

CHAPTER I

INTRODUCTION

1.1 Background

Recommender systems are information filtering systems that deal with the problem of information overload by filtering vital information fragment out of large amount of dynamically generated information according to user's preferences, interest, or observed behaviour about an item using rating [1]. In Netflix they use recommender system for personalized video ranker, top-n video ranker, trending now, continue watching, video similarity, row selection and ranking, evidence, search and related work [2]. Recommender system has advantages for provider also users, it's reduce transaction cost of finding and selecting items also have proved to improve quality and decision making [1]. Netflix declare that their recommendation system can reduce effectively customer churn and saves more than \$1 billion annually [3]. Other than Netflix, many companies provide recommendation system for their product, TripAdvisor and Yelp provide recommendations for hotels and restaurants, Amazon recommend consumers products, New York Times recommends news articles and Spotify for recommend songs and personalized song [3].

There are three methods to build recommender system, content-based (CB), collaborative filtering (CF), and hybrid recommendation system [4]. Collaborative filtering (CF) is the most successful and popular approach to build recommender system [5]. Content based filtering (CB) recommend similar items to user that are

rated positively in the past, for example, if user likes movies with genre comedy CB will recommend related movies [6]. Collaborative filtering (CF) has significant role in the recommendation process because this approach recommendation for each active user is received by comparing with preferences of other users who have rated the product in similar way to the active user [7]. Hybrid filtering is a combination of filtering approach between collaborative filtering and content based [6].

Collaborative filtering is divided into two approaches, memory based collaborative filtering and model based collaborative filtering. model based is the best approach to handle data sparsity, scalability and cold start problems. Examples of model based algorithms are alternating least squares, singular value decomposition, stochastic gradient descent and svd ++ [8].

In this study, The author will compare 2 model-based algorithms with several public datasets such as Movielens, Jester and Book Crossing to find which algorithm is best for handling sparsity data and which algorithm is suitable for the three datasets.

1.2 Problems

The traditional collaborative filtering could not handle sparsity as well as model based, the problem formulation is obtained that among the 2 model based algorithms to be tested (Alternating Least Squares and Singular Value Decomposition), which one is more suitable based on accuracy to overcome sparsity which will be applied to the Movielens , Jester and Book Crossing datasets.

1.3 Problems Limitation

The authors created problem limitations to prevent the widespread scope of problems in this study, the problem limitations of this study are as follows:

1. This research emphasize more on testing
2. This research will look for the accuracy and effectiveness of two algorithms in these model-based approach Alternating Least Squares and Singular Value Decomposition dealing with sparsity.
3. The dataset will be used is open dataset namely Movielens, Jester and Book Crossing which have different characteristics.
4. Testing method will be used RMSE and MAE

1.4 Research Purpose

This study aims to compare the two algorithms in model-based ALS and SVD to find which algorithm has the best accuracy in dealing with sparsity data on the different characteristic datasets.

1.5 Research Benefits

The benefit gained in this study for researchers is, to gain new knowledge about which algorithm in model based between ALS and SVD has better accuracy to implemented in public dataset Movielens, Jester and Book Crossing. While for others, this research expected to increase the contribution to the science.

1.6 Research Methodology

These are the stages performed in this study.

1.6.1 Literature Review

To collect various references to the Recommendation System, Collaborative Filtering, Model Based Approach , ALS , and SVD method. From several journals, proceedings, books, articles, and several other references. Also collect public dataset such as Movielens, Jester and Book Crossing.

1.6.2 Problem analysis

The analysis of the problems is based on the information obtained at the previous stage in order to obtain the right method to solve the problem in this study.

1.6.3 System Design

At this stage, the system is designed using pipelines to solve the problems contained in the analysis stage.

1.6.4 Implementation

At this stage, implementation is carried out based on the analysis that has been done in the form of programs (notebooks) in accordance with the design.

1.6.5 Testing

The next stage is testing the system has been made to test the quality of the Recommendation System using ALS and SVD. The techniques used to perform the test are Mean Absolute Error (MAE) and Root Mean Squared Error (RMSE)

1.7 Systematics Writing

Overall this thesis report consists of five (5) chapters in accordance with the thesis report guidelines applicable at Universitas AMIKOM Yogyakarta, the systematics of writing this report are as follows:

CHAPTER I INTRODUCTION

Contains the background, problem formulation, problem boundaries, research aims and objectives, research benefits, research methods, and writing systematics.

CHAPTER II RELATED WORKS

This chapter contains a description of the theories used and related to this research.

CHAPTER III RESEARCH METHODOLOGY

This chapter contains the analysis of the methods used, to improve the quality of the Recommendation System

CHAPTER IV IMPLEMENTATION AND DISCUSSION

This chapter contains a discussion of the implementation of the methods used as well as the analysis and design that has been done previously and testing the results obtained.

CHAPTER V CONCLUSION

This closing chapter contains the conclusions obtained by the author through the previous chapters and also answers the problem formulations in chapter 1, as well as suggestions for further research.

REFERENCES

This section contains a list of references that have been used in writing.