**Project Proposal**

**Introduction**

The purpose of this proposal is to give an overview of the recommendation system for books. In general, the recommendation system is used to predict the items that may be of interest to the user according to the user's preferences and behaviors. In this proposal, a recommendation system is used to recommend books based on book reviews and book title. This will build trust between the user and the service provider. This proposal will describe the data source, size and columns, the algorithm and tools used to build this recommendation system.

**Design**

Book recommendation based on book reviews and book title using the Good Readers dataset

**Data**

This data set contains of 10 columns and 11123 rows. The fields are:

1. **bookID:** A unique Identification number for each book.
2. **Title:** The name under which the book was published.
3. **Authors:** Names of the authors of the book. Multiple authors are delimited with
4. **average\_rating:** The average rating of the book received in total.
5. **Isbn:** Another unique number to identify the book, the International Standard Book Number.
6. **isbn13:** A 13-digit ISBN to identify the book, instead of the standard 11-digit ISBN.
7. **language\_code:** Helps understand what the primary language of the book is. For instance, ENG is standard for English.
8. **num\_pages:** Number of pages the book contains.
9. **ratings\_count:** Total number of ratings the book received.
10. **text\_reviews\_count:** Total number of written text reviews the book received.

**Algorithms**

I will build a book recommendation system using classification algorithms.

**Tools**

Data processing: pandas, Numpy.

Modelling: Classification KNN.

Visualization: Seaborn, Matplotlib.