

Zastosowanie algorytmów uczenia maszynowego w analizie predykcyjnej aktywności przeciwdrobnoustrojowej pochodnych imidazoliowych i wznowy raka płuca

mgr Andżelika Lorenc

Promotorzy, prof. dr hab. n. farm. Adam Bućński, dr hab. n. med. Anna Romaszko-Wojtowicz, prof. WUM

Streszczenie

Dynamiczny rozwój metod uczenia maszynowego stwarza nowe możliwości w zakresie modelowania złożonych zjawisk biologicznych i klinicznych. Celem niniejszej rozprawy doktorskiej była ocena przydatności sztucznych sieci neuronowych typu perceptron wielowarstwowy w modelowaniu problemów medycznych dwojakiej natury – aktywności przeciwdrobnoustrojowej związków imidazoliowych oraz w predykcji wznowy raka płuca na podstawie danych klinicznych. Rozprawa obejmuje trzy powiązane tematycznie publikacje, prezentujące rozwój metodologiczny od klasyfikacji binarnej, poprzez regresję wielowymiarową, aż do modelowania klinicznego.

W pierwszej z prezentowanych prac opracowano model perceptronu wielowarstwowego, wspartego metodą drzew decyzyjnych oraz analizą składowych głównych, celem klasyfikacji aktywności przeciwdrobnoustrojowej 140 nowo syntetyzowanych pochodnych imidazoliowych wobec *Klebsiella pneumoniae*. Na podstawie deskryptorów molekularnych opisujących cząsteczki zbudowano model umożliwiający skuteczne rozróżnienie związków aktywnych i nieaktywnych, co potwierdziły wysokie wartości miar klasyfikacyjnych.

Drugie badanie dotyczyło predykcji wartości minimalnego stężenia hamującego dla 126 pochodnych imidazoliowych wobec *Staphylococcus aureus* oraz *Candida albicans*. Zastosowano podejście hybrydowe obejmujące selekcję predyktorów z wykorzystaniem drzew decyzyjnych, redukcję wymiarowości metodą analizy głównych składowych oraz modelowanie przy użyciu perceptronu wielowarstwowego. Model umożliwił jednoczesną predykcję wartości lnMIC dla obu drobnoustrojów jednocześnie, potwierdzając skuteczność integracji algorytmów uczenia maszynowego również w złożonych problemach natury regresyjnej.

Trzecia praca obejmowała retrospektywną analizę danych klinicznych 2296 pacjentów z rakiem płuca hospitalizowanych w latach 2012-2017. Opracowany model perceptronu wielowarstwowego uczony był dwuetapowo z wykorzystaniem algorytmu wstecznej propagacji błędów oraz gradientów sprzężonych. Model osiągnął wysoką swoistość i pozwolił zidentyfikować najważniejsze czynniki predykcyjne wznowy, jednak jego czułość oceniono jako umiarkowaną.

Uzyskane w pracy wyniki potwierdzają, że sztuczne sieci neuronowe stanowią skuteczne narzędzie modelowania zarówno danych chemometrycznych, jak i danych klinicznych, mogąc wspierać projektowanie nowych związków biologicznie aktywnych oraz proces podejmowania decyzji medycznych.

Słowa kluczowe: uczenie maszynowe; sztuczne sieci neuronowe; aktywność przeciwdrobnoustrojowa; wznowa raka płuca