Technische Universiteit Delft
Faculty of Electrical Engineering, Mathematics and Computer Science

Netflix Challenge: Movie Rating Prediction

CSE-2525 Data Mining
Thomas Abeel, Gosia Migut

Prepared by
Yanqing Wu
Student ID 5142571
Kaggle ID yanqingwutudelft
Exchanged Computer Engineering
30 December 2019

Table of Contents

List of Figures

List of Tables

1 Introduction

The report, entitled ”Netflix Challenge: Movie Rating Prediction”, is prepared as my Challenge report for the course CSE2525-Data Mining at the Technische Universiteit Delft. The purpose of this report is to develop a recommendation system for predicting movie ratings. The goal of the recommendation system is to achieve Root Mean Square Error (RMSE) as small as possible on an unseen dataset.

1.1 Datasets

Table 1-1. The Provided Data Sets

Dataset	Features	Mean	Std	Min & Max
users	gender	0.72	0.45	0.00 1.00
	age	30.64	12.90	1.00 56.00
	profession	8.15	6.33	0.00 20.00
movies	year	1985.81	16.91	1919.00 2000.00
	title (string)	-	-	-
ratings	rating	3.58	1.12	1.00 5.00

In users - ‘gender’, ‘0’ and ‘1’ indicates female users and male users, respectively;
In movies - ‘year’, only non-zero entries are considered.

Three datasets are provided for training, as described in Table 1-1. There is a total of 910,190 ratings, which were given by 6,040 users and 3,706 movies. The rest ‘predictions.csv’ file is used for final testing, which contains only ‘userID‘ and ‘movieID’.

2 Methodology

yoyoyo

2.1 Algorithm I

heyheyhey

3 Results

Some more text.

4 Discussion

asdfasdfasdf

5 Conclusions and Future Work

From the analysis in the report body, it was concluded that...

References

Appendix A Title of First Appendix

Use the No Spacing style.

Appendix B Another Appendix

Again, use the no spacing style for appendices.