

**A Performance Analysis on Non-Personalized**

**Recommender Systems (Interim Report)**

**Submitted by: Li Guanlong**

**Matriculation Number: U1722033H**

**Supervisor: Assoc Prof Sun Aixin**

### School of Computer Science and Engineering

**2021**

# Introduction

## Motivations

Recommender system is an effective tool to deal with information overload. It is adopted in different businesses such as video streaming, e-commerce and social media to provide their user with a better experience.

## Objectives and Scope

This project utilized dataset provided by MovieLens, which is a website where users can rate movies in a scale of one to five stars. In this project, all algorithms are implemented and evaluated against MovieLens dataset.

In this project, recommender systems are considered to have two categories, which are (i) personalized and (ii) non-personalized. This project aims to provide an analysis on non-personalized recommender systems with the MovieLens dataset. The analysis includes (i) an in-depth overview of existing commonly used algorithms which are namely MostPop, RecentPop, DecayPop, and (ii) a new algorithm which is based on the timeline of each individual movie. With the analysis on the performance of different algorithms, this project aims to show the effectiveness of this new algorithm. All the algorithms discussed in this project are popularity based. Popularity of a movie is defined as the number of user interactions received by the movie.

For evaluation of the algorithms, the number of stars that user gives in a rating was not considered as a metric. Instead, the evaluation only focused on whether the user performs the action of rating, i.e., a user-movie interaction. For example, if the system makes a prediction on movies’ popularity on day t­­0 based on the data up to day t0 – 1. Then, it recommends a movie m to user u based on the popularity predicted. In other word, the recommender predicts that user u will rate movie m on day t­0. In this case, this recommendation is considered to be a correct one as long as user u does rate movie m on day t­0, regardless of the number of starts that the user gives. R-precision is chosen to be the measurement for evaluating the performance since all the algorithms discussed in this project are essentially top-N recommender.

## Algorithm Experimented

### MostPop

MostPop makes recommendation on t0 based on the total number of interactions a movie received from user in the training set.

### RecentPop

RecentPop focus on the recent interactions by introducing a Δt such that the recommendation on t0 is based on the total number of interactions a movie received from user during time period [t0 -Δt, t0 - 1] in the training set.

### DecayPop

DecayPop introduces a decay parameter on top of RecentPop. In DecayPop, the calculation is a weighted sum of interactions, with more recent interactions carries a higher weight. For instance, when Δtis n days, and in these n days, movie m is interacted for i1, i2, i3 … in times respectively, where i1 stands for the most recent day and in stands for the least recent day. Then the recommendation is based on the value of .

### Timeline Based Algorithm

This algorithm is a new algorithm defined in this project. The timeline of a movie refers to the number of interactions it received on each day of the past time period. For example, if a movie m is rated for [7, 6, 5, 4, 3, 2, 1] during the past 7 days, where the right side indicates most recent, and left side indicates least recent. Then, this array is called the timeline of m during past 7 days. Assume the most recent day to be t0 – 1, and the recommendation is for t0. The difference between this timeline based algorithm and the other three algorithms is that, the other algorithms are summation based, such that the predicated popularity of m on t0 would be a sum or weighted sum of elements in the array. However, the timeline based approach considers this array as something continuous and focus on the trend of the timeline. A simple example of such algorithm would be to predict the popularity of m on t0 to be 0, since the numbers of interaction have a trend to decrease by 1 per day. This algorithm is still being experimented and the details of this algorithm is still subject to change.

# Data and Result

## Experiment Setup

MovieLens dataset which is used in this project contains the user-movie interaction data between January 09, 1995 and November 21, 2019. Since recent data is a better representation of actual user behaviour, only the data in the most recent 10 years were used. The data between November 01, 2009 and October 31, 2014 were used as training set, which was to train MostPop, and also effective for the first Δt time period for RecentPop and DecayPop. The data between November 01, 2014 and October 31, 2019 were used as evaluation set. The evaluation was done on a daily basis. That is to say, MostPop, RecentPop, and DecayPop algorithms were run every day to produce a top-N recommendation on that day. Based on this top-N recommendation and the actual movie rated by users, a daily R-precision was calculated for each day. Then, monthly average and yearly average were calculated in each year and month interval.

## 2.2 Algorithm Exploration Results

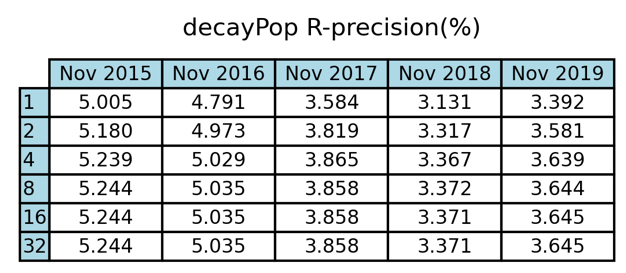
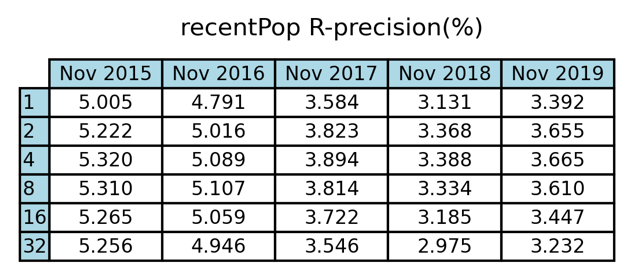
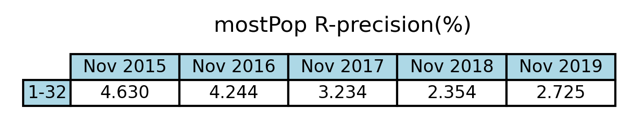
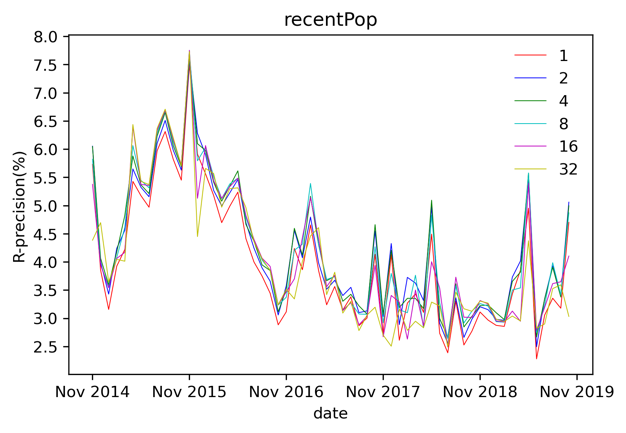
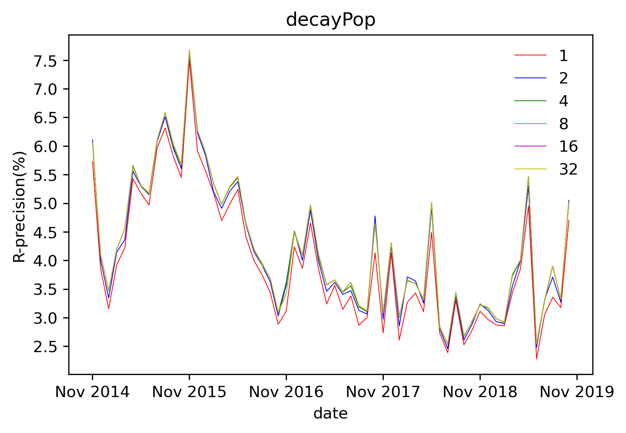
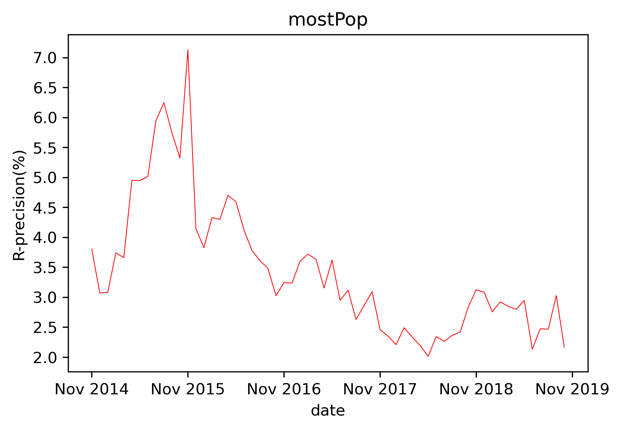


Figure 1 Monthly average R-precision for MostPop Figure 2 Monthly average R-precision for RecentPop

Figure 3 Monthly average R-precision for DecayPop. Table 1 Yearly average R-precision for MostPop

Table 2 Yearly average R-precision for RecentPop Table 3 Yearly average R-precision for DecayPop

Figure 1, 2, 3 shown above indicates the monthly average of R-precision for MostPop, RecentPop and DecayPop during evaluation. Similarly, Table 1, 2, 3 indicates the yearly average. For RecentPop and DecayPop, different values of Δt between 1 day and 32 days,were experimented. For the evaluation set which consists of the last 5 years of data, the separator between any two years is October 31 and November 01. For simplicity, in the rest of this report, year 2015 represents the time between Nov 2014 and Nov 2015, similarly for year 2016, 2017, 2018, 2019

From above figures and tables, it can be concluded that (i) The optimal Δt was 4 days for RecentPop, and 8 days for DecayPop, (ii) when using optimal Δt, RecentPop outperformed DecayPop and MostPop and (iii) R-precision calculated had a tendency to reduce as it moved towards a more recent time. In the rest of this report, Δt = 4 days is used for RecentPop, Δt = 8 days is used for DecayPop.

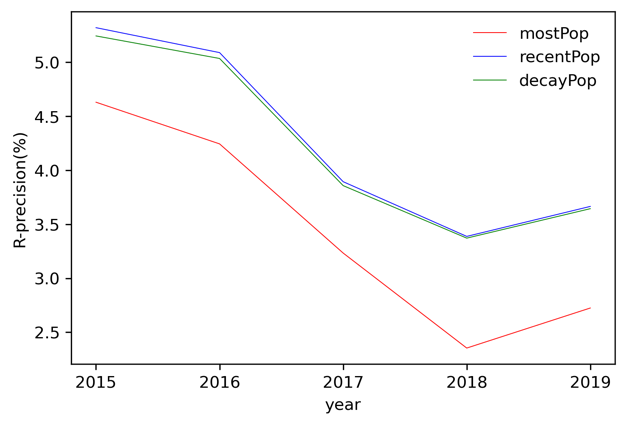
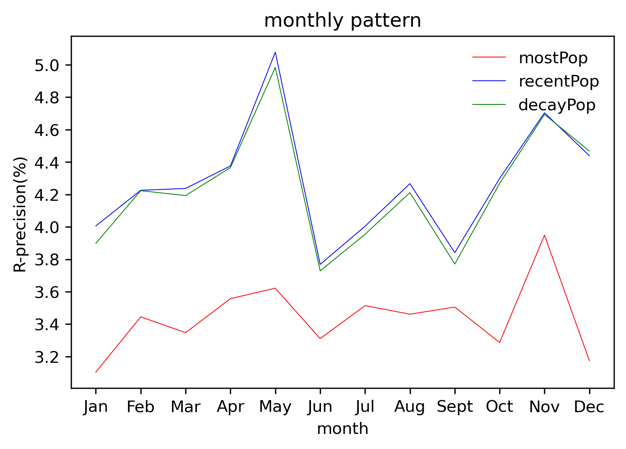
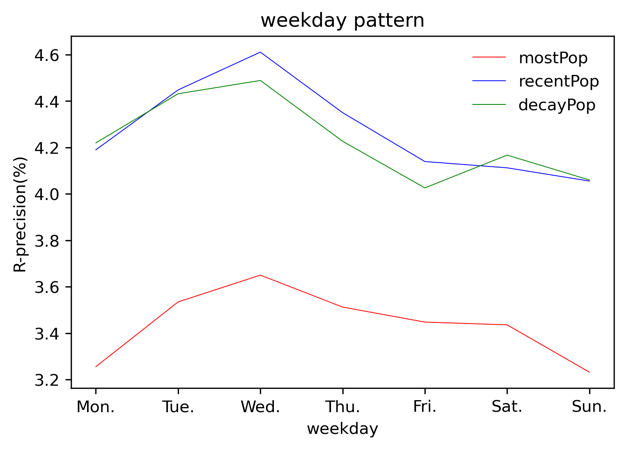


Figure 4 Yearly R-precision pattern Figure 5 Monthly R-precision pattern

Figure 6 Weekly R-precision pattern

Figure 4, 5, 6 shown above indicates the yearly, monthly and weekly pattern of R-precision calculated based on the evaluation set. Yearly, monthly and weekly pattern was calculated by taking average value of all the daily R-precision lies in each specific year, month and weekday, respectively. From above figures, it can be concluded that (i) apart from the small increase of yearly R-precision on 2019, the R-precision had a tendency to drop as it moved towards a more recent time, which confirms with the observation from Figure 1, 2, 3, (ii) MostPop, RecentPop and DecayPop generally had similar pattern, (iii) monthly R-precision pattern had two local maxima on May and November, and two minima on January and June, and (iv) weekly R-precision pattern reached maxima on Wednesday and minima on Monday and Sunday.

Bar chart

Description automatically generated with medium confidenceIcon

Description automatically generated

Figure 7 Yearly rating count Figure 8 Yearly new movie count

Icon

Description automatically generatedChart

Description automatically generated

Figure 9 Yearly active user count Figure 10 Monthly rating count

Chart, icon

Description automatically generatedChart, bar chart

Description automatically generated

Figure 11 Monthly new movie count Figure 12 Monthly active user count

Chart, icon, bar chart

Description automatically generated

Chart, bar chart

Description automatically generated

Figure 11 Weekly rating count Figure 12 Weekly new movie count

In order to find the reason behind the R-precision patterns, above figures were plotted. Figure 7, 9, 11 shows the total number of ratings, i.e., user-item interactions that falls in each year, month and weekday. Figure 8, 10, 12 shows the total number of new movies that are induced in each year, month and weekday.

From the figures, it can be concluded that (i) yearly and monthly patterns were not sensitive to the rating count or new movie count of the year and month, (ii)