

Abstract

Recommendation System for Books

The amount of information we encounter in the Digital Era is enormous. The past problem of not having enough information transformed into the issue of finding it in a vast ocean of data. Recommendation systems are tools to navigate users to valuable items through the abundance of choices. People often rely on opinions of others when they find themselves unable to make an informed decision. Suggestions presented to the user aim to help them in the decision-making process of what to buy, watch or where to go.

The objective of this thesis was to develop a recommendation system for books. The work consisted of three main stages: describing and analyzing the dataset, experimenting with different recommendation algorithms by comparing and evaluating their outcomes, implementing selected techniques in a web application to showcase the two most common use-cases of recommender engines.

Recommending books using user-assigned tags proved to be the most accurate solution to suggesting similar items for a specific book. Applied methods were able to predict ratings (on the 1-5 scale) of various users with an average of 0.6 point precision. The Singular Value Decomposition model delivered the best results in regard to the absolute values and generalization potential. The workflow devised during the whole process supported reproducibility of the results.

In the developed recommendation system the quality of data played a more significant role than its quantity, especially in the content-based methods. In the collaborative-filtering approach the optimization of parameters was an essential step preceding the comparison of final results. Additionally, the need for various evaluation metrics to fully assess models' performance was observed.

Keywords: recommendation system, reproducibility, natural language processing, collaborative filtering, content-based similarity