

Raport

Wstęp:

Geneza:

Po wstępnych badaniach związanych z artykułem „An interpretable mortality prediction model for COVID-19 patients”. Nasz zespół uznał, że projekt który wtedy wykonaliśmy zupełnie wyczerpała interesujące nas aspekty wspomnianej pracy. Ponadto uznaliśmy, że o ile sam model oczywiście był przydatny w medycynie, to nie miał on aż tak dużego znaczenia dla normalnej osoby. Z powyższych powodów postanowiliśmy przyjrzeć się uczeniu maszynowemu w kontekście koronawirusa pod zupełnie innym kątem.

O drugim artykule:

Postanowiliśmy zbadać artykuł „Predicting Mortality Due to SARS-CoV-2: A Mechanistic Score Relating Obesity and Diabetes to COVID-19 Outcomes in Mexico” między innymi ze względu na zupełnie inne podejście do zagadnienia predykcji śmiertelności czy też zachorowań. Podczas gdy w pierwszym artykule wyniki modelu opierały się tylko i wyłącznie na pomiarach parametrów krwi, obecności białek i innych typowo medycznych wskaźników, tak w drugim artykule predykcja bazowała na wszelakich chorobach współtowarzyszących.

Autorzy tego tekstu skupiali się na wyłapaniu zależności między chorobami współtowarzyszącymi, a reagowaniem na wirusa Covid-19, szczególnie skupiali się na cukrzycy oraz otyłości co sugeruje tytuł artykułu.

O naszej inspiracji:

Wyniki pracy meksykańskiego zespołu oraz zupełnie inne podejście do przewidywania zagrożenia chorobą zainspirowały nas do stworzenia aplikacji, która na podstawie chorób współistniejących przewidywałaby ryzyko zachorowania na koronawirusa. W ten sposób dopełnilibyśmy meksykański artykuł poprzez dostarczenie prostej w obsłudze przez każdą osobę aplikacji. W odróżnieniu od większości modeli każdy człowiek byłby w stanie poznać ryzyko zachorowania na podstawie swojego wieku oraz chorób towarzyszących, a więc danych o których posiada wiedzę nie wychodząc z domu.

Opis Aplikacji

O modelu:

Stworzyliśmy model liczący prawdopodobieństwo zakażenia Covid-19. W tym celu użyliśmy narzędzia ranger, który implementuje los losowy. Niewątpliwą zaletą modelu jest to, że jest zrozumiały i wytłumaczalny, jednak nie osiąga tak wysokich wyników jakie byśmy chcieli, więc w dalszej pracy nad projektem popracujemy nad ulepszeniem tego modelu lub użyjemy innego narzędzia.

O obsłudze:

GUI aplikacji podzielone jest na dwa panele: główny oraz boczny.

Panel główny przełączać można między 3 zakładkami na których znajdują się wizualizacje tworzone na podstawie danych wybranych na panelu bocznym.

Panel boczny składa się z suwaka wyboru wieku oraz pytań dotyczących chorób przewlekłych. Na podstawie tych danych generowane są wizualizacje w panelu głównym.

O wynikach wizualizacjach:

Aplikacja udostępnia użytkownikowi trzy wizualizacje oraz współczynnik ryzyka zachorowania na koronawirusa.

Pierwsza wizualizacja przedstawia w jaki sposób poszczególne zmienne wpływają na wynik modelu (tak zwany breakdown plot). Dzięki tej wizualizacji jesteśmy w stanie zapewnić wyjaśnialność naszego modelu.

Drugi wykres prezentuje ryzyko zachorowania przy wybranych chorobach współistniejących w zależności od wieku chorego.

Ostatnia wizualizacja prezentuje natomiast wykres shap dla podanych danych. Czym jest wykres shap? SHapley Additive exPlanations to funkcje, które mówią nam, dlaczego dana obserwacja należy do danej klasy, albo ma daną wartość.