**CMPE 258-02**

**Project Progress/Milestone**

**Deep Learning based E-Commerce Product Recommendation** By

Veena Vyshnavi Garre - 017545671 (veenavyshnavi.garre@sjsu.edu) & Viyani Sushmitha Maria Joseph - 017542473 (viyanisushmitha.mariajoseph@sjsu.edu)

**Project Goal -** To Build an E-Commerce Product Recommendation System that provides personalized product suggestions to users by predicting their preferences and future purchases based on their interactions with items.

**Targeted Outcome/Result for Goal -** Predicts the next likely purchase or top-N recommended products. Recommendation system delivers top recommendations to users, ensuring the model outperforms and achieves maximum accuracy in results.

**GitHub Code Url -**

<https://github.com/Veenasjsu/CMPE258ProductRecommendation/blob/main/Product_Recommendation.ipynb>

1. **Achievements so far**
   1. **Data Preparation**
      1. Loaded and preprocessed the Amazon Reviews 2023 dataset and metadata.
      2. Created mappings for user\_id and parent\_asin (product id) into numerical indices for compatibility with PyTorch embeddings.
   2. **Baseline Neural Collaborative Filtering (NCF) Model**
      1. Developed a Neural Collaborative Filtering (NCF) model with user/item embeddings and a multilayer perceptron (MLP) for learning user-item interaction.
   3. **Custom Dataset and DataLoader**
      1. Implemented a PyTorch dataset to handle user-item-rating triplets.
      2. Configured DataLoaders for efficient data batching and shuffling.
   4. **Training and Evaluation**
      1. Built a training loop with logging for loss and accuracy metrics.
      2. Implemented a separate evaluation function to calculate test loss and Mean Absolute Error (MAE). Since it is a regression task to predict or recommend the top-N product recommendation choose the performance metric MAE .
   5. **Recommendations and Visualization**
      1. Developed a recommendation function to predict top-K products for a user.
      2. Built a visualization module for displaying recommended items with titles and images from metadata.
2. **Current Status of Planned Modules**

| **Module** | **Status** | **Comments** |
| --- | --- | --- |
| Data Preprocessing | Fully Functional | Mappings for users and items are complete; DataFrames are prepared. |
| NCF Model Implementation | Fully Functional | Supports configurable embedding size and hidden layers. |
| Custom Dataset and DataLoader | Fully Functional | Handles user-item-rating triplets effectively. |
| Training Loop | Fully Functional | Supports logging of metrics per epoch. |
| Evaluation | Fully Functional | Calculates test loss and MAE after each epoch. |
| Recommendation Functionality | Fully Functional | Generates top-K recommendations for users based on trained embeddings. |
| Visualization | Fully Functional | Displays item images and titles but may require optimization for datasets with missing metadata.The chosen dataset has metadata. |
| Latency Computation | Yet to Implement | NA |

1. **Baseline Modules**
2. **Challenges & Plans to Outcome**

* Working on this NCF was computationally time consuming , requiring powerful hardware for training and inference.
* Overcomed using lightening AI platform paid GPU resource on hourly basis for resource requirement.
* Challenges in Interpreting Results - Metrics like MAE and loss provided limited insight into the quality of recommendations. Planned to overcome it with Personalization with more sophisticated evaluation methods. Hence planned to calculate the hit rate that would provide better insights and quality of recommendation.

**References**