UNIVERZITA KOMENSKÉHO V BRATISLAVE
FAKULTA MATEMATIKY, FYZIKY A INFORMATIKY



VYUŽITIE NEURÓNOVÝCH SIETÍ PRI MODELOVANÍ AKCIOVÝCH VÝNOSOV

DIPLOMOVÁ PRÁCA

2021 Bc. Ján PITOŇÁK

UNIVERZITA KOMENSKÉHO V BRATISLAVE
FAKULTA MATEMATIKY, FYZIKY A INFORMATIKY

VYUŽITIE NEURÓNOVÝCH SIETÍ PRI MODELOVANÍ AKCIOVÝCH VÝNOSOV

DIPLOMOVÁ PRÁCA

Študijný program:	Ekonomická a finančná matematika
Študijný odbor:	1113 Matematika
Školiace pracovisko:	Katedra aplikovanej matematiky a štatistiky
Vedúci práce:	Mgr. Katarína Lučivjanská, PhD.

Bratislava 2021 **Bc. Ján PITOŇÁK**



65466855

Univerzita Komenského v Bratislave
Fakulta matematiky, fyziky a informatiky

ZADANIE ZÁVEREČNEJ PRÁCE

Meno a priezvisko študenta: Bc. Ján Pitoňák
Študijný program: ekonomicko-finančná matematika a modelovanie (Jednoodborové štúdium, magisterský II. st., denná forma)
Študijný odbor: matematika
Typ záverečnej práce: diplomová
Jazyk záverečnej práce: slovenský
Sekundárny jazyk: anglický

Názov: Využitie neurónových sietí pri modelovaní akciových výnosov
Neural networks for stock return modeling

Anotácia: Schopnosť neurónových sietí zachytiť nelineárne správanie a vzťahy medzi vysvetľovanými a vysvetľujúcimi premennými sa osvedčila vo viacerých oblastiach a financie sú jednou z ďalších prirodzených aplikácií.

Vedúci: Mgr. Katarína Lučivjanská, PhD.
Katedra: FMFI.KAMŠ - Katedra aplikovanej matematiky a štatistiky
Vedúci katedry: prof. RNDr. Marek Fila, DrSc.

Dátum zadania: 17.12.2019

Dátum schválenia: 21.12.2019 prof. RNDr. Daniel Ševčovič, DrSc.
garant študijného programu

.....
študent

.....
vedúci práce

Poďakovanie Touto cestou sa chcem poďakovať mojej vedúcej diplomovej práce Mgr. Kataríne Lučivjanskej, PhD za ochotu, pomoc, odborné rady a podnetné pripomienky, ktoré mi pomohli pri písaní tejto práce. Mgr. Katarína Lučivjanská, PhD bola od prvého momentu ochotná podeliť sa so svojimi a v prípade akýchkoľvek problémov ma vedela nasmerovať správnym smerom. Chcem tiež oceniť jej priateľský a ľudský prístup počas celej spolupráce. Ďakujem aj svojej rodine a priateľom za ich trpezlivosť a podporu.

Abstrakt v štátnom jazyku

PITOŇÁK, Ján: Využitie neurónových sietí pri modelovaní akciových výnosov [Diplomová práca], Univerzita Komenského v Bratislave, Fakulta matematiky, fyziky a informatiky, Katedra aplikovanej matematiky a štatistiky; školiteľ: Mgr. Katarína Lučivjanská, PhD., Bratislava, 2021, 83 s.

Cieľom tejto diplomovej práce je modelovať budúci vývoj finančných trhov, pomocou modelovania ako cien, tak aj výnosov. Na základe týchto predikcií sa následne pokúsiť dosiahnuť čo najväčší zisk. V práci vybudujeme potrebný teoretický základ, potrebný pre pochopenie neurónových sietí a uvedieme prehľad niekoľkých prác publikovaných na túto tému. Siete natrénujeme na vybranej podvzorke akcií z indexu S&P 500. Pre všetky siete tiež otestujeme niekoľko hyperparametrov, na základe ktorých sa budeme snažiť nájsť najlepšiu sieť. Štatisticky popíšeme výkonnosť našich sietí a porovnáme použitie jednotlivých parametrov a aj dosiahnuté výsledky pri predikovaní cien s výsledkami pri predikovaní výnosov.

Kľúčové slová: Neurónové siete, LSTM, Modelovanie výnosov, Predikovanie akciových trhov

Abstract

PITOŇÁK, Ján: Neural networks for stock return modeling [Master Thesis], Comenius University in Bratislava, Faculty of Mathematics, Physics and Informatics, Department of Applied Mathematics and Statistics; Supervisor: Mgr. Katarína Lučivjanská, PhD., Bratislava, 2021, 83 p.

This master thesis aims to model the future development of financial markets, by modeling both prices and revenues. Based on these predictions, we try to maximize our profit on the market. In this thesis, we build the necessary theoretical ground of neural networks, crucial to understand how they work. We also include an overview of several previously published papers with the same topic as ours. We train networks on selected subsample stocks from the S&P 500 index. We test several hyperparameters for all networks based on which we will try to find the best model for each stock. Models' performance will be described using statistical parameters and visualization tools. Comparison of predicting both stock prices and stock returns take part at the end of each subchapter with results.

Keywords: Neural networks, LSTM, Returns modeling, Stock market prediction

Obsah

Úvod **8**

1 Neurónové siete **9**

1.1 Strojové učenie 9

1.2 Umelé neurónové siete 10

1.2.1 Perceptrón 11

1.3 Hlboké dopredné neurónové siete 12

1.4 Stratové a aktivačné funkcie 13

1.5 Trénovanie neurónovej siete 17

1.5.1 RMSProp 17

1.6 Algoritmus spätnej propagácie 18

1.6.1 Derivácia zloženej funkcie a jednoduchá neurónová sieť 18

1.6.2 Hlboká sieť 19

1.7 Rekurentné neurónové siete 21

2 Prehľad publikovaných prác **23**

2.1 Predikovanie cien 23

2.2 Predikovanie výnosov 25

3 Dáta **26**

3.1 Podvzorka pre denné dáta 26

3.2 Podvzorka pre hodinové dáta 28

3.3 Technické indikátory 29

4 Architektúra sietí **32**

4.1 Voľba predikovaného výstupu 33

4.2 Výber najlepšej architektúry 33

4.3 Výber počtu dát z minulosti 34

4.4 Výber technických indikátorov 35

4.5 Vyhodnocovanie modelov 35

4.5.1 Ziskovosť modelu 35

4.5.2 Sharpov pomer 36

5 Výsledky na denných dátach **37**

5.1 Predikcia cien 37

5.1.1 Výsledky na validačnej sade 37

5.1.2 Výsledky na testovacej sade 40

5.2 Predikcia výnosov 48

5.2.1 Výsledky na validačnej sade 48

5.2.2 Výsledky na testovacej sade 52

6 Výsledky na hodinových dátach **61**

6.1 Predikcia cien 61

6.1.1 Výsledky na validačnej sade 61

6.1.2 Výsledky na testovacej sade 63

6.2 Predikcia výnosov 69

6.2.1 Výsledky na validačnej sade 69

6.2.2 Výsledky na testovacej sade 72

Záver **77**

Literatúra **79**

Príloha **83**

Úvod

Akciové trhy historicky predstavujú aktívum s najväčšou mierou výnosnosti [6]. Sú preto lákadlom nielen pre veľkých investorov, aj pre malých sporiteľov, ktorí hľadajú zhodnotenie svojich úspor. Všetci účastníci trhu majú spoločný cieľ odhadnúť, kam sa bude trh v nejakom časovom horizonte pohybovať. Azda najjednoduchšou úlohou je predikovanie celého trhu, respektíve akciového indexu, na dlhé obdobie. V takomto prípade ak povieme, že trh bude rásť, tak s najväčšou pravdepodobnosťou budeme mať pravdu. Čím viac sa ale budeme zameriavať na menej agregované aktíva alebo na kratší časový horizont, tým sa táto úloha stáva náročnejšou. V našej práci sa budeme snažiť o predpovedanie vývoja akcií jednotlivých spoločností na nasledujúcu hodinu a deň. Hoci finančná literatúra predpokladá nezávislosť takýchto výnosov [23], my sa pokúsime pomocou neuorónvých sietí nájsť nejakú nelineárnu závislosť od predchádzajúcich dát.

V prvej kapitole si zadefinujeme pojem strojového učenia a uvedieme teoretický základ potrebný k pochopeniu neurónových sietí a ich fungovaniu. Tiež si zavedieme pojem učenia sa a objasníme, ako sa učia neurónové siete. Uvedieme aj základný optimalizačný algoritmus, ktorý sa využíva pri učení týchto sietí.

V ďalšej kapitole uvedieme prehľad prác, ktoré boli na túto tému publikované. Zhrnieme výsledky, ktoré títo autori dosiahli a pozrieme sa na ich prístup.

Objasnenie, s akými dátami budeme pracovať, bude súčasťou tretej kapitoly. Taktiež vyberieme podvzorku dát z celého akciového indexu S&P 500 tak, aby vierohodne reprezentovala celý index. V tejto kapitole si zadefinujeme aj stratégiu vstupovania do pozícií, pomocou ktorej budeme vyhodnocovať jednotlivé predikcie.

Štvrtá kapitola bude venovaná konkrétnej architektúre našich sietí a rôznym parametrom, ktoré jej budeme nastavovať.

Kapiola s číslom 5 bude venovaná výsledkom, ktoré sme dosiahli pre denné dáta. Posledná v poradí, ale určite nie, čo do významu, bude kapitola venovaná výsledkom pre hodinové dáta. V oboch týchto kapitolách zanalyzujeme výsledky jednotlivých sietí a pokúsime sa nájsť tie, ktoré fungujú najlepšie. Porovnáme aj výsledky, dosiahnuté pre predikovanie cien, s výsledkami, ktoré dosiahneme pri predikovaní výnosov.

1 Neurónové siete

V prvej kapitole sa oboznámime s konceptom neurónových sietí. Uvedieme základné druhy sietí a uvedieme si príklady využitia jednotlivých druhov. Bližšie sa pozrieme na rekurentné siete, ktoré sú základom tejto práce. Počas tejto kapitoly budeme vychádzať prevažne z [12], [17].

1.1 Strojové učenie

V praxi sa stretávame s problémami, pre ktoré je veľmi obtiažne až nemožné explicitne naprogramovať program, ktorý by bol schopný reagovať na nové podnety. Základnou myšlienkou strojového učenia, ako sa vysporiadať s týmto problémom, je zmena prístupu k riešeniu problémov. Pri normálnom programovaní sa algoritmu dodajú dáta a pravidlá, na základe ktorých algoritmus vyhodnotí vstup. Myšlienkou strojového učenia ale je, že algoritmu dodáme iba vstupné dáta a očakávané odpovede a necháme algoritmus určiť si vlastné pravidlá. Keďže v tomto prípade explicitne neprogramujeme daný algoritmus, ale nechávame stroj takpovediac učiť sa, hovoríme o strojovom učení. Takéto učenie bolo formálne zadefinované v [24]:

Definícia 1.1. *Hovoríme, že program sa učí zo skúsenosti* **E**, *vzhľadom na triedu úloh* **T** *a na mieru chybovosti* **P**, *ak sa jeho výkon, v úlohách* **T** *meraný v* **P**, *s pribúdajúcimi skúsenosťami* **E** *zlepšuje.*

Strojové učenie tak našlo svoje uplatnenie napríklad pri rozpoznávaní, aký objekt sa nachádza na obrázku, autonómnych vozidlách alebo pri rôznych predikciách, ako sú ceny bytov alebo spotreby energií.

Strojové učenie môžeme rozdeliť na tri základné kategórie:

- Učenie s učiteľom (Supervised learning): v tomto prípade algoritmus spracúva dvojice vstup-očakávaný výstup. Známym problémom je napríklad rozoznávanie rukou písaných číslic, kde na vstupe dostávame maticu pixelov a označenie, zodpovedajúceho čísla, ktoré predstavuje. Okrem takejto klasifikácie sa učenie s učiteľom využíva aj na regresné úlohy. Najznámejším príkladom je azda predikcia ceny bytov s ohľadom ich rozlohu. Práve riešenie predikčných úloh bude základom tejto práce.

- Učenie bez učiteľa (Unsupervised learning): o takomto spôsobe hovoríme, ak nemáme dáta označené očakávaným výstupom, namiesto toho, sa algoritmus snaží sám roztriediť dáta podľa nejakých spoločných vzorov. Príkladom, kde sa možno stretnúť s učením bez učiteľa, je rozdelenie zákazníkov do skupín, podľa predchádzajúcich transakcií.

- Učenie pomocou posilovania (Reinforcement learning): jedná sa o špeciálny druh strojového učenia, v ktorom učenie prebieha pomocou agenta, ktorý je daný do prostredia, kde vykonáva nejakú úlohu, za ktorú je následne odmeňovaný. Do povedomia odbornej obce sa zapísalo, keď takýto algoritmus porazil šachových veľmajstrov. V takomto prípade je algoritmus odmeňovaný za vyhratie partie alebo vyhodenie súperovej figúrky, a naopak, trestaný, ak mu súper figúrku vyhodí. Cieľom algoritmu je tak maximalizovať odmenu za hru. Takto definovaný algoritmus potom hrá milóny hier, sám proti sebe, počas ktorých sa učí, vďaka čomu sa stáva prakticky neporaziteľným [29].

1.2 Umelé neurónové siete

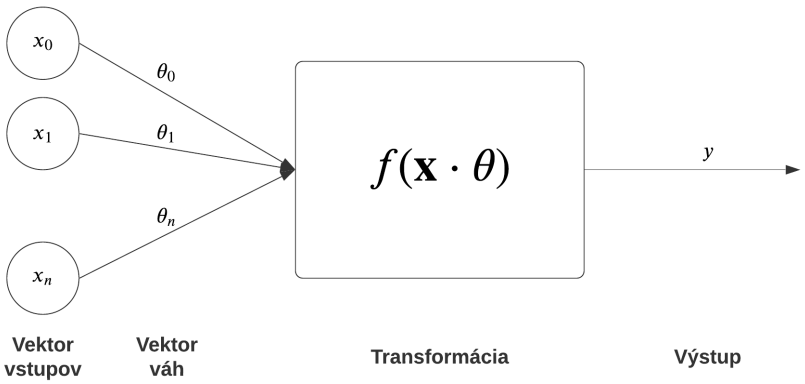
Keďže najlepším nástrojom na spracovanie a vyhodnocovanie informácií je bezpochyby ľudský mozog, uvedieme stručný opis jeho fungovania, pričom budeme vychádzať z [25]. Mozog je schopný spracovať a vyhodnotiť obrovské množstvo informácií v zlomkoch sekundy. Ideou za vznikom umelých neurónových sietí je preto snaha napodobniť činnosť ľudského mozgu. Ten je tvorený obrovským množstvom navzájom poprepájaných neurónov, v ktorých prebiehajú zložité procesy, vďaka ktorým sa môžu informácie spracovávať a posúvať ďalej. V odbornej literatúre sa uvádza, že v mozgu sa nachádza približne 10^{11} neurónov [22], vďaka ktorým sme schopní vnímať svet a rozmýšľať.

Dva neuróny medzi sebou môžu byť prepojené pomocou synapsií, ktoré prenášajú signál z jedného neurónu do druhého. V skutočnosti sa má za to, že učenie prebieha práve na týchto synapsiách, a to spôsobom, že signál na synapsii sa môže buď zoslabiť alebo zosilniť. Sila, akou synapsia vplýva na signál, sa nazýva váha synapsie. Učenie je teda zmena váh synapsií. Podobný prístup sa snažia implementovať aj umelé neurónové siete.

1.2.1 Perceptrón

Pojem perceptrón zaviedol americký psychológ Frank Rosenblatt, ktorý sa vo svojej praxi venoval kognitívnym vedám a neurobiológii. Rosenblatt vo svojej práci nadviazal na prácu McCullocha a Pittsa, ktorí predstavili prvý model umelej neurónovej siete, ktorá klasifikovala pravdivostnú hodnotu logických operátorov ('a', 'alebo' a 'negácia'). Rosenblatt v článku [26] ukázal, že takúto sieť možno natrénovať na klasifikáciu dát. Perceptrón tak predstavuje model, ktorý pozostáva z neurónu, ktorý spracovával vstupné signály. Spôsob, akým menil váhy týchto signálov (synapsií), bol nasledovný: zaznamenali sa všetky vstupy a odpovede perceptrónu na ne. V prípade správnej odpovede, váhy nemeníme. V prípade nesprávnej odpovede, upravíme všetky vstupné váhy. V prípade očakávaného pozitívneho výsledku a negatívnom výstupe, všetky vstupné váhy zvýšime, v opačnom prípade nesprávnej klasifikácie, váhy znížime. Hoci je tento algoritmus veľmi heuristický a prácny, je základom viacerých stále používaných algoritmov učenia s učiteľom.

Formálne perceptrón spracúva vektor vstupov $\mathbf{x} = (x_0, x_1, ..., x_n)^T$, kde $x_0 = 1$ (intercept) a n označuje počet dát, pomocou Hadamardovho súčinu s vektorom váh $\theta = (\theta_0, \theta_1, ..., \theta_n)^T$. Na tento súčin je následne aplikovaná transformačná funkcia f, ktorú budeme nazývať aktivačná (v prípade regresných úloh takáto transformácia nemusí byť aplikovaná). Výstup perceptrónu bude vyzerať nasledovne $y = f(\mathbf{x}, \theta) = f\left(\sum_{i=0}^{n} x_i \cdot \theta_i\right)$.



Obr. 1: Model perceptrónu, do ktorého vstupuje vektor vstupov $\mathbf{x}$.

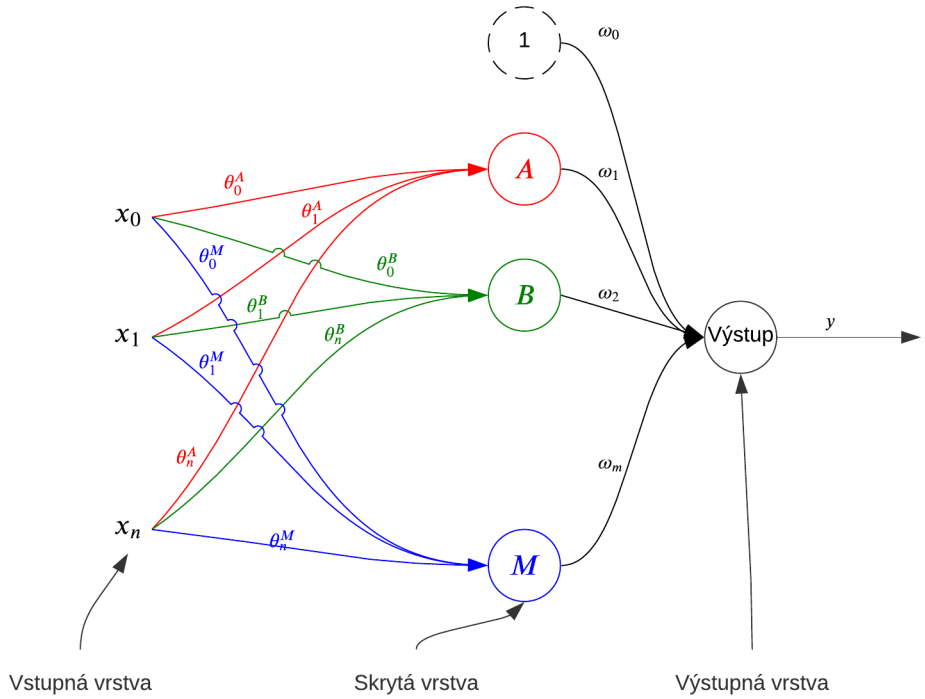
Jeden perceptrón sa líši od lineárnej alebo logistickej regresie iba v aktivačnej funkcii f. Preto sa v praxi spájajú viaceré perceptróny do vrstiev a jednotlivé vrstvy sa ešte skladajú na seba, čím sa nadobúda schopnosť učiť sa oveľa komplikovanejšie vzťahy.

Takéto perceptróny následne nazývame neuróny.

1.3 Hlboké dopredné neurónové siete

Dopredné neurónové siete sú základom hlbokého učena. Takéto siete majú uplatenenie hlavne v klasifikačných úlohách, ale ich obrovský prínos tkvie v tom, že predstavujú základ pre konvolučné siete, ktoré sú nezastupiteľné pri rozoznávaní predmetov na obrázkoch. Rovnako sú dôležité aj pre rekurentné neurónové siete, ktorým sa budeme venovať v podkapitole 1.7.

Princíp hlbokých dopredných sietí sme už naznačili aj na konci podkapitoly 1.2.1. Hlboké dopredné siete spájajú viacero vrstiev, z ktorých každá obsahuje aspoň jeden neurón, pričom informácie sa šíria iba smerom k nasledujúcim vrstvám a nie spätne alebo v rámci vrstvy. Takéto siete sa skladajú z prvej, vstupnej vrstvy (input layer), ktorá spracuje vstupné dáta, ku ktorým je pridaný intercept. Ďalšie vrstvy, okrem poslednej, sa nazývajú skryté (hidden layers). Poslednou vrstvou je výstupná (output layer), ktorá už poskytuje finálnu predikciu.



Obr. 2: Model siete s troma vrstvami.

Na takúto sieť sa môžeme pozerať ako na funkciu $F : \mathbb{R}^{n+1} \to \mathbb{R}$, kde F je zložená funkcia

$$F_\Theta(\mathbf{x}) = f^j_{\theta_j}(...f^3_{\theta_3}(f^2_{\theta_2}(f^1_{\theta_1}(\mathbf{x})))). \tag{1}$$

Napríklad na sieti z obrázku 2 by sme mali maticu váh prvej vrstvy:

$$\theta = \begin{pmatrix} \theta_0^A & \theta_1^A & ... & \theta_n^A \\ \theta_0^B & \theta_1^B & ... & \theta_n^B \\ ... & & & \\ \theta_0^M & \theta_1^M & ... & \theta_n^M \end{pmatrix}, \tag{2}$$

a vektor váh pre druhú, výstupnú vrstvu:

$$\omega = (\omega_0, \omega_1, \omega_2, ..., \omega_m)^T, \tag{3}$$

kde m označuje počet neurónov v prechádzajúcej vrstve. Potom zloženú funkciu, reprezentujúcu našu sieť, môžeme zapísať jednoducho, ako:

$$F(\mathbf{x}, \theta, \omega) = f^2(\omega^T f^1(\theta \mathbf{x})), \tag{4}$$

kde f^1 a f^2 sú aktivačné funkcie, o ktoré sú zadefinované v podkapitole 1.4. Otázkou ostáva, ako sieť natrénovať, teda nájsť správne θ a ω. Jednou z možností je analyticky spočítať gradient nejakej stratovej funkcie (bližšie o stratových funkciách v 1.4). Avšak výpočet gradientu pri neurónových sieťach spravidla nie je taký jednoduchý, ako napríklad pri lineárnej regresii. V praxi sa na jeho výpočet využíva algoritmus spätného šírenia chyby (back-propagation, viď podkapitola (1.6)). Zásadným problémom je aj to, že nemáme garantovanú konvexnosť úlohy, z tohto dôvodu sa na učenie viac hodí iteračný algoritmus než analytické riešenie.

1.4 Stratové a aktivačné funkcie

Pod pojmom stratová funkcia rozumieme takú funkciu, ktorá nejakým spôsobom meria, ako veľmi sa predikcia odchyľuje od skutočnosti. V tejto podkapitole si zadefinujeme niektoré základné typy stratových funkcií, ktoré budeme používať nielen pri klasifikačných úlohach, ale aj pri regresných.

V prípade klasifikačnej úlohy sa vlastne pýtame, s akou pravdepodobnosťou pri daných $\mathbf{x}$, bude naša predikovaná hodnota rovná jednej. V prípade binárnej klasifikácie, ak je

naša očakávaná hodnota $y = 1$, prirodzene chceme, aby predikovaná hodnota, $F(x_i) = p_i$, bola čo najbližšie k 1. Zaujíma nás, s akou pravdepodobnosťou, je každý vstup správne predikovaný. V prípade, že $y_i = 1$, tak potom pravdepodobnosť správnej predikcie je priamo p_i. Na druhú stranu, ak $y_i = 0$, pravdepodobnosť správnej predikcie získame ako $1 - p_i$. Pravdepodobnosť správnej predikcie môžeme zapísať ako $p_i^{y_i}(1 - p_i)^{1-y_i}$, je zrejmé, že ak $y_i = 1$, tak sa do úvahy berie iba p_i. Naopak, ak $y_i = 0$, tak aktívna časť bude len $(1 - p_i)$. Keďže pravdepodobnosť správnej predikcie pre vstup x_i je nezávislá od predikcie pre vstup x_{i-1}, preto pravdepodobnosť správnej predikcie pre všetky vstupy môžeme jednoducho zapísať ako súčin jednotlivých pravdepodobností:

$$\prod_{\forall i} p_i^{y_i}(1 - p_i)^{1-y_i}. \tag{5}$$

Takto naformulovanú úlohu, môžeme považovať za funkciu vierohodnosti. Kvôli jednoduchosti ju upravíme pomocou zápornej logaritmickej transformácie. Takto upravená funkcia bude stratovou funkciou v prípade klasifikačných úloh:

$$\begin{aligned} -\log\left(\prod_{\forall i} p_i^{y_i}(1 - p_i)^{1-y_i}\right) &= -\sum_{\forall i} \log\left(p_i^{y_i}(1 - p_i)^{1-y_i}\right) \\ &= -\sum_{\forall i} (y_i \log(p_i) + (1 - y_i)\log(1 - p_i)) \end{aligned} \tag{6}$$

Funkcia (6) sa nazýva aj binárna krížová entropia, ktorá bude stratovou funkciou v klasifikačných úlohách:

$$\min_{\Theta} J(\mathbf{x}, \Theta) = \min_{\Theta} -\sum_{\forall i} (y_i \log(F_\Theta(x_i)) + (1 - y_i)\log(1 - F_\Theta(x_i))) \tag{7}$$

Nasledujúce funkcie sa používajú prevažne v regresných úlohach a sú všeobecne známe, preto ich nebudeme odvádzať len, ich uvedieme. Prvou je stredná absolútna chyba (mean absolute error, MAE):

$$\min_{\Theta} J(\mathbf{x}, \Theta) = \min_{\Theta} \frac{1}{n} ||y_i - F_\Theta(x_i)||_1, \tag{8}$$

a druhou stredná kvadratická chyba (mean squared error, MSE):

$$\min_{\Theta} J(\mathbf{x}, \Theta) = \min_{\Theta} \frac{1}{n} ||y_i - F_\Theta(x_i)||^2. \tag{9}$$

MAE predstavuje priemerný rozdiel predikcie, od skutočnej hodnoty. Podobne sa dá nahliadať aj na MSE, avšak MSE oveľa viac penalizuje odľahlé hodnoty. Pri definovaní stratovej funkcie, ostáva ešte ozrejmiť, aké funkcie môžeme voliť za aktivačnú funkciu $f^i_{\theta_i}$ v zloženej funkcii (1). V skutočnosti to môže byť akákoľvek C^1 funkcia, avšak vo väčšine prípadov máme na tieto funkcie nejaké prirodzené obmedzenia. V praxi sa taktiež používa niekoľko základných funkcií, ktoré dobre fungujú na širokú škálu úloh.

Najjednoduchšou funkciou je afinné posunutie:

$$f^i = \theta^T \mathbf{x}. \tag{10}$$

Je zrejme, že funkcia 10 sa používa najma pri regresných úlohách, kde očakávaný výstup je ľubovolné reálne číslo (respektíve vektor). V prípade ale, že všetky aktivačné funkcie siete sú typu 10 a stratová funkcia je 8 alebo 9, tak sa úloha redukuje na riešenie lineárnej regresie, keďže máme $F_\Theta(\mathbf{x}) = \omega(\theta^T\mathbf{x}) = (\omega\theta^T)\mathbf{x} = \beta^T\mathbf{x}$.

Ďalšou známou aktivačnou funkciou je sigmoida, pri odvodení ktorej budeme postupovať podobne ako v [9]. Sigmoida sa používa v situáciách, kde očakávaný výstup y nadobúda iba hodnoty 0 alebo 1, teda y je z alternatívneho rozdelenia, čo je v podstate binomické rozdelenie pri jednom pokuse. Platí $y \sim Bin(1, p_x)$, kde $p_x = p(y = 1|x)$. Vieme potom, že $E(y) = p_x$. Zadefinujme si teraz funkciu logit:

$$logit(p) = \ln\left(\frac{p}{1-p}\right). \tag{11}$$

Logit je teda logaritmus šance. Táto funkcia mapuje pravdepodobnosť z $(0, 1) \to (-\infty, \infty)$. Podľa [9] sa $logit(E(y))$ dá namodelovať pomocou lineárneho modelu:

$$\begin{aligned} logit(p_x) &= \theta_0 + \theta_1 x_1 + ... + \theta_n x_n \\ \ln\left(\frac{p_x}{1-p_x}\right) &= \theta_0 + \theta_1 x_1 + ... + \theta_n x_n \\ \frac{p_x}{1-p_x} &= e^{\theta_0+\theta_1 x_1+...+\theta_n x_n} \\ p_x &= \frac{e^{\theta_0+\theta_1 x_1+...+\theta_n x_n}}{1+e^{\theta_0+\theta_1 x_1+...+\theta_n x_n}}, \end{aligned} \tag{12}$$

pričom funkciu 12 nazývame sigmoid. Častejšie sa sigmoid používa v mierne upravenej podobe, a to rozšírením funkcie 12 pomocou $\frac{\frac{1}{e^{\theta_0+\theta_1 x_1+...+\theta_n x_n}}}{\frac{1}{e^{\theta_0+\theta_1 x_1+...+\theta_n x_n}}}$, z čoho dostávame:

Definícia 1.2. *Majme* $z \in \mathbb{R}$, *kde* $z = \theta_0 + \theta_1 x_1 + ... + \theta_n x_n$. *Potom funkcu* $\sigma(z) : \mathbb{R} \to (0, 1)$:

$$\sigma(z) = \frac{1}{1 + e^{-z}}, \tag{13}$$

nazývame sigmoid.

vďaka čomu sme schopní modelovať očakávanú strednú hodnotu pre y. Ako vidno na obrázku 3, sigmoid na väčšine definičného oboru dosahuje hodnoty blízke 0 alebo 1, rastie iba v okolí $z = 0$. To môže spôsobovať problémy pri učení, kedže gradient je okrem okolia 0 takmer všade nulový.

Aktivačnou funkciou, ktorá bude dôležitá pri rekurentných sieťach v kapitole (bude doplnene), je funkcia hyperbolický tangens.

Definícia 1.3. *Majme* $z \in \mathbb{R}$, *kde* $z = \theta_0 + \theta_1 x_1 + ... + \theta_n x_n$. *Potom funkciu* $\tanh(z) : \mathbb{R} \to (-1, 1)$:

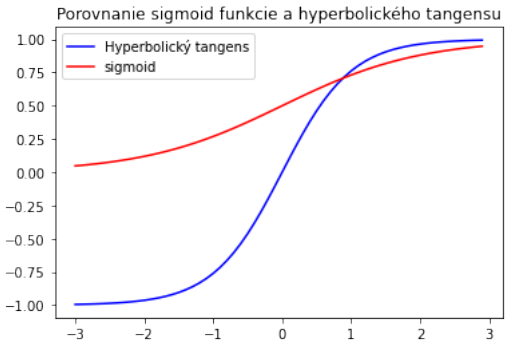
$$\tanh(z) = \frac{e^z - e^{-z}}{e^z + e^{-z}}, \tag{14}$$

nazývame hyperbolický tangens.

Hyperbolický tangens je v princípe transformovaná sigmoida, kde platí nasledujúci vzťah:

Lema 1.4. *Pre* $\forall z \in \mathbb{R}$ *platí:* $\tanh(z) = 2\sigma(2z) - 1$.

Platnosť lemy 1.4 možno jednoducho ukázať, ak si zoberieme sigmoidu zo vzťahu 12 a ten upravíme. Na obrázku 3 vidíme podobnosť týchto funkcií aj graficky.



Obr. 3: Znázornenie funkcií sigmoid a hyperbolického tangensu.

V súčasnosti je asi najpoužívanejšou aktivačnou funkcoiou ReLu (Rectified linear unit). Ide o spojitú, po častiach lineárnu funkciu.

Definícia 1.5. *Majme $x \in \mathbb{R}$. Potom funkciou ReLu nazývame funkciu:*

$$f(x) = \max(0, x). \tag{15}$$

V prípade, že už máme zadefinovanú nákladovú funkciu aj aktivačné funkcie, tak môžeme začať našu sieť trénovať. Ako sme spomenuli, analytické riešenie je aj vzhľadom na možný počet nastavení siete nepraktické, na výpočet gradientu sa používa metóda, ktorú si objasníme v podkapitole (1.6).

1.5 Trénovanie neurónovej siete

Po zadefinovaní stratovej funkcie $J(x, \Theta)$ potrebujeme nájsť optimálne váhy Θ. V strojovom učení sa na minimalizáciu stratovej funkcie používajú algoritmy prvého rádu, ktoré sú menej výpočtovo náročné oproti metódam druhého rádu, ale ich konvergencia je pomalšia. Navyše nemáme žiadne ohraničenia, tak sa problém redukuje na úlohu voľnej optimalizácie. Základnou metódou, je gradientná metóda s predpisom:

$$x^{k+1} = x^k - \lambda \nabla f(x), \tag{16}$$

kde λ je nejaká dostatočne malá konštanta a symbol $\nabla f(x)$, označuje gradient funkcie. Konvergencia tejto metódy je garantovaná iba pre silnokonvexné funkcie [11, str. 466] alebo pre ohraničené funkcie s Lipschitzovským gradientom [18, str. 91]. Hoci v prípade neurónových sietí nie je garantovaná ani jedna z týchto podmienok, v prípade dostatočne hlbokých neurónových sietí je zvyčajne postačujúce nájsť nejaké lokálne minimum [13]. Štandardom pri trénovaní neurónových sietí je použitie modifikovanej gradientnej metódy. My si popíšeme fungovanie algoritmu RMSProp. Tento algoritmus je vďaka svojej všestrannosti odporúčaný pre širokú škálu problémov v rámci hlbokého učenia [12], preto sme sa ho rozhodli použiť v našej práci.

1.5.1 RMSProp

Tento algoritmus [31] je modifikáciou algoritmu AdaGrad [15], ktorý upravuje parameter λ v predpise (16) tým, že uchováva informáciu o predchádzajúcich gradientoch. RMSProp, na rozdiel od AdaGradu využíva exponenciálny priemer predchádzajúcich gradientov, ako môžeme vidieť v algoritme 1. Tým sa zbavuje informácie o extrémnych minulých

hodnotách, ktoré nie sú relevantné, a v prípade nájdenia konvexného kužeľa je schopný veľmi rýchlo konvergovať.

Input: konštanta ϵ, konštanta poklesu ρ

Input: konštanta δ, na zabezpečenie prípustnosti delenia

while $x^{k+1} \neq x^k$ **do**

$r_j := \rho r_j + (1-\rho)\nabla_j f(x)^2$

$x_j^{k+1} := x_j^k - \frac{\epsilon}{\sqrt{\delta + r_j}} \nabla_j f(x).$

Algoritmus 1: Algoritmus RMSProp

RMSPRop sa ukázal ako vhodná voľba v prípade neurónových sietí a stal sa základným optimalizačným algoritmom, ktorý budeme využívať aj v našej práci.

1.6 Algoritmus spätnej propagácie

V kapitole (1.5) sme ukázali spôsob, akým budeme učiť neurónovú sieť. Na to však potrebujeme poznať gradient stratovej funkcie, ktorý nemusí byť jednoduché vypočítať analyticky. V trénovaní neurónových sietí sa na to používa práve algoritmus spätnej propagácie. Myšlienku si ukážeme na sieti z obrázku 2. Doprednou propagáciou sa označuje výstup siete, pre nejaký vstup $\mathbf{x} = (x_0, x_1, \ldots, x_n)^T$. Výstup siete spočítame pomocou výrazu (4), povezdme, nech je rovný 0,4. A nech skutočný výstup y je v tomto prípade $y = 1$. Keďže naším cieľom je, aby rozdiel medzi predikovaným a skutočným výstupom bol čo najmenší, tak si kladieme otázku, či 0.4 je dostatočne dobrá predikcia. Prirodzene, sa nám predikovaná hodnota zdá nízka a preto ju budeme chcieť zväčšiť. To môžeme docieliť úpravou váh ω a Θ. Budeme chcieť vedieť, ako máme dané váhy upraviť, aby sme dostali čo najlepšiu klasifikáciu pre všetky dáta. Táto podkapitola vychádza z [17] a [25], kde možno nájsť podrobnejšie odvodenie tejto metódy.

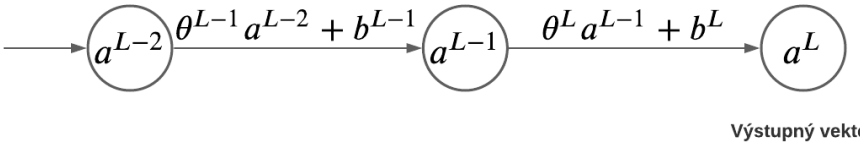
1.6.1 Derivácia zloženej funkcie a jednoduchá neurónová sieť

Uvažujme sieť zloženú z jedného neurónu v každej vrstve, ako je znázornené na obrázku 4. Výstup L-tej vrstvy dostaneme ako:

$$a^L = f^L(\theta^L a^{L-1} + b^L), \tag{17}$$

kde f^L predstavuje aktivačnú funkciu a b^L intercept. Chceme, aby výraz (17) správne predikoval príslušnú hodnotu y pre daný vstup. Pre zjednodušenie budeme ďalej označovať $\theta^L a^{L-1} + b^L$ ako z^L. Na predchádzajúce vrstvy sa tiež môžeme pozerať ako na transformáciu výstupov predchádzajúcich vrstiev. Môžeme tak napísať:

$$a^L = f^L(\theta^L(f^{L-1}(\theta^{L-1}((f^{L-2}(\theta^{L-2}(\dots a^0) + b^{L-2}))) + b^{L-1})) + b^L) \tag{18}$$



Obr. 4: Znázornenie jednoduchej siete, kde horný index označuje poradie vrstvy a b označuje intercept.

Ako sme už uviedli, bude nás zaujímať vplyv jednotlivých váh, na presnosť predikcie. Tú vyjadríme pomocou nejakej stratovej funkcie $J(x, \Theta)$. Chceme vyjadriť závislosť zmeny $J(x, \Theta)$ od napríklad θ^L, inými slovami, chceme parciálnu deriváciu $\frac{\partial J}{\partial \theta^L}$, na to potrebujeme derivovať zloženú funkciu:

$$\frac{\partial J}{\partial \theta^L} = \frac{\partial J}{\partial a^L}\frac{\partial a^L}{\partial z^L}\frac{\partial z^L}{\partial \theta^L}. \tag{19}$$

Vo výraze (19) časť $\frac{\partial z^L}{\partial \theta^L}$ bude rovná a^{L-1}, to znamená, že vplyv váhy θ^L bude priamo úmerný veľkosti predchádzajúceho neurónu. Ak by sme chceli vyjadriť vplyv predchádzajúceho neurónu, tak by sme dostali obdobný výraz, len s vymenenými a^{L-1} a θ^L. Hoci hodnotu θ^L chceme zistiť, táto derivácia nám hovorí, ako budú jednotlivé vrstvy vplývať na konečnú hodnotu. V prípade vyjadrenia vplyvu interceptu b dostávame výraz podobný (19):

$$\frac{\partial J}{\partial b^L} = \frac{\partial J}{\partial a^L}\frac{\partial a^L}{\partial z^L}\frac{\partial z^L}{\partial b^L}, \tag{20}$$

v tomto prípade ale platí, že $\frac{\partial z^L}{\partial b^L} = 1$. Analogicky vyjadríme aj vplyv ostatných parametrov.

1.6.2 Hlboká sieť

V prípade jednoduchej siete je výpočet pomerne priamočiary, hlavný rozdiel v porovnaní s hlbokou sieťou je, že neuróny nie sú ovplyvnené len jednou váhou, ale viacerými, ako

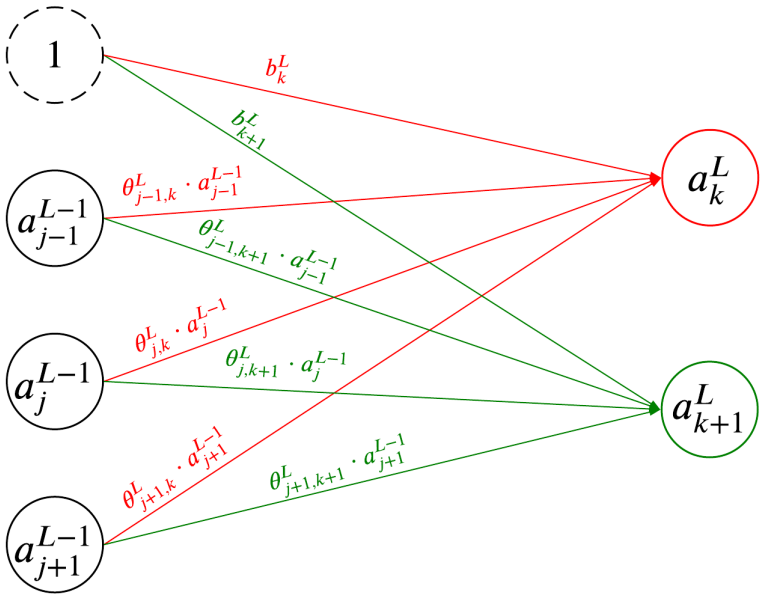
je znázornené na obrázku 5. Výstup pre L-tú vrstvu tak spočítame nasledovne: najprv vypočítame hodnoty pre z_k^L a z_{k+1}^L podobne, ako v jednoduchej sieti:

$$\begin{pmatrix} z_k^L \\ z_{k+1}^L \end{pmatrix} = \begin{pmatrix} \theta_{j-1,k}^L \cdot a_{j-1}^{L-1} + \theta_{j,k}^L \cdot a_j^{L-1} + \theta_{j+1,k}^L \cdot a_{j+1}^{L-1} + b_k^L \\ \theta_{j-1,k+1}^L \cdot a_{j-1}^{L-1} + \theta_{j,k+1}^L \cdot a_j^{L-1} + \theta_{j+1,k+1}^L \cdot a_{j+1}^{L-1} + b_{k+1}^L \end{pmatrix}, \tag{21}$$

následne na každý prvok vektora z (21) aplikujeme príslušnú aktivačnú funkciu, opäť obdobne ako pri jednoduchej sieti:

$$\begin{pmatrix} a_k^L \\ a_{k+1}^L \end{pmatrix} = \begin{pmatrix} f_k^L(z_k^L) \\ f_{k+1}^L(z_{k+1}^L) \end{pmatrix}. \tag{22}$$

Na presnosť predikcie nám vplývajú oba neuróny z L-tej vrstvy.



Obr. 5: Znázornenie váh pre sieť s viacerými vrstvami.

Ak sa pozrieme na vplyv váh na jednotlivé neuróny:

$$\frac{\partial J}{\partial \theta_{\circ,\bullet}^L} = \frac{\partial J}{\partial a_\bullet^L} \frac{\partial a_\bullet^L}{\partial z_\bullet^L} \frac{\partial z_\bullet^L}{\partial \theta_{\circ,\bullet}^L}, \tag{23}$$

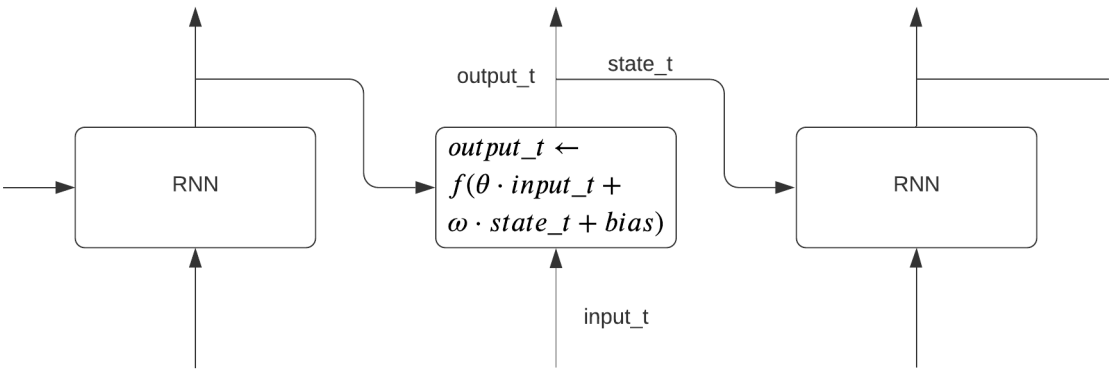
vidíme, že vplyv jednotlivých váh sa zásadne nezmenil. Zmena však nastane, ak budeme chcieť vyjadriť vplyv prechádzajúcich neurónov na stratovú funkciu $J(x, \Theta)$. Každý z neurónov z predchádzajúcej vrstvy vstupuje do viacerých neurónov z vrstvy L, a teda deriváciu nákladovej funkcie J, podľa a_j^{L-1}, vyjadríme ako:

$$\frac{\partial J}{\partial a_j^{L-1}} = \frac{\partial J}{\partial a_k^L} \frac{\partial a_k^L}{\partial z_k^L} \frac{\partial z_k^L}{\partial a_j^{L-1}} + \frac{\partial J}{\partial a_{k+1}^L} \frac{\partial a_{k+1}^L}{\partial z_{k+1}^L} \frac{\partial z_{k+1}^L}{\partial a_j^{L-1}}. \tag{24}$$

V tomto prípade využijeme pravidlo o derivácii súčtu, a sčítame jednotlivé derivácie vo vrstve L. Tento postup opakujeme spätne až k prvým vrstvám a získame tak všetky požadované gradienty. Výhodou v tomto prístupe je, že pri postupe k nižším vrstvám budeme väčšinu zložiek zloženej derivácie poznať.

1.7 Rekurentné neurónové siete

Ako sme spomenuli v úvode, v práci budeme používať rekurentné neurónové siete (RNN) a v tejto podkapitole budeme vychádzať z [12], [17] a [25]. Dôvod, prečo sú RNN vhodné na spracovanie sekvencií je, že na rozdiel od sietí z podkapitoly 1.3 disponujú určitým druhom pamäte. Rekurentné siete postupne prechádzajú a spracúvajú vstupné dáta a zároveň si pamätajú to, čo už spracovali. Inými slovami, jeden vstup nie je spracovaný v jedinom kroku, ale v akomsi internom cykle, čo je znázornené na obrázku 6. Zjednodušený algoritmus fungovania je stručne popísaný pomocou algoritmu 2.



Obr. 6: Znázornenie princípu RNN.

Problémom takýchto jednoduchých sietí je fenomén miznúceho gradientu, tento problém je popísaný v [10]. Riešenie sa našlo navrhnutím LSTM (Long short-term memory) a GRU (Gated recurrent unit) sietí. LSTM ukladá a odovzdáva informácie nielen z predchádzajúceho stavu, ale aj zo stavov predtým. Táto vlastnosť zabezpečí, že nedochádza k strate informácie, toto udržanie informácie sa deje pomocou pridania premennej c_t, ktorá bude informáciu uschovávať tak, ako je znázornené na obrázku 7. Podobne funguje aj GRU

Input: $X \in \mathbb{R}^{t \times n}$ `// Matica vstupov`

$no_outputs \leftarrow m$ `// Veľkosť výstupu`

Data: $no_steps \leftarrow t$, $no_inputs \leftarrow n$,

$state_t \leftarrow \mathbf{0} \in \mathbb{R}^{no_outputs}$,

$\theta = random_matrix(no_outputs, no_inputs)$,

$\omega = random_matrix(no_outputs, no_outputs)$,

$bias = random_matrix(no_outputs, 1)$

for $t \leftarrow 1$ **to** n **do**

  $input_t \leftarrow X[t, :]$,

  $output_t \leftarrow f(\theta \cdot input_t + \omega \cdot state_t + bias)$, `// kde` f `je aktivačná funkcia`

  $state_t \leftarrow output_t$

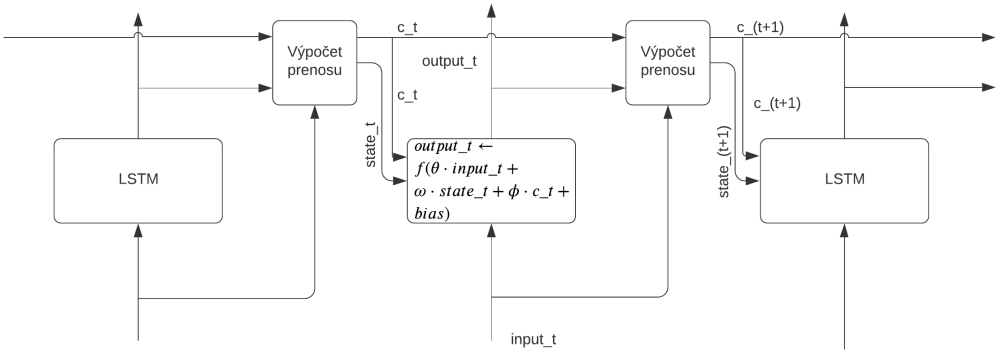
return $output_t$

Algoritmus 2: Algoritmus fungovania RNN. V tomto prípade algoritmus vracia iba poslednú hodnotu výstupu, ktorá obsahuje informáciu aj o predchádzajúcich iteráciách. V prípade, ak by sme potrebovali aj čiastkové výstupy, ukladali by sme do pamäte aj jednotlivé otuput_t. Vrátenie celej sekvencie je užitočné v prípade stohovania viacerých rekurentných vrstiev.

sieť, ktorá ale na rozdiel od LSTM neprodukuje tri výstupy, ale len dva, pričom zabúda výstup c_t.

Na obrázku 7 je znázornená sieť LSTM. Vidíme, že na vstupe t (označenej ako input_t) sa spracúva viacero rôznych vstupov. Output_t spočítame prirodzene pomocou daného vstupu, okrem toho ale vstupuje do výpočtu aj informácia z predchádzajúceho stavu vo forme c_t. V proceste spracovania, označeného ako Výpočet prenosu, LSTM spracuje výstup z častu t-1 a túto informáciu posunie na spracovanie v stave t ako state_t. Informácia z predchádzajúcich stavov sa tiež spracuje, napríklad sa zabudne zbytočná informácia, a takto sa posunie na spracovanie v stave t, ako c_t.

V našej práci budeme prevažne využívať LSTM siete a ich stohovanie, ale pozrieme sa aj na GRU siete. Architektúra sietí je podrobnejšie rozpísaná v kapitole 4. Poznamenajme ešte, že modelovanie v práci bude prebiehať v prostredí jazyka Python, špeciálne pomocou knižnice hlbokého učenia TensorFlow [7].



Obr. 7: Znázornenie princípu LSTM. Upravené z [12].

2 Prehľad publikovaných prác

Na tému využitia metód strojového učenia a predikovania vývoja akciových trhov bolo publikované veľké množstvo článkov. V tejto časti práce si predstavíme niektoré z nich. Práce môžeme rozdeliť podľa toho, akú metódu strojového učenia použili, alebo podľa toho, či sa zamerali na predikovanie priamo ceny aktíva alebo výnosu.

2.1 Predikovanie cien

Jedným z prvých, ktorý sa téme využitia neurónových sietí na predikciu cien aktív venoval, bol Schöneburg vo svojom článku z roku 1990 [27], v ktorom aplikoval neurónové siete na predikovanie cien akcií spoločností BASF, Commerzbank a Mercedes. Vstupné dáta pozostávali z linearizácie ceny v daný deň tak, že vstup pozostával z ceny v daný deň, absolútnou zmenou oproti predchádzajúcemu dňu, smer zmeny, smer zmeny dva dni dozadu a veľkosť najväčšej zmeny oprosti predchádzajúcemu dňu. Trénovacia sada pozostávala zo 42 dní a predikcie boli robené na najviac 56 dní. Pričom najlepšia presnosť - 79% bola pri akcií BASF. Pri akcií Comerzbank sa dosiahla presnosť 74% a najhoršie v prípade Mercedesu 56%.

Ding a Qin vo svojej práci [14] predikovali otváraciu, maximálnu a minimálnu cenu pomocou LSTM rekurentných neurónových sietí. Vo svojom modeli ako vstup zvolili 20 predchádzajúcich dní a predikovali nasledujúci deň pre tri aktíva, a to Shangai Index, Akcie PetroChina a akcie ZTE. V práci netestovali žiadne hyperparametre a preto dáta rozdelili len na trénovaciu a testovaciu sadu. Správnosť predikcie merali pomocou MAE

metriky (8), kde si zadefinovali presnosť ako 1 - MAE. Najlepšie výsledky dosiahli v prípade akcie spoločnosti PetroChnia a to 0.986 pre otváraciu cenu, 0.979 pre minimálnu a 0.98 pre maximálnu.

Kolektív autorov v článku [8], ktorý vyšiel v roku 2012 v časopise Journal of Emerging Trends in Computing and Information Sciences, sa zameral na predikciu cien akcií pomocou hlbokých dopredných neurónových sietí. Pracovali so sieťou s troma vrstvami a testovali rôzne počty neurónov v každej vrstve, kde najlepšie výsledky dosiahli pri počte 10-17-1. Ako vstup do siete zobrali otváracie, maximálne, minimálne a zatváracie ceny za dva predchádzajúce dni, a tiež príslušné objemy. Na vstup zobrali aj fundamentálne parametre za posledné dva roky, a to cenu počas oznámenia výsledkov, účtovnú hodnotu majetku a finančný stav firmy. Okrem týchto technických a fundamentálnych parametrov pridali aj aj parameter správ a noviniek za posledné dva dni. Avšak je potrebné povedať, že autori v nej neuvádzajú žiadnu metriku, ako vyhodnocovali modely, a taktiež nie je uvedené, o aké dáta sa jedná, ani ich zdroj.

Yu a Yan sa v článku z roku 2020 [34], ktorý bol vydaný v žurnáli Neural Computing and Applications, tiež zamerali na porovnanie LSTM sietí a iných predikčných metód. Porovnávané boli Arima, metóda oporných nadrovín a jednoduchá neurónová sieť. Predikovaných bolo 6 akciových indexov, tri čínske, dva zo Spojených Štátov a jeden japonský index. V tomto prípade je zaujímavý spôsob prípravy dát. Vstupné dáta upravili pomocou metódy PSR (Predictive state representation - Prediktívne znázornenie stavu) [19], kde sa za vstup zobrala posledná zatváracia cena a vektor predchádzajúcich zatváracích cien s dopredu definovaným posunutím. Využitie metódy PSR prinieslo zlepšenie na indexe S&P 500 o 3% a na ostatných indexoch o 1%. LSTM sieť obsahovala jednu skrytú vrstvu a počet neurónov bol testovaný ako hyperparameter. Vyhodnotnocovali sa tri metriky, správnosť určenia trendu, odmocnina z MSE, MAE vyjadrené ako percento z predikovanej hodnoty a Pearsonov korelačný koeficient. Autori ukázali dominanciu LSTM sietí oproti ostatným prístupom, keď vo všetkých prípadoch, okrem indexu S&P 500, porazili ostatné metódy vo všetkých metrikách. V prípade S&P 500 najlepšie výsledky dosiahla jednoduchá neurónová sieť.

2.2 Predikovanie výnosov

Li a Mei v článku [21] z roku 2020 uverejnenom v žurnále Research in International Business and Finance preikovali logaritmické výnosy dvoch čínskych akciových indexov SSE 50 a CSI 300. Za vstupné dáta si zobrali logaritmické rozdiely zatváracej, maximálnej a minimálnej ceny, a technické indikátory pozostávajúce z kĺzavých priemerov, indexu relatívnej sily, momentum a volatilitu. Autori sa pokúsili predikovať výnosy pomocou lineárnej regresie Enet a neurónových sietí. Výkonnosť predikcií merali pomocou koeficientu determinácie uvedeného v percentách. Pri indexe SSE 50 dosiahli najväčší koeficient determinácie práve neurónové siete a to 1.5 % na vzorke mimo trénovacích dát, naopak, lineárna regresia dosiahla R^2 iba -22.3 %. Pre index CSI 300 dostávame obdobné výsledky, keď neurónová sieť dosiahla koeficient determinácie 3.26 % a lineárna regresia -19 %.

Feng, He a Polson v práci s názvom Deep Learning for Predicting Asset Returns [16] porovnali prínos hlbokého učenia oproti klasickým metódam predikcie. Predikovaným výstupom boli výnosy indexu S&P 500. Na vstup zobrali 14 parametrov, ktoré pozostávali z fundamentov a ukazovateľov celkového makroekonomického vývoja. Následne pridali 4 scenáre, kde v prvom bolo len základných 14 parametrov, v druhom aj ich druhé mocniny, v treťom pridaný indikátor pozitívneho výnosu a v štvrtom sú skombinované všetky predchádzajúce vstupy. Autori porovnávali 4 možnosti nastavenia neurónovej siete a obyčajnú lineárnu regresiu, Lasso metódu PLS a ElasNet. V prípade predikovania mesačných výnosov najlepšie výsledky dosiahla sieť s troma skrytými vrstvami a s počtom neurónov 16-8-4. Táto sieť dosiahla na mimotrénovacej sade R^2 rovné 0.01111. Najhoršie výsledky dosiahli klasické regresie, ktoré ani v jednom prípade nemali kladný koeficient determinácie. Pre trojmesačné predikcie dostali obdobné výsledky, keď pre štvrtý scenár dosiahla tá istá sieť R^2 rovný 0.01266. Opäť výrazne najhoršia bola klasická regresia. Ako posledné predikovali ročné výnsoy a aj v tomto prípade dosiahli výsledky podobné ako v predchádzajúcich, kde sieť 16-8-4 dosiahla R^2 na úrovni 0.03356 a najhoršie dopadla lineárna regresia.

3 Dáta

V tejto kapitole si popíšeme, na základe akých dát budeme predpovedať budúci vývoj, čo bude naším vstupom do siete a ukážeme, ako sme vyberali podvzorku. V práci používame hodinové a denné dáta pre akcie z indexu S&P500, ktoré sme získali zo stránky marketstack.com [3]. Dáta rozdelíme nasledovným spôsobom, 70% najstarších dát bude slúžiť ako trénovacia sada, nasledujúcich 15% dát ako validačná sada, a najnovších 15% dát bude tvoriť testovaciu sadu. Dáta boli stiahnuté dňa 16.11.2020, a posledný záznam pre každú akciu bude zo dňa 10.11.2020 zo 16:00 hodiny.

Vstupné dáta budú pozostávať z otváracej, minimálnej, maximálnej a zatváracej ceny pre každú hodinu alebo deň, všetky upravené o vyplatené dividendy a delenia akcie. Ďalej z týchto cien odvodíme technické indikátory, ktoré sú popísané v podkapitole 3.3. Indikátory budeme testovať ako hyperparmetre, ktoré sú vysvetlené v časti 4.2.

3.1 Podvzorka pre denné dáta

Podvzorka pre denné dáta obsahuje 100 náhodne vybraných akcií tak, aby ich fundamentálne parametre zodpovedali parametrom celého indexu. Pod fundamentálnymi parametrami spoločnosti (vo všeobecnosti sa dá hovoriť o fundamentálnych parametroch ľubovoľného aktíva) rozumieme vyjadrenie rastu, likvidity, finančného zdravia, zadĺženia, vnútornej hodnoty alebo iného popisu spoločnosti pomocou účtovníctva spoločnosti a odvodených ukazovateľov. Všetky fundamentálne parametre pre jednotlivé spoločnosti v tejto práci pochádzajú zo stránky [4] a obsahujú dáta ku koncu druhého kvartálu roku 2020. My sa pozrieme na nasledujúce parametre:

- **Cena/Zisk** hovorí, koľko dolárov zisku pripadá na jednu akciu, čím je hodnota vyššia, tým možno hovoriť o, v určitom zmysle, drahšej akcii, pretože z aktuálneho zisku sa hodnota spoločnosti splatí za dlhší čas.

- **Zisk na akciu** zasa poskytuje informáciu o ziskovosti spoločnosti. Chceme, aby bol čo najvyšší, avšak tento ukazovateľ môže byť klamlivý, keďže nezohľadňuje počet vydaných akcií.

- **EBITDA** predstavuje zisk pred odpismi, amortizáciou, úrokmi a zdanením, tento ukazovateľ umožňuje lepšie porovnávať ziskovosť pre investora, keďže abstrahuje

od nehotovostných položiek a od položiek ako dane a úroky, ktoré sa po príchode cudzieho kapitálu môžu zásadne meniť.

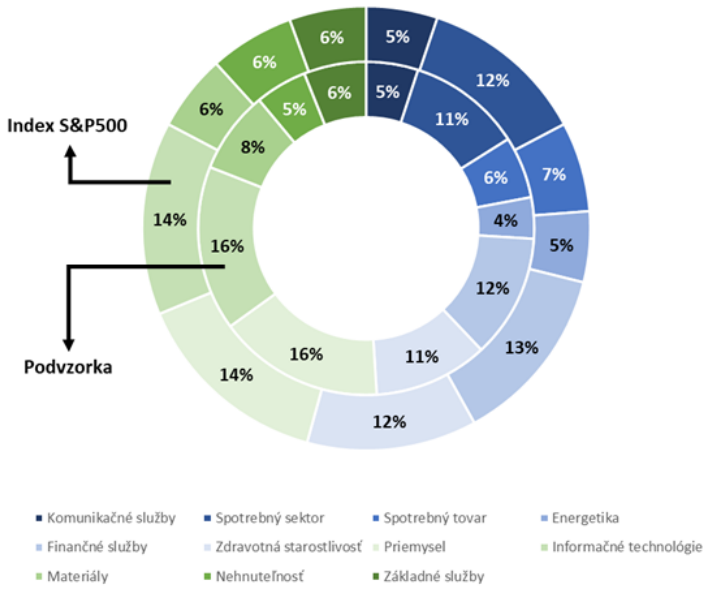
- **Cena/tržby** zasa hovorí o tom, koľko investor zaplatí za jednotku tržby, za lepšie sa považuje nižšie číslo, avšak v prípade nekvalitných firiem s malými tržbami to poskytuje klamlivý obraz.

- **Cena/účtovná hodnota** hovorí, koľkokrát je spoločnosť drahšia ako jej majetok očistený od záväzkov. Čím je tento pomer väčší, tým by mala byť akcia menej výhodná.

Priemer	Cena/ Zisk	Zisk na akciu	Trhová kapitalizácia v mld.	EBITDA v mld.	Cena/ Tržby	Cena/ Účtovná hodnota
Podvzorka	26.68	3.91	57.6	4.01	4.51	14.09
S&P 500	25.25	3.81	53	3.84	4.06	15.77

Tabuľka 1: Porovnanie vybraných fundamentálnych parametrov podvzorky pre denné dáta a indexu S&P 500.

V tabuľke 1 máme popis fundamentálnych parametrov našej podvzorky. Parameter cena/zisk ukazuje, že naša podvzorka je len mierne drahšia, ako celý index. Tiež, čo sa týka ziskovosti na akciu, sa podvzorka zhoduje s indexom. Na trhovej kapitalizácii vidno, že priemerné aktívum v našej podvzorke je takmer rovnako drahé, ako priemerné aktívum v indexe. Ukazovateľ cena/tržby ukazuje, že priemerná spoločnosť podvzorky, ako aj indexu, generuje vzhľadom na cenu podobné predaje. Cena/účtovná hodnota hovorí, že spoločosti v našej podvzorke majú, vzhľadom na svoju cenu, trocha menej aktív, ako je priemer. Je zrejmé, že priemerné fundamentálne parametre našej podvzorky v tomto prípade zodpovedajú parametrom celého indexu. Môžeme skonštatovať, že podvzorka nie je výrazne vychýlená oproti celému indexu. Zastúpenie jednotlivých sektorov v indexe a v podvzorke je na obrázku 8. Môžeme si všimnúť, že podiel sektorov v podvzorke veľmi dobre zodpovedá podielu sektorov v celom indexe.



Obr. 8: Znázornenie podielu jednotlivých sektorov v podvzorke a celkom indexe S&P500.

3.2 Podvzorka pre hodinové dáta

Podvzorka pre hodinové dáta bude menšia, ako pre denné dáta, keďže k dispozícii sme mali len malý počet akcií s dostatočným počtom dát. Vezmeme preto 20 akcií, pre ktoré sme vedeli získať dostatočný počet dát z minulosti. Z dôvodu nedostatku dát pre väčšinu akcií v indexe, je naša vzorka vychýlená, v zmysle fundamentálnych parametrov, oproti hodnotám pre celý index, ako môžeme vidieť v tabuľke 2. Parameter cena/zisk ukazuje, že naša podvzorka je podstatne drahšia, ako celý index. Podvzorka sa na základe zisku na akciu zdá značne ziskovejšia ako index. Na trhovej kapitalizácií vidno, že priemerné aktívum v našej podvzorke je viac ako trikrát také drahé, ako priemerné aktívum v indexe. Z EBITDA vidíme, že spoločnosti vo vzorke sú trojnásobne ziskovejšie. Ukazovateľ cena/tržby je najmenej vychýlený oproti celému indexu, a teda priemerná spoločnosť podvzorky, ako aj indexu, generuje vzhľadom na cenu podobné predaje. Cena/účtovná hodnota hovorí, že spoločnosti v našej podvzorke majú, vzhľadom na svoju cenu, výrazne menej aktív ako je priemer. Je zrejmé, že naša vzorka obsahuje akcie, ktoré sa zdajú lepšie, ako priemerná akcia indexu.

Priemer	Cena/ Zisk	Zisk na akciu	Trhová kapitalizácia v mld.	EBITDA v mld.	Cena/ Tržby	Cena/ Účtovná hodnota
Podvzorka	39.87	5.40	171	12	5.69	8.28
S&P 500	25.25	3.81	53	4	4.06	15.77

Tabuľka 2: Porovnanie vybraných fundamentálnych parametrov podvzorky a indexu S&P 500.

3.3 Technické indikátory

Technický indikátor je spravidla nejaký ukazovateľ, odvodený od ceny akcie, ktorý pomáha obchodníkom určiť ďalší vývoj akcie. Patria medzi kontroverzné nástroje používané pri analýze akcií, keďže neberú do úvahy žiadne ekonomické pozadie. V tejto časti si zadefinujeme technické indikátory, ktoré budeme používať v našich sieťach.

- **Kĺzavý priemer (SMA)**: Tento indikátor získame ako priemer za posledných n dní v čase t:

$$SMA(t,n) = \frac{p_t + ... + p_{t-n}}{n}, \tag{25}$$

kde p_t je zatváracia cena v čase t. Tento indikátor budeme používať ako kombináciu troch priemerov, s rôznymi n. Za rastový signál budeme považovať nasledovné zoradenie priemerov: najkratší má najväčšiu hodnotu, nasleduje stredný a najdlhší má najmenšiu. Takéto usporiadanie naznačuje, že priemerná cena za obdobie vo vzdialenejšej minulosti bola nižšia, ako v bližšej.

- **Bollinger Bands® (BB)**: Jedná sa o pridanie štandardnej odchýlky ku kĺzavému priemeru. BB teda pozostáva z hornej a dolnej hranice, ktoré dostaneme ako:

$$BB_{upper}(t,n) = SMA(t,n) + 2\sigma(t,n), \tag{26}$$

$$BB_{lower}(t,n) = SMA(t,n) - 2\sigma(t,n), \tag{27}$$

kde $\sigma(t,n)$ predstavuje štandardnú odchýlku v čase t, za posledných n dní. Myšlienka za týmito pásmi vyplýva zo štatistických intervalov spoľahlivosti, kde by v prípade normálneho rozdelenia mali hodnoty na 95% padnúť do tohto intervalu.

- **Konvergencia/divergencia kĺzavých priemerov (MACD)**: Ide o, v našom prípade, prvý momentum oscilátor, ktorý vyjadruje silu trendu. Pozostáva z troch častí, a to MACD línia, signálna línia a MACD histogram. Tieto sa vypočítajú na základe exponenciálnych kĺzavých priemerov:

$$EMA(t,n) = \gamma(p_t + (1-\gamma)p_{t-1} + ... + (1-\gamma)^{n-1}p_{t-n}), \tag{28}$$

kde γ je konštanta vyhladzovania a zvyčajne sa volí ako $\gamma = 2/(n+1)$. Potom získame MACD ako:

$$MACD_t = EMA(t,12) - EMA(t,26), \tag{29}$$

$$SIGNAL_t = MACD_t + (1-\gamma)MACD_{t-1} + ... + (1-\gamma)^8 MACD_{t-9}, \tag{30}$$

$$MACD_t^{histogram} = MACD_t - SIGNAL_t. \tag{31}$$

Následne sa za signál na kúpu pokladá, ak MACD pretne signálnu líniu zospodu a MACD histogram sa dostane zo záporných do kladných hodnôt.

- **Priemerný index smeru (ADX)**: Ide o oscilátor, ktorý určuje silu trendu. Na jeho výpočet si potrebujeme zadefinovať premenné smeru pohybu (DM):

$$UpMove_t = h_t - h_{t-1}, \tag{32}$$

$$DownMove = l_{t-1} - l_t, \tag{33}$$

kde h_t je maximum počas obdobia t a l_t je minimum v období t. Potom definujeme +DM, -DM a TrueRange (TR) ako:

$$+DM_t = \begin{cases} UpMove_t, \text{ ak } UpMove_t > DownMove_t \text{ a } UpMove_t > 0 \\ 0, \text{ ináč} \end{cases}, \tag{34}$$

$$-DM_t = \begin{cases} DownMove_t, \text{ ak } UpMove_t < DownMove_t \text{ a } DownMove_t > 0 \\ 0, \text{ ináč}, \end{cases} \tag{35}$$

$$TR_t = max\{(h_t - l_t); |h_t - p_{t-1}|; |p_{t-1} - l_t|\}. \tag{36}$$

Následne sa tieto hodnoty exponenciálne vyhladia:

$$+DM(t,n) = \frac{n-1}{n} \sum_{i=t-n}^{n} (+DM_i) + (+DM_t), \tag{37}$$

$$-DM(t,n) = \frac{n-1}{n} \sum_{i=t-n}^{n} (-DM_i) + (-DM_t), \tag{38}$$

$$TR(t,n) = \frac{n-1}{n} \sum_{i=t-n}^{n} (TR_i) + (TR_t). \tag{39}$$

Akonáhle máme tieto hodnoty, spočítame smerový index $+DI(t,n)$ a $-DI(t,n)$:

$$+DI(t,n) = 100\frac{+DM(t,n)}{TR(t,n)}, \tag{40}$$

$$-DI(t,n) = 100\frac{-DM(t,n)}{TR(t,n)}. \tag{41}$$

Následne spočítame priemerný index smeru ADX ako:

$$ADX(t,n) = 100\frac{|(+DI) - (-DI)|}{(+DI) + (-DI)}. \tag{42}$$

Hodnoty indikátora ADX sa pohybujú medzi 0 a 100 a má sa za to, že hodnoty menšie ako 20 znamenajú slabý trend a hodnoty väčšie ako 40, naopak, trend silný.

- **Index relatívnej sily (RSI)**: Radí sa opäť medzi momentum oscilátory a určuje prekúpenosť, respektíve prepredanosť aktíva. Vypočítame ho nasledovne, nech $UP(t,n)$ je priemerný vzrast ceny počas vzrastových dní za posledných n dní a $DOWN(t,n)$ zasa priemerný pokles. Index relatívnej sily je potom daný ako:

$$RS(t,n) = \frac{UP(t,n)}{DOWN(t,n)}, \tag{43}$$

$$RSI(t,n) = 100 - \frac{100}{1 + RS(t,n)}. \tag{44}$$

Je zrejmé, že $RSI(t,n) \in [0,100]$. Za prekúpené (obchodované za vyššiu ako vnútornú cenu) aktívum sa považuje vtedy, ak hodnoty RSI presahujú hodnotu 70. Naopak, za prepredané (obchodované za nižšiu ako vnútornú cenu) sú považované aktíva s hodnotami RSI pod 30.

- **Stochastika (STOCH)**: Opäť sa jedná o momentum oscilátor, ktorý hovorí o prekúpenosti alebo prepredanosti aktíva. Skladá sa z dvoch častí, a to rýchla (%D)

a pomalá (%K). Vypočítame ich nasledovným spôsobom:

$$\%K_t = \frac{100(p_t - L_{14})}{H_{14} - L_{14}}, \tag{45}$$

$$\%D(t,n) = \frac{K_t + K_{t-1} + ... + K_{t-n}}{n}, \tag{46}$$

kde hodnoty H_{14} a L_{14} predstavujú maximálnu a minimálnu cenu za posledných 14 dní. Pomocou tohto indikátora porovnávame, ako sú tieto hodnoty vysoké a nízke. Hodnoty nad 80 sú považované za prekúpené a hodnoty pod 20 za prepredané. Ak sa obe nachádzajú v extrémnych hodnotách, tak sa sleduje aj ich kríženie, ak %D pretne zospodu %K, tak hovoríme o nákupnom signále.

Indikátory, ktoré budeme používať, môžeme rozdeliť do troch kategórií. Prvou sú indikátory, ktoré nám hovoria priamo o pohybe ceny akcie a na ich základe sa snažíme odhadovať ďalší vývoj. Zo spomenutých indikátorov sem môžeme zaradiť kĺzavé priemery a Boilnger Bands.

Druhou kategóriou sú momentum oscilátory, ktoré pomáhajú odhadovať silu, už aktuálneho trendu. Pomocou nich určujeme, či napríklad rast ceny má ešte silu pokračovať, alebo už je slabý a môžeme očakávať zmenu. Ide predovšetkým o konvergenciu/divergenciu kĺzavých priemerov (MACD) a priemerný index smeru (ADX).

Index relatívnej sily (RSI) a stochastiku (STOCH) zaraďujeme do poslednej kategórie, a to indikátory určujúce prekúpenosť a prepredanosť aktíva, teda snažiace sa dať do kontextu trhovú cenu a vnútornú hodnotu.

4 Architektúra sietí

V tejto kapitole si uvedieme, aké siete a ich nastavenia budeme testovať. V našej práci sa pozrieme na dve rôzne predikcie. V jednej budeme modelovať priamo cenu nasledujúceho tiku a v druhej logaritmický výnos. Taktiež budeme brať do úvahy tri množiny parametrov, a to architektúra siete, počet dát z minulosti, na ktoré sa dívame a technické indikátory.

4.1 Voľba predikovaného výstupu

Ako prvé budeme testovať dva prístupy k predikciám. Pokúsime sa predikovať priamo ceny akcie, čo bude predstavovať prvý prístup. Ďalším krokom bude predikcia logaritmických výnosov.

Prvým prístupom je predikovanie priamo ceny akcie v ďalšom tiku. V tomto prípade dáta po rozdelení na trénovaciu, validačnú a testovaciu sadu každú znormalizujeme pomocou min-max normalizácie:

$$X^n = \frac{X - \min(X)}{\max(X) - \min(X),} \tag{47}$$

kde X označuje vstupné dáta a X^n normalizované. V tomto prípade $min(X)$ a $max(X)$ pochádza z každej sady samostatne. Výhodou tohto prístupu je, že ceny sa zvyknú pohybovať v nejakom trende, ktorý môže byť natoľko silný, že pre sieť bude veľmi jednoduché predikovať nasledujúci pohyb. Na druhú stranu, problematickým miestom v tomto prístupe je absencia stacionárnych dát.

Druhý prístup predstavuje predikciu logaritmických výnosov. Tento prístup má výhodu v odstránení nestacionarity dát, ale naráža vo svojej podstate na teóriu finančnej matematiky. Konkrétne sa predpokladá, že výnosy sa riadia Brownovým pohybom, kde sú výnosy v čase t nezávislé od predchádzajúcich výnosov [23]. Našou snahou bude pozrieť sa na to, či neurónové siete budú schopné namodelovať nejakú nelinearitu v dátach, vďaka ktorej sa nám podarí vysvetliť aspoň časť variancie. V tomto prípade dáta nebudeme normalizovať, keďže logaritmické výnosy by sa mali riadiť normálnym rozdelením so strednou hodnotou 0. V tomto prípade zvolíme opačný prístup a pokúsime sa zväčšiť varianciu prenásobením konštantou 100 tak dostaneme dáta opäť z normálneho rozdelnia s núlovou strednou hodnotou, ale s oveľa väčšou varianciou. Myšlienka zväčšenia variancie dát je, že pre našu sieť bude jednoduchšie naučiť sa vzory, ak budú výrazné väčšie než v prípade, že variancia je len na úrovni bieleho šumu.

4.2 Výber najlepšej architektúry

Obzvlášť dôležitým krokom je voľba vhodnej architektúry. My budeme testovať 9 rôznych nastavení od najjednoduchších rekurentných sietí až po zložitejšie. Sieťam nastavíme počiatočný bod na incializačné váhy, tak aby mali všetky tak-povediac ten istý štartovací

bod a nastavíme 20% mieru výpadku.

Výpadok (dropout) sa používa v boji proti preučeniu [30]. Myšlienka za ním je, že s nejakou dopredu danou pravdepodobnosťou, sa náhodne nastaví váha neurónu na 0. Tým pádom, tento neurón už nemá žiadny vplyv na učenie a akoby zo siete "vypadol". V našom prípade 20% miera výpadku znamená, že pravdepodobnosť nastavenia váhy siete na nulu je 0.2.

Všetky siete budú mať poslednú vrstvu pozostávajúcu z vrstvy 'Dense', ktorá slúži priamo na predikciu finálnej hodnoty. Jednotlivé nastavenia sú uvedené v tabuľke 3. Počet neurónov v skrytých vrstvách budeme voliť podľa pravidla 2/3 z veľkosti vstupnej vrstvy, ako je to spomenuté v [33].

Index siete						
0	LSTM	Dense				
1	LSTM	LSTM	Dense			
2	LSTM	LSTM	LSTM	Dense		
3	LSTM	LSTM	LSTM	LSTM	Dense	
4	LSTM	LSTM	LSTM	LSTM	LSTM	Dense
5	LSTM	Dense	Dense			
6	LSTM	LSTM	Dense	Dense		
7	GRU	Dense				
8	GRU	GRU	Dense			

Tabuľka 3: Architektúra používaných sietí.

4.3 Výber počtu dát z minulosti

Budeme sa zaujímať aj o to, koľko dát z minulosti je pre predikciu relevantných. Postupne budeme prechádzať všetky architektúry z podkapitoly 4.2 s nastavením od 10 do 50 minulých tikov, a to s krokom vždy 10. Na vstupe tak budeme testovať 10, 20, 30, 40 a 50 minulých dát.

4.4 Výber technických indikátorov

Posledným hyperparametrom, ktorý budeme testovať, bude výber vhodných technických indikátorov. Tento hyperparameter je zaujímavý, pretože technické indikátory sú spravidla nejaká transformácia poslednej a prechádzajúcich cien. Z toho dôvodu by niektorí zástancovia nadradenosti neurónových sietí mohli namietať, že dostatočne hlboká neurónová sieť by sa mala tieto transformácie naučiť aj sama, v prípade, že sú dôležité. My najskôr otestujeme každý z vybraných indikátorov z podkapitoly 3.3. V prípade predikcie cien budeme testovať všetky spomenuté indikátory. Pri logaritmických výnosoch nebudeme testovať ADX, RSI a STOCH, keďže v tomto prípade dáta nenormalizujeme, a tieto indikátory nadobúdajú výrazne iné hodnoty, ako ostatné dáta. Z nich zoberieme ten najlepší vzhľadom na PL a Sharpovho pomer. K tomu následne otestujeme zvyšné indikátory a pozrieme sa, či sa zvýši zisk a Sharpov pomer na validačnej sade. V prípade, že sa zvýši, tak postup opakujeme, ak nie, tak skončíme.

4.5 Vyhodnocovanie modelov

Všetky modely budeme optimalizovať na trénovacej sade pomocou strednej absolútnej chyby (MAE) a sledovať budeme aj strednú štvorcovú chybu (MSE), ktoré sú definované v podkapitole 1.4. Dôvod, prečo sme zvolili MAE pred MSE je ten, že min-max normalizácia znormuje všetky hodnoty do intervalu $[0, 1]$ a na tomto intervale MSE vracia veľmi nízke hodnoty. Avšak výber hyperparametrov na validačnej sade už nebude prebiehať pomocou MAE ani MSE, ale zadefinujeme si finančné metriky, pomocou ktorých budeme merať koľko, a pri akom riziku vieme zarobiť pri obchodovaní s danou akciou.

4.5.1 Ziskovosť modelu

Na určenie ziskovosti si potrebujeme zadefinovať stratégiu, pomocou ktorej budeme obchodovať. Na začiatku budeme mať kapitál vo výške 100 000 $ a každý tik budeme rebalancovať. Vstupovať budeme môcť do nákupnej (buy, long), predávajúcej (sell, short) pozície, alebo nebudeme musieť vstúpiť do žiadnej. Z toho dôvodu si spočítame prvý kvartil absolútnych výnosov trénovacej sady, čo budeme považovať za našu hranicu na otvorenie pozície. Ak je predikovaný nárast väčší ako tento kvartil, tak uzatvárame nákupnú pozíciu. Naopak, ak je menšia, ako záporný kvartil, tak vstupujeme do predajnej pozície,

a ak sa nachádza v intervale medzi záprnou a kladnou hodnotou kvartilu, tak do pozície nevstupujeme.

V našom texte budeme pod pojmom kapitál alebo nadobudnutý kapitál rozumieť konečný stav peňazí, ktoré získame, inými slovami počiatočný kapitál + obrat z obchodovania. Pod pojmom zisk alebo zisk stratégie rozumieme iba čistý obrat z obchodovania zmenšený o počiatočný kapitál.

4.5.2 Sharpov pomer

Sharpov pomer je ukazovateľ finančnej výkonnosti, ktorý bol uvedený v roku 1966 Williamom F. Sharpeom v publikácii [28]. Tento pomer berie do úvahy to, aké veľké riziko podstupujeme na dosiahnutie nadmerného výnosu. Sharpov pomer sa definuje ako [23]:

$$S = \frac{E(R - R_f)}{\sigma}, \tag{48}$$

kde R označuje výnos investície, R_f bezrizikový výnos a σ označuje volatilitu investície. Čím je tento pomer vyšší, tým je investícia vzhľadom na svoju rizikovosť lepšia. Za bezrizikový výnos sa berú výnosy krátkodobných dlhopisov spoľahlivej krajiny, ako napríklad Nemecko, ktorých výnosová krivka je v súčasnosti záporná, tak za bezrizikový výnos zvolíme 0.

5 Výsledky na denných dátach

V piatej kapitole si predstavíme a popíšeme dosiahnuté výsledky, pričom v tejto kapitole za jeden tik považujeme jeden deň. Kapitola je rozdelená do podkapitol podľa typu predikovaného výstupu, ktoré sú ďalej členené na výsledky pre validačnú a testovaciu sadu.

5.1 Predikcia cien

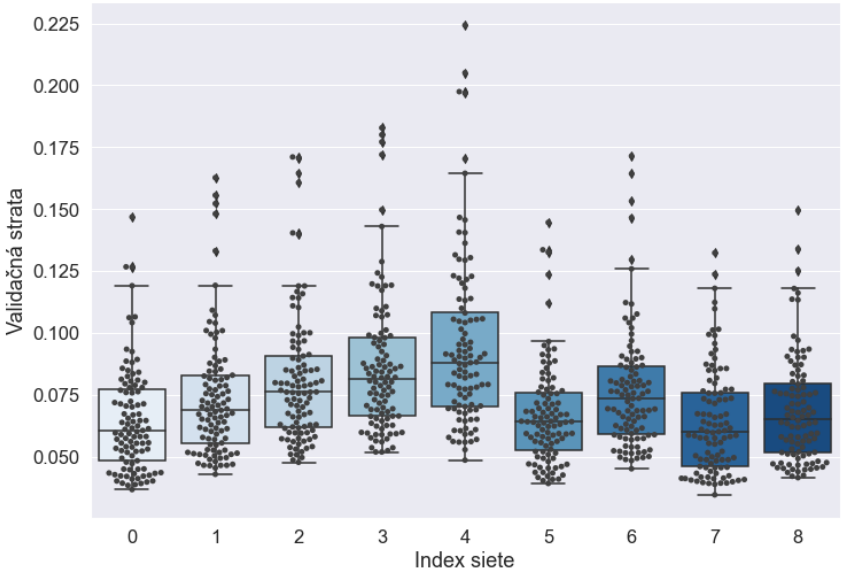
Pri predikcii cien sa spoliehame na to, že pri pohybe je trend, ktorý bude naša sieť schopná identifikovať a na základe ktorého budeme vedieť identifikovať nasledujúcu hodnotu. Na druhú stranu, táto nestacionarita dát je problematickým bodom tohto prístupu. V tejto podkapitole sa pozrieme, ako sa naša sieť popasovala s touto úlohou.

5.1.1 Výsledky na validačnej sade

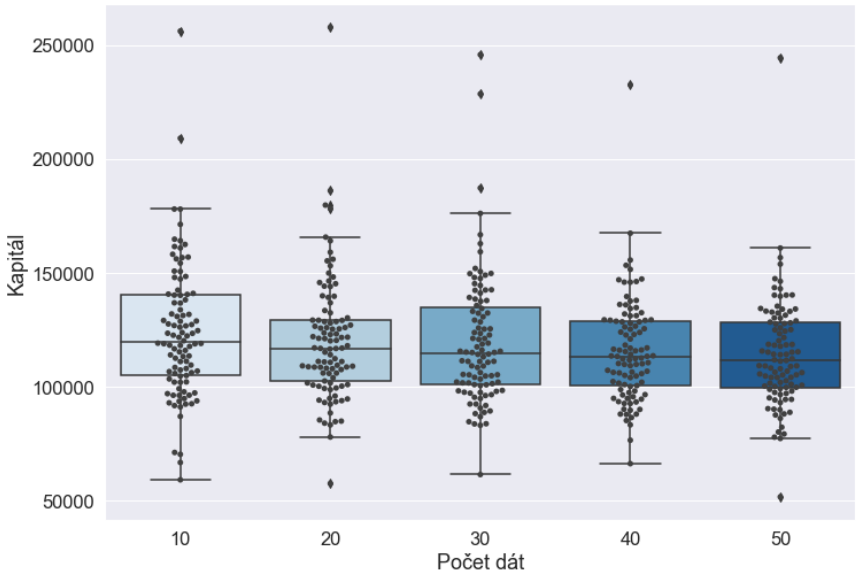
V tejto časti si uvedieme, ako naše siete fungovali alebo nefungoavli na validačnej sade, so zreteľom na testované hyperparametre. Pri posudzovaní toho, ako dobre sieť fungovala, sa budeme dívať na validačnú stratu - MAE, kapitál a Sharpov pomer. Na obrázku 9 sú znázornené krabicové grafy pre rôzne siete z podkapitoly 4.2 a validačné straty, aké tieto siete dosiahli. Javí sa, že jednoduchšie siete, s menším počtom vrstiev (indexy 0, 5 a 7) dosahujú nižšiu validačnú stratu, ako komplikovanejšie siete. Jedným z vysvetlení je, že čím je sieť hlbšia, tým má väčšiu náchylnosť k preučeniu. Tiež jednoduchšie GRU siete vykazujú nižšiu validačnú stratu, ako komplikovanejšie LSTM.

Zaujímavý je pre nás aj získaný kapitál na validačnej sade, keďže finálnu sieť vyberáme podľa získaného kapitálu a Sharpovho pomeru. Ak rozdelíme získaný kapitál podľa počtu dní, ktoré sme brali na vstupe, tak pri všetkých rôznych nastaveniach vstupných dní, sme dosiahli takmer rovnaké priemerné výsledky, ako je vidno na obrázku 10.

Najzaujímavejšou metrikou pre nás ale bude Sharpov pomer, ktorý poskytuje dostatočnú informáciu nielen o ziskovosti alebo stratovosti, ale aj o rizikovosti našej stratégie. Ako prvý si zoberieme Sharpov pomer a index siete, ktorý je na obrázku 11. Medzikvartilový rozsah všetkých sietí je veľmi podobný, a nedá sa jednoznačne povedať, ktorá sieť z tohto hľadiska vykazuje najlepšie hodnoty. Za povšimnutie stojí, že zložitejšie siete s



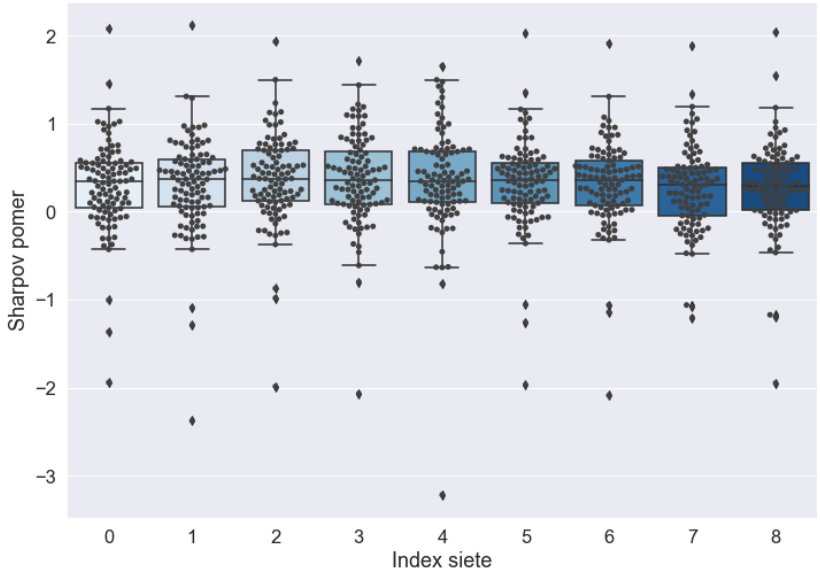
Obr. 9: Krabicový graf so znázornením validačnej straty - MAE, vzhľadom na index siete.



Obr. 10: Krabicový graf so znázornením nadobudnutého kapitálu a počtu dní, ktoré do siete vstupovali.

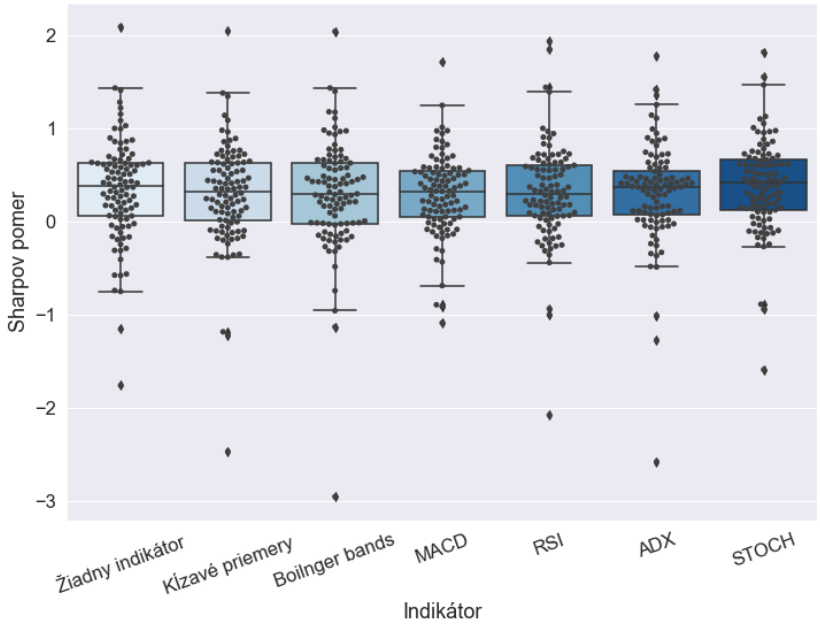
viacerými vrstvami (indexy 2, 3 a 4), majú najviac dát s vysokými hodnotami Sharpovho pomeru. Môžeme teda očakávať, že práve tieto siete, budú prevažne použité aj na testovacej sade.

Ďalším testovaným hyperparametrom boli technické indikátory. Pre prehľadnosť uvádzame výsledky len pre základné indikátory a nie aj pre ich kombinácie. Za základný model považujeme model bez indikátora a zaujíma nás, či pridaním nejakého technického in-



Obr. 11: Krabicový graf zobrazujúci Sharpov pomer vzhľadom na použitú sieť.

dikátora sa zlepší Sharpov pomer. Dosiahnutý Sharpov pomer vzhľadom na použitý indikátor je na obrázku 12. Opäť môžeme konštatovať, že žiadny z indikátorov výrazne nezlepšuje Sharpov pomer oproti základnému modelu. Spomedzi vetkých indikátorov iba indikátor STOCH, dokázal mierne zlepšiť Sharpov pomer. Môžeme teda povedať, že naša sieť sa bola schopná naučiť všetky potrebné informácie zo samotných cien.



Obr. 12: Krabicový graf zobrazujúci vzťah medzi nadobudnutým Sharpovým pomerom a použitým indikátorom.

5.1.2 Výsledky na testovacej sade

Najpodstatnejšie pre nás je zistenie, aké výsledky sme dosiahli na testovacích dátach. Na tejto sade chceme poraziť stratégiu "Buy & Hold", kde danú akciu kúpime na začiatku testovacieho obdobia, a držíme až do jeho skončenia. Našu stratégiu tak budeme porovnávať s jednoduchou stratégiou, ktorú by mohol investor aplikovať, ak by nepoznal našu stratégiu, budeme teda "Buy & Hold"považovať za akýsi benchmark. V tabuľke 4 je uvedený základný popis rozdelenia jednotlivých sledovaných metrík na testovacej sade. Testovacia strata (MAE) je väčšia ako MSE, čo je spôsobené použitím min-max normalizácie. Ak porovnáme zisk z našej stratégie a zisku stratégie "Buy & Hold", tak sa zdá, že naša stratégia prekonáva prístup "Buy & Hold"čo do veľkosti zisku, aj do obmedzenia straty. Štatistickú významnosť tohto zdania otestujeme pomocou párového Studentovho t-testu, na 5% hladine významnosti. Nulová hypotéza bude predpokladať, že profit v "Buy & Hold"je väčší alebo rovný ako zisk našej stratégie a alternatívna hypotéza pokrýva doplnok:

$$H_0\text{: } \mu_{zisk} \leq \mu_{Buy\&Hold} \text{ vs } H_1\text{: } \mu_{zisk} > \mu_{Buy\&Hold}. \tag{49}$$

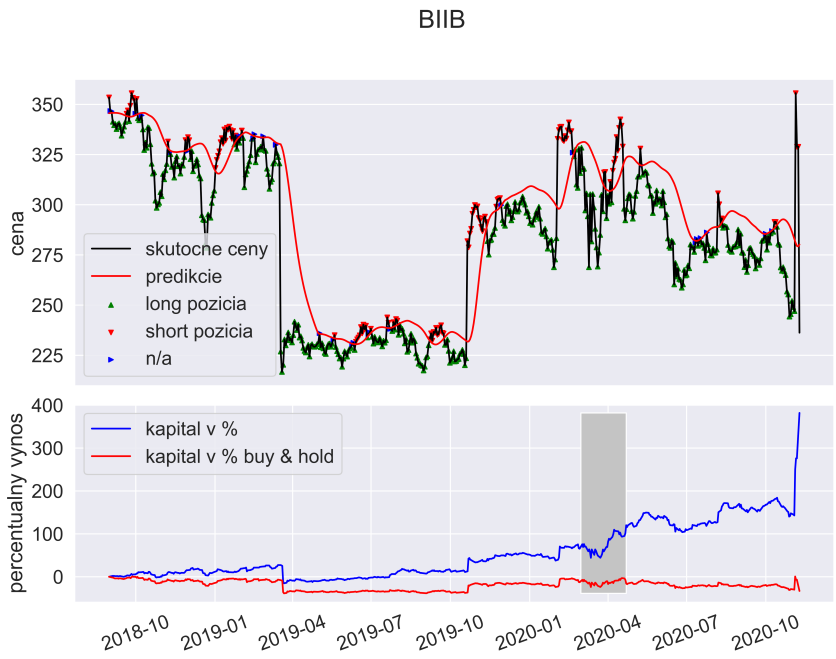
Dostávame testovaciu štatistiku $T = 1.06$, ktorú porovnáme so Studentovým rozdelením 99 stupňami voľnosti. Dostávame p-hodnotu 0.1451, ktorá je väčšia ako 0.05, teda nezamietame H_0 a priemerný rozdiel v zisku nie je štatisticky signifikantný.

V tabuľke 4 sa nachádza niekoľko výrazne vysokých alebo nízkych charakteristík. Maximálny zisk 381 947.87 $ bol nadobudnutý pre spoločnosť Biogen Inc (symbol: BIIB). Biogen je biotechnologická spoločnosť, ktorá vyvíja lieky na boj s neurologickými ochoreniami [1]. Zanalyzujme, čo sa dialo počas testovacieho obdobia pri tejto akcii. Na obrázku 13 je vývoj akcie počas testovacieho obdobia so znázornením pozícií, do ktorých sme vstupovali. Najväčší zisk bol zaznamenaný počas posledných dní v sade, ktoré boli správne predikované, čo sa týka smeru. V tomto období, od 4. novembra do 9. novembra 2020, FDA (Food and Drug Administration - agentúra v USA, kontrolujúca potraviny a lieky) najskôr oznámila, že má dostatok dát, aby schválila nový experimentálny liek od spoločnosti Biogen na liečbu Alzheimerovho ochorenia. Toto vyhlásenie vyústilo do prudkého rastu, keďže sa jednalo o prvý liek na túto chorobu [2]. Následne na to, ale komisia FDA hlasovala proti schváleniu, kde toto rozhodnutie prirodzene vyústilo k pádu.

Iným parametrom, ktorý nadobudol, naopak, nezvyčajne nízku hodnotu, je parameter

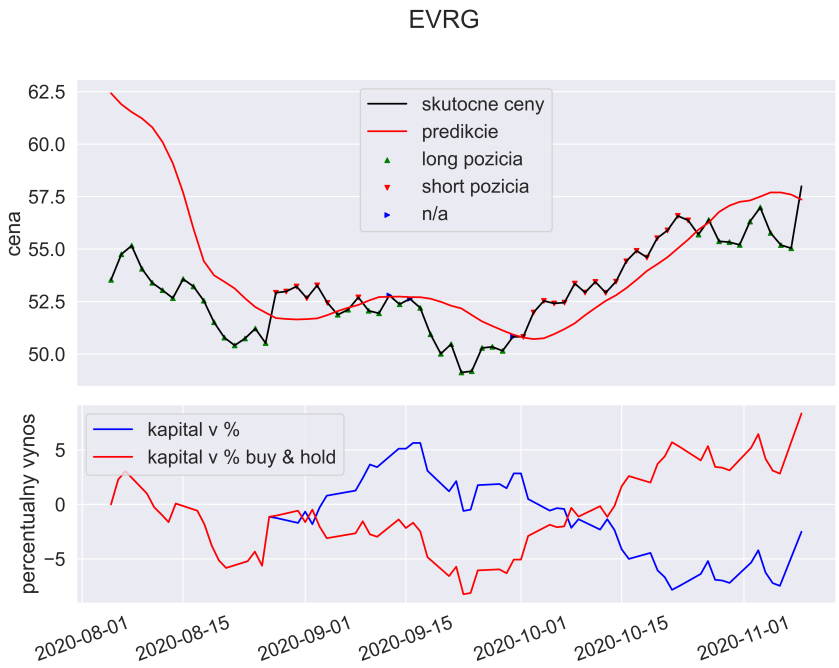
	Strata	MSE	Zisk	Sharopv pomer	R2	Zisk Buy Hold
Priemer	0.08	0.01	27 640.79	0.32	0.65	19 661.97
Štd. odchýlka	0.03	0.01	71 587.5	0.69	0.35	42 939.92
Min. hodnota	0.03	0.00	-65 665.42	-3.11	-0.87	-70 065.45
25%	0.06	0.01	-15 307.07	-0.07	0.62	-8 426.06
50%	0.08	0.01	10 443.27	0.30	0.76	16 252.14
75%	0.09	0.02	52 962.86	0.79	0.89	44 167.48
Max. hodnota	0.21	0.06	381 947.87	1.76	0.96	166 654.84

Tabuľka 4: Výsledky testovacej sady.



Obr. 13: V hornej časti sa nachádza graf znázorňujúci priebeh ceny akcie BIIB počas testovacieho obdobia. V dolnej časti je percentuálny kumulatívny vývoj ziskovosti stratégie podľa predikcie a podľa "Buy & Hold"stratégie, sivý pás znázorňuje obdobie prepadu v dôsledku koronavírusu.

R^2 s hodnotou -0.87. Táto hodnota bola dosiahnutá pre spoločnosť EVERGY (symbol: EVRG), ktorá sa venuje výrobe a distribúcii elektrickej energie v oblasti Kansasu a Missouri. Pri tejto akcii sme mali malý počet dát, a naša sieť nebola schopná dostatočne predikovať varianciu dát. Znázornenie predikcie je na obrázku 14, vidíme veľmi slabú varianciu v predikciách a tiež neschopnosť sieťe reagovať na zmenu trendu.



Obr. 14: Priebeh akcie EVRG počas testovacieho obdobia. Percentuálny vývoj kumulovanej ziskovosti stratégií počas testovacieho obdobia a porovnanie s "Buy & Hold"stratégiou vidíme v dolnej časti.

Zastavme sa na tomto mieste ešte pri zápornom R^2, keďže tento parameter zvyčajne nadobúda hodnoty z intervalu $< 0, 1 >$. V našom prípade počítame R^2 ako:

$$R^2 = 1 - \frac{\sum_i (y_i - F_\Theta(x_i))^2}{\sum_i (y_i - \bar{y})^2}, \tag{50}$$

kde sa čitateľ nazýva reziduálna suma štvorcov a menovateľ celková suma štvorcov. Za základný model sa považuje model v tvare $F_\Theta(x_i) = \bar{y}$, $\forall i$. V takom prípade dostávame $R^2 = 0$, ale v prípade, že reziduálna suma štvorcov je väčšia ako celková suma, tak dostávame $R^2 < 0$. Túto hodnotu interpretujeme tak, že model reprezentuje dáta horšie ako ich priemer [20].

Je prirodzené klásť si otázku, pre aké akcie tento prístup funguje a pre aké nie. Budeme si klásť otázku, či existuje vzťah medzi fundamentálnymi parametrami jednotlivých

spoločností a výkonnosťou našich sietí. Najskôr sa však vysporiadame s odľahlými hodnotami vo fundamentálnych parametroch. Na ich odstránenie použijeme metódu Winsorizingu [32], kde si spočítame hornú a dolnú hranicu:

$$horna = Q_3 + 1.5(Q_3 - Q_1), \tag{51}$$

$$dolna = Q_1 - 1.5(Q_3 - Q_1), \tag{52}$$

kde Q_1 označuje prvý kvartil a Q_3 tretí. Po spočítaní týchto hraníc, všetky hodnoty, ktoré sú väčšie ako horná hranica, nahradíme hornou hranicou. A rovnako všetky hodnoty menšie ako dolná hranica, nahradíme dolnou hranicou.

Po úprave fundamentalných parametrov spoločností sa pozrieme na dosiahnutý Sharpov pomer vzhľadom na jednotlivé parametre. Začneme pri jednoduchých regresiách s jednotlivými parametrami oddelene, pričom zostavíme nasledujúci model, pomocou ktorého sa snažíme zistiť, či vieme výkonnosť siete vysvetliť fundamentálnymi parametrami:

$$S = \beta_0 + \beta_1 \cdot parameter + \epsilon, \tag{53}$$

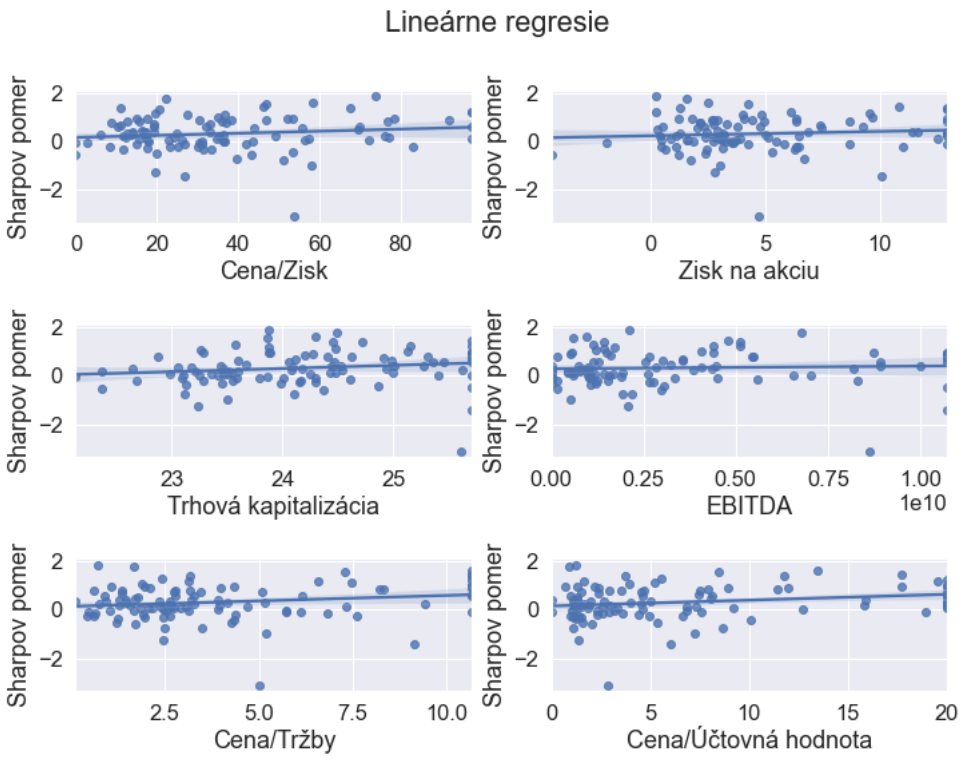
kde S označuje Sharpov pomer. V tabuľke 5 sú uvedené výsledky jednotlivých regresií. Ak sa pozrieme na hodnoty R^2, tak najviac variancie vysvetľuje parameter Cena/Účtovná hodnota. Koeficient pri tomto parametri je aj z pohľadu p-hodnoty považovaný za signifikantný na hladine významnosti 5%. Teda ak hodnota ukazovateľa stúpne o 1, tak očakávaná hodnota predikovaného Sharpovho pomeru stúpne o 0.02. Inými slovami, pre spoločnosti, ktoré sú výrazne drahšie ako je ich čistý majetok, očakávame lepšie hodnoty Sharpovho pomeru.

Parameter, ktorý spomedzi všetkých parametrov priniesol druhé najlepšie R^2, bol parameter Cena/Tržby, ktorý vysvetľuje 4% variance. Taktiež koeficient pri tomto parametri je stále signifikantný na 5% hladine významnosti. V tomto prípade môžeme povedať, že ak sa cena za jednotku tržby zvýši o jedna, tak Sharpov pomer pri stratégii vyplývajúcej z predikcií vzrastie o 0.05.

Pri ostatných parametroch sme dostali p-hodnoty, ktoré hovoria, že tieto parametre nie sú signifikantné na 5% hladine významnosti, a teda velkosť výsledného Sharpovho pomeru od nich nezávisí. Grafické znázornenie jednotlivých regresných nadrovín je na obrázku 15.

Parameter	R^2	β_0	p-hodnota β_0	β_1	p-hodnota β_1
Cena/Zisk	0.02	0.17	0.20	0.00	0.15
Zisk na akciu	0.01	0.24	0.04	0.02	0.33
log(kapitalizácie)	0.03	-2.87	0.12	0.13	0.09
log(EBITDA)	0.01	0.04	0.90	0.01	0.31
Cena/Tržby	0.04	0.14	0.23	0.05	0.05
Cena/ Účtovná hodnota	0.05	0.16	0.13	0.02	0.03

Tabuľka 5: Výsledky lineárnej regresie aplikovanej na jednotlivé parametre samostatne pre Sharpov pomer získaný pri predikovaní cien. Parameter β_0 označuje intercept v modeli a parameter β_1 koeficient pri vysvetľujúcej premennej.



Obr. 15: Grafické znázornenie lineárnych regresií z modelu popísanom v (53).

V prípade modelu, ktorý by pozostával zo všetkých parametrov, ktoré sme použili

jednotlivo a ktorých vyśledky sú v tabuľke 5, dostávame, že na 5 % hladine významnosti nie je žiadny parameter signifikantný. Koeficient determinácie tohto modelu vyšiel 0.07.

Podobne sa pozrieme aj na to, či vieme pre rôzne sektory povedať, ako pre ne bude fungovať náš model. Každá zo spoločností pôsobí v jednom z 11 sektorov (obrázok 8), na základe ktorého zostavíme model:

$$S = \beta_0 + \beta_1 \cdot \text{Materiály} + \beta_2 \cdot \text{Komunikačné služby} + \beta_3 \cdot \text{Spotrebný sektor} + \tag{54}$$

$$\beta_4 \cdot \text{Spotrebný tovar} + \beta_5 \cdot \text{Energetika} + \beta_6 \cdot \text{Finančné služby} + \tag{55}$$

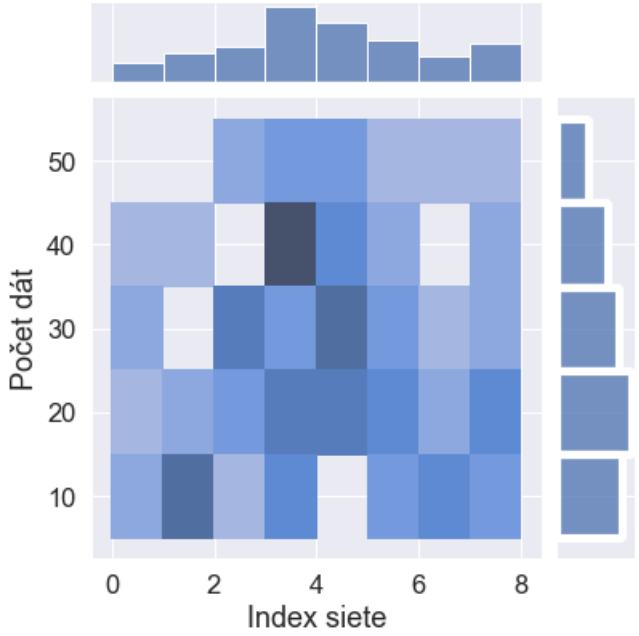
$$\beta_7 \cdot \text{Zdravotná starostlivosť} + \beta_8 \cdot \text{Priemysel} + \beta_9 \cdot \text{Reality} + \tag{56}$$

$$\beta_{10} \cdot \text{Informačné technológie} + \epsilon, \tag{57}$$

kde jednotlivé sektory predstavujú indikátorové premenné, a za základný sektor považujeme Verejné služby. V tomto prípade je s najnižšou p-hodnotou sektor Materiály (p-hodnota = 0.02), kde máme informáciu, že ak je spoločnosť z tohto sektoru, tak môžeme očakávať až o jeden bod horší Sharpov pomer, ako pri Verejných službách. Môžeme tiež povedať, že pre spoločnosti zo sektoru Materiály, nevieme prísť s tak úspešnými stratégiami ako v iných sektoroch. Pri ostatných parametroch dostávame nesignifikantné parametre, čo interpretujeme tak, že sektor, v ktorom spoločnosť pôsobí, nemá zásadný vplyv na výslednú výkonnosť predikcie. Uveďme ešte, že R^2 tohto modelu je 0.15.

Okrem fundamentálnych parametrov, nás zaujíma, aké hyperparametre boli použité na finálnych modeloch. Rozdelenie jednotlivých použitých sietí a počtu použitých dát sa nachádza na obrázku 16. Najpoužívanejšou bola sieť s indexom 3, čo predstavuje sieť so 4 LSTM vrstvami. Pozoruhodné je, že v podkapitole 5.1.1 sme ukázali, že táto sieť nebola na validačnej sade najlepšia spomedzi ostaných sietí. Ak si zoberieme počet minuých dát, tak najviac zastúpené boli modely, ktoré mali na vstupe 20 posledných dní a nasledovali modely s 10 a 30 poslednými dňami. Naopak, zdá sa, že 50 dát je pre takúto sieť už zbytočne veľa, a preto to bol najmenej používaný vstup. Napriek tomu, najčastejšia kombinácia parametrov bola 40 dní z minulosti a sieť zo 4 LSTM vrstiev, kde táto kombinácia bola použitá na 8 sieťach.

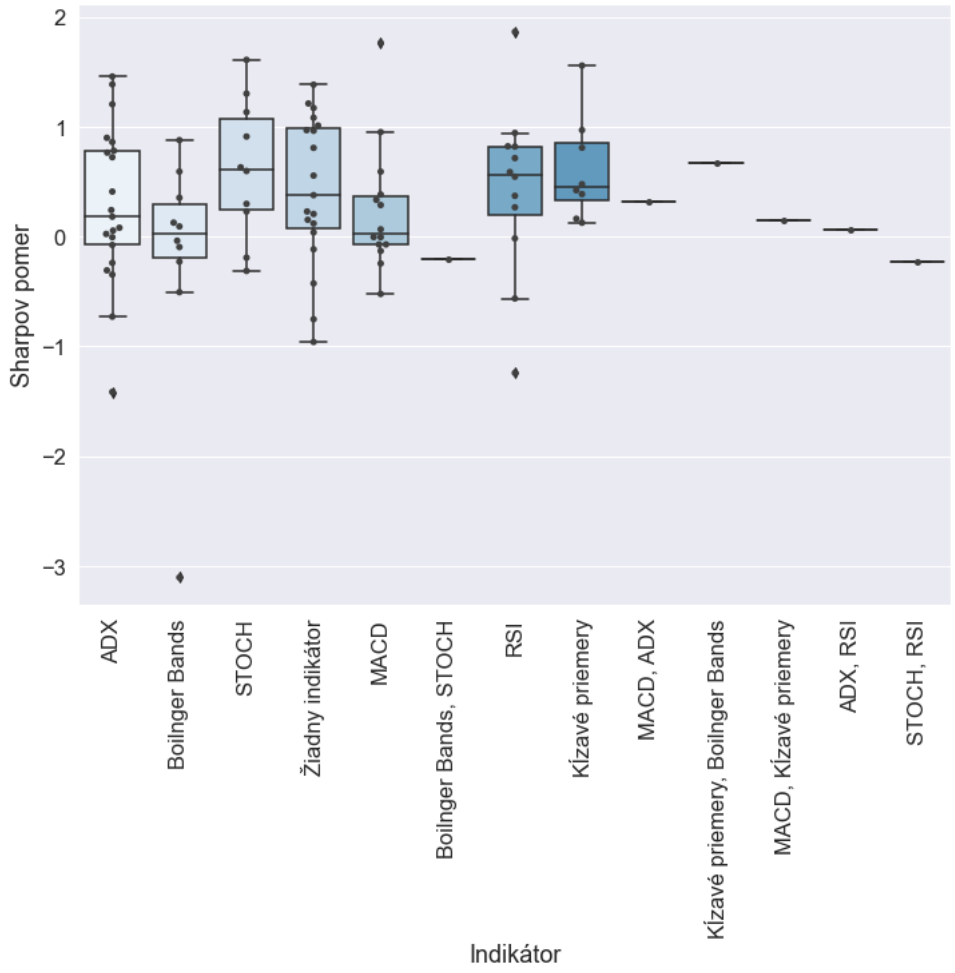
Nemenej dôležitým hyperparametrom sú aj technické indikátory. Krabicový graf na obrázku 17 zobrazuje, aký Sharpov pomer sme dosiahli vzhľadom na jednotlivé použité indikátory. Najpoužívanejšími indikátormi boli ADX, ktorý bol použitý v prípade 21



Obr. 16: Dvojrozmerný histogram so znázornením počtu rôznych komibnácií hyperparametrov, indexu siete a počtu vstupných dát. Tmavšia farba znamená väčší počet.

spoločností a druhým boli modely bez indikátora, kde toto nastavenie bolo použité v 19 prípadoch. Najlepšie hodnoty sme zaznamenali pri indikátore STOCH, kde z 10 prípadov, v ktorých bol použitý, boli stratové iba 2. Zaujímavé výsledky sú aj pri použití kĺzavých priemerov. Tento indikátor nedosiahol najlepšie výsledky, avšak ako jediný, ktorý bol použitý viacnásobne, mal vždy ziskovú stratégiu.

Ako sme uviedli v úvode podkapitoly, našu stratégiu budeme hodnotiť aj vo vzťahu k základnej "Buy & Hold"stratégii. Pri porovnaní počtu akcií s kladným ziskom je na tom lepšie "Buy & Hold", pri ktorej bolo 68 ziskových akcií, oproti 56 ziskovým pri stratégii založenej na predikovaných hodnotách. Grafické porovnanie ziskovosti oboch stratégií sa nachádza na obrázku 18, na ktorom je uvedená aj najväčšia ziskovosť na validačnej sade. Ziskovosť na validačnej sade je v porovnaní s testovacou sadou aj so stratégiou "Buy & Hold"výrazne vyššia. Toto výrazné vychýlenie je spôsobené tým, že za finálnu sieť sa zobrala tá, ktorá mala najväčšiu ziskovosť, a teda bola najviac pozitívne vychýlená. Okrem toho je zrejmé, že stratégia založená na predikciách, vykazuje podstatne viac veľmi ziskových spoločností, než "Buy & Hold". Zdá sa, že akcie, ktoré sú ziskové pri "Buy & Hold"stratégii na testovacej sade, budú ziskové aj pri našej stratégii. Výberový korelačný koeficient medzi týmito stratégiami vyšiel 0.22. Otestujeme, či tento koeficient

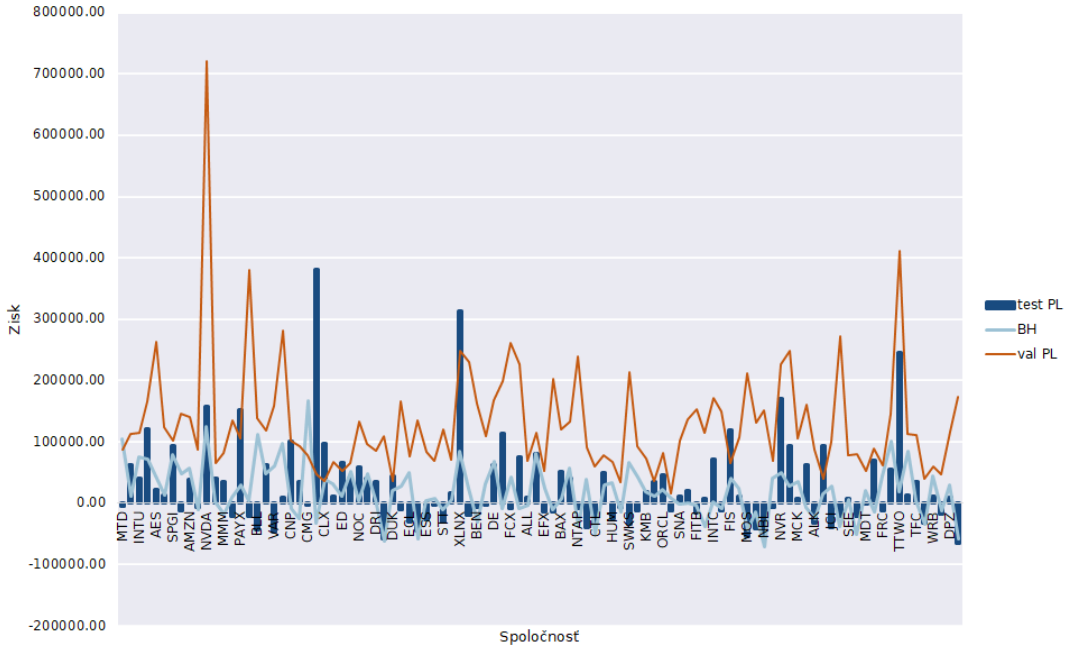


Obr. 17: Krabicový graf znázorňujúci vzťah medzi použitými indikátormi a Sharpovým pomerom.

je štatisticky signifikantne kladný. Zostavíme nasledovnú hypotézu:

$$H_0 : \rho_{PL,BH} \leq 0 \text{ vs. } H_1 : \rho_{PL,BH} > 0. \tag{58}$$

Dostávame p-hodnotu = 0.001, teda na 5% hladine významnosti zamietame H_0, čiže obe stratégie sú štatisticky významne korelované.



Obr. 18: Porovnanie ziskovosti stratégie "Buy & Hold"(BH), stratégie založenej na predikciách z neurónových sietí na testovacej sade (test PL) a validačnej sade (val PL) pre jednotlivé spoločnosti.

5.2 Predikcia výnosov

V tejto podkapitole si predstavíme výsledky, ktoré sme získali pri predikovaní výnosov. Výnosy síce zabezpečujú stacionaritu dát, avšak narážame na teóriu finančnej matetmatiky.

5.2.1 Výsledky na validačnej sade

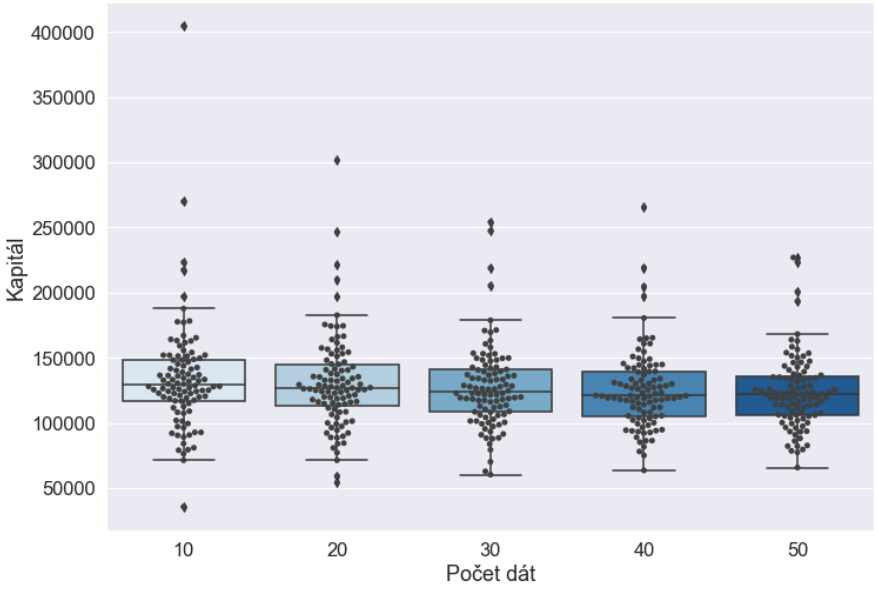
V tejto podkapitole budeme postupovať podobne ako v 5.1.1, kde si postupne predstavíme výsledky na validačnej sade. V prípade výnosov sa ale nezameriame výhradne na krabicové grafy, nakoľko výsledky sú v niektorých prípadoch veľmi podobné. Tabuľka 6 sumarizuje stratu na validačnej sade (strednú absolútnu chybu) podľa jednotlivých sietí. Všetky siete dosiahli takmer totožné priemerné hodnoty validačnej straty. Siete sa tiež zhodujú v rozptyle dát, kde nadobúdajú takmer identické hodnoty štandardných odchýlok. Tiež maximálna a minimálna hodnota naznačujú, že aj extrémne hodnoty sú medzi sieťami rovnako rozdelené. Tieto výsledky smerujú k záveru, že zložitejšie siete sa z dát nenaučia komplikovanejšie vzory.

Index	Priemerná hodnota	Maximálna hodnota	Minimálna hodnota	Štandardná odchýlka
0	1.063	4.236	0.615	0.441
1	1.060	4.215	0.614	0.439
2	1.060	4.210	0.613	0.439
3	1.059	4.211	0.612	0.439
4	1.058	4.342	0.611	0.446
5	1.061	4.221	0.614	0.440
6	1.060	4.210	0.613	0.439
7	1.064	4.238	0.618	0.442
8	1.057	4.224	0.613	0.439

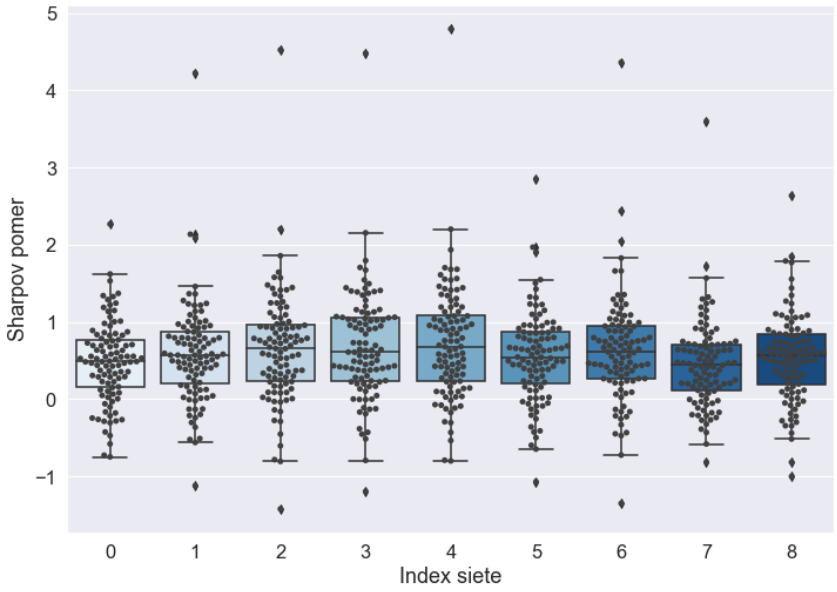
Tabuľka 6: Hodnoty validačnej straty vzhľadom na použitú sieť.

Zoberme si teraz nadobudnutý kapitál vzhľadom na počet dát. Na krabicovom grafe na obrázku 19 vidíme, že sieť so vstupom 10 minulých dní vykazuje najväčšiu priemernú hodnotu nadobudnutého kapitálu (priemerný zisk pre 10 minulých dní na vstupe 134 078$). Tento je o takmer 4 000$ väčší ako v prípade druhej najlepšej siete, so vstupom 20 minulých dní (priemerný zisk pre 20 minulých dní na vstupe 130 732$). Na druhej strane, pri 10 dňoch na vstupe je nadobudnutý kapitál najviac rozptýlený.

Prejdeme na kľúčovú metriku, Sharpov pomer. Opäť sa na prvé pozrieme na Sharpov pomer v závislosti od použitej siete. Na obrázku 20 je znázornený krabicový graf. Vidíme, že priemerná hodnota ani v jednej sieti nedosiahla Sharpov pomer väčší ako 1. Najlepší priemerný Sharpov pomer dosiahla sieť s indexom 4 (5 LSTM vrstiev) s hodnotou 0.71. V tejto sieti sa nachádza aj hodnota s najväčším dosiahnutným Sharpovým pomerom 4.79. Okrem toho si môžeme všimnúť, že medzikvartilový rozsah jednoduchších sietí (indexy 1, 5 a 7) je výrazne menší v porovnaní s ostatnými, zložitejším sieťam. Najhoršie výsledky sme dosiahli na sieti s indexom 7 (jedna GRU vrstva). Kde táto sieť dosiahla priemerný Sharpov pomer 0.45 a iba v 11 prípadoch bol Sharpov pomer väčší ako 1.



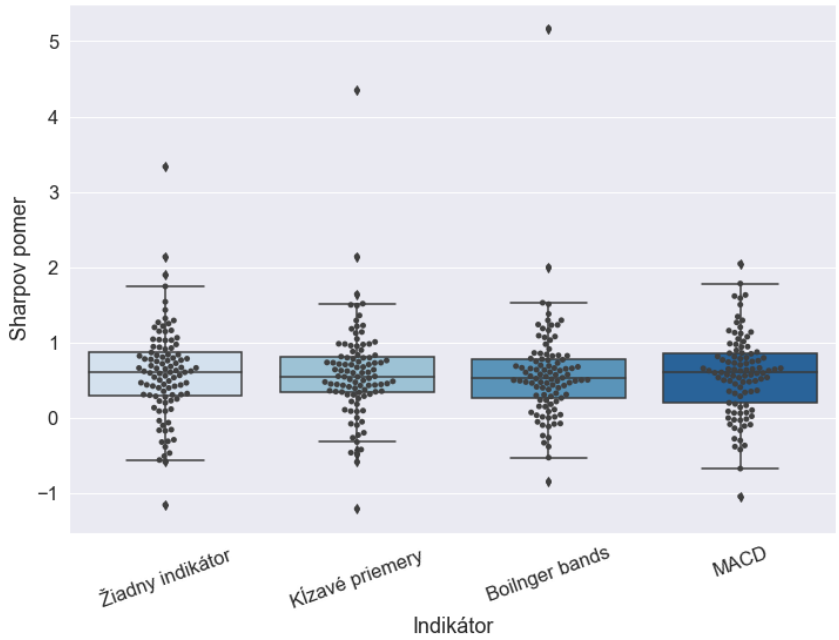
Obr. 19: Krabicový graf so znázornením nadobunutuého kapitálu a počtu dní, ktoré do siete vstupovali.



Obr. 20: Krabicový graf znázorňujúci Sharpov pomer vzhľadom na použitú sieť.

Na validačnej sade nás zaujíma aj vplyv technických indikátorov ako hyperparametra. Pri predikovaní výnosov sme použili iba tri technické parametre, ktoré boli z rovnakého intervalu ako naše dáta, keďže tie sme nenormovali. Obrázok 21 znázorňuje Sharpov pomer, ktorý sme dosiahli pri použití jednotlivých technických indikátorov. Modely s pridaným technickým indikátorom dosahovali takmer totožné výsledky v priemernom dosiahnutom

Sharpovom pomere aj v medzikvartilovom rozsahu ako modely bez indikátora. Tiež po pridaní indikátorov nezačali pribúdať extrémne hodnoty Sharpovho pomeru. To môžeme interpretovať tak, že sieť sa z pôvodných dát naučila všetko, čo potrebovala a ďalšie transformácie parametrov na vstupe neboli potrebné.



Obr. 21: Krabicový graf zobrazujúci vzťah medzi nadobudnutým Sharpovým pomerom a použitým indikátorom.

Porovnajme teraz tieto dva prístupy a ich výsledky na validačnej sade. Zásadným rozdielom je rozptýlenosť výsledkov. Výrazné rozdiely sú zrejmé pri porovnaní obrázku 9 a tabuľky 6, kde v prípade predikcie výnosov sme dosiahli takmer identické validačné straty pre rôzne siete. Na druhej strane, v prípade predikcií cien, bolo zrejmé, že jednoduchšie siete výrazne prekonali komplikovanejšie siete.

Pri porovnaní krabicových grafov zobrazujúcich nadobudnutý kapitál vzhľadom na počet vstupných dát, obrázky 10 a 19, tak vidíme, že priemerný dosiahnutý kapitál je podobný vo všetkých prípadoch. Zásadný rozdiel, ktorý je medzi týmito dvoma prístupmi, je v rozsahu nadobudnutého medzikvartilového rozsahu. Pri predikovaní cien máme výrazne viac rozptýlený nadobudnutý kapitál ako pri výnosoch.

Čo sa týka Sharpovho pomeru, tak v tomto prípade máme v oboch prípadoch rozptýlené dáta, ale kým v prípade predikcií výnosov máme výrazne viac pozitívne vychýlených hodnôt, tak pri cenách máme tieto hodnoty viac vychýlené do záporných extrémov. Čo

je ale v rozpore s tým, aké validačné straty sme dosahovali v týchto prípadoch. Jedno z vysvetlení je, že hoci pri výnosoch máme väčšiu validačnú chybu, lepšie vieme predikovať pohyb a tým pádom odhadnúť smer.

V oboch prípadoch sme pri použití technických indikátorov nezaznamenali výrazné zvýšenie Sharpovho pomeru. Iba v prípade použitia niektorých z nich, ako napríklad STOCH, pri cenách a Boilnger Bands pri výnosoch, sme zaznamenali viac pozitívnych odľahlých hodnôt. To môžeme interpretovať tak, že naša sieť sa vedela tieto vzory naučiť zo základných dát na vstupe, a teda ďalšia transformácia týchto dát bola redundantná.

5.2.2 Výsledky na testovacej sade

V tejto časti sa pozrieme na výsledky, ktoré sme dosiahli na testovacej sade pri predikovaní výnosov. Postupovať budeme ako v podkapitole 5.1.2, kde si prejdeme výsledky a pozrieme sa na to, čo funguje, a čo nie. V tabuľke 7 sa nachádza popis rozdelenia sledovaných metrík. Priemerný zisk našej stratégie sa aj v tomto prípade javí výrazne väčší, ako v prípade "Buy & Hold"stratégie. Kde pri stratégií založenej na predikciách sme v priemere na jednu akciu zarobili 33 523.10 $ oproti 19 661.97 $ pri "Buy & Hold". Signifikantnosť tohto rozdielu otestujeme pomocou párového Studentovho t-testu na 5 % hladine významnosti. Nulovú hypotézu formulujeme nasledovne: profit v "Buy & Hold"stratégii je väčší, alebo rovný ako profit v našej stratégii. Alternatívnu hypotézu naformulujeme opačne: zisk v "Buy & Hold"stratégii je menší, ako v našej:

$$H_0\text{: } \mu_{zisk} \leq \mu_{Buy\&Hold} \text{ vs } H_1\text{: } \mu_{zisk} > \mu_{Buy\&Hold}. \tag{59}$$

Testovacia štatistika v našom teste vyšla $T = 1.91$, ktorú opäť porovnávame so Studentovým rozdelením s 99 stupňami voľnosti. Dostaneme p-hodnotu rovnú $0.029 < 0.05$, a zamietame H_0. V tomto prípade môžeme povedať, že priemerný rozdiel v ziskovosti jednotlivých stratégií je štatisticky signifikantný, inými slovami, naša stratégia prináša na danej vzorke signifikantne väčší priemerný výnos, ako "Buy & Hold"stratégia.

Aj v tomto prípade sa pozrieme na niektoré príliš vysoké alebo nízke charakteristiky z tabuľky 7. Najväčšiu hodnotu Sharpovho pomeru (2) a zisku (430 546.56 $) sme dosiahli pri spoločnosti Intel (symbol: INTC), ktorej hlavnou činnosťou je výroba procesorov. Intel je typickým príkladom moat spoločnosti, keď si dlhodobo udržiaval až monopolné posta-

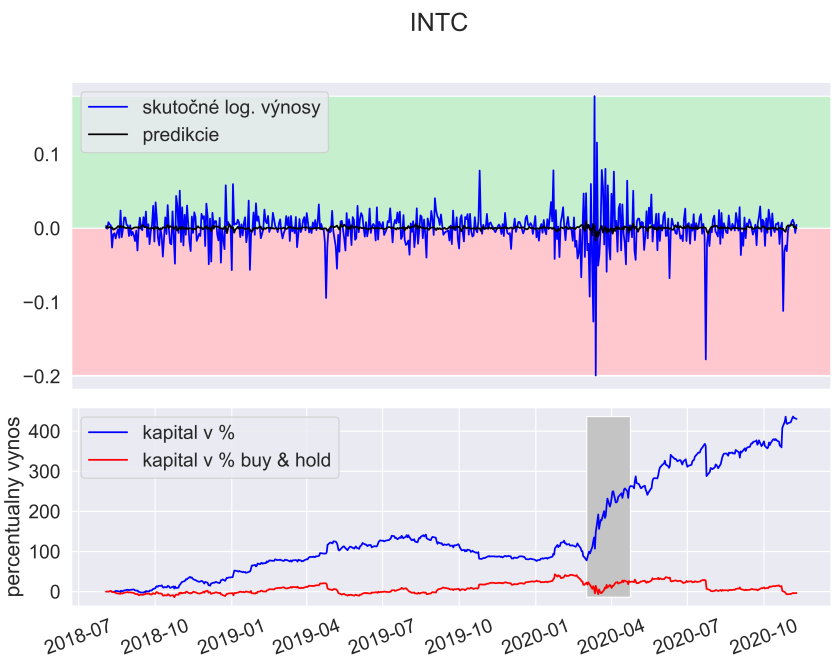
	Strata	MSE	Zisk	Sharpov pomer	R2	Zisk Buy & hold
Priemer	1.74	8.30	33 523.10	0.39	-0.004	19 661.97
Štd. odchýlka	0.55	6.92	76 009.22	0.68	0.026	42 939.92
Min. hodnota	1.04	2.40	-82 803.74	-1.88	-0.243	-70 065.45
25%	1.38	4.73	-16 970.02	-0.05	-0.004	-8 426.06
50%	1.62	6.28	19 163.15	0.43	-0.001	16 252.14
75%	1.94	8.71	71 260.48	0.86	0.002	44 167.48
Max. hodnota	4.71	48.90	430 546.56	2.00	0.031	166 654.84

Tabuľka 7: Výsledky testovacej sady.

venie v produkcii procesorov. Avšak toto postavenie v poslednom období začali ohrozovať ďalší producenti, ktorí boli schopní vyvinúť konkurencieschopné produkty. Pod najväčší tlak sa však Intel dostáva od druhej polovice roka 2020, keď Apple oznámil prechod k procesorom od spoločnosti ARM. To vyústilo až k prepadu ceny o takmer 25 % koncom júla roku 2020. Vývoj stratégie a znázornenie predikcií pre túto akciu sa nachádza na obrázku 22. Na tomto obrázku si môžme všimnúť aj to, akú malú časť variancie výnosov, bola naša sieť schopná modelovať.

Vidíme, že najväčší nárast kapitálu nastal v období prudkého prepadu v dôsledku korony. Naša sieť výkyvy počas tohto obdobia dokázala pomerne dobre predikovať, čo v kombinácii s extrémnymi pohybmi viedlo k veľkému nárastu. Na druhú stranu, ak sa pozrieme na prepad spomínaný v dôsledku prechodu Applu na procesory ARM (výrazný prepad v období od 2020-07), tak toto naša sieť nedokázala správne predikovať. Poznamenajme však, že ide o externý faktor, ktorý naša sieť, ani ostatní investori dopredu nevedeli (inak by sa jednalo o takzvaný Inside job), ak by sieť tento prepad správne predikovala, išlo by s najväčšou pravdepodobnosťou o náhodu.

Najmenší Sharpov pomer a zároveň aj najväčšiu stratu na testovacej sade mala spoločnosť Eastman Chemical Company (symbol: EMN). Spoločnosť sa venuje výrobe chemikálií pre ďalšie spracovanie a výrobu vláken. Priebeh vývoja kapitálu počas stratégie je na obrázku



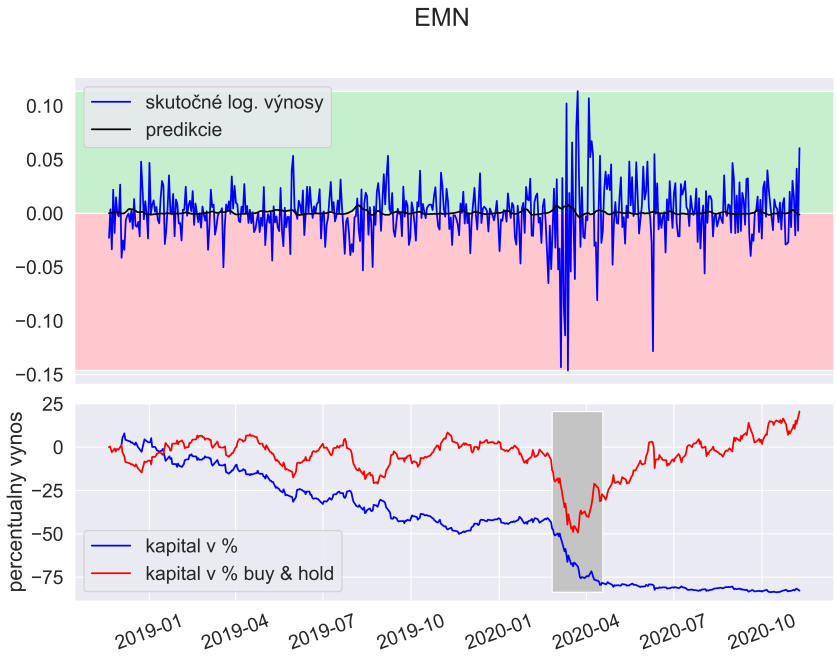
Obr. 22: V hornej časti sa nachádza graf znázorňujúci priebeh ceny akcie INTC počas testovacieho obdobia. V dolnej časti je percentuálny kumulovaný vývoj ziskovosti stratégie podľa predikcie a podľa "Buy & Hold"stratégie, sivý pás znázorňuje obdobie prepadu v dôsledku koronavírusu.

23, môžeme si všimnúť, že už v období pred korona krízou si naša sieť počínala zle a nebola schopná správne predikovať pohyby. Počas koronového obdobia sa vývoj kapitálu prepadol do ešte väčšej straty. Môžeme povedať, že v prípade tejto akcie naša sieť nebola schopná nájsť v dátach žiadnu závislosť, ktorá by nám umožnila predikovať budúce výnosy.

Pokúsime sa aj teraz vysvetliť výkonnosť siete pomocou windsorizovaných fundamentálnych parametrov. Zostavíme pre každý parameter z tabuľky 8 nasledujúci model:

$$S = \beta_0 + \beta_1 \cdot parameter + \epsilon, \tag{60}$$

kde S opäť označuje nadobudnutý Sharpov pomer na testovacej sade. Koeficienty determinácie (R^2) a odhady jednotlivých parametrov s p-hodnotami pre jednotlivé modely sú uvedené v tabuľke 8. Model s najväčším koeficientom determinácie a s najmenšou p-hodnotou je model zostavený pomocou logaritmu trhovej kapitalizácie spoločnosti. Týmto modelom sa nám podarilo vysvetliť až 11 % variancie vo výslednom Sharpovom pomere.



Obr. 23: V hornej časti sa nachádza graf znázorňujúci priebeh ceny akcie EMN počas testovacieho obdobia. V dolnej časti je percentuálny kumulovaný vývoj ziskovosti stratégie podľa predikcie a podľa "Buy & Hold"stratégie, sivý pás znázorňuje obdobie prepadu v dôsledku koronavírusu.

V tomto modeli dostávame hodnotu parametra β_1 rovnú 0.24, čo interpretujeme tak, že ak hodnota spoločnosti vzrastie o 1 %, očakávame nárast Sharpovho pomeru o 0.0024. Môžeme tiež očakávať, že pri drahších firmách dosiahneme väčšie hodnoty Sharpovho pomeru.

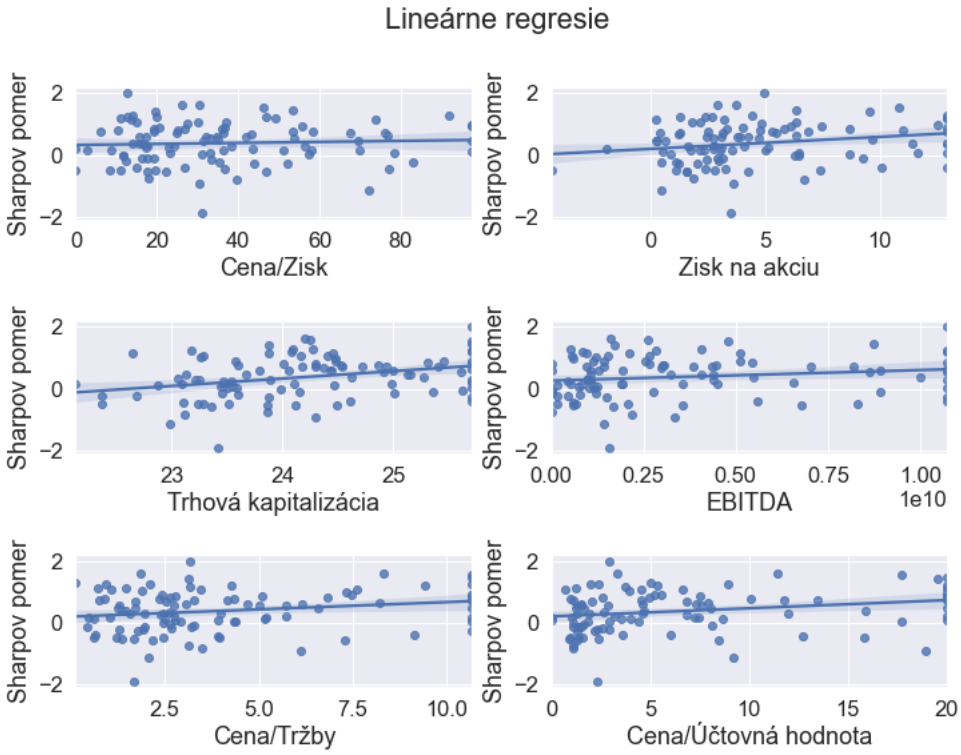
Cena/Účtovná hodnota bol parameter, ktorý dosiahol druhú najnižšiu p-hodnotu a tiež tento model dosiahol druhé najvyššie R^2. V tomto prípade môžeme interpretovať výsledok tak, že ak pomer ceny k majetku spoločnosti vzrastie o 1, tak Sharpov pomer sa zvýši o 0.03. Aj v tomto prípade môžeme podobne povedať, že očakávame väčší Sharpov pomer pre spoločnosti, ktoré sú drahšie, ako je hodnota ich účtovného majetku.

Na 5 % hladine významnosti boli ešte signifikantné koeficienty β_1 pri parametroch Zisk na akciu a Cena/Tržby. V oboch prípadoch koeficient determinácie nadobudol hodnotu 0.05, teda oba vysvetlili 5 % variancie. Pri oboch parametroch sme dostali kladné hodnoty pre β_1, čiže zvýšenie týchto parametrov prinesie vyšší Sharpov pomer.

V prípade regresie, kde na vstup dáme všetky uvažované parametre z regresií z tabuľky

Parameter	R^2	β_0	p-hodnota β_0	β_1	p-hodnota β_1
Cena/Zisk	0.00	0.33	0.01	0.00	0.56
Zisk na akciu	0.05	0.21	0.05	0.04	0.02
log(kapitalizácie)	0.11	-5.39	0.00	0.24	0.00
log(EBITDA)	0.02	0.00	0.99	0.02	0.14
Cena/Tržby	0.05	0.20	0.07	0.05	0.03
Cena/ Účtovná hodnota	0.06	0.21	0.03	0.03	0.01

Tabuľka 8: Výsledky lineárnej regresie aplikovanej na jednotlivé parametre samostatne pre Sharpov pomer získaný pri predikovaní výnosov. Parameter β_0 označuje intercept v modeli a parameter β_1 koeficient pri vysvetľujúcej premennej.



Obr. 24: Grafické znázornenie lineárnych regresií z modelu popísanom v (60).

8, dostávame na 5 % hladine významnosti štatisticky signifikantný koeficient logaritmus trhovej kapitalizácie. Ostatné koeficienty na tejto hladine vychádzajú ako nesignifikantné.

Koeficient determinácie tohto modelu vyšiel na úrovni 0.16.

Aj pri predikcií výnosov sa pozrieme na to, či sektor, v ktorom spoločnosť pôsobí, vie vysvetliť, aký Sharpov pomer môžeme očakávať. Zostavíme model s indikátorovými parametrami, ktoré predstavujú sektor, v ktorom spoločnosť pôsobí:

$$S = \beta_0 + \beta_1 \cdot \text{Materiály} + \beta_2 \cdot \text{Komunukačné služby} + \beta_3 \cdot \text{Spotrebný sektor} + \tag{61}$$

$$\beta_4 \cdot \text{Spotrebný tovar} + \beta_5 \cdot \text{Energetika} + \beta_6 \cdot \text{Finančné služby} + \tag{62}$$

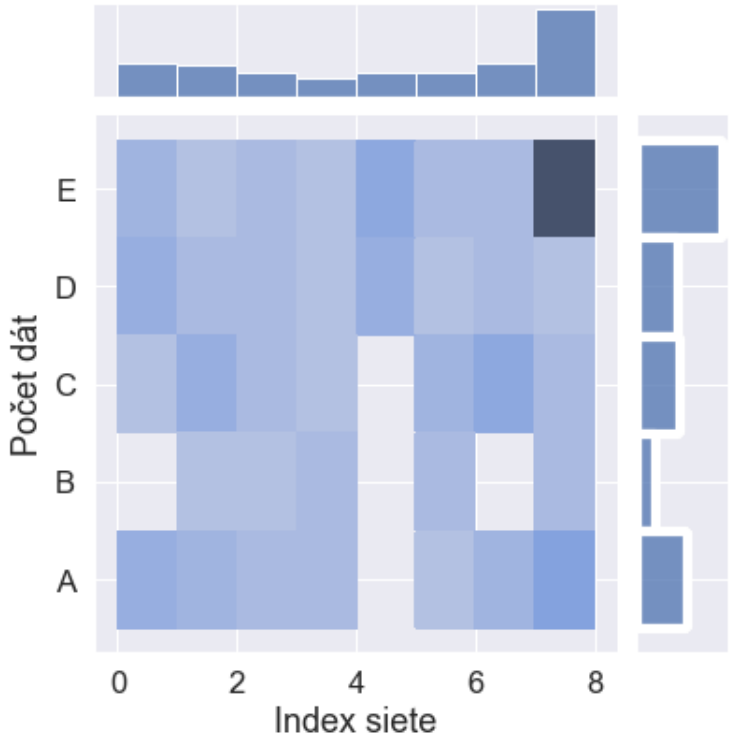
$$\beta_7 \cdot \text{Zdravotná starostlivosť} + \beta_8 \cdot \text{Priemysel} + \beta_9 \cdot \text{Reality} + \tag{63}$$

$$\beta_{10} \cdot \text{Informačné technológie} + \epsilon, \tag{64}$$

v tomto prípade za základný sektor považujeme Verejné služby. Pri tomto modeli dostávame iba jeden signifikantný koeficient na hladine významnosti 5 %, a to parameter, či daná spoločnosť pôsobí v sektore Realít. Dostávame tak hodnotu $\beta_9 = -0.86$ a príslušnú p-hodnotu 0.03. Môžeme teda povedať, že pre spoločnosti pôsobiace v sektore Realít prichádzame s menej úspešnými stratégiami, ako pre spoločnosti pôsobiace v iných sektoroch. Podobne ako pri predikciách cien, môžeme povedať, že sektor, v ktorom spoločnosť pôsobí, nemá zásadný vplyv na výsledný Sharpov pomer. Poznamenajme, že koeficient determinácie tohto modelu bol 0.17.

Pozrieme sa aj na hyperparametre, ktoré boli použité na testovacej sade. Histogramy početnosti použitia jednotlivých sietí a počtu vstupných dát sa nachádzajú na obrázku 25. Vzhľadom na rovnomerné rozloženie výsledkov na validačnej sade očakávame podobné rovnomerné rozloženie. Vidíme však, že kombinácia so vstupom 50 minulých dní a s indexom siete 8, bola použitá výrazne viac, ako ostatné kombinácie. Konkrétne takéto nastavenie bolo použité v 18 prípadoch. Použitie ostatných kombinácií bolo v súlade s naším očakávaním. Zaujímavé je tiež, že na validačnej sade boli siete so vstupom 20 minulých dní druhé najziskovejšie (viď. obrázok 19), ale na testovacej sade boli najmenej používané. Najpoužívanejšími sa tak stali siete, ktoré boli z testovaných najzložitejšie, pričom na validačnej sade tomu nič nenasvedčovalo.

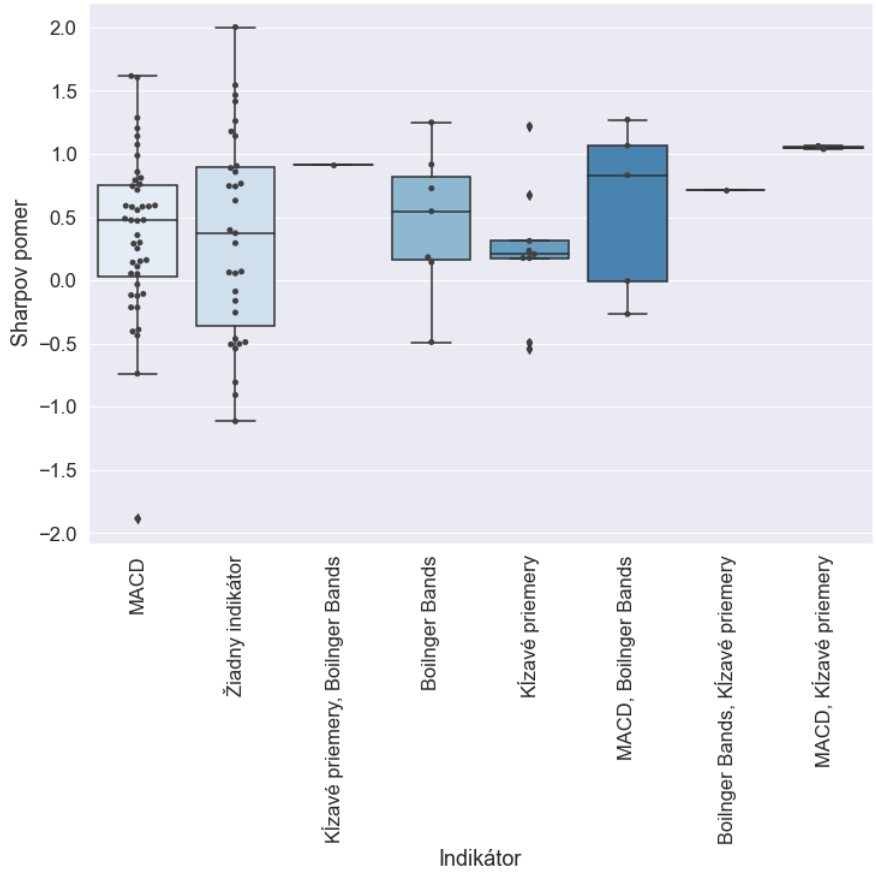
Krabicový graf na obrázku 26 zobrazuje vzťah medzi použitými technickými indikátormi a nadobudnutým Sharpovým pomerom na testovacej sade. Výrazne najpoužívanejším indikátorom bol MACD, ktorý bol použitý pre 44 spoločností. Poznamenajme, že na va-



Obr. 25: Dvojrozmerný histogram so znázornením počtu rôznych komibnácií hyperparametrov indexu siete a počtu vstupných dát. Tmavšia farba znamená väčší počet.

lidačnej sade sme po pridaní technických indikátorov k základnému modelu nezaznamenali výrazne zlepšenie v dosiahnutom Sharpovom pomere. Základný model bez indikátora bol na testovacej sade použitý ako druhý najviac používaný a použitý bol v prípade 31 spoločností.

Aj pre výnosy provnáme stratégiu založenú na predikciách neurónových sietí a základnú "Buy & Hold"stratégiu. Ako sme uviedli v podkapitole 5.1.2, v prípade stratégie "Buy & Hold"máme 68 akcií s kladným ziskom na testovacej sade. Pre stratégiu založenú na predikovaní výnosov máme 64 spoločností, pri ktorých sme dosiahli kladný zisk. Grafické porovnanie týchto stratégií sa nachádza na obrázku 27. Aj v tomto prípade môžeme vidieť viac extrémne ziskových spoločností pri stratégii založenej na predikciách neurónových sietí, než pri "Buy & Hold"stratégii. Aj v tomto prípade nás bude zaujímať, či existuje korelácia medzi spoločnosťami, pri ktorých sme dosiahli na testovacej sade zisk v "Buy & Hold"stratégii a pri ktorých pri stratégii založenej na predikciách. Výberový korelačný koeficient dostávame 0.37, a štatistickú signifikantnosť tohto koeficientu otestujeme pomocou nasledovnej hypotézy:



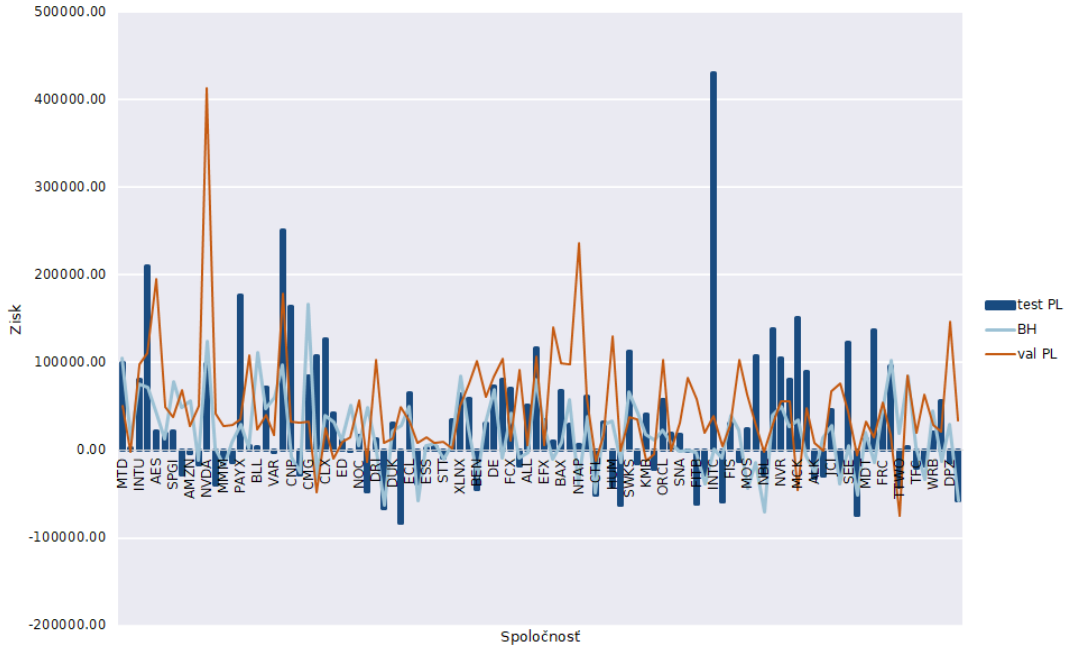
Obr. 26: Krabicový graf znázorňujúci vzťah medzi použitými indikátormi a Sharpovým pomerom.

$$H_0 : \rho_{PL,BH} \leq 0 \text{ vs. } H_1 : \rho_{PL,BH} > 0. \tag{65}$$

Dostávame p-hodnotu limitne blízku nule, teda na 5% hladine významnosti zamietame H_0, a môžeme tvrdiť, že obe stratégie sú štatisticky významne korelované.

Pri porovnaní výsledkov na testovacej sade, ktoré sme dosiahli modelovaním výnosov s výsledkami, pri modelovaní cien, je najzaujímavejšie, že hoci predikovanie cien vykazuje oveľa lepšie hodnoty stratových funkcií a koeficientov determinácie, tak na finančných stratégiách bolo predikovanie výnosov úspešnejšie. Zreteľne bol tento rozdiel viditeľný pri testovaní, či priemerný profit v našich stratégiách je väčší, ako profit v "Buy & Hold"stratégii. V tomto teste sa potvrdilo, že priemerný zisk pri stratégii založenej na predikciách výnosov je signifikantne väčší, ako v prípade "Buy & Hold"stratégie, avšak nepotvrdilo sa nám to v prípade predikcie cien a "Buy & Hold".

Aj pri vysvetľovaní nadobudnutého Sharpovho pomeru pomocou fundamentálnych



Obr. 27: Porovnanie ziskovosti stratégie "Buy & Hold"(BH), stratégie založenej na predikciách na testovacej sade (test PL) a validačnej sade (val PL) pre jednotlivé spoločnosti.

parametrov, sme vedeli získať väčšiu informáciu v prípade modelovania výnosov. V tomto prípade sme vedeli vysvetliť 11 % variancie Sharpovho pomeru na testovacej sade pomocou trhovej hodnoty spoločnosti, kde pri modelovaní cien, sme tento parameter dostali ako nesignifikantný na 5 % hladine významnosti. Pri modelovaní cien sa ukázal ako parameter s najnižšou p-hodnotou a najvyšším koeficientom determinácie Cena/Účtovná hodnota. Ale aj tento parameter pri modelovaní výnosov mal menšiu p-hodnotu a vyšší koeficient determinácie.

Avšak ako v prípade predikcií cien, ako aj v prípade predikcií výnosov, sme ukázali, že ak je pre spoločnosť zisková "Buy & Hold"stratégia, tak môžeme očakávať, že pre ňu budú ziskové aj stratégie založené na predikciách, a to bez ohľadu na predikovaný výstup.

6 Výsledky na hodinových dátach

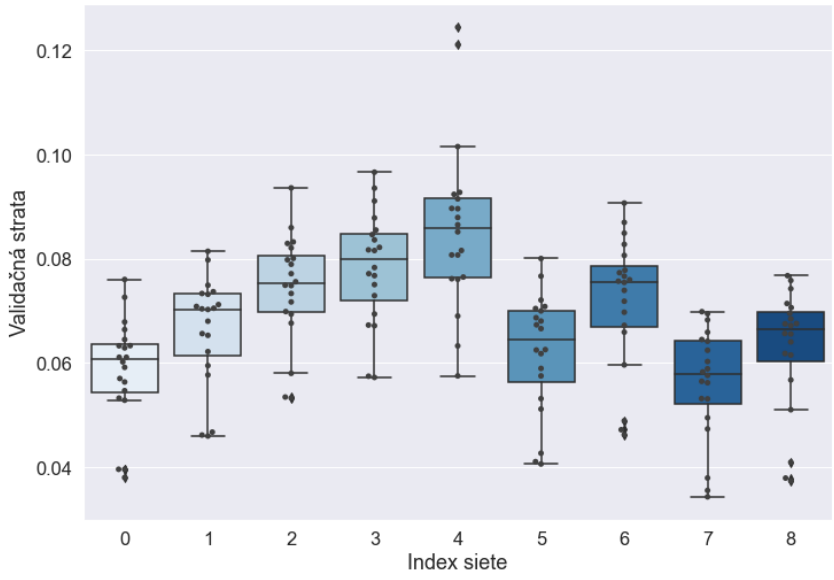
V tejto kapitole sa opäť pozrieme na výsledky ako v kapitole 5, avšak zameriame sa na hodinové výsledky. Kapitola je členená rovnako, ako prechádzajúca.

6.1 Predikcia cien

Aj v tejto časti, si ako prvé uvedieme dosiahnuté výsledky pri predikovaní cien.

6.1.1 Výsledky na validačnej sade

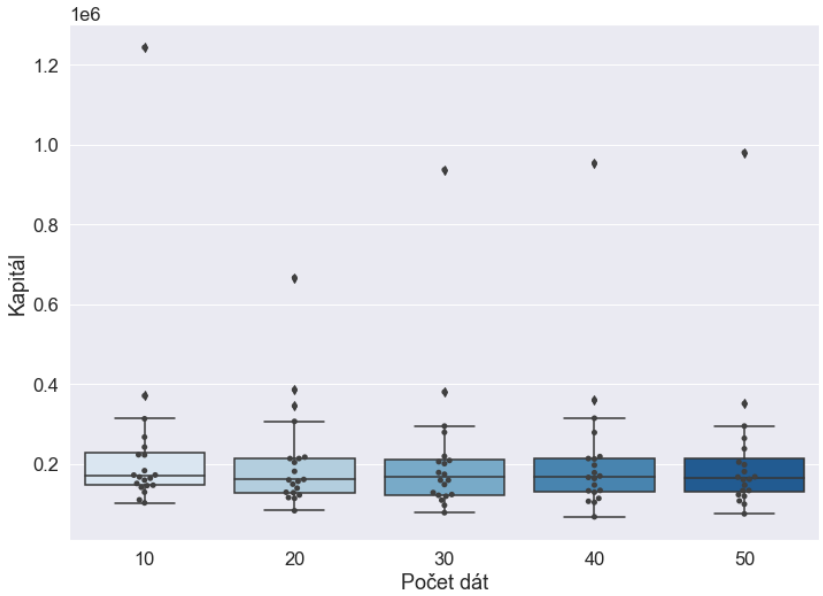
Pozrieme sa na výsledky, ktoré naša sieť dosiahla na validačnej sade, ukážeme, ako siete fungujú alebo nefungujú vzhľadom na testované hyperparamtre. Zoberme validačnú stratu, pomocou ktorej hľadáme najlepšie váhy v sieti, vzhľadom na index siete, ako je znázornené na obrázku 28. Možeme si všimnúť, že jednoduchšie siete, s menším počtom vrstiev, fungujú lepšie, s menšou validačnou stratou, ako komplikovanejšie siete. To môže byť spôsobené preučením komplikovanejších sietí na trénovacej sade. Zaujímavé je aj to, že jednoduchšie siete GRU fungujú lepšie, ako LSTM siete.



Obr. 28: Krabicový graf so znázornením validačnej straty - MAE, vzhľadom na index siete.

Nadobunutý kaptiál na validačnej sade rozdelený podľa počtu dní, ktoré do siete vstúpili je znázornený na obrázku 29. Najzaujívajšia je prítomnosť extémne ziskových odľahlých hodnôt, pre všetky testované počty dát na vstupe. Tiež na validačnej sade sme

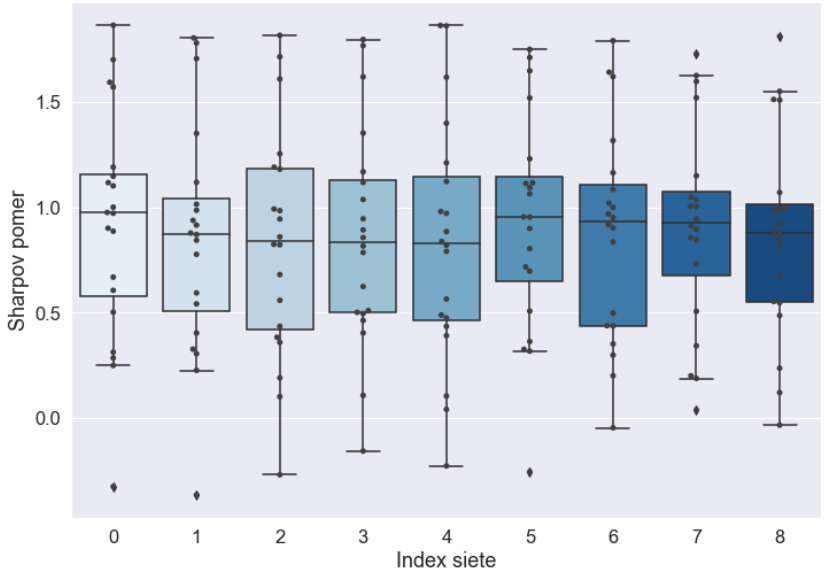
dosahovali pre všetky siete v priemere zisk. Tieto pozitívne výsledky sú spôsobené tým, že validačná sada obsahuje obdobie prepadu a zotavenia v dôsledku koronavírusu.



Obr. 29: Krabicový graf so znázornením nadobudnutého kapitálu a počtu dní, ktoré do siete vstupovali.

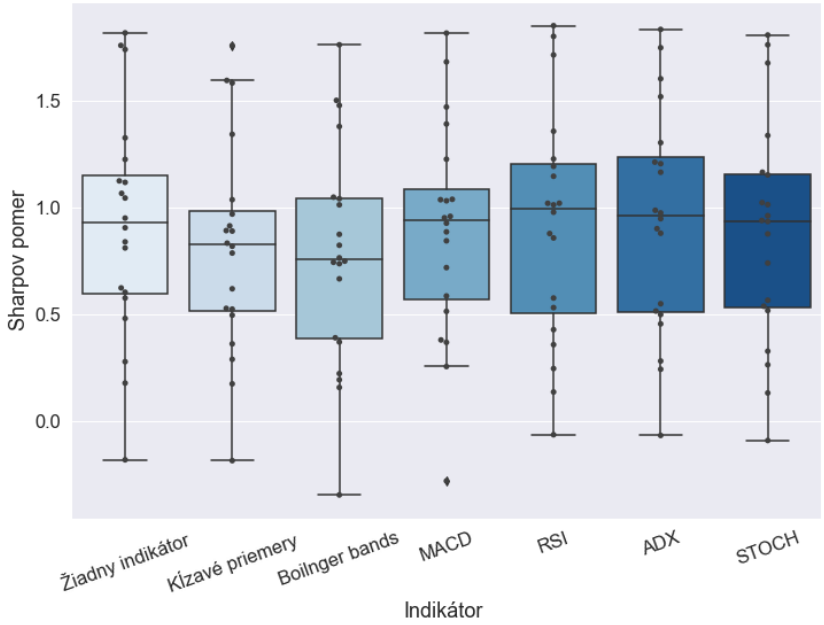
Na validačnej sade nás ale najviac zaujímajú finančné výsledky, na základe ktorých vyberieme konkrétnu sieť. Pozrieme sa na Sharpov pomer, ktorý poskytuje informáciu nielen o zhodnotení investície, ale aj o jej rizikovosti. Opäť sa na prvé pozrieme na výsledky podľa indexu siete, ktoré máme na obrázku 30. V prípade Sharpovho pomeru nevidno výrazné rozdiely medzi jednotlivými sieťami, kde priemerné hodnoty Sharpovho pomeru sa pohybujú medzi hodnotami 0.9 až 1. Zaujímavé je, že horné fúzy LSTM sietí (indexy 0 až 6) dosahujú viditeľne lepšie hodnoty, ako GRU siete, čo je v kontraste s obrázkom 28. Sharpov pomer podľa počtu dát je rovnomerne rozdistribuovaný medzi jednotlivé hodnoty minulých dát.

V prípade indikátorov sa aj v tomto prípade pozrieme len na základné indikátory, ktoré sme používali. Na obrázku 31 máme vzťah medzi jednotlivými indikátormi a Sharpovým pomerom. Za základnú hodnotu, benchmark, považujeme sieť bez pridaného indikátora. Môžeme si všimnúť, že kĺzavé priemery a Bolinger bands výsledky siete nevylepšili. MACD indikátor tiež zásadne nezlepšil priemerný, ale ani najlepší dosiahnutý Sharpov pomer, však výrazne zvýšil hodnotu spodného chvostu do kladných hodnôt, čo možno interpretovať tak, že pri použití tohto inidkátora nedošlo k stratovým stratégiám. Posledné



Obr. 30: Krabicový graf zobrazujúci Sharpov pomer vzhľadom na použitú sieť.

tri indikátory mierne vylepšili Sharpeho pomer na validačnej sade aj zvýšenie spodného chvosta.



Obr. 31: Boxplot so znázornením validačnej straty - MAE, vzhľadom na použitý indikátor.

6.1.2 Výsledky na testovacej sade

Keďže na validačnej sade sme pri hodinových dátach a predikciách cien dosahovali pomerne dobré výsledky, pozrime sa, ako budú naše siete fungovať na testovacej sade. Cha-

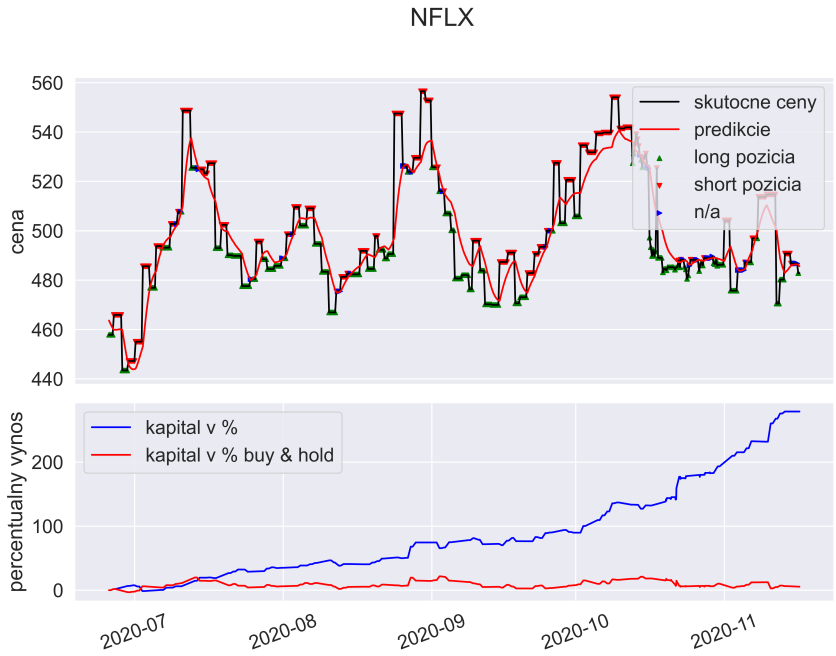
rakteristiky jednotlivých sledovaných parametrov, aké sme dosiahli na testovacej sade, sa nachádzajú v tabuľke 9. Naša stratégia, založená na predikciách pomocou neurónových sietí, výrazne prekonáva "Buy & Hold"stratégiu (p-hodnota v Studentovom t-teste analogickom k testom (49) a (59) vyšla rovná numerickej nule a teda na 5 % hladine významnosti zamietame H_0). Ďalej vidíme veľmi vysoké hodnoty aj v prípade koeficientu determinácie (R^2), kde mediánová hodnota vyšla 0.92, čiže pri polovici z vybraných akcií sa nám podarilo popísať viac ako 92 % variancie. Hodnotami, ktoré najviac zaujmú, sú nadobudnuté zisky pri našej stratégii.

	Strata	**MSE**	**Zisk**	**Sharpov pomer**	**R2**	**Zisk Buy & hold**
Priemer	0.06	0.01	103 807.50	1.82	0.88	10 498.75
Štd. odchýlka.	0.02	0.01	81 436.94	1.00	0.10	17 456.29
Min. hodnota	0.03	0.00	-2 996.94	-0.04	0.66	-23 647.97
25%	0.04	0.00	40 584.50	1.15	0.83	964.12
50%	0.05	0.00	91 386.86	1.80	0.92	10 584.94
75%	0.07	0.01	149 851.70	2.60	0.95	23 900.65
Max. hodnota	0.11	0.02	279 043.30	3.34	0.98	39 527.81

Tabuľka 9: Výsledky testovacej sady.

Pozrieme sa najskôr na akciu, ktorá má najväčší zisk. Ide o spoločnosť Netflix (symbol NFLX), ktorá je najväčšou streamingovou spoločnosťou na svete. Cena tejto spoločnosti bola počas testovacieho obdobia veľmi volatilná a pohybovala sa v rozmedzí od 440 $ do 560 $, a vďaka tejto volatilite sme boli schopní počas testovacieho obdobia nadobudnúť zisk na úrovni 279 043 $. Pohyby pri tejto akcii sa nám podarilo pomerne dobre predikovať, ako je vidno z obrázku 32. Sieť správne predikovala pohyby ako na rast ceny, tak aj na pokles. Sieť tak zachytila nielen celkový trend, ale aj jednotlivé oscilácie.

Obdobne dobre naša sieť zachytila aj varianciu pri spoločnosti s najvyšším Sharpovým pomerom (3.34). V tomto prípade išlo o spoločnosť Charter Communications, pôsobiacu v oblasti telekomunikácií a masmédií. Priebeh ceny tejto akcie a vývoj výnosu je znázornený



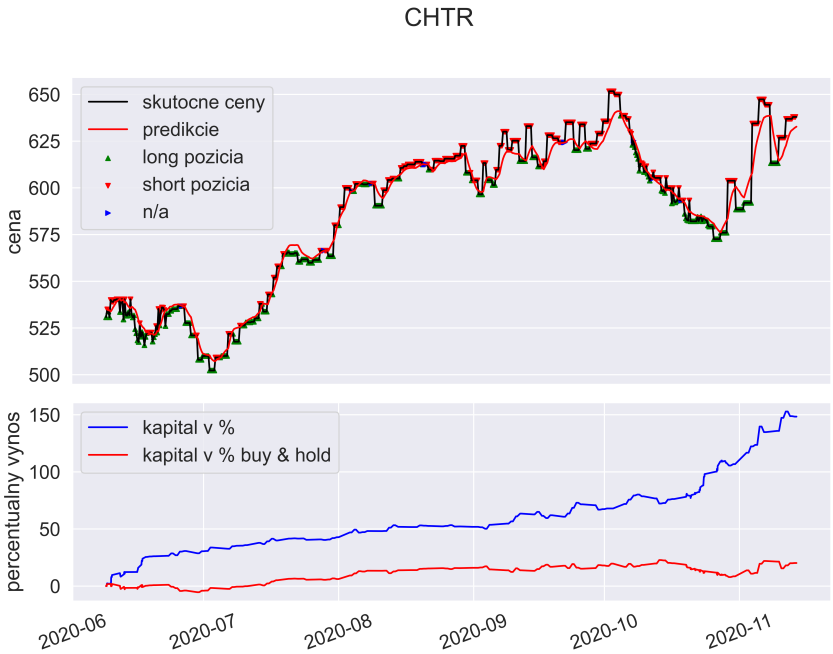
Obr. 32: V hornej časti sa nachádza graf znázorňujúci priebeh ceny akcie NFLX počas testovacieho obdobia. V dolnej časti je percentuálny kumulatívny vývoj ziskovosti stratégie podľa predikcie a podľa "Buy & Hold"stratégie.

na obrázku 33.

V prípade hodinových dát nebudeme vysvetľovať nadobudnutý Sharpov pomer pomocou fundamentálnych parametrov, keďže naša podvzorka obsahuje len malý počet rôznych spoločností.

Pozrieme sa na použité hyperparametre. V prípade hodinových dát a predikcie cien bola najpoužívanejšou sieťou sieť s indexom 7, teda sieť s jednou GRU vrstvou. Tá bola použitá v prípade 7 spoločností. Na druhú stranu, sieť s troma vrstvami LSTM (index 2) nebola použitá ani raz. Zaujímavé je použitie minulých dát. Najpoužívanejšími boli modely so 40 a 50 minulými dátami, kde oba boli použité v 7 prípadoch. Model s 10 minulými hodinami bol použitý v 4 prípadoch. Ostatné, 20 a 30 minulých dát, boli použité každý pre jednu spoloňosť. Najpoužívanejšou kombináciou bola sieť s indexom 7 a 50 minulých dát na vstupe.

Použité indikátory a nadobudnutý Sharpov pomer na testovacej sade sa nachádza na obrázku 34. Najpoužívanejším indikátorom bol MACD, ktorý bol použitý v 5-tich prípadoch. Tiež tri z týchto 5-tich akcií získali Sharpov pomer vyšší ako 2.5. Indikátory



Obr. 33: V hornej časti sa nachádza graf znázorňujúci priebeh ceny akcie CHTR počas testovacieho obdobia. V dolnej časti je percentuálny kumulatívny vývoj ziskovosti stratégie podľa predikcie a podľa "Buy & Hold"stratégie.

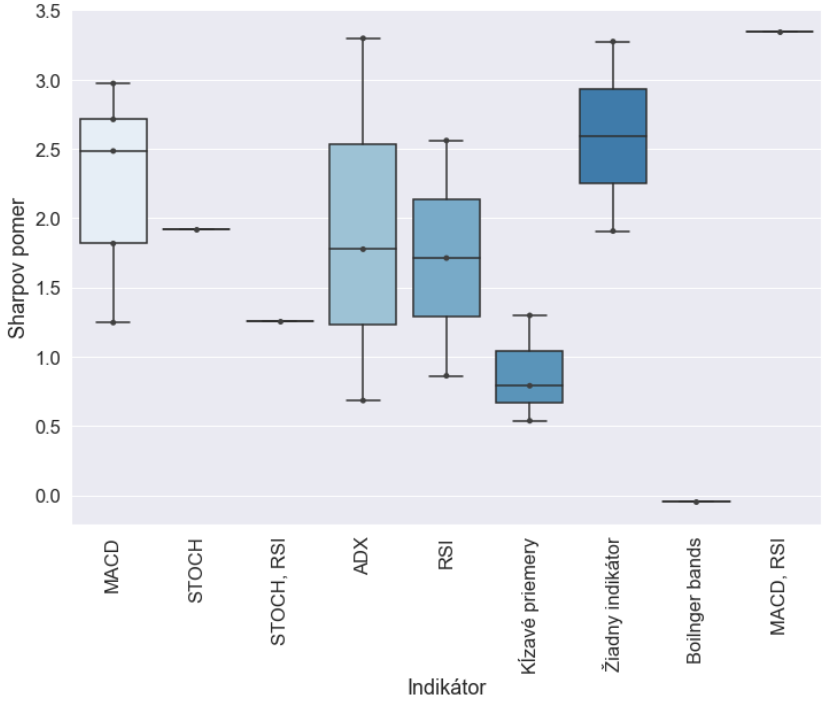
ADX, RSI a KĺzavéPriemery boli použité v troch prípadoch každý. ADX a RSI vykazujú takmer rovnaké priemerné hodnoty Sharpovho pomeru, avšak ADX prináša väčší rozptyl hodnôt. V porovnaní s nimi KĺzavéPriemery majú oveľa horší výkon vzhľadom na Sharpov pomer. Jediný indikátor, ktorý vyústil do záporného Sharpovho pomeru, bol Boilnger Bands. Ten bol použitý v jednom prípade, pričom Sharpov pomer bol -0.04.

Ako sme povedali na úvod tejto podkapitoly, naša stratégia výrazne prekonala "Buy & Hold"stratégiu. Porovnanie týchto stratégií a znázornenie najväčšieho zisku na validačnej sa nachádza na obrázku 35. Vidíme, že v našej stratégií bola stratová iba jedna akcia (Nike, symbol: NKE). Vo všetkých ostatných prípadoch výrazne prekonávame základnú "Buy & Hold"stratégiu. Otestujeme aj v tomto prípade koreláciu medzi ziskovosťou "Buy & Hold"s našou stratégiou. Výberový korelačný koeficient dostávame na úrovni -0.11, v tomto prípade sa nám nezdá, že ide o významnú koreláciu, otestujeme preto naše zdanie.

Zostavíme nasledujúcu hypotézu:

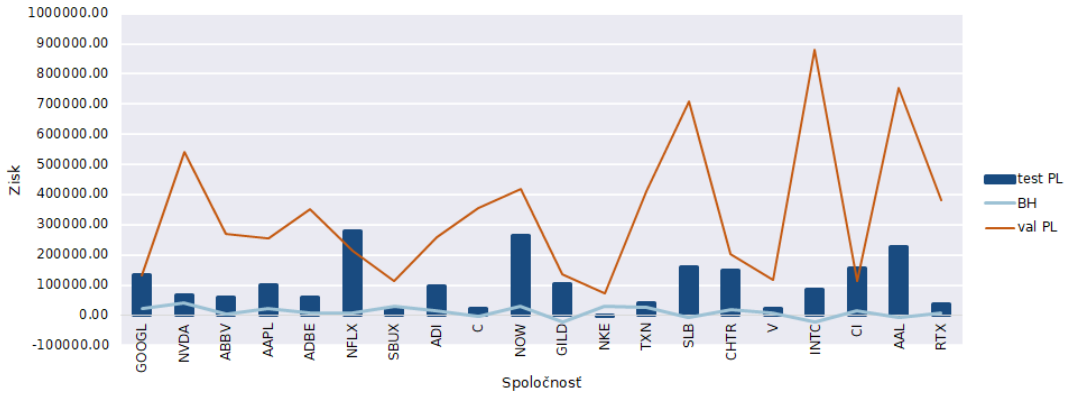
$$H_0 : \rho_{PL,BH} = 0 \text{ vs. } H_1 : \rho_{PL,BH} \neq 0. \tag{66}$$

Dostávame p-hodnotu rovnú 0.6382, skutočne tak korelácia medzi týmito dvoma stratégiami



Obr. 34: Krabicový graf znázorňujúci vzťah medzi použitými indikátormi a Sharpovým pomerom.

nie je štatisticky signifikantná.

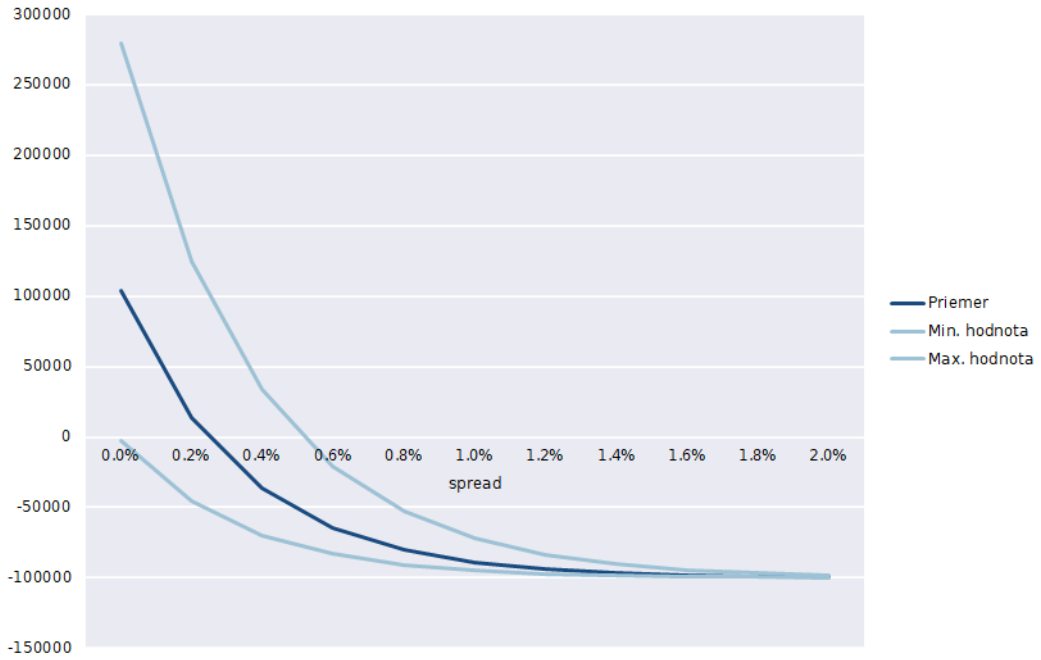


Obr. 35: Porovnanie ziskovosti stratégie "Buy & Hold"(BH), stratégie založenej na predikciách na testovacej sade (test PL) a validačnej sade (val PL) pre jednotlivé spoločnosti.

Poznamenajme, že sme nezobrali do úvahy transakčné náklady na otvorenie a uzatvorenie pozície, ktoré by zásadne zmenili výsledky. Preto pridáme k našim výsledkom ešte akúsi citlivostnú analýzu, v ktorej upravíme zatváracie ceny na Bid a Ask. Pojmom spread označme rozdiel, medzi hodnotou Bid a Ask. Pozrieme sa na citlivosť zisku na

testovacej sada, v závislosti od daného spreadu. Na obrázku 36 sa nachádza znázornenie vývoja zisku v závislosti od spreadu medzi Bid a Ask cenou. Priemerná hodnota zisku je veľmi závislá, od tohto spreadu a veľmi rýchlo sa priemerný zisk prepadá do straty. Tiež pri spreade vo výške 0.6 % už sú stratégie pri všetkých spoločnostiach stratové. Od jedno percentného spreadu priemerná stratégia v podstate prerobila celý vstupný kapitál, a od spreadu 1.6 % už tento kapitál prerobili všetky vybrané spoločnosti.

Vidíme, že napriek tomu, že sme dokázali pomerne dobre predikovať hodinové dáta, takéto časté rebalancovanie nie je nutne to najlepšie, a veľmi rýchlo môže, napriek správnym odhadom, vyústiť do straty.



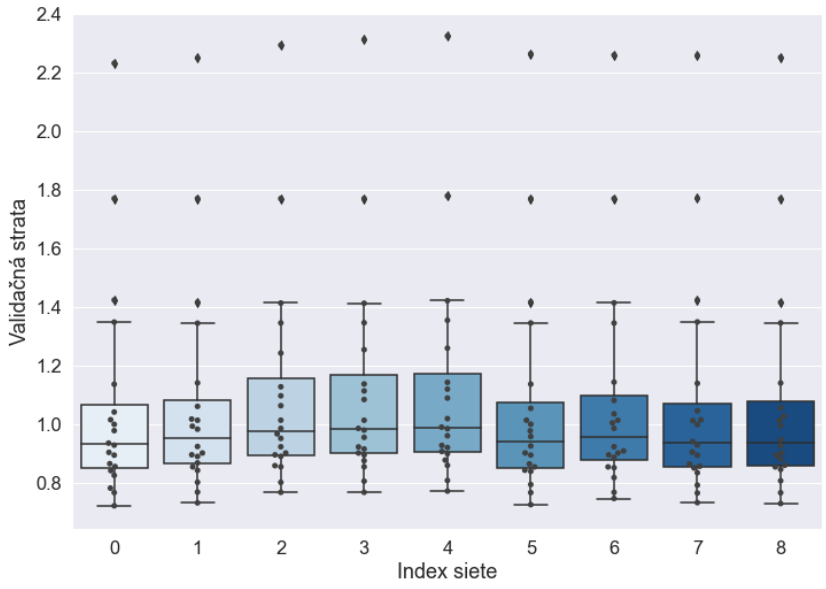
Obr. 36: Znázornenie veľkosti nadobudnutého zisku vzhľadom na veľkosť spreadu.

6.2 Predikcia výnosov

V tejto časti si predstavíme výsledky, ktoré sme na hodinových dátach dosiahli pri predikovaní výnosov. Taktiež nadobudnuté výsledky porovnáme s výsledkami dosiahnutými pri predikovaní cien na tomto datasete.

6.2.1 Výsledky na validačnej sade

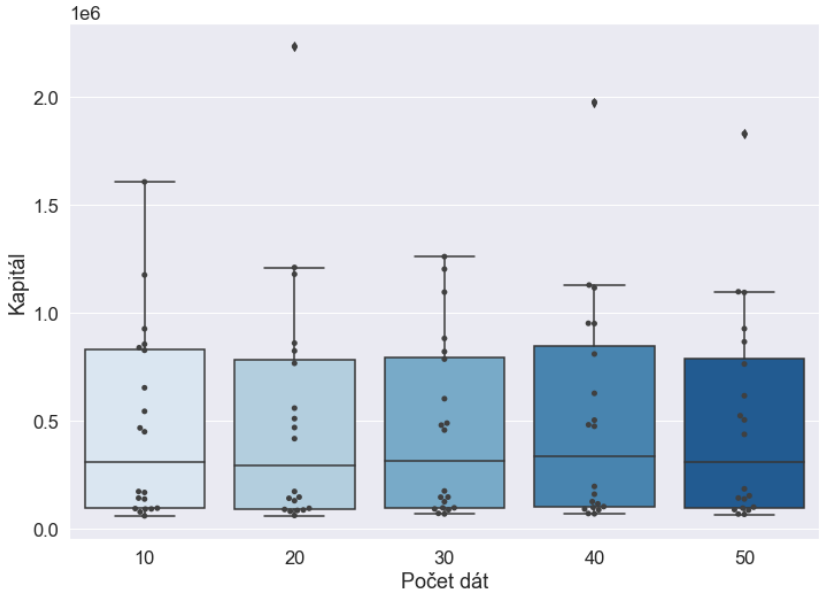
Podobne ako v predchádzajúcich kapitolách, si najskôr uvedieme výsledky na validačnej sade. Krabicový graf na obrázku 37 znázorňuje validačnú stratu vzhľadom na použitú sieť. V tomto prípade vidíme, že najlepšie výsledky sme dosahovali pre jednoduchšie siete, s indexom 0, 1, 5, 7 a 8. Pri týchto sieťach sme dosiahli nielen najnižšiu priemernú stratu ale aj medzikvartilový rozptyl týchto sietí nadobúda najnižšie hodnoty. Naopak, komplikovanejšie siete s viacerými LSTM vrstvami, dosiahli horšie výsledky, čo sa týka priemernej straty aj rozptýlenosti nadobudnutých hodnôt.



Obr. 37: Krabicový graf so znázornením validačnej straty - MAE, vzhľadom na index siete.

Ak sa pozrieme na nadobunutý kapitál na validačnej sade rozdelený podľa počtu dát, pre ktoré bol nadobudnutý, tak vidíme menšie rozdiely ako v prípade validačnej straty a indexu siete. Krabicový graf so znázornením kapitálu podľa počtu vstupných dát sa nachádza na obrázku 38. Najvyšší priemerný kapitál sme dosiahli pre modely s 10-imi a 40-imi hodinami na vstupe. Modely so vstupom 10 dát navyše zaznamenali aj najväčšiu

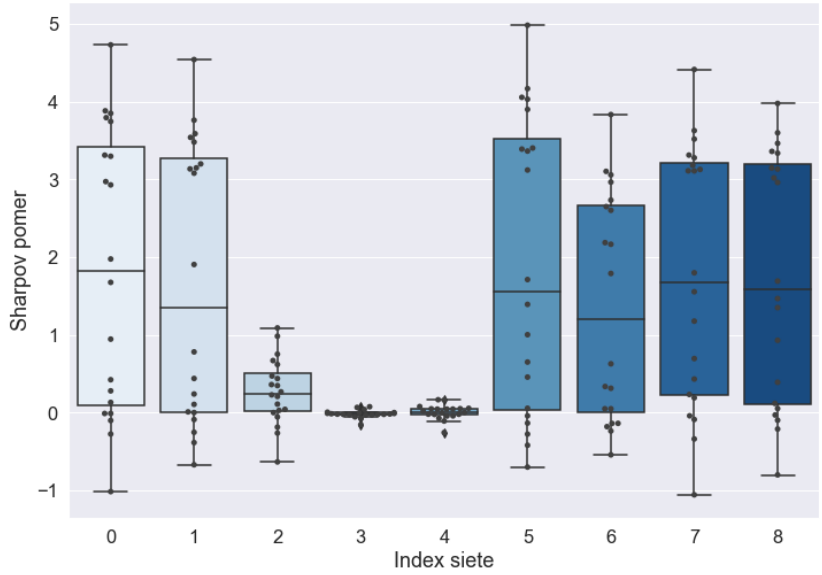
hodnotu pre horný fúz, teda pre takéto modely môžeme očakávať aj viac pozitívnych extrémov. Avšak všetky modely zaznamenali priemerný kapitál podstatne menší, ako bol počiatočný, čiže výkonnosť našich sietí pre všetky vstupy môžeme v tomto prípade označiť za slabú.



Obr. 38: Krabicový graf so znázornením nadobudnutého kapitálu a počtu dní, ktoré do siete vstupovali.

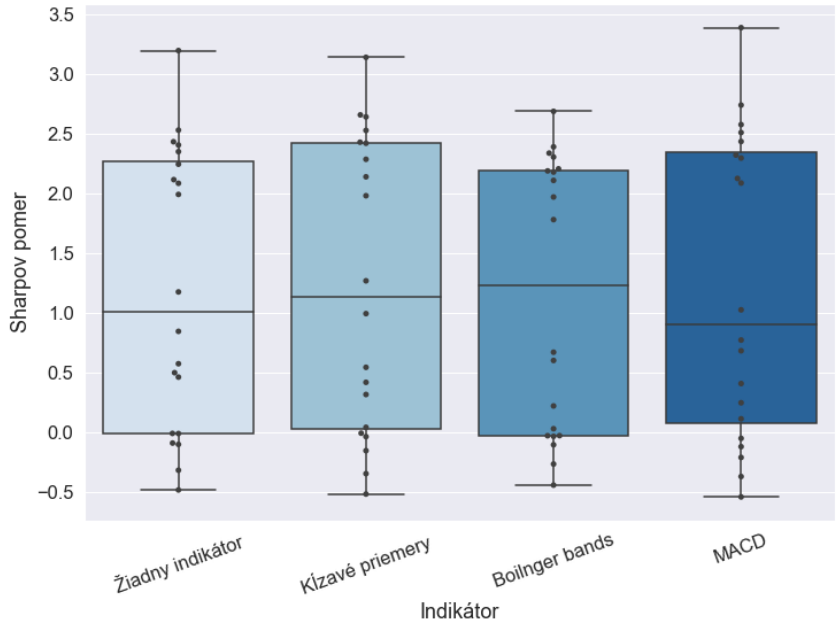
Metrikou, ktorá nám poskytne najviac informácií, je aj v tomto prípade Sharpov pomer. Na obrázku 39 máme krabicový graf znázorňujúci vzťah medzi indexom siete a nadobunutým Sharpovým pomerom. Najvyššiu priemernú hodnotu Sharpovho pomeru sme na validačnej sade dosiahli pre siete s indexom 0 a 7. Tieto siete ale vykazujú aj najväčší rozptyl dát. Zaujímavé výsledky dosiahli siete s indexmi 3 a 4, siete so 4-mi respektíve 5-timi LSTM vrstvami. Pre tieto siete sme vo všetkých prípadoch dosiahli Sharpov pomer nulový, alebo veľmi blízky nule. Čiže podobne ako pri validačnej strate, aj pri tejto metrike môžeme konštatovať, že jednoduchšie nielen dosiahli lepšie výsledky.

Posledným hyperparametrom, na ktorý sa pozrieme, sú použité technické indikátory. Opäť sa zameriame len na základné indikátory, ktoré sme použili. V prípade predikcie výnosov používame kĺzavé priemery, Boilnger bands a MACD. Krabicový graf, znázorňujúci vzťah medzi použitými indikátormi a validačným Sharpovým pomerom, sa nachádza na obrázku 40. Najväčší Sharpov pomer sme pri predikciách výnosov na hodinových dátach dosiahli pre siete, ktoré okrem základných vstupov mali ešte aj Boilnger bands, ako ďalší



Obr. 39: Krabicový graf zobrazujúci Sharpov pomer vzhľadom na použitú sieť.

parameter. Najhoršie výsledky, čo sa týka priemerného Sharpovho pomeru aj rozptýlenosti dát, sme dosiahli po pridaní indikátora MACD. Avšak rozdiely medzi jednotlivými indikátormi neboli natoľko výrazné, aby sme vedeli jednoznačne povedať, že pridanie niektorého z nich, výrazne zlepšilo Sharpov pomer na validačnej sade.



Obr. 40: Krabicový graf zobrazujúci vzťah medzi nadobudnutým Sharpovým pomerom a použitým indikátorom.

Porovnáme aj výsledky na validačnej sade pri predikovaní výnosov s výsledkami pri

predikovaní cien. Validačná strata pri predikovaní cien dosahovala aj vďaka transformácii výrazne nižšie hodnoty, ako v prípade predikcie výnosov. Pri cenách, ako aj pri výnosoch, sme za najlepšie a najhoršie prehlásili siete s rovnakým indexom. To naznačuje, že jednoduchšie siete, vďaka menšej náchylnosti k preučeniu, dávajú na validačnej sade lepšie výsledky. Túto spoločnú črtu dobre vidieť pri porovnaní obrázkov28 a 37.

Čo sa týka Sharpovho pomeru, ktorý sme dosiahli na validačnej sade pri týchto dvoch prístupoch, tak pri predikovaní výnosov sme dosiahli síce väčšie, ale menej konzistentné výsledky. Pri predikovaní cien, sa nadobudnutý Sharpov pomer pre všetky siete pohyboval okolo hodnoty 1, a pri predikovaní výnosov sme dosahovali pre najlepšie siete hodnoty v okolí 2 a pre najhoršie iba v okolí 0.

6.2.2 Výsledky na testovacej sade

V tejto časti sa pozrieme na to, aké výsledky sme pri predikovaní výnosov dosiahli na sade, ktorá nás najviac zaujíma, a to testovacia sada. Charakteristiky sledovaných parametrov sa nachádzajú v tabuľke 10. Priemerná hodnota zisku našej stratégie oproti zisku v základnej "Buy & Hold"stratégii je takmer trikrát väčšia. Avšak ak spravíme Studentov t-test analogický k testom (49) a (59) dostávame p-hodnotu 0.06 a na 5 % hladine významnosti nezamietame, že sa priemerný zisk pri oboch stratégiach rovná. Hodnoty koeficientu determinácie sú podľa očakávania výrazne nižšie, ako v prípade predikcie cien, ale napriek tomu sa zdá, že sú tieto hodnoty až priveľmi nízke a naše nielen nepredikovali výnosy veľmi dobre. Hodnotu, ktorá nás tiež môže znepokojovať, je najväčšia strata, ktorú sme dosiahli. Okrem toho aj mediánová hodnota ziskov je takmer rovnaká ako v prípade "Buy & Hold"stratégie, teda vyššia priemerná ziskovosť v prípade našej stratégie bola ovplyvnená výrazne vysokými odľahlými hodnotami.

Spoločnosťou, pre ktorú sme dosiahli najväčší zisk, bola ServiceNow (symbol NOW). Zaoberá sa poskytovaním cloudových služieb a ich správe pre nadnárodné spoločnosti. Počas testovacieho obdobia táto spoločnosť zaznamenala výrazný nárast ceny, kde vzrástla o takmer 30 %. Znázornenie logaritmických výnosov tejto akcie a vývoja percentuálnej zmeny kapitálu našej a "Buy & Hold"stratégie sa nachádza na obrázku 41. Naša sieť dobre odhadovala priemerné výnosy, čo vidno priamo na ich predikciách, ale aj na tom, že percentuálny kumulovaný kapitál počas takmer celého obdobia rástol. Výrazný nárast

	Strata	MSE	Zisk	Sharpov pomer	R2	Zisk Buy & hold
Priemer	0.55	1.29	29 684.24	0.47	-0.21	10 498.75
Štd. odchýlka	0.24	0.80	57 579.19	1.02	0.21	17 456.29
Min. hodnota	0.25	0.46	-31 754.80	-1.37	-0.52	-23 647.97
25%	0.34	0.72	-4 629.73	-0.07	-0.39	964.12
50%	0.52	1.17	10 436.37	0.33	-0.26	10 584.94
75%	0.66	1.61	52 333.30	1.13	0.00	23 900.65
Max. hodnota	1.14	3.78	200 742.54	2.65	0.01	39 527.81

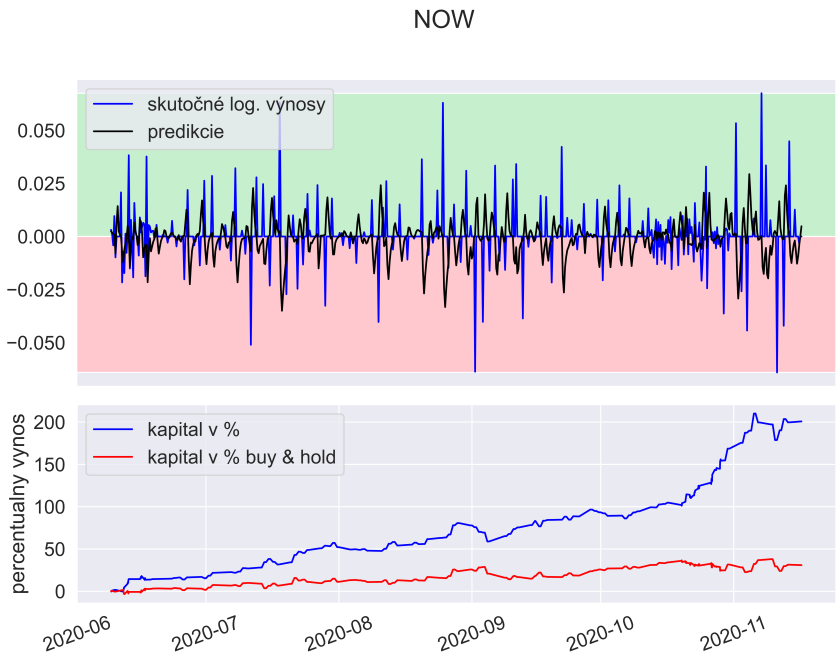
Tabuľka 10: Výsledky na testovacej sade.

kapitálu sme zaznamenali od obdobia 2020-11, kedy sa kumulovaný percentuálny kapitál pri našej stratégii zvýšil o takmer 50 percentuálnych bodov.

Najmenšiu hodnotu koeficientu determinácie sme dosiahli -0.52. Dosiahnutá bola pre spoločnosť AbbVie (symbol ABBV), pričom sa jedná o biotechnologickú spoločnosť. Znázornenie predikcií a skutočných výnosov, ako aj percentuálny vývoj kapitálu, sa nachádza na obrázku 42.Pri tejto akcii sme neboli schopní takmer vôbec správne zachytiť jednotlivé výnosy. Ku koncu testovacieho obdobia ale správne predikovala niekoľko hodín, vďaka čomu sa naša stratégia dostala do zisku a nakoniec aj porazila "Buy & Hold".

Pozrieme sa na použité hyperparametre, na testovacej sade. Jednoznačne najpoužívanejšou bola sieť s indexom 7, jedna GRU vrstva, ktorá bola použitá v prípade 8 spoločností. Ďalšie pomerne početne zastúpené boli siete s indexom 0 a 1, teda jedna a dve LSTM vrstvy, pričom každá bola použitá v 4 prípadoch. Vôbec neboli použité siete s indexami 3, 4 a 6, čo korešponduje aj s výsledkami na validačnej sade.

Čo sa týka počtu vstupných dát, tak s veľkým náskokom najpoužívanejšími boli modely s 10 hodinami na vstupe. Konkrétne toto nastavenie bolo použité pre 15 spoločnosí. Ďalej pri 4 spoločnostiach bolo na vstupe 20 minulých dát a pre jednu spoločnosť 50 minulých dát.

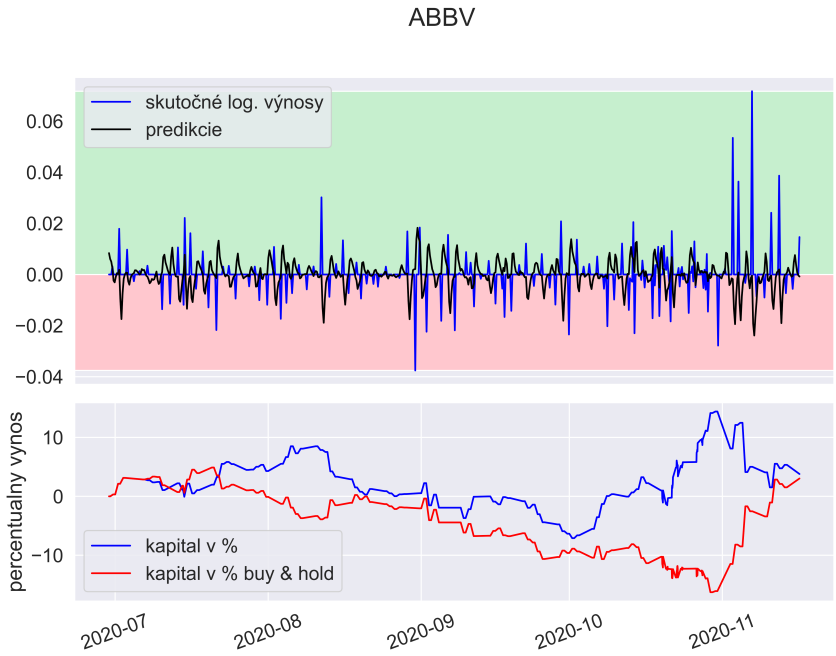


Obr. 41: V hornej časti sa nachádza graf znázorňujúci priebeh ceny akcie NOW počas testovacieho obdobia. V dolnej časti je percentuálny kumulatívny vývoj ziskovosti stratégie podľa predikcie a podľa "Buy & Hold"stratégie.

Kombináciou, ktorá bola použitá pre najviac spoločností, bola sieť s indexom 7 a s 10 dátami na vstupe. Táto kombinácia bola použitá pre 6 firiem.

Pri pohľade na technické indikátory použité na testovacej sade v prípade predikcie výnosov na hodinových dátach, zistíme, že boli použité len indikátory MACD a kĺzavé priemery. Použité indikátory a nadobudnutý Sharpov pomer sú znázornené na krabicovom grafe na obrázku 43. Modely bez indikátora boli použité v 15 prípadoch, preto je prirodzené, že pri nich máme aj najväčší rozptyl dát. Kĺzavé priemery a kĺzvé priemery s MACD boli použité oba v jednom prípade. Oba tieto modely ale dosiahli pomerne vysoké hodnoty Sharpovho pomeru, 1.75 respektíve 1.4.

Aj v prípade predikcie výnosov na hodinových dátach porovnáme zisk našej stratégie a zisk "Buy & Hold"stratégie. Porovnanie zisku našej stratégie a "Buy & Hold"stratégie sa nachádza na obrázku 44. V tomto prípade vidíme pri väčšine spoločností podobné zisky pri oboch stratégiách. Naša stratégia výrazne prekonáva "Buy & Hold"len v prípade 5-tich spoločností. Výberový korelačný koeficient v tomto prípade dostávame na úrovni 0.45, čo



Obr. 42: V hornej časti sa nachádza graf znázorňujúci priebeh ceny akcie ABBV počas testovacieho obdobia. V dolnej časti je percentuálny kumulatívny vývoj ziskovosti stratégie podľa predikcie a podľa "Buy & Hold"stratégie.

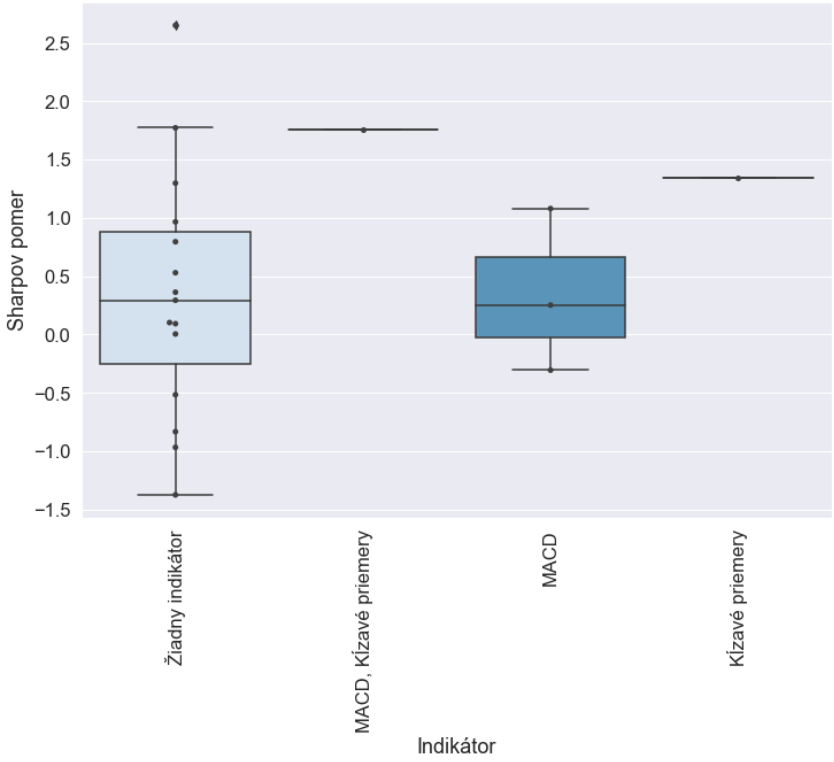
naznačuje štatisticky signifikantnú koreláciu medzi stratégiami. Zostavíme hypotézu:

$$H_0 : \rho_{PL,BH} \leq 0 \text{ vs. } H_1 : \rho_{PL,BH} > 0. \tag{67}$$

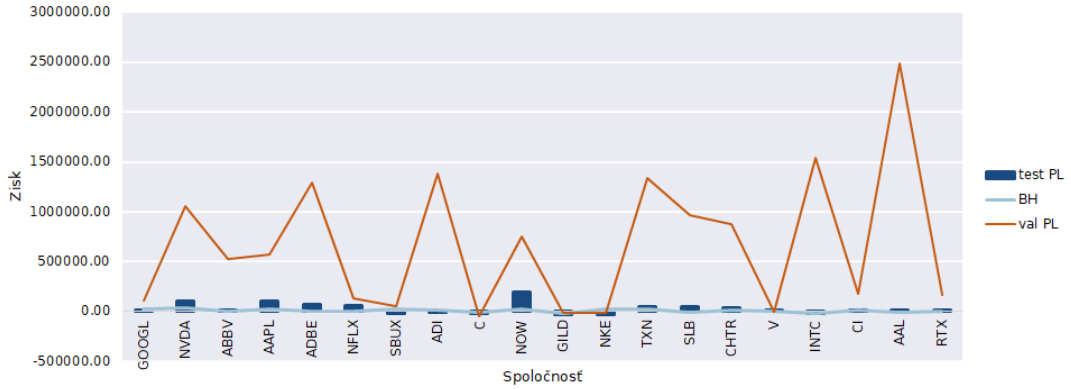
Dostávame p-hodnotu rovnú 0.04508, a na 5 % hladine významnosti zamietame H_0, čiže obe stratégie sú signifikantne štatisticky korelované.

Aj v prípade predikcie výnosov dôležitú úlohu zohrajú transakčné náklady. Avšak v tomto prípade máme ziskové stratégie iba pre prípad nulového spreadu. Už pri najmenšom spreade sa dostávame do straty na úrovni počiatočného kapitálu. Z toho môžeme súdiť, že predikovanie výnosov na hodinových dátach neprináša systematicky správne výsledky, ale ide skôr o náhodu.

Porovnajme výsledky predikcie cien a výnosov na hodinových dátach. Na rozdiel od denných dát, sa ukázalo, že na predikovanie hodinových dát sa viac hodí modelovať priamo ceny, a nie výnosy. Videli sme to napríklad pri testovaní, či je priemerný zisk našej stratégie signifikantne väčší ako priemerný zisk "Buy & Hold"stratégie, kde sme pri cenách dostali, že je signifikantne väčší, však nie pri výnosoch. Výrazne lepšie výsledky naznačovala už sama hodnota priemerného dosiahnutého zisku.



Obr. 43: Krabicový graf znázorňujúci vzťah medzi použitými indikátormi a Sharpovým pomerom.



Obr. 44: Porovnanie ziskovosti stratégie "Buy & Hold"(BH), stratégie založenej na predikciách na testovacej sade (test PL) a validačnej sade (val PL) pre jednotlivé spoločnosti.

Tiež citlivosť stratégií na veľkosť spreadu hovorí výrazne v prosprech predikovania cien. V ich prípade sme mali ešte akúsi "bezpečnú"rezervu pre veľkosť spreadu, kde sme dosahovali síce výrazne menší zisk, ale neprepadli sme sa hneď pri všetkých spoločnostiach do straty. Pri predikovaní výnosov sme o takýto vankúš prišli, keď po pridaní spreadu sa stratégie prakticky pre všetky spoločnosti prepadli do straty.

Záver

V práci sme sa venovali modelovaniu cien a výnosov akcií spoločností, ktoré patria do indexu S&P 500. Na modelovanie sme využili hlboké rekurentné neurónové siete.

Teoretický úvod do problematiky strojového učenia a špeciálne do neurónových sietí sme uviedli v prvej kapitole. Zadefinovali sme si základne pojmy potrebné pre pochopenie fungovania neurónových sietí. Taktiež sme uviedli akýsi historický vývoj týchto sietí a prvé myšlienky smerujúce k tejto prelomovej metóde. Azda matematicky najzložitejšou časťou tejto kapitoly boloodvodeniee trénovania sietí a algoritmus spätnej propagácie. Pre túto prácu bolo najdôležitejšie zadefinovanie a ukázanie fungovania rekurentných sietí, ktoré si ukladajú do pamäte aj už spracované vstupy. Takéto siete sú základom nášho ďalšieho modelovania.

V kapitole s poradovým číslom 2 sme uviedli prehľad prác, ktoré sa venovali rovnakej, alebo podobnej téme ako my. Pre všetky práce sme uviedli spôsob predikcie a dosiahnuté výsledky.

Tretia kapitola bola venovaná dátam, s ktorými sme pracovali. Predstavili sme podvzorky, na ktorých sme naše siete trénovali. Z dôvodu nedostatku hodinových dát, sme museli pre tieto vybrať inú, menšiu, podvzorku, ktorá nereprezentovala fundamentálne parametre celého indexu tak dobre, ako podvzorka pre denné dáta. Vstupom do našich modelov, boli okrem cien, aj technické indikátory, ktoré sme zadefinovali na konci tejto kapitoly.

Architektúre sietí a jednotlivým hyperparametrom bola venovaná kapitola 4. Okrem toho, sme si zadefinovali finančné stratégie, pomocou ktorých sme vyberali finálnu sieť pre jednotlivé spoločnosti.

V kapitolách 5 a 6 sme už predstavili konkrétne výsledky, ktoré sme dosiahli. Obe kapitoly boli členené na výsledky pre predikovanie cien a výnosov. Na konci každej podkapitoly venovanej predikovaniu výnosov sme ešte porovnali dosiahnuté výsledky pre predikovanie cien a výnosov.

Konkrétne 5. kapitola bola venovaná výsledkom na denných dátach. Na týchto dátach, sa napriek lepším štatickým a chybovostným metrikám, ako koeficient determinácie alebo stredná absolútna chyba, v prípade predikovania cien ukázali ako ziskovejšie predikcie výnosov. To sa ukázalo nielen na výške priemerného výnosu, ale aj na počte ziskových

spoločností. Treba ale povedať, že jednoduchá "Buy & Hold"stratégia bola pre obe predikcie zisková pre viac spoločností, ako naše stratégie, bez ohľadu na predikovaný výstup.

Výsledky na hodinových dátach sme rozobrali v kapitole 6. V tomto prípade sme dostali opačné výsledky, predikcia cien výrazne prekonala predikciu výnosov. Pri predikovaní cien na takto krátkom časovom úseku sme boli schopní predpovedať takmer presne budúce pohyby. Avšak pri predikovaní výnosov často išlo len o náhodu. Otázkou zostáva, aké výsledky by sme dostali na komplexnejšej sade spoločností, ktorá by lepšie kopírovala index.

V práci sme sa pokúsili predikovať budúci vývoj akciových trhov. Ukázali sme, že minimálne v prípade hodinových dát existuje potenciál na to, aby sme vedeli takpovediac odhadnúť budúcnosť. Problémom však ostáva nerobustnosť nášho prístupu, keďže každá spoločnosť si mohla vybrať sieť, ktorá pre ňu funguje najlepšie. Následne, ak sme sa pokúšali nájsť pre sadu spoločností napríklad z jedného odvetvia, alebo podľa jej veľkosti nejakú sieť, ktorá by bola pre ne najpoužívanejšia, sme takú sieť neboli schopní nájsť. Problematická bola nejednoznačnost výsledkov. Kde pre žiadnu kategóriu sa neobjavoval nejaký vzor, ktoré hyperparametre boli použité. Zdá sa teda, že takéto neurónové siete aspoň zatiaľ nenahradia finančných analytikov a názor skúseného odborníka má stále svoje nezastupiteľné miesto.

Stále populárnejšou metódou aj pre obchodovanie na finančných trchoch sa stáva učenie posilňovaním. Túto metódu sa nám v našej práci nepodarilo obsiahnúť, ale určite by bolo zaujímavé porovnať jej výsledky s našimi.

Literatúra

[1] Biogen stock price (biib). https://www.investing.com/equities/biogen-idec-inc.

[2] Biogen's stock jumps 42% after fda staff says it has enough data to support approving alzheimer's drug. https://www.cnbc.com/2020/11/04/biogens-stock-jumps-30percent-after-fda-staff-says-it-has-enough-data-to-support-approving-alzheimers-drug-.html.

[3] marketstack.com. `https://marketstack.com/`.

[4] Wharton research data services. `https://wrds-web.wharton.upenn.edu/wrds/ds/comp/funda/index.cfm?navId=80`.

[5] Zdrojový kód. `https://github.com/jan314159/diplomova_praca`.

[6] Investment returns by asset class (1985 to 2020). `https://themeasureofaplan.com/investment-returns-by-asset-class/`, Jan 2021.

[7] Martín Abadi, Ashish Agarwal, Paul Barham, Eugene Brevdo, Zhifeng Chen, Craig Citro, Greg S. Corrado, Andy Davis, Jeffrey Dean, Matthieu Devin, Sanjay Ghemawat, Ian Goodfellow, Andrew Harp, Geoffrey Irving, Michael Isard, Yangqing Jia, Rafal Jozefowicz, Lukasz Kaiser, Manjunath Kudlur, Josh Levenberg, Dandelion Mané, Rajat Monga, Sherry Moore, Derek Murray, Chris Olah, Mike Schuster, Jonathon Shlens, Benoit Steiner, Ilya Sutskever, Kunal Talwar, Paul Tucker, Vincent Vanhoucke, Vijay Vasudevan, Fernanda Viégas, Oriol Vinyals, Pete Warden, Martin Wattenberg, Martin Wicke, Yuan Yu, and Xiaoqiang Zheng. TensorFlow: Large-scale machine learning on heterogeneous systems. `https://www.tensorflow.org/`, 2015. Software available from tensorflow.org.

[8] Ayodele A Adebiyi, Charles K Ayo, Marion O Adebiyi, and Sunday O Otokiti. Stock price prediction using neural network with hybridized market indicators. *Journal of Emerging Trends in Computing and Information Sciences*, 3(1):1–9, 2012.

[9] A. Agresti. *Categorical Data Analysis*. Wiley Series in Probability and Statistics. Wiley, 2012.

[10] Yoshua Bengio, Patrice Simard, and Paolo Frasconi. Learning long-term dependencies with gradient descent is difficult. *IEEE transactions on neural networks*, 5(2):157–166, 1994.

[11] Stephen Boyd and Lieven Vandenberghe. *Convex Optimization*. Cambridge University Press, USA, 2004.

[12] Francois Chollet. *Deep learning with Python*. Manning Publications, New York, NY, 2018.

[13] Yann N Dauphin, Razvan Pascanu, Caglar Gulcehre, Kyunghyun Cho, Surya Ganguli, and Yoshua Bengio. Identifying and attacking the saddle point problem in high-dimensional non-convex optimization. In Z. Ghahramani, M. Welling, C. Cortes, N. Lawrence, and K. Q. Weinberger, editors, *Advances in Neural Information Processing Systems*, volume 27, pages 2933–2941. Curran Associates, Inc., 2014.

[14] Guangyu Ding and Liangxi Qin. Study on the prediction of stock price based on the associated network model of lstm. *International Journal of Machine Learning and Cybernetics*, 11(6):1307–1317, 2020.

[15] John Duchi, Elad Hazan, and Yoram Singer. Adaptive subgradient methods for online learning and stochastic optimization. *Journal of Machine Learning Research*, 12(61):2121–2159, 2011.

[16] Guanhao Feng, Jingyu He, and Nicholas G. Polson. Deep learning for predicting asset returns, 2018.

[17] Ian Goodfellow, Yoshua Bengio, and Aaron Courville. *Deep Learning*. MIT Press, 2016.

[18] Mária Hamala, Milan a Trnovská. *Nelineárne programovanie*. Epos, Bratislava, 2013.

[19] Michael R. James and Satinder Singh. Learning and discovery of predictive state representations in dynamical systems with reset. In *In ICML*, pages 417–424, 2004.

[20] Tarald O. Kvålseth. Cautionary note about r 2. *The American Statistician*, 39(4):279–285, 1985.

[21] Weiping Li and Feng Mei. Asset returns in deep learning methods: An empirical analysis on sse 50 and csi 300. *Research in International Business and Finance*, 54:101291, 2020.

[22] Jozef Maršala. *Systematická a funkčná neuroanatómia.* Osveta (Martin), 1985.

[23] Igor Melicherčík, Ladislava Olšarová, and Vladimír Úradníček. *Kapitoly z finančnej matematiky 1.* Bratia Sabovci, 2005.

[24] T.M. Mitchell. *Machine Learning.* McGraw-Hill International Editions. McGraw-Hill, 1997.

[25] Pavol Návrat, Mária Bieliková, L'ubica Beňušková, I Kapustník, and Milan Unger. Umelá inteligencia. *Bratislava: Vydavatel'stvo STU*, 2007.

[26] Frank Rosenblatt. The perceptron: a probabilistic model for information storage and organization in the brain. *Psychological review*, 65(6):386, 1958.

[27] Eberhard Schöneburg. Stock price prediction using neural networks: A project report. *Neurocomputing*, 2(1):17–27, 1990.

[28] William F. Sharpe. Mutual fund performance. *The Journal of Business*, 39(1):119–138, 1966.

[29] David Silver, Thomas Hubert, Julian Schrittwieser, Ioannis Antonoglou, Matthew Lai, Arthur Guez, Marc Lanctot, Laurent Sifre, Dharshan Kumaran, Thore Graepel, Timothy P. Lillicrap, Karen Simonyan, and Demis Hassabis. Mastering chess and shogi by self-play with a general reinforcement learning algorithm. *CoRR*, abs/1712.01815, 2017.

[30] Nitish Srivastava, Geoffrey Hinton, Alex Krizhevsky, Ilya Sutskever, and Ruslan Salakhutdinov. Dropout: a simple way to prevent neural networks from overfitting. *The journal of machine learning research*, 15(1):1929–1958, 2014.

[31] Tijmen Tieleman and Geoffrey Hinton. Lecture 6.5-rmsprop: Divide the gradient by a running average of its recent magnitude. *COURSERA: Neural networks for machine learning*, 4(2):26–31, 2012.

[32] John W. Tukey. *The Future of Data Analysis*, pages 408–452. Springer New York, New York, NY, 1992.

[33] Shuxiang Xu and Ling Chen. A novel approach for determining the optimal number of hidden layer neurons for fnn's and its application in data mining. 2008.

[34] Pengfei Yu and Xuesong Yan. Stock price prediction based on deep neural networks. *Neural Computing and Applications*, 32(6):1609–1628, 2020.

Príloha

Neoddeliteľnou súčasťou tejto práce je zdrojový kód. Kód je uložený a voľne prístupný na stránke GitHub, tento kód bol vytvorený pomocou jazyka Python vo verzii 3.83 a knižnice TensorFlow vo verzii 2.3.0.

Zdrojový kód Všetky kódy, použité pre túto prácu sú uložené a prístupné na [5].