

Wykorzystanie metod uczenia maszynowego do wykrywania mutacji w sekwencjach DNA oraz zmian sygnałów epigenetycznych istotnych dla regulacji ekspresji genów w wybranych chorobach

dr Michał Dąbrowski
Instytut Podstaw Informatyki Polskiej Akademii Nauk
m.dabrowski@ipipan.waw.pl

Opis tematyki

Celem tego interdyscyplinarnego tematu jest odkrycie istotnych zaburzeń w regionach regulatorowych, które zmieniają poziom ekspresji genów powodujących powstanie stanu chorobotwórczego. Dzięki rozległym badaniom molekularnym w ostatnich dwóch dekadach, wykazano, że analiza regionów kodujących białka nie może dać odpowiedzi na przyczyny wielu chorób powszechnie występujących w populacji ludzi. W niniejszym temacie, proponujemy na podstawie publicznie dostępnych baz danych oraz własnych, zweryfikować, które elementy w obszarach regulatorowych są istotne w patogenezie oraz opracować sieć powiązań między nimi.

Projekt zakłada analizę danych pochodzących z Next Generation Sequencing obejmujących sekwencje DNA, transkryptom, metylom, wybrane ChIP-seq.

Wyniki badań przyczynią się do odkrycia mechanizmów patogenezy i pomogą wyjaśnić dysfunkcje sieci kontrolujących regulację ekspresji genów i powiązań epigenetycznych w tych procesach.

Wymagania

- ukończone studia drugiego stopnia z dziedzin biologiczno-, matematyczno- lub informatyczno-pokrewnych
- podstawowa wiedza z zakresu analizy danych, statystyki
- dobra znajomość języka angielskiego
- mile widziana znajomość programowania w dowolnym języku np. R, Perl, Python