Universiti Teknologi MARA

Utilization of Data Mining for Television Programme's Genre Recommendation

Che Wan Mohd Arrazee bin Che Wan Alias

Thesis submitted in fulfillment of the requirements for
Bachelor of Science (Hons) Intelligent System
Faculty of Computer and Mathematical Sciences

August 2014

ACKNOWLEDGEMENT

In the name of Allah, The most Gracious and Most Merciful.

First, I would like to gives my gratitude to our creator Allah SWT because with his permission my project is successful. All praise to Allah for giving me a commitment and strength to finish this project.

Second, I would like to take this opportunity to give highest appreciation to my project supervisor, Dr. Shuzlina binti Abdul Rahman. Without her guidance and continuous support for the project. In her supervision, many thing that I have learn about idea, commitment and a way of being a good student.

Third, a special thanks to several lecturer such as Madam Sofianita binti Abdul Wahab, Madam Azliza binti Mohd Ali and Madam Nurzeatul Hamimah who are intentionally or unintentionally teach me many things.

Lastly, I would like to give thanks to my family and my friends who always gave a full support. I hope that Allah will always give his blessing to them.

ABSTRACT

Television is an entertainment media that has entertained many people for a decade. It has gone through many revolution and as we can see now that it already entered the 4K Resolution era. Regarding to its glory in the advancement of the technology, the television is lack of something which is content. We always go through a situation when there are a thousand channels available to us but there is nothing can attract us. It is a responsibility for the television company to identify the user preferences and provide an interesting channels. From that situation, we can see the existence of recommender system to identify user likeness. This project is a genre recommender for a programme television. It uses data mining technology to build a model for a recommendation. A collaborative filtering method has been chosen to develop the system. Naïve Bayesian Technique was employed to calculate the posterior probability that can help to improve the results. After a month of system development, the system has meet the second project objective. It can do a recommendation of programme's genre that is suitable on each 77 time slot which covers all days in whole week from 6.30 pm to 11.30 am. Based on analysis several recommendation for improvement was detected. The dimension of dataset need to be widen more for an improvement of the accuracy result. Another recommendation is to improve process time used to find the occurrences for each event by changing of data structure algorithm.

TABLE OF CONTENTS

CONTENTS **PAGE**

SUPERVISOR'S APPROVAL ii

STUDENT'S DECLARATION iii

ACKNOWLEDGEMENT iv

ABSTRACT v

TABLE OF CONTENTS vi

LIST OF FIGURES ix

LIST OF TABLES x

CHAPTER ONE: INTRODUCTION

1.1 Background of Study 1

1.2 Project Overview 2

1.3 Problem Statement 2

1.4 Research Question 3

1.5 Objective of the Project 3

1.6 Scopes of the Project 3

1.7 Significance of the Project 4

1.8 Summary 4

CHAPTER TWO: LITERATURE REVIEW

2.1 Recommender System 5

2.2 Content-based filtering system 6

2.3 Collaborative filtering system 6

2.4 The Dynamic User Model 9

2.5 Bayesian Network 10

2.6 Naive Bayes Clasifier 10

2.7 Summary 11

CHAPTER 1

INTRODUCTION

This chapter focuses on the basic overview of the research background as well as the problem that triggered the researcher to conduct the study.

1.1 Background of the study

Malaysia is moving into a new phase nowadays by trying to achieve developed country. As a result, there are many changes and standardization in many sectors to make sure that this country is on the path. One of the sectors that has a significant changes is broadcasting. We can see in previous years the broadcasting sectors is improving their quality in terms of service, media and content. In media especially, they are trying to upgrade into high definition media so that the audience can have a pleasure watching their television in crystal clear high definition.

With all the changes in high definition media, there is one thing that most of the television company in Malaysia is lacking off which is they cannot controlled over flooded television programs. It is an advantage for the audience but disadvantage for Television Companies because when there are so many programs which has a potential to optimally satisfy the audience, the audience will end up watching a limited number of television programs and ignoring the other ones(Ardissono et al., 2004). From that situation the ignored television programs will have a low rating and end up being discontinued with loss of profit.