Project Proposal – Book recommendations system

Team members:

Faiz Khwaja: Will be working with data preprocessing and building different ML models.

Lasya Manthripragada: Will be working on evaluating the ML models, and building the UI, and displaying results.

Project description:

We are building a recommendation system for books. Ideally, the input is a book that the user already likes, and the system should give a list of books that the user might like. This kind of system is used by kindle, google e-books etc., The system is trained on online dataset [1]. The dataset has books with very high rating, but the book has been rated inly once or twice. So, books with at least 50r ratings will be considered. Users who rated at least 200 books are out target, since we can rely on them.

Algorithms:

We are using various approaches. Like,

1. Popularity Based recommendation
2. User-item collaborative filtering recommendation
3. Correlation based recommendation
4. Nearest neighbor-based recommendation
5. Content based recommendation

We will compare all the approaches and will pick the best one.

Dataset:

The dataset [1] has been taken from IIF website. It consists of three csv files- Books, Users, and Book-ratings. We will be dropping the image columns in Books file. We will be dropping the null values, even without the null values the data is enough to train models on it. The location in Users tables has three values integrated, we will split the string using relevant delimiter.

Libraries and Tools:

We will be using Sklearn, numpy, seaborn, matplotlib, pandas, Keras, PowerBI, Django.

Results:

Our final deliverable would be a flask application, where a user can give the book that they like, and we will give books that they might like based on the input. If time permits, we will show the results, and analysis on PowerBI. We will be evaluating the models using the metrics learned in class.

References:

1. http://www2.informatik.uni-freiburg.de/~cziegler/BX/