Book Recommendation System
Capstone Project Summary

Instructions:
i) Please fill in all the required information.
ii) Avoid grammatical errors.

Team Member's Name, Email and Contribution:
NAME – Sourabh Pramanik Email – sourabhpramanik1998@gmail.com Ph No. - 9546980373
Please paste the GitHub Repo link.
Github Link:- https://github.com/PSourabh97/Book-Recommendation-System
Please write a short summary of your Capstone project and its components. Describe the problem statement, your approaches and your conclusions. (200-400 words) **Problem Statement-** In today's world there are lots of books present. Users read books and give ratings and reviews according to their interest and experience. In this project we will create a recommendation system which will provide genuine suggestion of the books to the user. **Data Description** In this project we have 3 datasets- 1. Books 2. Users 3. Ratings

Introduction-

Recommendation system is defined as the computer program that helps the user determine goods and content by predict the users rating of each item and presentation them the substance that they would rate highly. The recommendation system is containing three types that are: collaborative filtering, content based filtering and hybrid filtering.

Objective –

Building a Book recommendation system which can suggest books matching with user's interest.

Steps involved:

- Exploratory Data Analysis
  After reading the data we will perform Exploratory Data analysis.
- Null values Treatment
  Only 'Rating' dataset contain some null values in 'Age' column.
- Visualization
  We have done some visualization of Rating counts and Age group of users.
- Model
  We will use the Nearest Neighbour algorithm which is the similar as K nearest which is used for clustering based on Euclidian distance In this project we are using only the records of US and Canada people .Using only those books whose rating greater than 50.

Modelling-

Here we are using Collaborative Filtering Technique. We have prepared our dataset for modelling. we are using the Nearest Neighbours algorithm which is the similar as K nearest which is used for clustering based on Euclidian distance.

Nearest Neighbor Algorithm-

Nearest Neighbours is a simple algorithm widely used in predictive analysis to cluster data by assigning an item to a cluster by determining what other items are most similar to it. A typical use of the Nearest Neighbours algorithm follows these steps: Derive a similarity matrix from the items in the dataset.

Euclidian Distance –

The Euclidean distance between two points in Euclidean space is the length of a line segment between the two points.

Here we are specifying an algorithm which is brute means find the distance of every point to every other point. And we will specify "metric=cosine" so that the algorithm will calculate the cosine similarity between rating vectors. Finally, we are fitting the model.

Cosine Similarity-

Cosine similarity is a metric used to measure how similar the documents are irrespective of their size. Mathematically, it measures the cosine of the angle between two vectors projected in a multi-dimensional space. The cosine similarity is advantageous because even if the two similar documents are far apart by the Euclidean distance (due to the size of the document), chances are they may still be oriented closer together .The smaller the angle, higher the cosine similarity.

Conclusion -

In this project we have made the Recommendation system which use Collaborative Filtering .The main goal was speed of recommendation .Experimental result shows that the proposed method shows relevant Recommendations.

The proposed methods can be applied in other domain also like Movies, music, Videos etc.