Technická univerzita v Košiciach

Fakulta elektrotechniky a informatiky

Prediktívne modely v spracovaní dát z oblasti kozmického počasia

Bakalárska práca

2024 **Terézia Drengubiaková**

Technická univerzita v Košiciach

Fakulta elektrotechniky a informatiky

Prediktívne modely v spracovaní dát z oblasti kozmického počasia

Bakalárska 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 **Terézia Drengubiaková**

Abstrakt v SJ

Bakalárska práca sa zaoberá prediktívnymi modelmi na spracovanie dát z oblasti kozmického počasia. Kozmické počasie, najmä geomagnetické búrky, môžu výrazne ovplyvniť technológie na Zemi, a preto je ich spoľahlivá predikcia kľúčová. Práca popisuje metódy strojového učenia, umelé neurónové siete a hlboké učenie využívané na predpoveď geomagnetických búrok. Praktická časť zahŕňa analýzu historických dát, výber vhodných atribútov na modelovanie pomocou rôznych typov neurónových sietí a populárneho modelu typu transformer. Výsledky modelov sú vyhodnotené na základe rôznych metrík presnosti predikcie.

Kľúčové slová

Geomagnetické búrky, hlboké učenie, umelá neurónová sieť, prediktívne modely, LSTM, transformer

Abstrakt v AJ

The bachelor thesis deals with predictive models for processing data in the field of space weather. Space weather, especially geomagnetic storms, can significantly impact Earth's technologies, making reliable prediction crucial. The thesis describes machine learning methods, artificial neural networks, and deep learning used for predicting geomagnetic storms. The practical part includes the analysis of historical data, selection of suitable attributes for modeling using various types of neural networks and the popular transformer model. The results of the models are evaluated based on various prediction accuracy metrics.

Kľúčové slová v AJ

Geomagnetic storms, deep learning, artificial neural network, predictive models, LSTM, transformer

73234

TECHNICKÁ UNIVERZITA V KOŠICIACH

FAKULTA ELEKTROTECHNIKY A INFORMATIKY

Katedra kybernetiky a umelej inteligencie

ZADANIE BAKALÁRSKEJ PRÁCE

Študijný odbor: **Informatika**

Študijný program: **Hospodárska informatika**

Názov práce:

Prediktívne modely v spracovaní dát z oblasti kozmického počasia

Predictive models in space weather data processing

Študent: **Terézia Drengubiaková**

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

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

Konzultant práce: **Ing. Viera Krešňáková, PhD., RNDr. Šimon Mackovjak, PhD.**

Pracovisko konzultanta: **Ústav experimentálnej fyziky SAV**

Pokyny na vypracovanie bakalárskej práce:

1. Podať teoretický prehľad problematiky metód hlbokého učenia, s dôrazom na vhodné metódy pre predikciu indexov kozmického počasia.
2. Získať a predspracovať vhodné dáta pre riešenie zvoleného problému.
3. Navrhnúť, realizovať a vyhodnotiť experimenty s predikciou zvoleného indexu kozmického počasia.
4. Vypracovať dokumentáciu podľa pokynov katedry a vedúceho práce.

Jazyk, v ktorom sa práca vypracuje: slovenský

Termín pre odovzdanie práce: 24.05.2024

Dátum zadania bakalárskej práce: 31.10.2023

TECHNICKÁ UNIVERZITA V KOŠICIACH FAKULTA ELEKTROTECHNIKY A INFORMATIKY -11-

v.z. Pichidová

prof. Ing. Liberios Vokorokos, PhD.

dekan fakulty

Čestné vyhlásenie

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

Košice 24. 5. 2024 .....

Vlastnoručný podpis

Poďakovanie

Chcela by som poďakovať vedúcemu bakalárskej práce doc. Ing Petrovi Butkovi, PhD. a mojim konzultantom, pani Ing. Viere Krešňákovej, PhD. a RNDr. Šimonovi Mackovjakovi, PhD. za ich ústretovosť, vedenie a čas, ktorý mi venovali pri písaní tejto práce. Taktiež by som sa rada poďakovala svojej rodine za ich podporu počas celého štúdia.

Obsah

Úvod **1**

1 Umelá inteligencia **2**

1.1 Strojové učenie 3

1.2 Umelá neurónová sieť 3

1.3 Úvod do hlbokého učenia 6

1.4 Architektúra hlbokých neurónových sietí 7

1.4.1 Rekurentné neurónové siete 8

1.4.2 Long Short-Term Memory (LSTM) siete 9

1.4.3 Transformačná neurónová sieť 12

2 Kozmické počasie **15**

2.1 Geomagnetické búrky 15

2.2 Analýza súčasného stavu 17

3 Predikcia DST pomocou hlbokého učenia **19**

3.1 Pochopenie a príprava dát 19

3.2 Použité metriky 25

3.3 Modelovanie 27

3.3.1 RNN s využitím LSTM a Bidirectional LSTM vrstvy 28

3.3.2 Využitie architektúry transformer na spracovanie sekvencií . . 35

3.4 Vyhodnotenie 38

4 Záver **41**

Zoznam príloh **46**

Zoznam obrázkov

1 – 1 Porovnanie biologického a umelého neurónu 4
1 – 2 Architektúra neurónovej siete 6
1 – 3 Problém miznúcich gradientov 9
1 – 4 LSTM pamäťový blok s 1 bunkou 10
1 – 5 Rozdiel medzi RNN a LSTM modelom 11
1 – 6 Architektúra transformačného modelu. Zdroj: Vaswani et al. (2017) 13
2 – 1 Interakcia medzi slnečným vetrom a magnetickým poľom Zeme . . 16
3 – 1 Zoznam chýbajúcich hodnôt atribútov 20
3 – 2 Grafy zobrazujúci korelácie medzi atribútmi a DST 21
3 – 3 Graf zobrazujúci 10 najvhodnejších atribútov na predpoveď DST . 22
3 – 4 Vytvorenie nových atribútov 24
3 – 5 Rozdelenie datasetu 25
3 – 6 Sumarizácia modelu 29
3 – 7 Vyhodnotenie zvolených metrík pre atribút DST 31
3 – 8 Vyhodnotenie zvolených metrík pre atribút DST a BZ_GSM . . . 32
3 – 9 Vyhodnotenie zvolených metrík pre atribút DST a KP_index . . . 34
3 – 10 Vyhodnotenie zvolených metrík pre atribút DST, BZ_GSM a KP_index 35
3 – 11 Vyhodnotenie zvolených metrík pre atribút DST 38
3 – 12 Predikcie modelov v porovnaní so skutočnými hodnotami 40

Zoznam tabuliek

2 – 1 Rozdelenie geomagnetických búrok podľa intenzity 17
3 – 1 Konfúzna matica pre klasifikačný model 30
3 – 2 Výsledky klasifikačného modelu 30
3 – 3 Konfúzna matica pre klasifikačný model 32
3 – 4 Výsledky klasifikačného modelu 32
3 – 5 Konfúzna matica pre klasifikačný model 33
3 – 6 Výsledky klasifikačného modelu 33
3 – 7 Konfúzna matica pre klasifikačný model 34
3 – 8 Výsledky klasifikačného modelu 35
3 – 9 Konfúzna matica pre klasifikačný model 37
3 – 10Výsledky klasifikačného modelu 37
3 – 11Porovnanie metrík dvoch modelov 38

Zoznam symbolov a skratiek

ACC , z ang. *Accuracy*, úspešnosť

ACE z angl. *Advanced Composition Explorer*, Advanced Composition Explorer

Adam , z angl. *Adaptive Moment Estimation*, optimalizačná funkcia

AI z angl. *Artificial Intelligence*, umelá inteligencia

CEC z angl. *Constant Error Carousel*, konštantný chybový kruh

DST z angl. *Disturbance Storm Time*, čas búrkového nepokoja

FP z angl. *False Positive*, falošne pozivítne

FN z angl. *False Negative*, falošne negatívne

GSM z angl. *Geocentric Solar Wind*, geocentrický slnečný vietor

IMF z angl. *Interplanetary Magnetic Field*, medziplanetárne magnetické pole

LSTM z angl. *Long Short-Term Memory*, dlhá krátkodobá pamäť

MAE , z angl. *Mean Absolute Error*, stredná absolútna chyba

MCC , z angl. *Matthew's correlation coefficient*, Matthewsov korelačný koeficient

MSE , z agnl. *Mean Squared Error*, stredná kvadratická chyba

RNN z angl. *Recurrent Neural Network*, rekurentná neurónová sieť

TP z angl. *True Positive*, pravdivo pozitívne

TN z angl. *True Negative*, pravdivo pozitívne

Úvod

Geomagnetické búrky sú prírodné fenomény, ktoré vznikajú v dôsledku zmien v magnetickom poli Zeme. Majú výrazne negatívny vplyv na rôzne aspekty nášho života, narúšajúc fungovanie technologických systémov, ako sú elektroenergetické siete, komunikačné siete a satelitné navigačné systémy. Preto je nevyhnutné mať k dispozícii spoľahlivé nástroje na ich predpovedanie s dostatočným predstihom. Takéto predpovede umožňujú prijatie opatrení na minimalizáciu negatívnych dopadov týchto búrok.

V poslednom čase získavajú metódy hlbokého učenia čoraz väčšiu popularitu, a to nielen pri predpovedaní geomagnetických búrok. Hlavným motívom tejto práce je vytvorenie, natrénovanie a vyhodnotenie modelu, ktorý bude vedieť predikovať výskyt geomagnetických búrok s niekoľko hodinovým predstihom s čo najvyššou presnosťou na základe sekvencie historických dát.

Prvá kapitola sa zameriame na predstavenie strojového a hlbokého učenia, umelých neurónových sietí a architektúry použitých modelov. V druhej kapitole si podrobnejšie prejdeme geomagnetické búrky, aké sú ich príčiny vzniku a ako ich klasifikujeme. Následne sa pozrieme, aké sú dostupné riešenia a výsledky, ktoré prinášajú. V Kapitole 3 sa venujeme metrikám vyhodnocovania a predstavíme praktickú časť s výsledkami našej práce.

1 Umelá inteligencia

Ľudstvo sa už od nepamäti snaží neustále napredovať. Od priemyselnej revolúcie však došlo k veľkému rozvoju technológií. Mnohá komplikovaná a namáhavá ručná práca bola nahradená strojmi. Umelá inteligencia je taktiež jednou z technických inovácií.

Definícia inteligencie a jej aplikovateľnosť na stroje

Umelá inteligencia (AI) je fascinujúcim a rýchlo sa rozvíjajúcim odvetvím, ktoré sa zameriava na vytváranie strojov schopných vykonávať úlohy, ktoré by zvyčajne vyžadovali ľudskú inteligenciu. Napriek rozsiahlemu výskumu a vývoju v tejto oblasti, definícia AI zostáva prekvapivo obtiažne uchopiteľná a otvorená interpretáciám. Podľa Websterovho Slovníka *Webster's New World Dictionary (1988)* inteligencia zahŕňa schopnosť učiť sa alebo rozumieť skúsenostiam, schopnosť získavať a uchovávať znalosti, mentálnu schopnosť, schopnosť rýchlo a úspešne reagovať na nové situácie a použitie schopnosti racionálneho uvažovania pri riešení problémov, či efektívne riadenie správania.

V kontexte AI sa tieto definície stretávajú s výzvou, keď sa aplikujú na stroje. Významné je, že ak má byť koncept inteligencie aplikovateľný na neživé stroje, nemôže byť inteligencia vymedzená tak, že by stroje nemohli byť inteligentné z definície. Táto perspektíva nás vedie k zamysleniu sa nad tým, čo presne umožňuje strojom vykazovať správanie, ktoré by sme mohli charakterizovať ako inteligentné (Fetzer and Fetzer, 1990).

Uvažovanie o tom, čo robí stroje inteligentnými, nás vedie k hlbšiemu pochopeniu umelej inteligencie. Keďže AI neustále napreduje, naše pochopenie toho, čo to znamená, keď stroj preukazuje „inteligenciu“, sa neustále vyvíja a rozširuje. Tento pohyb vpred v oblasti AI nás núti prehodnotiť naše tradičné názory na inteligenciu, stroje a ich vzájomné pôsobenie v našom svete.

1.1 Strojové učenie

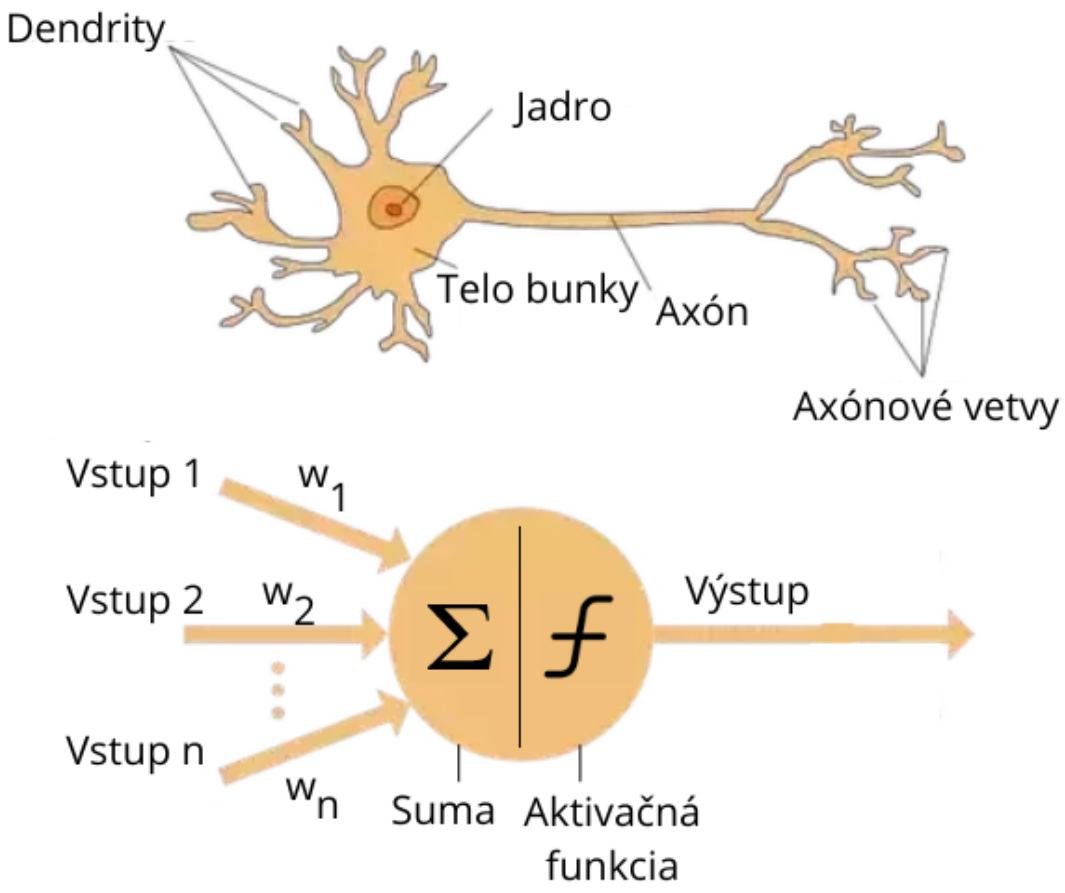
Strojové učenie predstavuje kľúčové odvetvie informatiky, ktoré vychádza z umelej inteligencie. Svoje uplatnenie primárne hľadá v oblastiach ako je predikcia a klasifikácia. Princípom strojového učenia je umožniť počítačom „učiť sa“ a zlepšovať svoj výkon pri rôznych úlohách prostredníctvom „skúseností“, čo v kontexte algoritmov často znamená schopnosť adaptácie na základe dostupných dát. Hranice medzi strojovým učením a tradičnými štatistickými prístupmi sú často nejasné, pretože definícia toho, čo sa považuje za strojové učenie, závisí nielen od použitých techník, ale aj od historického kontextu a interpretácie. Napriek tomu sa strojové učenie odlišuje svojim dôrazom na prediktívnu presnosť, často pracuje s rozsiahlymi dátovými súbormi, ktoré sú charakteristické vysokou dimenzionalitou. Tento prístup je neoceniteľný najmä v ére „Veľkých dát“, kde tradičné štatistické metódy narazia na svoje limity a strojové učenie ponúka rozsiahle riešenia pre komplexné výzvy spojené s analýzou obrovských a komplexných dátových súborov (Bi et al., 2019).

1.2 Umelá neurónová sieť

Neurónová sieť je vzájomne prepojený súbor jednoduchých spracovateľských prvkov, jednotiek alebo uzlov, ktorých funkčnosť je voľne založená na biologickom neuróne. Spracovateľská schopnosť siete je uložená vo váhach medzi jednotkami, ktoré sa získavajú procesom adaptácie alebo učenia zo sady trénovacích vzorov. Môžeme teda povedať, že neurónová sieť je model, ktorý sa inšpiruje fungovaním mozgu a jeho základnou jednotkou sú neuróny. Pre porovnanie sa môžeme pozrieť na Obrázok 1 – 1, kde vidíme podobnosť medzi biologickým a umelým neurónom. Či už hovoríme o jednej alebo druhej forme neurónovej siete, všimneme si, že neuróny sú spojené medzi sebou a každý má svoje váhy, ktoré určujú, aký význam má pre neurón signál z iných neurónov. Vstupné signály, ktoré sú podobné aktivácii neurónov v mozgu, sú násobené týmito váhami a sčítané. Ak výsledok presiahne istý prah, neurón sa aktivuje a vygeneruje výstup (Gurney, 2018).

Podľa Dongare et al. (2012) sa dozvedáme, že neurónová sieť pozostáva z nasledujúcich častí:

- vstupné hodnoty, ktorá prijíma vstupy problému *vstup* n,
- súbor synapsí charakterizovaných synaptickými váhami w,
- sčítacieho spoja alebo lineárneho kombinátora pre súčet vstupných signálov $\sum$,
- aktivačnej funkcie na obmedzenie amplitúdy výstupu neurónu na nejakú konečnú hodnotu,
- výstupná hodnota,



Obrázok 1 – 1 Porovnanie biologického a umelého neurónu

Neurónové siete môžu mať rôzne štruktúry a architektúry. Napríklad neuróny môžu byť usporiadané do vrstiev, pričom každá vrstva môže mať svoju špecifickú úlohu pri spracovaní vstupných údajov, viď Obrázok 1 – 2. Existuje viacero typov neurónových sietí, ale všeobecne sú známe ako dopredné, rekurentné alebo konvolučné siete. Tvorí ich viacero vrstiev neurónov, ktoré sú navzájom prepojené v sieti. Tieto vrstvy podľa Uzair and Jamil (2020) delíme na:

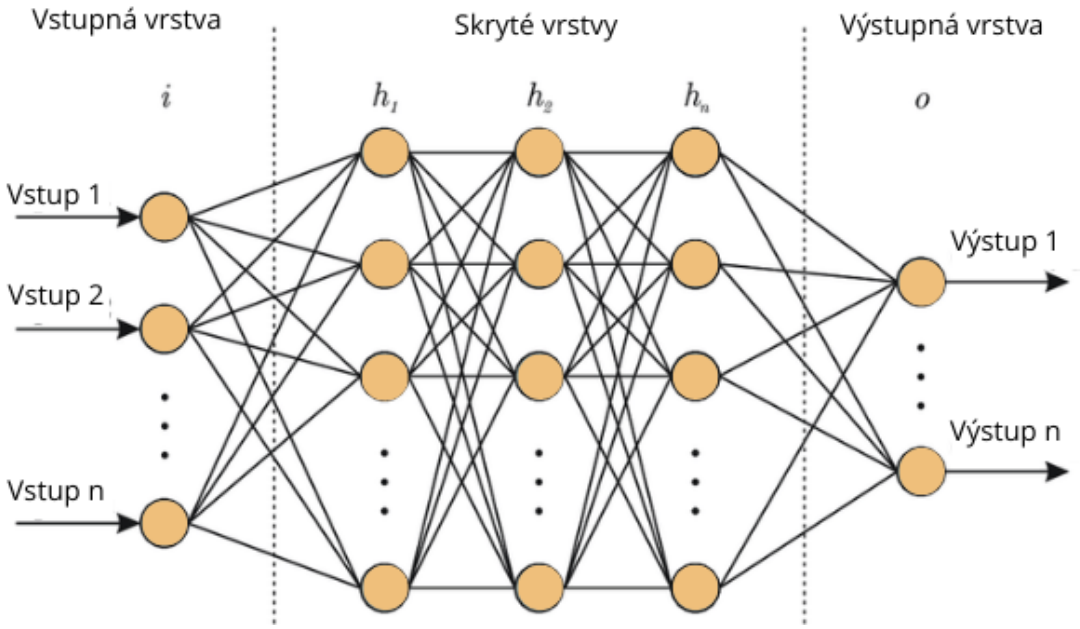
- Vstupná vrstva je vrstva v neurónovej sieti, ktorá slúži na komunikáciu s vonkajším prostredím. Keď je do vstupnej vrstvy siete zavedený vstup, výstupná vrstva začne produkovať výsledné vzory. Taktiež definuje podmienky, na ktorých závisí trénovanie siete.

- Skrytá vrstva je základne medziprepojovacia vrstva medzi vstupnou a výstupnou vrstvou siete. Počet skrytých vrstiev v sieti sa mení od problému k problému. Niektoré problémy potrebujú jedinú skrytú vrstvu, niektoré potrebujú viaceré skryté vrstvy. Skryté vrstvy sú zoskupením malých neurónov, ktoré prenášajú údaje a trénovaciu vrstvu do vrstiev.

- Výstupná vrstva je tiež spojená s vonkajším prostredím a reprezentuje výstup vstupu, ktorý bol poskytnutý sieti, vonkajšiemu prostrediu. Neuróny vo výstupných vrstvách majú priamu súvislosť s komplexnosťou problému alebo práce daného sieti.

Trénovanie neurónovej siete spočíva v upravovaní váh medzi neurónmi na základe trénovacích dát. Proces učenia umožňuje neurónovej sieti prispôsobiť sa a naučiť sa rozpoznávať vzory v dátach. To sa dosahuje opakovaným predstavovaním trénovacích vzorov sieti a porovnávaním jej výstupu s požadovanými výstupmi.

Cieľom trénovania je, aby neurónová sieť vedela generalizovať svoje znalosti a úspešne sa správala aj pri nových, neznámych dátach. Táto schopnosť generalizácie je kľúčová pre efektívne fungovanie neurónových sietí vo viacerých úlohách, ako je klasifikácia, predikcia alebo rozpoznávanie vzorov Gurney (2018).



Obrázok 1 – 2 Architektúra neurónovej siete

1.3 Úvod do hlbokého učenia

Hlboké učenie je podmnožinou strojového učenia. So strojovým učením sa dnes môžeme stretnúť na každom kroku; pri vyhľadávaní na internete, filtrovaní obsahu na sociálnych sieťach alebo pri spotrebiteľských produktoch ako sú kamery či smartfóny. Systémy strojového učenia používame na identifikáciu objektov na obrázkoch, prepis reči na text alebo pri už vyššie spomenutom vyhľadávaní či komercii. Čoraz častejšie sa však na tieto úkony využíva trieda techník nazývaná hlboké učenie (LeCun et al., 2015). Slovo hĺbka v spojení hlboké učenie znamená postupnosť vrstiev reprezentácií, nie je odkazom na nejaké hlbšie pochopenie pri využití tejto metódy (Chollet, 2019).

Spoločnou vlastnosťou všetkých modelov hlbokého učenia je už vyššie spomínané učenie reprezentácií. Niekedy sa to nazýva aj učenie vlastností. Tento termín označuje model, ktorý sa učí nové a lepšie reprezentácie oproti surovým dátam. Dôležité je, že modely hlbokého učenia nezískavajú konečnú reprezentáciu v jednom kroku, ale viacero reprezentácií, ktoré zodpovedajú transformáciám reprezentácií na via-

cerých úrovniach medzi skrytými vrstvami. To znamená, že každá vrstva sa učí koncepty z dát, na ktorých stavajú nasledujúce vrstvy; čím vyššia úroveň, tým abstraktnejšie koncepty sa učia. V jednoduchosti, neurónová sieť na detekciu tvarov sa najprv naučí rozoznávať jednoduché hrany v prvej vrstve a následne pridáva rozpoznávanie zložitejších tvarov, ktoré sa skladajú z daných hrán v nasledujúcich vrstvách.

Ďalšou spoločnou vlastnosťou modelov hlbokého učenia je, že následné transformácie medzi vrstvami sú nelineárne. To zvyšuje expresívnu silu modelu. Okrem toho, individuálne reprezentácie nie sú navrhované manuálne, ale učia sa prostredníctvom trénovacích dát. To robí modely hlbokého učenia veľmi flexibilnými (Emmert-Streib et al., 2020).

Hlboké neurónové siete sa v súčastnosti používajú v mnohých aplikáciách umelej inteligencie. Od aplikácie hlbokej neurónovej siete na rozpoznávanie reči a obrazov sa počet aplikácií využívajúcich danú sieť výrazne zvýšil. Môžeme ju nájsť v samo-riadiacich autách, pri detekcii rakoviny alebo pri hraní hier. V mnohých oblastiach je sieť schopná prekonať presnosť ľudí. Jej nadštandardná výkonnosť pochádza z jej schopnosti extrahovať vysokoúrovňové vlastnosti zo surových senzorických údajov po použití štatistického učenia nad veľkým množstvom údajov na získanie efektívnej reprezentácie vstupného priestoru. Nadštandardná presnosť však prichádza s vysokou výpočtovou zložitosťou, preto sú hlboké neurónové siete odkázané na spoľahlivý a výkonný hardware (Sze et al., 2017).

1.4 Architektúra hlbokých neurónových sietí

Existujú rôzne typy architektúr hlbokého učenia. Ich výber závisí od požadovanej úlohy. Niektoré z hlavných modelov hlbokého učenia podľa Smys et al. (2020) sú:

- Autoenkóder - umelá neurónová sieť, ktorá má schopnosť učiť sa rôzne kódovacie vzorce,
- *Deep Belief Network* - vhodná na zvládanie nekonvexných účelových funkcií

pri použití viacvrstvového vnímania,

- Konvolučná neurónová sieť (CNN) - typ doprednej neurónovej siete, kde sú jednotlivé neuróny usporiadané tak, aby reagovali na všetky oblasti v danej oblasti,
- Rekurentné neurónové siete (RNN) - typ neurónovej siete špeciálne navrhnutej na spracovanie sekvenčných dát,

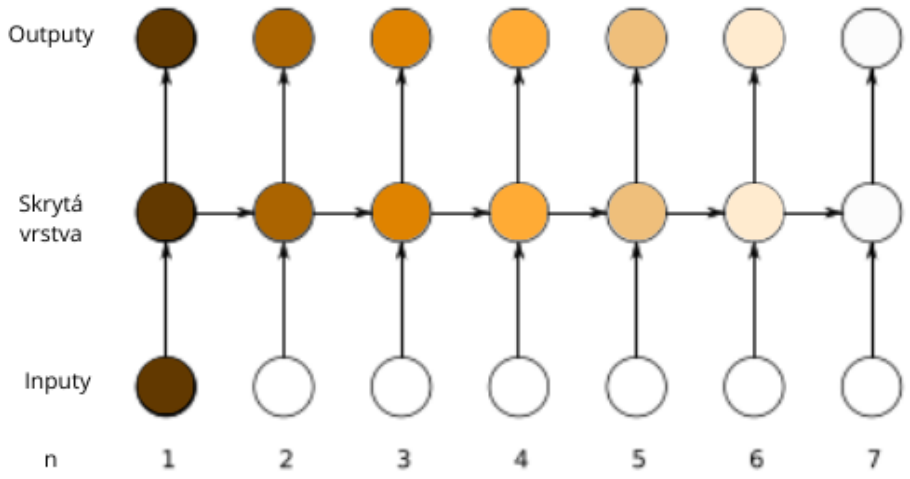
V tejto podkapitole si popíšeme typy neurónových sietí, ktoré budeme využívať v praktickej časti tejto práce.

1.4.1 Rekurentné neurónové siete

Rekurentné neurónové siete (RNN) sú triedou neurónových sietí, ktorá sa využíva na spracovanie sekvenčných dát. Sekvenčnými dátami môžeme nazvať tie dáta, ktoré zachytávajú postupný vývoj určitého javu a vykazujú úzku vzájomnú súvislosť medzi svojimi predchádzajúcimi a nasledujúcimi hodnotami. RNN sa skladá z vstupnej vrstvy, skrytej vrstvy a výstupnej vrstvy, pričom skrytá vrstva zohráva najdôležitejšiu úlohu. Pri tomto type neurónovej siete skrytá vrstva neobsahuje len výstup vstupnej vrstvy, ale aj výstup z jej predchádzajúcich udalostí v sekvencii. Oproti bežným neurónovým sieťam, ktoré navzájom spájajú len vrstvy, RNN spája aj neuróny medzi vrstvami. Môžeme použiť skrytú vrstvu ako úložný priestor pre celú sieť. Keď je rekurentná neurónová sieť rozvinutá, zistíme, že ju možno využiť na riadenie klasifikácie, kde si pamätá predchádzajúce informácie a aplikuje ich. To zásadne odlišuje RNN od tradičných dopredných neurónových sietí (Shang et al., 2021). RNN sú však obmedzené na pohľad do minulosti asi desať časových krokov. Toto je spôsobené tým, že spätný signál buď zaniká, alebo sa nadmerne zväčšuje . Tento problém bol riešený pomocou rekurentných neurónových sietí s dlhým a krátkodobým pamäťovým systémom (LSTM-RNN). LSTM siete sú do istej miery biologicky pravdepodobné a schopné naučiť sa viac ako 1 000 časových krokov, v závislosti od zložitosti postaveného modelu (Staudemeyer and Morris, 2019).

1.4.2 Long Short-Term Memory (LSTM) siete

Štandardné rekurentné neurónové siete nedokážu prekročiť viac ako 5 až 10 časových radov. Je to spôsobené tým, že pri spätnom šírení chýb sa chybová funkcia zväčšuje alebo zmenšuje každým časovým radom. Preto sa chyba nadmerne zväčšuje alebo zaniká. Vysoká chyba spôsobuje osciláciu váh, zatiaľ čo príliš nízka vedie k problému miznúcich gradientov, kedy sa učenie spomalí alebo zastaví úplne Staudemeyer and Morris (2019). Na Obrázku 1 – 3 môžeme vidieť, ako sa citlivosť časom n znižuje, keďže nové vstupy prepisujú aktivácie skrytej vrstvy a sieť tak postupne zabúda na prvé vstupy (zaniká). Tieňovanie jednotlivých uzlov v sieti ukazuje jej citlivosť na vstupy v čase, čím tmavší odtieň, tým vyššia citlivosť Graves and Graves (2012).

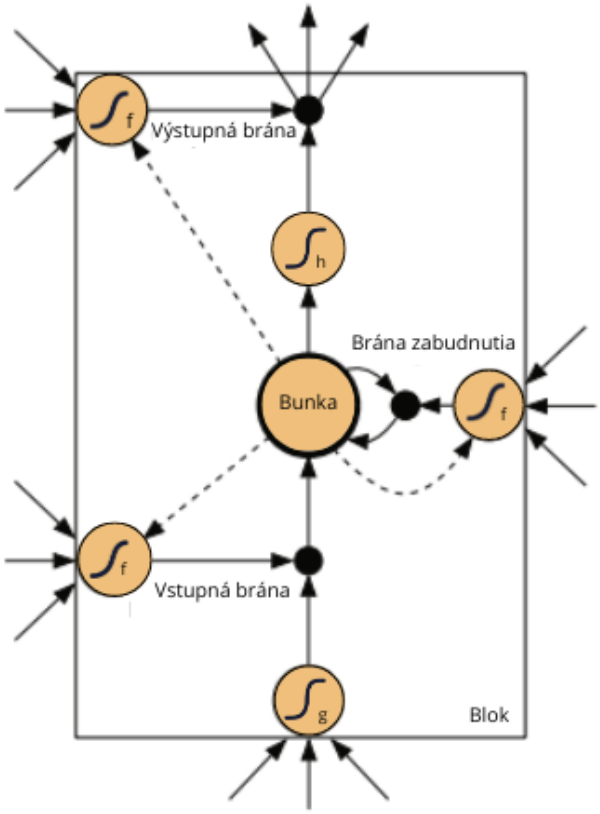


Obrázok 1 – 3 Problém miznúcich gradientov

Model LSTM je výkonný rekurentný neurónový systém špeciálne navrhnutý na prekonanie daného problému. Vieme sa mu vyhnúť použitím konštantného chybového kruhu (CEC), ktorý udržiava chybový signál v bunkách každej jednotky. Tieto bunky sú tak rekurentné samé o sebe, majú zaujímavejšiu architektúru tým, že CEC je rozšírený o ďalšie funkcie, a to o brány vstupu a výstupu, ktoré tvoria pamäťovú bunku. Samorekurentné spojenia značia spätnú väzbu s oneskorením jedného časového radu

Van Houdt et al. (2020).

Architektúra LSTM sa teda skladá z množiny rekurentne prepojených podsietí, známych ako pamäťové bloky. Každý blok obsahuje prinajmenšom jednu samorekurentne pamäťovú bunku a tri multiplikačné jednotky, vstupnú, výstupnú a zabudnutú bránu, ktoré poskytujú spojitý analóg zápisu, čítania a resetovania operácií v bunke.

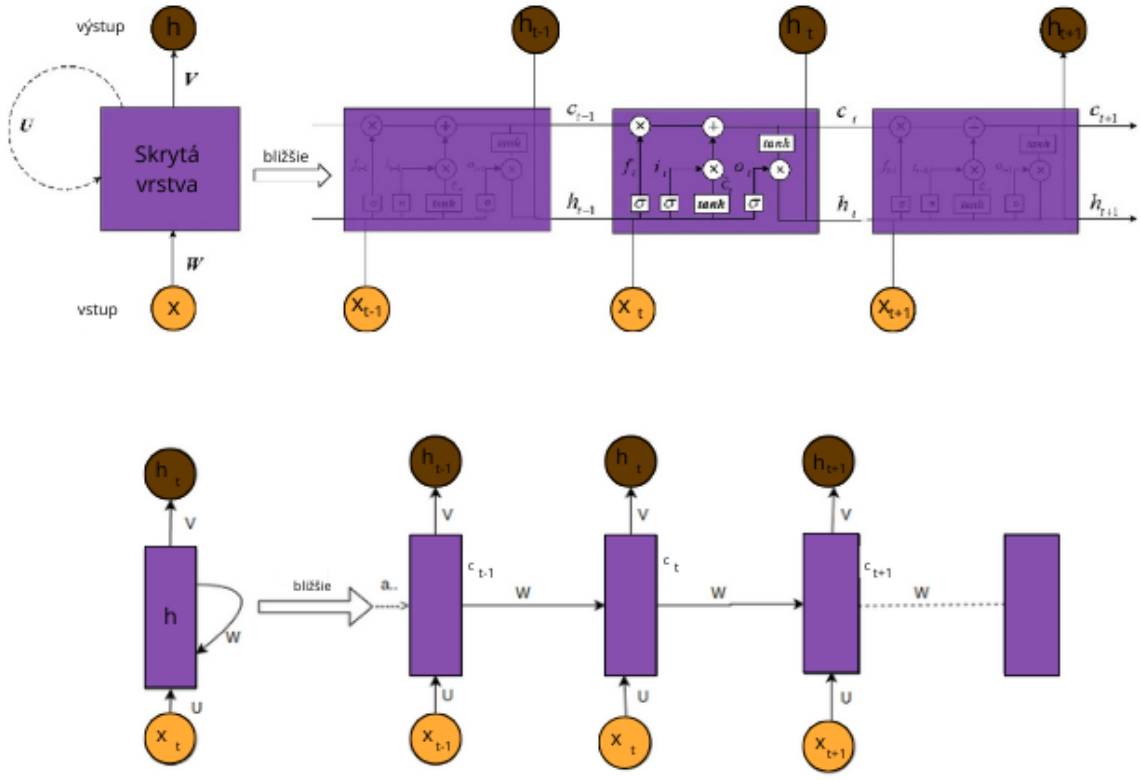


Obrázok 1 – 4 LSTM pamäťový blok s 1 bunkou

Pamäťový blok jednej bunky sa nachádza na Obrázku 1 – 4. Ako môžeme vidieť, tieto 3 brány sú nelineárne súčtové jednotky, ktoré zbierajú aktivitu zvonka aj z vnútra bloku, čím regulujú aktiváciu bunky prostredníctvom násobení, ktoré zobrazujú malé čierne kruhy. Vstupná a výstupná brána násobí vstup a výstup bunky. Na druhú stranu, brána zabudnutia násobí predchádzajúci stav bunky. V danej

bunke sa nepoužíva žiadna aplikačná funkcia, ktorá zabezpečuje zmenu výsledku na požadovaný interval. Aktivačná funkcia brány f je obvykle logistická sigmoida, čo znamená, že aktivity brán sú obvykle medzi 0 (zatvorená brána) a 1 (otvorená brána). Aktivačné funkcie vstupu a výstupu (g a h) bunky sú zvyčajne hyperbolický tangens alebo taktiež logistická sigmoida. Váhové spojenia z bunky do brán zobrazujú prerušované čiary. Všetky ostatné spojenia bloku sú neváhové alebo majú váhu rovnú 1 alebo 0. Výstupy z bloku do siete vychádzajú z násobenia výstupnej brány.

Siete LSTM sú v podstate rovnaké ako štandardné RNN, čo ich však odlišuje je, že sčítacie jednotky v skrytej vrstve sú nahradené pamäťovými blokmi (Graves and Graves, 2012). Túto skutočnosť si môžeme lepšie predstaviť za pomoci Obrázka 1 – 5 kde vidíme LSTM (hore) a bežnú RNN (dolu).



Obrázok 1 – 5 Rozdiel medzi RNN a LSTM modelom

1.4.3 Transformačná neurónová sieť

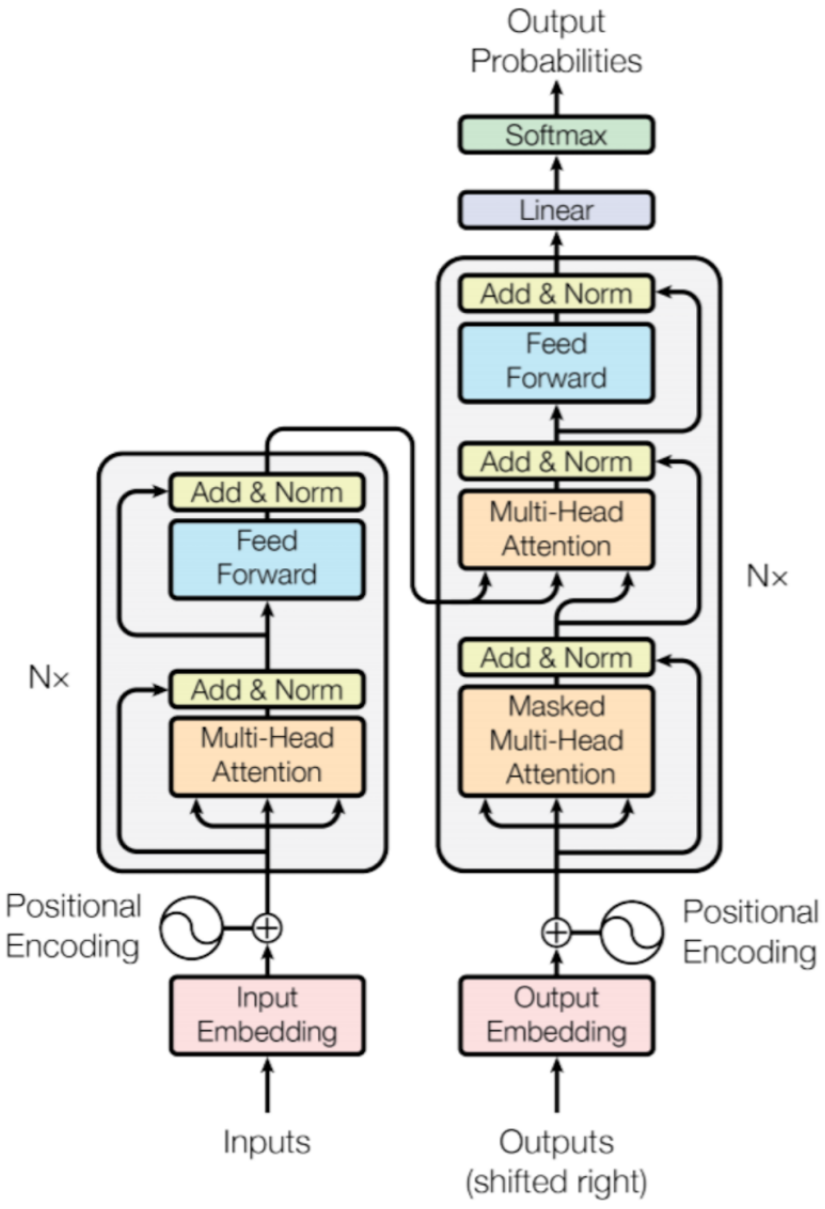
Transformačné neurónové siete sa stali v posledných rokoch veľmi populárnymi, nakoľko prekonávajú výkonnosť tradičných rekurentných a konvolučných neurónových sietí. Architektúra modelu transformátora sa spolieha na tzv. mechanizmus *attention* (pozornosť) a pozičného kódovania namiesto rekurzie(Vaswani et al., 2017). Tento prístup umožňuje oveľa vyššiu paralelizáciu ako rekurentné neurónové siete pre modelovanie sekvencií. *Attention* mechanizmus teda vynikol ako významná technika v učení sekvencií, kde je potrebná dlhodobá pamäť. Model transformátora sa zvyčajne skladá z dvoch hlavných blokov - kódera a dekódera. Pred vstupom dát do kódera je každá vstupná sekvencia mapovaná na vektor. Spracovaný vstup je následne odoslaný do bloku kódera. Kóder zabezpečuje mapovanie vstupnej sekvencie do abstraktnej spojitej reprezentácie. Dekóder má za úlohu spracovať túto reprezentáciu a postupne produkuje jediný výstup, pričom sú mu neustále pridávané predošlé výstupy (Afifi et al., 2023). Architektúru modelu môžeme vidieť na Obrázku 1 – 6.

Kóder

Kóder časť sa skladá z šiestich identických vrstiev označených ako $N = 6$. Každá z týchto vrstiev obsahuje dve podvrstvy. Prvá podvrstva je založená na mechanizme tzv. *multi-head self-attention*, zatiaľ čo druhá podvrstva predstavuje jednoduchú doprednú sieť, ktorá je plne prepojená. Okolo každej podvrstvy je implementované reziduálne spojenie, po ktorom nasleduje vrstva normalizácie. Funkcia každej podvrstvy je definovaná ako:

$$\text{LayerNorm}(x + \text{Sublayer}(x)), \tag{1.1}$$

kde Sublayer(x) označuje funkciu, ktorú realizuje daná podvrstva. Na podporu týchto reziduálnych spojení majú všetky podvrstvy v modeli, vrátane vrstiev pre vkladanie a výstupný rozmer veľkosti rovný 512.



Obrázok 1 – 6 Architektúra transformačného modelu. Zdroj: Vaswani et al. (2017)

Dekóder

Dekodér je rovnako zostavený z šiestich identických vrstiev, označených ako $N = 6$. V každej vrstve dekodéra sa nachádzajú tri podvrstvy, pričom tretia podvrstva je pridaná na vykonanie *multi-head attention* nad výstupmi zásobníka kódera. Ako v prípade kódera, aj tu sú implementované reziduálne spojenia okolo každej podvrstvy, nasledované vrstvou normalizácie. V dekodéri je špeciálne upravená podvrstva *self-attention* tak, aby zabránila pozíciám v zameriavaní sa na následujúce pozície.

Attention

Funkcia *attention* môže byť definovaná ako mapovacie zariadenie, ktoré priraďuje dopyt a množinu dvojíc kľúč-hodnota k vektorovému výstupu. Dopyt, kľúče, hodnoty a výstup predstavujú vektory. Výstup sa určuje ako vážený súčet hodnôt, pričom váhy pridelené jednotlivým hodnotám sú určené pomocou funkcie kompatibility, ktorá hodnotí súlad dotazu s odpovedajúcim kľúčom (Vaswani et al., 2017).

2 Kozmické počasie

Vesmírne počasie označuje stav slnečnej aktivity a slnečného vetra, magnetosféry, ionosféry a termosféry, ktoré môžu mať vplyv na fungovanie a spoľahlivosť technologických systémov vo vesmíre aj na zemi, a taktiež môžu ovplyvniť ľudský život a zdravie (Schwenn, 2006).

2.1 Geomagnetické búrky

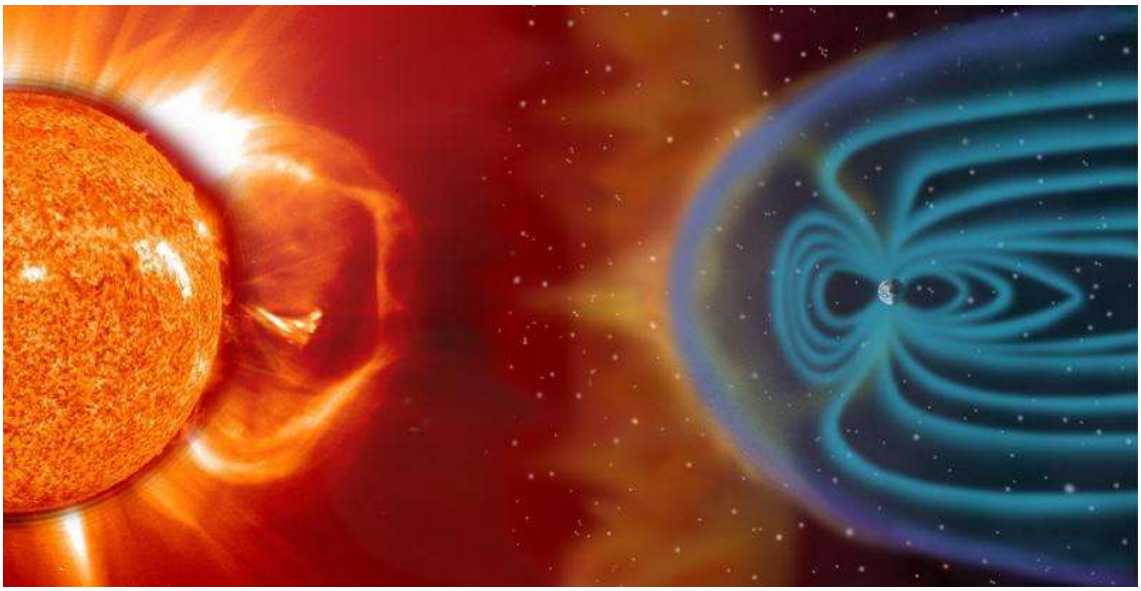
Geomagnetická búrka predstavuje výnimočné obdobie v dynamike Zeme, kedy náš planetárny magnetický štít - magnetosféra - čelí zvýšenej aktivite v dôsledku silného elektrického poľa, ktoré vzniká medzi planétami našej slnečnej sústavy. Toto pole, generované slnečným vetrom, je dostatočne silné a trvá dostatočne dlho na to, aby výrazne zvýšilo energiu v systéme spojenom medzi magnetosférou a ionosférou Zeme. Dôsledkom je posilnenie takzvaného prstencového prúdu - prúdu nabitých častíc, ktorý obklopuje Zem v určitej vzdialenosti od rovníka. Keď tento prúd dosiahne určitú intenzitu, prekračuje stanovené prahové hodnoty špecifického indexu, nazývaného DST, ktorý slúži na kvantifikovanie intenzity geomagnetickej búrky. Výsledkom je rad javov, od krásnych polárnych žiar po možné narušenie komunikačných systémov a elektrických sietí na Zemi (Gonzalez et al., 1994).

Vznik geomagnetickej búrky

Geomagnetické búrky podľa Gonzalez et al. (1994) vznikajú vplyvom slnečného vetra. Kľúčovým mechanizmom prenosu energie z tohto slnečného vetra je magnetická rekonencia medzi medziplanetárnym poľom (IMF) a magnetickým poľom Zeme. Tento proces môže viesť k uvoľneniu veľkého množstva energie a vytvoriť silné elektrické pole, ktoré sa tiahne cez celú magnetosféru Zeme.

Toto elektrické pole potom poháňa nabitú energiu do magnetosféry a ionosféry Zeme, čo spôsobuje zvýšenie energie častíc a prúdov okolo našej planéty, viď Ob-

rázok 2 – 1. Jedným z hlavných dôsledkov je posilnenie takzvaného prstencového prúdu - prúdu nabitých častíc, ktorý obieha Zem. Keď tento prúd dosiahne určitú silu, spôsobuje to, čo nazývame geomagnetickou búrkou (Gonzalez et al., 1994).



Obrázok 2 – 1 Interakcia medzi slnečným vetrom a magnetickým poľom Zeme

Klasifikácia geomagnetických búrok

Geomagnetické búrky sú klasifikované podľa hodnôt indexu geomagnetickej aktivity DST, ktorý meria zmeny geomagnetického poľa na Zemi. Škála DST sa pohybuje v rozmedzí približne +100 nT až -600 nT. V tabuľke 2 – 1 môžeme vidieť rozdelenie do hlavných skupín podľa intenzity.

Za pokojný čas považujeme jav, kedy je DST rovné 0, avšak tento dej nenastáva často. Podľa analýzy Magsat z roku 1979 je základnou hranicou hodnota približne -20 nT. Môžeme teda povedať, že negatívne hodnoty DST sú bežné, nemusia nutne signalizovať geomagnetickú búrku; preto pri určovaní intenzity geomagnetických búrok je dôležité zohľadniť špecifické stupnice rôznych autorov, ako napríklad tie, ktoré uvádza (Gonzalez et al., 1994):

Tabuľka 2 – 1 Rozdelenie geomagnetických búrok podľa intenzity

kategória	DST (nT)
slabá	⟨-30 ; -50)
stredná	⟨-50 ; -100)
silná	⟨-100 ; -250)
veľmi silná	-250 a menej

2.2 Analýza súčasného stavu

V súčasnom stave sa výrazne rozvíja využitie pokročilých techník strojového učenia a neurónových sietí v oblasti predpovedania vesmírneho počasia. Tradičné metódy predpovedania DST indexu, ako sú fyzikálne modely alebo jednoduchšie strojové učenie, majú svoje obmedzenia, najmä v presnosti a schopnosti efektívne zohľadniť komplexnú dynamiku slnečného vetra a jeho interakciu s magnetosférou Zeme.

V článku *Multi-Hour-Ahead Dst Index Prediction Using Multi-Fidelity Boosted Neural Networks* sa autori zaoberajú metodikou a výsledkami predpovedania indexu DST pomocou pokročilých techník strojového učenia. Autori najprv trénujú svoj model pomocou metódy strojového učenia známej ako Gated Recurrent Unit (GRU), ktorá je variantom rekurentných neurónových sietí. Toto trénovanie využíva parametre slnečného vetra na predpoveď DST pre silné búrkové obdobia s predstihom 1-6 hodín.

Štúdia zdôrazňuje kritický aspekt odhadu neistôt spojených s predpoveďami DST. To sa dosahuje metódou Accurate and Reliable Uncertainty Estimate (ACCRUE), ktorá tiež využíva sieť GRU. Unikátnym aspektom ich prístupu je metóda multi-fidelity boosting, ktorá iteratívne zlepšuje model zameraním sa na podmnožiny trénovacej sady, ktoré predstavujú najväčšie výzvy pri predpovedi, čím sa zvyšuje presnosť a spoľahlivosť modelu.

Výsledky štúdie sú sľubné, ukazujúce, že model s multi-fidelity boostingom môže predpovedať hodnoty DST 6 hodín vopred s chybou strednej kvadratickej odchýlky

13.54 nT, čo prekonáva jednoduchšie modely. Autori demonštrujú účinnosť modelu pomocou skutočných údajov o geomagnetických búrkach, pričom osobitne zdôrazňujú jeho výkon počas búrok na Halloween v rokoch 2003 a 2022. Tieto prípadové štúdie ukazujú, že model môže presne predpovedať načasovanie a intenzitu vrcholov DST, čo je kľúčové pre efektívnu predpoveď vesmírneho počasia.

Na záver, článok zdôrazňuje potenciál techník strojového učenia, najmä prístupu siete s multi-fidelity boostingom, v zlepšovaní presnosti a spoľahlivosti predpovedí vesmírneho počasia. Tento pokrok má významné dôsledky pre ochranu infraštruktúry vo vesmíre a na zemi pred účinkami geomagnetických búrok. Štúdia tiež položila základy pre ďalší výskum s cieľom ďalej zlepšiť tieto modely a rozšíriť ich prediktívne schopnosti (Hu et al., 2023).

3 Predikcia DST pomocou hlbokého učenia

V tejto kapitole sa zameráme na predikciu DST indexu, na ktorú využijeme viaceré typy modelov. Naším cieľom je nájsť najvhodnejší model a dopracovať sa k čo najlepšiemu výsledku. Aby sme dosiahli tento cieľ, budeme sa venovať pochopeniu a analýze dát, výberu najvhodnejších atribútov pre predikcie, najvhodnejšieho modelu a thresholdu. V závere našej predikcie rozdelíme na dve kategórie: geomagnetická búrka nastala alebo nenastala. Na trénovanie využijeme GPUs TeslaK40c, Quadro RTX4000 a Quadro P4000 s povolením KKUI TUKE[1].

Dataset OMNI je dostupný online[2]. Dataset bol získaný pomocou knižníc Aidapy a Heliopy v rámci výskumnej skupiny počas predchádzajúcich prác.

3.1 Pochopenie a príprava dát

Máme k dispozícii dataset s 520 443 záznamami a 55 atribútmi. Naším cieľom je určiť, či došlo k geomagnetickej búrke, čo indikuje atribút DST index, v datasete označený ako `DST`. Úlohou je predpovedať tento atribút a na základe jeho hodnoty klasifikovať pravdepodobnosť nadchádzajúcej búrky niekoľko hodín vopred. Kritická hodnota pre definíciu geomagnetickej búrky, podľa Kapitoly 2 a analýzy Magsat, je stanovená na -20nT. Hodnoty DST indexu nižšie ako tento prah indikujú výskyt búrky (klasifikované ako 1), zatiaľ čo hodnoty nad týmto prahom znamenajú normálny stav (klasifikované ako 0).

Ako prvé sme si overili, či sa v datasete nachádzajú chýbajúce hodnoty a ak áno, tak aké sú ich počty. Na Obrázku 3 – 1 môžeme vidieť, že pri niektorých atribútoch tieto čísla presahujú aj 60% (napríklad atribút `PR$FLX_4`). Čo je pre nás dôležité je fakt, že atribút `DST` neobsahuje žiadne chýbajúce hodnoty.

V ďalšom kroku sme využili korelačnú analýzu. Táto štatistická metóda sa používa na posúdenie vzťahov medzi dvoma alebo viacerými kvantitatívnymi premen-

[1]Pristup umožňuje: `https://datalab.kkui.fei.tuke.sk/gpu`

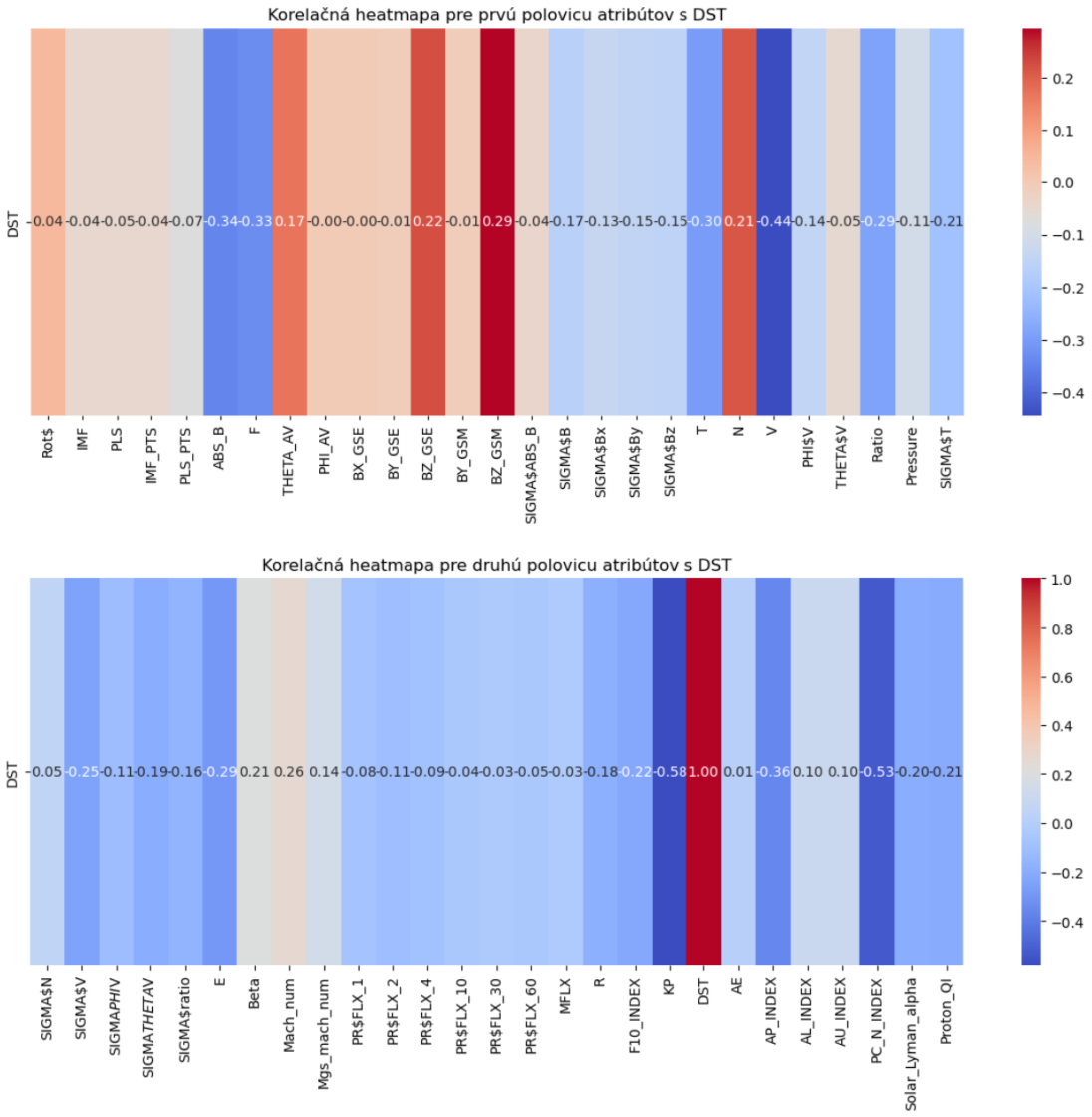
[2]Zdroj datasetu: `https://spdf.gsfc.nasa.gov/pub/data/omni/low_res_omni/`

time1	0	BY_GSE	124067	V	127822	E	151606	R	0
Rot$	0	BZ_GSE	124067	PHI$V	159323	Beta	176713	F10_INDEX	523
IMF	0	BY_GSM	124093	THETA$V	194992	Mach_num	160867	KP	0
PLS	0	BZ_GSM	124093	Ratio	228207	Mgs_mach_num	176708	DST	0
IMF_PTS	0	SIGMA$ABS_B	141086	Pressure	144153	PR$FLX_1	286726	AE	0
PLS_PTS	0	SIGMA$B	124223	SIGMA$T	164786	PR$FLX_2	323149	AP_INDEX	0
ABS_B	124067	SIGMA$Bx	124622	SIGMA$N	165563	PR$FLX_4	323150	AL_INDEX	0
F	124067	SIGMA$By	124609	SIGMA$V	159753	PR$FLX_10	155408	AU_INDEX	0
THETA_AV	124067	SIGMA$Bz	124606	SIGMAPHIV	174407	PR$FLX_30	155441	PC_N_INDEX	117482
PHI_AV	124067	T	161020	SIGMA$THETA$V	208894	PR$FLX_60	155455	Solar_Lyman_alpha	41
BX_GSE	124067	N	144145	SIGMA$ratio	228207	MFLX	0	Proton QI	160870

Obrázok 3 – 1 Zoznam chýbajúcich hodnôt atribútov

nými. Pomocou tejto analýzy vieme zistiť, či existuje štatisticky významný vzťah medzi premennými. Ak áno, zisťujeme aká je sila a smer tohto vzťahu. Vieme určiť, či je medzi premennými pozitívny alebo negatívny vzťah. Ak je vzťah pozitívny, znamená to, že ak jedna premenná rastie, rastie aj druhá, negatívny vzťah poukazuje na opačný jav ak jedna premenná rastie, druhá klesá. Daná analýza nám tiež umožňuje zistiť mieru sily tohto vzťahu. V našej práci sme ju vyjadrili pomocou Paersonovho korelačného koeficientu, ktorý môže nadobúdať hodnoty od -1 po 1, kde hodnota blízko -1 alebo 1 značí silný vzťah, zatiaľ čo hodnota blízko 0 naznačuje slabý alebo neexistujúci vzťah. Ak existuje silný vzťah medzi dvoma premennými, jednu premennú môžeme využiť na predikciu druhej premennej. To znamená, že korelačná analýza býva často využívaná pri vývoji regresných modelov, ktoré sa využívajú na predikčné modelovanie. Analýzou vieme vylúčiť nadbytočné premenné, ktoré by mohli viesť k problémom s multikolinearitou v regresných modeloch. Týmto spôsobom ich vieme odhaliť ešte pred zahájením modelovania a vylúčiť ich z nášho datasetu.

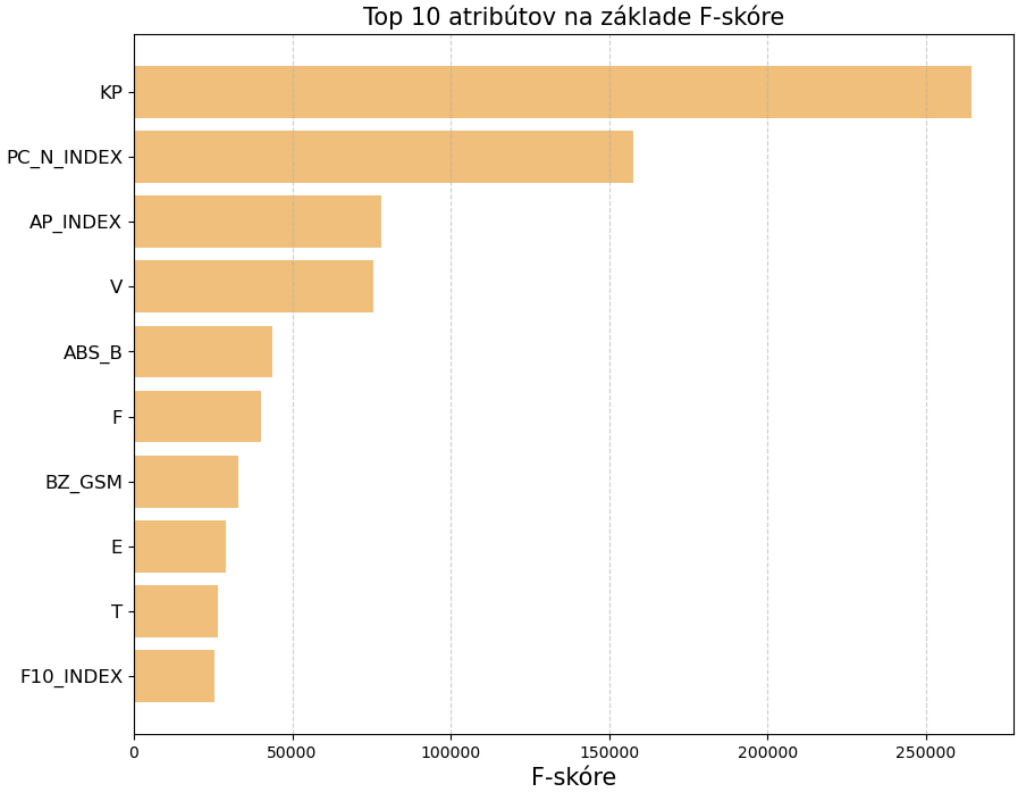
V našom prípade sme najsilnejšiu pozitívnu koreláciu zaznamenali u atribútu `BZ_GSM` na hodnote 0.29, najsilnejšiu negatívnu koreláciu vidíme pri atribúte `KP_index` na hodnote 0.58. Mapu korelácií s cieľovým atribútom možno vidieť na Obrázku 3 – 2, kde tmavšie teplé farby zobrazujú silnejšiu pozitívnu koreláciu a tmavšie studené naopak negatívnu.



Obrázok 3 – 2 Grafy zobrazujúci korelácie medzi atribútmi a DST

Pokračovali sme analýzou výberu najdôležitejších atribútov použitím metódy SelectKBest s F-regresiou. Cieľom bolo zaistiť využitie najvhodnejších atribútov pre naše modelovanie. To znamená, že aj táto metóda nám pomôže zúžiť výber atribútov na minimum, čím dokážeme zlepšiť výkon našej neurónovej siete a zabezpečiť, aby nedošlo k pretrénovaniu. SelectKBest nám vráti k počet najvhodnejších atribútov, ktoré dosiahnu najvyššie F-skóre, v našom prípade to znamená 10 najvhodnejších atribútov. Pre danú analýzu vieme využiť viacero typov funkcii na meranie

výkonnosti, my si vyberáme F-regresiu, nakoľko sa naša práca zaoberá predikciou. Výsledky tejto analýzy sú reprezentované na Obrázku 3 – 3.



Obrázok 3 – 3 Graf zobrazujúci 10 najvhodnejších atribútov na predpoveď DST

Z daných analýz sme došli k výsledku, že na trénovanie využijeme nasledovné atribúty, nakoľko nám vyšla najvyššia korelácia DST s `BZ_GSM` a `KP_index`. F-skóre nám taktiež potvrdilo, že dané atribúty budú vhodné pre trénovanie.

time1

Daný atribút sa nenachádza v analýzach, pretože sa jedná o časovú stopu. Počet záznamov v datasete znamená počet hodín, pre ktoré sú vytvorené meriania jednotlivých atribútov. Náš dataset je časovo štruktúrovaný tak, že prvý záznam začína

1. januárom 1963 o 00:00:00 v koordinovanom svetovom čase (UTC) a následne sa pridáva nový záznam každú hodinu.

DST

Ako sme už spomínali v Kapitole 2, DST index je hlavným indikátorom geomagnetických búrok, ktorý meria globálny stav zemského magnetického poľa v reálnom čase.

BZ_GSM

Zložka Bz poľa GSM je jednou zo súradníc magnetického poľa v globálnom slnečnom magnetickom súradnicovom systéme. Bz zložka tohto poľa ukazuje primeranú koreláciu s pozorovanými poliami blízko Zeme. Denná variácia tejto zložky je významná a preukázateľne korelačná s indexmi `KP_index` a `DST` (Jackson et al., 2018). Zberom dát sa zaoberá vesmírna sonda Advanced Composition Explorer (ACE). Os rotácie vesmírnej lode je orientovaná smerom k Slnku s presnosťou +/- 20 stupňov a nachádza sa v halo orbite okolo libračného bodu L1 medzi Zemou a Slnkom.

KP index

KP index index je dôležitý geomagnetický index monitorujúci subaurorálne geomagnetické poruchy na globálnej úrovni. Odvodený je z trojhodinových K indexov zo špecifických observatórií a poskytuje priemernú hodnotu týchto indexov z 13 prispievajúcich staníc. Používa sa na parametrizáciu empirických modelov geospace a koreluje s rôznymi parametrami v blízkosti Zeme. Je dôležitý pre monitorovanie vesmírneho počasia, poskytuje varovania v reálnom čase pred možnými vplyvmi a je k dispozícii pre vedecký výskum (Matzka et al., 2021).

Nakoľko je naším cieľom predikovať hodnotu v čase dopredu, do nášho datasetu sme museli pridať budúce hodnoty predikovaného atribútu `DST`. To sme docielili odvodením, kde sme jednotlivé hodnoty posunuli o žiadaný čas dopredu - 1 hodina.

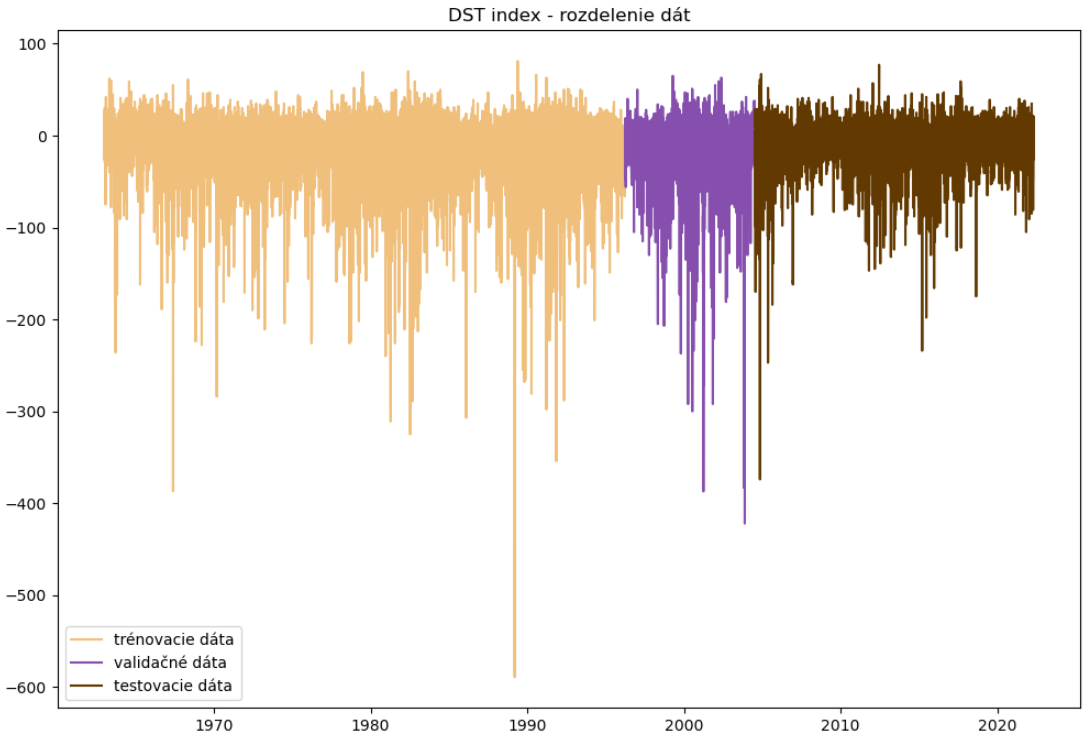
Nové atribúty sme nazvali `DST+1`, `DST+2`, `DST+3`, `DST+4`, `DST+5` a `DST+6`. To znamená, že `DST+1` obsahuje rovnakú hodnotu, ako `DST` v čase *n*, zatiaľ čo `DST+2` obsahuje hodnotu DST v čase *n+1*. Pre uľahčenie si pomôžeme Obrázkom 3 – 4, kde môžeme lepšie vidieť tento vzťah. Napríklad `DST` v prvom zázname (1.1.1963 01:00:00) je -5, zatiaľ čo `DST+2` je tiež -5, napriek tomu, že v treťom zázname (1.1.1963 03:00:00) vidíme hodnotu -3. Je to kvôli neskoršiemu využitiu Timeseries generátora, ktorý nám atribút na predikovanie posunie o jednu hodinu (záznam) dopredu. Vo finálnej podobe bude `DST+2` reprezentovať skutočnosť, pretože vďaka posunu, ktorý vytvorí generátor pre `DST+2` bude reprezentovať posun o 2 hodiny dopredu.

time1	DST	DST+1	DST+2	DST+3	DST+4	DST+5	DST+6
1963-01-01 01:00:00	-5.0	-5.0	-5.0	-3.0	-3.0	-6.0	-8.0
1963-01-01 02:00:00	-5.0	-5.0	-3.0	-3.0	-6.0	-8.0	-9.0
1963-01-01 03:00:00	-3.0	-3.0	-3.0	-6.0	-8.0	-9.0	-6.0
1963-01-01 04:00:00	-3.0	-3.0	-6.0	-8.0	-9.0	-6.0	-2.0
1963-01-01 05:00:00	-6.0	-6.0	-8.0	-9.0	-6.0	-2.0	-2.0
1963-01-01 06:00:00	-8.0	-8.0	-9.0	-6.0	-2.0	-2.0	-5.0
1963-01-01 07:00:00	-9.0	-9.0	-6.0	-2.0	-2.0	-5.0	-6.0

Obrázok 3 – 4 Vytvorenie nových atribútov

Následne je nutné dataset rozdeliť na trénovaciu a testovaciu sadu. Na toto rozdelenie sme zvolili dátum 19.11.2003, nakoľko sme chceli urobiť predikciu pre geomagnetickú búrku, ktorá nastala 20.-21.11.2003. Výsledný graf našej predikcie sa nachádza v časti vyhodnotenie. Tento dátum sme si zvolili pre porovnanie s článkom *Multi-Hour-Ahead Dst Index Prediction Using Multi-FidelityBoosted Neural Networks*, ktorý sa zaoberá danou doménou Hu et al. (2023). To znamená, že v trénovacej sade sa nachádzajú všetky záznamy od roku 1963 po zvolený dátum a v testovacej všetky zvyšné od 19.11.2003 00:00:00. V testovacej sade sme ponechali iba

atribúty potrebné na testovanie a to: `time1`, `BZ_GSM`, `KP_index`, `DST`, `DST+1`, `DST+2`, `DST+3`, `DST+4`, `DST+5` a `DST+6`. Týmto spôsobom sme si vytvorili 2 nové datasety, `omni_train.csv` s dĺžkou 358 407 záznamov a `omni_test.csv` s dĺžkou 162 036 záznamov. Atribúty trénovacej sady si vyberáme už pri samotnom modelovaní, kde vyberáme len tie podstatné pre náš model. Následne 20% z trénovacej sady využijeme na vytvorenie validačnej sady. Na Obrázku 3 – 5 môžeme vidieť rozdelenie hodnôt `DST` indexu na trénovaciu množinu - žltá, validačnú - fialová a testovaciu - hnedá.



Obrázok 3 – 5 Rozdelenie datasetu

3.2 Použité metriky

V tejto časti si popíšeme metriky, ktoré sme využili pri vyhodnocovaní našich modelov. Tieto metriky vyjadrujú úspešnosť modelov. Udávajú sa v percentách alebo v rozmedzí $\langle 0,1 \rangle$. Ich využitie nájdeme v celej rade strojového učenia. V našom

rozdeľovaní do dvoch tried využívame 4 nasledovné hodnoty:

- **TP** - Pravdivo pozitívne (angl. *True Positive*) - Počet správnych predpovedí triedy 1. Príklad: Spamový email je označený za spamový.
- **FP** - Falošne pozitívne (angl. *False Positive*) - Počet nesprávnych predpovedí triedy 1. Príklad: Nespamový email je označený za spamový.
- **TN** - Pravdivo negatívne (angl. *True Negative*) - Počet správnych predpovedí triedy 0. Príklad: Nespamový email je označený za nespamový.
- **FN** - Falošne negatívne (angl. *False Negative*) - Počet nesprávnych predpovedí triedy 0. Príklad: Spamový email je označený za nespamový.

Presnosť

Presnosť (angl. *Precision*) meria, aký podiel pozitívnych alebo negativných predpovedí bol správne určený. Vzorce na výpočet presností v našich triedach sú:

$$\text{presnosť}_0 = \frac{\text{TN}}{\text{TN} + \text{FN}} \quad \text{presnosť}_1 = \frac{\text{TP}}{\text{TP} + \text{FP}} \tag{3.1}$$

Návratnosť

Návratnosť (angl. *Recall*) je podiel správne určených pozitívnych prípadov a všetkých pozitívnych prípadov alebo naopak podiel správne určených negatívnych prípadov a všetkých negatívnych prípadov. Vypočítame ju vzorcami:

$$\text{návratnosť}_0 = \frac{\text{TN}}{\text{TN} + \text{FP}} \quad \text{návratnosť}_1 = \frac{\text{TP}}{\text{TP} + \text{FN}} \tag{3.2}$$

F1-skóre

F1-skóre je harmonický priemer presnosti a návratnosti. Je užitočný, keď je potrebná rovnováha medzi presnosťou a návratnosťou. Vzorec na výpočet je:

$$F_1 \text{ skóre} = 2 * \frac{\text{presnosť} * \text{návratnosť}}{\text{presnosť} + \text{návratnosť}} \tag{3.3}$$

Úspešnosť

Úspešnosť (angl. *Accuracy*) meria podiel správnych predpovedí k celkovému počtu predpovedí. Je najbežnejšou metrikou, avšak môže byť zavádzajúca pri nevyvážených triedach. Vzorec, ktorým ju môžeme vypočítať je:

$$\text{úspešnosť} = \frac{\text{TP} + \text{TN}}{\text{TP} + \text{TN} + \text{FP} + \text{FN}} \tag{3.4}$$

Matthewsov korelačný koeficient

Matthewsov korelačný koeficient (MCC) poskytuje hodnotu v rozmedzí od -1 do 1, kde 1 znamená perfektnú predpoveď, 0 znamená náhodnú predpoveď a -1 znamená úplne chybnú predpoveď. MCC je obzvlášť užitočný pri nevyvážených triedach, pretože berie do úvahy všetky hodnoty kontingenčnej matice (TP, TN, FP, FN). Vypočítame ho za pomoci vzorca:

$$\text{MCC} = \frac{(TP \times TN) - (FP \times FN)}{\sqrt{(TP + FP) \times (TP + FN) \times (TN + FP) \times (TN + FN)}} \tag{3.5}$$

3.3 Modelovanie

V tejto časti bakalárskej práce si predstavíme modely a postupy, ktoré sme využili. Naším cieľom je vytvoriť model, ktorý bude predpovedať 1 hodnotu DST o hodinu až šesť dopredu, na základe 6 vstupných hodnôt. Daná veľkosť vstupného okna bola zvolená pri predchádzajúcom výskume.

Timeseries Generátor

Timeseries Generátor je trieda knižnice Keras, ktorá nám umožňuje generovať jednotlivé úseky (batch) sekvenčných údajov na trénovanie, validáciu a testovanie re-

kurentných neurónových sietí. V našom prípade sme zvolili batch o veľkosti 256 a vstupnú sekvenciu 6 časových krokov. To znamená, že ak máme časovú sériu $[\mathbf{x_1}, \mathbf{x_2}, \mathbf{x_3}, \mathbf{x_4}, \mathbf{x_5}, \mathbf{x_6}, \mathbf{x_7}, \mathbf{x_8}]$ a dĺžku 6, generátor vytvorí vzorky:

- Prvá vstupná sekvencia: $[\mathbf{x_1}, \mathbf{x_2}, \mathbf{x_3}, \mathbf{x_4}, \mathbf{x_5}, \mathbf{x_6}]$ s cieľovou hodnotou $\mathbf{x_7}$
- Druhá vstupná sekvencia: $[\mathbf{x_2}, \mathbf{x_3}, \mathbf{x_4}, \mathbf{x_5}, \mathbf{x_6}, \mathbf{x_7}]$ s cieľovou hodnotou $\mathbf{x_8}$

3.3.1 RNN s využitím LSTM a Bidirectional LSTM vrstvy

Náš prvý model sa skladá z nasledovných vrstiev:

- **Bidirectional LSTM vrstva** Dvojsmerná LSTM vrstva s 128 jednotkami, ktorá vracia sekvencie a používa dropout a recurrent dropout 0.1,
- **LSTM vrstva** Obyčajná LSTM vrstva s 128 jednotkami, ktorá vracia sekvencie,
- **TimeDistributed Dense vrstva** Časovo distribuovaná hustá vrstva s jedným neurónom a lineárnou aktiváciou, aplikovaná na každú časť časovej sekvencie,
- **Flatten vrstva** Vrstva, ktorá splošťuje vstupy do jednorozmerného vektora,
- **Dense vrstva** Hustá vrstva s jedným neurónom a lineárnou aktiváciou, ktorá vracia finálny výstup.

Na kompiláciu využívame ako funkciu straty funkciu Mean Squared Error (MSE), optimalizátor Adam a metriku Mean Absolute Error (MAE). To znamená, že MSE penalizuje väčšie chyby viac, čím sa model snaží minimalizovať veľké odchýlky. Adam optimizátor adaptívne upravuje učebnú rýchlosť pre každú váhu a pomáha modelu konvergovať rýchlo a efektívne. MAE sleduje výkonnosť modelu, poskytuje jednoduchý a interpretovateľný spôsob, ako hodnotiť predikcie modelu z hľadiska priemerných absolútnych chýb. Následne ukladáme vždy najlepší model na základe va-

lidačného MAE. Okrem dropoutu, na zabránenie preučenia využívame metódu EarlyStopping, ktorá ukončí tréning, ak sa MAE nezlepší po 25 epochách. Sumarizáciu modelu môžeme vidieť na Obrázku 3 – 6.

```
_________________________________________________________________
 Layer (type)                 Output Shape              Param #
=================================================================
 input_1 (InputLayer)         [(None, 6, 1)]            0

 bidirectional (Bidirectiona  (None, 6, 256)           133120
 l)

 lstm_1 (LSTM)                (None, 6, 128)            197120

 time_distributed (TimeDistr  (None, 6, 1)             129
 ibuted)

 flatten (Flatten)            (None, 6)                 0

 dense_1 (Dense)              (None, 1)                 7

=================================================================
Total params: 330,376
Trainable params: 330,376
Non-trainable params: 0
```

Obrázok 3 – 6 Sumarizácia modelu

Autoregresný model na predikciu DST indexu iba za pomoci predchádzajúcich DST hodnôt

Ako prvé sme urobili predikcie iba za použitia atribútu DST, kde sme použili jeho predchádzajúce hodnoty o počte 6 záznamov. Model bol vytvorený na predikciu 1 až 6 hodín dopredu. Vľavo na Obrázku 3 – 7 môžeme vidieť priebeh metrík presnosti, návratnosti a F1-skóre, ktoré sme používali pri vyhodnocovaní výsledkov modelu. Na Obrázku 3 – 7 vpravo vidíme priebeh metrík ACC a MCC. Na osi x ležia body časovej predpovede a na osi y môžeme vidieť ich príslušné hodnoty. Hraničná hodnota,

ktorá určuje, či búrka nastala alebo nie, je pri všetkých modeloch rôzna. Na nájdenie optimálnej hranice sme použili funkciu, ktorá nám vráti najvhodnejší threshold z validačnej sady, ktorý následne použijeme na testovacej sade.

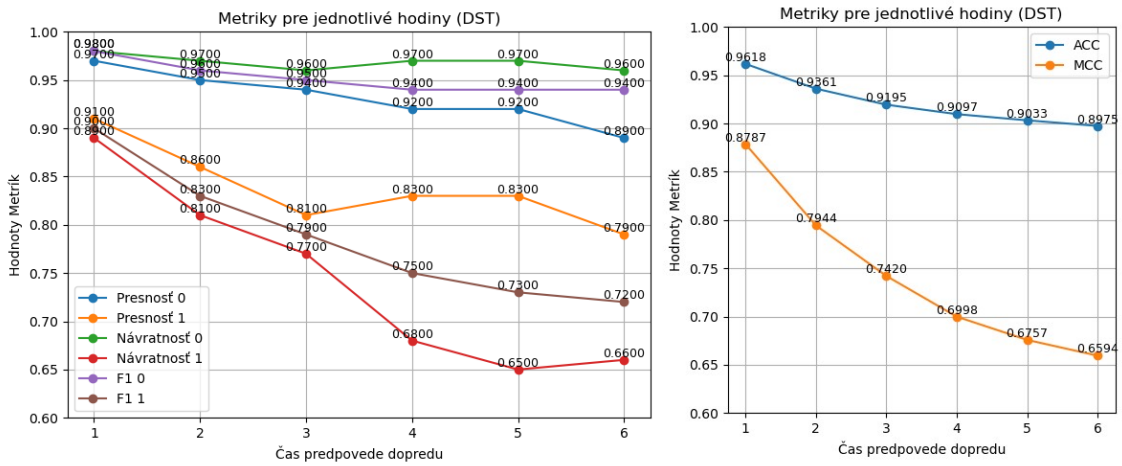
Z ľavého Obrázka 3 – 7 taktiež môžeme vyčítať, že modely udržujú dobrú presnosť a návratnosť pre triedu 0 aj na dlhšie časové horizonty. To znamená, že modely vedia s vysokou pravdepodobnosťou určiť, že búrka nenastane. Na druhú stranu, pri triede 1, môžeme vidieť, že predpovede na dlhší čas sa zhoršujú, čo naznačuje, že modely majú problém s určením skutočných pozitívnych prípadov. Podrobnejšie výsledky klasifikácie daného modelu na predpoveď 1 hodiny dopredu môžeme vidieť v kontingenčnej Tabuľke 3.3.1 a v klasifikačnom reporte v Tabuľke 3.3.1.

	predikované hodnoty	
skutočné hodnoty	0	1
0	127 264	2 721
1	3 462	28 583

Tabuľka 3 – 1 Konfúzna matica pre klasifikačný model

	presnosť	návratnosť	F1-skóre	počet
0	0.97	0.98	0.98	129 985
1	0.91	0.89	0.90	32 045
úspešnosť			0.96	162 030
macro avg	0.94	0.94	0.94	162 030
weighted avg	0.96	0.96	0.96	162 030

Tabuľka 3 – 2 Výsledky klasifikačného modelu



Obrázok 3 – 7 Vyhodnotenie zvolených metrík pre atribút DST

Model na predikciu DST indexu s pridaným atribútom BZ_GSM

Náš ďalší krok pozostáva z pridania ďalšieho atribútu do procesu vytvárania modelov. Aj pri tomto modeli sme zvolili rovnaké metriky na jeho vyhodnotenie. Vľavo na Obrázku 3 – 8 môžeme vidieť, že hodnoty pre triedu 0 zostávajú viac-menej rovnaké a malá zmena nastala pri triede 1. Môžeme si všimnúť, že presnosť pre model na predpoveď indexu DST o 3 hodiny dopredu sa zvýšila o 0.05, avšak návratnosť nám klesla. Opačný jav môžeme pozorovať pri modeli, ktorý predpovedá hodnoty na 6 hodín dopredu, kde nám klesla presnosť, no na druhú stranu nám narástla návratnosť. Na Obrázku 3 – 8 vpravo môžeme vidieť, že presnosť modelu sa zvýšila len veľmi minimálne. To isté vieme povedať aj o metrike MCC. Podrobnejšie výsledky klasifikácie daného modelu na predpoveď 1 hodiny dopredu môžeme vidieť v kontingenčnej Tabuľke 3.3.1 a v klasifikačnom reporte v Tabuľke 3.3.1. Môžeme si všimnúť, že klasifikačný report má lepší výsledok pre takmer všetky metriky oproti modelu s použitím atribútu DST.

Model na predikciu DST indexu s pridaným atribútom KP_index

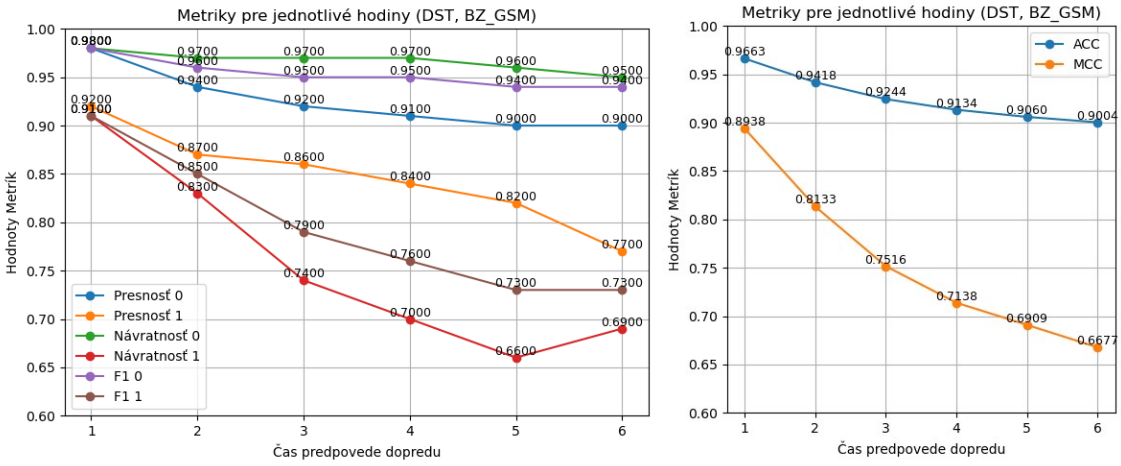
V ďalšom modeli sme vyskúšali nahradiť atribút `BZ_GSM` atribútom `KP_index`. Tu sme zaznamenali opäť zmenu pri triede 1, zatiaľ čo trieda 0 zostávala s približne

skutočné hodnoty	predikované hodnoty 0	1
0	127 292	2 693
1	2 760	29 285

Tabuľka 3 – 3 Konfúzna matica pre klasifikačný model

	presnosť	návratnosť	F1-skóre	počet
0	0.98	0.98	0.98	129 985
1	0.92	0.91	0.91	32 045
úspešnosť			0.97	162 030
macro avg	0.95	0.95	0.95	162 030
weighted avg	0.97	0.97	0.97	162 030

Tabuľka 3 – 4 Výsledky klasifikačného modelu



Obrázok 3 – 8 Vyhodnotenie zvolených metrík pre atribút DST a BZ_GSM

rovnakými výsledkami. Narozdiel od modelu s pridaným atribútom `BZ_GSM`, sme si všimli, že presnosť triedy 1 klesla takmer pri všetkých časových radoch až na predikciu 6 hodín dopredu, kde sme zaznamenali menší nárast. Na druhú stranu nám však stúpla návratnosť. Rozdiel je vidieť aj v predpovedi 2 hodín dopredu, kde nám

návratnosť prekonala hodnotu presnosti aj F1-skóre. Očakávali sme, že tento model prekoná model s atribútom `BZ_GSM`, nakoľko nám korelačná a F-skóre analýza ukázala, že daný atribút je vhodný na zaradenie do trénovania. Ako však môžeme vidieť na na Obrázku 3 – 9 vľavo, naše očakávania sa nenaplnili.

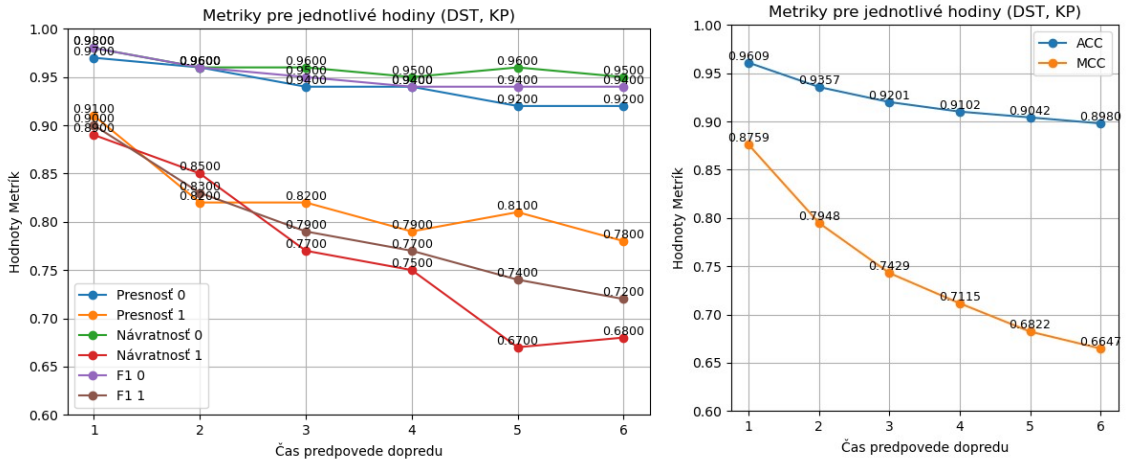
Metriky pre ACC a MCC sa nachádzajú medzi úrovňou modelu s atribútom DST a modelom s pridaným atribútom `BZ_GSM`. Vpravo na Obrázku 3 – 9 môžeme vidieť, že sme nezaznamenali žiadne výraznejšie zmeny. Podrobnejšie výsledky klasifikácie daného modelu na predpoveď 1 hodiny dopredu môžeme vidieť v kontingenčnej Tabuľke 3.3.1 a v klasifikačnom reporte v Tabuľke 3.3.1. V tabuľkách môžeme vidieť takmer totožné výsledky s modelom, ktorý využíva iba atribút `DST`.

	predikované hodnoty	
skutočné hodnoty	0	1
0	127 219	2 766
1	3 560	28 485

Tabuľka 3 – 5 Konfúzna matica pre klasifikačný model

	presnosť	návratnosť	F1-skóre	počet
0	0.97	0.98	0.98	129 985
1	0.91	0.89	0.90	32 045
úspešnosť			0.96	162 030
macro avg	0.94	0.93	0.94	162 030
weighted avg	0.96	0.96	0.96	162 030

Tabuľka 3 – 6 Výsledky klasifikačného modelu



Obrázok 3 – 9 Vyhodnotenie zvolených metrík pre atribút DST a KP_index

Model na predikciu DST indexu s pridanými atribútmi BZ_GSM a KP_index

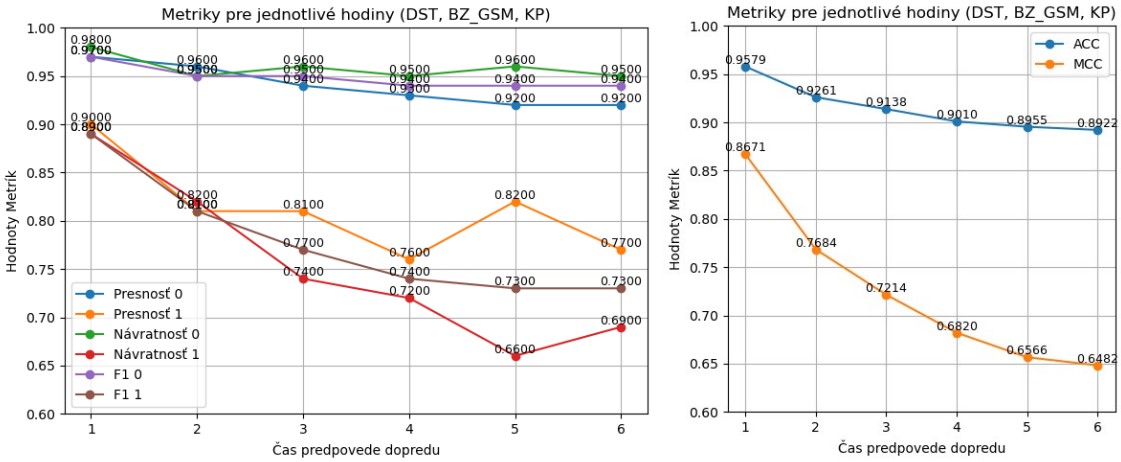
Ako poslednú kombináciu sme zvolili pridať atribút `BZ_GSM` spoločne s atribútom `KP_index`. O triede 0 opätovne môžeme povedať, že žiadne výrazné zmeny nenastali, avšak pri triede 1 sme pre presnosť zaznamenali najhoršie výsledky. Zaujímavé však je, že presnosť modelu na predikciu 5 hodín dopredu prekonala presnosť 3 a 4 hodinového modelu. Návratnosť zostáva približne rovnaká ako pre model, kde sme použili atribúty `DST` a `KP_index`. Vývoj týchto metrík môžeme vidieť vľavo na Obrázku 3 – 10. Metriky ACC a MCC sú pre tento model na najnižšej úrovni. Tento fakt môžeme vidieť na Obrázku 3 – 10 vpravo. Podrobnejšie výsledky klasifikácie daného modelu na predpoveď 1 hodiny dopredu môžeme vidieť v kontingenčnej Tabuľke 3.3.1 a v klasifikačnom reporte v Tabuľke 3.3.1. Podľa týchto tabuliek by sme daný model mohli označiť za najslabší.

	predikované hodnoty	
skutočné hodnoty	0	1
0	126 739	3 246
1	3 562	28 483

Tabuľka 3 – 7 Konfúzna matica pre klasifikačný model

	presnosť	návratnosť	F1-skóre	počet
0	0.97	0.98	0.97	129 985
1	0.90	0.89	0.89	32 045
úspešnosť			0.96	162 030
macro avg	0.94	0.93	0.93	162 030
weighted avg	0.96	0.96	0.96	162 030

Tabuľka 3 – 8 Výsledky klasifikačného modelu



Obrázok 3 – 10 Vyhodnotenie zvolených metrík pre atribút DST, BZ_GSM a KP_index

3.3.2 Využitie architektúry transformer na spracovanie sekvencií

Náš druhý model sa skladá z troch vrstiev - vstupná vrstva, kóder a výstupná vrstva. Kóder obsahuje:

- **Vstupné projekcie** Každý vstupný bod je premietnutý na vyšší rozmerový priestor (embed_dim=64), aby sa zvýšila kapacita modelu na zachytávanie zložitejších vzorov v dátach,
- **MultiHead Attention** Umožňuje paralelné spracovanie všetkých pozícií vstup-

nej sekvencie a efektívne modelovanie vzájomných závislostí, vďaka čomu transformer rýchlo a efektívne zvláda aj dlhé závislosti v dátach,

- **Feedforward sieť** Plne prepojená sieť, ktorá zahŕňa dve husté vrstvy s relu aktiváciou medzi nimi. Prvá vrstva expanduje dimenzie na ff_dim=256, čím zvyšuje modelovú kapacitu, zatiaľ čo druhá vrstva redukuje dimenzie späť na embed_dim.,
- **Dropout a Layer Normalization** Na predídenie pretrénovania model využíva dropout 0.1 a normalizáciu

Výstupná vrstva sa skladá z vrstvy **Flatten** a **Dense**, ako tomu bolo aj pri našom prvom modeli. Kompilácia zostáva rovnaká, využívame optimalizátor Adam, funkciu straty MSE a metriku MAE. Opätovne sme použili funkciu EarlyStopping, ktorá ukončí trénovanie, ak sa MAE nezlepší po 25 epochách a ModelChecKP_indexoint na uloženie najlepšieho modelu.

Model na predikciu DST indexu iba za pomoci predchádzajúcich DST hodnôt

V tejto časti sme si vybrali kvôli rozsahu práce iba 1 model. Zvyšné modely a ich metriky môžete nájsť v priečinku `Metriky_Modelov`, ktorý je dostupný online[3]. V priečinku sa nachádza aj súbor `vysledky_modelov.xlsx`, ktorý obsahuje metriky vo forme tabuliek.

Prvý rozdiel, ktorý vidíme pri metrikách daného modelu na Obrázku 3 – 11 vľavo, je ten, že metriky nám postupne klesajú, zatiaľ čo pri využití modelu s LSTM sme mohli sledovať, že pri niektorých vzdialenejších predpovediach hodnoty prekonali tie pri kratších predpovediach. Trieda 0 zostáva približne rovnaká ako pri ostatných modeloch, opäť vidíme väčšiu dynamiku pri triede 1. Presnosť triedy 1 nám mierne klesla pri predpovedi 4 a 5 hodín dopredu, avšak pri 3 a 6 hodinách sme zaznamenali

[3]Metriky modelov: `https://github.com/terezbiak/BP_Drengubiakova_Terezia/tree/main/modelovanie/Metriky_Modelov`

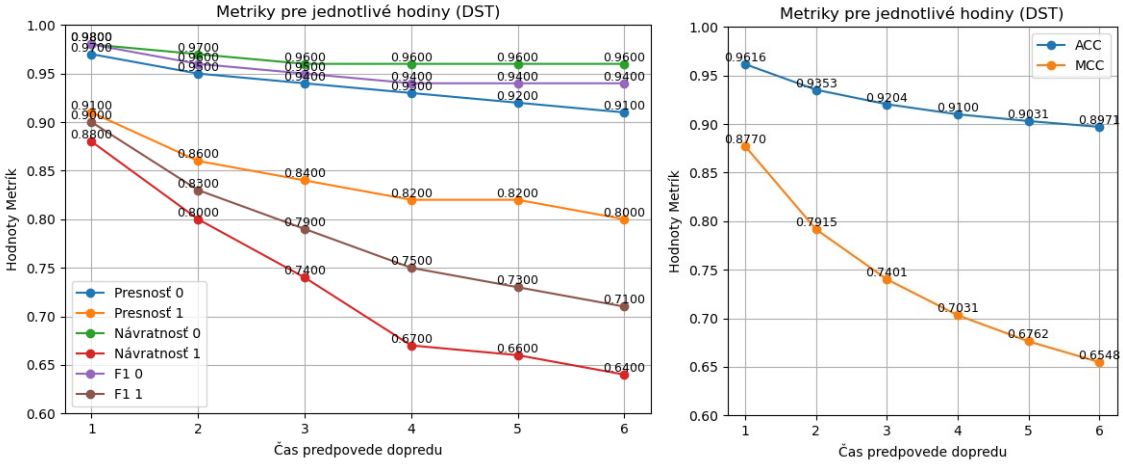
malé zlepšenie. Návratnosť je na tom približne rovnako, zlepšenie zachytávame iba pri predikcii 4 a 5 hodín dopredu. F1-skóre zostáva identické až na predikciu 6 hodín dopredu, kde vidíme pokles o 0,01. Pri metrikách ACC a MCC pozorujeme minimálny pokles. Ich priebeh možno vidieť na Obrázku 3 – 11 vpravo. Podrobnejšie výsledky klasifikácie daného modelu na predpoveď 1 hodiny dopredu môžeme vidieť v kontingenčnej Tabuľke 3.3.2 a v klasifikačnom reporte v Tabuľke 3.3.2.

	predikované hodnoty	
skutočné hodnoty	0	1
0	127 712	2 273
1	3 947	28 098

Tabuľka 3 – 9 Konfúzna matica pre klasifikačný model

	presnosť	návratnosť	F1-skóre	počet
0	0.97	0.98	0.98	129 985
1	0.93	0.88	0.90	32 045
úspešnosť			0.96	162 030
macro avg	0.95	0.93	0.94	162 030
weighted avg	0.96	0.96	0.96	162 030

Tabuľka 3 – 10 Výsledky klasifikačného modelu



Obrázok 3 – 11 Vyhodnotenie zvolených metrík pre atribút DST

3.4 Vyhodnotenie

V tejto časti práce si predstavíme výsledky práce. `Model 1` predstavuje RNN model s použitím atribútu `DST`, `Model 2` predstavuje Transformačný model s použitím atribútu DST. V Tabuľke 3.4 môžeme vidieť porovnanie týchto modelov. Mô-

metrika		Model 1	Model 2
úspešnosť		0.96	0.96
presnosť	0	0.97	0.97
	1	0.91	0.93
návratnosť	0	0.98	0.98
	1	0.89	0.88
F1-skóre	0	0.98	0.98
	1	0.90	0.90
MCC		0.88	0.88

Tabuľka 3 – 11 Porovnanie metrík dvoch modelov

žeme si všimnúť, že medzi jednotlivými modelmi nenastali žiadne zásadné rozdiely. `Model 2` síce vykazuje vyššiu presnosť pre triedu 1 o 2%, avšak návratnosť nám klesla o 1%. Na druhú stranu, na Obrázku 3 – 12 to vyzerá, že `Model 2` dosahuje trošku

lepšie predikcie. Na potvrdenie tohto tvrdenia by bolo potrebné preskúmať viacero udalostí (záznamy o geomagnetických búrkach) na testovacej sade. Zaujímavý je pre nás vývin MSE metriky, ktorá s rastúcim časom predikcie klesá. Ako budúcu prácu by sme odporúčali preskúmať daný jav.

Naše výsledky môžeme zhrnúť nasledovne (výsledky sú na škále od `DST+1` po `DST+6`):

LSTM Bidirectional Model

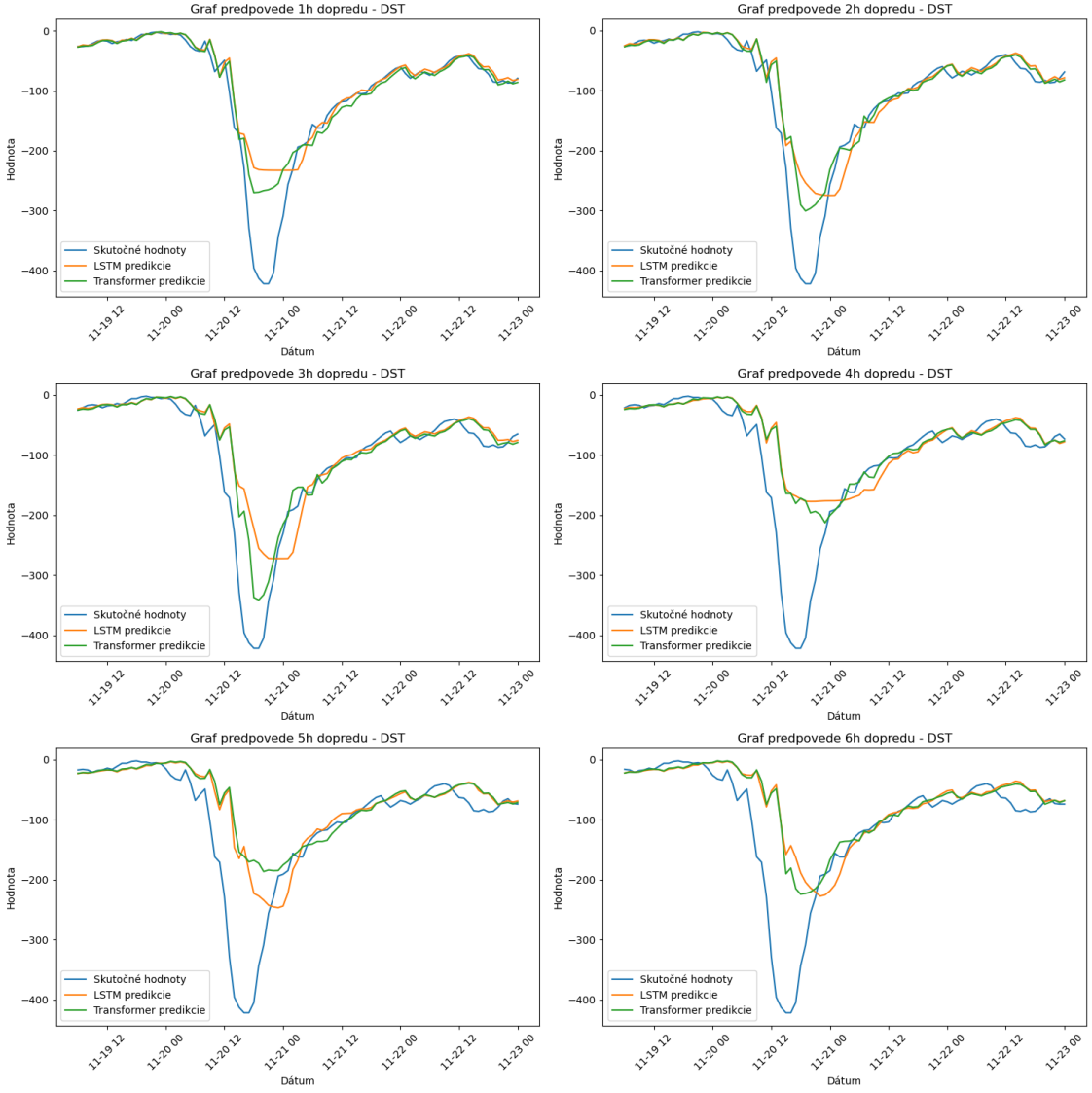
- **MSE** nachádza sa v rozmedzí od 98.73 do 53.52
- **Úspešnosť** klesá od 96% po 89%
- **Presnosť**
  - Trieda 0 stabilne vysoká od 97% po 92%
  - Trieda 1 klesá od 91% po 79%
- **Návratnosť**
  - Trieda 0 stabilne vysoká od 98% po 96%
  - Trieda 1 klesá od 89% po 66%

Transformer Model

- **MSE** nachádza sa v rozmedzí od 102.38 do 17.94
- **Úspešnosť** klesá od 96% po 90%
- **Presnosť**
  - Trieda 0: je stabilne vysoká od 97% po 91%
  - Trieda 1: klesá od 91% po 80%
- **Návratnosť**

– Trieda 0 je stabilne vysoká od 98%po 96%

– Trieda 1 klesá od 88% po 64%



Obrázok 3 – 12 Predikcie modelov v porovnaní so skutočnými hodnotami

4 Záver

V bakalárskej práci sme skúmali vznik a dopady geomagnetických búrok a ich možnými predpoveďami. Na predikcie sme využili metódy hlbokého učenia, presnejšie rekurentnú neurónovú sieť a model transformer.

Teoretická časť práce obsahuje základy strojového a hlbokého učenia, ktoré sú potrebné pre porozumenie praktickej časti. Obsahuje taktiež popis modelov, ktoré sme v tejto práci používali. Doménová časť sa venuje porozumeniu geomagnetických búrok a otázke, prečo je dôležitá ich predikcia. Taktiež sa v nej dozvedáme o súčasnom stave riešenia danej problematiky. Praktická časť popisuje fázy pochopenia cieľa, získania dát, ich prípravu a analýzu, modelovania a vyhodnotenia.

Naším cieľom bolo navrhnúť model, ktorý bude schopný predikovať geomagnetické búrky s časovým predstihom 1 až 6 hodín. Na predpoveď sme použili časové okno, ktoré zahrňuje údaje z posledných 6 hodín. Vytvorili sme a natrénovali dva modely: rekurentnú neurónovú sieť s LSTM a Bidirectional LSTM vrstvou, a model transformer. Vo fáze vyhodnotenia sme sa zamerali na modely využívajúce atribút `DST` pre predikciu indexu DST o 1 hodinu dopredu. Oba modely dosiahli úspešnosť na úrovni 96% a Matthewsov korelačný koeficient 88%. Rekurentná neurónová sieť vykázala vyššiu návratnosť pri klasifikácii výskytu geomagnetickej búrky, dosahujúc 89%, čo je o 1% viac ako model transformer. Na druhej strane, model transformer preukázal lepšiu presnosť pre túto triedu, dosahujúc 93%, čo je o 2% viac ako rekurentná neurónová sieť. V práci sa však venujeme aj modelom, ktoré využívajú pridané atribúty `KP_index` a `BZ_GSM`. Tu sme očakávali výraznejšie zlepšenie výsledkov, nakoľko nám dané atribúty vyšli ako vhodné na zaradenie do trénovania podľa korelačnej a F-skóre analýzy. Zistili sme, že dané atribúty nepriniesli očakávané zlepšenie. Príčinou môžu byť chýbajúce hodnoty atribútu `BZ_GSM`, avšak `KP_index` chýbajúce hodnoty neobsahuje. Naším odporúčaním je bližšie preskúmať túto problematiku, napríklad metódou vysvetliteľnosti, v budúcich prácach.

Taktiež sme zistili, že presnosť predikcie sa znižuje so zvyšujúcim sa časovým

predstihom, čo je prirodzené vzhľadom na komplexnosť a dynamiku geomagnetických búrok. Napriek tomu dosiahnuté výsledky preukazujú potenciál metód hlbokého učenia pri predikcii týchto fenoménov, čo má významné praktické dopady na ochranu technologických systémov a infraštruktúry.

Budúci výskum by sa mohol zamerať na optimalizáciu modelov a zahrnutie ďalších relevantných atribútov, ako aj na zlepšenie výpočtovej efektívnosti, čo by umožnilo ešte presnejšie a rýchlejšie predpovede. Taktiež by sme odporúčali preskúmať vývin metriky MSE, ktorá s rastúcim časom predikcie klesá.

Literatúra

Afifi, S., Sunny, F., Nikdast, M. and Pasricha, S. (2023). Tron: transformer neural network acceleration with non-coherent silicon photonics, *Proceedings of the Great Lakes Symposium on VLSI 2023*, pp. 15–21.

Bi, Q., Goodman, K. E., Kaminsky, J. and Lessler, J. (2019). What is machine learning? a primer for the epidemiologist, *American journal of epidemiology* **188**(12): 2222–2239.

Chollet, F. (2019). *Deep learning v jazyku Python: knihovny Keras, Tensorflow*, Grada Publishing.

Dongare, A., Kharde, R., Kachare, A. D. et al. (2012). Introduction to artificial neural network, *International Journal of Engineering and Innovative Technology (IJEIT)* **2**(1): 189–194.

Emmert-Streib, F., Yang, Z., Feng, H., Tripathi, S. and Dehmer, M. (2020). An introductory review of deep learning for prediction models with big data, *Frontiers in Artificial Intelligence* **3**: 4.

Fetzer, J. H. and Fetzer, J. H. (1990). *What is artificial intelligence?*, Springer.

Gonzalez, W., Joselyn, J.-A., Kamide, Y., Kroehl, H. W., Rostoker, G., Tsurutani, B. T. and Vasyliunas, V. (1994). What is a geomagnetic storm?, *Journal of Geophysical Research: Space Physics* **99**(A4): 5771–5792.

Graves, A. and Graves, A. (2012). Long short-term memory, *Supervised sequence labelling with recurrent neural networks* pp. 37–45.

Gurney, K. (2018). *An introduction to neural networks*, CRC press.

Hu, A., Camporeale, E. and Swiger, B. (2023). Multi-hour-ahead dst index prediction using multi-fidelity boosted neural networks, *Space Weather* **21**(4): e2022SW003286. e2022SW003286 2022SW003286.

Jackson, B., Hick, P., Buffington, A., Yu, H.-S. and Zhao, G. (2018). Bz determinations and forecasts using ucsd analysis techniques, *42nd cospar scientific assembly* **42**: D2–3.

LeCun, Y., Bengio, Y. and Hinton, G. (2015). Deep learning, *nature* **521**(7553): 436–444.

Matzka, J., Stolle, C., Yamazaki, Y., Bronkalla, O. and Morschhauser, A. (2021). The geomagnetic kp index and derived indices of geomagnetic activity, *Space weather* **19**(5): e2020SW002641.

Schwenn, R. (2006). Space weather: The solar perspective, *Living reviews in solar physics* **3**(1): 1–72.

Shang, K., Chen, Z., Liu, Z., Song, L., Zheng, W., Yang, B., Liu, S. and Yin, L. (2021). Haze prediction model using deep recurrent neural network, *Atmosphere* **12**(12): 1625.

Smys, S., Chen, J. I. Z. and Shakya, S. (2020). Survey on neural network architectures with deep learning, *Journal of Soft Computing Paradigm (JSCP)* **2**(03): 186–194.

Staudemeyer, R. C. and Morris, E. R. (2019). Understanding lstm–a tutorial into long short-term memory recurrent neural networks, *arXiv preprint arXiv:1909.09586* .

Sze, V., Chen, Y.-H., Yang, T.-J. and Emer, J. S. (2017). Efficient processing of deep neural networks: A tutorial and survey, *Proceedings of the IEEE* **105**(12): 2295–2329.

Uzair, M. and Jamil, N. (2020). Effects of hidden layers on the efficiency of neural networks, *2020 IEEE 23rd international multitopic conference (INMIC)*, IEEE, pp. 1–6.

Van Houdt, G., Mosquera, C. and Nápoles, G. (2020). A review on the long short-term memory model, *Artificial Intelligence Review* **53**(8): 5929–5955.

Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., Kaiser, Ł. and Polosukhin, I. (2017). Attention is all you need, *Advances in neural information processing systems* **30**.

Zoznam príloh

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

Príloha B Kód je dostupný ako BP_Drengubiakova_Terezia_main.zip alebo online - `https://github.com/terezbiak/BP_Terezia_Drengubiakova/tree/main`