Technická univerzita v Košiciach

Fakulta elektrotechniky a informatiky

Využitie hlbokého učenia pre spracovanie a analýzu ionosférických dát

Diplomová práca

2024 **Bc. Silvia Kostárová**

Technická univerzita v Košiciach

Fakulta elektrotechniky a informatiky

Využitie hlbokého učenia pre spracovanie a analýzu ionosférických dát

Diplomová práca

Študijný program:	Hospodárska informatika
Študijný odbor:	Informatika
Školiace pracovisko:	Katedra kybernetiky a umelej inteligencie (KKUI)
Školiteľ:	doc. Ing. Peter Butka, PhD.
Konzultant:	Ing. Viera Krešňáková, PhD.

Košice 2024 **Bc. Silvia Kostárová**

Abstrakt v SJ

Ionosféra predstavuje najkritickejšiu vrstvu zemnskej atmosféry pre prenos rádiového signálu medzi vesmírnymi misiami a pozemnými stanicami. Náhle zmeny v hustote iónov a elektrónov dokážu prudko modifikovať fázu a amplitúdu rádiového signálu. Tieto modifikácie nesú názov ionosférické scintilácie. Ich výskyt spôsobuje problémy pre využívanie globálnych navigačných satelitných systémov. Cieľom tejto diplomovej práce je vyvinutie modelov hlbokého učenia a efektívneho spôsobu spracovania ich výstupov na predikciu fázových ionosférických scintilácií. Použité boli ionosférické dáta z oblasti Kanady a dáta popisujúce geomagnetickú aktivitu v okolí Zeme. Dosiahnuté výsledky pri testovaní vytvorených modelov predstavujú sľubný začiatok cesty implementovania výstupov tejto práce do praxe.

Kľúčové slová

ionosféra, ionosférická scintilácia, sekvenčné dáta, hlboké učenie, neurónové siete, predikcia, vesmírny výskum

Abstrakt v AJ

The ionosphere is the most critical layer of the Earth's atmosphere for radio signal transmission between space missions and ground stations. Sudden changes in the density of ions and electrons can significantly modify the phase and amplitude of the radio signal. These modifications are called ionospheric scintillations. Their occurrence causes problems for the use of global navigation satellite systems. The aim of this thesis is to develop deep learning models and an efficient method of processing their outputs for the prediction of phase ionospheric scintillations. Ionospheric data from the Canadian region and data describing geomagnetic activity around the Earth were used. The results achieved during testing the models represent a promising beginning of the path of implementing the outputs of this work into practice.

Kľúčové slová v AJ

ionosphere, ionospheric scintillation, sequential data, deep learning, neural networks, prediction, space research

71603

TECHNICKÁ UNIVERZITA V KOŠICIACH

FAKULTA ELEKTROTECHNIKY A INFORMATIKY

Katedra kybernetiky a umelej inteligencie

ZADANIE DIPLOMOVEJ PRÁCE

Študijný odbor: **Informatika**

Študijný program: **Hospodárska informatika**

Názov práce:

Využitie hlbokého učenia pre spracovanie a analýzu ionosférických dát

Use of deep learning methods for processing and analysis of ionospheric data

Študent: **Bc. Silvia Kostárová**

Školiteľ: **doc. Ing. Peter Butka, PhD.**

Školiace pracovisko: **Katedra kybernetiky a umelej inteligencie**

Konzultant práce: **Ing. Viera Krešňáková, PhD.**

Pracovisko konzultanta: **Katedra kybernetiky a umelej inteligencie**

Pokyny na vypracovanie diplomovej práce:

1. Podať teoretický prehľad problematiky metód hlbokého učenia, s dôrazom na ich použitie pri spracovaní ionosférických dát.
2. Získať a predspracovať dáta pre predikciu ionosférických scintilácií.
3. Navrhnúť, realizovať a vyhodnotiť experimenty s predikciou ionosférických scintilácií pomocou metód hlbokého učenia.
4. Vypracovať dokumentáciu podľa pokynov katedry a vedúceho práce.

Jazyk, v ktorom sa práca vypracuje: slovenský

Termín pre odovzdanie práce: 19.04.2024

Dátum zadania diplomovej práce: 31.10.2023

TECHNICKÁ UNIVERZITA V KOŠICIACH FAKULTA ELEKTROTECHNIKY A INFORMATIKY -11-

v.z. Pribilová

prof. Ing. Liberios Vokorokos, PhD.

dekan fakulty

Čestné vyhlásenie

Vyhlasujem, že som diplomovú prácu vypracovala samostatne s použitím uvedenej odbornej literatúry.

Košice 19. 4. 2024 .....

Vlastnoručný podpis

Poďakovanie

Moje poďakovanie patrí vedúcemu tejto práce doc. Ing. Petrovi Butkovi, PhD., konzultatntke Ing. Viere Krešňákovej, PhD. a RNDr. Šimonovi Mackovjakovi, PhD. za ich ochotu, cenné rady a v neposlednom rade za možnosť spolupodieľať sa na výskume kozmického počasia.

Obsah

Úvod **1**

1 Hlboké učenie a neurónové siete **3**

1.1 Hlboké učenie . . . 3

1.2 Neurónové siete . . . 6

1.2.1 Aktivačné funkcie . . . 9

1.2.2 Chybové funkcie . . . 11

1.2.3 Optimalizácia učenia . . . 12

1.2.4 Regularizácia učenia . . . 15

1.3 Rekurentné neurónové siete . . . 16

1.3.1 LSTM . . . 18

1.3.2 BiLSTM . . . 19

2 Ionosférické dáta **20**

2.1 Ionosféra . . . 20

2.2 Ionosférická scintilácia . . . 22

2.3 Analýza súčasného stavu . . . 24

2.3.1 Kľúčové pojmy . . . 24

2.3.2 Analýza literatúry . . . 25

3 Automatická predikcia výskytu ionosférických scintilácií metódami hlbokého učenia **31**

3.1 Pochopenie cieľa . . . 32

3.2 Pochopenie dát . . . 35

3.2.1 Intenzita fázovej ionosférickej scintilácie . . . 35

3.2.2 Analýza ďalších parametrov . . . 38

3.3 Príprava dát . . . 41

3.3.1 Normalizácia . . . 42

3.3.2 Aplikovanie metódy posuvného okna . . . 42

3.3.3 Výber finálnych množín na trénovanie a testovanie 45
3.4 Modelovanie 46
3.4.1 Autoregresná predikcia o 15 minút dopredu 47
3.4.2 Post-processing 50
3.4.3 Autoregresné modely 59
3.4.4 Multivariačné modely 62
3.5 Vyhodnotenie 69
3.6 Nasadenie 71

4 Záver **73**

Zoznam príloh **80**

Zoznam obrázkov

1 – 1 Porovnanie strojového a hlbokého učenia. 4

1 – 2 Porovnanie výkonnosti hlbokého a strojového učenia v závislosti od množstva trénovacích dát. 5

1 – 3 Architektúra perceptrónu. 7

1 – 4 Ukážka architektúry doprednej neurónovej siete. 8

1 – 5 Graf funkcie ReLU. 10

1 – 6 Graf lineárnej funkcie, konkrétne ide o funkciu $f(x) = 2x$. 11

1 – 7 Spätné šírenie chýb. 14

1 – 8 Rozdiely v trénovaní modelov podľa sledovania chybovej funkcie. . 15

1 – 9 Porovnanie optimálne naučeného modelu s podučeným a preučeným modelom. 16

1 – 10 Deaktivácia neurónov podľa metódy Dropout (Srivastava et al., 2014). 16

1 – 11 Porovnanie architektúry rekurentnej neurónovej siete s architektúrou doprednej neurónovej siete. 17

1 – 12 Architektúra bloku LSTM. 18

1 – 13 Porovnanie architektúry LSTM s architektúrou BiLSTM. 19

2 – 1 Šírenie rádiových vĺn odrazom od ionosféry. 21

2 – 2 Zobrazenie ionosférickej scintilácie ako modifikovaného rádiového signálu v ionosfére. 22

2 – 3 Mapa výskytu ionosférických scintilácií (Kintner et al., 2009). . . 24

3 – 1 Matica zámen pre binárnu klasifikáciu. 33

3 – 2 Rozmiestnenie a veľkosti binov podľa počtu hodnôt σ_ϕ vytvorených z databázy CHAIN. 36

3 – 3 Ukážka priebehu intenzity ionosférickej scintilácie. 37

3 – 4 Počet záznamov parametra σ_ϕ pre roky 2013 až 2021 v rámci binu 13616. 37

3 – 5 Boxplot pre parameter *Ap index* pre roky 2013 až 2021. 39
3 – 6 Boxplot pre parameter *ASY/H* pre roky 2013 až 2021. 39
3 – 7 Boxplot pre parameter *Bz GSE* pre roky 2013 až 2021. 40
3 – 8 Boxplot pre parameter *PC(N)* pre roky 2013 až 2021 z údajov binu 13839. 40
3 – 9 Porovnanie vývoja Fisherovho skóre pre vybrané najdôležitejšie parametre. 41
3 – 10 Metóda posuvného okna. 43
3 – 11 Architektúra autoregresného modelu s rekurentnými vrstvami. . . . 49
3 – 12 Porovnanie predikcií autoregresného modelu zo 4. epochy so skutočnými testovacími hodnotami vybraných z binu 13725. 50
3 – 13 Aplikovanie vybraných *smoothing* metód na testovacie hodnoty σ_ϕ. 52
3 – 14 Aplikovanie vybraných *smoothing* metód na predikované hodnoty σ_ϕ. 52
3 – 15 Porovnanie testovacích hodnôt σ_ϕ a z nich vyjadrených CoV s predikovanými hodnotami a z nich vyjadrených CoV pre obdobie medzi 21. a 22. aprílom 2019. 55
3 – 16 Porovnanie testovacích hodnôt σ_ϕ a z nich vyjadrených CoV s predikovanými hodnotami a z nich vyjadrených CoV pre obdobie medzi 9. a 10. aprílom 2019. 56
3 – 17 Porovnanie vyhodnotení autoregresného modelu s predikciou 15 minút dopredu podľa rozdielnych nastavení hraničných hodnôt CoV. 58
3 – 18 Porovnanie autoregresných modelov pre rôzne časové posuny. . . . 61
3 – 19 Architektúra základného multivariačného modelu. 63
3 – 20 Porovnanie najlepších modelov každého druhu pre predikciu 15 minút dopredu. Zhora: predikcia perzistencie, autoregresný model, základný multivariačný model s parametrom *PC(N)*, rozšírený multivariačný model so všetkými parametrami. 66

3 – 21 Porovnanie predikcií a klasifikácií autoregresného a multivariačného modelu. 68
3 – 22 Porovnanie vývoja metrík pre rozličné nastavenia CoV autoregresného modelu s posunom 60 minút. 70

Zoznam tabuliek

2 – 1 Výsledky vyhodnotenia modelov práce McGranaghan et al. (2018). Stĺpec podmienky scintilácie obsahuje intervaly predikovaných hodnôt, z ktorých sa robila klasifikácia. 29
3 – 1 Zoznam použitých parametrov. 38
3 – 2 Počet záznamov vo finálnych množinách pre trénovanie a testovanie podľa jednotlivých modelov. 46
3 – 3 Metriky pre predikciu perzistencie a pre autoregresný model s časovým posunom 15 minút podľa rozličných nastavení CoV. Pre testovacie hodnoty bola hraničná hodnota CoV nastavená na 0,5. 57
3 – 4 Matica zámen pre vyhodnotenie predikcií autoregresného modelu s posunom 15 minút a s nastavenou hranicou CoV pre predikcie 0,01. . . 57
3 – 5 Výber najlepších autoregresných modelov podľa epochy, v ktorej boli dosiahnuté najlepšie výsledky pri klasifikácii podľa metriky TSS. . . . 59
3 – 6 Metriky pre predikciu perzistencie a pre autoregresný model s časovým posunom 30 minút podľa rozličných nastavení CoV. Pre testovacie hodnoty bola hraničná hodnota CoV nastavená na 0,5. 59
3 – 7 Metriky pre predikciu perzistencie a pre autoregresný model s časovým posunom 45 minút podľa rozličných nastavení CoV. Pre testovacie hodnoty bola hraničná hodnota CoV nastavená na 0,5. 60
3 – 8 Metriky pre predikciu perzistencie a pre autoregresný model s časovým posunom 60 minút podľa rozličných nastavení CoV. Pre testovacie hodnoty bola hraničná hodnota CoV nastavená na 0,5. 60
3 – 9 Výber multivariačných modelov podľa epochy, v ktorej dosiahli najlepšie výsledky pri klasifikácii. 64

3 – 10Metriky najlepšieho nastavenia CoV autoregresného modelu na predikciu 15 minút dopredu v porovnaní s metrikami multivariačných modelov podľa nastavení CoV. Pre skutočné testovacie hodnoty bola hraničná hodnota CoV nastavená na 0,5. 65
3 – 11Výsledky pre predikciu s posunom 60 minút dosiahnuté v práci (McGranaghan et al., 2018). 69
3 – 12Výsledky pre predikciu s posunom 60 minút dosiahnuté v tejto práci pomocou autoregresného modelu. 69

Slovník termínov

Adam Algoritmus adaptívnej optimalizácie momenta

ASPIS Autonomous Service for Prediction of Ionospheric Scintillations

BiLSTM Bidirectional Long-Short-Term Memory

CoV Coefficient of variation

EIA Ekvatoriálna ionizačná anomália

ESA European Space Agency

FN False Negative

FP False Positive

GNSS Global Navigation Satellite System

LSTM Long-Short-Term Memory

MAE Mean Absolute Error

MLP Multilayer Perceptron

MSE Mean Squared Error

P Precision

PECS Plan for European Cooperating States

R Recall

ReLU Recified Linear Unit

RNN Recurrent Neural Network

SVM Support Vector Machine

TEC Total Electron Content

TN True Negative

TP True Positive

TSS True Skill Score

Úvod

Ionosféra je vrstva zemskej atmosféry, ktorá je ovplyvňovaná prichádzajúcim žiarením zo Slnka. Je to oblasť s komplexnou dynamikou, ktorá výrazným spôsobom zasahuje do prenosu rádiových signálov. Na jednej strane tvorí odrazovú plochu pre ich šírenie po zemskom povrchu, no na druhej strane sa v nej odohrávajú javy, ktoré sú veľmi častou príčinou modifikácie, či dokonca až výpadku rádiového signálu. To môže mať negatívny vplyv, napr. pre globálne navigačné satelitné systémy. Pojmom ionosférické scintilácie sa označujú náhle zmeny v amplitúde a fáze rádiového signálu, ktoré vznikajú pri prechode ionosférou. Zdrojom vzniku narušení signálu sú nabité častice, ktorých počet sa počas zvýšenej slnečnej aktivity zvyšuje.

V tejto diplomovej práci sa zameriame na predikciu ionosférických scintilácií. Využijeme na to hlboké učenie, ktorého popularita spočíva okrem iného v schopnosti riešiť problémy, ktoré doteraz žiaden iný prístup riešiť nedokázal. Preto aplikujeme princípy hlbokého učenia na tvorbu modelov nazývaných neurónové siete, ktorých úlohou bude analyzovať ionosférické dáta a pokúsiť sa odhadnúť výskyt ionosférickej scintilácie. Keďže tento jav predstavuje značné riziko pre správnosť fungovania navigačných systémov, či komunikácie medzi vesmírnymi satelitmi a pozemnými prijímačmi, považujeme prínos informácií o možnom výskyte ionosférických scintilácií za nepopierateľný.

Táto diplomová práca vznikla vďaka spolupráci s Ústavom experimentálnej fyziky Slovenskej akadémie vied v Košiciach a spoločnosťou GlobalLogic Slovakia s.r.o. v rámci projektu ASPIS (*Autonomous Service for Prediction of Ionospheric Scintillations*) podporovaného cez program Európskej vesmírnej agentúry PECS (*Plan for European Cooperating States*). Hlavným cieľom projektu ASPIS je vytvorenie automatizovaného procesu na poskytovanie predikcií výskytu ionosférických scintilácií. Výstupy tejto diplomovej práce v podobe modelov a návrhov spracovania dát poslúžia pre projekt ASPIS k návrhu automatického systému na predikciu scintilácií. Pracovať budeme s dátami vo forme časových radov, ktoré budú použité ako vstupy

do modelov.

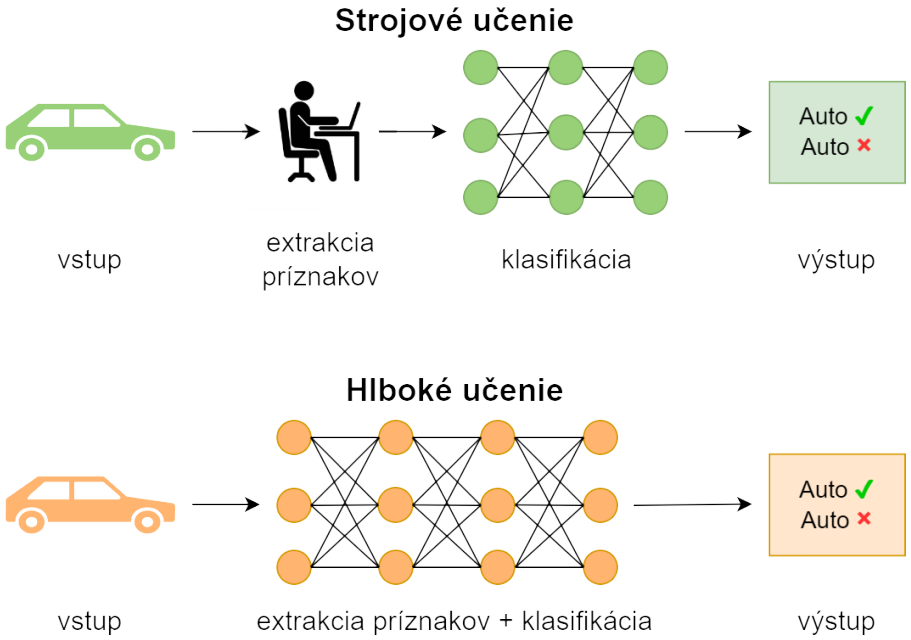
V prvej kapitole ponúkneme teoretický pohľad na hlboké učenie a neurónové siete. Spomenieme všetky aspekty hlbokého učenia, ktoré využijeme v praktickej časti. Druhá kapitola sa zaoberá doménou tejto práce, ktorou sú ionosférické dáta. V tejto kapitole sa budeme okrem iného venovať aj poznaniu súčasného stavu a analýze iných, už existujúcich riešení na túto problematiku. V tretej kapitole prejdeme do praktickej časti tejto diplomovej práce, kde budeme prostredníctvom metodiky CRISP-DM prechádzať procesom analýzy dát s cieľom objaviť užitočné znalosti a tým prispieť k výskumu kozmického počasia.

1 Hlboké učenie a neurónové siete

Pojem umelá inteligencia sa dostáva každým dňom čoraz viac do pozornosti, či už širokej verejnosti alebo médií, no v neposlednom rade sa tento pojem čoraz častejšie vyskytuje práve vo vedeckej a výskumnej sfére. Napríklad v oblastiach astronómie a astrofyziky bol v poslednej dobe zaznamenaný obrovský nárast v počte nazbieraných dát vďaka výkonnejším prístrojom a teleskopom, ktoré dokážu nazrieť stále viac hlbšie do vesmíru. S tým prichádzajú problémy spracovania a analýzy dát s cieľom identifikovať v dátach skryté vzory, na čo už nepostačujú samotné ľudské schopnosti. Práve kvôli tomu sa nielen v oblasti astrofyziky čoraz častejšie využíva umelá inteligencia v podobe strojového učenia (Sen et al., 2022). Keďže táto práca v sebe spája vyššie spomenuté aspekty, ktorými sú astrofyzikálne dáta a strojové, resp. hlboké učenie, je veľmi dôležité na začiatok ujasniť, čo umelá inteligencia, strojové učenie a hlboké učenie znamenajú a aké bude ich postavenie v tejto diplomovej práci.

1.1 Hlboké učenie

Umelá inteligencia (skr. AI, angl. *artificial intelligence*) je pojem zastrešujúci systémy, ktoré umožňujú napodobňovať ľudské správanie v oblasti rozhodovania, s cieľom riešiť zložité logické úlohy a problémy samostatne alebo s minimálnym zásahom od človeka (Janiesch et al., 2021). Zo začiatku išlo o formalizáciu znalostí do podoby pravidiel, pomocou ktorých potom stroj dokáže automaticky vyhodnotiť vstupné dáta. Keďže nie všetky znalosti je možné vyjadriť alebo sformulovať, takýto prístup mal svoje obmedzenia, s ktorými si ale dokáže poradiť strojové učenie (skr. ML, angl. *machine learning*) (Russell and Norvig, 2016). Strojové učenie, ktoré je podmnožinou umelej inteligencie, je založené na princípe automatického učenia a budovania modelov bez explicitného programovania. Klasické algoritmy strojového učenia si ale vyžadujú ľudský zásah pri výbere kľúčových vlastností a príznakov pre pochopenie rozdielov medzi vstupnými dátami (Goodfellow et al., 2016). Naproti tomu hlboké učenie (angl. *deep learning*) dokáže nájsť a extrahovať kľúčové príznaky samostatne

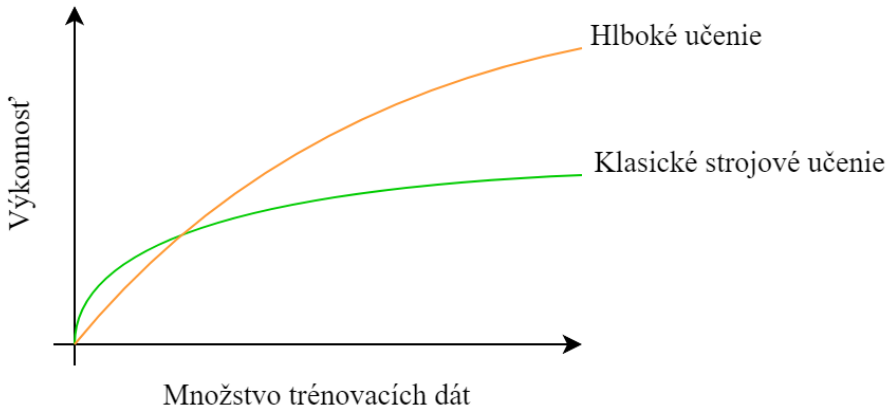


Obrázok 1 – 1 Porovnanie strojového a hlbokého učenia.

počas procesu učenia. Porovnanie obidvoch princípov je možné vidieť aj na Obrázku 1 – 1. Ak by sme mali vysvetliť vzťahy medzi umelou inteligenciou, strojovým učením a hlbokým učením, mohli by sme to zhrnúť nasledovne: strojové učenie ako celok spadá pod umelú inteligenciu a jednou z podmnožín strojového učenia je práve hlboké učenie.

Pri hlbokom učení prebieha proces učenia samostatne prostredníctvom oboznámenia sa so vzorovými príkladmi z vybranej oblasti. Jeho hlavnou výhodou je plne automatizovaný výber príznakov, čo výrazne uľahčuje tvorbu modelov oproti prístupom používaných čisto v strojovom učení (Moshayedi et al., 2022). Okrem toho človek takéto príznaky dokonca veľakrát sám o sebe ani nedokáže identifikovať, môžu zostať buď bez povšimnutia alebo môžu byť považované za nepodstatné.

Výhodou oproti strojovému učeniu je schopnosť spracovať veľké množstvo dát, vrátane tých komplexných. Hlboké učenie dokáže pracovať s rôznymi typmi dát, či už ide o zvukové, textové alebo aj obrazové dáta. Taktiež treba poznamenať, že výkonnosť hlbokého učenia je vo väčšine prípadov vyššia (Obrázok 1 – 2), vďaka



Obrázok 1 – 2 Porovnanie výkonnosti hlbokého a strojového učenia v závislosti od množstva trénovacích dát.

čomu je možné nielen dosiahnuť lepšie výsledky v porovnaní so strojovým učením, ale navyše umožňuje riešiť problémy, s ktorými si doposiaľ žiaden iný prístup nedokázal poradiť (Chollet, 2021). Na jednej strane síce hlboké učenie dokáže pracovať s veľkými dátami, no na druhej strane si aj vyžaduje veľké množstvo kvalitných trénovacích dát pre efektívne učenie. Získať vhodné a kvalitné dáta môže byť niekedy samo o sebe výzvou z pohľadu časovej náročnosti alebo problematickej dostupnosti (Talaei Khoei et al., 2023). Ďalšou výhodou hlbokého učenia je všestrannosť a opakovateľnosť. V princípe ide o opakované použitie predtrénovaného modelu, ktorý sa dotrénuje použitím dodatočných dát z oblasti konkrétneho problému, vďaka čomu nie je nutné začať proces trénovania úplne od nuly (Iman et al., 2023).

Výhody hlbokého učenia sú naozaj nepopierateľné. Aj keď má vysoké nároky na výpočtovú silu, práve pokrok v oblasti vývoja hardvéru a softvéru umožnil okrem iného dosiahnuť aj pokrok v hlbokom učení. Jeho využívanie sa stáva dostupnejšie vďaka rýchlejším procesorom a efektívnejším prístupom v oblasti paralelného spracovania procesov (LeCun, 2019). Je ale potrebné poznamenať aj jeho nevýhody, alebo lepšie povedané výzvy, ktorým je nutné čeliť. Veľmi podstatnou je interpretovateľnosť modelov hlbokého učenia, keďže kvôli komplexnosti výpočtov a zložitosti celého procesu učenia sa je v konečnom dôsledku ťažké odôvodniť ako a na základe

čoho dospel model k svojmu výsledku. Model, pri ktorom nie je možné interpretovať proces odvodenia výstupu, sa označuje aj ako čierna skrinka (angl. *black box*). Vysoká komplexnosť modelov hlbokého učenia so sebou prináša aj veľké množstvo hyperparametrov. Vhodnosť ich nastavenia sa veľakrát vyhodnotí až po natrénovaní modelu (Feurer and Hutter, 2019).

Modely hlbokého učenia nesú názov neurónové siete, keďže sa skladajú zo samostatne sa učiacich vrstiev. S narastajúcim počtom vrstiev rastie aj hĺbka modelu, čím sa dostávame k pôvodu pomenovania hlbokého učenia ako „hlbokého“. V nasledujúcich podkapitolách sú podrobnejšie vysvetlené princípy, na základe ktorých neurónové siete fungujú, z čoho sú zložené a aké sú ich vlastnosti.

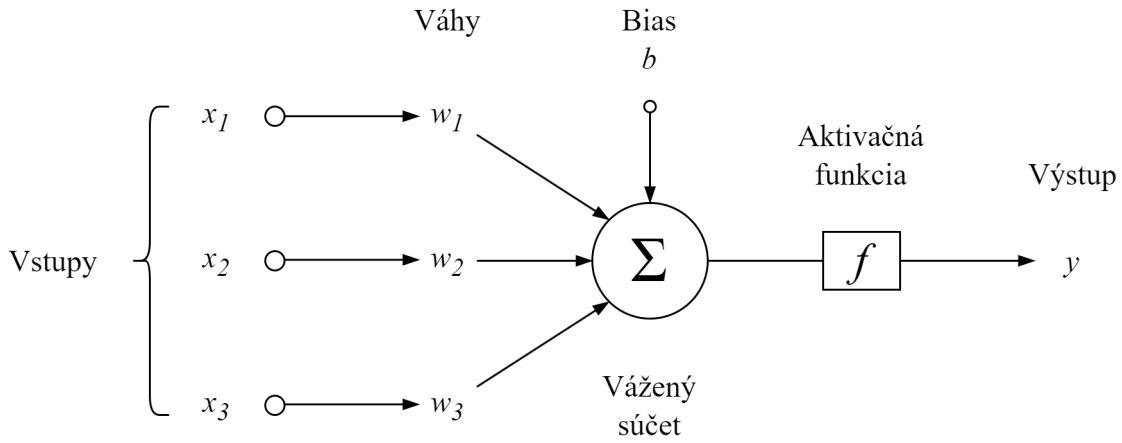
1.2 Neurónové siete

Základným stavebným prvkom neurónovej siete je perceptrón (Obrázok 1 – 3), niekedy označovaný aj ako umelý neurón na základe analógie s biologickými neurónmi v ľudskom mozgu, ktoré získavajú, spracovávajú a ukladajú informácie, ktoré sú prenášané prostredníctvom elektrického signálu. Autorom perceptrónu, ktorý sformalizoval ideu napodobnenia ľudskej schopnosti učiť sa do matematickej podoby, je psychológ F. Rosenblatt (1957). Podarilo sa mu demonštrovať, že perceptrón sa dokáže „učiť“ z dát vďaka upravovaniu váh pre vstupné dáta s cieľom čo najviac sa priblížiť očakávanému výstupu.

Úlohou perceptrónu je z prijatého vstupného vektora $\overline{x} = (x_1, x_2, ..., x_{n+1})$ vynásobením s príslušnými váhami z vektora $\overline{w} = (w_1, w_2, ..., w_{n+1})$ vyprodukovať výstup y pomocou nasledujúceho vzťahu:

$$y = f\left(\overline{w} \cdot \overline{x} + b\right) = f\left(b + \sum_{i=1}^{n+1} w_i x_i\right), \tag{1.1}$$

pričom f predstavuje aktivačnú funkciu, váhy z vektora $\overline{w}$ vyjadrujú vplyv príslušného vstupu z vektora $\overline{x}$ na celkový výstup a b je bias, ktorý sa pripočíta k váženému súčtu zo súčinu vstupných hodnôt s ich váhami, čím sa môže zlepšiť presnosť



Obrázok 1 – 3 Architektúra perceptrónu.

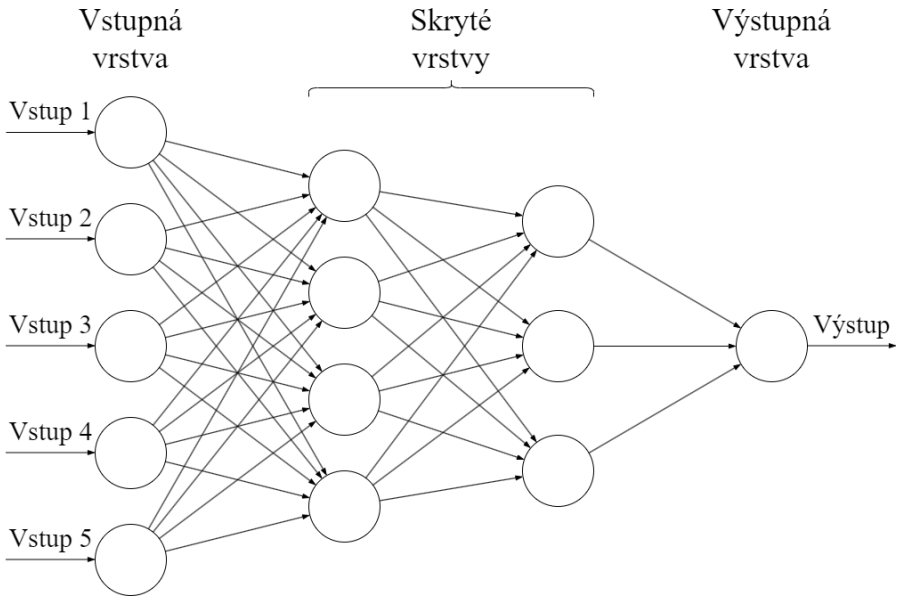
modelu (Nielsen, 2015).

Jeden perceptrón, ktorý v podstate predstavuje jednoduchú jednovrstvovú neurónovú sieť, dokáže riešiť iba lineárne separovateľné problémy. V reálnom svete má ale bohužiaľ väčšina problémov nelineárny charakter. Na ich riešenie sa preto využívajú neurónové siete zložené z vrstiev, kde každá vrstva pozostáva z viacerých umelých neurónov (Obrázok 1 – 4). V neurónových sieťach rozlišujeme tri typy vrstiev (Goodfellow et al., 2016):

- **Vstupná vrstva**, ktorej počet neurónov je rovný počtu vstupných hodnôt. Keďže k vstupným hodnotám je možný priamy prístup, o vstupnej vrstve možno hovoriť ako o viditeľnej vrstve.

- **Skrytá vrstva**, v rámci ktorej prebieha učenie. Neurónová sieť môže obsahovať viacero skrytých vrstiev, pričom počet neurónov v skrytej vrstve nie je nijako špecificky podmienený. Skryté vrstvy sa nachádzajú medzi vstupnou a výstupnou vrstvou. Ako už bolo spomenuté, počet skrytých vrstiev tvoriacich neurónovú sieť sa nazýva hĺbka modelu. Vstup pre každý neurón konkrétnej vrstvy je tvorený z výstupov neurónov predchádzajúcej vrstvy.

- **Výstupná vrstva**, ktorej výstup je zároveň výstupom celej neurónovej siete. Počet neurónov vo výstupnej vrstve závisí od typu riešenej úlohy. Ak neuró-



Obrázok 1 – 4 Ukážka architektúry doprednej neurónovej siete.

nová sieť rieši regresný problém alebo binárnu klasifikáciu, na výstupe je jeden neurón, ktorý nesie informáciu o jednej výslednej hodnote. Ak je predmetom riešenia klasifikačná úloha do viacerých tried, počet neurónov výstupnej vrstvy zodpovedá počtu všetkých tried danej úlohy.

Na spomínanom Obrázku 1 – 4 možno vidieť architektúru doprednej neurónovej siete s dvoma skrytými vrstvami. Pomenovanie doprednej neurónovej siete vychádza zo skutočnosti, že informácie sa v nej šíria iba smerom dopredu. Ak zoberieme funkcie $f^{(1)}$, $f^{(2)}$, $f^{(3)}$ a $f^{(4)}$, ktoré reprezentujú vrstvy v architektúre doprednej neurónovej siete (v tomto prípade uvažujme o architektúre na Obrázku 1 – 4) a skombinujeme ich do jednej funkcie $f(x)$, potom $f(x) = f^{(4)}\left(f^{(3)}\left(f^{(2)}\left(f^{(1)}(x)\right)\right)\right)$, kde $f^{(1)}$ je vstupná vrstva, $f^{(2)}$ je prvá skrytá vrstva, $f^{(3)}$ druhá skrytá vrstva a napokon $f^{(4)}$ je výstupná vrstva. Inak povedané, výstup z jednej vrstvy je použitý ako vstup do nasledujúcej vsrtvy (Nielsen, 2015).

1.2.1 Aktivačné funkcie

Aktivačné funkcie v neurónových sieťach zohrávajú veľmi kľúčovú úlohu a je nevyhnutné klásť dôraz na ich vhodný výber pre jednotlivé vrstvy. Doteraz sme hovorili o výstupoch umelých neurónov ako o výpočte vykonanom pomocou aktivačnej funkcie, do ktorej vstupujú prislúchajúce výstupné hodnoty z predošlých neurónov, váhy a bias. Ak by sme uvažovali o situácii, v ktorej by neurónová sieť neobsahovala aktivačné funkcie, potom by všetky neuróny vykonávali iba lineárnu transformáciu vstupov prostredníctvom váh a biasu, a výsledkom by bol iba lineárny regresný model, ktorý by neumožňoval riešiť zložité nelineárne úlohy.

Práve to je dôvod, prečo sú aktivačné funkcie v neurónových sieťach také dôležité. Ich úlohou je pridať neurónovým sieťam do výpočtov istú formu nelineárnosti pri spracovaní vstupov. Vďaka tomu je možné pre neurónovú sieť riešiť aj bežné problémy zo života, pri ktorých sú vzťahy medzi vstupmi a výstupmi komplexné a nelineárne (Ding et al., 2018). Vstupné hodnoty sa transformujú do určitého intervalu podľa typu aktivačnej funkcie. Aktivačná funkcia v konečnom dôsledku rozhoduje o tom, či sa neurón aktivuje alebo nie. (Sharma et al., 2017).

V súčasnosti je pre neurónové siete k dispozícii veľké množstvo aktivačných funkcií. Ich špecifické vlastnosti predurčujú, na aký typ úlohy a na akú vrstvu je ich aplikovanie adekvátne. V našom prípade sme v praktickej časti používali aktivačné funkcie ReLU a linear.

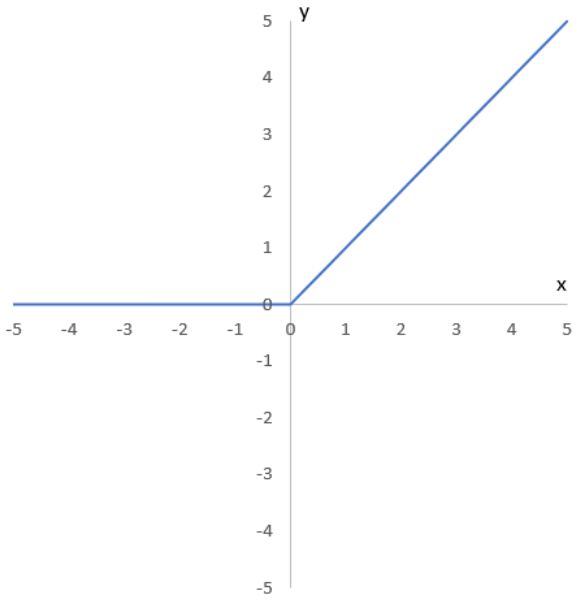
- **ReLU**

  Funkcia ReLU (angl. *Rectified Linear Unit*) je jednou z najpoužívanejších aktivačných funkcií pre skryté vrstvy v neurónových sieťach vďaka svojej efektívnosti a jednoduchosti. Jej matematický zápis

  $$f(x) = \max(0, x) \tag{1.2}$$

  hovorí o tom, že k aktivácii neurónu dôjde v prípade, ak bude hodnota na vstupe kladná. V takom prípade bude výstupom samotná vstupná kladná hod-

nota. V prípade záporného vstupu vráti nulu a neurón sa neaktivuje. Priebeh tejto funkcie možno vidieť aj na Obrázku 1 – 5.



Obrázok 1 – 5 Graf funkcie ReLU.

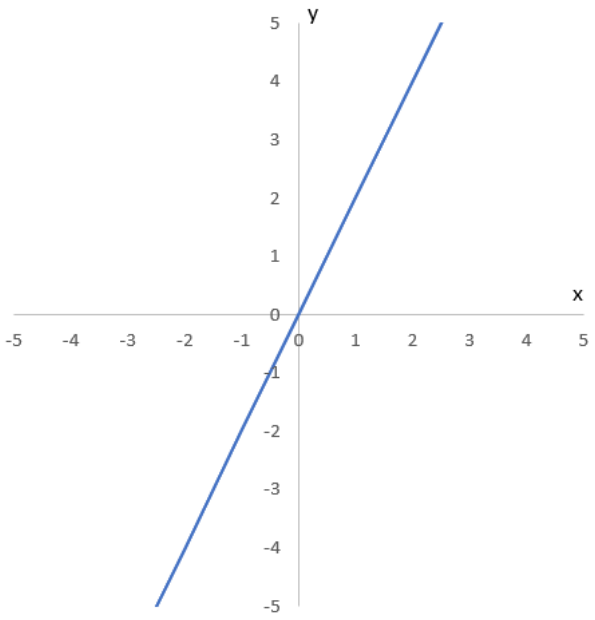
Oproti iným aktivačným funkciám má jednu veľkú výhodu týkajúcu sa výpočtovej náročnosti, keďže v sebe nezahŕňa žiadne exponenciálne alebo iné zložité výpočty. Navyše, deriváciou pre $x > 0$ je vždy konštantná hodnota 1, vďaka čomu sa pri ReLU nezvykne objavovať problém s miznúcim gradientom. Nevýhoda ReLU sa prejavuje v situácii, kedy je vstupom do funkcie záporná hodnota, pretože nulová výstupná hodnota spôsobí spomalenie a zníženie efektívnosti učenia (Ding et al., 2018).

- **Linear**

  Aktivačná funkcia linear je, ako to už aj z jej názvu vyplýva, lineárnou aktivačnou funkciou so zápisom

$$f(x) = ax, \tag{1.3}$$

  kde a je ľubovoľná konštanta (Obrázok 1 – 6). Na prvý pohľad sa môže zdať, že vloženie lineárnej funkcie do neurónovej siete nedáva zmysel. Jej použitie



Obrázok 1 – 6 Graf lineárnej funkcie, konkrétne ide o funkciu $f(x) = 2x$.

na všetkých vrstvách vrátane skrytých by malo za následok to, že aktivačná funkcia prvej vrstvy by bola totožná s aktivačnou funkciou poslednej vrstvy, čím by boli skryté vrstvy zbytočné a výsledný model by bol identický s lineárne regresným modelom (Feng and Lu, 2019). Na rozdiel od funkcie ReLU sa zvykne lineárna aktivačná funkcia používať len na výstupnej vrstve pri regresných modeloch, ktorých úlohou je predikcia spojitej hodnoty.

1.2.2 Chybové funkcie

Ak chceme dosiahnuť aby sa neurónová sieť učila a počas trénovania zlepšovala, musíme vedieť vyhodnotiť jej aktuálny stav. Preto sa používa chybová funkcia, ktorá vyjadruje rozdiel medzi priebežným výstupom neurónovej siete a očakávaným skutočným výstupom. Cieľom neurónovej siete je dosiahnuť čo najnižšiu hodnotu chybovej funkcie, alebo inými slovami, cieľom je dosiahnuť vytvorenie takých predikovaných hodnôt zo vstupu, ktoré sa budú nachádzať čo najbližšie pri skutočných hodnotách (Chollet, 2021).

Výber chybovej funkcie je podmienený typom riešenej úlohy. V prípade tejto práce budeme trénovať neurónové siete na predikciu spojitých hodnôt jedného z parametrov opisujúcich javy v ionosfére. Ako chybová funkcia sa pri regresnom probléme používa stredná kvadratická chyba (skr. MSE, angl. *Mean Squared Error*), ktorá je definovaná ako:

$$\mathrm{MSE} = \frac{1}{n} \sum_{i=1}^{n} (y_i - \widehat{y}_i)^2 \,. \tag{1.4}$$

Parameter n predstavuje počet hodnôt, y_i sú skutočné hodnoty a $\widehat{y}_i$ sú modelom predikované hodnoty. Inak povedané, pre každú skutočnú hodnotu s príslušnou predikovanou hodnotou sa vypočíta rozdiel na druhú a zo všetkých týchto rozdielov sa vypočíta priemer. Z toho vyplýva, že MSE kladie dôraz na väčšie chyby, keďže rozdiely sú umocňované.

1.2.3 Optimalizácia učenia

Uvažujme o situácií, v ktorej máme identifikovaný problém a navrhnuté riešenie založené na neurónových sieťach, podľa typu problému sme vybrali vhodnú chybovú funkciu, ktorú chceme počas učenia minimalizovať, začneme s trénovaním a dostaneme prvú hodnotu chybovej funkcie. Ako docielime, aby sa v ďalších krokoch táto hodnota zmenšovala? To je úlohou optimalizácie učenia.

Optimalizácia učenia neurónovej siete priamo súvisí s chybovou funkciou, keďže cieľom optimalizácie je minimalizovať chybovú funkciu pri učení prostredníctvom zmien vykonaných v nastavení váh a biasov takým spôsobom, aby sa dosiahla čo najvyššia presnosť. Nižšie sa nachádza bližší opis konkrétnych techník optimalizácie, ktoré boli použité v praktickej časti tejto práce.

- **Metóda gradientového zostupu**

  V metóde gradientového zostupu sa iteratívne menia nastavenia váh tak, aby sa hodnota funkcie v iterácii postupne približovala k minimálnej hodnote funkcie. Keďže gradient v danom bode funkcie vyjadruje smer, v ktorom dochádza k najväčšiemu rastu, na postupné znižovanie hodnoty funkcie sa posunú váhy

opačným smerom. Nové váhy sa vypočítajú ako:

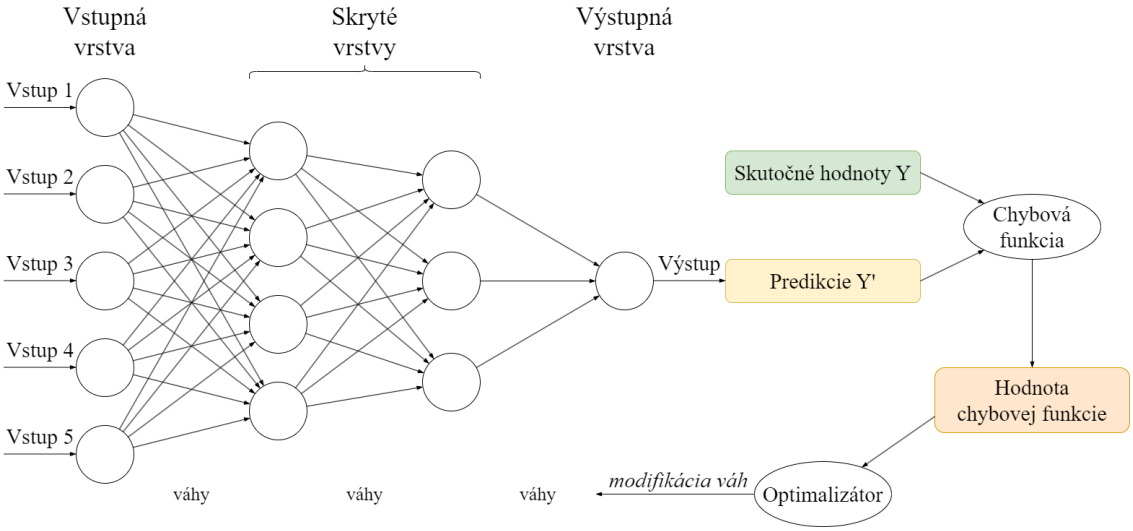
$$w_{t+1} = w_t - \epsilon_t \nabla_w J\left(w_t\right), \tag{1.5}$$

kde w sú váhy, J je chybová funkcia a ϵ_t predstavuje rýchlosť učenia (ako sa zmenia váhy pri každej iterácii) (Ruder, 2016). Pri metóde gradientového zostupu sa pri každej zmene váh vypočíta gradient na celej trénovacej množine. Postup, podľa ktorého dochádza k zmene nastavenia váh sa nazýva algoritmus spätného šírenia chyby.

- **Spätné šírenie chyby**

  Algoritmus spätného šírenia chýb (angl. *backpropagation*) sa skladá zo 4 krokov: dopredná fáza, výpočet chyby, spätná fáza a aktualizácia parametrov (Rojas and Rojas, 1996). Počas dopredného toku informácií sa v neurónovej sieti začína so vstupnou vrstvou a postupuje sa až k výstupnej vrstve. Každý neurón zo vstupov vypočíta vážený súčet, aplikuje naň aktivačnú funkciu a vypočítaný výstup posúva do ďalšej vrstvy, až kým sa nedostane do výstupnej vrstvy. Tá vypočíta výstup celej siete na základe aktuálnych váh.

  Výstup sa porovná so skutočnou trénovacou hodnotou a vypočíta sa chyba pomocou chybovej funkcie. Chybový signál sa propaguje späť cez sieť, od poslednej až po prvú vrstvu, pričom sa v každej vrstve k vypočítajú čiastkové gradienty pre každú váhu, ktorá spája neurón i v predchádzajúcej $k-1$ vrstve s neurónom j v aktuálnej vrstve k. Váhy sa aktualizujú podľa konkrétnej metódy, napríklad podľa gradientového zostupu, a teda v opačných smeroch voči smerom čiastkových gradientov. Pokračuje sa opakovaním celého postupu od dopredného toku. Celý proces sa zastaví po dosiahnutí maximálneho počtu iterácií alebo až kým sa nedosiahne požadovaná úroveň presnosti (LeCun et al., 2002). Pre lepší prehľad slúži zjednodušený nákres spätného šírenia chýb na Obrázku 1 – 7.



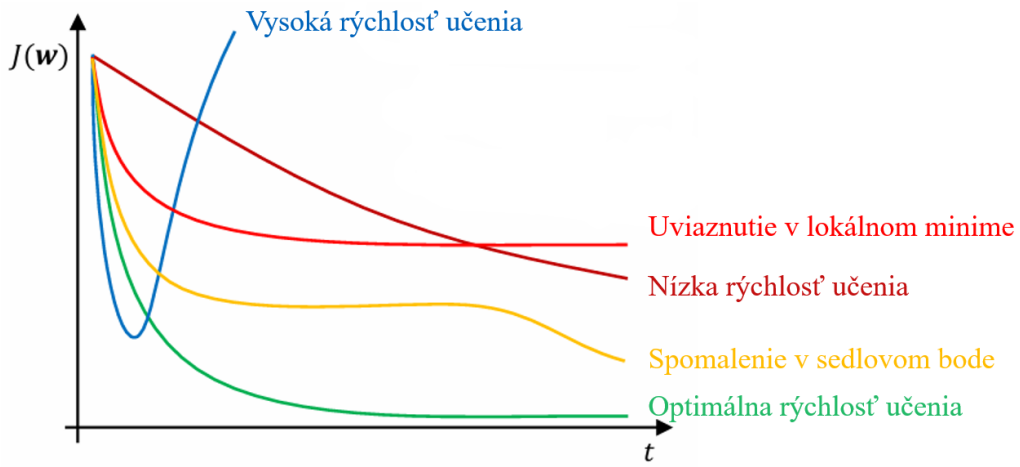
Obrázok 1 – 7 Spätné šírenie chýb.

- **Adam**

  Pri štandardnej metóde gradientového zostupu sa gradient počíta na celej trénovacej množine. Ak vezmeme do úvahy namiesto celej množiny iba jej náhodnú časť, hovoríme o stochaistickom gradientovom zostupe. Algoritmus adaptívnej optimalizácie momenta (angl. *Adaptive Moment Estimation*, skr. Adam), je optimalizátor, ktorý je rozšírením stochaistického gradientového zostupu (Kingma and Ba, 2014).

  Adam v sebe kombinuje výhody optimalizačných algortimov *Adaptive Gradient Algorithm*, skrátene AdaGrad (Duchi et al., 2011) a *Root Mean Squared Propagation*, skrátene RMSProp (Tieleman, 2012), čo z neho robí v súčasnosti najpopulárnejší a najpoužívanejší algoritmus optimalizácie. Medzi jeho výhody sa zaraďujú efektívnosť výpočtov, jednoduchá implementácia, nízka náročnosť na pamäť, či schopnosť riešiť problémy s veľkým množstvom dát alebo parametrov.

Počas trénovania neurónovej siete je nevyhnutné sledovať aktuálne hodnoty chybovej funkcie z dôvodu kontroly priebehu učenia. Práve vďaka chybovej funkcii je možné identifikovať rôzne problémy, ktoré môžu nastať pri učení (Obrázok 1 – 8).

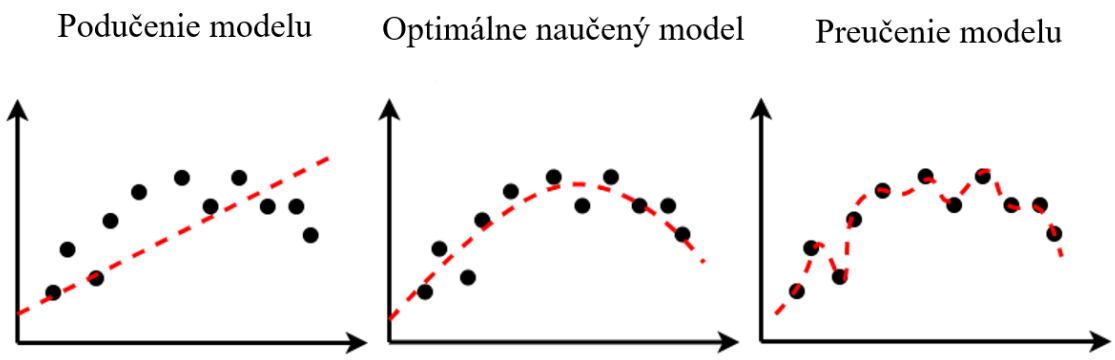


Obrázok 1 – 8 Rozdiely v trénovaní modelov podľa sledovania chybovej funkcie.

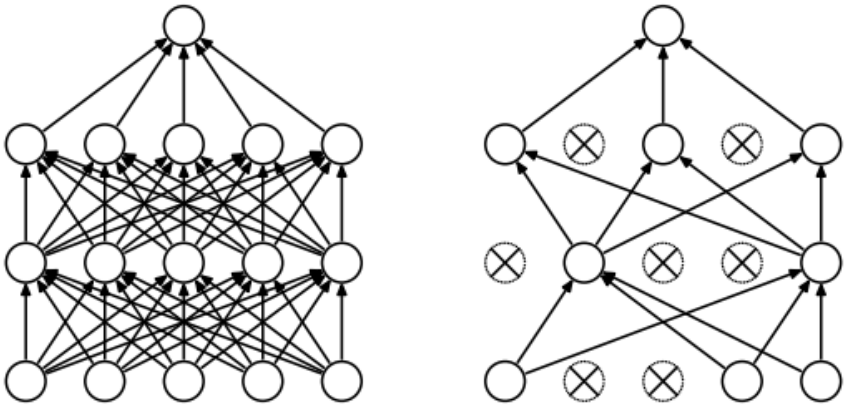
1.2.4 Regularizácia učenia

Cieľom regularizácie učenia je zabrániť modelu počas trénovania uchovávať v pamäti príliš špecifické detaily a šum z trénovacej množiny, aby nedošlo k situácii preučenia modelu (angl. *overfitting*) (Obrázok 1 – 9). V praxi to znamená, že hoci by preučený model na trénovacej množine dosiahol vysokú presnosť, pre testovaciu množinu (s iným šumom) by boli výsledky predikcie oveľa horšie, čím by sa model stal nepoužiteľný na iné dáta ako sú trénovacie. Existuje viacero regularizačných techník na zabránenie preučeniu, ktoré na zabezpečenie extrakcie dostatočne generalizovaných príznakov pri učení aplikujú potrebné kroky, aby model dosiahol rovnako dobré výsledky aj pre vstupy, ktoré ešte nikdy predtým nevidel (Goodfellow et al., 2016).

Jednou z možností je znížiť komplexnosť modelu, napríklad znížením počtu parametrov modelu. Ďalšími možnosťami sú penalizačné metódy penalizujúce váhy v modeli. Oveľa častejšie sa ale používa metóda *Dropout*, ktorá neobmedzuje komplexnosť modelu, ale znižuje varianciu učenia. Do neurónovej siete sa aplikuje ako samostatná vrstva, ktorá v každej iterácii deaktivuje náhodný výber neurónov spolu so všetkými spojeniami na ostatné neuróny (Obrázok 1 – 10). V každej iterácii dochádza k inému náhodnému výberu neurónov. Neuróny, ktoré neboli vo výbere a teda



Obrázok 1 – 9 Porovnanie optimálne naučeného modelu s podučeným a preučeným modelom.



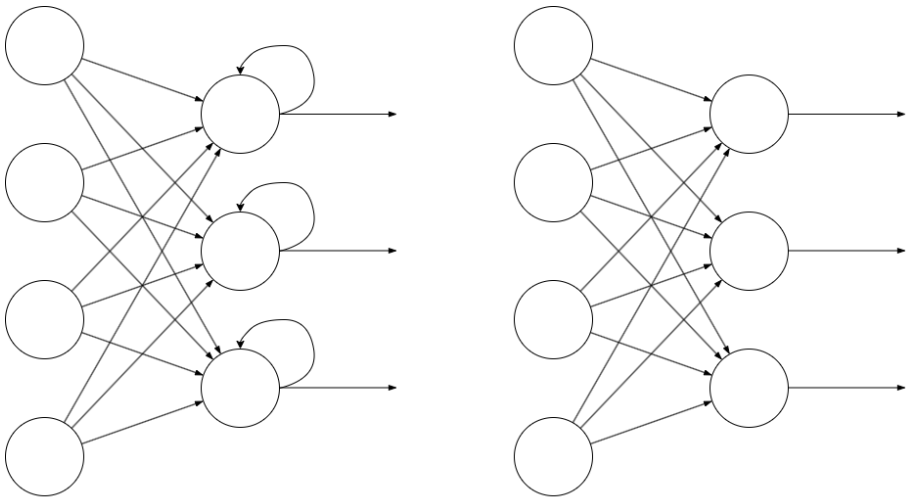
Obrázok 1 – 10 Deaktivácia neurónov podľa metódy Dropout (Srivastava et al., 2014).

neboli deaktivované, predstavujú súbor podsietí. Pri Dropout-e je potrebné nastaviť hyperparameter pravdepodobnosti deaktivácie neurónu (Srivastava et al., 2014).

1.3 Rekurentné neurónové siete

Na spracovanie sekvenčných dát sa používajú rekurentné neurónové siete (skr. RNN, angl. *Recurrent Neural Networks*). Základnou črtou sekvenčných dát je závislosť medzi jednotlivými prvkami v sekvencii. Prvkami môžu byť písmená, čísla alebo celé slová, z ktorých možno vytvoriť sekvencie vo forme časových radov alebo textov.

Kľúčová vlastnosť rekurentných neurónových sietí je ich schopnosť uchovávať v pamäti informácie z predchádzajúcich vstupov a použiť ich na spracovanie nasle-



Obrázok 1 – 11 Porovnanie architektúry rekurentnej neurónovej siete s architektúrou doprednej neurónovej siete.

dujúcich vstupov v rámci sekvencie (Yu et al., 2019). Umožňujú tak riešiť problémy, v ktorých je potrebné pochopiť kontext. Na to im slúži vnútorný cyklus, pomocou ktorého je možná úprava aktuálneho stavu siete na základe aktuálnych vstupných dát, avšak užitočnejším využitím cyklu je úprava aktuálneho stavu na základe predchádzajúcich stavov. Takýto cyklus v dopredných neurónových sieťach nie je k dispozícii. Na Obrázku 1 – 11 sú zobrazené rozdiely v architektúre medzi rekurentnou a doprednou neurónovou sieťou.

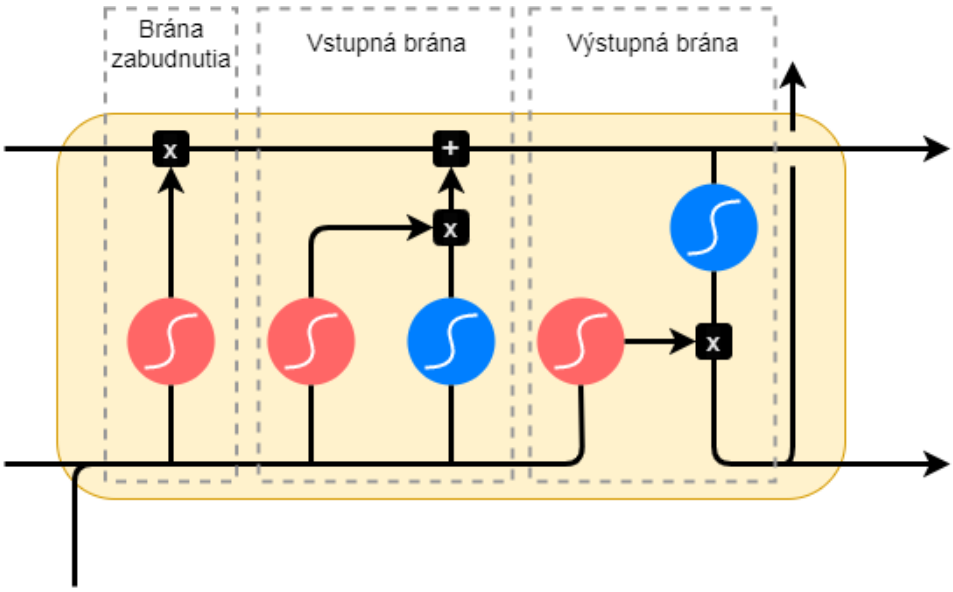
Značnou nevýhodou štandardných rekurentných neurónových sietí je ich relatívne krátkodobejšia pamäť. Takisto sa u nich zvyknú objavovať problémy s miznúcimi a explodujúcimi gradientmi (Chollet, 2021). Kvôli tomu boli navrhnuté špeciálne architektúry rekurentných sietí, ktoré použijeme aj my v praktickej časti. Ide konkrétne o architektúry LSTM a BiLSTM, ktoré sú veľmi populárne pri práci so sekvenčnými dátami.

1.3.1 LSTM

Rekurentná neurónová sieť LSTM (angl. *Long short-term memory*) je sieťou, ktorá má implementovanú špecifickú vnútornú štruktúru (Obrázok 1 – 12) vo forme bloku (bunky) pozostávajúceho z troch brán (Hochreiter and Schmidhuber, 1997):

- **vstupná brána**, ktorá vyberá dôležité informácie a ukladá ich do stavu bunky,
- **brána zabudnutia** rozhodujúca o ponechaní alebo zahodení informácií v pamäti, reguluje uchovávanie informácií,
- **výstupná brána**, ktorej úlohou je rozhodnúť o hodnote na výstupe.

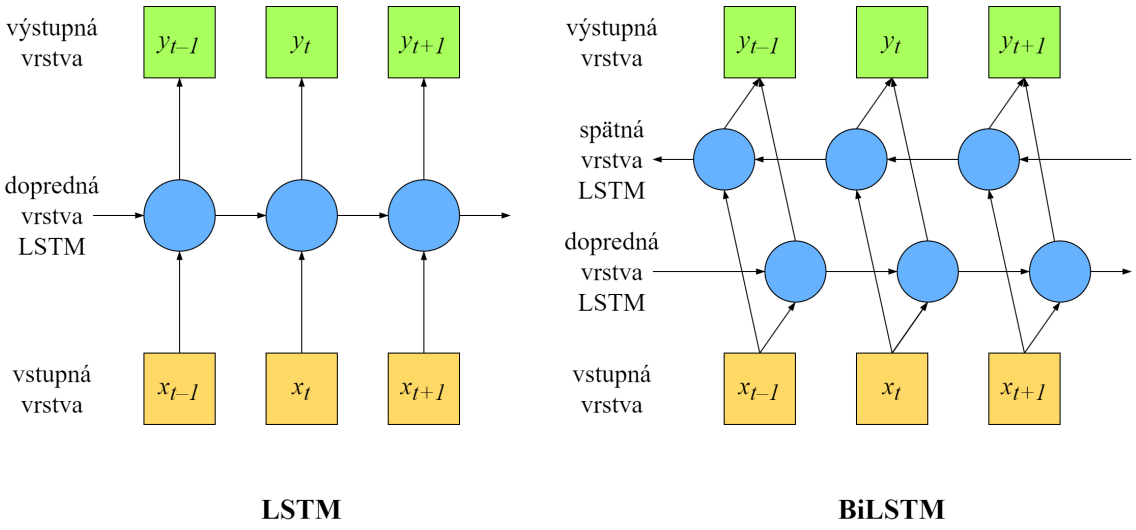
Tieto tri brány majú spoločný cieľ, ktorým je kontrolovať tok informácií cez sieť. Architektúra LSTM ponúka okrem lepšieho výkonu aj dlhšiu pamäť, vďaka čomu dokáže spracovávať dlhé sekvencie napríklad vo forme textov alebo hodnôt zaznamenaných meraní v priebehu času (Yu et al., 2019).



Obrázok 1 – 12 Architektúra bloku LSTM.

1.3.2 BiLSTM

BiLSTM (angl. *Bidirectional long short-term memory)* je ďalším typom špeciálnej architektúry rekurentnej neurónovej siete. Skladá sa totiž z dvoch samostatných skrytých vrstiev LSTM (Obrázok 1 – 13). Tie umožňujú spracovanie informácií v obidvoch smeroch, čo z architektúry BiLSTM robí užitočný nástroj pre spracovávanie prirodzeného jazyka (Schuster and Paliwal, 1997).



Obrázok 1 – 13 Porovnanie architektúry LSTM s architektúrou BiLSTM.

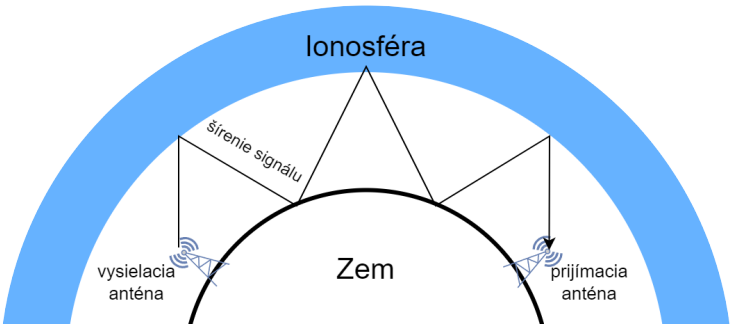
2 Ionosférické dáta

2.1 Ionosféra

Zemská atmosféra sa z pohľadu zmeny teploty s nadmorskou výškou skladá z piatich vrstiev zoradených od najbližšej po najvzdialenejšiu vrstvu k zemskému povrchu nasledovne: troposféra, stratosféra, mezosféra, termosféra a exosféra. Okrem uvedeného obvyklého delenia sa na rozloženie atmosféry možno pozrieť aj z iných uhlov pohľadu. Ak totiž uvažujeme o časticiach, ktoré sa v atmosfére nachádzajú, ich množstve a náboji, nachádzame v oblasti od 80 km do 1000 km vrstvu iónov a elektrónov nazývanú ionosféra. Prechádza súbežne cez termosféru a oblasti mezosféry a exosféry, môžeme ju preto zaradiť do takzvanej hornej atmosféry Zeme. Táto oblasť je zaujímavá tým, že tvorí rozhranie medzi procesmi, ktoré sa odohrávajú vo vesmíre a procesmi odohrávajúcimi sa na Zemi. Zhora na ňu pôsobí slnečné žiarenie a kozmické počasie, zdola zasa atmosférické počasie a klíma v blízkosti zemského povrchu. Kvôli týmto nepretržitým vplyvom predstavuje ionosféra veľmi dynamické prostredie (Heelis and Maute, 2020).

Hlavnou príčinou výskytu nabitých častíc v ionosfére je už spomínané ultrafialové (UV) a röntgenové žiarenie pochádzajúce zo Slnka, ktoré smeruje k Zemi a prechádza atmosférou. Ionosféra toto žiarenie absorbuje, čím dochádza k fotoionizácií. Ide o proces, pri ktorom atómy a molekuly absorbujú elektromagnetické žiarenie z dopadajúceho fotónu, v dôsledku čoho sa z atómov a molekúl uvoľňujú elektróny. Výsledkom fotoionizácie je teda vytvorenie iónov (častice, ktoré stratili alebo získali elektrón) a voľných elektrónov. Keďže celý tento proces závisí od slnečného žiarenia, počas dňa sa ionosféra a percentuálne zastúpenie iónov v nej mení. Kvôli absencii prichádzajúcej energie zo Slnka počas noci sa početnosť iónov v ionosfére znižuje. To má za príčinu zaniknutie, resp. oslabenie spodnej časti ionosféry, pričom vrchná časť ionosféry pretrváva aj počas noci (Pfaff, 2012).

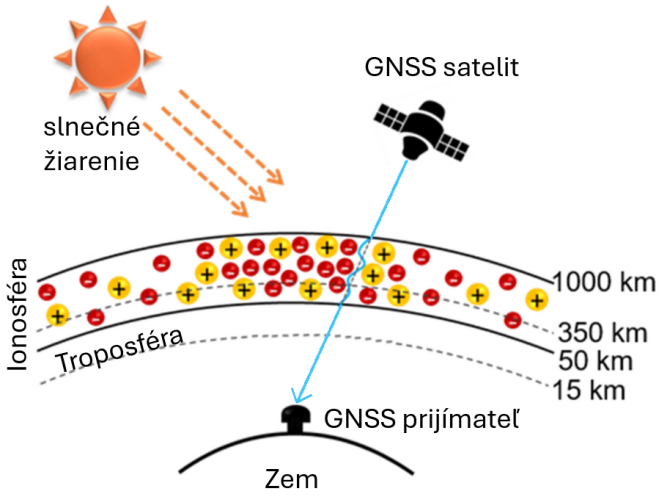
Existencia ionosféry zabezpečuje ochranu na Zemi pred škodlivými účinkami UV žiarenia, keďže ho dokáže čiastočne absorbovať (Rishbeth, 1998). Jej ďalšou vlast-



Obrázok 2 – 1 Šírenie rádiových vĺn odrazom od ionosféry.

nosťou je schopnosť odrážať krátke rádiové vlny určitých frekvencií vysielaných zo Zeme, ktoré presmeruje naspäť k Zemi. Tento proces je možné opakovať, vďaka čomu sa rozširuje celkový dosah signálu (Obrázok 2 – 1). Tento princíp sa využíval ako spôsob komunikácie na veľké vzdialenosti. Problémom bola nespoľahlivosť z dôvodu zmien nastávajúcich v ionosfére, ktoré súvisia so zmenami slnečnej aktivity v závislosti od slnečného cyklu, alebo kvôli oslabeniu časti ionosféry počas noci. V súčasnosti sa používa skôr na lokálny a regionálny prenos signálu (Rawer, 2013).

Ak sa pozrieme na komunikáciu s objektmi za zemskou atmosférou, aj tu zohráva ionosféra kľúčovú úlohu. Prenos rádiového signálu medzi satelitmi vo vesmíre a pozemnými stanicami sa využíva pre telekomunikáciu, prenos dát, navigačné systémy, monitorovanie Zeme, skúmanie vesmíru a mnoho ďalších. Ide o veľmi dôležité aplikácie v praxi, ktoré každodenne využívame. Práve pre globálne navigačné satelitné systémy (skr. GNSS, angl. *Global Navigation Satellite System*), ktoré sa používajú na vysoko presný výpočet polohy a času, majú zmeny nastávajúce v ionosfére negatívny charakter. Náhla zmena hustoty iónov a elektrónov môže ovplyvniť amplitúdu a fázu rádiového signálu takým spôsobom, že dokáže signál modifikovať, či dokonca sa môže tento signál vytratiť, v dôsledku čoho sa presnosť výpočtov znižuje a pre používateľov navigačných systémov prináša problémy s určovaním polohy (Hlubek et al., 2014). Tento jav, kedy v ionosfére nastávajú náhle modifikácie a poruchy rádiového signálu, sa nazýva ionosférická scintilácia (Obrázok 2 – 2).



Obrázok 2 – 2 Zobrazenie ionosférickej scintilácie ako modifikovaného rádiového signálu v ionosfére.

2.2 Ionosférická scintilácia

Jav ionosférickej scintilácie bol prvýkrát pozorovaný v rámci merania rádiových signálov hviezd počas výskytu polárnej žiary (Maxwell and Little, 1952). Neskôr, vďaka vyslaniu a uvedeniu satelitov do prevádzky, boli ionosférické scintilácie pozorovateľné prostredníctvom signálov vysielaných zo satelitov (Forte et al., 2022). Dopady ionosférických scintilácií na našu spoločnosť odvtedy za posledné desaťročia iba narastali a v súčasnosti sú výrazným fenoménom kozmického počasia so širokým spektrom dopadov pre námornú a pozemnú prepravu, letectvo a prevádzku satelitov na orbite (Pi et al., 2017).

Pri ionosférickej scintilácii dochádza k silnému kolísaniu fázy alebo amplitúdy rádiového signálu pri jeho prechode ionosférou. Môžeme preto hovoriť o dvoch druhoch scintilácií:

- **amplitúdové scintilácie**, pri ktorých môže dochádzať k slabnutiu signálu (Davies, 1990),
- **fázové scintilácie**, ktoré predstavujú problém pre prijímanie signálu. Pri silných fázových scintiláciách môže dôjsť u vysielaného signálu k Dopplerovmu

javu, ktorý spôsobuje posun frekvencie signálu, čím prijímač nedokáže tento signál prijať alebo z neho správne extrahovať informácie (Rama Rao et al., 2006).

Amplitúdové scintilácie sa popisujú parametrom S_4, ktorý je definovaný ako normalizovaná štandardná odchýlka intenzity prijatého signálu (Briggs and Parkin, 1963) a počíta sa podľa vzťahu

$$S_4 = \sqrt{\frac{\langle I^2 \rangle - \langle I \rangle^2}{\langle I \rangle^2}}, \tag{2.1}$$

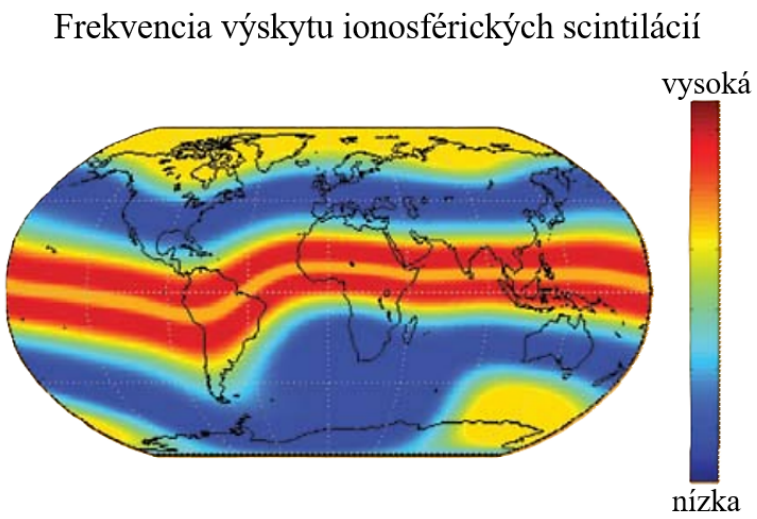
kde I predstavuje intenzitu signálu. Hodnota $\langle I \rangle$ reprezentuje priemernú hodnotu zo skúmaného časového intervalu, ktorý zvykne byť stanovený na 60 sekúnd. Amplitúdové scintilácie s hodnotami S_4 väčšími ako 0,3 sú považované za závažné scintilácie (Aarons, 1982).

Na popis intenzity fázovej scintilácie sa používa parameter σ_ϕ. Ten je definovaný ako štandardná odchýlka fázy signálu za určitý časový interval (Yeh and Liu, 1982). Počíta sa ako

$$\sigma_\phi = \sqrt{\langle \phi^2 \rangle - \langle \phi \rangle^2}, \tag{2.2}$$

kde ϕ je fáza signálu, ktorý bol zbavený trendu. Hodnota $\sigma_\phi = 0,1$ rad ($\approx 6°$) sa považuje za spodnú hraničnú hodnotu pre detekciu scintilácie (Jiao et al., 2013).

Výskyt ionosférických scintilácií záleží od polohy vzhľadom k magnetickému poľu Zeme (Obrázok 2 – 3). Z pohľadu scintilačnej aktivity môžeme na Zemi rozlišovať tri oblasti. Prvou je oblasť v okolí geomagnetického rovníka siahajúca až do $\pm 20°$ (nízkej) geomagnetickej šírky. Typické sú pre ňu amplitúdové scintilácie, ktorých intenzita výrazne narastá po západe slnka s dosiahnutím maxima približne pred polnocou. Zároveň ide o oblasť s najfrekventovanejším výskytom scintilácii (Priyadarshi, 2015). Pod druhý typ spadajú oblasti v rozmedzí $60°$ až $90°$ (vysokej) geomagnetickej šírky v okolí obidvoch pólov. V tejto oblasti sa vyskytujú fázové scintilácie, najviac počas solárneho maxima, počas zimy a počas výskytu polárnej žiary. Frekvencia ich výskytu je o niečo nižšia než v oblastiach prvého typu. Treba ale podotknúť, že počas



Obrázok 2 – 3 Mapa výskytu ionosférických scintilácií (Kintner et al., 2009).

geomagetických búrok dochádza k vzniku silných fázových scintilácií, ktoré môžu byť sprevádzané aj amplitúdovými scintiláciami (Aquino et al., 2005). Do tretej oblasti patria všetky zvyšné oblasti so strednou geomagnetickou šírkou, v ktorých je výskyt ionosférických scintilácií nízky a scintilácie, ktoré sa aj vyskytnú, majú oveľa nižšiu intenzitu (Spogli et al., 2009). Keďže pre rôzne oblasti sú typické rôzne druhy scintilácií, ktoré popisujú individuálne parametre, je dôležité rozlišovať dáta podľa oblastí, z ktorých pochádzajú.

Ionosféru do značnej miery ovplyvňuje slnečné žiarenie, ktoré je v prípade slnečných erupcií ešte intenzívnejšie a okrem ionosférických scintilácií môže spôsobiť aj geomagnetické búrky. Keďže sa ionosféra prelína aj s ďalšími atmosférickými vrstvami, dokonca aj s magnetosférou, nemožno vylučiť vplyv parametrov popisujúcich iné javy kozmického počasia na ionosférické scintilácie.

2.3 Analýza súčasného stavu

2.3.1 Kľúčové pojmy

Keďže už vieme, aký vplyv má na nami používané technológie ionosférická scintilácia a aké riziká predstavuje, našu pozornosť sme pri analýze súčasného stavu upriamili

zo širokej oblasti javov vyskytujúcich sa v ionosfére smerom k ionosférickým scintiláciám. Tie stáli aj za motiváciou tejto diplomovej práce, ktorej bola najmä podpora pri predikovaní výskytu ionosférických scintilácií. Náplňou tejto práce je využitie hlbokého učenia, a preto boli na vytvorenie predikčných modelov vybrané neurónové siete. Na základe uvedených informácií boli odvodené kľúčové slová, pomocou ktorých sme hľadali relevantné zdroje pre analýzu literatúry, resp. analýzu súčasného stavu v oblastiach analýzy ionosférických dát a predikcie výskytu ionosférických scintilácií. Použité kľúčové slová boli nasledovné:

- ionosféra (angl. *ionosphere*),
- ionosférické dáta (angl. *ionospheric data*),
- ionosférická scintilácia (angl. *ionospheric scintillation*),
- predikcia ionosférických scintilácií (angl. *prediction of ionospheric scintillation*),
- hlboké učenie (angl. *deep learning*),
- neurónové siete (angl. *neural network, artificial neural network*).

2.3.2 Analýza literatúry

Analýza ionosférických dát je široká oblasť, do ktorej spadá množstvo výskumov. V kontexte dolovania znalostí z dát (angl. *data mining*) sa najčastejšie jedná o úlohy predikovania hodnôt parametra *total electron content* (skr. TEC). Ďalej nasledujú výskumy zaoberajúce sa predikciou parametrov S_4 a σ_ϕ. Práve tieto posledné dva spomenuté sú parametrami popisujúcimi intenzitu ionosférických scintilácií, preto sme sa zaoberali práve analýzou publikácií, ktorých predmetom boli parametre S_4 alebo σ_ϕ. Ako už bolo spomenuté, výskyt ionosférických scintilácií je závislý najmä od polohy, preto budeme aj analyzované práce ďalej rozlišovať podľa miesta zdroja dát, ktoré boli predmetom výskumu. Predikcia ionosférických scintilácií je ale vo

všeobecnosti náročná, keďže sa odvíja od premenlivosti stavu ionosféry a slnečnej aktivity (Kintner et al., 2007).

V roku 1977 ponúkol R. K. Crane teoretický základ k ionosférickým scintiláciám. Vo svojej publikácii popisuje scintilácie z fyzikálneho hľadiska a uvádza matematické modely pre fyzikálne parametre ionosféry. Neskôr v roku 1980 boli publikované výsledky experimentov, počas ktorých sa autori snažili objasniť fyzikálne mechanizmy stojace za vznikom ionosférických scintilácií v oblastiach okolo rovníka a pólov (Fejer and Kelley, 1980). Objavenie týchto poznatkov spolu s radou ďalších vedeckých prác bolo kľúčovým faktorom v neskorších snahách modelovať scintilácie, či dokonca ich vedieť aj predpovedať.

V súčasnosti existuje hneď niekoľko modelov, ktoré sú založené na teórii šírenia rádiových vĺn, ktorých cieľom je modelovať efekt scintilácie (Priyadarshi, 2015). Napríklad model WBMOD (angl. *WideBand MODel*), dokáže na základe polohy, dátumu, času a geofyzikálnych podmienok odhadnúť dopad ionosférickej scintilácie prostredníctvom fyzikálnych parametrov intenzity scintilácie S_4 a σ_ϕ (Secan et al., 1995). Ďalším modelom, ktorý počíta hodnoty scintilačných indexov z globálneho hľadiska je model GISM (angl. *Global Ionospheric Scintillation Model*) (Béniguel, 2002). Problém, resp. obmedzenie takýchto teoretických modelov ale je, že nedokážu predikovať samotný výskyt scintilácií. V podstate iba modelujú používateľom zadanú situáciu a scenár v ionosfére. Tieto doposiaľ spomenuté modely neponúkli využitie procesu dolovania v dátach na objavovanie nových znalostí.

Pri hľadaní existujúcich riešení pre predikciu výskytu ionosférických scintilácií sa spočiatku môže javiť, že tento problém je rozsiahlo preskúmaný vzhľadom na veľké množstvo publikácií venujúcich sa tejto problematike. Háčik ale spočíva v tom, že drvivá väčšina z nich zameriava svoju pozornosť práve na amplitúdové ionosférické scintilácie v oblasti rovníka. Ich predikcia je ľahšie riešiteľným problémom z dôvodu ekvatoriálnej ionizačnej anomálie (skr. EIA). Táto anomália sa pravidelne opakuje, čo má za následok šancu na lepšiu predikciu a tak sa jej dostáva väčšej pozornosti výskumníkov (McGranaghan et al., 2018).

Ako prví využili podľa svojich vlastných slov dolovanie v dátach pre analýzu a predikciu ionosférických scintilácií v oblasti Brazílie Rezende et al. (2010). V svojej práci použili modely rozhodovacích stromov, ktoré vytvorili pomocou algortimu CART (angl. *Classification and Regression Tree*) a použitím metódy *bagging*. Na trénovanie a testovanie boli použité dáta z januára a marca 2000 a z februára, marca, novembra a decembra roku 2001. Tieto dáta pochádzali z mesta Sao Luis v Brazílii. Ich cieľom bolo predikovať hodnotu parametra S_4, ktorú považovali za hlavný znak ionosférickej scintilácie, najprv hodinu dopredu a následne skúšali predikciu aj viacero hodín dopredu. Na vstupe boli okrem parametra S_4 aj vertikálna driftová rýchlosť, index K_p a tok slnečného žiarenia. Pri hodinovej predikcii bola stredná kvadratická chyba približná hodnote 0,053. Po zvyšovaní predikcie o viac hodín dopredu stúpala aj chyba. Na záver boli autori s výsledkami spokojní, keďže sa im podarilo otestovať aj situácie, kedy výsledkom bola a aj mala byť nulová hodnota pre S_4 (scintilácia sa nevyskytla). Okrem toho ale poukazujú na nevýhodu, ktorou je nepoužiteľnosť modelu pre iné geografické šírky.

S ionosférickými dátami z ekvatoriálnej oblasti nad Brazíliou pracovali aj De Lima et al. (2015). Tieto dáta pochádzali taktiež zo Sao Luis v Brazílii, avšak zachytávali väčší časový rozsah. Išlo o roky 2001 až 2003 a 2010 až 2013. Ich cieľom bolo podobne ako v predchádzajúcej práci predikovať hodnotu parametra S_4. Na predikciu ale použili až 13 parametrov. Čo je v tomto prípade zaujímavé, je rozdelenie hodnôt parametra S_4 podľa hraničnej hodnoty 0,5. Pre dáta pod touto hodnotou parametra S_4 hovorili o slabej scintilácii a naopak, pre väčšiu hodnotu hovorili o silnej scintilácii. Týmto prístupom riešili problém ionosférických scintilácií hneď z dvoch pohľadov:

- predikcia hodnôt parametra S_4,
- binárna klasifikácia predikovanej hodnoty do 2 tried: slabá a silná scintilácia.

Tento prístup napovedá, že niekedy môžu byť samotné číselné hodnoty nedostatočne užitočné pre bežné použitie, keďže nemusia mať takú výpovednú hodnotu pre rie-

šenie problému, ako môže mať zaradenie do triedy. Pre predikciu použili neurónovú sieť *multilayer perceptron* (skr. MLP) s aplikovaným spätným šírením chýb pre nastavenie váh v sieti. Konkrétna použitá architektúra pozostávala z dvoch podsietí, z ktorých každá mala 13 vstupných neurónov, 2 skryté vrstvy obsahujúce 13 neurónov a 1 výstupný neurón. Tieto 2 výstupné neuróny skrytých vrstiev boli na vstupe pre výstupný neurón celej siete s lineárnou aktivačnou funkciou. Po natrénovaní modelu a získaní predikovaných hodnôt z testovacích dát boli tieto hodnoty klasifikované podľa sily scintilácie. Pri klasifikácii dosiahli pre metriku presnosti (*accuracy*) hodnotu 0,99.

Čo ale s fázovými ionosférickými scintiláciami? Ako už bolo spomenuté, pre vyššie geografické šírky neexistuje jav podobný EIA a predikcia je v takom prípade ťažšia. Doteraz najlepšie nájdené riešenie pre automatickú predikciu výskytu fázových ionosférických scintilácií, s ktorým sa budeme v praktickej časti tejto práce porovnávať, bolo založené na algoritme strojového učenia, konkrétne išlo o metódu podporných vektorov (skr. SVM, angl. *support vector machine*) (McGranaghan et al., 2018). Táto práca s názvom *New Capabilities for Prediction of High-Latitude Ionospheric Scintillation: A Novel Approach With Machine Learning* predstavuje prvý prípad použitia SVM na predpovedanie fázovej scintilácie. Keďže v svojom čase ešte nemali možnosť s niekým porovnávať dosiahnuté výsledky, porovnávali sa s predikciou perzistencie. Jej hlavnou myšlienkou je, že hodnota sledovaného parametra bude v skúmanej budúcnosti rovnaká ako je aj teraz. Pre tento výskum boli použité dáta získané z globálnych navigačných systémov. Po predikcii opäť nasledovala klasifikačná úloha, v tomto prípade boli ale použité triedy: je to scintilácia, nie je to scintilácia. Výsledky vyhodnocovali pomocou metriky TSS (angl. *True Skill Score*). Po klasifikácii použili spracovanie výstupov, kedy sledovali koľko po sebe idúcich bodov bolo označených ako scintilácia. Cieľom bolo minimalizovať vplyv rôznych výrazných odchýlok pri klasifikácii. Ich dosiahnuté výsledky sú uvedené v Tabuľke 2 – 1. Napr. v prvom riadku sú výsledky modelu, ktorý predikoval hodnotu parametra o 1 hodinu dopredu, pričom ak táto jedna hodnota bola vyššia ako 0,1, výsledok bol

predikcia	podmienky scintilácie	FN	FP	TN	TP	TSS
o 1 hod.	1 hodnota	1 071	32 188	89 718	3 252	0,49
o 1 hod.	3 hodnoty za sebou	638	18 218	106 156	1 217	0,51
o 1 hod.	6 hodnôt za sebou	267	7 934	117 626	402	0,54
o 1 hod.	priemer z 15 hodnôt	1 053	29 337	92 728	3 111	0,51
o 3 hod.	1 hodnota	940	37 865	83 961	3 463	0,48
o 3 hod.	3 hodnoty za sebou	824	16 865	107 316	1 224	0,46
o 3 hod.	6 hodnôt za sebou	434	8 376	117 055	364	0,39
o 3 hod.	priemer z 15 hodnôt	879	34 193	87 619	3 538	0,52

Tabuľka 2 – 1 Výsledky vyhodnotenia modelov práce McGranaghan et al. (2018). Stĺpec podmienky scintilácie obsahuje intervaly predikovaných hodnôt, z ktorých sa robila klasifikácia.

klasifikovaný ako scintilácia.

Ako sme už v úvode spomínali, tak táto diplomová práca je realizovaná v spolupráci s Ústavom experimentálnej fyziky Slovenskej akadémie vied v rámci prebiehajúceho projektu ASPIS (celým názvom *Feasibility study of data-driven Autonomous Service for Prediction of Ionospheric Scintillations*), ktorý je podporovaný cez program Európskej vesmírnej agentúry PECS. Cieľom projektu ASPIS je vyvinúť autonómny systém schopný poskytovať presné predpovede výskytu fázovej ionosférickej scintilácie, čo umožní zlepšiť satelitnú komunikáciu a navigáciu. Preto sme doterajšie výsledky a poznatky z tohto projektu zaradili medzi analýzu súčasného stavu, keďže poukazujú na aktuálny vývoj riešenia tohto problému a poskytujú odrazový most aj pre smerovanie tejto diplomovej práce. V čase začatia experimentov v praktickej časti tejto diplomovej práce boli predbežným výstupom projektu ASPIS autoregresné a multivariačné modely na predikciu parametra σ_ϕ. Výber vstupných parametrov pre multivariačné modely prebiehal prostredníctvom analýzy rozptylu ANOVA. Z vyše 50 parametrov sa ako najvhodnejšie ukázali *Ap index*, *ASY/H*, *PC(N) index* a *Bz*. Cieľom bolo pretransformovať predikované spojité hodnoty na binárne označujúce výskyt fázovej scintilácie. Doposiaľ na to bola používaná hra-

ničná hodnota 0,1 a následné spracovanie týchto označených dát podobne ako v práci McGranaghan et al. (2018). Problémom bolo vysoké množstvo falošne pozitívnych prípadov, preto sa vynára otázka, ako pristupovať ku klasifikácií týchto hodnôt efektívnejšie.

3 Automatická predikcia výskytu ionosférických scintilácií metódami hlbokého učenia

S príchodom nových technológií, ktoré nám uľahčujú život alebo umožňujú skúmať úplne nové možnosti, sa vynára aj jedno podstatné riziko. Ako spoločnosť sa totiž postupne čoraz viac stávame závislými od nových vymožeností, čo platí napríklad aj pre globálne navigačné satelitné systémy. Tie sú v súčasnosti najpoužívanejším nástrojom na určovanie presnej geografickej polohy a očakáva sa, že s príchodom autonómnych vozidiel, internetu vecí (IoT) a budovaním inteligentných miest budú nároky na GNSS ešte viac rásť (Egea-Roca et al., 2022).

Čoraz intenzívnejší výskum vesmíru, ktorého súčasťou je aj skúmanie javov ovplyvňujúcich život na Zemi, je uskutočňovaný vďaka komunikácii medzi pozemnými prijímacími stanicami a satelitmi vo vesmíre, podobne ako tomu je pri GNSS. Mohli by sme teda zhrnúť, že prenos signálu medzi vesmírom a Zemou je pre nás kriticky dôležitý. Okrem toho už vieme, že ho negatívne ovplyvňujú, modifikujú, či dokonca prerušujú náhle zmeny v ionosfére, čo označujeme pojmom ionsférická scintilácia. Informácia o možnom výskyte scintilácie preto môže byť užitočným varovným signálom pre používateľov spomínaných systémov.

V tejto časti sa budeme venovať predikcii výskytu ionosférických scintilácií pomocou metód hlbokého učenia. Keďže sa táto diplomová práca zaoberá analýzou vzorky dát s cieľom objaviť nové znalosti, z pohľadu konceptuálneho rámca použijeme známu metodológiu CRISP-DM (angl. *Cross Industry Standard Process for Data Mining*). Pozostáva zo 6 fáz (Wirth and Hipp, 2000):

1. **pochopenie problému** – pochopenie cieľa a pochopenie požiadaviek, stanovenie kritérií úspechu,
2. **pochopenie dát** – oboznámenie sa so zozbieranými dátami,
3. **príprava dát** – výber, čistenie a transformácia dát s cieľom pripraviť dáta pre modelovanie,

4. **modelovanie** – výber a aplikovanie metódy modelovania spolu s vyhodnotením na testovacej množine,

5. **vyhodnotenie** – porovnanie dosiahnutých výsledkov so stanovenými cieľmi,

6. **nasadenie** – popis plánu implementácie vytvorených modelov do praxe.

Na spracovanie dát a realizáciu experimentov bol použitý programovací jazyk Python. Pri práci boli použité aj jeho knižnice ako napríklad Pandas, Numpy, Math, Matplotlib, Datetime, Scikit-learn, Keras a TensorFlow.

3.1 Pochopenie cieľa

Hlavným cieľom tejto diplomovej práce je vytvorenie modelov schopných automaticky predikovať výskyt ionosférických scintilácií pomocou hlbokého učenia. Keďže táto práca bola tvorená súčasne s projektom ASPIS, našim cieľom bolo vyprodukovať výstupy, ktoré by predstavovali prínos pre spomínaný projekt. Preto bol aj náš cieľ podmienený cieľu projektu ASPIS, ktorý mal poskytovať predpovede výskytu fázovej ionosférickej scintilácie. Kvôli tomu sme našu pozornosť venovali špeciálne fázovým ionosférickým scintiláciam, ktoré ako už bolo spomenuté, sa vyskytujú vo vysokých geomagnetických šírkach v širšej oblasti okolo pólov. Náš cieľ teda môžeme pre zjednodušenie transformovať do troch krokov:

1. vytvorenie modelov neurónových sietí na predikciu parametra popisujúceho fázové ionosférické scintilácie σ_ϕ (regresná úloha),

2. návrh a aplikovanie vhodných metód spracovania výstupov neurónovej siete pre binárnu klasifikáciu,

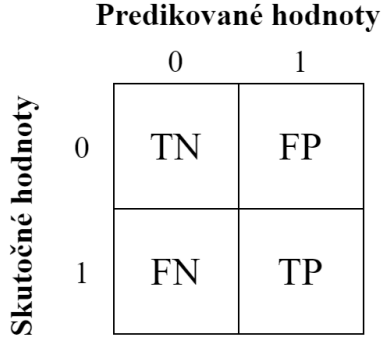
3. vytvorenie binárneho klasifikátora vo forme podmienky na klasifikovanie výstupov z predošlého kroku.

Pri vstupoch do modelov budeme uvažovať o použití hodnôt parametra σ_ϕ vo forme časových radov a o pridaní jedného alebo viacerých parametrov spomedzi geomagnetických indexov *Ap index*, *ASY/H*, *PC(N) index* a *Bz*.

Očakávaným výstupom tejto práce sú jednotlivé modely neurónových sietí, k nim príslušné spôsoby následného spracovania výstupov (angl. *post-processing*) a podmienky klasifikácie do pozitívnej triedy. Testovanie bude prebiehať na testovacej množine nezávislej od trénovacej množiny. Pri vyhodnotení modelov neurónových sietí budeme pozorovať metriku priemernej absolútnej chyby (angl. *mean absolute error*), ktorá je definovaná ako:

$$\mathrm{MAE} = \frac{1}{n}\sum_{i=1}^{n}|x_i - x|, \tag{3.1}$$

kde x je skutočná testovacia hodnota, x_i je predikovaná hodnota a n je počet testovaných hodnôt. Výber najlepšieho modelu ale nebude závisieť iba od najnižšej hodnoty metriky MAE, ale aj od jeho potenciálu umožniť predikované hodnoty klasifikovať ako scintilácie (1 – predikovaná hodnota je ionosférickou scintiláciou, 0 – predikovaná hodnota nie je ionosférickou scintiláciou). Pre vyhodnotenie kvality binárneho klasifikátora použijeme maticu zámen (angl. *confusion matrix*), ktorej rozloženie hodnôt je znázornené na Obrázku 3 – 1. Jednotlivé prípady, ktoré môžu medzi predikciami (1 alebo 0) a referenčnými testovacími hodnotami (1 alebo 0) nastať, sú:



Obrázok 3 – 1 Matica zámen pre binárnu klasifikáciu.

- **True Positive (TP)** – skutočne pozitívne prípady: scintilácia bola predikovaná a aj skutočne nastala,

- **False Negative (FN)** – falošne negatívne prípady: scintilácia nebola predikovaná avšak v skutočnosti nastala,
- **False Positive (FP)** – falošne pozitívne prípady: scintilácia bola predikovaná ale v skutočnosti nenastala,
- **True Negative (TN)** – skutočne negatívne prípady: scintilácia nebola predikovaná a v skutočnosti ani nenastala.

O hodnotách matice zámen možno uvažovať aj ako o zásahu (*hit*) pre TP, minutí (*miss*) pre FN a falošnom poplachu (*false alarm*) pre FP. Tieto pomenovania môžu uľahčiť chápanie dôležitosti nasledujúcich metrík pre vyhodnotenie binárnej klasifikácie v kontexte nášho problému predikcie ionosférických scintilácií, ktorými sú **precision** (skr. P), **recall** (návratnosť, skr. R), **F1 score** (skr. $F1$) a **true skill score** (skr. TSS). Ich výpočet sa riadi nasledujúcimi vzorcami:

$$P = \frac{TP}{TP + FP}, \tag{3.2}$$

$$R = \frac{TP}{TP + FN}, \tag{3.3}$$

$$F1 = \frac{2PR}{P + R}, \tag{3.4}$$

$$TSS = \frac{TP}{TP + FN} - \frac{FP}{FP + TN}. \tag{3.5}$$

Pri úlohách sledovania a detekcie ojedinelých, či skôr výnimočných ako štandardných javov, sa zvykne objavovať problém nevyváženej dátovej množiny, kedy jednotlivé triedy majú výrazne odlišné zastúpenie. Kvôli tomu sme ako hlavné metriky zvolili *precision* a *recall*. Skóre *F1* je vyjadrením ich harmonického priemeru. Pre umožnenie porovnania s inými prácami budeme počítať aj metriku *TSS*.

Naše výsledky budeme porovnávať s výsledkami predikcie perzistencie. Dôležitejšie bude ale porovnanie s prácou McGranaghan et al. (2018), ktorej výsledky

sú uvedené v Tabuľke 2 – 1. Okrem kvantitatívneho vyhodnotenia prostredníctvom metrík bude nevyhnutné dosiahnuté výsledky vyhodnotiť aj kvalitatívne, a to spôsobom grafického znázornenia predikcií a výsledkov klasifikácie, kde sa pozrieme na zmysluplnosť našich výstupov, ich kvalitu a ich porovnanie so skutočnými testovacími hodnotami.

Pre zhrnutie našich cieľov a našej motivácie uvádzame hypotézy, ktoré chceme týmto výskumom testovať:

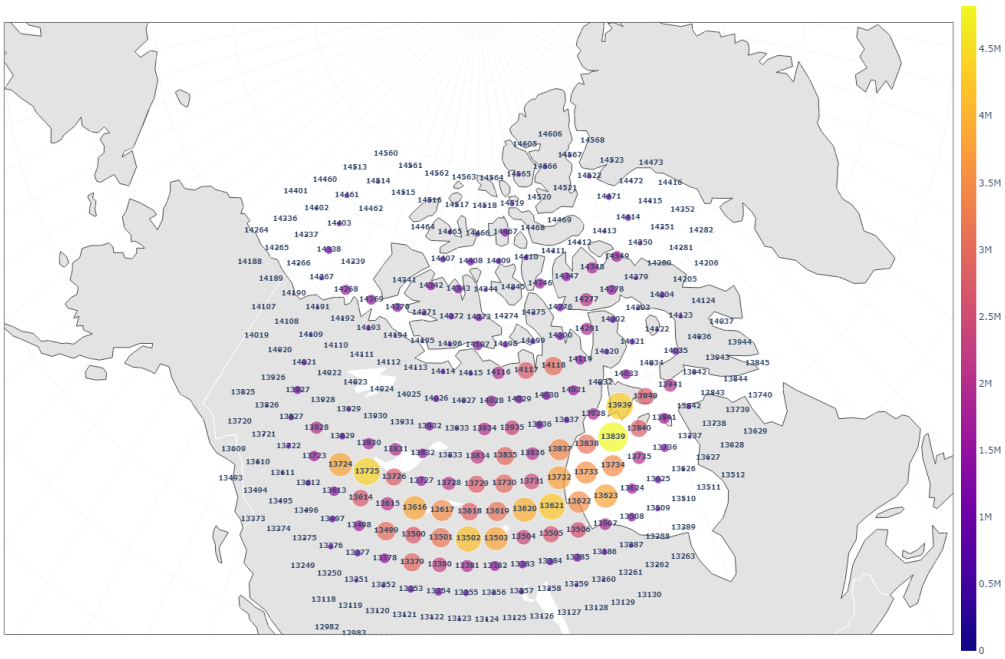
- Vhodným spracovaním výstupov modelov je možné znížiť počet falošne negatívnych prípadov so zachovaním vysokej miery návratnosti.
- Ak použijeme na vstupe okrem σ_ϕ aj ďalší parameter, dosiahneme lepšie výsledky ako iba s σ_ϕ.

3.2 Pochopenie dát

3.2.1 Intenzita fázovej ionosférickej scintilácie

Dáta ionosférických scintilácií, s ktorými sme pracovali, pochádzali z databázy kanadskej siete CHAIN [1] (angl. *Canadian High Arctic Ionospheric Network*), ktorá pozostáva z pozemných rádiových prístrojov rozmiestnených na severe Kanady. Do finálneho výberu boli použité merania z 11 staníc. Tie pokrývali obdobie od mája 2013 do decembra 2021 a obsahovali namerané hodnoty parametra σ_ϕ s časovými rozostupmi jednej minúty. Aby sa zohľadnil nielen časový ale aj priestorový pohľad na scintilácie, celá Zem bola rozdelená do 14 710 ekvidištantných oblastí s rozostupmi približne 200 kilometrov, ktoré budeme ďalej označovať ako biny. Dáta z 11 staníc boli priradené do binov podľa miesta, v ktorom priamka spájajúca satelit s pozemným prijímačom preťala ionosféru vo výške 350 km. Výsledkom tohto procesu boli dáta intenzity fázovej ionosférickej scintilácie v celkovo 294 binoch nad Kanadou (Obrázok 3 – 2), ktoré slúžili ako dátový vstup pre túto prácu.

[1] Dostupné na: `http://chain.physics.unb.ca/data/`



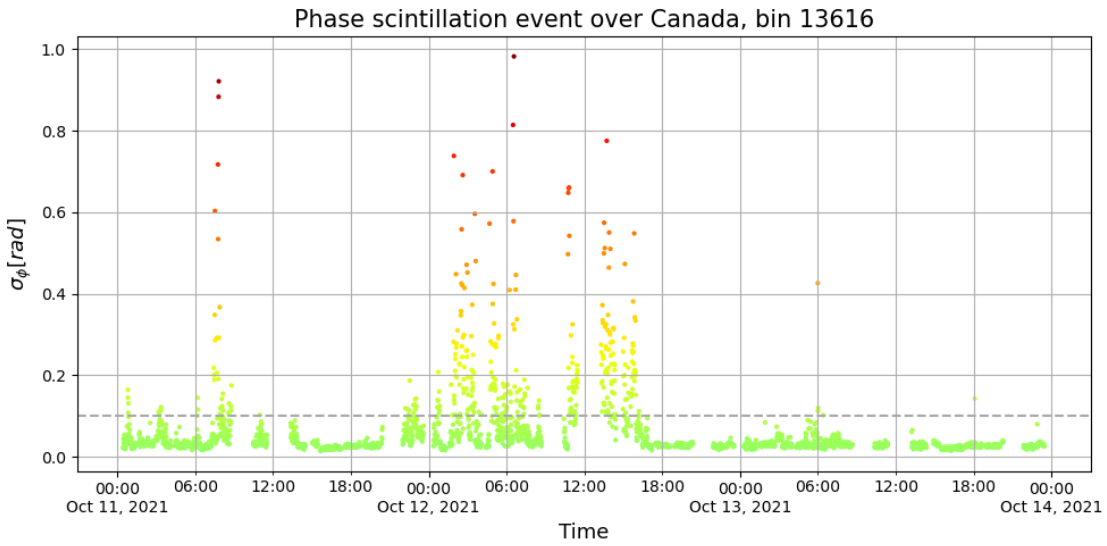
Obrázok 3 – 2 Rozmiestnenie a veľkosti binov podľa počtu hodnôt σ_ϕ vytvorených z databázy CHAIN.

Výstupom neurónových sietí bude predikcia spojitej hodnoty σ_ϕ, ktorá vyjadruje intenzitu fázovej ionosférickej scintilácie v radiánoch pre daný bin. Jej predikcia, teda budúca hodnota, bude vychádzať z priebehu časového radu jej minulých hodnôt, kvôli čomu považujeme σ_ϕ okrem výstupného parametra aj za vstupný parameter.

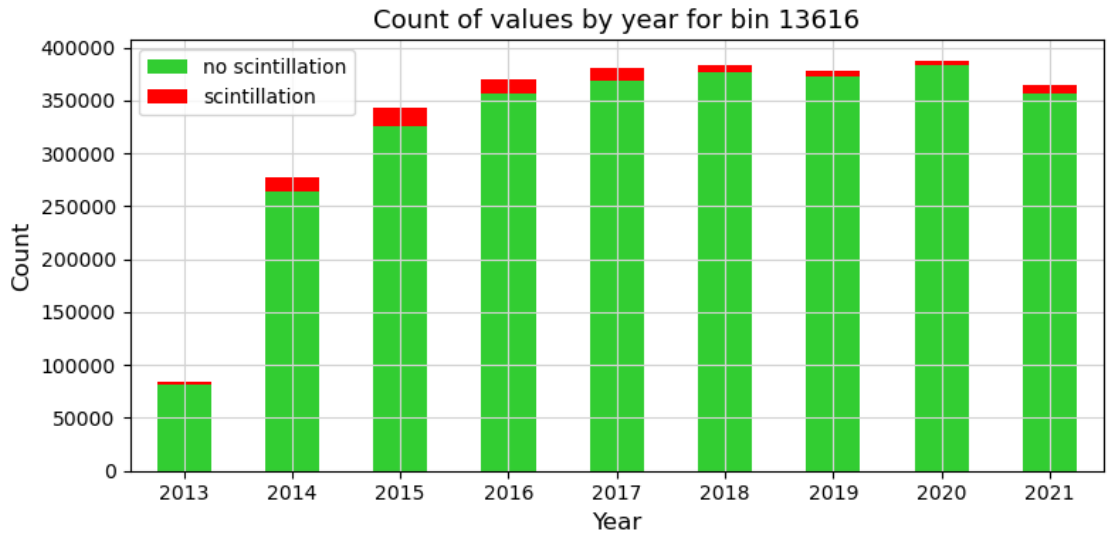
Ukážku priebehu ionosférickej scintilácie medzi 11. až 13. októbrom 2021 v oblasti binu 13616 možno vidieť na Obrázku 3 – 3. Prerušovaná priamka prechádzajúca hodnotou 0,1 znázorňuje hranicu, nad ktorou sú vyskytujúce sa namerané hodnoty považované za scintiláciu. Hraničná hodnota 0,1 bola stanovená na základe analýzy literatúry. Z vizualizácie počtu dostupných hodnôt σ_ϕ rozdelených podľa hranice 0,1 na hodnoty predstavujúce scintilácie a hodnoty, ktoré nie sú scintiláciami, môžeme vidieť, že výskyt scintilácií je naozaj zriedkavý (Obrázok 3 – 4).

Vráťme sa ale ešte na chvíľu k Obrázku 3 – 3. Ak sa pozrieme na časový interval, v ktorom sa vyskytujú hodnoty nad 0,1, môžeme si všimnúť, že v tom istom intervale je možno dokonca ešte viac hodnôt menších ako 0,1. To nás priviedlo k myšlienke

uvažovať o scintiláciách nie ako o nameraných hodnotách σ_ϕ väčších ako 0,1, ale ako o oblastiach alebo časových intervaloch s výskytom scintilácie, v ktorých by boli všetky hodnoty považované za scintilačné. Na túto myšlienku budeme nadväzovať aj neskôr pri príprave dát a modelovaní.



Obrázok 3 – 3 Ukážka priebehu intenzity ionosférickej scintilácie.



Obrázok 3 – 4 Počet záznamov parametra σ_ϕ pre roky 2013 až 2021 v rámci binu 13616.

Zo všetkých binov sme mali k dispozícii dokopy 247 999 339 nameraných hodnôt parametra σ_ϕ. V nasledujúcej podkapitole sa pozrieme, ako sú na tom parametre *Ap index*, *ASY/H*, *PC(N) index* a *Bz*.

3.2.2 Analýza ďalších parametrov

Zoznam ďalších parametrov, o ktorých budeme uvažovať ako o možných pridaných vstupoch do neurónových sietí, je uvedený v Tabuľke 3 – 1. Hodnoty týchto parametrov boli získané z databázy OMNIWeb, ktorá je spravovaná centrom *Goddard Space Flight Center* patriacim pod NASA. Zdrojmi údajov v tejto databáze sú viaceré satelity, ktorých úlohou je monitorovať kozmické počasie.

Parameter	**Jednotka**	**Rozlíšenie**	**Zdroj**
Ap index		1 hodina	OMNIWeb [2]
ASY/H	nT	1 minúta	OMNIWeb [3]
Bz GSE	nT	1 minúta	OMNIWeb [3]
PC(N)	mV/m	1 minúta	OMNIWeb [3]

Tabuľka 3 – 1 Zoznam použitých parametrov.

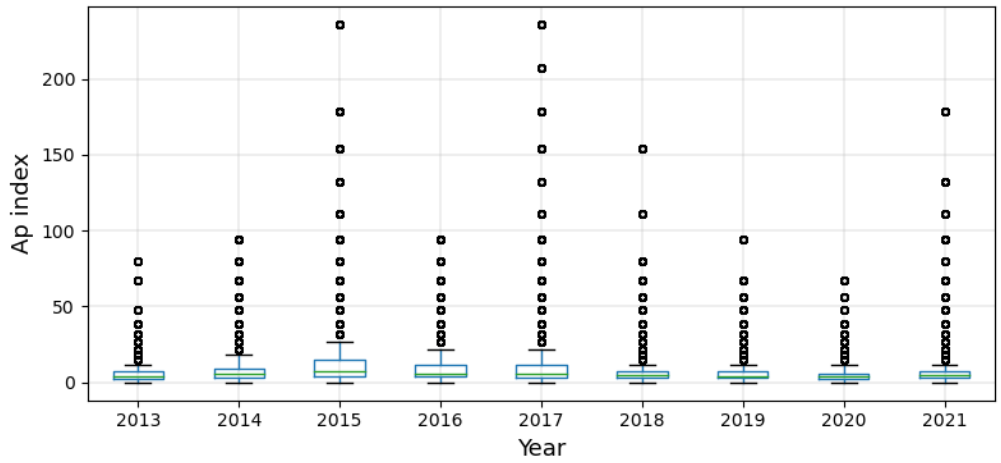
Hodnoty týchto parametrov boli podľa dátumu a času zaradené do binov. Je potrebné dodať, že pre niektoré existujúce hodnoty σ_ϕ neexistovali príslušné údaje o meraniach ostatných parametrov. Nepredstavovalo to ale veľký problém, keďže pre každý parameter sme po vytvorení binov mali k dispozícii vyše 200 miliónov záznamov v rovnakých dátumoch a časoch parametra σ_ϕ.

Ap index poskytuje informáciu o všeobecnom leveli dennej geomagnetickej aktivity. Vyjadruje celkový stav magnetického poľa Zeme z pohľadu výskytu rôznych narušení spôsobených slnečným žiarením. Čím vyššia je hodnota indexu, tým vyššia je aktivita počas jedného dňa. Z grafu na Obrázku 3 – 5 teda môžeme odvodiť,

[2] Dostupné na: `https://omniweb.gsfc.nasa.gov/form/dx1.html`

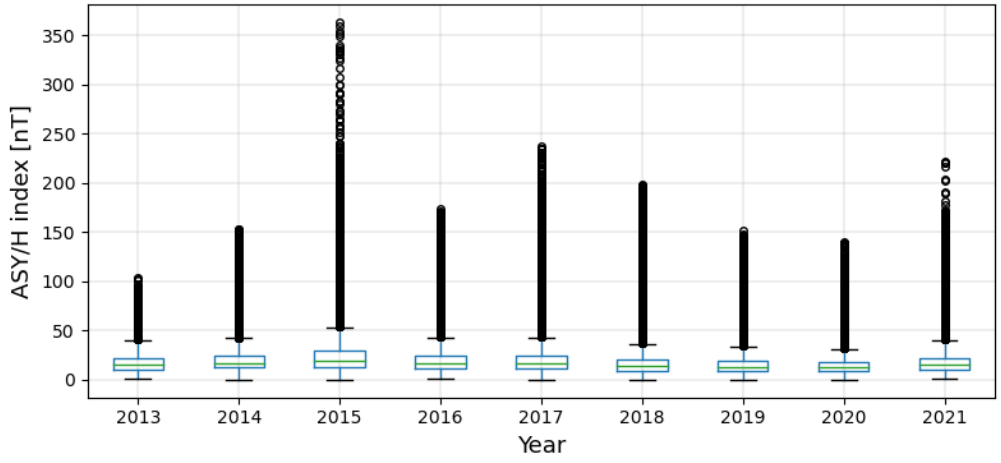
[3] Dostupné na: `https://omniweb.gsfc.nasa.gov/form/omni_min_def.html`

v ktorých rokoch boli dni s najvyšším levelom geomagnetickej aktivity. Údaje o tomto indexe máme k dispozícii v hodinových rozostupoch.



Obrázok 3 – 5 Boxplot pre parameter *Ap index* pre roky 2013 až 2021.

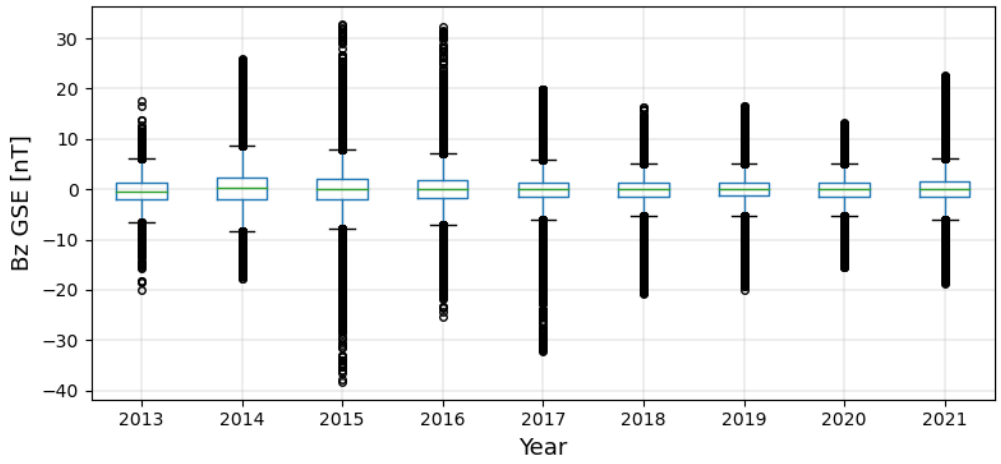
Index *ASY/H* slúži na opis asymetrických geomagnetických porúch a narušení v stredných zemepisných šírkach rovnobežných s osou magnetického dipólu Zeme. Vyjadruje sa v jednotkách nanotesla a ide o minútové dáta.



Obrázok 3 – 6 Boxplot pre parameter *ASY/H* pre roky 2013 až 2021.

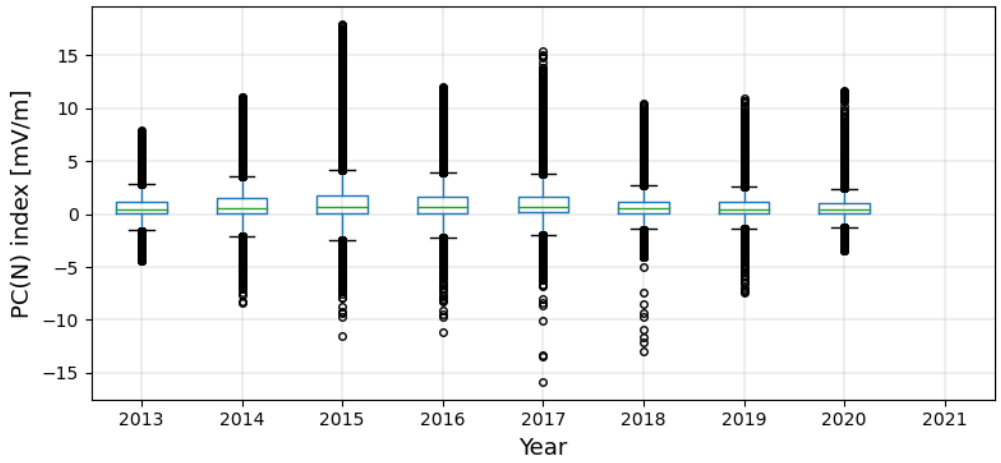
Hodnota *Bz* je jednou zo zložiek vyjadrujúcich silu medziplanetárneho magnetic-

kého poľa. Tento parameter je ovplyvňovaný vlnami a inými poruchami slnečného vetra. Taktiež je dôležitý pre aktivitu polárnej žiary. Udáva sa podobne ako *ASY/H* v jednotkách nanotesla.



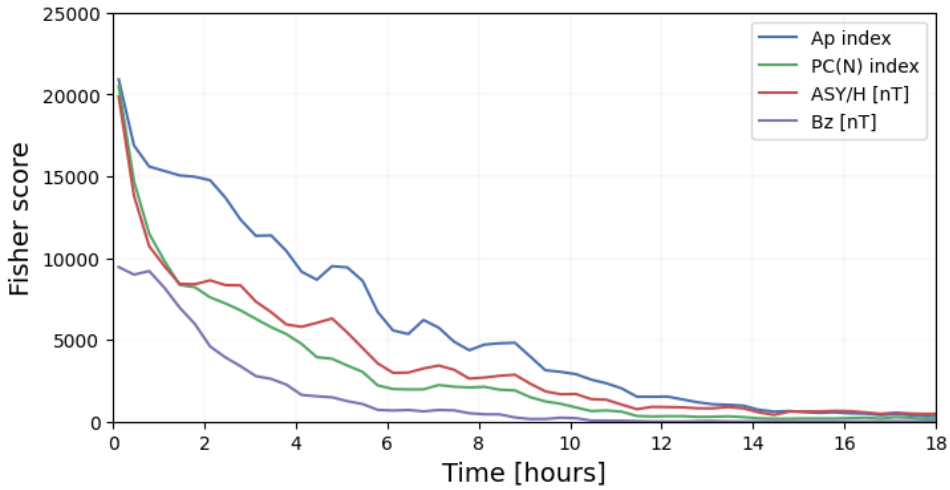
Obrázok 3 – 7 Boxplot pre parameter *Bz GSE* pre roky 2013 až 2021.

North Polar Cap Index, skrátene *PC(N)* je indexom, ktorý monitoruje geomagnetickú aktivitu v oblasti severného pólu vyjadrenú v jednotkách mV/m.



Obrázok 3 – 8 Boxplot pre parameter *PC(N)* pre roky 2013 až 2021 z údajov binu 13839.

Ako sme už spomenuli, samotný proces selekcie parametrov s potenciálom pomôcť pri predikcii scintilácií bol už v rámci potrieb projektu ASPIS vykonaný skôr, ešte pred začatím písania tejto práce. Z množiny 50 parametrov, ktoré prichádzali do úvahy, boli analýzou rozptylu ANOVA vybrané 4 parametre s najvyšším Fisherovým skóre, ktoré neboli medzi sebou navzájom závislé. Vývoj Fisherovho skóre medzi parametrami pre vstup do modelu a parametrom, ktorý je na výstupe, je možné vidieť na Obrázku 3 – 9.



Obrázok 3 – 9 Porovnanie vývoja Fisherovho skóre pre vybrané najdôležitejšie parametre.

3.3 Príprava dát

Prvým krokom prípravy dát bol výber trénovacej a testovacej množiny. Rozhodli sme sa dáta z roku 2019 použiť na testovanie modelov a na trénovanie modelov sme použili dáta z ostatných rokov. Takýmto spôsobom sme zabezpečili, aby testovacia množina nebola rozhádzaná a tvorená z údajov v náhodných dátumoch a časoch, keďže v procesoch spracovania výstupov modelov budeme potrebovať predikcie vo forme usporiadaných časových radov. Proces prípravy dát pre túto prácu zahŕňal normalizáciu, aplikovanie metódy posuvného okna na vytvorenie časových okien a výber finálnych trénovacích a testovacích podmnožín.

3.3.1 Normalizácia

Pod pojmom normalizácia sa rozumie proces preškálovania hodnôt jednotlivých parametrov do podoby vhodnej pre modelovanie a analýzu. Pri úlohách s viacerými vstupnými parametrami, ktoré majú rôzne škály hodnôt, je normalizácia nevyhnutným krokom prípravy dát. Z normalizovaných dát sa model dokáže lepšie naučiť potrebné vzory pre riešenie problému.

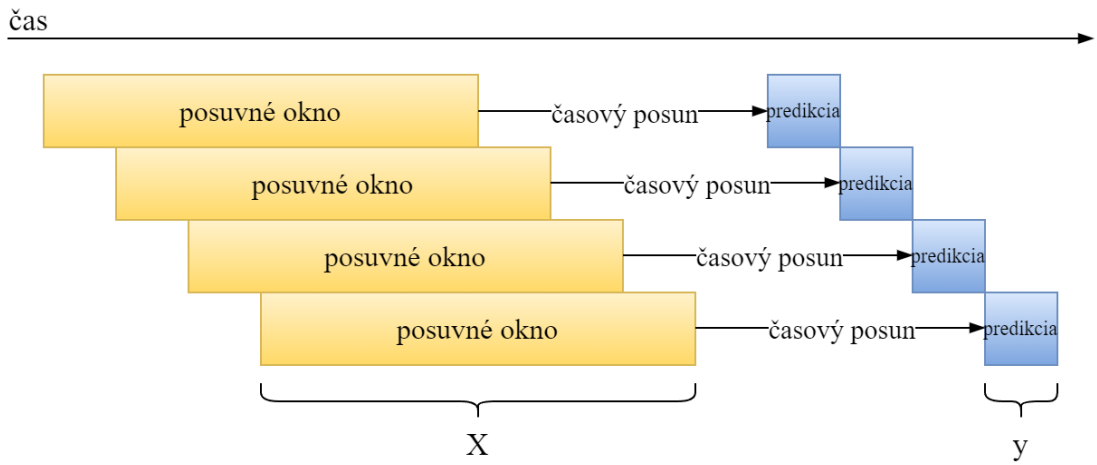
Na základe zistení z fázy pochopenia dát v predchádzajúcom kroku sme použili normalizáciu. Konkrétne išlo o aplikovanie min-max normalizácie na interval $\langle 0, 1 \rangle$ pre všetky parametre.

3.3.2 Aplikovanie metódy posuvného okna

Dáta, s ktorými pracujeme, predstavujú sekvencie numerických hodnôt zoradených podľa času od najstarších po najnovšie. Inými slovami, ide o prácu s časovými radmi (angl. *time series*), pri ktorých je veľký potenciál na využitie závislostí medzi jednotlivými prvkami časového radu k predikcii budúcich stavov.

Úloha, ktorú v tejto práci riešime, je predikcia budúcich hodnôt parametra σ_ϕ. Našim cieľom je preto vytvoriť časové okná a k nim priradiť príslušnú hodnotu z budúcnosti, pričom rozdiel medzi časom poslednej hodnoty v okne a časom priradenej hodnoty sa nazýva časový posun `shift`. Týmto spôsobom vieme vytvoriť úlohu kontrolovaného učenia, keďže k vstupom (časové okná) poznáme správne výstupy, resp. predikcie (hodnota o určitý časový posun). Príprava dát časových radov do okien, ktorú sme v tejto časti použili, sa volá metóda posuvného okna (angl. *sliding window approach*). Na pochopenie jej fungovania je potrebné sa oboznámiť s nasledujúcimi pojmami (Chu, 1995):

- **Časový krok**: Časová jednotka, v našom prípade sme kvôli práci najmä s minútovými dátami použili časový krok rovný 1 minúte.
- **Okno**: Označuje sa tiež ako časové alebo posuvné okno. Predstavuje výsek stanovenej dĺžky z časového radu, pričom tento výsek obsahuje nepretržitú



Obrázok 3 – 10 Metóda posuvného okna.

postupnosť hodnôt usporiadaných podľa času. Veľkosť okna `windows_size` je určená počtom minulých časových krokov, ktoré okno obsahuje. Hodnoty okna sú vstupnými dátami pre model.

- **Horizont**: Pojem horizont reprezentuje počet časových krokov, ktoré chceme predikovať. V kontexte strojového učenia ide o veľkosť výstupu modelu, čo je v podstate počet predikovaných hodnôt z jedného okna. V tejto práci budeme predikovať vždy iba jednu hodnotu.

Metóda posuvného okna je graficky zjednodušene vyjadrená na Obrázku 3 – 10. Postup tvorby okien aplikovaný na našu úlohu spočíva v prechádzaní časovým radom konkrétneho binu a hľadaní za sebou idúcich hodnôt s rozostupmi jednej minúty, ktorých počet je rovný `window_size`. Inými slovami, v prípade chýbajúcich hodnôt v časovom intervale veľkosti `window_size` sa takéto okno nevytvorí. Okno sa nevytvorí ani v prípade, že o `shift` minút dopredu nebude v dátach cieľová hodnota. Celý pôvodný časový rad údajov sa týmto spôsobom rozdelí na čiastočne prekrývajúce sa okná. Jeden trénovací príklad pozostáva z časového okna (vstupné dáta X) a cieľovej hodnoty (výstup y).

Veľmi dôležitým krokom pri metóde posuvného okna je zvolenie vhodnej veľkosti okna `window_size`. Pri výbere príliš veľkého okna hrozí, že sa kvôli chýbajúcim hod-

notám vytvorí malý počet okien. V prípade príliš malého okna je hrozba, že okno nebude obsahovať dostatočný počet hodnôt na extrahovanie príznakov potrebných pre učenie. Pri prvých experimentoch sa nám najviac osvedčilo používať okno veľkosti 45 minút, pričom okná 15, 30 a 60 minút dosahovali horšie výsledky podľa metriky MAE. Napriek väčšiemu oknu sa modely s oknom 60 minút nedokázali naučiť predikovať hodnoty lepšie ako modely s 45-minútovým oknom, pretože kvôli chýbajúcim hodnotám nebolo možné vytvoriť dostatočné množstvo trénovacích príkladov. Čo sa týka nastavenia časového posunu, teda parametra `shift`, zvolili sme hodnoty 15, 30, 45 a 60 minút. Po zvolení nastavení sme mohli spustiť proces generovania okien.

Pri modelovaní budeme navrhovať architektúru neurónovej siete pre autoregresné a multivariačné modely. V prípade autoregresných modelov sme potrebali vytvoriť časové okná iba z parametra σ_ϕ. Pre multivariačné modely sme pripravili časové okná pozostávajúce z hodnôt:

- σ_ϕ a *Ap index* pre model `ap_index`,
- σ_ϕ a *ASY/H* pre model `AsyH`,
- σ_ϕ a *Bz* pre model `BzGSE`,
- σ_ϕ a *PC(N)* pre model `PC`,
- σ_ϕ, *Ap index*, *ASY/H*, *Bz* a *PC(N)*, teda zo všetkých uvažovaných parametrov pre model, ktorý sme pomenovali `all_parameters`.

Na predikciu pomocou autoregresných modelov boli vygenerované časové okná s posunom 15, 30, 45 a 60 minút. Pri multivariačných modeloch je našim cieľom porovnať ich s autoregresnými modelmi a posúdiť, ktorý je lepší. Preto sme sa rozhodli ich porovnanie uskutočniť na predikcii 15 minút dopredu. Pre všetky multivariačné modely sme vytvorili časové okná s hodnotou `shift` 15 minút a pomocou nich vytvorené modely sme porovnali s autoregresným modelom na predikciu 15 minút dopredu.

Celý postup generovania časových okien sme aplikovali na pôvodné trénovacie a aj testovacie dáta po jednotlivých binoch. Vo výsledku sme po tomto kroku mali k dispozícii množiny trénovacích a testovacích príkladov zložených z časových okien a cieľových hodnôt, ktoré bolo potrebné ešte mierne upraviť.

3.3.3 Výber finálnych množín na trénovanie a testovanie

Kvôli nízkej početnosti výskytu ionosférických scintilácii v našich dátach, bolo nevyhnutné pristúpiť k ďalšiemu kroku v predspracovaní trénovacích dát na vyváženie počtov príkladov pre jednotlivé triedy.

Tieto triedy sme sa rozhodli vybalansovať takým spôsobom, že sme rozdelili všetky trénovacie hodnoty y, ktoré sú na výstupe modelov do dvoch skupín podľa toho, či predstavovali scintiláciu. Na určenie výskytu ionosférickej scintilácie pre danú hodnoty sme ale nepoužili prístup založený na hraničnej hodnote 0,1. Namiesto toho sme sa pokúsili zachytiť celé scintilačné okno použitím variačného koeficientu, ktorý budeme v tejto práci ďalej označovať skratkou CoV (angl. *Coefficient of Variation*). Inými slovami, pozreli sme sa na okolie okolo hodnoty y a zisťovali sme, či sa nachádza v okne s výskytom scintilácie. Variačný koeficient CoV sa vypočíta ako podiel smerodajnej odchýlky a priemeru hodnôt z dátovej množiny. Celý proces aplikovania a výpočtu CoV je uvedený neskôr pri modelovaní (Kapitola 3.4.2). Nateraz nám stačí vedieť, že pre každú hodnotu y sme vybrali hodnoty z časového intervalu 45 minút so stredom v bode y, z ktorých sme vypočítali CoV. Je nutné upozorniť, že tento interval hodnôt nie je totožný s posuvnými oknami, s ktorými sme pracovali už skôr, ide totiž o úplne odlišné okná. Po výpočte CoV pre všetky hodnoty y sme stanovili hraničnú hodnotu, tentokrát však pre CoV. Ak bol CoV väčší ako 0,5, potom hodnota y bola súčasťou oblasti s výskytom scintilácie.

Keď sme už hodnoty y mali označené, z trénovacích dát sme vybrali všetky príklady s označením „scintilácia“, ktorých počet nazveme n_{scinti}. Následne sme zo všetkých príkladov bez scintilácií náhodne vybrali n_{scinti} príkladov. Týmto spôsobom sme zabezpečili vyváženosť trénovacej množiny z pohľadu zastúpenia obidvoch tried.

Na testovaciu množinu tento proces nebol aplikovaný pre zachovanie jej pôvodného charakteru a rozloženia hodnôt.

V tomto bode sme mali pripravené trénovacie dáta na modelovanie a testovacie dáta na otestovanie modelov. Trénovacích dát však bolo až príliš veľa a z časového hľadiska by sa model na týchto dátach učil až zbytočne dlho. Z našich pokusov trénovať model na celej trénovacej množine sme zistili, že takýto model dosiahol horšie výsledky nielen pri MAE, ale aj pri klasifikácií, než model, ktorý bol trénovaný na vybranej podmnožine. Preto sme sa rozhodli do finálnej trénovacej množiny pre každý model posunúť iba 1/16 z aktuálnej trénovacej množiny. Testovaciu množinu sme ponechali v rovnakej podobe, aby sme mohli testovať predikcie počas celého obdobia roku 2019. Prehľad veľkostí všetkých množín možno vidieť v Tabuľke 3 – 2.

Model	**shift**	**Veľkosť train_data**	**Veľkosť test_data**
autoreg	15	385 748	4 043 728
autoreg	30	336 358	3 656 039
autoreg	45	338 420	3 695 372
autoreg	60	347 226	3 684 376
ap_index	15	385 748	4 043 728
AsyH	15	385 748	4 043 728
BzGSE	15	217 645	2 324 499
PC	15	354 486	3 981 169
all_parameters	15	198 925	2 287 633

Tabuľka 3 – 2 Počet záznamov vo finálnych množinách pre trénovanie a testovanie podľa jednotlivých modelov.

3.4 Modelovanie

Ukončením prípravy dát sme sa mohli presunúť k modelovaniu. Ako sme už spomínali, našim cieľom je natrénovať autoregresné a multivariačné modely neurónových sietí na predikciu σ_ϕ. Túto časť sme rozdelili do viacerých fáz. Tou prvou je predikcia

spojitej hodnoty σ_ϕ prostredníctvom autoregresných modelov s posunom 15 minút. Modely všetkých epoch použijeme na predikciu a predikované hodnoty následne analyzujeme. Pre spracovanie skutočných hodnôt a predikcií do podoby vhodnej na zachytenie okna scintilácie použijeme vybrané metódy vyhladzovania priebehu funkcie a variačný koeficient CoV. Výstupom tejto fázy bude výber modelu tej epochy, pri ktorej budú dosiahnuté najlepšie výsledky binárnej klasifikácie podľa metriky *TSS*. Druhým výstupom v tejto fáze bude aj návrh spracovania časových radov testovacích a predikovaných hodnôt σ_ϕ.

V ďalšej fáze budeme trénovať zvyšné autoregresné modely pre `shift` 30, 45 a 60 minút, aby sme zistili, ako veľmi sa zmenia výsledky s narastajúcim posunom. V poslednej fáze budeme trénovať multivariačné modely na predikciu o 15 minút a porovnáme ich s najlepším autoregresným modelom.

3.4.1 Autoregresná predikcia o 15 minút dopredu

Princíp autoregresnej predikcie je založený na myšlienke predikovať z minulých hodnôt konkrétneho parametra jeho budúce hodnoty. Použitie hlbokého učenia na túto problematiku je úplne na mieste, práve vďaka jeho vlastnostiam spracovať komplexné dáta, či pamätať si kontext v sekvencii. Úlohou je zostaviť a natrénovať architektúru neurónovej siete, ktorej výstupy dosiahnu čo najlepšie výsledky pri klasifikovaní.

Kvôli charakteru a vlastnostiam našich dát sme pre navrhovanú architektúru pristúpili k použitiu rekurentných neurónových sietí. Pracovali sme aj s verziou obsahujúcou konvolučnú vrstvu, avšak po jej odstránení dosiahol model pri testovaní lepšie výsledky, preto sme ju ďalej nepoužívali. Vizualizáciu najlepšej architektúry s implementovanou rekurentnou vrstvou `Bidirectional(LSTM)` možno vidieť na Obrázku 3 – 11 a skladá sa z nasledujúcich vrstiev:

- Na začiatku siete je vstupná vrstva pozostávajúca zo 45 neurónov na spracovanie vstupných dát, keďže naše vstupy predstavujú časové rady hodnôt

o veľkosti 45 minút.

- Za ňou nasleduje rekurentná vrstva `Bidirectional` ako prvá skrytá vrstva modelu. Celkovo má model 5 skrytých vrstiev, pričom prvá z nich pozostáva z obojsmernej vrstvy `LSTM`, kde je pre každý smer vyhradených 128 neurónov. Dokopy je táto vrstva tvorená 256 neurónmi a nastavenie jej parametrov `dropout` a `recurrent_droput` na hodnotu 0,1 zabezpečuje regularizáciu učenia náhodným vypnutím 10% neurónov.

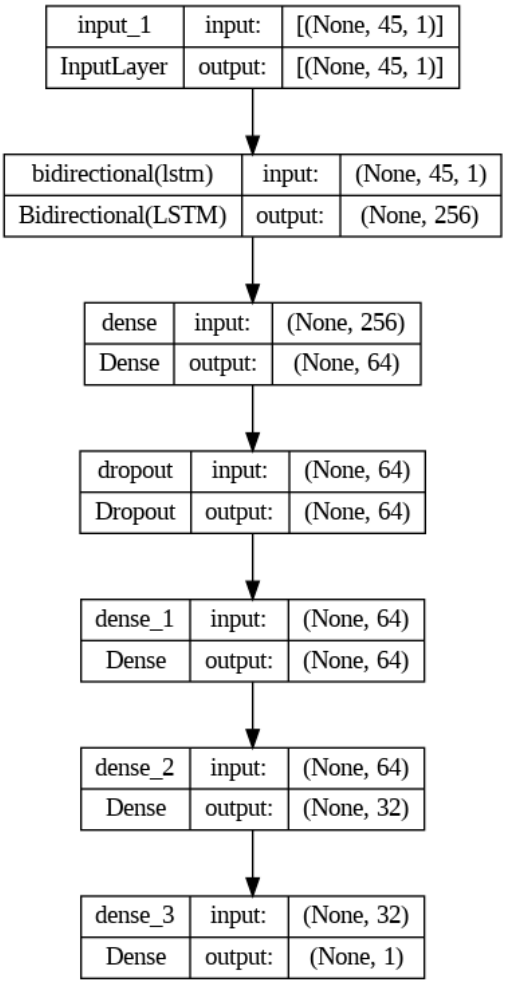
- Výstupy z predchádzajúcej vrstvy boli odoslané ako vstup do vrstvy `Dense` so 64 neurónmi, ktorej výpočet výstupov bol riadený aktivačnou funkciou `relu`.

- Ďalšou skrytou vrstvou je vrstva `Drouput` s nastavením náhodného vypnutia neurónov na 20%.

- Nasledujú dve po sebe idúce vrstvy `Dense` s aktivačnými funkciami `relu`, pričom prvá `Dense` vrstva má 64 a druhá 32 neurónov. Jej výstup smeruje na výstupnú vrstvu celého modelu.

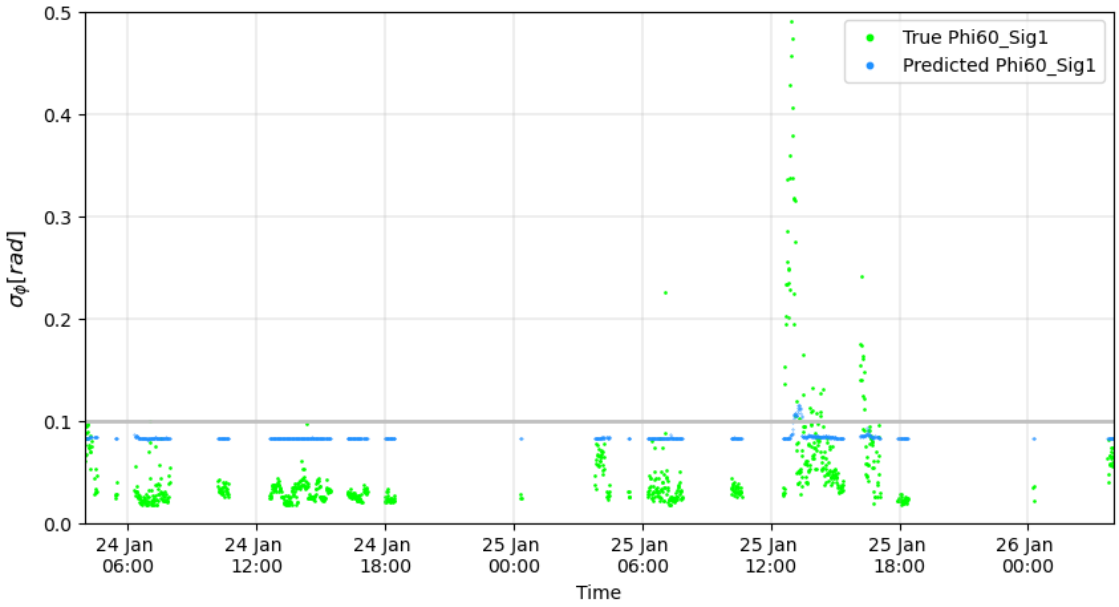
- Celkovo posledná vrstva neurónovej siete, taktiež označovaná ako výstupná vrstva, je vrstva `Dense`, ktorá pozostáva z 1 neurónu a aktivačnej funkcie `linear`, ktorá vyprodukuje hodnotu predstavujúcu predikciu parametra σ_ϕ.

Počet epoch, počas ktorých prebiehal proces trénovania modelu, sme nastavili na hodnotu 10. Predchádzali tomu experimenty, v ktorých sme zistili, že po 10 epochách sa už model z pohľadu veľkosti hodnoty chybovej funkcie nezlepšoval. Pred trénovaním sme ešte z trénovacej množiny vyčlenili 10% pre validačnú množinu. Po úspešnom natrénovaní modelov sme z testovacích dát vložených do modelov dostali predikované hodnoty. Na tie sme sa podrobnejšie pozreli s cieľom pochopiť ako sa model „správa“ pri predikcii.



Obrázok 3 – 11 Architektúra autoregresného modelu s rekurentnými vrstvami.

Na Obrázku 3 – 12 sú zobrazené testovacie hodnoty spolu s predikovanými hodnotami. Môžeme vidieť, že dňa 25. januára v čase medzi 12:00 a 18:00 došlo k scintilácií, keďže mnohé hodnoty parametra σ_ϕ (zelené body) ďaleko prevyšovali hodnotu 0,1. Predikované hodnoty (modré body) v tomto intervale mierne poskočili smerom nahor. V oblastiach bez scintilácie sa predikované hodnoty pohybujú okolo konkrétnej konštanty. Po porovnaní testovacích a predikovaných hodnôt celého roku 2019 z akéhokoľvek binu sme došli k záveru, že model pre oblasti s ionosférickou scintiláciou predikuje hodnoty, ktoré buď oscilujú alebo sa medzi sebou výrazným spôsobom líšia. Pre oblasti bez scintilácie je model naopak naučený predikovať hodnoty blížiace



Obrázok 3 – 12 Porovnanie predikcií autoregresného modelu zo 4. epochy so skutočnými testovacími hodnotami vybraných z binu 13725.

sa určitej konštante. Tieto zistenia boli pre nás veľmi užitočné pri návrhu spracovania výstupov modelu, čím sme sa ocitli opäť o krok bližšie k riešeniu klasifikačnej úlohy detekcie scintilácií.

3.4.2 Post-processing

V tejto časti uvádzame nápady pre spracovanie výstupov modelu, ktoré sú testované na autoregresnom modeli s predikciou 15 minút dopredu. Naše návrhy pozostávali z aplikovania metódy *smoothing* a aplikovania výpočtu variačného koeficientu, pomocou ktorých sme testovali ako dokážu tieto metódy oddeliť scintilácie od ostatných hodnôt.

Smoothing

Našou prvou ideou bolo upraviť priebeh časového radu do takej podoby, aby sme ho „zjemnili“, nízke hodnoty v oblastiach výskytu scintilácií zdvihli nahor aj za cenu potlačenia príliš vysokých hodnôt nadol, pričom by sme z globálneho hľadiska

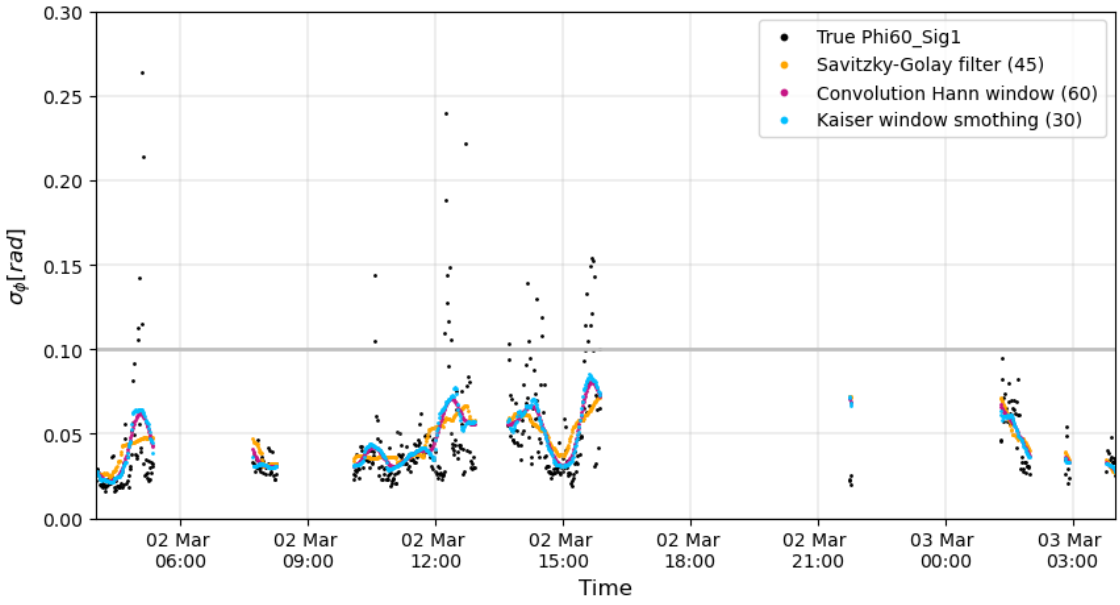
zachovali stúpanie a klesanie hodnôt pri scintiláciách. Takouto použitou metódou na spracovanie hodnôt σ_ϕ, testovacích a aj predikovaných, je metóda *smoothing* alebo v slovenskom preklade metóda vyhladzovania priebehu (napr. časového radu, funkcie). Túto metódu možno použiť napríklad na odstránenie nechceného šumu v dátach. My sme sa ju rozhodli aplikovať najprv na testovacie hodnoty (Obrázok 3 – 13), kde sú lepšie viditeľné rozdiely medzi hodnotami z dôvodu ich väčšieho rozsahu než aký majú predikcie (Obrázok 3 – 14). Použili sme nasledujúce prístupy, ktorých základná myšlienka spočíva v aplikovaní vyhladzovacieho filtra na hodnoty v časovom rade:

- **Convolution-Based Smoothing**. Ako to už z názvu vyplýva, ide o metódu založenú na konvolučnom filtri, ktorý je tvorený váhami. Pri prechádzaní časovým radom sú tieto váhy aplikované na okno, ktoré tvorí okolie hodnoty, ktorú chceme transformovať. Svojim spôsobom sú aj ďalšie uvedené metódy obdobou konvolúcie, avšak s tým rozdielom, že ich filtre sú omnoho viac komplexnejšie, než aké sa používajú pri bežnej konvolúcii. V tomto bode sme použili filter v podobe okna Hann (Testa et al., 2004) pomocou funkcie `scipy.signal.windows.hann`.

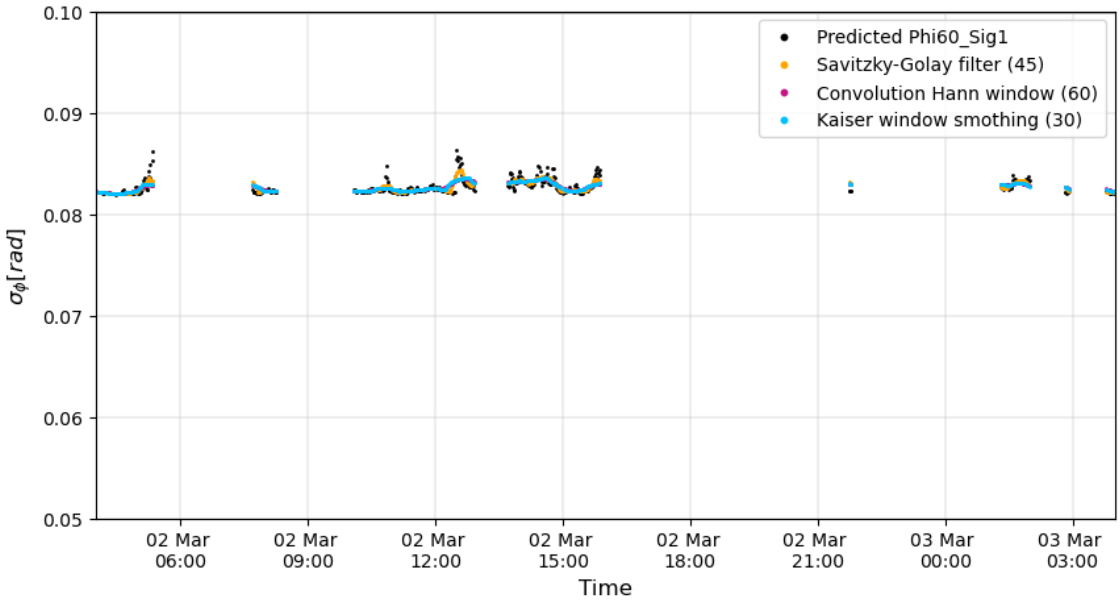
- **Savitzky-Golay filter**. Táto metóda vyhladzovania dát je založená na lokálnej aproximácii polynómu pomocou metódy najmenších štvorcov (Savitzky and Golay, 1964). Tento filter sme aplikovali pomocou funkcie `savgol_filter` z knižnize `scipy`, pričom veľkosť filtračného okna sme nastavili na 45.

- **Kaiser window smoothing**. Pri tejto metóde je filtrom okno Kaiser, niekedy označované aj ako Kaiser-Bessel (Theußl et al., 2000). Na jeho aplikovanie sme použili funkciu `kaiser` z knižnice `numpy`.

Na spomínaných Obrázkoch 3 – 13 a 3 – 14 môžeme vidieť transformáciu pôvodných hodnôt (čierne body) na upravené hodnoty. Pre jednotlivé metódy sme použili rôzne veľkosti okien, keďže pri použití rovnakých veľkostí boli výstupom veľmi podobné transformácie pri všetkých metódach. Horný obrázok reprezentuje skutočné



Obrázok 3 – 13 Aplikovanie vybraných *smoothing* metód na testovacie hodnoty σ_ϕ.



Obrázok 3 – 14 Aplikovanie vybraných *smoothing* metód na predikované hodnoty σ_ϕ.

testovacie hodnoty a dolný predikcie. Je potrebné si uvedomiť, že osi y majú nastavené rôzne škály, keďže ako už vieme, predikcie sa pohybujú okolo určitej konštanty, zatiaľ čo skutočné testovacie hodnoty nadobúdajú hodnoty zo širšej škály, ktorá po-

pisuje reálne priebehy scintilácií. Z tohto pohľadu by sa mohlo zdať, že je možné nájsť také hraničné hodnoty pre testovacie a predikované hodnoty zvlášť, podľa ktorých by sme dokázali oddeliť scintilačné okná od okien bez scintilácie. Pri testovaní tejto teórie sme ale pochopili, že opak je pravdou. Z globálneho hľadiska, teda z pohľadu na celý rok 2019, je veľmi ťažké nájsť hranice, pomocou ktorých by sme zachytili hoc len tie najintenzívnejšie scintilácie v celom roku. Pri metódach smoothingu berieme do úvahy iba lokálne vlastnosti časového radu. Pri predikciách sa môže napríklad stať, že hraničná hodnota pre detekciu scintilácií na uvedenom obrázku zachytí aj kopec ďalších oblastí, ktoré však len „povyskočili“ od konštantnej hodnoty bez väčších výkyvov. Preto sme sa rozhodli namiesto zmiernenia priebehu zobrať do úvahy priebežné zmeny, výkyvy a oscilácie s ohľadom na globálny priebeh.

Variačný koeficient

Variačný koeficient (skr. CoV, angl. *coefficient of variation*) je veličina používaná v štatistickej analýze, ktorá je definovaná ako:

$$\mathrm{CoV} = \frac{\sigma}{\mu}, \tag{3.6}$$

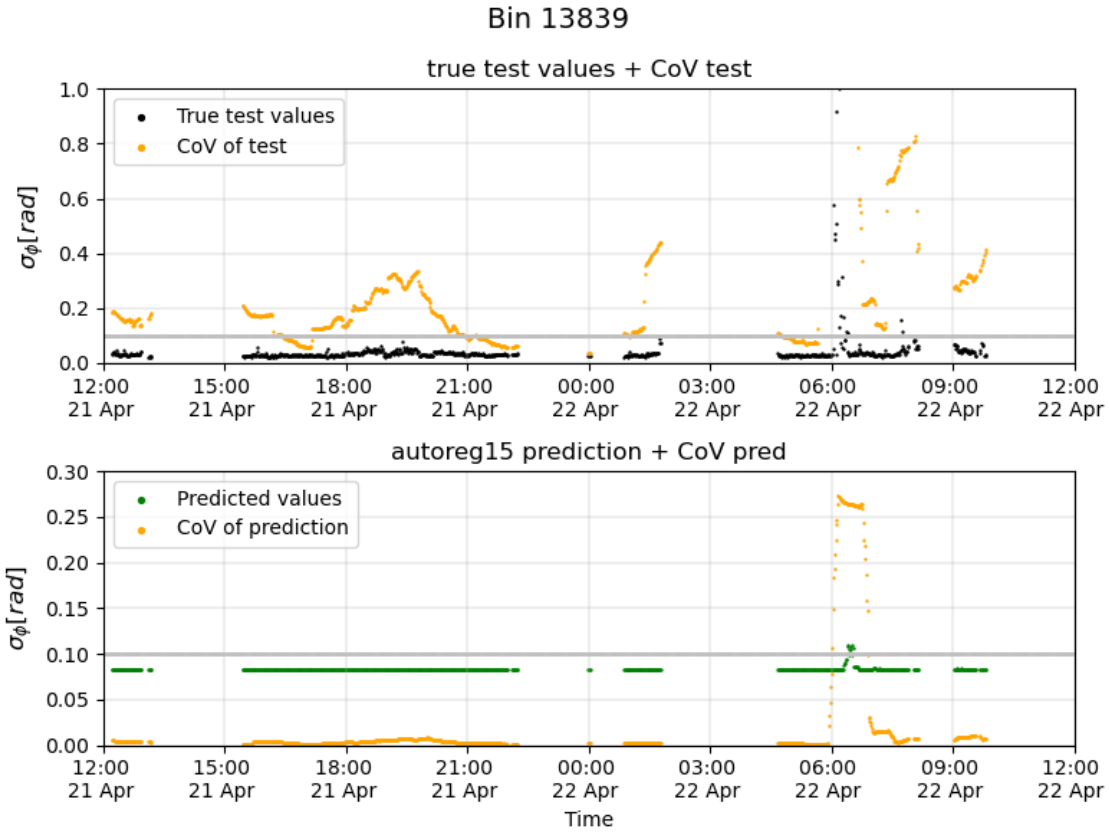
kde σ je smerodajná odchýlka a μ je priemer hodnôt zo súboru dát. Je vyjadrením variability vzhľadom k priemernej hodnote. Túto veličinu budeme počítať pre konkrétnu hodnotu y z hodnôt v jej okolí, z ktorých vypočítame smerodajnú odchýlku a predelíme ju priemerom všetkých hodnôt y, ktoré predstavujú parameter σ_ϕ v testovacej množine. Spravíme tak osobitne na testovacej množine a na modelom predikovaných hodnotách. Pri výpočte CoV na predikovaných hodnotách budeme ako priemer používať priemer z testovacej množiny.

Inými slovami, hodnota CoV pre konkrétny bod nám bude hovoriť o relatívnej veľkosti zmien v oblasti okolo zvoleného bodu. Čím vyššia hodnota CoV, tým väčšia je variancia. A to je presne to, čo sa deje pri predikciách intenzity ionosférických scintilácií pomocou nášho modelu.

Výpočet CoV pre testovaciu množinu a CoV pre predikované hodnoty sme robili pomocou nasledujúceho postupu:

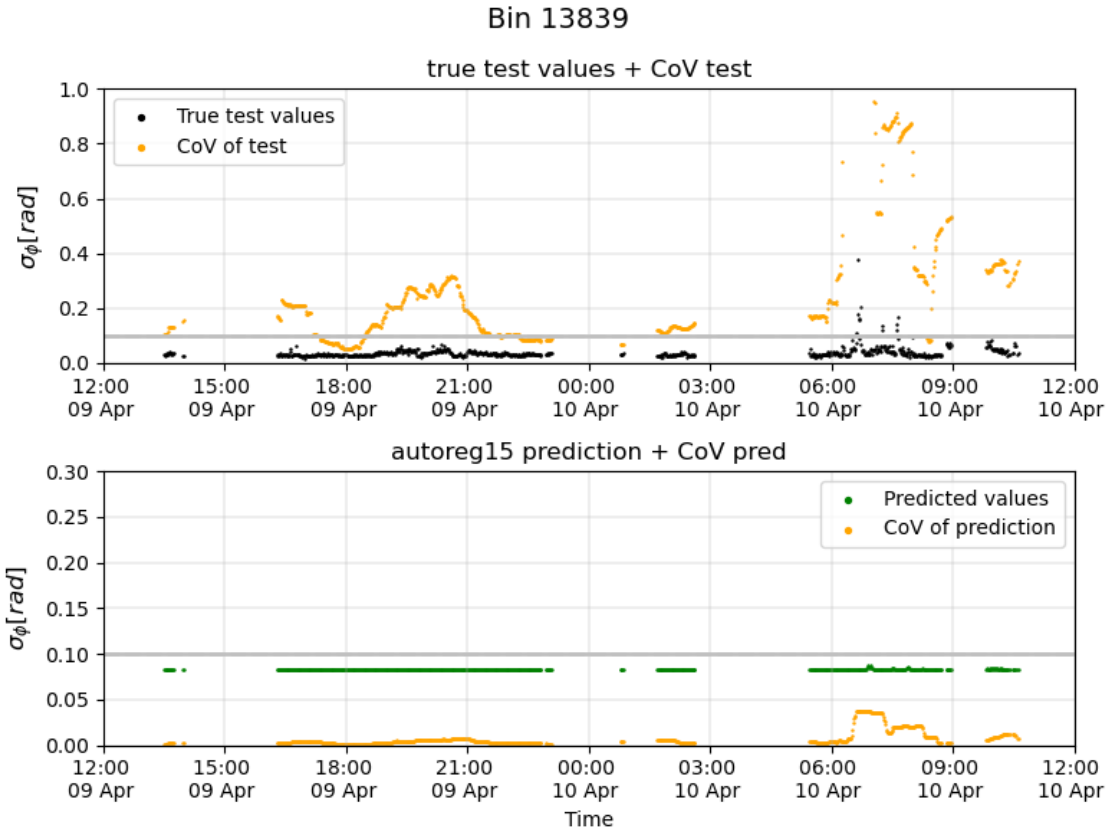
- zoberieme údaje o σ_ϕ (výstup modelu y) z jedného binu,
- okolo každého bodu y vytvoríme časové okno 45 minút,
- hodnôt v okne môže byť maximálne 45, minimálne 1, ktorou je samotné y,
- ak je v okne okrem y aspoň 1 ďalšia hodnota, z hodnôt nachádzajúcich sa v tomto okne vypočítame smerodajnú odchýlku,
- ak okno okrem y neobsahuje žiadne ďalšie hodnoty, nemôžeme vypočítať smerodajnú odchýlku a tým pádom ani CoV,
- týmto postupom získame pre výstupné hodnoty y veľkosti smerodajnej odchýlky,
- tie predelíme priemerom všetkých hodnôt y z testovacej množiny,
- pre každé y vyjadrujúce intenzitu ionosférickej scintilácie v testovacej množine alebo v množine predikovaných hodnôt, dostaneme týmto postupom hodnoty CoV (ak je možné ich vypočítať).

Výstup tohto procesu a jeho použiteľnosť sme otestovali na našich dátach z roku 2019 (Obrázok 3 – 15). Začneme testovacími hodnotami, ktoré sú zobrazené na hornom grafe. Dňa 22. apríla 2019 po 6. hodine možno pozorovať vysoké hodnoty intenzity scintilácie (čierne body), z čoho usudzujeme, že v tom období sa scintilácia vyskytla. Hodnoty CoV (oranžové body) pri prudkých zmenách medzi analyzovanými hodnotami vyskočili na vysoké hodnoty, z ktorých niektoré dokonca presahujú zobrazenú škálu osi y. Ak sa pozrieme na predikcie (zelené body) v spodnom grafe, tak môžeme v rovnakom časovom období u týchto hodnôt pozorovať prudké zmeny oproti predošlému priebehu časového radu. Taktiež aj príslušné hodnoty CoV (oranžové body) sú vysoké v porovnaní s predchádzajúcimi hodnotami.



Obrázok 3 – 15 Porovnanie testovacích hodnôt σ_ϕ a z nich vyjadrených CoV s predikovanými hodnotami a z nich vyjadrených CoV pre obdobie medzi 21. a 22. aprílom 2019.

Poďme sa ale pozrieť aj na predikované hodnoty, ktoré počas scintilácie stúpali a klesali v menšej miere (Obrázok 3 – 16). Konkrétne 10. apríla po 6. hodine je zaznamenaná ionosférická scintilácia, čo sme usúdili na základe testovacích hodnôt v prvom grafe. V tom istom čase sa predikované hodnoty iba mierne líšia od ostatných hodnôt mimo danej oblasti. Variačný koeficient CoV však tieto zmeny dokázal zachytiť, keďže jeho hodnota v oblasti s výskytom scintilácie narástla a bola vyššia než hodnoty CoV mimo scintilačné obdobie. Aj keď tento rozdiel nebol príliš veľký, pri vhodnom určení hraničnej hodnoty je možné túto scintiláciu zachytiť. Napríklad hodnota CoV 0,02 by dokázala oddeliť takmer celé scintilačné okno od ostatných hodnôt.



Obrázok 3 – 16 Porovnanie testovacích hodnôt σ_ϕ a z nich vyjadrených CoV s predikovanými hodnotami a z nich vyjadrených CoV pre obdobie medzi 9. a 10. aprílom 2019.

Na základe stanovenia hraničných hodnôt pre CoV testovacej množiny a CoV predikcií, môžeme hodnoty σ_ϕ rozdeliť na hodnoty v oblasti scintilácie (1) a hodnoty mimo oblasti scintilácie (0). Výber konkrétnych hraničných hodnôt CoV prebiehal prostredníctvom pozorovania vývoja CoV v oblastiach so scintiláciou. Cieľom nebolo nájsť úplne presné hodnoty, pri ktorých by boli dosiahnuté najlepšie výsledky z pohľadu vybraných metrík, ale približné hodnoty, ktoré by mohli byť aplikované aj na ďalšie modely a ďalšie testovacie dáta. Vyhodnotenie klasifikácie je uvedené v Tabuľke 3 – 3. Pri testovacích dátach sme stanovili CoV na hodnotu 0,5 a pri predikciách sme nastavenie CoV menili s cieľom posúdiť, ktoré nastavenie je najlepšie. Pre porovnanie sú v tabuľke uvedené aj výsledky predikcie perzistencie, ktorá funguje na princípe, že súčasná hodnota je taká istá ako bude predikovaná, resp.

Model	CoV	P	R	F1	TSS
predikcia perzistencie		0,21	0,21	0,21	0,19
autoregresný	**0,01**	**0,63**	**0,76**	**0,69**	**0,70**
	0,02	0,82	0,47	0,60	0,45
	0,05	0,89	0,19	0,32	0,19
	0,10	0,90	0,12	0,21	0,12

Tabuľka 3 – 3 Metriky pre predikciu perzistencie a pre autoregresný model s časovým posunom 15 minút podľa rozličných nastavení CoV. Pre testovacie hodnoty bola hraničná hodnota CoV nastavená na 0,5.

budúca. V tomto prípade sme predikciu perzistencie robili s posunom 15 minút, keďže aj analyzovaný autoregresný model je určený na predikciu hodnôt o 15 minút dopredu.

Pre nastavenie CoV na hodnotu 0,01 sme dosiahli najlepšie výsledky spomedzi testovaných nastavení podľa metrík *TSS* a návratnosti *R*. Pre toto konkrétne nastavenie je v Tabuľke 3 – 4 zobrazená matica zámen. Problém veľkého množstva falošne pozitívnych prípadov bol v porovnaní s predošlými prácami na predikciu výskytu ionosférických scintilácií zanedbateľný. Aj keď hodnoty metrík vyzerajú sľubne, kvalitatívne vyhodnotenie je stále nevyhnutnosťou.

	Predikcia 0	**Predikcia 1**
Skutočná hodnota 0	3 307 715	211 025
Skutočná hodnota 1	112 100	355 043

Tabuľka 3 – 4 Matica zámen pre vyhodnotenie predikcií autoregresného modelu s posunom 15 minút a s nastavenou hranicou CoV pre predikcie 0,01.

Na kvalitatívne vyhodnotenie prístupu založenom na výpočte CoV sme výsledky klasifikácie graficky znázornili spôsobom uvedeným na Obrázku 3 – 17. Grafy v každom riadku zobrazujú skutočné testovacie hodnoty parametra σ_ϕ. Ich farba ale závisí od správnosti označenia triedy podľa kritéria hraničnej hodnoty CoV. Body ozna-

čené zelenou boli správne klasifikované ako body patriace do scintilačnej oblasti, červenou sú tie, ktoré neboli zachytené prostredníctvom CoV, modrou sú skutočne negatívne a oranžovou falošne pozitívne prípady. V prvom riadku je predikcia perzistencie a v ďalších riadkoch sú vyjadrené predikcie autoregresného modelu. Poradie vyhodnotení na Obrázku 3 – 17 kopíruje Tabuľku 3 – 3.



Obrázok 3 – 17 Porovnanie vyhodnotení autoregresného modelu s predikciou 15 minút dopredu podľa rozdielnych nastavení hraničných hodnôt CoV.

3.4.3 Autoregresné modely

Ďalší proces zahŕňal tvorbu autoregresných modelov na predikcie s časovým posunom 30, 45 a 60 minút. Pri nich sme chceli sledovať, ako sa s narastajúcou hodnotou parametra `shift` bude meniť výkonnosť modelov. Trénovanie opäť prebiehalo počas 10 epoch, pričom za najlepšie modely pre každý `shift` sme označili tie, ktoré dosiahli pri testovaní najvyššiu hodnotu metriky *TSS*. Výber modelov je uvedený v Tabuľke 3 – 5.

Model	`shift`	Epocha	MAE
`autoreg`	15	4	0,055
`autoreg`	30	10	0,0061
`autoreg`	45	8	0,0064
`autoreg`	60	4	0,0065

Tabuľka 3 – 5 Výber najlepších autoregresných modelov podľa epochy, v ktorej boli dosiahnuté najlepšie výsledky pri klasifikácii podľa metriky TSS.

Nasledujúce Tabuľky 3 – 6, 3 – 7 a 3 – 8 obsahujú vyhodnotenie modelov podľa časového posunu pre úlohu klasifikácie. Najlepšie výsledky z hľadiska metriky *TSS* sú zvýraznené hrubým písmom.

Model	CoV	P	R	F1	TSS
predikcia perzistencie		0,15	0,15	0,15	0,13
autoregresný	0,01	0,25	0,79	0,38	0,48
	0,02	**0,31**	**0,72**	**0,43**	**0,51**
	0,05	0,49	0,48	0,49	0,42
	0,10	0,69	0,22	0,33	0,21

Tabuľka 3 – 6 Metriky pre predikciu perzistencie a pre autoregresný model s časovým posunom 30 minút podľa rozličných nastavení CoV. Pre testovacie hodnoty bola hraničná hodnota CoV nastavená na 0,5.

Model	CoV	P	R	F1	TSS
predikcia perzistencie		0,13	0,12	0,13	0,11
autoregresný	**0,01**	**0,31**	**0,60**	**0,40**	**0,43**
	0,02	0,36	0,53	0,42	0,41
	0,05	0,47	0,36	0,41	0,31
	0,10	0,59	0,16	0,25	0,15

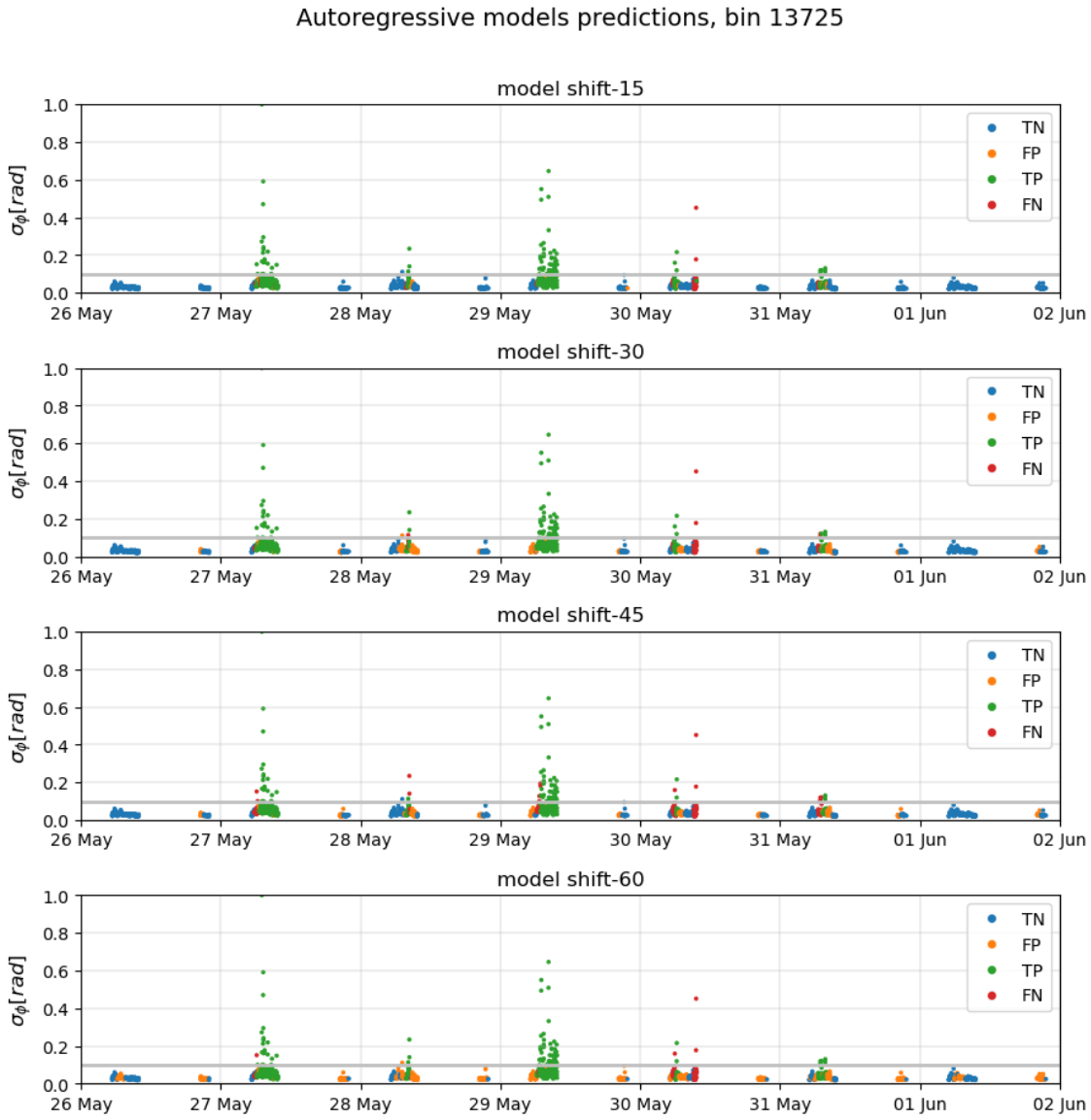
Tabuľka 3 – 7 Metriky pre predikciu perzistencie a pre autoregresný model s časovým posunom 45 minút podľa rozličných nastavení CoV. Pre testovacie hodnoty bola hraničná hodnota CoV nastavená na 0,5.

Model	CoV	P	R	F1	TSS
predikcia perzistencie		0,12	0,11	0,11	0,10
autoregresný	**0,01**	**0,23**	**0,67**	**0,34**	**0,39**
	0,02	0,35	0,46	0,40	0,35
	0,05	0,54	0,22	0,31	0,20
	0,10	0,62	0,13	0,22	0,12

Tabuľka 3 – 8 Metriky pre predikciu perzistencie a pre autoregresný model s časovým posunom 60 minút podľa rozličných nastavení CoV. Pre testovacie hodnoty bola hraničná hodnota CoV nastavená na 0,5.

Tabuľky metrík autoregresných modelov nám potvrdzujú, že s narastajúcim časovým posunom predikovanej hodnoty klesá aj výkonnosť modelov. Aby sme videli o aký veľký rozdiel ide v realite, porovnali sme predikcie výskytu ionosférických scintilácií všetkých autoregresných modelov pomocou grafického zobrazenia (Obrázok 3 – 18). Pre každý model sme vybrali najlepšie nastavenia hodnôt CoV z hľadiska vyhodnotenia pomocou metrík. Z uvedeného obrázku možno vyčítať, že hoci sú hodnoty metrík horšie pre predikcie s narastajúcim časovým posunom, vo výsledku sú najintenzívnejšie scintilácie zachytené a chyba je spôsobená práve nepresnou detek-

ciou hraníc intervalov s výskytom scintilácie. Počet falošne pozitívnych prípadov rastie spolu s časovým posunom.



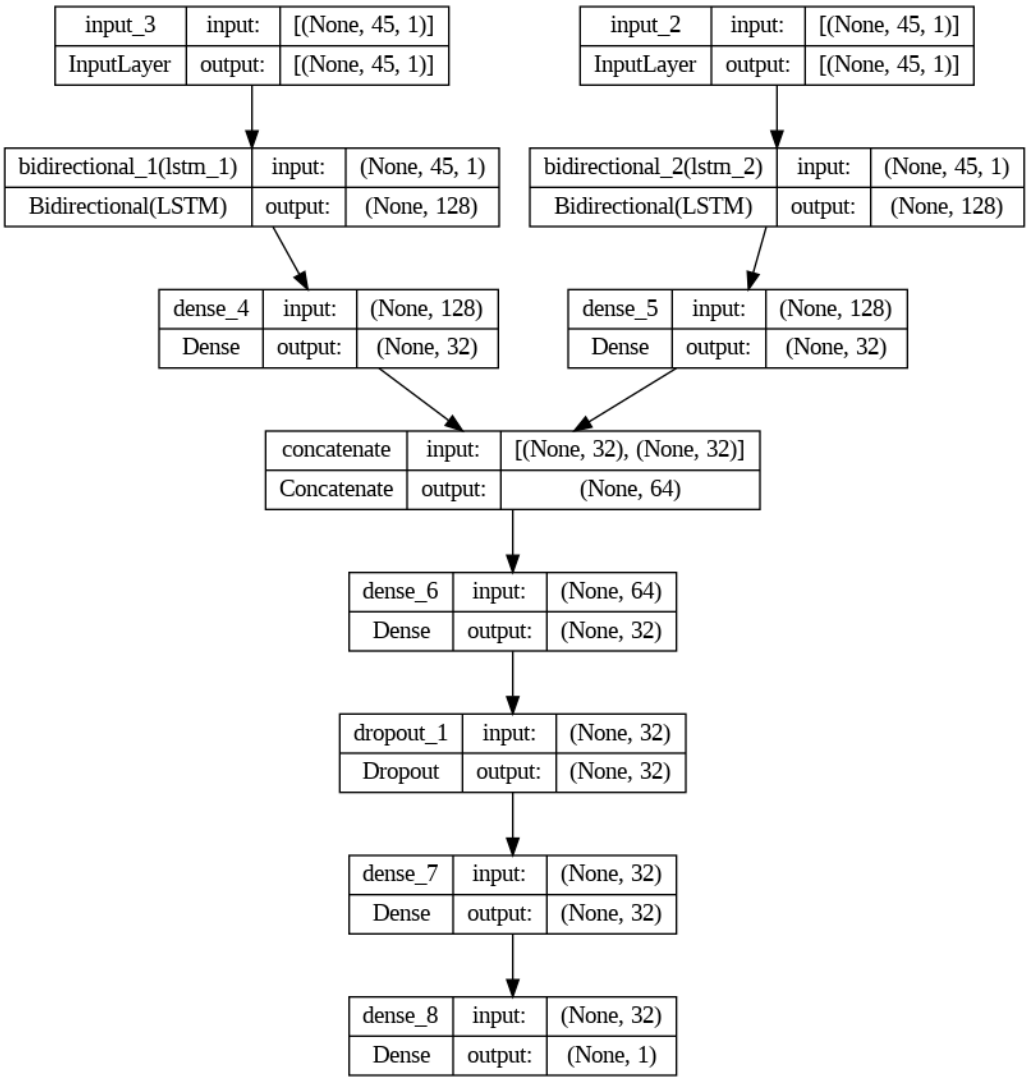
Obrázok 3 – 18 Porovnanie autoregresných modelov pre rôzne časové posuny.

3.4.4 Multivariačné modely

Multivariačná predikcia nazývaná tiež aj ako viacrozmerná, je predikcia výstupnej hodnoty jedného parametra z viacerých vstupných parametrov. V našom prípade uvažujeme o situácii, kedy na predikciu hodnôt parametra σ_ϕ použijeme okrem časových okien σ_ϕ aj časové okná ďalších parametrov.

Pri zostavovaní architektúry neurónových sietí pre multivariačné predikcie sme vychádzali z architektúry siete použitej pri autoregresných modeloch. Architektúra základného multivariačného modelu je zobrazená na Obrázku 3 – 19:

- Prvá vetva je vstupom pre hodnoty σ_ϕ. Jej vstupná vrstva má 45 neurónov a jej výstupy smerujú do obojsmernej rekurentnej vrstvy `LSTM`, kde je dokopy 128 neurónov, 64 pre každý smer.
- Ďalej nasleduje vrstva `Dense` s 32 neurónmi a aktivačnou funkciou `relu`.
- Z druhej vetvy, ktorá spracuje hodnoty pridaného parametra, sa rovnakým postupom cez vrstvy `Bidirectional` a `Dense` dostaneme na koniec vetvy. Pri tejto vetve je dôležité poznamenať, že v prípade parametrov minútových hodnôt je vo vstupnej vrstve priradených 45 neurónov. Ide o parametre *ASY/H*, *Bz* a *PC(N)*. V prípade parametrov, ktorých údaje máme k dispozícii za celé hodiny, má vstupná vrstva tejto vetvy iba 1 neurón. To je prípad parametra *Ap index*.
- Obidve vetvy spojíme do jednej vrstvy `concatenate`, ktorá spojí prichádzajúcich 32 vstupov z každej strany a vytvorí 64 výstupov.
- Nasleduje štandardná vrstva `Dense` s 32 neurónmi, ďalej vrstva `Dropout` s náhodným deaktivovaním 20% neurónov.
- Na záver je dvojica vrstiev `Dense` s 32 neurónmi a s 1 neurónom, ktorého výstup je výstupom celej siete.



Obrázok 3 – 19 Architektúra základného multivariačného modelu.

V prípade trénovania rozšíreného multivariačného modelu, ktorého vstupmi boli časové okná všetkých parametrov, bol začiatok architektúry siete rozdelený na 5 vetiev, ktoré sa po vrstvách `Bidirectional` a `Dense` spojili do jednej vetvy. Pripomenieme ešte, že multivariačné modely boli trénované na predikciu hodnoty intenzity scintilácie o 15 minút dopredu. Trénovanie prebiehalo opäť počas 10 epoch. Prehľad najlepších modelov z epochy s najlepšími výsledkami pri metrike *TSS* je uvedený v Tabuľke 3 – 9.

Model	shift	Epocha	MAE
ap_index	15	1	0,0358
AsyH	15	3	0,0369
BzGSE	15	2	0,0309
PC	15	9	0,0528
all_parameters	15	2	0,0134

Tabuľka 3 – 9 Výber multivariačných modelov podľa epochy, v ktorej dosiahli najlepšie výsledky pri klasifikácii.

V Tabuľke 3 – 10 sú uvedené metriky vyhodnotenia všetkých multivariačných modelov, opäť v závislosti od zvolenej hraničnej hodnoty pre CoV. Vidíme, že základné multivariačné modely dosiahli medzi sebou hodnoty *TSS* v rozmedzí 4 stotín. Ako najlepší sa ukázal multivariačný model na predikciu z *PC(N)* indexu. Nemožno ale povedať, že by ostatné modely zaostávali. Najvyššia návratnosť bola dosiahnutá pri modeli ap_index. Tento model teda dokázal zachytávať pozitívne príklady najlepšie spomedzi všetkých multivariačných modelov s rovnakými nastaveniami CoV.

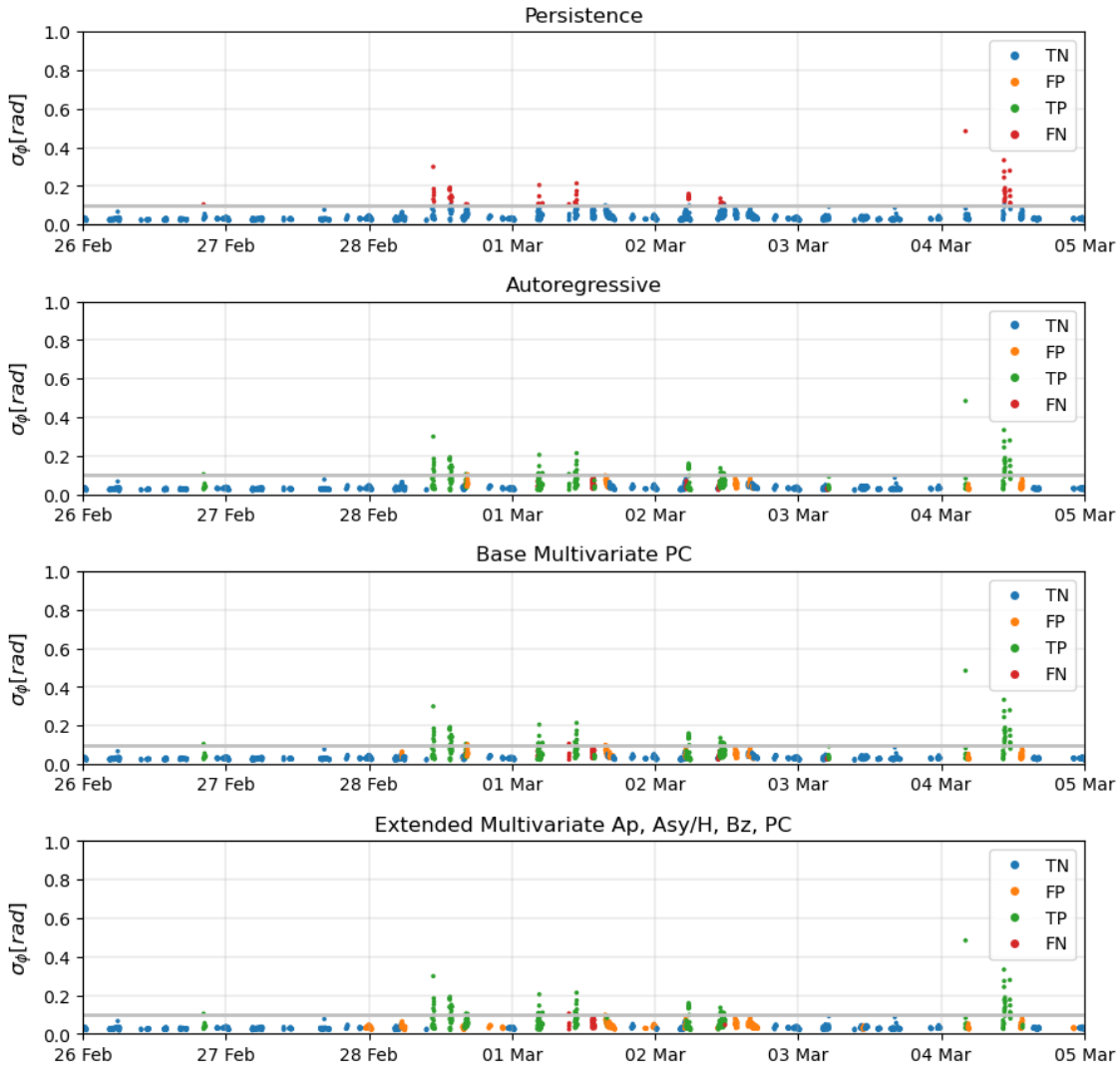
Čo sa týka rozšíreného multivariačného modelu, vidíme, že hoci hodnota *TSS* je najnižšia medzi týmito modelmi, tak návratnosť R presahuje hranicu 0,82 a je dokonca vyššia ako u najlepšieho autoregresného modelu.

Na porovnanie výkonnosti všetkých typov modelov, ktoré predikujú o 15 minút dopredu slúži graf na Obrázku 3 – 20. V ňom sú zahrnuté nasledujúce modely:

- model perzistencie,
- autoregresný model s hodnotou CoV 0,01,
- základný multivariačný model pre predikciu z *PC(N)* s hodnotou CoV 0,1,
- rozšírený multivariačný model s hodnotou CoV 0,05.

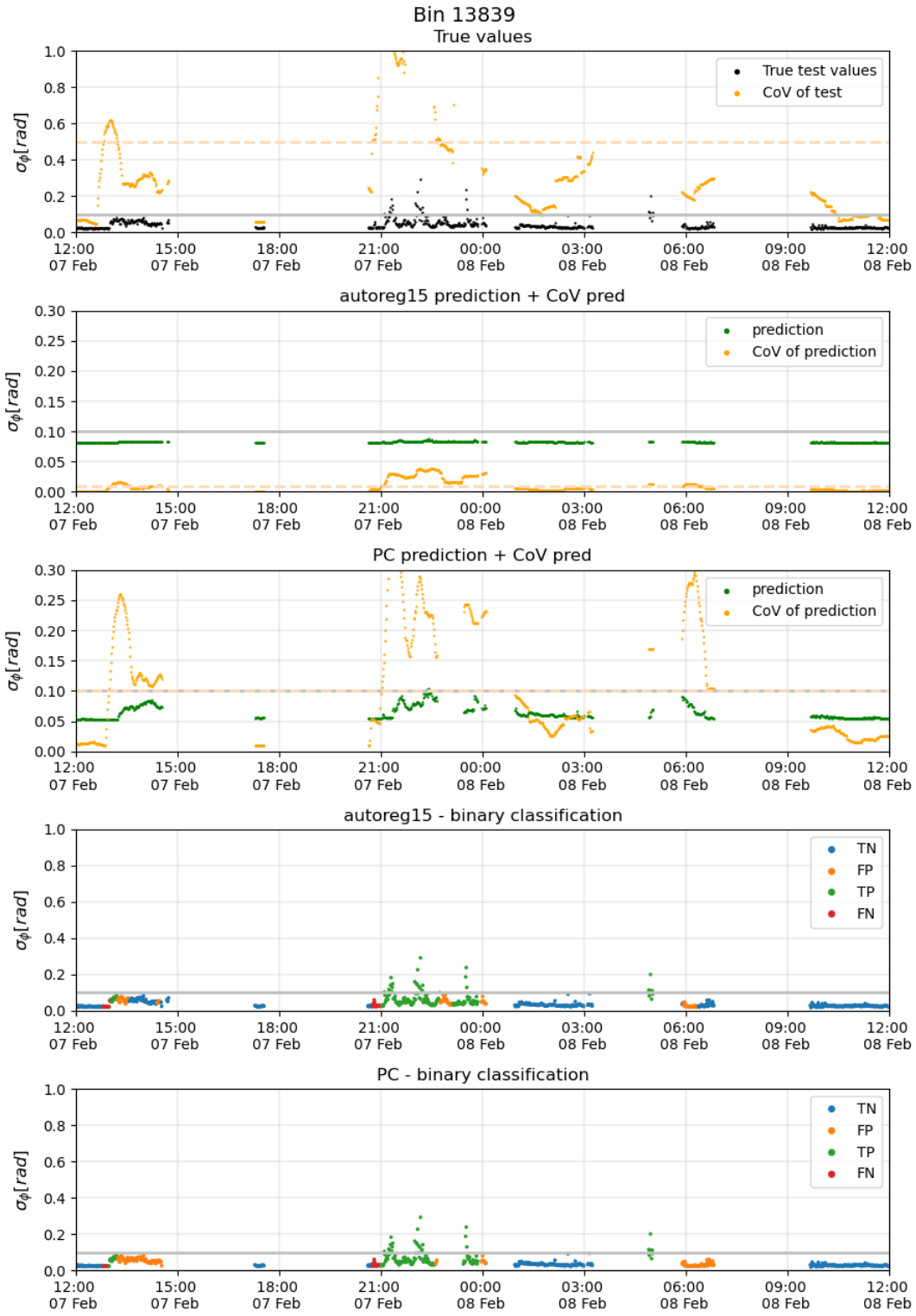
Model	CoV	P	R	F1	TSS
autoregresný	**0,01**	**0,63**	**0,76**	**0,69**	**0,70**
základný multivariačný	**0,05**	**0,40**	**0,84**	**0,54**	**0,67**
`ap_index`	0,10	0,61	0,68	0,64	0,62
	0,15	0,71	0,52	0,60	0,49
základný multivariačný	0,05	0,30	0,88	0,45	0,61
`AsyH`	**0,10**	**0,54**	**0,73**	**0,62**	**0,65**
	0,15	0,68	0,57	0,62	0,53
základný multivariačný	**0,05**	**0,38**	**0,82**	**0,52**	**0,64**
`BzGSE`	0,10	0,49	0,72	0,59	0,62
	0,15	0,59	0,58	0,59	0,53
základný multivariačný	0,05	0,27	0,90	0,41	0,58
`PC`	**0,10**	**0,50**	**0,78**	**0,61**	**0,68**
	0,15	0,67	0,61	0,64	0,57
rozšírený multivariačný	**0,05**	**0,31**	**0,82**	**0,45**	**0,59**
`all_parameters`	0,10	0,42	0,68	0,52	0,56
	0,15	0,50	0,55	0,53	0,48

Tabuľka 3 – 10 Metriky najlepšieho nastavenia CoV autoregresného modelu na predikciu 15 minút dopredu v porovnaní s metrikami multivariačných modelov podľa nastavení CoV. Pre skutočné testovacie hodnoty bola hraničná hodnota CoV nastavená na 0,5.



Obrázok 3 – 20 Porovnanie najlepších modelov každého druhu pre predikciu 15 minút dopredu. Zhora: predikcia perzistencie, autoregresný model, základný multivariačný model s parametrom *PC(N)*, rozšírený multivariačný model so všetkými parametrami.

Po vyhodnotení multivariačných modelov sme chceli zistiť, na čom je založený rozdiel medzi metrikami multivariačného modelu PC a autoregresného modelu. Preto sme sa pozreli na priebehy predikcií, priebehy CoV z predikcií a následne samotnú binárnu klasifikáciu (Obrázok 3 – 21). Zistili sme, že multivariačný model pristupuje k predikcii trochu iným spôsobom. V oblasti scintilácií sú predikcie výrazne kolísavé, čo má za následok vyššie hodnoty CoV. Naproti tomu sa autoregresný model správa tak, že predikuje hodnoty blízke konštante a v prípade, že sa vyskytne scintilácia, začne mierne oscilovať. Pre binárnu klasifikáciu to vyzerá ako lepší prístup. Ak by sme však chceli poznať aj intenzitu scintilácie, nie iba jej prítomnosť, multivariačný model by sa javil ako lepšia voľba. To však zostáva pre túto chvíľu nezodpovedanou otázkou a možným návrhom pre ďalší výskum.



Obrázok 3 – 21 Porovnanie predikcií a klasifikácií autoregresného a multivariačného modelu.

3.5 Vyhodnotenie

Pre porovnanie našich výsledkov s prácou (McGranaghan et al., 2018) sme ich výstupy vyjadrili aj pomocou nami používaných metrík. Tieto výsledky sú uvedené v Tabuľke 3 – 11. Prvý stĺpec označuje spôsob vyhodnotenia modelov, ktorý pre pripomenutie stanovuje podmienku pre klasifikáciu scintilácie ako:

- **1** – 1 hodnota nad 0,1,
- **2** – 3 za sebou idúce hodnoty nad 0,1,
- **3** – 6 za sebou idúcich hodnôt nad 0,1,
- **4** – priemer z 15 za sebou idúcich hodnôt väčší ako 0,1.

eval	FN	FP	TN	TP	TSS	P	R	F1
1	1 071	32 188	89 718	3 252	0,49	0,09	0,75	0,16
2	638	18 218	106 156	1 217	0,51	0,06	0,66	0,11
3	267	7 934	117 626	402	0,54	0,05	0,60	0,09
4	1 053	29 337	92 728	3 111	0,51	0,10	0,75	0,17

Tabuľka 3 – 11 Výsledky pre predikciu s posunom 60 minút dosiahnuté v práci (McGranaghan et al., 2018).

Naše najlepšie dosiahnuté výsledky podľa stanovených hodnôt CoV pre predikciu o 60 minút sú uvedené v Tabuľke 3 – 12.

eval	FN	FP	TN	TP	TSS	P	R	F1
CoV 0,01	133 995	920 642	2 360 857	268 882	0,39	0,23	0,67	0,34

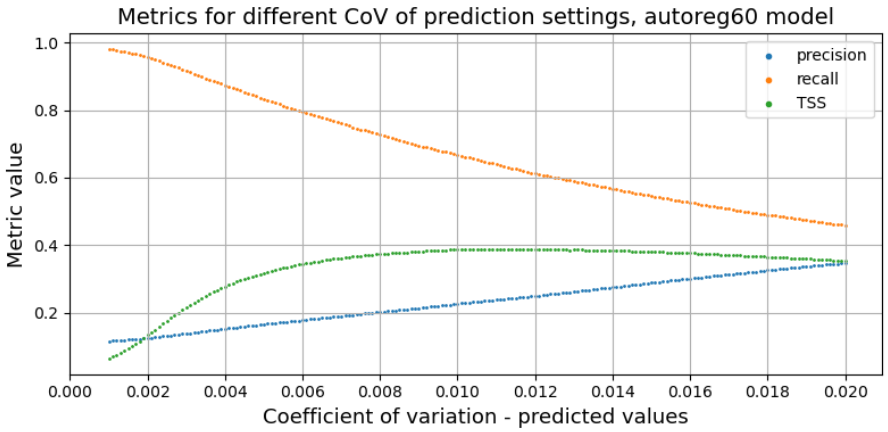
Tabuľka 3 – 12 Výsledky pre predikciu s posunom 60 minút dosiahnuté v tejto práci pomocou autoregresného modelu.

Hoci výsledky podľa metriky *TSS* môžu naznačovať, že model od McGranaghan et al. (2018) je lepší v porovnaní s naším modelom, je avšak dôležité pozrieť sa na

metriku presnosti P. Pri nej dosiahli hodnotu 0,1 iba pri jednom type vyhodnotenia. Naproti tomu náš model na predikciu s časovým posunom 60 minút dosiahol presnosť pre najlepšie TSS 0,23. Ich modely totiž produkovali veľké množstvo falošne pozitívnych príkladov, čím zachytili väčšinu skutočne pozitívnych príkladov, ale otázkou je, či takýto model má uplatnenie. Naproti tomu pri našom modeli vidíme, že aj keď produkuje falošne pozitívne príklady, tak ich umiestnenie sa nachádza práve pri scintilačných oknách.

Pre lepšiu predstavu výsledkov, ktoré dosahuje náš model, je určený Obrázok 3 – 22. Ak mala referenčná práca na výstupe najvyšší *recall* približne 0,75, tak v takom prípade dosiahla metrika *precision* hodnotu iba 0,09, resp. 0,1. Naproti tomu náš model pri rovnakej hodnote *recall* dosiahol približnú hodnotu *precision* až 0,2. Môžeme teda povedať, že pri rovnakej miere návratnosti dosahuje náš model lepšiu *precision*.

Dospeli sme k názoru, že pri porovnaní modelov je potrebné spýtať sa, čo od nich očakávame. Chceme zachytiť všetky scintilácie aj za cenu vysokého počtu falošne pozitívnych prípadov? Alebo chceme počet falošných poplachov radšej minimalizovať? Ak totiž chceme dosiahnuť istú stanovenú úroveň *recall* a pri tom minimalizovať počet FP, je náš model rozhodne lepšou voľbou.



Obrázok 3 – 22 Porovnanie vývoja metrík pre rozličné nastavenia CoV autoregresného modelu s posunom 60 minút.

Ak sa pozrieme na naše požadované ciele, môžeme zhodnotiť, že sa nám v tejto práci podarilo analyzovať a spracovať ionosférické dáta, navrhnúť a natrénovať modely pre predikciu intenzity ionosférickej scintilácie, no a v neposlednom rade sme pomocou variačného koeficientu navrhli spôsob vyhodnocovania výstupov neurónových sietí, aby sme dokázali povedať, ktorá predikovaná hodnota je súčasťou scintilačného okna.

Prístup spracovania výsledkov pomocou variačného koeficientu má svoje obmedzenia. Hraničné hodnoty CoV, ktoré sme počas vyhodnocovania uvádzali, predstavovali nami odhadované hodnoty po vykreslení dát vo forme grafov. Pre iné vstupné dáta, bude potrebné skontrolovať vývoj CoV.

Čo sa týka stanovených hypotéz, tak môžeme skonštatovať, že:

- Vhodným spracovaním výstupov modelov je naozaj možné znížiť počet falošne pozitívnych prípadov a zároveň zachovať vysokú mieru návratnosti.
- Pridaním ďalších vstupných parametrov sme síce nedosiahli lepšie výsledky prostredníctvom metriky *TSS*, no myslíme si, že tieto modely sú určite použiteľné a produkujú zmysluplné výsledky. Otázka pre budúce výskumy scintilácií môže byť, či neexistuje iný spôsob použitia hodnôt ďalších parametrov na predikciu výskytu scintilácií.

3.6 Nasadenie

Samotné nasadenie nebolo súčasťou a ani cieľom tejto diplomovej práce. Výskum v tejto diplomovej práci predstavuje kľúčovú súčasť medzičasom už úspešne ukončeného projektu ASPIS pre ESA. Vyprodukované výstupy tejto diplomovej práce zahŕňajúce vytvorené modely a navrhnuté postupy spracovania predikovaných hodnôt boli použité pre vývoj služby na predikciu ionosférických scintilácií, ktorá je dostupná na `https://aspis.services/#/`.

Pre potreby tejto služby bolo nevyhnutné vytvoriť ďalšie modely pre predikcie z okien s veľkosťou 15 minút, kvôli väčšiemu počtu vytvorených binov v reálnom

čase počas používania služby, než aký by bol počet vytvorených binov pri použití modelov s oknami 45 minút. Napriek tomu ale do budúcna nemožno vylúčiť nasadenie modelov s časovými oknami 45 minút do praxe, keďže počet meracích zariadení a nazbieraných dát bude do budúcna iba rásť.

Vo výsledku predstavujú výstupy tejto diplomovej práce a projektu ASPIS sľubný začiatok cesty ďalšieho rozvoja riešenia tejto problematiky a možného nasadenia celej služby do portálu *Space Weather Service Network* [4], ktorý spravuje ESA, čo by znamenalo zavedenie úplne prvej služby pre predikciu založenej na hlbokom učení do tohto portálu.

[4] `https://www.esa.int/Space_Safety/Space_Weather_Service_Network`

4 Záver

Na záver tejto diplomovej práce chceme zhrnúť tie najdôležitejšie body našej práce. Prvé dve kapitoly predstavovali oboznámenie sa s hlbokým učením a ionosférou, špeciálne s javom zvaným ionosférická scintilácia. Jej výskyt predstavuje jedno z najväčších rizík spomedzi javov kozmického počasia. Ich monitorovanie a poskytnutie informácií o ich možnom výskyte má pre súčasnú technologicky závislú dobu veľkú prioritu.

Cieľom tejto práce bolo vyvinutie automatizovaného prístupu aplikovateľného v reálnom čase na spracovanie ionosférických dát a predikciu scintilácií. Našou úlohou bolo v prvom rade analyzovať ionosférické dáta, ktoré sme mali k dispozícii zo siete kanadských meracích staníc CHAIN. Z nich sme pripravili časové okná, ktoré slúžili ako vstupy do predikčných modelov. Dáta sme vhodne upravili a pomocou nich sme natrénovali modely neurónových sietí na predikciu intenzity fázových ionosférických scintilácií. Podstatným prvkom architektúr neurónových sietí boli rekurentné vrstvy, ktoré sú silným nástrojom na spracovanie časových radov.

Kľúčovou časťou tejto práce bolo skúmanie vhodných spôsobov spracovania výstupov modelov na ich transformáciu do podoby vhodnej pre binárne klasifikovanie. Spomenuli sme metódy vyhladzovania časových priebehov, ktoré sa ale neukázali ako dostatočne užitočné. Následne sme vyskúšali ďalší spôsob založený na výpočte variačného koeficientu, pomocou ktorého sa nám podarilo zachytávať časové oblasti s výskytom ionosférickej scintilácie. Počet falošne pozitívnych príkladov sme výrazne stiahli nadol oproti ostatným výskumom, s ktorými sme sa porovnávali.

Napokon treba konštatovať, že aj pre samotný projekt ASPIS boli tieto výsledky natoľko užitočné, že sme ich sformulovali do podoby odborného článku určeného na publikáciu. Okrem toho boli výstupy tejto diplomovej práce aplikované aj v službe ASPIS: *Autonomous Service for Prediction of Ionospheric Scintillations*. Na úplný záver chceme skonštatovať, že uplatnenie strojového a hlbokého učenia vo vesmírnom výskume má obrovský potenciál, čoho dôkazom je aj táto práca.

Literatúra

Aarons, J. (1982). Global morphology of ionospheric scintillations, *Proceedings of the IEEE* **70**(4): 360–378.

Aquino, M., Moore, T., Dodson, A., Waugh, S., Souter, J. and Rodrigues, F. S. (2005). Implications of ionospheric scintillation for gnss users in northern europe, *The Journal of Navigation* **58**(2): 241–256.

Béniguel, Y. (2002). Global ionospheric propagation model (gim): A propagation model for scintillations of transmitted signals, *Radio Science* **37**(3): 1–14.

Briggs, B. and Parkin, I. (1963). On the variation of radio star and satellite scintillations with zenith angle, *Journal of Atmospheric and Terrestrial Physics* **25**(6): 339–366.

Chollet, F. (2021). *Deep learning with Python*, Simon and Schuster.

Chu, C.-S. J. (1995). Time series segmentation: A sliding window approach, *Information Sciences* **85**(1-3): 147–173.

Crane, R. K. (1977). Ionospheric scintillation, *Proceedings of the IEEE* **65**(2): 180–199.

Davies, K. (1990). *Ionospheric radio*, number 31, IET.

De Lima, G., Stephany, S., De Paula, E., Batista, I. and Abdu, M. (2015). Prediction of the level of ionospheric scintillation at equatorial latitudes in brazil using a neural network, *Space Weather* **13**(8): 446–457.

Ding, B., Qian, H. and Zhou, J. (2018). Activation functions and their characteristics in deep neural networks, *2018 Chinese control and decision conference (CCDC)*, IEEE, pp. 1836–1841.

Duchi, J., Hazan, E. and Singer, Y. (2011). Adaptive subgradient methods for online learning and stochastic optimization., *Journal of machine learning research* **12**(7).

Egea-Roca, D., Arizabaleta-Diez, M., Pany, T., Antreich, F., Lopez-Salcedo, J. A., Paonni, M. and Seco-Granados, G. (2022). Gnss user technology: State-of-the-art and future trends, *IEEE Access* **10**: 39939–39968.

Fejer, B. G. and Kelley, M. (1980). Ionospheric irregularities, *Reviews of Geophysics* **18**(2): 401–454.

Feng, J. and Lu, S. (2019). Performance analysis of various activation functions in artificial neural networks, *Journal of physics: conference series*, Vol. 1237, IOP Publishing, p. 022030.

Feurer, M. and Hutter, F. (2019). Hyperparameter optimization, *Automated machine learning: Methods, systems, challenges* pp. 3–33.

Forte, B., Fallows, R. A., Bisi, M. M., Zhang, J., Krankowski, A., Dabrowski, B., Rothkaehl, H. and Vocks, C. (2022). Interpretation of radio wave scintillation observed through lofar radio telescopes, *The Astrophysical Journal Supplement Series* **263**(2): 36.

Goodfellow, I., Bengio, Y. and Courville, A. (2016). *Deep Learning*, MIT Press. `http://www.deeplearningbook.org`.

Heelis, R. and Maute, A. (2020). Challenges to understanding the earth's ionosphere and thermosphere, *Journal of Geophysical Research: Space Physics* **125**(7): e2019JA027497.

Hlubek, N., Berdermann, J., Wilken, V., Gewies, S., Jakowski, N., Wassaie, M. and Damtie, B. (2014). Scintillations of the gps, glonass, and galileo signals at equatorial latitude, *Journal of Space Weather and Space Climate* **4**: A22.

Hochreiter, S. and Schmidhuber, J. (1997). Long short-term memory, *Neural computation* **9**(8): 1735–1780.

Iman, M., Arabnia, H. R. and Rasheed, K. (2023). A review of deep transfer learning and recent advancements, *Technologies* **11**(2): 40.

Janiesch, C., Zschech, P. and Heinrich, K. (2021). Machine learning and deep learning, *Electronic Markets* **31**(3): 685–695.

Jiao, Y., Morton, Y. T., Taylor, S. and Pelgrum, W. (2013). Characterization of high-latitude ionospheric scintillation of gps signals, *Radio Science* **48**(6): 698–708.

Kingma, D. P. and Ba, J. (2014). Adam: A method for stochastic optimization, *arXiv preprint arXiv:1412.6980* .

Kintner, P., Humphreys, T. and Hinks, J. (2009). Gnss and ionospheric scintillation, *Inside GNSS* **4**(4): 22–30.

Kintner, P. M., Ledvina, B. M. and De Paula, E. (2007). Gps and ionospheric scintillations, *Space weather* **5**(9).

LeCun, Y. (2019). 1.1 deep learning hardware: Past, present, and future, *2019 IEEE International Solid-State Circuits Conference-(ISSCC)*, IEEE, pp. 12–19.

LeCun, Y., Bottou, L., Orr, G. B. and Müller, K.-R. (2002). Efficient backprop, *Neural networks: Tricks of the trade*, Springer, pp. 9–50.

Maxwell, A. and Little, C. (1952). A radio-astronomical investigation of winds in the upper atmosphere, *Nature* **169**(4305): 746–747.

McGranaghan, R. M., Mannucci, A. J., Wilson, B., Mattmann, C. A. and Chadwick, R. (2018). New capabilities for prediction of high-latitude ionospheric scintillation: A novel approach with machine learning, *Space Weather* **16**(11): 1817–1846.

Moshayedi, A. J., Roy, A. S., Kolahdooz, A. and Shuxin, Y. (2022). Deep learning application pros and cons over algorithm deep learning application pros and cons over algorithm, *EAI Endorsed Transactions on AI and Robotics* **1**(1).

Nielsen, M. A. (2015). *Neural networks and deep learning*, Vol. 25, Determination press San Francisco, CA, USA.

Pfaff, R. F. (2012). The near-earth plasma environment, *Space science reviews* **168**: 23–112.

Pi, X., Iijima, B. A. and Lu, W. (2017). Effects of ionospheric scintillation on gnss-based positioning, *Navigation: Journal of The Institute of Navigation* **64**(1): 3–22.

Priyadarshi, S. (2015). A review of ionospheric scintillation models, *Surveys in geophysics* **36**: 295–324.

Rama Rao, P., Gopi Krishna, S., Niranjan, K. and Prasad, D. (2006). Study of spatial and temporal characteristics of l-band scintillations over the indian low-latitude region and their possible effects on gps navigation, *Annales Geophysicae*, Vol. 24, Copernicus Publications Göttingen, Germany, pp. 1567–1580.

Rawer, K. (2013). *Wave propagation in the ionosphere*, Vol. 5, Springer Science & Business Media.

Rezende, L. F. C. d., De Paula, E., Stephany, S., Kantor, I., Muella, M., De Siqueira, P. and Correa, K. (2010). Survey and prediction of the ionospheric scintillation using data mining techniques, *Space Weather* **8**(6).

Rishbeth, H. (1998). How the thermospheric circulation affects the ionospheric f2-layer, *Journal of Atmospheric and Solar-Terrestrial Physics* **60**(14): 1385–1402.

Rojas, R. and Rojas, R. (1996). The backpropagation algorithm, *Neural networks: a systematic introduction* pp. 149–182.

Rosenblatt, F. (1957). *The perceptron, a perceiving and recognizing automaton Project Para*, Cornell Aeronautical Laboratory.

Ruder, S. (2016). An overview of gradient descent optimization algorithms, *arXiv preprint arXiv:1609.04747* .

Russell, S. J. and Norvig, P. (2016). *Artificial intelligence: a modern approach*, Pearson.

Savitzky, A. and Golay, M. J. (1964). Smoothing and differentiation of data by simplified least squares procedures., *Analytical chemistry* **36**(8): 1627–1639.

Schuster, M. and Paliwal, K. K. (1997). Bidirectional recurrent neural networks, *IEEE transactions on Signal Processing* **45**(11): 2673–2681.

Secan, J., Bussey, R., Fremouw, E. and Basu, S. (1995). An improved model of equatorial scintillation, *Radio Science* **30**(3): 607–617.

Sen, S., Agarwal, S., Chakraborty, P. and Singh, K. P. (2022). Astronomical big data processing using machine learning: A comprehensive review, *Experimental Astronomy* **53**(1): 1–43.

Sharma, S., Sharma, S. and Athaiya, A. (2017). Activation functions in neural networks, *Towards Data Sci* **6**(12): 310–316.

Spogli, L., Alfonsi, L., De Franceschi, G., Romano, V., Aquino, M. and Dodson, A. (2009). Climatology of gps ionospheric scintillations over high and mid-latitude european regions, *Annales Geophysicae*, Vol. 27, Copernicus Publications Göttingen, Germany, pp. 3429–3437.

Srivastava, N., Hinton, G., Krizhevsky, A., Sutskever, I. and Salakhutdinov, R. (2014). Dropout: a simple way to prevent neural networks from overfitting, *The journal of machine learning research* **15**(1): 1929–1958.

Talaei Khoei, T., Ould Slimane, H. and Kaabouch, N. (2023). Deep learning: Systematic review, models, challenges, and research directions, *Neural Computing and Applications* **35**(31): 23103–23124.

Testa, A., Gallo, D. and Langella, R. (2004). On the processing of harmonics and interharmonics: Using hanning window in standard framework, *IEEE Transactions on Power Delivery* **19**(1): 28–34.

Theußl, T., Hauser, H. and Gröller, E. (2000). Mastering windows: Improving reconstruction, *Proceedings of the 2000 IEEE symposium on Volume visualization*, pp. 101–108.

Tieleman, T. (2012). Lecture 6.5-rmsprop: Divide the gradient by a running average of its recent magnitude, *COURSERA: Neural networks for machine learning* **4**(2): 26.

Wirth, R. and Hipp, J. (2000). Crisp-dm: Towards a standard process model for data mining, *Proceedings of the 4th international conference on the practical applications of knowledge discovery and data mining*, Vol. 1, Manchester, pp. 29–39.

Yeh, K. C. and Liu, C.-H. (1982). Radio wave scintillations in the ionosphere, *Proceedings of the IEEE* **70**(4): 324–360.

Yu, Y., Si, X., Hu, C. and Zhang, J. (2019). A review of recurrent neural networks: Lstm cells and network architectures, *Neural computation* **31**(7): 1235–1270.

Zoznam príloh

Príloha A CD médium – záverečná práca v elektronickej podobe, prílohy v elektronickej podobe.

Príloha B Odkaz na dokumentáciu a zdrojový kód, ktoré sú dostupné online: `https://github.com/skostarova/Diplomovy_projekt_Kostarova`