

Mladá veda

Young Science

Špeciálne vydanie

Mladá veda

Young Science

MEDZINÁRODNÝ VEDECKÝ ČASOPIS MLADÁ VEDA / YOUNG SCIENCE

Číslo 9, ročník 5., vydané v decembri 2017

ISSN 1339-3189

Kontakt: info@mladaveda.sk, tel.: +421 908 546 716, www.mladaveda.sk

Fotografia na obálke: Château de Chambord. © Branislav A. Švorc, foto.branisko.at

REDAKČNÁ RADA

doc. Ing. Peter Adamišín, PhD. (Katedra environmentálneho manažmentu, Prešovská univerzita, Prešov)

doc. Dr. Pavel Chromý, PhD. (Katedra sociálnej geografie a regionálneho rozvoje, Univerzita Karlova, Praha)

prof. Dr. Paul Robert Magocsi (Chair of Ukrainian Studies, University of Toronto; Royal Society of Canada)

Ing. Lucia Mikušová, PhD. (Ústav biochémie, výživy a ochrany zdravia, Slovenská technická univerzita, Bratislava)

doc. Ing. Peter Skok, CSc. (Ekomos s. r. o., Prešov)

prof. Ing. Róbert Štefko, Ph.D. (Katedra marketingu a medzinárodného obchodu, Prešovská univerzita, Prešov)

prof. PhDr. Peter Švorc, CSc., predseda (Inštitút histórie, Prešovská univerzita, Prešov)

doc. Ing. Petr Tománek, CSc. (Katedra verejnej ekonomiky, Vysoká škola báňská - Technická univerzita, Ostrava)

REDAKCIA

PhDr. Magdaléna Keresztesová, PhD. (Fakulta stredoeurópskych štúdií UKF, Nitra)

Mgr. Martin Hajduk (Inštitút histórie, Prešovská univerzita, Prešov)

RNDr. Richard Nikischer, Ph.D. (Ministerstvo pro místní rozvoj ČR, Praha)

Mgr. Branislav A. Švorc, PhD., šéfredaktor (Vydavateľstvo UNIVERSUM, Prešov)

PhDr. Veronika Trstianska, PhD. (Ústav stredoeurópskych jazykov a kultúr FSŠ UKF, Nitra)

Mgr. Veronika Zuskáčová (Geografický ústav, Masarykova univerzita, Brno)

VYDAVATEĽ

Vydavateľstvo UNIVERSUM, spol. s r. o.

www.universum-eu.sk

Javorinská 26, 080 01 Prešov

Slovenská republika

© Mladá veda / Young Science. Akékoľvek šírenie a rozmnožovanie textu, fotografií, údajov a iných informácií je možné len s písomným povolením redakcie.

PREDIKCE BUDOUCÍHO VÝVOJE CEN AKCIÍ POMOCÍ NEURONOVÝCH SÍTÍ NA ZÁKLADĚ DENNÍCH A MĚSÍČNÍCH DAT O CENĚ AKCIÍ

PREDICTION OF FUTURE DEVELOPMENT OF SHARE PRICES USING NEURAL NETWORKS BASED ON DAILY AND MONTHLY SHARE PRICE DATA

Simona Hašková¹

Autorka působí jako akademický pracovník – odborný asistent – na Ústavu znalectví a oceňování Vysoké školy technické a ekonomické v Českých Budějovicích. Ve svém výzkumu se věnuje podnikové ekonomice a managementu, mikroekonomii, konkrétně například tématům jako je hodnocení podnikatelského prostředí, manažerské optimalizační problémy, intelektuální potenciál manažerů, logistice skladových zásob a v poslední době také metodám komplexního hodnocení podniku a investičnímu rozhodování.

The author appears to be an academic staff – assistant professor – working for the School of Expertness and Valuation in Institute of Technology and Business in České Budějovice. In her research, she focuses on corporate management and economics, microeconomics, specifically on topics such as evaluation of business environment, manager's optimization problems and intellectual potential, inventory's logistics. Lately she has also been focusing on comprehensive evaluation of the company and investment decisions making.

Abstract

Stock prices evolve dynamically and nonlinearly. The price is influenced by many factors, such as global price indexes, aggregated economic activity, exchange rate, interest rate, etc. It is why the stock prices are difficult to predict correctly. However, successful price prediction can be very useful. It is no surprise, that the most accurate prediction attracts analysts from around the world. For the prediction, in this article, artificial neural networks are used. These networks have unquestionable benefits and are therefore suitable for prediction of time series. The aim of this article is to predict future development of share prices of company CEZ, a.s. on Prague Stock Exchange using daily and monthly data and then to compare collected results. For the analysis, share price data collected over 5 and half years are used. Statistica software by DELL is used for processing, specifically a tool using the time series (regression). Each time, 10,000 neural networks are generated, which are multi-layer perceptron networks

¹ Adresa pracoviště: Ing. Simona Hašková, Ph.D., Ústav znalectví a oceňování, Vysoká škola technická a ekonomická v Českých Budějovicích, Okružní 517/10, 370 01 České Budějovice
E-mail: haskova@mail.vstecb.cz

(MLP) or radial basic functions networks (RBF). From those only 5 best results are preserved. 62 work days are predicted for daily data and 2 upcoming months for monthly data. Afterwards, the validity of generated networks is valuated statistically and expertly. The results show that the difference between predictions based on daily and monthly data is between 2.5-3.5%. It is also confirmed that the volume of the entering data correlates with the quality of prediction.

Key words: prediction, stocks, artificial neural networks

Abstrakt

Cena akcií se vyvíjí nelineárně a dynamicky. Cenu ovlivňuje celá řada faktorů, v čele s globálními indexy cen, souhrnnou ekonomickou aktivitou, směnným kurzem, úrokovou mírou apod. Je proto poměrně složité úspěšně cenu akcií předpovídat. Úspěšná predikce budoucího vývoje těchto cen ovšem může být velmi užitečná. Nelze se divit, že právě co nejpřesnější predikce cen akcií láká výzkumníky z celého světa. V tomto příspěvku jsou pro predikci využity umělé neuronové sítě, které mají své nesporné výhody a jsou vhodné právě pro predikci časových řad. Cílem příspěvku je predikovat budoucí vývoj cen akcií společnosti ČEZ, a. s. na Burze cenných papírů Praha na základě denních a měsíčních dat o ceně akcií a následně porovnat získané výsledky. Analyzována jsou data o cenách akcií za více jak 5,5 roku. Pro zpracování je využit software Statistica společnosti DELL, konkrétně pak nástroj využívající časové řady (regresi). Generováno je vždy 10000 neuronových sítí typu vícevrstvé perceptronové sítě (MLP) nebo radiální základní funkce (RBF), z nichž vždy 5 nejlepších výsledků je uchováno. U denních dat je predikováno dalších 62 pracovních dní, u měsíčních dat jsou predikovány následující dva měsíce. Validita vygenerovaných sítí je poté hodnocena statisticky i expertním pohledem. Výsledek uvádí, že rozdíl mezi predikcemi realizovanými na základě denních a naproti tomu měsíčních dat se pohybuje mezi 2,5-3,55 %. Je také potvrzeno, že objem vstupních dat naprosto koreluje s kvalitou predikce.

Klíčová slova: predikce, akcie, umělé neuronové sítě

Úvod

Akciový trh je definován jako veřejný trh, na kterém dochází k obchodování s akciemi a deriváty společnosti za schválenou cenu akcií. Zahedi a Rounaghi (2015) uvádí, že takovýto typ trhu poskytuje makléřům a společnostem příležitost investovat na neutrální půdě a je jedním z hlavních ukazatelů ekonomického stavu země. Vui aj. (2013) představuje akciový trh, jako slibnou finanční investici, která může generovat velké bohatství. Nicméně volatilita akciového trhu ji činí velmi rizikovou investicí. Výzkumní pracovníci používají v počítačových vědách a ekonomice různé metody, které jim umožňují predikovat vývoj cen akcií, a to prostřednictvím získaných informací o pohybu cen akcií.

Přesná predikce cen akcií je velmi obtížná z mnoha důvodů. Nepředvídatelnost je obzvláště pozoruhodná v důsledku globální finanční krize. Také Qun, Xu a Zhang (2017) zmiňují, že zlepšení přesnosti prognózování ceny akcií je velice náročná úloha. Avšak přesná predikce pohybu cen akcií může hrát důležitou roli při pomoci investorům zlepšit návratnost akcií (Ticknor, 2013). Dle Sheelapriya a Murugesan (2016) je velice důležité poskytnout spolehlivý model, který překoná složitou predikci a dokáže zachytit vývoj cen akcií na finančním trhu.

Studie o předpovědi vývoje ceny akcií má velký teoretický význam a aplikační hodnotu. Tradiční metody prognózování akcií nelze aplikovat a dobře analyzovat, jelikož existují problémy, jako je nepřesnost předpovědi, pomalá tréninková rychlost apod. (Vochozka, 2012).

V uplynulých dvou desetiletích byla vyvinuta velká část výzkumu pro predikci vývoje cen akcií. Podle Ghiassi aj. (2005) tato analýza obsahuje zejména mnoho přístupů umělé inteligence, kdy jsou především využívány umělé neuronové sítě. Podle Hassan aj., (2007) neuronové sítě představují široce používanou softwarovou techniku právě pro prognózování akciového trhu. Mabu aj., (2009) rovněž uvádí, že umělé neuronové sítě byly nedávno využity pro aplikace na finančních trzích. Chang (2011) se zaměřuje na umělé neuronové sítě, které jsou velice často používané pro predikci cen akcií. Jeho studie zjistila, že v porovnání s jinými způsoby, jsou neuronové sítě stabilnější metodou pro předvídaní vývoje cen akcií.

Yamashita aj. (2005) prokázali, že multifunkční neuronové sítě mohou mít vyšší reprezentaci a generalizační schopnost než běžné neuronové sítě. Pomocí ceny akcií v časových řadách a dalších informací se tyto sítě mohou naučit předpovědět cenu následujícího dne. Výsledkem jejich simulací je, že navrhovaný systém má lepší přesnost než systém využívající konvenční neuronové sítě. Shi a Liu (2014) zase navrhli metodu predikce prostřednictvím modelu Elmanovy neuronové sítě. Výsledkem této studie byla vyšší přesnost, rychlejší síťová rychlost a krátká trénovací doba.

Göcken aj. (2016) se zaměřují na akciový trh v Turecku, přičemž konstatují, že složité a nejisté chování akciového trhu neumožňuje přesné stanovení vývoje cen akcií, a proto jsou modely prognóz velice žádoucí pro rozhodovací proces investora. Aby se zachytili vztah mezi technickými ukazateli a burzovním trhem ve sledovaném období, používají se zejména umělé neuronové sítě (ANN). Dle De Oliveira (2011) je přístup ANN považován za nejlepší prognostickou metodu s vysokou úrovní v oblasti prognózy vývoje cen akcií.

Lertyngyod a Benjamas (2016) představují predikční model, který předpovídá ceny akcií pomocí technik dolování dat. Tento výzkum umožní investorovi, aby se rozhodl informovaněji nakupovat a prodávat akcie v nejvhodnějším období. Prediktivní koncept v této práci předpokládá učení historických cenových modelů, ukazatelů a chování, a poté předpovídat budoucí vývoj cen akcií v jednotlivých obdobích. Výsledky experimentu ukazují, že předpovědní model pro týdenní (5 až 10 dní) predikci cen akcií je zlepšen díky použití umělé neuronové sítě, kde maximální přesnost modelu dosáhla 93.89% na predikci 10 dnů.

Cílem příspěvku je predikovat budoucí vývoj cen akcií společnosti ČEZ na Burze cenných papírů Praha pomocí neuronových sítí na základě denních a měsíčních dat o ceně akcií a následně porovnat získané výsledky.

Ve vztahu k cíli se nabízí hypotéza: „Rozdíl v predikovaných budoucích cenách akcií společnosti ČEZ na Burze cenných papírů Praha na základě denních a měsíčních dat bude na v následujících dvou měsících maximálně 1 %.

Data a metody

Společnost ČEZ je v rámci ČR jedním z nejvýznamnějších podnikatelských subjektů. Svou činnost, strukturu, vizi a aktivity charakterizuje takto (ČEZ, 2017): „Skupina ČEZ je integrovaným energetickým seskupením působícím v řadě zemí střední a jihovýchodní Evropy a v Turecku s centrálou v České republice. Hlavní předmět podnikání skupiny tvoří výroba, distribuce, obchod a prodej v oblasti elektřiny a tepla, obchod a prodej v oblasti zemního plynu a těžba uhlí. Společnosti Skupiny ČEZ zaměstnávají téměř 27 tisíc zaměstnanců.

Nejvýznamnějším akcionářem mateřské společnosti ČEZ, a. s., je Česká republika s podílem na základním kapitálu (ke dni 14. 6. 2017) téměř 70 %. Akcie ČEZ, a. s., jsou obchodovány na pražské a varšavské burze cenných papírů, kde jsou součástí burzovních indexů PX a WIG-CEE.

Posláním Skupiny ČEZ je zajišťovat bezpečnou, spolehlivou a pozitivní energii zákazníkům i celé společnosti, jejím cílem je přinášet inovace pro řešení energetických potřeb a přispívat k vyšší kvalitě života. Strategie reflektuje zásadní proměnu energetického trhu v Evropě. Skupina ČEZ chce provozovat energetická aktiva maximálně efektivním způsobem a přizpůsobit se rostoucímu podílu decentralizované a bezemisní výroby. Další prioritou je nabízet svým zákazníkům širokou paletu produktů a služeb v synergii s prodejem elektřiny a plynu. Třetí prioritou je aktivně investovat do perspektivních energetických aktiv se zaměřením na region střední Evropy a do podpory moderních technologií v raném stádiu vývoje.

V České republice společnosti Skupiny ČEZ těží a prodávají uhlí, vyrábějí a distribuují elektřinu a teplo, obchodují s elektřinou a dalšími komoditami, prodávají koncovým zákazníkům elektřinu, teplo a zemní plyn a poskytují další služby. Výrobní portfolio tvoří jaderné, uhelné, plynové, vodní, fotovoltaické, větrné a bioplynové zdroje.“

K dispozici jsou data o cenách akcií v období 2. 2. 2012 do 9. 10. 2017, celkem tedy 1442 denních záznamů (dále jako „denní data“). Jedná se o konečné ceny každého dne, v němž v daném časovém úseku byly akcie společnosti obchodovány. Měsíční zůstatky jsou určeny denní konečnou cenou posledního dne obchodování v kalendářním měsíci (dále jako „měsíční data“). K dispozici tak máme 69 záznamů. Data pocházejí z databáze Burzy cenných papírů Praha. Statistiky dat jsou uvedeny v tabulce číslo 1.

Statistika	Cena akcie (denní zůstatky)	Cena akcie (měsíční zůstatky)
Minimum (Trénovací)	364,1000	367,0000
Maximum (Trénovací)	840,0000	782,0000
Průměr (Trénovací)	554,9726	547,9857
Směrodatná odchylka (Trénovací)	117,0231	110,5054
Minimum (Testovací)	370,0000	416,0000
Maximum (Testovací)	816,0000	729,9000
Průměr (Testovací)	546,0833	570,6100
Směrodatná odchylka (Testovací)	117,6275	95,4912
Minimum (Validační)	372,0000	414,0000
Maximum (Validační)	818,5000	807,0000
Průměr (Validační)	540,9051	545,8600

Statistika	Cena akcie (denní zůstatky)	Cena akcie (měsíční zůstatky)
Směrodatná odchylka (Validační)	216,0132	218,2665
Minimum (Overall)	364,1000	367,0000
Maximum (Celkové)	840,0000	807,0000
Průměr (Overall)	551,5338	550,9565
Směrodatná odchylka (Celková)	115,9294	114,0616

Tabulka 1 – Statistiky vstupních dat zkoumaných datových souborů

Zdroj: Autor

Data obou souborů budou zpracována analogicky. Pro zpracování dat bude použit software Statistica ve verzi 12 od společnosti DELL. Bude využit nástroj data miningu neuronové sítě. Konkrétně budeme využívat časové řady (regresí).

Budeme generovat vícevrstvé perceptronové sítě (MLP) a neuronové sítě základní radiální funkce (RBF). Jako nezávislá proměnná bude čas. Za závislou proměnnou určíme cenu akcií společnosti. Časovou řadu rozdělíme na tři soubory – trénovací, testovací a validační. V první skupině bude 70 % vstupních údajů. Na základě trénovací množiny dat vygenerujeme neuronové struktury. Ve zbylých dvou množinách dat ponecháme vždy 15 % vstupních informací. Obě skupiny nám budou sloužit pro ověření spolehlivosti nalezené neuronové struktury, resp. nalezeného modelu. Zpoždění časové řady bude 1. Budeme generovat 10000 neuronových sítí. Z nich uchováme 5, které budou vykazovat nejlepší charakteristiky². Ve skryté vrstvě budeme mít nejméně dva neurony, maximálně pak 20. V případě radiální základní funkce bude ve skryté vrstvě nejméně 21 neuronů, nejvýše pak 30. Pro vícenásobnou perceptronovou síť budeme uvažovat tyto aktivační funkce ve skryté vrstvě a ve vrstvě výstupní:

- Lineární,
- Logistickou,
- Atanh,
- Exponenciální,
- Sinus.

Ostatní nastavení ponecháme defaultní (dle nástroje ANS – automatizované neuronové sítě). Výsledky, pokud nebudou výstupy adekvátní, můžeme následně korigovat úpravou vah jednotlivých neuronů ve struktuře.

Jakmile vygenerujeme neuronové sítě, budeme jejich validitu hodnotit expertně, nejen podle statistických charakteristik. Ideální nám poslouží konfrontace predikcí. Vývoj cen bude predikován na další 2 dnů, v nichž budou akcie obchodovány.

Následně porovnáme výsledky nejlepších predikcí obou datových souborů (tedy měsíčních a denních dat). Základnou pro odvození procenta odchylky bude nejlepší výsledek souboru denních dat.

² Budeme se orientovat pomocí metody nejmenších čtverců. Generování sítí ukončíme, pokud nebude docházet ke zlepšení, tj. ke snižování hodnoty součtu čtverců. Uchováme tedy ty neuronové struktury, jejichž součet čtverců reziduí ku skutečnému vývoji zlata bude co nejnižší (ideálně nula).

Výsledky

Jak již bylo uvedeno, budou zvlášť uvedeny výsledky vycházející z denních dat a měsíčních dat.

Denní data

Tabulka číslo 2 nabízí 5 neuronových sítí s nejlepšími charakteristikami z 10000 vygenerovaných struktur.

Index	Jméno sítě	Trénovací výkon	Testovací výkon	Validační Výkon	Trénovací chyba	Testovací chyba	Validační chyba	Trénovací algoritmus	Chybová funkce	Aktivace skryt.vr.	Výstupní akt.fce
1	MLP 1-2-1	0,997742	0,997614	0,997649	30,86339	31,51797	27,94235	BFGS (Quasi-Newton) 86	Sum.čtvr.	Tanh	Identita
2	MLP 1-2-1	0,997759	0,997617	0,997649	30,62266	31,4799	27,91572	BFGS (Quasi-Newton) 504	Sum.čtvr.	Tanh	Identita
3	MLP 1-7-1	0,997778	0,997605	0,997648	30,36993	31,63007	27,95397	BFGS (Quasi-Newton) 123	Sum.čtvr.	Logistická	Tanh
4	MLP 1-5-1	0,997778	0,997597	0,997648	30,37189	31,74181	27,91681	BFGS (Quasi-Newton) 143	Sum.čtvr.	Tanh	Tanh
5	MLP 1-4-1	0,997774	0,997603	0,997649	30,41342	31,67233	27,89103	BFGS (Quasi-Newton) 88	Sum.čtvr.	Logistická	Tanh

Tabulka 2 – Přehled uchovaných neuronových sítí denních dat

Zdroj: Autor

Všechny uchované neuronové sítě jsou vícevrstvémi perceptronovými neuronovými sítěmi (tedy MLP). To znamená, že neuronové sítě radiální základní funkce splňovaly parametry výkonu a chyby hůře než uchované neuronové sítě. Všechny vznikly pomocí algoritmu Quasi-Newton (vždy pomocí rozdílné verze). K aktivaci skryté vrstvy neuronů využívají dvou funkcí – konkrétně hyperbolického tangentu a logistické funkce. K aktivaci neuronů vnější vrstvy používají rovněž dvě funkce. V tomto případě je to funkce logistická a funkce Identity. Výkon neuronové sítě je popsán hodnotou korelačního koeficientu. Korelační koeficienty všech uchovaných sítí a všech množin dat jsou uvedeny v tabulce číslo 3.

Neuronová síť	Cena akcie Trénovací množina dat	Cena akcie Testovací množina dat	Cena akcie Validační množina dat
1.MLP 1-2-1	0,997742	0,997614	0,997649
2.MLP 1-2-1	0,997759	0,997617	0,997649
3.MLP 1-7-1	0,997778	0,997605	0,997648
4.MLP 1-5-1	0,997778	0,997597	0,997648
5.MLP 1-4-1	0,997774	0,997603	0,997649

Tabulka 3 – Výkon uchovaných neuronových sítí denních dat

Zdroj: Autor

Hledáme ideálně takovou neuronovou síť, která má korelační koeficient ideálně co nejbližší hodnotě 1. Výkon všech tří množin by měl být ideálně obdobný. To znamená, že struktura, která vznikla pomocí trénovací množiny dat, je validní a ověřená na dalších dvou datových souborech. Hodnota korelačních koeficientů všech neuronových struktur a množin dat je vždy vyšší než 0,997. Rozdíly mezi jednotlivými neuronovými sítěmi jsou minimální. Důležitá tak bude analýza statistik predikcí (blíže tabulka číslo 4).

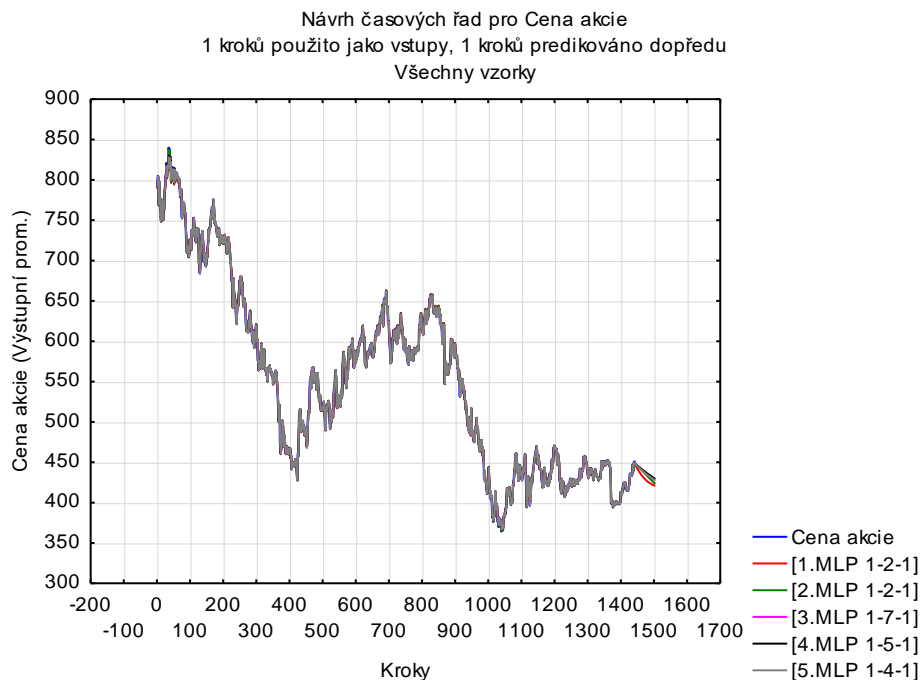
Statistiky	1.MLP 1-	2.MLP 1-	3.MLP 1-	4.MLP 1-	5.MLP 1-
Minimální predikce (Trénovací)	367,3786	364,5671	366,0114	365,8832	366,3944
Maximální predikce (Trénovací)	834,9579	837,4499	830,2773	830,0707	828,0633
Minimální predikce (Testovací)	375,7819	373,7538	374,6131	374,5315	374,9392
Maximální predikce (Testovací)	804,5249	805,9253	807,1347	807,1093	806,9573
Minimální predikce (Validační)	372,8233	370,5248	371,5806	371,4845	371,9272
Maximální predikce (Validační)	817,9336	819,7824	818,9782	819,0903	817,9395
Minimální rezidua (Trénovací)	-54,3914	-53,8817	-52,6843	-52,6893	-52,6833
Maximální rezidua (Trénovací)	33,7995	33,1885	33,0831	32,8209	32,8376
Minimální rezidua (Testovací)	-40,2349	-40,8202	-40,9406	-41,2066	-41,1995
Maximální rezidua (Testovací)	18,6165	18,9097	20,2742	20,2101	20,135
Minimální rezidua (Validační)	-34,459	-34,5719	-34,6799	-34,7063	-34,6833
Maximální rezidua (Validační)	23,0364	23,3467	23,9501	23,9698	23,9495
Minimální standardní rezidua (Trénovací)	-9,7906	-9,7369	-9,56	-9,5606	-9,553
Maximální standardní rezidua (Trénovací)	6,084	5,9974	6,0032	5,9555	5,9544
Minimální standardní rezidua (Testovací)	-7,1668	-7,2754	-7,2795	-7,3139	-7,3207
Maximální standardní rezidua (Testovací)	3,316	3,3703	3,6049	3,5872	3,5778
Minimální standardní rezidua (Validační)	-6,5188	-6,5433	-6,5593	-6,5686	-6,5673
Maximální standardní rezidua (Validační)	4,358	4,4188	4,5299	4,5366	4,5349

Tabulka 4 – Statistiky predikcí jednotlivých neuronových struktur denních dat

Zdroj: Autor

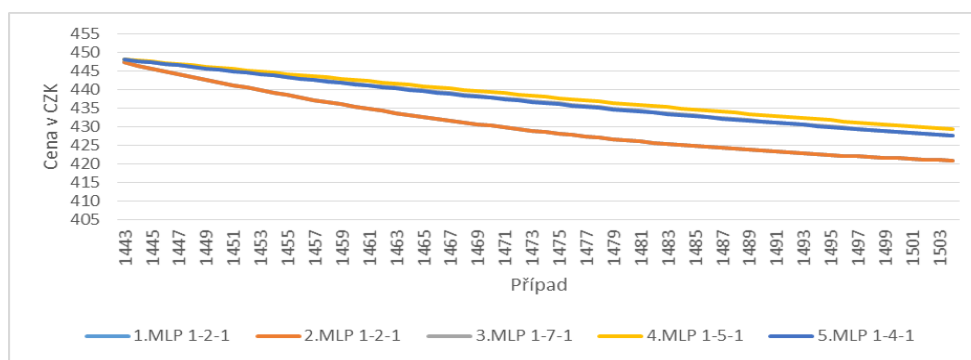
Pokud sledujeme statistiky predikcí jednotlivých neuronových sítí, docházíme nutně k závěru, že rozdíly mezi sítěmi jsou naprosto minimální, a to u všech statistik.

Správný výsledek nám může napovědět i grafický vývoj cen a predikcí. Obrázek číslo 1 nabízí grafické srovnání vývoje skutečných cen akcií společnosti ČEZ a predikcí vypočtených pomocí všech neuronových sítí.



Obrázek 1 – Časové řady s predikcí na 62 obchodovacích dnů
Zdroj: Autor

Z obrázku je patrné, že všechny sítě dokázaly kopírovat v minulých datech zcela přesně skutečný pohyb cen. To nahrává závěru, že jsou všechny sítě aplikovatelné v praxi. Na konci křivky je vložena predikce vývoje dalších 62 obchodních dnů. Obrázek číslo 2 nabízí detailní pohled právě na těchto 62 obchodních dnů. Případ číslo 1442 odpovídá dnu 9. 10. 2017.



Obrázek 2 – Predikce budoucích 62 obchodních dnů
Zdroj: Autor

Z grafu je patrný určitý rozptyl budoucích predikcí. Je však otázkou, jak je rozptyl významný. Nyní už vidíme zřejmý pokles ceny akcií, jenž nastane po 9. 10. 2017.

Měsíční data

Uchované neuronové struktury druhého datového souboru jsou uvedeny v tabulce číslo 5.

Index	Jméno sítě	Trénovací výkon	Testovací výkon	Validační výkon	Trénovací chyba	Testovací chyba	Validační chyba	Trénovací algoritmus	Chybová funkce	Aktivace skryt.vr.	Výstupní akt.fce
1	RBF 1-12-1	0,956492	0,93962	0,999056	450,0545	984,275	789,3047	RBFT	Sum.čtvr.	Gausova	Identita
2	RBF 1-12-1	0,958746	0,936236	0,998977	426,5114	1097,348	369,0861	RBFT	Sum.čtvr.	Gausova	Identita
3	RBF 1-13-1	0,959655	0,942661	0,998949	417,2865	957,137	480,9784	RBFT	Sum.čtvr.	Gausova	Identita
4	RBF 1-11-1	0,956401	0,948865	0,999179	450,9036	955,601	787,6628	RBFT	Sum.čtvr.	Gausova	Identita
5	RBF 1-14-1	0,960215	0,939536	0,99896	411,2249	1231,386	603,889	RBFT	Sum.čtvr.	Gausova	Identita

Tabulka 5 – Přehled uchovaných neuronových sítí (měsíčních dat)

Zdroj: Autor

V tomto případě jsou uchovány sítě jednoho typu. Jedná se neuronové struktury základní radiální sítě. Ve všech případech je použit totožný trénovací algoritmus (RBFT), chybová funkce (suma čtverců), funkce k aktivaci skryté vrstvy neuronů (Gausova křivka) a funkce k aktivaci neuronů ve výstupní vrstvě (identita).

Tabulka číslo 6 nabízí korelační koeficienty jednotlivých sítí a podmnožin dat.

	Korelační koeficienty (Vývoj ceny akcií-ČEZ)		
	Cena akcií Trénovací	Cena akcií Testovací	Cena akcií Validační
1.RBF 1-12-1	0,956492	0,939620	0,999056
2.RBF 1-12-1	0,958746	0,936236	0,998977
3.RBF 1-13-1	0,959655	0,942661	0,998949
4.RBF 1-11-1	0,956401	0,948865	0,999179
5.RBF 1-14-1	0,960215	0,939536	0,998960

Tabulka 6 – Výkon uchovaných neuronových sítí měsíčních dat

Zdroj: Autor

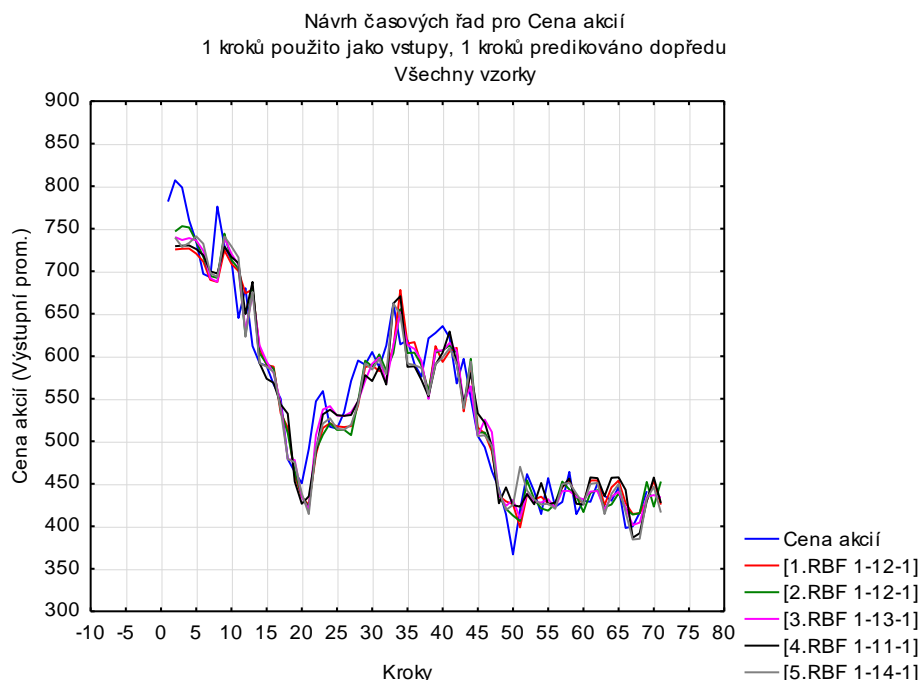
Z hodnot je zřejmý vysoký výkon všech uchovaných neuronových sítí. Nejnižší hodnota se pohybuje na hranici 0,93. V tomto případě se jedná o testovací množinu dat. Nejnižší hodnota trénovacího souboru dat je téměř 0,96. Validační množina dat se vykazuje ještě vyšší hodnotou korelačních koeficientů. U všech uchovaných neuronových sítí dosahuje téměř 0,999 a více. Optimálně hledáme u všech souborů dat jedné sítě maximálně vysokou hodnotu korelačních koeficientů, ideálně navíc ve všech souborech shodnout. V tomto případě vykazují všechny soubory poměrně vysoké hodnoty, avšak rozdíly mezi soubory jsou patrné. Tabulka číslo 7 nabízí statistiky predikcí podle jednotlivých neuronových sítí a souborů dat.

Statistiky	1.RBF 1-12-1	2.RBF 1-12-1	3.RBF 1-13-1	4.RBF 1-11-1	5.RBF 1-14-1
Minimální predikce (Trénovací)	398,8168	406,1358	401,5168	386,6341	384,4943
Maximální predikce (Trénovací)	726,659	751,5191	738,9587	730,2952	741,0764
Minimální predikce (Testovací)	415,4373	415,3396	404,416	391,853	385,3136
Maximální predikce (Testovací)	724,6652	744,2394	739,7569	728,8597	740,8257
Minimální predikce (Validační)	429,3836	420,6314	424,0088	445,8295	419,2661
Maximální predikce (Validační)	726,3731	752,9488	740,224	730,1027	739,6118
Minimální rezidua (Trénovací)	-66,497	-69,3407	-64,1227	-75,3936	-63,6308
Maximální rezidua (Trénovací)	88,6058	83,8271	88,3735	78,6849	82,8035
Minimální rezidua (Testovací)	-55,1931	-58,0869	-62,7698	-64,5763	-71,1287
Maximální rezidua (Testovací)	66,0527	68,6887	71,3187	68,2667	76,6054
Minimální rezidua (Validační)	-25,0456	-10,5153	-13,5069	-36,6564	-21,0177
Maximální rezidua (Validační)	81,4159	60,2372	66,776	77,4804	70,008
Minimální standardní rezidua (Trénovací)	-3,1345	-3,3576	-3,139	-3,5505	-3,1378
Maximální standardní rezidua (Trénovací)	4,1767	4,059	4,3262	3,7055	4,0833
Minimální standardní rezidua (Testovací)	-1,7592	-1,7535	-2,0289	-2,089	-2,027
Maximální standardní rezidua (Testovací)	2,1054	2,0735	2,3052	2,2084	2,183
Minimální standardní rezidua (Validační)	-0,8915	-0,5473	-0,6159	-1,3061	-0,8553
Maximální standardní rezidua (Validační)	2,8979	3,1355	3,0448	2,7607	2,8488

Tabulka 7 – Statistiky predikcí jednotlivých neuronových struktur měsíčních dat
 Zdroj: Autor

Z hodnot uvedených v tabulce je patrný rozdíl mezi jednotlivými soubory dat, i mezi jednotlivými sítěmi. I přesto se mohou všechny vygenerované a uchované neuronové sítě jevit jako použitelné pro predikci cen společnosti.

Více napovědět může ještě graf na obrázku číslo 3. Obrázek demonstruje vývoj cen akcií společnosti ČEZ v celém zkoumaném období. Navíc připojuje predikci vývoje na měsíc říjen a listopad 2017. Na rozdíl od grafu na obrázku číslo 1 je patrný rozdíl v průběhu cen odhadovaných pomocí neuronových sítí. Rozdíly mezi sítěmi jsou tak významnější. Průběh skutečných cen akcií popisuje modrá křivka. Ostatní barvy jsou vyhrazeny pro neuronové sítě. I přesto, že jsou mezi sítěmi na obrázku patrné rozdíly (a to i mezi skutečným průběhem ceny akcií a predikcemi), není možné určit jednu síť, která by byla ve svých predikcích přesnější než sítě ostatní. Rozhodnutí bude muset být učiněno na základě analýzy reziduí, kdy budeme hledat nejen ideální součet reziduí (tedy hodnotu 0), ale i minimální hodnoty reziduí.



Obrázek 3 – Časové řady s predikcí na říjen a listopad 2017
Zdroj: Autor

Komparace

Tabulka číslo 8 nabízí predikce podle výsledků denních dat od prosince roku 2016 do listopadu 2017. Abychom mohli provést srovnání obou výsledků, musíme srovnávat data ve stejném čase. Proto volíme okamžiky srovnání ve stejných intervalech, v nichž máme k dispozici výsledky získané pomocí měsíčních dat. Jedná se tak o 12 údajů, vždy končnou cenu vždy za poslední den obchodování v kalendářním měsíci zkráceného sledovaného období.

Datum	1.MLP 1-2-1	2.MLP 1-2-1	3.MLP 1-7-1	4.MLP 1-5-1	5.MLP 1-4-1	Maximum	Minimum	Rozdíl max - min	% z max. ceny
30.12.16	427,2346	427,1991	427,3468	427,3625	427,3667	427,3667	427,1991	0,1676	0,04%
31.01.17	423,2659	423,1529	423,3021	423,312	423,3358	423,3358	423,1529	0,1829	0,04%
28.02.17	452,1288	452,645	452,6047	452,677	452,5984	452,677	452,1288	0,5482	0,12%
31.03.17	438,3455	438,5789	438,6431	438,6795	438,6382	438,6795	438,3455	0,334	0,08%
28.04.17	431,1208	431,1759	431,3023	431,3244	431,3112	431,3244	431,1208	0,2036	0,05%
31.05.17	447,1901	447,6179	447,6083	447,6664	447,5987	447,6664	447,1901	0,4763	0,11%
30.06.17	403,6523	403,4048	403,2388	403,2255	403,3725	403,6523	403,2255	0,4268	0,11%
31.07.17	400,7199	400,3982	400,2299	400,2128	400,3821	400,7199	400,2128	0,5071	0,13%
31.08.17	418,3502	418,1775	418,2851	418,2886	418,3392	418,3502	418,1775	0,1727	0,04%
29.09.17	446,2044	446,6127	446,6103	446,6657	446,6006	446,6657	446,2044	0,4613	0,10%
31.10.17	436,6856	436,6856	442,4357	443,3264	442,2806	443,3264	436,6856	6,6408	1,50%
30.11.17	426,7581	426,7581	434,9907	436,5383	434,7774	436,5383	426,7581	9,7802	2,24%

Tabulka 8 – Analýza přesnosti výsledků neuronových sítí uchovaných na základě denních dat
Zdroj: Autor

Tabulka kromě dílčích predikcí vypočtených pomocí uchovaných neuronových sítí nabízí sumář maximálních a minimálních predikcí daného případu. Dále vypočítává rozdíl mezi minimální a maximální predikcí. Následně je tento rozdíl porovnán s maximální predikcí a vyjádřen procentním podílem rozdílu na maximální predikci. Zjišťujeme tak, že mezi minimální a maximální predikcí se pohybuje rozdíl mezi 0,04 a 2,24 %. Všechny sítě jsou tak použitelné. Analytik může očekávat, že se cena bude pohybovat vždy v rozmezí minima a maxima pro daný den.

Tabulka číslo 9 předkládá analogický výsledek pro měsíční data.

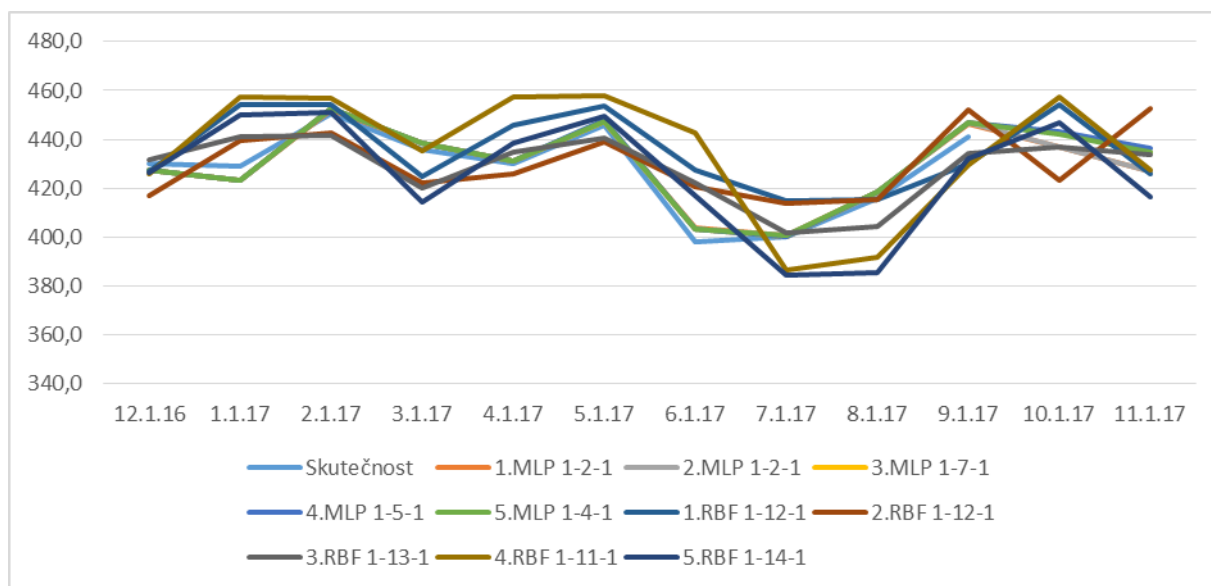
Datum	1.RBF 1-12-1	2.RBF 1-12-1	3.RBF 1-13-1	4.RBF 1-11-1	5.RBF 1-14-1	Maximum	Minimum	Rozdíl max - min	% z max. ceny
30.12.16	426,61	416,7125	431,6366	425,6644	426,1804	431,6366	416,7125	14,9241	3,46%
31.01.17	453,9456	439,4153	440,8238	457,5051	449,9177	457,5051	439,4153	18,0898	3,95%
28.02.17	453,9942	442,4859	441,4672	456,5889	451,0608	456,5889	441,4672	15,1217	3,31%
31.03.17	424,9564	422,2617	419,7804	435,4	414,3779	435,4	414,3779	21,0221	4,83%
28.04.17	445,8807	425,6627	434,7322	457,0981	438,2117	457,0981	425,6627	31,4354	6,88%
31.05.17	453,8812	438,8546	440,688	457,6376	449,6681	457,6376	438,8546	18,783	4,10%
30.06.17	427,3565	420,7822	422,3149	442,892	416,8162	442,892	416,8162	26,0758	5,89%
31.07.17	414,5496	413,8823	401,5168	386,6341	384,4943	414,5496	384,4943	30,0553	7,25%
31.08.17	415,4373	415,3396	404,416	391,853	385,3136	415,4373	385,3136	30,1237	7,25%
29.09.17	429,7974	452,1717	434,3781	429,7174	432,1784	452,1717	429,7174	22,4543	4,97%
31.10.17	453,9936	423,185	436,7207	457,3	446,5764	457,3	423,185	34,115	7,46%
30.11.17	425,5637	452,6425	433,7978	427,4354	416,1643	452,6425	416,1643	36,4782	8,06%

Tabulka 9 – Analýza přesnosti výsledků neuronových sítí uchovaných na základě měsíčních dat

Zdroj: Autor

Rozdíl mezi minimální a maximální predikcí pomocí neuronových sítí radiální základní funkce se pohybuje mezi 3,31 a 8,06 %. I podle těchto predikcí můžeme uvažovat, že se v budoucnu bude cena pohybovat v intervalu predikovaných minimálních a maximálních cen. Přitom v 30. 11. 2017 očekáváme maximum predikované na základě denních dat 436,5383 CZK, podle měsíčních dat 452,6425 CZK. Rozdíl tak činí více než 16 CZK. Minimum predikované ve stejný den podle neuronových struktur vzniklých na základě denních dat činí 426,7581 CZK, podle měsíčních dat 416,1643 CZK. Rozdíl tak dělá více než 10 CZK.

Graf na obrázku vykresluje pohyb cen ve sledovaných časových bodech zkoumaného období.



Obrázek 4 – Průběh skutečného vývoje cen akcií a predikcí v období prosinec 2016 – listopad 2017
Zdroj: Autor

Graf nabízí průběh skutečných cen společnosti ČEZ a predikce podle všech uchovaných neuronových sítí (vzniklých pomocí denních dat i dat měsíčních). Mezi predikcemi a skutečností je patrný rozdíl. Nejvíce se hodnotami i tvarem blíží skutečnosti síť 5. MLP 1-4-1, tedy síť sestavená na základě denních dat.

Jak již bylo naznačeno výše, důležité pro rozhodnutí o vhodnosti sítě (ale i pro potvrzení, či vyvrácení stanovené hypotézy) bude analýza reziduí (blíže tabulka číslo 10).

Datum	1. MLP 1-2-1	2. MLP 1-2-1	3. MLP 1-7-1	4. MLP 1-5-1	5. MLP 1-4-1	1. RBF 1-12-1	2. RBF 1-12-1	3. RBF 1-13-1	4. RBF 1-11-1	5. RBF 1-14-1
30.12.16	-2,8	-2,8	-2,7	-2,6	-2,6	-3,4	-13,3	1,6	-4,3	-3,8
31.01.17	-5,6	-5,7	-5,6	-5,6	-5,6	25,0	10,5	11,9	28,6	21,0
28.02.17	1,6	2,1	2,1	2,2	2,1	3,5	-8,0	-9,0	6,1	0,6
31.03.17	2,3	2,6	2,6	2,7	2,6	-11,0	-13,7	-16,2	-0,6	-21,6
28.04.17	0,9	1,0	1,1	1,1	1,1	15,7	-4,5	4,5	26,9	8,0
31.05.17	1,2	1,6	1,6	1,7	1,6	7,9	-7,1	-5,3	11,6	3,7
30.06.17	5,7	5,4	5,2	5,2	5,4	29,4	22,8	24,3	44,9	18,8
31.07.17	0,8	0,5	0,3	0,3	0,5	14,6	14,0	1,6	-13,3	-15,4
31.08.17	2,4	2,2	2,3	2,3	2,3	-0,6	-0,7	-11,6	-24,1	-30,7
29.09.17	4,9	5,3	5,3	5,4	5,3	-11,5	10,9	-6,9	-11,6	-9,1
Průměr	1,14	1,22	1,24	1,26	1,27	6,96	1,08	-0,50	6,42	-2,86
Minimum	-5,6	-5,7	-5,6	-5,6	-5,6	-11,5	-13,7	-16,2	-24,1	-30,7
Maximum	5,7	5,4	5,3	5,4	5,4	29,4	22,8	24,3	44,9	21,0
Vzdálenost	11,286	11,151	10,908	10,953	10,936	40,859	36,520	40,534	69,039	51,704

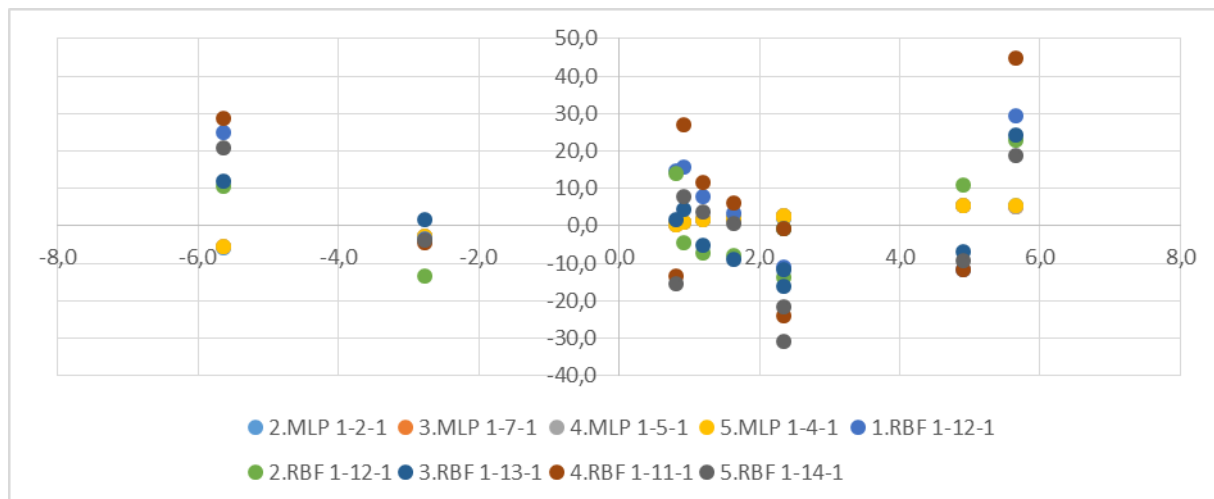
Tabulka 10 – Rezidua neuronových sítí (denních i měsíčních dat) v období prosinec 2016 až září 2017
Zdroj: Autor

Z tabulky je zřejmé, že síť MLP sestavené na základě denních dat vykazují:

- Menší rozptyl průměrných reziduí než síť RBF.

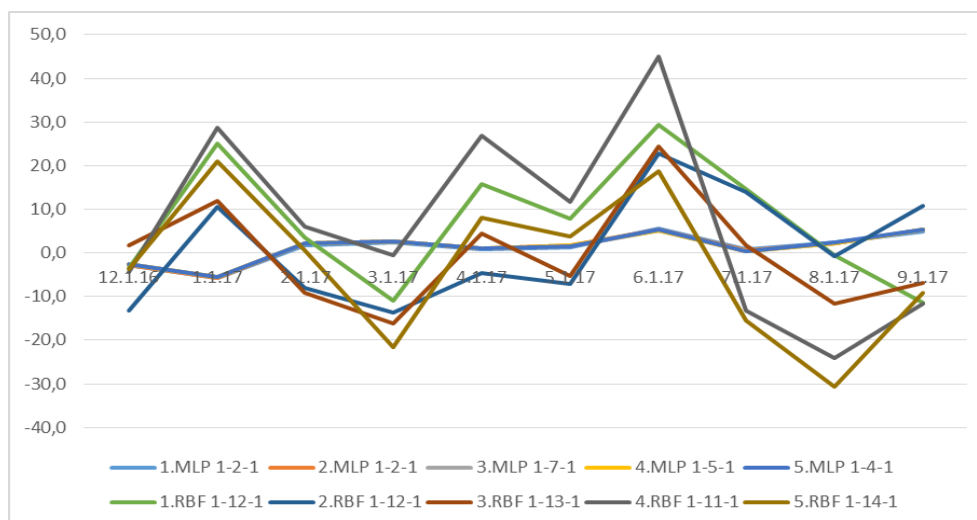
- Minima reziduí bližší nule než sítě RBF.
- Maxima reziduí bližší nule než sítě RBF.
- Menší vzdálenost minima a maxima reziduí na ose než sítě RBF.

Graf na obrázku číslo 5 vykresluje rezidua v podobě bodů v barvách jednotlivých neuronových sítí.



Obrázek 5 – Bodový graf reziduí neuronových sítí (denních i měsíčních dat) v období prosinec 2016 až září 2017
Zdroj: Autor

Z grafu je patrné větší rozptýlení reziduí predikcí sítí RBF. Zajímavý však je rovněž průběh reziduí v čase podle jednotlivých sítí (obrázek číslo 6).



Obrázek 6 – Vývoj reziduí neuronových sítí (denních i měsíčních dat) v období prosinec 2016 až září 2017
Zdroj: Autor

Z grafu je patrný velmi obdobný průběh reziduí sítí MLP. Rozdíly jsou minimální. Naopak rozdíly reziduí sítí RBF jsou patrné z grafu opticky.

Na úvod byla stanovena hypotéza: „Rozdíl v predikovaných budoucích cenách akcií společnosti ČEZ na Burze cenných papírů Praha na základě denních a měsíčních dat bude na v následujících dvou měsících maximálně 1 %.“ Na základě výše uvedeného můžeme konstatovat, že hypotéza byla vyvrácena. Rozdíl mezi predikcemi realizovanými sítěmi RBF (měsíční data) a MLP (denní data) je významnější než 1 %, konkrétně se pohybuje v rozmezí téměř 2,5 % až 3,55 %.

Závěr

Cílem příspěvku bylo predikovat budoucí vývoj cen akcií společnosti ČEZ na Burze cenných papírů Praha pomocí neuronových sítí na základě denních a měsíčních dat o ceně akcií a následně porovnat získané výsledky.

Na základě dvou souborů dat bylo vygenerováno a uchováno vždy pět neuronových struktur. V případě denních dat byly uchovány pouze sítě MLP. Na podkladu měsíčních dat byly uchovány pouze sítě RBF, přestože byla použita shodná metodiky. Může se tak jevit, že v případě nižšího počtu dat jsou přesnější sítě RBF. A naopak, v případě vyššího počtu dat jsou vhodnější sítě MLP.

V každém případě se potvrdilo na základě analýzy predikcí a reziduí, že objem vstupních dat má pro kvalitu predikcí naprosto zásadní význam. Při nižším počtu vstupních dat je rozdíl mezi minimální a maximální predikcí ceny akcií společnosti ČEZ k 30. 11. 2017 8,06 %. V případě vyššího počtu dat (denních dat) je rozdíl pouze 2,24 %. Naprosto tak koreluje objem vstupních dat a kvalita predikce.

Cíl příspěvku byl splněn.

Tento článek doporučil k publikování ve vědeckém časopise Mladá veda:

Mgr. Zdeněk Caha, MBA, Ph.D.

Seznam literatury

1. ČEZ, 2017. *ČEZ, a.s. – o společnosti* [online]. [cit. 2017-09-26], Dostupné z: <https://www.cez.cz/cs/o-spolecnosti/cez/profil-spolecnosti.html>
2. DE OLIVEIRA, F. A., L. E. ZARATE, M. D. REIS a C. N. NOBRE, 2011. The Use of Artificial Neural Networks in the Analysis and Prediction of Stock Prices. *Proceedings of the 2011 IEEE International Conference on Systems, Man and Cybernetics*, Anchorage, AK, s. 2151-2155. ISSN 1062-922X.
3. GHIASSI, M., H. SAIDANE a D. K. ZIMBRA, 2005. A dynamic artificial neural network model for forecasting time series events. *International Journal of Forecasting*. Roč. 21, č. 2, s. 341-362. ISSN 0169-2070.
4. GÖÇKEN, Mustafa, Mehmet ÖZÇALICI, Aslı BORU a Ayşe Tuğba DOSDOĞRU, 2016. Integrating metaheuristics and Artificial Neural Networks for improved stock price prediction. *Expert Systems with Applications*. Roč. 44, s. 320-331. ISSN 0957-4174.
5. HASSAN, Md. Rafiul, Baikunth NATH a Michael KIRLEY, 2007. A fusion model of HMM, ANN and GA for stock market forecasting. *Expert Systems with Applications*. Roč. 33, č. 1, s. 171-180. ISSN 0957-4174.
6. CHANG, Tsung-Sheng, 2011. A comparative study of artificial neural networks, and decision trees for digital game content stocks price prediction. *Expert Systems with Applications*. Roč. 38, č. 12, s. 14846-14851. ISSN 0957-4174.
7. LERTYINGYOD, Nunnapus a Weerachart BENJAMAS, 2016. Stock Price Trend Prediction using Artificial Neural Network Techniques. 2016 20th International Computer Science and Engineering Conference, Chiang Mai, Thailand, ISSN 2381-5035.
8. MABU, Shingo, Yan CHEN, Dongkyu SOHN, Kaoru SHIMADA a Kotaro HIRASAWA, 2009. Stock price prediction using neural networks with RasID-GA. *IEEJ Transactions on Electrical and Electronic Engineering*. Roč. 4, č. 3, s. 392-403. ISSN 1931-4973.

9. QUN, Zhuge, Lingyu XU a Gaowei ZHANG, 2017. LSTM Neural Network with Emotional Analysis for Prediction of Stock Price. *Engineering Letters*. Roč. 25, č. 2, s. 167-175. ISSN 1816-093X.
10. SHEELAPRIYA, G a R MURUGESAN, (2016). Stock price trend prediction using Bayesian regularised radial basis function network model. *Spanish Journal of Finance and Accounting – Revista Española de Financiación y Contabilidad*. Roč. 46, č. 2, s. 189-211. ISSN 0210-2412.
11. SHI, Hongyan a Xiaowei LIU, 2014. Application on stock price prediction of Elman neural networks based on principal component analysis method. *2014 11th International Computer Conference on Wavelet Active Media Technology and Information Processing*, China, s. 411-414. ISBN 978-1-4799-7208-1.
12. TICKNOR, Jonathan L., 2013. A Bayesian regularized artificial neural network for stock market forecasting. *Expert Systems with Applications*. Roč. 40, č. 14, s. 5501-5506. ISSN 0957-4174.
13. VOCHOZKA, Marek, 2012. Ekonomičeskaja dobavlenaja stoimost' i jeje informacionnaja značimost' dlja akcionerov predprijatija [Ekonomická přidaná hodnota a její vypovídající hodnota pro akcionáře podniku]. *Učenyje zapiski Petrozavodskogo gosudarstvennogo universiteta: Obščestvennyje i gumanitarnyje nauki*. Č. 5, s. 84-87. ISSN 1998-5053.
14. VUI, Chang Sim, Gan Kim SOON, Chin Kim ON, Rayner ALFRED a Patricia ANTHONY, 2013. A Review of Stock Market Prediction with Artificial Neural Network (ANN). *2013 IEEE International Conference on Control System, Computing and Engineering*, Penang, Malaysia, s. 477- +. ISBN 978-1-4799-1508-8.
15. YAMASHITA, T., K. HIRASAWA a JL HU, 2005. Multi-branch neural networks and its application to stock price prediction. *Proceedings of the Knowledge-Based Intelligent Information and Engineering Systems*, Melbourne, Australia, s. 1-7. ISSN 0302-9743.
16. ZAHEDI, Javad a Mohammad Mahdi ROUNAGHI, 2015. Application of artificial neural network models and principal component analysis method in predicting stock prices on Tehran Stock Exchange. *Physica A: Statistical Mechanics and its Applications*. Roč. 438, s. 178-187. ISSN 0378-4371.