

Cvičenie 3.

Na danom súbore kosatcov otestujte klasifikačné funkcie zaraďujúce na základe minimálnej vzdialenosti. Rozhodnite, ktorá metrika dáva najlepší výsledky.

1. Jednotlivé prípady súboru klasifikujte podľa minimálnej Hammingovej vzdialenosti. Pre túto metriku zobrazte množinu rovnako vzdialených bodov.
2. Jednotlivé prípady súboru klasifikujte podľa minimálnej Čebyševovej vzdialenosti. Pre túto metriku zobrazte množinu rovnako vzdialených bodov.
3. Jednotlivé prípady súboru klasifikujte podľa minimálnej Mahalanobisovej vzdialenosti. Táto možnosť je vo funkcii `classify`, s ktorou sme pracovali na minulom cvičení. Určte úspešnosť tejto klasifikácie.
4. Jednotlivé prípady súboru klasifikujte na základe minimálnej vzdialenosti použitím metriky Canberra. Určte úspešnosť tejto klasifikácie.
5. Jednotlivé prípady súboru klasifikujte na základe minimálnej vzdialenosti použitím uhlovej vzdialenosti. Určte úspešnosť tejto klasifikácie.
6. Pre klasifikáciu podľa minimálnej vzdialenosti použite metódu najbližšieho suseda. Určte úspešnosť tejto klasifikácie. Môžete skúsiť klasifikovať použitím k -najbližších susedov, kde $k = 2, 3, \dots$ a tiež prostredníctvom rôznych metrik. Pretože sa jedná o metódu zhlukovania pre jednotlivé príklady zobrazte dendogram.
7. Rozhodnite, ktorá z daných klasifikácií je najúspešnejšia.

Zobrazenie deliacich čiar

1. Pre oba typy lineárnych klasifikačných funkcií zobrazte na 2D a 3D úrovni deliace čiary a roviny pre jednotlivé dvojice a trojice znakov. 2D a 3D obrázky dvojíc a trojíc znakov sme robili na prvom cvičení, tak do nich stačí zobrazíť len deliace čiary a roviny.
2. Deliace čiary zobrazte tiež pre klasifikácie podľa minimálnych pravdepodobností.