

Porównanie walki wybranych krajów z koronawirusem

Autor:

Szymon Majorek

szy.majorek@gmail.com

Github: https://github.com/wizkiz/korona_science_hw3/blob/master/project.py

Colab: https://colab.research.google.com/drive/1uqaVvkqEs-9mh9W7sA7qCZJ_UlfsYHjmp

Spis treści

1. Wprowadzenie
2. Źródła danych
3. Omówienie indywidualnych państw
 - a. Włochy
 - b. USA
4. Porównanie
5. Wnioski

1. Wprowadzenie

Będę się przyglądał sytuacji dotyczącej koronawirusa w USA oraz we Włoszech. Będę głównie operował na ilości przypadków zachorowań w danym kraju, tylko napominając o śmierciach. Postaram się również przeprowadzić powierzchowną próbę przewidzenia ilości przypadków na przyszłość. Teoretycznie mógłbym do tego użyć uczenia maszynowego, ale ze względu na to, że nasz szereg czasowy ma dosyć małą ilość danych (jest to nowa sytuacja i dane zaczynają się najwcześniej z początkiem tego roku), będę do celów przewidywania tego, co może nastąpić w przyszłości używał dopasowywania funkcji zamiast uczenia maszynowego.

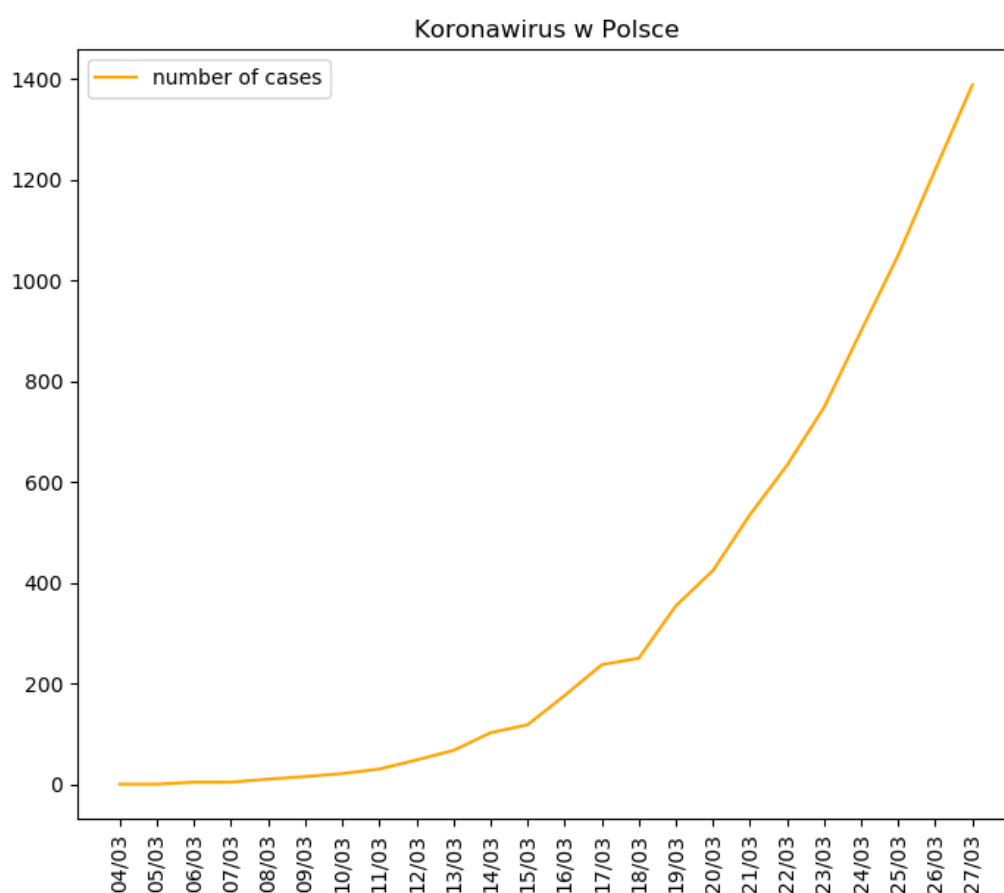


Figure 1 - Przykładowy wykres

Jak widać na figurze 1. wykres przypomina kształtem funkcję wykładniczą, możemy to wykorzystać do dopasowania parametrów A, B, C funkcji $f(x) = Ae^{Bx} + C$ w celu przewidywania przyszłych wyników. Trzeba jednak pamiętać, że taka metoda jest bardzo pogładowa i nie bierze pod uwagę bardzo wielu czynników takich jak dystansowanie międzyludzkie, czy ogólna zmiana nastawienia społeczeństwa mająca na celu minimalizację ilości zachorowań i spłaszczenie właśnie tej krzywej. Można więc uznać, że przewidywania które będę liczył będą najgorszym możliwym przypadkiem, kiedy zupełnie nic się nie zmieni i ilość przypadków, czy też śmierci, będzie niezmiennie podążała jedną ścieżką. Jedną z kwestii, które chciałbym spróbować poruszyć jest też ewentualny wpływ wprowadzenia kwarantanny na zachowanie się wykresu ukazującego ilość przypadków.

2. Źródło danych

W analizie są pobierane dane na temat koronawirusa – liczba przypadków oraz śmierci dla wybranego kraju oraz dane ogólne odnośnie tego samego kraju, z czego wykorzystywane będzie tylko liczba przypadków oraz liczba populacji. Do danych o koronawirusie używam api corona.lmao.ninja/v2/historical natomiast do danych o populacji używam restcountries.eu/rest/v2/name/

3. Omówienie poszczególnych państw

Na początek przyjrę się wynikom dla poszczególnych państw, będą to Włochy w których sytuacja obecnie uważana jest za najgorszą, USA, gdzie dopiero robi się tragicznie, za to o wiele szybciej niż we Włoszech oraz Polska, z oczywistych względów.

Chiny również są dobrym kandydatem do analizy, jednakże nie pasują do modelu pozostałej trójki, gdyż ich wykres, przedstawiony poniżej, już się wypłaszczył i nie przejawia już tendencji wykładniczej.

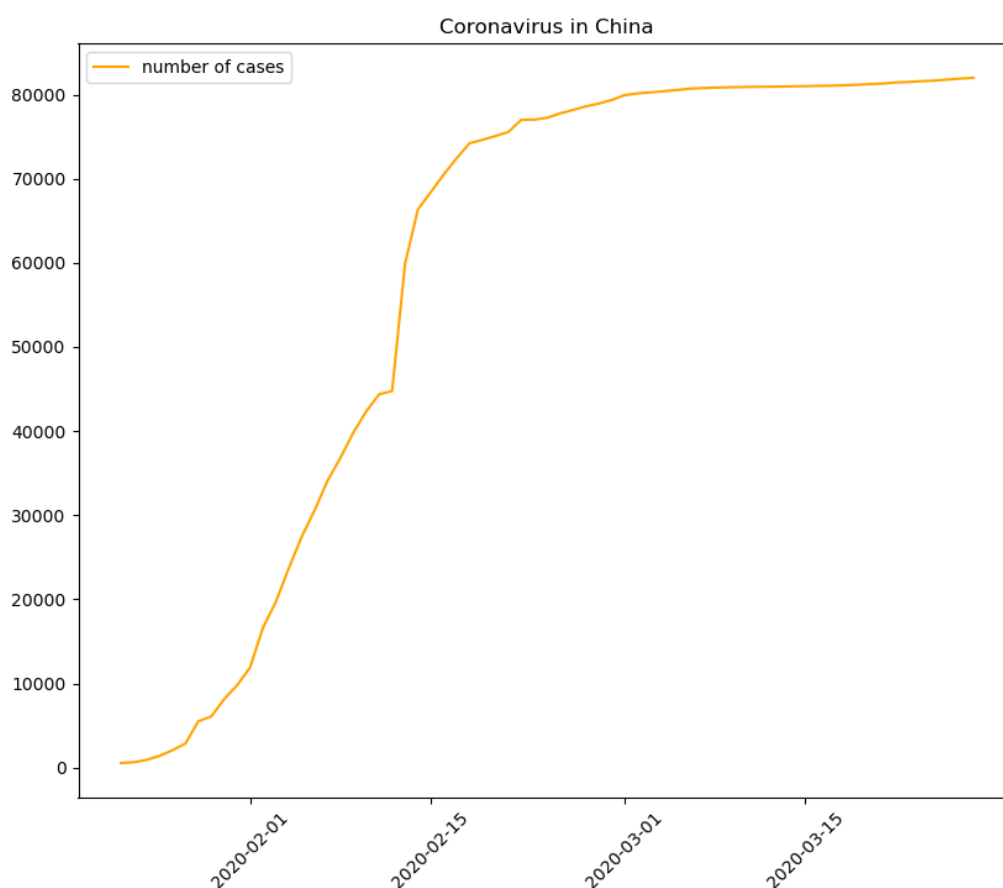


Figure 2 - Wykres koronawirusa w Chinach

3. a) Włochy

O sytuacji we Włoszech jest w mediach najgłośniejsze, i nie bez przyczyny, ich problemy z koronawirusem trwają w Europie najdłużej oraz są najpoważniejsze, w ostatnich dniach umiera tam dziennie niemalże 1000 osób.

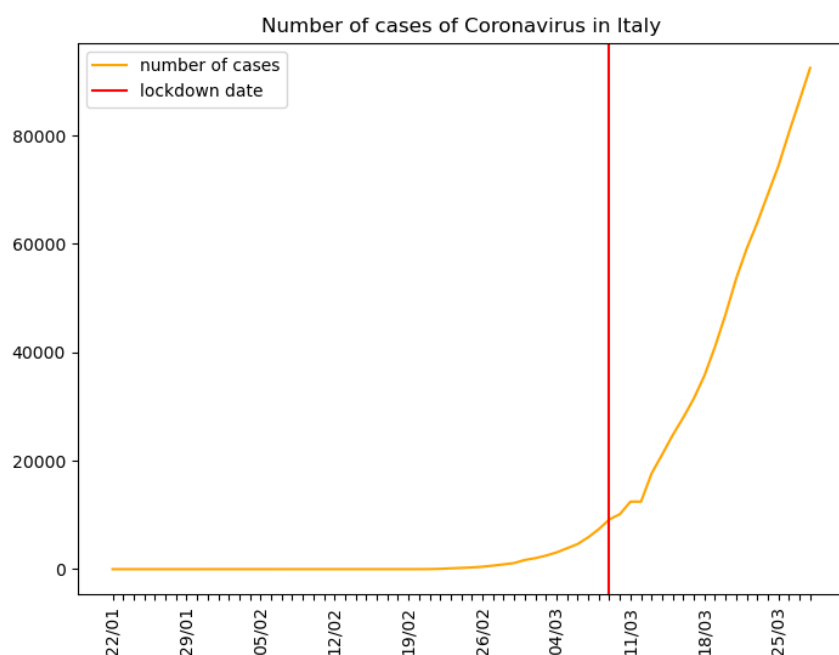


Figure 3 - Wykres ilości przypadków koronawirusa we Włoszech

Do wykresu została dodana linia daty nałożenia kwarantanny na cały kraj (9. Marca 2020). Nie widać wyraźnego spłaszczenia wykresu po wprowadzeniu kwarantanny. Może być to odbierane negatywnie, aczkolwiek drugą opcją jest to, że gdyby nie ona to wykres byłby niemalże pionowy. Ukazane wyżej dane możemy wykorzystać do próby przewidzenia tego co będzie dalej, wykorzystam do tego funkcję *curve_fit* zamiast typowego podejścia jakim byłoby uczenie maszynowe.

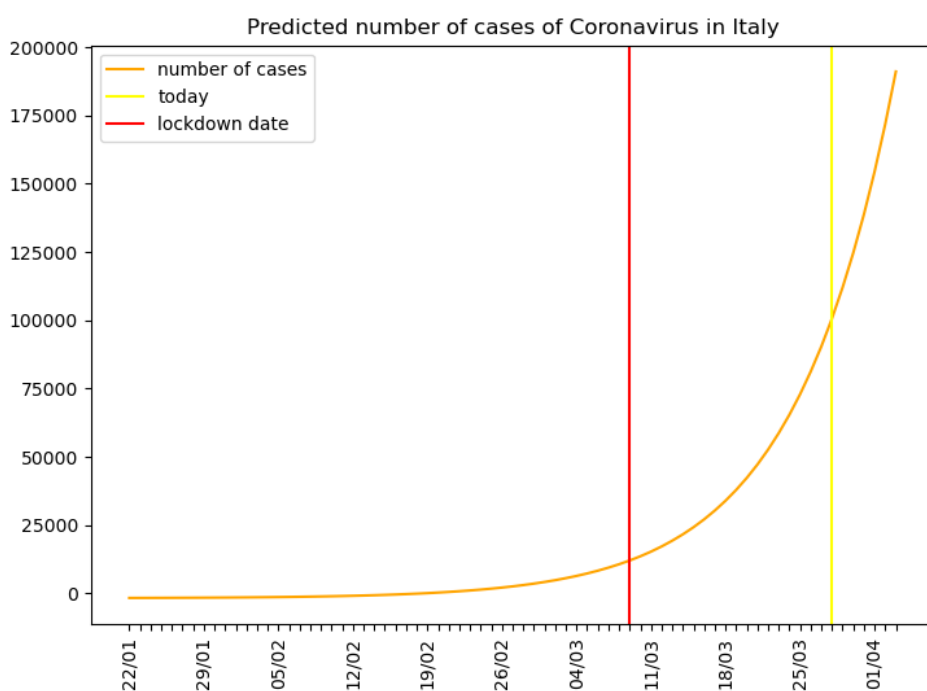


Figure 4 - Wykres przewidywanych przypadków koronawirusa we Włoszech

Z dopasowanej funkcji ($f(x) = 97.37 \cdot e^{0.1x} - 1743.73$) wynika, że jeżeli obecny trend się nie zmieni w znaczącym stopniu, w przeciągu nieco ponad tygodnia ilość przypadków we Włoszech osiągnie poziom 200000 co z kolei znacząco zwiększy śmiertelność, gdyż szpitale nie będą w stanie przyjąć tylu nowych pacjentów i zapewnić im potrzebnej opieki.

3. b) Stany Zjednoczone

Drugim z krajów któremu chciałbym się przyjrzeć są stany zjednoczone.

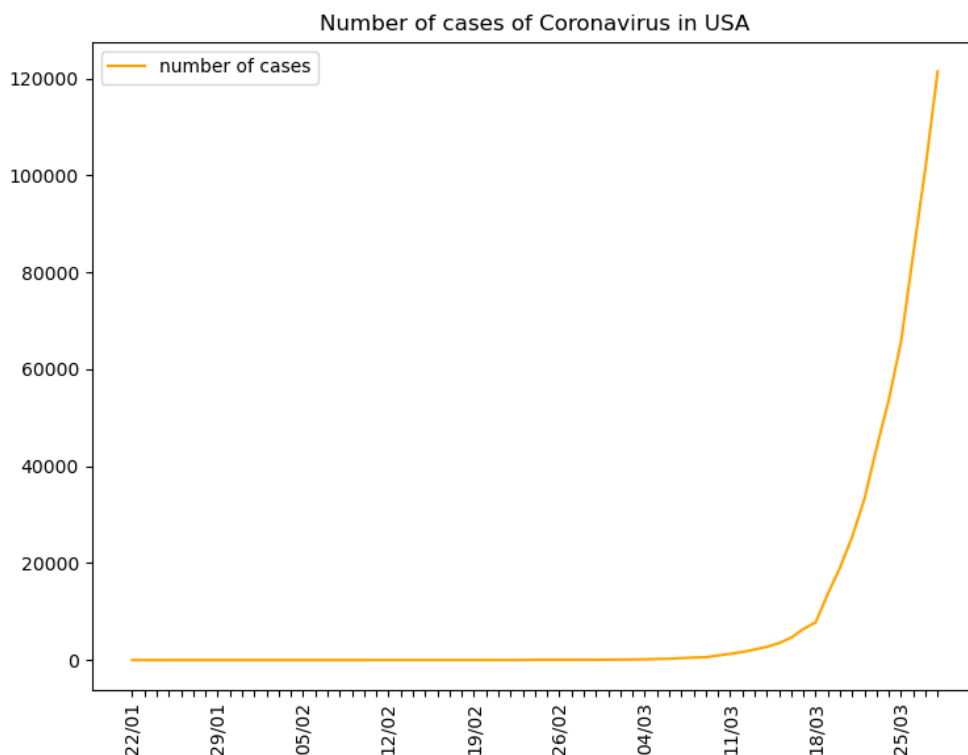


Figure 5 - Wykres przypadków koronawirusa w USA

Do danych w USA dopasowana została funkcja $f(x) = 0.04 \cdot e^{0.23x} - 435$. Jest ona bardzo płaska na początku, jednakże, gdy już zacznie rosnąć, rośnie bardzo szybko. Jest to widoczne w ostatnich dniach, gdzie na przykład przez ostatnią dobę przybyło niemalże 20 tysięcy przypadków. Co ciekawe w USA nie została wprowadzona kwarantanna mimo że przekroczyli już liczby osiągnięte przez Chiny.

Dopasowawszy funkcję do danych, możemy przyjrzeć się wykresowi próbującemu przewidzieć co będzie się działo odnośnie przypadków zachorowań w następnych dniach.

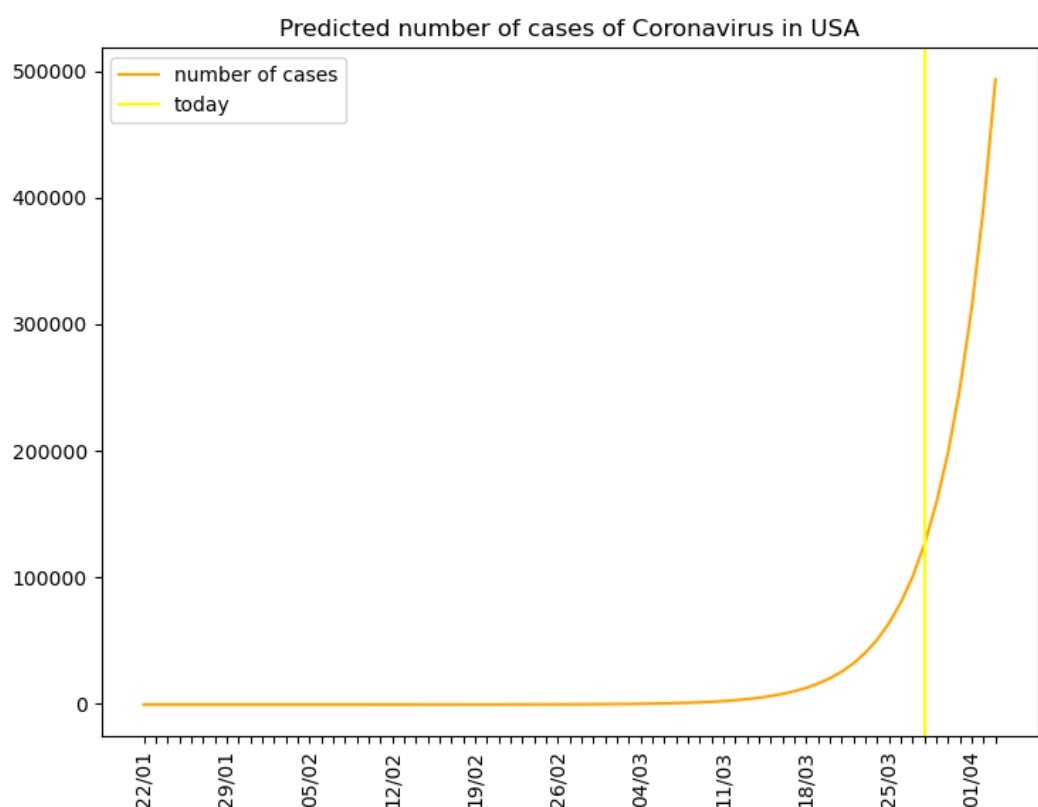


Figure 6 - Wykres przewidywanych przypadków koronawirusa we USA

Ze względu na duży współczynnik przy argumentie w dopasowanej funkcji, wykres jest o wiele bardziej pionowy, przez co liczba zachorowań rośnie dużo bardziej gwałtownie.

4. Porównanie

W celu porównania użyję nie ilości zachorowań, ale ilość zachorowań na milion mieszkańców, gdyż liczba mieszkańców Włoch oraz USA różni się diametralnie przez co bezpośrednie porównanie byłoby bez sensu. W celu policzenia ilości zachorowań na milion mieszkańców użyłem danych o populacji pobranych z API udostępnionego przez restcountries.eu.

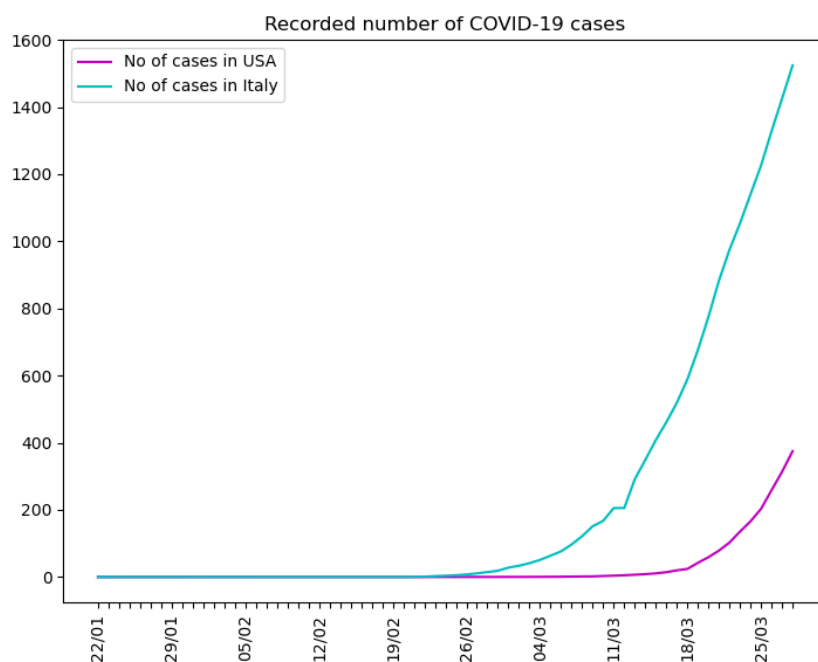


Figure 7 - Porównanie obecnych danych o ilości przypadków koronawirusa

Jak widać na wykresie sytuacja we Włoszech jest o wiele gorsza, jednakże wiemy również, że funkcja dla USA ma o wiele większy współczynnik przy argumentie, przez co będzie bardziej pionowa, można się więc spodziewać, że w przyszłości USA będzie miało więcej przypadków na milion mieszkańców niż Włochy.

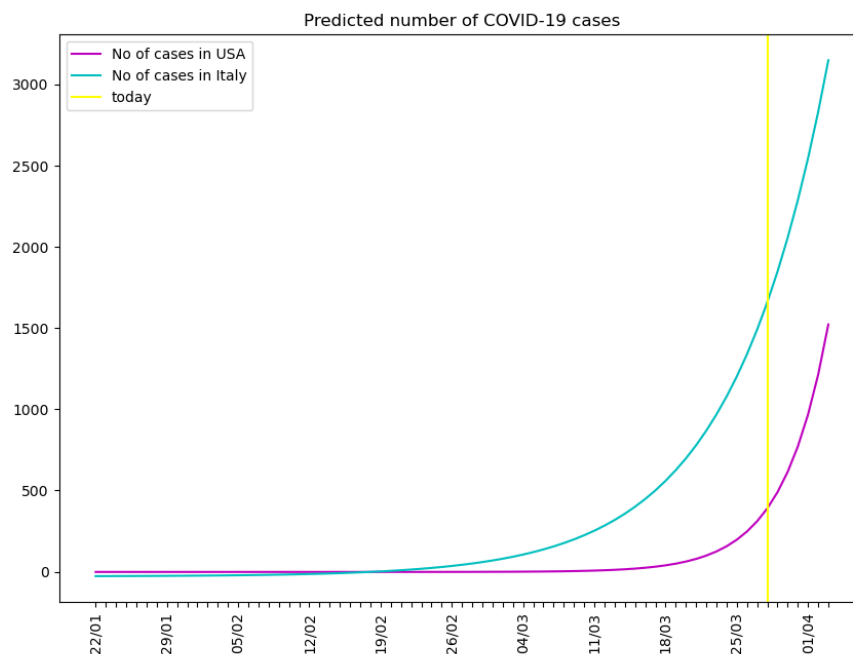


Figure 8 - Porównanie przewidywań dla USA i Włoch (ok. 7 dni)

W najbliższym tygodniu można się spodziewać, że liczba zachorowań na milion mieszkańców w USA znacząco się zbliży do tej we Włoszech, ale nie będzie wyższa. Jeżeli jednak spojrzymy na jeszcze kolejny tydzień, to sytuacja jest już inna.

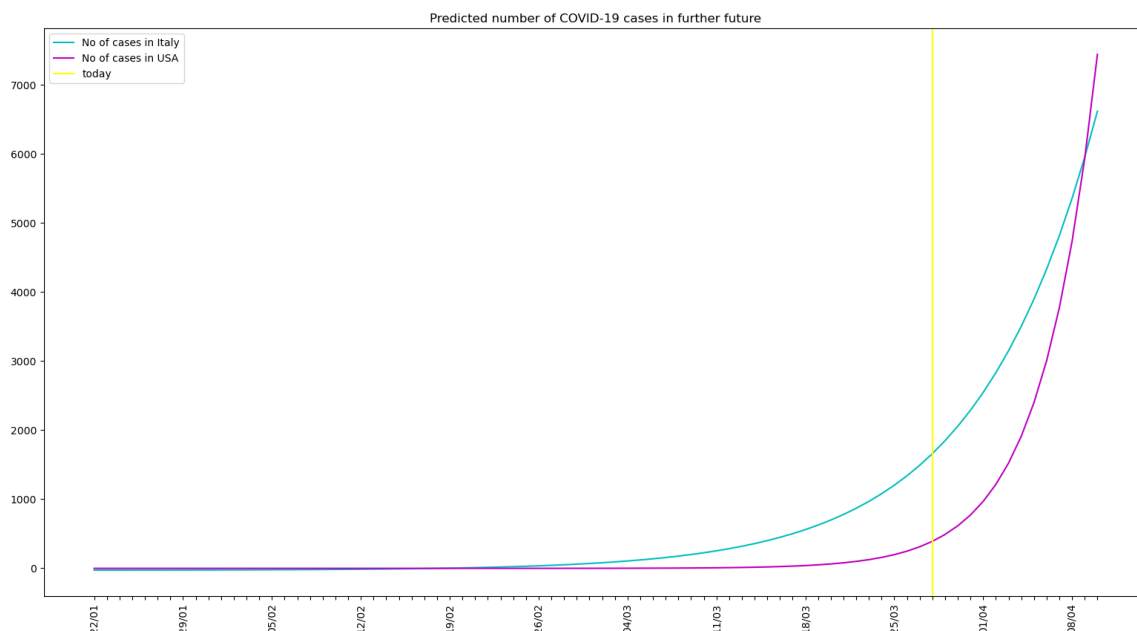


Figure 9 - Porównanie przewidywań dla USA i Włoch (ok. 14 dni)

Tutaj już widać moment w którym USA obejmie prowadzenie pod tym względem.

5. Wnioski

Mimo tego jak dramatyczna wydaje się być sytuacja we Włoszech, to co się będzie w niedalekiej przyszłości działo w USA może okazać się jeszcze gorsze, i to nie tylko jeśli chodzi o całkowitą liczbę zachorowań, ale również jeżeli chodzi o liczbę zachorowań przypadających na milion mieszkańców. Interesujące może być jak system zdrowotny, oraz sami ludzie w USA poradzą sobie z tak potężną falą zachorowań, kiedy okaże się, że USA jest kolejnym dużym (a może i największym) epicentrum jeżeli chodzi o koronawirusa.