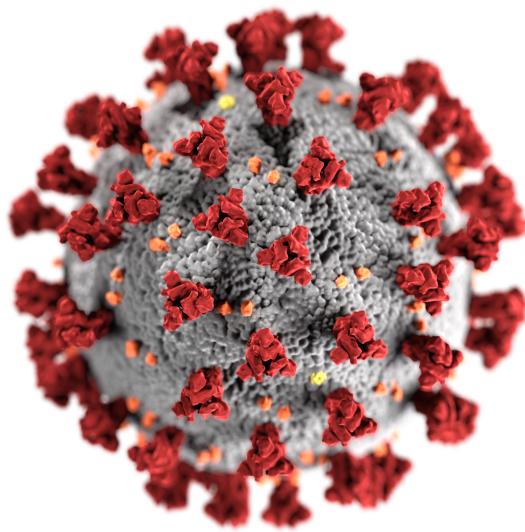


Instrukcja instalacji i obsługi
Modelowanie epidemii SARS-CoV-2

Wiktoria Wilman
Gustaw Walaś
Uniwersytet Przyrodniczy we Wrocławiu

2021



Spis treści

1 Wstęp	2
1.1 Prawa autorskie	2
1.2 Cel oprogramowania	2
2 Ogólna specyfikacja	2
3 Opis programu	2
3.1 Środowisko	2
3.2 Ograniczenia projektu	2
4 Opisy algorytmów	2
4.1 Model SIR	2
4.1.1 Wstępne założenia modelu	2
4.1.2 Niezbędne parametry wejściowe	3
4.2 Model Machine Learning	3
4.2.1 Założenia modelu	3
4.2.2 Niezbędne parametry wejściowe	4
5 Instalacja oprogramowania	4
5.1 Wymagania sprzętowe	4
5.2 Pod systemem Windows	4
5.3 Pod systemem Linux	5
6 Obsługa	5

1 Wstęp

1.1 Prawa autorskie

Projekt jest tworzony przez studentów I roku II stopnia Bioinformatyki: Gustawa Walasia i Wiktorię Wilman na polecenie Zleceniodawcy dr Wojciecha Jakubowskiego pod patronatem zajęć Pracownia informatyczna. Temat projektu został wybrany przez studentów i poziom jego wykonania będzie odzwierciedlał ich umiejętności w zakresie budowania oprogramowania zgodnie z zaleceniami Zleceniodawcy.

1.2 Cel oprogramowania

Celem projektu było otrzymanie aplikacji działającej na bazie modelu SIR, która będzie na podstawie otrzymanej bazy danych modelowała rozprzestrzenianie się epidemii wywołanej przez SARS-CoV-2 w konkretnym kraju i prezentowała to za pomocą wykresów.

2 Ogólna specyfikacja

Celem projektu jest otrzymanie modelu względnie jak najlepiej przewidującego rozprzestrzenianie się epidemii wywołanej przez SARS-CoV-2 na podstawie otrzymanej bazy danych dla konkretnego kraju. Zostały stworzone i porównane ze sobą Model SIR oraz Model ML.

3 Opis programu

3.1 Środowisko

Oprogramowanie zostało stworzone przy pomocy Python 3 i było testowane zarówno na środowisku Windows jak i Linux. Do tworzenia oprogramowania był wykorzystywany również Jupyter Notebook, w którym najlepiej było pracować z modelem ML i wizualizować wyniki.

3.2 Ograniczenia projektu

Stworzony model nie będzie nadawał się do wszystkich możliwych epidemii. Jest oparty na założeniach spełnianych przez epidemię SARS-CoV-2 i nie może mieć zastosowania przy pracach nad modelowaniem innych epidemii.

Jednocześnie możliwe jest dostosowanie modelu SIR jak i modelu ML do pracy na danych pochodzących z różnych krajów bądź kontynentów.

4 Opisy algorytmów

4.1 Model SIR

4.1.1 Wstępne założenia modelu

Wstępne nierzeczywiste założenia dla modelu SIR:

1. Wyzdrowienie z choroby = nabycie odporności (nawet u zmarłych);
2. Populacja składa się z 3 grup:

- 1 grupa - osoby podatne S ;
 - 2 grupa - osoby zainfekowane I ;
 - 3 grupa - osoby ozdrowiałe R ;
- Przyrost w grupie osób zainfekowanych jest proporcjonalny do liczby zainfekowanych i podatnych (r - przyrost zainfekowanych, $r \geq 0$ to stały parametr).
 - Współczynnik ozdrowienia i przenoszenia do grupy odpornych jest proporcjonalny do liczby osób zainfekowanych (a - zdrowienie zainfekowanych, $a \geq 0$ to stały parametr).
 - Okres inkubacji jest na tyle krótki, że zakładamy, że zainfekowany jest natychmiast chory.
 - Zakładamy, że każda para osób ma takie samo prawdopodobieństwo bezpośredniego kontaktu.
 - Współczynnik zdrowienia jako stosunek zdrowienia zainfekowanych do rozchorowywania się zdrowych obliczymy na podstawie poprzednich parametrów $p = a/r$
 - Współczynnik kontaktów jako stosunek rozchorowywania się zdrowych do zdrowienia zainfekowanych obliczymy na podstawie poprzednich parametrów $\sigma = r/a$, odwrotność współczynnika zdrowienia.
 - R_0 to bazowy współczynnik reprodukcji, przedstawia liczbę osób wtórnie zainfekowanych względem jednego pierwotnego osobnika pierwotnie zainfekowanego
 - Sposoby na zmniejszenie bazowego współczynnika reprodukcji:
 - Szczepienia
 - Odporność zbiorowa

4.1.2 Niezbędne parametry wejściowe

Najważniejszym parametrem jest baza danych, która musi zawierać:

- Ograniczenie lokalizacji jaką bierzemy pod uwagę (w zależności od kraju bądź opcji dla całego świata)
- Daty kolejnych pomiarów
- Podział na grupy według modelu

Na podstawie powyższych danych jesteśmy w stanie uzyskać wszystkie pozostałe niezbędne parametry.

4.2 Model Machine Learning

4.2.1 Założenia modelu

Nasz model tworzymy przy użyciu rekurencyjnej architektury sieci neuronowej. Model składa się z warstwy wejściowej, po której następują trzy warstwy LSTM, które wykorzystują tzw. dropout, aby zapobiec nadmiernemu dopasowaniu naszego modelu. Wynik jest warstwą z pojedynczym neuronem używającym funkcji aktywacji ReLU, ponieważ przewidujemy liczbę przypadków Corona, więc zapewniamy by wynik był liczbą dodatnią (0,).

4.2.2 Niezbędne parametry wejściowe

Najważniejszym parametrem jest baza danych, która musi zawierać:

- Ograniczenie lokalizacji jaką bierzemy pod uwagę (w zależności od kraju bądź opcji dla całego świata)
- Daty kolejnych pomiarów

Oprócz tego wymagane parametry dla modelu to:

- batch size - ilość danych branych na raz do epoki treningowej
- epochs - ilość sesji treningowych
- podział na dane treningowe i testowe - zapewnia możliwość nauczenia się, a następnie przetestowania predykcji

5 Instalacja oprogramowania

5.1 Wymagania sprzętowe

- procesor nie starszy niż Intel i5 lub Amd A10
- system operacyjny: Linux Bionic Beaver (lub nowszy) lub Windows 7 (lub nowszy)
- minimum 4GB RAM
- 1GB wolnego miejsca na dysku
- Python 3.7 (lub nowszy)

5.2 Pod systemem Windows

Włącz wiersz poleceń (CMD), a następnie:

1. Sprawdź wersję swojego Python'a przy pomocy

```
python3 --version
```

2. Zainstaluj pip

```
python get-pip.py --user
```

3. Zainstaluj wymagane biblioteki

```
python -m pip install --user numpy scipy matplotlib
```

```
jupyter pandas datetime sklearn tensorflow
```

4. Zainstaluj Jupyter Notebook

```
pip install notebook
```

5.3 Pod systemem Linux

Włącz terminal, a następnie:

1. Sprawdź wersję swojego Python'a przy pomocy

```
python3 --version
```

2. Zainstaluj pip

```
python get-pip.py --user
```

3. Zainstaluj wymagane biblioteki

```
python -m pip install --user numpy scipy matplotlib ipython
```

```
jupyter pandas datetime sklearn tensorflow
```

4. Zainstaluj Jupyter Notebook

```
pip install notebook
```

6 Obsługa

Aby włączyć program należy włączyć CMD w Windowsie (terminal w Linuxie) i wpisać polecenie

```
jupyter notebook ścieżka_do_pliku/Program_modelowanie_epidemii.ipynb
```

W odpowiedzi na komendę odpali nam się Jupyter Notebook z kodem programu. W celu przeprowadzenia modelowania należy wpisać w cudzysłowie jedną z angielskich nazw krajów do zmiennej "kraj".

Przykład dla Zimbabwe:

```
kraj='Zimbabwe'
```

Przykład dla Polski:

```
kraj='Poland'
```

Następnie należy z górnego paska opcję "Cell" -> "Run All" by otrzymać wyniki dla danego kraju. Aby zmienić kraj docelowy należy zmienić zawartość zmiennej "kraj".