

VILNIAUS UNIVERSITETAS
KAUNO HUMANITARINIS FAKULTETAS

INFORMATIKOS KATEDRA

Verslo informacijos sistemų studijų programa
Kodas 62103S138

AURIMAS SKIRGAILA

MAGISTRO BAIGIAMASIS DARBAS

**AUTOASOCIATYVINIŲ NEURONINIŲ TINKLŲ TAIKYMAS
VERTYBINIŲ POPIERIŲ KAINOS PROGNOZEI**

Kaunas, 2008

VILNIAUS UNIVERSITETAS
KAUNO HUMANITARINIS FAKULTETAS

INFORMATIKOS KATEDRA

AURIMAS SKIRGAILA

MAGISTRO BAIGIAMASIS DARBAS

**AUTOASOCIATYVINIŲ NEURONINIŲ TINKLŲ TAIKYMAS
VERTYBINIŲ POPIERIŲ KAINOS PROGNOZEI**

Leidžiama ginti _____

Magistrantas _____
(parašas)

Darbo vadovas _____
(parašas)

(darbo vadovo mokslo laipsnis, mokslo
opedagoginis vardas, vardas ir pavardė)

Darbo įteikimo data _____

Registracijos Nr. _____

Kaunas, 2008

TURINYS

TURINYS.....	3
SANTRUMPŲ SĄRAŠAS	5
PAVEIKSLŲ SĄRAŠAS.....	5
LENTELIŲ SĄRAŠAS.....	6
SUMMARY	7
ĮVADAS.....	8
1. VERTYBINIŲ POPIERIŲ RINKOS FAKTORIAI	12
1.1 Vertybinių popierių rinkos prognozavimo galimybės.....	14
1.1.1 Efektyvios rinkos hipotezė	14
1.1.2 Chaoso teorija	15
1.2 Fundamentinių rodiklių atranka	15
1.2.1 Pelnas vienai akcijai (EPS).....	16
1.2.2 Akcijos kainos ir EPS santykis (P/E)	16
1.2.3 Grynojo pelno ir nuosavo (akcinio) kapitalo santykis (ROE)	17
1.2.4 Bendrasis likvidumas.....	17
1.2.5 Institucinių investuotojų dalyvavimas	17
1.2.6 Kompanijos akcijų dalis tenkanti įmonės darbuotojams.	17
1.2.7 Kompanijos darbuotojų aktyvumas perkant/parduodant savos kompanijos akcijas.	17
1.2.8 Fundamentinių rodiklių analizės išvados	18
1.3 Techninių rodiklių atranka	18
1.3.1 Lyginamosios jėgos indeksas (RSI).....	18
1.3.2 Slankiųjų vidurkių kirtimosi indikatorius (MACD)	22
1.3.3 Slankusis tikrojo diapazono reikšmių vidurkis (ATR)	25
1.3.4 Bollingerio juostos.....	26
1.3.5 Techninis indikatorius Gaubiančiosios kreivės (Envelopes).....	28
1.3.6 Techninių rodiklių analizės išvados ir apibendrinimas	29
1.4 Vertybinių popierių rinkos prognozavimo galimybių analizės išvados.....	29
2. INTELEKTINIŲ SISTEMŲ TAIKYMAS VERTYBINIŲ POPIERIŲ RINKOJE.....	30
2.1 Principinių komponentų taikymas klasterizavimui	30
2.2 Neuroninių tinklų taikymas klasterizavimui	32
2.2.1 Neuroninių tinklų apmokymo algoritmai	35
2.2.1.1 Paklaidos skaičiavimo atgal algoritmas (error back propagation)	36
2.2.1.2 Apmokymo algoritmų palyginimas.....	37
2.2.2 Neuroninių tinklų trūkumai	37
2.3 Autoasociatyviniai neuroniniai tinklai.....	38
2.4 Algoritmų taikymas vertybinių popierių klasifikavimui.....	39
2.5 Autoassociative ANN.....	40
2.6 Algoritmų apibendrinimas ir išvados	41
3. EKSPERIMENTINIS PCA IR AANN METODŲ TYRIMAS	43
3.1 PCA tyrimas	43
3.1.1 Programinės įrangos parinkimas.....	43
3.1.2 Statistika.....	43
3.2 AANN tyrimas	47
3.2.1 Programinės įrangos parinkimas.....	47
3.2.2 Duomenų paruošimas	48
3.2.3 Tyrimo vykdymas	49
3.2.4 Rezultatai	50
3.2.5 Atvirkštinis eksperimentas.....	51
3.2.6 Rezultatai	52
3.3 Techniniais rodikliais grįstas PCA tyrimas.....	52

3.3.1	Duomenų paruošimas	53
3.3.2	Tyrimo vykdymas	54
3.3.3	Rezultatai po 2 mėnesių.....	55
3.3.4	Rezultatai po 10 mėnesių.....	58
3.4	Eksperimentinio tyrimo išvados.....	60
IŠVADOS IR PASIŪLYMAI.....		61
LITERATŪRA.....		62
1	PRIEDAS	65
2	PRIEDAS	70

SANTRUMPŲ SĄRAŠAS

AANN – autoassociative artificial neural networks (autoasociatyviniai neuroniniai tinklai)
ATP – arbitrage pricing theory (arbitražinės kainodaros teorija)
ATR – average true range (vidutinis tikrasis diapazonas)
BB – Bollinger bands (Boilingerio juostos)
Cross-validation – kryžminis tikrinimas
EMH – efficient market hypothesis (efektyvios rinkos hipotezė)
EPS – Earnings per share (pelnas vienai akcijai)
IAS - International Accounting Standards (tarptautiniai apskaitos standartai)
IFRS - International Financial Reporting Standards (tarptautiniai finansinių raportų standartais)
MA – Moving average (slankusis vidurkis)
MACD – Moving average convergence divergence (slankiojo vidurkio konvergencija ir divergencija)
MSE – mean squared error (vidutinė kvadratinė paklaida)
NLPCA – non-linear principal component analysis (netiesinė principinių komponentų analizė)
P/E – Price/EPS (akcijos kainos ir jos EPS santykis)
PCA – principal component analysis (principinių komponentų analizė)
ROE – Return on Equity (Nuosavo kapitalo pelningumas)
RSI – Relative strength index (santykinio stiprumo indeksas)

PAVEIKSLŲ SĄRAŠAS

1 pav. Kompanijos TDMresearch portfelio ir S&P500 indekso palyginimas.....	12
2 pav. RSI indeksas	20
3 pav. RSI tendų rinkoje	21
4 pav. RSI ir stochastinis RSI kylančių tendų rinkoje	21
5 pav. RSI ir stochastinis RSI besileidžiančių tendų rinkoje	22
6 pav. MACD indikatorius	23
7 pav. Divergencija	24
8 pav. Konvergencija.....	25
9 pav. ATR indikatorius	26
10 pav. Boilingerio juostos.....	27
11 pav. Gaubiančiosios juostos	28
12 pav. PCA metodo schema	30
13 pav. Neuronų sandara.....	33
14 pav. Dirbtinio neurono funkcinė schema	34
15 pav. Neuroninis tinklas su vienu paslėptu sluoksniu.....	35

16 pav. Dirbtinio neurono įėjimo/išėjimo priklausomybės	35
17 pav. Neuroninio tinklo, naudojamo akcijų kainos prognozavimui, struktūra	36
18 pav. Kryžminis patikrinimas	38
19 pav. Autoasociatyvinio neuroninio tinklo struktūra	39
20 pav. Autoassociative_ANN darbo algoritmas	41
21 pav. Akcijų klasterizavimas naudojant „Statistica“	44
22 pav. A klasterio akcijos	45
23 pav. B klasterio akcijos	45
24 pav. C klasterio akcijos	45
25 pav. D klasterio išskyrimas	46
26 pav. D klasterio akcijos	47
27 pav. Neuroninio tinklo treniravimas.....	50
28 pav. Neuroninio tinklo įėjimo ir išėjimo kintamųjų palyginimas	50
29 pav. Akcijų pasiskirstymas AANN sudaryto „butelio kakliuko“ erdvėje	51
30 pav. Menka klasterizacija AANN sudarytoje erdvėje	52
31 pav. Originalus ir rekonstruoti signalai	52
32 pav. Pagrindinis PickStock langas	54
33 pav. 3 principinių komponentių portfelių pelningumo palyginimas po 2 mėn.	56
34 pav. WDC ir BMM palyginimas su S&P500 indeksu (2 mėn.)	56
35 pav. WDC ir BMM palyginimas su S&P500 indeksu (10 mėn.)	58
36 pav. 3 principinių komponentių portfelių pelningumas po 10 mėn.....	59

LENTELIŲ SĄRAŠAS

PCA, ANN ir NLPCA palyginamojo testo rezultatai	14
Apmokymo algoritmų palyginimas.....	37
Portfeliai sudaryti naudojant 5 (ir daugiau) principinių komponentių	55
Portfeliai sudaryti naudojant 3 principines komponentes	57
Portfeliai sudaryti naudojant 1 principinę komponentę	57
Portfeliai sudaryti naudojant 5 (ir daugiau) principinių komponentių	58
Portfeliai sudaryti naudojant 3 principines komponentes	59

SKIRGAILA, Aurimas. (2008) *Stock forecasting by applying associative Neural Networks*. MBA Graduation paper. Kaunas: Vilnius University, Kaunas Faculty of Humanities, Department of Informatics. 64p.

SUMMARY

This is a survey on the application of auto associative neural networks and principal component analysis in clustering stocks. Main principles of these two methods are presented, reviewing the current usage of AANN and PCA and future outlook. An experiment is being carried out by building two stock portfolios using PCA. The portfolios are being monitored within one year.

The main goal of the survey is to estimate the abilities of application of auto-associative neural networks stock forecasting in the US stock market.

In order to reach the goal, the following tasks have been set:

- To analyze the probability of general market prediction; analyze fundamental and technical factors, select the most suitable ones for further investigation.
- To consider different implementations of artificial neural networks, select the most suitable ones for stock market forecasting
- To compare various stock forecasting software solutions based on neural networks or different intelligent systems.
- According to the chosen methods and software, perform the historical stock data analysis, build investment portfolios.
- To analyze the performance of portfolios on the time basis, compare the efficiency level of different methods applied.

The US stock market has been selected as the most popular market with the highest efficiency of economical laws. A set of 8 fundamental keys has been selected for the further investigation. The PCA and the AANN have been selected to compare the efficiency of different intelligent systems application on stock market.

PCA based portfolios were built using “PickStock” software. “Statistica” application and own created “autoassociative_ANN” script was used in order to separate respectively PCA and ANN based clusters of stocks. All portfolios created by “PickStock” were highly profitable, while stocks in the clusters built by “Statistica” were not diverging, and “autoassociative_ANN” was not able to extract any apparent cluster. This leads to conclusion, that AANN is still promising way to for stocks clustering and forecasting, however it requires further investigation based on technical factors.

IVADAS

Egzistuoja įvairiausių strategijų, kurios naudojamos investuojant į vertybinių popierių rinką. Kartais sėkmę vertybinių popierių rinkoje patiria netgi investuotojai, kurie vertybinių popierių portfelį formuoja visiškai akiai, neatlikdami jokios rinkos analizės. Tačiau statistika rodo, kad apie 90% individualių investuotojų, kurie nesinaudoja jokia analize ir pirkimo/pardavimo operacijas atlieka vadovaudamasi nuojauta, per metus praranda savo indėlius ir yra išmetami iš rinkos (Simutis, 2006).

Stabilų pelną vertybinių popierių rinkoje pasiekia daugiausiai tie investuotojai, kurie nemažą dėmesį skiria bendrai vertybinių popierių rinkos analizei ir atskirų kompanijų finansinės veiklos rodiklių studijavimui.

Šios veiklos supaprastinimui yra kuriamos įvairios programinės įrangos sistemos vertinančios kompanijas skirtingais metodais pagal įvairius rodiklius.

JAV vertybinių popierių birža tarsi aukso kasykla traukia investuotojus iš viso pasaulio, taigi šios intelektualinės sistemos atrenkančios akcijas pagal tam tikrus rodiklius ar prognozuojančios akcijų kitimo tendencijas yra labai paklausios. Sistemos paklausumą didžiaja dalimi lemia jos istoriniai pelningumo rodikliai.

JAV finansų sektorius į dirbtinių neuroninių tinklų tyrimus investavo daugiausiai lėšų po JAV gynybos departamento (Ward, 2006). Nobelio premijos 2001 ir 2002 metais ekonomikos srityje buvo skirtos už darbus neigiančius efektyvios rinkos hipotezę, taip dar kartą pažymint, koks svarbi yra ši tema.

Vertinant finansinių rinkų prognozavimo ir prognozavimo modelių taikymo galimybes galima išskirti du kontraversiškus požiūrius šia tematika:

- efektyvios rinkos požiūris,
- prekeivio ar investuotojo požiūris.(Norgaard, 2003)

Pagal efektyvios rinkos hipotezę galima teigti, jog akcijų ir valiutų kursų svyravimai yra neprognozuojamas, atsitiktinis reiškinys. Tokioje rinkoje nėra neįvertintų ar perversintų akcijų, o rinkos prognozavimo uždavinys prilyginimas amžinojo variklio paieškai.

Visiškai priešingas yra investuotojų požiūris į rinką. Daugelis jų įrodo, kad su savais instrumentais sugeba pasiekti aukštesnį pelną nei rinkos vidurkis. Garsiausi investuotojai, tokie kaip Warren Buffet ir Peter Lych turėdavo išpūdingus stabilius rodiklius – W.Buffet siekė vidutiniškai 23,6% metinio pelno, P.Lych – 29%, kai tuo tarpu vidutinis S&P500 tuo laikotarpiu metinis augimas siekė 11,8%. Gerus rodiklius rodo ir dirbtinio intelekto sistemos. Pavyzdžiui vidutinis metinis tdmresearch.com suformuoto portfelio augimas yra 19,3%, kai to laikotarpio S&P500

indeksas padidėja vos 8,3%. Be abejonės, tdmresearch.com neviešina kokiais konkrečiais metodais yra pasirenkamos perkamos/parduodamos akcijos, tačiau pats faktas, kad naudojant dirbtinio intelekto sistemas vertybinių popierių rinkoje galima pasiekti puikių rezultatų masina siekti to paties. (Simutis, 2006)

Kadangi problema yra išties aktuali ir įdomi, šia tematika yra paskelbta daug publikacijų mokslinėje spaudoje, tačiau dėl dviejų visiškai skirtingų požiūrių, vienos nuomonės neprieinama jau dešimtmečius.

Vilniaus universitete išsamūs autoasociatyvinių neuroninių tinklų taikymo vertybinių popierių kainų prognozei tyrimai atlikti nebuvo. Pasak prof. R.Simučio, buvo tik šiek tiek pasižiūrta bandant sugrupuoti kompanijas į atskirus klasterius naudojant principines komponentes. (Simutis, 2006)

Atlikus tolimesnę informacijos paiešką kituose Lietuvos elektroniniuose informacijos šaltiniuose nerasta laisvai prieinamos informacijos apie panašius tyrimus. Prieita išvados, kad tokie tyrimai arba atliekami nebuvo arba jų rezultatai viešai neskelbiami.

Darbo tyrimo **objektas** yra vertybinių popierių rinkų analizė taikant klasterizavimo metodus.

Darbo **tikslas** – išanalizuoti autoasociatyvinių neuroninių tinklų taikymo galimybes JAV akcijų rinkoje kotiruojamų akcijų klasifikavimui.

Siekiant iškelto tikslo, yra sprendžiami tokie **uždaviniai**:

1. Išnagrinėti bendras akcijų rinkos klasterizavimo galimybes, išnagrinėti fundamentinius ir techninius rodiklius, atrinkti kurie iš jų gali būti svarbūs tolimesnei analizei
2. Išnagrinėti skirtingas dirbtinių neuroninių tinklų realizacijas, parinkti tinkamiausias vertybinių popierių klasterizavimo uždaviniams spręsti
3. Palyginti įvairias programines sistemas atliekančias neuroninių tinklų skaičiavimus, prognozuojančias akcijų kainas, klasterizuojančias akcijas
4. Naudojantis nagrinėtais algoritmais, atrinkti kelis kriterijus kurie gali suformuoti akcijų klasterius
5. Pagal atrinktus algoritmus ir programinę įrangą atlikti istorinę akcijų rodiklių analizę, suformuoti investicinius portfelius
6. Palyginti gautus rezultatus su pradiniais, patikrinti bandytų metodų veiksmingumą

Pirmojo etapo metu iš aibės kriterijų atrenkami svarbiausi kriterijai kurie bus naudojami eksperimentinio tyrimo metu.

Antrajame etape detaliai nagrinėjami įvairūs neuroniniai tinklai, svarstomos jų panaudojimo vertybinių popierių klasterizavimui galimybės.

Trečiajame etape atliekama lyginamoji programų analizė, parenkamos ar sukuriamos programos, kurios bus naudojamos eksperimentinio tyrimo metu.

Vykdant ketvirtąjį etapą pirmuosius kelis kartus, vertybinių popierių rangavimo kriterijai bus parenkami atsitiktiniu būdu. Vėliau kriterijai bus atrenkami atsižvelgiant į praėjusių iteracijų rezultatus. Pažymėtinas šio etapo neapibrėžtinumas.

Penktojo etapo rezultatas – suformuoti geriausių ir prasčiausių akcijų portfeliai. Skirtingų akcijų skaičius ir jų dalis portfelyje iš anksto nežinomi, jie bus apskaičiuoti eksperimentinio tyrimo metu.

Šeštajame etape vertinami portfelių rezultatai, lyginamas skirtingų metodikų veiksmingumas, tikrinama ar pasitvirtino hipotezė.

Tiriamą vertybinių popierių sąrašą sudarys JAV vertybinių popierių biržoje prekiaujamos akcijos.

Vertybiniai popieriai bus ranguojami pagal įvairius jų vertinimo kriterijus, kriterijų kombinacijas, pradžioje juos pasirenkant atsitiktinai, vėliau pasirinkimai bus įtakojami tarpinių tyrimo rezultatų.

Uždavinių sprendimui bus naudojamos šios metodikos:

- Principinių komponentų metodas
- Autoasociatyvinių neuroninių tinklų metodas

Kiekviena duomenų imtis turi būti apdorota naudojantis kiekviena aukščiau aprašyta metodika.

Atliekant tyrimą bus naudojamas kiekybinis tyrimo metodas — kiekybinė duomenų analizė: matematinio, statistinio apdorojimo metodai. Skaičiavimai atliekami šiomis programomis:

- Autoassociative_ANN darbo metu sukurta programa autoasociatyvinių neuroninių tinklų klasterizavimui, veikia MATLAB terpėje.
- PickStock – pervertintų/neįvertintų akcijų klasterių atrinkimas naudojant PCA.
- Statistica – statistinės analizės priemonių paketas realizuojantis PCA.

Kiekvienos duomenų imties parinkimas, apdorojimas ir analizė atliekama kas mėnesį. Numatomas tyrimo laikas – 1,5 metų (2006/10 – 2008/04).

Nustatyta, kad autoasociatyvinių neuroninių tinklų taikymas vertybinių popierių klasterizavimui yra teoriškai pagrįstas metodas, kurio konkrečios taikymo galimybės reikalauja tolimesnių tyrimų, nes pirminiai eksperimentiniai tyrimai neišskyrė aiškios akcijų klasterizacijos.

Straipsnis šia tematika „Autoasociatyvinių neuroninių tinklų taikymas vertybinių popierių kainų prognozei“ buvo pristatytas studentų mokslinėje-praktinėje konferencijoje „Akademinių

jaunimo siekiai: Ekonomikos, Vadybos ir Technologijų įžvalgos“ 2008 05 09, Klaipėda, Vakarų Lietuvos verslo kolegija.

Eksperimentinio tyrimo metu nepatogumu sukėlė tai, kad kompanija investor.reuters.com nutraukė PowerScreener palaikymą. PowerScreener teikiamais istoriniais duomenimis buvo grįstas kuriamos programos „Autoassociative_ANN“ veikimas, taigi teko ieškoti alternatyvaus duomenų šaltinio.

Neuroniniais tinklais grįsti algoritmai turi taikymo apribojimų apie kuriuos kalbama skyriuje 2.4 „Algoritmų taikymas vertybinių popierių klasifikavimui“.

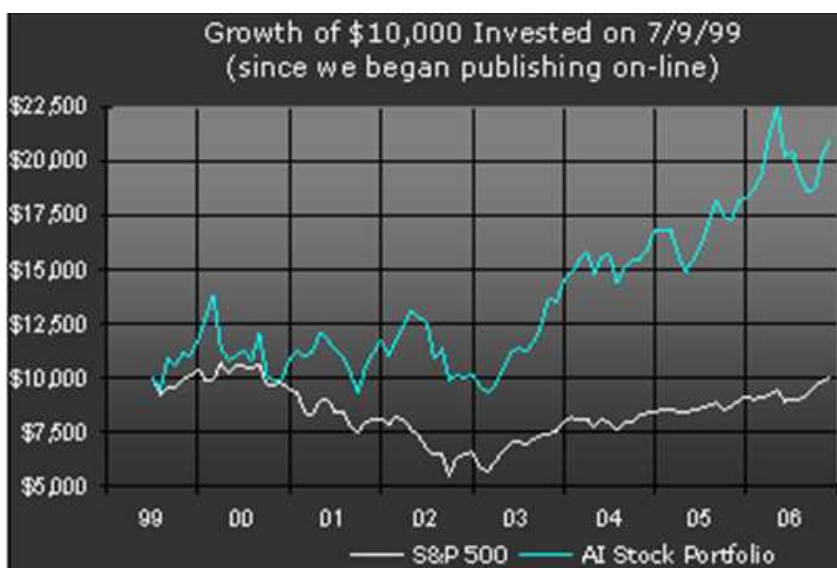
Darbą sudaro įvadas, skyriai, išvados, naudotos literatūros sąrašas, priedai. Darbo apimtis – 64 puslapiai, jame yra 7 lentelės ir 36 paveikslai. Bibliografinį aprašą sudaro 32 šaltiniai.

1. VERTYBINIŲ POPIERIŲ RINKOS FAKTORIAI

Pirmajame darbo etape literatūros pagrindą sudarė įvairūs šaltiniai sprendžiantys duomenų klasifikavimo uždavinius skirtingais metodais ir įvairiose srityse. Pereinant prie intelektinių sistemų taikymo finansų rinkose, buvo analizuota doc. R.Simučio dėstomo modulio „Intelektualios sistemos finansų rinkoje“ paskaitų medžiaga (Simutis, 2006). Joje išsamiai analizuojami skirtingi vertybinių popierių prognozės metodai, daugiausiai dėmesio sutelkiant ties intelektinėmis sistemomis. Pateikiama teorinė medžiaga yra realizuojama ir programiškai, MATLAB programinės įrangos pagalba.

Gilesnei teorinei pažinčiai gerai pasitarnavo George T. Albanis ir Roy T. Batchelor straipsnis „Predicting high performance stocks using dimensionality reduction techniques based on neural networks“ („akcijų kainų prognozė naudojant dimensijų mažinimo metodus grįstus neuroniniais tinklais“ – angl.). Svarbiausia šio straipsnio dalis atliekamam tyrimui yra palyginamieji skirtingų metodikų – principinių komponentų bei neuroninių tinklų testai atliekami įvairiais aspektais. Straipsnyje išryškėja netiesinių metodų (neuroninių tinklų) pranašumas prieš tiesinius (principines komponentes), pavaizduojama kintamųjų rinkinio parinkimo svarba, deja konkretūs kintamųjų rinkiniai nenurodomi.

Su realiu neuroninių tinklų taikymu finansų rinkose supažindina tdmresearch.com internetiniame puslapyje patalpinti straipsniai – „Why Our Investment Strategy Works“ bei „What is Neural Network Forecast?“. Ši kompanija teikia portfelio formavimo paslaugas naudodama jų pačių sukurtas prognozavimo sistemas grįstas neuroniniais tinklais. Reklaminiais tikslais pateikiami įspūdingi veiklos rodikliai ženkliai viršijantys S&P500 indekso metinį augimą (1 pav.).



Šaltinis: What is Neural Network Forecast? Why does it seem to change from time to time?

1 pav. Kompanijos TDMresearch portfelio ir S&P500 indekso palyginimas

Kompanijos paslaugos mokamos, taigi natūralu, kad sukurti algoritmai viešinami nėra, tačiau pats faktas, kad, prognozuojant akcijų kainas naudojant neuroninius tinklus, buvo pasiekta gerų rezultatų, suteikia papildomo stimulo ir motyvacijos magistriniam darbui.

N. Basalto, R. Bellotti ir kiti straipsnyje „Clustering stock market companies via chaotic map synchronization“ („Kompanijų akcijų klasterizavimas sinchronizuojant chaosinius žemėlapius“ – angl.) apžvelgia akcijų klasterizavimo galimybes. Vėlgi prieinama prie teigiamos išvados, kad sudarinėjant portfelį remiantis klasterizuotomis akcijomis, galima sumažinti rizikos laipsnį, lengviau atskirti klasterio tendencijas. Tačiau, kaip ir anksčiau apžvelgtuose straipsniuose, nėra užsimenama apie panaudotus konkrečius kintamuosius.

Lou Mendelsohn straipsnis „Using Neural Networks For Financial Forecasting“ („Neuroninių tinklų panaudojimas finansų prognozei“ – angl.) reziumuoja paties autoriaus atliktą neuroninių tinklų tyrimą. Straipsnyje akcentuojama, kad neuroninių tinklų analizė dėl įvairiausių priežasčių negali būti kaip vienintelis įrankis apsprendžiantis prekybos akcijomis veiksmus. Naudojant neurotinklus šablono atpažinimui didelę įtaką turi žmogiškasis faktorius – skirtingi asmenys turi skirtingą rizikos supratimą, todėl jie tas pačias praeities situacijas („šablonus“) vertins skirtingai, dėl to neuroninis tinklas bus apmokytas skirtingai. Siūlomos kelių skirtingai apmokytų neuroninių tinklų sistemos. Žvelgiant į ateitį, numatomas neuroninių tinklų, ekspertinių sistemų ir genetinių algoritmų populiarumo augimas, tačiau dar kartą pabrėžiama, kad bet kuri šių sistemų tėra tik vienas įrankis, kurio nepakanka pastovaus pelno užtikrinimui. (L. Mendelsohn)

Gilesnei autoasociatyvinių neuroninių tinklų analizei buvo nagrinėtas autoriaus Mark A. Kramer straipsnis „Nonlinear principal component analysis using autoassociative neural networks“ (Netiesinių principinių komponentų analizė naudojant autoasociatyvinius neuroninius tinklus – angl.). Straipsnyje iškeliamas duomenų multidimensiškumo problema, aptariami dabartiniai būdai šiai problemai spręsti – PCA. Metodus iliustruojamas praktiniais PCA panaudojimo atvejais – duomenų tikrinimas, kokybės kontrolė, duomenų vizualizacija, prognozės. Kaip pagrindinis netiesinių PCA pranašumas (NLPCA) prieš tiesines PCA yra iškeliamas didesnis netiesinių PCA tikslumas, didesnis suspaudimo laipsnis (galimybė suspausti duomenis į mažiau principinių komponentų išlaikant tą patį klaidų santykį). Smulkiai išdėstomas netiesinių PCA veikimas paremtas dirbtiniais neuroniniais tinklais. Pabaigoje testuojamas PCA ir netiesinių PCA veikimas.

PCA, ANN ir NLPCA palyginamojo testo rezultatai

Metodas	Klaidų lygis	FPE
PCA	27.8	0.0708
ANN, be tarpinio lygio neuronų	26.4	0.0708
NLPCA, tarpinio lygio neuronai		
2	10.5	0.0318
3	1.35	0.00444
4	0.348	0.00124
6	0.336	0.00142
8	0.307	0.00154
10	0.302	0.00183

Šaltinis: Sudaryta autoriaus pagal KRAMER, M. (1991) Nonlinear principal component analysis using autoassociative neural networks, p. 241

Akivaizdžiausiai NLPCA pranašumas matomas lyginant klaidų skaičių – PCA ir dirbtinis neuroninis tinklas be pereinamųjų sluoksnių veikia skaičiavimus atliko su labai aukštu klaidų lygiu, o NLPCA su 3 ir daugiau principinių komponentų pasirodė žymiai geriau. (Kramer, 1991)

1.1 Vertybinių popierių rinkos prognozavimo galimybės

Pirmas natūraliai kylantis klausimas – ar apskirtai įmanoma finansinių rinkų prognozė? Kaip darbo įvade buvo minėta, egzistuoja dvi kontraversiškos nuomonės šiuo klausimu – efektyvios rinkos hipotezė ir investuotojų požiūris.

1.1.1 Efektyvios rinkos hipotezė

Efektyvios rinkos hipotezė (ERH) teigia, jog bet kuriuo metu akcijos kaina atspindi visą žinomą informaciją apie tą akciją. Kadangi šią informaciją žino dauguma rinkos dalyvių, kainos svyravimai yra atsitiktiniai, ir priklausomi tik nuo naujos atsitiktinės informacijos atsiradimo. Taigi, akcijos kainos kitimas yra atsitiktinis procesas ir todėl investuotojui negali būti įmanoma įveikti rinkos.

Nors praktikoje šis griežtas teiginys atrodo neteisingas, ERH iki šiol nėra nei atmesta nei patvirtinta. Remiantis neuroniniais tinklais, šios hipotezės taip pat negalima nei atmesti, nei patvirtinti, kadangi neuroninio tinklo efektyvumas labai priklauso nuo to kaip jis buvo ištreniruotas (Lawrence, 1997).

ERH yra svarbi, kadangi ji prieštarauja kitoms analizės formoms. Jei išties neįmanoma įveikti rinkos, tai tada teoriškai tiek techninė, tiek fundamentinė analizė negali duoti geresnių rezultatų nei atsitiktinis akcijų parinkimas. Tačiau faktas, kad nemažai rinkos dalyvių sugeba aplenkėti rinką reiškia, jog ERH praktikoje yra nepasitvirtina. ERH galėtų būti teisinga idealiomis

sąlygomis, jei visa informacija būtų prieinama visiems rinkos dalyviams tuo pačiu metu, tačiau šiandien rinkoje yra ir išskirtinių žaidėjų, besinaudojančių vidine informacija ar kitomis priemonėmis (Lawrence, 1997).

1.1.2 Chaoso teorija

Netiesinių dinaminų sistemų, kaip pavyzdžiui, akcijų rinka, modeliavimui gali būti naudojama chaoso teorija. Chaoso teorija analizuoja procesus darydama prielaidą, jog dalis proceso vyksta atsitiktinai, tačiau kita jo dalis yra determinuota. Chaosas yra netiesinis procesas kuris atrodo esąs atsitiktinis. Chaoso teorija yra bandymas parodyti, jog vizualiai atsitiktiniame procese gali egzistuoti tam tikra tvarka. Darant prielaidą jog vertybinių popierių rinka yra chaotiška, o ne vien atsitiktinė, chaoso teorija paneigtų ERH.

2001m. Nobelio premijos ekonomiką srityje laureatas Joseph E. Stiglitz teigia, kad rinkos dalyviai turi skirtingą informaciją apie rinką, ir šie skirtumai daro didelę įtaką ekonomikos funkcionavimui. Jis savo darbu, kaip ir 2002m. Nobelio premijos ekonomikos srityje laureatas Daniel Kahneman, neigia efektyvios rinkos hipotezę. (Stiglitz, 2002)

Chaotiška sistema iš esmės yra deterministinių ir atsitiktinių procesų darinys. Deterministinius procesus galima aprašyti naudojant regresiniais metodais, o atsitiktiniai procesai gali būtų aprašomi paskirstymo funkcijom. Taigi, naudojant vien deterministinius arba vien statistinius metodus, negalima pilnai apibūdinti chaotiškos sistemos. Neuroniniai tinklai gali dirbti vienu metu tiek su deterministinėm, tiek su atsitiktinėm funkcijom, todėl jie gali būti naudojami chaotiškų sistemų modeliavimui.

1.2 Fundamentinių rodiklių atranka

Fundamentinioji analizė – visuma finansinės analizės būdų, metodų, kuriais siekiama nustatyti „tikrąją“ finansinių vienetų (vertybinių popierių) vertę. Priešingai nei techninė analizė fundamentalioji iš esmės remiasi ne biržų kursų duomenimis, bet įmonės veiklos, ūkio šakos duomenimis, vadinamaisiais fundamentaliaisiais duomenimis (pamatine informacija). Fundamentaliosios analizės tikslas – nustatyti vertę, prognozuoti jos kitimą.

Tokiai analizei svarbi balansų analizė, bendrovės finansiniai duomenys bei veiklos rodikliai; analizuojami aktyvai ir pasyvai, skolos nuosavybės santykis, dividendai, pajamos akcijai, pardavimų rodikliai, rinkos dalis, produktai, marketingas, vadyba (pvz., kokybės vadyba), dividendų padengimas (assets and dividend coverage). („Pinigai daro pinigus“, 2007)

1.2.1 Pelnas vienai akcijai (EPS)

Pelnas vienai akcijai (angl. Earnings per share, EPS) – finansinis rodiklis, lygus įmonės grynojo pelno daliai padalintai iš vidutinio skaičiaus paprastųjų akcijų.

Pelnas vienai akcijai yra vienas iš pagrindinių rodiklių, naudojamų lyginti įmonės, kurios akcijos laisvai platinamos Vertybinių popierių biržoje, patrauklumą investuotojams. Tarptautiniai finansinės atskaitomybės standartai IAS/IFRS ir US-GAAP nustato pelno apskaičiavimo vienai akcijai ir atskaitomybės reikalavimus metinėje atskaitomybėje.

$$EPS = \frac{I_N - D_P}{S_A}$$

čia I_N – ataskaitinio laikotarpio tiesioginis pelnas,

D_P – per ataskaitinį periodą apskaičiuoti dividendai pagal privilegijuotas akcijas,

S_A – ataskaitiniu periodu esantis vidutinis paprastų akcijų skaičius.

Skaičiuojant S_A įtraukiami visi įprastų akcijų skaičiaus pasikeitimai per apskaitomą periodą.

Tuo atveju, kai apskaičiuotas EPS yra žemiau nulio, galima kalbėti apie nuostolį vienai akcijai.

Pelnas vienai akcijai apskaičiuojamas įvertinant apyvartoje esantį akcijų skaičių, vadinamas akcijos bazinis pelnas. Tai pat yra ir rodiklis apytikris pelnas akcijai (angl. diluted EPS), įvertinami galimi akcijų skaičiaus ir pelno pasikeitimai konvertuojant kitus vertybinius popierius į įprastas akcijas (sakykim, išleistų obligacijų keitimą į akcijas, ar dividendų išmokas pagal naujai privilegijuotas akcijas).

Apskaičiuojant EPS vietoj realaus švaraus pelno gali būti naudojamas laukiamas švarus pelnas, finansinių analitikų duomenys. Tokiu atveju kalbama apie akcijos laukiamą pelną (angl. forward EPS). („Pinigai daro pinigus“, 2007)

1.2.2 Akcijos kainos ir EPS santykis (P/E)

Akcijos kainos ir jos EPS santykis yra vienas populiariausių ir svarbiausių rodiklių investuotojui. Parodo kiek kartų akcijos kaina yra didesnė už jei tenkančią pelno dalį. Vadinasi, kuo didesnis P/E santykis, tuo pelningesnė akcija. Tai konservatyvus parametras, prieš dešimtmetį geru rodikliu buvo skaitoma reikšmė 10-15. Dabar kai kada investuotojai perka akcijas kai šis rodiklis didesnis už 500. („Pinigai daro pinigus“, 2007)

1.2.3 Grynojo pelno ir nuosavo (akcinio) kapitalo santykis (ROE)

Grynojo pelno ir nuosavo (akcinio) kapitalo santykis (ROE, return on equity). Renkantis akcijas ekspertai dažnai mėgsta kai šis rodiklis viršija 10 %.

1.2.4 Bendrasis likvidumas

Bendrasis likvidumas(Current ratio) = Trumpalaikis turtas / Trumpalaikiai įsiskolinimai

Pagal nusistovėjusią nuomonę laikoma, jog šis santykis turėtų būti artimas 2. Jei bendrasis likvidumas yra žemesnis nei 1, tai reiškia, jog įmonė nesugeba laiku atsiskaityti su kreditoriais, jei santykis smarkiai aukštesnis už 2, gali būti, kad įmonė nesugeba efektyviai panaudoti savo turimo turto.

1.2.5 Institucinių investuotojų dalyvavimas

Didelę dalį pasiūlos ir paklausos rinkoje sudaro instituciniai investuotojai (pvz., fondai, bankai, draudimo kompanijos). Kompanijai nebūtina turėti daug investuotojų, bet jų investicijos – tai stiprūs balsai už palaikymą. Pagrindinę taisyklę, pagal O'neal, sudaro akcijų paieška, kurios turi ne mažiau kaip 3 – 10 institucinių investuotojų ir kurių veiklos rodikliai aukštesni už vidutinį.

Bet institucinių investuotojų perviršis gali ir pakenkti. Kuomet įmonės akcijas valdos per daug investicinių investuotojų, galima sakyti, jog investuoti jau vėlu. Jeigu instituciniai investuotojai valdo 70 - 80% išleistos akcijos apyvartos, tai vargu ar galima tikėtis žymesnio kainos kilimo. Negana to, bet kurios nepalankios investicijos sugeba iššaukti aktyvų akcijų išpardavimą.

O'neal mano, jog idealus akcijos pirkimo periodas yra tada, kai ją visiškai neseniai įsigijo keletas svarbių institucinių investuotojų, bet ji dar nėra labai populiari, kad visos institucinės organizacijos pradėtų ją aktyviai pirkti.(„Pinigai daro pinigus“, 2003)

1.2.6 Kompanijos akcijų dalis tenkanti įmonės darbuotojams.

Investiciniai fondai stengiasi pirkti akcijas, kai šis rodiklis viršija 10-20 procentų. Šis rodiklis patikimas garantas, kad firmos darbuotojai suinteresuoti ir tikintys firmos klestėjimu.(Simutis, 2006)

1.2.7 Kompanijos darbuotojų aktyvumas perkant/parduodant savos kompanijos akcijas.

Jei kompanijos darbuotojai aktyviai perka savo kompanijos akcijas, tai didelė tikimybė, kad akcijų kaina gali artimiausiu metu kilti.

1.2.8 Fundamentinių rodiklių analizės išvados

Akivaizdu, kad vieno rodiklio neužtenka norint įvertinti konkrečią akciją, ar palyginti ją su kitoms. Be to, ir skirtingų rodiklių naudojimas negali vienareikšmiškai apibūdinti akcijos - įvairių rodiklių reikšmės atskiroms akcijoms yra skirtingos ir kontraversiškos.(Simutis, 2006)

Efektyviam fundamentinių rodiklių panaudojimui jie yra grupuojami pagal ekspertų metodikas (CANSLIM rinkinys - <http://www.investors.com/learn/C.asp>) arba savarankiškai ranguojami nustatant kiekvieno rodiklio svarbą. Galimas tiek tiesinis, tiek netiesinis rangavimas (Kohoneno žemėlapiai, saviorganizuojantys neuroniniai tinklai).

William J. O'Neil, Chairman & Founder, Investors Business Daily pateikiamas toks svarbiausių techninių ir fundamentinių rodiklių šešetas (SmartSelect - <http://www.investors.com/ibddifference/>):

1. EPS
2. Lyginamoji jėga (Relative Price Strength - RS)
3. Pramonės šakos lyginamoji jėga (Industry Group Relative Strength)
4. Pardavimai + pelno marža (Sales+Profit Margins+ROE - SMR)
5. Pirkimų/pardavimų santykis (Accumulation/Distribution)
6. Bendra penkių aukščiau esančių rodiklių kombinacija (Composite Rating)

1.3 Techninių rodiklių atranka

Techninė analizė apima vertybinių popierių rinkos konjunktūros ir kursų kitimo tendencijų nagrinėjimą, bei būsimų rinkos tendencijų prognozavimą. Šios analizės galimybės vertybinių popierių rinkose vertinamos labai prieštaringai, nuo per daug optimistinių nuostatų iki pesimistinių.

Per ilgametę vertybinių popierių rinkos istorija yra sukurta eilė techninių indikatorių, kuriais bandoma apibūdinti akcijų kainų pokyčiai bei juos prognozuoti.(Simutis, 2006)

1.3.1 Lyginamosios jėgos indeksas (RSI)

Techninis indikatorius Lyginamosios Jėgos Indeksas (Relative Strength Index, RSI) – tai osciliatorius, sekantis iš paskos kainos ir svyruojantis diapazone nuo 0 iki 100. Įvesdamas Relative Strength Index, J. Wilder rekomendavo naudoti jo 14 periodų variantą. Vėliau paplito ir 9 bei 25 periodų indikatoriai.

Vienas iš populiarių Relative Strength Index indikatoriaus analizės metodų yra divergencijų, su kuriomis kaina sudaro naują maksimumą, o RSI nepavyksta įveikti prieš tai buvusio

maksimumo, paieška. Tokia divergencija byloja apie kainų apsisukimo galimybę. Jeigu po to indikatorius pasuka į apačią ir nusileidžia žemiau savo dugno, jis užbaigia vadinamąjį „nepavykusį mostą“ (failure swing). Šis nepavykęs mostas yra laikomas greito kainų apsisukimo įrodymu.

Grafikų analizėje išskiriami tokie Relative Strength Index signalai:

- Viršūnės ir pagrindai
- Relative Strength Index indikatorius viršūnės paprastai formuojasi aukščiau 70, o pagrindai – žemiau 30, ir dažniausiai aplenkia viršūnių ir pagrindų darinius kainų grafike.
- Grafiniai modeliai
- Relative Strength Index dažnai sudaro grafinius modelius – tokius kaip „galva ir pečiai“ arba trikampiai, kurių kainų grafike gali ir nesimatyti.
- Nepavykęs mostas (palaikymo arba pasipriešinimo lygio proveržis)
- Atsitinka, kai Relative Strength Index pakyla aukščiau prieš tai buvusio maksimumo (piko) arba nusileidžia žemiau prieš tai buvusio minimumo (dugno).
- Palaikymo ir pasipriešinimo lygiai
- Relative Strength Index indikatorius grafike palaikymo ir pasipriešinimo lygiai matomi net aiškiau nei kainų grafike.
- Divergencijos

Kaip jau buvo sakyta aukščiau, divergencijos atsiranda, kai kaina pasiekia naują maksimumą (minimumą), bet jis nepasitvirtina nauju maksimumu (minimumu) RSI grafike. Tada dažniausiai įvyksta kainų korekcija RSI indikatorius judėjimo kryptimi.



Šaltinis: spekuliantai.lt („Pinigai daro pinigus“, 2003)

2 pav. RSI indeksas

RSI Apskaičiavimas:

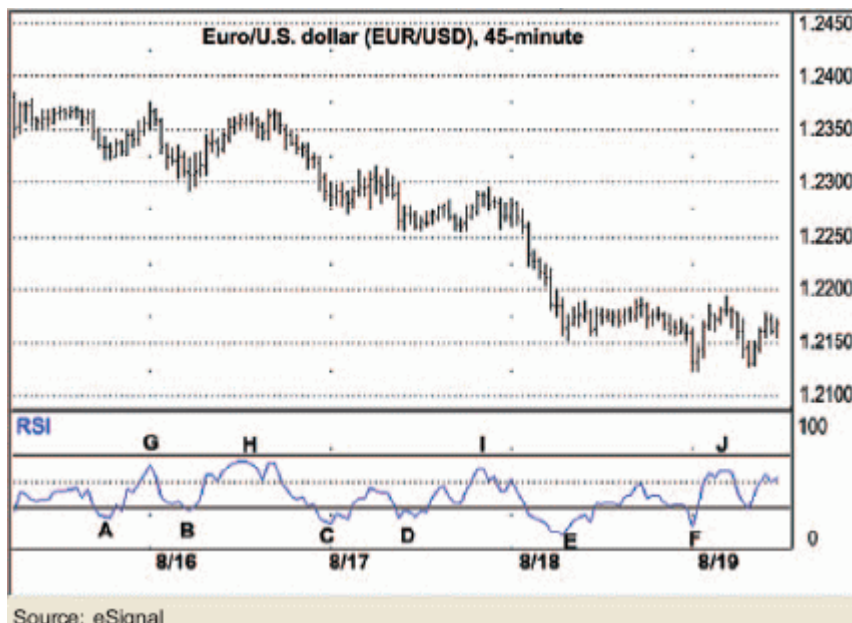
$$RSI = 100 - (100 / (1 + U/D))$$

čia U — vidutinis pozityvių kainos pokyčių skaičius

D — vidutinis negatyvių kainos pokyčių skaičius

Jei RSI yra tarp 70 ir 30 pirkti nepatartina – tuo metu bulių ir meškų jėgos yra lygios – patariama laukti. Atsargus prekyautojas vietoj 70 ir 30 lygių gali nustatyti kritinėmis reikšmėmis lygius 80 ir 20. Bet tada, kai periodo reikšmės bus didelės, RSI grafikas per retai kirs kritinius lygius.

RSI nebus efektyvus trendų rinkoje (rinka, kur kaina juda tam tikra kryptimi: aukštyn/žemyn). Pavyzdžiui, jeigu kaina kryptingai juda žemyn, RSI yra perpardavimo zonoje (taškai A, B, C, D, E, F), o tai reiškia, kad pagal indikatorius signalus, reikia pirkti:



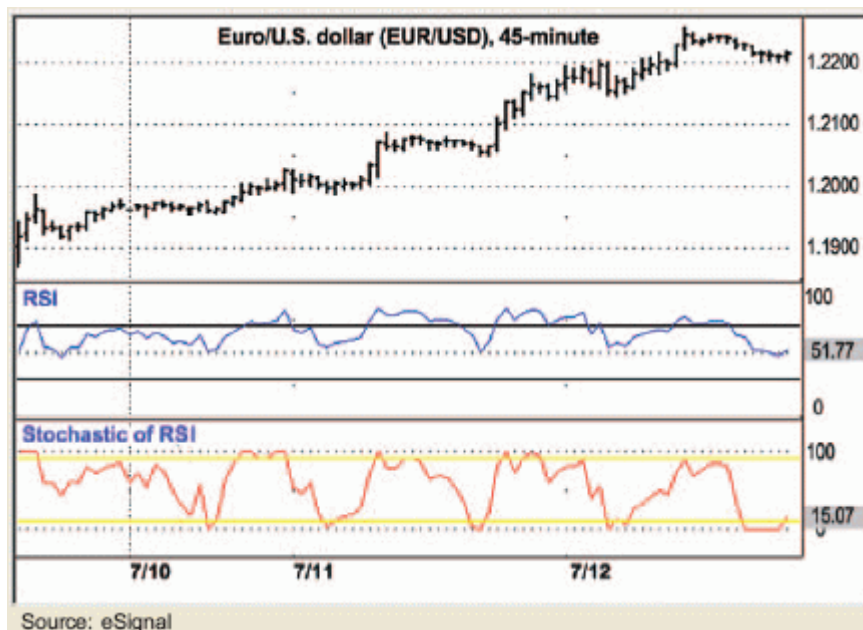
Šaltinis: spekuliantai.lt („Pinigai daro pinigus“, 2003)

3 pav. RSI trendų rinkoje

O pirkti tektų prieš tendą, kas, be abejo, nepriimtina. Tokiu būdu indikatorius negeneruoja signalų pardavimams, kurių panaudojimas leistų išgauti pelną iš poros judėjimo mažėjančiame (šiuo atveju) trende.

Šios problemos sprendimas apima stochastinio osciliatoriaus RSI arba kitų trendų indikatorių (MACD, SMA/EMA, DMI) panaudojimą.

Žemiau pateiktame grafike (pav. 4) matomas kylantis trendas: RSI formuoja daug klaidingų perpirkimo signalų ir ne vieno perpardavimo, stochastinis RSI, atvirkščiai – duoda kelis gerus signalus pirkimui:



Šaltinis: spekuliantai.lt („Pinigai daro pinigus“, 2003)

4 pav. RSI ir stochastinis RSI kylančių trendų rinkoje

O štai (pav. 5) mažėjančio trendo pavyzdys (USDCHF M5, kovo 16), kur RSI nerodo perpirkimo. Stochastinis RSI duoda ryškius signalus pardavimui:



Šaltinis: spekuliantai.lt („Pinigai daro pinigus“, 2003)

5 pav. RSI ir stochastinis RSI besileidžiančių trendų rinkoje

1.3.2 Slankiųjų vidurkių kirtimosi indikatorius (MACD)

Techninis slankiųjų vidurkių kirtimosi indikatorius (Moving Average Convergence/Divergence, sutrumpintai MACD) — tai dinamiškas indikatorius, parodantis dviejų slankiųjų vidutinių kainų sąryšį. MACD vaizduojamas kaip skirtumas tarp dviejų eksponentinių slankiųjų vidurkių (EMA) su 12 ir 26 dienų periodais. Tam, kad galima būtų aiškiai pavaizduoti palankius pirkimo arba pardavimo momentus MACD historigrame (HISTOGRAM) brėžiama signalinė linija (SIGNAL) – 9 dienų periodo slankusis vidurkis:



Šaltinis: spekuliantai.lt („Pinigai daro pinigus“, 2003)

6 pav. MACD indikatorius

MACD efektyviausias, kai rinkos svyravimo amplitudė „prekybos koridoriuje“ yra didelė. Dažniausiai naudojami MACD signalai — susikirtimai, perpirkimo/perpardavimo būsenos ir išsišakojimai.

Perpirkimas/perpardavimas: kai signalinė linija yra virš histogamos (t.y. MACD auga), tai reiškia, kad tiriamojo instrumento kaina greičiausiai per daug pakelta ir greitai sugrįš prie realiesnio lygmens.

Pagrindinė prekyavimo su MACD taisyklė paremta indikatoriaus (HISTOGRAM) persikirtimu su signaline linija (SIGNAL): kai Moving Average Convergence/Divergence nusileidžia žemiau signalinės linijos – reikia parduoti, o kai kyla aukščiau jos – pirkti.

MACD nulinės linijos perkirtimai aukštyn/žemyn taip pat naudojami kaip pirkimo/pardavimo signalas.

Pažymėsiu, kad prekiauti su MACD reikėtų ne žemesniuose nei M30 – H1 laikotarpiuose. Trumpiesiems periodams, tokiems kaip M1, M5, M15 – MACD netinka.

Galima išskirti šiuos MACD signalus:

- reikšmės, sekančios signalinės linijos (SIGNAL) maksimumą arba minimumą – signalas sandorio sudarymui;

- jeigu MACD histograma (HISTOGRAM) yra žemiau nulio, trendas yra krintantis, jei aukščiau – kylantis.
- signalinės linijos konvergencija/divergencija – stipriausias trendo silpnumo požymis: geriau sulaukti aiškesnio signalo.
- jeigu MACD yra žemiau nulio ir konvergencijos nėra, o MACD histograma (HISTOGRAM) perkerta signalinę liniją (SIGNAL) iš apačios į viršų, galima tikėtis kainų atoslūgio į viršų.
- jeigu MACD yra aukščiau nulio ir divergencijos nėra, o MACD histograma (HISTOGRAM) perkerta signalinę liniją (SIGNAL) iš viršaus į apačią, galima tikėtis kainų atoslūgio į apačią.

Divergencija (pav. 7) atsiranda, kai naujas kainos pikas nėra patvirtintas naujo indikatorius piko, t. y. kainos pikas aukštesnis už prieš tai buvusį, o atitinkamas indikatorius piko žemesnis už prieš tai buvusį. Tai byloja apie trendo silpnumą. („Pinigai daro pinigus“, 2003)



Šaltinis: spekuliantai.lt („Pinigai daro pinigus“, 2003)

7 pav. Divergencija

Konvergencija (pav. 8) atsiranda, kai naujas kainos dugnas (angl. market bottom) nėra patvirtintas naujo indikatorius dugno, t. y. sekantis kainos dugnas yra žemesnis už prieš tai buvusį, o sekantis indikatorius dugnas - aukštesnis už prieš tai buvusį. Tai taip pat byloja apie trendo silpnumą. („Pinigai daro pinigus“, 2003)



Šaltinis: spekuliantai.lt(„Pinigai daro pinigus“, 2003)

8 pav. Konvergencija

1.3.3 Slankusis tikrojo diapazono reikšmių vidurkis (ATR)

Techninis indikatorius Average True Range, ATR – tai rinkos kintamumo (angl. volatility) rodiklis. Šis indikatorius naudojamas kaip daugelio kitų indikatorių ir prekybos sistemų sudedamoji.

Average True Range indikatorius dažnai pasiekia aukštų reikšmių rinkos dugne po staigaus kainų kritimo, sukulto paniškų pardavimų. Žemos indikatoriaus reikšmės dažnai atitinka teigiamus horizontalaus judėjimo periodus, kurie stebimi rinkos viršūnėse ir konsolidacijos metu. Šį indikatorių galima interpretuoti pagal tas pačias taisykles, kaip ir kitus kintamumo indikatorius.

True Range (tikrasis diapazonas)– tai didžiausias iš šių trijų dydžių:

- skirtumas tarp einamojo maksimumo ir minimumo;
- skirtumas tarp prieš tai buvusios uždarymo kainos ir einamojo maksimumo;
- skirtumas tarp prieš tai buvusios uždarymo kainos ir einamojo minimumo.

Average True Range indikatorius - tai slankusis tikrojo diapazono reikšmių vidurkis:



Šaltinis: spekuliantai.lt („Pinigai daro pinigus“, 2003)

9 pav. ATR indiktorius

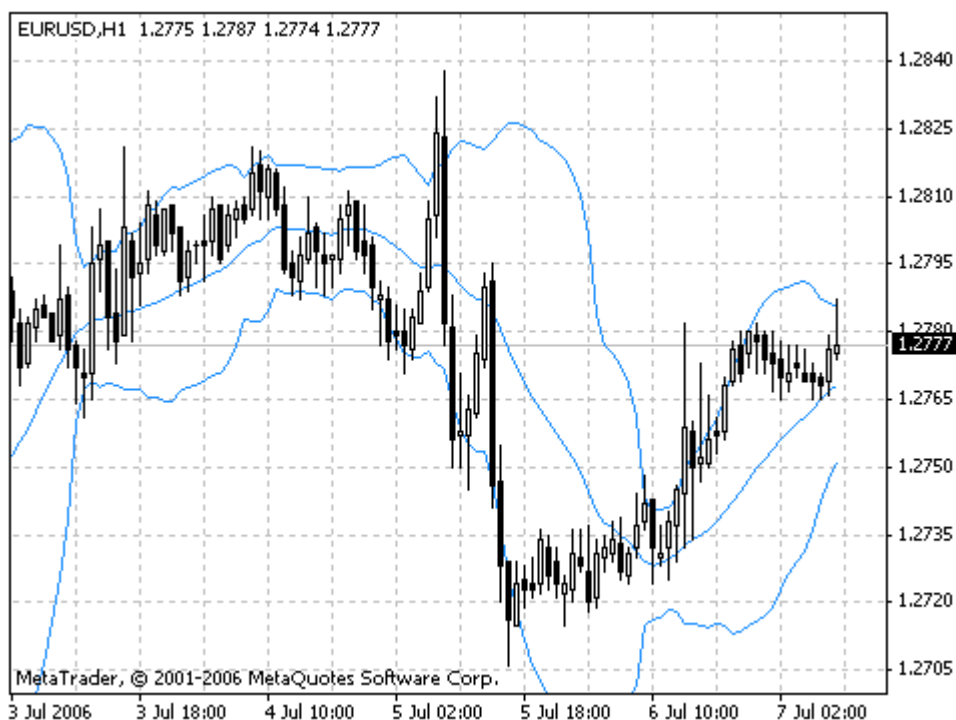
Prognozavimo su ATR principas yra toks: kuo aukštesnė indikatoriaus reikšmė, tuo didesnė trendo pasikeitimo galimybė; kuo jo reikšmė mažesnė, tuo silpnesnis trendo kryptingumas. („Pinigai daro pinigus“, 2003)

1.3.4 Bollingerio juostos

BB (Bollinger Bands, BB) panašios į Envelopes ir interpretuojamos vienodos sistemos pagalba:

Jie skiriasi tuo, kad Gaubiančiųjų kreivių (Envelopes) ribos yra aukščiau ir žemiau slankiojo vidurkio kreivės fiksuotame, procentais išreikštame atstume, o Bollinger Bands ribos remiasi atstumais, kurie lygūs tam tikram standartinių nukrypimų skaičiui (Standart deviation).

Dėl to, kad standartinio nukrypimo dydis priklauso nuo kintamumo, kreivės pačios reguliuoja savo plotį: jis didėja, kai rinka nepastovi ir mažėja stabilesnių periodų metu:



Šaltinis: spekuliantai.lt („Pinigai daro pinigus“, 2003)

10 pav. Boilingerio juostos

Boilingerio juostų panaudojimas:

1. Po kreivės linijų susispaudimo greičiausiai įvyksta stiprus rinkos judėjimas;
2. Rinkos išėjimas už kreivės ribų perspėja apie greitą apsisukimą (paprastai korekcinį atoslūgį);
3. Labai dažnai rinka, priėjusi prie slankiojo vidurkio (MA) sugrįžta atgal, o jau po to įvyksta jos pramušimas.
4. BB divergencija stebima, kai stiprėja veikiantis arba prasideda naujasis trendas. Divergencija, kai apimtys yra padidėjusios – geras trendo patvirtinimas. Slankusis vidurkis yra geras palaikymo lygis bulių rinkoje ir geras pasipriešinimo lygis meškų rinkoje.

Bollingerio juostų analizės pagrindu sprendimas priimamas, kai kaina arba pakyla aukščiau viršutinės pasipriešinimo linijos Bollinger Bands, arba nusileidžia žemiau apatinės palaikymo lygio linijos BB. Jeigu kainos grafikas svyruoja tarp šių dviejų linijų, tai reiškia, kad patikimų pirkimo/pardavimo signalų BB analizės pagrindu nėra. Pozicijos atidarymo sprendimas priimamas tik tada, kai kainos grafikas kerta Bollinger Bands liniją tam, kad sugrįžtų prie normalios būklės.

Kartais išėjimas už Bollinger Bands ribų reiškia „netikrą pramušimą“, t.y. kai kainos tik pabandė naują lygmenį ir iš karto sugrįžo atgal. Šiuo atveju pas jus atsiranda galimybė dirbti prieš trendą, tačiau atidžiai įvertinkite – ar tikrai pramušimas yra netikras. Geru patvirtinimu tokiais atvejais tampa apimtys rodiklis, kuris, jeigu pramušimas netikras, turi staigiai sumažėti.

Be abejo, Forex'e neįmanoma tiksliai nustatyti apimčių, o tai neleis visiems apimties indikatoriams (pavyzdžiui, Čaikino osciliatoriui) rodyti teisingas reikšmes. Todėl galima panaudoti tą patį RSI indikatorių, apie kurį mes kalbėjome praeitą kartą, kad galima būtų spręsti apie galimą apsisukimą (būtent apsisukimo patvirtinimui su netikru BB pramušimu), arba bendrą fundamentinį foną – neigatyvios arba atvirkščiai labai teigiamos ataskaitos pasirodymas gali akivaizdžiai patvirtinti, kad pramušimas – tikras. („Pinigai daro pinigus“, 2003)

1.3.5 Techninis indikatorius Gaubiančiosios kreivės (Envelopes)

Šis indikatorius sudarytas iš dviejų slankiųjų vidurkių, vienas iš kurių paslinkęs į viršų, o kitas – į apačią. Optimalaus santykinio kreivės poslinkio dydžio parinkimas apsprendžiamas rinkos kintamumo: kuo jis didesnis – tuo didesnis poslinkis:



Šaltinis: spekuliantai.lt („Pinigai daro pinigus“, 2003)

11 pav. Gaubiančiosios juostos

Envelopes apsprendžia viršutines ir apatines normalaus kainų svyravimo diapazono ribas. Signalas pardavimui atsiranda tada, kai kaina pasiekia viršutinę kreivės ribą, o signalas pirkimui – kai pasiekia apatinę.

Techninio indikatoriaus Envelopes panaudojimas pagrįstas natūralia rinkos elgesio logika: kai dėl labai aršių pirkėjų ar pardavėjų spaudimo kainos pasiekia ekstremalias reikšmes (t.y. viršutinės arba apatinės kreivės ribos), jos dažnai stabilizuojasi, grįždamos prie realesnių lygmenų.

Kaip jau buvo sakyta, toks pat principas naudojamas interpretuojant Bollinger Bands.(„Pinigai daro pinigus“, 2003)

1.3.6 Techninių rodiklių analizės išvados ir apibendrinimas

Nėra bendros nuomonės dėl techninių rodiklių panaudojimo vertybinių popierių rinkose – vieni giriasi puikiais rezultatais, pasiektais naudojantis šia metodika, kitiems tai panašu į būrimą iš kavos tirščių.

Vienas sudėtingiausių uždavinių sprendžiant nagrinėjant techninius rodiklius, tai jų formalizavimas naudojant skaitinio intelekto metodus (neryškios aibės, neuroniniai tinklai, klasifikatoriai).

1.4 Vertybinių popierių rinkos prognozavimo galimybių analizės išvados

Atlikus rinkos efektyvumo prielaidų analizę nuspręsta remtis praktiškai įrodyta chaoso rinkos teorija leidžiančia tikėtis teigiamų tyrimo rezultatų.

Išnagrinėjus fundamentinius ir techninius akcijų rodiklius, nuspręsta toliau remtis fundamentiniais rodikliais, kadangi neišspręsta techninių rodiklių formalizavimo problema. Vertybinių popierių klasterizavimo uždaviniui parinkti šie fundamentiniai rodikliai:

- EPS
- EPS augimo tempas
- ROE
- Akcijų dalis tenkanti įmonės darbuotojams
- Bendrasis likvidumo koeficientas (current ratio)
- Akcijos rinkos kainos ir jos buhalterinės vertės santykis (price to book)
- Institucinių investuotojų dalyvavimas
- P/E santykis

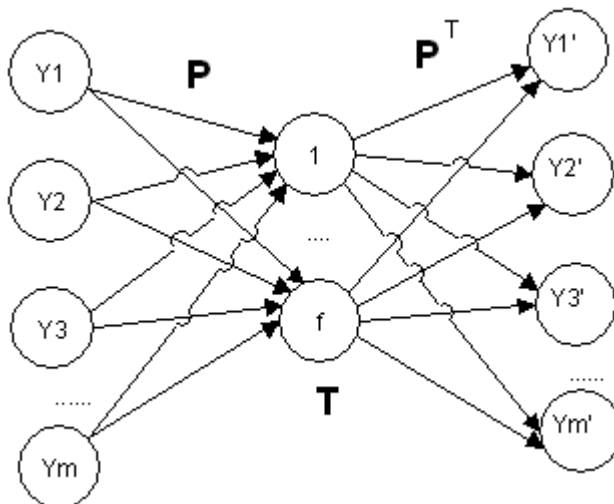
2. INTELEKTINIŲ SISTEMŲ TAIKYMAS VERTYBINIŲ POPIERIŲ RINKOJE

Kaip minėta anksčiau, vertybinių popierių rinkose reikia sekti ir įvertinti labai didelį kintamųjų skaičių. Pažangūs investiciniai fondai savo finansiniams sprendimams paremti seka 50-100 kintamųjų, kurie charakterizuoja bendros rinkos būseną, bei atskiros kompanijos veiklą.

Investuotojams yra labai sudėtinga pastoviai sekti ir interpretuoti šią informaciją. Principinių komponentių formavimo (principal component analysis, PCA) algoritmai yra speciali naujai išvystyta technika, leidžianti su mažiausiais informacijos nuostoliais transformuoti m -matę vertybinių popierių rinkos būseną į dvimatę ar trimatę „sintetinių“ arba „pseudo“- požymių erdvę, kurią investuotojas gali žymiai paprasčiau interpretuoti.

2.1 Principinių komponentių taikymas klasterizavimui

Principinių komponentių formavimo (principal component analysis, PCA) algoritmai yra speciali naujai išvystyta technika, leidžianti su mažiausiais informacijos nuostoliais transformuoti m -matę vertybinių popierių rinkos būseną į dvimatę ar trimatę „sintetinių“ arba „pseudo“- požymių erdvę, kurią investuotojas gali žymiai paprasčiau interpretuoti.



Šaltinis: R.Simutis. Paskaitų konspektai (Simutis, 2006)

12 pav. PCA metodo schema

Principinių komponentių formavimas atliekamas šiais žingsniais:

1. Iš pradinių duomenų yra suformuojama $m \times n$ matrica, kur m – stebimų kintamųjų skaičius, o n – atliktų stebėjimų skaičius.

2. Duomenų matrica normalizuojama – iš kiekvieno eilutės nario atimamas eilutės vidurkis taip suformuojant eilutę, kurios vidurkis lygus nuliui.
3. Apskaičiuojama kovariacinė matrica.
4. Surandami kovariacinės matricos eigenvektoriai ir tikrosios reikšmės. (Shlens, 2005)

Tegu Y yra vertybinių popierių rinkos duomenų matrica, kurios dimensija yra $n \times m$ (n – rinkoje atliktų stebėjimų skaičius, pvz., valandos, dienos, savaitių *DJI* kitimas, o m – yra stebimų kintamųjų skaičius, pvz. *DJI*, *S&P500*, *Advance/Decline line* ir t.t.).

Principinių komponentų analizė yra optimalus matricos Y faktorizavimas į dvi matricas, taip vadinamą žymių matricą T (scores matrix, dimensija $n \times f$) ir užkrovimo matricą P (loadings matrix, dimensija $m \times f$) bei liekanų matricą E (matrix of residuals, $n \times m$):

$$Y = T P^T + E$$

Kur f yra principinių faktorių skaičius ($f < m$).

Šis faktorizavimas turi būti atliekamas taip, kad liekanų matricos Euklidinė norma, būti minimali pasirinktam faktorių skaičiui. Teorijoje yra žinoma, kad ši sąlyga išpildoma tada, kai P yra Y kovariacinės matricos pirmų f -didžiausių nuosavų reikšmių nuosavų vektorių matrica.

PCA algoritmą galime pavaizduoti kaip tam tikros m -matės kintamųjų erdvės suspaudimą į f -matę erdvę. Suspaudimas šiuo atveju atliekamas naudojant tiesinę užkrovimo matricą P . Žymių matrica šiuo atveju apskaičiuojama pagal sekančią išraišką:

$$T = Y P$$

Užkrovimo matrica P turi būti paskaičiuota taip, kad vėl išskleidus kintamuosius į originalią m -matę erdvę naudojant išraišką:

$$Y' = T P^{-1}$$

Paklaidų kvadratų suma tarp originalių kintamųjų ir “rekonstruotų” kintamųjų būtų minimali, t.y.

$$E = ||Y - Y'|| \rightarrow \min$$

Realizuojant šį algoritmą, kintamieji pradžioje spaudžiami į vieną “pseudo” kintamąjį, o vėliau “rekonstravimo” netikslumai mažinami paeiliui didinant kintamųjų skaičių. (Simutis, 2006)

Principinės komponentės taip pat gali būti efektyviai panaudojamos atrenkant akcijas-kandidatus, kurios toliau stebimos ir atliekama prekybos signalų paieška. Šiuo atveju procedūros taikymo tvarka yra panaši į saveorganizuojančių žemėlapių kūrimą ir susideda iš sekančių žingsnių:

1. Pasirenkami fundamentalūs kintamieji, kurie gerai apibūdina kompanijos būklę;

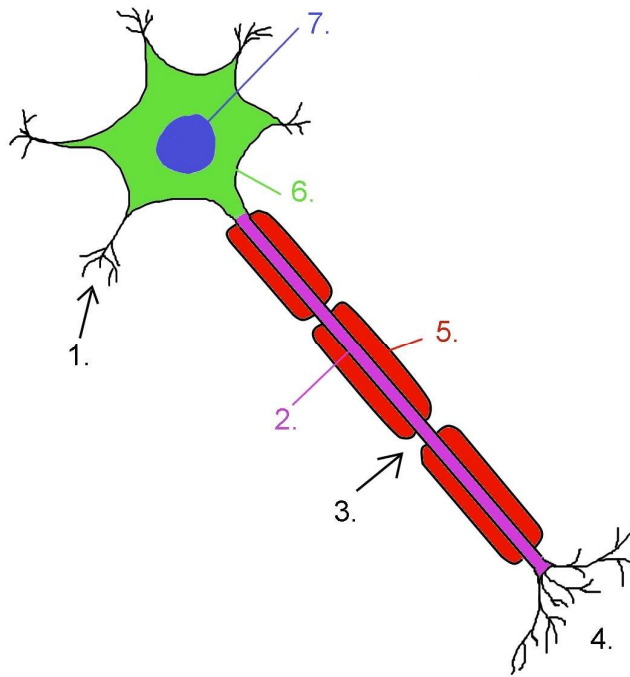
2. Iš duomenų bazių atrenkamos kompanijos su minėtais fundamentaliais kintamaisiais;
3. Atrinktų kompanijų būsenos atvaizduojamos dvimatėje principinių komponentių erdvėje, ir apibrėžiami klasteriai, kuriose randasi geros (kurių kaina pastarąjį ketvirtį/mėnesį/savaitę kilo) ir blogos kompanijos (kurių kaina pastarąjį ketvirtį/mėnesį/savaitę krito);
4. Atrenkant akcijas tolimesniems stebėjimams, pasirenkamos tik tos akcijos kurios papuolė į „gerus“ klasterius, taip pas stebimas šių akcijų „judėjimas“ dvimatėje principinių komponentių erdvėje.

Sudarant principines komponentes yra remiamasi prielaida, kad duomenys tarpusavyje yra susieti tiesiniais ryšiais. (Shlens, 2005). Tačiau vertybinių popierių rinkas dėl sudėtingų tarpusavio priklausomybių tikslingiau būtų analizuoti naudojant netiesines transformacijas. Šiems tikslams šiuo metu bandoma pritaikyti autoasociatyvinius neuroninius tinklus.(Simutis, 2006)

2.2 Neuroninių tinklų taikymas klasterizavimui

Taikant principinių komponentių metodus naudojamos įvairios originalių kintamųjų tiesinės transformacijos. Tačiau vertybinių popierių rinkas dėl sudėtingų tarpusavio priklausomybių tikslingiau būtų analizuoti naudojant netiesines transformacijas. Šiems tikslams šiuo metu bandoma pritaikyti neuroninius tinklus. Dirbtinis neuroninis tinklas suprantamas, kaip didelės apimties netiesinė sistema, susidedanti iš tam tikro skaičiaus paprastų, netiesinių duomenų apdorojimo elementų.(Simutis, 2006)

Dirbtiniuose neuroniniuose tinkluose tam tikra prasme bandoma kopijuoti biologinio neurono darbą. Manoma, kad žmogaus smegenys ir nervinė sistema yra sudaryta iš apie 25 milijardų neuronų. Pav.13 vaizduoja pavienį neuroną ir neuronų jungimosi schemą. Neuronas gauna informaciją dendrito pagalba. Sinapsės reguliuoja informacijos mainus tarp kitų tinklo neuronų. Nors neuronas turi tik vieną išėjimą (aksoną), jis gali turėti daugiau kaip 104 įėjimų. Priklausomai nuo įėjimų sumos ir neurono būsenos, neurono išėjime gaunamas signalas, kuris aksono pagalba perduodamas kitiems neuronams.



Šaltinis: Laisvoji enciklopedija. Vikipedija, (2005)

13 pav. Neurono sandara

Pagal 13 pav., neuronas yra sudarytas iš:

- Dendritas (1)
- Aksonas (2)
- Ranvje sąsmauka (3)
- Aksono „šluotelė“ (4)
- Mielininis dangalas (Švano ląstelės, 5)
- Ląstelė (6)
- Ląstelės branduolys (7)

Biologinė informacija yra perduodama tokio tipo tinklais, ir perdavimo metu gali būti modifikuojama. Pagrindinė šių procesų savybė yra ta, kad informacija gali būti perduodama ir apdorojama paraleliai. Pasitelkiant biologinio neurono analogiją galime suformuoti dirbtinį neuroną, kurio struktūra pavaizduota pav. 14.

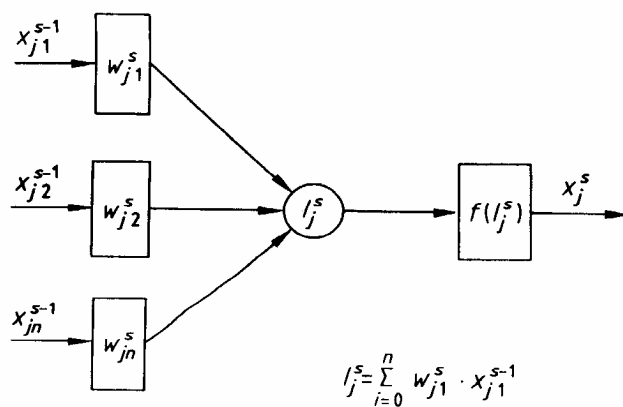
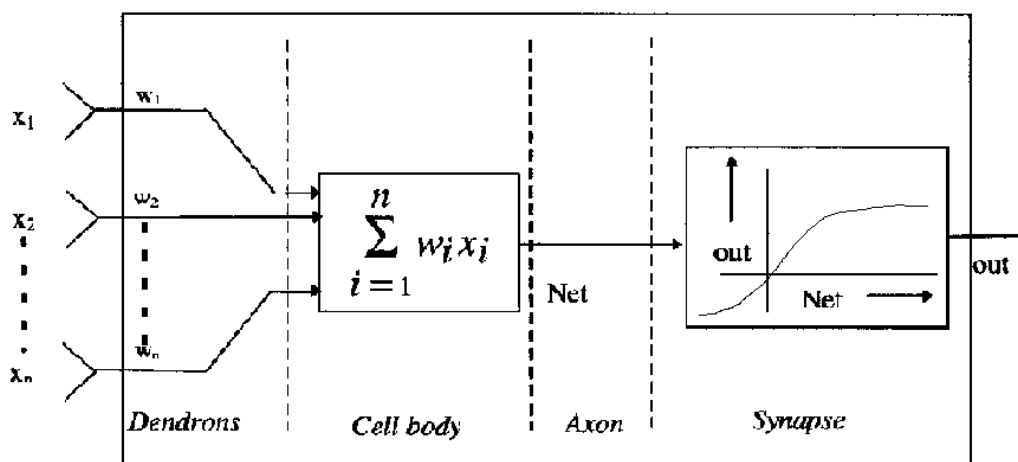
Šio neurono išėjimo signalas apskaičiuojamas iš sekančios išraiškos:

$$x_j^s = f\left[\sum_{i=0}^n (w_{ji}^s \cdot x_i^{s-1})\right] = f(I_j^s)$$

čia x_{js}^s - s sluoksniu j neurono išėjimas,

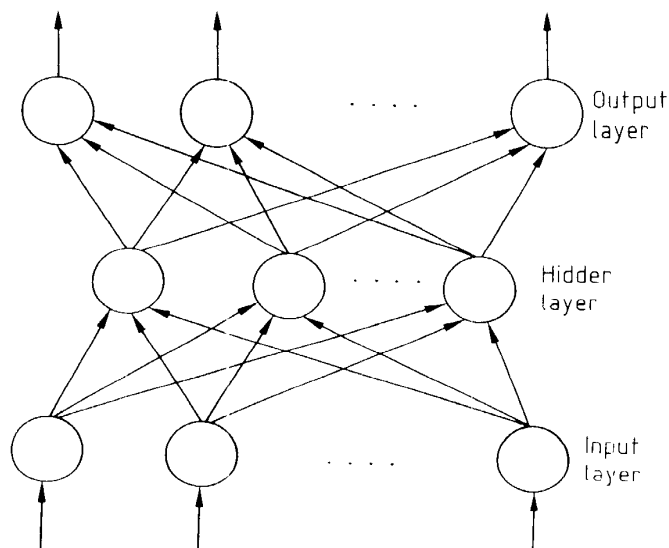
w_{ji}^s - svorinis koeficientas jungtyje nuo i-jo neurono s-1 sluoksnyje iki j-jo neurono s sluoksnyje, I_j^s - j-jo neurono įėjimų suma lygyje s.

Iš pavienių neuronų analogiškai gali būti **formuojami neuroniniai tinklai**. Neuronų perdavimo funkcijos $f(I_j^s)$ pavaizduotos pav. 14.



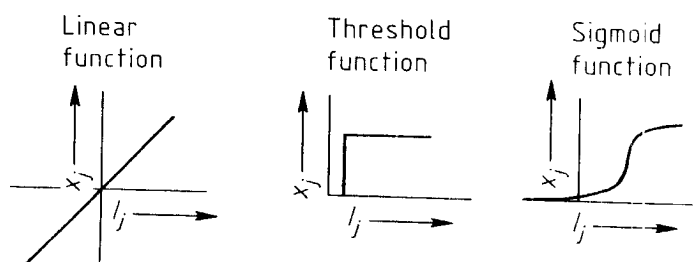
Šaltinis: SIMUTIS, R. (1996) Intelektinės sistemos finansų rinkose

14 pav. Dirbtinio neurono funkcinė schema



Šaltinis: SIMUTIS, R. (1996) Intelektinės sistemos finansų rinkose

15 pav. Neuroninis tinklas su vienu paslėptu sluoksniu



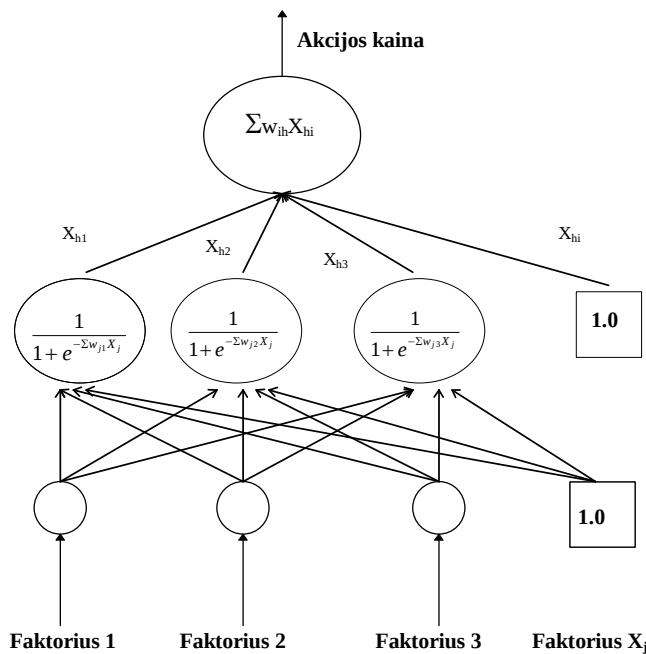
Šaltinis: SIMUTIS, R. (1996) Intelektinės sistemos finansų rinkose

16 pav. Dirbtinio neurono įėjimo/išėjimo priklausomybės

Taigi pasirinkus dirbtinių neuroninių tinklų struktūrą, bei neuronų perdavimo funkcijas ir derinant svorius wjis galima taip suformuoti tinklo išėjimą, kad jis realizuotų sudėtingas netiesines funkcijas tarp įėjimo ir išėjimo. Neuroninio tinklo svoriniai koeficientai parenkami pagal tam tikrus racionalius algoritmus. Plačiausiai paplitę yra taip vadinami gradientiniai klaidos skleidimo atgal algoritmai (back propagation algorithms) ir atsitiktinės paieškos metodai.(Simutis, 2006)

2.2.1 Neuroninių tinklų apmokymo algoritmai

Dirbtinius neuroninius tinklus mes naudosime ieškant akcijų kainą veikiančių esminių faktorių ir trumpalaikę kainos prognozei. Šiems uždaviniams spręsti patogų naudoti neuroninius tinklus su vienu paslėptu sluoksniu ir sigmoidine perdavimo funkcija šio lygio neuronuose. Tinklo išėjimas turi prognozuoti akcijos kainą, o išėjimo neurono perdavimo funkcija yra svorinis sumatorius. Šio tinklo schema pavaizduota pav.17.



Šaltinis: SIMUTIS, R. (1996) Intelektinės sistemos finansų rinkose

17 pav. Neuroninio tinklo, naudojamo akcijų kainos prognozavimui, struktūra

Prieš atliekant neuroninio tinklo apmokymą suformuojami įėjimo faktorių ir akcijos kainos duomenų masyvai. Kadangi įėjimo duomenys gali turėti labai skirtingas skaitines reikšmes, efektyviam tinklo apmokymui būtina atlikti įėjimo dydžių normavimą (Pvz. X_j/X_{j_max}). Atskirai parenkami duomenys apmokyto tinklo testavimui. Panagrinėsime paprasčiausius tokio tinklo apmokymo algoritmus.

2.2.1.1 Paklaidos skaičiavimo atgal algoritmas (error back propagation)

Atsitiktinės paieškos metodai yra gana paprasti, tačiau reikalauja daug skaičiavimo resursų. Kai kuriais atvejais žymiai efektyvesni yra gradientiniai apmokymo algoritmai. Jie ypač pasiteisina apmokant arba treniruojant neuroninius tinklus, nes dirbtiniams neuroniniams tinklams galima analitiškai apskaičiuoti gradientus, naudojant paklaidos skaičiavimo atgal algoritmą (error back propagation).

Taikant gradientinius metodus modeliavimo paklaidos minimizavimui, modelio parametrai (mūsų atveju -neuroninio tinklo svoriai) keičiami pagal sekančią išraišką:

$$\mathbf{w}_{i+1} = \mathbf{w}_i - \alpha \frac{\partial E}{\partial \mathbf{w}_i}$$

Neuroninio tinklo svoriams išraišką $\frac{\partial E}{\partial \mathbf{w}_i}$ galima paskaičiuoti analitiškai pagal sudėtinių funkcijų diferencijavimo taisyklę.

Ši savybė leidžia greitai suskaičiuoti gradientus netgi sudėtingiems neuroniniams tinklams, todėl šis treniravimo algoritmas yra dažniausiai taikomas įvairiuose programų paketuose.

2.2.1.2 Apmokymo algoritmų palyginimas

Skirtingų apmokymo algoritmų palyginimui yra sudaryta lentelė, kurioje algoritmai yra vertinami pagal skirtingas savybes nuo geriausio (++) iki prasčiausio (--).

2 lentelė

Apmokymo algoritmų palyginimas

Algoritmas	Vykdymo laikas	Patikimumas/tvirtumas	Call	Memory
Back-propagation	--	+	--	-
Rekursinis back-propagation	--	+	--	++
Levenbergo- Marquardt	++	++	++	--
Optimizuotas Levenbergo- Marquardt	+	++	++	0
Rekursinis spėjimo (Gauso- Niutono)	0	-	--	++

Šaltinis: sudaryta autoriaus pagal Neural Network Based System Identification TOOLBOX

Tolimesniam darbui tikslingiausia būtų naudoti Levenbergo-Marquardt arba atminčiai optimizuotą Levenbergo-Marquardt algoritmą.

2.2.2 Neuroninių tinklų trūkumai

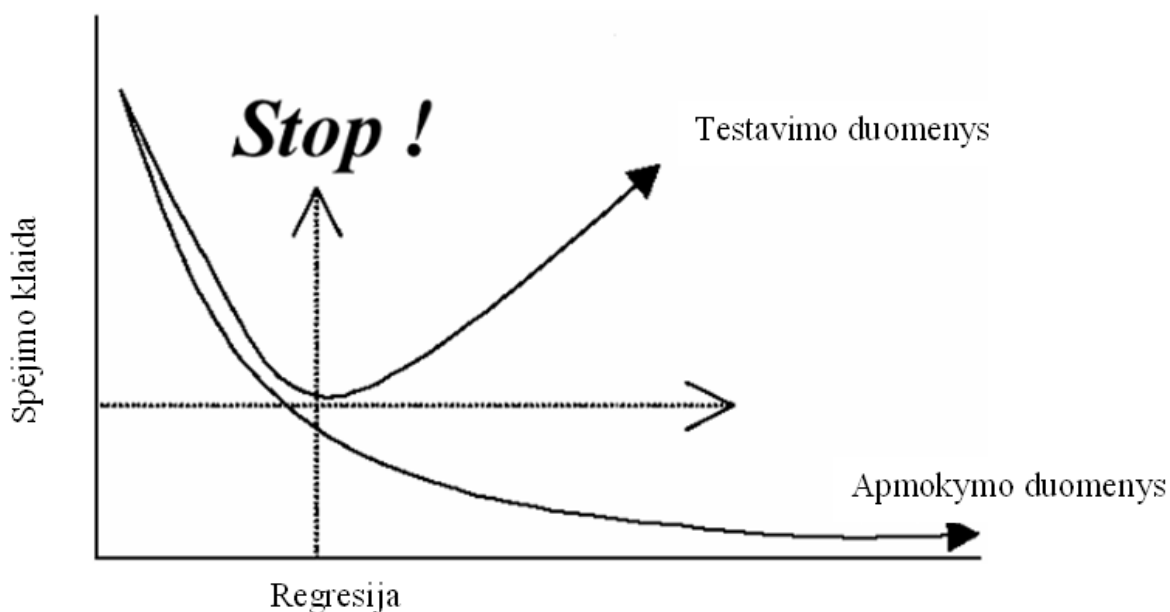
Sprendžiant neuroninių tinklų uždavinius susiduriama su šiais jų trūkumais:

- Pertreniravimo problema
- Neuroninio tinklo treniravimas užima daug laiko
- Neuroninio tinklo modeliavimo problema (sunku parinkti reikiamą paslėptų mazgų skaičių)
- Sunku pritaikyti neuroninius tinklus konkretiems uždaviniams (Al-Zobaydi, 2006)

Svarbus uždavinys treniruojant neuroninius tinklus yra numatyti, kada reikia nutraukti tolimesnį treniravimą. Neuroniniai tinklai turi svarbią savybę generalizuoti turimus duomenis. Pertreniravimas įvyksta kai sistema įsimena tam tikrą šabloną ir taip praranda savo generalizavimo savybes, o būtent prognozavimo sistemoms pati svarbiausia savybė ir yra ta, kad jos galėtų

korektiškai dirbti su pirmą kartą matomais duomenimis. Pertreniruotas neuroninis tinklas labai gerai veikia su apmokymo duomenimis, tačiau tampa visai bevertis jam padavus naujus pradinius duomenis. Pertreniravimą gali sąlygoti per didelis paslėptų mazgų kiekis arba pernelyg didelis apmokymo iteracijų skaičius. Pertreniravimo problemos galima išvengti naudojant skirtingas mokymosi ir testavimo duomenų imtis, arba kitus įvairius metodus, kaip pavyzdžiui kryžminis patikrinimas (cross-validation) (Lawrence, 1997).

Kryžminio patikrinimo metu neuroninis tinklas yra mokomas minimizuojant klaidos funkciją MSE mokymo imties atžvilgiu ir tuo pačiu tikrinama kaip prognozavimo paklaida MSE testavimo imčiai. Neuroninio tinklo apmokymas nutraukiamas kai testinės imties paklaida ima didėti, šiuo atveju tinklas pasižymi geriausiomis generalizavimo savybėmis (Simutis, 2004).



Šaltinis: sudaryta autoriaus pagal ALMEIDA, J. (1999) A Very Quick Guide to Developing and Training Feed Forward Neural Networks

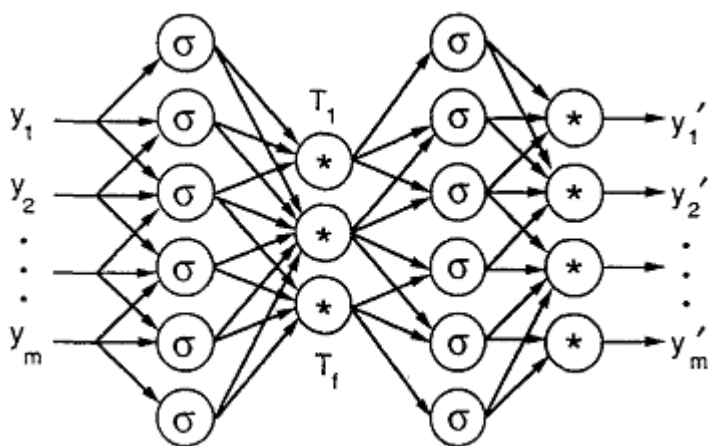
18 pav. Kryžminis patikrinimas

2.3 Autoasociatyviniai neuroniniai tinklai

Autoasociatyvinis dirbtinis neuroninis tinklas (Autoassociative Artificial Neural Network) yra dirbtinis neuroninis tinklas treniruotas taip, kad sugebėtų įėjimo vektorių suspausti į kelis kintamuosius ir kuo tiksliau iš jų atgaminti pradinius duomenis. Treniruojant autoasociatyvinį neuroninį tinklą jo viduryje suformuojamas 2-3 neuronų sluoksnis į kurį suspaudžiama informacija.

Greta autoasociatyvinių neuroninių tinklų yra sutinkami ir hetero-asociatyviniai. Pastarojo neuroninio tinklo išėjimai dažniausiai skiriasi nuo įėjimo vektoriaus. (Norgaard, 2003)

Autoasociatyvinio neuroninio tinklo struktūra yra pavaizduota 19 pav.:



Šaltinis: KRAMER, M. (1991) Nonlinear principal component analysis using autoassociative neural networks, p. 236

19 pav. Autoasociatyvinio neuroninio tinklo struktūra

Pradinis duomenų vektorius Y yra transformuojamas į pseudo kintamuosius, iš kurių formuojami principiniai kintamieji. Šis vidurinis sluoksnis, kuriame yra suspausti duomenys vadinamas butelio kakliuko (angl. - bottleneck) sluoksniu. Vėliau iš šio sluoksnio yra vėl suformuojamas pseudo kintamųjų sluoksnis, o iš jo – išėjimo vektorius. Išėjimo vektorius lyginimas su įėjimo. Jei vektorių panašumas netenkina, neuroninis tinklas yra apmokomas, ir įėjimo vektorius yra dar kartą perspaudžiamas per tinklą. Ši procedūra – neuroninio tinklo apmokymas - yra kartojama tol, kol išėjimo vektorius tampa pakankamai panašus į įėjimo, arba kol pasiekiamas tam tikras iteracijų skaičius. (Simutis, 2006)

Autoasociatyvinio neuroninio tinklo formavimas ir apmokymas yra skaičiavimams reikli operacija, todėl būtina numatyti rezultatų paklaidos tolerancijos ribas.

Pradinės informacijos suspaudimas viduriniame sluoksnyje leidžia naudoti autoasociatyvinius neuroninius tinklus analogiškai PCA metodui.

2.4 Algoritmų taikymas vertybinių popierių klasifikavimui

Apie 95% prekybos laiko akcijų ir valiutų kainų kitimas yra artimas triukšmui ir tik apie 5% laiko kainų kitimas įgauna tendencijas, kurios turi determinuotų procesų elementų. Kadangi akcijų rinkose prekiaujama tūkstančių kompanijų akcijomis, labai tikėtina, kad rinkose kiekvienu metu yra dešimtys kompanijų, kurių kainų kitime yra determinizmo elementų ir šį kitimą, pasirinkus atitinkamus modelius, galima efektyviai prognozuoti. Čia svarbiausia nustatyti, kokių akcijų kitime atsiranda determinizmo elementų (Simutis, 2004).

Panašiu principu remiasi ir kompanijos TradeTrek siūlomas paslauga – TradeTrek Neuro-Predictor. Pagrindinis jos uždavinys – nustatyti, kokios akcijos turi determinizmo elementų ir jas pasiūlyti vartotojui. TradeTrek Neuro-Predictor numato akcijos kainos kitimo prognozę 80% tikslumu, t.y. paliekama vos 20% tikimybė jog akcija nepateks į apskaičiuotus rėmus, kita vertus,

šie akcijos kitimo rėmai gali būti labai platūs. Programa remiasi regresinių neuroninių tinklų algoritmu ("Artificial Intellingence Applied to Stock Trading", 2008).

2.5 Autoassociative_ANN

„Autoassociative_ANN“ – MATLAB aplinkoje veikianti programa atliekanti duomenų perspaudimą autoasociatyvinių neuroninių tinklų pagalba ir grafiškai atvaizduojanti „butelio kakliuko“ sluoksnio neuronus dvimatėje koordinatinių plokštumoje.

Programa remiasi dviem laisvai prieinamom MATLAB aplinkoje veikiančiom programom dscale.m ir makebottle.m.

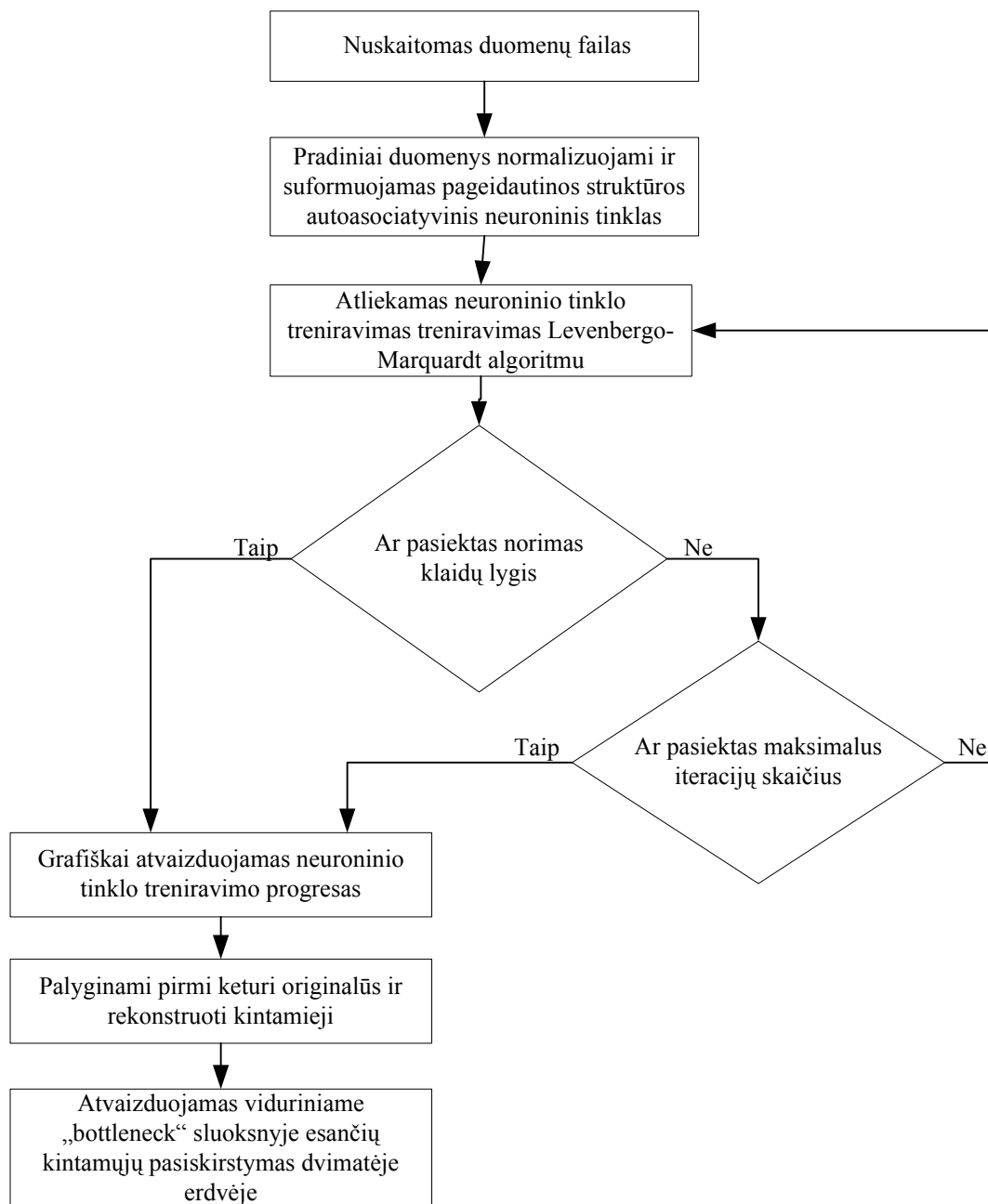
Failas dscale.m atlieka pradinio duomenų vektoriaus normalizavimą.

Failas makebottle.m atlieka neuroninio tinklo formavimą pagal nurodytą neuronų skaičių paslėptame bei viduriniame „bottleneck“ sluoksniuose.

Programos „Autoassociative_ANN“ veikimo algoritmas:

1. Nuskaitomas duomenų failas.
2. Dscale.m pagalba pradiniai duomenys normalizuojami ir makebottle.m pagalba suformuojamas norimos struktūros autoasociatyvinis neuroninis tinklas.
3. Tinklas treniruojamas pasirinktą iteracijų skaičių.
4. Pasiekus norimą klaidų lygį arba maksimalų iteracijų skaičių treniravimas nutraukiamas.
5. Grafiškai atvaizduojamas neuroninio tinklo treniravimo progresas.
6. Palyginami pirmi keturi originalūs ir rekonstruoti kintamieji.
7. Atvaizduojamas viduriniame „bottleneck“ sluoksnyje esančių kintamųjų pasiskirstymas dvimatėje erdvėje.

Programos blokinė schema pavaizduota 20 pav.



Šaltinis: sukurta autoriaus

20 pav. Autoassociative_ANN darbo algoritmas

Programos išeities tekstas pateikiamas 2 priede.

2.6 Algoritmų apibendrinimas ir išvados

Autoasociatyviniai neuroniniai tinklai ir principinės komponentės gali būti sėkmingai panaudoti tam pačiam tikslui – duomenų klasterizavimui.

Nors abiem atvejais gaunami rezultatai stipriai priklauso nuo kelių dedamųjų (pradinis kintamųjų skaičius, siekiamas kintamųjų skaičius) autoasociatyvinių neuroninių tinklų metodas yra žymiai tikslesnis už PCA, todėl tolimesniuose tyrimuose didžiausias dėmesys turi būti sutelktas būtent į autoasociatyvinių neuroninių tinklų taikymą.

Sukurta MATLAB aplinkoje dirbanti programa Autoassociative_ANN. Programa AANN pagalba atlieka finansinių rodiklių spaudimą į principines komponentes ir jas grafiškai atvaizduoja taip leisdama vartotojui išskirti galimus akcijų klasterius.

3. EKSPERIMENTINIS PCA IR AANN METODŲ TYRIMAS

Eksperimentinio tyrimo metu, naudojant įvairius faktorius, bus siekiama palyginti principinių komponentų ir autoasociatyvinių neuroninių tinklų veikimo galimybes ir rezultatus vertybinių popierių rinkoje.

3.1 PCA tyrimas

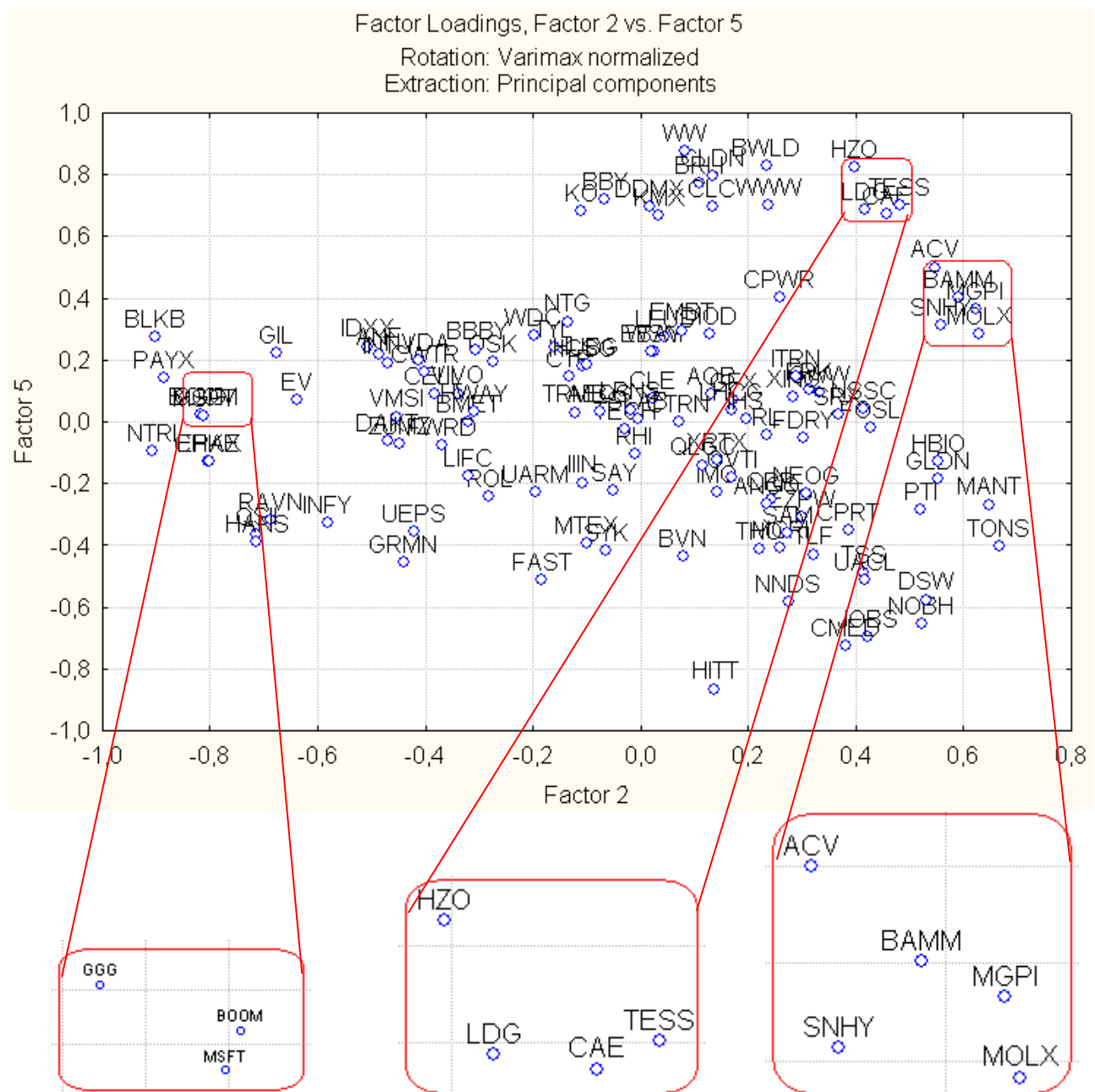
PCA tyrimo tikslas – naudojant įvairius analitiniame skyriuje atrinktų fundamentaliųjų faktorių derinius suformuoti akcijų klasterius, patikrinti ar atskiruose klasteriuose esančių akcijų kainos kinta panašia linkme.

3.1.1 Programinės įrangos parinkimas

Komercinis programinės įrangos paketas „Statistica“ buvo parinktas dėl savo populiarumo, plačių panaudojimo galimybių, geros dokumentacijos bei galimybės šia programa laisvai naudotis universitete. Be to ši programa gali dirbti su to beveik paties formato duomenimis, kaip ir vėlesniuose eksperimento etapuose bandytos programos. „Statistica“ nėra orientuota į finansų rinkas, tačiau principinių komponentų formavimas finansiniams uždaviniams niekuo nesiskiria nuo kitų uždavinių.

3.1.2 Statistica

Programinės įrangos paketas „Statistica“ taip pat turi priemonių duomenų išskaidymui į klasterius principinių komponentų pagalba. Tyrimui buvo panaudoti išsamūs 2007 06 13 JAV akcijų rinkos duomenys gauti PowerScreener įrankiu. Pradiniai duomenys buvo suspausti į penkis faktorius, iš jų parenkami du faktoriai ir atvaizduojami koordinatinių plokštumoje. Atrenkamos ryškiausios akcijų grupės – klasteriai.



Šaltinis: sudaryta autoriaus

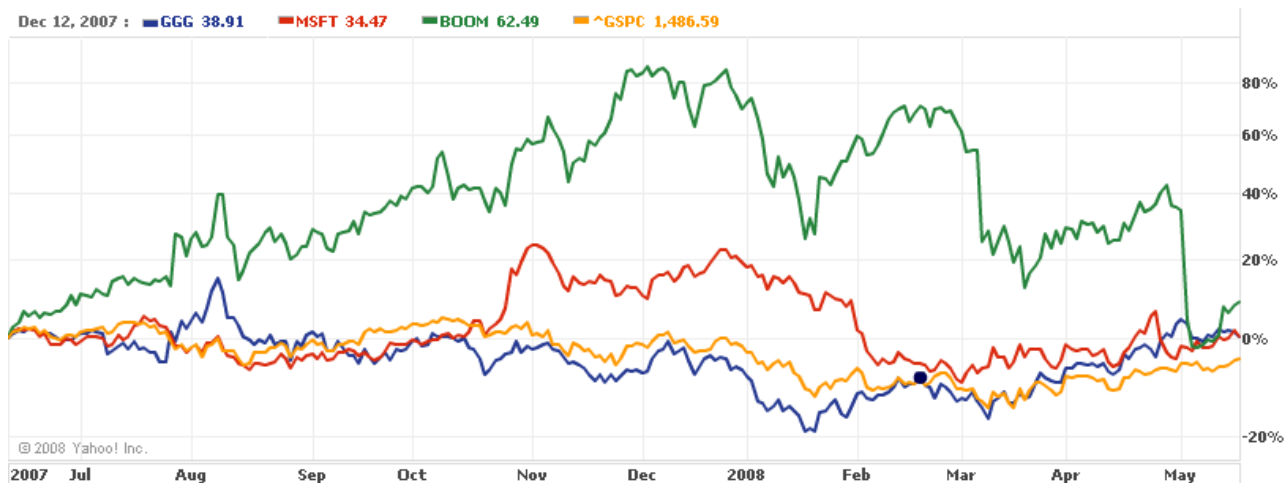
21 pav. Akcijų klasterizavimas naudojant „Statistica“

Matome, kad nuo bendro akcijų debesies kiek atsiskyrė kelios grupelės akcijų. Remiantis teorija, jų kainų judėjimas turėtų būti panašus.

Klasteris A: GGG, BOOM, MSFT

Klasteris B: HZO, LDG, CAE, TESS

Klasteris C: ACV, BAMB, MGPI, MOLX, SNHY



Šaltinis: sukurta autoriaus

22 pav. A klasterio akcijos



Šaltinis: sukurta autoriaus

23 pav. B klasterio akcijos



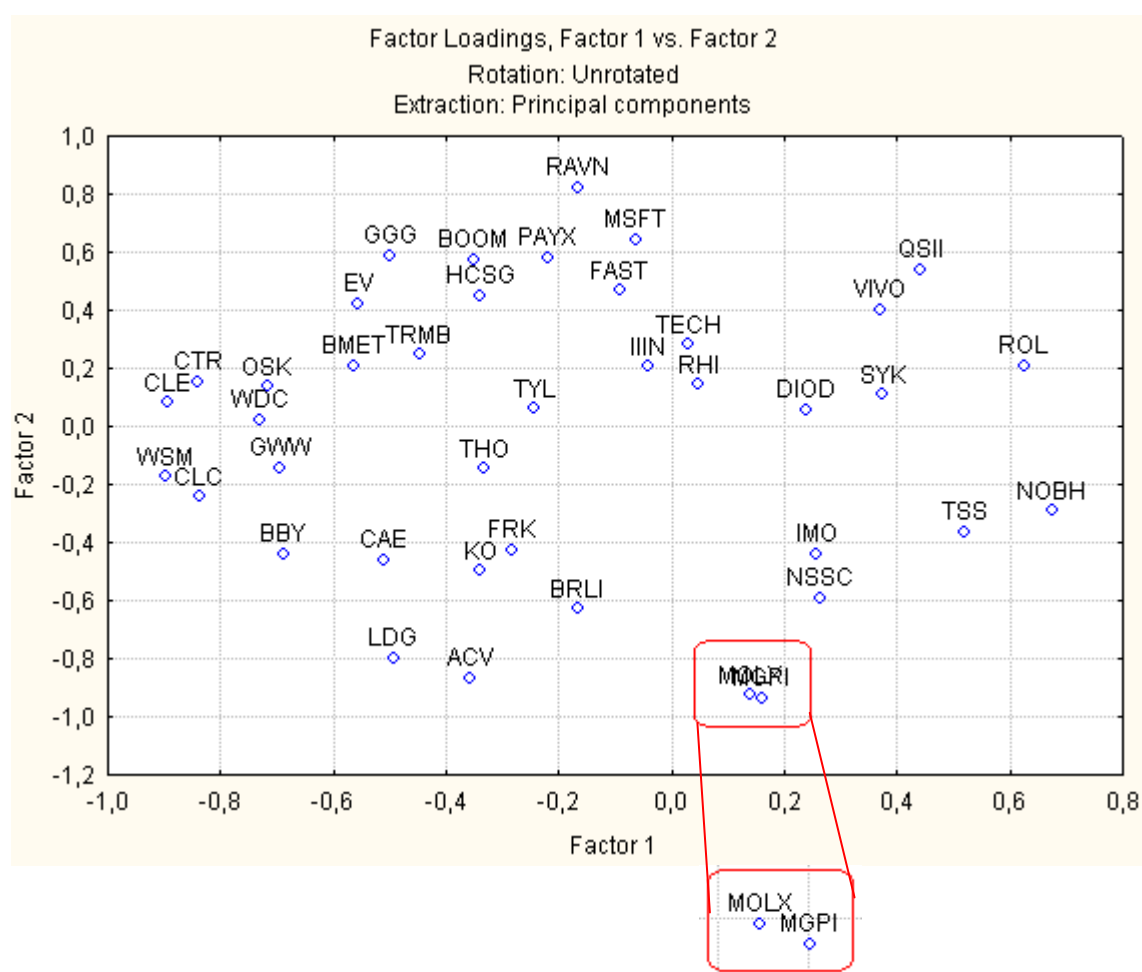
Šaltinis: sukurta autoriaus

24 pav. C klasterio akcijos

Klasterius A ir B sudarančios akcijos kito gan panašiai – A klasteryje susikaupė akcijos per metus laiko nepakitusios daugiau nei 10%, B klasteris – „blogų“ akcijų klasteris – visų jų kainos

smarkiai krito. Pirmieji du klasteriai tarsi ir patvirtintų hipotezę, tačiau trečiojo klasterio akcijų svyravimai yra labai diverguoti – t.y. pusė akcijų kilo, kita pusė – smarkiai krito, todėl negalime taikyti šio metodo klasterių atskyrimui.

Paminėtina, kad klasterizavimas buvo atliekamas spaudžiant įėjimo vektorių į 5 principines komponentes, tačiau analizuojant tik po dvi iš jų. Toks spaudimas užtikrina mažesnę duomenų praradimą, tačiau analizuodami tik dvi komponentes iš penkių, prarandame dalį sudėtinės informacijos, todėl eksperimentas buvo pakartotas spaudžiant įėjimo vektorių į dvi principines komponentes ir atrinktas akcijas apjungiant į klasterį D.



Šaltinis: sudaryta autoriaus

25 pav. D klasterio išskyrimas



Šaltinis: sudaryta autoriaus

26 pav. D klasterio akcijos

Glaudžiai susijusios MOLX ir MGPI aiškiai nutolusios nuo likusių akcijų, todėl tikėtina, kad jos kis panašiai, tačiau realus šių akcijų kitimas buvo visai nepanašus. Išvada – principinių komponentių metodas prognozuojant akcijų kainų pakyčius yra neefektyvus.

3.2 AANN tyrimas

Nepasitvirtinus akcijų klasterizavimui PCA metodu, atliekamas eksperimentas naudojant tuos pačius fundamentinius rodiklius AANN metodu, kuris teoriškai yra efektyvesnis klasterizuojant finansinius duomenis.

3.2.1 Programinės įrangos parinkimas

Esama nemažai komercinių produktų neuroninių tinklų generavimui skirtų darbui su vertybiniais popieriais. Vieni jų – Tradecision, NeuroShell – yra gerai išbaigtos, daug žadančios, daug rekomendacijų turinčios programos, kurių licencijos kaina siekia virš 1000Lt, deja dažniausiai jų negalima parsisiųsti netgi pabandymui. Kiti – Pitstock, Alyuda NeuroIntelligence – nebrangios, dažnai nemokamas apribotas versijas turinčios programos, deja dažniausiai susiduriama su sudėtingu programos valdymu, abejotiniais rezultatais. Tačiau, tiek vienu, tiek kitu atveju gamintojai pabrėžia, jog jų kūriniai yra skirti vartojimui tik kaip pagalbinė priemonė greta kitų pažangių vertybinių popierių rinkos analizės metodų.

PCA analizei naudota „Statistica“ nuo aštuntosios versijos taip pat turi įskiepių neuroninių tinklų uždaviniams spręsti, tačiau jie yra mokami.

AANN eksperimentui tinkamos laisvai prieinamos įrangos rasta nebuvo, todėl buvo MATLAB aplinkoje buvo sukurta AANN realizuojanti Autoassociative_ANN. Programa spaudžia duomenis į pasirinkta skaičių principinių komponentių, po to jas atvaizduoja koordinačių sistemoje.

Taip pat grafiškai atvaizduojami pradiniai ir „perdirbti“ duomenys, tokiu būdu galima akivaizdžiai įsitikinti AANN perspaudimo tikslumu.

3.2.2 Duomenų paruošimas

Duomenys turi būti pateikti tekstiniam faile. Vienoje eilutėje yra aprašomi vieną akciją apibūdinantys rodikliai. Pagal nutylėjimą rodiklių turi būti 10, tačiau šį skaičių galima keisti šiek tiek pageduojant programos išeities kodą.

Rodikliai turi būti atskirti tabuliacijos simboliais.

Duomenis galima imti iš įvairių šaltinių, tačiau patogiausias yra PowerScreener internetinis puslapis:

<http://www.investor.reuters.com/GoTo.aspx?nss=www&.t=/opinion/find/wresults>

Čia xls formatu galime parsisiųsti informaciją apie vienos dienos 645 įmonių 16 rodiklių.

Kadangi darbo tikslas yra surasti toki rodiklių rinkinį su kuriuo būtų įmanoma išskirti akcijas į atskirus klasterius, reikės atlikti skaičiavimus imant vis skirtingas rodiklių grupes. Kombinatoriniais apskaičiavimais, 16'ai rodiklių sugrupuoti po 10, egzistuoja virš 8000 skirtingų būdų:

$$\frac{16!}{10!(16-10)!} = 8008$$

Vadinasi eksperimentą reiktų atlikti 8008 kartus, be to tikslinga būtų atlikti skaičiavimus ir nebūtinai su 10 kintamųjų. Siekiant sumažinti eksperimentų skaičių, reiktų apibrėžti tik pačius svarbiausius rodiklius nagrinėtus skyriuje Fundamentiniai rodikliai.

Akcijų parinkimas yra dar vienas sudėtingas uždavinys ruošiant eksperimentinę bazę. Principinių komponentų koordinačių sistemoje pilnas 645 akcijų sąrašas gali atrodyti triukšmingai, be to neuroninio tinklo treniravimas yra imlus skaičiavimams uždavinys, todėl su daugiau akcijų (įskaitant ir tas, kurios mums nėra aktualios) jis bus vykdomas atitinkamai ilgesnį laiko tarpą (konkretus skaičiavimo laikas labai priklauso ir nuo kompiuterio procesoriaus). Dėl šių priežasčių būtų tikslinga pašalinti akcijas, kurios turi ne visus užpildytus rodiklius, akcijas, kurių kaina mažesnė nei 10 JAV dolerių ir pan.

Pradiniai duomenys yra gaunami iš trečiosios šalies, todėl negalima užtikrinti, kad duomenys bus visada prieinami. Šiuo atveju taip ir atsitiko – investor.reuters.com nutraukė PowerScreener palaikymą, ir todėl, norint tęsti darbą su „Autoassociative_ANN“ programa reikia ieškoti naujo duomenų šaltinio arba alternatyvios programos. Alternatyvios programos buvo apžvelgtos ankstesniame skyriuje. Naujo duomenų šaltinio paieška yra daug laiko užimantis uždavinys, kadangi reikia praktiškai išbandyti dauguma analizuojamų įrankių. Didžioji jų dalis

parsiunčia tik istorines akcijų kainas be kitų fundamentinių rodiklių. Be to dažniausiai duomenų nusiurbimui naudoja ir taip visiems prieinamus serverius kaip Yahoo, Medved ir pan. Dalis programų, kurios vis tik parsiunčia fundamentinius rodiklius, juos pateikia savo „standartizuotais“ formatais, todėl ir ši informacija tampa faktiškai bevertė.

„HSQuote Plus Dowloader“ [<http://www.1free-historical-stock-quotes-downloader.com/>] – nemokama riboto veikimo versija leidžiant parsisiųsti fundamentinius rodiklius xls formatu, tačiau yra gana ribotas pačių rodiklių pasirinkimas – siūlomi tik P/E, EPS, Short Ratio, Divident Rate, Price/Book, Book value ir dar keli kriterijai. Vis dėlto, programos sukuriamas rezultatų failas gali būti nesunkiai pritaikytas Autoassociative_ANN programai.

3.2.3 Tyrimo vykdymas

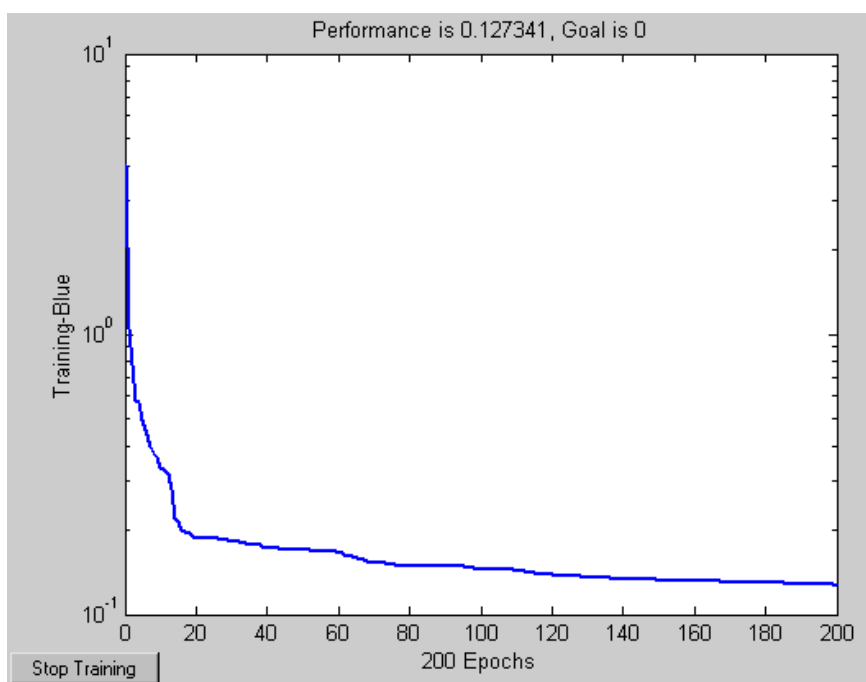
Paruošti bandomieji duomenų failai paeiliui apdorojami Autoassociative_ANN programa:

- Stebimas klaidų kiekis.
- Analizuojamas akcijų principinių komponentų pasiskirstymas koordinačių plokštumoje.
- Pastebėjus grupavimąsi į atskiras grupes, fiksuojamas esamas rodiklių rinkinys, išsaugomas grupių pasiskirstymo atvaizdas.

Ciklas kartojamas tol, kol išrenkamas rodiklių rinkinys aiškiausiai suskirstantis akcijas į atskiras grupes. Rezultatų, kurių metu nebuvo išskirta grupių registracija nėra reikalinga, kadangi tai tik padidintų laiko sąnaudas.

Tolimesniems stebėjimams sudaromas akcijų esančių tam tikruose klasteriuose sąrašas. Akcijos (tuo pačiu ir jų grupės) stebimos toliau – stebimas jų judėjimas koordinačių plokštumoje, kainos kitimas.

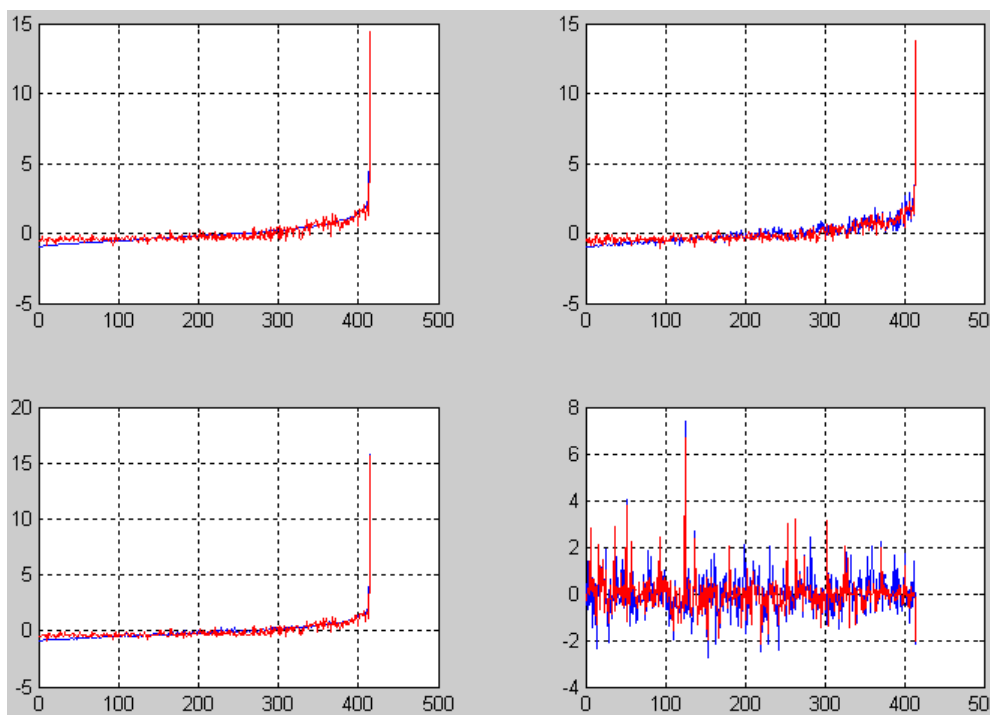
3.2.4 Rezultatai



Šaltinis: sudaryta autoriaus

27 pav. Neuroninio tinklo treniravimas

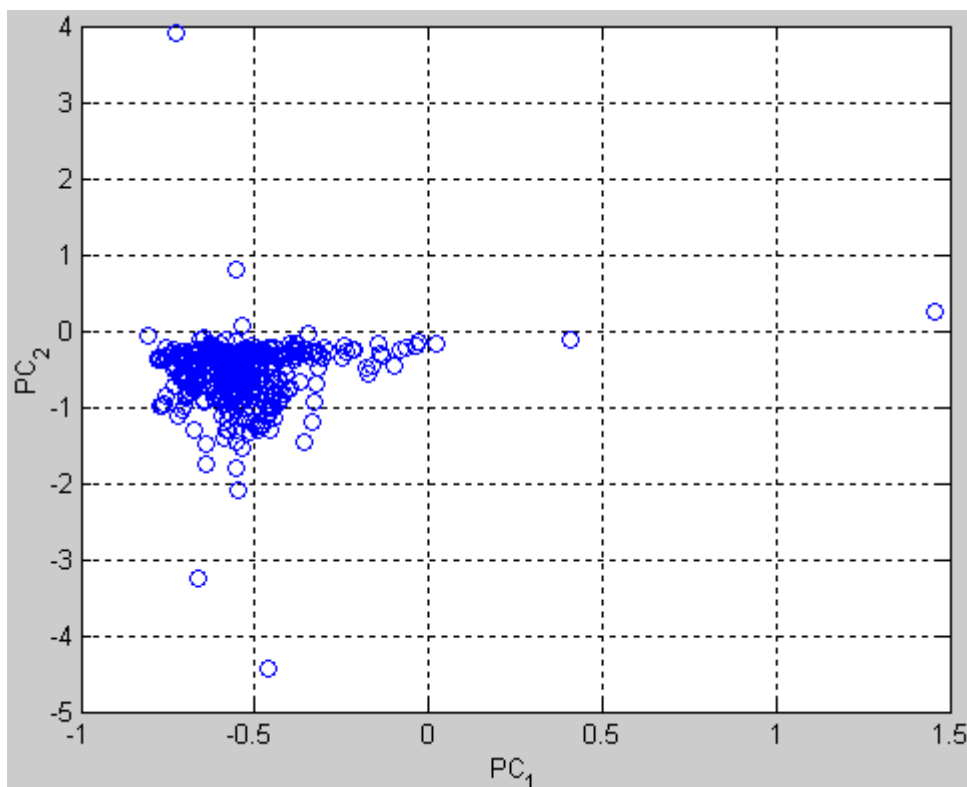
Paveiksle matomas neuroninio tinklo treniravimosi progresas. Matyti, kad po 20 treniruočių klaidų skaičiui sumažėja iki toleruotinos ribos.



Šaltinis: sudaryta autoriaus

28 pav. Neuroninio tinklo įėjimo ir išėjimo kintamųjų palyginimas

Apie gerai ištreniruotą neuroninį tinklą byloja ir įėjimo duomenų palyginimas su išėjimo duomenimis – visuose grafikuose (pav. 30) įėjimo ir išėjimo grafikas faktiškai sutampa.



Šaltinis: sudaryta autoriaus

29 pav. Akcijų pasiskirstymas AANN sudaryto „butelio kakliuko“ erdvėje

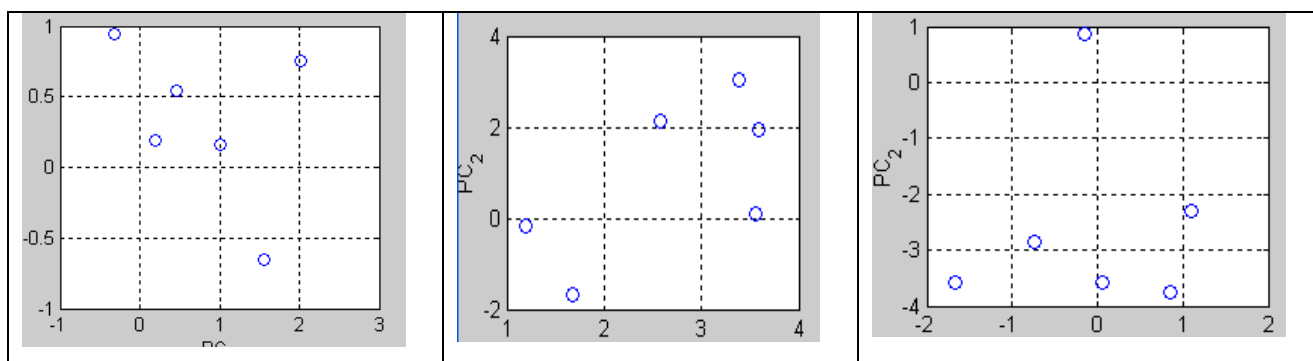
Pateiktoje schemoje negalima išskirti atskirų akcijų grupių – klasterių, todėl tyrimą reiktų tęsti kaitaliojant kriterijus bei kompanijas tol, kol išsiskirs kelios grupės.

3.2.5 Atvirkštinis eksperimentas

Kadangi atliktų AANN bandymų metu nebuvo išskirta aiškaus akcijų grupavimosi į grupes, buvo imtasi „reverse engineering“ – teorijos tikrinimo pagal jau žinomus rezultatus. Po sėkmingo PCA tyrimo nutarta neuroniniam tinklui paduoti jau žinomas geriausias (WDC, TYL, MSFT) ir prasčiausias (MGP, LFUS, BAMM) akcijas ir stebėti, ar jos susigrupuos į atskirus klasterius.

Hipotezė: geriausios ir prasčiausios akcijos susiburs į atskirus klasterius

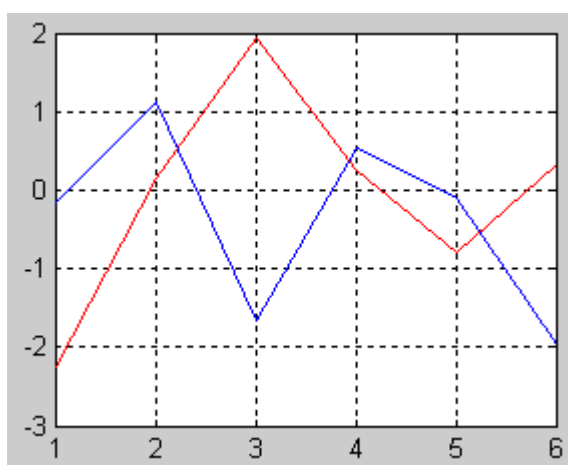
3.2.6 Rezultatai



Šaltinis: sudaryta autoriaus

30 pav. Menka klasterizacija AANN sudarytoje erdvėje

Varijuojant skirtingais įėjimo kintamaisiais (fundamentiniais rodikliais) gauti keli pagal AANN sudarytų principinių komponentių („bottleneck“) žemėlapiai. Kaip matyti pateiktuose rezultatuose, akcijos vėlgi nesiburia į grupes, todėl hipotezės patvirtinti negalima. Be to, dėl per mažo duomenų kiekio, neuroninis tinklas nesugeba gerai apsimokyti, todėl gauti rezultatai nėra tikslūs. Pav. 33 aiškiai matosi originalaus ir perspausto signalo nesutapimai.



Šaltinis: sudaryta autoriaus

31 pav. Originalus ir rekonstruoti signalai

3.3 Techniniais rodikliais grįstas PCA tyrimas

Kadangi bandymai klasterizuoti akcijas pagal fundamentinius rodiklius naudojant PCA ir AANN metodus buvo nesėkmingi, buvo nuspręsta išplėsti eksperimentą naudojant komercinę programinę įrangą dirbančią su techniniais faktoriais.

Esama nemažai programinės įrangos atliekančios duomenų suspaudimą į principines komponentes, tačiau tik nedidelė šių programų dalis yra orientuota į finansų rinkas. Programinės įrangos, kuri nėra tiesiogiai skirta vertybinių popierių analizei vartojimas labai komplikuoja

duomenų paruošimą bei rezultatų apdorojimą ir analizę. Dėl šios priežasties buvo pasirinktas komercinis produktas – „PickStock“.

„PickStock“ (© 2003, Bell Software and Services, Inc.) yra kompiuterinė programa principinių komponentų pagalba ieškanti galimai neįvertintų ir pervertintų akcijų.

Programa paremta arbitražinio vertinimo teorija („arbitrage pricing theory“). Ši teorija teigia, jog akcijos kaina gali būti suskaičiuota kaip tiesinė pagrindinių ekonominių rodiklių kombinacija. Šiuo metu teorija nėra įrodyta, dėl jos vyksta ginčai, tačiau programa PickStock besiremddama šia APT jau dabar leidžia ją panaudoti su tikrais duomenimis.

„PickStock“ PCA pagalba spaudžia istorinius akcijų kainos ir prekybos apimtys duomenis į nurodytą skaičių principinių komponentų bei formuoja galimai pervertintų ir galimai neįvertintų akcijų klasterius.

Programos licencijos kaina 49 JAV doleriai, tačiau gamintojo svetainėje yra pateikiama ir bandomoji versija. Vienintelis bandomosios versijos apribojimas – vienu metu programa gali apdoroti ne daugiau nei 10 skirtingų akcijų. Šį apribojimą galima „neutralizuoti“ suskaidant turimus duomenis į atskirus failus taip, kad viename faile būtų ne daugiau 10 skirtingų akcijų.

„PickStock“ programa neturi jokio įrankio automatiniam duomenų parsisiuntimui iš interneto, taigi duomenų paruošimas yra paliktas vartotojui.

3.3.1 Duomenų paruošimas

Duomenys turi būti pateikiami tekstiniu formatu, ASCII koduote. Kiekvienoje eilutės turi būti tokia informacija (eilės tvarka) : akcijos pavadinimas, prekybos diena (yyyymmdd formatu), atidarymo kaina, aukščiausia dienos kaina, žemiausia kaina, uždarymo kaina, prekybos apimtis. Duomenys turi būti išrikiuoti pagal akcijos pavadinimą, bei datą mažėjimo tvarka.

Duomenų surinkimui buvo pasirinkta modifikuota interneto naršyklė „Opera“, o duomenų šaltinis – <http://finance.yahoo.com> .

Duomenys – 12 mėnesių (nuo 2006 04 03 iki 2007 04 02) kasdienė akcijos kainos ir prekybos apimtys informacija. Šie duomenys pasiekiami adresu <http://ichart.finance.yahoo.com/table.csv?s=XXX&a=03&b=03&c=2006&d=03&e=02&f=2007&g=d&ignore=.csv>, kur XXX – ieškomos akcijos santrumpa.

Tyrimui informacija buvo surinkta apie atsitiktiniu būdu parinktų 50 laisvai kotiruojamų JAV įmonių akcijų (ACV, ACV, BAMM, BBY, BMM, BOOM, BRLL, CAE, CLC, CPWR, CTR, DAKT, DIOD, EV, EZPW, FAST, FOSL, FRK, GGG, GWW, HCSG, HWAY, IDXX, IINN, IMO, INTU, KO, LDG, LFUS, MGPI, MOLX, MSFT, NOBH, NSSC, OSK, PAYX, QSII, RAVN, RHI, ROL, SYK, TECH, THO, TYL, TRMB, TSS, VIVO, WDC, WSM, WWW).

3.3.2 Tyrimo vykdymas

„PickStock“ programos lange nurodome kelią iki duomenų failo, prekybos pradžios ir pabaigos datas bei pasirenkame principinių komponentių skaičių.

Kadangi įėjime, kiekvienai akcijai turime 5 kintamuosius, jų „spaudimas“ į 5 ar daugiau principinių komponentių neprarastų jokios informacijos.

PickStock version 1.6

This software has been registered for use on this machine. An unlimited number of stocks may be analyzed.

Save license to A: drive. Select this only if you make hardware changes or if you want to move your licensed copy of PickStock to another machine.

General Overview

Stock price file to process: C:\PickStock\50.csv Browse ?

Use the shortcut version? (requires the same set of stocks used in the last long version, existing basis vectors from the last long version are used) ☐ Yes, use shortcut ☒ No, use long version ?

First date used in stock price file: 20060403 Calendar Aid Last date used in stock price file: 20070402 Calendar Aid ?

☐ Fraction of variance explained by model: .99500 ?

☒ Number of principal components to use in the model: 5 ?

☒ Use cross validation to further optimize the number of components. ?

Maximum allowed consecutive missed days: 3 ?

Maximum allowed relative price change between consecutive days: .50000 ?

Minimum number of trading days required: 8 ?

Transformation method: ?

☐ No transformation (not recommended for stock analysis; useful for other technical applications).

☐ Log transformation of prices, normalized by average of each stock price logarithm.

☒ Log transformation of stock prices, normalized by each stock price logarithm on the earliest common date among the set of selected stocks.

☐ Log of stock prices, normalized by each stock price logarithm on the earliest common date among the set of selected stocks. Each of the transformed prices is then weighted by the square root of the product of volume and price on the last day of trading.

OK

Exit

Šaltinis: sudaryta autoriaus

32 pav. Pagrindinis PickStock langas

Rezultatai yra gaunami tekstinio failo formatu. Rezultatų akcijos yra įvertintos koeficientais nurodančiais jų neįvertinimo lygį.

Pateikiamos akcijos surikiuotos nuo labiausiai neįvertintos iki labiausiai pervertintos:

Duomenys gauti iš: C:\PickStock\100.csv, 20070402

WDC	-0.13616	19.08750	16.80000
TYL	-0.12551	14.27143	12.68000
MSFT	-0.11514	30.93400	27.74000
DIOD	-0.11083	38.42374	34.59000

...

MGPI	0.10232	18.17799	20.25000
LFUS	0.13594	35.40052	40.97000
FRK	0.21374	52.75834	67.10000

Irašas „WDC -0,13616“ reiškia, kad, akcija WDC yra nuvertinta 13,6 %, todėl tai gan stiprus signalas ją pirkti.

Irašas „FRK 0,21374“ reiškia, kad, akcija FRK yra perversinta 21,37 %, todėl tai stiprus signalas ją parduoti trumpam.

Eksperimento metu suformuosiu tris portfelių poras. Kiekviena pora bus sudaryta iš portfelio kuriame akcijos turi būti perkamos – portfelis „Pirkti ir laikyti“, bei portfelio su akcijomis kurios turi būti parduodamos trumpam – portfelis „Parduoti trumpam“. Pirmasis portfelis pelningas bus laikomas tuo atveju, jei jame esančių akcijų vertė pakils daugiau nei S&P500 indeksas, antrasis bus pelningas, jei S&P500 indekso kitimas bus aukštesnis nei portfelio.

Portfelių poros bus sudaromos PickStock programoje parenkant skirtingą principinių komponentių skaičių. Tam, kad įsitikintume korektišku programos veikimu, pirmoji portfelių pora bus generuojama naudojant 5 principines komponentes. Remiantis teorija, naudojant 3 principines komponentes informacijos nuostoliai turi būti nedideli, taigi šiuo metodu bus randamos antrojo portfelio akcijos. Perleidžiant duomenis per vieną principinę komponentę, duomenų praradimas yra neišvengiamas, tačiau, kad tuo įsitikintume suformuosime trečiąjį portfelį naudodami vieną principinę komponentę.

3.3.3 Rezultatai po 2 mėnesių

Portfeliai sudaryti naudojant 5 (ir daugiau) principinių komponentių:

3 lentelė

Portfeliai sudaryti naudojant 5 (ir daugiau) principinių komponentių

Pirkti ir laikyti		
Akcija	PickStock santykis	Pokytis
WDC	-0,07423	15 %
PAYX	-0,06531	5,20 %
RHI	-0,04965	-2 %
Portfelio pokytis		7,15 %

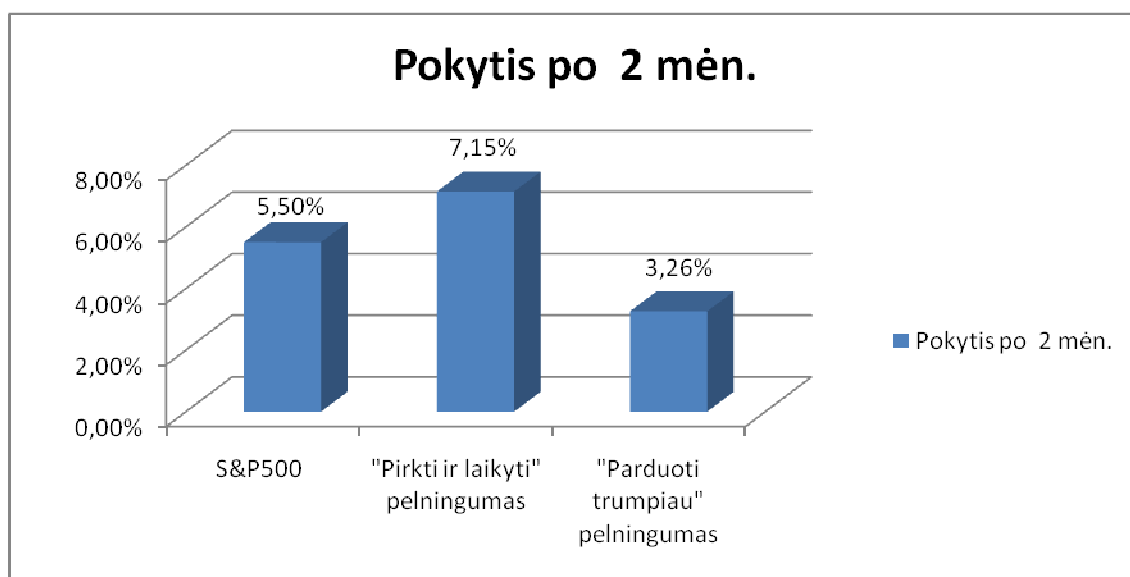
Parduoti trumpam		
Akcija	PickStock santykis	Pokytis
WSM	0,03745	-12 %
NSSC	0,0427	15,80 %
BAMM	0,33934	-4,70 %
Portfelio pokytis		3,26 %

Šaltinis: sudaryta autoriaus

Bendras portfelių balansas imant lygiom dalim 3 labiausiai neįvertintas ir tris labiausiai pervertintas akcijas yra lygus 5,205 %.

$$(7,15 \% (\text{Top3}) - (-3,26\%) (\text{Worst3})) * 0,5 = 5,205 \%$$

Atsižvelgiant į S&P500 indeksą, kuris per tą patį laikotarpį pakilo 5,5 % bendras abiejų portfelių pelningumas nuo jo atsilieka.



Šaltinis: sukurta autoriaus

33 pav. 3 principinių komponentų portfelių pelningumo palyginimas po 2 mėn.

Kiek kitokia situacija gaunama, jei labiau atkreipiamas dėmesys į PickStock apskaičiuotą akcijos neįvertinimo koeficientą. Akcija WDC, kuri laikoma labiausiai neįvertinta, per 2,5 mėn. pakilo 15%, o labiausiai pervertinta akcija BAAM nukrito 4,70%.



Šaltinis: sukurta autoriaus

34 pav. WDC ir BAMM palyginimas su S&P500 indeksu (2 mėn.)

Portfeliai sudaryti naudojant 3 principines komponentes:

4 lentelė

Portfeliai sudaryti naudojant 3 principines komponentes

Pirkti ir laikyti			Parduoti trumpam		
Akcija	PickStock santykis	Pokytis	Akcija	PickStock santykis	Pokytis
WDC	-0,13616	15 %	MGPI	0,10232	-21%
TYL	-0,12551	-4,20 %	LFUS	0,13594	-14%
MSFT	-0,11514	6,10 %	FRK	0,21374	1%
Portfelio pokytis		5,88 %	Portfelio pokytis		8,5 %

Šaltinis: sudaryta autoriaus

Šiuo atveju pervertintų akcijų portfelio „Parduoti trumpam“ pelningumas atrodo puikiai, o neįvertintų akcijų portfelio rodiklis yra panašus gautam pirmajame etape su 5 principinėm komponentėm. Bendras šių portfelių pelningumas yra 7,19 %, o tai jau lenkia rinkos vidurkį.

Analizuojant rezultatus, gautus atlikus transformaciją per tris principines komponentes, reiktų pastebėti, kad jie smarkiai skiriasi, nuo transformacijos kuomet principinių komponentių skaičius lygus kintamųjų skaičiui. Vos viena akcija (WDC) patenka į abejus portfelius.

Vadinasi, vykdant transformacijas buvo prarasta didelė dalis originalių duomenų ir todėl negalima paskliauti šiais rezultatais.

Portfeliai sudaryti naudojant 1 principinę komponentę:

5 lentelė

Portfeliai sudaryti naudojant 1 principinę komponentę

Pirkti ir laikyti			Parduoti trumpam		
Akcija	PickStock santykis	Pokytis	Akcija	PickStock santykis	Pokytis
NSSC	-1,86365	15,80 %	BRLI	0,31883	5,10 %
RAVN	-0,2868	24,90 %	FOSL	0,33708	14,90 %
FAST	-0,26722	20,30 %	BAMM	0,33934	-4,70 %
Portfelio pokytis		17,37 %	Portfelio pokytis		5,07 %

Šaltinis: sudaryta autoriaus

Kaip jau minėta anksčiau, yra iš anksto žinoma, jog ši portfelių pora suformuota praradus labai daug informacijos, todėl, kad ir kokie puikūs būtų pirmojo portfelio rezultatai (17,37 % per 2,5 mėn.), jais pasitikėti negalime.

Šis portfelis buvo sudarytas tam, jog įsitikinti, kad PCA metu naudojant vieną principinę komponentę bus prarasta dalis informacijos. Hipotezė pasitvirtino – lyginant su akcijų sąrašu gautu naudojant 5 ir daugiau komponentių, pastarasis skiriasi žymiai – sutampa vos viena akcija.

3.3.4 Rezultatai po 10 mėnesių

Pasirinkti portfeliai buvo palyginti ir po 10 mėnesių (2008-02-05).

Portfeliai sudaryti naudojant 5 (ir daugiau) principinių komponentių, po 10 mėn.

5 lentelė

Portfeliai sudaryti naudojant 5 (ir daugiau) principinių komponentių

Pirkti ir laikyti		
Akcija	PickStock santykis	Pokytis
WDC	-0,07423	65%
PAYX	-0,06531	-10%
RHI	-0,04965	-23%
Portfelio pelningumas		11 %

Parduoti trumpam		
Akcija	PickStock santykis	Pokytis
WSM	0,03745	-21%
NSSC	0,0427	12%
BAMM	0,33934	-39%
Portfelio pelningumas		-0,32

Šaltinis: sudaryta autoriaus



Šaltinis: sudaryta autoriaus

35 pav. WDC ir BAMM palyginimas su S&P500 indeksu (10 mėn.)

Geriausios (WDC – viršutinis grafikas) ir prasčiausios (BAMM – apatinis grafikas) akcijų kitimas visiškai atitiko prognozes sudarytas principinių komponentių metodu. Vidurinis – SP500 indeksas liko beveik nepakitęs. Kadangi programos PickStock pagalba buvo analizuoti vienerių metų duomenys, tyrimą reikės baigti ir galutinį matavimą atlikti po maždaug vienerių metų nuo principinių komponentių sudarymo – 2008 04 01, tačiau jau dabar turimas rezultatas leidžia tikėtis gerų šio metodo rezultatų.

Tikslesnis pasirinkto modelio vertinimas gaunamas imant kelias „geriausias“ ir „blogiausias“ akcijas. Be to, reikia atsižvelgti ir į programos apskaičiuotą tikimybės koeficientą – „PickStock santykis“:

Portfelis (Top3):

$$\frac{0,074 \times 0,65}{0,074 + 0,065 + 0,049} + \frac{0,065 \times (-0,1)}{0,074 + 0,065 + 0,049} + \frac{0,049 \times (-0,23)}{0,074 + 0,065 + 0,049} = 0,11$$

Portfelis (Worst 3):

$$\frac{0,33934 \times (-0,39)}{0,33934 + 0,0427 + 0,03745} + \frac{0,0427 \times 0,12}{0,33934 + 0,0427 + 0,03745} + \frac{0,03745 \times (-0,21)}{0,33934 + 0,0427 + 0,03745} = -0,32$$

Matyti, kad Top3 portfelis per 10 mėnesių padidėjo 11%, o Worst3 nuvertėjo 32%. Nors rezultatai nėra tokie įspūdingi, kaip skaičiuojant tik pačią geriausią ir prasčiausią akciją, šis metodas išlieka pelningas, ypač atsižvelgiant į tai, kad S&P500 indeksas per šį laiką nukrito 1,7%.

Portfeliai sudaryti naudojant 3 principines komponentes, po 10 mėn.

7 lentelė

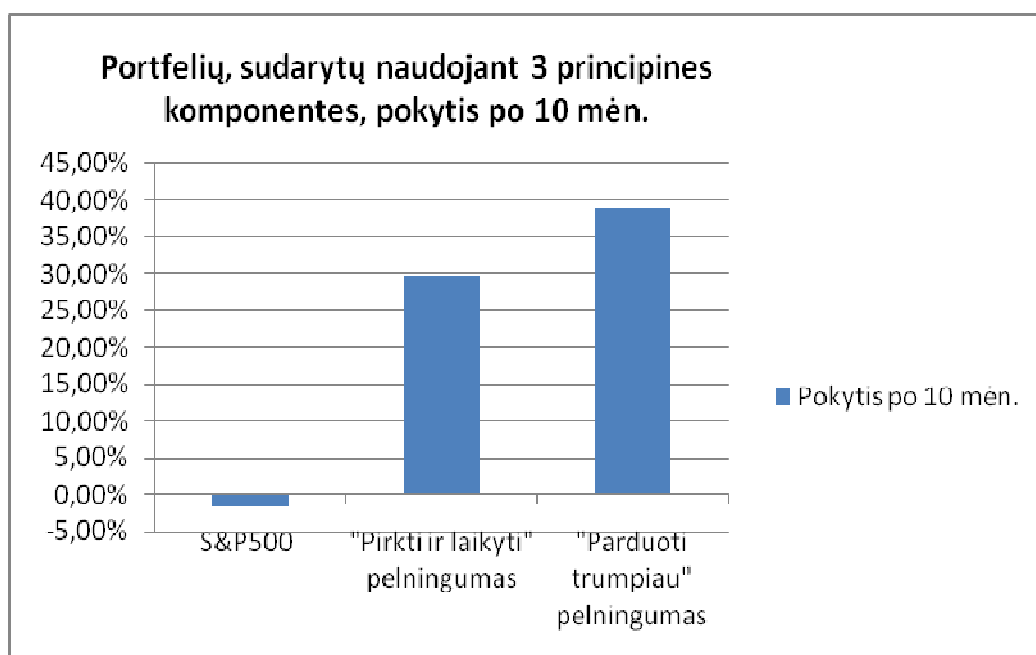
Portfeliai sudaryti naudojant 3 principines komponentes

Pirkti ir laikyti		
Akcija	PickStock santykis	Pokytis
WDC	-0,13616	65%
TYL	-0,12551	10%
MSFT	-0,11514	10%
Portfelio pelningumas		29,8%

Parduoti trumpam		
Akcija	PickStock santykis	Pokytis
MGPI	0,10232	-24%
LFUS	0,13594	-51%
FRK	0,21374	Nėra duomenų
Portfelio pokytis		-39%

Šaltinis: sudaryta autoriaus

Portfelių sudarytų po transformacijų su 3 principinėm komponentėm rezultatai vėl pranoko lūkesčius – Top3 portfelis stabiliai augo – 29,8% o Worst3 portfelis staigiai krito – -39% per 10 mėnesių.



Šaltinis: sudaryta autoriaus

36 pav. 3 principinių komponentių portfelių pelningumas po 10 mėn.

Portfelis sudarytas spaudžiant kintamuosius į 1 principinę komponentę tikrintas nebebuvo, kadangi, kaip anksčiau jau buvo minėta, – toks portfelis netgi teoriškai yra netikslus.

3.4 Eksperimentinio tyrimo išvados

„Statistica“ programos pagalba atlikus akcijų klasterizavimą su įvairiais fundamentinių rodiklių rinkiniais naudojant PCA metodą buvo išskirti keli ryškesni akcijų klasteriai, tačiau laiko bėgyje klasteryje esančių akcijų kitimas nebuvo panašus, todėl prieita išvados, kad klasterizavimas naudojant PCA buvo netikslus ir tyrimas buvo tęsiamas naudojant AANN.

Atlikus akcijų klasterizavimą su įvairiais fundamentinių rodiklių rinkiniais naudojant autoasociatyvinius neuroninius tinklus buvo įsitikinta aukštu metodo efektyvumu spaudžiant duomenis į mažesnę erdvę, tačiau atskirų klasterių išskirta nebuvo, todėl hipotezė teigianti, jog AANN metodas gali klasterizuoti akcijas tolimesnei jų analizei nebuvo patvirtinta.

Naudojantis komercine „PickStock“ programa buvo atliktas akcijų klasterizavimas pagal techninius rodiklius – istorinė akcijos kaina ir prekybos apimtys – naudojant PCA metodą. Sudarius akcijų portfelius pagal programos suformuotus klasterius, laiko bėgyje jie pasirodė esą pelningi, bei aplenkė S&P 500 indeksą. Nors korektiškas „PickStock PCA“ skaičiavimas kelia abejonių, visi sugeneruoti pirktinų akcijų portfeliai dviejų mėnesių bėgyje pasirodė pelningi. Norint galutinai įsitikinti šios programinės įrangos veiksmingumu, reiktų toliau atlikti stebėjimus su turimais duomenimis, taipogi būtų tikslinga išplėsti tyrimo imtį apimant daugiau akcijų, bei varijuojant laikotarpį. Praėjus 10 mėnesių nuo duomenų rinkimo ir portfelių formavimo dienos galima teigti, jog Principinių komponentių metodas pasiteisino – visi keturi portfeliai buvo pelningi. Abejonių kelia tik tai, kad teoriškai nepatikimesni portfeliai sudaryti naudojantis tik 3 principinėm komponentėm negu atnešė didesnę pelną nei teoriškai tiksliau sumodeliuotieji.

Žvelgiant iš šalies, silpnai atrodo pradiniai duomenys – egzistuoja tik akcijos kaina dienos bėgyje (atidarymo, aukščiausia, žemiausia, uždarymo) ir prekybos apimtys. Visiškai nėra atsižvelgiama į kitus fundamentinius ar techninius rodiklius. Įvesdami kainos priklausomybę nuo fundamentinių ar techninių parametrų neišvengiamai susidurtume su netiesinėm priklausomybėm, poreikiu reaguoti į gaunamus tarpinius rezultatus – apsimokymu. Tokiais atvejais teoriškai tikslingiau būtų naudoti autoasociatyvinius neuroninius tinklus.

Teigiami rezultatai naudojant PCA metodą su techniniais rodikliais leidžia tikėtis, kad pagal techninius rodiklius autoasociatyvinių neuroninių tinklų pagalba sudaryti akcijų klasteriai taip pat būtų sėkmingi.

IŠVADOS IR PASIŪLYMAI

1. Išanalizavus fundamentinius ir techninius vertybinių popierių rinką apibūdinančius faktorius, tolimesniame darbe nuspręsta remtis fundamentinių rodiklių aštuonetu.
2. Įvertintos principinių komponentių klasterizavimo galimybės transformuojant n-matę požymių erdvę į dvimatę ar trimatę, bei šios metodikos taikymas sprendžiant prognozavimo uždavinius. Įvertinti neuroninių tinklų privalumai prieš principines komponentes finansinių uždavinių sprendimo atžvilgiu. Tolimesniam darbui parinktas autoasociatyvinių neuroninių tinklų metodas kaip artimiausia neuroninių tinklų realizacija principinėms komponentėms.
3. Išbandžius įvairias programas realizuojančias PCA ir AANN, tolimesniam darbui buvo sukurta „autoassociative_ANN“ programa, bei parinktos komercinės programos – „PickStock“, „Statistica“.
4. Naudojantis „Statistica“, buvo suformuoti akcijų klasteriai, kuriems priklausančių akcijų kitimas turėtų būti panašus. Pagal tuos pačius fundamentinius rodiklius naudojant AANN klasterių suformuoti nepavyko.
5. Sėkmingų klasterių išskyrimas pagal fundamentinius rodiklius naudojantis PCA bei AANN metodais nepavyko, tuo tarpu pagal techninius rodiklius su PCA metodu veikianti „PickStock“ programa suformavo klasterius pasiekusius geresnių rezultatų nei nulinė prognozė, todėl tolimesnius tyrimus reiktų tęsti atliekant techninių rodiklių klasterizavimą AANN pagalba.
6. Autoasociatyvinių neuroninių tinklų taikymas vertybinių popierių rinkoje yra palyginti nauja kryptis, todėl kol kas sunku numatyti realias AANN efektyvaus panaudojimo galimybes.

LITERATŪRA

1. Advanced Neural Network Investing [interaktyvus]. [Žiūrėta 2008 m. gegužės 11 d.]. Prieiga per internetą: <<http://neuralinvesting.com/Tour/ANNITourTopics.aspx>>
2. ALBANIS, George T; BATCHELOR, Roy T. (2000) Predicting High Performance Stocks using Dimensionality Reduction Techniques based on Neural Networks [interaktyvus]. [Žiūrėta 2006m. spalio 17 d.]. Prieiga per internetą: <<http://www.staff.city.ac.uk/r.a.batchelor/Dimreds.pdf>>
3. ALMEIDA, J. Very Quick Guide to Developing and Training Feed Forward Neural Networks [interaktyvus]. [Žiūrėta 2008 m. gegužės 11 d.]. Prieiga per internetą: <<http://www.microcortex.com/InfoLinks/ANNDocumentation/VeryQuickGuidetoNN/>>
4. AL-ZOBAYDI, Ayad T. MS/Ayad-Marwan Network. A new architecture of an Artificial Neural Network. De Montfort University, Leicester, U.K. [interaktyvus]. [Žiūrėta 2007 m. birželio 20 d.]. Prieiga per internetą: <<http://www.imcl-conference.org/IMCL2006/pdf/zobaydi.pdf>>
5. Artificial Intellingence Applied to Stock Trading [interaktyvus]. [Žiūrėta 2008 m. kovo 1 d.]. Prieiga per internetą: <http://www.tradetrek.com/Education/ai/ai_stock_trading01.asp>
6. BASALTO, Nicolas; BELLOTTI, Roberto. (2004) Clustering stock market companies via chaotic map synchronization [interaktyvus]. [Žiūrėta 2007 m. birželio 20 d.]. Prieiga per internetą: <http://arxiv.org/PS_cache/cond-mat/pdf/0404/0404497.pdf>
7. DENG, Min. (2007) Death of the Efficient Market Hypothesis. Iš 20th Australasian Finance & Banking Conference 2007 [interaktyvus]. [Žiūrėta 2008 m. sausio 10 d.]. Prieiga per internetą: <<http://ssrn.com/abstract=1006716>>
8. Forex technical and fundamental indicators [interaktyvus]. [Žiūrėta 2007 m. birželio 20 d.]. Prieiga per internetą: <<http://forexindicators.blogya.de/forexindicators/>>
9. GIRARD, Stephane; IOVLEFF, Serge. Auto-Associative Models, Nonlinear Principal Component Analysis, Manifolds and Projection Pursuit [interaktyvus]. [Žiūrėta 2008 m. gegužės 11 d.]. Prieiga per internetą: <<http://pca.narod.ru/8MainGorbanKeglWunschZin.pdf>>
10. HINTON, G.E.; SALAKHUTDINOV, R.R. (2006) Reducing the Dimensionality of Data with Neural Networks [interaktyvus]. [Žiūrėta 2008 m. gegužės 11 d.]. Prieiga per internetą: <<http://www.cs.toronto.edu/~hinton/science.pdf>>
11. KERSCHEN, Gaetan; GOLINVAL, Jean-Claude. (2003) Feature extraction using auto-associative neural networks [interaktyvus]. Prieiga per internetą: <<http://stacks.iop.org/SMS/13/211>>

12. KRAMER, Mark. (1991) Nonlinear principal component analysis using autoassociative neural networks. Journal of the American Institute of Chemical Engineers, Vol 37, p. 233-243.
13. LAWRENCE, Ramon (1997). Using Neural Networks to Forecast Stock Market Prices [interaktyvus]. [Žiūrėta 2008 m. kovo 1 d.]. Prieiga per internetą: < <http://people.ok.ubc.ca/rlawrenc/research/Papers/nn.pdf> >
14. MONDELSON, Lou. (2000) Using Neural Networks For Financial Forecasting [interaktyvus]. [Žiūrėta 2007 m. birželio 20 d.]. Prieiga per internetą: <http://www.tradertech.com/financial_forecasting.asp>
15. Neural Network construction and research [interaktyvus]. [Žiūrėta 2008 m. gegužės 11 d.]. Prieiga per internetą: < http://pitstock.com/products/neural_net2.html >
16. Neural Networks: Understand to Win. © 2001-2008 Alyuda Research [interaktyvus]. [Žiūrėta 2008 m. gegužės 11 d.]. Prieiga per internetą: <http://www.tradecision.com/support/neural_networks_edge.htm >
17. NORGAARD, Magnus. Neural Network Based System Identification TOOLBOX [interaktyvus vartotojo žinydas]. [Žiūrėta 2008 m. kovo 1 d.]. Prieiga per internetą: <<http://www.iau.dtu.dk/research/control/nnsysid.html>>
18. PickStock dokumentacija, (©2003, Bell Software and Services, Inc.) [interaktyvus]. [Žiūrėta 2008 m. vasario 12 d.] Prieiga per internetą: <<http://www.bellsws.com/Documentation.htm>>
19. Pinigai daro pinigus [interaktyvus]. [Žiūrėta 2007 m. birželio 20 d.]. Prieiga per internetą: <<http://www.spekuliantai.lt/>>
20. PowerScreener Lite [interaktyvi duomenų bazė]. [Žiūrėta 2006 m. birželio 20 d.]. Prieiga per internetą: <www.investor.reuters.com/nscreen/builder.asp>
21. Q10: What is Neural Network Forecast? Why does it seem to change from time to time? [interaktyvus]. [Žiūrėta 2006 m. gruodžio 1 d.]. Prieiga per internetą: <<http://www.tradetrek.com/faqs/q10.asp> >
22. SHLENS, Jonathon. 2005, A Tutorial on Principal Component Analysis [interaktyvus]. [Žiūrėta 2006 m. gruodžio 1 d.]. Prieiga per internetą <<http://www.cs.cmu.edu/~elaw/papers/pca.pdf>>
23. SIMUTIS, Rimvydas. (2006) Intelektinės sistemos vertybinių popierių rinkose. Paskaitų konspektai.
24. SIMUTIS, Rimvydas. Finansinių rinkų prognozavimo priemonių kūrimo ypatumai. Informacinės technologijos verslui - 2004 : konferencijos pranešimų medžiaga. Kaunas, Technologija, 2004, p. 194-200. ISBN 9955-09-649-7.

25. SKIRGAILA, Aurimas. (2008) Autoasociatyvinių neuroninių tinklų taikymas vertybinių popierių kainų prognozei. „Akademinio jaunimo siekiai: Ekonomikos, Vadybos ir Technologijų išvalgos“. Studentų mokslinė-praktinė konferencija 2008 05 09, Klaipėda, Vakarų Lietuvos verslo kolegija, p. 160-163.
26. STIGLITZ, Joseph (2003) There is no invisible hand. Friday, December 20, 2002. The Guardian
27. Techninė analizė – indikatoriai [interaktyvus]. [Žiūrėta 2007 m. birželio 20 d.] Prieiga per internetą: <<http://www.metaquotes.net/techanalysis/indicators/>>
28. Terminų Žodynas (2008). [Žiūrėta 2008 m. kovo 1 d.] Prieiga per internetą: <<http://www.verslumas.lt/?serv=glossary&what=&cat=24&page=7>>
29. VAGA, T. „Profiting from Chaos“ [interaktyvus]. [Žiūrėta 2008 m. gegužės 11 d.]. Prieiga per internetą: <<http://books.google.com/books?id=hjUMHEHpp38C&hl=lt>>
30. Vikipedija, laisvoji enciklopedija [interaktyvi enciklopedija]. [Žiūrėta 2007 m. birželio 20 d.]. Prieiga per internetą: <<http://lt.wikipedia.org/>>
31. WARD, S., SHELARD, M. Successful Trading Using Artificial Intelligence [interaktyvus]. [Žiūrėta 2008 m. gegužės 11 d.]. Prieiga per internetą: <<http://www.neuroshell.com/Successful%20Trading%20Using%20Artificial%20Intelligence.pdf>>
32. Why Our Investment Strategy Works [interaktyvus]. [Žiūrėta 2008 m. gruodžio 2 d.]. Prieiga per internetą: <<http://tdmresearch.com/WhyItWorks.htm>>

AUTOASOCIATYVINIŲ NEURONINIŲ TINKLŲ TAIKYMAS VERTYBINIŲ POPIERIŲ KAINŲ PROGNOZEI

Aurimas Skirgaila

Vilniaus universitetas, Kauno humanitarinis fakultetas

Darbo vadovas: prof. Rimantas Simutis

Anotacija

Straipsnyje apžvelgiamas autoasociatyvinių neuroninių tinklų bei principinių komponentių taikymas vertybinių popierių kainų prognozei. Glaustai supažindinama su esminiais šių metodų veikimo principais, apžvelgiamas dabartinis šių metodų panaudojimas akcijų kainų prognozėms, bei ateities perspektyvos. Eksperimentinio tyrimo metu suformuojami akcijų portfeliai ir stebimas jų vertės kitimas metų bėgyje.

PAGRINDINIAI ŽODŽIAI: akcijos, vertybiniai popieriai, autoasociatyviniai neuroniniai tinklai, techninė analizė, fundamentinė analizė, principinės komponentės.

Abstract

This article is a survey on the application of auto associative neural networks and principal component analysis in forecasting stock prices. Main principles of these two methods are presented, reviewing the current usage of AANN and PCA and future outlook. An experiment is being carried out by building two stock portfolios using PCA. The portfolios are being monitored within one year.

KEYWORDS: stock market, prediction, auto associative neural networks (AANN), principal component analysis (PCA), technical analysis, fundamental analysis.

Įvadas

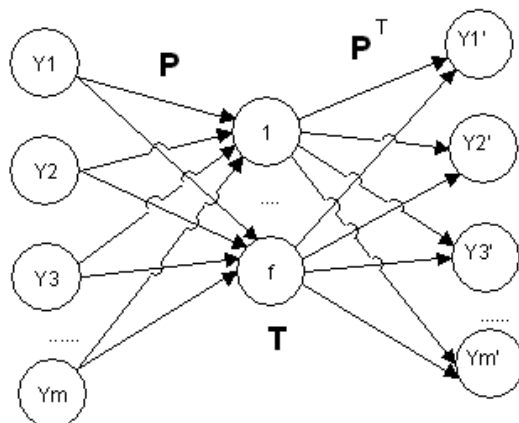
Vertybinių popierių rinkose stabilus pelnas pasiekiamas skiriant didelį dėmesį bendrai rinkos analizei ir atskirų kompanijų finansinės veiklos rodyklių studijavimui.

Pažangūs investiciniai fondai savo finansiniams sprendimams paremti seka 50-100 kintamųjų, kurie charakterizuoja bendros rinkos būseną, bei atskiros kompanijos veiklą. Investuotojams yra labai sudėtinga pastoviai sekti ir interpretuoti šią informaciją.

Principinių komponentių formavimo (principal component analysis, PCA) bei auto asociatyviniai neuroniniai tinklai (auto associative neural networks, AANN) yra naujai išvystytos technikos, leidžiančios su mažiausiais informacijos nuostoliais transformuoti m -matę vertybinių popierių rinkos būseną į dvimatę ar trimatę „sintetinių“ arba „pseudo“- požymių erdvę, kurią investuotojas gali žymiai paprasčiau interpretuoti.

Principinės komponentės

Principinių komponentių formavimo (principal component analysis, PCA) algoritmai yra speciali naujai išvystyta technika, leidžianti su mažiausiais informacijos nuostoliais transformuoti m -matę vertybinių popierių rinkos būseną į dvimatę ar trimatę „sintetinių“ arba „pseudo“- požymių erdvę, kurią investuotojas gali žymiai paprasčiau interpretuoti.



Šaltinis: R.Simutis. Paskaitų konspektai [1]

Pav. 1 PCA metodo schema

Principinės komponentės taip pat gali būti efektyviai panaudojamos atrenkant akcijas-kandidatus, kurios toliau stebimos ir atliekama prekybos signalų paieška. Šiuo atveju procedūros taikymo tvarka yra panaši į saveorganizuojančių žemėlapių kūrimą ir susideda iš sekančių žingsnių:

1. Pasirenkami fundamentalūs kintamieji, kurie gerai apibūdina kompanijos būklę;
2. Iš duomenų bazių atrenkamos kompanijos su minėtais fundamentaliais kintamaisiais;
3. Atrinktų kompanijų būsenos atvaizduojamos dvimatėje principinių komponentių erdvėje, ir apibrėžiami klasteriai, kuriose randasi geros (kurių kaina pastarąjį ketvirtį/mėnesį/savaitę kilo) ir blogos kompanijos (kurių kaina pastarąjį ketvirtį/mėnesį/savaitę krito);
4. Atrenkant akcijas tolimesniems stebėjimams, pasirenkamos tik tos akcijos kurios papuolė į „gerus“ klasterius, taip pas stebimas šių akcijų „judėjimas“ dvimatėje principinių komponentių erdvėje.

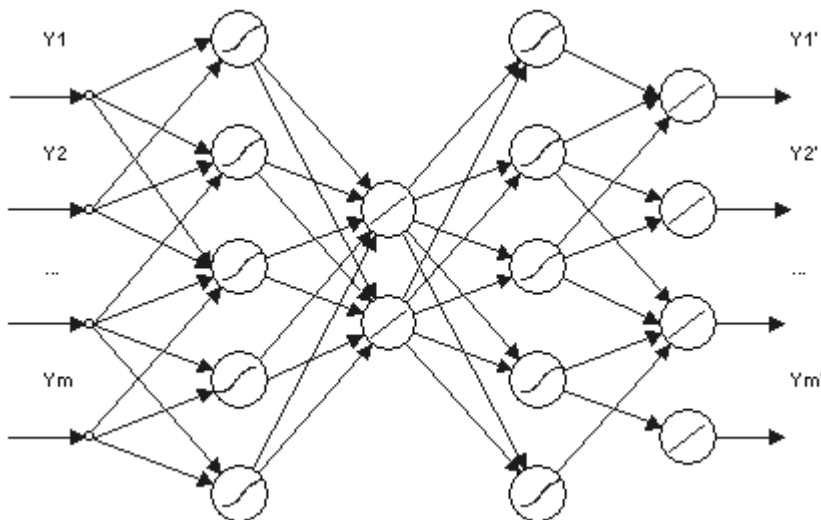
Taikant principinių komponentių metodus naudojamos įvairios originalių kintamųjų tiesinės transformacijos. Tačiau vertybinių popierių rinkas dėl sudėtingų tarpusavio priklausomybių tikslingiau būtų analizuoti naudojant netiesines transformacijas. Šiems tikslams šiuo metu bandoma pritaikyti autoasociatyvinius neuroninius tinklus.[1]

Autoasociatyviniai neuroniniai tinklai

Taikant principinių komponentių metodus naudojamos įvairios originalių kintamųjų tiesinės transformacijos. Tačiau vertybinių popierių rinkas dėl sudėtingų tarpusavio priklausomybių tikslingiau būtų analizuoti naudojant netiesines transformacijas. Šiems tikslams šiuo metu bandoma pritaikyti autoasociatyvinius neuroninius tinklus.

Autoasociatyvinis dirbtinis neuroninis tinklas (Autoassociative Artificial Neural Network) yra dirbtinis neuroninis tinklas treniruotas taip, kad sugebėtų įvesties vektorių suspausti į kelis kintamuosius ir kuo tiksliau iš jų atgaminti pradinis duomenis.

Autoasociatyvinio neuroninio tinklo struktūra yra pavaizduota 2 pav.



Šaltinis: R.Simutis. Paskaitų konspektai [1]

Pav. 2 Autoasociatyvinio neuroninio tinklo struktūra

Pradinis duomenų vektorius Y yra transformuojamas į pseudo kintamuosius, iš kurių formuojami principiniai kintamieji. Šis vidurinis sluoksnis, kuriame yra suspausti duomenys vadinamas butelio kakliuko (angl. - bottleneck) sluoksniu. Vėliau iš šio sluoksnio yra vėl suformuojamas pseudo kintamųjų sluoksnis, o iš jo – išėjimo vektorius. Išėjimo vektorius lyginamas su įėjimo. Jei vektorių panašumas netenkina, neuroninis tinklas yra apmokoma, ir įėjimo vektorius yra dar kartą perspaudžiamas per tinklą. Ši procedūra – neuroninio tinklo apmokymas - yra kartojama tol, kol išėjimo vektorius tampa pakankamai panašus į įėjimo, arba kol pasiekiamas tam tikras iteracijų skaičius.[1]

Autoasociatyvinio neuroninio tinklo formavimas ir apmokymas yra skaičiavimams reikli operacija, todėl būtina numatyti rezultatų paklaidos tolerancijos ribas.

1 PRIEDAS (TĘSINYS)

Esama įvairių neuroninių tinklų treniravimo metodų. Dauguma jų pagrįsti palaipsniui mažinant klaidų skaičių išėjimo sluoksnyje.

Konkretus neuroninio tinklo apmokymo metodas neturi reikšmės, - svarbu tiesiog, kad tinklas būtų tinkamai ištreniruotas, todėl atskiri metodai plačiau nagrinėjami nebus.

PCA prieš AANN

Autoasociatyviniai neuroniniai tinklai ir principinės komponentės gali būti sėkmingai panaudoti tam pačiam tikslui - daugelio kintamųjų suspaudimui į kelis.

Nors abiem atvejais gaunami rezultatai stipriai priklauso nuo kelių dedamųjų (pradinis kintamųjų skaičius, siekiamas kintamųjų skaičius) autoasociatyvinių neuroninių tinklų metodas yra žymiai tikslesnis už PCA, todėl tolimesniuose tyrimuose didžiausias dėmesys turi būti sutelktas būtent į autoasociatyvinių neuroninių tinklų taikymą.

PCA ir AANN praktikoje

Esama nemažai programinės įrangos atliekančios duomenų suspaudimą į principines komponentes, tačiau tik nedidelė šių programų dalis yra orientuota į finansų rinkas. Programinės įrangos, kuri nėra tiesiogiai skirta vertybinių popierių analizei vartojimas labai komplikuoja duomenų paruošimą bei rezultatų apdorojimą ir analizę. Dėl šios priežasties buvo pasirinktas komercinis produktas – PickStock.

PickStock (© 2003, Bell Software and Services, Inc.) yra kompiuterinė programa principinių komponentių pagalba ieškanti galimai neįvertintų ir pervertintų akcijų.

Programa paremta arbitražinio vertinimo teorija ("arbitrage pricing theory"). Ši teorija teigia, jog akcijos kaina gali būti suskaičiuota kaip tiesinė pagrindinių ekonominių rodiklių kombinacija. Šiuo metu teorija nėra įrodyta, dėl jos vyksta ginčai, tačiau programa PickStock besiremddama šia APT jau dabar leidžia ją panaudoti su tikrais duomenimis. Programos pagrindą sudaro ARPACK – sistema eigenvektorių skaičiavimui.

Visos akcijos yra apibūdinamos penkiais kriterijais - atidarymo kaina, aukščiausia dienos kaina, žemiausia kaina, uždarymo kaina, prekybos apimtis. Penkių kriterijų spaudimas į vieną principinę komponentę pradangintų didžiąją dalį informacijos, „spaudimas“ į penkias – išlaikytą tą pačią originalią informaciją. Rezultatai yra gaunami tekstinio failo formatu - akcijos yra įvertintos koeficientais nurodančiais santykį tarp dabartinės kainos ir apskaičiuotos vertės. Pavyzdžiui, įrašas „WDC -0,13616“ reiškia, kad, akcija WDC yra nuvertinta 13,6 %, todėl tai gan stiprus signalas ją pirkti. Įrašas „FRK 0,21374“ reiškia, kad, akcija FRK yra pervertinta 21,37 %, todėl tai stiprus signalas ją parduoti trumpam.

Eksperimento metu buvo suformuotos trys portfelių poros. Kiekviena pora sudaryta iš portfelio kuriame akcijos turi būti perkamos, bei portfelio su akcijomis kurios turi būti parduodamos trumpam.

Portfelių poros sudaromos PickStock programoje parenkant skirtingą principinių komponentių skaičių. Tam, kad įsitikintume korektišku programos veikimu, pirmoji portfelių pora bus generuojama naudojant 5 principines komponentes. Remiantis teorija, naudojant 3 principines komponentes informacijos nuostoliai turi būti nedideli, taigi šiuo metodu bus randamos antrojo portfelio akcijos. Perleidžiant duomenis per vieną principinę komponentę, duomenų praradimas yra neišvengiamas, tačiau, kad tuo įsitikintume suformuosime trečiąjį portfelį naudodami vieną principinę komponentę.

Portfeliai sudaryti naudojant 5 (ir daugiau) principinių komponentių, po 10 mėn:

Akcija	PickStock santykis	Pokytis
WDC	-0,07423	65%
PAYX	-0,06531	-10%
RHI	-0,04965	-23%
Vidurkis		11 %

Akcija	PickStock santykis	Pokytis
WSM	0,03745	-21%
NSSC	0,0427	12%
BAMM	0,33934	-39%
Vidurkis		-0,32

Šaltinis: sudaryta autoriaus

Lentelė 1. Portfeliai sudaryti naudojant 5 (ir daugiau) principinių komponentių



Šaltinis: sudaryta autoriaus

Pav. 3 WDC ir BMM palyginimas su S&P500 indeksu

Geriausios (WDC – viršutinis grafikas) ir prasčiausios (BMM – apatinis grafikas) akcijų kitimas visiškai atitiko prognozes sudarytas principinių komponentų metodu. Vidurinis – SP500 indeksas liko beveik nepakitęs. Tikslus pasirinkto modelio vertinimas gaunamas imant kelias „geriausias“ ir „blogiausias“ akcijas. Be to, reikia atsižvelgti ir į programos apskaičiuotą tikimybės koeficientą – „PickStock santykis“:

Portfelis (Top3).

$$\frac{0,074 \times 0,65}{0,074 + 0,065 + 0,049} + \frac{0,065 \times (-0,1)}{0,074 + 0,065 + 0,049} + \frac{0,049 \times (-0,23)}{0,074 + 0,065 + 0,049} = 0,11$$

Portfelis (Worst 3)

$$\frac{0,33934 \times (-0,39)}{0,33934 + 0,0427 + 0,03745} + \frac{0,0427 \times 0,12}{0,33934 + 0,0427 + 0,03745} + \frac{0,03745 \times (-0,21)}{0,33934 + 0,0427 + 0,03745} = -0,32$$

Matyti, kad Top3 portfelis per 10 mėnesių padidėjo 11%, o Worst3 nuvertėjo 32%. Nors rezultatai nėra tokie įspūdingi, kaip skaičiuojant tik pačią geriausią ir prasčiausią akciją, šis metodas išlieka pelningas, ypač atsižvelgiant į tai, kad S&P500 indeksas per šį laiką nukrito 1,7%.

Išvados

Praėjus 10 mėnesių nuo duomenų rinkimo ir portfelių formavimo dienos galima teigti, jog principinių komponentų metodas pasiteisino – visi keturi portfeliai buvo pelningi. Abejonių kelia tik tai, kad teoriškai nepatikimesni portfeliai sudaryti naudojantis tik 3 principinėm komponentėm negu atnešė didesnę pelną nei teoriškai tiksliau sumodeliuotieji.

Žvelgiant iš šalies, silpnai atrodo pradiniai duomenys – egzistuoja tik akcijos kaina dienos bėgyje (atidarymo, aukščiausia, žemiausia, uždarymo) ir prekybos apimtis. Visiškai nėra atsižvelgiama į fundamentinius ar techninius rodiklius. Įvesdami kainos priklausomybę nuo fundamentinių ar techninių parametrų neišvengiamai susidurtume su netiesinėm priklausomybėm, poreikiu reaguoti į gaunamus tarpinius rezultatus – apsimokymu. Tokiais atvejais tikslingiau būtų naudoti autoasociatyvinius neuroninius tinklus.

Autoasociatyvinių neuroninių tinklų taikymas vertybinių popierių rinkoje yra palyginti nauja kryptis ir įvairių mokslinio tyrimo darbų objektas.

Realias jų efektyvaus panaudojimo galimybes kol kas sunku numatyti.

Literatūra

1. R. Simutis. Paskaitų konspektai. Intelektinės sistemos vertybinių popierių rinkose
2. Why Our Investment Strategy Works. <http://tdmresearch.com/WhyItWorks.htm> (žiūrėta 2006-12-02)
3. Q10: What is Neural Network Forecast? Why does it seem to change from time to time? <http://www.tradetrek.com/faqs/q10.asp> (žiūrėta 2006-12-01)
4. George T. Albanis ir Roy T. Batchelor. Predicting High Performance Stocks using Dimensionality Reduction Techniques based on Neural Networks <http://www.staff.city.ac.uk/r.a.batchelor/Dimreds.pdf> (žiūrėta 2006-10-17)
5. N. Basalto, R. Bellotti ir kt. Clustering stock market companies via chaotic map synchronization. http://arxiv.org/PS_cache/cond-mat/pdf/0404/0404497.pdf
6. Lou Mendelsohn „Using Neural Networks For Financial Forecasting“. http://www.tradertech.com/financial_forecasting.asp (žiūrėta 2007-06-20)
7. Kramer, M. A. Nonlinear principal component analysis using autoassociative neural networks. Journal of the American Institute of Chemical Engineers, Vol 37, 1991, p. 233-243.
8. Pinigai daro pinigų. <http://www.spekuliantai.lt/> (žiūrėta 2007-06-20)
9. Forex technical and fundamental indicators. <http://forexindicators.blogya.de/forexindicators/> (žiūrėta 2007-06-20)
10. Techninė analizė – indikatoriai. <http://www.metaquotes.net/techanalysis/indicators/> (žiūrėta 2007-06-20)
11. Vikipedija, laisvoji enciklopedija. <http://lt.wikipedia.org/> (žiūrėta 2007-06-20)
12. PickStock dokumentacija, (©2003, Bell Software and Services, Inc.) <http://www.bellsys.com/Documentation.htm> (žiūrėta 2008-02-12)
13. <http://www.verslumas.lt/?serv=glossary&what=&cat=24&page=7> (žiūrėta 2008-03-01)

STOCK FORECAST USING AUTO ASSOCIATIVE NEURAL NETWORKS

Stable profit in the stock trading market can be achieved closely analyzing the general market and financial indicators of particular companies. Advanced investors are monitoring 50-100 indicators that represent general market trends and individual company's data. It's a hard job. PCA and AANN are techniques that allow compressing multiple n-series variables into human friendly 2D or 3D series.

PCA transformation is a linear process, while AANN is a non-linear self-organizing process. Due to complex relations between the financial indicators, the AANN is more preferable way to execute the transformation. Using PickStock application, two stock portfolios of three most underpriced and three most overpriced stocks were built. Results of the "best" and "worst" stocks after one year were astonishing - best stock gained 65% up, while the worst dropped by 39%. For mitigating the risk of single stock failure, it would be wise to compare portfolios containing different sets of stocks - the Top3 portfolio has generated 11% profit, while the Worst3 remained almost the same. Despite the fact, that all the portfolios were profitable, the weak point of the experiment is data - only the daily price and volume and were evaluated, thus forgetting about fundamental and technical indicators. Conclusion: although the method works fine for building the portfolio, it is most suitable for extracting a cluster of attractive stocks, which are worth further investigation.

PROGRAMOS AUTOASSOCIATIVE_ANN IŠEITIES KODAS

```
% Netiesiniu principiniu komponenciu formavimas naudojant autoasociatyvinius neuroninius tinklus
% =====
close all, clear all
% uzkraunamas duomeniu failas duomt.txt PowerScreener Lite duomenimis
Input_out=load('duomt.txt');
Input=Input_out(:,1:10);
[Input_s,sc1t]=dscale(Input');
Input_sc=Input_s';
% Nustatoma AANN struktura - 10 neuronu pasleptame sluoksnyje ir 2 neuronai "butelio kakliuko"
sluoksnyje
Ann_Struc=[10 2];
c{1}='tansig'; c{2}='tansig';
% AANN tinklo konstravimas
net=makebottle(Input_sc,Ann_Struc,c);
% sukonstruoto tinklo treniravimas
net=init(net);
net.trainParam.epochs = 200; % treniravimo iteraciju skaicius
net=train(net,Input_sc',Input_sc');
Y=sim(net,Input_sc');

%-----
% Visu imoniu rodikliu principines komponentes
figure(1)
plot(Y(1,:), 'r-'), hold on
plot(Y(2,:), 'b-'), grid

% Originaliu ir atstatytu indikatoriu palyginimas
%-----
figure(2)
subplot(2,2,1), plot(Input_sc(:,1)), hold on, plot(Y(3,:), 'r'), grid
subplot(2,2,2), plot(Input_sc(:,2)), hold on, plot(Y(4,:), 'r'), grid
subplot(2,2,3), plot(Input_sc(:,3)), hold on, plot(Y(5,:), 'r'), grid
subplot(2,2,4), plot(Input_sc(:,4)), hold on, plot(Y(6,:), 'r'), grid
% error of reconstruction
Total_RMSE=mean(mean(abs(Input_sc(:,1:4)-Y(3:6,:))))

%Imones atspindinciu principiniu komponencios atvaizdavimas
figure(3)
plot(Y(1,:), Y(2,:), 'ob'), grid
xlabel('PC_1'), ylabel('PC_2')
```

Failas **dscale.m** atlieka pradinio duomenų vektorių normalizavimą.

```
function [X,Xscale]=dscale(X,Xscale)
% DSCALE
% -----
% [Xs,Xscale]=dscale(X) scales data to zero mean and variance 1.
%
% Xs=dscale(X,Xscale) scales data using the scaling parameters in
% Xscale: Xs(k,:) = [X(k,:) - Xscale(k,1)]/Xscale(k,2)
% Typically used to scale a test data set with the same scaling parameters
% that were used for scaling the training data to zero mean and variance 1.
%
% INPUTS:
% X - Data matrix.
% (dimension is # of data vectors in matrix * # of data points)
% Xscale - See below.
%
% OUTPUTS:
% Xs - Scaled data matrix.
% Xscale - Matrix containing sample mean (column 1) and standard
% deviation (column 2) for each data vector in X.
%
% See the function WRESCALE on how to rescale the weights of the
% trained network.

% Written by Magnus Norgaard, IAU/IMM, Technical University of Denmark
% LastEditDate: Jan. 8 2000
```

2 PRIEDAS (TĘSINYS)

```
if nargin==0,
    error('DSCALE called with no arguments.');
```

```
end
[r,N] = size(X);          % r = # of data vectors, N = # of data
if nargin==1,
    Xscale = [mean(X')' std(X')']; % Col. 1 contains mean values, col 2 the std's
else
    if (size(Xscale,1)~=r | size(Xscale,2)~=2)
        error('Dimension mismatch between "X" and "Xscale"');
    end
end
for k=1:r,
    X(k,:) = (X(k,:) - Xscale(k,1))/Xscale(k,2);
End
```

Failas **makebottle.m** atlieka neuroninio tinklo formavimą pagal nurodytą neuronų skaičių paslėptame bei viduriniame „bottleneck“ sluoksniuose.

```
function net=makebottle(X,v,s)
%FUNCTION NET=MAKEBOTTLE(X,V,S)
%Returns a bottleneck network ready for initialization
%and training. This constructs it in the standard approach, where
%the network has 5 layers,
%
%input-encoding-bottleneck-decoding-output
%
%The input:
% X=data set. Input as number of points x dimension
% v=vector of 2 numbers for the sizes of the encoding and bottleneck layers.
% (It is possible to define each layer's nodes. See bottle1.m)
% s=Cell consisting of 2 strings, corresp to transfer function at each hidden layer.
%
%EXAMPLE: Data set X is three dimensions, reduce it to two using
% 5 nodes in the hidden layers, with nonlinear transfer functions at
% the hidden layers, so the ending network is 3-5-2-5-3
%
% c{1}='tansig'; c{2}='tansig';
% net=makebottle(X,[5,2],c);
%
%AFTER construction, the following commands initialize
% and train the network: (Default training: Leven.-Marq.)
%
%net=init(net);
%net=train(net,X',X');
%
%To simulate the network using data in P (numpts x dim)
% Y=sim(net,P');
% THE FIRST K ROWS OF Y correspond to output at
% the bottleneck! The remaining rows are output
% at the end of the bottleneck.
%
%
%The bottleneck neural net is used to perform nonlinear
%principle components analysis. For the primary literature
%reference, see: Kramer, "Nonlinear Principal Component
%Analysis Using Autoassociative Neural Networks", AIChE Journal,
%1991, 37(2), 233-243
%
%Written April 2001
%Doug Hundley
%Whitman College, Mathematics Department
%hundledr@whitman.edu
net=network;
net.numInputs=1;
net.numLayers=4;
net.biasConnect=[1;0;1;1]; %Need a bias at the output?
net.inputConnect(1)=1;
net.layerConnect(2,1)=1;
net.layerConnect(3,2)=1;
net.layerConnect(4,3)=1;
net.outputConnect(4)=1;
```

```

net.outputConnect(2)=1;
net.targetConnect(4)=1;
[numpts,dim]=size(X);
M=minmax(X');
net.inputs{1}.size=dim;
net.inputs{1}.range=M;
net.layers{1}.size=v(1); %Matlab has 4 layers
net.layers{3}.size=v(1);
net.layers{1}.transferFcn=s{1};
net.layers{3}.transferFcn=s{2};
net.layers{2}.size=v(2);
net.layers{2}.transferFcn='purelin';
net.layers{4}.size=dim;
net.layers{4}.transferFcn='purelin';
net.layers{1}.initFcn='initnw';
net.layers{2}.initFcn='initnw';
net.layers{3}.initFcn='initnw';
net.layers{4}.initFcn='initnw';
%Now initialize the functions for the network
net.initFcn='initlay';
net.performFcn='mse';
net.trainFcn='trainlm';

```