

Technická univerzita v Košiciach
Fakulta elektrotechniky a informatiky
Katedra kybernetiky a umelej inteligencie

Predikcia geomagnetických búrok pomocou hlbokého učenia

Diplomová práca

Systémová príručka

Vedúci diplomovej práce:

doc. Ing. Peter Butka, PhD.

Autor:

Bc. Veronika Motúzová

Konzultanti diplomovej práce:

Ing. Viera Maslej Krešňáková, PhD. a RNDr. Šimon Mackovjak, PhD.

Košice 2023

Obsah

1	Funkcia programu	1
2	Analýza riešenia	1
3	Popis programu	1
3.1	Popis riešenia	2
3.2	Popis algoritmov a údajových štruktúr, globálnych premenných . . .	2
3.3	Zoznam zdrojových textov	4
4	Zhodnotenie riešenia	4

1 Funkcia programu

Program bol vytvorený na predikciu geomagnetických búrok, pomocou metód hlbokého učenia, a to konkrétne rekurentnej neurónovej siete. Predikuje sa atribút DST index, na základe ktorého klasifikujeme stav na geomagnetickú búrku (1) alebo normálny stav(0).

Program bol vyvíjaný v programovacom jazyku **Python**, čo je v súčasnosti jeden z popredných jazykov využívaných v oblasti hlbokého učenia. Program načíta potrebné knižnice na prácu, prijme vstup od používateľa, upraví dáta na požadovaný tvar, načíta natrénovaný model, pomocou ktorého určí hodnotu DST indexu s predstihom 6 hodín, ktorá sa vypíše používateľovi do konzoly.

2 Analýza riešenia

Hlavnou úlohou pri vývoji riešenia bolo navrhnuť RNN sieť, ktorá bude dosahovať čo najvyššiu úspešnosť predikcie. Model RNN bol navrhnutý a optimalizovaný v priebehu troch kôl experimentov.

3 Popis programu

Program je zložený z viacerých častí. Hlavná časť je v súbore typu **jupyter notebook**, nachádza sa v priečinku **6_nasadenie**, dostupné online¹, v súbore **user.ipynb**. Je potrebné spustiť ho v programe podporujúcom tento typ súboru, napr. **JupyterLab**. Zdrojový kód je písaný v blokoch, ktoré na seba logicky naväzujú. Pre správne fungovanie je potrebné spustiť bloky v poradí, v akom sú napísané. Dôležité sú aj súbory **6_6_pridane_atr.hdf5** a **6_dozadu_6_dopredu.hdf5**. Tieto súbory obsahujú natrénovaný model, ktorý sa v súbore **user.ipynb** načíta. Tieto súbory sú uložené v spoločnom priečinku.

¹Nasadenie:https://github.com/motuziik/DP_Veronika_Motuzova/tree/main/6_nasadenie

3.1 Popis riešenia

Celý technický postup návrhu modelov je v priečinku s názvom `4_modelovanie`, dostupné online².

3.2 Popis algoritmov a údajových štruktúr, globálnych premenných

Hlavný skript je tvorený z malých blokov, ktorých funkcionality si teraz popíšeme:

- Blok 1 : importuje všetky potrebné knižnice alebo jednotlivé konkrétne funkcie z nich.

```
from tensorflow import keras
from keras.preprocessing.sequence import TimeseriesGenerator
from keras.models import Model
import numpy as np
```

- Blok 2 : načítanie vstupných dát od používateľa pre prvý model.

```
DST = [-10,13,16,-20,-25,-40,-30]
# používateľ zadá sekvenciu aspoň 7 historických hodnôt DST indexu
```

- Blok 3 : úprava vstupných dát.

```
generator = TimeseriesGenerator(DST, [0,0,0,0,0,0,0], length=6, batch_size=256)
```

- Blok 4 : načítanie natrénovaného predikčného modelu zo súboru.

```
model = keras.models.load_model('6_dozadu_6_dopredu.hdf5')
```

²Modelovanie: https://github.com/motuziik/DP_Veronika_Motuzova/tree/main/4_modelovanie

- Blok 5 : predikcia hodnoty DST indexu s predstihom 6 hodín.

```
DST_6 = model.predict(generator)
```

- Blok 6 : výpis výslednej predikovanej hodnoty.

```
print(DST_6)
```

- Blok 7 : načítanie vstupov od používateľa pre druhý model, spojenie vstupov do jedného poľa.

```
DST = np.array([[ -10], [13], [16], [ -20], [ -25], [ -40], [ -30]])  
# používateľ zadá sekvenciu aspoň 7 historických hodnôt všetkých 3 atribútov  
F10_INDEX = np.array([[177.69], [177.69], [178.39], [178.39], [178.39], [178.39], [178.39]])  
BZ_GSM = np.array([[1.79], [2.29], [1.60], [1.29], [1.89], [3.20], [0.40]])  
  
inputs = np.concatenate((DST, F10_INDEX, BZ_GSM), axis=1)
```

- Blok 8 : kontrolný výpis vstupu.

```
print(inputs)
```

- Blok 9 : úprava vstupu pre potreby modelu.

```
generator = TimeseriesGenerator(inputs, [0,0,0,0,0,0,0], length=6, batch_size=256)
```

- Blok 10 : načítanie natrénovaného predikčného modelu zo súboru.

```
model = keras.models.load_model('6_6_pridane_atr.hdf5')
```

- Blok 11 : predikcia hodnoty DST indexu s predstihom 6 hodín.

```
DST_6 = model.predict(generator)
```

- Blok 12 : výpis výslednej predikovanej hodnoty.

```
print(DST_6)
```

3.3 Zoznam zdrojových textov

Priložené DVD obsahuje súbor `DP_Motuzova.zip`, v ktorom sú všetky zdrojové kódy zatriedené do pod-adresárov jednotlivých fáz. Súbor je dostupný aj online³. V každom pod-adresári sú zdrojové kódy a potrebné dátové súbory pre každú fázu, v `requirements.txt` sú spísané balíky, ktoré potrebujeme nainštalovať pred spustením.

4 Zhodnotenie riešenia

Riešenie danej úlohy bolo splnené. Úspešne sme natrénovali rekurentnú neurónovú sieť, ktorá dokáže detekovať meteor z astronomických snímok. K práci vznikol aj článok s cieľom publikácie v blízkej budúcnosti. V budúcnosti by práca mohla pokračovať vývojom a optimalizáciou iných architektúr rekurentných neurónových sietí, prípadne by sa na predikciu mohla použiť iná kombinácia vstupných atribútov. Užitočná by určite bola aj optimalizácia hyperparametrov siete. Dátová množina, ktorej sa výskum venuje, sa s plynúcim časom neustále zväčšuje, čo bude viesť k zlepšeniu predikcie v budúcich výskumoch. Práca vznikla v spolupráci s Ústavom experimentálnej fyziky Slovenskej akadémie vied, v.v.i. a poslúži na ich výskumné účely.

³https://github.com/motuziik/DP_Veronika_Motuzova