

Vysoká škola ekonomická v Praze

Fakulta financí a účetnictví

DIPLOMOVÁ PRÁCE

2019

Bc. Marek Svoboda

VYSOKÁ ŠKOLA EKONOMICKÁ V PRAZE

Fakulta financí a účetnictví

Katedra bankovníctví a pojišťovnictví

Studijní obor: Bankovníctví a pojišťovnictví



Aplikace umělé inteligence při obchodování na kapitálových trzích

Diplomová práce

Autor diplomové práce: Bc. Marek Svoboda

Vedoucí diplomové práce: Ing. Milan Fičura, Ph. D.

Rok obhajoby: 2019

Čestné prohlášení

Tímto prohlašuji, že jsem diplomovou práci na téma *Aplikace umělé inteligence při obchodování na kapitálových trzích* vypracoval samostatně, s využitím literatury a zdrojů, na které se v práci odkazuji.

V Praze dne

Podpis

Poděkování

Touto cestou bych rád poděkoval Ing. Milanovi Fičurovi, Ph.D. za cenné rady, připomínky a pomoc, které mi během psaní diplomové práce poskytoval.

Abstrakt

Tato diplomová práce zkoumá možnost využití metody umělé inteligence neuronových sítí pro predikci vývoje časových řad na akciových trzích. Byly sestaveny tři modely neuronových sítí, z nichž první představuje rekurentní LSTM neuronovou síť, sestavenou prostřednictvím programu Matlab. Zbylé dva modely neuronových sítí představují feed-forward neuronovou síť, kdy jeden z nich byl sestaven na základě modelu NAR a druhý na základě modelu NARX. Tyto dva modely byly také sestaveny prostřednictvím programu Matlab. Jako vstupní data slouží sedm akciových indexů, a to tři indexy z USA, dva z Evropy a dva z Asie. K vyhodnocení predikční schopnosti zkoumaných modelů je využito statistických metrik a metriky výnosnosti. Model rekurentní LSTM sítě dosahoval kladné výnosnosti, oproti strategii nákupu na počátku investičního období a prodeje na konci investičního období, u tří ze sedmi zkoumaných indexů. Nicméně, míry chybovosti mezi reálným vývojem a predikovaným vývojem akciových indexů byly velmi nízké, a proto je tento model hodnocen za využitelný pro predikci vývoje zkoumaných akciových indexů. Zbylé dva modely byly hodnoceny z hlediska chybovosti a dalších statistických metrik, kde za zvoleném období dosahovaly velmi kvalitních výsledků. Mohou být tedy také hodnoceny za vhodné pro predikci vývoje zkoumaných akciových indexů. Hlavním cílem této práce je zjistit, zda lze s dostatečnou přesností předpovídat budoucí vývoj zkoumaných akciových indexů.

Klíčová slova

Umělá inteligence, Neuronové sítě, LSTM, Akciové indexy

Abstract

The objective of this master thesis is to examine utilization of the neural network method of artificial intelligence to predict time series development within stock markets. Three neural network models were formed, from which the first one represents the recurrent LSTM neural network, formed using the Matlab software. The remaining two neural network models represent a feed-forward neural network, where one of them was based on the NAR model and the second one on the NARX model. Likewise, the two models were formed via the Matlab platform. There are seven stock indexes utilized as data inputs, namely three indexes from the USA, two from Europe and two from Asia. In order to evaluate the prediction capabilities of the examined models, statistical and profitability metrics are used. The recurrent LSTM network model reached positive profitability, contrary to the purchase strategy at the beginning of the investment period and to the sales at the end of the investment period. This applies to three out of the seven examined indexes. Nevertheless, the error rates between the real and the predicted development of stock indexes were very low, and therefore this model is evaluated as exploitable for the usage of prediction of examined stock indexes. The remaining two models were assessed in terms of the error rate and other statistical metrics, as they reached fine results in the selected period. Thus, they may be as well evaluated as suitable for prediction of the development of examined stock indexes. The principal objective of this thesis is to determine, whether it is possible to predict the future development of examined stock indexes with an adequate accuracy.

Key words

Artificial Intelligence, Neural Networks, LSTM, Stock indexes

Obsah

1	Úvod.....	10
2	Umělá inteligence, strojové a hluboké učení	12
2.1	Umělá inteligence.....	12
2.2	Strojové učení.....	12
2.3	Hluboké učení	13
3	Neuronové sítě	14
3.1	Průběh neuronové sítě	15
3.1.1	Aktivační funkce	15
3.1.2	Forwardpropagation	17
3.1.3	Ztrátová funkce	17
3.1.4	Backpropagation	18
3.1.5	Gradient descent.....	18
3.2	Princip učení neuronových sítí	19
3.3	Metody učení	20
3.3.1	Učení s učitelem	20
3.3.2	Učení bez učitele	21
3.3.3	Analýza hlavních komponent	23
3.3.4	Kombinace učení s učitelem a bez učitele (Semi-supervised learning - SSL)	23
3.3.5	Zpětnovazebné učení (Reinforcement learning)	23
3.4	Vývoj technologie neuronových sítí.....	24
3.4.1	Historie neuronových sítí.....	24
3.4.2	Neuronové sítě v současnosti.....	24
3.4.3	Využití neuronových sítí ve financích.....	26
3.5	Silné stránky neuronových sítí.....	27
3.6	Slabé stránky neuronových sítí.....	28
3.7	Typy neuronových sítí	29
3.7.1	Feed-Forward neuronové sítě	29

3.7.2	Rekurentní neuronové sítě.....	32
3.7.3	Symetricky propojené neuronové sítě	36
3.8	Další metody umělé inteligence	38
3.8.1	Metoda podpůrných vektorů.....	38
3.8.2	Algoritmus K-nejbližších sousedů	38
3.8.3	Decision trees.....	39
3.8.4	Random forrests	40
4	Rešerše.....	41
4.1	Forecasting Stock Market with Neural Networks	41
4.2	Deep learning with long short-term memory networks for financial market predictions	45
4.3	Stock Market Prediction Using Artificial Neural Networks	46
4.4	Stock Market Index Prediction with Neural Network during Financial Crises: A Review on Bist-100	47
4.5	Stock market index prediction using artificial neural networks trained on foreign markets.....	48
5	Praktická část	50
5.1	Indexy	50
5.1.1	Dow Jones Industrial Average	50
5.1.2	S&P 500	50
5.1.3	NASDAQ Composite	51
5.1.4	DAX	51
5.1.5	CAC 40.....	51
5.1.6	HANG SENG	51
5.1.7	NIKKEI 225.....	51
5.2	Sestavení modelů.....	52
5.2.1	Úprava dat.....	52
5.2.2	Rozdělení na tréninková, validační a testovací data.....	53
5.2.3	Hyperparametry	54
5.2.4	Ukazatele vyhodnocení modelů	54
5.3	Predikční modely.....	55

5.3.1	Model rekurentní LSTM neuronové sítě	55
5.3.2	Model NAR	60
5.3.3	Model NARX.....	64
6	Závěr.....	70
7	Použité zdroje	72
8	Seznam obrázků	80
9	Seznam tabulek	81

1 Úvod

V současnosti obor umělé inteligence a strojového učení čelí velkému rozmachu a s jeho reálnou aplikací se lidé setkávají na každodenní bázi, ať už se jedná o běžné či komerční využití. Známým příkladem může být osobní asistent Siri od společnosti Apple, který na základě hlasové interakce dokáže nalézt potřebné informace, určit požadovanou trasu cesty, přidat událost do kalendáře či poslat textovou zprávu. V rámci komerčního využití se může jednat o cloudovou platformu HANA od společnosti SAP, kterou společnost využívají ke správě svých informačních databází. Tato platforma je schopna přijímat obrovské množství informací, které společností jejich zákazníci poskytují skrze své obchodní aktivity, a to z mnoha zdrojů.

Tato práce se zabývá konkrétní metodou umělé inteligence, a to neuronovými sítěmi. Tato metoda má své využití v mnoho oborech, až už se jedná o autonomní vozidla, chatboty v zákaznické podpoře, identifikaci osob skrze rozeznání obličeje či diagnózu vážných chorob ve zdravotnictví. Metoda neuronových sítí nachází mnoho využití i ve finančním sektoru, a to mimo jiné při vývoji obchodních systémů s cennými papíry, vyhodnocení kreditního rizika u spotřebitelských půjček, hypoték a kreditních karet, prognóze vývoje ratingů u dluhopisů, modelaci vývoje měnových kurzů a predikcí vývoje cen na akciových trzích.

V této je metoda neuronových sítí aplikována na poslední ze zmíněných využití ve finančním sektoru, a to na predikci vývoje cen na akciových trzích. Pro určení správné predikce vývoje cen akciových titulů je zapotřebí obrovské množství dat, které se pohybuje podle nepravidelných vzorů. Z důvodu obtížné identifikace těchto nepravidelných cenových pohybů skrze standardní statistické metody, může být využito právě metody neuronových sítí. Tato metoda dokáže při poskytnutí dostatečně výkonné výpočetní techniky pracovat s obrovským množstvím dat a je skrze neurony obsažených ve skrytých vrstvách sítí schopna pracovat s nelineárními vzory pohybu cenových dat.

Cílem této práce je posoudit využitelnost zvolených modelů neuronových sítí při predikci vývoje cen na akciových trzích. Konkrétně budou zvoleny tři modely neuronových sítí, a to model rekurentních LSTM neuronové sítě a modely NAR a NARX pro feed-forward neuronovou síť. Pro posouzení využitelnosti aplikace této metody na různých trzích po světě je za zkoumaná data zvolen vývoj cenových indexů v USA, Evropě a Asii. K posouzení predikční schopnosti těchto modelů bude u modelů NAR a NARX využito statistických metrik chybovosti mezi skutečným vývojem akciových indexů a vývojem, který předpověděla neuronová síť. K ověření neskreslené predikce bude využito metody regrese a autokorelace. V rámci modelu rekurentní LSTM neuronové sítě bude schopnost predikce ověřena skrze statistické metriky chybovosti. Dále bude porovnána výnosnost predikce vývoje cen, který navrhla neuronová síť s výnosem oproti běžné strategii Buy and Hold.

Práce je členěná na teoretickou část, část rešerše dřívějších výzkumů a praktickou část. Teoretická část práce se nejprve zabývá úvodem do umělé inteligence a do strojového a hlubokého učení. Poté je popsáno,

jak neuronové sítě obecně fungují a také jsou uvedeny jednotlivé části neuronové sítě. Dále se práce zabývá popisem různých metod, na základě kterých se neuronová síť učí. Další kapitolu tvoří popis vývoje technologie neuronových sítí, což následuje kapitola o silných a slabých stránkách této metody. Dále jsou uvedeny různé typy neuronových sítí a popis na jakém principu fungují. V poslední kapitole praktické části této práce jsou uvedeny další metody umělé inteligence. V rámci části této práce, kde jsou obsaženy řešerše vztahujícím se k dříve proběhlým zkoumáním využitelnosti metody neuronových sítí na kapitálových trzích, je posouzeno pět prací zabývajících se tímto tématem. V praktické části této práce jsou nejprve uvedeny zkoumané akciové indexy. Poté postup na základě, kterého byla zvolena architektura zkoumaných modelů a následně testování a vyhodnocení predikční schopnosti těchto modelů neuronové sítě.

2 Umělá inteligence, strojové a hluboké učení

V první části této práce je obecně popsána umělá inteligence, metoda strojového učení a metoda hlubokého učení. Dále jsou v této části uvedeny možné reálné aplikace těchto metod a odkazy na články zabývající se touto problematikou.

2.1 Umělá inteligence

Umělá inteligence označuje nejširší pojem pro pokročilou počítačovou inteligenci. V roce 1956 na Dartmouth Artificial Intelligence Conference byla tato technologie interpretována jako: Každý okruh vědomostí nebo jakýkoliv jiný prvek inteligence, který může být popsán tak dokonale, že je ho stroj schopný simulovat. Okruh umělé inteligence může zahrnovat cokoli od počítače hrajícího šachy či systému, který rozezná a interpretuje řeč spolu se schopností na ni odpovídat. Tato technologie může být rozřazena do tří skupin: Omezená umělá inteligence, obecná umělá inteligence a super inteligentní umělá inteligence. (Reese, 2017)

Příkladem umělé inteligence může být počítačový program od IBM Deep Blue, který v roce 1996 porazil Garry Kasparova ve hře šachy (Greenemeier, 2017), anebo program DeepMind Alpha Go od Google, který v roce 2016 porazil Lee Sedola ve hře Go. Charakteristika omezené umělé inteligence je ta, že vyniká v jedné konkrétní úloze, zatímco obecná umělá inteligence se více chová jako člověk a je schopna vyřešit širokou škálu úloh. (Petr, 2017)

Super inteligentní umělá inteligence posouvá tuto oblast ještě dále a je označována za technologii, která je daleko chytřejší než nejlepší mozky světa prakticky v každém okruhu lidské činnosti, zahrnující i vědeckou činnost, obecné povědomí a sociální dovednosti.

2.2 Strojové učení

Strojové učení označuje podobor umělé inteligence. Hlavní charakteristikou je, že stroj získá data a začne se z nich učit. Systémy strojového učení jsou schopny nově nabyté vědomosti rychle použít a trénovat za pomoci velkých souborů dat, za účelem rozpoznávání obličejů, hlasu, objektů, překladu a mnoho dalších úkolů. Oproti ručně naprogramovanému softwaru, který plní daný úkol podle specifických instrukcí, strojové učení umožňuje systému se svépomocí naučit rozpoznávat vzory v datech a tvořit předpovědi. Oba programy Deep Blue a DeepMind patří mezi umělou inteligenci. Nicméně Deep Blue nepatří do systémů strojového učení, jelikož jeho fungování je závislé na naprogramovaných pravidlech. DeepMind je zařazován mezi systémy strojového učení, jelikož se tahy hry Go naučil pomocí vstupního souboru dat. (Rouse, 2018)

2.3 Hluboké učení

Hluboké učení je označováno za podobor strojového učení. Neuronová síť, simulující lidské chování, využívá k řešení reálných problémů některé techniky strojového učení. Hluboké učení může být velmi nákladné a vyžaduje pro tréninkovou fázi obrovské množství dat, a to z důvodu potřeby naučení se širokého množství parametrů. K ilustraci toho jak probíhá učební proces algoritmu hlubokého učení lze využít příkladu zobrazení kočky.(Velocity, 2017) Do algoritmu se vloží velké množství snímků kočky za účelem naučení a rozeznání i malých detailů. Algoritmus je poté schopen rozlišit jestli jde o lišku, geparda či pantera. Vyhledávání na základě textu, odhalení podvodů a spamu, rozpoznání rukopisu či hlasu, překlad do různých jazyků a prognóza vývoje titulů na kapitálových trzích jsou dalšími příklady využití hlubokého učení. (Reese, 2017)

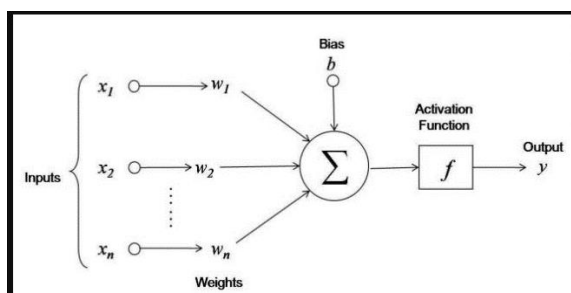
3 Neuronové sítě

V oblasti informačních technologií jsou neuronové sítě definovány jako hardwarový nebo softwarový systém navržený na základě neuronových operací v lidském mozku. Technologie neuronových sítí spadá do kategorie plýtkých neuronových sítí a hlubokého učení, patřící do oblasti umělé inteligence. Komerční aplikace této technologie se obecně zaměřuje na řešení komplexních úloh anebo na nacházení vztahů v nelineárně rozložených souborech dat. Mezi široké možnosti komerčního využití této technologie mimo jiné patří rozpoznání podpisového vzoru při využití bankovních služeb (Srinivasan, 2006), transkripce hlasové nahrávky do textu (Google, 2019), předpověď počasí (Abhishek, 2012), autonomní vozidla (Kubota, 2019) anebo predikce vývoje cen při obchodování na kapitálových trzích. (Gill, 2019)

Neuronová síť obvykle zahrnuje velké množství souběžně pracujících procesorů, uspořádaných do několika vrstev. První vrstva obsahuje hrubé vstupní informace (analogie k optickým nervům zpracovávající vizuální informace). Každá následující vrstva přijímá výstup z úrovně, která ji předchází, a ne z hrubé vstupní informace (stejně jako neurony vzdálenější od optického nervu přijímají signály od neuronů, které jsou k nim blíže). Výsledkem poslední vrstvy neuronové sítě je výstup systému. Každý neuron zpracovávající data, obsahuje svůj vlastní okruh znalostí, zahrnující již dříve zpracovanou informaci a jakékoliv naprogramované anebo sítí vyvinuté pravidlo. Jednotlivé neurony jsou vzájemně silně propojeny, což znamená, že každý neuron úrovně n je propojen s řadou dalších neuronů na úrovni $n-1$ (vstupy) a také s neurony na úrovni $n+1$, jelikož poskytuje vstup pro tyto neurony. Vrstva výstupu tvořící výsledek, může být složena z několika neuronů. Zásadní pro neuronové sítě je jejich přizpůsobivost. Na základě učení z prvotních testů se sami modifikují, a díky tomu poskytují kvalitnější informace pro následující běhy. Základní model učení se zaměřuje na vážení vstupních dat. Skrze tento postup neuron vyhodnocuje důležitost vstupní informace přicházející z každého předcházejícího neuronu. Ty vstupy, které přispívají k získání správného výstupu, jsou ohodnoceny vyššími vahami. (Rouse, 2018)

Výše popsaný systém neuronové sítě může být viděn na obrázku č. 1.

Obrázek 1 Základní fungování neuronové sítě



Zdroj: Mishra, 2017

V rovnici níže n vyjadřuje počet vstupních proměnných neuronu. x_1, x_2 až x_n jednotlivě znázorňuje vstupní proměnné a w_1, w_2 a w_n jsou příslušné váhy jednotlivých vstupních proměnných. Bias neuronu je označován b , $f(m)$ transferová funkce a m funkce souhrnu. Znakem y je označován požadovaný výstup systému.

$$y = f\left(\sum_{n=1}^n x_n w_n + b\right) = f(m)$$

3.1 Průběh neuronové sítě

3.1.1 Aktivační funkce

Součástí každého neuronu v síti je aktivační funkce, která určuje výstup neuronu. Její využití spočívá v určení a pracování s lineárními a nelineárními vztahy uvnitř sítě. (Missinglink.ai, 2017)

Binární step funkce – K aktivaci této funkce dochází, pokud je překročena určitá hranice. Pokud je vstup nad anebo pod danou hranicí, neuron je aktivován a posílá signál do následující vrstvy. Nevýhodou step funkce je, že nedovoluje pracovat s několika výstupy. To znamená, že nedokáže rozřadit vstupy do více kategorií.

Rovnice binární step funkce:

$$f(x) = \begin{cases} 0 & \text{pokud } x < 0 \\ 1 & \text{pokud } x \geq 0 \end{cases}$$

Lineární aktivační funkce

Rovnice lineární aktivační funkce:

$$f(x) = w_i * x_i$$

Do této funkce je vložen vstup, který je vynásoben danou váhou přiřazenou pro daný neuron a na tomto základě vytvoří výstupní signál. Výhoda lineární aktivační funkce nad binární step funkcí je ta, že dovoluje pracovat s více výstupy. Nicméně tato funkce trpí dvěma hlavními problémy:

V tréninkové fázi modelu není možné využít backpropagation (gradient descent) – derivace funkce je konstanta, která nemá žádný vztah vůči vstupu x . Tím pádem funkce neumožňuje vrácení se zpět a určit, které váhy vstupních neuronů mohou poskytnout kvalitnější prognózu.

Všechny vrstvy neuronové sítě jsou sloučeny do jedné – Jelikož je poslední vrstva sítě lineární funkcí první vrstvy, v případě lineární aktivační funkce nezáleží, kolik vrstev neuronová síť obsahuje. Lineární aktivační funkce tedy způsobí, že neuronová síť obsahuje pouze jednu vrstvu. Neuronová síť obsahující lineární aktivační funkci odpovídá lineárnímu regresnímu modelu. Kvůli tomu má jen omezené možnosti ve zpracování komplexních parametrů vstupních dat.

Nelineární aktivační funkce

Většina současných modelů neuronových sítí využívá nelineární aktivační funkce. Umožňují modelu komplexně propojit vstupy a výstupy sítě, což je podstatné pro učící proces sítě a modelaci komplexních dat jako časových posloupností, snímků, videí, hlasových záznamů a dalších nelineárně rozložených souborů dat.

Pomocí nelineárních funkce lze vyřešit následující nedostatky lineárních aktivačních funkcí:

- Umožňují backpropagaci, jelikož obsahují derivovatelnou funkci navázanou na vstupy.
- Umožňují vytvořit několik vrstev neuronů a tedy vytvořit hlubokou neuronovou síť. Několik skrytých vrstev neuronů umožňuje síti pracovat s komplexními datovými soubory s vysokou přesností. (Missinglink.ai,2017)

Mezi základní čtyři nelineární aktivační funkce patří (Walia,2017):

a) Logistická funkce:

Její hodnoty jsou v rozmezí 0 -1 a má podobu S-zakřivené křivky. Je velmi jednoduše použitelná, nicméně disponuje několika nevýhodami. Mezi ně patří problém mizejícího gradientu a pomalá konvergence funkce. Dále tím že se střed hodnot výstupu se nenachází v nule, což ztěžuje optimalizaci. Díky tomu se aktualizovaný gradient může posouvat různými směry.

Rovnice logistické funkce:

$$f(x) = \frac{1}{1 + e^{-x}}$$

b) Hyperbolická tangent funkce (Tanh)

Jelikož se hodnoty této funkce nachází v rozmezí -1 – 1, je vyřešen problém výstupu. Tato metoda je pro optimalizaci vhodnější než logistická funkce, a proto je v praxi více používána. Nicméně stále trpí problémem mizejícího gradientu.

Rovnice Tanh funkce:

$$f(x) = \tanh(x) = \frac{2}{1 + e^{-2x}} - 1$$

c) ReLU

Tato funkce je velmi jednoduchá a lehce použitelná. Jelikož řeší problém mizejícího gradientu je v dnešní době často využívána k tvorbě modelů hlubokého učení. Její nevýhodou ale je, že může být použita pouze ve skrytých vrstvách neuronové sítě.

$$f(x) = \begin{cases} 0 & \text{pokud } 0 > x \\ x & \text{pokud } x \geq 0 \end{cases}$$

d) Radial basis funkce (RBF)

RBF funkce označuje funkci, která každému vstupu přiřazuje reálnou hodnotu a hodnota vzniklá jako výstup funkce leží vždy v absolutní hodnotě (jedná se o míru vzdálenosti, která nemůže být negativní). Obvykle se používá euklidovská vzdálenost, tedy přímka spojující dva body v euklidovském prostoru. RBF funkce se používají jako aktivační funkce k aproximaci funkcí v neuronové síti. Na základě následující rovnice je sestavena RBF síť:

$$y(x) = \sum_{i=1}^N w_i \phi(\|x - x_i\|)$$

Aproximace funkce je vzhledem k váhám w diferencovatelná. Tyto váhy jsou naučeny skrze iterativní proces, což je běžná metoda používající se v práci s neuronovými sítěmi. (Nicholson, 2019)

3.1.2 Forwardpropagation - První fáze forwardpropagation nastává tehdy, když jsou do sítě vloženy tréninková data, která projdou celou neuronovou sítí, za účelem kalkulace prognózy budoucího vývoje. Neurony dané vrstvy obdrží vstupní data z vrstvy předcházející, která transformují a zašlou do vrstvy následující. Jakmile data projdou všemi vrstvami a všechny neurony provedou své kalkulace, výstupní vrstva stanoví konečnou prognózu (label) budoucího vývoje. (Missinglink.ai, 2017)

3.1.3 Ztrátová funkce – Za účelem odhadu ztráty (chyby) a ohodnocení správnosti a porovnání vytvořené predikce budoucího vývoje se správným výsledkem, se využívá ztrátové funkce (v případě učení s učitelem, kde známe očekávané budoucí hodnoty). Ideálně by chyba měla dosahovat nulové hodnoty, tedy žádného rozdílu mezi odhadovanou a očekávanou hodnotou. Díky tréninkové fázi modelu se váhy propojující neurony mění, až do té doby než je dosaženo přesné predikce. (Missinglink.ai, 2017)

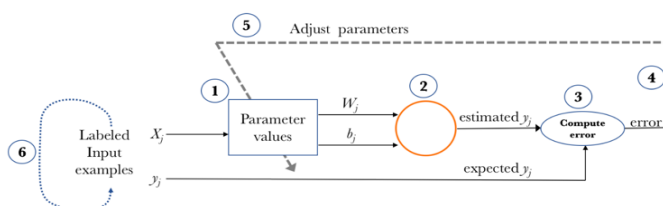
3.1.4 Backpropagation – Jakmile je chyba spočítána, tato informace je zaslána sítí zpět. S počátkem ve vrstvě výstupu se informace o chybě šíří zpět do všech neuronů ve skryté vrstvě. Neurony ve skryté vrstvě obdrží pouze část celkové informace o chybě, což záleží na jejich relativním podílu, kterým každý neuron přispívá k vytvoření výstupu. Tento proces je vrstvu po vrstvě několikrát zopakován, dokud všechny neurony v síti neobdrží informaci o ztrátě, která vyjadřuje jejich relativní podíl na celkové chybě. (Missinglink.ai, 2017)

3.1.5 Gradient descent – Nyní když je informace o chybě zaslána sítí zpět, mohou být upraveny váhy propojení mezi neurony. Cílem je při tvoření další predikce dosáhnout co nejnižší hodnoty chyby. Pro tento účel se využívá gradient descent. Za pomoci výpočtu derivace (gradientu) ztrátové funkce, (určuje jakým způsobem dosáhnout celkového minima chyby) se v malých krocích mění příslušné váhy propojení. Tento postup je použit pro každé opakování u všech souborů dat, která jsou do sítě vložena. (Missinglink.ai, 2017)

3.2 Princip učení neuronových sítí

Neuronová síť je tvořena ze vzájemně propojených neuronů a současně každé propojení neuronové sítě je spojeno s váhou, která určuje míru významnosti tohoto propojení násobenou vstupní hodnotou. Každý neuron sítě obsahuje aktivační funkci, která určuje výstup neuronu. Aktivační funkce je použita k vyjádření nelinearity modelovací schopnosti sítě. Proces učení sítě zpracovávat hodnoty parametrů je označován za trénování sítě. Trénovací fáze je pravou podstatou hlubokého učení a je popsána jako iterativní proces postupující vpřed a nazpět skrze vrstvy neuronů. Postup vpřed je označován za „forward-propagaci“ a postup nazpět za back-propagaci. Za první fázi neuronové sítě je označována forward-propagace, tedy proces kdy jsou do sítě vložena tréninková data, která projdou celou sítí za účelem výpočtu prognózovaných hodnot výstupu (labels). Vstupní data prochází sítí takovým způsobem, aby všechny neurony zohlednily informace, které získaly z předchozí vrstvy, a takto upravené je předaly neuronům ve vrstvě následující. Jakmile data projdou všemi vrstvami, a neurony provedou své kalkulace, výstupní vrstva dokáže ze vstupních dat určit prognózu hodnot výstupu. V případě metody učení s učitelem se k odhadnutí chybovosti a k porovnání a změření souladu výsledku prognózy, vzhledem k správnému řešení, použije ztrátová funkce. Ideálně by ztráta měla být nulová, což znamená žádný nesoulad mezi prognózovanou a očekávanou hodnotou. Proto se v tréninkové fázi sítě váhy propojení mezi neurony postupně mění, dokud není dosaženo optimální prognózy výstupu. Jakmile je chybovost spočítána, tato informace je zaslána skrze back-propagaci zpět. S počátkem ve výstupní vrstvě, je tato informace rozšířena zpět do všech skrytých vrstev, které přímo působí na výstup. Nicméně, na základě váhy jakou každý jednotlivý neuron přispěl k původnímu výstupu, obdrží neurony ve skryté vrstvě pouze část informace celkové chyby. Tento postup je opakován vrstvu po vrstvě, dokud všechny neurony v síti neobdrží informaci o chybě, kterou přispívají k celkové chybovosti. (Torres, 2018)

Obrázek 2 Princip učení neuronové sítě



Zdroj: Torres, 2018

Nyní když je informace o chybovosti rozšířena zpět, mohou být upraveny váhy propojení mezi neurony. Snahou je, aby při příští tvorbě predikce skrze síť, bylo dosaženo ztráty co nejbližší k nule. Pro tento účel se používá metoda nazývaná gradient descent. Za pomoci kalkulace derivace (gradient) ztrátové funkce, která umožňuje zjistit kterým směrem (descent) se ubírat k dosažení minimální chyby, tato metoda pozměňuje dané váhy. Tato metoda je použita na každou skupinu dat, kterou vložíme do sítě k iteraci. (Torres, 2018)

3.3 Metody učení

V rámci strojového učení se berou v úvahu hlavní dva typy úloh, a to učení s učitelem a učení bez učitele. Hlavním rozdílem mezi těmito typy úloh je ten, že u učení s učitelem existují dané hodnoty výstupu, kterých by skrze umělou inteligenci mělo být dosaženo. Po vložení vzorku dat a požadovaného výstupu je tedy cílem učení s učitelem naučení sítě takové funkce, která nejlépe popisuje vztah mezi vstupem a výstupem. Naproti tomu učení bez učitele nevyužívá předem určené hodnoty výstupů, takže jeho cílem je odvodit přirozenou strukturu, která se mezi poskytnutými daty nachází. Dalšími popisovány metodami strojové učení jsou Analýza hlavních komponent, Kombinace učení s učitelem a bez učitele a Zpětnovazebné učení.

3.3.1 Učení s učitelem

Učení s učitelem nachází typické využití v klasifikaci dat za účelem propojení vstupů s požadovanými výstupy, či u regrese, kde je zapotřebí propojit vstupy se spojitými výstupy. Mezi běžné algoritmy v rámci učení s učitelem patří logistická regrese, naive bayes, metoda podpůrných vektorů, umělá inteligence neuronových sítí či metoda random forests. Cílem regrese i klasifikace je najít ve vstupních datech určité vztahy či strukturu, které efektivně umožní získat správný výstup. Nicméně správnost výstupu je zcela určena skrze vložená tréninková data, takže pokud předem určené hodnoty výstupu model považuje za správné, stále to neznamená, že tyto hodnoty budou využitelné v reálných situacích. Chybná očekávání výstupu tedy omezují efektivnost modelu. Při použití metody učení s učitelem se musí vzít v úvahu komplexnost modelu a dilema vychýlení a rozptylu. Oba tyto aspekty spolu vzájemně souvisí. Komplexnost modelu se vztahuje ke komplexnosti funkce, kterou se neuronová síť pokouší naučit. Vhodná úroveň komplexnosti modelu je obecně určena skrze povahu tréninkových dat. Pokud je využito malého množství dat anebo pokud jsou rovnoměrně rozloženy v různých možných scénářích, měl by být zvolen model s malou komplexností. To zejména proto, že pokud by byl zvolen model s vysokou komplexností, došlo by při použití malého množství dat k jeho přeoptymalizovatelnosti. Přeoptymalizovatelnost se vztahuje ke schopnosti naučení se funkce, která funguje spolu s tréninkovými daty, ale není obecná ve vztahu k dalším datovým charakteristikám. Jinými slovy dochází pouze k naučení se simulace tréninkových dat bez zohlednění jejich aktuálního trendu či struktury, které vedou k výstupu. Také dilema vychýlení a rozptylu ovlivňuje obecnost modelu. V každém modelu existuje rovnováha mezi vychýlením, což je termín pro konstantní chybu a rozptylem, který měří rozsah chyby mezi různými soubory tréninkových dat. Tedy vysoká odchylka a nízký rozptyl může být charakteristikou pro model, který konstantně chybí ve 20 % případů. Oproti tomu malá odchylka a vysoký rozptyl určuje model, který tvoří chybné předpoklady kdekoli mezi 5 % až 50 %, což záleží na použitých datech. Typické pro tyto dvě charakteristiky je, že se pohybují v opačném směru. Tedy zvyšující se odchylka povede k nižšímu rozptylu a naopak. Obecně, zvyšující se odchylka (a snižující se rozptyl) povede k modelům s relativně pevnou úrovní výkonnosti. Navíc za účelem tvorby modelu, který zohledňuje co nejvíce hledisek, by velikost

rozptylu dat v modelu měla odpovídat velikosti a komplexnosti tréninkových dat. Tedy na učení se z malých datových souborů, by se měly využívat modely s nízkou odchylkou. Zatímco velké komplexní datové soubory využijí za účelem naučení se celé struktury dat modely s vyšší odchylkou. (Soni, 2018)

3.3.2 Učení bez učitele

Učení bez učitele patří mezi metodu strojového učení, která má za úkol najít strukturu závislosti mezi vstupními daty. Data vložená do algoritmu bez učitele nemají danou očekávanou výstupní hodnotu. To znamená, že jsou vloženy pouze vstupní proměnné X bez odpovídajícího výstupu a algoritmus musí v datech sám objevit jejich strukturu. (Jhunjhunwala, 2019)

3.3.2.1 Clustrování

Mezi jeden z technik učení bez učitele patří clustering, tedy seskupení dat do jednotlivých clusterů. V rámci souboru dat je clusterovací algoritmus použit k rozřazení jednotlivých dat do daných skupin. V rámci teorie se předpokládá, že data se stejnými vlastnostmi budou rozřazena do stejných skupin a v ostatních skupinách se budou nacházet data s odlišnými vlastnostmi. Clustrování je velmi jednoduše použitelná a efektivní technika, která do dat umožňuje hlubší pohled. Tato metoda nejčastěji nachází své využití zejména v biologii, medicíně a průzkumu trhů. (Rokach, 2005)

Hierarchické clustrování

Tyto metody clustrování sestavují dané clustery pomocí rekursivního rozdělování buďto přístupem postupujícím od shora dolů anebo naopak od spodu nahoru.

Tyto metody mohou být rozčleněny na následující:

- Aglomerativní hierarchické clustrování – Každý soubor dat představuje jednotlivý cluster, kdy jsou všechny clustery spojeny dohromady. Tento postup pokračuje, dokud není dosaženo požadované clusterové struktury.
- Rozdělovací hierarchické clustrování – Všechny soubory dat patří do jednoho clusteru, který je v souladu s touto metodou rozdělen na další pod clustery. Tento postup pokračuje, dokud není dosaženo požadované clusterové struktury.

Výsledkem hierarchické metody clustrování je dendrogram, který představuje hierarchické vztahy mezi jednotlivými soubory dat. Jeho hlavním využitím je nalezení nejlepšího způsobu alokace datových souborů do clusterů. Seskupení nebo rozdělení clusterů je prováděno na základě zvolené metriky podobnosti, tak aby došlo k optimalizaci zvoleného kritéria (příkladem může být suma čtverců). (Rokach, 2005)

Reálné využití tohoto typu clustrování může být mimo jiné nalezeno ve zdravotnictví při hledání zdrojů virové nákazy, v biologii při určení evoluce jednotlivých živočišných druhů anebo ve velkoobchodě při segmentaci zákazníků. (Kilitcioglu, 2018)

K – means clustrování

Tento typ clustrování také patří mezi metody učení bez učitele, za použití dat neobsahující informaci o požadovaném výstupu. Cílem tohoto algoritmu je rozdělit data do skupin, jejichž počet je určen proměnou K . Algoritmus využívá interačního postupu a na základě zvolených parametrů přiřazuje daná data do jedné z K skupin. Data jsou tedy clustrovány na základě podobnosti ve svých parametrech. Výstup K-means clustrovacího algoritmu je následující:

- Centroid K clusterů, který může být použit k určení výstupní informace pro nová data
- Výstupní informace pro trénovací data, kdy každé z dat je přiřazeno k jednotlivému clusteru

Spíše než zvolení skupin dat před tím než jsou data zpracována, clustrování umožňuje nalézt skupiny, které vznikly přirozeně. Každý centroid clusteru představuje skupinu parametrů, které určují výslednou podobu skupiny. (Trevino, 2016) Praktické využití tohoto přístupu clustrování může být nalezeno při identifikaci lokalit s vysokou kriminalitou, detekci pojistných podvodů či při klasifikaci dokumentů. (Raghupathi, 2018)

3.3.2.2 Autoencodery

Autoencodery označují jednoduchou třívrstvou neuronovou síť, kde jsou výstupní neurony přímo propojené se vstupními neurony. Obecně je počet skrytých vrstev mnohem menší než počet viditelných vrstev. Úlohou tréninkové fáze sítě je minimalizovat chybovost sítě, tedy nalezení efektivní struktury ve vstupních datech. (Sayantini, 2019)

Autoencodery označují neuronovou síť, která využívá metodu učení bez učitele za využití backpropagace. Používají se pro snížení velikosti vstupů a pokud je zapotřebí původních dat, lze je z komprimovaných dat získat zpět.

Architektura autoencoderů může být popsána následovně:

- Encoder: tato část sítě komprimuje vstupní data. Vrstvy encoderů tedy zašifrují vstupní snímek do komprimované podoby ve zmenšené velikosti a vytvoří jeho zkreslenou podobu.
- Code: tato část sítě představuje zkomprimovaný vstup, který vstupuje do decoderu. Rozhoduje, které parametry pozorovaných dat nesou podstatnou informaci, a které mohou být vyřazeny.
- Decoder: tato vrstva sítě rozšifruje zašifrovaný snímek do jeho původní velikosti. Rozšifrované verze snímku obsahuje pouze data, která byla v předchozí části vyhodnocena jako relevantní.

Praktické využití autoencoderů může být nalezeno při převádění černobílých snímku na barevné, odstranění datového šumu, redukování datových rozměrů anebo odstranění vodotisku na snímku. (Sayantini, 2019)

3.3.3 Analýza hlavních komponent

Analýza hlavních komponent patří mezi statistické postupy převádějící soubor pozorování korelovaných proměnných do souboru lineárně nekorelovaných proměnných, nazývaných hlavní komponenty. Počet hlavních komponent se buď rovná, nebo je menší než počet původních proměnných. Jinými slovy se tento postup snaží určit podprostor, ve kterém data přibližně leží. (Wang, 2014)

Příkladem může být soubor dat obsahující sto indikátorů, který slouží k vytvoření modelu. Z důvodu zvýšení rozptylu a celkové komplexnosti modelu by bylo kvůli přeoptymalizovatelnosti použití všech sta indikátorů vysoce neefektivní. Namísto toho se pomocí analýzy hlavních komponent vyhledá deset nejvíce korelovaných proměnných, které jsou za účelem vytvoření hlavních komponent lineárně zkombinovány. Tyto hlavní komponenty mohou dále sloužit jako hlavní charakteristiky modelu. (Walia, 2018)

3.3.4 Kombinace učení s učitelem a bez učitele (Semi-supervised learning - SSL)

Jedná se o metodu strojového učení, která je kombinací metod učení se s učitelem a učení se bez učitele. U učící metody s učitelem se použitý algoritmus trénuje na základě datového souboru, kde každý údaj zahrnuje výstupní informaci. Na základě těchto informací algoritmus identifikuje vztahy mezi cílovanou proměnou (target) a zbytkem datového souboru. Oproti tomu u metody učení bez učitele se algoritmus učí na základě datového souboru, ve kterém nejsou zahrnuty informace výstupních proměnných. U SSL metody se algoritmus učí na základě obou zmíněných metod, kdy použitý datový soubor obsahuje jak předem určené hodnoty výstupu, tak i data, kdy tyto předem určené hodnoty výstupu zahrnuté nejsou. Využití této metody je tedy vhodné pokud je zapotřebí sestavit přesný model a není k dispozici dostatečné množství dat, kde jsou předem určené hodnoty výstupu. Tato metoda mimo jiné nachází své reálné využití při analýze hlasu, rozeznávání obrazových snímků anebo u bioinformatiky. (Castle, 2018)

3.3.5 Zpětnovazebné učení (Reinforcement learning)

Tento přístup strojového učení inspirován behaviorální psychologií a je podobný procesu, kdy se dítě učí nový úkol. Oproti ostatním metodám strojového učení se odlišuje tím, že algoritmus nemá přímo nastaveno, jak vykonat zadanou úlohu, ale učící se sám na základě svých chyb a úspěšně splněných úloh. Příkladem může být autonomní vozidlo, či program hrající partii šachu, kdy je „odměna“ získána za splnění konkrétní úlohy, což může být přijetí do požadované destinaci anebo výhra v šachové partii. Naopak penalizace nastává tehdy, pokud algoritmus svou úlohu nesplní, tedy pokud například sjede z cesty nebo ve hře dostane šach-mat. V průběhu učící fáze se algoritmus rozhoduje s cílem dosažení maximální odměny a minimální penalizace. Výhoda tohoto přístupu umělé inteligence je ta, že dovoluje algoritmu naučit se postup v požadované úloze bez naprogramování postupu ke splnění konkrétní úlohy. (Techopedia, 2019)

3.4 Vývoj technologie neuronových sítí

Neuronové sítě a metody strojového učení jsou v dnešní době již součástí každodenního pracovního procesu v mnoha oblastech, a to zejména v oblasti lékařského výzkumu, autonomních vozidel, online prodeji, u virtuální asistenci či v obchodování na kapitálových trzích. Následující část se zaměřuje na počátky vzniku a časový vývoj této technologie. Dále také na současné využití této technologie v mnoha oblastech.

3.4.1 Historie neuronových sítí

Za základy pro pokročilou technologii budoucnosti v podobě umělé inteligence neuronových sítí se v roce 1943 zasadili Warren McCulloch a Walter Pitts. Tito dva vědci z oblasti kybernetiky vyvinuli matematický model umělé inteligence neuronových sítí za využití prahové logiky k imitaci fungování neuronů v lidském mozku. (McCulloch, 1943) Poté v roce 1958 americký psycholog, zajímající se o umělou inteligenci, Frank Rosenblatt vytvořil „Perceptron“ model, který jako první svého druhu byl schopen provést rozpoznávání vzorů. Nicméně Marvin Minsky a Seymour Papert našli v uvedeném modelu mnoho chyb, které za použití back-propagace vyřešil v roce 1974 Paul Werbos. (Werbos, 1974)

V roce 1990 byla pro predikci vývoje časových řad vytvořena rekurentní neuronová síť. Za jejího představitele je považován Jeffrey Elman. (Elman, 1990) Velmi populární LSTM neuronové sítě byly představeny roku 1997 akademickými pracovníky Hochreiter and Schmidhuber. (Hochreiter, 1997) Mezi další pokračovatele ve vývoji rekurentních neuronových sítí a hlubokých feed-forward sítí se řadí vědecký tým Jürgena Schmidhubera, který je často označován za otce moderní umělé inteligence. Tento tým mezi lety 2009 a 2012 rozvinuly zmíněné typy neuronových sítí, které v oblasti rozpoznávání vzorů a strojového učení zvítězily v osmi mezinárodních soutěžích. (Joshi, 2018)

3.4.2 Neuronové sítě v současnosti

Vývoj umělé inteligence neuronových sítí v současnosti prochází exponenciálním růstem a to i díky rozvoji rozšířené reality, strojového učení a umělé inteligence. Spojení neuronových sítí s ostatními technologiemi umožňuje jejich aplikaci v různých podobách.

Jednou z aplikací neuronových sítí v kombinaci s ostatními technologiemi jsou chatboti, kteří v současnosti nachází široké uplatnění v zákaznické podpoře u většiny klíčových organizací. Před vynalezením chatbotů musely firmy najmout celé oddělení zákaznické podpory, jehož náplní bylo zodpovídání klientských dotazů. Nicméně při použití chatbotů je celý proces automatizován a poskytuje plnou zákaznickou podporu při minimálním lidském zásahu. (Accenture, 2016)

Dále to této kategorie patří i virtuální asistenti Siri, Google Assistant a Cortana kteří napodobují lidskou konverzaci a vykonávají jednoduché úkoly, mezi které patří objednání taxi, správa upomínek, poskytnutí předpovědi počasí či nákup vstupenky do kina. (Martindale, 2018)

Své uplatnění našly neuronové sítě i v online maloobchodu, kdy v kombinaci s technologií strojového učení, předpovídají poptávku po zboží na základě minulých a současných nákupů. (Qi, 2001) Navigační služby jako Google Maps využívají neuronové sítě spolu s technologií GPS k poskytnutí optimálního návrhu cesty. (Chansarkar, 2000) Neuronové sítě spolu s technologií hlubokého učení shromažďují informace ohledně vytiženosti cesty a celkovou dopravní situaci k návrhu nejvíce efektivního spojení. Neuronové sítě také nabízejí využití v odvětví autonomních vozidel, kdy data shromážděna skrze strojové učení a neuronové sítě jsou použita k jejich testování. Dále technologičtí giganti jako Facebook, Google a Apple využívají tyto technologie pro identifikaci osob skrze rozeznání obličeje. (Ibarz, 2017)

Z rozvoje neuronových sítí také těží sektor zdravotnictví. Výzkumy ukazují, že technologie neuronových sítí může být použita k diagnóze a určení postupu léčby u vážných chorob jako je rakovina. (De Perio, 2019) Navíc by tato technologie mohla být využita pro objevení nových léků, zabraňujícím smrtelným chorobám. Dále některé z nových aplikací neuronových sítí také zahrnují předpověď zemětřesení využívajících seismogramy, či vytvoření uměleckých děl na základě existujících maleb. (Joshi, 2018)

Technologie neuronových sítí také nachází své uplatnění v několika finančních oblastech. Patří mezi ně predikce vývoje akciových cen, vývoj systémů pro obchodování s cennými papíry, prognózování ratingů u dluhopisů (Singleton, 1993), modelace vývoje kurzů zahraničních měn (Pacelli, 2011) či ve vyhodnocení rizika defaultu u půjček, hypoték a kreditních karet. (Kvamme, 2018) Finanční analytici se často snaží skrze metodu náhodné procházky vysvětlit vývoj finančních dat. Nicméně velké množství finančních dat se pohybuje podle nepravidelných vzorů, které za pomoci standardních statistických metod lze velmi obtížně identifikovat. (Swamy, 1997)

Tento úkol řeší právě neuronové sítě, které poskytují empirický rámec pro nalezení vzorových struktur ve vývoji časových řad, s následnou aplikací v odhadu vývoje cen akciových titulů.

3.4.3 Využití neuronových sítí ve financích

Největší uplatnění v oboru financí nachází neuronové sítě v tradingu a finančním prognózování. Neuronové sítě jsou využívány k řešení široké škály reálných úloh, a to například v odhadu budoucích cen, oceňování a zajištění skrze deriváty, odhadu vývoje zahraničních měnových kurzů, předpovídání defaultu, při výběru a odhadu výkonnosti akciových titulů, k určení složení a optimalizaci portfolia a také k ohodnocení finanční volatility.

3.4.3.1 Optimalizace portfolia

Při výběru vhodné investice musí investor vždy volit mezi optimálním rizikem a výnosem, takže je nezbytné stanovit vhodný poměr mezi těmito dvěma faktory. Optimalizace portfolia skrze analýzu těchto faktorů, maximalizace výnosů při daném riziku a v případě nutnosti rebalancování portfolia je při investování zásadní úlohou. Steiner a Wittkemper (1997) ve své práci zkonstruovali model pro optimalizaci struktury portfolia na bázi každodenního obchodování. Zatímco rozhodování u výběru akcií je odvozeno skrze nelineární dynamický kapitálový tržní model, odhady a předpovědi jsou tvořeny na základě modelu neuronových sítí. Za použití cen akcií německých společností od roku 1990 do roku 1994, tento model vedl k portfoliu, které překonalo výsledky tržního portfolia o 60 %. Hung, Liang a Liu (1996) při optimalizaci výkonnosti portfolia propojili arbitrážní teorii oceňování aktiv (APT) s modelem neuronových sítí a uvedli, že takto propojená strategie předčila výkonnost benchmarku a také výkonnost tradičního ARIMA modelu. (Kamruzzaman, 2006)

3.4.3.2 Odhad vývoje zahraničních měnových kurzů

Modelování vývoje zahraničních měnových kurzů je pro investiční společnosti podstatnou úlohou, a to zejména proto, že na přesné predikci měnových kurzů záleží investiční rozhodování. Vzhledem k tomu, že vývoj zahraničních měnových kurzů je nestacionární a nepravidelný, se tato předpověď stává velmi náročnou. Yao a Tan (2000) vyvinuli model neuronové sítě, ve kterém použili šest jednoduchých indikátorů a aplikovali ho na předpověď vývoje šesti různých měn vůči americkému dolaru. Ve výsledku bylo při porovnání výkonnosti s ARIMA modelem dosaženo daleko kvalitnějších výsledků. Dále Kamruzzaman a Sarker (2003) vyvinuli tři různé neuronové sítě za účelem předpovědi vývoje zahraničních kurzů a zjistili, že všechny tři modely dosáhly lepší výkonnosti než pět různých modelů založených na tradičních statistických metodách. (Kamruzzaman, 2006)

3.5 Silné stránky neuronových sítí

Neuronové sítě se vyznačují mnoha silnými stránkami, díky kterým jsou u vývojářů umělé inteligence více preferované než jejich alternativy. Mezi ty nejpodstatnější z nich patří následující:

- **Vysoká výkonnost při řešení úloh s mnoha proměnnými**

Při řešení úloh s pevně zadanými pravidly a požadavky a omezeným rámcem možných vstupů, nepředstavuje dosažení správných výsledků skrze výpočetní technologie náročný úkol. Příkladem může být obyčejná kalkulačka, kdy matematické zákonitosti nemohou být nikdy porušeny a lze se podle nich relativně jednoduše řídit. Nicméně identifikace vzorů řeči, diagnóza nemoci či předpověď budoucího vývoje finančních časových řad vyžaduje použití daleko více proměnných, kterým výpočetní technika musí porozumět. V rámci předpovědi vývoje finančních časových řad musí být bráno v úvahu mnoho tržních či netržních faktorů, které na vývoj řady přímo či nepřímo působí. Technologie neuronových sítí v této úloze historicky vykazuje velmi dobrou výkonnost a často dosahuje lepších výsledků než běžné statistické postupy. (Landman, 2019)

- **Použitelnost**

Neuronová síť se také vyznačuje silnou flexibilitou, a pokud je zpočátku dostatečně efektivně vyvinuta, může být použita téměř na cokoliv. Pomáhá například zaměstnancům identifikovat příčiny omezujících jejich produktivitu anebo vylepšení dopravních schémat pro jejich efektivnější řízení. Klíčovou funkcí neuronové sítě je efektivní učení. Tedy pokud existuje systém, který je schopen se naučit rozeznávat vzorové chování, může být použit pro jednoduché rozeznání vzorového chování v jakékoliv oblasti. (Landman, 2019)

- **Ukládání informací v celé síti**

Informace jsou stejně jako u tradičního programování ukládány v rámci celé sítě, a ne pouze na jednom místě v dané databázi. Pokud tedy na jednom místě sítě dojde ke ztrátě části informace, neznamená to omezení funkčnosti sítě jako celku. (Mijwel, 2018)

- **Možnost zpracování neúplné informace**

Neuronová síť po tréninkové fázi disponuje schopností vytvořit z dat výstup i za použití neúplné informace. Výkonnost sítě poté zejména závisí na důležitosti této chybějící informace. (Mijwel, 2018)

- **Tolerance chybovosti**

Zkreslení či chyba u jednoho či více neuronů nezabraňuje síti generovat výstup. Tento prvek činí z neuronových sítí systém tolerantní k chybovosti. (Mijwel, 2018)

3.6 Slabé stránky neuronových sítí

Technologie neuronových sítí disponuje i slabými stránkami, které ji omezují ve využití celkového potenciálu. Mezi tyto slabé stránky patří následující:

- **Náročnost datových vstupů**

Za účelem zdokonalení neuronové sítě musí všechny nejprve projít procesem učení, díky kterému začínají rozpoznávat vzory. Přestože současná technologie umožňuje zefektivnit učicí proces sítě lépe než kdykoliv dříve, stále je k dostatečné funkčnosti algoritmu nutné vložit nadměrný objem dat. V závislosti na typu využití sítě je zapotřebí tisíců rozdílných souborů dat, což při optimalizaci neuronové sítě způsobuje značnou časovou náročnost a limituje možnosti jejího využití. (Landman, 2019)

- **Finanční a časová nákladnost**

Vývoj neuronových sítí se vyznačuje značnou finanční i časovou nákladností. Výpočetní procesy potřebné ke zpracování všech souborů dat a proměnných vyžadují vysokou CPU a GPU výkonnost, která je nad rámec běžných systémů. Vývoj funkčního systému se tedy stává velmi nákladný, což může vést k odrazení mnoha vývojářů. (Landman, 2019)

- **Náročnost a neprůhlednost**

Při konstrukci algoritmu se vyvinutí neuronové sítě vyznačuje vyšší složitostí než při použití obecného pravidla. Sama náročnost naučení se jak vyvinout neuronovou síť způsobuje, že mnoho vývojářů této snahy zanechá. Navíc kvůli komplikovanosti neuronové sítě nelze přesně vysvětlit chování algoritmu, což přispívá k neprůhlednosti sítě v rámci vysvětlení získaných výsledků. Přesnost výstupů může být sice posouzena, avšak i pro profesionály se stává velmi náročné určit, jakým přesným způsobem bylo těchto výstupů dosaženo. (Landman, 2019)

- **Dlouhodobé využití**

Díky neuronovým sítím bylo v oblasti umělé inteligence již dosaženo značného pokroku. Nicméně při posouzení jejich dlouhodobého využití nemusí být tak výkonné jako jejich alternativy (Kernelové metody, klasická umělá inteligence). Jejich omezená efektivnost a komplikovanost odradila mnoho vývojářů od jejich vývoje. (Landman, 2019)

3.7 Typy neuronových sítí

V následující části jsou uvedeny tři základní architektury neuronových sítí, na jejichž základu je vyvíjeno několik typů sítí, které se používají v současnosti. První architekturou jsou Feed-forward neuronové sítě, do které spadá mimo jiné Konvoluční neuronová síť, která se používá u počítačového rozpoznávání snímků a obrazů s reálnou aplikací u autonomních vozidel. Další základní architekturou jsou Rekurentní neuronové sítě, široce využívané ve strojovém učení lidské řeči či prognóze vývoje cen na kapitálových trzích. Třetí základní architekturou neuronových sítí je Symetricky propojená neuronová síť. Tato architektura sítě se velmi podobá rekurentním neuronovým sítím a patří do ní Hopfieldova neuronová síť a Boltzmann machines.“

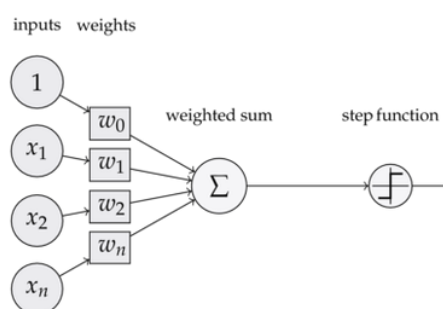
3.7.1 Feed-Forward neuronové sítě

Tato architektura neuronových sítí je považována za nejvíce obecnou. První vrstva sítě představuje vstup a poslední vrstva sítě výstup. Pokud se v síti nachází více než jedna skrytá vrstva, síť je označena za hlubokou neuronovou síť. Zpracování dat v neuronech každé vrstvy probíhá skrze nelineární funkci za použití zpracovaných dat z neuronů ve vrstvě předcházející. (Le, 2018)

3.7.1.1 Perceptron

V problematice strojového učení je perceptron algoritmus využívající metodu učení s učitelem, za účelem binárního členění dat do skupin a znázorňuje základní jednotku neuronu. Představuje typ lineárního členění, tedy třídící algoritmus, který své předpovědi tvoří na základě lineární funkce, kombinovanou se souborem vah přiřazené k charakteristikám vstupu.

Obrázek 3 Architektura perceptronu



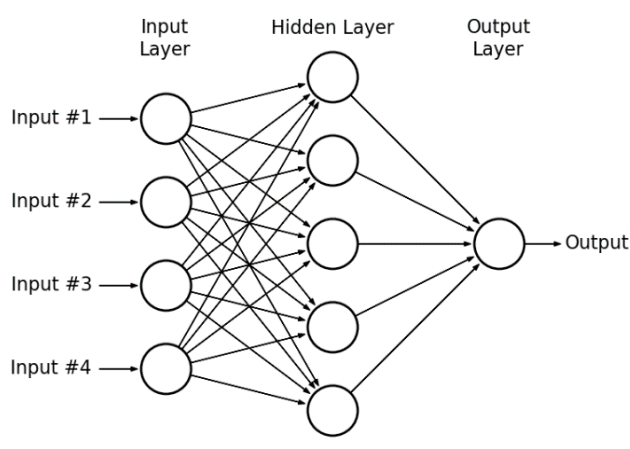
Zdroj: JustinB, 2017

Lineární členění dat definuje, jak mají být tréninková data rozčleněna do odpovídajících skupin. Pokud jsou data členěna do dvou skupin, poté také tréninková data musí být rozčleněna do dvou skupiny. (JustinB, 2017)

3.7.1.2 Několikavrstvý perceptron

Za nejjednodušší typ neuronové sítě je označován jednovrstvý perceptron, skládající se z jedné vrstvy výstupních neuronů, kdy vstupy jsou po přidělení příslušných vah přímo transformovány na výstup. Tímto způsobem funguje nejjednodušší feed-forward neuronová síť v podobě perceptronu. Nicméně v případě několikavrstvého perceptronu se jedná o více vrstev, ve kterých probíhají příslušné výpočetní operace, navzájem propojené feed-forward způsobem. Každý neuron dané vrstvy je přímo propojen s neuronem vrstvy následující. Pro mnoha případů v rámci reálné aplikace je za aktivační funkci použita logistická funkce. (Di Persio, 2016)

Obrázek 4 Architektura několikavrstvého perceptronu

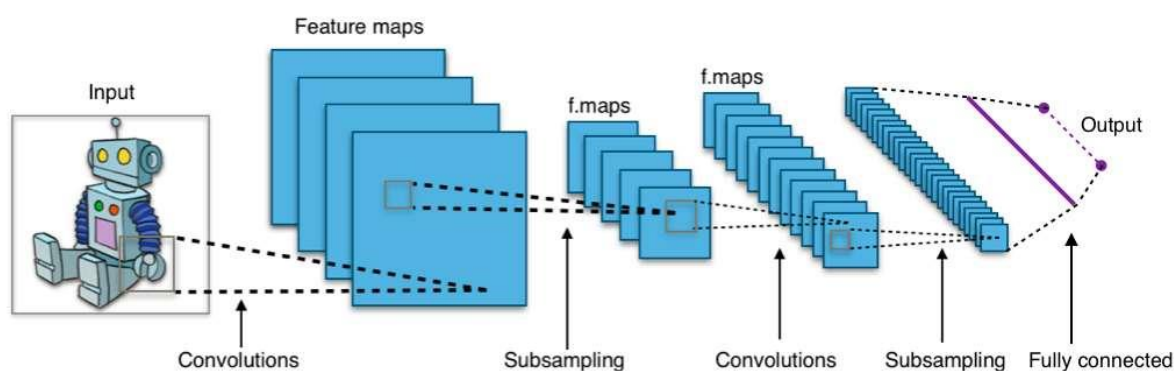


Zdroj: Zahran, 2015

3.7.1.3 Konvoluční neuronové sítě

Označení konvoluční neuronová síť je odvozeno od konvolučního operátoru (Dominguez, 2015). Jejich primární účel je spojován se získáváním prvků z vloženého snímku. Tento typ sítě zachovává prostorový vztah mezi pixely díky učením se jednotlivých prvků snímku. Často se využívají v oblastech jako identifikace obličejů, kdy snímač vyhodnotí vložený snímek v nízkém rozlišení, za účelem vyřazení oblastí, ve kterých se nenachází obličej. Poté důkladně zpracuje oblasti snímku ve vyšším rozlišení obličeje pro přesnější detekci. Další využití můžeme nalézt v oblasti autonomních vozidel, k zajištění bezpečnosti cestujících a ostatních vozidel. Díky možnosti vyhodnocení vstupů skrze mnoha parametrů poskytují konvoluční vrstvy vysokou všestrannost. (Shah, 2017)

Obrázek 5 Architektura konvoluční neuronové sítě



Zdroj: Yung, 2017

Konvoluční sítě využívají systém podobný několikavrstvému perceptronu, který byl navržen ke snížení náročnosti zpracování. Vrstvy tohoto typu sítě se skládají z vrstvy vstupů, vrstvy výstupů a skryté vrstvy zahrnující několik konvolučních vrstev. Odstraněním omezení a zvýšením efektivnosti při zpracování vede k systému, který při zpracování snímků a porozumění lidské řeči přináší daleko lepší výsledky. Běžné neuronové sítě nejsou typicky využívány ke zpracování snímků a z tohoto důvodu musí být snímky vkládány v částech s nižším rozlišením. Naproti tomu uspořádání neuronů u konvoluční sítě více připomíná uspořádání neuronů v čelním laloku lidského mozku, tedy v oblasti určené ke zpracování vizuálních stimulů. Vrstvy neuronů jsou uspořádány takovým způsobem, aby bylo pokryto celé zorné pole, což zabraňuje neucelenému zpracování snímku, ke kterému dochází u běžných neuronových sítí. (Rouse, 2018)

Vstup konvoluční neuronové sítě představuje soubor hodnot pixelů snímku. Skrytá vrstva sítě se skládá z několika různých vrstev sloužících k získání parametrů snímku a plně propojená vrstva je použita k rozeznání objektů vyobrazených na obrazovém snímku. Jádrem celé konvoluční neuronové sítě tvoří konvoluční operátor. CNN síť se skládá ze čtyř vrstev, a to konvoluční vrstvy, ReLU vrstvy, „Pooling“ vrstvy a plně propojené vrstvy. Konvoluční vrstva aplikuje na soubor pixelů obrazového snímku filtrační matici a provede konvoluční operace k získání preaktivačních výstupů. Preaktivační výstupy konvolučních

vrstev se nazývají mapy vlastností. Další vrstvou sítě je ReLU vrstva, která v rámci sítě zachytí nelinearitu v datech. V rámci aktivačních funkcí se ale také používá funkce tanh či logistická funkce. Všem pixelům obsahující záporné hodnoty přiřadí nulovou hodnotu. Ke zjištění skrytých charakteristik a vzorů v obrazovém snímku je původní snímek vložen do konvolučních a ReLU vrstev. V další vrstvě sítě se provádí podvzorkování získaných map vlastností (pooling), která sníží jejich rozměry. Tato vrstva používá několik různých filtrů za účelem identifikace různých částí snímku jako hrany, rohy a tělo snímku. Podvzorkovaná mapa vlastností je dále rozložena do vektorů, načemž při jejich spojení vznikne plně propojená vrstva, která je použita k rozeznání obrazových snímků. (Sharma, 2018)

3.7.2 Rekurentní neuronové sítě

U této architektury sítě je mezi jednotlivými vrstvami využíváno přímé cyklické propojení. To znamená, že při úpravě sítě je možné se vrátit nazpět a dále pokračovat ve směru následující posloupnosti. Vyznačují se komplikovanou dynamikou, což zvyšuje náročnost tréninkové fáze. Tento typ sítě má schopnost si ve své skryté vrstvě zapamatovat danou informaci po delší časový úsek. Nicméně kvůli náročnosti trénovací fáze není jednoduché využít jejich potenciál. (Le, 2018)

Rekurentní neuronové sítě se využívají v oblasti hlubokého učení a při vývoji modelů, které simulují aktivitu neuronů v lidském mozku. Nacházejí vysoké využití u případů, kdy je podstatné předpovědět následující vývoj a liší se od ostatních typů neuronových sítí tím, že ke zpracování časových řad používají zpětnou vazbu, čímž předávají informaci do konečného výstupu, což také může být časová řada. Tato zpětná vazba umožňuje síti dané informace uchovat, což je často označováno za paměť neuronové sítě. (Valkov, 2017)

Rekurentní neuronová síť (RNN) označuje typ neuronových sítí, kde výstup předchozího kroku je použit jako vstup kroku následujícího. U běžných neuronových sítí jsou všechny vstupy a výstupy na sobě vzájemně nezávislé. Nicméně pokud je zapotřebí např. předpovědět následující slovo věty, je nutné znát slova předcházející, k čemuž slouží paměť. Tuto problematiku díky využití skrytých vrstev vyřešily rekurentní neuronové sítě. Hlavní a nejpodstatnější složkou RNN jsou skryté vrstvy, které si pamatují informace z poskytnuté časové posloupnosti. (Rouse, 2018)

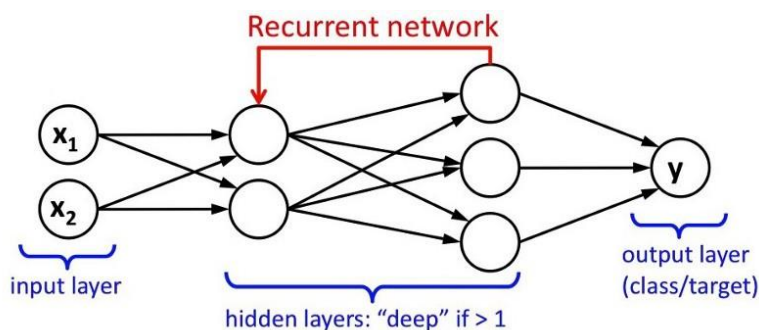
RNN disponuje pamětí, která zaznamenává všechny informace o provedených kalkulacích sítě. Za účelem získání výstupu aplikuje na každý ze vstupů určených pro splnění totožné úlohy stejný soubor parametrů, a tím oproti ostatním typům neuronových sítí snižuje komplexnost těchto parametrů.

Jak RNN fungují je ukázáno na následujícím příkladu:

Předpokládá se hluboká neuronová síť, obsahující vstupní vrstvu, tři skryté vrstvy a jednu vrstvu výstupu. Poté tak jako u ostatních neuronových sítí každá ze skrytých vrstev obsahuje svůj vlastní soubor vah a bias. To znamená, že každá z těchto skrytých vrstev je na sobě nezávislá, což jim neumožňuje si zapamatovat předchozí hodnoty výstupů. Poté RNN skrze přiřazení stejných hodnot vah a biasu do všech skrytých vrstev

sítě převede nezávislé aktivační funkce na aktivační funkce závislé. Tím sníží zvyšující se komplexitu parametrů a zapamatuje si každou hodnotu výstupu díky jejímu využití jako hodnoty vstupu v následující skryté vrstvě. Jelikož jsou hodnoty vah a biasu pro všechny tři skryté vrstvy totožné, mohou být tyto vrstvy sloučeny do jedné rekurentní vrstvy. (aishwarya.27, 2019)

Obrázek 6 Architektura rekurentní neuronové sítě



Zdroj: Valkov, 2017

RNN se využívají například u jazykových modelů, kde předpovídají následující písmeno ve slově, či následující slovo ve větě na základě písmen či slov, které jim předcházejí. Přesvědčujících výsledků dosáhl tento typ sítě při experimentu, kdy do sítě byly jako vstup použita Shakespearova díla za účelem vytvoření podobného díla. Tento typ sítě tedy umožňuje umělé inteligenci simulovat lidskou kreativitu, skrze pochopení významu a gramatiky naučených ze vstupních tréninkových souborů. (Rouse, 2018)

Pokud je vyvinuta taková síť, která správně odpovídá modelu známé posloupnosti hodnot, může být použita k predikci budoucích výsledků. Známým využitím pro tento přístup jsou predikce vývoje cen titulů na kapitálových trzích. (Shah, 2017)

Výhody rekurentní neuronové sítě

- RNN si je schopna zapamatovat všechny zpracované informace, což je možné využít u tvorbě predikcí vývoje časových řad
- RNN je možné použít spolu s konvolučními vrstvami za účelem efektivnějšího výběrů jednotlivých pixelů snímků

Nevýhody rekurentní neuronové sítě:

- Problém explodujícího (mizejícího) gradientu
- Vysoká náročnost tréninkové fáze sítě
- RNN není schopna zpracovat rozsáhlé časové řady, pokud používá Tanh nebo ReLU aktivační funkce

(aishwarya.27, 2019)

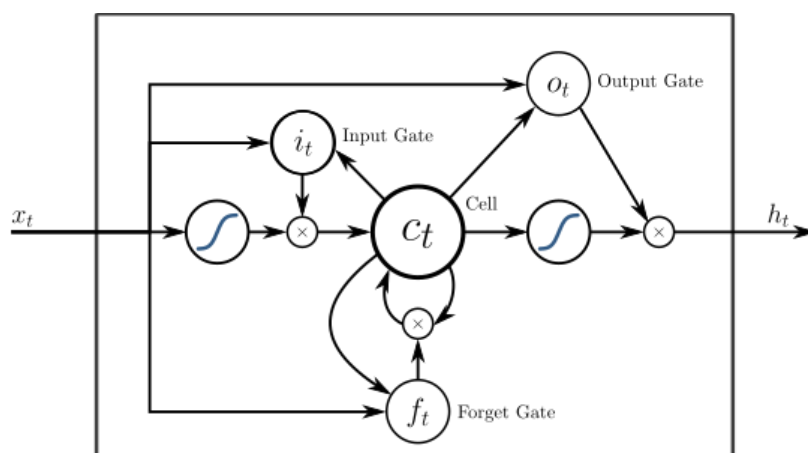
3.7.2.1 Problém mizejícího (explodujícího) gradientu

Rekurentní neuronové sítě se pokládají za zastaralé a na počátku 90. let se problém mizejícího gradientu stal jednou z hlavních překážek k zvyšování jejich výkonnosti. Tak jako přímka mezi dvěma body vyjadřuje změnu veličiny x spolu se změnou veličiny y , tak gradient vyjadřuje změnu v hodnotě chyby s ohledem na změnu váhy. Pokud neznáme gradient, nelze pozměnit váhy takovým způsobem, aby byla snížena chyba, a tím pádem se síť přestává učit. Rekurentní neuronové sítě se snaží nalézt propojení mezi konečným výstupem a událostí již mnoho kroků dozadu než dojde ke zpoždění, protože je velmi obtížné určit významnost časově vzdálených vstupů. To částečně nastává proto, že přitom jak informace prochází sítí je na mnoha úrovních znásobena. Tak jako v případě složeného úročení, pokud je násobeno hodnotou o něco vyšší než jedna, výsledná hodnota se stává velmi vysoká. Nicméně také naopak násobení hodnotou menší než jedna vytváří problém. Pokud na každé sázce jednoho dolaru vyhrájeme pouze 97 centů, poté se lehce může dojít k bankrotu. Díky tomu že vrstvy a časové kroky neuronové sítě jsou vzájemně propojeny díky multiplikaci, velikost vah gradientu se může lehce dostat na velmi nízké hodnoty, kdy téměř vymizí nebo naopak do velmi vysokých hodnot, kdy jsou váhy až moc významné. Nicméně významné hodnoty gradientu jsou relativně jednoduše řešitelné, protože mohou být omezeny. V opačném případě se ale hodnoty stávají moc nízké na to, aby s nimi počítač mohl při učení sítě pracovat, což je těžší úloha k řešení. Analogicky k případu, kdy gradient prochází mnoha vrstvami a postupně mizí, tak pokud je na data opakovaně aplikována logistická funkce, dochází k jejich zploštění až do té úrovně, kdy nelze pozorovat sklon funkce. (Nicholson, 2019)

3.7.2.2 Long Short-Term Memory rekurentní neuronové sítě (LSTM)

Long Short-Term Memory (LSTM) představuje specifický typ architektury rekurentní neuronové sítě, který byl navržen tak aby modeloval datové posloupnosti a jejich dlouhodobé dopady s vyšší přesností než obecné rekurentní neuronové sítě. Jedním z důvodů konstrukce LSTM sítí řešení problému mizejícího gradientu. (Sak, 2014)

Obrázek 7 Architektura LSTM neuronové sítě



Zdroj: Graves, 2017

Nevýhodou standardních rekurentních neuronových sítí představuje problém mizejícího gradientu, kvůli kterému neuronová síť ztrácí na výkonnosti, protože nemůže být patřičně trénována. K tomuto problému zejména dochází u více vrstevných neuronových sítí, které se používají ke zpracování velmi komplexních dat. Rekurentní neuronové sítě vytvořené na základě LSTM jednotek rozzařují data do paměťových buněk podle toho, jestli se jedná o data krátkodobá či dlouhodobá. Tento proces umožňuje rekurentní sítí vyhodnotit, která data jsou podstatná, a tedy by měly být zapamatována a vrácena zpět do sítě, a která data mohou být zapomenuta. (Olah, 2015)

LSTM nepředstavují podstatně odlišnou architekturu sítě než rekurentní sítě, ale liší se tím, že využívají k výpočtům ve skryté vrstvě rozdílnou funkci. LSTM neuron se skládá z pěti podstatných složek, které dovolují modelovat jak dlouhodobá tak krátkodobá data: „cell state, hidden state, input gate, forget gate and output gate“. Hlavní výhodou LSTM je jejich schopnost zapamatování si informace a práce s dlouhodobými datovými posloupnostmi. (Lazzeri, 2018)

LSTM sítě zvládají uchovat informační souvislosti mezi vstupy, přidáním smyčky, která dovoluje informaci kolovat od jednoho kroku k druhému. Při čtení textu je každé slovo chápáno na základě pochopení předchozích slov a ne tak, že je vše zapomenuto a u každé slovo je pochopeno bez předchozích souvislostí. Podobně tak i LSTM sítě berou v predikcích do úvahy minulé hodnoty vstupů. Na druhou stranu, v delším časovém období se stává stále méně pravděpodobné, že nový výstup bude záviset na velmi starém vstupu. Informace o závislosti na základě časové vzdálenosti je další důležitou informací, kterou by se neuronová síť měla naučit. Pro tyto účely se LSTM sítě učí za jakých podmínek si má síť informaci pamatovat a za jakých zapomenout skrze “forget gate“. (Moawad, 2018)

3.7.3 Symetricky propojené neuronové sítě

Tato architektura neuronových sítí je velmi podobná rekurentním neuronovým sítím s tím rozdílem, že ve spojeních mezi jednotlivými částmi sítě využívají v obou směrech stejných vah. Symetrické neuronové sítě se tedy mnoho lépe analyzují. Jsou také více omezené ve své činnosti, protože se řídí energií funkce. Tato architektura sítě bez skryté vrstvy je označována za Hopfieldovu neuronovou síť. Naopak pokud síť obsahuje skryté vrstvy, bývá označována za Boltzmann machines. (Le, 2018)

3.7.3.1 Hopfieldova síť

Nelineární části rekurentní sítě lze velmi těžko analyzovat, jelikož se mohou pohybovat v mnoha různých směrech. Mohou zůstat stabilní, oscilovat nebo se pohybovat v nesystematických trajektoriích, které pak nelze predikovat do budoucnosti. Neurony Hopfieldovy sítě disponují k vyhodnocování binárním prahem s vzájemným rekurentním propojením. V roce 1982 John Hopfield zjistil, že pokud jsou propojení symetrická, dojde k výskytu globální energie funkce. Každé binární nastavení sítě disponuje určitou energií a rozhodování skrze pravidlo využívající binární práh pomáhá síti dosáhnout minimální energie funkce. Praktickou cestou jak tento typ výpočtů využít, je využití paměti neuronové sítě jako energetického minima a v každém případě kdy si síť zapamatuje konfiguraci, snaží se vytvořit nové energetické minimum. Díky použití energetického minima představující paměť se získá obsahově adresovatelná paměť. Nalézt údaj pak lze jen za potřebí znalosti pouhé části obsahu a síť se stává odolnou vůči hardwarovému poškození. (Le, 2018)

Využití Hopfieldovy sítě může být nalezeno ve zpracování obrazových snímků a hlasu, vyhledávání v databázích či rozzaření obrazových snímků ve zdravotnictví. (MCELROY, 2019)

3.7.3.2 Restricted Boltzman Machines (RBM)

Obecně se za RBM označuje stochastická síť jejíž struktura je složena z viditelné vrstvy a skryté vrstvy. Viditelná vrstva obsahuje zkoumaná data a je spojena se skrytou vrstvou. Skrytá vrstva slouží k naučení se jakým způsobem ze vstupních dat získat jejich potřebné charakteristiky. Původně byl tento typ neuronové sítě určen k práci s binárními daty a až později se naučil pracovat i dalšími typy dat. V rámci RBM jsou spojení mezi neurony dvojsměrná a symetrická, což znamená, že toky informací sítě proudí oběma směry. Dále z důvodu zjednodušení interpretace výstupu, nejsou neurony stejné vrstvy mezi sebou propojeny. (Assis, 2018)

a) Boltzmann Machines

Boltzmann Machines označují generativní stochastické modely hlubokého učení, obsahující pouze dva typy neuronů, a to viditelné a skryté. V tomto typu neuronové sítě nejsou žádné neurony výstupů, což určuje stochastickou vlastnost modelu. Výstupem tedy nejsou hodnoty 1 a 0, díky kterým jsou naučeny vzory v datech, tak jak je typické u běžných neuronových sítí.

Další čím se tento typ neuronové sítě odlišuje od ostatních běžných typů sítí, je že disponuje propojením i mezi vstupními neurony. Ať už se jedná o viditelné či skryté neurony sítě, všechny jsou vzájemně propojeny. To jim umožňuje mezi sebou sdílet informace a generovat následná data. Při poskytnutí vstupních dat je tedy tento algoritmus schopen zachytit všechny jejich parametry, vzory a korelace. Boltzmann machines jsou tedy řazeny mezi neuronové sítě učící se na základě metody bez učitele. (Sharma, 2018)

b) Restricted Boltzmann Machines

RBM označují dvouvrstvé neuronové sítě (viditelné a skryté vrstvy) s generativními schopnostmi, jelikož mají schopnost se ze vstupního souboru dat naučit jeho pravděpodobnostní rozdělení. Využití RBM může být nalezeno v klasifikaci, regresi, společném filtrování a učení se datových charakteristik. V rámci tohoto typu sítě je každý neuron viditelné vrstvy propojen s neurony ve skryté vrstvě, nicméně nedochází k propojení neuronů ležících ve stejné vrstvě. Zmíněné omezení umožňuje efektivnější tréninkový algoritmus, než který je použit u obecného typu Boltzmann machines. (Sharma, 2018)

3.8 Další metody umělé inteligence

V následující části budou zmíněny další možné metody umělé inteligence s příklady jejich reálného využití. V této části bude zmíněna Metoda podpůrných vektorů, Algoritmus K-nejbližších sousedů, metoda „Decision Trees“, která je spojena s metodou „Decision Forest“.

3.8.1 Metoda podpůrných vektorů

Metoda podpůrných vektorů označuje typ algoritmu hlubokého učení, který pro rozřídění či regresi souborů dat využívá učení s učitelem.

V oblasti umělé inteligence a strojového učení systémy využívající učení s učitelem pracují jak se vstupními, tak výstupními daty. Tyto výstupní data jsou dále rozčleněna, což poskytuje základ pro jejich budoucí zpracování. Metoda podpůrných vektorů se využívá k rozřazení souborů dat do jednotlivých skupin, a to tak že na základě vzorových struktur algoritmus vytváří linie, kterými dané vzory rozřídí. Stejně jako ostatní systémy učení s učitelem, metoda podpůrných vektorů vyžaduje k trénování modelu označená data, tudíž soubory dat jsou označovány za účelem rozřídění. Tréninkové data pro metodu podpůrných vektorů jsou odděleně rozřazena do různých bodů a seskupena do oddělených skupin. Po zpracování řady tréninkových příkladů se mohou začít tyto systémy učit bez učitele. Cílem algoritmu je dosažení co nejlepšího rozdělení dat s maximální hranicí okolo linií. Metoda podpůrných vektorů bývá využívána při rozpoznávání textu či snímků. Umí pracovat s ručně psaným textem nebo nalézá využití v biologických laboratořích. Systémy učení s učitelem i bez učitele se mimo jiné používají u chat botů, autonomních vozidel a k rozpoznávání obličejů. (Rouse, 2017)

3.8.2 Algoritmus K-nejbližších sousedů

Algoritmus K-nejbližších sousedů vychází z předpokladu, že data s podobnými vlastnostmi se nachází blízko sebe. Označuje přístup ke klasifikaci dat, který odhaduje s jakou pravděpodobností je určitý datový bod součástí skupiny, na základě toho do jaké skupiny patří datové body nejbližší k němu. Metoda K – nejbližších sousedů patří mezi „lazy learner“ algoritmy, jelikož předem negeneruje model pro datový soubor. Jediné kalkulace nastávají tehdy, když je algoritmus dotázán k volbě nejbližších datových bodů. To činí z algoritmu velmi jednoduchou metodu pro data mining. Algoritmus sleduje daný bod, kde se data nachází. Dále se snaží se určit, jestli tento bod patří do skupiny A nebo B, podle toho do jaké skupiny patří body v jeho blízkosti. Pokud je většina bodů v jeho blízkosti ve skupině A, značí to vyšší pravděpodobnost, že daný bod se nachází spíše ve skupině A než B. (Techopedia, 2019)

Algoritmus se rozhoduje o blízkosti bodů na základě kalkulací, které měří vzdálenost mezi jednotlivými body. Pro měření této vzdálenosti mohou být použity následující metriky:

- Euklidovská vzdálenost: $d_{ij} = \sqrt{\sum_{k=1}^n (x_{ik} - x_{jk})^2}$

Jedná se o odmocninu sumy druhé mocniny rozdílů mezi jednotlivými body. Tato metrika je nejčastěji používána. (Teknomo, 2015)

- Manhattanská vzdálenost: $d = \sum_{i=1}^n |x_i - y_i|$

Jedná se o sumu rozdílů jednotlivých bodů. (Improvedoutcomes, 2004)

- Mahalanobis vzdálenost: $D_M(\vec{x}) = \sqrt{(\vec{x} - \vec{\mu})^T S^{-1} (\vec{x} - \vec{\mu})}$

Jedná se o vzdálenost mezi dvěma body v prostoru s více proměnnými, definovaný podle rovnice výše kde $\vec{x}$ značí vektor souboru pozorování, $\vec{\mu}$ vektor střední hodnoty ze souboru pozorování a S kovariační maticí. (Stephanie, 2017)

3.8.3 Decision trees

Hlavní myšlenky stojící za podstatou „Decision Trees“ byly vymyšleny před více než 70 lety a v současnosti se řadí mezi nejsilnější nástroje strojového učení. Bývají vývojáři užívány ve vědecko-datových soutěžích, operačním řízení a k řešení mnoha dalších problémů, které vyžadují automatické rozhodování. „Decision Trees“ využívají „white box“ metodu, což znamená, že výsledná rozhodnutí mohou být jednoduše vysvětlena na rozdíl od neuronových sítí, jejichž struktura se vyznačuje vysokou komplexitou. (Dezhic, 2017)

V oblasti systémů, které využívají metodu učení s učitelem, jsou označovány za jednu z nejjednodušších metod. Postupuje se tak, že na každý uzel stromu je aplikován test klasifikace a výsledek dotazování je postoupen na další uzel dané větve. Typ testu je z

Tento postup pokračuje, dokud se dotazování nedostane ke svému konečnému uzlu. Konečný uzel je spojen s hodnotou, a to buď daným označením u klasifikačních stromů či numerickou hodnotou u regresním stromů. Dosažená hodnota konečného uzlu skrze dotazování představuje výstup stromu. (Suter, 2019)

3.8.4 Random forrests

Metoda „Random Forest“ představuje model složený z mnoha „Decision Trees“. Spíše než pouhé zprůměrování predikcí obdržných z jednotlivých stromů, používá tento model dva klíčové koncepty.

- **Náhodné rozzaření dat z tréninkového souboru při konstrukci stromů**

V tréninkové fázi se každý strom, který je součástí „random forest“ učí na základě náhodného vzorku dat. Vzorky dat jsou použity skrze metodu „bootstrapping“, což znamená, že některá data budou v rámci jednoho stromu použita několikrát. Myšlenka spočívá v trénování každého stromu za pomoci různých vzorků dat, s tím že každý jednotlivý „strom“ může na základě určitého vzorku dat vykazovat vysoký rozptyl, nicméně celý „les“ by měl vykazovat rozptyl nižší, a to při nezvyšující se úrovni biasu. V testovací fázi jsou vytvořeny predikce skrze zprůměrování predikcí vytvořených jednotlivými „stromy“. (Koehrsen, 2018)

- **Určení náhodné podskupiny parametrů při rozdělování uzlů**

Další hlavní koncepcí v rámci metody „random forest“ je, že při rozdělování uzlů u jednotlivých „stromů“ je brána v úvahu pouze podskupina ze všech parametrů. Obecně se při rozdělování uzlů skrze klasifikaci používá odmocnina z celkového počtu parametrů. Tedy pokud máme k dispozici 16 parametrů, pouze 4 budou brány v úvahu. Metoda Random Forest může ve své tréninkové fázi využít v každém z uzlů všech dostupných parametrů, což se používá u regrese.

Metoda „Random Forest“ tedy kombinuje stovky nebo i tisíce „Decision Trees“, kdy každý z nich prochází tréninkovou fázi na různém souboru pozorování a rozdělování uzlů je prováděno vzhledem k omezenému počtu parametrů. Konečné hodnoty predikce metody „Random Forest“ jsou vytvořeny skrze zprůměrování predikcí vytvořených každým jednotlivým „stromem“. (Koehrsen, 2018)

Využití této metody může být nalezeno v detekci podvodných transakcí v bankovním sektoru, v oblasti medicíny v určení vhodné kombinace látek při tvorbě léků nebo při tvorbě analýzy chování akciových trhů. (Anurag, 2018)

4 Rešerše

V následující části je popsáno pět prací, které se zabývaly využitím neuronových sítí na kapitálových trzích. Dále je také zhodnocena využitelnost zvolených typů neuronových sítí a interpretovány jejich výsledky.

4.1 Forecasting Stock Market with Neural Networks (Lin, 2009)

Cílem této zkoumané práce je zhodnotit výnosnost obchodování na akciových trzích za použití neuronové sítě. V této studii jsou predikce neuronové sítě transformovány do jednoduché obchodní strategie, jejíž výnosnost je porovnávána oproti obchodní strategii kup a drž. Neuronová síť byla použita na analýzu dat z Taiwan Weighted Index a indexu S&P 500 v USA.

Metodologie

Při analýze byla nejprve shromážděna odpovídající data cen z hlavních indexů a spočítána výnosnost za předcházejících devět dní. Tato informace byla standardizována a použita k predikci vývoje cen za použití neuronové sítě. Poté co neuronová síť prošla tréninkovou fází, byly příslušné výsledky vyhodnoceny. Následně byla výnosnost investice tohoto modelu porovnávána s výnosností investice odvozené podle strategie kup a drž.

Zdroj dat

Data shromážděna v této práci zahrnují základní informace o vývoji cen pro Taiwan Weighted Index během období od 1. ledna 1970 do 31. prosince 2005 a pro index S&P 500 v USA za stejnou časovou periodu. Základní informace o vývoji cen zahrnují otevírací/zavírací cenu a denní minima/maxima cen.

Investiční strategie

a) Strategie kup a drž

Jako nákupní cena je brána zavírací cena indexu prvního dne daného roku, tedy na počátku investiční periody. K nákupu titulů byla použita plná částka disponibilních aktiv a držena až do konce investičního období

b) Full Trading Strategie

V rámci této strategie byl při nastání obchodního signálu nakoupen/prodán akciový index v hodnotě plné disponibilní částky určené pro nákup titulů. Tato strategie je aplikována v modelu neuronové sítě.

Model neuronové sítě

V rámci této práce byl použit model neuronové sítě s back-propagací. Stručný popis struktury neuronové sítě, zahrnující vrstvu vstupu, vrstvu výstupu a skrytou vrstvu je uveden níže.

Vrstva vstupu

Jako vstupní data neuronové sítě byl použit výnos z investice za předchozích devět dní (R_{t-1} , R_{t-2} , ..., R_{t-9}).

Rovnice pro výpočet má následující podobu:

$$R_t = \frac{P_t}{P_{t-1}} \quad R_t = \text{výnosnost z investice v den } t$$

Jelikož jsou vkládány hodnoty s rozdílným rozsahem, je zapotřebí předejít situaci, kdy významnost proměnných s vyšším rozsahem převyší významnost proměnných s menším rozsahem. Tento problém je vyřešen skrze standardizaci rozsahu proměnných.

Vrstva výstupu

K určení rozhodnutí o načasování obchodování byl použit FK indikátor, signalizující nákupní a prodejní signály. Rozpětí hodnot FK indikátoru sahá od 0 do 1. Pokud cena akcie roste, hodnota FK indikátoru se posouvá k 0 a pokud cena klesá tak se hodnota indikátoru posune opačným směrem k 1. Tento indikátor je v této studii použit k identifikaci nákupních a prodejních signálů.

V rámci této práce jsou stanoveny hranice pro nákup a prodej, a to na hodnotách 0,2 a 0,8. Pokud se tedy hodnota FK indikátoru dostane pod hodnotu 0,2, systém vyhodnotí výstup jako nákupní signál (hodnota = 1) a pokud hodnota FK indikátoru překročí hodnotu 0,8, systém vyhodnotí výstup jako prodejní signál (hodnota = -1). U ostatních hodnot obchodování nenastane (hodnota = 0). Za nejvhodnější transferovou vrstvu výstupní vrstvy byla vybrána hyperbolická tangent funkce.

Skrytá vrstva

Skrytá vrstva slouží k propojení neuronů, které na sebe díky ní vzájemně působí. Na základě výsledků z této zkoumané práce je v běžných případech využití jedné skryté vrstvy dostatečné.

K určení vhodného množství neuronů pro každou skrytou vrstvu sítě neexistuje exaktní rovnice. Ve zkoumané práci je počet neuronů v každé vrstvě, kde první číslo značí počet neuronů vstupní vrstvy, druhé číslo počet neuronů skryté vrstvy a poslední číslo počet neuronů vrstvy výstupu, navržen následovně:

Struktura sítě		
	9 - 3 - 1	
	9 - 4 - 1	
	9 - 5 - 1	

Pro vrstvu výstupu byla vybrána hyperbolická tangent funkce. Po několika testováních bylo zjištěno, že nejlepších výsledků u skryté vrstvy dosahuje logistická funkce.

Měření výkonnosti investice

V rámci této práce byla použita investiční strategie, kdy k provedení nákupu/prodeje dochází po nastání prvotního signálu. Tato strategie spočívá v tom, že pokud se objeví nákupní signál, žádné další nákupy nebudou uzavřeny, dokud se neobjeví signál pro prodej. Stejně tak pokud se nejprve objeví prodejní signál, žádné další prodeje nebudou uskutečněny, dokud se neobjeví signál pro nákup.

Rovnice pro výpočet výnosnosti z investice:

$$\text{Výnosnost} = \frac{\text{Prodejní cena} - \text{Poplatek za prodej} - \text{Kapitálová daň}}{\text{Nákupní cena} + \text{Poplatek za nákup}}$$

Rovnice pro výpočet výnosnosti u strategie kup a drž:

$$\text{Výnosnost} = \frac{\text{Prodejní cena v poslední den období} - \text{Poplatek za prodej} - \text{Kapitálová daň}}{\text{Nákupní cena v první den období} + \text{Poplatek za nákup}}$$

Rovnice pro výpočet ročního výnosu:

$$\text{Výnosnost} = \sqrt[n]{\prod_{k=1}^n R_k} - 1$$

R_k = Výnosnost investice, n = doba držení cenného papíru

Předpoklady praktické části

V rámci této práce byla použita neuronová síť za účelem predikce Taiwan Weighted Index a index S&P v USA. Praktická část práce se řídí následujícími předpoklady:

- 1) Pokud se vyskytne obchodní signál, je obchodováno za uzavírací cenu dne
- 2) Pokud dojde ke stejnému obchodnímu signálu, poté co byla transakce provedena, neobchoduje se, dokud nedojde k opačnému obchodnímu signálu
- 3) Úroveň poplatku za každý nákup či prodej činí 0,1425 %
- 4) Daň z kapitálového zisku činí 0,3 % u každého prodeje akcie
- 5) Pokud na konci zkoumaného období existují akcie, které nebyly obchodovány, dojde k jejich prodeji za poslední uzavírací cenu zkoumaného období
- 6) Potenciální dividendy získané za dlouhodobou držbu nejsou v této práci zohledněny
- 7) Všechna výkonnost investic použita v této práci je brána jako výkonnost za minulé období. Není tedy brána jako indikátor budoucí výkonnosti.

Empirické výsledky

V rámci zkoumané práce byl použit vývoj hlavních indexů za období od 1. ledna 1970 do 31. prosince 2005 a porovnán s výsledkem strategie kup a drž. Kromě této strategie byl k porovnání výnosností také využit model na bázi neuronové sítě, který byl trénován na stejné periodě.

První porovnání výnosnosti investic se týká full trading strategie obchodovanou skrze neuronovou síť se strategií kup a drž. Obě tyto strategie byly aplikovány na akciový index Taiwan General Index. Výsledky ukazují, že roční výnosnost z full trading strategie je podstatně vyšší než výnosnost strategie kup a drž. Roční výnosnost obdržená při aplikaci neuronové sítě ve struktuře 9-3-1 je 17,79 %, ve struktuře 9-4-1 je 14,63 % a 9-5-1 13,61 %. Všechny tři tyto struktury neuronové sítě podstatně převyšují roční výnosnost strategie kup a drž, která činí 11,98 %.

Dále byla ve zkoumané práci porovnána roční výnosnost amerického indexu S&P 500 při aplikaci totožných strategií jako u prvního porovnání výnosností. Výsledek opět ukázal, že roční výnosnost full trading strategie na bázi neuronové sítě je podstatně vyšší, než roční výnosnost strategie kup a drž. Roční výnos obdržený z obchodování po aplikaci neuronové sítě ve struktuře 9-3-1 je 10,214 %, ze struktury 9-4-1 8,325 % a ze struktury 9-5-1 8,61 %. Oproti tomu strategie kup a drž dosáhla roční výnosnosti 7,46 %.

Oba výsledky zkoumané práce tedy ukázaly, že spojení neuronové sítě a full trading strategie je schopno detekovat obchodní signály a tím pádem dosahovat lepší výnosnosti než běžná obchodní strategie kup a drž.

4.2 Deep learning with long short-term memory networks for financial market predictions (Fischer, 2017)

LSTM neuronové sítě patří mezi nejmodernější technologie pro učení datových posloupností a nachází velmi dobré uplatnění u predikcí vývoje finančních časových řad. Tato zkoumaná práce porovnává výnosnosti investic za použití investiční strategie skrze metodu LSTM neuronové sítě, random forest, obecné neuronové sítě a logistické regrese. Zmíněné metody jsou aplikovány na časovou řadu vývoje akciových cen vybraných titulů indexu S&P 500 od roku 1992 do roku 2015.

Ve zkoumané práci jsou analyzovány vlastnosti portfolia skládající se ze $2k$ počtu akcií. V rámci investiční strategie se nakoupí akcie k nejvíce kapitalizovaných společností a prodají akcie k nejméně kapitalizovaných společností, kdy $k \in \{10, 50, 100, 150, 200\}$. K porovnání výkonnosti LSTM neuronové sítě s ostatními přístupy slouží průměrná střední hodnota denního výnosu a průměrný roční Sharpeho poměr před zahrnutím transakčních nákladů.

V rámci této práce byla při porovnávání daných indikátorů potvrzena využitelnost LSTM architektury neuronové sítě. Bez ohledu na velikost portfolia k , LSTM neuronové sítě dosahují v rámci porovnávaných indikátorů nejlepších výsledků. Průměrná střední hodnota denních výnosů před zahrnutím transakčních nákladů dosahuje 0,46 % pro $k = 10$. Ostatní metody dosahují nižších výnosností, a to 0,43 % u random forest, 0,32 u obecné neuronové sítě a 0,26 % u logistické regrese. Také u portfolii, která obsahují vyšší počet titulů, bylo dosaženo v rámci metody LSTM nejlepších výsledků. Výjimkou je portfolio s $k=200$, které dosáhlo stejné výnosnosti jako metoda random forest.

V rámci indikátoru Sharpeho poměru dosahuje LSTM neuronová síť v porovnání s ostatními metodami nejvyšší výkonnosti až do $k = 100$. Pro množství titulů $k = 150$ a $k = 200$ dosahuje lehce lepších výsledků metoda random forest. Sharpeho poměr u $k = 150$ dosahuje u random forest hodnoty 3,82 oproti hodnotě 3,68 u metody LSTM a u $k = 200$ za použití random forest hodnoty 3,58 oproti hodnotě 3,30 u LSTM. V porovnání s ostatními metodami dosáhla LSTM síť vyšší výkonnosti.

Dosažené výsledky této zkoumané práce a následné porovnání výkonnosti metody LSTM neuronové sítě s dalšími metodami, potvrzují využitelnost tohoto typu sítě při aplikaci na finančních trzích.

4.3 Stock Market Prediction Using Artificial Neural Networks (Kutlu,2015)

Cílem této zkoumané práce je za pomoci neuronové sítě predikovat vývoj akciové indexu Istanbul Stock Exchange. V této práci jsou jako vstupy použity hodnota akciového indexu předchozího dne, měnový kurz TL/USD předchozího dne, úroková sazba předchozího dne a pět dummy proměnných, každá zastupující pět pracovních dnů v týdnu. Odpovídající data byla shromážděna za období 417 dní od 1. července 2001 do 28. února 2003. Pro tréninkovou fázi je použito 90 % dat a pro testovací fázi 10 % dat.

Pro tréninkovou a testovací část této práce je použita neuronová síť na základě několikavrstvému perceptronu a obecné feed-forward sítě. K porovnání výkonnosti jsou použity oba zmíněné typy neuronových sítí, obsahující jednu, dvě a čtyři skryté vrstvy. Výsledky jsou dále porovnány s výkonností strategie na základě pěti a deseti denních klouzavých průměrů.

V této zkoumané práci je tedy použito šest verzí neuronové sítě. K měření výkonnosti modelu je použit koeficient determinace (R^2), za účelem měření přesnosti predikce na základě tréninkových dat. Dalším indikátorem k měření přesnosti predikce slouží střední procentuální chyba odhadu.

V rámci porovnání přesnosti modelů neuronových sítí pomocí koeficientu determinace, dosáhly nejlepších hodnot u obou zkoumaných typů neuronové sítě s jednou vrstvou. U více vrstvého perceptronu dosáhla jednovrstvá neuronová síť $R^2 = 0,81$, dvouvrstvá $R^2 = 0,79$ a čtyřvrstvá $R^2 = 0,78$. U jednovrstvé obecné feed-forward sítě bylo dosaženo $R^2 = 0,82$ a u dvouvrstvé i čtyřvrstvé $R^2 = 0,81$.

V rámci ukazatele střední procentuální chyby odhadu bylo k šesti verzím neuronových sítí přidáno porovnání s výsledky pěti a deseti denního klouzavého průměru. Celkově modely neuronových sítí u všech šesti verzí dosahovaly nižších hodnot chyby než modely klouzavých průměrů. Nejnižší chyby dosáhla jednovrstvá neuronová síť na bázi obecné feed-forward sítě, a to 1,59 %. Dobrého výsledku také dosáhla jednovrstvý perceptron, a to 1,62 %. Oproti tomu pěti denní klouzavý průměr dosáhl chyby 2,17 % a deseti denní 3,03 %.

Díky této práci byl učiněn závěr, že modely neuronových sítí jsou při tvorbě predikcí přesnější než modely využívající klouzavé průměry. Pro tvorbu predikcí se ukázal vhodnějším model na základě obecné feed-forward sítě.

4.4 Stock Market Index Prediction with Neural Network during Financial Crises: A Review on Bist-100 (Yavuz, 2015)

Další zkoumaná práce se zabývá posouzením predikční funkčnosti neuronových sítí během finanční krize. Jako akciový index je použit turecký index BIST-100, kde je predikována hodnota tohoto indexu pro následující den a pro následující týden, za využití feed-forward neuronové sítě s back-propagací. Použitá data se vztahují k periodě 30 měsíců, a to od července 2007 do prosince 2009, tedy na období světové finanční krize. Mezi zkoumané proměnné je zahrnuta cena zlata, ropy, úroková sazby, spotřebitelský cenový index, měnový kurz, peněžní nabídka a objem transakcí na BIST. Neuronová síť je trénována na datech za 12 měsíců (červenec 2007 – červen 2008) a testována na datech za období 18 měsíců (červenec 2008 – prosinec 2009). Použitá neuronová síť se skládá z jedné vstupní vrstvy, dvou skrytých vrstev a jedné vrstvy výstupů. Počet neuronů v každé vrstvě je v závislosti na pořadí uvedených vrstvách sítě 7-9-7-2.

Pro porovnání přesnosti odhadu vývoje reálných data indexu BIST-100 s predikovanými hodnotami skrze neuronovou síť byl použit t-test. Pro ověření normálního rozdělení v datových souborech byl použit Kolmogorov-Smirnov (K-S) test. K-S test pro hodnotu indexu následujícího dne dosáhl hodnoty 0,79 a u odhadu hodnoty indexu skrze neuronovou síť 0,749. Jelikož tyto hodnoty leží na hladině významnosti vyšší než 0,05, oba porovnávané datové soubory mají normální rozdělení. V rámci výsledku t-testu, dosáhla t-statistika hodnoty -0,453 s hodnotou významnosti 0,656. Tento výsledek značí, že na hladině významnosti 5 % není mezi predikovanými a reálnými hodnotami významný statistický rozdíl.

V rámci predikování hodnoty indexu pro následující týden dosáhl K-S test u reálných dat hodnoty 0,846 a u odhadu skrze neuronovou síť hodnoty 0,727. Jelikož jsou hodnoty významnosti vyšší než 0,05, oba porovnávané soubory mají normální rozdělení. V rámci t-testu bylo dosaženo t-statistiky -0,114 s hodnotou významnosti 0,911. Tento výsledek značí, že na hladině významnosti 5 % není mezi predikovanými a reálnými hodnotami významný statistický rozdíl.

V rámci této zkoumané práce technologie neuronových sítí potvrdila svou využitelnost při odhadu hodnoty indexu jak pro následující den, tak i pro následující týden. Predikované hodnoty v porovnání vývoje reálných hodnot dosáhly velmi dobrých výsledků. Může být tedy učiněn závěr, že v rámci této práce byla neuronová síť schopna predikovat směr vývoje akciového indexu BIST-100 během období finanční krize.

4.5 Stock market index prediction using artificial neural networks trained on foreign markets (Karlsson, 2015)

Tato zkoumaná práce se zabývá otázkou, jestli jsou neuronové sítě schopny predikovat vývoj akciových indexů na geograficky odlišných akciových trzích. Práce se zabývá vývojem cen pěti indexů, a to indexy DAX (Německo), Nikkei (Japonsko), OMX 30 Copenhagen (Dánsko), OMX 30 Stockholm (Švédsko) a S&P 500 (USA) v časové periodě 100 pracovních dní. Hodnoty indexů mohou být chápány jako průměrná hodnota nejvlivnějších akciových titulů na daném trhu. Indexy jsou vážené, tudíž je nutno brát v úvahu, vyšší vliv větších společností na hodnotu akciového indexu.

Akciové indexy v této zkoumané práci jsou vybrány na základě geografické polohy. Švédsko, Dánsko a Německo se od sebe nachází v relativně krátké vzdálenosti, a pokud je učiněn předpoklad, že poloha trhu má vliv na jeho vývoj, měly by se tyto trhy vyvíjet podobně. Oproti tomu Japonsko a USA jsou trhy geograficky vzdálené, a proto by se podle tohoto předpokladu měly vyvíjet odlišně.

Jelikož v daných zemích existuje několik akciových indexů, byl vybrán ten index s nejvyšší kapitalizací a širokou diverzifikací oborů zahrnutých společností. Dále musí být bráno v úvahu, že indexy Nikkei a S&P 500 se skládají s podstatně vyššího počtu společností než indexy DAX, OMX 30 Stockholm a OMX 30 Copenhagen.

Neuronová síť byla v rámci každého trhu trénována v periodě od 1. ledna 2014 do 1. června 2014. Jelikož analýza ukázala, že starší data ztrácejí na predikci vývoje budoucích akciových cen vliv, byla pro tréninkovou fázi použita data pouze za pět měsíců. Neuronová síť byla dále na všech pěti trzích testována v periodě od 2. června 2014 do 1. prosince 2014.

Na základě předpokladu vhodnosti feed-forward neuronové v práci s datovými posloupnostmi, byl pro zkoumanou práci zvolen tento typ sítě. Neuronová síť se skládá z jedné vrstvy vstupů obsahující čtyři neurony, jedné skryté vrstvy obsahující deset neuronů a jedné vrstvy výstupu obsahující jeden neuron.

Jako vstupy byly vybrány za den t data denního minima ceny, denního maxima ceny, otevírací a zavírací cena. Na základě těchto vstupů je predikován výstup a to zavírací cena dne $t + 1$. To znamená, že v rámci této práce jsou tvořeny krátkodobé predikce. Každý den neuronová síť předpoví zavírací cenu dne následujícího. Pro další predikci neuronová síť obdrží správné hodnoty a na jejich základě vytvoří další predikci.

K vyhodnocení celkové výkonnosti predikce neuronové sítě na daném trhu byla použita suma chyby. Suma chyby se vyjadřuje v absolutní hodnotě, což zabraňuje vyrovnání chyb s opačným znaménkem.

Dále byl pro vyhodnocení výkonnosti neuronové sítě vypočten rating p . Tento rating výkonnosti je založen na poměru sumy chyb neuronové sítě aplikované na zahraniční trhu, oproti sumě chyby neuronové sítě aplikované na trhu domácím.

$$p = \frac{\sum_0^T e_m(t)}{\sum_0^T e_d(t)}$$

Hodnota $em(t)$ vyjadřuje chybu neuronové sítě na zahraničním trhu a $ed(t)$ chybu neuronové sítě na trhu domácím v čase t . V rámci zkoumané práce $T = 100$.

Jelikož každé trénink neuronové sítě vytvoří novou neuronovou síť, k vyhodnocení práce bylo na každém trhu trénováno 100 sítí. Například tréninková fáze 100 sítí byla uskutečněna na německém indexu DAX a její výsledky testovány na všech trzích včetně německého. Podle tohoto postupu se postupovalo na každém trhu.

Nejvyššího rozdílu v rámci ukazatele p bylo dosaženo u neuronové sítě trénované na datech indexu Nikkei a dále testované na indexu OMX 30 Copenhagen, kdy $p = 1,0113$. To znamená, že výkonnost sítě dosahující nejhorší výsledky je o 1,13 % horší než výkonnost sítě, kde testovací fáze byla vykonána skrze neuronovou síť použitou na datech domácího indexu (data stejného indexu jako u tréninkové fáze). V osmi případech ze dvaceti, neuronová síť testována na zahraničním indexu dosáhla jen lehce lepších výsledků než testování na datech domácího trhu. Nejlepších výsledků dosáhla neuronová síť trénována na datech německého indexu a testována na datech japonského indexu, jelikož dosáhla výkonnosti o 1,46 % vyšší než testování na datech domácího indexu.

V rámci zkoumané práce byl učiněn závěr, že neuronové sítě testované na zahraničním trhu dosáhly porovnatelné výkonnosti jako neuronové sítě testované na domácím trhu, tedy na stejném trhu jako byly trénovány.

5 Praktická část

Účelem praktické části práce je posoudit využitelnost aplikovaných typů neuronových sítí v predikci vývoje na akciových trzích. Zvolené typy neuronových sítí jsou aplikovány na sedm akciových indexů, které determinují vývoj akciových trhů na různých částech světa. K vytvoření a posouzení výsledků aplikace neuronových sítí je použit program Matlab. Zdrojem cenových dat posuzovaných akciových indexů je server Yahoo Finance.

5.1 Indexy

Pro posouzení využitelnosti neuronových sítí v predikci vývoje cen na akciových trzích byly v praktické části této práce použita cenová data sedmi světových indexů. Z důvodu ověření využitelnosti na různých trzích po světě byly v praktické části zahrnuty tři největší akciové indexy nejsilnější světové ekonomiky USA, a to indexy Dow Jones Industrial Average, S&P 500 a NASDAQ Composite. Pro evropské trhy byla použita data německého indexu DAX a francouzského indexu CAC 40. Dále pro asijské trhy byla využita data hong kongského indexu HANG SENG a japonského indexu NIKKEI 225.

5.1.1 Dow Jones Industrial Average

V případě Dow Jones Industrial Average (DJIA) se jedná o akciový index, který sleduje 30 největších, veřejně vlastněných společností, obchodujících se na amerických burzách NYSE a NASDAQ, která nejlépe reprezentují dané odvětví ekonomiky. Index byl vytvořen v roce 1896 a je pojmenován po jeho zakladateli Charlesi Dow a jeho obchodnímu partnerovi Edwardu Jones. Tento akciový index patří k jednomu z nejstarších na světě, zahrnující společnosti jako Walt Disney, Exxon Mobil či Microsoft. Když se na televizních obrazovkách v USA objeví zpráva o dnešním růstu trhu, většinou se tato informace vztahuje k indexu Dow Jones. (Ganti, 2019)

5.1.2 S&P 500

Akciový index S&P 500 sleduje 500 nejvíce tržně kapitalizovaných společností ekonomiky USA a představuje výkonnost tohoto trhu. Pro investory slouží jako benchmark celého trhu, ke kterému jsou všechny ostatní investice porovnávány. Zahrnuje akciové společnosti z celé škály odvětví, mezi které mimo jiné patří telekomunikační služby, automobilový průmysl, maloobchod, potravinářský průmysl, energetika, technologie a bankovní sektor. Všechny společnosti zahrnuté do tohoto indexu jsou zalistovány na americké burze NYSE. V porovnání s indexem Dow Jones je do S&P 500 zahrnuto více největších společností ekonomiky z hlediska tržní kapitalizace. Nicméně v obsahu technologických společností převládá index NASDAQ, který zahrnuje i soukromě vlastněné společnosti. (Amadeo, 2019)

5.1.3 NASDAQ Composite

Index NASDAQ Composite sleduje výkonnost okolo 3000 akciových titulů obchodovaných na burze NASDAQ. Využívá se zejména jako indikátor výkonnosti malých i velkých společností technologického sektoru. Index byl založen roku 1971 a jsou do něj zahrnuty společnosti jako Apple, Microsoft, Amazon či Facebook. (Capital, 2019)

5.1.4 DAX

Akciový index DAX představuje 30 největších a nejvíce likvidních německých společností obchodovaných na burze ve Frankfurtu. Ceny akciových titulů sloužící k výpočtu hodnoty indexu prochází skrze elektronický obchodní systém XETRA. Index byl založen roku 1988 a společnosti do něj zahrnuté představují 75 % celkové tržní kapitalizace obchodované na burze ve Frankfurtu. Mimo jiné jsou do indexu zahrnuty společnosti jako Lufthansa, RWE a Deutsche Bank. (Chen, 2018)

5.1.5 CAC 40

Index CAC 40 představuje hlavní akciový index společností ve Francii. Slouží zejména jako indikátor francouzského trhu, představující výkonnost 40 nejvíce kapitalizovaných a nejlikvidnějších akcií zalistovaných na burze Euronext v Paříži. Index CAC 40 patří mezi hlavní globální indexy akciového trhu a je podle něj posuzována výkonost evropské ekonomiky jako celku. Byl založen roku 1987 a zahrnuje společnosti jako L'Oreal, Renault, Michelin a BNP Paribas. (Capital, 2019)

5.1.6 HANG SENG

Hang Seng index slouží jako hlavní indikátor vývoje akciového trhu v Hong Kongu. Do indexu patří 50 tamních firem s největší kapitalizací, obchodované na akciové burze v Hong Kongu a nezahrnuje pouze trh Hong Kongu ale i společnosti v Číně. Index je vlastněn Hang Seng Bank, a byl založen roku 1969. Patří do něj společnosti jako Industrial and Commercial Bank of China, China Mobile a Tencent. (MarketBusinessNews, 2019)

5.1.7 NIKKEI 225

Index Nikkei 225 představuje hlavní akciový index japonské ekonomiky. Je do něj zahrnuto 225 společností s nejvyšší kapitalizací zalistovaných na burze v Tokyu. Index zejména slouží jako měřítko japonského akciového trhu a výkonnosti celé japonské ekonomiky. Index je počítán od roku 1950 a zahrnuje společnosti ze 36 různých odvětví japonské ekonomiky. Patří do něj společnosti jako Sony, Canon, Nissan či Panasonic. Od ostatních akciových indexů se liší tím, že společnosti nejsou zalistovány podle tržní kapitalizace ale podle ceny akcie. Ceny akcií jsou denominovány v japonském jenu a kvůli pořadí v indexu přehodnocovány každý rok v září. (CFI, 2019)

5.2 Sestavení modelů

V následující části je popsáno, jak jsou použita data upravena, aby model dosahoval vyšší výkonnosti. Dále rozdělení celkového souboru dat na jednotlivé podsoubory hyperparametrů modelu, které tvoří jeho architekturu. Nakonec jsou určeny metriky podle jakých ukazatelů jsou modely vyhodnoceny.

5.2.1 Úprava dat

Tržní data typicky nesou informaci ve formátu High, Low, Open, Close, Volume. V rámci této práce je využito uzavíracích cen (Close) jednotlivých akciových indexů za každý den, kdy byly tyto indexy ve zvoleném časovém horizontu obchodovány. Modely neuronových sítí se učí patřičná propojení mezi vstupními proměnnými a výstupní proměnou. Nicméně, rozsah a distribuce se pro každou proměnou může lišit a to způsobuje zvýšenou náročnost modelace zkoumaného problému. Vstupní proměnné mohou mít různé jednotky, kdy příkladem u dat akciového trhu může být cena a obchodovaný objem akcie. Objem obchodované akcie se pohybuje v daleko vyšších jednotkách než její cena a přiřadí tedy tomuto parametru daleko vyšší váhu. Model s vysokými vahami může vykazovat slabou výkonnost během učící fáze a vést k vyšší celkové chybovosti. Aby se tomuto problému předešlo, je v modelech zahrnutých v této práci použita metoda standardizace a normalizace.

Normalizace dat

Metoda normalizace datového souboru se používá k úpravě rozmezí, ve kterém se data pohybují, tak aby jejich hodnoty ležely mezi 0 a 1. Ve svém výpočtu vyžaduje přesný odhad minima a maxima ze zkoumaných dat, která jsou k dispozici. Výpočet aplikovaný k normalizaci dat je následující, kdy x představuje vstupní hodnotu, max hodnotu maxima datového souboru a min hodnotu minima datového souboru:

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

Při výpočtu se musí brát v úvahu, že pokud vstupní hodnota x leží mimo rozmezí maxima a minima, výsledná hodnota po normalizaci nemůže ležet mezi hodnotami 0 a 1.

Standardizace dat

Metoda standardizace datového souboru se používá ke změně distribuce použitých dat, tak aby jejich střední hodnota nabývala nuly a směrodatná odchylka čísla jedna. Tento postup bývá také označován za vycentrování datového souboru. Podobně jako metoda normalizace, standardizace může být u některých algoritmů strojového učení přímo vyžadována. Zejména jako v této práci, kdy se vstupní hodnoty modelu pohybují v různých rozmezích. Tato metoda ke svému výpočtu požaduje přesný odhad střední hodnoty a směrodatné odchylky pozorovaných hodnot, které jsou v této práci získány z použitých souborů dat.

$$y = \frac{(x - \mu)}{\sigma}$$

V rovnici standardizace dat x představuje vstupní hodnotu, μ střední hodnotu a σ směrodatnou odchylku.

5.2.2 Rozdělení na tréninková, validační a testovací data

V rámci strojového učení bývá běžnou úlohou sestavit funkční algoritmy, které jsou schopny se učit a vytvářet predikce ze zkoumaných dat. Správnost těchto predikcí je závislá na vytvoření matematického modelu ze vstupních dat, jejichž zdrojem bývá několik datových souborů. V rámci této práce je u LSTM neuronové sítě využito rozdělení původního datového souboru na dva, a to jeden pro tréninkovou fázi algoritmu a druhý pro testovací fázi algoritmu. V rámci NAR a NARX modelů neuronové sítě je původní datový soubor rozdělen na soubory tři, a to tréninkový, validační a testovací.

Tréninkový datový soubor

V první fázi je k modelům přiřazen tréninkový datový soubor, který představuje soubor příkladů za účelem nastavení parametrů vah ke spojení jednotlivých neuronů sítě. Modely jsou trénovány na datovém souboru využívající metodu učení se s učitelem (gradient descent). V praxi to znamená, že se tréninkový soubor skládá z několika vstupních vektorů a odpovídajícímu výstupnímu vektoru. Modely díky vstupním vektorům s využitím patřičného učícího algoritmu vytvoří výsledek, který je porovnán s výstupním vektorem. Na tomto základě jsou dále upraveny váhy propojující jednotlivé neurony.

Validační datový soubor

Po tréninkové fázi nastává fáze validační, kdy je model schopen vytvářet predikce, ze kterých vychází validační datový soubor. Validační datový soubor poskytuje vyhodnocení správnosti přiřazení modelu k tréninkovému datovému souboru a slouží k úpravě hyperparametrů modelu. V této práci u modelů NAR a NARX tvoří validační datový soubor 20 % z celkového datového souboru.

Testovací datový soubor

Testovací datový soubor se používá ke konečnému vyhodnocení správnosti přiřazení tréninkovému datového souboru ke zvolenému modelu neuronové sítě. Tento datový soubor je nezávislý na tréninkovém datovém souboru, nicméně pokud jsou tréninková data správně přiřazená k danému modelu, znamená to, že jsou také k modelu správně přiřazena data testovací. Hlavním účelem testovacího datového souboru je vyhodnocení výkonnosti zvoleného modelu neuronové sítě.

5.2.3 Hyperparametry

Za hyperparametry modelu jsou označovány proměnné, které určují strukturu neuronové sítě (např. počet skrytých vrstev) a proměnné, které určují způsob průběhu tréninkové fáze (např. learning rate. Nastavení těchto hyperparametrů je určeno před tréninkovou fází sítě. Hyperparametry sítě v této práci jsou zejména stanoveny na základě manuálního hledání nejvhodnějšího nastavení.

- Typ neuronové sítě
- Počet skrytých vrstev
- Počet neuronů ve skrytých vrstvách
- Časová zpoždění
- Aktivační funkce –umožňuje modelu naučit se nelineární vztahy mezi vstupními daty.
- Learning rate – určuje jak rychle je síť schopna aktualizovat své parametry
- Počet epoch – počet epoch označuje počet kroků, ve kterých jsou tréninková data použita v síti během tréninkové fáze
- Optimalizační algoritmus

5.2.4 Ukazatele vyhodnocení modelů

Pro vyhodnocení predikční schopnosti a výnosnosti modelů jsou v této práci použity následující tři indikátory. Vzhledem k dostupným datům jsou indikátory míry výnosnosti a RMSE použity k posouzení využitelnosti LSTM neuronové sítě. Pomocí indikátoru MSE spolu s regresním a autokorelačními grafy byla vyhodnocena predikční schopnost feed-forward neuronových sítí, konstruovaných skrze modely NAR a NARX.

Střední čtvercová chyba:

$$a) RMSE = \sqrt{\sum \frac{(y_{predikce} - y_{skutečné})^2}{Počet\ dat}}$$

$$b) MSE = \sum \frac{(y_{predikce} - y_{skutečné})^2}{Počet\ dat}$$

Míra výnosnosti:

$$R = \frac{V_1 - V_0}{V_0}$$

V_0 = hodnota investice v začátku sledovaného období, V_1 = hodnota investice na konci sledovaného období

5.3 Predikční modely

Za vytváření predikcí může být označen druh dynamického filtrování, ve kterém jsou použity minulé hodnoty jedné nebo více časových řad k predikci hodnot budoucích. Dynamické neuronové sítě, které jsou v této práci aplikovány, byly použity pro predikci vývoje nelineárních časových řad akciových indexů. V následující části je zkoumána využitelnost rekurentní LSTM neuronové sítě, NAR a NARX modelů.

5.3.1 Model rekurentní LSTM neuronové sítě

Prvním ze zkoumaných typů sítí je rekurentní LSTM neuronová síť, pomocí které bude testována její využitelnost pro předpověď budoucích hodnot zvolených akciových indexů. Na jednotlivých indexech tedy bude vyhodnocena predikční schopnost sítě vzhledem ke zvoleným metrikám a porovnáno na kterém z trhů bylo dosaženo nejvyšší výnosnosti. Jelikož je v této práci zkoumáno sedm různých indexů, bude pro grafickou ilustraci průběhu testování LSTM sítě využit pouze jeden z indexů, a to index S&P 500.

5.3.1.1 Data

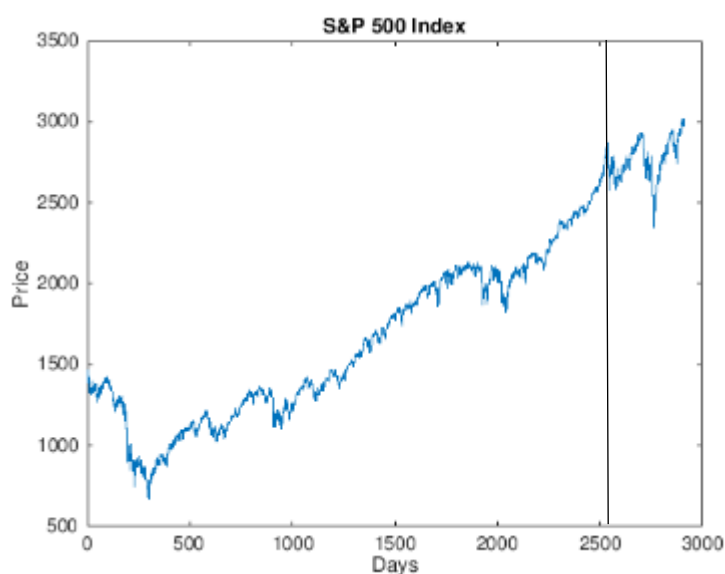
Všechny datové soubory byly rozděleny v poměru 90 % dat pro tréninkovou fázi a 10 % dat pro fázi testovací v následujících obdobích:

Tréninková data: 1. 1. 2008 - 31. 12. 2018

Testovací data: 1. 1. 2019 – 26. 7. 2019

V modelu byly zvoleny vstupní data uzavíracích cen (Close) jednotlivých indexů pro každý den, ve kterém byly tyto indexy obchodovány. Rozdělení datového souboru může být viděno na grafu níže:

Obrázek 8 Vývoj indexu S&P 500



Zdroj: mathworks.com

5.3.1.2 Hyperparametry

Po načtení datových souborů a jejich rozdělení na tréninkový soubor a testovací soubor byla data tréninkového souboru standardizována v souladu s dříve uvedeným vzorcem pro standardizaci. Poté byla zvolena architektura neuronové sítě viz následující tabulka:

Tabulka 1 Hyperparametry LSTM neuronové sítě

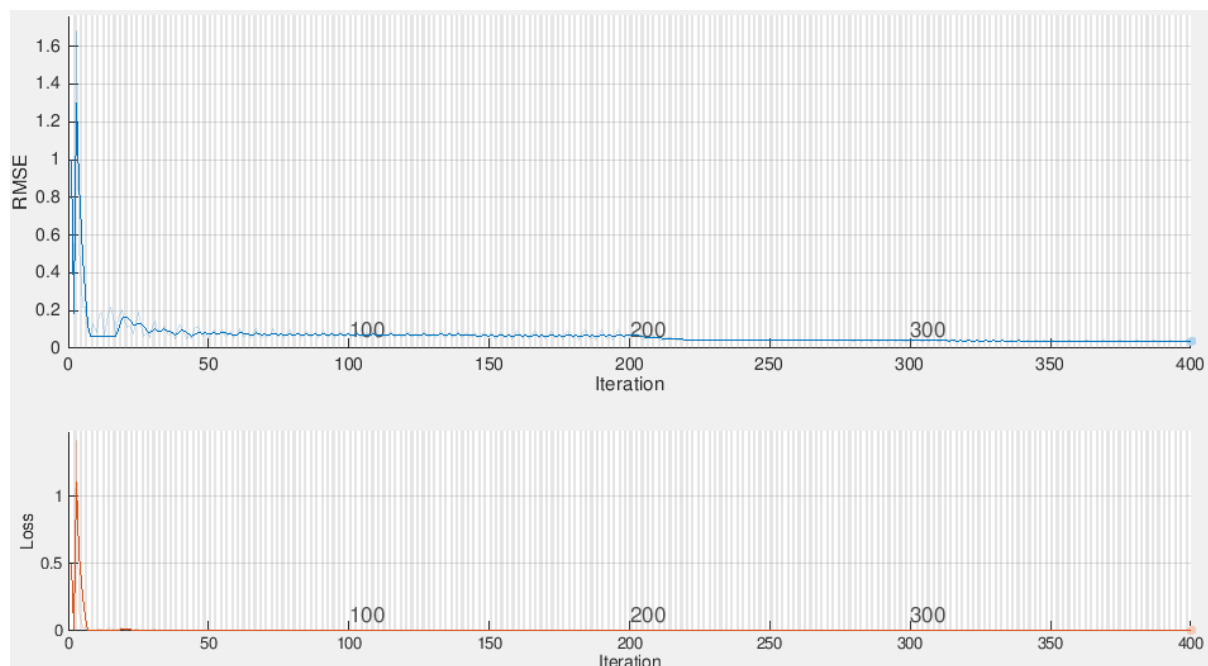
Hyperparametry LSTM neuronové sítě	
Typ neuronové sítě	Rekurentní LSTM síť
Počet skrytých vrstev	1
Počet neuronů ve skryté vrstvě	200
Optimalizační algoritmus	adam
Maximální počet epoch	100, 150, 200, 250, 300, 400
Learning rate	0,005
Aktivační funkce	Logistická

Na základě výsledků testování byla vybrána LSTM síť obsahující jednu skrytou vrstvu, a to z důvodu přeučení při aplikaci více skrytých vrstev. Počet neuronů ve skryté vrstvě byl stanoven na hodnotě 200. Pro tréninkovou fázi sítě byl zvolen optimalizační algoritmus „adam“. V počtu epoch se výsledky u jednotlivých akciových indexů lišily, a proto byl testován různý počet epoch. Počáteční learning rate byl stanoven na hodnotě 0,005, který po dosažení poloviny počtu epoch byl vynásoben hodnotou 0,2. Aktivační funkce ve skryté vrstvě vychází z povahy LSTM sítě a jedná se o logistickou funkci.

5.3.1.3 Tréninková fáze

S tímto nastavením byla síť trénována, čehož průběh může být viděn na následujícím grafu:

Obrázek 9 Tréninková fáze sítě



Zdroj: *mathworks.com*

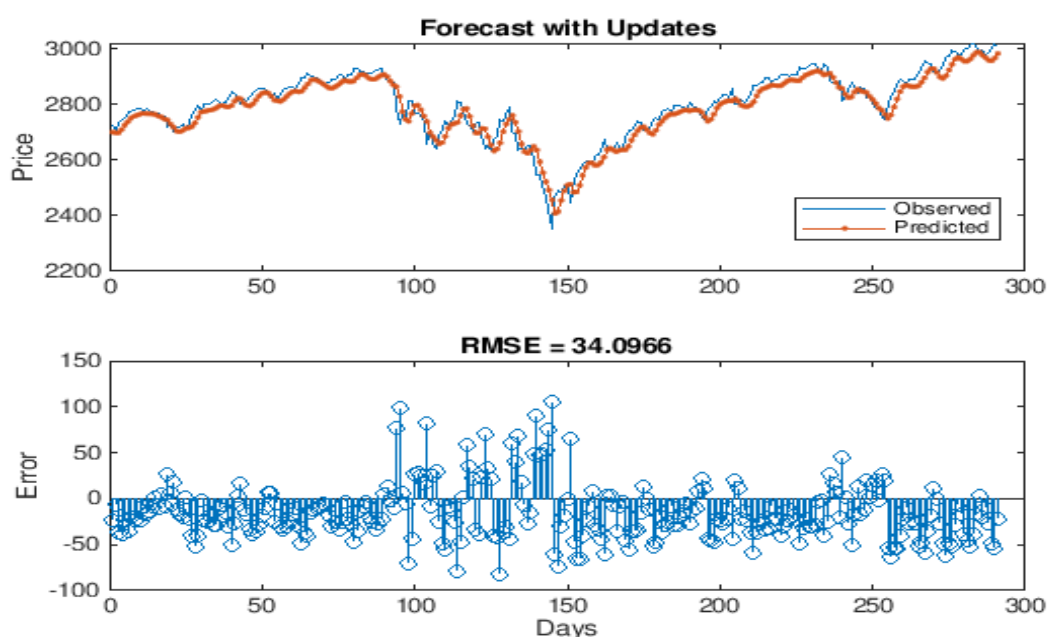
Pro index S&P 500 byla na základě zkoumání zvolena architektura sítě s počtem 400 epoch. Na grafu výše může být viděna klesající chyba modelu při rostoucím počtu tréninkových epoch.

Následně stejným způsobem jako tréninková data jsou standardizována i data testovacího souboru. Nyní jsou data připravena k testovací fázi a je vytvořena predikce. K následnému posouzení výsledku jsou data vzniklé predikce odstandardizovány.

5.3.1.4 Vyhodnocení

Na prvním z grafů níže může být viděna predikční schopnost modelu, kdy modrá křivka znázorňuje použitá testovací data a oranžová křivka vzniklou predikci sítě. Výsledek je u indexu S&P 500 vyhodnocen jako uspokojivý a identického závěru bylo dosaženo i u zbylých šesti indexů. LSTM síť dokázala kopírovat hlavní trendy zkoumaných dat, nicméně dochází zde ke zpoždění ve předpovědi. Tato LSTM síť tedy nedokáže data predikovat s předstihem.

Obrázek 10 Výsledek predikce LSTM neuronové sítě



Zdroj: mathworks.com

V druhém z grafů je zobrazena střední čtvercová chyba predikce v porovnání s reálnými daty, která dosáhla relativně nízké úrovně $RMSE = 34,0966$. V následující tabulce je zobrazena nejnižší dosažená střední čtvercová chyba u ostatních zkoumaných indexů. Tento indikátor chybovosti nicméně nelze porovnávat napříč zkoumanými indexy, z důvodu různého cenového rozpětí zkoumaných indexů.

Tabulka 2 Optimální počet epoch a výsledná RMSE u LSTM neuronové sítě

Index	Optimální počet epoch	RMSE
S&P 500	400	34,1
Dow Jones	200	302,3
NASDAQ	250	108,6
DAX	250	113,5
CAC 40	200	43,5
HANG SENG	150	314,6
NIKKEI 225	300	249,4

V tabulce níže je nejprve vypočtena míra výnosnosti LSTM neuronové sítě na jednotlivých akciových indexech z predikovaných dat. Tyto výnosnosti jsou dále porovnány se skutečnou výnosností za použití jednoduché obchodní strategie Buy and Hold. Rozdíl značí, zda bylo na daném indexu výhodnější použít LSTM neuronovou síť anebo stačilo v počátečním období akciový index nakoupit a držet do konečného data zkoumaného období.

Tabulka 3 Porovnání výnosové míry investice u LSTM neuronové sítě

Index	LSTM neuronová síť (%)	BUY and HOLD (%)	Rozdíl (%)
S&P 500	10,50	10,27	0,23
Dow Jones	9,99	10,03	-0,04
NASDAQ	10,13	10,40	-0,27
DAX	-2,84	-2,41	-0,43
CAC 40	3,38	2,65	0,73
HANG SENG	-5,33	-6,92	1,58
NIKKEI 225	-2,36	-2,00	-0,35

Z tabulky výše může být pozorováno, že LSTM neuronová síť dosáhla nejlepšího výsledku na asijském indexu HANG SENG, kde předčila výnosnost strategie BUY and HOLD o 1,58 %. Nejhoršího výsledku naopak síť dosáhla na německém indexu DAX, kde za zvolenou strategii zaostávala o 0,43 %. Celkově LSTM neuronová síť dosáhla pozitivního výsledku pouze u tří indexů ze sedmi. Nemůže být tedy tvrzeno, že by použití tohoto typu neuronové sítě přinášelo lepší výsledky než pouhý nákup a prodej indexu. Nicméně, výnosnosti obou typů strategií se významně neliší. Může být tedy tvrzeno, že zvolená neuronová síť dokáže správně predikovat vývoj zvolených indexů.

5.3.2 Model NAR

V následující modelaci neuronové sítě pro předpověď vývoje akciových indexů, bylo využito toolboxu programu Matlab. V tomto případě je se jedná o model NAR, který označuje nelineární autoregresivní neuronovou síť. Tento model funguje na principu předpovědi budoucích hodnot časové řady $y(t)$, za využití minulých hodnot této časové řady. Znak d označuje časové zpoždění, v jakém jsou data vkládána do modelu.

$$y(t) = f(y(t-1), \dots, y(t-d))$$

Jelikož zkoumáme predikční schopnost tohoto modelu pro sedm akciových indexů, bude pro grafickou ilustraci použit asijský index HANG SENG.

5.3.2.1 Data

Vstupní datový soubor byl rozdělen v poměru 60 % dat pro tréninkovou fázi, 20 % dat pro fázi validační a 20 % dat pro fázi testovací v následujících obdobích:

Tréninková data: 1. 1. 2008 – 30. 11. 2014

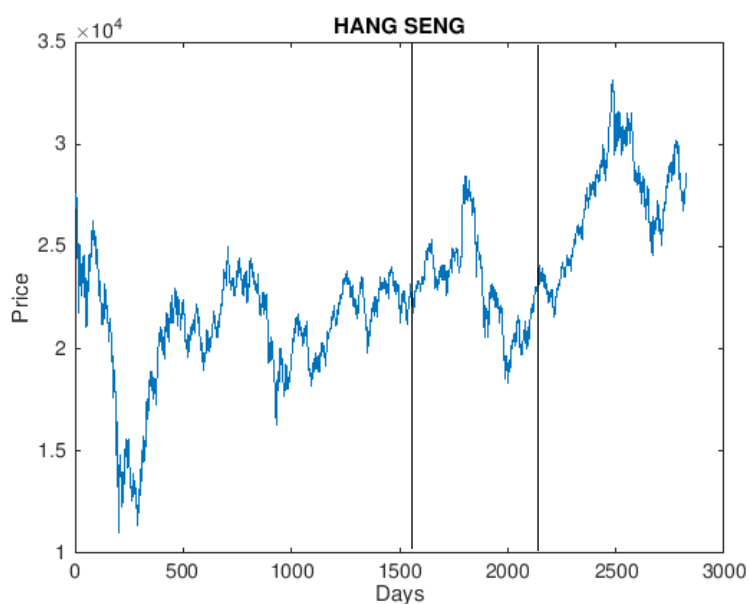
Validační data: 1. 12. 2014 – 31. 03. 2017

Testovací data: 1. 4. 2017 – 26. 07. 2019

Jako vstupní data byly v modelu zvoleny uzavírací ceny (Close) jednotlivých indexů pro každý den, ve kterém byly tyto indexy obchodovány.

Rozdělení datového souboru pro vstupní data uzavíracích cen (Close) může být viděno na grafu níže:

Obrázek 11 Vývoj indexu HANG SENG



Zdroj: mathworks.com

5.3.2.2 Hyperparametry

Po načtení datových souborů a jejich rozdělení na tréninkový, validační a testovací soubor byla data souborů standardizována v souladu s dříve uvedeným vzorcem pro standardizaci. Poté byla zvolena architektura neuronové sítě, kdy bylo přihlíženo k nastavení modelu NAR viz následující tabulka:

Tabulka 4 Hyperparametry NAR neuronové sítě

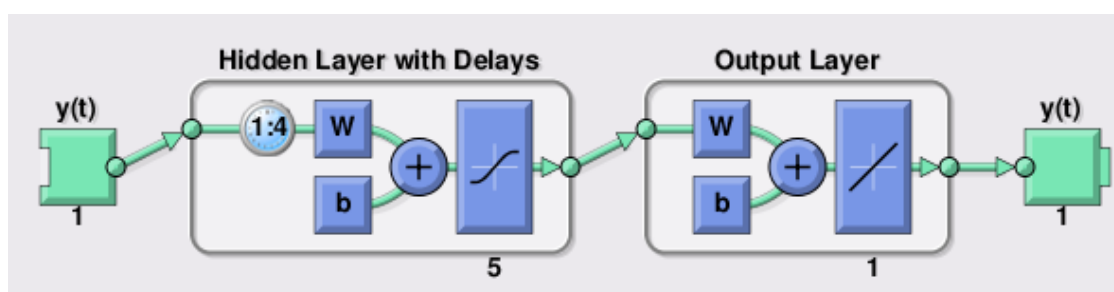
Hyperparametry NAR neuronové sítě	
Typ neuronové sítě	Feed-forward síť
Počet skrytých vrstev	1
Počet neuronů ve skryté vrstvě	5, 10, 15, 20
Optimalizační algoritmus	Levenberg-Marquardt
Časová zpoždění	2, 4, 6, 8
Aktivační funkce ve skryté vrstvě	Logistická
Aktivační funkce ve výstupní vrstvě	Lineární

V tomto případě byla použita feed-forward neuronová síť obsahující jednu skrytou vrstvu, a to z důvodu přeučení při aplikaci více skrytých vrstev. Vhodný počet neuronů ve skryté vrstvě byl testován mezi hodnotami 5, 10, 15 a 20 v kombinaci se vhodným časovým zpožděním na hodnotách 2, 4, 6 a 8. Pro tréninkovou fázi sítě byl zvolen optimalizační algoritmus „Levenberg-Marquardt“. Aktivační funkce ve skryté vrstvě byla zvolena logistická funkce a ve výstupní vrstvě funkce lineární.

5.3.2.3 Tréninková fáze

S tímto nastavením byla síť trénována, kdy na základě testování byl za nejvhodnější počet neuronů u indexu HANG SENG zvolen počet 5 s časovým zpožděním 4 kroků.

Obrázek 12 Architektura NAR neuronové sítě



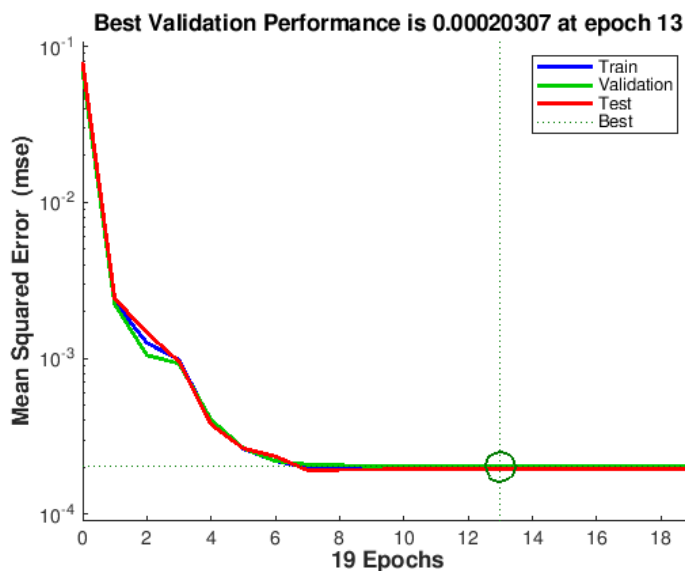
Zdroj: mathworks.com

Na obrázku výše může být viděna architektura NAR sítě použité v tréninkové fázi u indexu HANG SENG.

5.3.2.4 Vyhodnocení

Graf uvedený níže ukazuje pokles střední čtvercové chyby v učící fáze sítě. Trénink sítě skončí, jakmile validační chyba přestane klesat, a to v tomto případě nastalo ve třinácté epoše. Střední čtvercová chyba sítě dosáhla hodnoty 0,00020307, což je hodnota velmi nízká.

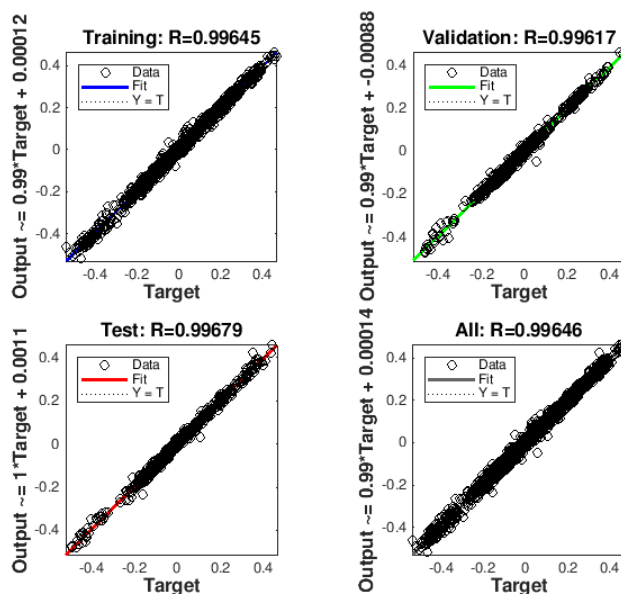
Obrázek 13 Výkonnost tréninkové fáze u NAR neuronové sítě



Zdroj: mathworks.com

U dalšího grafu může být pozorováno, že se podařilo dosáhnout velmi vysoké hodnoty regrese. To značí, že se této neuronové síti podařilo dosáhnout téměř dokonalé kladné korelace mezi predikovanými daty, které síť vytvořila a výstupními (Target) daty, které byly do sítě vloženy.

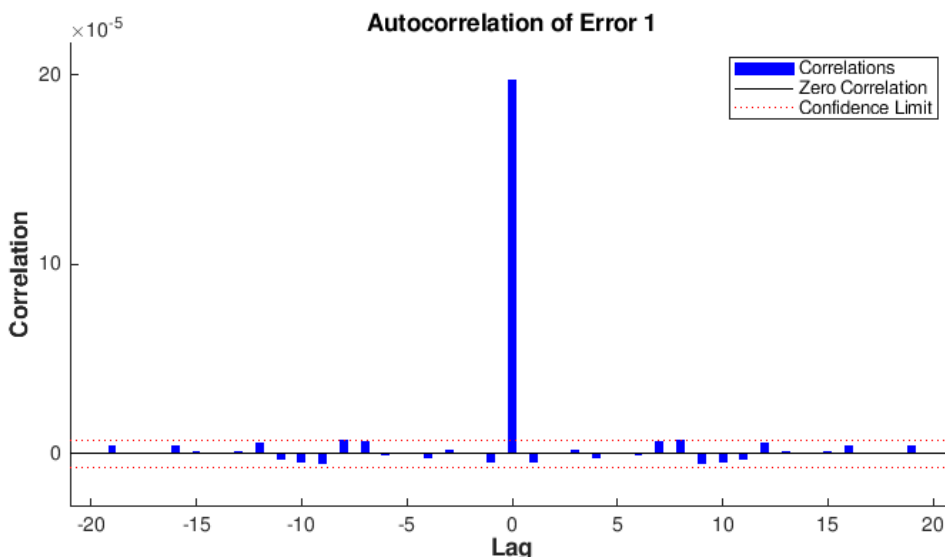
Obrázek 14 Regrese NAR neuronové sítě



Zdroj: mathworks.com

V následujícím grafu je vyobrazena autokorelační funkce chyby, která popisuje vztah mezi chybami predikce v čase. U dostatečně kvalitního predikčního modelu by se měla objevit pouze jedna nenulová hodnota v nulovém zpoždění, čehož zde bylo dosaženo. Znamená to, že chyby predikce byly spolu navzájem naprosto nekorelované. V tomto případě spadají korelace do 95% konfidenčního intervalu okolo nuly, takže predikční model může být vyhodnocen jako kvalitní.

Obrázek 15 Autokorelace chyby NAR neuronové sítě



Zdroj: mathworks.com

V tabulce níže může být viděn optimální počet neuronů skryté vrstvy a optimální časové zpoždění, za kterých zvolený model u zkoumaných indexů dosahuje nejnižší střední čtvercové chyby.

Tabulka 5 Optimální počet neuronů a optimální časové zpoždění NAR neuronové sítě

Optimální struktura NAR sítě a MSE			
Index	Optimální počet neuronů	Optimální časové zpoždění	MSE
S&P 500	20	2	0,000048902
Dow Jones	10	4	0,000049195
NASDAQ	15	6	0,000037646
DAX	20	4	0,000098974
CAC 40	5	2	0,00030291
HANG SENG	5	4	0,00020307
NIKKEI 225	5	6	0,00013481

Zvolený model dosahuje velmi nízkých hodnot této metriky chybovosti u všech zkoumaných akciových indexů a dosáhl také uspokojivých výsledků vzhledem k uvedeným výsledkům u grafů regrese a autokorelace. Tento model tedy může být vyhodnocen jako dostatečně kvalitní pro predikci vývoje zkoumaných akciových indexů.

5.3.3 Model NARX

V následující modelaci neuronové sítě pro předpověď vývoje akciových indexů, bylo využito toolboxu programu Matlab. V tomto případě je se jedná o model NARX, který označuje nelineární autoregresivní neuronovou síť za využití exogenních vstupů. Tento model funguje na principu předpovědi budoucích hodnot časové řady $y(t)$ za využití minulých hodnot této časové řady a také minulých hodnot dalších časových řad $x(t)$. Znak d označuje časové zpoždění v jakém jsou data vkládána do modelu.

$$y(t) = f(y(t-1), \dots, y(t-d), x(t-1), \dots, (t-d))$$

Typicky se tento model využívá k predikci vývoje cen finančních instrumentů jako jsou akcie nebo akciové indexy. Jelikož zkoumáme predikční schopnost tohoto modelu pro sedm akciových indexů, bude pro grafickou ilustraci použit index DAX.

5.3.3.1 Data

Všechny vstupní datové soubory byly rozděleny v poměru 60 % dat pro tréninkovou fázi, 20 % dat pro fázi validační a 20 % dat pro fázi testovací v následujících obdobích:

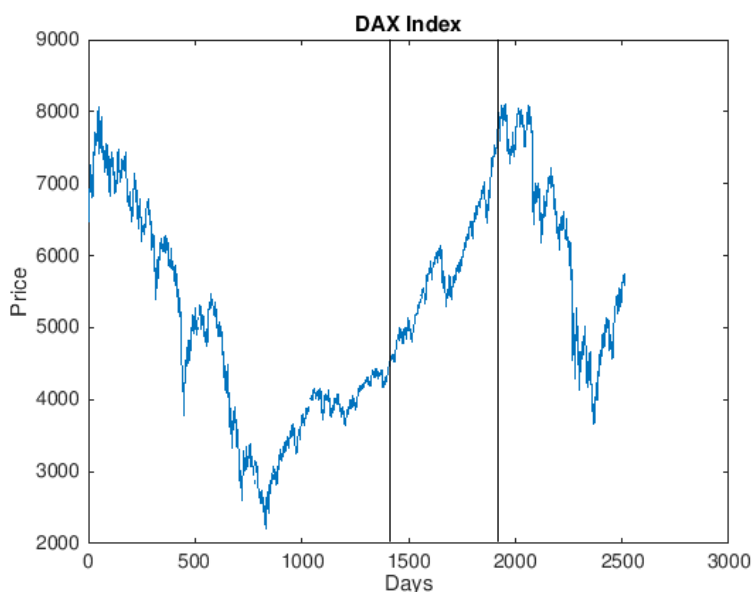
Tréninková data: 1. 1. 2000 – 31. 10. 2005

Validační data: 1. 11. 2005 – 31. 10. 2007

Testovací data: 1. 11. 2007 – 30. 09. 2009

Jako vstupní data byly v modelu zvoleny uzavírací ceny (Close) jednotlivých indexů pro každý den, ve kterém byly tyto indexy obchodovány. Rozdělení datového souboru pro vstupní data uzavíracích cen (Close) může být viděno na grafu níže:

Obrázek 16 Vývoj indexu DAX



Zdroj: mathworks.com

Jako výstupní data (Target), na základě kterých model porovnává úspěšnost svého odhadu, slouží uzavírací ceny (Close) každého dne, kdy byly dané indexy obchodovány. K správné funkčnosti tohoto modelu je požadováno stejného počtu dat vstupů v souhrnu pro tréninkovou, validační a testovací fázi a stejného počtu dat výstupu (Target), které jsou také rozděleny na tréninkové, validační a testovací datové soubory.

Tréninková data: 1. 10. 2009 – 31.8.2015

Validační data: 1. 9. 2015 – 31. 8. 2017

Testovací data: 1. 9. 2017– 26. 7. 2019

5.3.3.2 Hyperparametry

Po načtení datových souborů a jejich rozdělení na tréninkový, validační a testovací soubor byla data souborů standardizována v souladu s dříve uvedeným vzorcem pro standardizaci. Poté byla zvolena architektura neuronové sítě, kdy bylo přihlíženo k nastavení modelu NARX viz následující tabulka:

Tabulka 6 Hyperparametry NARX neuronové sítě

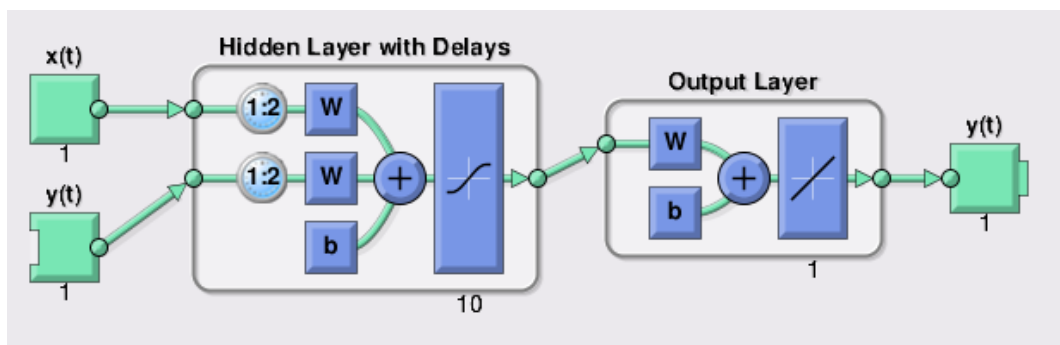
Hyperparametry NARX neuronové sítě	
Typ neuronové sítě	Feed-forward síť
Počet skrytých vrstev	1
Počet neuronů ve skryté vrstvě	5, 10, 15, 20
Optimalizační algoritmus	Levenberg-Marquardt
Časová zpoždění	2, 4, 6, 8
Aktivační funkce ve skryté vrstvě	Logistická
Aktivační funkce ve výstupní funkce	Lineární

V tomto případě byla použita feed-forward neuronová síť obsahující jednu skrytou vrstvu, a to z důvodu přeučení při aplikaci více skrytých vrstev. Vhodný počet neuronů ve skryté vrstvě byl testován mezi hodnotami 5, 10, 15 a 20 v kombinaci se vhodným časovým zpožděním testováno na hodnátach 2, 4, 6 a 8. Pro tréninkovou fázi sítě byl zvolen optimalizační algoritmus „Levenberg-Marquardt“. Aktivační funkce ve skryté vrstvě byla zvolena logistická funkce a ve výstupní vrstvě funkce lineární.

5.3.3.3 Tréninková fáze

S tímto nastavením byla síť trénována, kdy na základě testování byl za nejvhodnější počet neuronů u indexu DAX zvolen počet 10 s časovým zpožděním 2 kroků.

Obrázek 17 Architektura NARX neuronové sítě

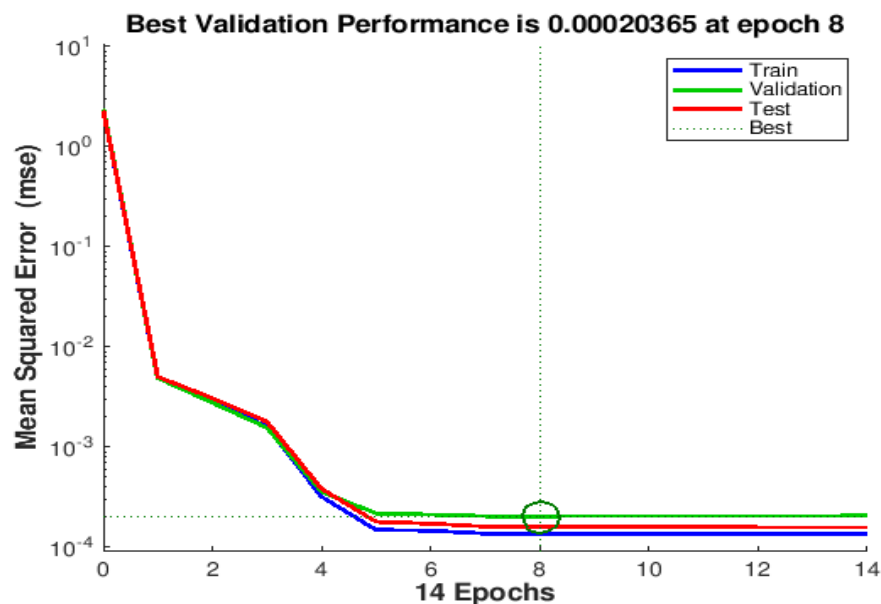


V diagramu výše může být viděna architektura sítě použité v tréninkové fázi u indexu DAX.

5.3.3.4 Vyhodnocení

Graf uvedený níže ukazuje pokles střední čtvercové chyby v učící fázi sítě. Trénink sítě skončí, jakmile validační chyba přestane klesat, a to v tomto případě nastalo ve čtrnácté epoše. Střední čtvercová chyba sítě dosáhla hodnoty 0,00020365, což je hodnota velmi nízká.

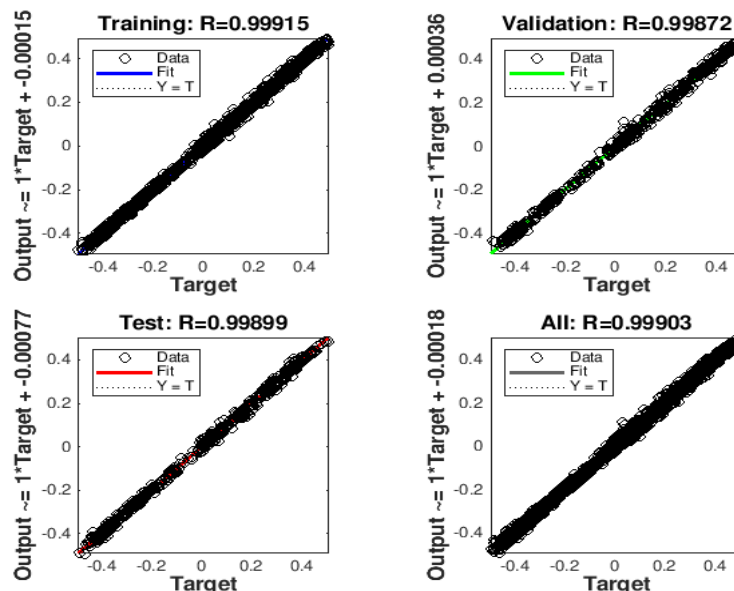
Obrázek 18 Výkonnost tréninkové fáze u NARX neuronové sítě



Zdroj: mathworks.com

U dalšího grafu může být pozorováno, že se podařilo dosáhnout velmi vysoké hodnoty regrese. To značí, že se této neuronové síti podařilo dosáhnout téměř dokonalé kladné korelace mezi predikovanými daty, které síť vytvořila a výstupními (Target) daty, které byly do sítě vloženy.

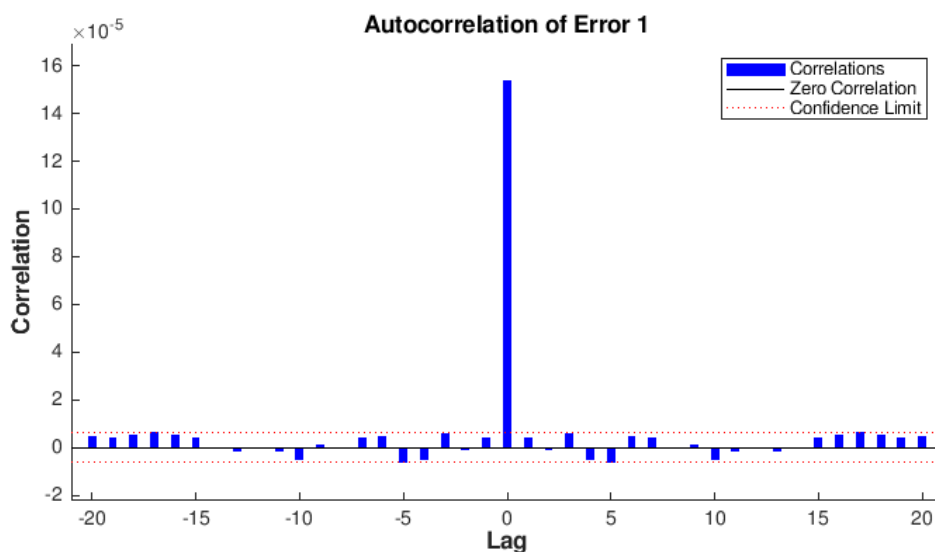
Obrázek 19 Regrese NARX neuronové sítě



Zdroj: mathworks.com

V následujícím grafu je vyobrazena autokorelační funkce chyby, která popisuje vztah mezi chybami predikce v čase. U dostatečně kvalitního predikčního modelu by se měla objevit pouze jedna nenulová hodnota v nulovém zpoždění, čehož zde bylo dosaženo. Znamená to, že chyby predikce byly spolu navzájem naprosto nekorelované. V tomto případě podobně jako u modelu NAR spadají korelace do 95% konfidenčního intervalu okolo nuly, takže predikční model může být vyhodnocen jako kvalitní.

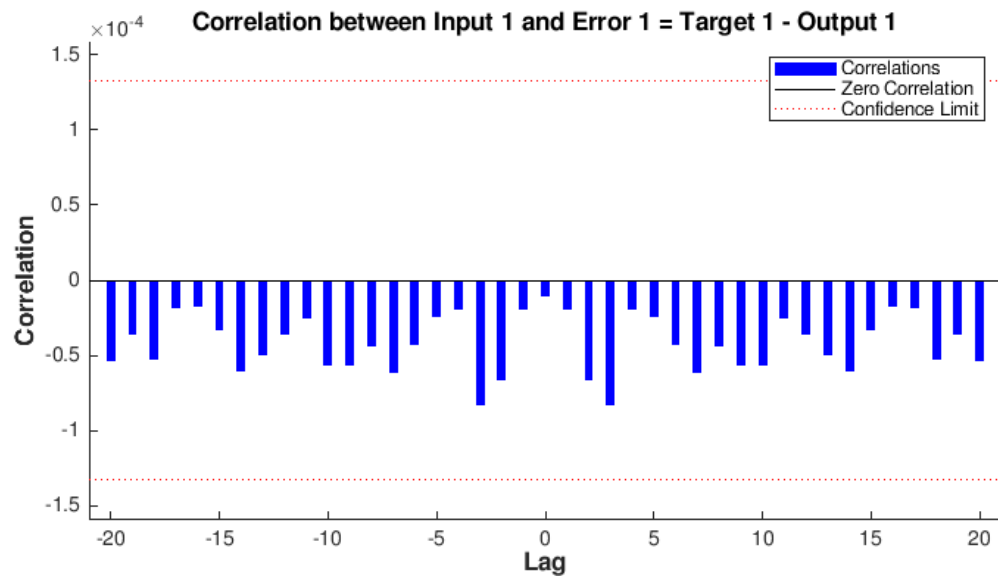
Obrázek 20 Autokorelace chyby NARX neuronové sítě



Zdroj: mathworks.com

Další graf křížové korelace vstupu a chyby predikce také slouží k potvrzení predikční schopnosti vytvořené neuronové sítě.

Obrázek 21 Autokorelace vstupu a chyby predikce NARX neuronové sítě



Zdroj: mathworks.com

Tento graf zobrazuje jak je chyba sítě korelována se vstupní časovou řadou $x(t)$. V uvedeném případě všechny korelace spadají do pásem konfidečního intervalu okolo nuly. Může být tedy tvrzeno, že i v tomto měřítku model splňuje předpoklady k úspěšnému vytváření predikcí.

V tabulce níže může být viděn optimální počet neuronů skryté vrstvy a optimální časové zpoždění, za kterých zvolený model u zkoumaných indexů dosahuje nejnižší střední čtvercové chyby.

Tabulka 7 Optimální počet neuronů a optimální časové zpoždění NARX neuronové sítě

Optimální struktura NARX sítě a MSE

Index	Optimální počet neuronů	Optimální časové zpoždění	MSE
S&P 500	10	6	0,000086245
Dow Jones	10	6	0,00007167
NASDAQ	15	4	0,000072764
DAX	10	2	0,00020365
CAC 40	15	4	0,00028716
HANG SENG	5	4	0,00024227
NIKKEI 225	15	6	0,00012448

Zvolený model dosahuje velmi nízkých hodnot této metriky chybovosti u všech zkoumaných akciových indexů a dosáhl také uspokojivých výsledků vzhledem k uvedeným výsledkům u grafů regrese, autokorelace chyb predikce a křížové korelace vstupních dat s chybami predikce. Tento model tedy může být vyhodnocen jako dostatečně kvalitní pro predikci vývoje zkoumaných akciových indexů.

6 Závěr

Cílem této práce bylo ověřit aplikovatelnost umělé inteligence při obchodování na kapitálových trzích. Pro tyto účely byla vybrána metoda umělé inteligence neuronových sítí. V rámci této metody byly v praktické části této práce skrze program Matlab sestaveny tři modely neuronových sítí. Jako první byla sestavena rekurentní LSTM neuronová síť, dále feed-forward neuronová síť na základě modelu NAR a jako poslední feed-forward neuronová síť na základě modelu NARX. Účelem sestavení těchto modelů bylo ověřit jejich schopnost predikce cenového vývoje na kapitálových trzích. Za zdrojová data pro tyto modely byly vybrány uzavírací ceny každého obchodovaného dne ve zvolené periodě sedmi akciových indexů. Nejprve tři akciové indexy z USA, následně dva akciové indexy z Evropy a jako poslední dva akciové indexy z Asie. Ve zvolených modelech byla testována různá architektura neuronové sítě, a to přes výběr optimálních hyperparametrů pro následnou predikci vývoje cen akciových indexů. Mezi tyto testované hyperparametry mimo další patřil zejména počet skrytých vrstev neuronové sítě, počet neuronů ve skrytých vrstvách, optimalizační algoritmus, maximální počet epoch a časové zpoždění vložení dat do modelu. Pro tréninkovou fázi rekurentní LSTM neuronové sítě byl datový soubor rozdělen na část tréninkovou a testovací a u feed-forward neuronové sítě u modelů NAR a NARX na část tréninkovou, validační a testovací. Dále pro optimální struktura dat pro vložení do modelů byla zkoumaná data standardizována.

V rámci vyhodnocení predikční schopnosti modelu rekurentní LSTM neuronové sítě bylo využito ukazatele RMSE k porovnání chyby predikce neuronové sítě vůči skutečnému cenovému vývoji zkoumaných akciových indexů. Tento ukazatel dosahoval relativně nízkých hodnot pro všechny ze zkoumaných akciových indexů. V rámci této práce byl vývoj tohoto ukazatele graficky ilustrován na indexu S&P 500, kde dosahoval hodnoty $RMSE = 34,0966$, což je hodnota relativně nízká. Tato nízká chybovost byla dále podpořena grafickou ilustrací predikce cenového vývoje, kterou vytvořila neuronová síť vůči skutečnému cenovému vývoji zkoumaného indexu. V této grafické ilustraci byla pozorována správná predikční schopnost neuronové sítě, jelikož zachycuje hlavní trendy vývoje zkoumaného indexu. Nicméně zde dochází k časovému zpoždění predikce cenového vývoje neuronové sítě oproti reálnému cenovému vývoji. Predikční schopnost rekurentní LSTM sítě byla v rámci tohoto ukazatele vyhodnocena za uspokojivou, nicméně potenciální investor využívající tuto strategii musí brát v úvahu časové zpoždění obchodních signálů. Totožný závěr byl učiněn i u zbylých šesti zkoumaných indexů. V rámci vyhodnocení výnosnosti, při využití predikce LSTM neuronové v porovnání s výnosností podle strategie BUY & HOLD za predikované časové období, dosáhla LSTM pozitivního výsledků u tří ze sedmi indexů. Nevyšší výnosnosti dosáhla neuronová síť u indexu HANG SENG, kdy předčila výnosnost strategie BUY & HOLD o 1,58 % a nejnižšího u indexu DAX, kdy za zvolenou strategii zaostávala o 0,43 %. Nemůže být tedy tvrzeno, že by použití tohoto typu neuronové sítě přinášelo výrazně vyšší výsledky než pouhý nákup a prodej indexu. Nicméně, výnosnosti obou typů strategií se významně nelišily. V souhrnu byl učiněn závěr, že zvolená neuronová síť dokázala v porovnání s reálným vývojem trhu správně predikovat vývoj zvolených indexů.

V rámci vyhodnocení predikční schopnosti modelu feed-forward NAR neuronové sítě bylo využito ukazatele MSE k porovnání chyby predikce neuronové sítě vůči skutečnému cenovému vývoji zkoumaných akciových indexů. Tento ukazatel dosahoval velmi nízkých hodnot pro všechny ze zkoumaných akciových indexů. V rámci této práce byl vývoj tohoto ukazatele graficky ilustrován na indexu HANG SENG, kde dosahoval hodnoty $MSE = 0,00020307$, což je hodnota velmi nízká. Z hlediska regrese dosahovaly pozorované akciové indexy velmi vysokých hodnot, což značí kladnou korelaci mezi predikovanými daty, která síť vytvořila a výstupními daty, které byly do sítě vloženy. Dále byla graficky ilustrována autokorelační funkce chyby predikce, která popisuje vztah mezi chybami v čase. V tomto případě spadaly korelace do 95% konfidenčního intervalu okolo nuly, takže predikční model byl u indexu HANG SENG vyhodnocen za kvalitní. Totožných závěrů bylo dosaženo i pro autokorelační funkci chyby predikce u ostatních zkoumaných akciových indexů. Vzhledem k těmto výsledkům byl použitý model feed-forward NAR neuronové sítě vyhodnocen za dostatečně kvalitní pro predikci vývoje zkoumaných akciových indexů.

V rámci vyhodnocení predikční schopnosti modelu feed-forward NARX neuronové sítě bylo také využito ukazatele MSE k porovnání chyby predikce neuronové sítě vůči skutečnému cenovému vývoji zkoumaných akciových indexů. Tento ukazatel dosahoval velmi nízkých hodnot pro všechny ze zkoumaných akciových indexů. V rámci této práce byl vývoj tohoto ukazatele graficky ilustrován na indexu DAX, kde dosahoval hodnoty $MSE = 0,00020365$, což je hodnota velmi nízká. Z hlediska regrese dosahovaly pozorované akciové indexy podobně jako u modelu NAR velmi vysokých hodnot, což značí kladnou korelaci mezi predikovanými daty, která síť vytvořila a výstupními daty, které byly do sítě vloženy. Dále byla graficky ilustrována autokorelační funkce chyby predikce, která popisuje vztah mezi chybami v čase. V tomto případě spadaly korelace do 95% konfidenčního limitu okolo nuly, takže predikční model byl u indexu DAX vyhodnocen za kvalitní. Totožných závěrů bylo dosaženo i pro autokorelační funkci chyby predikce u ostatních zkoumaných akciových indexů. Nakonec byla u indexu DAX graficky ilustrována křížová korelace časové řady vstupů $x(t)$ a chyby predikce. V uvedeném případě všechny korelace spadají do pásem konfidenčního limitu okolo nuly. Může být tedy tvrzeno, že i v tomto měřítku model splňuje předpoklady k úspěšnému vytváření predikcí. Stejných výsledků bylo dosaženo i u zbylých šesti akciových indexů. Vzhledem k výše uvedeným vyhodnocením byl i model feed-forward NARX neuronové sítě vyhodnocen za dostatečně vhodný pro predikci vývoje zkoumaných akciových indexů.

7 Použité zdroje

7 Types of Neural Network Activation Functions: How to Choose? [online]. 2018 [cit. 2019-08-13]. Dostupné z: <https://missinglink.ai/guides/neural-network-concepts/7-types-neural-network-activation-functions-right/>

Chatbots in costumer service [online]. 2016 [cit. 2019-08-14]. Dostupné z: https://www.accenture.com/t00010101t000000_w_/br-pt/acnmedia/pdf-45/accenture-chatbots-customer-service.pdf

ABHISHEK, Kumar, Saswata GHOSH, Kumar GHOSH a M.P. SINGH. *Weather Forecasting Model using Artificial Neural Network* [online]. 2012 [cit. 2019-08-14]. Dostupné z: <https://www.sciencedirect.com/science/article/pii/S221201731200326X>

AISHWARYA.27. *Introduction to Recurrent Neural Network* [online]. 2011 [cit. 2019-08-14]. Dostupné z: <https://www.geeksforgeeks.org/introduction-to-recurrent-neural-network/>

AMADEO, Kimberly. *The S&P 500 and How It Works* [online]. 25.6.2019 [cit. 2019-08-14]. Dostupné z: <https://www.thebalance.com/what-is-the-sandp-500-3305888>

ANURAG. *Random Forest Analysis in ML and when to use it* [online]. 17.8.2018 [cit. 2019-08-15]. Dostupné z: <https://www.newgenapps.com/blog/random-forest-analysis-in-ml-and-when-to-use-it>

ASSIS, Carlos, Adriano C. M. PEREIRA a Eduardo G. CARRANO. *Restricted Boltzmann Machines for the Prediction of Trends in Financial Time Series* [online]. 1.7.2018 [cit. 2019-08-14]. Dostupné z: https://www.researchgate.net/publication/328400175_Restricted_Boltzmann_Machines_for_the_Prediction_of_Trends_in_Financial_Time_Series

CASTLE, Nikki. *What is Semi-Supervised Learning?* [online]. [cit. 2019-08-13]. Dostupné z: <https://www.datascience.com/blog/what-is-semi-supervised-learning>

CAPITAL. *CAC 40 index* [online]. 2019 [cit. 2019-08-14]. Dostupné z: <https://capital.com/cac-40-index-definition>

CAPITAL. *Nasdaq Composite* [online]. 2019 [cit. 2019-08-14]. Dostupné z: <https://capital.com/nasdaq-composite-defintion>

CFI. *Nikkei Index* [online]. 2019 [cit. 2019-08-14]. Dostupné z: <https://corporatefinanceinstitute.com/resources/knowledge/trading-investing/nikkei-index/>

ELMAN, Jeffrey. *Finding structure in time* [online]. 1990 [cit. 2019-08-13]. Dostupné z: <https://crl.ucsd.edu/~elman/Papers/fsit.pdf>

- FISCHER, Thomas a Christopher KRAUSS. *Deep learning with long short-term memory networks for financial market predictions* [online]. 11.1.2017 [cit. 2019-08-14]. Dostupné z: <https://www.econstor.eu/bitstream/10419/157808/1/886576210.pdf>
- DE PERIO, Antonie. *Forecasting aggregate retail sales:: a comparison of artificial neural networks and traditional methods* [online]. 2019 [cit. 2019-08-14]. Dostupné z: <https://towardsdatascience.com/histopathological-cancer-detection-with-deep-neural-networks-3399be879671>
- DEZHIC, Egor. *Understanding Decision Trees* [online]. 26.6.2017 [cit. 2019-08-14]. Dostupné z: <https://becominghuman.ai/understanding-decision-trees-43032111380f>
- DI PERSIO, Luca a Oleksandr. *Artificial Neural Networks architectures for stock price prediction: comparisons and applications* [online]. 2016 [cit. 2019-08-14]. Dostupné z: <https://iris.univr.it/retrieve/handle/11562/955101/60620/Artificial%20Neural%20Networks%20architectures%20for%20stock%20price%20prediction%20comparisons%20and%20applications.pdf>
- DOMINGUEZ, Alejandro. *A History of the Convolution Operation* [online]. 2015 [cit. 2019-08-14]. Dostupné z: <https://pulse.embs.org/january-2015/history-convolution-operation/>
- GANTI, AKHILESH. *Dow Jones Industrial Average (DJIA)* [online]. 1.4.2019 [cit. 2019-08-14]. Dostupné z: <https://www.investopedia.com/terms/d/djia.asp>
- GILL, Navdeep Singh. *Artificial Neural Networks and Neural Networks Applications* [online]. 1.3.2019 [cit. 2019-08-13]. Dostupné z: <https://www.xenonstack.com/blog/artificial-neural-network-applications/>
- GOOGLE. *Cloud Speech-to-Text* [online]. 2006 [cit. 2019-08-14]. Dostupné z: <https://cloud.google.com/speech-to-text/>
- GRAVES,, Alex, Abdel-rahman MOHAMED a Geoffrey HINTON. *Peephole Long Short-Term Memory* [online]. 2017 [cit. 2019-08-14]. Dostupné z: https://commons.wikimedia.org/wiki/File:Peephole_Long_Short-Term_Memory.svg
- GREENEMEIER, Larry. *20 Years after Deep Blue: How AI Has Advanced Since Conquering Chess* [online]. 2017 [cit. 2019-08-14]. Dostupné z: <https://www.scientificamerican.com/article/20-years-after-deep-blue-how-ai-has-advanced-since-conquering-chess/>
- HOCHREITER, Sepp a Jürgen SCHMIDHUBER. *Long Short-Term Memory". Neural Computation* [online]. 1997 [cit. 2019-08-13]. Dostupné z: <https://www.bioinf.jku.at/publications/older/2604.pdf>

CHANSARKAR, Mangesh. *Neural Networks in GPS Navigation* [online]. 2000 [cit. 2019-08-14]. Dostupné z:

https://www.researchgate.net/publication/225637806_Neural_Networks_in_GPS_Navigation

CHEN, JAMES. *DAX* [online]. 8.4.2018 [cit. 2019-08-14]. Dostupné z:

<https://www.investopedia.com/terms/d/dax.asp>

IBARZ, Julian. *Updating Google Maps with Deep Learning and Street View* [online]. 3.3.2017 [cit. 2019-08-14]. Dostupné z: <https://ai.googleblog.com/2017/05/updating-google-maps-with-deep-learning.html>

IMPROVEDOUTCOMES. *Manhattan Distance Metric* [online]. 2004 [cit. 2019-08-14].

Dostupné z:

http://www.improvedoutcomes.com/docs/WebSiteDocs/Clustering/Clustering_Parameters/Manhattan_Distance_Metric.htm

JOSHI, NAVEEN. *The evolution of neural networks* [online]. 16.12.2018 [cit. 2019-08-14].

Dostupné z: <https://www.allerin.com/blog/the-evolution-of-neural-networks>

JHUNJHUNWALA, Abhishek. *An overview of different unsupervised learning techniques* [online]. 23.7.2019 [cit. 2019-08-13]. Dostupné z: <https://towardsdatascience.com/an-overview-of-different-unsupervised-learning-techniques-facb1e1f3a27>

JUSTINB. *Introduction to Perceptron: Neural Network* [online]. 12.9.2017 [cit. 2019-08-14].

Dostupné z: <https://blog.knoldus.com/introduction-to-perceptron-neural-network/>

KVAMME, Håvard, Nikolai SELLEREITE, Kjersti AAS a Steffen. *Predicting Mortgage Default using Convolutional Neural Networks* [online]. 2018 [cit. 2019-08-14]. Dostupné z:

https://www.researchgate.net/publication/323271558_Predicting_Mortgage_Default_using_Convolutional_Neural_Networks

KAMRUZZAMAN, Joarder, Rezaul BEGG a Ruhul A. SARKER. *Artificial Neural Networks in Finance and Manufacturing* [online]. 2006 [cit. 2019-08-14]. Dostupné z:

<https://books.google.cz/books?id=fUy9AQAAQBAJ&pg=PA14&lpg=PA14&dq=future+of+neural+networks+in+finance&source=bl&ots=Ukt4nCnnCy&sig=ACfU3U0YRfCR3Gm2882HOp0eQhTw1KWedA&hl=cs&sa=X&ved=2ahUKEwiOkcnw5NHhAhWwyKYKHU2QCUMQ6AEwA3oEACAcQAQ#v=onepage&q=future%20of%20neural%20networks%20in%20finance&f=false>

KARLSSON, SIMON a MARCUS NORDBERG. *Stock market index prediction using artificial neural networks trained on foreign markets* [online]. 2015 [cit. 2019-08-14]. Dostupné z:

<http://www.diva-portal.org/smash/get/diva2:812198/FULLTEXT01.pdf>

KILITCIOGLU, Doruk. *Hierarchical Clustering and its Applications* [online]. 26.10.2018 [cit. 2019-08-13]. Dostupné z: <https://towardsdatascience.com/hierarchical-clustering-and-its-applications-41c1ad4441a6>

- KOEHRSEN, Will. *An Implementation and Explanation of the Random Forest in Python* [online]. 30.8.2018 [cit. 2019-08-14]. Dostupné z: <https://towardsdatascience.com/an-implementation-and-explanation-of-the-random-forest-in-python-77bf308a9b76>
- KUBOTA, Taylor. *Neural network helps autonomous car learn to handle the unknown* [online]. 30.3.2019 [cit. 2019-08-14]. Dostupné z: <https://www.therobotreport.com/neural-network-autonomous-car-unknown/>
- KUTLU, Birgul a Meltem ÖZTURAN. *Stock Market Prediction Using Artificial Neural Networks* [online]. 10.4.2015 [cit. 2019-08-14]. Dostupné z: https://www.researchgate.net/publication/228407623_Stock_Market_Prediction_Using_Artificial_Neural_Networks
- LANDMAN, Frank. *Everything You Need to Know About the Future of Neural Networks* [online]. [cit. 2019-08-14]. Dostupné z: <https://readwrite.com/2019/01/25/everything-you-need-to-know-about-the-future-of-neural-networks/>
- LAZZERI, Francesca. *Neural Networks for Forecasting Financial and Economic Time Series* [online]. 21.9.2018 [cit. 2019-08-14]. Dostupné z: <https://medium.com/microsoftazure/neural-networks-for-forecasting-financial-and-economic-time-series-6aca370ff412>
- LE, James. *THE 8 NEURAL NETWORK ARCHITECTURES MACHINE LEARNING RESEARCHERS NEED TO LEARN* [online]. 6.1.2018 [cit. 2019-08-14]. Dostupné z: <https://jameskle.com/writes/neural-networks>
- LI, Cheng a Bingyu WANG. *Principal Components Analysis* [online]. 3.11.2014 [cit. 2019-08-13]. Dostupné z: <https://www.edureka.co/blog/autoencoders-tutorial/>
- LIN, Tsong-Wuu a Chan-Chien YU. *Forecasting Stock Market with Neural Networks* [online]. 16.1.2009 [cit. 2019-08-14]. Dostupné z: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=1327544
- MARKETBUSINESSNEWS. *Hang Seng Index – Definition And Meaning* [online]. 2019 [cit. 2019-08-14]. Dostupné z: <https://marketbusinessnews.com/hang-seng-index-definition-meaning/>
- MARTINDALE, Jon. *Cortana vs. Siri vs. Google Assistant* [online]. 8.12.2018 [cit. 2019-08-14]. Dostupné z: <https://www.digitaltrends.com/computing/cortana-vs-siri-vs-google-now/>
- MATHWROKS. *Matlab* [online]. 2019 [cit. 2019-08-14]. Dostupné z: www.mathworks.com
- MCCULLOCH, Warren S. a Walter PITTS. *A logical calculus of the ideas immanent in nervous activity* [online]. The University of Chicago Press 1943, 1943 [cit. 2019-08-13]. ISBN 1522-9602. Dostupné z: <https://link.springer.com/article/10.1007/BF02478259>

MCELROY, MATTHEW. *APPLICATIONS OF THE HOPFIELD NETWORKS* [online]. 2019 [cit. 2019-08-14]. Dostupné z: <http://www.doc.ic.ac.uk/~mtm15/>

MIJWEL, Maad M. *Artificial Neural Networks Advantages and Disadvantages* [online]. 1.1.2018 [cit. 2019-08-14]. Dostupné z: https://www.researchgate.net/publication/323665827_Artificial_Neural_Networks_Advantages_and_Disadvantages

MISHRA, Rachit. *Deep Learning using TensorFlow and NLTK — Training the Neural Network* [online]. 11.8.2017 [cit. 2019-08-14]. Dostupné z: <https://becominghuman.ai/deep-learning-using-tensorflow-and-nltk-training-the-neural-network-part-2-db5c13c02504>

MOAWAD, Assaad. *The magic of LSTM neural networks* [online]. 2.2.2018 [cit. 2019-08-14]. Dostupné z: <https://medium.com/datathings/the-magic-of-lstm-neural-networks-6775e8b540cd>

NICHOLSON, Chris. *Radial Basis Function Networks: 1.1.2019* [online]. [cit. 2019-08-13]. Dostupné z: <https://skymind.com/wiki/radial-basis-function-network-rbf>

NICHOLSON, Chris. *A Beginner's Guide to LSTMs and Recurrent Neural Networks* [online]. 2019 [cit. 2019-08-14]. Dostupné z: <https://skymind.com/wiki/lstm#feedforward>

OLAH, Christopher. *Understanding LSTM Networks* [online]. 27.8.2015 [cit. 2019-08-14]. Dostupné z: <http://colah.github.io/posts/2015-08-Understanding-LSTMs/>

PACELLI, Vincenzo, Vitoantonio BEVILACQUA a Michele AZZOLLINI. *An Artificial Neural Network Model to Forecast Exchange Rates* [online]. 2011 [cit. 2019-08-14]. Dostupné z: https://www.researchgate.net/publication/220062551_An_Artificial_Neural_Network_Model_to_Forecast_Exchange_Rates

PETR, Jaroslav. *Nejlepší stratég na světě? Samouk AlphaGo Zero!* [online]. 23.3.2017 [cit. 2019-08-13]. Dostupné z: <https://vtm.zive.cz/clanky/nejlepsi-strateg-na-svete-samouk-alphago-zero/sc-870-a-190102/default.aspx>

QI, Min, Min QI a Ilan ALON. *Forecasting aggregate retail sales:: a comparison of artificial neural networks and traditional methods* [online]. 2001 [cit. 2019-08-14]. Dostupné z: <https://www.sciencedirect.com/science/article/pii/S0969698900000114>

REESE, Hope. *Understanding the differences between AI, machine learning, and deep learning* [online]. 23.2.2017 [cit. 2019-08-13]. Dostupné z: <https://www.techrepublic.com/article/understanding-the-differences-between-ai-machine-learning-and-deep-learning/>

RAGHUPATHI, Kaushik R. *10 Interesting Use Cases for the K-Means Algorithm* [online]. 27.3.2018 [cit. 2019-08-13]. Dostupné z: <https://dzone.com/articles/10-interesting-use-cases-for-the-k-means-algorithm>

ROKACH, Lior a Oded Z. MAIMON. *DATA MINING AND KNOWLEDGE DISCOVERY HANDBOOK* [online]. Springer Science & Business Media, 2005 [cit. 2019-08-13]. ISBN 9780387254654. Dostupné z:

<https://pdfs.semanticscholar.org/ff4b/91c4de7c263f7cfcfe0ae4e806af73149214.pdf>

ROUSE, Margaret. *Machine learning (ML): 1.5.2018* [online]. [cit. 2019-08-13]. Dostupné z: <https://searchenterpriseai.techtarget.com/definition/machine-learning-ML>

ROUSE, Margaret. *Artificial neural network (ANN)* [online]. 1.7.2018 [cit. 2019-08-13]. Dostupné z: <https://searchenterpriseai.techtarget.com/definition/neural-network>

ROUSE, Margaret. *Convolutional neural network* [online]. 16.11.2017 [cit. 2019-08-14]. Dostupné z: <https://searchenterpriseai.techtarget.com/definition/convolutional-neural-network>

ROUSE, Margaret. *Recurrent neural networks* [online]. 1.7.2017 [cit. 2019-08-14]. Dostupné z: <https://searchenterpriseai.techtarget.com/definition/recurrent-neural-networks>

ROUSE, Margaret. *Support vector machine (SVM)* [online]. 1.11.2017 [cit. 2019-08-14]. Dostupné z: <https://whatis.techtarget.com/definition/support-vector-machine-SVM>

SAK, Hasim, Andrew SENIOR a Francoise BEAUFAYS. *Long Short-Term Memory Recurrent Neural Network Architectures for Large Scale Acoustic Modeling* [online]. 2014 [cit. 2019-08-14]. Dostupné z: <https://static.googleusercontent.com/media/research.google.com/cs//pubs/archive/43905.pdf>

SAYANTINI. *Autoencoders Tutorial : A Beginner's Guide to Autoencoders* [online]. 22.5.2019 [cit. 2019-08-13]. Dostupné z: <https://www.edureka.co/blog/autoencoders-tutorial/>

SAYANTINI. *Restricted Boltzmann Machine Tutorial – Introduction to Deep Learning Concepts* [online]. 2019 [cit. 2019-08-14]. Dostupné z: <https://www.edureka.co/blog/restricted-boltzmann-machine-tutorial/>

SHAH, Jay. *Neural Networks for Beginners: Popular Types and Applications* [online]. 16.11.2017 [cit. 2019-08-14]. Dostupné z: <https://blog.statsbot.co/neural-networks-for-beginners-d99f2235efca>

SHARMA, Aditya. *Demystifying Restricted Boltzmann Machines* [online]. 2.10.2018 [cit. 2019-08-14]. Dostupné z: <https://adityashrm21.github.io/Restricted-Boltzmann-Machines/>

SHARMA, Anirudh. *Is there a difference between neural networks and convolutional neural networks?* [online]. 12.7.2018 [cit. 2019-08-15]. Dostupné z: <https://www.quora.com/Is-there-a-difference-between-neural-networks-and-convolutional-neural-networks>

SINGLETON, J. C. a A. J. *Bond Rating with Neural Networks* [online]. 1993 [cit. 2019-08-14]. Dostupné z:

https://www.researchgate.net/publication/239065973_Bond_Rating_with_Neural_Networks

SONI, Devin. *Supervised vs. Unsupervised Learning* [online]. 22.7.2018 [cit. 2019-08-13]. Dostupné z:

<https://towardsdatascience.com/supervised-vs-unsupervised-learning-14f68e32ea8d>

SRINIVASAN, D., W.S. NG a A.C. LIEW. *Neural-Network-Based Signature Recognition for Harmonic Source Identification* [online]. 2006 [cit. 2019-08-14]. Dostupné z:

<https://ieeexplore.ieee.org/abstract/document/1564224>

STEPHANIE. *Mahalanobis Distance: Simple Definition, Examples* [online]. 21.11.2017 [cit. 2019-08-14]. Dostupné z: <https://www.statisticshowto.datasciencecentral.com/mahalanobis-distance/>

SUTER, David a Qince LI. *Artificial Intelligence Decision Trees* [online]. 7.1.2019 [cit. 2019-08-14]. Dostupné z: https://cs.adelaide.edu.au/~dsuter/Harbin_course/DecisionTrees.pdf

SWAMY, Krishna a Mary PASHLEY. *Neural Network Applications in Finance* [online]. 1997 [cit. 2019-08-14]. Dostupné z: <https://www.researchgate.net/publication/282829832>

TECHOPEDIA. *K-Nearest Neighbor (K-NN)* [online]. 2019 [cit. 2019-08-14]. Dostupné z: <https://www.techopedia.com/definition/32066/k-nearest-neighbor-k-nn>

TECHOPDIA. *Reinforcement Learning* [online]. 2019 [cit. 2019-08-13]. Dostupné z: <https://www.techopedia.com/definition/32055/reinforcement-learning>

TEKNOMO, Kardi. *Euclidian Distance* [online]. 2015 [cit. 2019-08-14]. Dostupné z: <https://people.revoledu.com/kardi/tutorial/Similarity/EuclideanDistance.html>

TORRES, Jordi. *Learning process of a neural network How Do Artificial Neural Networks Learn?* [online]. 21.8.2018 [cit. 2019-08-13]. Dostupné z: <https://towardsdatascience.com/how-do-artificial-neural-networks-learn-773e46399fc7>

TREVINO, Andrea. *Introduction to K-means Clustering* [online]. 6.10.2016 [cit. 2019-08-13]. Dostupné z: <https://www.datascience.com/blog/k-means-clustering>

VALKOV, Venelin. *Making a Predictive Keyboard using Recurrent Neural Networks — TensorFlow for Hackers (Part V)* [online]. 25.5.2017 [cit. 2019-08-14]. Dostupné z: <https://medium.com/@curiously/making-a-predictive-keyboard-using-recurrent-neural-networks-tensorflow-for-hackers-part-v-3f238d824218>

VELOCITY. *Understanding Machine Learning—with pictures of cats and dogs*. [online]. 14.9.2017 [cit. 2019-08-14]. Dostupné z: <https://medium.com/velocity-vlcty/understanding-machine-learning-with-pictures-of-cats-and-dogs-536e1c712d52>

WALIA, Anish Singh. *Activation functions and it's types-Which is better?* [online]. 29.5.2017 [cit. 2019-08-13]. Dostupné z: <https://towardsdatascience.com/activation-functions-and-its-types-which-is-better-a9a5310cc8f>

WALIA, Anish Singh. *Principal Component Analysis – Unsupervised Learning* [online]. 5.2.2018 [cit. 2019-08-13]. Dostupné z: <https://www.edureka.co/blog/autoencoders-tutorial/>

WERBOS, Paul. *Beyond regression : new tools for prediction and analysis in the behavioral sciences* [online]. 1974 [cit. 2019-08-13]. Dostupné z: https://www.researchgate.net/publication/35657389_Beyond_regression_new_tools_for_prediction_and_analysis_in_the_behavioral_sciences

YAVUZ, Mehmet, Necati OZDEMIR, Şakir SAKARYA a Aslan deniz KARAOGLAN. *Stock Market Index Prediction with Neural Network during Financial Crises: A Review on Bist-100* [online]. 1.7.2015 [cit. 2019-08-14]. Dostupné z: https://www.researchgate.net/publication/274616740_Stock_Market_Index_Prediction_with_Neural_Network_during_Financial_Crises_A_Review_on_Bist-100

YUNG, Jessica. *Explaining Tensorflow Code for a Convolutional Neural Network* [online]. 2017 [cit. 2019-08-14]. Dostupné z: <https://www.jessicayung.com/explaining-tensorflow-code-for-a-convolutional-neural-network/>

ZAHRAN, Mohamend. *ASSESSMENT OF ARTIFICIAL NEURAL NETWORK FOR BATHYMETRY ESTIMATION USING HIGH RESOLUTION SATELLITE IMAGERY IN SHALLOW LAKES: CASE STUDY EL BURULLUS LAKE*. [online]. 2015 [cit. 2019-08-14]. Dostupné z: https://www.researchgate.net/figure/A-hypothetical-example-of-Multilayer-Perceptron-Network_fig4_303875065

8 Seznam obrázků

Obrázek 1 Základní fungování neuronové sítě.....	14
Obrázek 2 Princip učení neuronové sítě.....	19
Obrázek 3 Architektura perceptronu.....	29
Obrázek 4 Architektura několikvrstvého perceptronu	30
Obrázek 5 Architektura konvoluční neuronové sítě.....	31
Obrázek 6 Architektura rekurentní neuronové sítě	33
Obrázek 7 Architektura LSTM neuronové sítě.....	34
Obrázek 8 Vývoj indexu S&P 500.....	55
Obrázek 9 Tréninková fáze sítě	57
Obrázek 10 Výsledek predikce LSTM neuronové sítě	58
Obrázek 11 Vývoj indexu HANG SENG.....	60
Obrázek 12 Architektura NAR neuronové sítě.....	61
Obrázek 13 Výkonnost tréninkové fáze u NAR neuronové sítě.....	62
Obrázek 14 Regrese NAR neuronové sítě.....	62
Obrázek 15 Autokorelace chyby NAR neuronové sítě	63
Obrázek 16 Vývoj indexu DAX	64
Obrázek 17 Architektura NARX neuronové sítě	66
Obrázek 18 Výkonnost tréninkové fáze u NARX neuronové sítě	66
Obrázek 19 Regrese NARX neuronové sítě	67
Obrázek 20 Autokorelace chyby NARX neuronové sítě	67
Obrázek 21 Autokorelace vstupu a chyby predikce NARX neuronové sítě.....	68

9 Seznam tabulek

Tabulka 1 Hyperparametry LSTM neuronové sítě	56
Tabulka 2 Optimální počet epoch a výsledná RMSE u LSTM neuronové sítě	58
Tabulka 3 Porovnání výnosové míry investice u LSTM neuronové sítě	59
Tabulka 4 Hyperparametry NAR neuronové sítě	61
Tabulka 5 Optimální počet neuronů a optimální časové zpoždění NAR neuronové sítě	63
Tabulka 6 Hyperparametry NARX neuronové sítě.....	65
Tabulka 7 Optimální počet neuronů a optimální časové zpoždění NARX neuronové sítě	69