

Movie Recommendation System Using Genre Correlation

Long Trinh - ltrinh19@earlham.edu

Department of Computer Science, Earlham College



1. Introduction

Recommendation systems are important emerging research technologies in computer science. They can be used to help customers find products they want. Recommendation systems use past users' preferences to predict future users' preferences and make suggestions to users. Recommendation system algorithms have been applied to movies to build movie recommendation system algorithms. In order to improve the accuracy when recommending movies to users, I design a movie recommendation system using genre correlation that combines with content-based method and uses the computation of recommendation points.

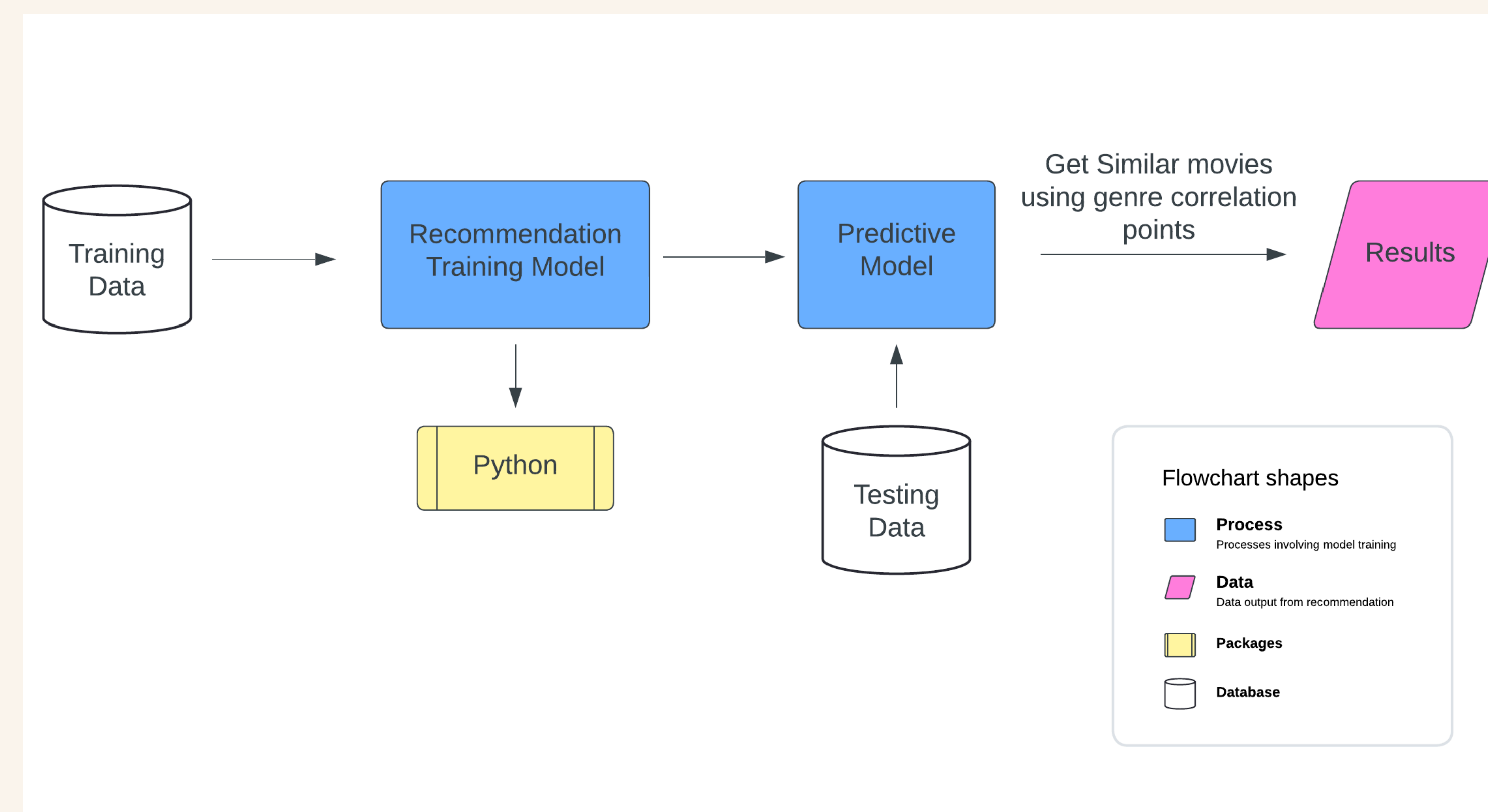
2. Design

An algorithm of using genre correlation in a content-based system is built. The prediction that the algorithm builds will be compared to the one that the system produces when treating the dataset and evaluated in terms of prediction accuracy.

3. Dataset

The MovieLens 100K dataset is a commonly used dataset in research for developing and evaluating movie recommendation systems. It contains information about movies, users, and ratings, making it suitable for training and testing recommendation algorithms. The dataset includes 100,000 ratings on a scale of 1-5 from 943 users on 1682 movies, with each user rating at least 20 movies.

4. Methods



The movie dataset is typically divided into two sets: a training dataset and a test dataset. Initially, we use the movie training dataset to develop a model that predicts the movies a user is likely to watch. This model is trained based on various features such as movie genres, ratings, and user preferences.

Once the model is trained, we apply it to the test dataset to predict the movies that users in the test dataset are likely to enjoy. This prediction is based on the patterns and insights learned from the training dataset. We then compare these predictions with the actual movies that the users in the test dataset liked.

This process helps evaluate the effectiveness of the model in accurately predicting user preferences. By comparing the predicted movies with the actual preferences of the users, we can assess the performance of the model and make any necessary adjustments to improve its accuracy.

5. Experiment and evaluation

The project will evaluate the recommendation system based on its accuracy, measured by the percentage of correctly recommended movies. It will use a content-based genre correlation algorithm to analyze movie genres and make predictions based on user preferences. This evaluation will help assess the algorithm's effectiveness and guide improvements for better user satisfaction.

6. Future work

Enhancing the recommendation system by incorporating additional movie features like director, actors, release year, and user demographics for improved accuracy. Implementing advanced algorithms such as collaborative filtering or matrix factorization could further enhance performance. Developing mechanisms for user feedback and contextual information could lead to more personalized recommendations. Improving scalability, efficiency, and exploring new evaluation metrics could also be valuable.

7. Acknowledgement

I would like to thank Dr. David Barbella and Dr. Charlie Peck, professor of Computer Science at Earlham College for support and detailed feedback on my project topic and research proposal.