

Předpověd protein-ligand vazebných míst je důležitým úkolem, dovolujícím nám pochopit interakce mezi proteinem a ligandem, jejichž pochopení je nezbytné při návrhu léčiv a rozvoji některých oblastí biologie. Ačkoliv již byly vytvořeny nástroje strojového učení pro predikci vazebných míst, tak doposud se vytvořené metody zajímaly pouze o predikci ze 3D struktury proteinu, která ale není pro většinu proteinů známá. Proto se v naší práci zajímáme o předpověď ze znalosti pouhé sekvence reziduí představující protein. Srovnáváme zde možné přístupy k řešení tohoto problému. Srovnáváme reprezentaci reziduí pomocí jejich chemicko-fyzikálních vlastností s reprezentací používající metody z rozpoznávání přirozeného jazyka. Dále porovnáваме zvolené metody strojového učení. Na závěr porovnááme naše výsledky s P2Rank metodou, jakožto s nejmodernější metodou používající k předpovědi protein-ligand vazebných míst 3D strukturu.