MOVIE RECOMMENDER SYSTEMS

Building a movie recommender system using collaborative filtering

Submitted in the partial fulfillment for the award of the degree of

BACHELOR OF ENGINEERING

IN

C.S.E-AI&ML

Submitted by:

Ala Sai Vardhan Reddy	**18BCS6182**
Mulla Thanuj Kumar Reddy	**18BCS6189**
Dandala Bhaghat Reddy	**18BCS6196**
Ramagiri Nivas	**18BCS6132**
Rohit Kota	**18BCS6058**
Devendra Reddy Bakka	**18BCS6175**

Under the Supervision of:

ER. Shifali Sharma

CHANDIGARH UNIVERSITY, GHARUAN, MOHALI - 140413, PUNJAB

December 6, 2020

Table of Contents

Abstract

Recently, the building of recommender systems becomes a significant research area. The recommender systems are used in a variety of areas including music, movies, books, news, search queries, and commercial products. Collaborative Filtering algorithm is one of the popular successful techniques of RS, which aims to find users closely similar to the active one in order to recommend items. In this project we aim to actively take the ratings from the users as an input for this algorithm. The more users participate to provide reviews the more effective this algorithm goes on becoming.

List of Symbols

Symbol	*Description*
Y	*Number of movies × Number of users matrix*
X	*Feature vector of the movies*
Theta	*Parameter vector of the users*
J	*Cost function*
params	*Vector where the cost function is unrolled*
X_ grad	*Gradient of feature vector of the movies*
Theta_grad	*Gradient of parameter vector of the users*

1 INTRODUCTION

1.1Problem Definition:

Providing relevant recommendations to users based on their likes and dislikes is quite a challenge. In this movie recommender systems we will make recommendations based on the user's preferences.

1.2 Project Overview/Specifications:

In this project we will be performing collaborative filtering on the MovieLens 100k Dataset from GroupLens Research. This dataset consists of ratings on a scale of 1 to 5. The dataset has nu = 943 users, and nm = 1682 movies. We will be building a cost function and the gradient for the algorithm and we will use fmincg.m to learn the parameters for collaborative filtering.

1.3 Hardware Specification:

- Processor: Intel i3 or above
- Memory: 2 GB RAM or above

1.4 Software Specification:

- Operating system: Windows 7 or above
- Octave / Matlab GUI
- Notepad

2 LITERATURE REVIEW

A series of recent studies has indicated that Recommendation System are playing vital role in various areas. It is used in many real life applications like Entertainment, E-Commerce, Services, Social Media etc. As, in entertainment it plays a crucial role in selecting our favorite song or movie, feels the taste of our choice and interests. The literature review shows that recommender systems are defined as the supporting systems which help users to find information, products, or services (such as books, movies, music, digital products, web sites, and TV programs) by aggregating and analyzing suggestions from other users, which mean reviews from various authorities, and user attributes

The most important responsibility for every recommendation system is how to make the appropriate personalized recommendation for different customers rapidly and effectively. The context of the movies is also considered while recommending. The user - user relationship as well as user - item relationship plays a role in the recommendation. Recommendations functionality displays a list of items to a user. The items are listed in the order of usefulness to the user.

This paper begins with a short review of the literature regarding the Collaborative Filtering. Collaborative Filtering is the process of filtering or evaluating items using the opinions of other people. This filtering is done by using profiles.

Benefits Of Collaborative filtering :

- Collaborative filtering Algorithms does not store huge amounts of term frequency data for each user and document, it creates user profiles and item profile.
- Collaborative filtering techniques collect and establish profiles, and determine the relationships among the data according to similarity models.
- Collaborative filtering Algorithms do not require contents to be analyzed.

Challenges Faced by Collaborative filtering:

- Increase in the number of users and items
- Similar items
- False recommendations

The performance of collaborative filtering algorithms dependence on the type of data sets. Many collaborative filtering algorithms have been designed specially to suite the characteristic of data set, different algorithms are used when there are many users than items and when there are many items when compared to user.

Collaborative filtering algorithms can be used for various goals such as measuring the accuracy of collaborative filtering or for measuring large errors that occur due to difference in actual rating and predicted rating or for measuring the degree to which recommendations cover the entire range of items. But the most important goals of recommendation algorithm is to attain high level of user satisfaction .the user must be satisfied by using recommendation system.

2.1 Literature Review Summary

Table 2.1: Literature review summary

Year and citation	Purpose of study	Intermediate representation	Granularity Level	Match detection technique	Types of clone	Type of vulnerabilities	Data set	Evaluation parameters
2010	Recommendations based on recorded information on the users' preferences.	Collaborative filtering	Genre of movies	-	E-commerce shopping sites, music steaming sites	False reviews from users.	MovieLens 100k Dataset from GroupLens Research	-

Many measures are available to measure the performance of collaborative filtering algorithms. An efficient choice has to be made to select the best possible measure for evaluating the performance of collaborative filtering algorithms.

The advantage of collaborative filtering is that recommendations for a new user are based on the preferences made by group of users who has similar taste or preferences. Collaborative filtering techniques have been proposed to decrease the processing time and the data latency. The recommendation systems can predict user behavior patterns without any knowledge of the user in advance, and accuracy can be evaluated by comparing the prediction and reality.

3 PROBLEM FORMULATION

Collaborative filtering is one of the most important applications of Machine learning. The great this about this algorithm is it can be modified and used in many different ways for various purposes. If we look at tech giants like Netflix, Amazon, E-bay the major part of their revenue is generated by this collaborative filtering algorithm. How? Most of these companies give various recommendations to the users based on their likes and dislikes and keep the users engaged on their platform. These recommendations are only possible because of the collaborative filtering algorithm.

Now let us take an example:
Suppose there are 5 films and 4 users. The ratings of the users for the movies are given below

Movie	Alice (1)	Bob (2)	Carol (3)	Dave (4)
Love at last	5	5	0	0
Romance forever	5	?	?	0
Cute puppies of love	?	4	0	?
Nonstop car chases	0	0	5	4
Swords vs. karate	0	0	5	?

So let the first three movies fall into romantic genre and the rest into action and by the looks of the table Alice and Bob seems to like romantic movies and hate action movies at the same time we have Carol and Dave who seems to like action movies but hate romantic movies.

In here we can also see some movies are not rated by the users so our job to build the recommender systems in such a way that these missing ratings should be filled based on the interests of users and recommend the movies to them based on the ratings.

4 RESEARCH OBJECTIVES

The proposed research is aimed to carry out work leading to the development of an approach for **Movie Recommender Systems**. The proposed aim will be achieved by dividing the work into following objectives.

- **General Objective**:

The general objective of this project is to provide a recommender systems that not provides recommendations but also both perfect and justifiable recommendations, providing the ability to a user, to verify the reasoning behind a recommendation.

- **Immediate Objectives:**

The Machine learning model which we have built is used for recommending movies. We have built this by using Collaborative Filtering Algorithm.
The entire code of this project is written in Octave/Matlab.
This is the same recommender system used in Netflix, Amazon prime, Hotstar, etc.

- **Ultimate Objective:**

This Machine learning model is built by using Collaborative Filtering Algorithm. This can be used in various other fields like all e-commerce websites, online book stores, Audio streaming websites and applications.

5 METHODOLOGY

Collaborative filtering is a machine learning algorithm where predictions are made based on the likes and dislikes of the similar users.

The first part of the script ex8_cofi.m will load the dataset ex8_movies.mat, providing the variables Y and R in your Octave/MATLAB environment. The matrix Y (a num movies × num users matrix) stores the ratings y (i,j) (from 1 to 5). The matrix R is an binary-valued indicator matrix, where R(i, j) = 1 if user j gave a rating to movie i, and R(i, j) = 0 otherwise. The objective of collaborative filtering is to predict movie ratings for the movies that users have not yet rated, that is, the entries with R(i, j) = 0. This will allow us to recommend the movies with the highest predicted ratings to the user.

The first part of implementation of the collaborative filtering algorithm is writing a cost funntion(without regularization).

$$J(x^{(1)},...,x^{(n_m)},\theta^{(1)},...,\theta^{(n_u)}) = \frac{1}{2}\sum_{(i,j):r(i,j)=1}((\theta^{(j)})^T x^{(i)} - y^{(i,j)})^2.$$

This is the cost function which we will be implementing.

The next step is to write the gradient of this cost function.
The gradients of this cost function is given by

$$\frac{\partial J}{\partial x_k^{(i)}} = \sum_{j:r(i,j)=1}((\theta^{(j)})^T x^{(i)} - y^{(i,j)})\theta_k^{(j)}$$

$$\frac{\partial J}{\partial \theta_k^{(j)}} = \sum_{i:r(i,j)=1}((\theta^{(j)})^T x^{(i)} - y^{(i,j)})x_k^{(i)}.$$

The implementation of our gradients will be checked.

The next step is writing a regularized cost function that performs better than the non regularized one.

$$J(x^{(1)},...,x^{(n_m)},\theta^{(1)},...,\theta^{(n_u)}) = \frac{1}{2}\sum_{(i,j):r(i,j)=1}((\theta^{(j)})^T x^{(i)} - y^{(i,j)})^2 + \left(\frac{\lambda}{2}\sum_{j=1}^{n_u}\sum_{k=1}^{n}(\theta_k^{(j)})^2\right) + \left(\frac{\lambda}{2}\sum_{i=1}^{n_m}\sum_{k=1}^{n}(x_k^{(i)})^2\right).$$

Now it is time to regularize the gradient

$$\frac{\partial J}{\partial x_k^{(i)}} = \sum_{j:r(i,j)=1}((\theta^{(j)})^T x^{(i)} - y^{(i,j)})\theta_k^{(j)} + \lambda x_k^{(i)}$$

$$\frac{\partial J}{\partial \theta_k^{(j)}} = \sum_{i:r(i,j)=1}((\theta^{(j)})^T x^{(i)} - y^{(i,j)})x_k^{(i)} + \lambda\theta_k^{(j)}.$$

All of our implementations are finished now it is time to train our model and pass the inputs.

6 TENTATIVE CHAPTER PLAN FOR THE PROPOSED WORK

CHAPTER 1: INTRODUCTION

This chapter will cover the overview of movie recommender system.Basically, recommender systems are defined as the supporting systems which help users to find content, products, or services etc.

CHAPTER 2: LITERATURE REVIEW

A movie recommender system in this paper which goes far away from just recommending movies. It achieves both perfect and justifiable recommendations, providing the ability to a user, to verify the reasoning behind a recommendation.

CHAPTER 2: BACKGROUND OF PROPOSED METHOD

This chapter will provide introduction to the concepts which are necessary to understand the proposed system.

CHAPTER 4: METHODOLOGY

Collaborative filtering technique is the most widely used technique and is able to handle scalability issue. Its provides better recommendation for the users. In this project it covers collaborative filtering approach on MovieLens 100k Dataset from GroupLens Research to provide recommendation of movies.

CHAPTER 5: EXPERIMENTAL SETUP

The basic setup required for the complete functioning of this project on a PC is a perfectly working Octave/Matlab GUI.

CHAPTER 6: RESULTS AND DISCUSSION

The result will be the recommendation of the movies based on the input provided.

CHAPTER 7: CONCLUSION AND FUTURE SCOPE

Movie Recommendation systems proved themselves to be a best solution for addressing problem of the information overload. They help in taking choices by preserving time and energy. Future work will focus on enhancement of the existing methods and algorithms used so that the recommendation systems predictions and recommendations quality can be improved. Further this collaborative filtering algorithm can also be used in various e-commerce and music streaming applications and websites.

7 REFERENCES

[1] Jian Hei, Kai Chen, Yi Zhou, “Collaborstive filtering and deep learning based recommendation systems,” in School of Engineering and Applied Science, Aston University, Birmingham, B4 7ET, UK

[2] Daniel E. Acuna, Tulakan Ruangrong and Konrad Kording**, “**Science Concierge: A Fast Content-Based Recommendation System for Scientific Publications**”**