

1. INTRODUCTION

In this modern era, there is so much information available on the internet that it makes it difficult for internet users to choose the available products or services. To overcome this problem, researchers developed a recommender system to provide relevant recommendations to make it easier for users to choose the product or service [1]. Recommender systems have been widely used in various fields. One of the biggest contributions of recommender systems is in book recommendations [2]. The book recommender system functions to provide a number of book recommendations based on user preferences [3]. With the help of these recommendations, users are able to access many books with less effort [4].

Collaborative Filtering (CF) is one of the methods in recommender systems that is most commonly used today [5][6]. CF works by predicting item ratings for certain users based on items that have been previously rated by other similar users [7]. Several researchers have developed recommender systems using CF in various fields. In book recommendations, researchers [3] developed a CF-based book recommender system. Meanwhile, researchers [4] developed an improved CF-based system. Researchers [8] developed a user-based system using KNN. In e-commerce field, researcher [9] designed a product recommender system using combination of PCA K-Means with user-based CF.

However, a CF-based recommender system has several weaknesses, one of which is that the system is too focused on providing relevant recommendations for its users. Generally, recommender systems ignore less relevant popular items or new items with few ratings and only focus on items that have enough ratings to be recommended [10][11]. This causes the item recommendations given by system to be less diverse and makes the scope of recommended items very narrow. So there needs to be a recommender system that can provide diverse recommendations with high accuracy [7].

Several studies have been conducted in addressing the lack of item diversity in recommender system. Research [7] used the diversity balancing method and succeeded balancing recommendation item diversity. In Research [12], a movie recommender system was developed using Two-stage Collaborative Filtering, successfully increasing item diversity while maintaining accuracy, precision, recall, personal diversity, and aggregate diversity. In research [10], researchers successfully increased the Personal Diversity of the recommendation results in the recommender system by combining Content-based and Collaborative Filtering methods.

In this research, the authors are interested in implementing diversity balancing in a book recommendation system. Therefore, we propose a book recommendation system based on Two-stage Collaborative Filtering with the Diversity Balancing method. This approach was chosen because it can increase diversity while maintaining good accuracy of CF results. Two-stage Collaborative Filtering is a CF method that has two main stages. The first stage is the generating a list of candidate items recommended to users. In this research, we use K-Means CF because it works well in sparse data. The second stage focuses on diversity balancing of the recommendation results. Hence, the recommendations provided by the system can contain less popular items but are still relevant to user preferences.

There are 3 datasets used in this research, i.e. ratings, users, and books datasets obtained from the Book-Crossing Community. System accuracy is measured based on precision and recall. While diversity of system is measured based on Personal Diversity and Aggregate Diversity.