

Streszczenie

System rekomendacji książek

Ilość informacji, które napotykamy w erze cyfrowej, jest ogromna. Dotychczasowy problem braku odpowiednich informacji został zastąpiony przez kwestię znalezienia ich w ogromnym zasobach danych. Systemy rekomendacji są narzędziami prowadzącymi użytkowników do wartościowych przedmiotów w obliczu licznych możliwości wyboru. Ludzie często polegają na opiniach innych osób, gdy nie są w stanie podjąć pewnej decyzji. Sugestie przedstawione użytkownikowi pomagają w procesie decyzyjnym dotyczącego tego co kupić, obejrzeć lub gdzie się udać.

Celem tej pracy było opracowanie systemu rekomendacji książek. Praca składała się z trzech głównych etapów: opisywania i analizowania zbioru danych, eksperymentowania z różnymi algorytmami rekomendacji poprzez porównywanie i ocenianie ich wyników, użycie wybranych technik w aplikacji internetowej w celu zaprezentowania dwóch najczęstszych przypadków użycia metod rekomendacji.

Rekomendacje książek bazujące na przyporządkowanych przez użytkowników tagach okazały się najdokładniejszym rozwiązaniem sugerującym podobne pozycje dla konkretnej książki. Zastosowane metody były w stanie przewidzieć oceny różnych użytkowników podawane w skali 1-5 ze średnią dokładnością wynoszącą 0,6. Model bazujący na rozkładzie według wartości osobliwej (Singular Value Decomposition) osiągnął najlepsze wyniki biorąc pod uwagę wszystkie miary. Opracowany system pracy wspierał odtwarzalność uzyskanych wyników.

W opracowanym systemie rekomendacji jakość danych okazała się znacznie ważniejsza niż jej ilość, szczególnie w metodach opartych na treści. W podejściu korzystającym z filtrowania społecznościowego (collaborative filtering) optymalizacja parametrów była niezbędna do późniejszego porównania końcowych wyników. Dodatkowo, wyraźnie widoczna była potrzeba stosowania różnych metryk do pełnej oceny wydajności modeli.

Słowa kluczowe: system rekomendacji, odtwarzalność, przetwarzanie języka naturalnego, filtrowanie społecznościowe, podobieństwo bazujące na zawartości