Problem Statement and Goals
TTE RecSys

Yinying Huo

Table 1: Revision History

Date	Developer(s)	Change
Jan. 17 2025	Yinying Huo	First Draft
Apr. 11 2025	Yinying Huo	Revision 1

1 Problem Statement

1.1 Problem

Recommendation systems play a pivotal role in enhancing user experience by tailoring content to individual preferences. Book recommendation systems, in particular, help users discover content aligned with their tastes, enriching their reading experiences. The Two Tower embeddings (TTE) model architecture typically consists of a query tower and an item tower. In the context of book recommendations, the query tower encodes user-specific information. The item tower processes book-specific information, including title, authors, and publication details, to generate item embeddings. TTE is particularly well-suited for studying relationships between two entity types, such as readers and books. Its structure effectively models interactions between user preferences and book attributes, making it a natural choice for personalized book recommendation systems.

1.2 Inputs and Outputs

The training dataset I will be using is from https://www.kaggle.com/datasets/arashnic/book-recommendation-dataset?select=Ratings.csv. It consists of about 271k books and 340k user book ratings with some user information.

After the training, the input of the model will be the user's information, and the output will be the recommended book for this specific user.

1.3 Stakeholders

Students who like to learn about how to build a recommendation system.

1.4 Environment

This project will be able to run on multiple platforms, including macOS and Windows. It will be designed to run on personal computing devices without the need for high-performance computing resources.

2 Goals

- Develop a functional two-tower recommendation system specifically designed for book recommendations

- Create a clear and modular codebase that facilitates easy understanding and modification

- Ensure the system is lightweight and can run on personal computing devices

3 Stretch Goals

support importing custom datasets, enabling users to adapt it to different domains beyond book recommendations.

4 Challenge Level and Extras

Challenge Level: Non-resaerch project
Extras: User Manual

Appendix — Reflection

1. What pain points did you experience during this deliverable, and how did you resolve them?

   Finding a good dataset was challenging. All public datasets have very limited user information, probably due to privacy concerns.