

**Jihočeská univerzita v Českých Budějovicích
Přírodovědecká fakulta**

Bakalářská práce

2012

Vlastimil Matějka

**Jihočeská univerzita v Českých Budějovicích
Přírodovědecká fakulta**

Technická analýza – burzovních dat

Bakalářská práce

Vlastimil Matějka

Školitel: Ing. Jiří Jelínek, CSc.

České Budějovice 2012

Bibliografické údaje:

Matějka, V.,2012: Technická analýza – burzovních dat. [Technical analysis – stock data. Bc. Thesis, in Czech.] – 36 p., Faculty of Science, University of South Bohemia, České Budějovice, Czech Republic.

Abstrakt:

Práce se zabývá predikcí v oblasti budoucího vývoje na akciové burze. Pomocí neuronové sítě a indikátorů technické analýzy se v této práci budu snažit odhadnout pohyb trendů na burze.

Abstract:

This work deals prediction in future developments in stock market. Using neural network and indicators technical analysis in this work i will try estimate move trends in stock market.

Klíčová slova:

neuronová síť, technická analýza, akcie, RapidMiner

Key words:

neural network, technical analysis, stock, RapidMiner

Prohlášení

Prohlašuji, že svoji bakalářskou práci jsem vypracoval samostatně pouze s použitím pramenu a literatury uvedených v seznamu citované literatury. Prohlašuji, že v souladu s § 47b zákona c. 111/1998 Sb. v platném znění souhlasím se zveřejněním své bakalářské, a to v nezkrácené podobě, elektronickou cestou ve veřejně přístupné části databáze STAG provozované Jihočeskou univerzitou v Českých Budějovicích na jejích internetových stránkách, a to se zachováním mého autorského práva k odevzdanému textu této kvalifikační práce. Souhlasím dále s tím, aby toutéž elektronickou cestou byly v souladu s uvedeným ustanovením zákona c. 111/1998 Sb. zveřejněny posudky školitele a oponentů práce i záznam o průběhu a výsledku obhajoby kvalifikační práce. Rovněž souhlasím s porovnáním textu mé kvalifikační práce s databází kvalifikačních prací Theses.cz provozovanou Národním registrem vysokoškolských kvalifikačních prací a systémem na odhalování plagiátu.

V Českých Budějovicích, dne 25. 4. 2012

Podpis:

Poděkování

Rád bych poděkoval Ing. Jiřímu Jelínkovi, CSc. za cenné rady při zpracování bakalářské práce.

Obsah

1. Úvod a cíle práce.....	1
1.1 Úvod.....	1
1.2 Cíle práce	1
2. Teoretická východiska.....	2
2.1 Základní pojmy	2
2.1.1 Akcie.....	2
2.1.2 Trhy	2
2.2 Analýzy trhu	3
2.2.1 Psychologická analýza.....	3
2.2.2 Fundamentální analýza	3
2.2.3 Technická analýza	3
2.2.4 Dowova teorie (Dow Theory).....	3
2.3 Indikátory technické analýzy	6
2.3.1 Klouzavé průměry	6
2.3.2 Oscilátory	8
2.3.3 Objemové indikátory	10
2.3.4 Sentiment indikátory	12
2.3.5 Indikátory šíře a výkonosti trhu.....	13
2.4 Neuronové sítě	14
2.4.1 Umělý neuron.....	14
2.4.2 Modely neuronových sítí.....	15
2.4.3 Učení neuronových sítí.....	16
3. Praktické experimenty.....	18
3.1 Výběr dat.....	18
3.2 Aplikace pro přípravu dat	18
3.2.1 Načtení dat.....	18
3.2.2 Uložení dat.....	18
3.2.3 Metody zpracování dat	19
3.2.4 Ovládání neuronové sítě	19
3.3 Experimenty s Neuronovými sítěmi	20
3.3.1 Predikce trendu	20
3.3.2 Predikce hodnoty indikátoru.....	22

4. Závěry práce.....	24
4.1 Experimenty s neuronovými sítěmi	24
4.1.1 Predikce trendu	24
4.1.2 Predikce hodnoty indikátoru.....	29
4.2 Závěr	34
5. Literatura	35
6. Přílohy	36

1. Úvod a cíle práce

1.1 Úvod

Práce se zabývá problematikou, předpovídání budoucího vývoje akcií na akciovém trhu. Toto téma je v době nestability na světových finančních trzích, kdy se jen těžko nechá odhadnout pohyb akcií velmi aktuální a pro mě velmi zajímavé.

Tuto problematiku se budu snažit v této práci řešit pomocí vybraných indikátorů technické analýzy a implementace neuronové sítě. Nejdříve budou v práci popsány základní pojmy a typy a analýz trhu. Následně budou specifikovány a rozčleněny jednotlivé indikátory technické analýzy, které budou implementovány v java aplikaci. Tato aplikace bude čerpat data z burzy a připravovat je pomocí metod technické analýzy pro neuronovou síť.

V závěru práce bude vyhodnocení experimentů prováděných pomocí technických indikátorů a neuronových sítí.

1.2 Cíle práce

1. Charakterizovat technickou analýzu burzovních dat - základní pojmy a metody s důrazem na metody umělé inteligence, zejména neuronové sítě.
2. Postupy výběru a předzpracování burzovních dat pro technickou analýzu s užitím neuronových sítí, aplikace v Javě pro získání a přípravu těchto dat.
3. Technická analýza burzovních dat a užitím neuronových sítí s důrazem na získání výstupů prakticky využitelných v obchodování s akciemi; implementace neuronové sítě.
4. Analýza a vyhodnocení získaných výsledků, závěry a doporučení.

2. Teoretická východiska

2.1 Základní pojmy

2.1.1 Akcie

Akcie patří do skupiny cenných papírů. S akciemi se obchoduje na akciových burzách. Majitelem akcie se může stát jakákoliv fyzická osoba.

Majitel akcie se stává automaticky akcionářem a získává tak podíl na majetku firmy. Akcionář má také nárok na podíl ze zisku. Velikost podílu na zisku je závislá na počtu akcií, které vlastní. Akcionář může hlasovat na valné hromadě společnosti.

2.1.2 Trhy

Finanční trhy se dělí na čtyři základní typy měnový, akciový, komoditní a trh s dluhopisy.

Měnový trh je trh, na kterém se obchoduje s měnami jednotlivých států. Kurzy měn jednotlivých států jsou odvislé od jejich ekonomické výkonnosti, zadlužení a spousty dalších faktorů. Měnový trh má suverénně nejvíc účastníků ze všech výše uvedených trhů.

Trh s akciemi se zabývá obchodováním s již dříve zmíněnými akciemi. Setkávají se zde tři strany emitent neboli ten kdo akcie vydává, investor ten kdo obchoduje na burze a prostředník tedy burza.

Komoditní trh je místo kde se obchoduje s takzvanými komoditami. Patří sem drahé kovy, obiloviny, ropa, zemní plyn ale třeba i káva a spousta dalších druhů rostlin a dalších surovin. Nepatří sem věci vyrobené člověkem jako například stroje, budovy a podobně. Jako na akciovém trhu zde probíhá nákup a prodej komodit.

Posledním hlavním trhem je **dluhopisový**. Obchoduje se s dluhopisy a dalo by se říci, že se prodávají a kupují závazky. Nejdůležitějšími údaji na dluhopise je datum splatnosti a částka, dlužná částka. Každý majitel dluhopisu má právo, aby v datu splatnosti mu byla dlužná částka uhrazena.

2.2 Analýzy trhu

2.2.1 Psychologická analýza

Jak lze již z názvu odvodit zabývá se především lidmi samotnými. Zaměřuje se na to, co ovlivňuje chování účastníků trhu, davovou psychologií a podobně. Používá se spíše jako doplňková analýza k technické a fundamentální analýze.

2.2.2 Fundamentální analýza

Fundamentální analýza vychází z toho, že data na burze nejsou vše, co ovlivňuje kurz akcie a zabývá se vnějšími vlivy, jako jsou například zisky, tržby a podobně. Fundamentální analýza se snaží najít vnitřní hodnotu akcie. Vnitřní hodnota akcie nám ukazuje či je daná akcie nadhodnocená či naopak podhodnocená neboli pravou hodnotu akcie.

2.2.3 Technická analýza

Technická analýza je nejstarší ze tří výše uvedených metod analýzy trhu. Vychází většinou z veřejných burzovních dat a soustřeďuje se zejména na hledání souvislostí v jejich historickém vývoji, aby dokázala předpovědět budoucí vývoj. Oproti oběma předchozím typům analýz se technická analýza nezabývá ani psychologií účastníků na trhu, ani vnějšími vlivy na akcie, nýbrž předpokládá, že všechny tyto faktory ovlivňující cenu akcie jsou již zahrnuty v její ceně. Jako u předchozích dvou analýz i tato analýza má své příznivce i odpůrce. Obě strany mají své argumenty pro a proti.

Základní předpoklady technické analýzy jsou:

1. Tržní ceny zahrnují veškeré potřebné informace o akciích
2. Tržní ceny se pohybují v trendech
3. Vývoj na trhu se opakuje.

2.2.4 Dowova teorie (Dow Theory)

Charles Dow se narodil 6. listopadu 1851 ve Sterlingu ve státě Connecticut. Jeho teorie vznikla z článků a studií, které publikoval ve vlastních novinách, z nichž později vznikl The Wall Street Journal přesněji v červenci 1889. Články dali vzniknout teorii díky, které je dodnes Charles Dow považován za otce technické analýzy.

Dowova teorie se zabývá, zjednodušeně řečeno tím, kdy akcie prodat či koupit. Charles Dow pro tyto účely definoval dva základní indexy, které jsou v podstatě základem teorie. Tyto indexy jsou DJIA (Dow Jones Industrial Average) a DJRA (Dow Jones Railroad Average).

Dow Jones Industrial Average (DJIA)

V roce 1896 přišel Dow s myšlenkou, že původní index Railroad Average z roku 1884 je schopen popsat pouze část ekonomiky, neboť obsahuje jenom úzký vzorek akcií, a proto je třeba vytvořit ještě další index, jenž by obraz celé ekonomiky rozšířil, zpřesnil a potvrdil. Tímto indexem se stal Dow Jones Industrial Average (DJIA).¹ Tento index je používán zejména na americké burze a je složen z 30 akcií společností, s nimiž je nejvíce obchodováno na burze. Každá akcie má přitom jinou váhu, která je odvislá od ceny akcie. Logicky největší váhu má akcie s největší cenou.

Dow Jones Railroad Average (DJRA)

Vznikl v roce 1884 a vycházel z 11 akcií železničních firem. Později byl přejmenován na současný název DJTA (Dow Jones Transportation average) a to kvůli rozšíření počtu akcií na celkových dvacet a počítání průměru nejen ze železničních společností, ale i z ostatních přepravních firem. Většina z dvaceti akcií transportních firem patří dodnes železničním přepravním. Pro výpočet se berou závěrečné ceny akcií na burze.

Pro jasné potvrzení trendu se musí oba indexy jak DJIA tak DJTA pohybovat ve stejném směru. Objem obchodování na trhu také potvrzuje nastoupený trend trhu například, když objemy obchodování klesají tak lze očekávat medvědí trh.

Trendy

Budoucí vývoj kurzů (tj. pokračování trendu nebo jeho změnu) lze odvodit z minulé tržní situace.² Z hlediska pohybu trendu rozlišujeme dva druhy trendů.

1. Rostoucí (Býčí trh)
2. Klesající (Medvědí trh)

¹ VESELÁ, Jitka. *Investování na kapitálových trzích*. 2., aktualiz. vyd. Praha: Wolters Kluwer Česká republika, 2011, 789 s. ISBN 978-80-7357-647-9 (Váz.). s. 426

² VESELÁ, Jitka. *Investování na kapitálových trzích*. 2., aktualiz. vyd. Praha: Wolters Kluwer Česká republika, 2011, 789 s. ISBN 978-80-7357-647-9 (Váz.). s. 428

Býčí trh

Dalo by se říci, že se jedná o jakýkoliv trh, který roste. Musí však být zachováno pravidlo, že každé další maximum a minimum je větší než předchozí. Jedná se o trh kde další minimum je větší než předchozí a kde další maximum je větší než předchozí.

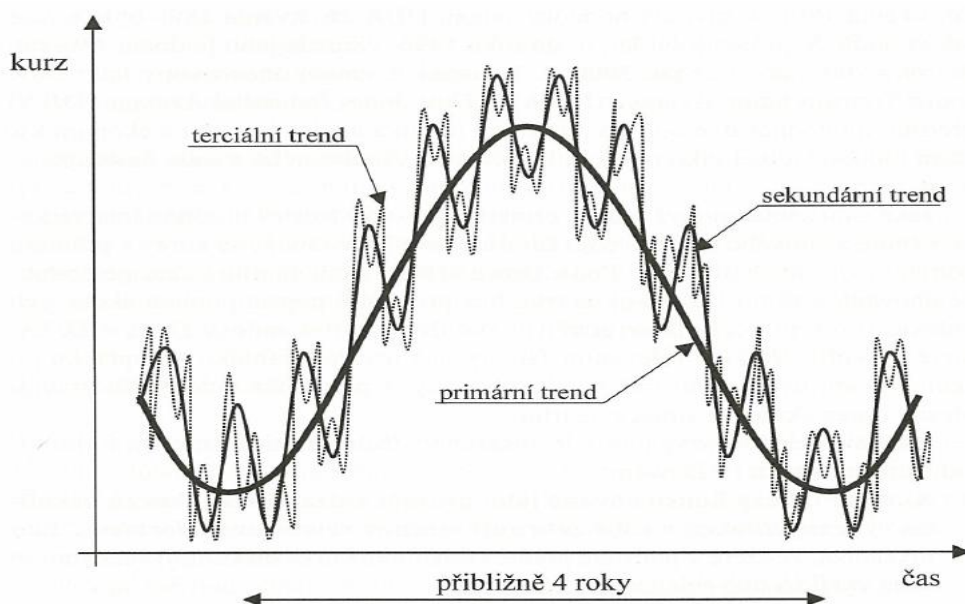
Medvědí trh

Je opakem Býčího trhu, jedná se o trh, který klesá. Jedná se o trend, kde každé další maximum a minimum je menší než předchozí.

Porušení trendu znamená, že je narušen systém poklesu maxim a minim. Při takovéto situaci lze očekávat změnu trendu z Býčího trhu na Medvědí trh a naopak.

Další možností jak rozlišovat trend na trhu je dle času trvání. Trendy dělíme na tři základní druhy.

1. Primární trend
2. Sekundární trend
3. Terciální trend



Obrázek rozlišení trendů³

³ Pramen: VESELÁ, Jitka. *Investování na kapitálových trzích*. 2., aktualiz. vyd. Praha: Wolters Kluwer Česká republika, 2011, 789 s. ISBN 978-80-7357-647-9 (Váz.). s. 428

Primární trend

Jedná se o hlavní trend. Z časového hlediska se jedná o dlouhodobý trend. Jehož délka se pohybuje v řádech let.

Sekundární trend

Jedná se o kratší trend než primární. Délka trendu bývá v řádech týdnů až měsíců. Vyjadřuje většinou poklesy na Býčím trhu a růsty na Medvědím trhu.

Terciální trend

Terciální trend je nejkratší ze všech tří trendů, které se rozlišují. Jeho délka trvání bývá v délce od několika hodin přes dny až maximálně v řádech pár týdnů. Ukazuje jen krátké výkyvy na trzích.

2.3 Indikátory technické analýzy

Pod tímto názvem si lze představit soubor indikátorů, které se na základě tržních dat jako jsou ceny akcií, objemy obchodování a podobně snaží odhadnout budoucí vývoj trendů na trhu. Indikátory jsou tak obchodníkovi, který využívá technickou analýzu nápomocny k tomu, aby se správně rozhodl kdy akcie nakoupit či prodat. Jelikož je indikátorů velký počet a fungují na různých principech, dělíme je do pěti základních skupin.

1. Klouzavé průměry a metody na nich založené
2. Oscilátory
3. Objemové indikátory
4. Sentiment indikátory
5. Indikátory šíře a výkonosti trhu

2.3.1 Klouzavé průměry

Patří v technické analýze k nejpoužívanějším indikátorům. Jedná se o průměr hodnot za určitou časovou jednotku od hodiny až po roky. Většinou se používají pro výpočet zavírací ceny na trhu. Podle postupu výpočtu dělíme klouzavé průměry takto:

1. Jednoduchý
2. Vážený

3. Exponenciální

4. Trojúhelníkový

5. Proměnlivý

Jednoduchý klouzavý průměr

Postup výpočtu je následující jednotlivé hodnoty za časovou periodu se sečtou a výsledek součtu se vydělí počtem hodnot. Problematické u tohoto průměru je to, že nezohledňuje ze kdy daná hodnota je, tudíž všechny hodnoty v dané periodě mají stejnou váhu, přestože nejnovější hodnoty se od těch nejstarších mohou značně lišit, proto výsledek může být zkreslený.

Vážený klouzavý průměr

Zavádí do aritmetického průměru váhy, které slouží pro přiřazení hodnotám různou důležitost. Podle stáří jsou těmto hodnotám přiřazovány různé velikosti vah. Nejnovější hodnoty potom mají největší váhu a logicky nejstarší tu nejmenší.

Exponenciální klouzavý průměr

Exponenciální klouzavý průměr dává největší váhu nejnovější hodnotě a starší hodnoty, tudíž mají nižší váhu. Aktuální přiřazovaná hodnota z kurzu má potom takzvanou vyhlazovací funkci. Tato hodnota je buď pevně stanovená analytikem, nebo je závislá na délce periody, kterou jsme zvolili.

Trojúhelníkový klouzavý průměr

Jedná se o kombinaci jednoduchých klouzavých průměrů. Největší váha je přiřazována prostředním hodnotám z délky periody.

Proměnlivý klouzavý průměr

Proměnlivý klouzavý průměr vychází z exponenciálního průměru. Podobně jako exponenciální průměr funguje na bázi vah a časové periody, nicméně rozdíl je v tom, že upravuje exponenciální poměr. Proměnlivý klouzavý průměr je nejcitlivějším ze všech dříve zmíněných klouzavých průměrů.

2.3.2 Oscilátory

Funkcí oscilátorů je měření síly trendu, jeho kolísavost a pravděpodobnost jeho změny. Z hlediska matematického se většinou nejedná o žádné složité vzorce většinou založeny na jednoduchých početních metodách, tudíž se celkem rychle a lehce počítají. *Pro tuto podskupinu technických indikátorů, jak již naznačuje její název, je charakteristické, že jejich hodnota zpravidla kolísá buď kolem nějaké úrovně, nebo v rámci nějakého pásma. Vzhledem ke své konstrukci jsou oscilátory mnohdy při splnění určitých podmínek schopny přinést nákupní a prodejní signály⁴.* Mezi oscilátory patří například Momentum, Relativ strength index, Williams %R indikátor, Rate of Change.

Monentum

Podobně jako indikátor RSI je jeho autorem J. Welles Wilder. Používá se pro zjištění rychlosti a síly trendu.

Matematický výpočet

$$Mom = c_x - c_{x-n}$$

Kde c_x je aktuální cena c_{x-n} je cenou o N dní dříve.

RSI (Relative strength index)

RSI neboli index relativní síly je oscilátor, který vyvinul v červnu roku 1978 J. Welles Wilder. Funkce indexu spočívá ve zjištění síly trhu a tak zjistit o jaký trh se jedná. Výsledné hodnoty se pohybují mezi 0 – 100. Jestliže se hodnota pohybuje pod číslem 30, jedná se o trh překoupený. Naopak, když se hodnota pohybuje nad číslem 70, jedná se o trh přeprodaný. Tyto hodnoty jsou pro investora signálem pro prodej či nákup daných akcií.

Matematický výpočet

$$RS = \frac{x}{y}$$

Kde x je průměr kladných změn v kurzu a y je průměr záporných změn v kurzu.

$$RSI = 100 - \left[\frac{100}{1 + RS} \right]$$

⁴ Pramen: VESELÁ, Jitka. *Investování na kapitálových trzích*. 2., aktualiz. vyd. Praha: Wolters Kluwer Česká republika, 2011, 789 s. ISBN 978-80-7357-647-9 (Váz.). s. 447

Williams %R indikátor

Indikátor byl vytvořen Larrym Williamsem v roce 1966. Výsledné hodnoty se pohybují mezi 0 – 100. Hodnoty mezi 80 – 100 znamenají překoupený trh a hodnoty mezi 0 – 20 znamenají přeprodaný trh. Signály od tohoto indikátoru by se měli brát v úvahu, jenom když jsou ve směru trendu.

Matematický výpočet

$$\%R = \frac{z_x - h_{x-n}}{h_{x-n} - l_{x-n}}$$

Kde z_x je aktuální cena, h_{x-n} je nejvyšší cena N dní zpět a l_{x-n} je nejnižší cena N dní zpět.



Obrázek průběhu Williams %R indikátoru, který vidíme ve spodní části.⁵

ROC (Rate of Change)

Úkolem tohoto oscilátoru je odhalení krátkodobých (maximálně dny) pohybů v trendu, což může předeslat i oslabení trendu. To lze z rate of change vyvodit tak, že dokud indikátor dosahuje nových a nových vrcholů znamená to, že trend sílí. Naopak, když dosahujeme nových maxim a najednou se objeví menší hodnota, než minulé maximum lze očekávat, že se trend změní. Když se tato událost bude opakovat lze s celkem vysokou jistotou tvrdit, že z Býčího trhu se stává trh Medvědí.

⁵ Pramen:

http://stockcharts.com/school/data/media/chart_school/technical_indicators_and_overlays/williams_r/wpr-4-tjxfail.png

Matematický výpočet

$$ROC = \frac{z_d - z_{d-n}}{z_{d-n}}$$

Kde z_d je dnešní zavírací cena a z_{d-n} je zavírací cenou N dní dříve.

2.3.3 Objemové indikátory

Objemové indikátory pracují vždy s údaji o objemu obchodů, které často kombinují s údaji o vývoji kurzu. Podskupina objemových indikátorů je schopna do jisté míry informovat o likviditě trhu, zájmu či nezájmu investorů.⁶ Objemové indikátory pracují na těchto základních předpokladech, které se zabývají vztahem kurzů a objemem obchodování na trhu.

1. Rostoucí kurzy s rostoucími objemy jsou dobré známky pro Býčí trh
2. Klesající kurz s rostoucími objemy znamenají velký počet prodávajících, což je potvrzení nastoupeného trendu
3. Klesající kurz ve spojení s klesajícími objemy obchodování znamená oslabení stávajícího trendu

Mezi hlavní objemové indikátory patří PVI, NVI, OBV, PVT.

PVT (Price and volume trend)

Vychází z On Balance Volume indikátoru, ale jednotlivým objemům přiřazuje váhy podle toho jako moc se změnila cena. Díky zavedení vah, které pomohli zohlednit, jak velká změna nastala se tento indikátor stal přesnější oproti OBV.

Matematický výpočet

$$PVT = PVT_{t-1} + \frac{c_t - c_{t-1}}{c_{t-1}} * Vt$$

Kde c_t je dnešní zavírací cena a c_{t-1} – včerejší zavírací cena.

OBV (On Balance Volume)

OBV indikátor vyvinul Joe Granville v roce 1963 a publikoval ho ve své knize. Jeho výpočet probíhá tak, že dnešní zavírací cena cenného papíru se porovná se včerejší na trhu a jestliže

⁶ Pramen: VESELÁ, Jitka. *Investování na kapitálových trzích*. 2., aktualiz. vyd. Praha: Wolters Kluwer Česká republika, 2011, 789 s. ISBN 978-80-7357-647-9 (Váz.). s. 459

je dnešní cena větší tak se dnešní denní objem přičte k celkovému součtu obchodů, když je cena menší tak se dnešní denní objem obchodování odečte od celkového součtu obchodů a když jsou si rovny tak nenastane žádná matematická operace.

Matematický výpočet

dnešní zavírací cena > včerejší zavírací cena

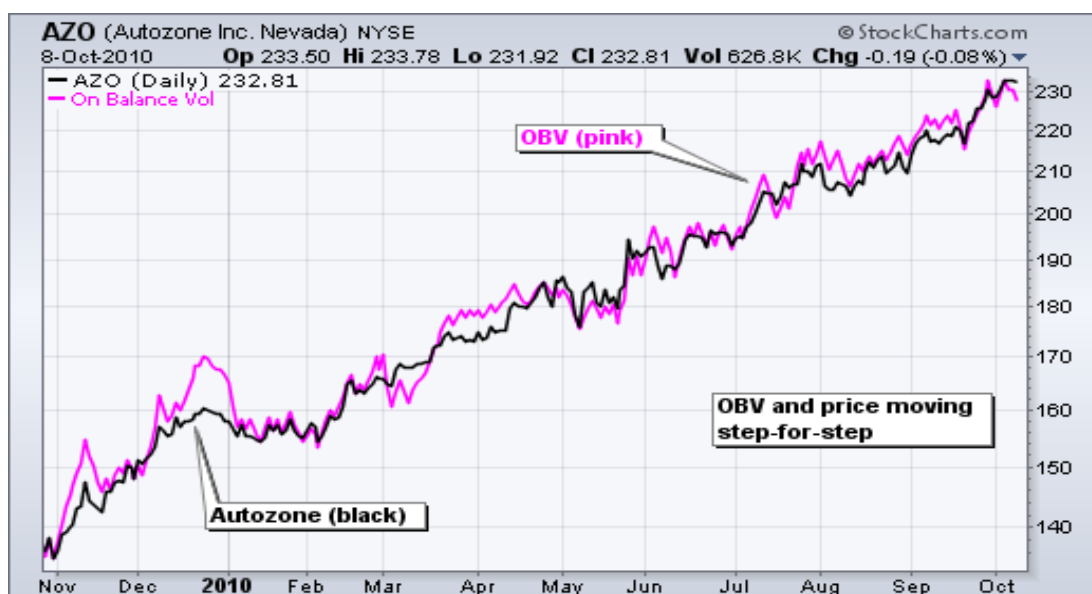
$OBV = OBV + \text{dnešní objem obchodování}$

dnešní zavírací cena < včerejší zavírací cena

$OBV = OBV - \text{dnešní objem obchodování}$

dnešní zavírací cena = včerejší zavírací cena

$OBV = OBV$



Obrázek průběhu On Balance Volume⁷

NVI (Negative Volume Index)

Index se používá v období klesajících objemů. Pomáhá identifikovat medvědí trh.

Matematický výpočet

Dnešní objem < včerejší objem $NVI = NVI_{t-1} + \left(\frac{c_t - c_{t-1}}{c_{t-1}} * NVI_{t-1} \right)$

⁷ Pramen: <http://stockcharts.com/h-sc/ui?s=IBM&p=D&yr=0&mn=6&dy=0&id=p32660999348&listNum=30&a=216754758>

PVI (Positive Volume Index)

Index se používá v období rostoucích objemů. Pomáhá identifikovat býčí trh.

Matematický výpočet

Dnešní objem > včerejší objem $PVI = PVI_{t-1} + \left(\frac{c_t - c_{t-1}}{c_{t-1}} * PVI_{t-1}\right)$

Dnešní objem <= včerejší objem $PVI_t = PVI_{t-1}$

2.3.4 Sentiment indikátory

Cílem sentiment indikátorů je zachytit chování, mínění a nálady účastníků obchodování, které jsou v krátkodobém časovém horizontu významným kurzotvorným faktorem. V pohledu sentiment indikátorů a psychologické analýzy na situaci na trhu je možné vysledovat mnoho společných rysů.⁸ Sentiment indikátory se dělí do dvou hlavních skupin anticyklické sentiment indikátory a cyklické sentiment indikátory.

Anticyklické indikátory

Tato skupina indikátorů se snaží odhadnout chování drobných investorů na trhu. K těmto indikátorům se vztahuje jedno základní pravidlo většina drobných investorů reaguje opožděně nebo špatně na vývoj trendu, proto zkušení investoři se chovají opačně, než naznačuje tato skupina indikátorů. Mezi anticyklické indikátory patří

1. Put/Call Ratio
2. Short Sales Ratio
3. Odd-lot Theory
4. Doporučení investičních poradců

Put/Call Ratio

Jedná se o sentiment indikátor, který se počítá jako podíl mezi nákupními obcemi a prodejními obcemi.

⁸ Pramen: VESELÁ, Jitka. *Investování na kapitálových trzích*. 2., aktualiz. vyd. Praha: Wolters Kluwer Česká republika, 2011, 789 s. ISBN 978-80-7357-647-9 (Váz.). s. 461

Short Sales Ratio

Vyjadřuje poměr tak zvaných prázdných prodejů. Jedná se o prodej akcií, které daný investor nevlastní nýbrž má je pouze zapůjčené. Tyto prodeje jsou pouhé spekulace a bývají předzvěstí medvědího trhu.

Odd-lot Theory

Tento indikátor je vyjádřen jako podíl mezi nákupními a prodejními příkazy v odd lot. Na burze se obchoduje v minimálních jednotkách. Odd lot jsou neúplné jednotky s kterými za poplatek obchodují menší investoři.

Doporučení investičních poradců

Jedná se o indikátor vycházející z předpokladu, že investiční poradci následují nastoupený trend a nesledují budoucnost.

Cyklické indikátory

Oproti anticyklickým indikátorům se tato skupina indikátorů soustřeďuje pouze na profesionální investory. Základní pravidlo se tak mění, oproti anticyklickým indikátorům vyplátí se následovat tyto indikátory. Mezi cyklické indikátory patří:

1. Barron's Confidence index
2. Struktura portfolia fondů

Barron's Confidence index

Index funguje na dvou předpokladech, že obchodníci na dluhopisových trzích dokáží rychleji odhadnout změnu trendu na trhu a že se tyto změny na depozitním trhu se zpožděním přenesou i na trh akciový.

Struktura portfolia fondů

Indikátor funguje na bázi podílu likvidních prostředků v aktivech fondů a z tohoto podílu se pak odvozuje budoucí vývoj na poptávku po akciích fondu.

2.3.5 Indikátory šíře a výkonosti trhu

Tyto indikátory sledují celkovou situaci na trhu. Zabývají se pohyby na trhu jako počtem klesajících akcií, počtem rostoucích akcií. Tedy nezabývají se jednotlivým akciemi, ale cel-

kem. Nejznámějším indikátorem patřící do této skupiny je Advance/Decline Line (A/D Line).

Advance/Decline Line (A/D Line)

Existují dva hlavní druhy A/D Line a to nekumulativní a kumulativní.

Nekumulativní A/D Line je matematicky vyjádřena jako denní rozdíl mezi počtem rostoucích a klesajících akcií. Tento indikátor šíře trhu je značně kolísavý, proto bývá ještě dělen celkovým počtem obchodovaných akcií nebo místo rozdílu mezi rostoucími a klesajícími akciemi používá podíl. Také se pro jeho vyhlazení používá klouzavý průměr.

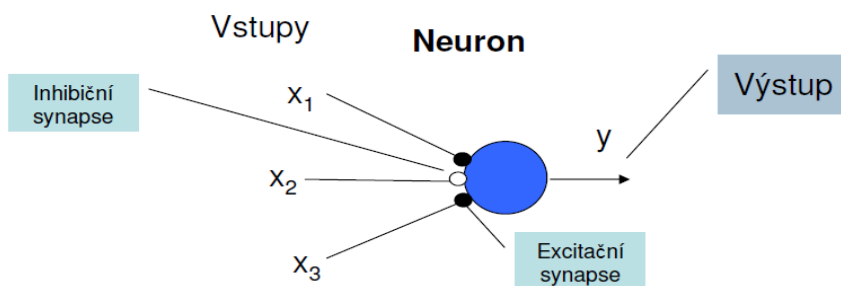
Kumulativní A/D Line je kumulativním součtem denního či týdenního rozdílu mezi rostoucími a klesajícími akciemi.

Z vývoje tohoto indikátoru a pohybu akcií lze odvodit budoucí vývoj trendu trhu. Například, když indikátor i akciový index roste lze očekávat i v budoucnu pokračující býčí trh.

2.4 Neuronové sítě

2.4.1 Umělý neuron

Umělý neuron je základem modelu neuronové sítě. Jeho konstrukce je inspirována biologickým neuronem. Prvním model neuronu vznikl již v roce 1943 díky spolupráci Warrena McCullocha a Waltera Pittse. Schéma tohoto neuronu vypadá následovně.



Obrazek McCulloch a Pitts neuronu⁹

Perceptron

V roce 1957 vzniká nový model neuronu zvaný perceptron. Autorem toho modelu je Frank Rosenblatt. Používá se v modelech jednovrstvých a více vrstevných neuronových sítích. Perceptron měl však jeden problém a to je, že nedokázal řešit funkce XOR. To zapříčinilo tažení proti neuronovým sítím. Tento problém byl vyřešen později při „vzkříšení“

⁹ Pramen: Neuronové sítě a Výpočetní inteligence přednášky z předmětu Neuronové sítě 2011 na PRF JCU dostupné z: <http://moodle.prf.jcu.cz>

neuronových sítí v 80 letech, pomocí vícevrstvé neuronové sítě perceptronů a objevení učícího algoritmu backpropagation o tom přesněji v kapitole o učení neuronových sítí.

Matematické vyjádření nelineární výstupní funkce perceptronu: $y = S(\sum_{i=1}^N w_i x_i + \Theta)$

RBF neuron

Používá se v sítích se dvěma vrstvami neuronů a dopředným šířením signálu. U RBF neuronů se změnila také metrika kde oproti perceptronům, které používají skalární součin tento model neuronu, používá eukleidovskou metriku.

výpočet vnitřního potenciálu

Výpočet výstupu

$$y = G\left(\sqrt{\sum_{i=1}^N (x_i - w_i)^2}\right)$$

$$y = \sum_{i=1}^n \binom{n}{k} w_i y_i^*$$

2.4.2 Modely neuronových sítí

Vícevrstvá perceptronová síť

Jedná se o vůbec nejpoužívanější model neuronových sítí. Základem této sítě je již dříve zmíněný perceptron. Tato síť obsahuje tři vrstvy a těmi jsou vstupní vrstva, skrytá vrstva a poslední vrstva výstupní. Nejpoužívanější metodou pro učení těchto sítí je backpropagation zmíněná v kapitole o učení neuronových sítí.

Hopfieldova síť

John Hopfield přišel s tímto modelem neuronové sítě na počátku 80. let. Základním stavebním prvkem sítě je energetická funkce. *Hopfieldova síť má tolik neuronů jako je počet jejích vstupů. Každý neuron je napojen na vstup sítě pouze jediným svým vstupem. Výstup každého neuronu je veden zpět na vstupy ostatních neuronů přes váhy w_{ij} , a tak se vytváří uzavřená smyčka. Tímto vzájemným uspořádáním vzniká symetrická síť.*¹⁰

Učení Hopfieldovy sítě probíhá ve dvou krocích. V prvním kroku proběhne nastavení vah podle vstupních vzorů. Druhým krokem je předkládání vzorů z trénovací množiny a opětované provádění prvního kroku dokud není trénovací množina vyčerpána.

¹⁰ Pramen: ŠNOREK, Miroslav. *Neuronové sítě a neuropočítače*. Vyd. 1. Praha: ČVUT, 1996, 124 s. ISBN 80-010-1455-X. s. 18

Vybavení Hopfieldovy sítě má 3 kroky. Nejdříve jsou inicializovány počáteční stavy podle předložených vstupů. Dalším krokem je iterace, která probíhá do doby, než se přestanou měnit váhy neuronů. Posledním krokem je předložení nového vzoru či konec procesu.

Kohonenova síť

Tento model sítě přichází ve stejném období jako Hopfieldovy sítě a to na počátku 80. let kdy neuronové sítě zažívají po období stagnace svojí renesanci. Patří do sítí, které se označují názvem samoorganizující sítě. Každý neuron v síti má tolik vstupů jako je celkový počet vstupů sítě. Princip fungování neuronů v síti se od sítí založených na perceptronech liší, protože každý neuron je schopen provádět akorát výpočet odchylky mezi předložený vzorem a vzorem, který má daný neuron. Kohonenova síť má jen dvě vrstvy a to vrstvu vstupní a výstupní a každý neuron je v podstatě výstupem této sítě.

Učení této sítě je učením bez učitele. Na začátku jsou neurony uspořádány náhodně. Učící algoritmus se snaží neurony v mřížce uspořádat tak, aby tyto neurony byli schopny klasifikace vstupních dat. Celá neuronová množina je tedy přeskládávána tak, aby co nejlépe odpovídala trénovací množině dat.

2.4.3 Učení neuronových sítí

Základní pojmy v učení neuronových sítí

Vzor je pár vstupní a výstupní vektor. **Trénovací množina** je skupina vzorů vybraná pro učení sítě. **Testovací množina** slouží k testování neuronové sítě. **Validační množina** slouží k testování sítě během učení. **Vybavení** je předložení jednoho vzoru neuronové síti a výpočet. **Iterace** je vybavení a provedení jednoho kroku učení. **Epocha** je provedení iterace na všech vzorech. **Přeučení** je situace, kdy neuronová síť ztrácí schopnost generalizace (zevšeobecnění) neboli dochází k nárůstu chyby na netrénovaných datech.

učení bez učitele

Při učení bez učitele nejsou neuronové sítě známy požadované výstupy. Tak jak je tomu při učení s učitelem. Celý proces učení si tak neuronová síť řídí sama. Neuronová síť tak hledá souvislosti neboli podobné vlastnosti v datech podle, kterých se je pak snaží rozřadit do skupin (shluků). Jako příklad sítě, která je učena bez učitele bych uvedl SOM (Self Organizing Map) samoorganizující mapy.

učení s učitelem

Učení s učitelem je založeno na minimalizaci chyby mezi momentální a požadovanou odezvou neuronové sítě. Roli učitele zde hraje požadovaná odezva sítě, kterou se snažíme neuronovou síť naučit. V průběhu učení změnami vah minimalizujeme chybu.¹¹

Algoritmus zpětného šíření (Backpropagation)

Jedním z nejpoužívanějších algoritmů pro učení neuronových sítí je algoritmus zpětného šíření (Backpropagation). Objev tohoto algoritmu byl pro neuronové sítě velmi významný, neboť v polovině osmdesátých let odstartoval, po téměř dvacetileté odmlce, novou vlnu rozvoje tohoto vědního oboru.¹²

Základním principem je zpětné šíření chyby a podle její velikosti následné úpravy vah a probíhá pomocí učení s učitelem. Při postup učení se nejdříve se náhodně inicializují váhy neuronů v síti. V dalším kroku přiložíme na vstup neuronové sítě páry vstup a výstup z trénovací množiny dat. Následně jsou vypočteny výstupy sítě. Poté vypočteme odchylku a podle velikosti odchylky upravíme váhy. Pokud je průměrná chyba i po projití celé trénovací množiny větší než požadovaná opakujeme celý proces. Naopak když je menší než stanovené kritérium můžeme proces ukončit.

Učení Backpropagation probíhá pomocí gradientního algoritmu. Takto vypadá delta pravidlo pro úpravu vah neuronů v neuronové síti.

Delta pravidlo pro úpravu vah

$$\Delta w_{ij}(t + 1) = \eta \delta_i(t) x_j + \alpha \Delta w_{ij}(t)$$

w_{ij} – změna váhy mezi i-tým a j-tým neuronem, η – rychlost učení Eta, x_j – j-tý vstup neuron, $\alpha \Delta w_{ij}(t)$ – moment Alfa, $\eta \delta_i(t) x_j$ - chyba připadající na j-tý neuron

Největším nepřítelem Backpropagation je uváznutí v lokálním minimum. Řešením tohoto problému je úprava parametru α v delta pravidle pomocí, kterého jsou upravovány váhy. To nám pomůže překonat lokální minimum. Dalšími možnostmi řešení lokálních minim jsou ještě úprava rychlosti učení v delta pravidle, úprava struktury trénovací množiny, náhodný výběr vzorů z trénovací množiny a použití gradientní metody s proměnným krokem.

¹¹ Pramen: Neuronové sítě a Výpočetní inteligence přednášky z předmětu Neuronové sítě 2011 na PRF JCU

¹² Pramen: ŠNOREK, Miroslav. *Neuronové sítě a neuropočítače*. Vyd. 1. Praha: ČVUT, 1996, 124 s. ISBN 80-010-1455-X. s. 29

3. Praktické experimenty

3.1 Výběr dat

Zdrojem burzovních dat jsem vybral americkou internetovou burzu Nasdaq. Z této burzy lze volně získat data akcií i několik let zpátky ve formě souboru s příponou csv to bylo jedna z hlavních předností, proč jsem si tuto burzu vybral. Soubor stažený z této stránky obsahuje závěrací ceny, otevírací ceny, denní maximum, denní minimum a objem obchodování.

3.2 Aplikace pro přípravu dat

Aplikace je napsána v programovacím jazyce Java a JavaFX. Jako vývojové prostředí jsem zvolil NetBeans, které je volně šiřitelné a mezi programátory hodně používané.

3.2.1 Načtení dat

Aplikace je schopna načíst data již ve dříve zmíněném formátu csv pomocí výběrového menu. Přesný formát tohoto souboru vidíme v následující tabulce.

	A	B	C	D	E	F	G	H
1	date,"close","volume","open","high","low"							
2	2012/03/16,"185.0500","4924986.0000","183.2800","185.6800","182.3500"							
3	2012/03/15,"184.4300","4158918.0000","182.0200","184.4300","180.3000"							
4	2012/03/14,"182.2600","3694639.0000","183.6500","184.3199","181.1400"							
5	2012/03/13,"184.5900","4405691.0000","183.9200","184.8700","180.7700"							
6	2012/03/12,"183.3900","2586637.0000","184.1300","185.4000","182.2000"							
7	2012/03/09,"184.3200","4510556.0000","186.7900","187.2000","183.4400"							
8	2012/03/08,"187.6400","4221035.0000","184.1700","188.3800","183.8000"							
9	2012/03/07,"183.7700","4950883.0000","182.6500","185.5000","182.4000"							

Obrázek tabulky vstupních dat

3.2.2 Uložení dat

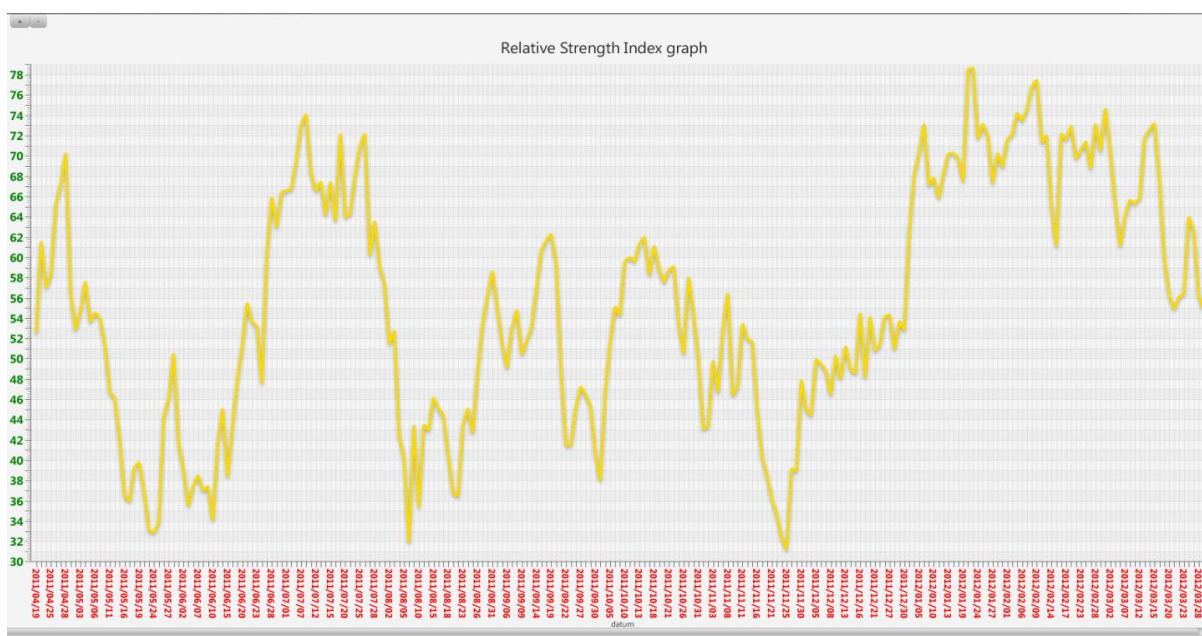
Zpracovaná data lze uložit do formátu csv. Při uložení jsou vynechána data, kde chybí hodnota některého z indikátorů technické analýzy, protože chybějící hodnoty stěžují učení vzorů neuronové sítě. Ukázku formátu uložených dat vidíme v následující tabulce.

	A	B	C	D	E
1	datum	roc	rsi	ema	sma
2	29.3.2012	-1.683501	55.671211	32.226492	32.2125
3	28.3.2012	0.4681647	56.870539	32.250157	32.285500
4	27.3.2012	1.6567677	62.825759	32.263526	32.343500
5	26.3.2012	1.8119337	64.155807	32.206532	32.3585
6	23.3.2012	0.5339195	57.138632	32.121317	32.3035
7	22.3.2012	1.3943907	56.935131	32.146054	32.301500
8	21.3.2012	0.3459119	55.825872	32.179621	32.303000
9	20.3.2012	-0.265003	57.118259	32.239537	32.295999

Obrázek tabulky uložených dat.

3.2.3 Metody zpracování dat

Celkově program používá čtyři indikátory ke zpracování a těmi jsou ze skupiny oscilátorů relative strength index a rate of change, ze skupiny klouzavých průměrů jsou to exponential average a simple moving average. Byly vybrány tyto indikátory, protože klouzavé průměry a oscilátory patří k nejpoužívanějším typům indikátorů technické analýzy na akciových trzích a implementované indikátory jsou jejich typickými zástupci. Tyto metody pracují s daty z načteného csv souboru. U všech metod zpracování se musí nastavit délka jejich period či vstupních hodnot, aby mohl být proveden jejich výpočet. Aplikace je schopna po dokončení zpracování všech metod vykreslit k těmto indikátorům také grafy. Ukázkou grafu relative strength indexu vytvořeného aplikací pro přípravu dat vidíme na následujícím obrázku.



Graf Relative Strength Index vytvořený aplikací.

3.2.4 Ovládání neuronové sítě

Pro implementaci neuronových sítí pro své experimenty jsem zvolil zdarma dostupný program RapidMiner. Vybral jsem RapidMiner, protože s tímto programem jsem již dříve pracoval během výuky Neuronových sítí a Data Miningu na univerzitě.

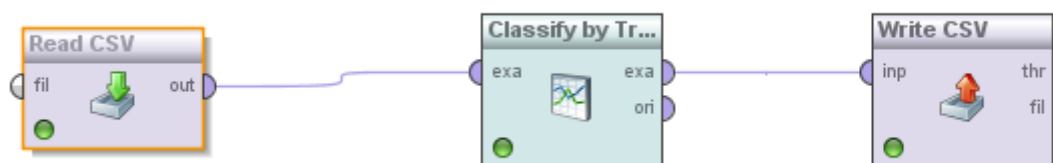
Díky importování knihoven RapidMineru do aplikace, je z programu možno spustit proces a získat výsledky z daného procesu. V aplikaci je implementováno celkem pět procesů, které jsou rozděleny do dvou logických celků, které představují dva experimenty. O experimentech se dočtete v následující kapitole.

3.3 Experimenty s Neuronovými sítěmi

Ve své práci provádím celkem dva experimenty. Prvním experimentem je pomocí neuronové sítě a technických indikátorů se pokusit predikovat budoucí pohyb trendu na burze. V druhém pokuse se snažím pomocí neuronové sítě a pohyblivého okna odhadnout budoucí hodnotu indikátoru exponenciálního průměru.

3.3.1 Predikce trendu

Pro provedení tohoto experimentu je zapotřebí celkem tři procesů.

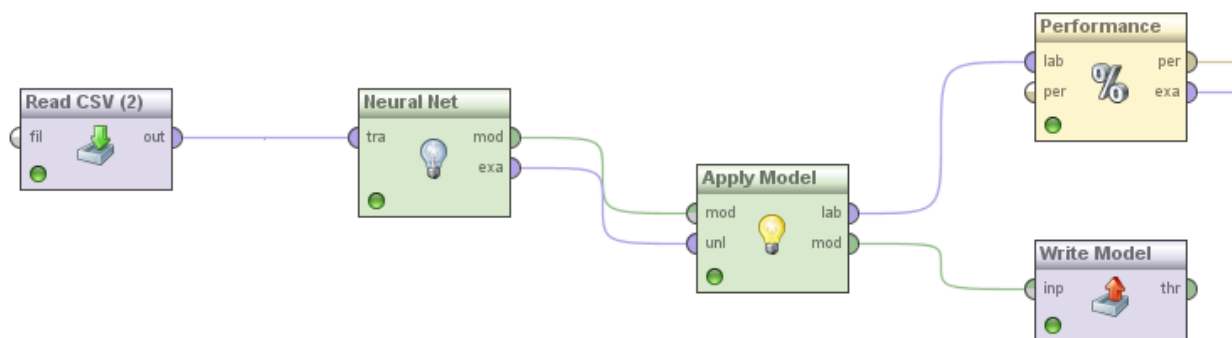


Obrázek schéma procesu přípravy dat.

U **prvního procesu**, jehož schéma vidíme výše, je nejdříve potřeba si zvolit jeden sloupec (hodnoty indikátoru) ze zdrojového souboru, který má název data.csv a nastavit je roli jako label. Poté pomocí Classify by trend u jednotlivých vzorů, získáme místo zvoleného sloupce sloupec label obsahující dvě hodnoty up a down. Tyto dvě hodnoty nám tedy představují pohyb trendu. Ukázkou dat získaných z tohoto procesu vidíme v následující tabulce. Kde vidíme, že pro zjištění trendu byl zvolen Exponenciální klouzavý průměr, který v datech získaných z procesu chybí a byl nahrazen sloupcem label. Proces takto vytvořená data uloží do csv souboru, který byl pojmenován dataTrend.

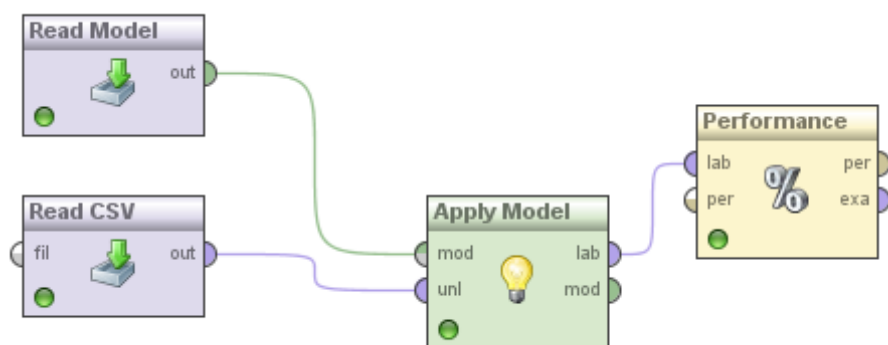
1	roc	rsi	sma	datum	label
2	2.3025263	65.519009	31.852000	9.3.2012	down
3	1.8129770	65.875331	31.801000	8.3.2012	down
4	1.8879999	64.336634	31.737000	7.3.2012	down
5	0.8630334	61.629467	31.679999	6.3.2012	down
6	5.8236272	65.626425	31.668499	5.3.2012	up
7	6.0330578	70.410310	31.613500	2.3.2012	down

Obrázek tabulky dat z procesu pro doplnění trendu.



Obrázek schéma procesu pro učení sítě.

Druhý proces, jehož složení vidíme výše, nám slouží k učení neuronové sítě. Vstupní data jsou čerpána ze souboru data1Tren.csv a jsou ve formátu, který má tři atributy (indikátory) a jeden sloupec sloužící jako label obsahující hodnoty up a down. Při načtení vstupních dat je nutno zvolit roli sloupce label jako label. Následně pak probíhá učení neuronové sítě. U neuronové sítě je také nutno nastavit počet neuronů ve skryté vrstvě a také počet trénovacích cyklů. Poté už jen proběhne uložení modelu a výpočet úspěšnosti učení sítě učení neboli velikosti klasifikační chyby.



Obrázek procesu pro predikci.

Třetí a zároveň poslední **proces**, který je v prvním experimentu zapotřebí. Má za úlohu načíst model sítě předcházejícího procesu učení. Pak už jen proces načte data z csv souboru data1Pred.csv, který obsahuje hodnoty tří indikátorů a datum a u nich provede predikci trendu. Pro potřeby ověření úspěšnosti predikce jsem já v tomto experimentu tyto data doplnil ještě o sloupec, který také obsahuje hodnoty up a down, ale jedná se o skutečné hodnoty. Tomuto sloupci byla nastavena role label. V následující tabulce, která slouží jen pro ukázkou, nejdříve vidíme predikovaný trend ve sloupci prediction (label) a skutečný trend ve sloupci label. V tabulce pod daty pak vidíme přesnost provedené predikce.

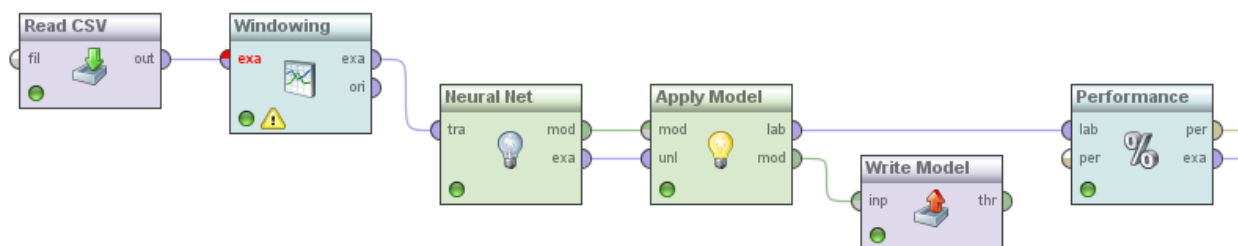
Row No.	datum	label	confidence(...	confidence(...	prediction(label)	roc	rsi	sma
1	30.3.2012	up	0.499	0.501	up	-1.572	57.522	32.178
2	29.3.2012	down	0.229	0.771	up	-1.684	55.671	32.212
3	28.3.2012	up	0.288	0.712	up	0.468	56.871	32.286
4	27.3.2012	up	0.858	0.142	down	1.657	62.826	32.344
5	26.3.2012	down	0.880	0.120	down	1.812	64.156	32.358

Obrázek tabulky dat získaných po predikci

accuracy: 78,57%	True up	True down	Class precission
Prediction up	5	2	71,43%
Prediction down	3	5	62,50%
Class recall	62,50%	71,43%	

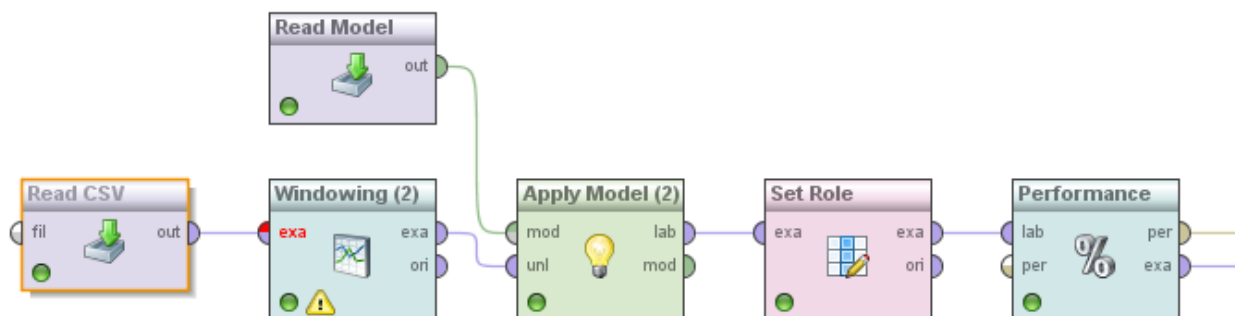
3.3.2 Predikce hodnoty indikátoru

Tento experiment obsahuje dva procesy a to proces učení a proces pro predikci.



Obrázek schéma procesu pro učení.

První proces jak již bylo napsáno, slouží k učení a jeho schéma vidíme na obrázku. Proces načítá data ze souboru data2Tren.csv souboru a pomocí operátoru Windowing je zpracovává. Při nastavení windowing horizon na hodnotu jedna funguje operátor následovně, vezme předchozí hodnotu zvoleného indikátoru a přiřazuje ji jako label hodnotu k následujícímu řádku. Tak nám vzniká nový sloupec label. Po dokončení práce operátoru Windowing, přichází na řadu učení sítě. Podobně jako v prvním experimentu je třeba nastavit počet neuronů ve skryté vrstvě a počet trénovacích cyklů. Po skončení učení sítě už jen proběhne uložení modelu naučené sítě a výpočet úspěšnosti procesu.



Obrázek schéma procesu pro predikci.

Druhý proces nám slouží k predikci indikátoru (exponenciálního klouzavého průměru). Model procesu vidíme výše. Proces načte model procesu učení a také data2Pred.csv soubor obsahující, pět atributů včetně data, které má nastavenou hodnotu id. Následně pak už jen proběhne predikce indikátoru a výpočet úspěšnosti. V následující tabulce je výstup z procesu, kde je vidět sloupec prediction label a zvýrazněn je i sloupec ema-0, který jsme se snažili predikovat.

Row No.	datum	prediction(label)	ema-0	roc-0	rsi-0	sma-0
1	21.3.2012	32.262	32.180	0.346	55.826	32.303
2	20.3.2012	32.295	32.240	-0.265	57.118	32.296
3	19.3.2012	32.282	32.295	-0.279	60.553	32.253
4	16.3.2012	32.128	32.316	2.710	67.800	32.212
5	15.3.2012	31.970	32.253	3.075	72.889	32.160
6	14.3.2012	31.849	32.120	4.530	72.268	32.104

Obrázek procesu dat získaných po predikci.

Všechny potřebné csv soubory pro procesy jsou umístěny v projektu Java aplikace včetně výše popsaných procesů.

4. Závěry práce

4.1 Experimenty s neuronovými sítěmi

Při svých experimentech jsem vycházel z burzovních dat společnosti Microsoft a to dva roky zpět od data 18. Dubna 2012. Data jsem si vybral aktuální, i přestože v současnosti lze jen těžko predikovat pohyby na trhu díky nedávné krizi. Zajímalo mě, jak si s touto situací poradí vybrané indikátory technické analýzy a neuronové sítě.

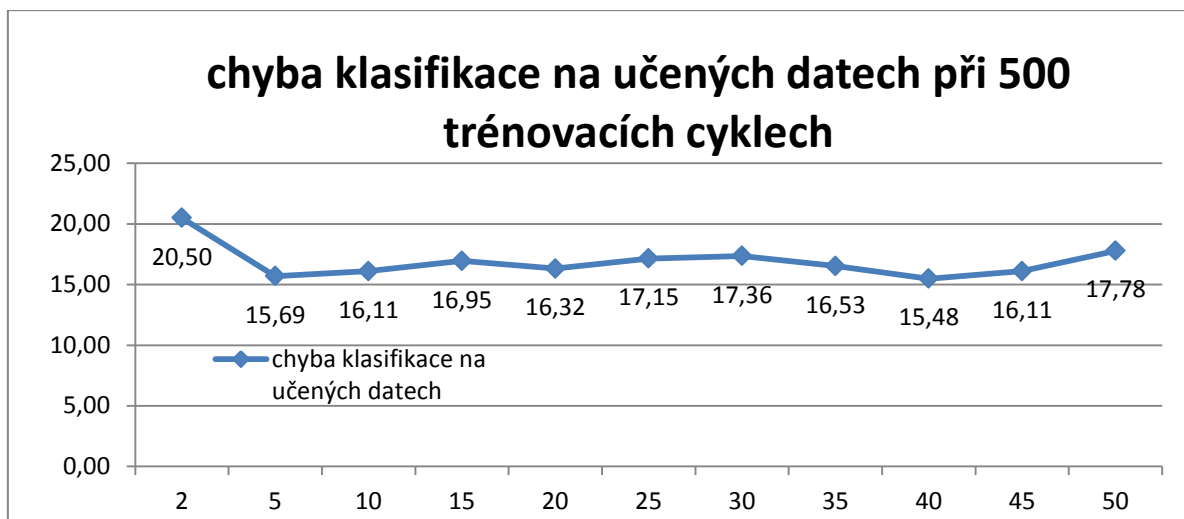
4.1.1 Predikce trendu

Postup experimentu

Při zpracování hodnot byli, nastaveny vstupní hodnoty pro zpracování následovně. Vstupní hodnota Relative strength index byla nastavena na čtrnáct. Vstupní hodnota Rate of change byla nastavena na dvanáct. Hodnoty exponenciálního a jednoduchého klouzavého průměru byli nastaveny na deset. Při nastavení vstupních hodnot jsem bral v potaz délku vstupního souboru a také doporučené hodnoty pro jejich používání. Po dokončení zpracování a přípravy dat procesem pro doplnění trendu byli data rozdělena do dvou csv souborů data1Tren.csv a data1Pred.csv. Trénovací soubor data1Tren.csv je určený pro učení sítě. Soubor data1Pred.csv pak obsahuje patnáct nejnovějších hodnot indikátoru pro predikci trendu. Pak byli tyto dva soubory, předloženy procesům pro učení a pro predikci. Nejdříve pak musí být spuštěn proces učení a pak logicky proces pro predikci.

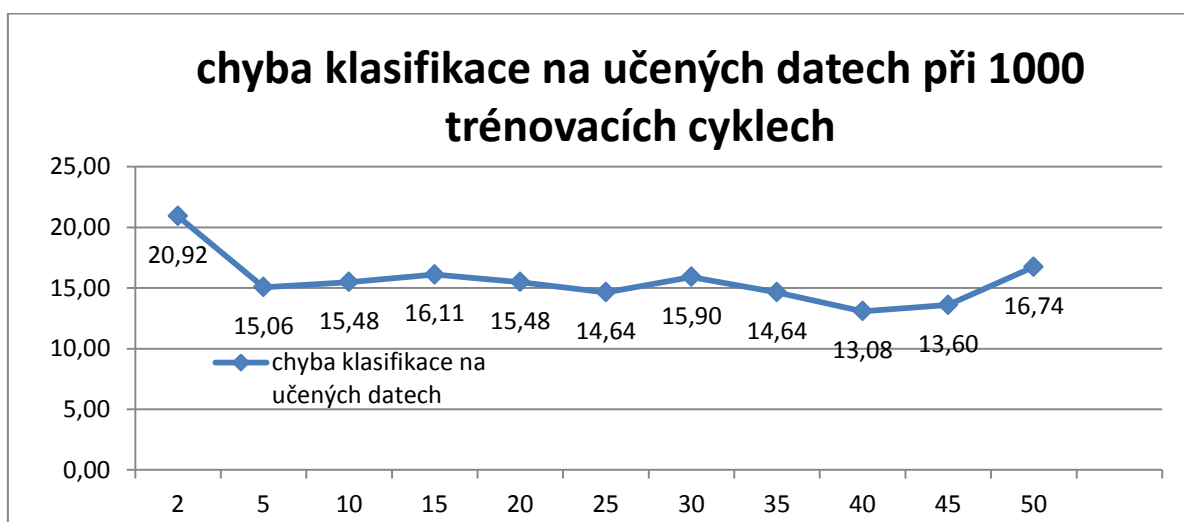
Experimenty s nastavením sítě

Při hledání optimální konfigurace neuronové sítě jsem zkoušel různé počty neuronů, více skrytých vrstev a různé počty trénovacích cyklů. Hlavním kritériem podle, které jsem danou konfiguraci sítě hodnotil, je její chyba klasifikace při učení a přesnost při predikci na netrénovaných datech. V následujících grafech jsou vidět úspěšnosti při učení různých nastavení počtu neuronů, počtu skrytých vrstev a počtu trénovacích cyklů v neuronové síti.



Graf chyby klasifikace na učených datech.

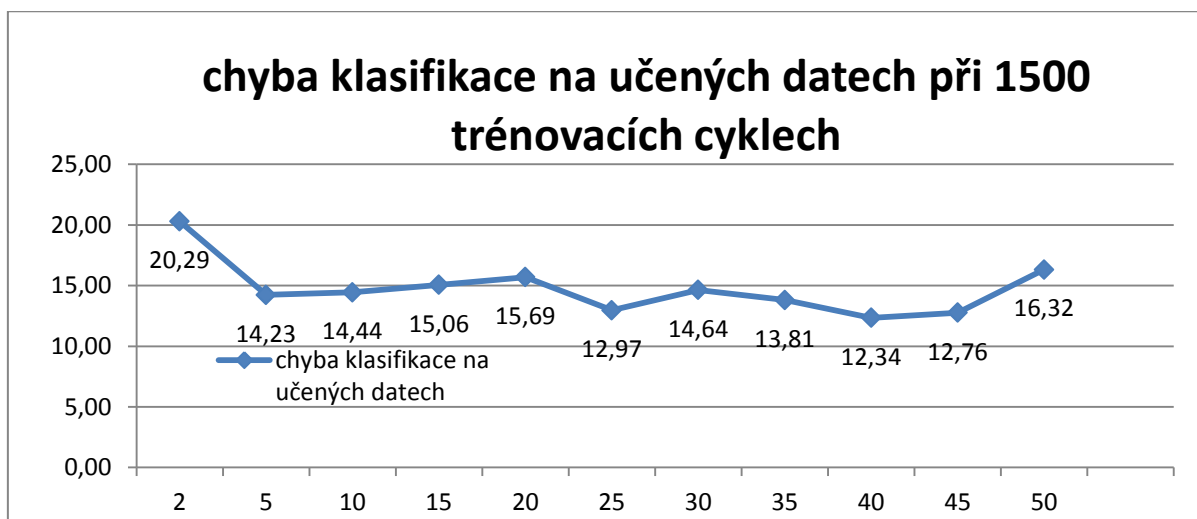
Na grafu výše vidíme velikost chyby klasifikace procesu učení na vybraných datech. Z grafu vyplývá, že při pěti a čtyřiceti neuronech ve skryté vrstvě je klasifikační chyba nejmenší. Naopak vidíme, že dva neurony ve skryté vrstvě mají největší chybu klasifikace ze všech ostatních počtů neuronů prezentovaných na grafu. Při počtech neuronů nad padesát byli výsledky obdobné a nepodařilo se mi již naleznout menší velikost chyby kvalifikace než u dvou či čtyřiceti neuronů.



Graf chyby klasifikace na učených datech.

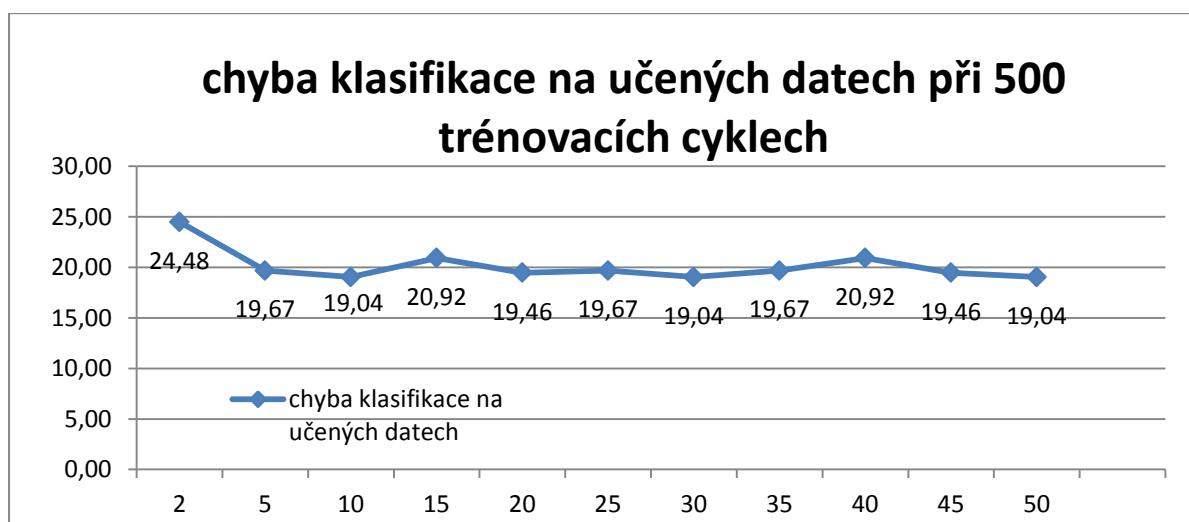
Při zvýšení počtu trénovacích cyklů na 1000. Na grafu vidíme pokles velikosti chyby u všech počtů neuronů. Z grafu lze také vyčíst, že pět neuronů ve skryté vrstvě již zdaleka není při tomto počtu trénovacích cyklů jednou z nejlepších variant. Nejlepší variantou počtu

neuronů nám, ale pořád zůstává čtyřicet neuronů ve skryté vrstvě. Nejhorší variantou počtu neuronů ve skryté vrstvě jsou dva s chybou klasifikace blíží se 21%.



Graf chyby klasifikace na učených datech.

Při zvýšení trénovacích cyklů na 1500. Vidíme v grafu snížení klasifikační chyby na trénovaných datech. Nejlepším počtem neuronů ve skryté vrstvě zůstává čtyřicet, protože i při 1500 trénovacích cyklech má nejmenší klasifikační chybu. Naopak nejhorším stále zůstává počet dvou neuronů ve skryté vrstvě, který má klasifikační chybu 20,29%.



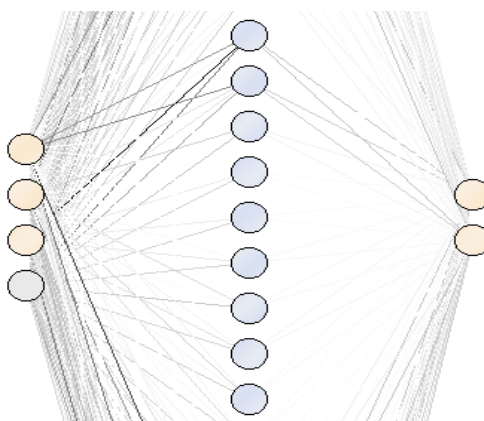
Graf chyby klasifikace na učených datech.

Posledním pokusem je přidání ještě jedné skryté vrstvy se stejným počtem neuronů a nastavení 500 trénovacích cyklů. Z grafu je jasně zřejmé, že chyby klasifikace jsou oproti předchozím pokusům o dost větší. Nejmenší chyby v klasifikaci dosahujeme při deseti a

padesáti neuronech. Naopak největší stále při malém počtu dvou neuronů. Je tedy zřetelné, že přidání skryté vrstvy v tomto případě není dobrým nápadem.

Všechny předchozí nastavení dosáhli stejného výsledku predikce, nicméně se lišili v přesnosti (confidence up a confidence down) což by u jiných dat mohlo způsobit rozdíly v predikovaných výsledcích. O nastavení počtu trénovacích cyklů lze polemizovat. Samozřejmě lze nastavit i větší počet trénovacích cyklů než bylo vyzkoušeno, ale při vyšším počtu lze také dosáhnout přetrénování sítě a tím značného zhoršení predikce trendu sítě. Proto bych volil spíše hodnotu počtu trénovacích cyklů na hodnotu 500, která je i standardně v RapidMineru nastavena.

Optimální nastavení



Obrázek schéma neuronů v síti.

Při pěti stech trénovacích cyklech jak jsme viděli na grafu v předchozí kapitole, vyšlo nejlépe nastavení s jednou skrytou vrstvou se čtyřiceti neurony. Na obrázku schéma neuronů v síti vidíme tedy, že naše síť má čtyři neurony ve vstupní vrstvě. Pak zde vidíme část z celkového počtu čtyřiceti neuronů ve skryté vrstvě. Ve výstupní vrstvě se nacházejí dva neurony.

Vyhodnocení experimentu

Optimální nastavení pro predikci a učení je zmíněno výše. V první tabulce níže vidíme velikost klasifikační chyby při učení. Na následující tabulce vidíme, jak dopadla predikce a v třetí tabulce je úspěšnost predikce. Při experimentování jsem došel k závěru, že síť neumí dostatečně rychle reagovat při predikci na krátkodobé změny v trendu. Dlouhodobě však dokáže pohyb trendu dobře odhadnout. Přesnost predikce je také ovlivněna nastavením vstupní hodnoty výpočtu indikátoru. V tomto experimentu proces používá pro definování

trendu exponenciální klouzavý průměr. Pro určení trendu tedy hodnot up nebo down by mohl být samozřejmě vybrán i jiný implementovaný indikátor technické analýzy například jednoduchý klouzavý průměr by mohl být také vhodný.

Classification error: 15,69%	True up	True down	Class precission
Prediction up	163	12	93.14%
Prediction down	63	240	79,21%
Class recall	72,12%	95,24%	

Row No.	datum	label	confidence(...	confidence(...	prediction(l...	roc	rsi	sma
1	18.4.2012	up	0.979	0.021	up	-3.457	46.160	31.008
2	17.4.2012	up	0.972	0.028	up	-2.117	49.743	31.088
3	16.4.2012	down	0.985	0.015	up	-3.464	44.912	31.174
4	13.4.2012	up	0.988	0.012	up	-5.258	41.093	31.292
5	12.4.2012	up	0.988	0.012	up	-4.956	42.809	31.423
6	11.4.2012	up	0.997	0.003	up	-5.201	32.874	31.544
7	10.4.2012	up	0.997	0.003	up	-4.766	33.961	31.762
8	9.4.2012	up	0.997	0.003	up	-2.538	40.178	31.974
9	5.4.2012	up	0.995	0.005	up	-1.469	45.313	32.064
10	4.4.2012	up	0.996	0.004	up	-3.075	40.063	32.112
11	3.4.2012	up	0.989	0.011	up	-2.025	50.708	32.182
12	2.4.2012	up	0.953	0.047	up	-1.705	57.511	32.187
13	30.3.2012	down	0.959	0.041	up	-1.572	56.975	32.178
14	29.3.2012	down	0.979	0.021	up	-1.684	54.939	32.212

Obrázek výsledku predikce.

accuracy: 78,57%	True up	True down	Class precission
Prediction up	11	3	78,57%
Prediction down	0	0	0,00%
Class recall	100,00%	0%	

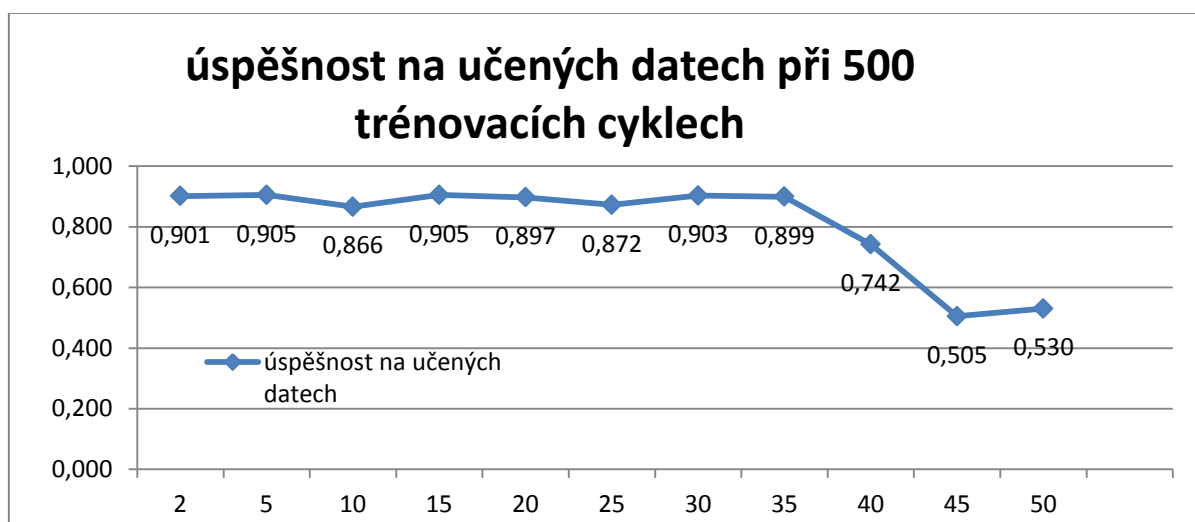
4.1.2 Predikce hodnoty indikátoru

Postup experimentu

Při zpracování hodnot ze zdrojového souboru s daty, který je pro oba experimenty stejný byli, nastaveny vstupní hodnoty pro zpracování následovně. Vstupní hodnota Relative strength index byla nastavena na čtrnáct. Vstupní hodnota Rate of change byla nastavena na dvanáct. Hodnoty exponenciálního a jednoduchého klouzavého průměru byli nastaveny na deset. Poté získaný csv soubor s hodnotami indikátorů byl rozdělen do dvou csv souborů. První soubor data2Tren.csv je určeným pro proces učení a druhý soubor data2Pred.csv je určen pro proces predikování hodnot. V operátoru Windowing v procesu pro učení sítě byli nastaveny hodnoty následovně. Hodnota horizon byla nastavena na hodnotu jedna. Window size na velikost jedna a hodnota stepsize také na hodnotu jedna. Dále byla zaškrtnuta položka create label, která nám vlastně vyjadřuje pro jaký sloupec (indikátor) budeme windowing používat. Po předložení souborů procesům pro učení a predikci poté získáme výstupy.

Experimenty s nastavením sítě

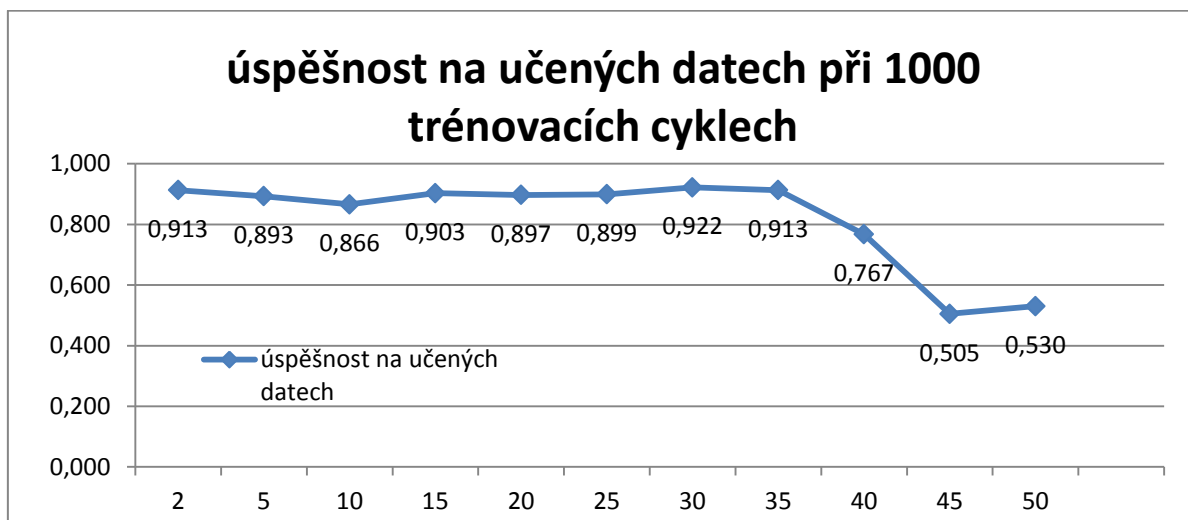
Při nastavení sítě jsem zkoušel různé počty neuronů, trénovacích cyklů a skrytých vrstev. V následujících grafech uvidíte úspěšnosti těchto nastavení při učení neuronové sítě. Grafy mají na ose x počet neuronů a křivka znázorňuje onu úspěšnost naučení.



Graf úspěšnosti učení.

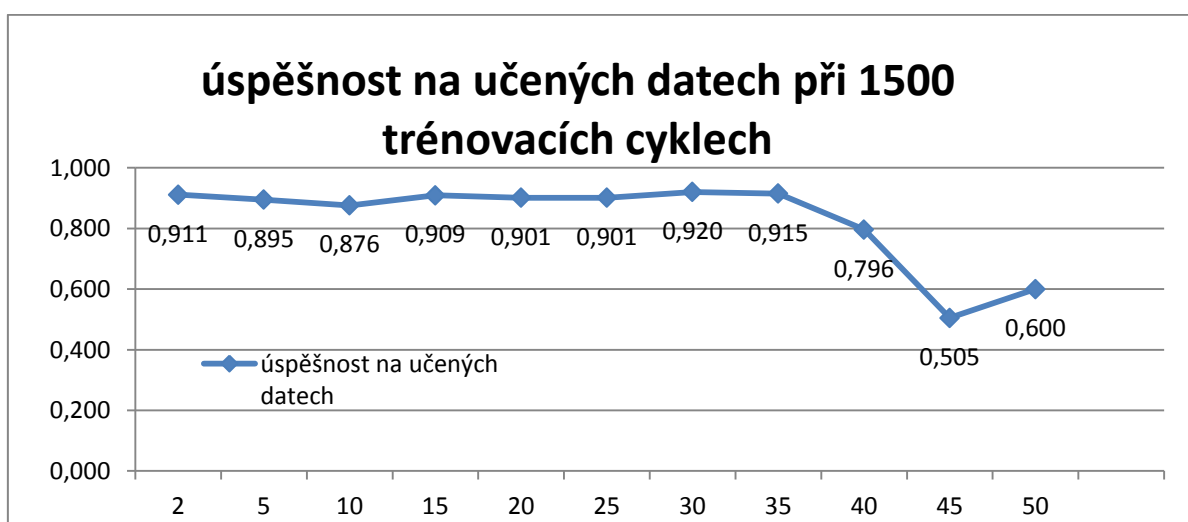
Na tomto grafu vidíme úspěšnost učení neuronové sítě při pěti stech trénovacích cyklech a jedné skryté vrstvě. V grafu vidíme, že do počtu 35 neuronů ve skryté vrstvě se úspěšnost

naučení sítě příliš nemění a rozdíly se pohybují v tisícinách či setinách. Poté úspěšnost učení výrazně klesá a pak stagnuje a takto to pokračuje i při větším počtu neuronů, proto nemělo cenu uvádět i vyšší počty neuronů. Nejlepšího výsledku jak je z grafu patrné jsme tedy dosáhli při pěti neuronech ve skryté vrstvě.



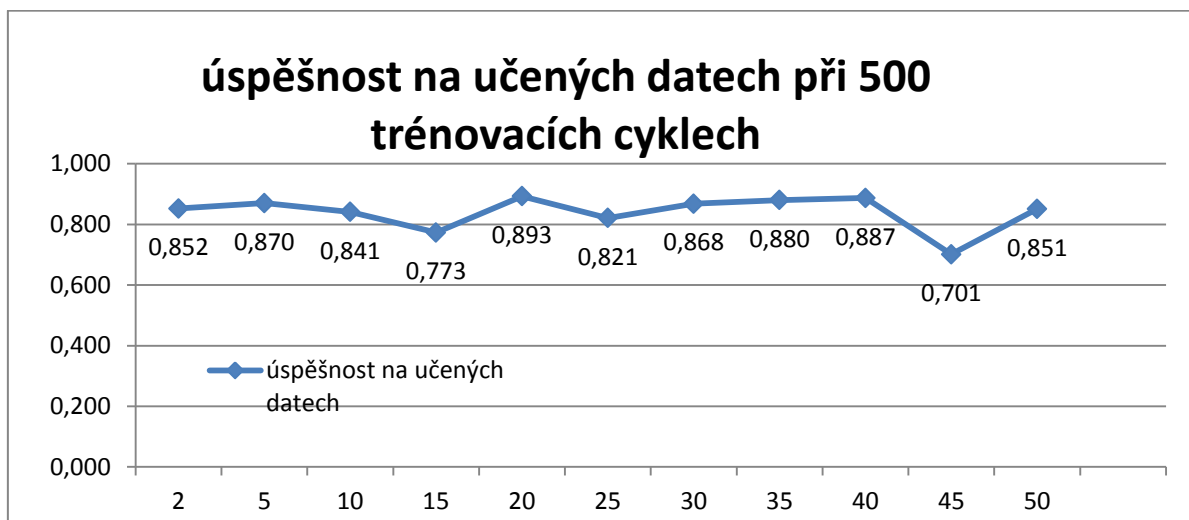
Graf úspěšnosti učení.

Zvýšení počtu trénovacích cyklů na tisíc při zachování počtu skrytých vrstev nepřineslo až tak dramatické změny. Došlo k minimálním pohybům velikosti chyby jak nahoru tak dolu u jednotlivých počtů neuronů. Při tomto nastavení opět vidíme, že od počtu 35 neuronů se úspěšnost učení zhoršila. Naopak nejlepšího výsledku na učících datech dosáhlo 30 neuronů ve skryté vrstvě.



Graf úspěšnosti učení.

Při počtu 1500 trénovacích cyklů se opět oproti předchozímu nastavení nic zásadního nestalo. Nejlepší počet neuronů ve skryté vrstvě zůstal 30 a opět vidíme pokles úspěšnosti predikce přes 35 neuronů ve skryté vrstvě. Oproti minulému procesu však většina hodnot úspěšnosti učení vzrostla.

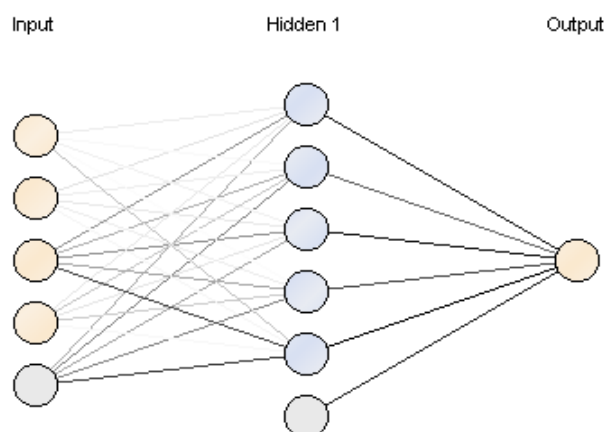


Graf úspěšnosti učení.

Při nastavení dvou skrytých vrstev se stejným počtem neuronů a pěti stech trénovacích cyklů. Vidíme, že ani jedna hodnota úspěšnosti učení u počtu neuronů nepřesáhla hodnotu 0,9. Zvolení dvou skrytých vrstev o stejném počtu neuronů se ukázalo být nejhorším možným, jak lze vidět na grafu výše.

Jak vidíme na grafech, rozdíly úspěšnosti učení na trénovacích datech při různém počtu trénovacích cyklů, nebyly nijak dramaticky rozdílné. To se odrazilo i ve výsledcích predikce, kde byly dosaženy stejné výsledky. U počtu neuronů 40, 45 a 50 tyto výsledky predikce nebyli na neučných datech nijak oslnivé, ale to se dalo předpokládat již z úspěšnosti učení na trénovací množině.

Optimální nastavení sítě



Obrázek schéma neuronů v síti.

Ideální nalezené nastavení sítě má jednu skrytou vrstvu a v ní 6 neuronů. Počet trénovacích cyklů byl nastaven na 500. Jak již bylo dokázáno v předchozích experimentech rozdíly v úspěšnosti učení a predikce s různým počtem trénovacích cyklů nebyly nijak veliké, proto jsem se rozhodl pro spíš menší počet trénovacích cyklů a to jednak z hlediska výpočetní náročnosti tak již zmiňovaných úspěšností na trénovací množině. V RapidMineru byla teda nastavena tedy hodnota neuronů na 5. Šestý neuron, jak vidíme na schématu neuronů v síti, je tzv. threshold node neboli prahový uzel. Dále ze schématu je také zřejmé, že máme pět neuronů ve vstupní a jeden neuron ve výstupní vrstvě.

Vyhodnocení experimentu

Při experimentování se nikdy nepodařilo přesně určit budoucí hodnotu indikátoru, ale ukázalo se, že lze pomocí tohoto experimentu dobře předvídat budoucí pohyb hodnoty indikátoru. Na obrázku vlevo vidíme úspěšnost predikce trendu neuronové sítě na trénovací množině. Obrázek vpravo nám ukazuje úspěšnost predikce trendu na neučených datech. Na grafu pak vidíme modře body skutečného exponenciálního klouzavého průměru a červené jsou body predikované neuronovou sítí. Vidíme, že síť predikovala vyšší hodnoty, než skutečně byli. Směr pohybu hodnoty indikátoru zejména u prvních třech dat, byl predikován přesně. Pak je vidět, že nastala změna pohybu indikátoru, na kterou síť nereagovala při predikci zcela správně.

prediction_trend_accuracy

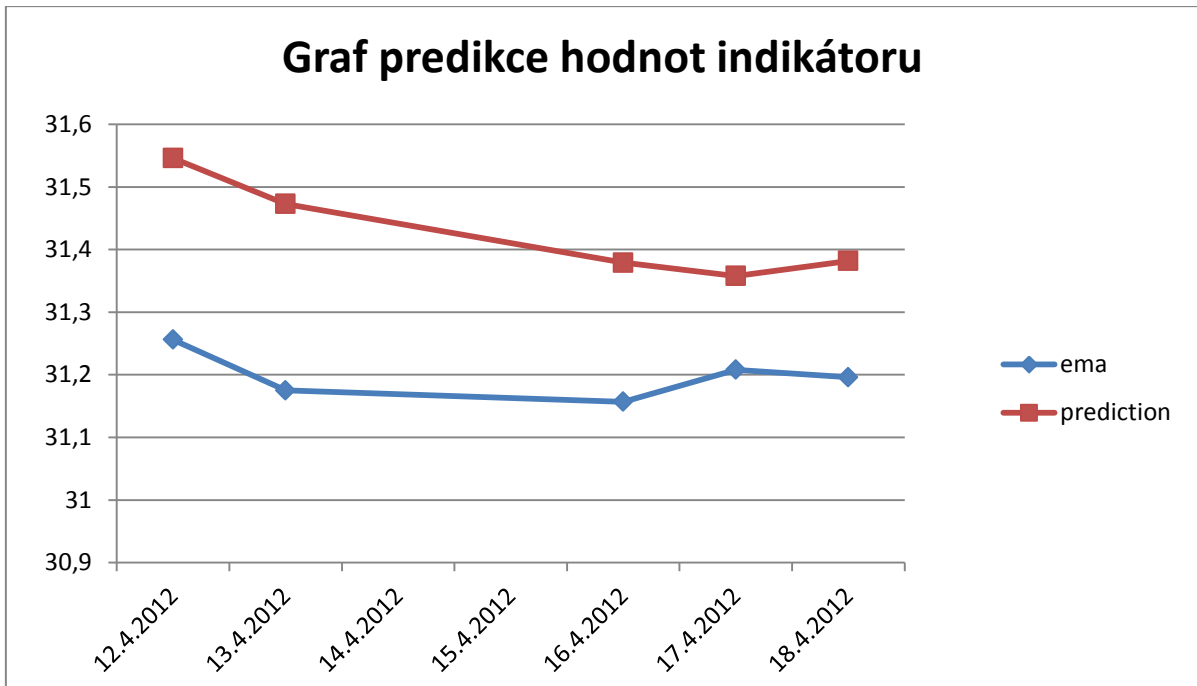
prediction_trend_accuracy: 0.905

Obrázek úspěšnosti učícího procesu.

prediction_trend_accuracy

prediction_trend_accuracy: 0.750

Obrázek úspěšnost procesu pro predikci.



Graf výsledků predikce.

4.2 Závěr

Cíle práce

Prvním cílem bylo charakterizovat technickou analýzu burzovních dat její základní pojmy a metody s důrazem na metody umělé inteligence, zejména neuronové sítě. Tento cíl byl splněn. V práci, je popsána technická analýza i její indikátory, které jsou přehledně rozděleny do jednotlivých skupin.

Dalším cílem bylo zvolení postupů výběru a předzpracování burzovních dat pro technickou analýzu s užitím neuronových sítí, vytvoření aplikace v Javě pro získání a přípravu těchto dat. I tento cíl lze považovat za splněný. Byla vytvořena aplikace v programovacím jazyce Java, ve které jsou implementovány čtyři indikátory technické analýzy. Aplikace je tak schopna zpracovat burzovní data.

Třetím cílem bylo použít technickou analýzu burzovních dat s využitím neuronových sítí a důrazem na získání výstupů prakticky využitelných v obchodování s akciemi a implementace neuronové sítě. V práci jsou prezentovány celkem dva experimenty s neuronovými sítě, z kterých se podařilo získat odhad budoucího trendu obchodování na burze a predikce hodnoty exponenciálního průměru. Oba dva experimenty vycházejí z dat indikátorů technické analýzy. Tím byl splněn cíl využití neuronových sítí a indikátorů technické analýzy k získání využitelných výsledků v obchodování s akciemi.

Posledním cílem je provést Analýzu a vyhodnocení získaných výsledků, závěry a doporučení. V práci se povedlo pomocí indikátorů technické analýzy a neuronových sítí získat data použitelná na akciových trzích. V prvním experimentu, který se zabýval predikcí trendu akcie, se povedlo s přesností 78,57 % predikovat budoucí trend. Síť sice nedovedla dostatečně přesně reagovat na krátkodobé změny v trendu, ale dlouhodobější trend dokázala predikovat správně. Druhý experiment se pokoušel odhadnout budoucí hodnotu exponenciálního klouzavého průměru. To se bohužel nepovedlo, ale lze s určitou přesností pomocí tohoto experimentu predikovat budoucí směr pohybu hodnoty indikátoru. I tento poslední cíl zadání bakalářské práce tak lze považovat za splněný.

5. Literatura

[1] VESELÁ, Jitka. *Investování na kapitálových trzích*. 2., aktualiz. vyd. Praha: Wolters Kluwer Česká republika, 2011, 789 s. ISBN 978-80-7357-647-9 (VÁZ.).

[2] STOCKCHARTS.COM. *Relative Strength Index (RSI)* [online]. 2012-04-02 [cit. 2012-04-02].

Dostupné z:

http://stockcharts.com/school/doku.php?id=chart_school:technical_indicators:relative_strength_index_

[3] MAREK, Josef. *TECHNICKÁ ANALÝZA VYBRANÝCH AKCIOVÝCH TITULŮ NA ČESKÉM A ZAHRANIČNÍM KAPITÁLOVÉM TRHU*. Brno, 2011. Bakalářská. MASARYKOVA UNIVERZITA PŘÍRODOVĚDECKÁ FAKULTA. Vedoucí práce Ing. Petr Málek.

[4] STOCKCHARTS.COM. *Williams %R* [online]. 2012-04-02 [cit. 2012-04-02]. Dostupné z:

http://stockcharts.com/school/doku.php?id=chart_school:technical_indicators:williams_r

[5] HÖFEROVÁ, Lucie. *Statistika a spekulace na burze*. Olomouc, 2008. Bakalářská. UNIVERZITA PALACKÉHO V OLOMOUCI PŘÍRODOVĚDECKÁ FAKULTA. Vedoucí práce Mgr. Jaroslav Marek, Ph.D.

[6] KLÍMA, David. Neuronové sítě typu RBF [online]. 2002 [cit. 2012-04-02]. Dostupné z:

<http://www.volny.cz/klimad/skola/spa/skrbek%20web/rbf/rbf.html>

[7] ŠNOREK, Miroslav. *Neuronové sítě a neuropočítače*. Vyd. 1. Praha: ČVUT, 1996, 124 s. ISBN 80-010-1455-X.

[7] Rapid-i [online]. 21.1.2012 [cit. 2012-03-02]. Integrating RapidMiner into your

application. Dostupné z WWW: <http://rapidi.com/wiki/index.php?title=Integrating_RapidMiner_into_your_application>

6. Přílohy

[A] CD s vlastní aplikací a jejími zdrojovými kódy, s dokumentací k vytvořené Java aplikaci a procesy v RapidMineru.

[B] CD s textem bakalářské práce v elektronické podobě a vlastní java aplikací včetně dokumentací.