

Artificiella neurala nätverk

En undersökning av neurala nätverk
som tillämpning för den finansiella marknaden

J A C K H A



**KTH Datavetenskap
och kommunikation**

Artificiella neurala nätverk

En undersökning av neurala nätverk
som tillämpning för den finansiella marknaden

J A C K H A

Examensarbete i datalogi om 15 högskolepoäng
vid Programmet för datateknik
Kungliga Tekniska Högskolan år 2010
Handledare på CSC var Henrik Eriksson
Examinator var Mads Dam

URL: [www.csc.kth.se/utbildning/kandidatexjobb/datateknik/2010/
ha_jack_K10028.pdf](http://www.csc.kth.se/utbildning/kandidatexjobb/datateknik/2010/ha_jack_K10028.pdf)

Kungliga tekniska högskolan
Skolan för datavetenskap och kommunikation

KTH CSC
100 44 Stockholm

URL: www.kth.se/csc

SAMMANFATTNING

Den finansiella marknaden präglas ofta av volatilitet som många spekulanter gärna vill utnyttja för att tjäna pengar på. Priset på aktier, finansiella instrument eller råvaror kan gå upp och ned under perioder och frågan många ställer sig är: finns det några mönster? I dagens finansiella marknaden används många olika handelssystem flitigt av aktörer såsom olika institutioner och företag. Den största utmaningen ligger i att kunna förutspå hur priset på de finansiella tillgångarna ändras. Genom lyckade prediktioner kan du som aktör på den finansiella marknaden generera stora vinster men misslyckade prediktioner kan dock leda till stora förluster. Det huvudsakliga syftet med denna uppsats är att ge en beskrivande bild över området aktiehandel med artificiella neurala nätverk som prediktionsmodell. Samtidigt ska vi med en analyserande inställning studera några valda arbeten där man implementerat artificiella neurala nätverk för att predicera aktiemarknaden.

Denna uppsats kan delas in i tre delar där den första delen är menad till att på ett enkelt och pedagogiskt förfaringssätt leda läsaren in i ämnet där problemets bakgrund beskrivs. Den andra delen beskriver artificiella neurala nätverk i mycket stora drag där meningen är att ge läsaren en enkel bild av vad artificiella neurala nätverk är. Den sista delen skall vi studera två masteruppsatser som behandlar artificiella neurala nätverk där empiriska studier har utförts. Resultaten av de empiriska studierna är av blandad karaktär vilket tillsammans med egna studier och logisk förfaringssätt leder till slutsatsen att det finns mycket stor potential i artificiella neurala nätverk som prediktionsmodell för den finansiella marknaden men fördelarna urholkas ifall andra aktörer har kännedom av din teknik.

ABSTRACT

The financial markets are often characterized by volatility that many speculators are willingly to exploit to make money. The price of shares, securities and commodities can go up and down at times and the question many ask themselves is: are there any patterns? In today's financial market different trading systems are frequently used by various actors such as institutions and companies. The biggest challenge is to predict how the price of financial assets changes. You, as an actor in the financial market can generate large profits from successful predictions, but it can also lead to huge losses if the outcome of the predictions would be wrong. The main purpose of this paper is to provide a descriptive picture of the area of equity trading with artificial neural network as a prediction model. In addition to that, we intend to make an analytical approach to study two selected works, which have made empirical research about artificial neural networks for prediction of the stock markets.

This paper can be divided into three parts, where the first part is meant to be a simple and pedagogical guide for the reader into the subject where the problem context is described. The second part describes the artificial neural network in very broad terms, in which the purpose is to give the reader a humble picture of what artificial neural networks is. The last section, we study two masters theses dealing with artificial neural networks, in which empirical studies have been applied. The results of the empirical studies are of mixed character which, together with our own studies and logical procedure, leads to the conclusion that there is great potential in using artificial neural network as prediction model for the financial market but leaving the idea undermined if other actors know your technique.

INNEHÅLLSFÖRTECKNING

1. Introduktion.....	5
1.1 Inledning.....	5
1.2 Bakgrund.....	5
1.3 Problemdiskussion.....	6
1.4 Problemformulering.....	7
1.5 Olika metoder för att predicera aktiemarknaden.....	7
2. Artificiella neurala nätverk.....	8
2.1 Inledning.....	8
2.2 Matematisk beskrivning av ANN.....	10
2.3 ANN modeller.....	12
2.4 Inlärning.....	13
2.5 Fördelar med artificiella neurala nätverk.....	13
2.6 Kritik och nackdelar med artificiella neurala nätverk.....	13
3. Studerade artiklar.....	14
3.1 Stock prediction – A Neural Network Approach.....	14
3.2 Using Neural Networks and Genetic Algorithms to predict stock market returns.....	15
4. Slutsats och avslutande diskussion.....	17
4.1 Artificiella neurala nätverk – är det lönsamt?.....	17
5. Referenser.....	18
5.1 Referenslista.....	18

1. INTRODUKTION

1.1 INLEDNING

Priset på en aktie skall i tanken och teorin spegla den förväntning marknaden har på det underliggande företags framtida kassaflöde. Inom den finansiella skolan finns det ett flertal olika metoder för att värdera ett företag (aktiepriset). Några exempel är utdelningsbaserade-, kassaflödesbaserade och tillgångsbaserade modeller.[1] Men denna studie ska utgå ifrån en förenklad modell: tillgång och efterfrågan.

Det som styr ett börsnoterat företags värde (aktiepriset) är tillgången och efterfrågan. Detta kan förstås genom ett deduktivt förhållningssätt till problemet. De myriader av faktorer som styr tillgången och efterfrågan är dock svårare att identifiera. Exempel på faktorer som direkt och indirekt påverkar efterfrågan är:

- Nyhetsflöde som direkt eller indirekt rör företaget t.ex. uttalanden och rapporter eller händelser som rör den bransch/marknad som företaget opererar inom
- Konjunkturcykler och räntor som centralbankerna påverkar
- Makroekonomiska faktorer
- Politiska faktorer, geopolitiska faktorer
- Klimat och naturfenomen
- Teknikens utveckling
- Samhällstrender
- Psykologiska faktorer, rykten och spekulationer

De data som den här uppsatsen arbetar med är både historiska och aktuella som framför allt är lättillgängliga. Ett krav på att dessa data går att transformera till numerisk data kommer att ställas, då det är den form av inputs vi kan arbeta på. Det handlar om att finna trender och mönster i det aktuella klimatet, främst baserade på empiriska och kvantitativa metoder för att sedan utnyttjas för prediktion av framtida händelser. Möjligheten att göra detta bestrids av den så kallade "random walk"-hypotesen. Det är en finansiell teori som säger att såväl enskilda aktiers pris som börsens rörelser är omöjliga att förutspå och speglar en slumpmässighet.[2] Ekonomer generellt har historiskt accepterat denna hypotes men det finns studier som anses ha motbevisat den, i varje fall till en viss grad.[3] För att vår studie skall vara av något värde utgår vi istället från "non random walk"-hypotesen. Det fundamentala antagande som uppsatsen kommer att bygga på är att marknaden går att förutspå, åtminstone till en viss grad.

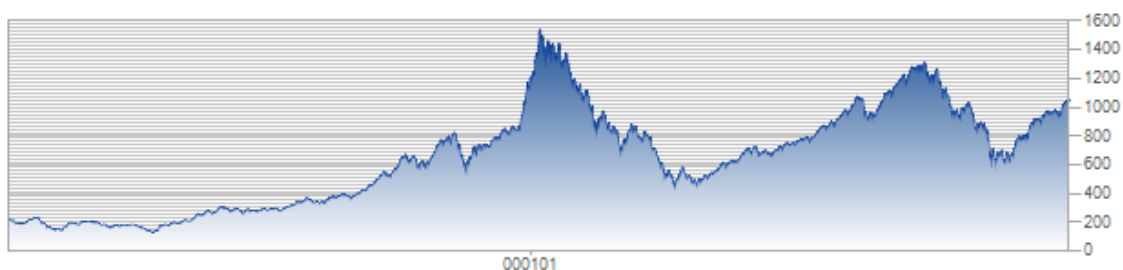


Fig 1. OMX Stockholm 30 Index från 1990 till 2010. Finns det några mönster?

1.2 BAKGRUND

Den finansiella marknaden fyller en viktig samhällsfunktion där du som investerare erbjuds många investeringsmöjligheter på ett smidigt sätt. Företagen kan enkelt ta in nya likvida medel av potentiella investerare direkt från börsen. På senare tid (åren 1990 – 2010) har teknikens genombrott öppnat de finansiella marknaderna även för privata investerare. Handeln har ökat och stora transaktioner sker kontinuerligt. De faktorer som påverkar transaktionerna (*handeln*) är många och antagligen omöjlig att överskåda. I en transaktion gynnas förhoppningsvis båda parter men ofta så handlar det om nollsummespel, där det finns vinnare och förlorare.

"Traders" benämns personer som handlar på den finansiella marknaden med ett kortare tidsperspektiv (allt från några sekunder upp till några månader) än vanliga investerare. Dagens moderna traders vill ofta ha verktyg och hjälpmedel för att kunna få fördelar i sina affärer (*trades*). Traders försöker att prognostisera

framtida priser och händelser för att kunna tjäna pengar på sin handel. Man har krav på snabb och automatisk orderläggning. Predicering kan ske genom fundamentala analyser (*kvalitativa*) eller under mer tekniska former (*kvantitativa*). Algoritmhandel baseras främst på kvantitativa data och är det område vi skall studera i detta arbete. Området är stort och vi har därför fokuserat på artificiella neurala nätverk.

1.3 PROBLEMDISKUSSION

Utmaningen ligger som i allt annat i vårt vardagliga liv att försöka förutspå och prognostisera framtida händelser, i vårt fall den finansiella marknaden (*aktiemarknaden företrädesvis*). Det finns en rad motiv för att göra detta. Den främsta fördelen med att lyckas förutspå börsen är den finansiella belöningen. Prestige och utmaningen i sig, att försöka bevisa att marknaden går att förutspå, kan också verka tillfredställande.

För att förstå problemet vidare ställs det krav på att man har relevanta kunskaper inom finansiell ekonomi och dess olika teorier och hypoteser. "Efficient Market Hypothesis" (EMH) är en av dessa hypoteser som är vida accepterad av många ekonomer idag. [4] EMH är den grundläggande utgångspunkten för "random walk"-hypotesen. Det EMH framför är att det inte är möjligt att predicera en specifik akties framtida pris och att priset för en aktie återspeglar all den information som finns tillgänglig vid det tillfället. Då det kommer ny information kommer det att direkt absorberas av aktiemarknaden och priser korrigeras efter detta. Det finns skäl att ifrågasätta EMH där flertalet anomalier har påträffats. Kritiken är av både empiriskt och teoretiskt slag och kommer primärt från olika investerare och forskare.

Behavioristerna berör särskilt kognitiva faktorer som orsak till irrationella beteenden och värderingar i den finansiella marknaden. Man utgår från människans ofullkomlighet och hennes benägenhet att göra misstag, som i sin tur leder till felvärderingar och uppkomsten av mönster i den finansiella marknaden.

En anomaly för EMH, som emellanåt uppstår är bubblor. Exempelvis IT-bubblan som vi fick erfara kring millennieskiftet. Marknaden visar då en form av ignorans för den underliggande tillgångens egentliga värde och agerar istället med stor överflöd av irrationalitet. För läsare som vill fördjupa sig i empirisk kritik mot EMH rekommenderas följande artiklar och litteratur:

- ∴ Francis Nicholson. Price-Earnings Ratios in Relation to Investment Results. *Financial Analysts Journal*. Jan/Feb 1968: 105-109.
- ∴ Sanjoy Basu. Investment Performance of Common Stocks in Relation to Their Price-Earnings Ratios: A test of the Efficient Markets Hypothesis. *Journal of Finance*. 1977. 32:663-682.
- ∴ Rosenberg B, Reid K, Lanstein R. Persuasive Evidence of Market Inefficiency. *Journal of Portfolio Management*. 1985. 13:9-17.
- ∴ Lo, Andrew. *A Non-Random Walk Down Wall Street*. Princeton: Princeton University Press, 1999.

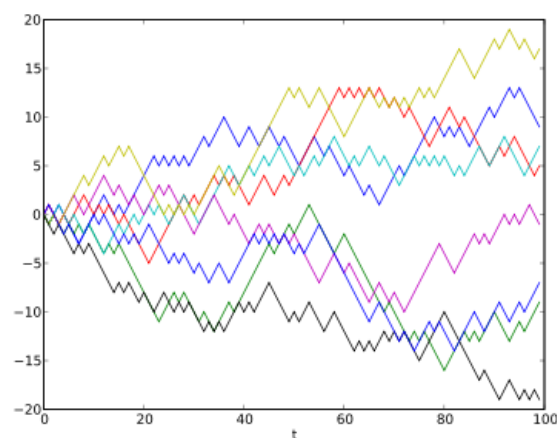


Fig 2. Åtta slumpmässiga grafer där alla börjar från 0 och går med tidsaxeln t.¹

¹ Bilden tagen från Wikipedia: http://en.wikipedia.org/wiki/Random_walk (2010-04-29).

1.4 PROBLEMFÖRMULERING

Studiens främsta syfte är att ge en översikt över artificiella neurala nätverk som verktyg för den finansiella marknaden. Eftersom området är mycket omfattande och studiens resurser små kommer fokus att ligga på den beskrivande bilden. Studien vill ur ett kritiskt perspektiv granska och studera artificiella neurala nätverk och skall utgå från publicerade artiklar och studier. Bland artiklar som granskas vill vi särskilt nämna:

- Efstathios Kalyvas. USING NEURAL NETWORKS AND GENETIC ALGORITHMS TO PREDICT STOCK MARKET RETURNS. *University of Manchester*. October 2001.
- Karl Nygren. Stock prediction – A Neural Network Approach. *Royal Institute of technology*. March 2004.

Övriga artiklar anges i referenslistan. En enkel diskussion om huruvida neurala nätverk inom finansmarknaden är lönsam och om det är möjligt att bygga ett intelligent algoritmiskt handelsystem som ger långsiktig lönsamhet ska tas upp. Uppenbarligen finns det ingen anledning att investera stora resurser i algoritmhandel ifall det inte är lönsamt i förhållande till en vanlig buy-and-hold strategi när det gäller den vanliga aktiemarknaden.

1.5 OLIKA METODER FÖR ATT PREDICERA AKTIEMARKNADEN

De två populäraste metoderna för att förutsäga framtida aktiepriser är *teknisk analys* och *fundamental analys*. Dessa två metoder används främst av småsparare och lekmän men det finns även "proffs" som använder sig utav dessa två metoder. Definitionen av teknisk analys är ofta av varierande form, men generellt sätt kan man placera definitionerna inom en gemensam ram. Aktiespararna förklarar teknisk analys som "*Teknisk analys är en analysmetod som främst analyserar tillgång och efterfrågan på värdepapper baserat på historiska kursdata och handelsvolym i aktier, terminer, optioner, räntepapper och valutor*".^[5] Teknisk analys är följaktligen ett samlingsbegrepp över ett flertal olika metoder som gemensamt utgår från historiska kursutvecklingar på finansiella tillgångar för att prognostisera framtida utvecklingen. Kärnhypotesen är att marknaden till största delen följer vissa mönster som baseras på psykologi och med hjälp av teknisk analys kan man finna trenderna.

Fundamental analys utgår istället från ett bolags redovisade uppgifter såsom redovisningar av olika slag. Information som är intressanta är företagets resultat, omsättning, kassaflöde och så vidare. Det handlar i grund och botten om en form av företagsvärdering. Förutom dessa två metoder kan vi skapa ytterligare två kategorier, *traditionell tidsserie analys* och *maskininlärningsmetoder*.

Traditionell tidsserieanalys behandlar matematiska modeller av tidsvarierande fenomen, vilket man försöker applicera på aktiemarknaden. Tidsserieanalys omfattar metoder för att bryta ned dataserier i tiden, för att erhålla statistisk information som man kan ha fördel av för att förutspå aktiemarknaden exempelvis. Kort förklarar: det handlar om att förutse framtida händelser baserad på den statistiska informationen man utvinner genom tidsserieanalysen.

Det är maskininlärningsmetoder som uppsatsen ska beröra där artificiella neurala nätverk ingår. Även denna kategori omfattar många olika metoder och tekniker. Den gemensamma egenskapen metoderna under maskininlärningsmetoder har är förmågan att "lära" sig, främst genom statistiska metoder för regression eller klassificering. Beskrivningarna av de fyra olika kategorierna här är knapphändiga och därför rekommenderas läsaren att på egen hand studera dessa alternativ ifall intresse finnes .

2. ARTIFICIELLA NEURALA NÄTVERK

2.1 INLEDNING

ARTIFICIELLA NEURALA NÄTVERK (ANN) är ett försök att imitera en hjärnas funktioner och tillvägagångssätt att bearbeta information och ta beslut. En skillnad mellan att använda ANN för att predicera marknaden jämfört med vanlig klassisk teknisk analys är att ANN inte baseras på redan fördefinierade begrepp. Ett ANN lär sig själv att känna igen och hitta mönster i stora data mängder. Det är en av fördelarna med att använda sig utav ANN. Vår hjärnas neurala nätverk består av följande element:

- **Neuroner** (*processing elements*) – cellerna som utgör själva hjärnan, bearbetar och lagrar information.
- **Nätverk** (*networks*) – en grupp av neuroner.
- **Dendriter** (*inputs*) – utskott från nervceller som leder impulser in mot neuronerna, mottagare av information.
- **Axoner** (*outputs*) – utsignalskopplingar till andra neuronerna och synapser.
- **Synapser** (*weights*) – finns i kopplingarna mellan neuronerna. Dessa kan ses som väljare som antingen hindrar eller släpper igenom flödet av information mellan neuronerna.

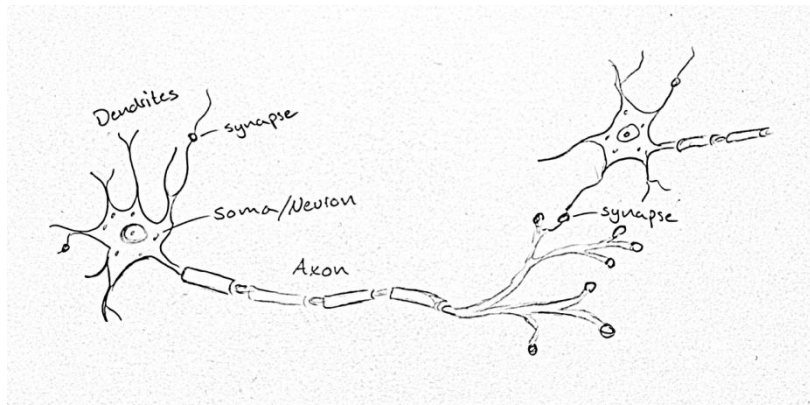


Fig 3. Biologisk neuralt nätverk.

Visionen med ANN är att efterlikna den biologiska funktionen av en mänsklig hjärna. Men det är viktigt att notera att ANN inte är en modell av en hjärna. Hur en mänsklig hjärna lär sig och fattar beslut är av en mycket större komplexitet än dagens ANN kan åstadkomma. Ett ANN är byggt för att på ett enkelt och effektivt sätt hitta mönster, både kontinuerliga och tillfälliga i tidsrummet.[6]

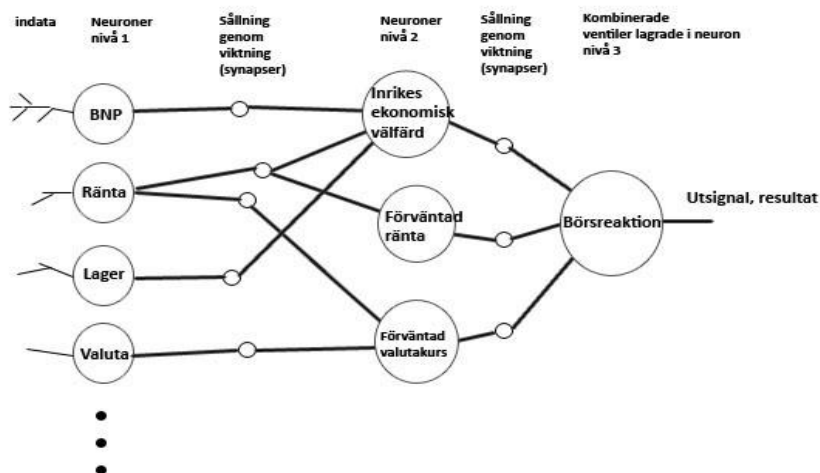


Fig 4. Ett artificiellt neuralt nätverk med tre lager av neuronerna.

ANN kan mycket förenklat beskrivas som en responsprocess med betingad respons. Ett exempel är den studie som Ivan Pavlov utförde på sina hundar.[7] Pavlov upptäckte att hans hundar utsöndrade saliv när de fick mat. Han testade då att före varje matning ringa med en klocka och upprepade detta några gånger. Klockan utgjorde då en betingad stimulus som följdes av matsservering strax efteråt och påverkade hundhjärnans synapser. Därför bildades det saliv i hundarnas mun när man ringde på klockan oavsett om hundarna fick mat eller inte.

DET ÄR AV STOR VIKT ATT MAN VÄLJER INDATA OMSORGSFULLT då det är utifrån dessa indata som ett ANN kommer att bearbeta och slutligen utmyнна i ett resultat. Som vi nämnt tidigare i introduktionen är det otroligt många faktorer som påverkar en akties pris och börsens rörelser och man bör sälla ut de mest relevanta för börsens rörelse. Relevanta indata kan vara av följande slag:

- **Bruttonationalprodukt (BNP, GDP)**
- **Arbetslöshetsstatistik**
- **Lagerstatistik**
- **Räntor**
- **Valutakurser**
- **Transaktioner och order av tillgångar**
- **Aktuella börskurser och historik**

Här bör man än idag hjälpa sitt ANN på traven med att gallra. Vår hjärna är än idag (2010) överlägsna datorerna i den meningen att vi har den förmågan att kunna förstå och se komplexa mönster och samband, vilket dagens teknik fortfarande har svårt med. För en vanlig person är det relativt intuitivt att statistik om nya examinerade ekonomer har en tämligen marginell om ens någon påverkan på börsens kortsiktiga riktning. Indata måste ha en viss validitet, med andra ord en stark koppling till finansmarknaden. Man skall inte välja indata med alltför långsökt koppling till börsen. Detta av den enkla anledningen att vi får alldeles för många parametrar att hålla reda på. Beräkningarna kan lätt bli allt för omfattande och långsamma. Å andra sidan bör man heller inte välja allt för få parametrar som indata, då det finns en risk att man inte får med essentiella faktorer som faktiskt har en viss påverkan på börsens riktning.

DET ÄR SYNAPSERNA SOM VÄGER DESSA FAKTORER och med detta avgör faktorernas betydelse. En faktor som påverkar börsen stort kan ges en större vikt. Räntorna eller räntebesked kan sålunda påverka börsen i större grad än vad lagerstatistiken för vete gör och därför ges en högre vikt än vikten för lagerbeskedet. Synapserna kan på ett enkelt sätt sägas agera som tröskel och man definierar en viss minimigräns på den viktade summan där synapserna tillåts släppa igenom information. Vikterna av dessa faktorer kan beräknas utifrån exempelvis genetiska algoritmer eller den mer populära backpropagationalgoritmen.

Ett ANN system "vet" inte vad som är rätt eller fel i initieringsfasen. Detta måste man tala om för det. Feedbackprocessen är av central betydelse i konstruktionen av ANN. Den ingår i själva inlärningsprocessen. En ansenlig mängd historiska data fordras för att inlärningsprocessen skall fungera bra. Detta av att den enkla anledningen att man behöver material att studera, och ur studierna kommer erfarenhet. Den mängd data man har tillgänglig har man för vana att dela upp i tre delar där den första delen av data används för träning, den andra delen för testning och den tredje delen för out-of-sample validering. Det vill säga validering av att ens system fungerar.

2.2 MATEMATISK BESKRIVNING AV ANN

McCulloch-Pitts Modell av ett artificiellt neuron

Den tidiga modellen av ett neuron presenterades utav Warren McCulloch och Walter Pitts 1943. Det är en enkel och fortfarande aktuell modell som har ett precist matematisk definition. Neuronets utsignal Y antas vara en funktion av den viktade summan av insignalerna x_i .

$$\zeta = \sum w_i x_i \quad Y = \sigma(\zeta)$$

Där ζ är den viktade summan av alla inputs och $\sigma(\zeta)$ är en funktion (aktiveringsfunktionen) av den viktade summan. Detta ger oss:

$$Y = \sigma(\sum w_i x_i)$$

Detta kan också skrivas som en skalärprodukt då \mathbf{w} och \mathbf{x} är en vektor enligt följande:

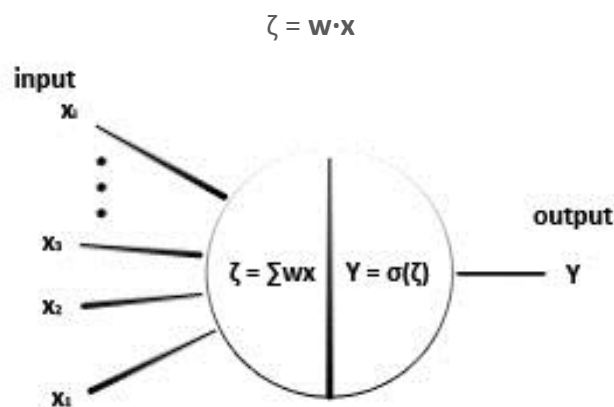


Fig 5. En modell av ett artificiellt neuron med aktiverings funktion

AKTIVERINGSFUNKTIONER

Nästan enbart tre olika former av aktiveringsfunktioner används inom ANN idag: tröskelfunktioner (stegfunktion), styckvis linjära funktioner och sigmoid funktioner. Här är de vanligaste aktiveringsfunktionerna.

- (i) **Tröskelfunktion** (stegfunktionen) är den enklaste formen. De har bara två olika utfall: A_1 och A_0 där dessa oftast är av siffrorna 1 respektive 0. Beroende på hur stort värde de summerade inputen har ger den två olika utfall. Enligt bilden så gäller att ifall den summerade värdet är större eller lika med noll ges signalen A_1 , annars A_0 . För stegfunktionen gäller:

$$\sigma(\zeta) = \begin{cases} 0, & \zeta < 0 \\ 1, & \zeta \geq 0 \end{cases}$$

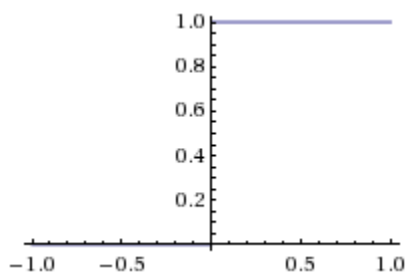


Fig 6. Tröskel (steg) funktion

(x from -1 to 1)

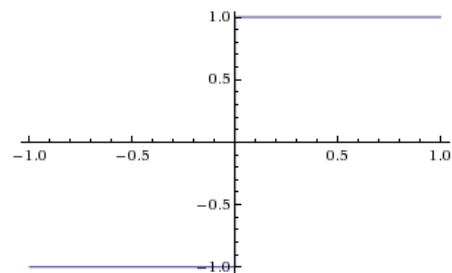


Fig 7. Signfunktion.

(ii) och för **signfunktionen**:

$$\sigma(\zeta) = \begin{cases} -1, & \zeta < 0 \\ 1, & \zeta \geq 0 \end{cases}$$

(iii) För den **linjära funktionen** gäller:

$$\sigma(\zeta) = \zeta$$

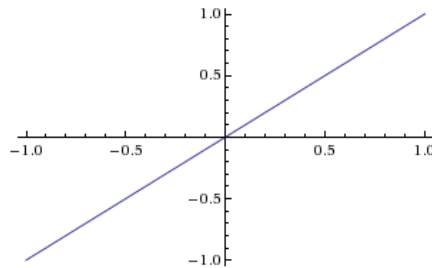


Fig 8. Linear function.

(iv) **Styckvis linjära funktionen** består av rätta linjestycken, som delvis är linjär i ett angivet intervall och delvis likt step-funktion.

$$\sigma(\zeta) = \begin{cases} 1, & \zeta \geq \frac{1}{2} \\ \zeta + \frac{1}{2}, & -\frac{1}{2} < \zeta < \frac{1}{2} \\ 0, & \zeta \leq -\frac{1}{2} \end{cases}$$

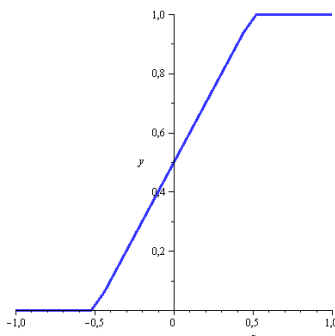


Fig 9. Styckvis linjär funktion.

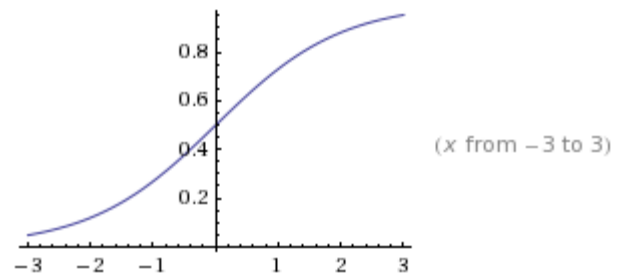


Fig 10. Log-Sigmoidfunktion.

(v) En mjukare form S-form har **Sigmoidfunktionen** (i vårt fall har vi valt att presentera en kontinuerlig log-Sigmoid funktion). Det kallas för log-Sigmoid av den anledningen att Sigmoid även kan konstrueras med hjälp av hyperbolisk tangent funktion. Formeln lyder enligt följande:

$$\sigma(\zeta) = \frac{1}{1+e^{-\beta\zeta}} \quad \text{där } \beta \text{ är en lutningsparameter.}$$

2.3 ANN MODELLER

DET FINNS ETT ANTAL OLIKA FORMER AV ANN MODELLER. Dessa inbegriper en del av dem [8]:

- Feed-Forwardnätverk.
- Radial Basis Functionnätverk.
- Recurrentnätverk.
- Echo-Statennätverk.
- Hopfieldnätverk.
- Self-Organizing Maps.
- Competitive Models.
- ART Models.
- Boltzmann Machines.
- Committee of Machines.

Vi ska här beskriva två av de, för uppsatsen, mest betydande modellerna. Några av de andra kan uppfattas som specialfall av dessa två. Det vill säga att recurrentnätverk omfattar andra modeller som är speciellt utformade varianter av recurrentnätverk.

Feedforwardnätverk är den enklaste formen av ANN. Här skickas informationen enbart framåt i nätverket och informationen kan inte gå sidledes eller bakåt i nätverket. Följaktligen har vi ingen feedback process. Detta kan matematiskt beskrivas enligt:

$$W_{ij} = 0 \text{ om } i = j$$

$$W_{ij} = 0 \text{ om } \text{lager}(j) \leq \text{lager}(i)$$

Recurrentnätverk är till skillnad från feedforwardnätverk ett nätverk där information kan skickas bakåt och sidledes mellan nivåerna av lager. Detta ger nätverket en förmåga att uppvisa dynamiskt tidsmässiga beteenden eller mönster.

Förmågan att kunna skicka feedbackloopar från outputen till inputen ger ett omfattande påverkan på nätverkets inlärningsförmåga. Generellt fungerar inlärningen genom en process av feedbackloopar, som upprepas flertal gånger tills outputen konvergerar mot ett konstant värde. Men det kan likväl uppstå en form av kaotiskt beteende där outputen aldrig stabiliseras. Det brukar korthugget benämnas att nätverket inte är *stabil*.

I regel tar det längre tid att träna ett *recurrent neuralt nätverk* än *feedforward nätverk*. [9] Detta har sin grund i att uträkningar tar betydande tid och att flertalet loopar krävs innan nätverket är tränat. Men det finns flera möjliga tillvägagångssätt att förbättra beräkningseffektiviteten för recurrentnätverk. [10-12]

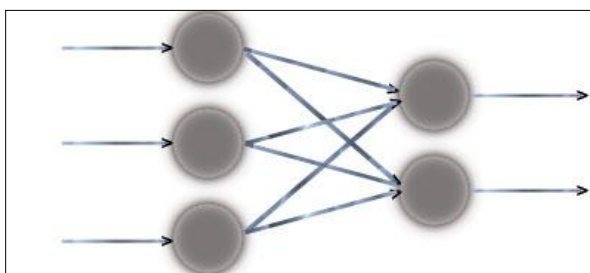


Fig 11. Ett exempel på ett feedforwardnätverk.

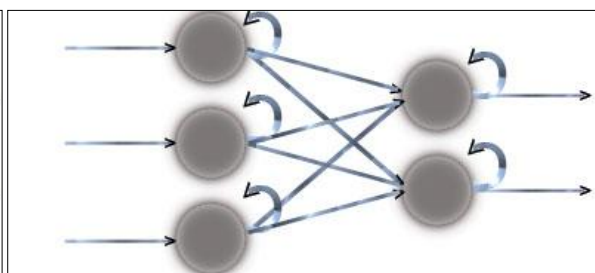


Fig 12. Ett exempel på ett recurrentnätverk.

2.4 INLÄRNING

Det unika med ett system byggt på ANN är att det kan tränas. Vi har tidigare påpekat detta, att ANN kan lära sig att känna igen mönster i stora mängder av data. I vårt fall vill vi utnyttja detta för att hitta *trender* (mönster) på börsdata. Man kan dela in nätverks träningen i två huvudsakliga kategorier, stegvis- och batch (sats eller hög) träning. Stegvis träning innebär att vikterna i nätverket modifieras efter varje gång indata-sampel (en kort tidsserie av data) presenteras för nätverket. Batch träning innebär att modifiering av vikterna inte sker förrän all data har presenterats för nätverket. Det har publicerats en mängd olika förslag på inlärningsmetoder. Vi presenterar här en lista över några vanliga och viktiga metoder:

- Hebb-regeln (Hebbian learning).
- Delta-regeln
- Backpropagation-algoritmen
- Kohonen-regeln
- Error-correction inläring
- Genetiska algoritmer

Att presentera dessa inlärningsmetoder i detalj blir allt för omfattande för denna grundläggande uppsats vilket vi därför hänvisar till egna studier kring dessa metoder. Det viktigaste att förstå är att träningsalgoritmer utför i generella drag en uppdatering av vikterna i ett ANN.

2.5 FÖRDELAR MED ARTIFICIELLA NEURALA NÄTVERK

Det som talar för användningen av ANN som verktyg för att analysera finansmarknaden är främst dess fördomsfrihet, dess förmåga att lära sig och att generalisera. Vi vet av erfarenhet att världen är föränderlig vilket parallellt innebär att börsen är föränderlig. Det är här en av ANNs egenskaper blir till fördel då ett system byggt på ANN kan lära sig och anpassa sig mot den föränderliga finansmarknaden på ett effektivt sätt. Det kan inte ett klassiskt analysverktyg som teknisk analys eller traditionella tidsserieanalys göra, då det grundar sig i fördefinierade fördomar och regler, där instruktioner är centralt. En annan stark fördel med ANN är dess förmåga att upptäcka icke-linjära samband mellan kombinationer av data.[7]

2.6 KRITIK OCH NACKDELAR MED ARTIFICIELLA NEURALA NÄTVERK

Jämfört med exempelvis klassisk teknisk analys kan det krävas en del programmeringskunskaper och resurser för att kunna utnyttja ANN och underhålla systemet fullt ut. Det är en hög inläringströskel i förhållande till andra kommersiella metoder. Ett annat problem med ANN är att det kan ta lång tid att träna det. Det kan också vara svårt att se vad som händer inne i nätverket. En så kallad black-box egenskap karakteriserar ANN. Överträning kan inträffa vilket innebär att systemet börjar beskriva slumpmässiga händelser och så kallad brus istället för att enbart beskriva det underliggande problemet. Även en form av expertis krävs för att kunna sortera ut "rätt" indata för systemet.

3. STUDERADE ARTIKLAR

Vi ska i detta avsnitt studera två akademiska uppsatser som behandlar ämnet artificiella neurala nätverk. Med ett kritiskt förhållningssätt ska vi studera och granska artiklarna för att ta lärdom av de empiriska experimenten som utförts av författarna och slutligen dra en slutsats utifrån deras resultat. Huvudfrågan är huruvida ANN är användbar och lönsam som prediktionsinstrument för aktiemarknaden.

3.1 STOCK PREDICTION – A NEURAL NETWORK APPROACH

SAMMANFATTNING

Karl Nygrens masteruppsats [13] beskriver neurala nätverk i överlag och han har presenterat en empirisk studie där han använt sig utav ANN med ECNN (Error correction neural networks) som specifik modell. Syftet med uppsatsen var att visa att man faktiskt kan med hjälp av neurala nätverk predicera marknaden i ett visst mån. Han utförde studien på bland annat den svenska börsindex (Karl Nygren refererar till SXGE, vilket är symbolen för Swiss Performance Index SPI vilket kan te sig märkligt) och två stora svenska aktiebolag, Ericsson B och Volvo B som referensobjekt. Resultaten var i överlag bättre för börsindex än för de två enskilda aktierna. Han jämförde detta med en metod som kallades för den naiva prediktions metoden.

Valet av dessa aktier och börsindex som studiereferens var på grund av deras goda likviditet och låga *spread*. Här har man valt att definiera *spread* som skillnaden mellan en akties/index högsta och lägsta kurs under en dag. I vanligt tal brukar man hellre syfta på skillnaden mellan bud- och säljpriser, när man använder ordet *spread*. Det är den skillnaden en *marketmaker* tjänar i en transaktion.

För all prediktion som Karl Nygren utförde använde han sig utav främst fyra tidserier av data för sin modell.

- Stängningskurs, det pris en aktie eller index hade på den senast genomförda transaktionen under dagen
- Högsta pris, det högsta priset en aktie eller index köptes/såldes för under dagen
- Lägsta pris, det lägsta priset en aktie eller index köptes/såldes för under dagen
- Volym, totalt antal handlade aktier under en dag

Förutom ovan nämnda tidsserier används för aktierna respektive Index. För Ericsson B och Volvo B användes även dessa externa tidsserier.

- Dow Jones aktie Index
- Svensk aktie Index
- 3-månaders ränta, Sverige
- 5-års ränta, Sverige
- Valutakurs, SEK/USD
- Valutakurs, SEK/DEM²

För den svenska aktieindexen användes dessa externa tidsserier utöver de fyra grundläggande tidsserierna som nämndes överst.

- S&P 500
- SX-16
- Nikkei 225
- Dow Jones aktie index
- German DAX
- Guldpris [\$/Oz]
- 3-månaders ränta, Sverige
- 5-års ränta, Sverige
- Valutakurs, SEK/USD
- Valutakurs, SEK/DEM

² D-mark, Tysklands föregående valuta innan konverteringen till EURO.

Nygren initierade sina vikter slumpmässigt, tränade sedan sitt nätverk genom error-correction inlärningsmetoden och till detta använder han programmet Simulation Environment for Neural Networks (SENN), från Siemens, för att bygga sitt system. Nygren kommer fram till att det kanske viktigaste delen i en implementation av ANN är att åstadkomma stabilitet över tiden. Han är medveten om att varken hans dagliga eller veckovisa modell är optimerad och förbättringar krävs för att åstadkomma denna stabilitet.

DISKUSSION

Uppsatsen av Nygren är i överlag mycket välskriven och beskriver ANN mycket övergripande där man som läsare enkelt och snabbt sätter sig in i ämnet utan överflöd av information. På ett pedagogiskt sätt leder Nygren läsaren genom ämnet som gradvis blir tekniskt djupare och svårhanterligare. När det kommer till det empiriska experimentet utförd av Nygren så finns en önskan att en förklaring eller redogörelse gavs över varför han valt faktorerna som angivits ovan som indata och varför dessa har en påverkan på index och aktierna. För en ekonomisk insatt person kan det vara intuitivt men en redogörelse bör finnas till som stöd för läsaren och för studiens validitet. En förklaring har tidigare givits om att en faktor ges en lägre vikt ifall dessa inte påverkar den underliggande tillgången (exempelvis Volvo B) avsevärt men det är av stor betydelse att man väljer "rätt" faktorer. Detta av den anledningen att ett ANN kan bli allt för komplext och långsamt ifall man tar med för många faktorer. Om man tar med mindre betydelsefulla faktorer så blir beräkningarna i systemet mer komplext och långsamt.

Kritik kan riktas mot att man inte redovisar träningsproceduren. Det skulle kunna ha lämnats som en bilaga. För att dra en riktig kvalificerad slutsats bör man även göra en empirisk studie på marknaden som ännu inte har ägt rum. Men detta har vi overseende över då arbetet är begränsad under en kort period medan en sådan studie kräver åtminstone några månaders testning.

3.2 USING NEURAL NETWORKS AND GENETIC ALGORITHMS TO PREDICT STOCK MARKET RETURNS

SAMMANFATTNING

Kalyvas Efstathios uppsats [14] huvudsakliga mål är att predicera den kortsiktiga framtida utvecklingen av aktiemarknaden. De två aktieindex som agerar studieobjekt för hans arbete är FTSE500 (London marknaden) och S&P500 (New York marknaden). Efstathios delar upp arbetet i två huvudsakliga delar där han först undersöker prediktionsmöjligheten av aktiemarknaden samtidigt som han försöker tillhandahålla bevis för att marknaden inte fluktuerar slumpmässigt. Detta är som jag påpekat tidigare i 1.3 en premis man måste utgå från för att kunna skapa ett system som kan predicera aktiemarknaden. Den andra delen applicerar Efstathios två prediktionsmodeller, autoregressive, feed-forward neuralt nätverk och mäter dess prestationsförmågor. Det som är intressant för oss är studierna om feed-forward nätverket. Efstathios använder sig utav en genetisk algoritm för att hitta den optimala topologin (vikt fördelningen).

Uppsatsen av Efstathios tar upp förutom traditionell tidsserieanalys, maskininlärningsmetoder även teknisk analys och fundamental analys kortfattat. Detta ger läsaren en överblick över ämnet och en omfångsrikare attityd till aktieprediktion. En allomfattande inledning ges där de flesta ekonomiska termer och teorier som är vital för förståelsen av den datalogiska studien om ANN beskrivs. Innan Efstathios börjar med experimenten, där bland annat feed-forward nätverket implementeras, beskriver han ingående hur data skall behandlas för att vara kompatibel med de modeller som kommer att användas. Detta innebär bland annat en definiering av data där det dessutom beskrivs i statistiska termer.

För att testa slumpmässigheten i tidsserierna som Efstathios hämtar från Datastream International använder han två empiriska tester, "Run" – test och "BDS" – test.[15] Tidsserierna omfattar intervallen mellan 1998-01-04 till 2000-12-12 och tidsserierna av FTSE500 index respektive UK T-Bill rates består av 3275 observationer. Tidsserierna för S&P500 Index och US T-Bill rates omfattas av 3277 observationer. Resultaten av av båda testerna gav indikationer på att tidsserierna inte fluktuerade helt slumpmässigt. Dock så kunde man inte med "run"-testen indikera att S&P500 serierna var av "non-random"-karaktär vilket "BDS"-testen gjorde. Slutsatsen Efstathios kommer fram utifrån sina studier är att marknaden inte fluktuerar slumpmässigt vilket öppnar möjligheten att skapa ett system som kan predicera marknaden.

Till sitt experiment använder Efstathios "Neural Networks Toolbox" i Matlab 5.2 för att träna, validera och testa sin feed-forward nätverk. Även den genetiska algoritmen är implementerad i Matlab 5.2. Experimentet delade han in i tre faser där den första fasen innebar att en genetisk algoritm söker genom hela neurala nätverk av olika strukturer för att skapa den optimala viktfordelningen. Den andra fasen innebar

kortfattat träning och validering av nätverket. Den sista delen testar Efstathios nätverket på ännu icke använd data och mäter nätverkets prestation. Han använder sig utav fyra olika mätinstrument för att utvärdera sina modeller där tre av dessa är benchmarking mot det naiva prediktionsmetoden och det fjärde är det genomsnittliga absoluta prediktions felet.

Utfallet av experimentet var inte så god som man kunnat hoppats på, då den neurala nätverket inte lyckades prestera bättre än den naiva prediktionsmetoden. De huvudsakliga slutsatserna som Efstathios drar är bland annat att man förmodligen inte bör använda sig utav daglig data som innerhåller höga nivåer av brus vilket försvårar arbetet med att hitta mönster i tidsserierna. Experimentet ansträngde sig i att hitta mönster i relativt långa tidsperioder vilket Efstathios uppmärksammar. Det han noterar då är att det kanske inte existerar mönster i dessa intervaller utan finns lokalt under kortare kronologiska perioder.

DISKUSSION

Efstathios uppsats är mycket välarbetad med stor fokus på den tekniska biten. Den väsentliga skillnaden mot Nygrens arbete är att Efstathios arbete är aningen svårare att ta till sig av den anledningen att Efstathios valt att gå djupare ned i de tekniska detaljerna. Kunskaper om traditionell tidsserieanalys är till fördel för att man på bästa sätt ska följa med i arbetet. De empiriska experimenten av Efstathios är relativt invecklad, då man valt att använda sig utav genetisk algoritm för att skapa en optimal topologi och sedan selekterar de mest anpassade nätverken. Efstathios valde att avgränsa indata till att enbart omfatta de fyra nämnda tidsserierna, vilket är hans antagande att det räcker för att predicera framtida utveckling av respektive index (FTSE500 & S&P500). Han valde enbart UK T-bill rates och US T-bill rates, som *fundamental* indata för sina experiment. Förmodligen är det en av orsakerna till den bristande resultatet. Att predicera två stora index som FTSE500 och S&P500 är förmodligen av en mer komplex art än att predicera den svenska börsen som Nygren gjort. Detta av den anledning att det med stor sannolikhet är större och fler konkurrenter inom FTSE500 index och S&P500 index än i den svenska. Exempelvis stod hög frekvent trading för 73 procent av handelsvolymen med aktier i USA 2009.[16]

4. SLUTSATS OCH AVSLUTANDE DISKUSSION

4.1 ARTIFICIELLA NEURALA NÄTVERK – ÄR DET LÖNSAMT?

Det råder inget tvivel om att det finns en stor potential i ANN som verktyg för att hitta mönster och trender. För att lyckas hitta mönster i stora mängder data är kravet att det inte råder slumpmässighet i den mängden. Det föreligger fortfarande delade meningar om aktiemarknaden fluktuerar slumpmässigt, varför EMH fortfarande är aktuell och erkänd bland många ekonomer idag. Utifrån exempelvis Efstathios studie som påvisar att STE500 index och S&P500 index inte fluktuerar slumpmässigt öppnar det upp möjligheten att applicera ANN för att predicera utvecklingen. Exempelvis lyckades Nygren implementera ett ANN med lyckat resultat på den svenska aktiebörsen, om än diskutierbar hur lukrativt det är. Emellertid måste vi påminnas om att "trading" på aktiemarknaden handlar om ett nollsummespel i generella drag, där det finns en vinnande respektive förlorande part.

Vi kan även dra slutsatsen att ANN som prediktionsmodell är lönsam endast om marknaden inte är homogen i sitt bruk av prediktionsmodeller eller system. Detta av ren förnuft, då alla (eller en stor del) aktörer som använder sig utav samma modell (hur löjligt det än låter så bör vi resonera kring detta) kommer att agera på samma sätt vilket då leder till en överreaktion åt något håll (pris: upp eller ned). Fördelarna med systemet/modellen kommer att urholkas och inte vara fördelar längre. Att skapa ett system som kan överprestera handlar i huvudsak om att skapa en fördel mot andra aktörer, vilket i sig talar för att man måste konstruera ett unikt system som inte ligger i andras vetskaper. Detta är en förutsättning för att ens system eller modell ska vara lönsam på längre sikt. Om du som aktör använder dig utav ett system som andra har kunskap om kan de agera mot dig och tjäna dina pengar. Sammanfattningsvis har vi kommit fram till slutsatsen att inga system som är lönsamma finns i allmänhetens vetskaper. De lukrativa systemen kommer att förbli dolda och nyttjade av framgångsrika aktörer. Så snart dess egenskaper demaskeras till allmänhetens vetskaper kommer dess fördelar ej vara fördelar längre. Att det existerar och går att konstruera lönsamma system av ANN är mycket sannolikt av den orsaken att man kan konstruera oändligt många system av ANN där kombinationer av variabler och parametrar kan ge oändligt många unika system. Det kan handla om vilka indata man valt och vilka inlärningsmetoder man använder.

5. REFERENSER

5.1 REFERENSLISTA

- [1] Brealey, Myers, Allen. *Principles of Corporate Finance*. 8th Edition. New York: McGraw-Hill Book Company, 2005.
- [2] Wikipedia: http://en.wikipedia.org/wiki/Random_walk_hypothesis (2010-04-20).
- [3] Lo, Andrew. *A Non-Random Walk Down Wall Street*. Princeton: Princeton University Press, 1999.
- [4] Wikipedia: http://en.wikipedia.org/wiki/Efficient-market_hypothesis (2010-04-20).
- [5] Aktiespararna: <http://www.aktiespararna.se/lar-dig-mer/Fordjupningar/Teknisk-analys/Del-1-Gor-din-egen-prognos-med-teknisk-borsanalys/> (2010-04-30).
- [6] Perry J. Kaufman. 3rd Edition. *Trading systems and methods*. New York: John Wiley & Sons Inc, 1998: 493-494.
- [7] Pavlov, I. P. *Conditional Reflexes*. New York: Dover publication, 1927/1960.
- [8] http://en.wikibooks.org/wiki/Artificial_Neural_Networks (2010-04-20).
- [9] Michael Negnevitsky. 2nd Edition. *Artificial Intelligence – A Guide to Intelligent Systems*. Addison Wesley, 2004: 165-215.
- [10] Caudill, M. Neural network training tips and techniques. *AI Expert*. January 1991:56-61.
- [11] Jacobs, R.A. Increased rates of convergence through learning rate adaption. *Neural Networks*. 1988:1:295-307.
- [12] Stubbs, D.F. Six ways to improve back-propagation results. *Journal of Neural Network Computing*. Spring 1990:64-67.
- [13] Karl Nygren. Stock prediction – A Neural Network Approach. *Royal Institute of technology*. March 2004.
- [14] Efstathios Kalyvas. USING NEURAL NETWORKS AND GENETIC ALGORITHMS TO PREDICT STOCK MARKET RETURNS. *University of Manchester*. October 2001.
- [15] Ibid. s. 47-51.
- [16] AdvancedTrading.com: <http://advancedtrading.com/algorithms/showArticle.jhtml?articleID=218401501> (2010-05-02).

