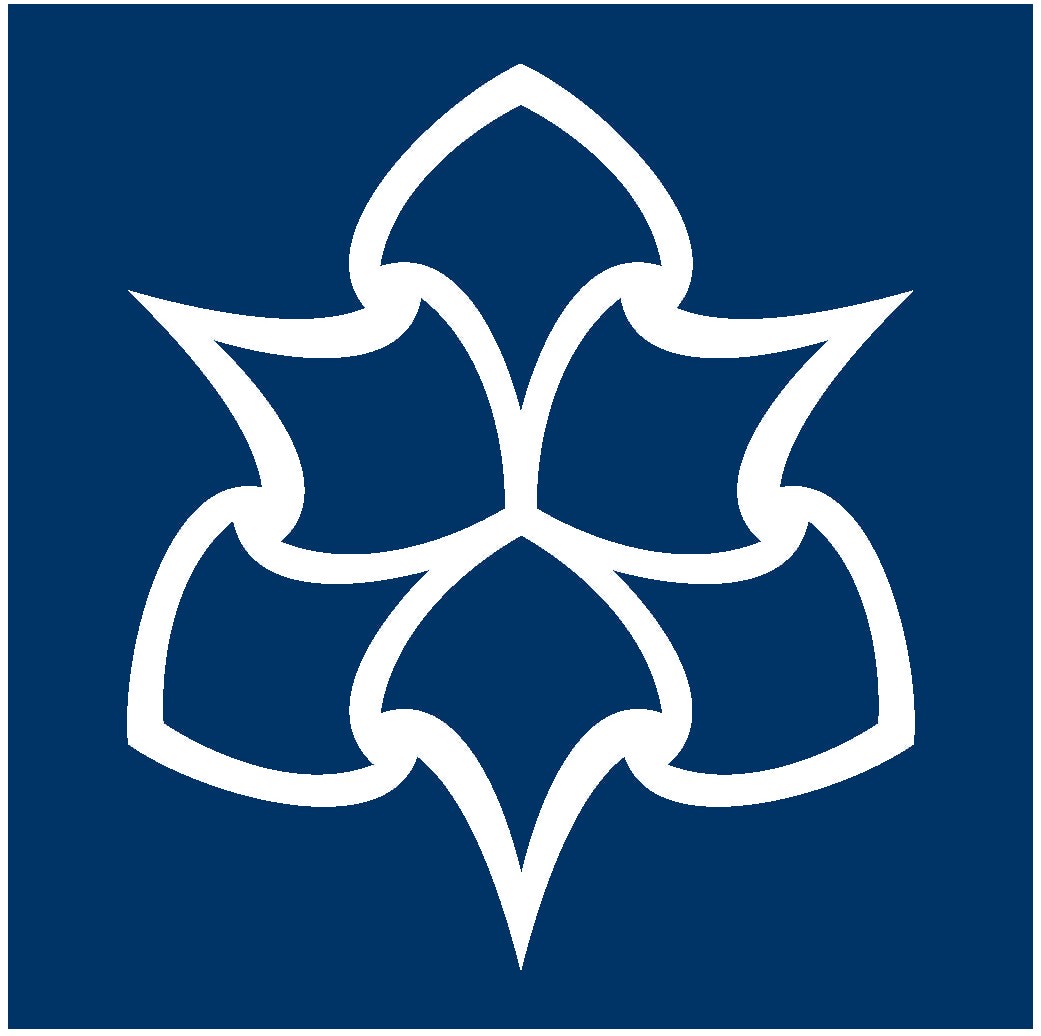
THE TITLE OF THE PROJECT

A DISSERTATION SUBMITTED TO MANCHESTER METROPOLITAN UNIVERSITY

FOR THE DEGREE OF MASTER OF SCIENCE

IN THE FACULTY OF SCIENCE AND ENGINEERING



2020

By

Swapnil Galhate

Department of Computing and Mathematics

Contents

[Abstract v](#_Toc40898999)

[Declaration vi](#_Toc40899000)

[Acknowledgements vii](#_Toc40899001)

[Abbreviations viii](#_Toc40899002)

[First Chapter - probably “Introduction” 9](#_Toc40899003)

[1.1 Referencing 9](#_Toc40899004)

[1.3 Equations 10](#_Toc40899005)

[Second Chapter - probably “Background” 4](#_Toc40899006)

[References 5](#_Toc40899007)

[The second appendix 7](#_Toc40899008)

List of Tables

1.1 Quality measurements used by CQA . . . . . . . . . . . . . . . . . . 2List of Figures

1.1 Transform coding. . . . . . . . . . . . . . . . . . . . . . . . . . . . . 2

# Abstract

The abstract is a formal description of the reason for doing this work, the methods used and results found, and the conclusions drawn. It should not be more than one page in length.

# Declaration

No part of this project has been submitted in support of an application for any other degree or qualification at this or any other institute of learning. Apart from those parts of the project containing citations to the work of others, this project is my own unaided work. This work has been carried out in accordance with the Manchester Metropolitan University research ethics procedures, and has received ethical approval number Your EthOS Number.

Signed:

Date:

# Acknowledgements

The acknowledgements thank the people who helped you do the work. It’s traditional to include your supervisor....

# Abbreviations

PCA Principal Components Analysis

LTA Three Letter Acronym

## 1. Introduction

The advent of digital platforms has revolutionized the way consumers interact with various industries, with e-commerce being one of the sectors experiencing the most significant transformation. Today, consumers have access to a virtually limitless array of products available at their fingertips. While this increased choice presents consumers with an unparalleled level of convenience, it also poses a new challenge: navigating an overwhelming number of options. Recommendation systems, by tailoring suggestions to individual users' tastes and preferences, have emerged as a critical tool in addressing this issue, facilitating personalized shopping experiences, increasing customer satisfaction, and driving business growth.

This dissertation aims to carry out an in-depth comparative analysis of two widely adopted recommendation system approaches within the context of e-commerce: Content-Based Filtering (CBF) and K-means Clustering Model-Based Collaborative Filtering (CF). Our research goal is to evaluate the performance of these methodologies under various scenarios, providing insights into their characteristics, advantages, and potential limitations in an e-commerce environment.

Content-Based Filtering (CBF) recommends items by comparing the content of the items and a user profile. For instance, in an e-commerce context, the content can include product descriptions, categories, or customer reviews. CBF is particularly effective when rich item descriptions are available, ensuring personalized recommendations that align with a user's known preferences. However, its scope can sometimes be limited as it typically recommends items similar to those the user has already purchased or interacted with.

On the other hand, Model-Based Collaborative Filtering (CF), particularly K-means Clustering, makes recommendations based on the preferences of similar users. It clusters users based on their past purchasing behaviours and recommends items liked or purchased by other users in the same cluster. This technique can leverage the collective behaviour of users to uncover and suggest items that the user has not interacted with, thereby potentially expanding the user's shopping horizons. Nevertheless, it has its own challenges, such as handling new users due to the 'cold start' problem, dealing with sparse data, and selecting the optimal number of clusters.

In this study, we will implement both CBF and K-means Clustering Model-Based CF, using real-world data from an e-commerce platform. We will evaluate their performances based on a range of metrics, including precision, recall, F1 score, and within-cluster sum of squares, considering the unique demands of an e-commerce environment. Through a thorough comparison and analysis, we hope to provide valuable insights into how these methodologies can be best applied in an e-commerce setting.

The ensuing chapters of this dissertation will present a detailed theoretical background on CBF and CF, describe our experimental setup, discuss the results, and draw conclusions based on our findings. We believe that this research will contribute to improving the personalized shopping experiences offered by e-commerce platforms and help businesses better understand and cater to their customers' preferences.1

*CHAPTER 1. FIRST CHAPTER - PROBABLY “INTRODUCTION”* 2

1.2 Figures and Tables

A figure caption should appear below each figure, and a table caption should appear below each table. Insert figures and tables after they are cited in the text; make sure that you cite them all. The figure or table, together with the caption (which should give the number of the figure or table), should be centred and referred to within the text as “Figure 1.1 shows...” or “...as shown in Table 1.1”. There should be a blank line above and below each figure or table.

Input Compressed

Frame Entropy Frame

Transform

Figure 1.1: Transform coding.

|  |  |
| --- | --- |
| Number | Measurement |
| M0 | Difference in Y values |
| M1 | Maximum of TI |
| M2 | RMS of TI |
| M3 | Range of TI values |
| M4 | RMS of SI |

Table 1.1: Quality measurements used by CQA

If preparing your report on Word, use its caption handling facility to enter figure and table headings. This will allow you to auto-number the headings and also to generate a List of Figures and a List of Tables at the beginning of your report.

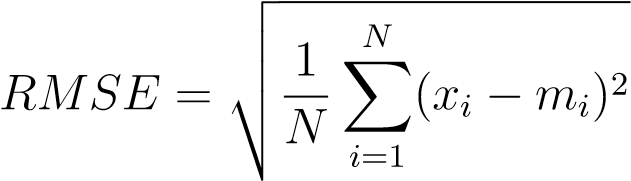
### 1.3 Equations

Number equations consecutively within each chapter (e.g. the first equation in Chapter 1 should be numbered (1.1), the first equation in Chapter 2 should be numbered (2.1) etc.). Equation numbers, within parentheses, should be positioned flush right as in (1.1). The equation should be centred and included in the sentence within your text which brackets it. There should be a blank line above and below each equation.

The following is an example of the correct use and formatting of an equation:

*CHAPTER 1. FIRST CHAPTER - PROBABLY “INTRODUCTION” 3*

Using the root mean square error,

*,* (1.1)

as a measure of accuracy, where *xi* and *mi* are respectively the *i*-th elements of the observation and reference,....

Chapter 2

## Second Chapter - probably “Background”

Citizens UK 2018

Some text here...

I’m not going to put files for all the chapters you’ll need...

## References

Chen, C., R. Jafari, and N. Kehtarnavaz (2015). “UTD-MHAD: A multimodal dataset for human action recognition utilizing a depth camera and a wearable inertial sensor”. In: *Image Processing (ICIP), 2015 IEEE International Conference on*. IEEE, pp. 168–172.

Citizens UK (2018). *Better hate crime protections for all*. http:\www.citizensuk. org/hate-crime-law-commission. Accessed: 2020-05-20.

Frith, U. (2012). “Why we need cognitive explanations of autism”. In: *The Quarterly Journal of Experimental Psychology* 65 (i*First*), pp. 1–20.

Ryoo, M. S. and L. Matthies (2013). “First-Person Activity Recognition: What Are They Doing to Me?” In: *Computer Vision and Pattern Recognition (CVPR), 2013* *IEEE Conference on*. DOI: [10.1109/CVPR.2013.352.](https://doi.org/10.1109/CVPR.2013.352)

Appendix A

The first appendix

Include whatever is necessary here - design materials perhaps?

Appendix B

## The second appendix

Include whatever is necessary here - code perhaps?