



# Projekt UmWPI:

Przewidywanie hepatoksyczności +  
interpretowalność

Nikodem Kwaśniak

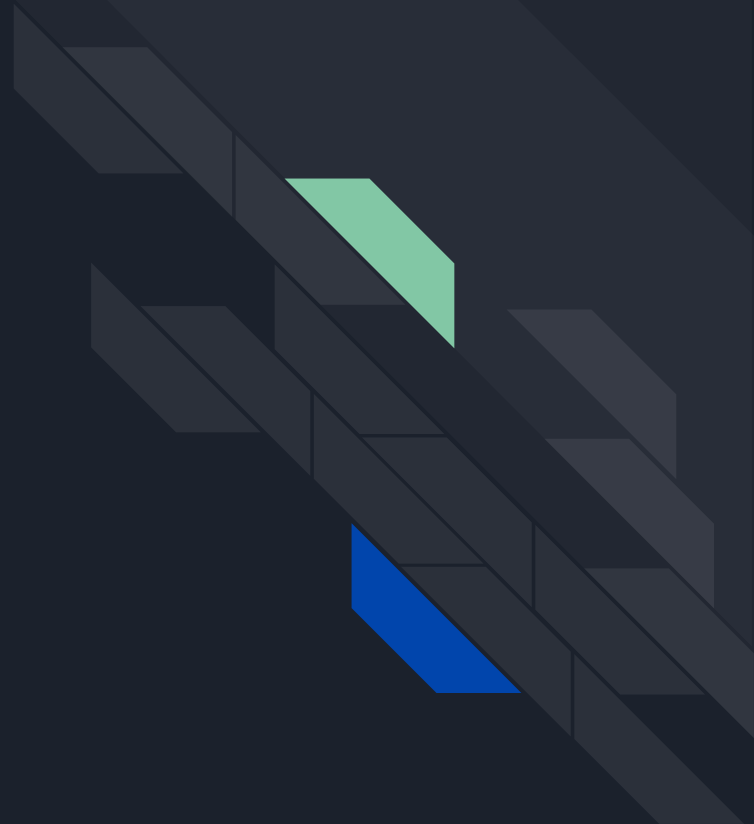
# SPIS TREŚCI

Wprowadzenie

Opis problemu

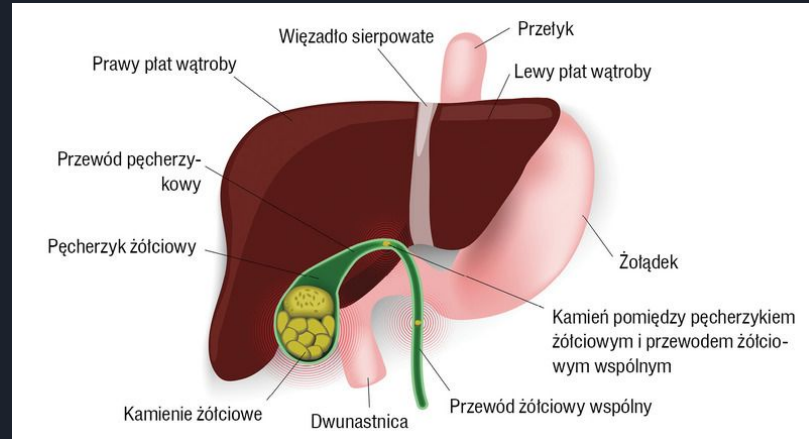
Spodziewany efekt

Planowana implementacja



# Czym jest hepatoksyczność?

**Hepatoksyczność** - cecha niektórych substancji chemicznych zawarta w lekach, których efektem ubocznym są uszkodzenia wątroby tj. martwica hepatocytów, czyli komórek wątroby i w końcowym rezultacie zaburzeń czynności wątroby, ale również mogą występować objawy uszkodzeń przewodów żółciowych wewnątrzwątrobowych, komórek śródbłonka naczyniowego a także komórek gwiaździstych.





## Opis problemu

Znalezienie odpowiedniej metody, a także dobór odpowiednich jej parametrów w celu określenia czy substancja jest hepatoksyczna.

## Spodziewany efekt

Poprzez odpowiedni wybór parametrów dla metod, na podstawie wykresów korelacji, a także zależności pomiędzy substancjami a ich składnikami model osiągnie skuteczność na poziomie 75%+, a także będzie można go zinterpretować.



# Planowana implementacja

Dane użyte do projektu: pliki .csv w postaci fingerprintów  
- podstrukturalnych Klekota&Roth.

Problemem może być to, że jest zbyt mała ilość rekordów,  
poza tym nie ma więcej problemów, danych nie trzeba  
oczyścić ze zbędnych rekordów.

Dane wejściowe fingerprinty, dane wyjściowe: liczby  
która określa "ALT" lub "pTD50"(mediana dawki  
toksycznej u 50% wszystkich gatunków)

Stos technologiczny: pandas, numpy, sklearn, matplotlib,  
python 3



# Planowana implementacja

W literaturze pojawiały się tematy z użyciem algorytmów regresyjnych uwzględniające dane w postaci fingerprintów w postaci Klekota&Ruth.

Metody, które postaram się zaimplementować:

- Liniowa regresja
- Regresja grzbietowa
- SVR



# Planowana implementacja

Do zmierzenia skuteczności posłużą podstawowe metody z biblioteki sklearn takie jak:

średni błąd bezwzględny

średni błąd kwadratowy

współczynnik determinacji  $R^2$





**Dziękuję za uwagę**