Roamio Recommendation system

Daniel Flint
Ara institute of Canterbury
daniel.flint385@outlook.com

Angus Gieseg
Pogo Tourism Ltd
ag@amio.kiwi

Amit Sarkar - Supervisor
Ara institute of Canterbury
amit.sarkar@ara.ac.nz

ABSTRACT

This paper explains the functionality and needs of the project completed during mu capstone project. More details and an elaboration of the functionality of the systems created will be explored and their importance to the company will be explained. This was a python project that utilized machine learning techniques in order to create a working recommendation engine which can be incorporated into the new tourism system due to be released soon.

INTRODUCTION

The purpose of the system created it to use a new user's data of which type of activities they like to connect them a range of attractions which are related to their personal choices.

The relevance of each attraction is decided by the model based of the user's answers, activities the user has rated and other users with similar interests.

BACKGROUND

Roamio currently as a recommendation model that is item based so this project was designed to create an updated version which is capable of handling user to user connectivity, multi labelled processing and collaborative filtering.

The Roamio system has users answer a range of questions to help identify what kind of activities that user may like then using a machine learning system to identify other users will similarly interests, the original user is given a range of activities using this information. This approach is taken to make unique recommendations that are personalized to each user while also allowing feedback to promote the best experiences.

PROCESS

To make personalized recommendations multiple machine learning techniques were utilized to work together as one working system. The final system had two models working together to successfully make recommendations with a third model working to classify the different activities

First when a new user answer's any of the questions the system will start learning the users likes and dislikes. This is used to find other users with similar interests relating to the original users by implementing a K-Nearest Neighbour method which finds a set number of users with similar interests.

Once the similar users have been identified the system uses collaborative filtering to filter out all unrelated data to the original user so that a method call Matrix factorization can be done.

This method uses the review scores from all the related users to predict the potential rating the main user would give to an activity that they have not reviewed. Positive predictions make up the results that are returned to the user.

Finally, a TF-IDF model is used to help classify all the activities available so that the other models can successfully identify which activity to recommend.

With all the models working together the system can make full recommendations that over time will become more focused and accurate as the database of users grows.
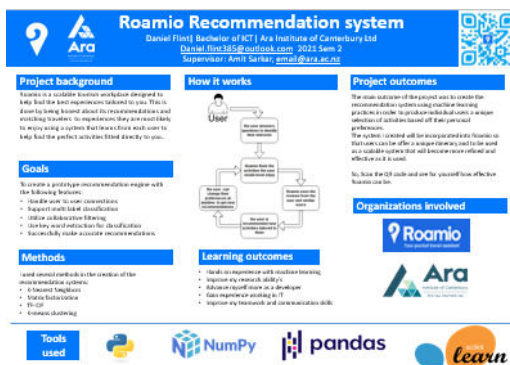
THE OUTCOME

The final product completed the project goals and is a flexible and scalable system. It made accurate recommendations when using the test data and can be further improved with more user information, the scalability of the system is achieved through the use of data frames created using the Pandas library and has made any expansion easy. The system successfully made accurate recommendation and showed to be efficient in handling data by filtering out all irrelevant data.

CONCLUSION

The final model was a success, it made accurate recommendations and is a solid improvement when compared to the current systems. it's a model that will be effective as the system grows and worked well following the goals and ambitions of the company.

I'm happy with overall outcome of the project and working on it was a great way to learn new practices and get new experience as an IT professional