

Instrukcje do sprawozdania

Objaśnienie wartości zawartych w dane_leki.xlsx

Zmienna zależna:

parametr logK HSA – stała równowagowa tworzenia się kompleksu w roztworze; miara siły interakcji między reagentami; wyraża powinowactwo leku do albuminy surowicy człowieka.

Albumina (HSA, ang. human serum albumin) to jedno z dwóch głównych białek osocza odpowiedzialnych za wiązanie leków. Białka osocza są odpowiedzialne za utrzymanie równowagi kwasowo-zasadowej, prawidłowego ciśnienia osmotycznego oraz transport substancji nierozpuszczalnych w wodzie (takich jak endogenne hormony sterydowe lub kwasy tłuszczowe). Czynniki wpływające na stopień wiązania leku z białkami: stężenie leku, powinowactwo leku do białek, oddziaływania cząsteczki leku z „kieszeniami białek”. Lek związany z białkami jest nieaktywny farmakologicznie, nie przenika przez błony biologiczne i nie ulega metabolizmowi, więc zmniejszenie stopnia wiązania leku z białkami osocza skutkuje wzrostem siły działania i skróceniem czasu działania leku.

Zmienne niezależne:

logK CTAB – retencja (czas retencji = czas uwalniania) w fazie pseudostacjonarnej CTAB (bromek heksadecylotrimetyloamoniowy) z wykorzystaniem metody micelarnej chromatografii elektrokinetycznej (MEKC); micelle utworzone w CTAB mają strukturę podobną do HSA.

Deskryptory CATS – dostarczają dodatkowych informacji o strukturze cząsteczki oraz mogą dostarczyć użytecznych informacji odzwierciedlających zachowanie leku w regionie wiążącym HSA; kodują informację o częstościach par atomów, które mogą być potencjalnymi miejscami wiązania leku

CATS3D_09_AL – łączy informacje o lipofilowości i akceptorze wiązań wodorowych.

CATS3D_00_AA i CATS 3D_00_DD – ważone tylko przez dawcę wiązania wodorowego (D), siłę akceptora (A). Wpływ wiązania wodorowego jest dostrzegany jako jeden z krytycznych czynników determinujących interakcję między miejscem II HSA, a niektórymi typami ligandów – małymi, zwykle aromatycznymi kwasami karboksylowymi.

Student ma za zadanie stworzyć model MLR(*Multiple Linear Regression*) z wykorzystaniem danych wystanych przez prowadzącego(dane_leki.xlsx, dane są autoskalowane). W sprawozdaniu zawarty ma być:

- Kod w formacie .ipynb lub .py
- Wykres korelacji między zmiennymi
- Równanie modelu
- Obliczone wartości statystyk oceniające zdolności prognostyczne, jakość dopasowania oraz istotność statystyczną modelu(R^2 , RMSEC, F, Q^2_{EX} , RMSE_{EX})
- Wykres Williamsa z zaznaczoną wartością graniczną + interpretacja
- Krótką interpretację otrzymanych wyników
- Q^2_{CVLOO} , RMSE_{CVLOO} (Walidacja krzyżowa metodą Leave One Out)(Na dodatkowe 0,5 pkt na następnej weyściówce)

Sprawozdanie wraz z kodem proszę wysyłać na mail: rafal.ziniewicz@phdstud.ug.edu.pl