

1. *Neurale netwerken in marketing*

R.C.P. VAN DER VEER, B. WIERENGA EN J.C.H.W. KLUYTMANS

SAMENVATTING

Doel van dit artikel is een introductie te geven van neurale netwerken voor diegenen die in marketing of marktonderzoek werkzaam zijn. Eerst wordt ingegaan op de (biologische) achtergrond van neurale netwerken. Vervolgens worden diverse aspecten van neurale netwerken besproken: leren aan de hand van voorbeelden, parameters, 'backpropagation netwerken', prototype-gebaseerde netwerken. Daarna wordt een vergelijking getrokken met traditionele statistiek en artificiële intelligentie.

Het artikel vervolgt met een discussie over mogelijke toepassingen. Tot slot worden zaken besproken die samenhangen met toepassing in de praktijk, ook omvattend de beschrijving van een aantal beschikbare programma's.

1. INLEIDING

Neurale netwerken trekken vandaag de dag de aandacht in vele vakgebieden. Ook in marketing zien we de eerste toepassingen verschijnen: Hruschka (1991), Mazanec (1993), Hoptroff (1992), Kluytmans, Wierenga & Spigt (1993) en Wierenga en Kluytmans (1994).

Neurale netwerken zijn betrekkelijk eenvoudige wiskundige algorithmen die de input variabelen van een bepaald systeem (in marketing bijvoorbeeld: de marketing mix-instrumenten of de kenmerken van klanten) in verband brengen met de output van dat systeem (in marketing bijvoorbeeld: aankopen van klanten). De architectuur van neurale netwerken biedt flexibele presentatiemogelijkheden die een grote diversiteit aan situaties aankunnen. In principe kan een neuraal netwerk worden gezien als een lerend mechanisme: het netwerk 'leert' de connecties tussen inputs en outputs uit gepresenteerde 'cases'. Voor nieuwe cases kan het aldus getrainde netwerk de uitkomsten voorspellen.

In dit artikel wordt eerst ingegaan op de (biologische) achtergrond van neurale netwerken¹. Vervolgens worden diverse aspecten van neurale netwerken besproken waarbij ze onder andere worden vergeleken met traditionele methoden zoals multivariate technieken. Vervolgens komen de toepassingsmogelijkheden van neurale netwerken in marketing met een aantal voorbeelden aan de orde.

In de daaropvolgende paragraaf worden de resultaten vermeld van exploratief onderzoek naar de mogelijkheden van neurale netwerken bij marketing tijdreeksen.

Daarna volgt een overzicht van neurale netwerken tools en wordt het artikel afgesloten met enkele conclusies.

2. ACHTERGROND

De structuur van het biologisch zenuwstelsel van levende organismen vormt de grondslag voor de neurale netwerk theorieën. Met de ontwikkeling van dergelijke theorieën houden onderzoekers op het gebied van de kunstmatige intelligentie zich bezig. Men tracht modellen te ontwikkelen van de hersenen op basis van uiteenlopende experimenten. De hersenen bestaan uit een groot aantal kleine elementjes (zenuwcellen/neuronen), die via talloze verbindingen met elkaar in contact staan.

De grote drijfveer hierbij is om te komen tot de volledige doorgronding van het brein. Evenals het brein beschikken kunstmatige neurale netwerken immers over zelf-organiserend vermogen. Het is een uitdaging om een machine te ontwikkelen met de kracht van het brein. Deze machine zou in staat zijn tot het snel en doeltreffend uitvoeren van een aantal taken, zoals patroonherkenning in omgevingen met zeer veel ruis (waarnemen), het aansturen van complexe mechanismen met zeer veel vrijheidsgraden (motorische coördinatie), het over lange tijd bewaren van grote hoeveelheden informatie die bijzonder snel actief kan worden gemaakt (geheugen) en het oplossen van ingewikkelde en slecht gedefinieerde problemen door het met elkaar in verband brengen van relevante informatie (denken). Deze machine zou daarnaast in staat zijn om deze vaardigheden al doende te verwerven (leren) (Murre, 1991).

De specifieke structuur van het brein is waarschijnlijk cruciaal voor zijn indrukwekkende prestaties. Deze structuur is nog maar gedeeltelijk ontrafeld en de bestaande modellen doen nog lang geen recht aan de complexiteit van het brein. Het brein bestaat uit miljarden neuronen die biljoenen synaptische verbindingen hebben. Ieder neuron is te beschouwen als een complexe chemische fabriek die op een zeer groot aantal manieren beïnvloed kan worden en dus een enorm potentieel aan informatieverwerkende mechanismen heeft.

De individuele elementen in een netwerk van zenuwcel-achtige elementen worden geprikkeld en geremd. Naargelang de sterkte van deze impulsen en de weerstandswaarde van een bepaald element, zal dat element zelf in zekere mate geactiveerd worden en andere cellen via zijn verbindingen gaan beïnvloeden. Het gedrag van het neurale netwerk wordt bepaald door de sterkte van de verbindingen en de manier waarop de elementen met elkaar zijn verbonden.

De kennis van een neuraal netwerk ligt dus opgeslagen in het raamwerk van verbindingen. Het is niet zo dat elke verbinding overeenkomt met een brokje kennis. Iedere verbinding kan een rol spelen bij tal van vormen van kennis en elk stukje kennis kan over vele verbindingen verspreid zijn. Men spreekt van 'parallel distributed processing' (Rumelhart & McClelland, 1986), wat aangeeft dat vele elementen tegelijkertijd actief kunnen zijn en dat de verwerking van de informatie over een groot deel van het netwerk verspreid kan zijn.

De verbindingen tussen de elementen, waarin de kennis ligt opgeslagen, zijn veranderbaar. Bij iedere nieuwe waarneming van de omgeving worden de verbindingen in het netwerk heroverwogen. Er wordt 'geleerd' van nieuwe ervaringen. De neurale netwerken passen zich aan de omgeving en de veranderingen hierin aan. Vandaar dat zij geschaard kunnen worden onder de noemer 'adaptieve', of 'lerende' technieken.

Neurale netwerken worden reeds in een groot aantal uiteenlopende vakgebieden toegepast. Een aantal voorbeelden zijn kwaliteitszorg (Verdenius, 1993), prognostiseren op financiële markten (Refenez, et al., 1993) (Bastaens & van den Bergh, 1992), productie management (Epping & Nitters, 1990), het lezen en uitspreken van geschreven tekst (Sejnowski & Rosenberg, 1987), het maken van weersvoorspellingen, het filteren van ruis bij ECG's (Beale & Jackson, 1990), het matchen van gewenste eigenschappen met bestaande eigenschappen bij relatie-bemiddeling en personeelswerving (SMR).

3. INLEIDING NEURALE NETWERKEN

Neurale netwerken worden succesvol toegepast in situaties waarin moeilijke verbanden en vervuilde gegevens een rol spelen, zoals patroonherkenning in visuele of audio-gegevens. Enkele voorbeelden van zakelijke toepassingen zijn krediet-beoordeling (Collins et al, 1989), koersvoorspelling (Duives et al, 1991) en fraude-analyse. Met name binnen de marketing heeft men veelal te maken met ingewikkelde en veranderlijke verbanden. Daarnaast zijn de gegevens vaak onvolledig en vervuild. Neurale netwerken lijken daarmee bij uitstek geschikt voor marketingtoepassingen.

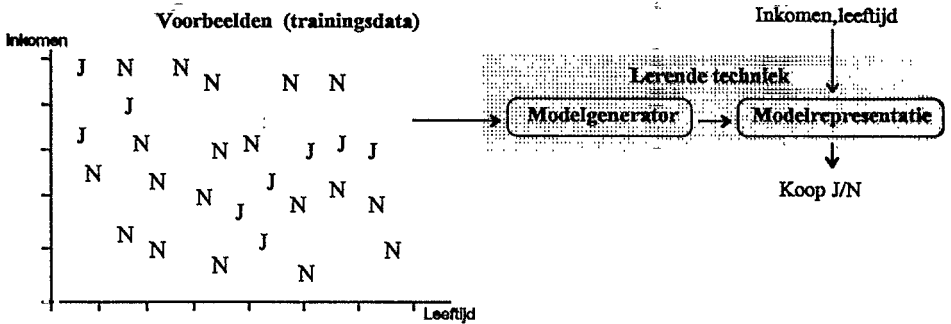
3.1 Leren aan de hand van voorbeelden

Om goede marketing te bedrijven is kennis van de markt (koopgedrag, stimulusgevoeligheid) onontbeerlijk. Marktgedrag is in de regel complex en wijzigt zichzelf continu, met als gevolg dat gedegen kennis helaas niet voor het oprapen ligt. Deze kennis zit echter wel verborgen in gegevens met registraties van gedrag, bijvoorbeeld een klantendatabase voorzien van persoonlijke kenmerken en gegevens over de relatie met die klant, zoals hoe vaak welke producten worden afgenomen. Door de steeds grotere beschikbaarheid van gegevens en de technologische ontwikkelingen bestaat er een toenemende belangstelling voor lerende technieken, zoals neurale netwerken die kennis uit gegevens kunnen extraheren.

Onder een lerende techniek wordt verstaan: een techniek die zonder noemenswaardige tussenkomst zelfstandig kennis uit gegevens kan extraheren in de vorm van een model. Een lerende techniek kan dus een gedrag modelleren, aan de hand van voorbeelden van het gedrag. Zo is het voor een marketeer bijvoorbeeld interessant om een model te hebben dat voorspelt of een klant uit een bepaalde bevolkingsgroep een bepaald produkt zal kopen, gegeven de specifieke kenmerken van die klant. Het grote voordeel van lerende technieken is dat met weinig menselijke inspanning en kennis eenvoudig een model kan worden opgesteld.

We gaan uit van de volgende case: er is een fabrikant van tuingereedschappen die een mailing, met een aanbieding van een maaimachine, wil sturen aan een aantal prospects.

De vraag is nu: 'Welke prospects zullen een maaimachine kopen?'. Als er gegevens beschikbaar zijn van welke klanten er in het verleden al dan niet een maaimachine hebben gekocht, dan kan een lerende techniek worden gebruikt om op basis van die gegevens een model te genereren. Deze voorbeeldgegevens vormen de trainingsdata, waarbij slechts twee variabelen bekend zijn: leeftijd en inkomen. De volgende figuur toont de trainingsdata met voor kopers een 'J' en voor niet-kopers een 'N'.



Figuur 1. Lerende techniek leert kopers en niet-kopers onderscheiden.

3.2 De werking van neurale netwerken

In de loop van de tijd is de term neurale netwerken een verzamelnaam geworden voor technieken die een architectuur hebben die grofweg vergelijkbaar is met het zenuwstelsel: een verzameling knopen die onderling massaal verbonden zijn en elk één uitvoer hebben. Daarbij verschilt per type netwerk hoe de gegevensstroom is, welke verwerking de knopen uitvoeren en hoe het netwerk moet leren door zijn verbindingen aan te passen (Rumelhart & McClelland, 1986). De belangrijkste eigenschap van neurale netwerken in het verband van dit stuk is dat ze autonoom een model van gedragspatronen op kunnen stellen en dus kunnen leren.

Parameters van netwerken

Aan neurale netwerken zijn de volgende aspecten of parameters te onderscheiden:

1. Het aantal knopen of neuronen

Een netwerk bestaat uit een stelsel van knopen, meestal variërend in aantal van minder dan 10 tot enkele tientallen. Een eerste stap in het opstellen van een netwerkmodel is het vaststellen van de knopen die verbonden zijn met de buitenwereld (invoer- en uitvoerlaag) en welke variabelen door welke knopen worden gerepresenteerd.

2. Verbindingen tussen de knopen

Elke knoop ontvangt invoer uit een aantal andere knopen en verstuurt zelf signalen naar een volgende set van knopen. Deze signalen worden verstuurd via een vast stelsel van verbindingen met voor elke verbinding een bijbehorend gewicht. Dit gewicht bepaalt de invloed van de ene knoop op de andere: als het gewicht op de verbinding van knoop A naar B groot is, dan heeft een signaal van A naar B een hoge activerende werking terwijl een negatief gewicht juist een remmend effect heeft. Bij de mees-

te netwerken kunnen de aktivatie-signalen maar één richting op wijzen: zogenaamde feed-forward netwerken. De knopen zijn meestal zo met elkaar verbonden dat er lagen van knopen ontstaan: de knopen in laag 1 zijn dan allemaal verbonden met alle knopen uit laag 2 en die in laag 2 met die in 3.

3. Aktivatie-toestand per knoop

Elke knoop berekent op grond van zijn invoersignalen zijn aktivatie-toestand die weer wordt doorgegeven aan volgende knopen. Een voorbeeld van een eenvoudige aktivatie-functie is: sommeer alle invoersignalen en vergelijk het resultaat met een grenswaarde; als het resultaat groter is zet dan de aktivatie-toestand op 1 en anders op 0. De invoer per verbinding wordt berekend door het inkomende aktivatie-signaal te vermenigvuldigen met het gewicht dat bij die verbinding hoort.

4. Leerregel

Een systeem dat bestaat uit een groot aantal eenheden, die elk voor zich eenvoudige berekeningen uitvoeren, kan in staat zijn complexe computaties uit te voeren. Het probleem is echter om op grond van een set gegevens dit systeem te konstrueren. De leerregel geeft aan hoe van een bepaald neurale netwerk, met een vooraf gekozen aantal knopen en verbindingen, de gewichten worden aangepast. Meestal gebeurt dit door een groot aantal keren alle voorbeelden uit de trainingsset te behandelen en elke keer de gewichten een klein beetje aan te passen.

De volgende twee belangrijkste soorten leerregels kunnen worden onderscheiden:

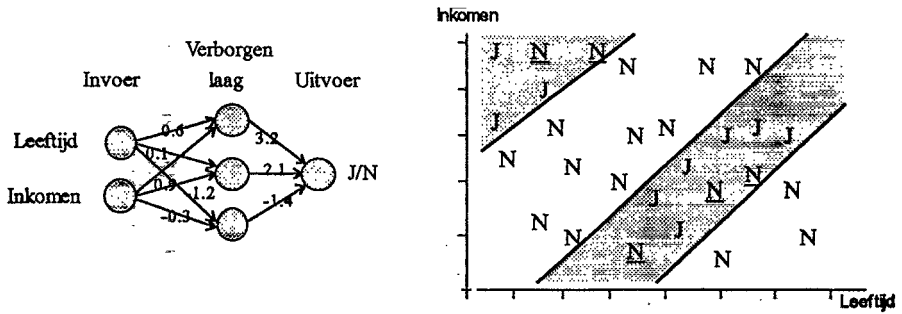
Supervised-leren: tijdens de trainings-fase is bekend wat voor elk voorbeeld de gewenste uitvoer is en leert het netwerk deze uitvoer te reproduceren.

Unsupervised-leren: hierbij leert het neurale netwerk zelf de onderverdelingen maken in de aangeboden voorbeelden (vergelijk clustering).

Backpropagation netwerken

De meest gebruikte soort neurale netwerken zijn die waarin elke knoop een zogenaamd *perceptron* is (Rosenblatt, 1962): een eenvoudige verwerkingseenheid die een lineaire scheidingfunctie toepast op een reeks variabelen. Een stelsel van perceptrons is dus vergelijkbaar met een complex lineair model. Er zijn technieken die in staat zijn om de parameters van een perceptron-netwerk (de *gewichten*) te schatten, waarvan de belangrijkste *Error Backpropagation* is. Een veelgebruikte vorm van een perceptronstelsel is de 3-laags *feedforward* structuur: de uitvoerlaag bevat een knoop voor elke uitvoer-variabele en de zogenaamde 'verborgen laag' bestaat uit een vast aantal knopen die de uitvoerlaag verbindt met elke invoer-variabele, de invoerlaag. Elke verbinding is voorzien van een gewicht en gezamenlijk vormen deze gewichten de model-representatie. Als in de praktijk wordt gesproken over een *backpropagation* netwerk, dan wordt doorgaans deze structuur bedoeld. Als een dergelijk netwerk zou worden toegepast op het voorbeeldprobleem, dan zou het volgende model kunnen ontstaan (grijs wordt als 'J' geklassificeerd en wit als 'N'). (Zie figuur 2, p. 14).

Het trainingsalgoritme stelt zelfstandig het netwerk zo af dat een zo goed mogelijk model ontstaat. Het algoritme heeft als taak de gewichten tussen de lagen bij te stellen zodat de aangeboden invoerpatronen worden afgebeeld op de juiste waarden van de uitvoer-neuronen. Dit wordt bereikt door de gemaakte fout op het uitvoerniveau te herleiden tot fouten op lagere niveaus en op basis daarvan de gewichten tussen de lagen aan te passen. Backpropagation is dus een vorm van supervised leren. Het enige wat de



Figuur 2. Een backpropagation-netwerk voor het voorbeeldprobleem.

bouwer moet doen is voorbereken van de gegevens en kiezen voor een aantal knopen in de hidden layer. Het backpropagation netwerk is in verschillende varianten succesvol toegepast op vele terreinen en vormt het meest gebruikte neurale netwerk.

Backpropagation-netwerken hebben twee bottlenecks: training en datapreparatie. De training van een netwerk is in de regel een langdurig proces, waarbij tevens de mogelijkheid bestaat dat de training niet convergeert; het trainingsalgoritme is dan 'het spoor bijster' en is niet staat om de optimale eindtoestand te bereiken. De trainingstijd en de kans op non-convergentie lopen op met het aantal variabelen en de complexiteit van de verbanden. De tweede bottleneck vormt de datapreparatie: het voorbereken van gegevens zodat ze bruikbaar zijn voor het netwerk. Bijvoorbeeld: voor een categorie variabele (nominale variabele) moet voor elke mogelijke categorie een nieuwe variabele worden aangemaakt, waardoor het betrekken van 'grote nominalen' zoals *postcode* vrijwel onmogelijk wordt.

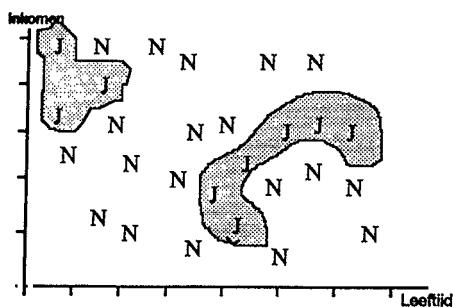
Een volgend punt is dat een maximaal presteren van het netwerk niet alleen afhankelijk is van keuzen met betrekking tot de representatie van de invoer maar ook van die voor de architectuur van het netwerk (aantal lagen en knopen) en de leerparameters (bijvoorbeeld de leersnelheid). Omdat er (nog) geen formele analyses van het leergedrag van een netwerk bij een bepaalde dataset mogelijk zijn, worden in de praktijk keuzes voor leer- en architectuurparameters gebaseerd op heuristieken en zijn deze moeilijk te automatiseren.

Prototype-gebaseerde netwerken

Naast het generen van een wiskundig model aan de hand van voorbeelden, is het ook mogelijk een model te representeren door middel van de voorbeelden zelf. Klassifikatie van een gegeven geval werkt dan als volgt: zoek het meest gelijkende voorbeeld en ken de klasse toe die dat voorbeeld heeft. Het is echter niet triviaal wat 'meest gelijkend' inhoudt; de overeenkomst voor bepaalde variabelen is bijvoorbeeld belangrijker dan die voor andere. Daarnaast mag er niet op worden vertrouwd dat elk voorbeeld prototypisch is, zeker niet als het gaat om marketinggegevens. Neurale netwerken die zelfstandig voorbeelden vertalen in prototypes en op basis daarvan kunnen klassificeren worden prototype-gebaseerde netwerken genoemd.

Een groot voordeel van prototype-gebaseerde netwerken is dat training eenvoudig is. De trainingsgegevens moeten worden vertaald in een aantal prototypische gevallen; er hoeft geen mathematisch model te worden opgesteld. Omdat er geen abstraktie wordt gemaakt is de nauwkeurigheid van prototype-gebaseerde netwerken hoger dan die van bijvoorbeeld backpropagation netwerken. Een nadeel is dat een dergelijk netwerk traag kan worden, vooral als het aantal prototypes groot is. Echter, door algoritmische verbeteringen en de toenemende opslag- en rekenkapaciteit van de betaalbare moderne computer, komen dergelijke systemen steeds meer binnen handbereik.

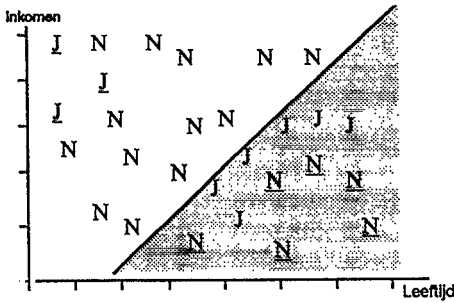
Er zijn diverse prototype-gebaseerde technieken ontwikkeld. Kohonen (1982) introduceerde de *Selforganizing feature map*, waarin de prototypes worden afgebeeld op een klein aantal (vaak twee) dimensies. ART staat voor *Adaptive Resonance Theory* en is ontwikkeld door Grossberg (Grossberg, 1976). ART is een tweelaags netwerk dat wordt getraind door een algoritme genaamd *competitive learning*. Indien nodig, past een ART-netwerk automatisch zijn structuur aan. Binnen Sentient Machine Research is het *Resonant Field Computing* paradigma ontwikkeld (den Uyl, 1986). Met dergelijke netwerken worden prototypes gegenereerd door gebruik te maken van fragmentarische opslag van voorbeelden. RFC is geïmplementeerd in een module onder de naam *Data-Detective*, welke tevens in staat is een uitspraak te onderbouwen door de bijbehorende oorspronkelijke voorbeelden te tonen. Voor een illustratie, zie figuur 3.



Figuur 3. Afbeelding van een RFC model.

3.3 Neurale netwerken en statistiek

De statistiek omvat methodes om te toetsen of bepaalde typen modellen opgaan voor een gegevensset, alsmede technieken om de onbekende parameters van een model te schatten. Een modelleringstrajekt houdt in: 1) aannames maken over de variabelen in het domein, 2) aannames toetsen, 3) een bepaald type model kiezen dat daarmee overeenstemt en de verlangde taak vervult en 4) de onbekende parameters schatten. De bekendste technieken voor parameter-schatting zijn regressie-analyse en de klassifikatie-variant van regressie: diskriminant-analyse. Een standaard diskriminant-analyse op het maachineprobleem zou het volgend model op kunnen leveren, waarbij geldt dat hoe verder een geval van de scheidingslijn af ligt, des te sterker de uitspraak is dat het tot een bepaalde klasse behoort:



Figuur 4. Resultaat van diskriminant-analyse.

In bovenstaande figuur worden alle gevallen in het gearceerde gebied als kopers geklassificeerd. De foute klassifikaties zijn onderstreept. De lage score illustreert dat een enkele diskriminant-functie niet in staat is om complexe en non-lineaire verbanden vast te leggen. Het probleem met de non-lineariteit houdt in dat een 'krom' scheidingsvlak alleen maar gemodelleerd kan worden als de statisticus van te voren weet wat het karakter van het scheidingsvlak is. Er moet dus een aanname worden gemaakt over de aard van de verbanden. Het probleem met de complexiteit is dat een diskriminant-functie één enkel gebied isoleert, terwijl in werkelijkheid bijvoorbeeld maaimachinekopers verschillende profielen hebben. Ook hiervoor dient de statisticus aannames te maken die door statistische toetsing moeten worden onderbouwd, waarna het model zal worden gevormd door een stelsel van diskriminant-functies.

Standaard statistische technieken kunnen niet tot de lerende technieken gerekend worden, aangezien de statisticus in feite het model opstelt, en de technieken enkel de ontbrekende parameters schatten: de technieken zijn niet zelfstandig. Aangezien gedragspatronen binnen een markt in de regel complex en dynamisch zijn, geldt dat bij toepassing van statistische technieken voor marketing, de kennis en inzet van de statisticus een grote bottleneck kan vormen. Omdat het voor een statisticus moeilijk kan zijn volledig inzicht te krijgen in de gedragspatronen, bestaat tevens het gevaar dat verkeerde inschattingen worden gemaakt. Daarnaast geldt dat statistische technieken slecht om kunnen gaan met ontbrekende waarden en weinig bestand zijn tegen vervuilde gegevens (ruis). Vooral bij marketing-problemen is er een behoefte aan technieken die vrijwel autonoom een model op kunnen stellen en redelijk bestand zijn tegen ruis en ontbrekende waarden.

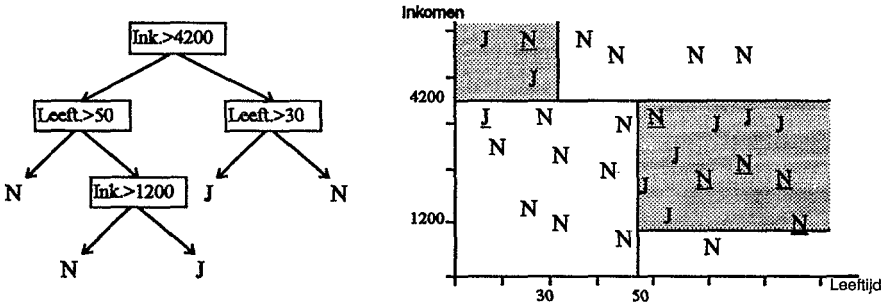
Wat betreft de zelfstandigheid van statistische technieken zijn de laatste ontwikkelingen binnen de non-parametrische statistiek hoopgevend. Een voorbeeld van deze ontwikkelingen wordt gevormd door de non-parametrische adaptieve regressiemethoden (Denteneer, 1993). Non-parametrische regressietechnieken onderscheiden zich door het ontbreken van aannames over de vorm van de regressiefunctie. Ze vertonen daarmee overeenkomst met neurale netwerken.

3.4. Neurale netwerken en beslisbomen

Het automatisch afleiden van kennisregels uit voorbeelden staat binnen de Kunstmatige Intelligentie bekend als *machine learning*. Een groot voordeel van regel-representatie is

het grote expressie-vermogen en de leesbaarheid die om twee reden goed is: *als-dan* regels zijn eenvoudig te begrijpen en er wordt geredeneerd met hanteerbare begrippen: "ALS klanttype=prospect EN provincie=Gelderland DAN mail=ja".

De bekendste machine learning technieken hebben betrekking op het afleiden van beslisbomen: NPPA, AQ en ID3 (Quinlan,1986). Een beslisboom is een boomstructuur die op elk niveau een opsplitsing maakt van één variabele in verschillende takken. Als een bepaald geval geklassificeerd moet worden dan moet per niveau worden gekeken in welke opsplitsing de waarde van de variabele in kwestie valt, waarna via die tak zich een nieuwe opsplitsing aandient. Een beslisboom-model voor het voorbeeldprobleem, zou er als volgt uit kunnen zien:



Figuur 5. Beslisboom-model en de bijbehorende afbeelding.

In de figuur is te zien dat de scheidingsvlakken recht zijn, hetgeen een groot nadeel is van beslisbomen: grillige verbanden, zoals die vaak voorkomen in marktgedrag, kunnen moeilijk worden gemodelleerd. Daarnaast kunnen beslisboom-generatie technieken slecht omgaan met ruis en ontbrekende waarden, typisch twee problemen die zich in marketing voordoen.

Als voordeel van beslisbomen wordt vaak genoemd dat de leesbaarheid goed is, hetgeen in de praktijk echter tegenvalt. ALS-DAN regels zijn op zichzelf te begrijpen, maar het totale model wordt niet opgebouwd vanuit een menselijk perspectief, hetgeen bizarre beslisstructuren en enorme bomen tot gevolg kan hebben. Daarnaast kunnen verbanden in de praktijk zo subtiel zijn en zich op zo'n laag niveau bevinden dat een toegankelijk representatie zowiezo onmogelijk is. De toegankelijkheid van een model is in dat geval alleen goed indien de complexiteit ervan beperkt wordt, hetgeen ten koste gaat van de nauwkeurigheid.

Een ander nadeel van regelgebaseerde systemen in het algemeen is dat het gedrag nogal gevoelig is voor kleine afwijkingen. De uitkomst van een konditie in een regel is tweewaardig (goed of fout) dus is een regelstructuur dat ook, met als gevolg dat bijvoorbeeld iemand met een inkomen van 4201,- als een absolute klant wordt gezien en iemand met een inkomen van 4199,- als een absolute non-klant. In de vakliteratuur wordt dit verschijnsel aangeduid met brittleness (broosheid): het systeem kan niet tegen een stootje (lees: ruis).

4. TOEPASSINGSMOGELIJKHEDEN VAN NEURALE NETWERKEN IN MARKETING

Zoals hiervoor is aangegeven, lenen neurale netwerken zich vooral voor het zoeken van structuren in grote databestanden met het oog op klassifikatie of voorspelling. In marketing beschikken we veelvuldig over dergelijke grote databestanden: verkoop- en marktaandeelgegevens voor de merken in een produktcategorie verzameld bij de detailhandel (tegenwoordig beschikbaar op weekbasis); de gegevensbestanden zoals die via consumentenpanels worden verzameld: steekproeven van duizenden consumenten, die rapporteren over hun aankopen met betrekking tot een groot scala aan produkten; in de sfeer van direct marketing: zeer grote bestanden met kenmerken van consumenten en hun koopgedrag ten aanzien van diverse produkten en diensten. Ook in het media-onderzoek hebben we vaak met zeer grote steekproeven te maken, terwijl in de business-to-business door het gebruik van verkoopsupportsystemen bedrijven gemakkelijk grote databestanden kunnen opbouwen over klanten en prospects en hun koopgedrag.

Tot nu toe werden deze data geanalyseerd met de eerdergenoemde min of meer conventionele statistische technieken zoals multiële regressie, discriminant-analyse, faktor-analyse, cluster-analyse, enz. Vaak moeten hierbij veronderstellingen worden gemaakt over de variabelen die twijfelachtig zijn (bijvoorbeeld normaal verdeeld, interval schaal gemeten, lineaire relaties, onderling onafhankelijk, enz.). Ook treden vaak complexe interacties tussen de variabelen op.

Al deze beperkingen zijn veel minder ernstig bij neurale netwerken. Een neurale netwerk is vooral geschikt voor voorspellings- en selektiedoeleinden; datgene wat geleerd is bij een groot aantal trainingscases, bijvoorbeeld over de samenhang tussen persoonskenmerken en het wel of niet bezitten van een bepaald produkt, bijvoorbeeld een Mercedes, kan worden gebruikt om voor een nieuwe case te voorspellen of de betreffende persoon wel of niet een potentiële klant is voor een Mercedes.

Daar staat tegenover dat een neurale netwerk weinig inzicht geeft in de vraag welke variabelen nu het meest verklaren van de variantie in een bepaalde variabele. Statistische theorieën waarmee kan worden bepaald of bepaalde variabelen in het model wel of niet significant zijn, ontbreken (vooralsnog) bij neurale netwerken.

Toch zijn de voorspellingsmogelijkheden die neurale netwerken bieden al voldoende reden om hun bruikbaarheid in marketing te onderzoeken.

Diverse voorbeelden van toepassingen van neurale netwerken in marketing kunnen worden genoemd.

Het segmenteren van groepen van consumenten op basis van gemeenschappelijke eigenschappen is van belangrijke strategische waarde in de marketing. Neurale netwerken kunnen omgaan met de segmentatie problematiek en beschikken zelfs over kwaliteiten om een a priori en een a posteriori benadering te combineren in een enkel model (Mazanec, 1993).

De effectiviteit van de marketing kan tot uitdrukking gebracht worden in de ROI per geïnvesteerde gulden en de tevredenheid van de consument. Het is daarom belangrijk

om consumenten te kunnen onderscheiden met een relatief hoge verwachtingswaarde van consumenten met een relatief lage verwachtingswaarde. Het budget kan dan effectiever worden aangewend. Ook voor deze taak kunnen neurale netwerken worden ingezet (Wurtz & Curry, 1993).

In de direct marketing is het selekteren van consumenten met een hoge kans op een positieve reactie op een mailing een belangrijke taak. Dit kan gebeuren aan de hand van gegevens over de manier waarop men in het verleden op vergelijkbare mailings heeft gereageerd (Garavaglia, 1993).

Voor het voorspellen van kijkcijfers kunnen eveneens neurale netwerken gebruikt worden. Deze voorspellingen zijn van belang voor optimalisatie van uitzendtijden van reclames (Fitzsimons, Khabaza & Shearer, 1993) (Boyero, Triguero Ruiz, 1993).

Een minder voor de hand liggende toepassing is wellicht het bepalen van consumenten-percepties. Hierbij wordt getracht grip te krijgen op het besluitvormingsproces van de consument (Moutinho, Curry, 1993).

Een andere selectietaak is het selekteren van sales promotie instrumenten. Dit kan gebeuren met behulp van scannerdata of op basis van expertkennis omtrent de mate waarin verschillende sales promotion instrumenten verondersteld worden aan te sluiten bij marketing doelstellingen (Kluytmans, Wierenga & Spigt, 1993).

Neurale netwerken, evenals andere methoden uit de kunstmatige intelligentie, worden soms met veel aplomb gepresenteerd als de oplossing voor alle problemen. Het lijkt verstandiger om op een nuchtere wijze te onderzoeken wat neurale netwerken wel en niet kunnen binnen marketing en waar hun belangrijkste bijdrage ligt.

Een groot project in dit verband in Nederland is een door het Direct Marketing Instituut Nederland (DMIN) geïnitieerd onderzoek², waarin diverse bedrijven met grote databestanden deelnemen om de toepasbaarheid van neurale netwerken (en andere adaptieve technieken) in de praktijk te onderzoeken. Hier gaat het vooral om zogenaamde cross-sectie-gegevens. Daarnaast is het interessant om na te gaan wat neurale netwerken te bieden hebben bij het andere type data dat we in marketing vaak tegenkomen: zogenaamde tijdreeksdata. Exploratief onderzoek op dit terrein komt in de volgende paragraaf aan de orde.

5. NEURALE NETWERKEN IN TIJDREEKS-ONDERZOEK

Het onderzoek naar de mogelijkheden van neurale netwerken bij tijdreeks-onderzoek begint op gang te komen. Een voorbeeld is het artikel van Hruschka (1991) die met 60 (maand) gegevens werkte van een merkartikel op de Oostenrijkse consumentenmarkt en de aanpassing van een neuraal netwerk aan deze gegevens vergeleek met die van een regressiemodel. Zijn conclusie is: 'Die Netzwerke führen zu deutlich besseren Ergebnissen'.

In een exploratief onderzoek naar de mogelijkheden van neurale netwerken hebben wij voor een markt van niet duurzame consumentenprodukten in een Westeuropees land nagegaan hoe goed de marktaandeelen van de belangrijkste merken kunnen worden voorspeld met een neuraal netwerk in vergelijking met de (traditionele) multi-pele regressiemethode.

In deze markt werden de grootste vijf merken in beschouwing genomen. Voor het gemak beperken we verder de markt tot de verkopen van deze vijf merken samen.

Voor de verklaring van het marktaandeel van een merk in een bepaalde periode werden de volgende variabelen gehanteerd:

- relatieve prijs;
- reclame-aandeel;
- gewogen distributie;
- gewogen out-of-stocks;
- trend.

Bij het multiple regressiemodel werd het marktaandeel door middel van een lineair model uit deze variabelen verklaard. Dezelfde variabelen fungeerden ook als input voor het neurale netwerk; de output variabele was daar uiteraard het marktaandeel.

De beschikbare gegevens waren tweemaandelijks Nielsen data vanaf de zesde periode dat 1984 tot en met de tweede periode van 1993 (totaal 51 datapunten). Het schatten van het regressiemodel, respectievelijk het trainen van het neurale netwerk geschiedde aan de hand van de gegevens van 1984 tot en met 1991 (43 datapunten). Vervolgens werd de voorspelkracht van beide methoden getest op de (verse) gegevens van begin 1992 tot en met de tweede periode van 1993.

Tabel 1. Gemiddelde voorspelfout neurale netwerk versus multiple regressie voor 5 merken over 8 perioden (1992 t/m 2e periode 1993).

	Neuraal Netwerk	Multipole Regressie
Merk A	0.96	1.08
Merk B	0.68	0.75
Merk C	1.02	2.88
Merk D	1.05	1.74
Merk E	2.10	2.63
Gemiddeld	1.16	1.82

Tabel 1 geeft de gemiddelde voorspelfout over deze acht perioden voor de vijf merken, voor het neurale netwerk en de multiple regressie.

Het blijkt dat voor alle vijf merken het neurale netwerk gemiddeld beter voorspelt dan het regressiemodel. De betere prestatie van het netwerk is ook statistisch significant ($p < 0.005$): in 31 van de 40 individuele cases ligt de voorspelling van het netwerk dicht bij het werkelijke marktaandeel dan die van het regressiemodel.

Opvallend is de relatief goede prestatie van het netwerk ten opzichte van de regressie bij merk C. Dit is een sterk merk, dat in de periode 1992/93 relatief in prijs omhoog is gegaan ten opzichte van de concurrenten. Het regressiemodel voorspelt voor deze testperiode in tegenstelling tot het neurale netwerk systematisch een te laag marktaandeel voor merk C.

Blijkbaar kent het regressiemodel een te groot gewicht toe aan de prijsstijging, terwijl het netwerk beter de sterke overall positie van merk C uit het verleden heeft opgepikt.

Uiteraard moeten we ons realiseren dat deze vergelijking van neurale netwerk en regressie slechts één produktklasse betreft. Toch is het opvallend dat blijkbaar met een betrekkelijk klein aantal verklarende variabelen het neurale netwerk al tot opvallend betere voorspellingen komt dan het regressiemodel (de gemiddelde fout bij regressie was 56% groter dan bij het neurale netwerk).

Tijdreeksgegevens zullen in de toekomst steeds rijker worden: gegevens over meer variabelen en met hogere frekwentie. Dit lijkt alleen maar verder in het voordeel te werken van neurale netwerken. Meer empirisch onderzoek met gegevens uit diverse markten is echter nodig voordat definitievere uitspraken kunnen worden gedaan.

Wierenga en Kluytmans (1994) hebben een simulatiestudie uitgevoerd om de prestaties van een neurale netwerk in een tijdreeksituatie te onderzoeken onder verschillende omstandigheden. Eén van de beschouwde factoren is de modelspecificatie (lineair tegenover exponentieel). Hun conclusie is dat, zoals verwacht, bij een lineair verband in de data tussen marketing instrumenten en verkopen het neurale netwerk het nauwlijks (maar wel significant) beter doet dan regressie. Bij een niet lineair model presteert het neurale netwerk echter duidelijk beter, ook als we het regressiemodel terugschatten met een log-transformatie. Bij 'echte' data hebben we uiteraard de informatie over de juiste terugtransformatie niet. Ook dit resultaat geeft dus aan dat er alle reden is om goed naar neurale netwerken te kijken als analyse- en vooral voorspelmethode in marketing.

6. TOOLS VOOR NEURALE NETWERKEN

Op basis van een gedefinieerde taak, zoals bijvoorbeeld mailingselectie of omzetvoorspelling, kan men een model gaan bouwen. Allereerst dienden de gegevens dan te worden geselecteerd en geprepareerd, waarna het model dient te worden getraind en geëvalueerd.

Dataselectie en -preparatie: Als men een verzameling variabelen heeft, dan kan het wenselijk zijn een aantal variabelen weg te laten en/of een aantal variabelen te combineren om zo een meer relevante indikator over te houden met meer voorspellingskracht. Dit proces kan gebaseerd zijn op marktkennis, maar ook kan het zijn dat sommige gegevens op zichzelf te weinig voorkomen om als prediktor gebruikt te worden. Daarom is het belangrijk dat voordat er aan modelbouw begonnen wordt er eerst een beperkte statistische analyse van de steekproef gedaan wordt. Verder stellen bepaalde technieken eisen aan de wijze waarop de data wordt aangeboden. Op basis van kennis, statistiek en technieken worden de gegevens verder geselecteerd en bewerkt. De steekproef moet ook onderverdeeld worden in een aantal gelijkwaardige sets om zo modelontwikkeling en validatie gescheiden te houden.

Training en evaluatie: In de meest gebruikelijke vorm wordt het model geëvalueerd op een aparte validatieset, die meestal een (disjunkte) steekproef is uit de gehele set van

gegevens. Soms wordt hierbij nog aanvullende informatie gegeven over de betrouwbaarheid van het model op basis van statistische kengetallen.

De kosten voor een systeem lopen uiteen, van een paar duizend gulden voor PC-gebaseerde algemene modeleringsprogramma's, tot f 100.000,- of meer voor gespecialiseerde marketingsystemen, inclusief uitgebreide consultancy. Belangrijk is dat daarbij een zeer uiteenlopend niveau van expertise van de gebruikers verwacht wordt. Voor gebruikers zonder analytische achtergrond is het niet eenvoudig een operationeel model te ontwikkelen met deze pakketten. De fabrikanten bieden weinig ondersteuning in de vorm van gebruikersopleiding of consultancy. Dit kan grote consequenties hebben voor de additionele investeringen die nodig zijn in cursussen, ervaringsopbouw en gebruik als operationeel model.

Een volledig overzicht van tools zou meer dan 8 pagina's omvatten en alleen ruimte laten voor een beknopte omschrijving van de functionaliteit (zie bijvoorbeeld het overzicht in AI Expert, Februari 1993). Er is hier gekozen voor een opzet die niet pretendeert volledig te zijn maar inzicht probeert te verschaffen in de mogelijkheden die momenteel op de markt worden geboden.

NeuroShell: van Ward Systems Group, prijsindicatie f 1000,-, werkt op een PC onder Windows of DOS. NeuroShell bevat een 5-tal verschillende netwerk-technieken waaronder backpropagation, recurrent backpropagation, kohonen en een tweetal 'statistische' netwerken (PPN en GRNN). Eenmaal ontwikkelde neurale netwerk modellen kunnen worden toegepast in een andere software omgeving.

Brainmaker: van California Scientific Software, prijsindicatie f 2500,-, werkt op een PC onder Windows of DOS. Een van de snelste PC-software-simulators voor backpropagation dat tevens het enige type netwerk is dat het pakket bevat. Er zijn veel mogelijkheden om het model te optimaliseren. De mogelijkheden voor dataselectie en datapreparatie zijn echter beperkt en niet 100% betrouwbaar. Ook de mogelijkheden voor modevaluatie zijn beperkt en eigen rapportages produceren is onmogelijk.

NeuralWorks Profesional II / Plus: van NeuralWare, prijsindicatie f 5000,- (PC-versie), werkt op PC's (onder DOS, echter wel grafisch georiënteerd), Macintosh, VAX en diverse UNIX workstations. NeuralWorks is het pakket met de meest uitgebreide selectie aan netwerk-technieken en ook een van de oudste aanbieders op de markt. De gebruiker heeft zeer veel mogelijkheden om tot in detail de modelvorming te sturen. Er zijn ook een aantal aanvullende modules om hier als gebruiker volledig vat op te krijgen (tegen een aanzienlijke meerprijs). Het pakket is om deze redenen veel gezien op universiteiten en onderzoekslaboratoria. Dataselectie is redelijk rudimentair.

4Thought: van Livingstones B.V. Strategic Business Systems, prijsindicatie f 18.000,-, voor PC's onder Windows. Als techniek wordt een variant van backpropagation gebruikt. Er zijn mogelijkheden voor eigen rapportages. Het systeem is duidelijk marketing gericht. Er is een mogelijkheid om het netwerkmodel te exporteren.

DataDetective is een intern produkt van SMR. Het is PC-gebaseerd met een Windows-interface (en mogelijkheden voor een Dos-interface). Het is een gereedschap dat speciaal ontwikkeld is om SMR's eigen techniek, Resonant Field Computing (RFC), opti-

maal te kunnen benutten. Het is op dit moment een gereedschap dat custom-made wordt toegepast binnen applicaties zoals Customer Marketing Rating Software, Talent-View en BlindDate. Het wordt ook projectgewijs toegepast bij de uitvoering van DM-taken zoals bijvoorbeeld mailingacties en doelgroepselecties.

7. CONCLUSIES

In het voorgaande is geprobeerd de lezer een beeld te geven van neurale netwerken en hun (potentiële) toepassingsmogelijkheden in marketing.

Neurale netwerken stellen zelfstandig een model op van een gedrag, op basis van voorbeelden van dat gedrag. Standaard statistiek vergt veel meer expertise en tijd bij de modelvorming en kan moeilijker omgaan met complexe non-lineaire verbanden in de data. Kennisbomen leveren enigszins meer leesbare modellen op die echter niet bestand zijn tegen vervuilde en inkomplete gegevens, die typisch zijn in database marketing.

De uitkomsten van het onderzoek tot nu toe wijzen in de richting dat neurale netwerken voor marketing beloften lijken te bieden. Zowel voor cross-sektie- als tijdreeksdata wordt dit momenteel verder onderzocht.

In marketing wordt veelvuldig over grote databestanden beschikt die zich goed lenen voor neurale netwerk-toepassingen. Hopelijk zal dit artikel een aantal lezers stimuleren hiermee aan de gang te gaan.

NOTEN

1. Een gedeelte van dit artikel is gebaseerd op het rapport 'Adaptieve Data-Analyse in Direct Marketing' dat door de eerste auteur (samen met collega's van SMR) is uitgebracht aan het Direct Marketing Instituut Nederland (1994).
2. Dit onderzoek wordt uitgevoerd door SMR.

LITERATUUR

- Anderson J.A., Pellionisz A., and Rosenfeld E., 1990. Introduction. *NeuroComputing 2*. MIT Press, London.
- Bailey, D. & D. Thompson, "How to Develop Neural Network Applications", in *AI-Expert*, June 1990, pp. 38-45.
- Bastaens, D.E. & W.M. van den Bergh, "Analyse van Rendementen op de Aandelenindex met behulp van Neurale Netwerken", in *VBA Journaal*, vol.8, september 1992, pp. 30-38.
- Beale, R. & T. Jackson, *Neural computing: An Introduction*, 1990, Bristol, New York, Adam Hilger, pp. 100.
- Boyer, E.P. & F.A. Triguero Ruiz, "A Method to Predict the Television Audience Based on Neural Networks", in *Proceedings of the Second International Conference on AI Applications on Wall Street*, April 1993, New York, pp. 174-182.
- Collins, E., S. Ghosh & C. Scofield, 1989, *An Application of a Multiple Neural Network Learning System to Emulation of Mortgage Underwriting Judgements*, Nestor Inc., Providence, RI.
- Denteneer, D., 1993, Niet-parametrische regressie: neuraal of statistisch? *Kennis en Methoden 44*, 21-34.
- Duives, P.C., Huysman, S.A., Mol, M.J. & J.K. Vegter, 1991, Neurale Netwerken in Beurshandel en Koersvoorspellingen. *Journal of Software Research*, 3, nummer 2.

- Epping, W.J.M. & G. Nitters, "A Neural Network for Analysis and Improvement of Gas Well Production", in *Proceedings of the 1990 Summer Computer Simulation Conference*, July 1990, Society for Computer Simulation, San Diego, pp. 791-796.
- Fitzsimons, M., T. Khabaza & C. Shearer, "The Applications of Rule Induction and Neural Networks for Television Audience Prediction", in *Proceedings ESOMAR-EMAC-AFM Symposium on 'Information Based Decision Making in Marketing'*, November 1993, Paris, pp. 69-82.
- Garavaglia, S., "Direct Mail Response Modeling Using a Counterpropagation Neural Network", in *Proceedings of the Second International Conference on AI Applications on Wall Street*, April 1993, New York, pp. 165-173.
- Grossberg, S., 1976, Adaptive pattern classification and universal recoding: I. Parallel development and coding of neural feature detectors. *Biological Cybernetics* 23, 121-134.
- Hoptroff, R.G., "The Principles and Practice of Time Series Forecasting and Business Modelling using Neural Nets", in *Neural Computing Applications*, vol.1, 1992.
- Hruschka, H., "Einsatz kunstlicher neuronaler Netzwerke zur Datenanalyse im Marketing", in *Zeitschrift fur Forschung und Entwicklung*, vol.4, 1991 IV, pp. 217-225.
- Kluytmans, J., B. Wierenga & M. Spigt, "Developing a Neural Network for the Selection of Sales Promotion Instruments in Different Marketing Situations", in *Proceedings of the Second International Conference on AI Applications on Wall Street*, April 1993, New York, pp. 160-164.
- Kohonen, T. 1982, "Self-organized formation of topologically correct feature maps. *Biological Cybernetics*, 43, pp. 59-69. In: *Neural Computing I*, Anderson et al., 1990, MIT Press: Cambridge.
- Mazanec, J.A., "A priori and a posteriori Segmentation: Heading for Unification with Neural Network Modeling", in *Proceedings of the 22nd Annual Conference of the European Marketing Academy*, May 1993, Barcelona, pp. 789-817.
- Moutinho, L. & B. Curry, "Consumer Perceptions of ATMs: an Application of Neural Networks", in *Proceedings of the 22nd Annual Conference of the European Marketing Academy*, May 1993, Barcelona, pp. 975-1000.
- Murre, J.M.J., "Leren in Neurale Netwerken", in *Informatie*, Jaargang 33, nr.6, pp. 376-388.
- Quinlan, J.R., 1986, Induction of Decision Trees. *Machine Learning* 1, pp. 81-107.
- Refenez, A.N., M. Azena-Barac, L. Chen & S.A. Karoussos, "Currency Exchange Rate Prediction and Neural Network Design Strategies", in *Neural Computing & Applications*, 1993, pp. 46-58.
- Rosenblatt, F., 1962, The perceptron: a probabilistic model for information storage in the brain. *Psychological review*, 65: 386-408.
- Rumelhart, D.E. & J.L. McClelland, *Parallel distributed Processing: Explorations in the Microstructure of Cognition*, 1986, Cambridge, Mass., MIT Press.
- Sejnowski, T.J. & C.R. Rosenberg, "Parallel networks that learn to pronounce English text", in *Complex Systems*, 1987, pp. 145-168.
- Simpson P.K., 1989. *Artificial Neural Systems*. Pergamon Press, San Diego.
- Uyl, M.J. den, "Representing Magnitude by Memory Resonance" in *Proceedings of the 6th Annual Conference of the Cognitive Science Society*, 1986, pp. 63-71.
- Verdenius, F., "Neurale Netwerken Toepassen in Kwaliteit", in *Sigma*, Jaargang 93, nr.3, pp. 21-27.
- Wierenga, B. & J. Kluytmans, "Neural Nets versus Marketing Models in time Series analysis: a Simulation study", in *Proceedings of the 23rd EMAC conference*, Maastricht, mei 1994, pp. 1139-1153.
- Wurtz, W. & J. Curry, "Customer Marketing Ratings", in *Conference Workbook on the 10th National Conference/Exposition on Database Marketing*, pp. 204-215.