

Využití technik umělé inteligence na finančních trzích

Usage of Artificial Intelligence Techniques in the Financial Markets

Bc. Jana Raková



Univerzita Tomáše Bati ve Zlíně

Fakulta aplikované informatiky

akademický rok: 2012/2013

ZADÁNÍ DIPLOMOVÉ PRÁCE

(PROJEKTU, UMĚLECKÉHO DÍLA, UMĚLECKÉHO VÝKONU)

Jméno a příjmení: **Bc. Jana Raková**

Osobní číslo: **A11499**

Studijní program: **N3902 Inženýrská informatika**

Studijní obor: **Informační technologie**

Forma studia: **kombinovaná**

Téma práce: **Využití technik umělé inteligence na finančních trzích**

Zásady pro vypracování:

1. Popište finanční trhy se zaměřením na trh futures.
2. Popište postupy technické analýzy.
3. Zpracujte popis obchodních indikátorů.
4. Využijte umělé neuronové sítě, zvolte vhodnou trénovací a testovací množinu pro vybraný typ vzoru cenového grafu.
5. Posudte vhodnost využití umělých neuronových sítí pro predikci trendu vývoje cen.

Rozsah diplomové práce:

Rozsah příloh:

Forma zpracování diplomové práce: **tištěná/elektronická**

Seznam odborné literatury:

1. ZELINKA, Ivan. Evoluční výpočetní techniky: principy a aplikace. 1. vyd. Praha: BEN – technická literatura, 2009, 534 s. ISBN 978-80-7300-218-3.
2. FANTA, Jiří. Technologie umělé inteligence na kapitálových trzích. Praha: Karolinum, 1999. 89 s. ISBN 80-7184-866-2.
3. FANTA, Jiří. Technická analýza kapitálových trhů. 1. vyd. Praha: Karolinum, 1996. 102 s. ISBN 80-7184-308-3.
4. MURPHY, J. Technical Analysis of the Financial Markets: A Comprehensive Guide to Trading Methods and Applications. New York Institute of Finance. 1999. 576 s. ISBN 978-0735200661.
5. ZELINKA, Ivan. Umělá inteligence I: Neuronové sítě a genetické algoritmy. 1. vyd. Brno: VUT v Brně. 1998. 126 s. ISBN 80-214-1163-5.
6. BOSE, N a Ping LIANG. Neural network fundamentals with graphs, algorithms, and applications. New York: McGraw-Hill. 1996. 478 s. ISBN 00-700-6618-3.
7. NOVÁK, Mirko, Josef FABER a Olga KUFUDAKI. Neuronové sítě a informační systémy živých organismů. Praha. Grada. 1993. 265 s. ISBN 80-58424-95-9.
8. BÍLA, Jiří. Umělá inteligence a neuronové sítě v aplikacích. ČVUT. 1996. ISBN 80-01-01275-1.

Vedoucí diplomové práce:

doc. Ing. Zuzana Komínková Oplatková, Ph.D.

Ústav informatiky a umělé inteligence

Datum zadání diplomové práce:

26. července 2013

Termín odevzdání diplomové práce:

3. září 2013

Ve Zlíně dne 16. července 2013

prof. Ing. Vladimír Vašek, CSc.
děkan



doc. Mgr. Roman Jašek, Ph.D.
ředitel ústavu

ABSTRAKT

Tato práce se zabývá využitím neuronové sítě na finančních trzích. Neuronová síť feedforward byla naučena a použita k rozpoznávání vzoru v časovém grafu. Pro diplomovou práci byl vybrán vzor, který se nazývá DoubleTop (DT). Pokud je detekován tento vzor, obvykle dochází ke změně trendu, která může poskytnout dobrou obchodní příležitost. Z vybraného vzoru DT bylo vytvořeno několik trénovacích množin, pomocí nichž byla síť naučena. Testování sítě bylo provedeno na předložených datech z reálného finančního trhu a vyzkoušeno vyhledávání vzoru DT.

Klíčová slova: finanční trh, neuronová síť, trénovací množina, testovací množina, DoubleTop

ABSTRACT

This work deals with the neural network exploitation in financial markets. The feedforward neural network was taught and used for the model detection in a time diagram. For this thesis the model was chosen which is named DoubleTop (DT). In case this model is detected it usually comes to a trend change by which a good business opportunity can be provided. Out of the chosen DT model several training sets were created by the help of which the network was taught. The network testing was provided on the submitted data from the real financial market and searching of the DT model was tested.

Keywords:

financial market, neural network, training set, testing set, DoubleTop

Poděkování

Děkuji paní doc. Ing. Zuzaně Komínkové Oplatkové, Ph.D. vedoucí mé práce za cenné rady a připomínky a také za čas věnovaný mé práci.

Dále děkuji Ing. Kateřině Kollarikové za její podporu.

Prohlašuji, že

- beru na vědomí, že odevzdáním diplomové/bakalářské práce souhlasím se zveřejněním své práce podle zákona č. 111/1998 Sb. o vysokých školách a o změně a doplnění dalších zákonů (zákon o vysokých školách), ve znění pozdějších právních předpisů, bez ohledu na výsledek obhajoby;
- beru na vědomí, že diplomová/bakalářská práce bude uložena v elektronické podobě v univerzitním informačním systému dostupná k prezenčnímu nahlédnutí, že jeden výtisk diplomové/bakalářské práce bude uložen v příruční knihovně Fakulty aplikované informatiky Univerzity Tomáše Bati ve Zlíně a jeden výtisk bude uložen u vedoucího práce;
- byl/a jsem seznámen/a s tím, že na moji diplomovou/bakalářskou práci se plně vztahuje zákon č. 121/2000 Sb. o právu autorském, o právech souvisejících s právem autorským a o změně některých zákonů (autorský zákon) ve znění pozdějších právních předpisů, zejm. § 35 odst. 3;
- beru na vědomí, že podle § 60 odst. 1 autorského zákona má UTB ve Zlíně právo na uzavření licenční smlouvy o užití školního díla v rozsahu § 12 odst. 4 autorského zákona;
- beru na vědomí, že podle § 60 odst. 2 a 3 autorského zákona mohu užít své dílo – diplomovou/bakalářskou práci nebo poskytnout licenci k jejímu využití jen s předchozím písemným souhlasem Univerzity Tomáše Bati ve Zlíně, která je oprávněna v takovém případě ode mne požadovat přiměřený příspěvek na úhradu nákladů, které byly Univerzitou Tomáše Bati ve Zlíně na vytvoření díla vynaloženy (až do jejich skutečné výše);
- beru na vědomí, že pokud bylo k vypracování diplomové/bakalářské práce využito softwaru poskytnutého Univerzitou Tomáše Bati ve Zlíně nebo jinými subjekty pouze ke studijním a výzkumným účelům (tedy pouze k nekomerčnímu využití), nelze výsledky diplomové/bakalářské práce využít ke komerčním účelům;
- beru na vědomí, že pokud je výstupem diplomové/bakalářské práce jakýkoliv softwarový produkt, považují se za součást práce rovněž i zdrojové kódy, popř. soubory, ze kterých se projekt skládá. Neodevzdání této součásti může být důvodem k neobhájení práce.

Prohlašuji,

- že jsem na diplomové práci pracoval samostatně a použitou literaturu jsem citoval. V případě publikace výsledků budu uveden jako spoluautor.
- že odevzdaná verze diplomové práce a verze elektronická nahraná do IS/STAG jsou totožné.

Ve Zlíně 3. září 2013

.....
podpis diplomanta

OBSAH

ÚVOD.....	8
I TEORETICKÁ ČÁST.....	9
1 FINANČNÍ TRHY	10
1.1 STRUKTURA FINANČNÍHO TRHU	10
1.1.1 Primární trh	10
1.1.2 Sekundární trh	10
1.1.3 Burza	11
1.2 KOMODITY	11
1.2.1 Historie komoditního obchodování	11
1.3 OBCHODOVÁNÍ.....	12
1.3.1 Fundamentální analýza.....	12
1.3.2 Technická analýza	12
2 INDIKÁTORY.....	18
2.1 INDIKÁTORY	18
2.1.1 Trendové indikátory	18
2.1.2 Oscilátory	20
2.2 VZORY - PATTERNY	22
3 UMĚLÁ INTELIGENCE A FINANČNÍ TRHY	24
4 NEURONOVÉ SÍTĚ.....	25
4.1 NEURONOVÉ SÍTĚ BEZ UČITELE	27
4.2 NEURONOVÉ SÍTĚ S UČITELEM.....	28
4.2.1 Vícevrstvé neuronové sítě	28
4.2.2 Backpropagation	29
4.2.3 Trénovací množina.....	30
II PRAKTICKÁ ČÁST	32
5 ÚVODEM.....	33
5.1 POUŽITÝ SOFTWARE	33
5.2 VSTUPY	33
5.2.1 Testovací množina	34
5.2.2 Trénovací množina.....	35
5.3 TRÉNOVÁNÍ SÍTĚ	38
5.3.1 Síť feedforward s algoritmem Backpropagation.....	40
5.3.2 Síť feedforward s algoritmem Levenberg-Marquardt.....	42
ZÁVĚR	50
CONCLUSION	51
SEZNAM POUŽITÉ LITERATURY.....	52
SEZNAM POUŽITÝCH SYMBOLŮ A ZKRATEK.....	54
SEZNAM OBRÁZKŮ	55
SEZNAM TABULEK.....	57
SEZNAM PŘÍLOH.....	58

ÚVOD

Velký rozvoj umělé inteligence začal v minulém století. Britský matematik Alan Turing sestavil v roce 1943 počítač Colossus, dešifrující německé zprávy za druhé světové války, které byly zašifrovány velmi důmyslným způsobem. Další cesta umělé inteligence vedla k pokusu napodobit činnost lidského mozku – vzniku modelu biologických neuronů a odtud neuronových sítí. V roce 1943 McCulloch a Pitts vytvořili model umělého neuronu, který se používá dodnes. Na základě jejich práce F. Rosenblatt sestavil v roce 1958 první neuronovou síť – Perceptron. Tato síť byla schopna řešit jenom lineárně separabilní problémy a to vedlo k poklesu zájmu o umělou inteligenci. V polovině 80. let vznikla práce Learning Internal Representation by Error Propagation od autorů D. Rumelharta, G. Hinton a R. Wiliamse o vícevrstvých sítích, schopných řešit lineárně neseperabilní problémy. Byly vytvořeny další typy sítí – Hopfieldova, Kohonenova, ART. [1]

V současnosti se rozsah využití neuronových sítí stále více rozšiřuje.

Tato práce pojednává o propojení umělé inteligence s finančními trhy. Na světových finančních trzích se pro obchodování stále více používá umělá inteligence.

V teoretické části bude věnována pozornost finančním trhům, jejich historii, rozdělení, způsobu obchodování. Bude popsána fundamentální a technická analýza. Technická analýza zkoumá dynamiku trhu pomocí statistiky, grafů a indikátorů za účelem určení trendu vývoje cen. Předpokládá existenci grafických formací, které se opakují a signalizují změnu dynamiky. Studie grafických formací bude použita v praktické části této práce.

V praktické části bude naučena neuronová síť na jednu grafickou formaci technické analýzy s cílem vyzkoušet schopnost sítě vyhledat tuto formaci v časovém grafu reálného trhu.

I. TEORETICKÁ ČÁST

1 FINANČNÍ TRHY

Finanční trh je souhrn investičních instrumentů, institucí a postupů, kde dochází k přesunu volných finančních prostředků z přebytkové strany na stranu deficitní. Přebytkovou stranu tvoří subjekty, které mají v dané chvíli volné finanční prostředky, pro něž nemají okamžité uplatnění. Mohou to být firmy, banky, fondy. Deficitní stranu tvoří subjekty, které mají nedostatek finančních prostředků pro svoje záměry a investice, jako firmy, státy, banky. To zajišťuje efektivní využití všech volných finančních prostředků.

Investiční instrumenty jsou aktiva, která přinášejí nárok na nějaký budoucí příjem, jako jsou dividendy, úroky a jiné. [9]

Investiční instrumenty tvoří dvě skupiny:

1. Finanční instrumenty - mezi finanční instrumenty patří různé druhy cenných papírů, finanční deriváty, pojišťovací kontrakty.
2. Reálné instrumenty (aktiva) - jsou to různé formy investic do drahých kovů, uměleckých sbírek, drahokamů.

Finanční trh ukazuje stav ekonomiky, její dosavadní vývoj a budoucí perspektivy. Role finančního trhu je ve fungující ekonomice nezastupitelná. [9]

1.1 Struktura finančního trhu

1.1.1 Primární trh

Subjekty, které potřebují finanční prostředky, emitují cenné papíry za účelem jejich prodeje. Na tomto typu trhu se prodávají právě emitované cenné papíry, buďto přímo samotným emitentem nebo investiční bankou. Primární trh může být veřejný, cenný papír je nabízen široké investorské veřejnosti nebo neveřejný, kde jsou cenné papíry nabízeny úzkému kruhu investorů, kteří znají ekonomickou situaci emitenta a drží je obvykle po celou dobu životnosti. [9]

1.1.2 Sekundární trh

Zde se obchoduje s cennými papíry, které byly alespoň jednou prodány. Emitent již z prodeje nezískává finanční prostředky. Likviditu cennému papíru zajišťuje burza.

Sekundární trh se dělí na dva trhy:

1. Veřejný sekundární trh - součástí veřejného sekundárního trhu je burzovní trh.
2. Neveřejný sekundární trh - má podobu individuálně dohodnutých obchodů přímo mezi zájemci a prodávajícími.

Dále existuje i mimoburzovní trh. Podmínky obchodování na něm nejsou tak přísně regulovány jako na burze. Příkladem mimoburzovního trhu je například Forex, kde se obchoduje s měnovými páry. [9]

1.1.3 Burza

Na burze se obchoduje s nejbonitnějšími akciemi, dluhopisy, podílovými listy nebo komoditními kontrakty. Burza má přesně vymezená pravidla a předpisy, vymezený způsob obchodování. Je také přísně regulována a má dohled příslušných orgánů dané země. V USA se o dohled stará komise CFTC *Commodity Futures Trading Commission*, která dohlíží na burzy, brokerské společnosti a všechny, kteří jsou spojeni s komoditním obchodováním. Obchodovat na burze mohou přímí členové burzy nebo zprostředkovatelé - brokeri. Ti převezmou příkaz k nákupu nebo prodeji a za poplatek zajistí vyplnění příkazu. Největší americké burzy jsou NYBOT, CBOT. V Evropě je to EUREX, LIFFE a další. [9]

1.2 Komodity

Komodity jsou názvem pro suroviny, jako je například ropa, bavlna, maso a pšenice. S komoditami se dále obchoduje. Toto obchodování se anglicky nazývá *futures trading*.

1.2.1 Historie komoditního obchodování

Komoditní obchodování sahá podle historických pramenů až do staré Číny, asi před 6000 lety. První doložené komoditní obchody jsou z Japonska, někdy v 17. století. Model současného obchodování je velmi podobný tomu dřívějšímu.

Zemědělci, kteří pěstovali rýži, potřebovali určitou finanční částku na osivo, péči o rostliny a také na sklizeň ještě před tím, než získali finance prodejem své úrody. Na druhé straně existoval obchodník, který věděl, že tuto rýži bude potřebovat. Vzájemně se domluvili na „předprodeji“ rýže za předem dohodnutou cenu, která bude dodána v konkrétním termínu. Tato dohoda byla zajištěna smlouvou, které se dnes říká *futures kontrakt*. Na základě smlouvy navštívil zemědělec banku, aby získal půjčku. Později, někoho napadlo, že se dá

obchodovat se samotnými *futures kontrakty*. Nakonec se ukázalo, že *futures kontrakty* jsou skvělým způsobem pro rychlé zbohatnutí.

V současné době se obchoduje se všemi nerostnými surovinami a plodinami na komoditních burzách. Zboží se nedodává ve stanovený den, ale podstatou obchodování je zisk daný rozdílem mezi nákupní a prodejní cenou. [9]

1.3 Obchodování

Jak bylo uvedeno výše, při obchodování je důležitý rozdíl mezi nákupní a prodejní cenou. Tím je realizován zisk. Jak zjistit, kterým směrem se bude cena posouvat?

1.3.1 Fundamentální analýza

Existuje celá řada vlivů, které působí na změnu ceny aktiv. Mohou to být politické vlivy, hospodářské, vlivy počasí a další. Fundamentální analýza předvídá změnu ceny aktiva na základě těchto událostí. Politické vlivy mohou být globální a vnitrostátní. Vlivy počasí a z toho plynoucí sezónnost mají za následek pohyb na komoditních trzích (zemědělské produkty).

Nevýhody fundamentální analýzy:

1. Není schopna zachytit jednotlivé negativní ekonomické informace, zejména ty, které jsou zpožděné, např. účetní výkazy.
2. Neumožňuje odhadnout pokles trhu na základě dezinformací a poplašných zpráv. Tyto zprávy jsou často protikladné. [3]

1.3.2 Technická analýza

Technická analýza se věnuje sběru, sledování, měření a vyhodnocování:

- ceny
- objemu
- změnou těchto parametrů v čase

Technická analýza se nezajímá o fundamentální faktory. Techničtí analytici jsou přesvědčeni, že na pohyb aktiva mají vliv i faktory psychologické a racionálně nepředvídatelné.

Základem technické analýzy je sledování finančních trhů po velmi dlouhou dobu, desítky až stovky let. Změny na trzích reflektují poměr nabídky a poptávky. Když převyšuje nabídka - ceny klesají a když převyšuje poptávka - ceny rostou. Technická analýza má za cíl odhadnout, kam se bude cena aktiva v budoucnosti posouvat a nesleduje, proč cena roste nebo klesá. Základem je studium cenových grafů - *charting*. Obchodníci své chování opakují a tak se v cenových grafech vytvářejí vzory - *patterny*, které jsou rozpoznatelné a které s určitou pravděpodobností předchází změně trendu. Pokud je predikce správná, obchodník bude vydělávat a naopak. [3]

Existují tři podmínky, na kterých je technický přístup založen:

1. Všechny faktory postihují pohyb trhu.
2. Vše, co ovlivní cenu (vliv politický, ekonomický a jiné) už bylo promítnuto do cenového grafu. Toto je nejdůležitější předpoklad technické analýzy.
3. Ceny se pohybují v trendech.

Pokud se trh pohybuje ve směru trendu, může být na rozdíl od chaosu účinně analyzován. Pohyb ceny zaprvé naznačuje, že aktuální směr trendu bude pokračovat a neobráť se do opačného směru a zadruhé aktuální trend bude pokračovat až do té doby, než nastoupí opačný trend. [3]

Historie se opakuje.

Trh se často opakuje, byť pokaždé trochu jinak. Cenové vzory fungují již mnoho let a zobrazují poměrně přesně ekonomický a psychologický stav trhu. Není důvod nevěřit v jejich funkci i v budoucnu. [9]

Výhody technické analýzy:

Může pomoci se správným načasováním vstupu do trhu a výstupu z něj. „ Je důležité vědět, jaká je historie ceny aktiva. Jestliže již cena nějakého aktiva výrazně vzrostla v *uptrendu*, mohlo by být rozumné počkat si na zpětný pohyb. A podobně když jeho cena již dlouho klesá v *downtrendu*, mohlo by se vyplatit počkat si na návrat zájmu kupujících a na obrat trendu.“ [9]

Nevýhody technické analýzy:

Technickou analýzu je možno interpretovat různě. Je to velmi subjektivní věc, obvykle je závislá na množství zkušeností a schopností toho, kdo ji provádí. Dva analytici neudělají ze stejného grafu stejné závěry. Technická analýza poskytuje zpožděnou reakci na pohyb ceny, než se vzor vykreslí, trend již může být v pohybu a vstup do něj je rizikový. Také není zaručeno, že vykreslený vzor bude následovat změnu trendu. Existuje pouze určitá pravděpodobnost, že tomu tak bude. [3,9]

Grafy

Graf je základní zdroj informace pro technickou analýzu. Existuje celá řada grafů, nejvíce se používají časové grafy, kde se cena aktiva vykresluje do úseček, které jsou od sebe vzdáleny zvolený čas. Tomu se říká časový rámec - *timeframe*. Časové rámce mohou být rychlé - sekundové, minutové nebo pomalé nebo také měsíční, roční. Pro obchodování lze časové rámce libovolně nastavit.

Existují dva nejběžnější typy grafické formy:

- svícové grafy - *candlestick bar*
- úsečkové grafy - *barchart*

Na Obr. 1 je znázorněn svícový graf, ve kterém každá jednotlivá svíce se vykreslí za takový čas, jaký je zvolen časový rámec. [3,4,9]



Obr. 1 Graf Candlestick

Na Obr. 2 je znázorněn stejný graf, ale v úsečkové podobě. Grafy obsahují stejnou informaci, ale grafická interpretace je rozdílná. Při rozhodování jaký časový rámec je vhodné použít, je třeba brát v úvahu, že nižší *timeframe* obsahuje hodně šumu a spoustu drobných pohybů, které mohou být zavádějící, a neposkytuje pohled na trh v dlouhodobějším časovém horizontu.

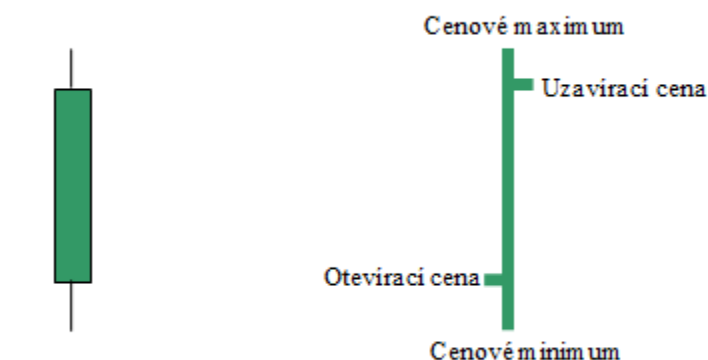


Obr. 2 Graf Barchart

Na Obr. 3 zobrazeny jsou úsečky, z nichž se tvoří grafy. Vlevo je úsečka svícového grafu a vpravo úsečkového. Úsečky v grafu mohou nabývat dvou barev podle toho, jestli došlo k růstu nebo poklesu ceny.

Zároveň každá úsečka udává čtyři informace:

1. Cenu, za kterou trh otevřel ve zvoleném *timeframe* (Otevírací cena-OPEN).
2. Nejvyšší cenu ve zvoleném *timeframe* (HIGH).
3. Nejnižší cenu ve zvoleném *timeframe* (LOW).
4. Cenu, za kterou trh uzavřel ve zvoleném *timeframe* (Uzavírací cena-CLOSE).



Obr. 3 Úsečky v grafu

Trendy

Trhy se obecně nepohybují v přímé linii, pohybují se nahoru a dolů. Trend je tedy směr pohybu trhu. Trend může být dlouhodobý, střednědobý a krátkodobý. Obchodovat lze do směru i proti směru trendu. Při obchodování je lepší být na straně silného pohybu. Nejvíce se využívá obchodování do směru trendu. [3,4]

Existují tři směry trendů:

1. Býčí graf - *Uptrend*

Na *Obr. 4* je znázorněn graf *Uptrend*. Existuje-li tento trend, obchodníci předpokládají růst ceny aktiva. Nakupují se záměrem a následně prodat za vyšší cenu.



Obr. 4 Graf Uptrend

2. Medvědí graf - *Downtrend*

Na *Obr. 5* je znázorněn graf *Downtrend*. Zde je situace opačná, obchodníci prodávají aktiva za vyšší cenu s úmyslem nakoupit za nižší.



Obr. 5 Graf Downtrend

3. Graf *Chop* se pohybuje do strany nebo také graf netendující

Na *Obr. 6* je znázorněn graf *Chop*. Když se vytvoří „Chop“, těžko se hledají obchodní příležitosti. [3,9]



Obr. 6 Graf Chop

2 INDIKÁTORY

Žádný obchodník neví, jak se trh bude měnit, vyvíjet. Vlivy, které na trh působí, jsou ovlivňovány mnoha faktory, takže spolehlivě předvídat jeho pohyb je nemožné. Přesto obchodníci potřebují podnět, na jehož základě do trhu vstupují. Ve skutečnosti potřebují s nějakou pravděpodobností odhadnout budoucí směr. [4]

Je možno použít:

- Indikátory
- Vzory (*patterny*) a jiné vizuální formace

2.1 Indikátory

Indikátory pomáhají obchodníkům při vstupu na trh a výstupu z trhu. Zachycují delší časové období a tím snáze identifikují střednědobý a dlouhodobý trend. Existuje mnoho druhů indikátorů. Dále jsou popsány klouzavý průměr, exponenciální klouzavý průměr a MACD indikátor. Na *Obr. 7*, *Obr. 8*, *Obr. 9*, *Obr. 10* jsou znázorněny stejné grafy s použitím různých indikátorů. [9]

2.1.1 Trendové indikátory

a) Klouzavý průměr MA - *Moving Average*

Patří k nejstarším indikátorům. Na jeho základě jsou vytvořeny další indikátory. Je to průměr několika posledních cen daného aktiva, zjištěný z cenového grafu. Pro výpočet můžeme použít ceny otevírací, uzavírací, maxima, minima a jejich kombinace. Nejlepší je vzít uzavírací cenu, protože má lepší vypovídací schopnost. Důležitým bodem je počet časových období, použitých při výpočtu klouzavého průměru (1). [9]

$$MA = (P_1 + \dots + P_n)/N \quad (1)$$

P_n – cena z intervalu n dnů

N – počet dnů

b) Exponenciální klouzavý průměr EMA – *Exponential Moving Average*

Váhy MA jsou rozděleny exponenciálně. Největší váhu mají nejčerstvější hodnoty MA, nejmenší váhu mají nejstarší hodnoty MA. Výpočet se provádí podle vzorce (2). [9]

$$EMA = K * P + (1 - K) \quad (2)$$

$$K = 2 / (P + 1)$$

P - perioda



Obr. 7 Exponenciální klouzavý průměr

c) MACD indikátor – *Moving Average Convergence Divergence*

Je založen na bázi MA a spolehlivě potvrzuje trend. Nedoporučuje se používat, pokud směřuje trh do strany. Vztah pro výpočet MACD indikátoru (3). [9]

$$MACD = EMA(n1) - EMA(n2) \quad (3)$$

MA - klouzavý průměr

$n1$ - počet posledních dnů

$n2$ - počet posledních dnů



Obr. 8 MACD indikátor

2.1.2 Oscilátory

Tyto indikátory ukazují, kdy je trh připraven se obrátit.

a) **RSI** - *Relative Strenght Index*

Měří posilování nebo naopak oslabování trhu. Pokud trh posiluje, je to signál k nástupu do dlouhé pozice, pokud trh oslabuje, je to nástup do krátké pozice, vztah pro výpočet je (4).

$$RSI = 100 - \left[\frac{100}{1 + RS} \right] \quad (4)$$

RS - podíl průměrných kladných hodnot změn ceny a podíl průměrných záporných hodnot změn ceny v dané časové periodě. [9]



Obr. 9 RSI

b) **CCI** - *Commodity Channel Index*

Měří kolísání ceny kolem jejího statistického průměru. Vysoká hodnota indexu říká, že cena je neobvykle vysoko vzhledem k její průměrné hodnotě a nízká hodnota upozorňuje na nízkou cenu. Indikátor velmi dobře potvrzuje nový nebo pokračující trend. Reaguje pomalu a tak je schopen vyloučit falešné signály jiných indikátorů. Index je popsán vztahem (5). [9]

$$CCI = (TP - MATP) / (MDTP * 0,015) \quad (5)$$

TP - *Typical Price* ((nejvyšší, nejnižší a uzavírací cena dne)/3)

MATP - jednoduchý MA s periodou ve vztahu k *TP*

MDTP - vypočtená střední odchylka od *TP*



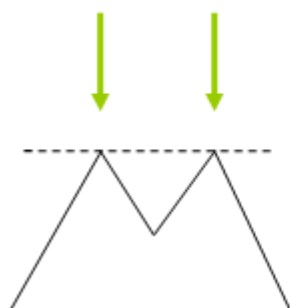
Obr. 10 Commodity Channel Index

2.2 Vzory - *patterny*

Jsou to grafické obrazce, které vytváří cenový graf. Generace obchodníků vypožadovaly, že určité konkrétní obrazce předchází s nějakou pravděpodobností změně trendu. Tak jako je mnoho typů indikátorů, existuje mnoho typů vzorů jako například dvojité dno, dvojitý vrchol, b2-reversal. Pro tuto práci byl vybrán dvojitý vrchol, který se může snadno vizuálně identifikovat v cenovém grafu. [4]

1. Dvojitý vrchol - *Double top*

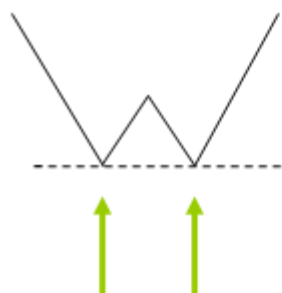
Typ obrazce dvojitý vrchol/dno, se nevyskytuje příliš často, zato je poměrně spolehlivý. Pokud se zobrazí vrchol, nastává pravděpodobnost poklesu trendu a pokud se zobrazí dno, s velkou pravděpodobností dojde k růstu ceny. Oba vrcholy musí být na stejné úrovni. Na *Obr. 11* vpravo ukazuje reálnou situaci z cenového grafu, kdy po vykreslení dvojitého vrcholu došlo k poklesu ceny. [4]



Obr. 11 Dvojitý vrchol

2. Dvojité dno - Double bottom

Stejná situace nastává také u dvojitého dna, jenom vzor je zrcadlově obrácený. Tady dochází naopak k růstu ceny. Dvojité dno se v trhu vyskytuje méně často. [4]



Obr. 12 Dvojité dno

3 UMĚLÁ INTELIGENCE A FINANČNÍ TRHY

Umělá inteligence je vědní obor, který simuluje lidské myšlení a chování. Počítačové systémy používají algoritmy a přesná pravidla, lidé se obvykle neřídí striktně podle pravidel a přesto dosahují svých cílů velmi efektivně.

„Doménami umělé inteligence jsou expertní úlohy (finanční analýzy, vědecké analýzy, lékařská diagnostika, inženýrství), formální úlohy (matematika, hry) a ostatní úlohy (přirozený jazyk, robotika, vnímání, apod.).“ [2]

V oblasti finančních trhů se používají zejména genetické algoritmy, neuronové sítě a fuzzy logika.

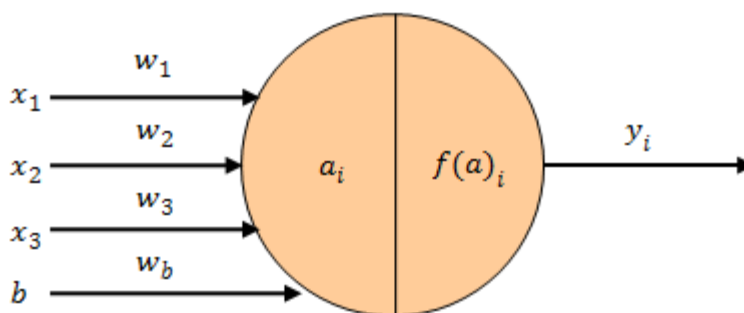
Genetické algoritmy řeší optimalizační úlohy. Proti klasickým optimalizačním metodám, které mají striktní pravidla, pracují genetické algoritmy s náhodnými veličinami. Na finančních trzích se dají genetické algoritmy použít k hledání optimálního složení portfolia aktiv.

Neuronové sítě jsou vhodné pro klasifikační a predikční úlohy. Jejich výhodou je lepší identifikace nelineárních vztahů mezi vstupními a výstupními hodnotami. Mají schopnost generalizace a učení podle příkladů, proto jsou vhodné pro použití při rozeznávání vzorů *pattern recognition*.

Fuzzy logika využívá technologii neostrých množin. Fuzzy logika přiřazuje výroky hodnotu pravda, nepravda, částečná pravda, ne zcela úplná pravda. Pracuje v šedé zóně mezi „ano“ a „ne“. Fuzzy technologie se dá využít na finančních trzích jako expertních systémů k podpoře rozhodování investorů nebo k predikci změny ceny. [2]

4 NEURONOVÉ SÍTĚ

Neuronová síť je výpočetní model, který velmi hrubě napodobuje činnost lidského mozku. Základem neuronové sítě je umělý neuron. Na *Obr. 13* je zobrazen umělý neuron, který má vstupy x_i , aktivační funkci a_i , váhy w_i , přenosovou funkci f_i a výstup y_i . Propojením vstupů a výstupů vzniká neuronová síť. Typ propojení se nazývá topologie sítě. [8]



Obr. 13 Umělý neuron

Vstupy do neuronů jsou číselné:

- kvalitativní vstupy
- kvantitativní vstupy

Kvalitativní vstupy - nabývají číselných hodnot. Mohou být rozsahu $< 0,1 >$ nebo $< -1,1 >$. Čísla se vážou k projevu určité vlastnosti daného objektu. Pokud vlastnost neexistuje, pak vstup bude 0 nebo -1 , pokud vlastnost existuje ve 100%, pak je vstup 1.

Kvantitativní vstupy - mohou být skutečně naměřené hodnoty. Každý vstup x_i může být upraven vahou w_i , která udává důležitost toho kterého vstupu. Výpočty hodnot vah a jejich ladění jsou podstatnou částí učení neuronových sítí. Na začátku učení mají váhy defaultní hodnoty.

- Aktivační funkce - je to vnitřní potenciál neuronu. Podle vztahu (6,7).

$$a = \sum_{i=1}^{n+1} x_i w_i \quad (6)$$

$$a = \sum_{i=1}^{n+1} x_i w_i + b w_b \quad (7)$$

x_i - vstupy

w_i - váhy

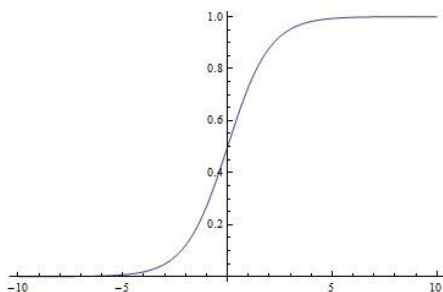
b - práh

w_b - váha prahu

- Přenosová funkce

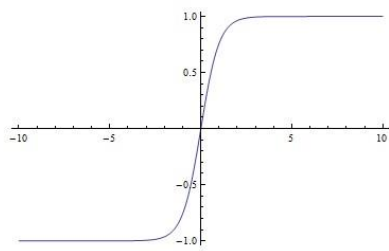
Převádí hodnotu vstupu na výstup. Používají se lineární a nelineární přenosové funkce. Mezi nelineární patří logistická sigmoida (8), hyperbolický tangens (9) a funkce Perceptron (10). Mezi lineární patří lineární (12) a lineární omezená funkce (11).

a) Logistická sigmoida



$$f(a) = \frac{1}{1 + e^{-\lambda a}} \quad (8)$$

b) Hyperbolický tangens



$$f(a) = \frac{e^a - e^{-a}}{e^a + e^{-a}} \quad (9)$$

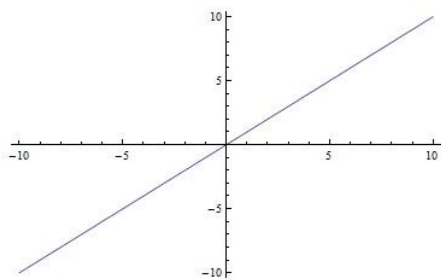
c) Perceptron

$$f(a) = a \quad f(a) = 0 \quad \text{pro } \forall a > 0 \quad (10)$$

d) Lineární omezená funkce

$$f(a) = \gamma \quad \text{pro } \forall a = h_1 \wedge f(a) = -\gamma \quad \text{pro } \forall a < -h_1 \wedge f(a) = ka \quad \text{pro } \forall a \in \langle -h_1, h_1 \rangle \quad (11)$$

e) Lineární funkce



$$f(a) = ka \quad (12)$$

- Výstupy - obvykle se v celé síti používají stejné typy vstupů a výstupů.

4.1 Neuronové sítě bez učitele

Učení bez učitele je proces, kdy síť zná vstupy, ale nezná výstup. Vstupní vzory si síť sama třídí do skupin podle typického vzoru nebo si přizpůsobí topologii podle vstupu. Jako příklad může sloužit Kohonenova a Hopfieldova síť. [6]

4.2 Neuronové sítě s učitelem

O tomto typu sítě mluvíme tehdy, když známe požadovaný originál a chceme, aby se výstup sítě tomuto originálu co nejvíce blížil. Síť je „seznámena“ s originálem a na základě aktuálního nastavení je zjištěn výsledek. Rozdíl mezi výsledkem a originálem se nazývá chyba učení. Pokud je chyba větší, než je přípustné, upravíme hodnoty vah a opakujeme učení sítě až do okamžiku, kdy je hodnota chyby vyhovující. Síť se učí změnou nastavení vah. První síť, která byla schopna učení, se nazývá Perceptron. Tato síť má jednu vstupní pevnou vrstvu a jednu výstupní, ve které se mohou měnit váhy. Dnes již existují vícevrstvé neuronové sítě, vrstvy jsou mezi sebou propojeny od vstupu na výstup. Jako učicí algoritmus se používá Backpropagation, jako základní a dále Levenberg-Marquardtův a Gauss-Newtonův algoritmus. Má fázi aktivační a adaptační. Aktivační fáze je vybavovací a adaptační fáze je učicí. [6,7]

4.2.1 Vícevrstvé neuronové sítě

Typ vícevrstvých neuronových sítí je nejvíce rozšířený model v mnoha praktických aplikacích. Je znám pod různými jmény, jako např. vícevrstvý Perceptron nebo Feedforward network. Kromě vstupní a výstupní vrstvy může mít jednu nebo více skrytých vrstev. Počet neuronů ve vstupní vrstvě je definován počtem vstupů a počet neuronů ve výstupní vrstvě je rovněž definován počtem výstupů. Ve skrytých vrstvách se počet neuronů nastavuje při inicializaci sítě. Také se nastavuje počet skrytých vrstev.

Výstup ze sítě je (13) [8]:

$$y = \sum_{i=1}^{nh} w_i^2 \sigma \left(\sum_{j=1}^n w_{i,j}^1 y_j + b_{j,i}^1 \right) + b^2 \quad (13)$$

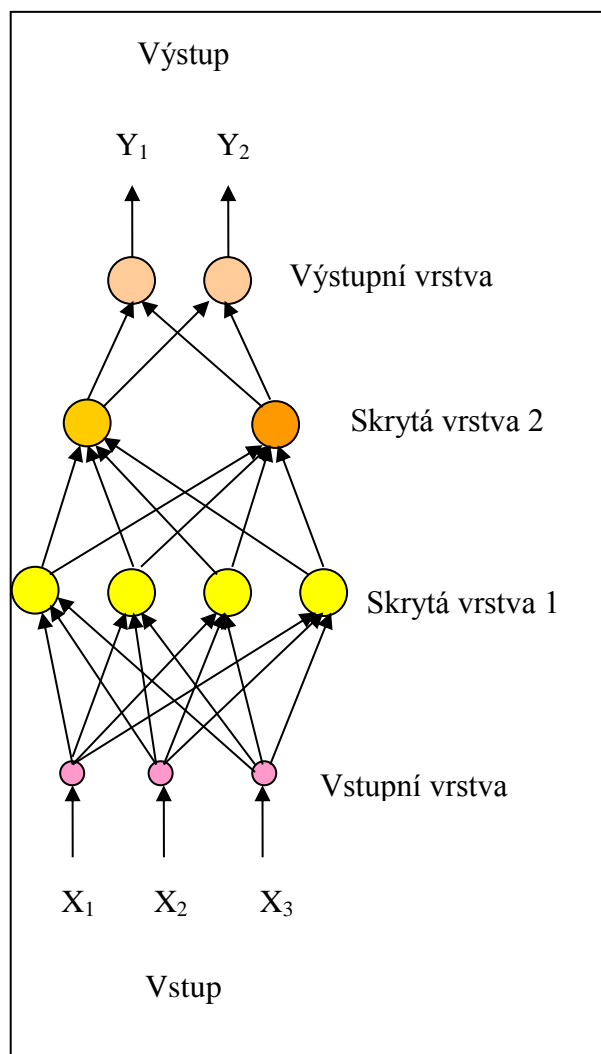
w_i - váhy

σ - přenosová funkce

b - práh

nh - počet skrytých neuronů

y_j - výstup



Obr. 14 Vícevrstvá síť

4.2.2 Backpropagation

Backpropagation nastavuje nebo opravuje jednotlivé váhy zpětným chodem tak, aby jejich hodnoty byly z hlediska řešeného problému co možná optimální. Nastavení vah probíhá v opačném směru, než jakým se šíří vstupní informace. Algoritmus hledá globální minimum chybové funkce. [5]

a) Aktivační fáze

V aktivační fázi se vstupní informace modifikuje momentální množinou vah a přenosovými funkcemi a dostane se na výstup. Při inicializaci sítě se váhy nastavují na

vhodnou hodnotu generátorem náhodných čísel. Kromě generátoru se dají použít genetické algoritmy a simulované žíhání.

b) Adaptační fáze

V této fázi se nastavují váhy sítě tak, aby rozdíl mezi originálem a výstupem ze sítě byl menší než dovolená chyba. Rozdíl je použit pro výpočet nových vah. Při nastavování vah se začíná u spojů, které vstupují do výstupní vrstvy a postupuje se směrem k nižším vrstvám. Po dosažení vstupní vrstvy se opakuje aktivační fáze. Opakování se nazývá iterace. Rozdíl originálu a výstupu se nazývá globální chyba (15). Musí být menší, než je dovolená chyba. [5,8]

- Chyba za vektor (14)

$$E(a) = \sum_{j=1}^m [y_j(k) - d_j(k)]^2 \quad (14)$$

$d_j(k)$ - je požadovaný výstupní vzor

$y_j(k)$ - je odezva na vstupní vzor

- Globální chyba za epochu (15)

$$E_T = \sum_{k=1}^N E(k) \quad (15)$$

N - je počet dvojic vstup-výstup [5,8]

4.2.3 Trénovací množina

Na trénovací množinu jsou kladeny určité nároky. Musí obsahovat všechny podstatné charakteristiky modelovaného objektu a musí být dostatečně široká. Špatně připravená trénovací množina vede ke špatným výsledkům. [5]

Problémy, které mohou nastat:

a) Počet vzorků

Vzorky mají být voleny tak, aby výběr pokryl daný problém. V trénovací množině by měly být zastoupeny všechny třídy, které se v problému vyskytují. Počet vzorků z jednotlivých tříd by měl být úměrný velikosti třídy. Informace, kterou vzorek nese, musí přímo souviset s daným problémem. I vzorky, které jsou „špatné“ musí být zastoupeny. [5]

b) Délka a posun vektorů

Čím je vektor delší, tím podrobněji popisuje problém. Na druhé straně délka vektoru prodlužuje dobu učení a narůstají váhy. Posuv vektoru se používá u časových řad. [5]

c) Transformace

Vstupní hodnoty je potřeba transformovat do intervalu $< 0,1 >$ nebo $< -1,1 >$ protože přenosové funkce se pohybují v těchto intervalech. Transformace se provádí tak, že vstupní hodnoty se dělí konstantou. Jde o lineární transformaci, která se používá v případě, kdy jsou vstupní hodnoty rozprostřeny rovnoměrně. Pokud tomu tak není, je potřeba použít transformaci nelineární. [5]

d) Hraniční případy

V některých případech je těžké rozhodnout, do které třídy vzorek patří. Pak je dobré mít vyšší počet takových vzorků v trénovací množině. [5]

II. PRAKTICKÁ ČÁST

5 ÚVODEM

V literatuře od Jiřího Fanty [2,3] jsou popsány způsoby využití umělé inteligence pro predikci. Autor - Jiří Fanta použil vícevrstvý Perceptron, který natrénovat pomocí akcií ČEZ. Dalšími vstupy byly indikátory technické analýzy *Momentum*, *Price Rate of Change* a *Moving Average Simple*. *Momentum* je indikátor, který měří sílu a rychlost trendu, *Price Rate of Change* měří procentuální změnu mezi poslední minulou cenou a cenou v minulosti, *Moving Average Simple* je aritmetický průměr posledních hodnot.

Podle autora zde [2,3] dávají indikátory *Momentum*, *Price Rate of Change* a *Moving Average Simple* signál k nákupu nebo prodeji. V knize *Technical Analysis of the Financial Markets: A Comprehensive Guide to Trading Methods and Applications* [4] jsou popsány prostředky technické analýzy a mezi nimi také různé grafické formace v cenových grafech, po nichž následuje změna trendu. Neuronové sítě se dobře dají použít k *pattern recognition*, proto byly v této práci k průzkumu a ověření možností detekce formací v časovém grafu.

5.1 Použitý software

Existuje několik matematických integrovaných prostředí, vhodných pro práci s neuronovými sítěmi. Patří k nim např. Statistica, Matlab a také Mathematica. V této práci byla použita Mathematica 8.

Mathematica je software vytvořený firmou Wolfram Research. Za jeho vznikem stojí idea vytvořit samostatný systém, který by zvládl všechny aspekty technického programování. Používá se v ekonomii, matematice, fyzice, geografii, informatice, na středních a vysokých školách. Mathematica pracuje s daty jakéhokoli typu, finančními, geografickými, chemickými, astronomickými. Data lze exportovat a importovat v dvou stech podporovaných formátech. [10]

5.2 Vstupy

Jako vstup do neuronové sítě byl použit vzor dvojité vrchol *double top*, jak je ukázáno na Obr. 15. Každá svíce vzoru má uzavírací cenu, která byla brána jako výchozí hodnota pro tvorbu testovací a trénovací množiny.

Obr. 15 Vrchol *Double Top*

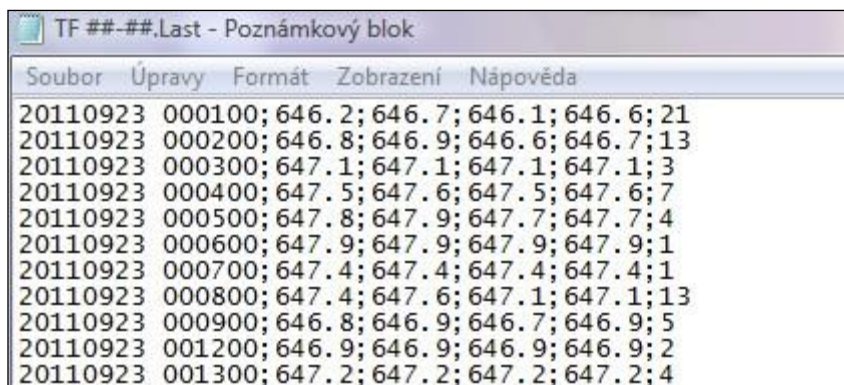
5.2.1 Testovací množina

Testovací množina byla vytvořena z časového grafu indexu Russel 2000, který je obchodován na americké burze NYBOT v New Yorku. K získání dat od poskytovatele byla použita „sim“ verze obchodního software NinjaTrader, který slouží k on-line obchodování na burze. Tento software je k dispozici na internetu, pro simulované obchodování zdarma. Umožňuje backtestování historických dat, tvorbu mechanických strategií a také omezeně reálné obchodování. Úvodní stránka programu je na *Obr. 16*. [11]



Obr. 16 Úvodní stránka Ninja Traderu

Z nainportovaných dat byl vytvořen graf, ve kterém byly nalezeny hledané vzory DT pro tvorbu tréninkové množiny a také data byla vyexportována v textové podobě pro použití v Mathematice. První sloupec obsahuje datum ve formátu rok-měsíc-den, druhý sloupec je čas, třetí sloupec je otevírací cena, čtvrtý nejvyšší dosažená cena, pátý nejnižší cena, šestý uzavírací cena každé jednotlivé svíce a poslední je volume. Pro natrénování sítě byla použita uzavírací cena tj. šestý sloupec *Obr. 17*.



TF ##-##.Last - Poznámkový blok					
Soubor Úpravy Formát Zobrazení nápověda					
20110923	000100;	646.2;	646.7;	646.1;	646.6; 21
20110923	000200;	646.8;	646.9;	646.6;	646.7; 13
20110923	000300;	647.1;	647.1;	647.1;	647.1; 3
20110923	000400;	647.5;	647.6;	647.5;	647.6; 7
20110923	000500;	647.8;	647.9;	647.7;	647.7; 4
20110923	000600;	647.9;	647.9;	647.9;	647.9; 1
20110923	000700;	647.4;	647.4;	647.4;	647.4; 1
20110923	000800;	647.4;	647.6;	647.1;	647.1; 13
20110923	000900;	646.8;	646.9;	646.7;	646.9; 5
20110923	001200;	646.9;	646.9;	646.9;	646.9; 2
20110923	001300;	647.2;	647.2;	647.2;	647.2; 4

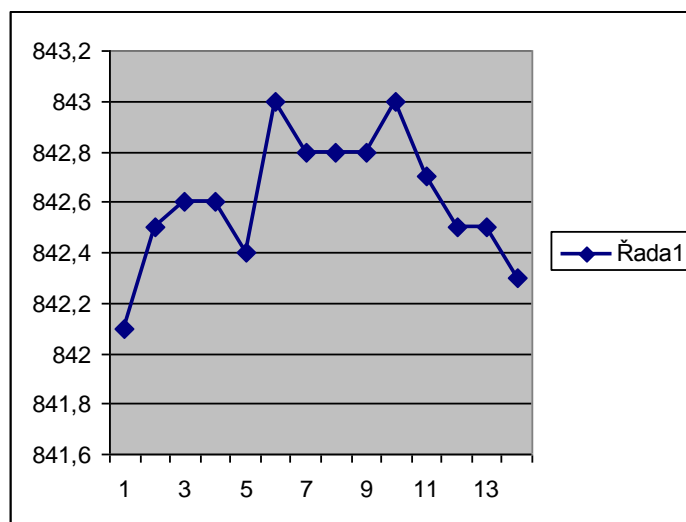
Obr. 17 Textová podoba dat

5.2.2 Trénovací množina

V časovém grafu indexu Russel 2000 byly vyhledány vzory Double Top a také to, co se nedá považovat za Double Top. Množina byla vytvořena z uzavíracích cen jednotlivých svíček a byla vložena do Excelu. Jednu část tvoří validní vzory, tj. formace Double Top a druhou část nevalidní vzory, tj. formace, které určitě nejsou Double Top. Bylo vyhledáno 20 validních vzorů a 20 vzorů nevalidních. Na *Obr. 18* je validní vzor DT, který se vykreslil v grafu ze dne 10. 9. 2012, na *Obr. 19* je stejný vzor jako graf Excelu.



Obr. 18 Validní vzor

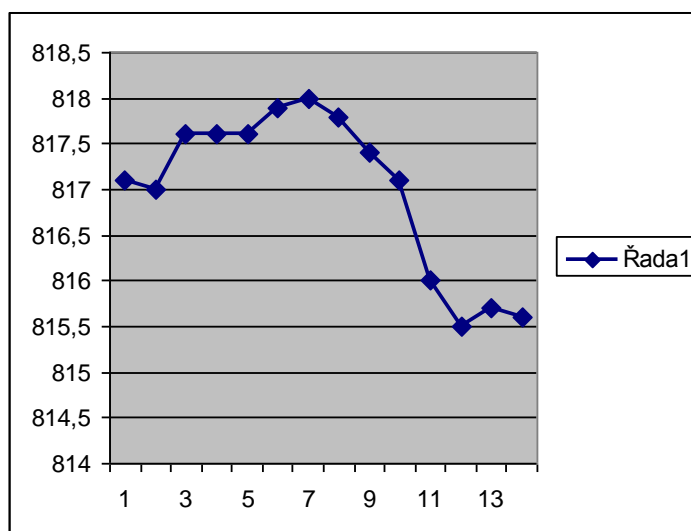


Obr. 19 Validní vzor v Excelu

Na *Obr. 20* je nevalidní vzor, který byl nalezen v grafu ze dne 5. 12. 2012, na *Obr. 21* je stejný vzor jako graf Excelu.



Obr. 20 Nevalidní vzor



Obr. 21 Nevalidní vzor v Excelu

Z časového grafu byly naskenovány všechny formace, které byly použity k vytvoření trénovací množiny.

Poté bylo nutno uzavírací ceny převést do intervalu $<0,1>$ pomocí Excelu a importovat do Mathematicy. Pro převod cen do intervalu $<0,1>$ byl použit matematický vztah (16):

$$y = \frac{(x - \min)}{(\max - \min)} \quad (16)$$

x – uzavírací cena jednotlivé svíce ve vzoru

y – hodnota x převedená do interval $<0,1>$

max – maximální uzavírací cena obsažená ve vzoru

min – minimální uzavírací cena obsažená ve vzoru

5.3 Trénování sítě

Nejprve byla použita neuronová síť feedforward s učícím algoritmem Backpropagation a ve druhé variantě síť feedforward s učícím algoritmem Levenberg-Marquardt.

Data pro učení neuronové sítě tvoří 20 vzorů validních a 20 vzorů nevalidních. Každý jednotlivý vzor obsahuje různý počet svíci. Vstupní množina neuronové sítě musí obsahovat stejný počet neuronů, proto byly jednotlivé vzory doplněny nulami tak, aby počet vstupních neuronů se rovnal 17, což je počet svíci v nejdelším vzoru.

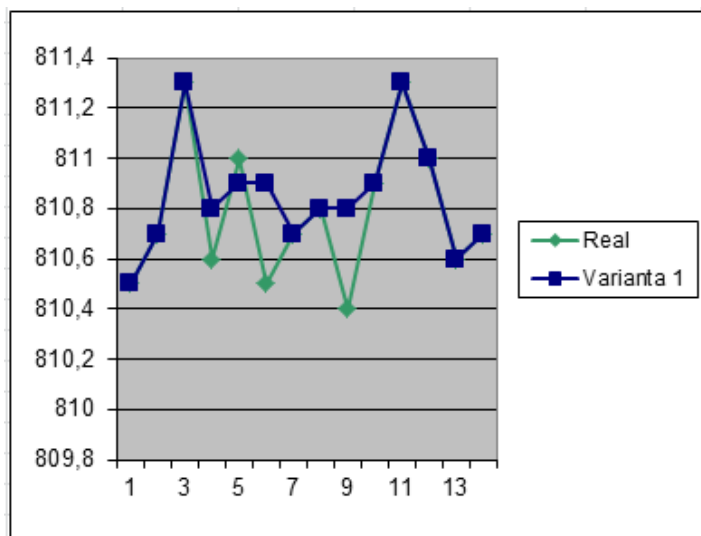
Byly připraveny tyto typy množin:

1. trénovací množina z reálných dat

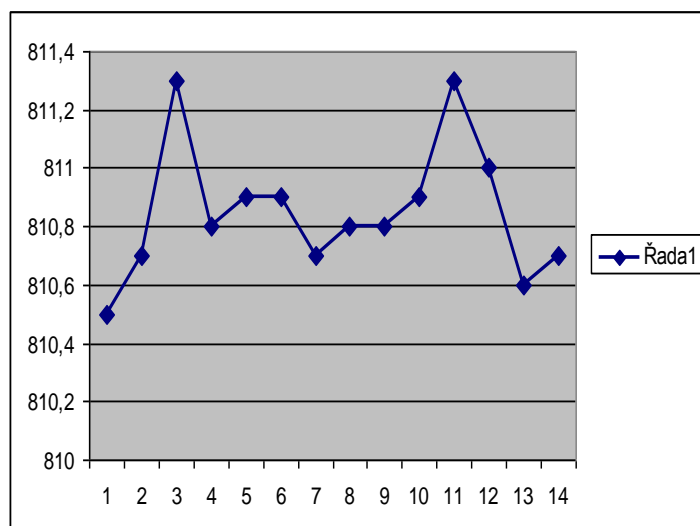
Zde jsou použity skutečné uzavírací ceny jednotlivých svíci. Tento způsob dával špatné výsledky, protože síť nacházela mnoho vzorů, které se v reálném grafu nevyskytovaly.

2. trénovací množina z upravených dat varianta 1

Původní vzory byly zjednodušeny, jak je zobrazeno na *Obr. 22*. V tomto případě nastal pokles nalezených vzorů.



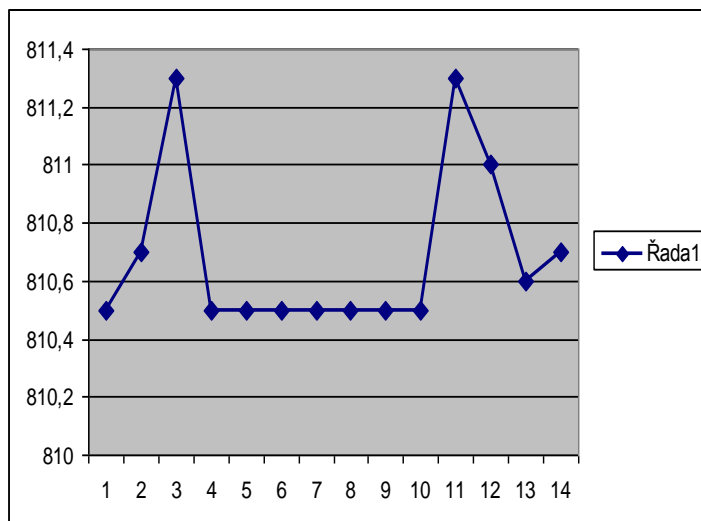
Obr. 22 Zjednodušení vzoru



Obr. 23 Vzor varianta 1

3. trénovací množina z upravených dat varianta 2

Po zlepšení výsledků (zmenšení počtu nalezených vzorů) při použití trénovací množiny varianty 1 byly původní vzory zjednodušeny ještě více, jako na *Obr. 24*.



Obr. 24 Vzor varianta 2

5.3.1 Síť feedforward s algoritmem Backpropagation

Na počátku byla použita síť feedforward a trénovací množina, která je vytvořena z reálných dat. Pro naučení sítě byl použit učící algoritmus Backpropagation. Jako zdroj dat byly použity uzavírací ceny dané časové řady převedené do textového souboru. Síť měla jednu skrytou vrstvu se 4 neurony a bylo použito 700 iterací. Počet neuronů ve skryté vrstvě byl určen podle vztahu (17) ze skript Umělá inteligence I od Ivana Zelinky [5]:

$$N_{skryt} = \sqrt{N_{vstup} * N_{výstup}} \quad (17)$$

po dosazení $N_{skryt} = \sqrt{(17 * 1)} = 4,12 \approx 4$

N_{skryt} - počet neuronů ve skryté vrstvě

N_{vstup} - počet neuronů ve vstupní vrstvě

$N_{výstup}$ - počet neuronů ve výstupní vrstvě

Síť našla 2769 vzorů. Tato hodnota neodpovídá skutečnosti. Bylo zřejmé, že je nutno vyzkoušet více variant nastavení sítě, což představuje použití různý počet skrytých vrstev a neuronů a zároveň také více typů vzorů trénovací množiny.

Počet neuronů v dalších skrytých vrstvách (18,19) byl vypočítán podle [5]:

$$N_{Skryt\ 1} = N_{Výstup} * (\sqrt[3]{N_{Vstup} / N_{Výstup}})^2 \quad (18)$$

po dosazení $N_{Skryt\ 1} = 1 * (\sqrt[3]{17/1})^2 = 6,6 \approx 7$

$$N_{Skryt\ 2} = N_{Výstup} * (\sqrt[3]{N_{Vstup} / N_{Výstup}}) \quad (19)$$

po dosazení $N_{Skryt\ 2} = 1 * (\sqrt[3]{17/1}) = 2,57 \approx 3$

V dalším kroku byla použita stejná síť, ale se dvěma skrytými vrstvami se čtyřmi neurony v první skryté vrstvě a sedmi neurony ve druhé skryté vrstvě. Výsledkem bylo nalezení 3141 vzorů.

Dále bylo použito nastavení sítě se třemi skrytými vrstvami s počtem neuronů 4, 7 a 3. Nalezeno 7045 vzorů.

Předchozí nastavení sítě a použité reálné vzory dávají velký počet nalezených hodnot, neodpovídajících realitě. Proto bylo přistoupeno k použití zjednodušeného vzoru – varianta 1. V následující *Tab. 1* je shrnutí nastavení sítě a dosažených výsledků.

Typ tréninkové množiny	Počet skrytých vrstev	Počet neuronů ve skrytých vrstvách	Počet nalezených vzorů
Reálná data	1	4	2769
Reálná data	2	4,7	2643
Reálná data	3	4,7,3	7045
Varianta 1	1	4	2341
Varianta 1	2	4,3	3188
Varianta 1	3	4,7,3	5063

Tab. 1 Porovnání výsledků reálných vzorů a varianty 1

Je zřejmé, že zvyšující se počet skrytých vrstev spíše zhoršuje dosažené výsledky a také zjednodušení tréninkového vzoru vede k výsledkům, které více ukazují cestu ke skutečnosti.

Algoritmus Backpropagation potřebuje velký počet iterací pro naučení sítě a tím i dlouhou dobu běhu, což není pro použití v reálné aplikaci vhodné, proto byl dále použit učící algoritmus Levenberg-Marquardt.

5.3.2 Sít' feedforward s algoritmem Levenberg-Marquardt

Levenberg-Marquardt je zlepšený algoritmus Backpropagation. Potřebuje k naučení sítě mnohem méně iterací a tak je daleko vhodnější. Tento algoritmus je defaultně v Mathematice nastaven.

Aby bylo možno posoudit, jaké nastavení sítě je nejvhodnější, byly použity stejné tréninkové množiny jako u sítě s Backpropagationem tj. reálné vzory a také varianta 1.

Sít' byla nastavena na jednu skrytou vrstvu s 4 neurony, jako učící algoritmus byl použit Levenberg-Marquardtův, tréninková množina ze skutečných dat a 45 iterací. Výsledkem bylo 5047 nalezených vzorů. Tato hodnota opět nebyla reálná. Bylo vyzkoušeno více variant nastavení sítě, jak ukazuje *Tab. 2*.

Typ tréninkové množiny	Počet skrytých vrstev	Počet neuronů ve skrytých vrstvách	Počet nalezených vzorů
Reálná data	1	4	5047
Reálná data	2	4,4	4993
Reálná data	3	4,7,3	5150
Varianta 1	1	4	2623
Varianta 1	2	4,3	3913
Varianta 1	3	4,7,3	3846

Tab. 2 Porovnání výsledků reálných vzorů a varianty 1

Ukazuje se, že nejlepší výsledky jsou získány, když se do tréninkové množiny použije zjednodušená varianta vzoru. Další cesta byla tímto směrem. Varianta 1 byla dále upravena na variantu 2 podle *Obr. 24*.

Síť byla natrénována pomocí vzoru varianty 2. Použitý algoritmus Levenberg-Marquardt a vyzkoušeno nastavení sítě podle *Tab. 3*, kde je pro porovnání uvedena i varianta 1.

Typ tréninkové množiny	Počet skrytých vrstev	Počet neuronů ve skrytých vrstvách	Počet nalezených vzorů
Varianta 1	1	4	2623
Varianta 1	2	4,3	3913
Varianta 1	3	4,7,3	3846
Varianta 2	1	4	1901
Varianta 2	2	4,3	3631
Varianta 2	3	4,7,3	2601

Tab. 3 Porovnání varianty 1 a varianty 2

Nejlepších výsledků bylo dosaženo s jednovrstvou sítí a trénovací množinou varianty 2.

Přesto je počet nalezených vzorů příliš velký, v datech, které byly použity k analýze, se nachází kolem jednoho sta vzorů. Bylo potřeba zjistit, co vlastně síť našla. Na *Obr. 25* je část výsledků z Mathematicy. Každé číslo udává pozici nalezeného vzoru v datech.

```
pozice = Position[klasifikaceOutRound, {1}]

{{15}, {16}, {49}, {50}, {51}, {73}, {105}, {106}, {107}, {121}, {122}, {123}, {124},
{125}, {126}, {127}, {128}, {148}, {149}, {150}, {151}, {156}, {159}, {161}, {162},
{163}, {164}, {165}, {166}, {182}, {186}, {187}, {188}, {204}, {205}, {206}, {207},
{216}, {217}, {218}, {236}, {237}, {238}, {240}, {257}, {258}, {259}, {260}, {261},
{263}, {264}, {265}, {271}, {284}, {285}, {286}, {287}, {288}, {289}, {290}, {318},
{319}, {320}, {321}, {322}, {323}, {324}, {325}, {326}, {327}, {328}, {332}, {333},
{334}, {358}, {359}, {360}, {364}, {365}, {366}, {376}, {377}, {378}, {379}, {380},
{393}, {394}, {395}, {396}, {397}, {398}, {399}, {400}, {403}, {404}, {406}, {407},
{409}, {410}, {411}, {412}, {413}, {414}, {421}, {426}, {427}, {449}, {451},
{452}, {453}, {454}, {455}, {457}, {461}, {462}, {463}, {464}, {465}, {480},
{482}, {483}, {488}, {489}, {490}, {491}, {492}, {494}, {495}, {533}, {539},
{540}, {541}, {542}, {557}, {559}, {560}, {562}, {573}, {574}, {575}, {576},
```

Obr. 25 Pozice nalezeného vzoru

První vzor byl zjištěn na pozici 15 a 16. Na *Obr. 26* je pozice 15 a na *Obr. 27* je pozice 16.

Je patrné, že v následujícím čase došlo k poklesu cen. Mohla by to být obchodní příležitost, ale hledaný vzor to není, protože síť nenašla vzor DT, ale dvě svíce jdoucí po sobě se stejnou uzavírací cenou.



Obr. 26 Vzor nalezený na pozici 15



Obr. 27 Nalezený vzor na pozici 16

Příští další vzor je na pozici 49, 50, 51. V případě pozice 49 je v čase 17.19 uzavírací cena 907,1 a tatáž uzavírací cena se objeví v čase 17.21. Na *Obr. 28* jsou označeny šipkami. Ovšem svíce označené šipkami nejsou vrcholy DT a uzavírací cena mezi nimi by měla poklesnout, což se v tomto případě nestalo. V pozicích 50 a 51 je to podobný případ, jak je ukázáno na *Obr. 29* a *Obr. 30*.



Obr. 28 Nalezený vzor na pozici 49



Obr. 29 Nalezený vzor na pozici 50

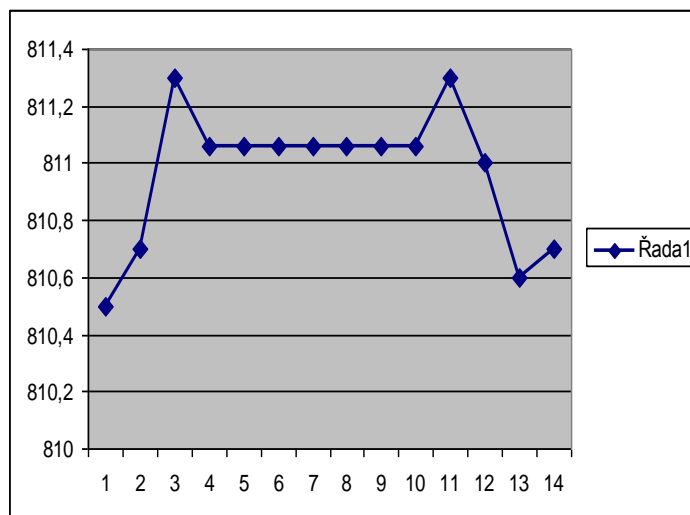


Obr. 30 Nalezený vzor na pozici 51

Podrobnějším rozbořem dalších pozic bylo zjištěno, že ve většině případů síť našla dvě svíce stejné uzavírací ceny v blízkosti sebe, někdy vzdálené jen jednu svíci od sebe. Síť poměrně spolehlivě najde formaci, ve které jsou obsaženy dvě stejné uzavírací ceny, avšak nebere ohled na uzavírací ceny svíci, které jsou mezi „vrcholy“ a také nalezená formace není vzor DT. Dále dochází k překrývání formací a jejich počet je příliš velký. *Obr. 28, Obr. 29 a Obr. 30.*

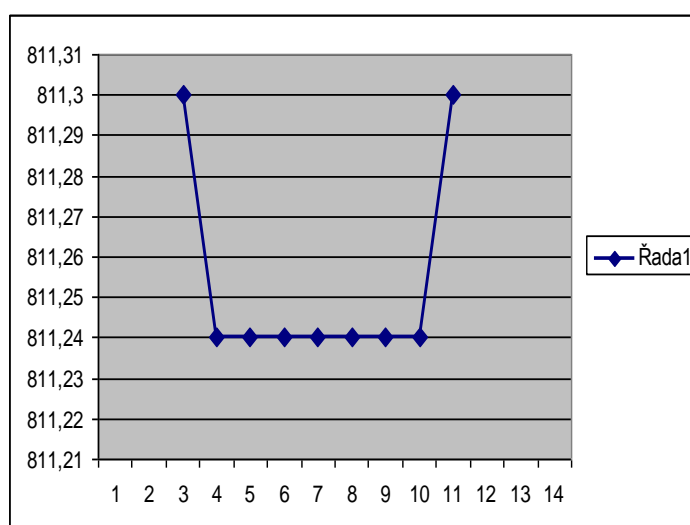
Aby síť lépe vyhledávala požadované vzory, bylo nutno dále upravit trénovací množinu. V dalších vzorech byl posílen význam svíci mezi „vrcholy“ tím, že hodnoty uzavíracích

cen byly zvětšeny tak, aby měly vyšší hodnotu bez ohledu na skutečnost. Vrcholům byla přiřazena hodnota 1 a svícím mezi nimi druhá nejvyšší hodnota ve vzoru na *Obr. 31*, který je možno porovnat s *Obr. 24*.



Obr. 31 Úprava vzoru

Tato úprava zlepšení nepřinesla, bylo nalezeno 2944 vzorů. Dále bylo přistoupeno ke zjednodušení tohoto vzoru tak, jak je znázorněno na *Obr. 32*. Byly vynechány krajní hodnoty ve vzoru a zůstaly zachovány jen vrcholy a to, co je mezi nimi.



Obr. 32 Vzor s vynechanými krajními hodnotami

Výsledkem bylo radikální snížení počtu nalezených vzorů na 123. Vzory se nacházely na pozici 13, 105, 160 a další pozice jsou znázorněny na *Obr. 33*, které se již více blíží realitě.

```
pozice = Position[klasifikaceOutRound, {1}]
{{13}, {105}, {160}, {161}, {200}, {201}, {202}, {269}, {357}, {358}, {370}, {739},
{740}, {792}, {793}, {794}, {820}, {958}, {1031}, {1061}, {1187}, {1223}, {1305},
{1408}, {1458}, {1530}, {1627}, {1677}, {1678}, {1703}, {1776}, {1812}, {1877},
{1978}, {1888}, {1889}, {1986}, {1987}, {1988}, {2152}, {2264}, {2513}, {2597},
{2771}, {2787}, {2788}, {2789}, {2860}, {2863}, {3043}, {3140}, {3190}, {3191},
{3492}, {3493}, {3546}, {3602}, {3743}, {3745}, {3746}, {3807}, {3831}, {3832},
{3892}, {3893}, {3894}, {3895}, {3971}, {3972}, {3991}, {3992}, {3993}, {4038},
{4141}, {4142}, {4265}, {4277}, {4401}, {4415}, {4430}, {4431}, {4432}, {4557},
{4563}, {4705}, {4723}, {4735}, {4742}, {4761}, {5332}, {5396}, {5615}, {5654},
{5655}, {5656}, {5770}, {5864}, {5912}, {5976}, {6125}, {6162}, {6204}, {6331},
{6344}, {6455}, {6456}, {6457}, {6458}, {6569}, {6629}, {6640}, {6690}, {6785},
{6889}, {6890}, {6891}, {6984}, {7000}, {7233}, {7338}, {7502}, {7516}, {7681}}
```

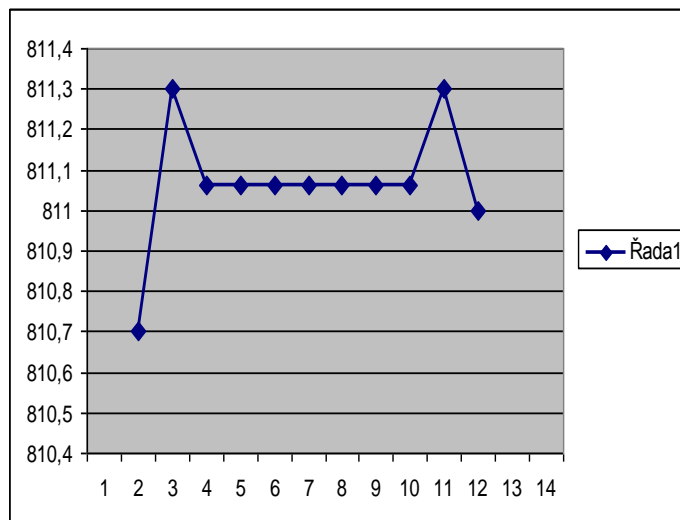
Obr. 33 Nalezené pozice vzoru v Mathematice

Bylo nutno posoudit kvalitu nalezených vzorů. Zde nedošlo ke zlepšení. Nalezené vzory nevytváří DT, přestože se zde nachází svíce se stejnou uzavírací cenou. Vzor na pozici 160 *Obr. 34* je zajímavý tím, že po jeho vykreslení dojde k poklesu ceny, ale nesplňuje podmínku DT.



Obr. 34 Vzor na pozici 160

Zjednodušení vzoru na *Obr. 32* vedlo ke snížení počtu nalezení. Nyní bylo potřeba, aby se nalezené vzory co nejvíce blížily hledanému vzoru DT. Byla vytvořena další trénovací množina tak, že do vzoru byly přidány krajní hodnoty podle *Obr. 35*.



Obr. 35 Vzor s přidanými krajními hodnotami

Počet nalezených vzorů byl 790 a úprava trénovací množiny nevedla k lepšímu vyhledání vzoru.

ZÁVĚR

Diplomová práce měla za cíl posoudit využití umělé inteligence na finančních trzích.

V časovém grafu indexu Russel 2000 byl vyhledán vzor DoubleTop (DT) a z něj bylo vytvořeno několik variant trénovacích množin. Pro nalezení správného vzoru jsou potřebné zkušenosti s technickou analýzou trhu. Bylo vytvořeno 6 typů trénovacích množin. První typ trénovací množiny byl reálný vzor, další typy byly zjednodušené a vycházely z reálného vzoru. V druhém typu trénovací množiny byla upravena data mezi vrcholy – zmenšené rozdíly mezi daty. Ve třetím typu bylo provedeno srovnání dat uprostřed vzoru na stejnou hodnotu. Ve čtvrté trénovací množině bylo provedeno posílení významu dat mezi vrcholy. V páté došlo k odstranění krajních hodnot, kvůli maximálnímu zjednodušení vzoru. Do šesté množiny naopak byly přidány krajní hodnoty ke vzoru pět. Nejlepší výsledky byly v páté trénovací množině, protože počet nalezených vzorů se blížil reálnému stavu. Při změně trénovací množiny musí být zachovány vrcholy, v opačném případě vzor tréninkové množiny neodpovídá vzoru DT.

Také z grafu byla vytvořena testovací množina z dat, která byla získána prostřednictvím aplikace Ninja Trader.

Pomocí těchto dat byla naučena neuronová síť feedforward. Byly použity dva učící algoritmy – Backpropagation a Levenberg-Marquardt. Více se osvědčilo použití algoritmu Levenberg-Marquardt, protože k naučení sítě potřebuje 45 iterací. Bylo vyzkoušeno různé nastavení sítě. Byla použita síť se 3, 2 a 1 skrytou vrstvou s různým počtem neuronů ve skrytých vrstvách. Bylo zjištěno, že nejlepší nastavení sítě je s jednou vrstvou a 4 neurony.

Nejlepšího výsledku bylo dosaženo pátou tréninkovou množinou a sítí s jednou skrytou vrstvou se čtyřmi neurony. Počet nalezených vzorů se blížil reálným hodnotám, ale nalezený vzor graficky neodpovídal vzoru DT. Výsledky prokázaly obtížnost nalezení hledaného vzoru.

CONCLUSION

The assessment of using of the artificial intelligence in financial markets was the aim of this thesis.

In the Russel 2000 index time diagram the model DoubleTop (DT) was searched for and several options of the training sets were created out of them. For the ascertainment of the right model the experience with the technical market analyses is needed. There were 6 types of the training sets created there. As the first type of the training set the real model was used, the other types were simplified and came out of the real model. In the second type of the training set the data between the top points were modified - differences between the data were decreased there. The comparison of the data in the middle of the model for the same value was provided there in the third type. In the fourth training set the strengthening of the data significance between the top points was provided. It came to the elimination of the edge values there in the fifth set in order to simplify the model maximally. On the contrary to it there were the edge values added to the fifth model in the sixth set. The best results were achieved in the fifth training set as the amount of the found models came close to the real condition. When the training set changing the top points must be preserved; failing it the training sets don't conform to the DT model.

The testing set of the data was created also from the diagram which was acquired through the Ninja Trader application.

By the help of these data the feedforward neural network was instructed. There were two instruction algorithms used there – Backpropagation and a Levenberg-Marquardt. The application of the Levenberg-Marquardt algorithm was better approved as it needs 45 iterations for the network instructing. Various network adjustments were tested. The network with 3, 2 and 1 hidden layers with various amounts of neurons in the hidden layers was used. It was found out the best network adjustment is with one hidden layer and 4 neurons.

The best results were acquired by the fifth training set and the network with one hidden layer of four neurons. The amount of the found models was close to the real values, but the found model didn't conform with the DT model graphically. The results proved its difficulty of finding a suitable model it was searched for.

SEZNAM POUŽITÉ LITERATURY

Knižní zdroje:

- [1] ZELINKA, Ivan. *Evoluční výpočetní techniky: principy a aplikace*. 1. vyd. Praha: BEN - technická literatura, 2009, 534 s. ISBN 978-80-7300-218-3.
- [2] FANTA, Jiří. *Technologie umělé inteligence na kapitálových trzích*. 1. vyd. Praha: Karolinum, 1999, 89 s. ISBN 80-718-4866-2.
- [3] FANTA, Jiří. *Technická analýza kapitálových trhů*. 1. vyd. Praha: Univerzita Karlova, 1996, 102 s. ISBN 80-718-4308-3.
- [4] MURPHY, J. *Technical Analysis of the Financial Markets: A Comprehensive Guide to Trading Methods and Applications*. New York Institute of Finance. 1999. 576 s. ISBN 978-0735200661.
- [5] ZELINKA, Ivan. *Umělá inteligence I: Neuronové sítě a genetické algoritmy*. 1. vyd. Brno: VUT v Brně, 1998, 126 s. ISBN 80-214-1163-5.
- [6] BOSE, N a Ping LIANG. *Neural network fundamentals with graphs, algorithms, and applications: Neuronové sítě a genetické algoritmy*. 1. vyd. New York: McGraw-Hill, 1996, 478 s. ISBN 00-700-6618-3.
- [7] NOVÁK, Mirko, Josef FABER a Olga KUFUDAKI. *Neuronové sítě a informační systémy živých organismů*. Praha. Grada. 1993. 265 s. ISBN 80-58424-95-9.
- [8] BÍLA, Jiří. *Umělá inteligence a neuronové sítě v aplikacích*. Vyd. 1. Praha: Vydavatelství ČVUT, 1995, 115 s. ISBN 80-010-1275-1.
- [9] VESELÁ, Jitka. *Investování na kapitálových trzích*. Vyd. 1. Praha: ASPI, 2007, 703 s. ISBN 978-80-7357-297-6.

Elektronické zdroje:

- [10] Wolfram. *Wolfram Research, makers of Mathematica, the only fully integrated technical computing software*. [online]. 2013 [cit. 2013-08-15]. Dostupné z: <http://www.wolfram.com/>.

- [11] NinjaTrader. *Ninja Trader stock, futures and forex charting software and online trading platform*. [online]. 2013 [cit. 2013-08-15]. Dostupné z: <http://www.ninjatrader.com/>.

SEZNAM POUŽITÝCH SYMBOLŮ A ZKRATEK

CBOT	Chicago Board of Trade – Chicagská burza.
CCI	Commodity Channel Index – indikátor.
CFTC	Commodity Futures Trading Commission – komise pro dohled nad obchodováním na burzách.
ČEZ	České energetické závody.
DT	Double Top – dvojitý vrchol.
EMA	Exponential Moving Average – exponenciální klouzavý průměr.
EUREX	Elektronická burza specializovaná na obchody s finančními deriváty a opcemi.
LIFFE	London International Financial Futures Exchange – Londýnská burza.
MA	Moving Average – klouzavý průměr.
MACD	Moving Average Convergence Divergence – indikátor.
MATP	Jednoduchý klouzavý průměr s periodou ve vztahu k typické ceně.
MDTP	Vypočtená střední odchylka od typické ceny.
NYBOT	New York Board of Trade – Newyorská burza.
RSI	Relative Strenght Index – indikátor.
TP	Typical Price – typická cena.

SEZNAM OBRÁZKŮ

Obr. 1 Graf Candlestick	14
Obr. 2 Graf Barchart	15
Obr. 3 Úsečky v grafu.....	16
Obr. 4 Graf Uptrend.....	16
Obr. 5 Graf Downtrend.....	17
Obr. 6 Graf Chop	17
Obr. 7 Exponenciální klouzavý průměr	19
Obr. 8 MACD indikátor.....	20
Obr. 9 RSI.....	21
Obr. 10 Commodity Channel Index.....	22
Obr. 11 Dvojitý vrchol.....	23
Obr. 12 Dvojité dno	23
Obr. 13 Umělý neuron	25
Obr. 14 Vícevrstvá síť.....	29
Obr. 15 Vrchol <i>Double Top</i>	34
Obr. 16 Úvodní stránka Ninja Traderu	34
Obr. 17 Textová podoba dat	35
Obr. 18 Validní vzor	36
Obr. 19 Validní vzor v Excelu	36
Obr. 20 Nevalidní vzor	37
Obr. 21 Nevalidní vzor v Excelu	37
Obr. 22 Zjednodušení vzoru	39
Obr. 23 Vzor varianta 1	39
Obr. 24 Vzor varianta 2	40
Obr. 25 Pozice nalezeného vzoru	44
Obr. 26 Vzor nalezený na pozici 15	44
Obr. 27 Nalezený vzor na pozici 16	45
Obr. 28 Nalezený vzor na pozici 49	45
Obr. 29 Nalezený vzor na pozici 50	46
Obr. 30 Nalezený vzor na pozici 51	46
Obr. 31 Úprava vzoru	47
Obr. 32 Vzor s vynechanými krajními hodnotami.....	47

Obr. 33 Nalezené pozice vzoru v Mathematice	48
Obr. 34 Vzor na pozici 160.....	48
Obr. 35 Vzor s přidáním krajními hodnotami.....	49

SEZNAM TABULEK

Tab. 1 Porovnání výsledků reálných vzorů a varianty 1	42
Tab. 2 Porovnání výsledků reálných vzorů a varianty 1	43
Tab. 3 Porovnání varianty 1 a varianty 2	43

SEZNAM PŘÍLOH

Příloha PI: Validní skeny.

Příloha PII: Nevalidní skeny.

Příloha PIII: CD-ROM obsahující vlastní práci.