



VYSOKÉ UČENÍ TECHNICKÉ V BRNĚ
BRNO UNIVERSITY OF TECHNOLOGY



FAKULTA INFORMAČNÍCH TECHNOLOGIÍ
ÚSTAV POČÍTAČOVÉ GRAFIKY A MULTIMÉDIÍ
FACULTY OF INFORMATION TECHNOLOGY
DEPARTMENT OF COMPUTER GRAPHICS AND MULTIMEDIA

ALGORITMICKÉ OBCHODOVÁNÍ NA BURZE S VYU- ŽITÍM UMĚLÝCH NEURONOVÝCH SÍTÍ

ALGORITHMIC TRADING USING ARTIFICIAL NEURAL NETWORKS

DIPLOMOVÁ PRÁCE

MASTER'S THESIS

AUTOR PRÁCE

AUTHOR

Bc. JAN ŠEDA

VEDOUCÍ PRÁCE

SUPERVISOR

Ing. IGOR SZÓKE, Ph.D.

BRNO 2016

Abstrakt

Určení vývoje ceny na světových trzích je aktuální problematikou, která v posledních dekádách nabývá na významu. Důležitou roli v tom sehraává rozvoj výpočetní techniky. V této práci je navržen mechanismus pro predikci budoucí ceny na trhu. Na základě toho je pak sestavena obchodní strategie. Jádrem predikčního systému používá pro svou činnost umělé neuronové sítě. Vstupem sítě jsou pak vybrané indikátory technické analýzy trhu. Obchodní systém byl implementován a úspěšně ověřen na historických datech.

Abstract

The capability to be able to determine the future progression on the world's stock exchange is an important issue, which has become discernible in the last decades. An important role of this progression lies within the fast advancements in computerized technology. Aforementioned document describes a mechanism used for prediction of the future price of a certain stock. The strategy of trading is built upon this mechanism, and the core of this prediction system is an artificial neural network. Inputs used in this network are indicators derived from technical analysis. This trading system was implemented into historical trades and successfully tested.

Klíčová slova

burza, umělé neuronové sítě, technická analýza, indikátory, predikce vývoje, časové řady

Keywords

stock exchange, artificial neural networks, technical analysis, indicators, prediction, time series

Citace

ŠEDA, Jan. *Algoritmické obchodování na burze s využitím umělých neuronových sítí*. Brno, 2016. Diplomová práce. Vysoké učení technické v Brně, Fakulta informačních technologií. Vedoucí práce Szóke Igor.

Algoritmické obchodování na burze s využitím umělých neuronových sítí

Prohlášení

Prohlašuji, že jsem tuto diplomovou práci vypracoval samostatně pod vedením pana Ing. Igora Szókeho Ph.D. Uvedl jsem všechny literární prameny a publikace, ze kterých jsem čerpal.

.....

Jan Šeda

27. května 2016

Poděkování

Velmi rád bych poděkoval svému konzultantovi Ing. Igoru Szókemu Ph.D. za jeho odborné vedení a ochotnou pomoc s problémy, které vyvstaly při řešení této práce.

© Jan Šeda, 2016.

Tato práce vznikla jako školní dílo na Vysokém učení technickém v Brně, Fakultě informačních technologií. Práce je chráněna autorským zákonem a její užití bez udělení oprávnění autorem je nezákonné, s výjimkou zákonem definovaných případů.

Obsah

1	Úvod	3
2	Burzy	4
2.1	Historie	4
2.2	Kde je možné obchodovat	4
2.3	S čím je možné obchodovat	5
2.4	Přístup na burzu	6
2.5	Způsoby obchodování	6
2.6	Reprezentace burzovních dat	6
2.7	Slovník pojmů	7
3	Technická analýza	9
3.1	Technická analýza není všelék	10
3.2	Indikátory	10
3.2.1	Dělení indikátorů	10
3.2.2	Jednoduchý klouzavý průměr	11
3.2.3	Exponenciální klouzavý průměr	12
3.2.4	MACD	13
3.2.5	CCI	15
3.2.6	RSI	16
3.3	Shrnutí	16
4	Neuronové sítě	18
4.1	Biologický neuron	18
4.2	Formální neuron	19
4.2.1	Aktivační funkce neuronu	20
4.3	Architektura neuronové sítě	20
4.4	Učení neuronových sítí	21
4.4.1	Učení bez učitele	22
4.4.2	Učení s učitelem	22
4.5	Vícevrstvá perceptronová síť	22
4.5.1	Architektura sítě	22
4.5.2	Vybavování ve vícevrstvé perceptronové síti	23
4.5.3	Učení ve vícevrstvé perceptronové síti	23

5 Praktická část	25
5.1 Burzy s digitalní měnou	25
5.2 Sběr dat	26
5.2.1 Návrh databáze	26
5.2.2 Komunikační rozhraní	28
5.2.3 Sběrač dat	29
5.3 Výpočet indikátorů a reprezentace dat	30
5.4 Predikční systém	32
5.4.1 Vstupy	33
5.4.2 Normalizace	33
5.4.3 Parametrizace sítě	34
5.4.4 Výstupy	34
5.4.5 Dosažené výsledky	34
5.4.6 Poznatky	35
5.5 Strategie	36
5.6 Použité nástroje	38
6 Závěr	40
Literatura	42
Přílohy	44
Seznam příloh	45
A Obsah CD	46
B Metriky	47
C Grafické vzory	48
C.0.1 Základní vzory	48
C.0.2 Pokročilé grafické vzory	50
D Časové rozložení sběračů	53
E Plakát	55

Kapitola 1

Úvod

Dokument popisuje tvorbu automatického obchodního systému. V práci jsou rozebrány teoretické základy, na niž je potom vystaven vlastní obchodní systém, jehož jádro tvoří umělé neuronové sítě.

Obchodování na burzách je velmi komplexní proces. Navíc jej každý obchodník pohybující se v tomto prostředí chápe odlišně. Cíl je ovšem pro všechny stejný a je jím zisk¹

Základní princip řídicí se úslovím "*levně koupit, draze prodat*", je velmi jednoduchý a všeobecně známý. Aplikace toho principu požaduje po obchodníkovi schopnost relativně přesně předpovědět budoucí cenový vývoj obchodované věci, neboli určit, jaký bude budoucí poměr nabídky a poptávky. Pro naplnění těchto myšlenek vznikla v moderním prostředí burz velká řada teorií a možných přístupů. Všechny tyto metody, ať už jakkoliv úspěšné, spojuje fakt, že žádná z nich nedokázala složité a velmi proměnlivé prostředí burz přesně popsat. Nikdo nedokáže se stoprocentní jistotou předpovědět budoucí vývoj. Toto tvrzení je platné, ať se jedná o jakkoliv vzdálenou budoucnost. Nedá se ani říci, že pro bližší budoucnost by existovala větší pravděpodobnost úspěšné předpovědi.

I přesto, že nelze přesně určit budoucí vývoj burzovních cen, existují přístupy, které se v historii ukázaly jako ziskové. Účelem práce je tyto přístupy najít, analyzovat, vylepšit a popřípadě spojit více přístupů dohromady, aplikovat a zhodnotit dosažené výsledky. Cílem není konstrukce přesného predikčního nástroje. Hlavní snahou bude vymyslet systém, ve kterém převáží zisky nad ztrátami. Budou brány na zřetel pouze ty přístupy, které je možné přenést do počítače. To ovšem neznamená, že by byly jediné možné.

Čtenář by měl mít na paměti, že žádný ze zmiňovaných přístupů nelze formálně verifikovat. Tyto metody se pouze jeví úspěšné vzhledem k jejich aplikaci v minulosti. Historické výsledky na poli burzovního obchodování ve své podstatě nic neříká o úspěšnosti v budoucnosti.

V kapitole 2 je blíže popsáno spleť prostředí burz. Dále jsou vysvětleny některé základní pojmy z finančního prostředí, jejichž znalost je nezbytná pro další výklad. Kapitola 3 o technické analýze trhu popisuje jeden ze způsobů obchodování. Vzhledem k tomu, že v této práci jsou použity umělé neuronové sítě, jsou nezbytné základy tohoto nástroje popsány v kapitole 4. Kapitola 5 popisuje navržený obchodní systém, který vychází z teoretické přípravy z předešlých kapitol. V této části jsou rovněž zaznamenány hlavní experimenty s obchodním modelem. Závěrečná kapitola 6 shrnuje dosažené výsledky a další možnosti vývoje.

¹Ať už je řeč o čistém finančním zisku, nebo zisku v podobě pořízení požadované aktiva za co nejnižší cenu.

Kapitola 2

Burzy

Tato kapitola popisuje burzovní prostředí, pro které je konstruován výsledný obchodní systém. Je tedy nezbytné správně pochopit hlavní principy burzy. V návaznosti na to je zaveden slovník pojmů používaný obchodníky. Informace poskytnuté v kapitole byly čerpány z [3] a [1].

2.1 Historie

Slovo burza pochází z latinského slova *burse*, které se překládá jako *kožený měšec*. V průběhu času se toto slovo stalo označením pro směnné prostředí. Důležitou roli hraje rodina belgických obchodníků *Van der Buerse*¹. Hostince v Bruggách provozované touto rodinou sloužily na konci 14. století jako místo, kde se potkávali obchodníci z celé Evropy a uzavírali zde obchody. Prostředí těchto hostinců se začalo označovat jako burzy.

Prvním oficiální burzou, na které došlo rovněž k první emisi obligací, se stala v roce 1531 instituce v Antverpách. Bylo zde možné nakupovat a prodávat směnky, zlaté a stříbrné mince. Postupem času začala vznikat další podobná místa. Roku 1745 vzniká londýnská burza a o něco později v roce 1792 vzniká slavná newyorská burza. Z jejich rozvojem vznikají i nové obchodní techniky a poprvé se objevuje přístup označovaný jako spekulace.

Rozvoji burz pomohl i technologický pokrok. Důležitý v této souvislosti je vynález a zavedení telefonu a telegrafu. Roku 1929 přichází 24. října černý čtvrtek a krach na newyorské burze, po kterém následovala celosvětová hospodářská krize. V poválečném období se začíná burzovní prostředí formovat do podoby, kterou známe dnes. Hlavní faktory ovlivňující tento vývoj jsou rozvoj výpočetní techniky a globalizace ekonomiky.

2.2 Kde je možné obchodovat

V současné době existuje všude po světě řada větších a menších burz, na kterých je možné obchodovat. Mezi nejznámější a největší burzy patří:

- **NYSE** New York Stock Exchange
- **NASDAQ** National Association of Securities Dealers Automated Quotations
- **LSE** London Stock Exchange

¹Rodinný erb obsahuje tři mešce.

- **TSE** Tokyo Stock Exchange

Na těchto burzách se pak prostřednictvím moderních technologií setkávají v jednu chvíli tisíce obchodníků z celého světa. S ohledem na tuto práci je dobré poznamenat, že jak malé tak velké burzy mohou být vhodné pro obchodování. Záleží hlavně na zvoleném přístupu k obchodování. Velké burzy s sebou nesou výhodu vysoké likvidity na druhou stranu jejich vývoj je zákonitě ovlivněn daleko širším spektrem faktorů. Výběr burzy je jednou z kritických částí, protože ne na všechny burzy se dá přistupovat prostřednictvím počítače.

2.3 S čím je možné obchodovat

Obchody se soustřeďují na tzv. podkladová aktiva nebo z nich odvozené finanční deriváty. Způsob obchodování je rozdílný. U podkladových aktiv v době nákupu dochází k reálnému vypořádání. V okamžiku koupě dojde k úhradě na základě smluvené ceny a my se stáváme majiteli. Mezi základní podkladová aktiva se řadí:

- **Akcie**
- **Akciové indexy**
- **ETFs** Exchange Trated Funds
- **Komodity**
- **Dluhopisy**

U derivátů patřících do skupiny termínovaných obchodů je situace odlišná. Nákupem derivátu dochází k uzavření obchodu a nastavení podmínek ceny a množství, v jakém bude obchod realizován. K reálnému vypořádání dojde až v budoucnu. Přesné datum dokončení celého obchodu je dopředu stanoveno. Známé a často obchodované deriváty jsou:

- **Futures**
- **Forwards**
- **Opce**
- **Spready**

Trochu bokem stojí mezinárodní obchodní systém Forex, který je určen pro směnu měnových párů. Střední kurzy tohoto trhu jsou považovány za oficiální kurzy pro dané měnové páry. Ceny jsou odvozeny od síly ekonomik, inflacemi a úrokovými sazbami². Zajímavosti o Forexu:

- Pro obchodování stačí účet fundovaný pouze 200\$. Vzhledem k tomu, že pro komoditní trhy se základní vstupní kapitál pohybuje kolem 5000\$, může být Forex vhodný pro první kroky začínajícího obchodníka.
- Forex funguje 24 hodin denně na rozdíl od ostatních.
- Je rovněž doporučován pro automatické obchodní systémy pro svou stabilitu. Například v publikaci [12].
- Je to nejlikvidnější trh s největším objemem obchodů.

²Zde jsem si vědom jisté nepřesnosti, protože Forex není burza. Vzhledem k povaze této práci se dá tato nepřesnost akceptovat.

2.4 Přístup na burzu

Na burzu není možné vstupovat přímo. Obchod musí provádět licencovaný subjekt. Vzhledem k cenám pořízení licence jsou obchody nejčastěji realizovány přes prostředníka, který je majitelem licence. Těmto prostředníkům se v obchodní praxi vžilo označení brooker. Výběr vhodného brookera je jedno z prvních vážných rozhodnutí, které musíme udělat pokud chceme vstoupit na burzu. Každý brooker si účtuje svoje služby podle jiných kritérií. Burzy a trhy, na kterých mohou jednotliví brookeři operovat se rovněž liší. Stejně tak je odlišný přenos příkazů mezi brookerem a jeho klientem. Příkazy je možné zadávat prostřednictvím počítače, telefonu nebo osobně.

2.5 Způsoby obchodování

V současné době existují dva hlavní způsoby obchodování na burzách. Vždy se jedná o celou filozofii, jak posuzovat aktuální vývoj cen. Klíčem pro klasifikaci je způsob rozhodování, zda nakoupit, nebo prodat. Jedná se o tyto přístupy:

- **Fundamentální** - Fundamentální obchodování je založeno na úsudku obchodníka, který posuzuje celkový kontext situace. Například obchodník posuzuje situaci v nějaké odvětví a na základě toho se rozhodne investovat do akcií určité firmy.
- **Technická analýza** - Technická analýza naproti tomu posuzuje nákup a prodej aktiva čistě na základě matematických vzorců. Blíže je problematika technické analýzy rozebrána v kapitole 3.

Třetím typem je pak kombinace dvou hlavních směrů.

Způsob obchodování lze rozlišovat ještě na základě doby, po kterou držíme dané aktivum. Pro účely této práce rozlišujeme pouze dvě kategorie. První je tzv. *intradenní* obchodování, kdy se nákup i prodej uskuteční v rámci jednoho obchodního dne. Aktivum tedy není drženo přes noc. Do druhé kategorie, označované jako *poziční*, budeme zařazovat vše ostatní.

2.6 Reprezentace burzovních dat

Základní jednotkou určující vývoj cen je jeden uzavřený obchod. O proběhlých obchodech jsou dostupné tyto informace:

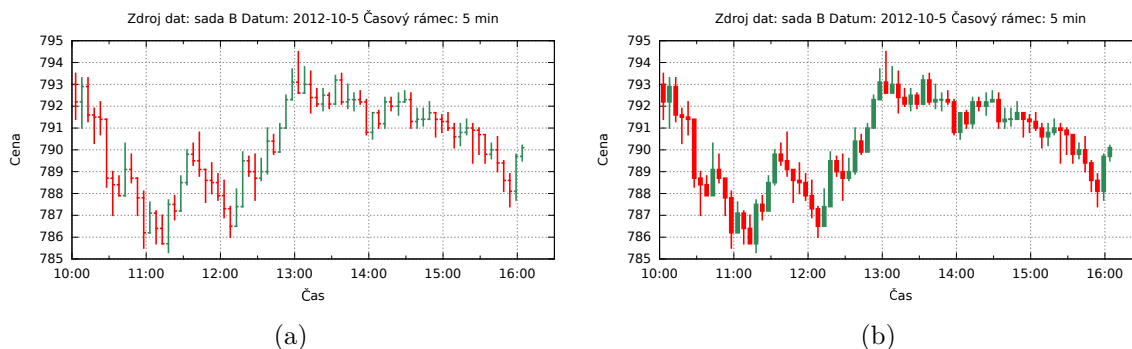
- **přesný čas** - kdy byl obchod uzavřen
- **cena** - za jakou cenu bylo aktivum koupeno (respektive prodáno)
- **objem** - s jakým množstvím bylo obchodováno

Z pozice dalšího hráče na trhu není tedy možné dohledat, kdo se obchodu účastnil a další z toho plynoucí informace.

Typicky jsou data zobrazována do grafu, jako např. na obrázku 2.1a, kde je použit tzv. sloupcový graf. Svíčkový graf na obrázku 2.1b je pouze jinou formou popsaná stejná situace.

Svíčka nebo sloupec pak zachycují časové období (rámce), v tomto případě pět minut a podávají čtyři základní informace. Pro tyto informace je zažité anglické označení:

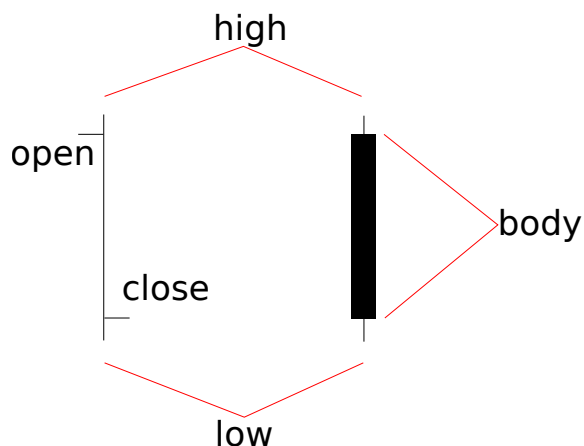
- **open** - otevírací cena časového rámce



Obrázek 2.1: Ukázka různých variant zobrazení burzovních dat

- **high** - nejvyšší cena ve sledovaném období
- **low** - nejnižší cena ve sledovaném období
- **close** - uzavírací cena ve sledovaném období

Popis sloupce a svíčky je na obrázku 2.2. Jak je vidět vyjadřovací síla je stejná. U sloupce je patrné zda došlo k poklesu nebo růstu ceny. V případě svíčky se tato situace rozlišuje barvou těla.



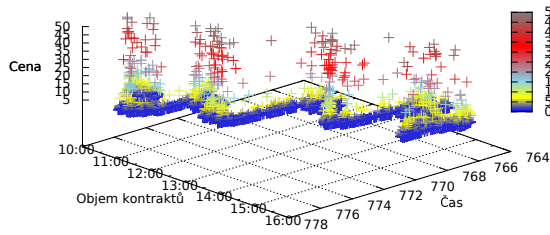
Obrázek 2.2: Popis sloupce a svíčky

Grafy na obrázcích 2.1b a 2.1a jsou nejtýpější formu zobrazení dat, kde je zachycen vývoj cen v závislosti na čase. Protože máme údaj o objemu obchodů, je možné všechna tato data zobrazit do jednoho 3D grafu, který poskytuje rozšířenou perspektivu na data. Možné 3D zobrazení zachycuje obrázek 2.3, který ukazuje stejnou situaci ze dvou různých úhlů.

2.7 Slovník pojmů

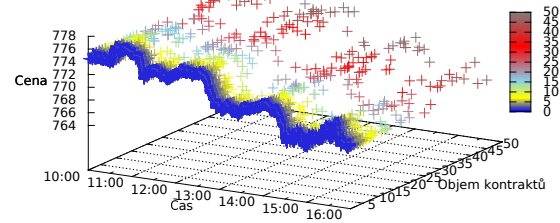
Slovník burzovních obchodníků je velice rozsáhlý. I když se většinou jedná o jednoduché věci nemusí být z názvu přímo patrné. Vzhledem k tomu, že se s těmito pojmy v omezené míře bude pracovat i v tomto textu, je zařazen slovník základních pojmenování a označení.

Zdroj dat: sada B Datum: 15.12.2012 Časový rámec: 5 min



(a)

Zdroj dat: sada B Datum: 15.12.2012 Časový rámec: 5 min



(b)

Obrázek 2.3: Zobrazení obchodů v závilosti na objemu

- **support** - nejnižší cena, za kterou jsou obchodníci ochotni aktivum prodat
- **resistance** - nejvyšší cena, za kterou jsou obchodníci ochotni aktivum koupit
- **býčí trh** - cena aktiva roste
- **medvědí trh** - cena aktiva klesá
- **trend** - dlouhodobý vývoj ceny jedním směrem
- **typická cena** - cena počítaná pro sloupec (respektive svíčku)

$$TP = \frac{high + low + close}{3} \quad (2.1)$$

- **dlouhá pozice** - situace, kdy se nakupuje aktivum s předpokladem, že jeho cena poroste
- **krátká pozice** - situace, kdy se aktivum prodává s předpokladem, že jeho cena bude padat. Důležité je že aktivum není potřeba v době prodeje vlastnit. Takovou operaci však není možno provést na všech trzích, pouze např. na komoditních trzích.
- volatilita - kolísavost trhu
- spread - cenový rozdíl mezi nabídkou a poptávkou
- překoupení a přeprodání - situace, ve kterých dochází k obratu trendu

Kapitola 3

Technická analýza

Jak bylo uvedeno v kapitole 2 jedním ze způsobů pro lepší pochopení aktuálního cenového dění na burzách je technická analýza. Tato oblast vychází z několika myšlenek, které byly definovány v knize [6]:

- Cena vyjadřuje vše. Není tedy nutné se zabývat studiem fundamentálních informací.
- Historie se opakuje. Cenový vývoj tvoří formace (respektive obrazce), které se v čase opakují. Tyto formace pak před svým vznikem vysílají jasné signály o svém vzniku, které se snažíme pomocí technické analýzy zachytit.
- Ceny mají tendenci se pohybovat v trendech. Cena může růst, klesat, nebo jde takzvaně do strany tzn. neroste ani neklesá. V momentě, kdy se vývoj v delším časovém úseku drží v jednom z těchto směrů, jde o tzv. trend. Předpokladem technické analýzy je, že trendy mají setrvačnou schopnost, kdy se snaží držet stále ve vytyčeném směru vývoje ceny.
- Obchodníka nezajímá, proč je cena taková, jaká je? Obchodníka zajímá pouze výše aktuální ceny a její vztah k minulým cenám.

O výše zmíně principy se pak opíráme při tvorbě nástrojů technické analýzy. Ty pak mají obchodníka informovat, jak se na trhu chovat. Zda je dobré otevřít pozici, uzavřít ji, nebo čekat. Vyhodou technické analýzy je, že ji lze bez rozdílu aplikovat na všechna aktiva i bez ohledu na časový rámec, přes který nazíráme na trh.

Technická analýza zkoumá trhy pomocí dvou hlavních nástrojů. Prvním jsou indikátory, které podle matematického vzorce transformují cenový vývoj většinou s ohledem na historii. Indikátory tvoří významnou část technické analýzy a mají i významnou roli pro tvorbu obchodního systému prezentovaného v této práci. Blíže jsou indikátory popsány v podkapitole 3.2.

Druhým důležitým nástrojem pro rozbor trhu, který používá technická analýza, jsou grafické vzory. Grafické cenové obrazce nejsou tak přímočaré pro použití v počítačích. Tento nástroj je rozebrán v pouze v rámci přílohy C, protože není součástí výsledného obchodního systému.

Co se týče technické analýzy je potřeba mít na paměti fakt, jež často vzpomínán ve všech publikacích zaměřených na toto téma. Obchodování pomocí technické analýzy je spíše uměním než vědou.

3.1 Technická analýza není všelék

Jako každý přístup i technická analýza má svá pro a proti. Některé problémy technické analýzy byly popsány v knize [6]. Shrnutí slabých míst:

- Historie se opakuje. Neopakuje se, ale úplně přesně, což vede k problému rozpoznání klíčových cenových formací.
- Cenové formace je možno interpretovat více způsoby.
- Vždy pouze část formací a parametrizací indikátorů funguje pro danou situaci na trhu.
- I když dostáváme z technické analýzy správné informace o vytvoření trendu, velmi často je dostáváme opožděně.
- Z hlediska obchodníka jako člověka vyvstává problém výběru správného nástroje z nepřeborné množiny.

3.2 Indikátory

Indikátory jsou jednou z hlavních částí technické analýzy. V historii bylo vytvořeno nepřeborné množství indikátorů, jejichž počet jde do tisíců. V této kapitole budou uvedeny povětšinou základní indikátory a pouze ty, s nimiž pracuje výsledný obchodní systém. Důvody pro použití základních indikátorů jsou dva.

Prvním důvodem je obecné tvrzení objevující se v mnoha knihách, např. [6] a [12]. Tvrdí, že se v obchodnické praxi nejvíce osvědčily ty nejjednodušší přístupy. Hledat pomocnou berličku v podobě složitých indikátorů se nevyplácí. S tím se dá souhlasit. Čím složitější indikátor, tím úžeji je ve výsledku zaměřen. Tím silněji jde proti pravidlu technické analýzy, které říká, že *cena je všeurčující*. Složitý indikátor potom důležitou část informace obsažené v ceně musí logicky zanedbávat a deformovat. Tím pádem je možné se domnívat, že jeho aplikace bude velmi omezená.

Druhým důvodem pro výběr těchto indikátorů je fakt, že jsou obecně mezi obchodníky považovány za vhodné a u každého indikátoru je uvedena práce, která dokládá jeho úspěšné použití na větším vzorku dat.

Čtenář by měl mít neustále na paměti, že i když situace vyobrazené v mnoha knihách ukazují perfektní aplikaci. Indikátory nejsou křišťálovou koulí. Skutečností je, i když možná těžko přijímanou, že indikátor pouze přepočítává aktuální cenu vzhledem k historii. Sám indikátor o budoucím vývoji neříká vůbec nic a poskytuje pouze jinou perspektivu dané situaci na trhu.

3.2.1 Dělení indikátorů

Indikátory je možné klasifikovat do několika kategorií. Základní znalost jednotlivých kategorií a jejich zástupců je velmi žádoucí. Vede k lepšímu pochopení určení aplikace indikátorů, a zabraňuje tak případným ztrátám. Nutno ovšem podotknout, že zde zmíněné dělení, byť vychází z odborné literatury [6] je spíše orientační, protože řada jiných zdrojů, např. [11], klasifikuje indikátory podle odlišného klíče.

- **Monetární indikátory.** Účelem je identifikovat investiční prostředí, např. na základě úrokové míry. Přinášejí informaci o bezpečnosti investování (respektive stabilitě instrumentu) a výši možných zisků.

- **Trendové indikátory.** Jak je patrné již z názvu tento typ indikátorů je určen pro rozpoznání dlouhodobých směrových pohybů. Do této kategorie spadají indikátory založené na klouzavých průměrech. Jmenovitě to pak jsou například *EMA* a *MACD*.
- **Netrendové indikátory.** Do netrendových indikátorů řadíme především tzv. oscilátory, které poskytují informace o tom, kdy dojde k obratu trendu. Typickým zástupcem je například *RSI*.
- **Indikátory volatility.** Popisují velikost cenových výchylek za určený časový úsek. Relevantní je u těchto indikátorů pouze velikost nikoliv směr změny. Indikátorem spadajícím do této kategorie je *Bollinger bands*.
- **Indikátory sentimentu.** Indikátory mapují očekávání investorů ještě dříve, než se stačí projevit v ceně. Logika zde použitá říká, že pokud všichni obchodníci věří, že cena bude růst, neexistují už žádní obchodníci, kteří by cenu mohli pozvednout.
- **Indikátory momenta.** V případě, že jde trh tzv. do strany, jinak řečeno nenachází se v žádném trendu, přichází na řadu indikátory momenta. Využívanou schopností je včasné upozornění na přeprodání nebo překoupení trhu.

Pro úplnost je třeba zmínit, že kniha [6] je zaměřena na obchodování na Forexu, ale nebude to chápáno jako překážka. Technickou analýzu je možné použít na všech trzích, přičemž i tato práce se bude orientovat především na Forex. ¹.

3.2.2 Jednoduchý klouzavý průměr

Tento indikátor je naprostým základem a je jako první zmiňován snad ve všech učebnicích zaměřených na technické obchodování. V anglických textech je označován jako *simple moving average*. Indikátor klouzavého průměru má mnoho mutací. Velmi podrobný přehled všech možných forem nabízí kniha [16] v kapitole *Moving Averages for Financial Data Smoothing*.

Jak již bylo zmíněno v části 3.2.1 tento indikátor spadá do kategorie takzvaných trendových indikátorů. Výpočet *SMA* se v této práci řídí obecně užívaným vzorcem 3.1, který je možno nalézt v práci [8] nebo [5].

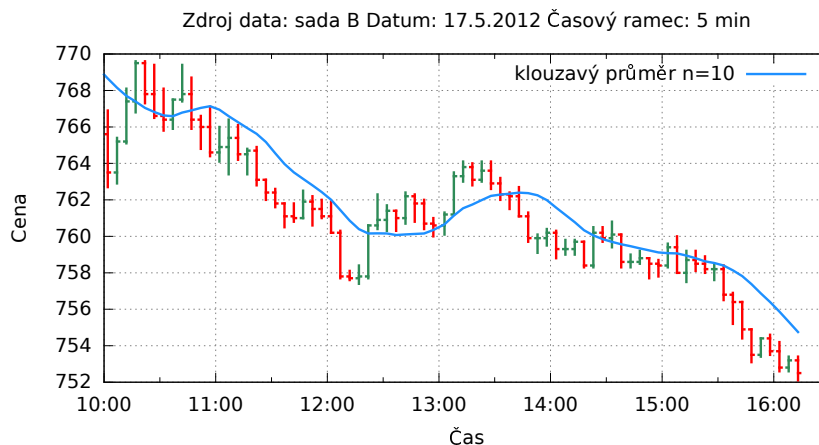
$$SMA_t(n) = \frac{P_{t-n} + P_{t-(n-1)} + \dots + P_t}{n} \quad (3.1)$$

Vstupní proměnná n ve vzorci 3.1 je perioda klouzavého průměru, která určuje, z kolika historických svíček bude průměr počítán. P_t proměnná značí cenu v čase t . Tuto cenu odečítáme ze svíček pro všechny prvky *SMA* vždy stejným ale libovolným způsobem. Vstupní cenou tedy můžou být ceny *low, high* atd.

Efekt aplikace indikátoru je patrný z obrázku 3.1. Obecně dojde k vyhlazení křivky. S narůstající periodou křivka silněji reprezentuje dlouhodobější trend. Na druhou stranu s narůstající periodou křivka slaběji a později reaguje na aktuální změny trendu.

Obchodování založené pouze na jednoduchých klouzavých průměrech pak ukazuje obrázek 3.2. Křížení průměrů značí změnu trendu. Směr nového trendu pak určuje zda průměr s menší periodou protne druhou křivku zespodu, nebo zeshora. Problémem u tohoto způsobu obchodování je nalezení vhodné parametrizace. V minulosti vznikly teorie pro správné zvolení periody. Navzdory mnoha teoriím se obchodování pouze na základě jednoduchých klouzavých průměrů v obchodní praxi moc neujalo.

¹Důvody k tomto rozhodnutí jsou popsány dále v textu.



Obrázek 3.1: Ukázka výpočtu klouzavého průměru



Obrázek 3.2: Ukázka obchodování pomocí klouzavého průměru

3.2.3 Exponenciální klouzavý průměr

Klouzavý exponenciální průměr (v angličtině *exponential moving average*) je další indikátor z rodiny klouzavých průměrů. Rovněž se jedná o trendový indikátor. Snahou je aktuálním cenám přiřadit větší váhu oproti starším. Tím se zrychuje reakce na aktuální změny trendu oproti jednoduchému klouzavému průměru. Pro účely této práce je využíván vzorec 3.2, který byl použit v práci [5], kde α vyjadřuje stupeň poklesu váhy pro starší ceny. Výpočet tohoto vyhlazovacího faktoru je $\alpha = \frac{2}{n+1}$. Proměnná P_t je pak, obdobně jako u jednoduchého klouzavého průměru, cena v čase t .

$$EMA_t(n) = \alpha \cdot P_t + (1 - \alpha) \cdot EMA_{t-1} \quad (3.2)$$

Ukázka výpočtu a porovnání EMA a SMA se stejným vstupním parametrem n je na obrázku 3.3, kde je dobře patrné, že EMA reaguje daleko pružněji na změnu trendu v porovnání s SMA . Způsob obchodování založený primárně na expoenciálním klouzavém

průměru by byl velmi podobný jako u jednoduchého klouzavého průměru. Vycházel by z křížení křivek.



Obrázek 3.3: Ukázka práce indikátoru exponenciálního klouzavého průměru

3.2.4 MACD

Jak bylo vidět u předchozích variant klouzavých průměrů při jejich správné kombinaci a parametrizaci jsou schopny, tyto indikátory generovat velmi spolehlivé signály. Způsob správné kombinace řeší právě indikátor *MACD*. Název pochází z anglického *Moving Average Convergence Divergence*. Jedná se o indikátor mezi obchodníky velmi oblíbený a v silně trendujících trzích i velmi úspěšný.

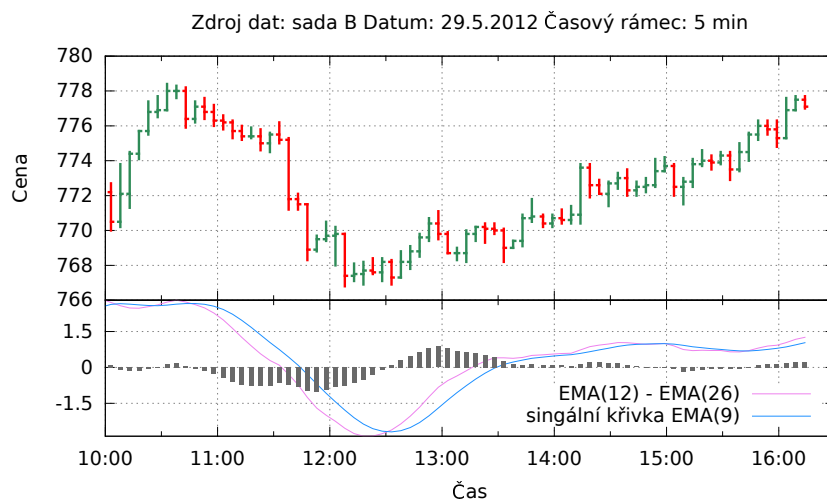
$$MACD_t = EMA_t(p_2) - EMA_t(p_1) \quad (3.3)$$

$$Signal = EMA_t^{MACD_t}(p_3) \quad (3.4)$$

Ze vzorce je patrné, že indikátor *MACD* je založen na trendových indikátorech což ho logicky rovněž řadí mezi trendové. Proměnné p_1, p_2, p_3 jsou periody exponenciálních průměrů. Parametrizace kombinací 26 – 12 – 9 je nejčastěji zmiňována v literatuře. Práce využívající tuto parametrizaci jsou například [4] nebo [17]. Jinou často užívanou parametrizací je 5 – 34 – 7.

Indikátor se zobrazuje do grafu v podobě dvou křivek a jednoho histogramu. První křivku, počítanou podle vzorce 3.3, tvoří rozdíl exponenciálních průměrů s periodami 12 a 26, druhá křivka je tzv. signální a tvoří jí typicky exponenciální průměr s periodou 9. Vstupními hodnotami ovšem nejsou ceny, ale hodnoty *MACD* počítané podle 3.3. Poslední zobrazovaný údaj u indikátoru *MACD* je histogram spočtený jako rozdíl první křivky a signální křivky. Původně byl *MACD* tvořen pouze první zmiňovanou křivkou. V roce 1986 byl dokonalem Thomas Aspray a doplněn o signální křivku. Způsob výpočtu byl převzat z knihy [4]. Názorná ukázka je na obrázku 3.4.

U tohoto indikátoru se na chvíli zastavíme a rozebereme ho trochu důkladněji. Důvodem je fakt, že indikátor *MACD* je použit i v rámci obchodního systému prezentovaného v této



Obrázek 3.4: Ukázka indikátoru *MACD*

práci. Používá se, protože oblíbený mezi obchodníky. Navíc byl úspěšně použit v kvalifikační práci [14] a [8]. V práci [8] se navíc při správném nastavení ukázal jako nejziskovější.

Překřížení

Základním pravidlem pro obchodování s indikátorem *MACD* je reagovat na překřížení *equilibria* (nulové linky) histogramem. Respektive být pozorný v situacích, kdy signální a *MACD* křivka změní vůči sobě pozice. Tato situace je dobře patrná na obrázku 3.4 v čase 11:00, kdy po dobře čitelném překřížení dochází k silnému propadu.

Překoupení a přeprodání trhu

Na překoupení nebo přeprodání trhu upozorňuje indikátor silným růstem *MACD* křivky v daném směru. Přesné podmínky pro překoupení a přeprodání trhu je nutné stanovit individuálně pro každé obchodované aktivum.

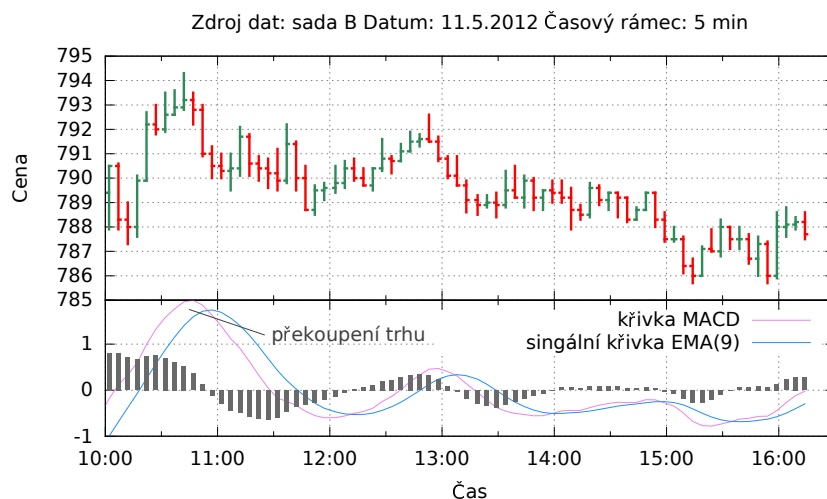
Divergence

Divergence nastává v situacích, kdy se *MACD* odchýlí od podkladového aktiva. Přesněji řečeno pokud podkladové aktivum vytvoří nové maximum (respektive minimum) a křivka *MACD* nezareaguje stejným způsobem. Tento jev nastává typicky při obratu trendu.

Rozbor nekorektního použití

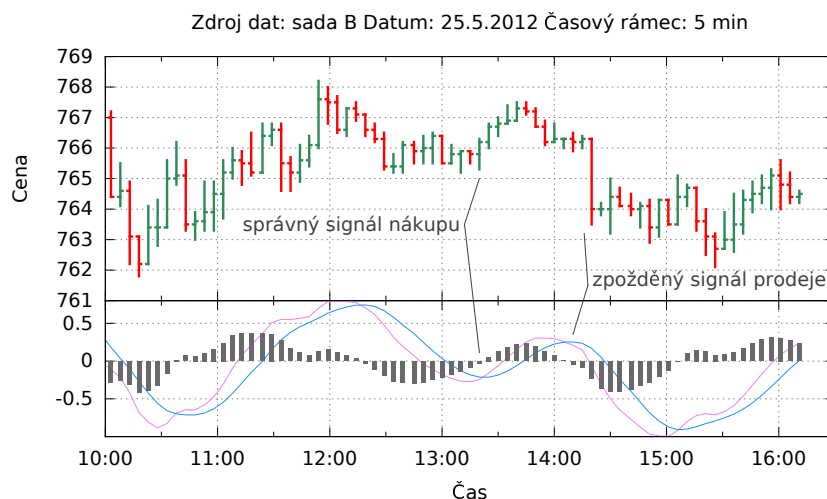
Pro dobré pochopení je potřeba rozumět i případům, kdy indikátor nedává správné signály. *MACD* je trendový indikátor, nebude tedy dobře fungovat, pokud jde graf tzv. do strany. Situace, kdy vývoj trhu silně netrenduje, je na obrázku 3.6². Je zde zachycená situace, při níž sice indikátor poskytuje správný vstup do pozice, ale signál výstupu přichází příliš pozdě. To by nejspíš vedlo k celkové ztrátě. Jak již bylo zmíněno *MACD* patří do skupiny trendových indikátorů. U trendu se obecně předpokládá, že mají schopnost vytrvat ve

²Legenda grafu je kvůli místu odstraněna záměrně. Křivky mají stejný význam jako v obrázku 3.4



Obrázek 3.5: Ukázka indikace překoupení/přeprodání trhu

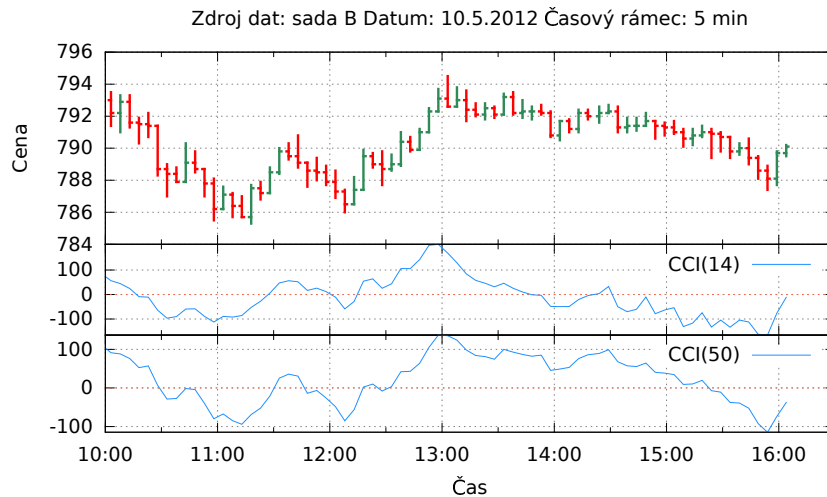
svém směru. Na tomto principu jsou konstruovány i trendové indikátory, které samozřejmě nemohou v netrendových dobách správně fungovat.



Obrázek 3.6: Ukázka špatné signalizace MACD

3.2.5 CCI

Název pochází z anglického *Commodity Channel Index*. I když má indikátor ve svém názvu *Commodity*, lze jej použít pro libovolné podkladové aktivum. Indikátor *CCI* podává informaci o tom, jak cena kolísá kolem svého statistického průměru. Pokud tedy hodnoty *CCI* rostou znamená to, že je cena vysoko nad svým dlouhodobým průměrem. Analogicky jsou interpretovány nízké hodnoty *CCI*. Překročení hranice 100 (respektive -100) je obecně chápáno jako signál pro možnou změnu trendu. Přesnou představu o výpočtu dokresluje vzorec 3.7, který se shoduje s principy výpočtů prezentovaných v publikacích [7] a [6].



Obrázek 3.7: Ukázka indikátoru CCI s různou parametrizací

$$TP = \frac{High + Low + Close}{3} \quad (3.5)$$

$$Mean\ Deviation = \frac{\sum_{i=0}^{n-1} |TP_{t-i} - SMA_n^{TP}|}{n} \quad (3.6)$$

$$CCI = \frac{TP - SMA_n^{TP}}{0.015 * Mean\ Deviation} \quad (3.7)$$

Parametrizace zvolená pro výpočet, který zachycuje obrázek 3.7, je převzata z publikace [12]. Rozhodnutí použít tento indikátor vychází z jeho úspěšného využívání v této knize.

3.2.6 RSI

Označení indikátoru pochází z anglického *Relative strength index*. Jde o indikátor ze skupiny oscilátorů. Rozpětí oscilačních hodnot je mezi 0 až 100. Výpočet probíhá podle vzorec z knihy [4]. Názorná ukázka je na obrázku 3.8. Obchodování s tímto indikátorem je založeno na hledání divergencí, kdy podkladové aktivum vytvoří nové maximum (respektive minimum) a křivka RSI nezareaguje stejným způsobem.

$$U = EMA_{close}^{gains}(p_1) \quad (3.8)$$

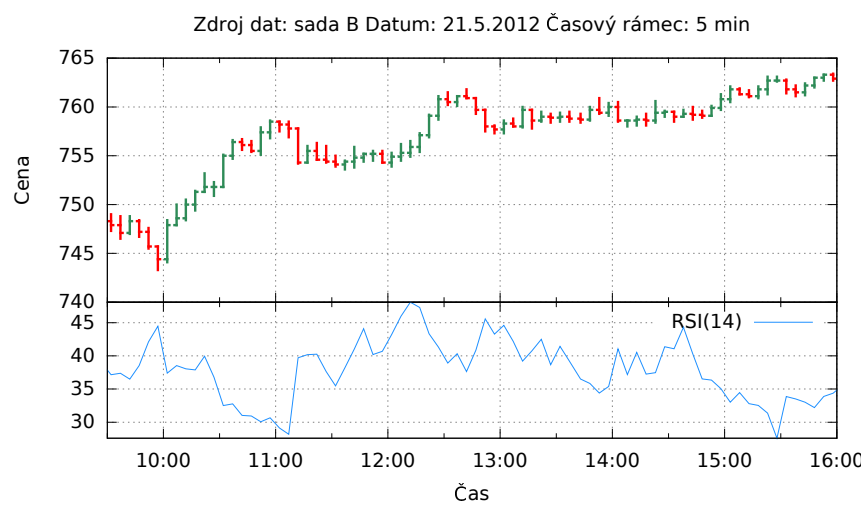
$$D = EMA_{close}^{losses}(p_1) \quad (3.9)$$

$$RS = \frac{U}{D} \quad (3.10)$$

$$RSI = 100 - \left(100 * \frac{1}{1 + RS}\right) \quad (3.11)$$

3.3 Shrnutí

V této kapitole byly představeny možnosti technické analýzy trhu. Byly zde vysvětleny základní myšlenky technické analýzy a popsány principy obchodování na základě jejich



Obrázek 3.8: Ukázka indikátoru RSI(14)

výstupů.

Kapitola 4

Neuronové sítě

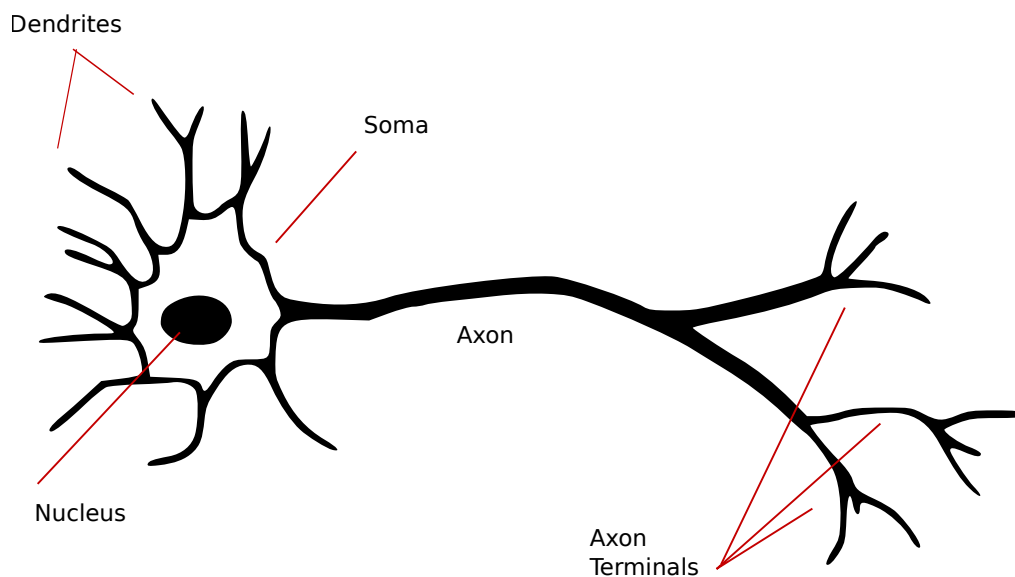
Tato kapitola vychází z informací uvedených v [18], [9] a [10]. Umělé neuronové sítě jsou výpočetním modelem pro strojové učení, které nacházejí využití především ve složitých úlohách, jejichž řešení nelze uspokojivě popsat algoritmem. U umělých neuronových sítí se rozlišují dvě pracovní fáze. První z nich je fáze *adaptivní*, během níž se síť učí. Druhou pracovní fází je fáze *aktivní*. Při ní síť vykonává naučenou činnost. To se obvykle označuje jako *vybavování*. Síti je předložen vstup, čímž ve vstupní vrstvě sítě vznikne nerovnovážený stav. Hodnoty uložené v síti se v důsledku toho začnou měnit (aktualizovat) až do chvíle, kdy síť dosáhne znovu stabilního rovnovážného stavu. Ve výstupní vrstvě sítě je pak k dispozici požadovaná odezva na poskytnutý vstup.

Typické oblasti použití neuronových sítí jsou rozpoznávání (klasifikace), predikce, expertní systémy, asociace, optimalizace a další. V této práci bude neuronová síť použita pro rozpoznávání. To lze charakterizovat jako rozhodování o tom, do které z předem známých kategorií zařadit předmět popsany vektorem hodnot, který slouží jako vstup sítě.

4.1 Biologický neuron

Neurony jsou z biologického hlediska základními stavebními prvky nervových soustav živých živočichů. Z pohledu fyziologie pojem neuron obecně představuje živou buňku, jejíž funkcí je sbírat, uchovávat, zpracovávat a přenášet informace. Schéma biologického neuronu je k vidění na obrázku 4.1. Výpočetním centrem neuronu je jeho tělo (soma). To zajišťuje zpracování elektrochemického signálu z dendritů, které slouží jako vstupní větve neuronu. Počet dendritů jednoho neuronu se pohybuje kolem desítky tisíc. Zpracovaná informace z neuronu vychází jediným výstupem (axonem), jenž může být bohatě větven. Výstupní signál, který je prostřednictvím axonu distribuovaný dále, odpovídá vstupům z dendritů zpracovaným uvnitř neuronu. Lze tedy říci, že výstup neuronu závisí na vstupech, které do něj skrze dendrity přicházejí.

Pro správnou funkčnost mozku je zapotřebí, aby neurony byly navzájem vhodným způsobem propojeny. Takto vzniká takzvaná neuronová síť. Propojení mezi neurony je realizováno prostřednictvím dendritů – ty se pomocí speciálních výběžků (synapsí) připojují na dendrity jiných neuronů, přímo na somu těchto neuronů nebo výjimečně přímo na axony. Zároveň jsou neurony v síti uspořádány tak, aby nefunkčnost jednoho neuronu neohrozila správné chování celku. Právě biologické neuronové sítě jsou inspirací pro sítě umělé snažící se jejich chování a vlastnosti napodobit.



Obrázek 4.1: Model biologického neuronu. Zdroj [13]

4.2 Formální neuron

Základem matematického modelu umělé neuronové sítě je formální neuron (dále jen neuron). Strukturu formálního neuronu schematicky znázorňuje obrázek 4.2.

Lze jej charakterizovat jako výpočetní jednotku s n vstupy, které modelují funkci dendritů u biologického neuronu. Ke každému ze vstupů přísluší i synaptická váha, která je udána obecně reálným číslem (v souladu s neurofyziologickým předobrazem může být i záporné). Tak jako biologický neuron, i formální neuron má pouze jediný výstup, jehož velikost je dána hodnotou funkce vážené sumy vstupů. Matematický zápis výpočtu výstupní hodnoty neuronu je ukázán v rovnici 4.1

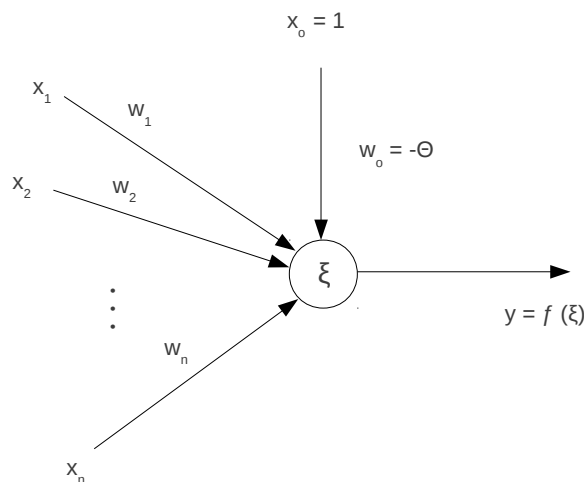
$$y = S\left(\sum_{i=1}^N w_i x_i + \Theta\right), \quad (4.1)$$

kde:

- x_i jsou vstupy neuronu (celkem je jich n)
- w_i jsou synaptické váhy jednotlivých vstupů
- S je přenosová funkce neuronu
- Θ je práh

Často se používá formální úprava, kde je práh neuronu nahrazen vahou w_0 dalšího formálního vstupu x_0 , který má konstantní jednotkovou hodnotu. Tato váha je označována jako tzv. *bias*. Výraz v závorce je označován jako vnitřní potenciál neuronu. Hodnota potenciálu je argumentem aktivační funkce neuronu, jejíž výsledek je distribuován na výstup. Dle povahy vstupních dat lze neurony dělit na:

- **neurony binární**, které zpracovávají dvouhodnotová vstupní data



Obrázek 4.2: Model formálního neuronu. Zdroj: autor dle [10]

- **neurony vícehodnotové (spojité)**, které zpracovávají vícebitovou vstupní informaci

4.2.1 Aktivační funkce neuronu

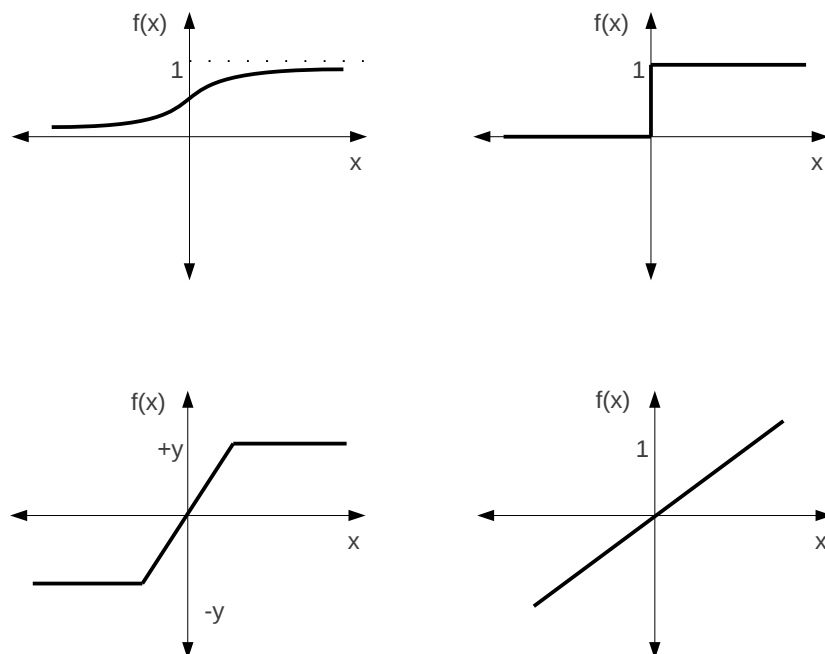
Výstup neuronu je dán hodnotou jeho aktivační funkce v bodě vnitřního potenciálu neuronu (viz výše). O aktivační (někdy označované i přenosové) funkci neuronu pak lze tvrdit, že převádí vstup (vnitřní potenciál) do předem definovaného oboru výstupních hodnot. Ten je většinou (ale ne vždy) vymezen intervalem $[0,1]$. Nejčastěji se používají následující čtyři přenosové funkce:

- lineární
- skoková aktivační funkce
- omezená
- sigmoida

Grafy těchto funkcí jsou znázorněny na obrázku 4.3. Aktivační funkce určuje, zda je neuron binární (skoková aktivační funkce), nebo spojitý (lineární, omezená, sigmoida).

4.3 Architektura neuronové sítě

Samotný neuron je dostačující pouze pro nejjednodušší úlohy. Pro řešení složitějších úkolů je potřeba uspořádat neurony do větších celků – *neuronových sítí*. Počet neuronů v síti a způsob, jakým jsou mezi sebou tyto neurony propojeny se označuje jako architektura (topologie) sítě. Ta je obvykle znázorňována grafem, jehož uzly odpovídají neuronům a hrany spojům mezi nimi. V rámci neuronové sítě lze z hlediska funkce a využití rozlišovat:



Obrázek 4.3: Grafy nejpoužívanějších aktivačních funkcí. Zdroj: autor dle [9]

- neurony vstupní
- neurony skryté (pracovní)
- neurony výstupní

4.4 Učení neuronových sítí

Před použitím neuronové sítě pro řešení daného problému je nutné, aby se síť pro tuto úlohu adaptovala. Tento proces se nazývá učení neuronové sítě a obvykle je realizován nastavováním vah mezi uzly sítě. Nezbytnou podmínkou k započatí učení je existence vzorku dat, který dostatečně přesně popisuje všechny platné vstupy sítě. Tato data se označují pojmem *trénovací množina*. Před zahájením učení je potřeba provést inicializaci hodnot jednotlivých vah sítě. V praxi jsou vahám při inicializaci přiřazeny buďto náhodné hodnoty, nebo hodnoty vhodně zvolené podle podobného případu. Během samotného procesu učení jsou síti postupně předkládány vzory z trénovací množiny, na něž síť poskytuje odezvu (výstup) a podle zvoleného učicího algoritmu jsou nastavovány váhy. Popsaný proces se iterativně opakuje dokud není dosaženo kritéria pro ukončení učení. Jedna iterace, při níž jsou neuronové síti předloženy všechny vzory z trénovací množiny, se nazývá *epocha*. Existují dva typy učení neuronové sítě: s učitelem a bez učitele.

4.4.1 Učení bez učitele

Při učení bez učitele není k dispozici žádné vnější kritérium správnosti výstupu. Síť se během procesu učení adaptuje tak, aby dokázala vzory z trénovací množiny sama třídit dle jejich podobnosti. Na základě této charakteristiky je učení bez učitele označováno jako *samoorganizace*. Mezi metody pro učení bez učitele patří například Hebbovské učení, kompetice nebo min-max učení.

4.4.2 Učení s učitelem

Zatímco pro učení bez učitele jsou zapotřebí pouze vektory vstupních hodnot, v nichž síť hledá podobnosti, učení s učitelem vyžaduje i existenci vektoru, který popisuje požadovaný výstup. S jeho využitím se v síti během učení nastavují váhy v závislosti na tom, jak se výstup poskytnutý neuronovou sítí liší od požadovaného výstupu daného vektorem vstupních hodnot. Pro učení s učitelem se používají algoritmy, které zaručují postupné snižování chyby – rozdílu mezi skutečným a požadovaným výstupem. Zřejmě nejznámější metodou pro učení s učitelem je učení algoritmem zpětného šíření. Mezi další metody patří: posilované učení, učení s korekcí chyby a další.

4.5 Vícevrstvá perceptronová síť

Zřejmě nejznámějším a nejpoužívanějším typem neuronové sítě je vícevrstvá perceptronová síť. Její základní prvek, perceptron, byl popsán F. Rosenblattem v roce 1957. Avšak algoritmus jejího učení, metoda zpětného šíření, byl popsán až v roce 1986 v [15]. S jeho využitím lze s pomocí sítě tří neuronů vyřešit problém XOR. Vícevrstvá perceptronová síť využívá učení s učitelem a pro její trénování je tedy potřeba znát požadovaný výstup. Aktivační funkcí sítě je sigmoida. Hodnota výstupu jednoho perceptronu v síti je tedy dána rovnicí

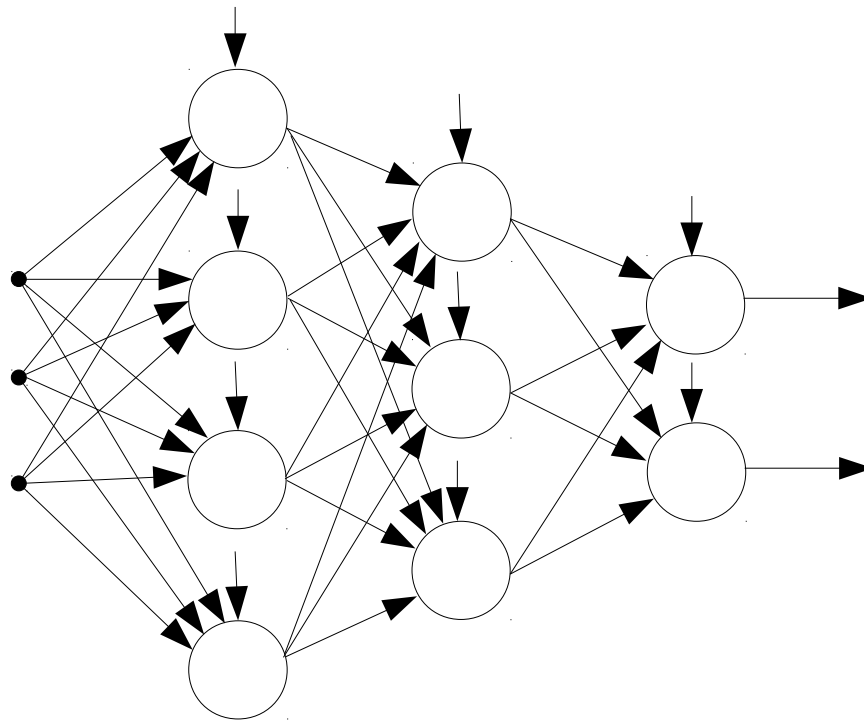
$$f(\xi) = \frac{1}{1 + e^{-\lambda\xi}} \quad (4.2)$$

kde λ je *parametr strmosti* určující míru změny hodnoty sigmoidy v okolí počátku a ξ je vnitřní potenciál perceptronu.

4.5.1 Architektura sítě

Vícevrstvá perceptronová síť je složena z několika vrstev neuronů. Jedná se o síť dopřednou – výstupy neuronů jsou distribuovány do všech neuronů následující vrstvy a tvoří tak *úplný bipartitní graf*. Na základě popsaného dělení neuronů v neuronové síti (viz 4.3) lze potom o architektuře vícevrstvé perceptronové sítě tvrdit, že udává počet vstupů, množství skrytých vrstev sítě a neuronů v nich a počet výstupů sítě. Ten se rovná počtu klasifikovaných tříd. Graf znázorňující architekturu třívrstvé sítě se třemi vstupy a dvěma výstupy je na obrázku 4.4.

Volba počtu neuronů v jednotlivých vrstvách sítě je netriviálním problémem. Při malém počtu neuronů dochází ke špatnému rozpoznávání předkládaných vzorů. Příliš velký počet neuronů na druhou stranu prodlužuje dobu učení a při nevhodně zvolené velikosti trénovací množiny obvykle vede ke zhoršení schopnosti generalizace. Tento jev je označován jako *přeučení* (overfitting). V praxi je potřeba volit počet vrstev sítě a počet neuronů v každé z nich v závislosti na charakteru úlohy, kterou má neuronová síť řešit.



Obrázek 4.4: Graf znázorňující architekturu třívrstvé perceptronové sítě se třemi vstupy, čtyřmi neurony v první skryté vrstvě, třemi neurony ve druhé skryté vrstvě a dvěma výstupy. Zdroj: autor dle [10]

4.5.2 Vybavování ve vícevrstvé perceptronové síti

Ve fázi vybavování jsou neuronům vstupní vrstvy předány hodnoty vstupního vektoru reprezentujícího daný vzor. Odezva na hodnoty předané vstupní vrstvě se pak přes váhy neuronů v jednotlivých vrstvách s využitím aktivační funkce v síti dopředně šíří až k výstupní vrstvě. Získané hodnoty výstupních neuronů odpovídají požadované odezvě na vstup a tvoří výstupní vektor. Obecně jsou výstupní hodnoty reálnými čísly a pokud tedy mají odpovídat jednotlivým třídám při klasifikační úloze, je nutné je vhodně kvantifikovat (například prahováním).

4.5.3 Učení ve vícevrstvé perceptronové síti

Učení sítě obecně je procesem, při kterém se hledá takové nastavení synaptických vah, aby odchylka (chyba) mezi poskytnutými a požadovanými vstupy sítě byla pro zvolenou trénovací množinu minimální. Chyba sítě je definována rovnicí

$$E = \sum_k E_k, \quad (4.3)$$

kde k je indexem k -tého vzoru v trénovací množině a E_k je chyba, která tomuto vzoru odpovídá. Pro její výpočet je použit vztah

$$E_k = \frac{1}{2} \sum_j (y_j - d_{kj}), \quad (4.4)$$

kde j odpovídá j -tému neuronu výstupní vrstvy a d_{kj} je j -tá hodnota požadovaného výstupního vektoru pro k -tý vzor.

Před zahájením učení jsou váhy včetně biasů inicializovány na malé náhodné hodnoty. Na základě heuristik se doporučuje, aby hodnoty spadaly např. do intervalu $\langle -\frac{2}{n}, \frac{2}{n} \rangle$, kde n určuje počet vstupů neuronu, pro něž jsou váhy nastavovány.

Po inicializaci jsou síti předkládány jednotlivé vzory z trénovací množiny a pro každý z nich se vypočítá chyba dle 4.4. Na konci každé epochy je pak vypočtena celková chyba ze vztahu 4.3 a podle ní jsou upraveny hodnoty vah v síti. Úprava hodnot je provedena na základě vzorce:

$$w_{ij}(t+1) = w_{ij}(t) + \Delta w_{ij}, \quad (4.5)$$

kde změny vah jsou vypočteny z rovnice

$$\Delta w_{ij} = -\eta \frac{\partial E}{\partial w_{ij}} \quad (4.6)$$

η se nazývá parametr učení a je volen v intervalu $\langle 0, 1 \rangle$. Výpočet $\frac{\partial E}{\partial w_{ij}}$ je předmětem algoritmu zpětného šíření chyby (*backpropagation*) a je odvozen v [10]. Minimalizace chyby probíhá gradientní metodou a hrozí tak její uváznutí v lokálním minimu.

Kapitola 5

Praktická část

V této kapitole je rozebrán vlastní obchodní systém. Je zde popsána konstrukce systému a výsledky simulace na historických datech. Jádrem celého systému je neuronová síť, která se stará o předpověď budoucího vývoje na burze. Obchodní systém využívající služeb predikčního systému následně určí, na základě kterých předpovědí vstoupit do pozice a vystoupit z ní. Systém a jeho součásti byly navrženy pro reálné použití na trzích s digitálními měnami. V následujících podkapitolách jsou popsány významné části systému.

5.1 Burzy s digitální měnou

Vzhledem ke snaze nasadit systém v reálném prostředí, bylo potřeba rozhodnout, kde bude systém nasazen. Toto rozhodnutí je poměrně zásadní, protože se od něj odvíjí vše, počínaje sběrem dat a konče sestavením obchodní strategie. Možností je celá řada, viz rozdělení burz v teoretické části. Každá možnost má své klady a zápory. Po zvážení všech možností byly zvoleny burzy s digitální měnou, kde je hlavním obchodním artiklem bitcoin. Důvodem jsou nízké bariéry vstupu a to jak finanční, tak technické. Pro představu o obchodování prostřednictvím významného světového brokera Interactive Brokers, který navíc poskytuje vlastní API, je aktuální výše minimálního požadovaného zůstatku na účtu 10 000 USD. Naproti tomu na trzích s digitální měnou tyto limity nejsou a jediným limitem je stanovení minimální obchodovatelné jednotky. Například na burze Kraken je toto minimum nastaveno na 0.01 podkladové měny. V případě bitcoinu to při aktuálních cenách znamená 4.4 USD. Druhá zmíněná bariéra je čistě technická a týká se aplikačního uživatelského rozhraní brokera nebo burzy. V mnoha případech aplikační uživatelské rozhraní není dostupné, nebo je značně komplikované. Výjimku právě tvoří burzy s digitální měnou, které v drtivé většině případů mají svá aplikační rozhraní, která jsou stavěna nad HTTPS protokolem. Dokumentace a stabilita těchto rozhraní jsou na dobré úrovni. Pro praktickou část byly vybrány následující burzy:

- Bitfinex
- Bitstamp
- Btcc
- Btce
- Kraken
- Okcoin

Procesor	Paměť	Diskový prostor	Operační systém
2.4 GHz	2GB	20 GB	Debian 8

Tabulka 5.1: Popis server

5.2 Sběr dat

První problém, který bylo třeba vyřešit, je ukládání dat z burz. Tato data jsou pro funkci celého systému stěžejním bodem. Informace, které je možné získat, jsou:

- *detaily aktuálně uzavřených obchodů*
Přesně se jedná o cenu, množství, typ a čas, ve kterém obchod nastal. Typ udává, kdo celou transakci inicioval ve smyslu nabízející a poptávající.
- *aktuální stav knihy příkazů* Zde jsou nabídky obchodníků jak ze strany nabídky tak poptávky, které čekají na svůj odpovídající protějšek.

K perzistentnímu ukládání burzovních dat bylo možno přistoupit různě. Na začátku systém ukládal data do textových souborů. Přestože bylo řešení jednoduché a přímočaré, záhy se ukázalo, že je málo flexibilní, a pro ukládání dat byla zvolena databáze.

Při návrhu databáze a sběračů burzovních dat bylo potřeba řešit následující otázky:

- Jak se vypořádat s nesourodým rozhraním jednotlivých burz?
- Jak navrhnout sběrač?
- Jak zajistit co největší konzistenci dat?
- Jak optimalizovat funkci sběrače s ohledem na zatížení serveru z pohledu výpočetního výkonu?
- Jak optimalizovat funkci sběrače s ohledem na zatížení serveru z pohledu síťového přenosu?
- Jak optimalizovat funkci sběrače s ohledem na zatížení databáze?
- Jak ukládat data s ohledem na velikost diskového prostoru serveru?
- Jak organizovat data pocházející z více burz pro různé měnové páry?
- Jak zajistit aby data neunikala?

Řešení každé z těchto otázek je věnován následující text. Pro doplnění uvádím, že celý systém běží na zpoplatněném soukromém serveru. Technický popis severu je v tabulce 5.1.

5.2.1 Návrh databáze

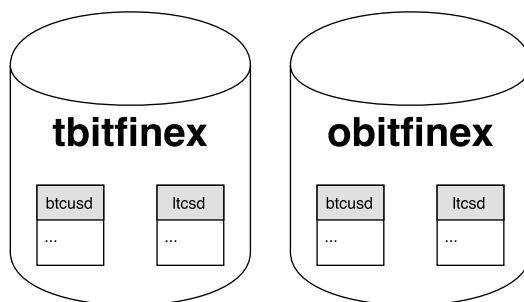
Při návrhu organizace a struktury databáze bylo potřeba rozhodnout následující problémy:

- Budou se data ukládat agregována nebo s maximální možnou granularitou?
- Jaká má být struktura tabulek a jejich příslušnost do databází?

- Jak budou řešena přístupová práva, aby byla zajištěna co největší bezpečnost a konzistence dat?
- Budou se ukládat údaje, které se dají dopočítat?

V otázce, zda ukládat data v surové nebo agregované podobě, bylo rozhodnuto, že je třeba záznamy z burzy uchovávat v původní podobě. Agregací totiž dochází ke ztrátě informací, který by se mohly v budoucnu ukázat jako klíčové. Pravidlo má jednu výjimku, a to ukládání informací o obchodech na čínských burzách, kde je celkový objem obchodů řádově vyšší než ve zbytku světa. Během sekundy zde proběhne více jak 10 obchodů. Vzhledem k tomu, že jsou tyto obchody s pohledu ceny prakticky stejné, jsou zde data agregována na úrovni jedné vteřiny. Tím se ušetří diskový prostor a granularita zůstane zachována. Výpočet ceny v jedné takovéto agregované vteřině je průměrem všech cen a objem obchodů je sumou.

Problémy organizace databáze a řešení přístupových práv jsou velmi úzce provázány. Vzhledem k tomu, že je potřeba uchovávat údaje o obchodech a otisky příkazových knih z šesti burz, kde na každé z nich existují různé měnové páry, byla zvolena struktura, zachycená na obrázku 5.1 a vysvětlena na příkladu burzy Bitfinex. Pro uchovávání dat z jedné burzy jsou vytvořeny dvě databáze, které mají stejnou příponu obsahující jméno burzy – tomto případě *Bitfinex*. Předponou, která obě databáze odlišuje, je pro obchody *t* (z anglického "trades") a pro příkazy *o* (z anglického spojení "orderbook"). Tabulky potom nesou název podle měnového páru.



Obrázek 5.1: Databáze uchovávající data z burzy Bitfinex

Tabulka pro ukládání obchodů obsahuje identifikátor, který je předáván přímo burzou, dále pak jednotkovou cenu, za kterou bylo prodáno uvedené množství. Typ obchodu udává v jakém směru došlo ke spárování nabídky a poptávky. Časové razítko je pak čas uskutečněný obchodu ve standardním *Epoch time* formátu, který se udává v počtu vteřin od 1.1. 1970. Schéma tabulky je na obrázku 5.2.

Obrázek 5.3 přibližuje tabulku zachycující knihu příkazů. Ve skutečnosti databázová tabulka neobsahuje pole, ale každá buňka pole je expandována na dvojici (*cena, množství*). Potom jeden řádek tabulky popisuje otisky knihy příkazů v daném čase. V každém otisku je zachyceno dvacet nespárovaných příkazů ze strany nabídky a dvacet ze strany poptávky.

Pro přístup k jednotlivým databázím existují vždy dva uživatelé. Jeden s právy, umožňujícími zapisovat do tabulky, a druhý uživatel, který je oprávněn pouze číst. Přístupové údaje uživatele oprávněného k zápis má pak pouze jeden jediný konkrétní sběrač.

Vzhledem k omezeným diskovým možnostem serveru bylo rozhodnuto, že všechny údaje, které je možné dopočítat, nebudou do databáze ukládány. Důvodem je zřejmá snaha šetřit

btcsud
id
cena
množství
typ
časové razítko

Obrázek 5.2: Schéma databazové tabulky pro ukládání obchodů

btcsud
id
časové razítko
poptavka[20]
nabídka[20]

Obrázek 5.3: Schéma databazové tabulky pro snímek knihy příkazů

paměťovým prostorem. Výpočetní výkon severu je dostatečný, aby si poradil s dopočítáním údajů, jako je například klouzavý průměr atp.

5.2.2 Komunikační rozhraní

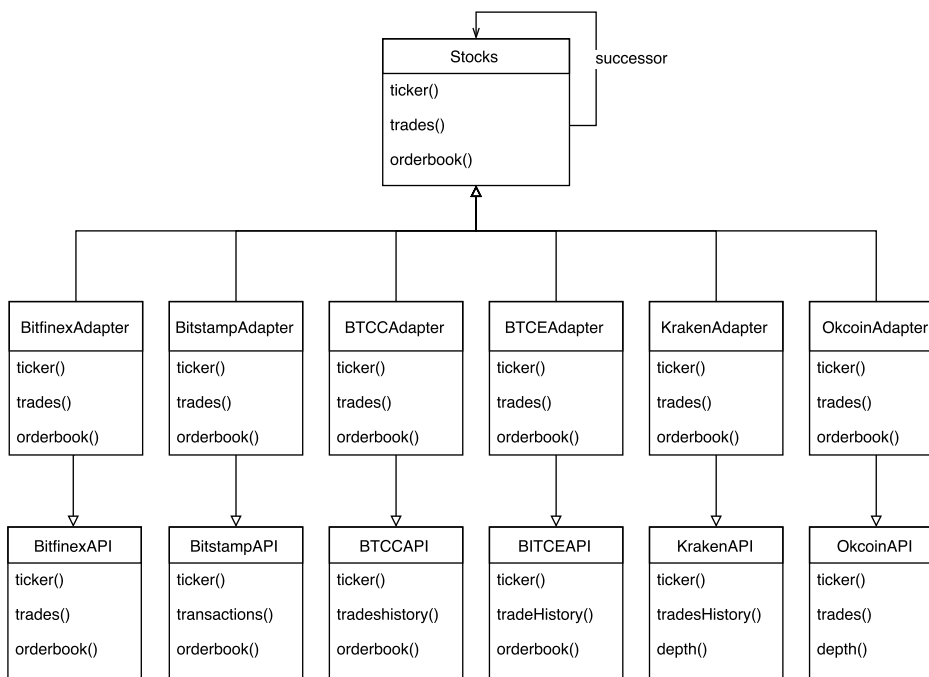
Interakce obchodního systému a burzy zajišťuje komunikační rozhraní. Tvorba komunikačního rozhraní byla rozdělena do tří fází:

1. implementace aplikačního rozhraní pro každou burzu zvlášť,
 - (a) veřejné rozhraní
 - (b) privátní rozhraní
2. unifikace aplikačního rozhraní,
3. spojení sjednocených aplikačních rozhraní do jednoho celku.

První fáze, při které bylo implementováno rozhraní vybraných burz, může být rozdělena na dva další bloky. První se týká implementace veřejného rozhraní obsahující takové funkce, které umožňují zjišťovat aktuální cenu, obchody uskutečněné v nedávné minulosti a číst knihu příkazů čekajících na zpracování. Veřejné rozhraní bylo implementováno pro všechny vybrané burzy. Oproti tomu privátní rozhraní poskytuje přístup ke konkrétním účtům a umožňuje se dotazovat na zůstatek, otevřené pozice, účetní knihu a především umožňuje zadávat příkazy na nákup a prodej. Implementace tohoto rozhraní vyžaduje správné pochopení podpisových mechanismů, který byl na každé burze nastaven jinak. Privátní rozhraní bylo naprogramováno pro Bitfinex, Btce, Kraken a Okcoin. Pro zbylé burzy implementováno nebylo z důvodu, že toto rozhraní není stěžejní z pohledu zadání diplomové práce. Pro úplnost je nutno zmínit, že uživatelské aplikační rozhraní je postaveno nad protokolem HTTP, respektive jeho šifrované podobě.

Úkolem druhé fáze bylo sjednotit různorodá veřejná rozhraní všech burz. Nejednotnost rozhraní nespočívala pouze v jiném pojmenování funkcí, ale především v rozdílném chování a formátu odpovědí. Pro sjednocení byl použit návrhový vzor *Adaptér*. Jeho aplikace je zachycena na obrázku 5.4.

Třetí fáze byla orientovaná na spojení unifikovaných rozhraní do jednoho celku. Výhodou tohoto spojení je možnost parametricky udávat nejenom měnový pár, nad kterým má být operace provedena, ale i burzu. To má pozitivní dopad na konstrukci sběrače burzovních dat, který je díky tomu mnohem jednodušší. Pro dosažení tohoto cíle byl použit návrhový vzor *Chain of responsibility*. Výsledný objektový návrh je na obrázku 5.4.



Obrázek 5.4: Struktura modulu pro komunikaci s burzami

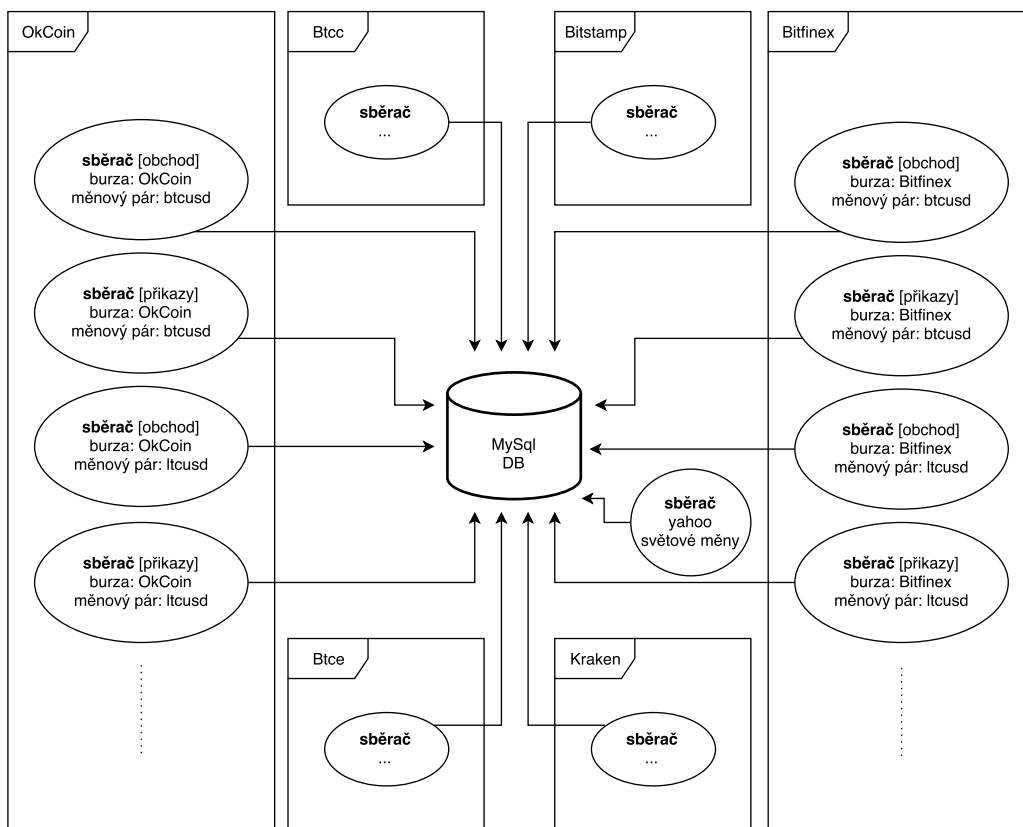
5.2.3 Sběrač dat

Sběrač dat je koncipován jako samostatně běžící proces, který sbírá data pro vybranou dvojici burza a měnový pár. Tím, že každá dvojice, má svůj vlastní proces pro záznam informací, je docíleno zvýšení bezpečnosti a stability celého systému, protože při výpadku jednoho z procesů jsou dopady minimální a zbylé sběrače fungují dále. Přestože se to může zdát jako nedůležité, praxe ukázala tento přístup jako velmi vhodný. K výpadkům nejčastěji dochází při komunikaci s čínskými servery. Příčiny výpadků si můžeme pouze domýšlet. Větší frekvence výpadků čínských serverů může být způsobena:

- velká vytíženost,
- čínský firewall,
- větší síťovou vzdáleností.

V systému existují dva hlavní typy sběračů a jeden vedlejší. Hlavními jsou sběrač pro ukládání údajů o proběhlých obchodech a pro snímání otisků příkazových knih. Vedlejším

typem je pak samostatný sběrač pro uchovávání vývoje cen světových měn. Tyto informace mohou nabídnout možnost srovnání pro hledání korelací ve vývoji. Schéma organizace sběračů je zachyceno na obrázku 5.5.



Obrázek 5.5: Schéma organizace sběračů data

Posledním problémem je časové rozložení sběračů, respektive periody, kdy sběrače čtou údaje o nových obchodech tak, aby byla optimalizována zátěž procesru, sítě i samotné databáze. To znamená zajistit, aby:

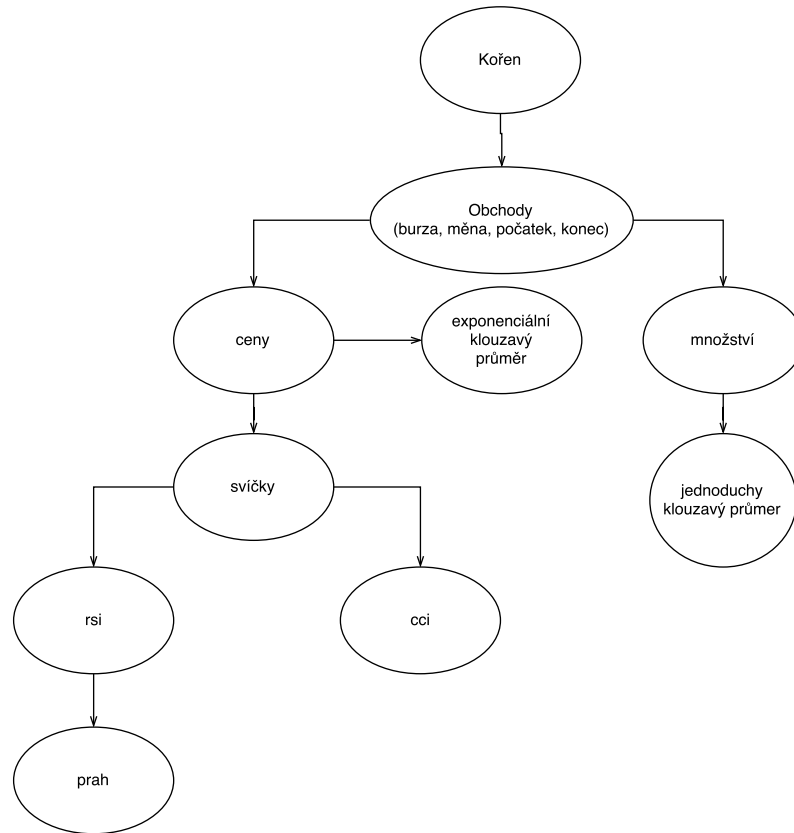
- nedocházelo k úniku dat,
- se najednou do databáze nenahrávalo příliš mnoho data.

Při zvažování všech možných variant nastavení jsem se řídil dokumentací Mysql. Konkrétně časové rozložení je v tabulce D.1. Sběr otisků příkazových knih probíhá periodicky každých třicet vteřin.

5.3 Výpočet indikátorů a reprezentace dat

Data uložená v databázi jsou v podobě detailů jednotlivých obchodů, jak bylo popsáno v sekci 5.2 o sběru dat. Tyta data je možno transformovat pomocí různých funkcí od indikátorů až po prahování a normalizaci. Z praktického hlediska bylo důležité navrhnout datovou strukturu a její výpočetní model tak, aby bylo možno datové transformace flexibilně řetězit. Možnot řetězení transformací je zachycena na schématickém obrázku 5.6. Při

tom bylo potřeba si uchovat informaci jak, výsledná data vznikala, respektive z čeho byla spočítána.

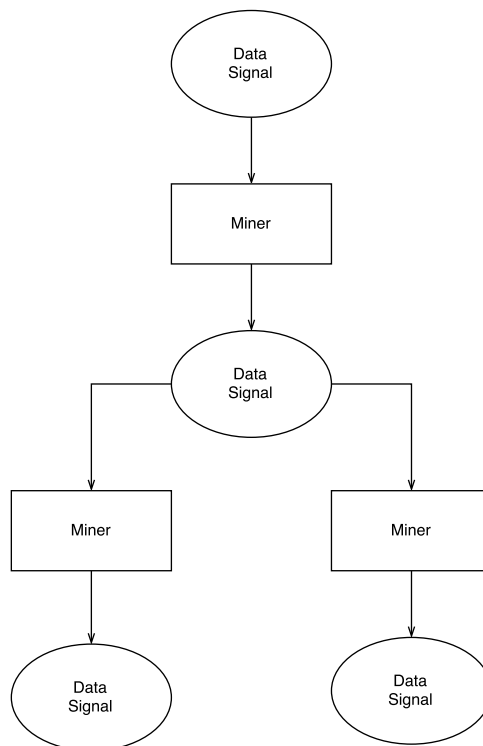


Obrázek 5.6: Ukázka cest výpočtu nad daty z databáze

Hlavní problémy vyplývající z výše zmíněných požadavků jsou:

- Jak minimalizovat zátěž databáze?
- Jak si udržet povědomí o postupné transformaci data prostřednictvím funkcí?
- Jak navrhnout celé řešení modulární tak, aby bylo možno snadno přidávat další funkce pro výpočty indikátorů?
- Jak zohlednit fakt, že některé funkce pracují pouze nad svíčkovým grafem?

Řešením byla vhodně zvolená datová struktura. Základní návrh je strom, který je obohacen o funkční prvky, které se starají o transformaci dat. Uzly stromu tvoří objekty modelující signál. Každý signál je popsán množinou dále nedělitelných prvků. Jedná se o dvourozměrné atomické jednotky, které mají svůj konkrétní čas a hodnotu. Iniciální signály jsou v tomto případě *cena* a *množství*, které jsou extrahovány z historie obchodů. Pro snížení zátěže databáze jsou záznamy o obchodech čteny pouze na začátku celého výpočtu. Následně dojde k separaci signálu ceny a množství. Každý signál ve stromu má svého rodiče a rovněž své synovské signály. Pomocí toho je vždy možné zjistit, jak byl daný signál vypočítán. Signál je pak schopen nad sebou samým vykonat požadovanou transformaci a vytvořit svého potomka. To umožňuje řetězení transformací. Výpočetní model v podobě stromu, kde jsou dva typy uzlů datové (signál) a výpočetní, je popsán na obrázku 5.7.



Obrázek 5.7: Výpočetní model

Výpočetní uzly jsou v tomto případě označeny anglickým slovem *miner*. Běh výpočtu je navržen tak, aby požadovaný signál byl vypočítán až v případě, že se na něj uživatel dotazuje. Jinými slovy nedochází k žádnému předpočítání.

Z pohledu architektury softwaru byl požadováno, aby řešení bylo modulární, jinak řečeno, aby bylo možno snadno přidávat další transformační funkce. To se pozděj ukázalo jako velmi rozumný požadavek. Řešením tohoto problému byl návrhový vzor *Chain of responsibility*. Na obrázku 5.8 je vidět použití návrhového vzoru a všechny transformační funkce, které jsou v systému implementovány¹. Zodpovědnost za správné řetězení transformací byla ponechána na uživateli.

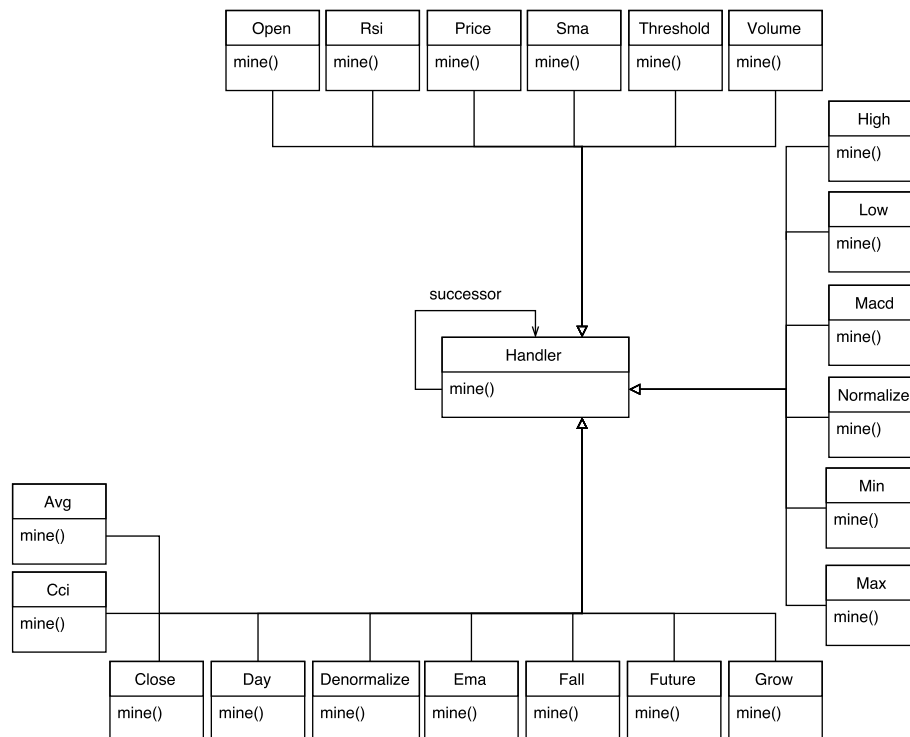
Praktickým výsledkem byl vznik rozhraní, které představuje pseudodotazovací jazyk na data. Ukázkou takového dotazu v jazyce Python je možné vidět na 5.1, kde je výsledný signál spočítán ze signálu ceny transformovaným klouzavým průměr s periodou 10. Tento signál je následně normalizován na úrovni dne a na závěr prahován s hodnotou prahu 0.7.

$$root[price()][sma(10)][normalize(d)][threshold(0.7)] \quad (5.1)$$

5.4 Predikční systém

Základní myšlenka tohoto predikčního systému vychází z podobného konceptu jaký byl uveden v práci [14]. Systém, který je zde prezentován, je výsledkem mnoha iterací, při kterých jsem se snažil najít nejlepší parametrizaci.

¹Všechny indikátory jsem od základu implementovat a nepoužíval jsem žádné předpřipravené moduly.



Obrázek 5.8: Schéma organizace transformátorů dat

5.4.1 Vstupy

Základ tvoří jedna centrální neuronová síť, do které vstupují hodnoty získané z technické analýzy burzovního vývoje. Konkrétně se jedná o tyto hodnoty:

- typická cena,
- jednoduchý klouzavý průměr s periodou 10,
- exponenciální klouzavý průměr s periodou 20,
- *MACD* indikátor s nastavením 26 – 12 – 9,
- indikátor *CCI* s periodou 50,
- indikátor *RSI* s periodou 14.

Od každého vstupu je bráno 10 historických hodnot. Indikátory jsou počítány nad 15 minutovým svíčkami.

5.4.2 Normalizace

Vstupní data jsou normalizována do rozsahu 0 až 1 podle vzorce 5.2. Proměnná h je hodnota, kterou je potřeba převést do požadovaného rozsahu. Hodnoty max a min jsou maximální a minimální hodnoty, kterých nenormalizovaná proměnná h může nabývat.

$$\frac{h - min}{max - min} \quad (5.2)$$

Při testování systému byly vyzkoušeny dva přístupy normalizace. První hledal maxima a minima na celém rozsahu vstupní množiny. Druhý přístup normalizoval hodnoty pouze v rámci dnů. Hlavní nevýhodou prvního přístupu je, že příliš vyhladí pohyby na úrovni dní, což je v případě intradenního obchodování nežádoucí. Z tohoto důvodu byl použit druhý zmiňovaný přístup, kdy jsou vstupní data normalizována v každém dni zvlášť.

5.4.3 Parametrizace sítě

Sít má standardně jednu vstupní a jednu výstupní vrstvu. Mezi nimi jsou dvě skryté vrstvy. První skrytá vrstva obsahuje 150 neuronů. Druhá vrstva má 100 neuronů. Správný počet neuronů ve skrytých vrstvách byl stanoven na základě experimentování. Aktivační funkce je nastavena pro všechny stejně. Nejlépe se chovala sigmoidální funkce například v porovnání s tzv. *relu*.

5.4.4 Výstupy

Výstupem systému je jeden signál, který udává pravděpodobnost toho, zda bude cena růst nebo padat v blízkém časovém horizontu. V aktuální konfiguraci to znamená, zda bude průměrná cena v budoucích třech hodinách natolik rozdílná, aby bylo možno vydělat na jedné pozici jeden dolar. Výstup sítě, který je mezi 0 a 1, lze chápat jako pravděpodobnost změny, na které bude možno vydělat alespoň jeden dolar. Nyní se čtenář asi zamýšlí proč zrovna jeden dolar. Je nutno si uvědomit, že vydělat jeden dolar znamená, že cena se musí změnit o jeden dolar a výši poplatků. Ty jsou na bitcoinových burzách nastaveny různě. S trochou nadhledu lze říci, že poplatky účtované při otevření a uzavření pozice je možno vyjádřit vzorcem 5.3.

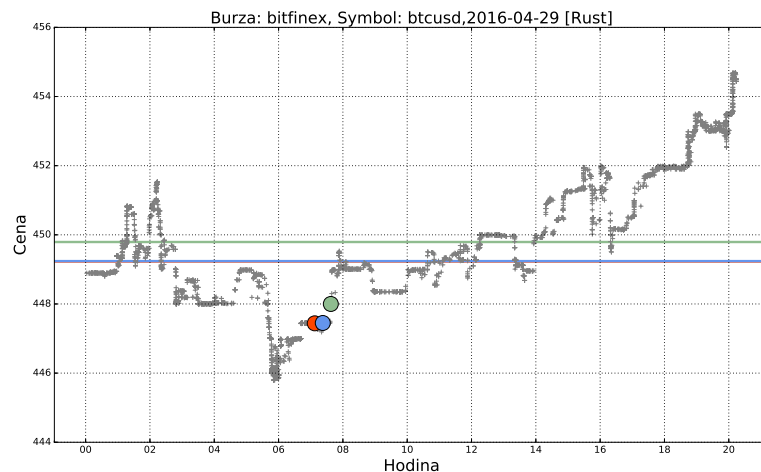
$$cena * \frac{4}{1000} \tag{5.3}$$

Pro představu – při ceně 400 USD za jeden bitcoin znamená vydělat jeden dolar změnu $1 + 1.6$ v požadovaném směru. Při bližším pohledu na bitcoinové burzy a velikost denních pohybů je tato hodnota optimální pro intradenní obchodování. Pro co nejlepší představu o trzích jsem si zřídil účet, ručně investoval a pečlivě sledovat vývoj cen v průběhu letního semestru. Praktická zkušenost pouze potvrzuje správnost volby výstupního parametru. Nakonec popisu výstupu sítě je potřeba zdůraznit, že sít není trénována na predikování konkrétní hodnoty. Sít se učí, zda bude v budoucnu změna tak výrazná, že na ní půjde vydělat alespoň jeden dolar.

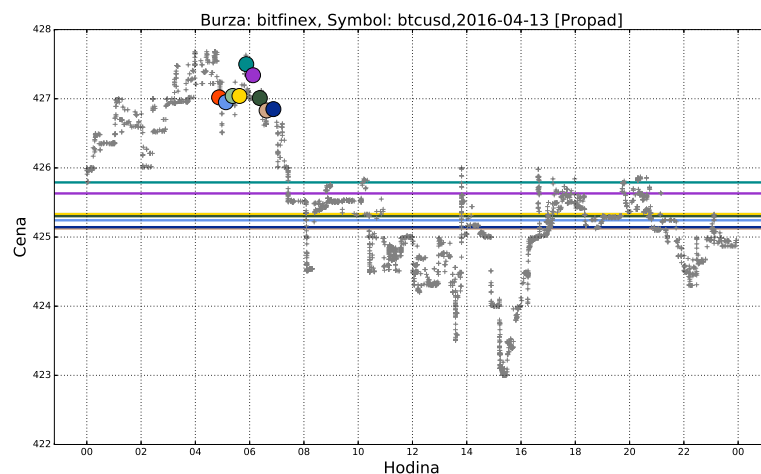
5.4.5 Dosažené výsledky

Na obrázcích 5.9 a 5.10 jsou ukázány predikce růstu a propadu na reálných pohybech cen. Barevné body označují signál ze systému pro vstup do pozice. Stejnou barvou je pak v grafu vykreslena přímka. Ta značí bod, od něhož se začíná vydělávat. Zde zobrazené signály jsou takové, které systém označil s pravděpodobností úspěchu větší než 0.95. Hodnota tohoto parametru se v průběhu experimentů se systémem ukázala jako nejvhodnější.

Důkladné vyhodnocení schopností celého systému bude provedeno v sekci 5.5 popisující strategii.



Obrázek 5.9: Ukázka signálu predikce růstu



Obrázek 5.10: Ukázka signálu predikce propadu

5.4.6 Poznátky

Predikční systém byl vyvíjen a v průběhu tohoto procesu byly učiněny některé zajímavé poznatky:

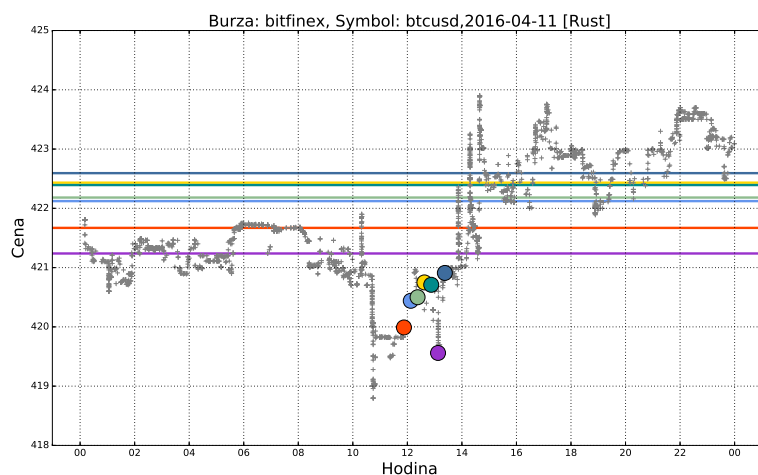
- Normalizace vstupu mezi hodnoty -1 a 1 se příliš neosvědčila. Daleko lépe modul neuronových sítí *Keras* fungoval při normalizaci mezi hodnoty 0 a 1. Tento poznatek je zajímavý, protože je v rozporu s všeobecnými doporučeními.
- Hodina přidaná jako vstup do neuronové sítě se na výsledku nijak výrazně neprojeví. Z praktického pozorování burz by se mohlo zdát, že tento parametr by mohl mít významný dopad, ovšem nemá.
- Přidání signálu objemu na vstup sítě se výrazněji neprojevuje. Podle textu [12] by mělo mít přidání tohoto signálu významný dopad.
- Předpověď konkrétní ceny není vhodným přístupem. Daleko lepším přístupem se ukázalo predikovat růst nebo propad obecně. Navíc zde odpadá problém s následnou

denormalizační výstupu sítě.

Tyto poznatky asi nelze nijak generalizovat, ale tvoří jeden z nejzajímavějších výstupů práce.

5.5 Strategie

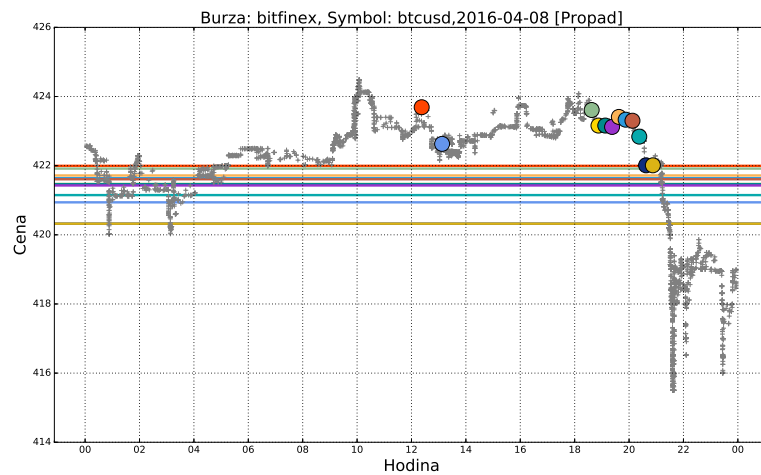
Jak již bylo zmíněno v předešlém textu ústředním bodem celého systému je neuronová síť. Do této sítě vstupují hodnoty indikátorů technické analýzy. Výstupem sítě je pak pravděpodobnost toho, zda bude vývoj ceny pokračovat zamýšleným směrem. Tento výstup je filtrován a v úvahu jsou brány pouze signály na vstup do pozice, kde je si predikční systém maximálně jistý. Číselně vyjádřeno jde o hodnotu 0.95 a vyšší. Strategie stojí na myšlence, že většina takto přefiltrovaných signálů pro vstup do pozice bude úspěšná. Ruku v ruce s touto domněnkou jde i snaha vydělat pouze 0.75 dolaru na jednu pozici, která má jistě větší šanci na úspěch než vyšší částky. V případě, že predikční systém predikoval falešný signál potom pozice zůstává otevřená do půlnoci a tam je automaticky uzavřena. Vyjádřeno terminologií obchodníka – tzv. *stop loss* zde není definován konkrétní částkou, ale časem. S ohledem na automatické uzavírání otevřených pozic o půlnoci strategie odmítá signály vstupu, které přijdou po desáté hodině večerní.



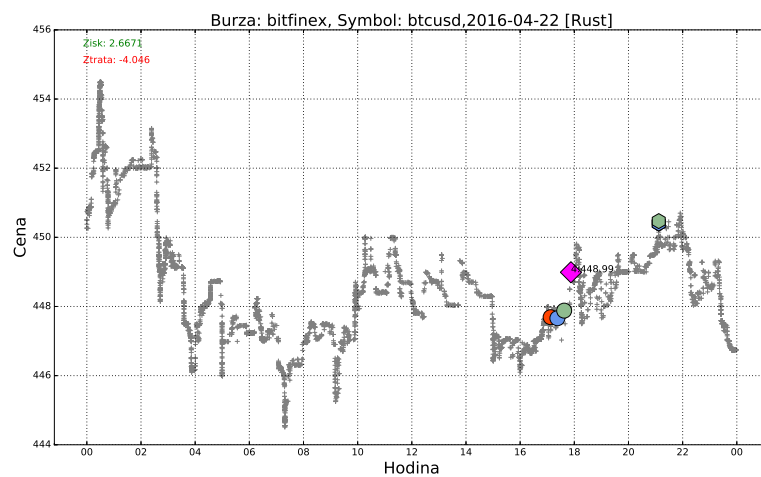
Obrázek 5.11: Ukázka signálů predikce růstu

Je potřeba zmínit, že nejspolehlivější signály pro vstup přicházejí ve skupinách. Tím je myšleno, že v krátké době dojde několik signálů pro vstup. Tyto situace jsou zachyceny na obrázcích 5.11 5.12. Při snaze zlepšit dosažené výsledky bylo vyzkoušeno sjednocení těchto signálů do jednoho a odfiltrování původně samostatně stojících signálů. Tento přístup se ve výsledku neukázal jako správný. Ve výsledném řešení byly ponechány všechny signály, kterým predikční systém přiřadil pravděpodobnost 0.95. Typické ukázky aplikace strategie jsou na obrázcích 5.13 5.14. Barevná dvojice kruhu a šestihranu značí vstup a výstup. Filové kosočtverce značí otevřené pozice, které se nepodařilo uzavřít na požadované ceně a jsou uzavřeny o půlnoci.

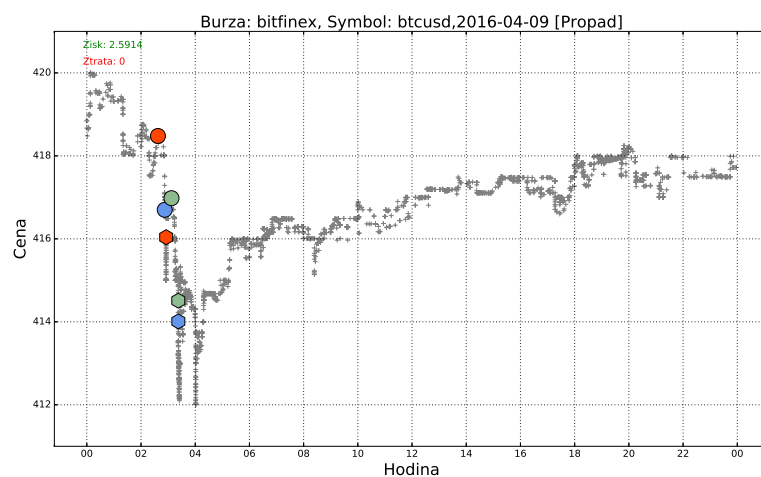
Výsledky dosažené aplikací strategie na měsíc duben tohoto roku jsou zachyceny v grafu 5.15a a 5.15b pro systém predikující růst. Pro úplnost dodejme, že duben byl měsícem, kterému nevládl jeden trend, ale růst a propad cen se průběžně střídaly. Pro systém predikující propad jsou výsledky zobrazeny v grafech 5.16a a 5.16b. Na obrázcích 5.15a a 5.16a jsou



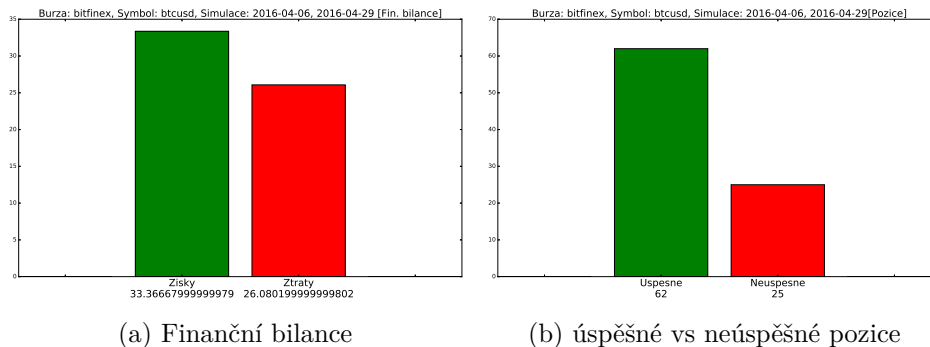
Obrázek 5.12: Ukázka signálů predikce propadu



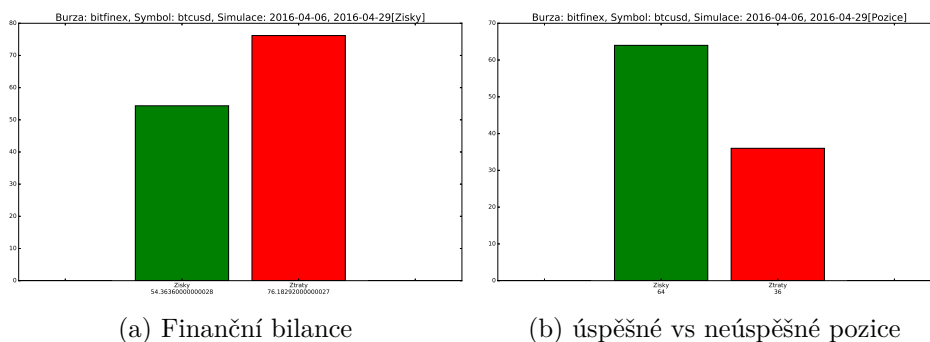
Obrázek 5.13: Ukázka typických denních signálů predikce růstu



Obrázek 5.14: Ukázka typických denních signálů predikce propadu



Obrázek 5.15: Hodnocení aplikace strategie pro predikci růstu



Obrázek 5.16: Hodnocení aplikace strategie pro predikci propadu

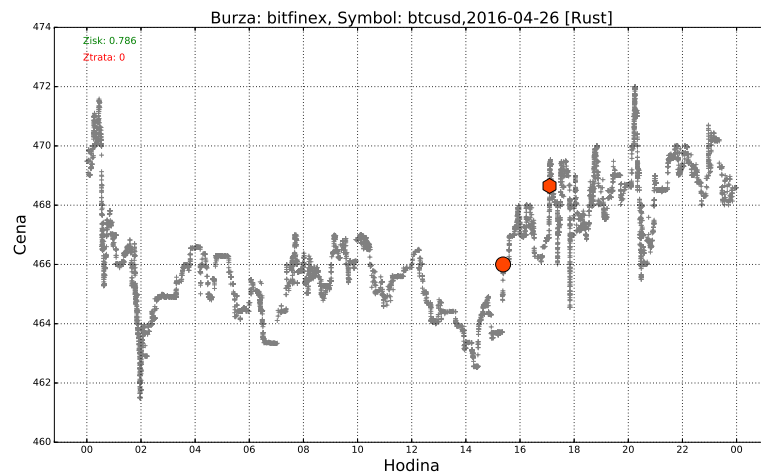
zachyceny celkové zisky a ztráty za testované období. V grafech 5.15b a 5.16b jsou zachyceny počty pozic, které byly úspěšně uzavřeny tzn. s výtěžkem a neúspěšně tzn. bez požadovaného výtěžku o půlnoci.

Z grafů je vidět, že systém pro predikci růstu je v kombinaci s touto strategií úspěšnější. Celkově větší úspěšnost systém pro predikci růstu systém vykazoval v průběh celého experimentování. Při analýze fungování obchodního systému jako celku je zřejmý hlavní problém, kterým je velmi nedostatečné využívání všech pohybů na trhu. Příklad nedostatečného využívání je zachycen na obrázku 5.17.

5.6 Použité nástroje

Nosnou platformou pro celý systém je *Python* verze 3.5. Standardní moduly pythonu byly doplněny velkým balíkem *anaconda3* pro podporu pokročilých výpočtů. V rámci tohoto balíku byl doinstalován modul *numpy*, který je fundamentálním modulem vědeckých výpočtů. Grafy prezentované v této práci byly automaticky vytvořeny pomocí modulu *matplotlib* a nástroje *gnuplot*.

Jedním ze stěžejních bodů práce je použití umělých neuronových sítí. Najít vhodný modul pro práci s neuronovými sítěmi v *Pythonu* nebylo jednoduché. Prvním zkušným modulem byl *neurolab*, který měl problémy s učení. Celý proces trval neúměrně dlouho k počtu trénovacích vzorů a požadované přesnosti a byl dosti nestabilní. Druhý modul pro strojové učení *Keras* pracuje bez problému a navíc má daleko širší možnosti parametrizace. Modul *Keras* využívá pro svou práci balíček *Theano* nebo *TensorFlow*. V tomto případě se používá balíček *Theano*.



Obrázek 5.17: Ukázka nevyužitých příležitostí k obchodování

Pro ukládání dat byla zvolena relační databáze *MySQL*. Pro čtení a zápis dat byl využit modul *mysqldb*. Tento modul je ze své podstaty velmi rychlý a nenáročný, protože se jedná pouze obal pro jazyk *Python* jehož jádro je napsáno v jazyce *C*.

Verzování zdrojových kódů zajišťoval systém *Git*, který se velmi dobře doplňoval s editorem *PyCharm*. Přechod na editor *PyCharm* velmi akceleroval celý vývojový proces a byly pomocí něj vytvořeny všechny zdrojové texty této práce. Doplňujícími nástroji se pak staly *gitstat* a *cloc* pro získání statistik. Funkci celého obchodního softwaru zajišťuje operační systém *Debian* běžící na virtuálním serveru. S pomocí těchto nástrojů byla realizována práce popsaná v předešlém textu.

Kapitola 6

Závěr

Hodnocení dosažených výsledků lze rozdělit na dvě části. Z pohledu oficiálního zadání byly všechny požadavky kompletně splněny. Byly získány teoretické znalosti z oblasti neuronových sítí a obchodování na burzách založené na technické analýze. Nad rámec svých povinností jsem si zařídil reálný obchodní účet a příležitostně jsem obchodoval, abych porozuměl burzovnímu světu. Tyto poznatky byly využity při konstrukci obchodního systému, který byl úspěšně otestován na historických datech. Objektivní posouzení obchodního systému je velmi zrádné. Přestože systém má určité schopnosti predikce vývoje burzovních cen a stojí na teoretických základech, nelze jej prohlásit za univerzální. Celkově vzato se systém učí, jaké trendy a zákonitosti vládly burze v minulosti. Není zde žádná garance, že se hlavní trendy pohybující cenou nezmění. Změnu zákonitostí burzy může vyvolat například vnější fundamentální impuls, který technická analýza nedokáže dostatečně rychle reflektovat. Ve výsledku se funkce celého systému opírá o víru, že trendy z minulosti budou burzu ovládat i v budoucnosti. V důsledku toho ani sebelepší testování a hledání optimální parametrizace nezaručuje jistotu výdělku. Pozitivním v tomto směru je fakt, že jedním z hlavních hybatelů cenou je lidská psychika, která je v obecném smyslu dosti stálá ve svém fungování.

Druhou část hodnocení dosažených výsledků tvoří naplnění vlastních cílů. Zde bylo hlavním úkolem otestovat systém v provozu na reálných burzách. K tomuto cíli bylo velmi významně nakročeno, ale naplněn nebyl. Splněny byly ale některé dílčí podcíle:

- implementace a otestování komunikačního rozhraní včetně privátních funkcí,
- dobře navržená databáze pro ukládání údajů z burzy,
- správně fungující systém sběračů burzovních dat.

Hlavní překážkou pro splnění cíle byla velká časová náročnost na integraci privátního rozhraní do obchodního systému.

Budoucí vývoj by se měl určitě zaměřit na následující body:

- Reálné nasazení a otestování v ostrém provozu.
- Využití údajů z knihy příkazů, které ve fungování popsaného obchodního systému nefigurují.
- Kombinace více neuronových sítí s různými výstupy, a například dlouhodobé a krátkodobé předpovědi.

Hlavním bodem zůstává reálné nasazení, protože jen to nejlépe ověří robustnost celého řešení.

Závěrem lze konstatovat, že problematika predikce vývoje cen na světových burzách je velmi komplexní. Jediným systémem, který by mohl znamenat opravdový průlom je nasazení umělé inteligence. V tomto ohledu dochází aktuálně k velmi bouřlivému vývoji. Velké firmy jako Google, IBM nebo Microsoft již nyní nabízejí veřejně svá řešení. Jak se tento technologický pokrok nakonec promítne v oblasti predikce cen na burzách je otázkou budoucnosti.

Literatura

- [1] S čím vším můžeme obchodovat na burze?
<http://www.tradeandfinance.eu/clanky/zaklady-obchodovani/prirucka-uspesneho-obchodovani-na-burze-kapitola-6-s-cim-vsime-muzeme-obchodovat/>.
- [2] Candlesticks Charts Explained.
http://www.actionforex.com/index2.php?option=com_content&task=view&id=7114, 2006.
- [3] Historie burz. <http://www.penize.cz/15865-historie-burz>, 2009.
- [4] Edwards, R.: *Technical analysis of stock trends*. Boca Raton, FL New York: CRC Press AMACOM, American Management Association, 2007, ISBN 978-0814408643.
- [5] Hansun, S.: A new approach of moving average method in time series analysis. In *New Media Studies (CoNMedia), 2013 Conference on*, Nov 2013, s. 1–4, doi:10.1109/CoNMedia.2013.6708545.
- [6] Hartman, O.; Turek, L.: *První kroky na FOREXu : Jak obchodovat a uspět na měnových trzích*. Brno: Computer Press, 2009, ISBN 978-80-251-2006-4.
- [7] Hayes, R.; Wu, J.; Chaysiri, R.; aj.: Effects of time horizon and asset condition on the profitability of technical trading rules. *Journal of Economics and Finance*, 2014: s. 1–19, ISSN 1055-0925, doi:10.1007/s12197-014-9291-5.
- [8] Kučera, R.: *Technická analýza a její využití na mezinárodním měnovém trhu [online]*. Diplomová práce, Masarykova univerzita, Ekonomicko-správní fakulta, 2010 [cit. 2014-12-30].
- [9] M. Šnorek, M. Jiřina: *Neuronové sítě a neuropočítače*. Vydavatelství ČVUT, 1996, ISBN 80-210-01455-X, 124 s.
- [10] Mařík, V.: *Umělá inteligence*. Praha: Academia, 1993, ISBN 80-200-1044-0.
- [11] Murphy, J.: *Technical analysis of the financial markets : a comprehensive guide to trading methods and applications*. New York: New York Institute of Finance, 1999, ISBN 978-0735200661.
- [12] Nesnídal, T.; Podhajský, P.: *Jak se stát intradenním finančníkem*. Praha: Centrum finančního vzdělávání, 2008, ISBN 978-80-903874-4-7.
- [13] Orlinsky, K.: The Stochastic Scientist: Rewiring the Brain.
<http://stochasticscientist.blogspot.com/2010/01/rewiring-brain.html>.

- [14] Putna, L.: *Predikce vývoje kurzu pomocí umělých neuronových sítí [online]*. Diplomová práce, FIT VUT v Brně, 2011.
- [15] Rumelhart, David E. and McClelland, James L. and PDP Research Group, Corporate (editor): *Parallel Distributed Processing: Explorations in the Microstructure of Cognition*. Cambridge, MA, USA: MIT Press, 1986, ISBN 0-262-68053-X, 318–363 s.
- [16] Skersys, T.: *Information and software technologies : 19th International Conference, ICIST 2013, Kaunas, Lithuania, October 2013. Proceedings*. Heidelberg: Springer, 2013, ISBN 978-3-642-41947-8.
- [17] Vasiliou, D.; Eriotis, N.; Papathanasiou, S.: How rewarding is technical analysis? Evidence from Athens Stock Exchange. *Operational Research*.
- [18] Šíma, J.: *Teoretické otázky neuronových sítí*. Praha: Matfyzpress, 1996, ISBN 80-85863-18-9.

Přílohy

Seznam příloh

A	Obsah CD	46
B	Metriky	47
C	Grafické vzory	48
	C.0.1 Základní vzory	48
	C.0.2 Pokročilé grafické vzory	50
D	Časové rozložení sběračů	53
E	Plakát	55

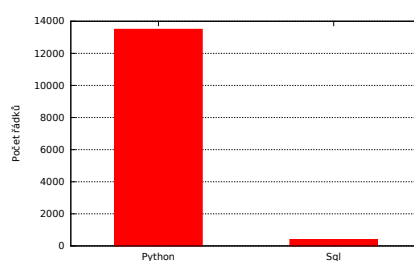
Příloha A

Obsah CD

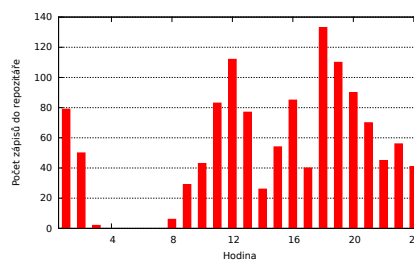
- CD
 - `text` - zdrojové soubory tohoto textu.
 - `src` - zdrojové soubory obchodního systému
 - `db` - zdrojové soubory sloužící pro inicializaci databáze MySQL
 - `instalation` - záznam popisující co vše bylo do čistého OS Debian doinstalováno
 - `modules` - adresář obsahující hlavní moduly obchodního systému
 - `collectors`
 - `db`
 - `graphs`
 - `indicators`
 - `miners`
 - `predictor`
 - `quants`
 - `signals`
 - `simulations`
 - `stocks`
 - `tests` - podrobné testy vše částí obchodního systému
 - `gitstats` - statistiky vývoje exportovaných ze systému Git
 - `poster` - plakátek
 - `video` - video prezentace

Příloha B

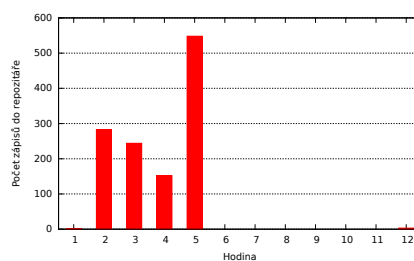
Metriky



Obrázek B.1: Jazyky a počty řádků kódu



Obrázek B.2: Souhrnný přehled zápisů do repozitáře Git v průběhu dne



Obrázek B.3: Souhrnný přehled zápisů do repozitáře Git v průběhu roku

Příloha C

Grafické vzory

Obrazcům, které vytváří podkladové aktivum vývojem svých cen v průběhu času, se říká grafické vzory. Můžeme je rozdělit na základní a pokročilé. Klíčem pro klasifikaci bude počet sloupců (respektive svíček) v daném časovém rámci, které tyto vzory obsahují. V této práci všechny obrazce obsahující víc jak jeden sloupec spadají do kategorie pokročilých.

C.0.1 Základní vzory

Základní grafické vzory používají, pro posouzení vývoje informace pouze z jednoho sloupce (respektive svíčky). Nejde samozřejmě o samonosný nástroj, který by byl sám o sobě schopen správně predikovat vývoj s vysokou pravděpodobností. Spíše jde o další střípek, který může pomoci lépe pochopit burzovní dění. Účelem zde není podat vyčerpávající výčet všech možných jednosloupcových grafických vzorů. Smyslem je představit nejzákladnější a nejznámější formace a především způsob jejich chápání.

Při vypracování bylo čerpáno z diplomové práce [8] a z webových stránek [2]. Stejně jako v práci [8] i zde byly použity původní názvy pro označení formací. Důvodem je místy prakticky nemožný překlad a fakt, že i oficiální česká literatura v takovýchto situacích používá původní označení. Pro snazší popis bude použito terminologie svíčkových grafů, i přestože jsou zde zobrazeny sloupcové grafy.

Vzory jsou zobrazovány v reálných grafech. V některých případech nejsou obrazce tak čistě čitelné jak v ryze teoretických popisech. Nicméně předvedení vzoru tímto způsobem je lepší pro realistickou představu o jeho fungování.

Marubozu

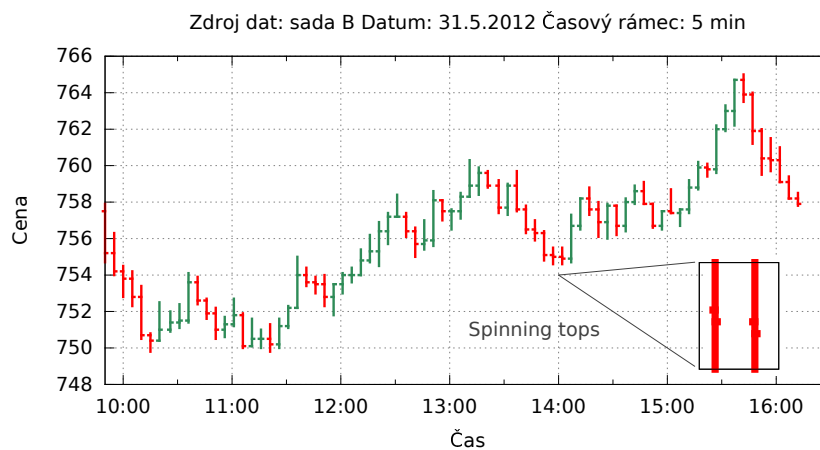
Vzor *Marubozu* je typický svým mohutným tělem a velmi malými nebo žádnými stíny. Tento obrazec se často nachází na počátku silného trendu. Útvar je vidět na obrázku C.1 hned dvakrát. První útvar se objevil v čase 10:00 a signalizoval ustanovení silného vzestupného trendu. Druhý případ se vyskytl v čase 10:40, kdy pouze potvrzuje sílu trendu.

Spinning tops

Obrazec *Spinning tops* je v provedení s *Marubozu* opačný. Charakteristické jsou dlouhé stíny a velmi krátké tělo. Objevuje se v situacích, kdy trh kolísá a nemá jasně ustanovený trend. Rovněž naznačuje možnost obratu. Situace kdy tento obrazec byl signálem obratu je na obrázku C.2.



Obrázek C.1: Ukázka grafického vzoru Marubozu



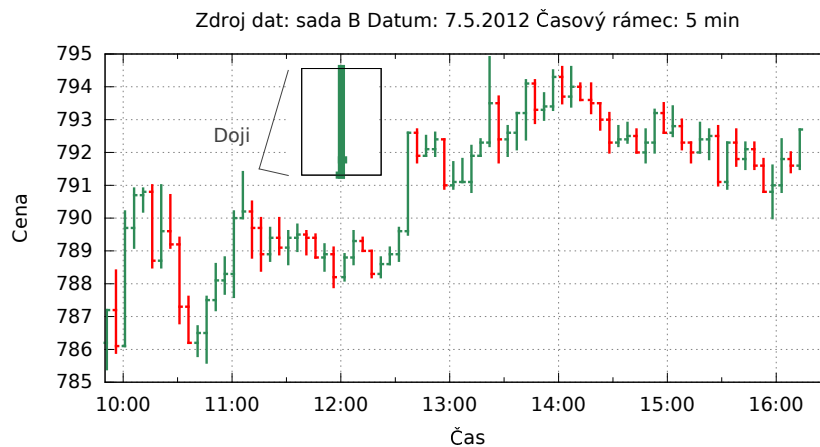
Obrázek C.2: Ukázka grafického vzoru Spinning tops

Doji

Vzor označovaný jako *Doji* má stejnou nebo velmi podobnou otevírací a uzavírací cenu a dlouhý stín. Stejně jako *Spinning tops* může být signálem pro obrat trendu, jak je vidět na obrázku C.3.

Hammer

Výskyt obrazce *Hammer* je typický pro přechod do býčího neboli stoupavého trendu. To znamená, že jej primárně budeme hledat v medvědím neboli klesajícím trendu. Charakteristikou *Hammer* je krátké tělo a stín směrem dolů. Obrázek C.4 ukazuje jeden z případů, kdy *Hammer* byl signálem pro obrat vzhůru.



Obrázek C.3: Ukázka grafického vzoru Doji



Obrázek C.4: Ukázka grafického vzoru Hammer

Hangin man

Tento obrazec je vizuálně stejný jako vzor *Hammer*. Rozdíl je v tom, že *Hangin man* hledáme ve stoupajícím trendu a jeho nalezení je pak signálem pro obrat směrem dolů. Popisem *Hangin man* je krátké tělo a stín směrem dolů. Příklad toho vzoru je zachycen na obrázku C.5.

C.0.2 Pokročilé grafické vzory

Pokročilé grafické vzory se liší od základních především v tím, že jde o formace více sloupců oproti základním jednosloupcových obrazců. Opět zde nebude poskytnut kompletní výčet všech vzorů. Znovu půjde pouze o základní výběr složitějších vzorů, které mají ilustrovat tento způsob pohledu na dění na trhu. I když se jedná o základní obrazce, mají svou platnost, protože jsou lépe identifikovatelné a mají větší procento výskytu. Snažší iden-

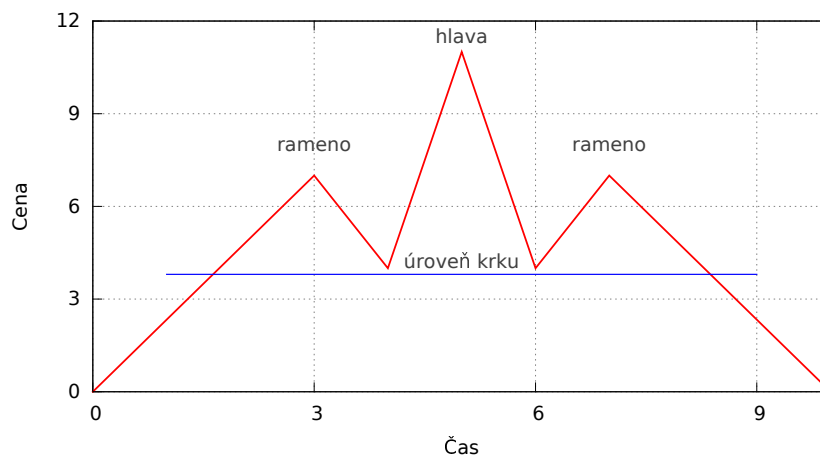


Obrázek C.5: Ukázka grafického vzoru Hanging man

tifikovatelnost bude s výhodou využita při přenosu do počítačového programu. Formace zde popsané je možno nalézt v jakémkoliv časovém rámci. Informace zde poskytnuté byly čerpány především z [6]. Detailnější informace pak nabídne kniha [4].

Hlava a ramena

Jedná se o jeden z nejznámějších obrazců, který se typicky vyskytuje na konci silných vzestupných trendů. Schématickou ukázkou lze vidět na obrázku C.6.¹

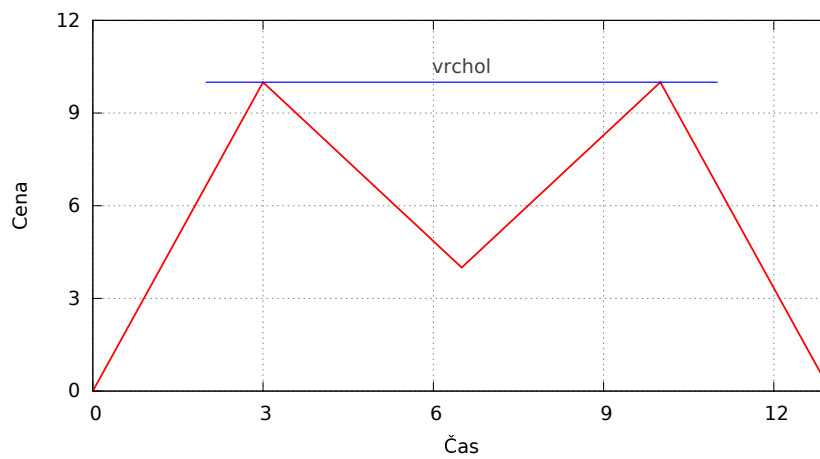


Obrázek C.6: Ukázka grafického vzoru *hlava a rameno*

¹Bude více rozvedeno v DIP.

Dvojitý vrchol

Jedná se o častý a rovněž velmi spolehlvý vzor, který se vyskytuje napříč všemi časovými pásmy na všech podkladových aktivech. Základní tvar je vidět na obrázku C.7.²



Obrázek C.7: Ukázka grafického vzoru *dvojitý vrchol*

²Bude více rozvedeno v DIP.

Příloha D

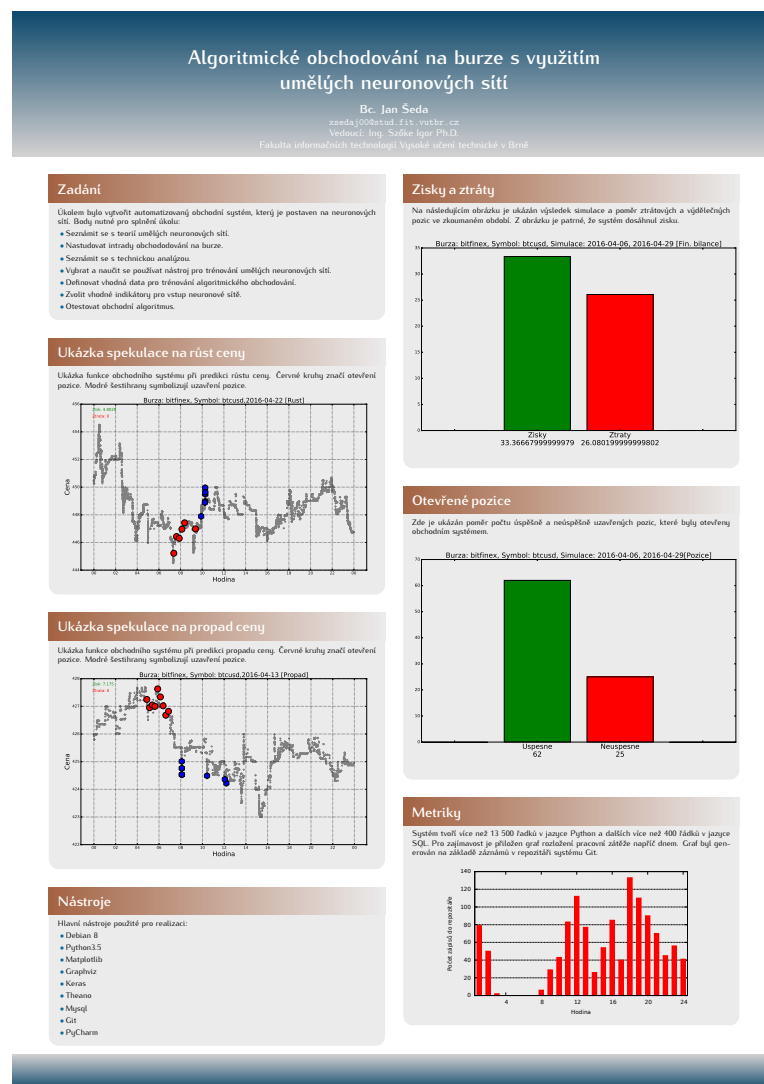
Časové rozložení sběračů

7 sekund	
Okcoin	btcny ltcny
Btcc	btcny
2 minuty	
Okcoin	btcusd ltcusd
Kraken	btceur
Btce	btcusd
15 minut	
Kraken	btcusd ltcusd
Btce	ltcbtc
	ltcusd
Bitfinex	btcusd
45 minut	
Bitstamp	btcusd
1 hodina	
Btcc	ltcny ltcbtc
4 hodiny	
Bitfinex	ltcusd ltcbtc
Btce	btczur btceur lczur ltceur
Kraken	ltceur
Kraken	btcltc
12 hodin	
Btce	usdrur eurusd
Btce	eurrur

Tabulka D.1: Časové skupiny

Příloha E

Plakát



Obrázek E.1: Plakátek