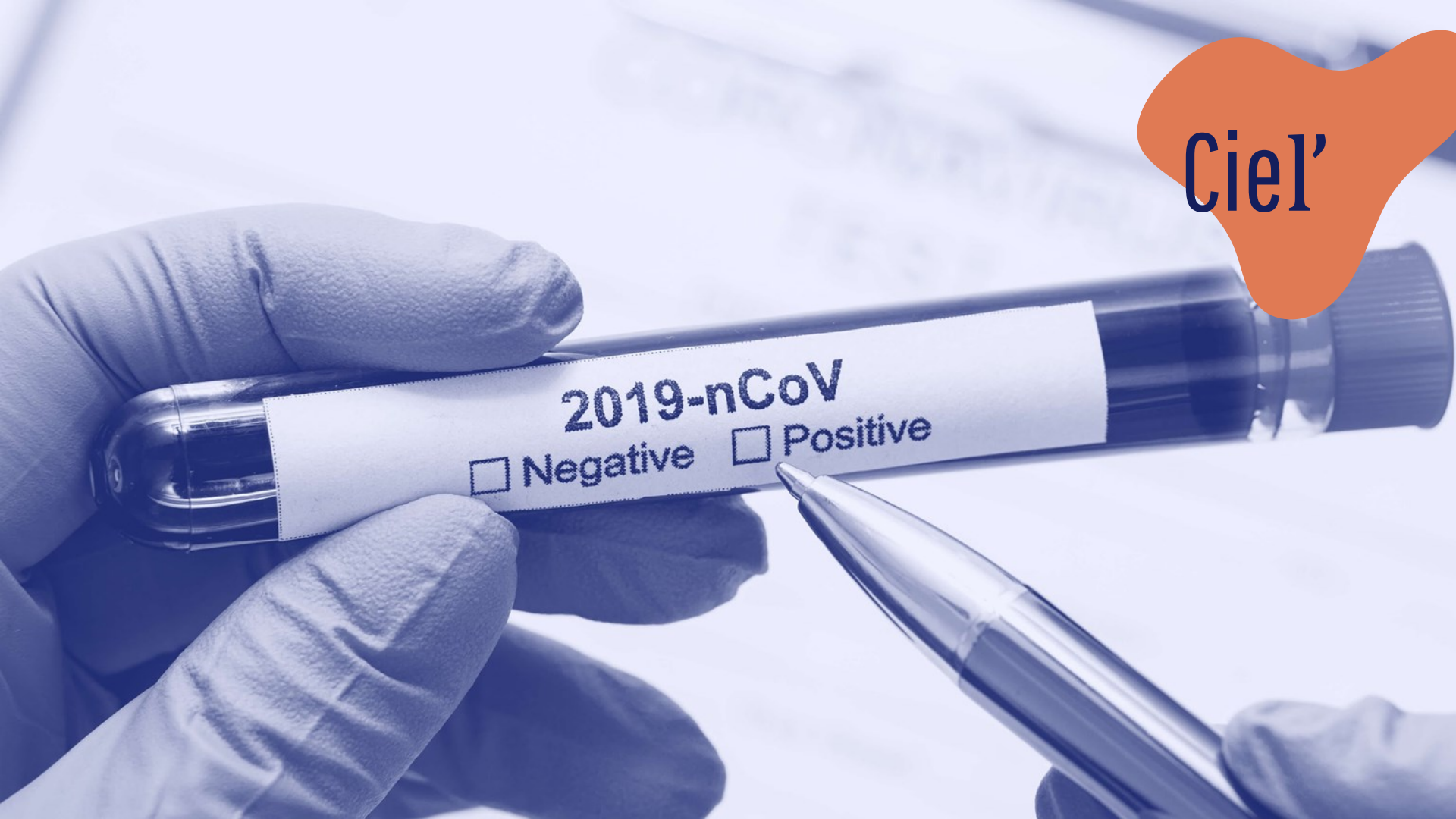
Neurónové siete

COVID-19



Ciel'

01

Dataset

02

Predspracovanie dát

Kód

03

Učenie

Vytvorenie neurónovej siete

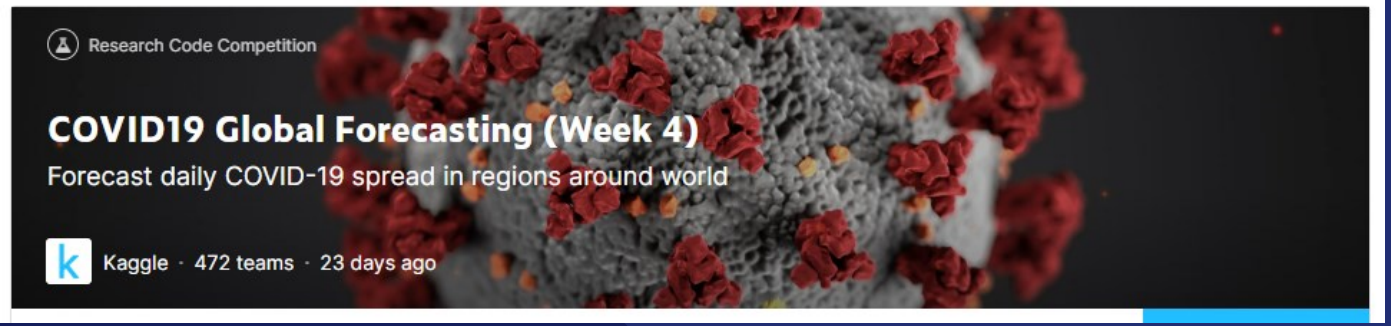
04

Výstup a výsledné dáta

Porovnania

Vizualizácia

Dataset


Research Code Competition
COVID19 Global Forecasting (Week 4)
Forecast daily COVID-19 spread in regions around world
Kaggle · 472 teams · 23 days ago

Súbory

01

train.csv
Trénovacie dáta
Od 22.1.2020 do 17.4.2020

02

test.csv
Dátumy, ktoré treba predpovedať
Od 2.4.2020 do 14.5.2020

03

submission.csv
Výsledné dáta

Kód – Import základných knižníc na prácu s datasetom

```
import numpy as np
import pandas as pd

import operator
import matplotlib
import matplotlib.pyplot as plt

from sklearn.linear_model import LinearRegression
from sklearn.preprocessing import PolynomialFeatures, StandardScaler


df_train = pd.read_csv("train.csv", parse_dates=['Date'])
df_test = pd.read_csv("test.csv", parse_dates=['Date'])
```

Učenie #1

```
df_train.head()
x_train = df_train[[]]

# Pre predikciu novych pripadov vyuzijem Linearnu regresiu
for country in df_train['Country_Region'].unique():
    # print('Robim model pre krajinu: {}'.format(country))
    df_train_country = df_train[df_train['Country_Region'] == country]
    df_test_country = df_test[df_test['Country_Region'] == country]
```

Učenie #2

```
# zaujima ma, ci ma krajina provinciu, ak ano -> vetva else
    if df_train_country['Province_State'].isnull().unique().any():
        # indexujem krajiny v csv subore - od a do
        x_train = np.array(range(len(df_train_country))).reshape((-1, 1))
        #print(x_train)
        y_train = df_train_country['ConfirmedCases']
        # print(y_train)

        # ucenie siete
        model = Pipeline([('poly', PolynomialFeatures(degree=2, include_bias=True)), ('linear', LinearRegression(fit_intercept=False))])
```

Učenie #3

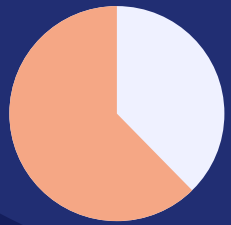
```
Y = np.array([y_train])
Yr = np.flip(Y, axis=1)
Yr = Yr[0,::-1]
Yrs = Yr.reshape(-1,1)
x_train_scale = StandardScaler().fit_transform(x_train)
y_train_scale = StandardScaler().fit_transform(Yrs)

model.fit(x_train_scale, y_train_scale)
```

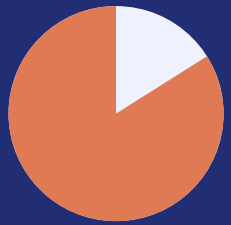
```
# zacnem indexovat dni od 0
x_test = np.array(range(len(df_test_country))).reshape((-1, 1))
prediction = model.predict(x_test)

# negativne pripady nahradim nulou
prediction[prediction < 0] = 0

# Vramci ConfirmedCases pridam sa novy stlpec v dataframe df_test kam vlozim predikciu k danej krajine
df_test.loc[df_test['Country_Region'] == country, 'ConfirmedCases'] = prediction
```

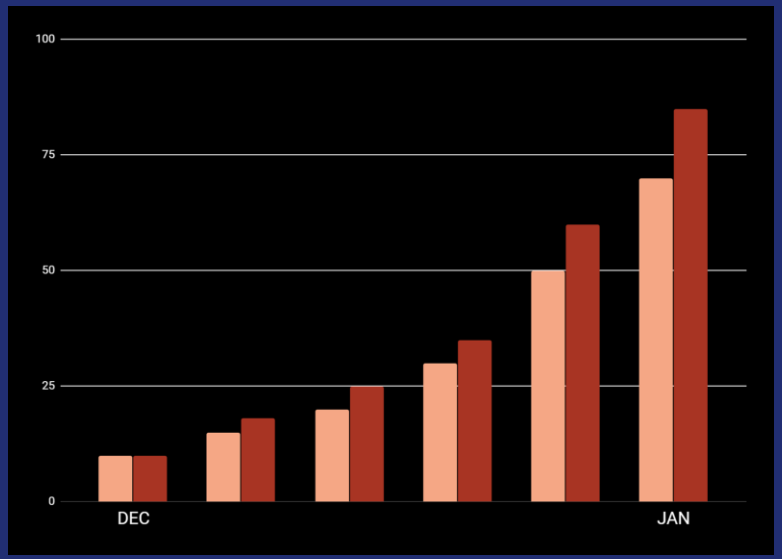
Učenie #4

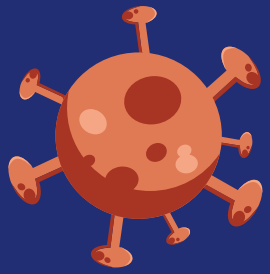
```
else:
    for state in df_train_country['Province_State'].unique():
        df_train_state = df_train_country[df_train_country['Province_State'] == state]
        df_test_state = df_test_country[df_test_country['Province_State'] == state]

        x_train = np.array(range(len(df_train_state))).reshape(-1,1)
        y_train = df_train_state['ConfirmedCases']

        model = Pipeline([('poly', PolynomialFeatures(degree=2, include_bias=True)), ('linear', LinearRegression(fit_intercept=False))])

        Y = np.array([y_train])
        Yr = np.flip(Y, axis=1)
        Yr = Yr[0,::-1]
        Yrs = Yr.reshape(-1,1)
        x_train_scale = StandardScaler().fit_transform(x_train)
        y_train_scale = StandardScaler().fit_transform(Yrs)

        model.fit(x_train_scale, y_train_scale)

        x_test = np.array(range(len(df_test_state))).reshape((-1,1))
        prediction = model.predict(x_test)

        prediction[prediction < 0] = 0
        df_test.loc[(df_test['Country_Region'] == country) & (df_test['Province_State'] == state), 'ConfirmedCases'] = prediction
```

Vizualizácia









Výstup

Vloženie naučených dát do submission.csv

```
# len vlozim data do pripraveneho csv suboru
df_submit = pd.read_csv('submission.csv')
df_submit['ConfirmedCases'] = df_test['ConfirmedCases'].astype('int')
df_submit.to_csv('submission.csv', index=False)
```

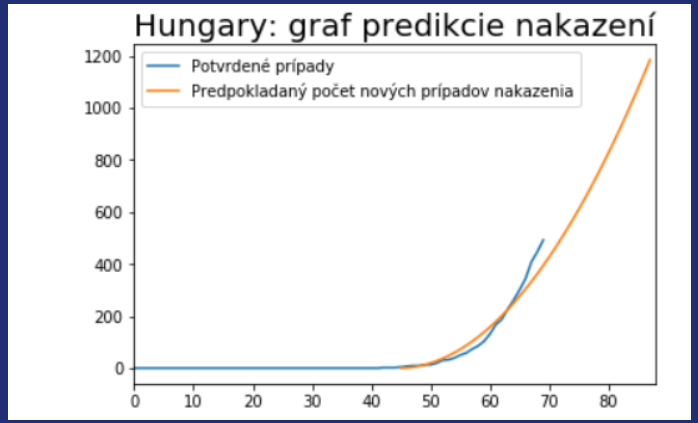
Graf - kód

```
if(country == “Slovakia"):
        plt.plot(x_train[:70,0], Yrs[:70,0])
        plt.xlim(0,10)

        #print(prediction)
        plt.plot(x_test+45, prediction)
        plt.xlim(0,88)
        plt.legend(['Potvrdené prípady', 'Predpokladaný počet nových prípadov nakazenia'])
        plt.title(country+': graf predikcie nakazení', size=20)

        plt.show()
```


Slovakia: graf predikcie nakazení
Potvrdené prípady
Predpokladaný počet nových prípadov nakazenia
0
200
400
600
800
1000
0
10
20
30
40
50
60
70
80


Albania: graf predikcie nakazení
Potvrdené prípady
Predpokladaný počet nových prípadov nakazenia
0
200
400
600
800
1000
0
10
20
30
40
50
60
70
80


India: graf predikcie nakazení
Potvrdené prípady
Predpokladaný počet nových prípadov nakazenia
0
200
400
600
800
1000
1200
1400
0
10
20
30
40
50
60
70
80


Hungary: graf predikcie nakazení
Potvrdené prípady
Predpokladaný počet nových prípadov nakazenia
0
200
400
600
800
1000
1200
0
10
20
30
40
50
60
70
80

Vizualizázia Slovenska v porovnaní s reálnymi údajmi

	ForecastId	Province_State	Country_Region	Date	ConfirmedCases
9360	9361	NaN	Slovakia	01.05.2020	541
9361	9362	NaN	Slovakia	02.05.2020	578
9362	9363	NaN	Slovakia	03.05.2020	617
9363	9364	NaN	Slovakia	04.05.2020	656
9364	9365	NaN	Slovakia	05.05.2020	697
9365	9366	NaN	Slovakia	06.05.2020	739
9366	9367	NaN	Slovakia	07.05.2020	783
9367	9368	NaN	Slovakia	08.05.2020	827
9368	9369	NaN	Slovakia	09.05.2020	873
9369	9370	NaN	Slovakia	10.05.2020	920
9370	9371	NaN	Slovakia	11.05.2020	969
9371	9372	NaN	Slovakia	12.05.2020	1018
9372	9373	NaN	Slovakia	13.05.2020	1069
9373	9374	NaN	Slovakia	14.05.2020	1121

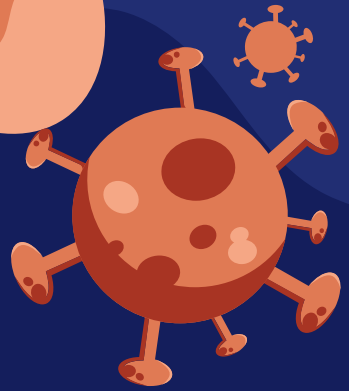
	Datum	Pocet nakazenych
0	01.05.2020	1403
1	02.05.2020	1407
2	03.05.2020	1408
3	04.05.2020	1413
4	05.05.2020	1421
5	06.05.2020	1429
6	07.05.2020	1445

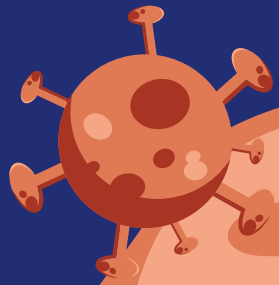
Záver

ZDROJE

- https://www.researchgate.net/publication/340394734_Neural_network_based_country_wise_risk_prediction_of_COVID-19
- https://github.com/aatishb/covid/blob/master/curvefit.ipynb
- https://meltingasphalt.com/interactive/outbreak/
- https://www.washingtonpost.com/graphics/2020/world/corona-simulator/
- https://www.youtube.com/watch?v=BtN-goy9VOY
- http://gabgoh.github.io/COVID/index.html

Ďakujem za pozornosť





Diskusia

Fonts & colors used

This presentation has been made using the following fonts:

Pathway Gothic One
(https://fonts.google.com/specimen/Pathway+Gothic+One)

Roboto Slab
(https://fonts.google.com/specimen/Hind)

#f5a785 #212e73 #eff1ff #e07a54


slidesgo