**Analiza danych medycznych**

Dane pochodzić będą z projektu dotyczącego leczenia białaczki.

Dane obejmują zarówno informacje związane z klasyczną diagnostyką (standardową procedurą diagnostyczną) jak również wyniki badań specjalistycznych (cytometria, cytogenetyka, sekwencjonowanie). Dane – w większości - są już przetworzone do postaci tabelarycznej (wiersze-pacjenci; kolumny – cechy opisujące pacjentów).

Cel analizy jest eksploracyjny, poszukujemy nowej wiedzy na temat zależności występujących w danych. Budujemy klasyfikatory i przeprowadzamy analizę przeżycia. Identyfikujemy najważniejsze zależności, szukamy najważniejszych cech.

Użyte metody analityczne muszą mieć możliwość objaśniania swoich decyzji np. drzewa decyzyjne, reguły, rough set, lasy losowe (z rankingiem cech). Możliwe jest również użycie metod nieinterpretowalnych i nałożenie na nie metod objaśniających.

Dane są całkowicie anonimowe.

**Możliwe efekty**: praca inżynierska, praca magisterska, publikacja naukowa, oprogramowanie naukowe dla lekarzy jako zestaw widoków (raportów) na wyniki analizy, aplikacja webowa.

**Opiekun**: Marek Sikora (współpraca Ł. Wróbel, J. Henzel)