**EMSE 6586 & CSCI 6907**

**DBMS for Data Analytics: Final Project Proposal**

**Review2Rec: Text Review-Based Recommendation System**

**Manikanta Allanki**

**G42549699**

**Ms. Data Analytics**

**George Washington University**

**Shrishail Terni**

**G28972385**

**Ms. Data Analytics**

**George Washington University**

**Shashank Boppana**

**G33584567**

**Ms. Data Analytics**

**George Washington University**

**Project Description:**

Review2Rec is a Recommendation system Leveraging textual reviews to provide the product recommendations for the Amazon Reviews. Our Approach was to find the similarities based on the metric from vector embedding of the reviews. The Project includes doing the text Analysis on text reviews, data pre-processing, and then building the recommendation engine based on the processed text reviews.

The dataset for our research work is Kaggle Amazon Kindle-reviews data set which has users reviews at least 5 products or and a product was reviewed by at least 5 users. Where each review is transformed to a embedding vector using pre trained model/transformers, and then searching the users with similar vectors and building the recommendation engine.

**Structure of the Project:**

A diagram of a data flow

Description automatically generated with medium confidence

**Task sheet:**

|  |  |
| --- | --- |
| **Task** | **Owner** |
| Data processing/sampling | Shrishail Terni |
| Summarizing/ processing text | Shrishail Terni |
| Creating the embedding and vector database | Manikanta Allanki, Shrishail Terni, Shashank Boppana |
| Building Recommendation Engine(Item similarity) | Shashank Boppana |
| Building Recommendation Engine(user similarity) | Manikanta Allanki |
| Validation | Manikanta Allanki, Shrishail Terni, Shashank Boppana |

**Github**

Project Link: <https://github.com/ManikantaGWUID/Review2Rec>

|  |  |
| --- | --- |
| **Name** | **User Name** |
| Manikanta Allanki | ManikantaGWUID |
| Shrishail Terni | ShriGWU |
| Shashank Boppana | G33584567 Shashank GWU |

References:

<https://www.kaggle.com/datasets/bharadwaj6/kindle-reviews/data?select=kindle_reviews.csv>