

Skupiny po 4.

1. Predspracovanie textu a základná analýza výskytu slov

Predspracujte dokumenty zvolenej textovej množiny a použite rôzne metódy redukcie príznakového priestoru, vypočítajte a vhodne vizualizujte štatistiky termov.

2. Analýza sentimentu

Predspracujte textové dokumenty, vytvorte model analýzy sentimentu a vyhodnoťte jeho presnosť. Použite vami zvolené dáta a metódu.

3. Klasifikácia dokumentov

Predspracujte textové dokumenty, vytvorte klasifikačný model a vyhodnoťte jeho presnosť. Použite vami zvolené dáta a metódu.

4. Zhľukovanie dokumentov

Predspracujte textové dokumenty, aplikujte metódu zhľukovania, vyextrahujte popis zhľukov a výsledky vhodne vizualizujte. Použite vami zvolené dáta a metódu.

5. Indexovanie a vyhľadávanie v distribuovaných vyhľadávacích strojoch

Nainštalujte zvolený distribuovaný stroj, indexujte vybranú množinu dokumentov a otestujte rôzne typy dopytov.

6. Extrahovanie kľúčových slov z dokumentov

Predspracujte textové dokumenty, aplikujte metódu extrahovania kľúčových slov a výsledky vhodne vizualizujte. Použite vami zvolené dáta a metódu.

7. Detekcia tém v dokumentoch

Predspracujte textové dokumenty, aplikujte metódu detekcie tém a výsledky vhodne vizualizujte. Použite vami zvolené dáta a metódu.

8. Odporúčacie systémy

Na vhodne zvolenej množine dát vytvorte model pre odporúčanie a vyhodnoťte jeho presnosť. Použite vami zvolenú metódu.

9. Detekcia komunít

Na vhodne zvolenej množine dát vytvorte model extrahujúci komunity používateľov a výsledky vhodne vizualizujte. Použite vami zvolenú metódu.