren, jedoch einschließlich Training zur Lösung unterschiedlicher Anwendungsprobleme. Die Ausprägung aller Verbindungsgewichte repräsentiert den gelernten Zusammenhang der Datenmuster. Es wurde mathematisch bewiesen, dass ein KNN mit endlich vielen Neuronen in zwei versteckten Schichten jede mathematische Funktion und logische Regel mit beliebiger Genauigkeit approximativ erlernen kann. Ein KNN mit unveränderter Architektur ist somit in der Lage, jeden Systemzusammenhang allein durch Variation der Beträge der Verbindungsgewichte aus den Trainingsdaten zu erlernen; vgl. Stahlbock (2002).

Die **Adaptionfähigkeit** software- oder hardwaretechnisch implementierter KNN an neue Daten und veränderte Problemstellungen ist durch wiederholtes Trainieren oder

Nachtrainieren mit einzelnen Datensätzen ohne Neuprogrammierung des Verfahrens gegeben. Dabei können neuartige oder veränderte funktionale Zusammenhänge adaptiert werden, die in der herkömmlichen Statistik nur durch einen Wechsel der Verfahrensklasse realisierbar wären.

Modellbildung ohne Strukturwissen in KNN

Die Komplexität realer Systeme verhindert die Aufstellung einfacher Hypothesen über empirische Ursache-Wirkungs-Beziehungen. Im Sinne einer „black-box“ sind zwar die Eingangsvariablen der Werbeausgaben etc. sowie das Systemresultat des Umsatzes beobachtbar, nicht aber die Systembeziehungen seiner Entstehung. Die Vorteile der Lernfähigkeit aus Beispielen und der Modellierungsflexibilität von KNN erlauben die **Modellbildung** und Problemlösung **ohne vorheriges Strukturwissen** oder **Hypothesenbildung** über die Struktur der Systemzusammenhänge. Dies wird im Vergleich zu Expertensystemen der Künstlichen Intelligenz (KI) und statistischen Verfahren deutlich:

Expertensystem	Statistisches Verfahren der Diskriminanzanalyse
WENN <i>Geschlecht</i> = „männlich“ UND <i>Einkommen</i> < „40000“ ODER <i>Einkommen</i> > „40000“ UND <i>Alter</i> < „18“ DANN !Abweisen Kredit	$Y_k = -2,24 + 1,23x_{\text{Geschlecht}} - 0,28x_{\text{Einkommen}} + 8,24x_{\text{Alter}}$

Abb. 3: Entscheidungsmodellierung der Kreditvergabe durch unterschiedliche Verfahren

Ein Expertensystem der Kreditvergabe trifft Entscheidungen auf der Basis der zuvor durch den Sachbearbeiter explizit formulierten Regeln. Die Formulierung der Regeln erfordert die bewusste Kenntnis der kausalen Zusammenhänge und ihrer Intensität in der Realität durch den Experten. Herkömmliche statistische Verfahren erfordern dagegen eine Hypothese über die Art der funktionalen Beziehung und der Skalierung der Variablen zur Auswahl einer passenden Verfahrensvariante. Diese Verfahrensvariante wird anschließend parametrisiert und somit die Intensität der funktionalen Beziehung aus den Beobachtungen ermittelt. Bei KNN muss hingegen a priori keine Form des zu modellierenden Zusammenhangs bekannt sein. Das KNN generiert das zutreffende Modell des funktionalen Zusammenhangs **selbstständig** aus realen Beobachtungen im Lernprozess, ohne vorherige Hypothesenbildung; vgl. Stahlbock (2002).

Fehlerrobustheit und Fehlertoleranz in KNN

Werden einem KNN fehlerbehaftete Daten präsentiert, produziert es dennoch annähernd richtige Ausgaben. Insbesondere verrauschte, mit Störungen überlagerte Daten sowie fehlende Daten stellen ein prominentes Problem der betrieblichen Realität dar. Im obigen Beispiel des Expertensystems führt das Fehlen eines Wertes der Eingangsvariablen, z.B. *Geschlecht*, zu einem Programmabbruch ohne Ergebnis. KNN geben in diesem Fall suboptimale aber zulässige Ausgaben und erlauben eine **fehlerrobuste** Ergebnisfindung. Ähnlich nimmt die Leistungsfähigkeit auch bei verrauschten, inkorrekten, inkonsistenten oder widersprüchlichen Eingabedaten aus fehlerhafter Datenerhebung oder -verarbeitung nur graduell ab und endet nicht abrupt, was als „**graceful degradation**“ bezeichnet wird. Dies erlaubt eine gewisse Resistenz gegen Messungenauigkeiten in der Datenerhebung, z.B. des Einkommens oder des Alters, und wird z.T. auch als Verarbeitung oder Schließen aus **unscharfen Informationen** interpretiert. Die **Fehlertoleranz** von KNN geht darüber hinaus und erlaubt bei hardware- oder softwaretechnischer Implementierung trotz Ausfall einer oder weniger Neuronen oder Verbindungen die Weiterverarbeitung von Daten, da die verteilte Wissensrepräsentation graduelle Informationsverluste kompensieren kann; vgl. Alex (1998).

Abbildung von Nichtlinearitäten in KNN

Die Wirkung einer Variablen ist i.d.R. durch **interdependente Wirkungsverknüpfungen** abhängig vom Niveau anderer Variablen. Nichtlineare Zusammenhänge, z.B. als multiplikative Interaktionseffekte zwischen unabhängigen Variablen, treten in empirischen betriebswirtschaftlichen Problemstellungen regelmäßig auf. Eine Erhöhung der Werbeausgaben führt nicht unendlich zu einer konstanten Erhöhung des Umsatzes im Sinne

exponentieller, degressiver oder s-förmiger Kurvenverläufe. Lineare Zusammenhänge implizieren z.B. konstante Grenzerlöse und -kosten und widersprechen somit selbst grundlegenden betriebswirtschaftlichen Modellen. In diesen Fällen können konventionelle, lineare Verfahren der Statistik nur Näherungslösungen bieten. Durch Überlagerung der Ausgaben aller Neuronen im KNN können **nichtlineare funktionale Zusammenhänge** und lineare Zusammenhänge unmittelbar abgebildet werden.

Nachteile von KNN

Aus der verteilten Repräsentation des Wissens in den Verbindungsgewichten leitet sich der Nachteil einer **fehlenden Erklärungskomponente** eines KNN in speziellen Entscheidungsproblemen ab, da eine direkte Interpretation des im KNN kodierten Wissens unmöglich ist. Eine Interpretation des expliziten Wissens im Expertensystem in Abb. 3 ist unmittelbar möglich, während aus der Analyse der Aktivierungszustände, Verbindungsgewichte und Schwellenwerte des KNN zur Kreditvergabe in Abb. 2 keine Aussagen getroffen werden können. Dies kann insbesondere bei wichtigen taktischen oder strategischen Einzelentscheidungen den Einsatz von KNN verbieten, wenn durch die Auswahl einer suboptimalen Handlungsalternative hohe Kosten entstehen, der Grund für die Wahl der Alternative jedoch nicht erklärt werden kann.

Darüber hinaus stellt die neuronale Informationsverarbeitung ein neues und komplexes, in der KI als subsymbolisch bezeichnetes, Paradigma dar. Zur erfolgreichen Anwendung von KNN ist daher ein spezifisches und **umfangreiches Methodenwissen** erforderlich.

KNN als universelles Entscheidungsverfahren?

Die Verfahrensklasse der KNN befindet sich in einer Phase der marktwirtschaftlichen Erprobung. Erste Anwendungserfolge lassen nicht auf eine grundsätzliche Dominanz von KNN gegenüber herkömmlichen Verfahren schließen. Der hohen Genauigkeit der KNN, z.B. in der Kreditbewilligung, steht eine hohe Verfahrenskomplexität gegenüber, die unter Berücksichtigung der Transaktionskosten ihrer Anwendung einen wirtschaftlichen Einsatz nicht immer rechtfertigt. Obwohl hohe Freiheitsgrade und die Komplexität in der Modellierung der KNN durch Automatisierung in Softwaresimulatoren verringert werden, bleibt ein erhebliches Methodenwissen erforderlich. Die Anwendung von KNN in betrieblichen Anwendungsbereichen erfordert daher eine spezifische **Kosten-Nutzen-Abwägung**. Für lineare Problemstellungen mit geringer Komplexität können herkömmliche Verfahren der linearen Statistik weiterhin effizient eingesetzt werden. KNN könnten hingegen als methodische Erweiterung multivariater Verfahren der Statistik bei nichtlinearen, hochdimensionalen Problemstellungen erprobt werden, in denen kein Strukturwissen über Kausalzusammenhänge vorhanden ist.

Dipl.-Kfm. Sven F. Crone, Hamburg

Literaturempfehlungen:

- Alex, B. (1998): Künstliche Neuronale Netze in Management-Informationssystemen – Grundlagen und Einsatzmöglichkeiten, Wiesbaden.
- Backhaus, K.; Erichson, B.; Plinke, W.; Weiber, R. (1996): Multivariate Analysemethoden – eine anwendungsorientierte Einführung, 8. verb. Aufl., Berlin.
- Bishop, C. M. (1995): Neural Networks for Pattern Recognition, London.
- Reed, R. D.; Marks II, R. J. (1998): Neural Smithing – Supervised Learning in Feedforward Artificial Neural Networks, Cambridge.
- Stahlbock, R. (2002): Evolutionäre Entwicklung künstlicher neuronaler Netze zur Lösung betriebswirtschaftlicher Klassifikationsprobleme, Berlin.
- Voß, S.; Gutenschwager, K. (2001): Informationsmanagement, Berlin.
- Zell, A. (1997): Simulation neuronaler Netze, Bonn.