

# Implementation and Analysis of Collaborative Filtering Algorithm on MovieLens Latest Small Dataset

Xinqi Zhu 5150297, Ryan Mckay

<sup>1</sup>School of Computer Science and Engineering – University of New South Wales

***Abstract.** We study the problem of implementing the recommender system on user-item rating system, i.e. user-movie rating system. We focus on the algorithm of collaborative filtering to make prediction of the users' preference on movies. In this report, we show that how our system is implemented and how to get a reasonable prediction accuracy based on experiments. Different methods are used to evaluate the performance of our recommender system.*

## 1. Introduction

Recommender systems have been widely used over the Internet and many Internet giants have their own version of recommender system to guide customers toward products they are more likely to be interested in. One of the most successful recommendation algorithms is collaborative filtering.

In this project, our goal is to implement a version of collaborative filtering algorithm on a MovieLens user-movie rating dataset and do sufficient experiments to make the system work properly and give a reasonable prediction accuracy.

Also we apply multiple evaluation methods to our system and see how the value of parameters change the performance of the system.

## 2. First Page

The first page must display the paper title, the name and address of the authors, the abstract in English and “resumo” in Portuguese (“resumos” are required only for papers written in Portuguese). The title must be centered over the whole page, in 16 point boldface font and with 12 points of space before itself. Author names must be centered in 12 point font, bold, all of them disposed in the same line, separated by commas and with 12 points of space after the title. Addresses must be centered in 12 point font, also with 12 points of space after the authors' names. E-mail addresses should be written using font Courier New, 10 point nominal size, with 6 points of space before and 6 points of space after.

The abstract and “resumo” (if is the case) must be in 12 point Times font, indented 0.8cm on both sides. The word **Abstract** and **Resumo**, should be written in boldface and must precede the text.

## 3. CD-ROMs and Printed Proceedings

In some conferences, the papers are published on CD-ROM while only the abstract is published in the printed Proceedings. In this case, authors are invited to prepare two final versions of the paper. One, complete, to be published on the CD and the other, containing only the first page, with abstract and “resumo” (for papers in Portuguese).

## 4. Sections and Paragraphs

Section titles must be in boldface, 13pt, flush left. There should be an extra 12 pt of space before each title. Section numbering is optional. The first paragraph of each section should not be indented, while the first lines of subsequent paragraphs should be indented by 1.27 cm.

### 4.1. Subsections

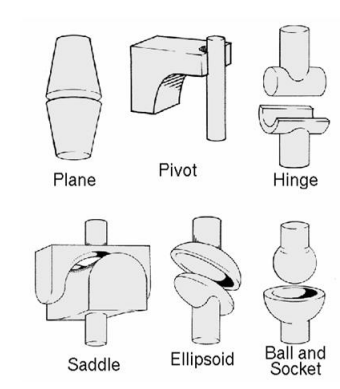
The subsection titles must be in boldface, 12pt, flush left.

## 5. Figures and Captions

Figure and table captions should be centered if less than one line (Figure 1), otherwise justified and indented by 0.8cm on both margins, as shown in Figure 2. The caption font must be Helvetica, 10 point, boldface, with 6 points of space before and after each caption.



**Figure 1. A typical figure**



**Figure 2. This figure is an example of a figure caption taking more than one line and justified considering margins mentioned in Section 5.**

In tables, try to avoid the use of colored or shaded backgrounds, and avoid thick, doubled, or unnecessary framing lines. When reporting empirical data, do not use more

decimal digits than warranted by their precision and reproducibility. Table caption must be placed before the table (see Table 1) and the font used must also be Helvetica, 10 point, boldface, with 6 points of space before and after each caption.

**Tabela 1. Variables to be considered on the evaluation of interaction techniques**

	Value 1	Value 2
Case 1	$1.0 \pm 0.1$	$1.75 \times 10^{-5} \pm 5 \times 10^{-7}$
Case 2	0.003(1)	100.0

## 6. Images

All images and illustrations should be in black-and-white, or gray tones, excepting for the papers that will be electronically available (on CD-ROMs, internet, etc.). The image resolution on paper should be about 600 dpi for black-and-white images, and 150-300 dpi for grayscale images. Do not include images with excessive resolution, as they may take hours to print, without any visible difference in the result.

## 7. References

Bibliographic references must be unambiguous and uniform. We recommend giving the author names references in brackets, e.g. [?], [?], and [?].

The references must be listed using 12 point font size, with 6 points of space before each reference. The first line of each reference should not be indented, while the subsequent should be indented by 0.5 cm.