

Cvičení 6

Pracovní list

1. Ze studijních materiálů si stáhněte soubor „mix.csv“ a vypočítejte pomocí PaDELu 2D deskriptory. Vypočítejte model pro tréninkovou sadu („mix_train.csv“) pro predikci pKa z těchto deskriptorů s automatickým výběrem deskriptorů. Model validujte pomocí testovací sady („mix_test.csv“).
2. Vytvořte náhodný pro predikci pKa z datového souboru „mix.csv“ z PaDEL deskriptorů. Detekujte odlehlé hodnoty.
3. Porovnejte výsledky z regrese a z náhodných lesů.