**Capstone Project Submission**

**Instructions:**

i) Please fill in all the required information.

ii) Avoid grammatical errors.

|  |
| --- |
| **Team Member’s Name, Email and Contribution:** |
| * Manoj Patil M   Email- [smmanoj208@gmail.com](mailto:smmanoj208@gmail.com)   * Data inspection * Data Cleaning * EDA on all three datasets * Checked distribution and handling missing values * Recommendation for New Users (Cold Start) on basis of weighted average * Collaborative filtering and evaluation * Memory-based * Content-Baes Filtering with evaluation * Analyzing the results of model based Recommendation system Content Based Filtering Recommendation * Conclusion with improvements points * Project summary * Saksham Tripathi   Email-[saksham757474@gmail.com](file:///I:\python\alma%20better\capstone%20project\supervised%20learning%20(regression)\saksham757474@gmail.com)   * Data inspection * Data cleaning * EDA on all three datasets * Recommendation for New Users (Cold Start)   1) Popularity based Recommendation for new users  2) Recommendation for New Users (on basis of weighted average   * Content-Based Filtering (for title, authors, publishers, **Tags**and**books purchase history**) * Conclusion with improvements points * Gulzar   [Email-gulzarkhan9980@gmail.com](mailto:Email-gulzarkhan9980@gmail.com)   * Data inspection * Data Cleaning * EDA on all three datasets * Checked distribution of numerical features * Visualization for top books and authors (with weighted average rating) * Making system of recommendation for new users (for cold start problem) * Collaborative filtering * Model-based (with svds) * Memory-based * Item-Item based (two method) * User-Item based (with evaluation) * Content-Based Filtering (for title, authors, publishers, **Tags**and**books-purchase history**)   + - Conclusion with improvements points  |  |  | | --- | --- | |  |  |  * Bindu Kovvada   Email- [bindukovvada187@gmail.com](file:///I:\python\alma%20better\capstone%20project\supervised%20learning%20(regression)\bindukovvada187@gmail.com)   * Data Inspection Exploratory Data Analysis on the below:          1) Users Dataset   2) Books Dataset  3) Ratings Dataset   * Merging all the Datasets Recommendation for New Users (Cold Start)   1) Popularity based Recommendation for new users  2) Recommendation for New Users ( on basis of weighted average)   * Model Based Collaborative Filtering Recommendation using SVD and NMF Models * Optimization of SVD Algorithm * Analyzing the results of model based Recommendation system Content Based Filtering Recommendation * Technical Documentation * Deepak Kumar Gautam   Email- [deepakpracheta@gmail.com](file:///I:\python\alma%20better\capstone%20project\supervised%20learning%20(regression)\deepakpracheta@gmail.com)   * Data inspection * Data Cleaning * EDA on all three datasets * Checked distribution of numerical features * Visualization for top books and  authors(with weighted average rating) * Making system of recommendation  for new users(for cold start problem) * Collaborative filtering * Content-Based Filtering(for title, authors, publishers, **Tags**and**books purchase history**) * Made PowerPoint Presentation * Conclusion with improvements points |
| **Please paste the GitHub Repo link.** |
| Github Link:- [github\_link](https://github.com/Manojpatil123/Capstone-project-UnSupervised_machinelearning_Book-recommendation) |
| **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:**  During the last few decades, with the rise of Youtube, Amazon, Netflix, and many other such web services, recommender systems have taken more and more place in our lives. From e-commerce (suggest to buyers articles that could interest them) to online advertisement (suggest to users the right contents, matching their preferences), recommender systems are today unavoidable in our daily online journeys.  In a very general way, recommender systems are algorithms aimed at suggesting relevant items to users (items being movies to watch, text to read, products to buy, or anything else depending on industries). Recommender systems are really critical in some industries as they can generate a huge amount of income when they are efficient or also be a way to stand out significantly from competitors. The main objective is to create a book recommendation system for users.  **Approach:**   * First, we load data set into Panda’s frame and initialize all the library which are required for doing EDA. * Then we did inspection of data on a basic level. * Then we did data cleaning by removing null values, duplicate values and outliers. * Then we used the matplotlib and seaborn to do Exploratory Data Analysis on sample data by plotting different graphs like count plot, pie chart, lmplot, bar plot, boxplot, subplot and heat map from this we got useful insights and correlation between target column and other features * Recommendation for New Users (Cold Start) on basis of weighted average * Collaborative filtering and evaluation * Memory-based * Content-Baes Filtering with evaluation * Analyzing the results of model based Recommendation system Content Based Filtering Recommendation * Conclusion with improvements points   **Conclusion:**  Alas! We have reached the end of our notebook.  For this project our client is an online book selling firm. They now need assistance in developing a model to recommend another books on the basis of customer purchase-history and other information which are given in the datasets.  Building a model to recommend another books is extremely beneficial to the company because it can increase their sales via recommend relevant books to their customers and optimise its business model and revenue accordingly.   * For modelling, it was observed that for **model based** collaborative filtering SVD technique worked way better than NMF with lower Mean Absolute Error (MAE) . * Amongst the memory based approach, **item-item CF performed better** than **user-item CF** because of lower computation. * Content-based recommendation on the basis of **Tags** are also doing good in terms of results. * Customers of age between 20 to 30 are more likely to buy books. * Customers who are in USA are more likely to buy books than others. * Our overall top selling authors are Agatha Cristie, William Shakespeare and Stephen King. * If we look at the ratings distribution, most of the books have high ratings with maximum books being rated 8. Ratings below 5 are few in number. * Our overall top selling publishers are Harlequin, Silhouette and Pocket. * Our overall top selling books are The Lovely Bones: A Novel, Wild Animus and The Da Vinci Code, The Red Tent (Bestselling Backlist). . |