

CSP-571 Project Plan and Detail

Members

- Rajeswari Depala(A20526535)
- Priyanka Basani(A20513566)
- Naga Durga Dhanalakshmi Nallaka(A20525715)
- Tarun Sai Varanasi(A20526965)
- Veera Manohar Reddy Alavalapati(A20526070)

TASK NAME	DUE DATE	ASSIGNED TO
Data Collection	12-02-2023	
MovieLens Dataset : GroupLens Website	12-02-2023	Rajeswari
Data Preparation	01-03-2023	All
Describing the data	12-02-2023	Rajeswari
Checking the missing values, duplicate entries	27-02-2023	Dhanalakshmi
Removing the extraneous data and outliers	27-02-2023	Dhana Lakshmi
Fill all the missing values to NA	27-02-2023	Dhana Lakshmi
Transform the data by merging different datasets	01-03-2023	Veera Manohar
Split the dataset into training and testing sets	01-03-2023	Veera Manohar
Data Exploration / Visualization	25-03-2023	All
Visualization of individual features	04-03-2023	Tarun Sai
Distribution of the number of ratings by users	04-03-2023	Tarun Sai
Distribution of the number of ratings by movies	04-03-2023	Tarun Sai
Most viewers tended to award	17-03-2023	Rajeswari
Convert the timestamp to rating year	17-03-2023	Rajeswari
Sum of ratings with tree maps	22-03-2023	Priyanka
Genre by mean rating	22-03-2023	Priyanka
Sum of movie rating er genre	22-03-2023	Priyanka
Scatter plots of movie age and ratings	25-03-2023	Veera Manohar
Determine the significance of potential Age Effects	25-03-2023	Veera Manohar
Data Modeling	07-04-2023	All
Train the model using the training dataset	01-04-2023	Priyanka
Define the RMSE/loss function for train and test dataset	02-04-2023	Priyanka
Collaborative filtering for the task	02-04-2023	Priyanka/Tarun Sai
Benchmarking model: Median model	04-04-2023	Dhana Lakshmi
Create the prediction	04-04-2023	Dhana Lakshmi
Benchmarking model with Regularization	07-04-2023	Rajeswari
Review model test and select final model	07-04-2023	Rajeswari
Evaluating the model	14-04-2023	All
Final model with validation	09-04-2023	Tarun Sai
Model features with regularization	09-04-2023	Tarun Sai
Table the train and test results against the validation set	12-04-2023	Veera Manohar
Kable function to generate a table for final results for best RMSE	14-04-2023	Veera Manohar
Project Presentation and Report	25-04-2023	All
Summary of key research and findings	17-04-2023	Rajeswari
Demonstration of data processing pipeline	17-04-2023	Priyanka
Write the conclusions and insights	19-04-2023	Tarun Sai

Writing the project Report	22-04-2023	All
Create the ultimate version of the presentation	25-04-2023	All