National College of Ireland

# Configuration Manual

MSc Research Project
MSc in Artificial Intelligence

Harishbabu Udatha
Student ID: X22192701

School of Computing
National College of Ireland

Supervisor: Prof Mayank Jain

**National College of Ireland**

**MSc Project Submission Sheet**

**School of Computing**

**Student Name:** Harishbabu Udatha

**Student ID:** X22192701

**Program:** MSc in Artificial Intelligence **Year: 1**

**Module:** …Research in Computing……

**Supervisor:** ……Prof Mayank Jain……

**Submission Due Date:** ……31/01/2024……

**Project Title:** Hyperparameter Optimized KNN Models for Recommendation Systems

**Word Count:** **866** **Page Count:** **7**

I hereby certify that the information contained in this (my submission) is information pertaining to research I conducted for this project. All information other than my own contribution will be fully referenced and listed in the relevant bibliography section at the rear of the project.
ALL internet material must be referenced in the bibliography section. Students are required to use the Referencing Standard specified in the report template. To use other author's written or electronic work is illegal (plagiarism) and may result in disciplinary action.

**Signature:** Harishbabu Udatha
……

**Date:** 31/01/2024
……

**PLEASE READ THE FOLLOWING INSTRUCTIONS AND CHECKLIST**

| | |
|---|---|
| Attach a completed copy of this sheet to each project (including multiple copies) | ✓ |
| **Attach a Moodle submission receipt of the online project submission,** to each project (including multiple copies). | ✓ |
| **You must ensure that you retain a HARD COPY of the project**, both for your own reference and in case a project is lost or mislaid. It is not sufficient to keep a copy on computer. | ✓ |

Assignments that are submitted to the Programme Coordinator Office must be placed into the assignment box located outside the office.

| **Office Use Only** | |
|---|---|
| Signature: | |
| Date: | |
| Penalty Applied (if applicable): | |

# Configuration Manual

Harishbabu Udatha
Student ID: X22192701

# 1 Introduction

This setup guide is part of the research paper that is part of the MSC AI in Business curriculum. This guide contains a full explanation of the procedures used to perform the study. It also describes the hardware configuration and technologies used in the system to get the results, making it possible for other researchers to replicate the study's conclusions.

# 2 System Configuration

This section discusses the system configuration required for executing the analysis.

## 2.1 Hardware Requirements

This section discusses the system configuration required for executing the analysis

- Version 10 of the Windows operating system, 64-bit
- 2.40GHz 2.50GHz Intel(R) Core(TM) i5-6300U CPU
- Memory: 8 GB
- Windows 10 Pro

## 2.2 Software

The system setup needed to carry out the analysis is covered in this section.

Visual Code for Microsoft (VS Code)

Google Collaborative

To run code, you can use Microsoft VS Code or Google Colab.

# 3 Upload Colab Notebook to Google Colab

Here are the steps to upload a Colab notebook into Google Colab:

**1. Access Google Colab:**
- - Go to [colab.research.google.com](https://colab.research.google.com/) to access Google Colab.

**2. Sign into Google:**

- Sign in with your Google account if you're not already signed in.

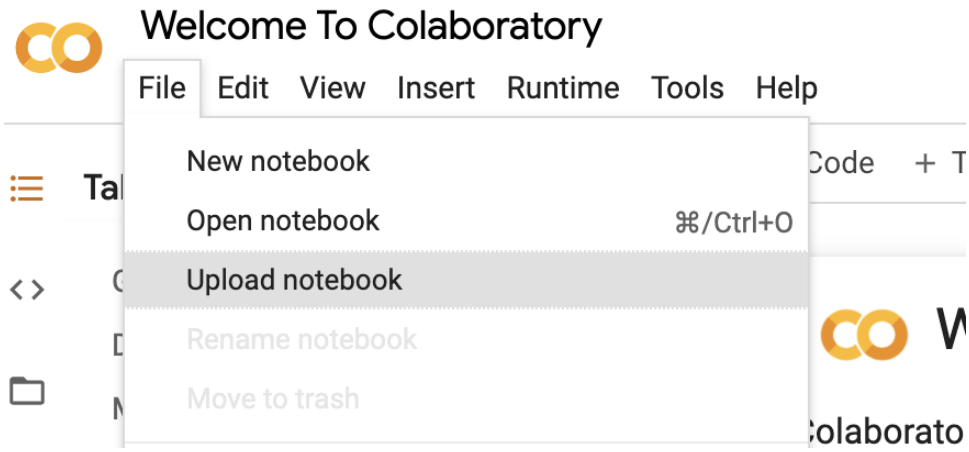
**3. Upload the Notebook:**

- Select the "File" tab from the Colab interface's top tabs.

Select the tab labelled "Upload notebook".

- Select the notebook file you wish to upload by clicking the "Choose file" button and locating it on your local computer.

- After making your selection, click "Open" or "Upload" to begin the upload.




**4. Access the Uploaded Notebook:**

- The notebook file will be shown in the "Files" sidebar after the upload is finished.

- To access the uploaded notebook in Google Colab, click the filename.

And that's it! A Colab notebook has been successfully posted to Google Colab. You can now execute cells, edit the code, work directly on it in the Colab environment, and make use of Google's processing power.

# 4 Upload Dataset

It's easy to upload datasets to Google Colab:

**1. Mount Google Drive:**

- Using Google Drive is the simplest approach to access datasets in Colab. To access your Google Drive files in Colab, mount your drive.

Please paste the dataset path into zip_file_path after uploading the dataset to your disc.

```
zip_file_path = "/content/drive/MyDrive/ml-25m.zip"
extraction_path = "/content/dataset"
```

# 5 Run Notebook

This is how to use a Colab Notebook:

**1. Runtime Settings:**
- Based on your needs, choose the hardware accelerator and runtime type.

- To change these parameters, choose "Runtime" -> "Change runtime type".

**2. Running Cells:**

To execute a single cell, pick it and press the "Play" icon or use the keyboard shortcut "Shift + Enter."

Make sure that each cell is performed in the correct order, particularly if there is interdependence amongst them.

**3. Save Notebook:**

- Use the keyboard shortcut "Ctrl + S" (Cmd + S for Mac) or "File" -> "Save" to periodically save your work.

**4. Install Libraries (if needed):**

- Install any extra libraries that your notebook might need by typing "!pip install package_name" in a code cell.

# 6 Visualizing Results

Make use of suitable visualisation libraries such as Matplotlib, Seaborn, or Plotly to visualise data, measurements, or outcomes.

Examine charts, graphs, and other material produced in order to comprehend the analysis.

- Store the notebook's outputs, such as tables, graphs, and logs, for documentation or additional analysis.

**Dataset**

| | movieId | title | genres |
|---|---|---|---|
| 0 | 1 | Toy Story (1995) | Adventure\|Animation\|Children\|Comedy\|Fantasy |
| 1 | 2 | Jumanji (1995) | Adventure\|Children\|Fantasy |
| 2 | 3 | Grumpier Old Men (1995) | Comedy\|Romance |
| 3 | 4 | Waiting to Exhale (1995) | Comedy\|Drama\|Romance |
| 4 | 5 | Father of the Bride Part II (1995) | Comedy |
| ... | ... | ... | ... |
| 62418 | 209157 | We (2018) | Drama |
| 62419 | 209159 | Window of the Soul (2001) | Documentary |
| 62420 | 209163 | Bad Poems (2018) | Comedy\|Drama |
| 62421 | 209169 | A Girl Thing (2001) | (no genres listed) |
| 62422 | 209171 | Women of Devil's Island (1962) | Action\|Adventure\|Drama |

62423 rows × 3 columns

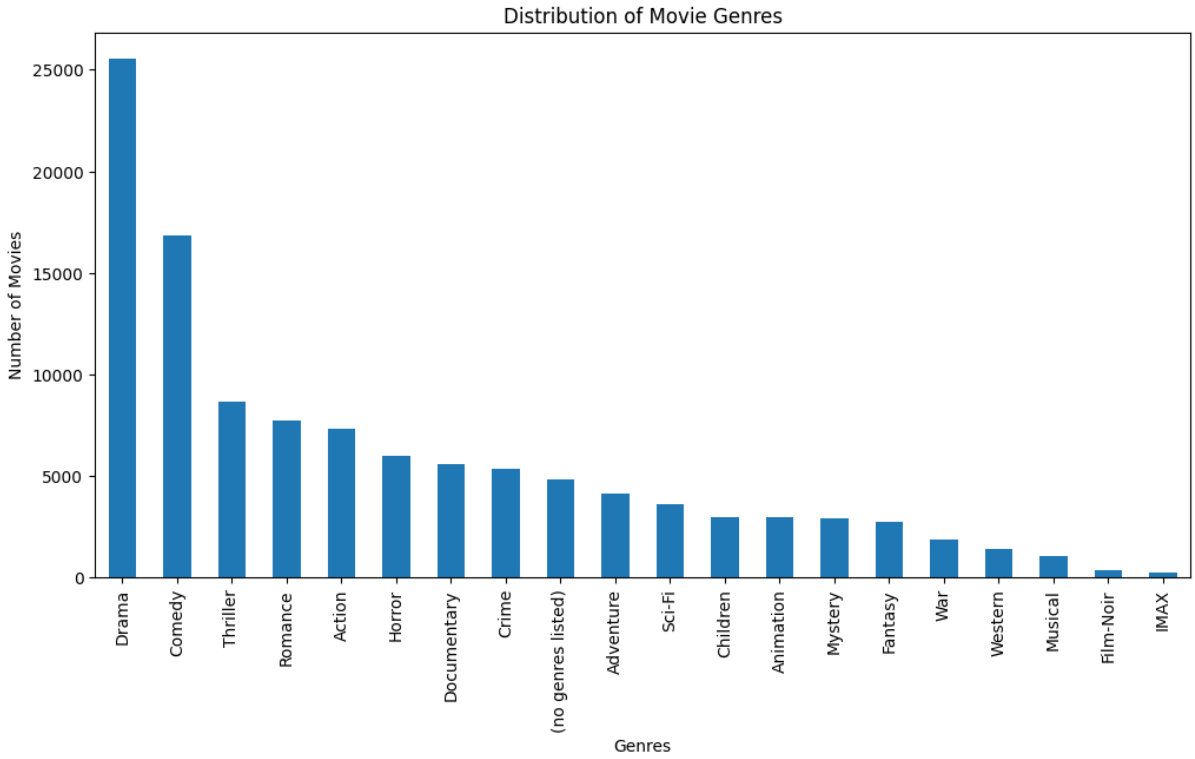
**Dataset (movies.csv)**

| | userId | movieId | rating | timestamp |
|---|---|---|---|---|
| 0 | 1 | 296 | 5.0 | 1147880044 |
| 1 | 1 | 306 | 3.5 | 1147868817 |
| 2 | 1 | 307 | 5.0 | 1147868828 |
| 3 | 1 | 665 | 5.0 | 1147878820 |
| 4 | 1 | 899 | 3.5 | 1147868510 |
| ... | ... | ... | ... | ... |
| 25000090 | 162541 | 50872 | 4.5 | 1240953372 |
| 25000091 | 162541 | 55768 | 2.5 | 1240951998 |
| 25000092 | 162541 | 56176 | 2.0 | 1240950697 |
| 25000093 | 162541 | 58559 | 4.0 | 1240953434 |
| 25000094 | 162541 | 63876 | 5.0 | 1240952515 |

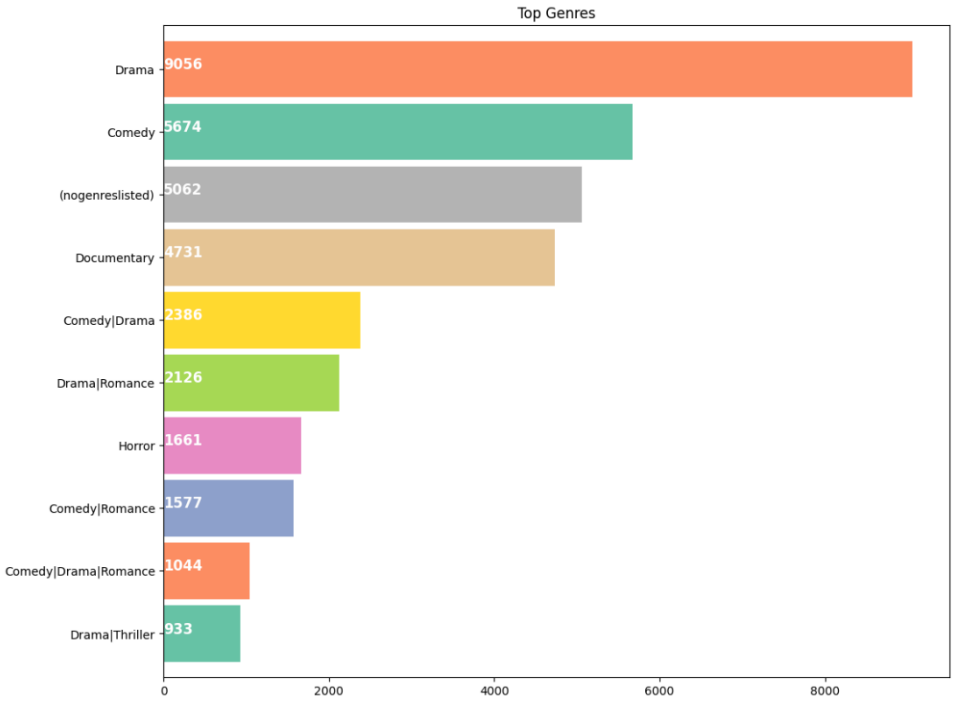
25000095 rows × 4 columns

**Dataset (rating.csv)**

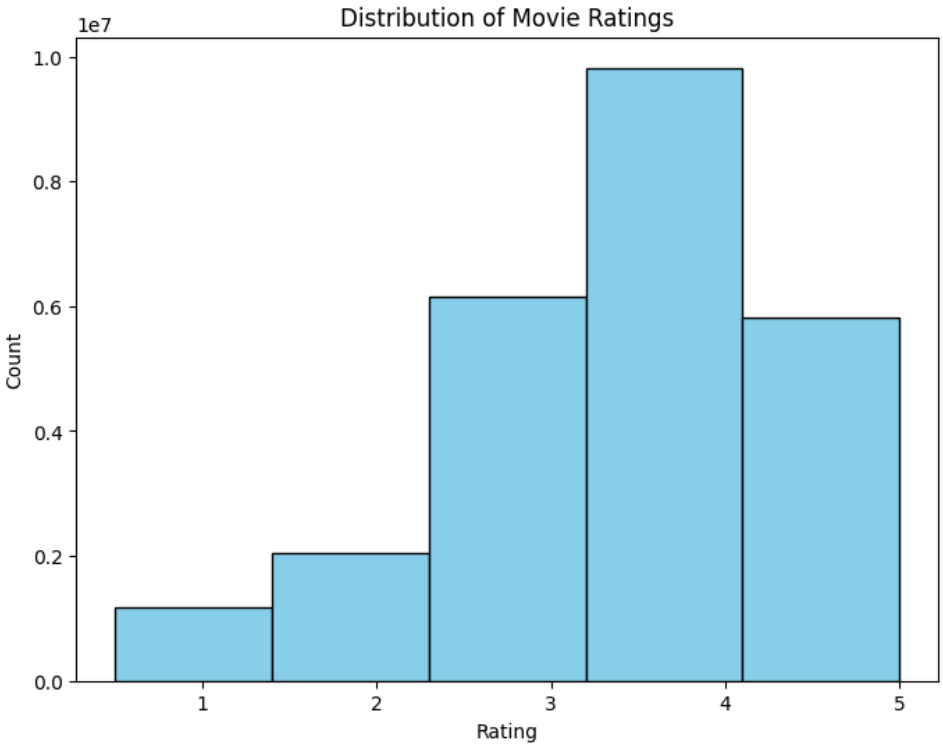
**Exploratory Data Analysis**




**Distribution of movie genres**




**Visualization of Top Genres**




**Histogram plot of movie ratings**

## Model Prediction

```
Movie: Postman, The (Postino, Il) (1994), Similarity: 1.0000
Ratings for Postman, The (Postino, Il) (1994):
      userId  rating
517    72941     5.0
1076  110006     4.0
2251    8270     5.0
2321    8105     5.0

Movie: Living Out Loud (1998), Similarity: 1.0000
No ratings found for Living Out Loud (1998).

Movie: Sleepless in Seattle (1993), Similarity: 1.0000
Ratings for Sleepless in Seattle (1993):
      userId  rating
490    71832     4.0
618     8688     4.0
1902  141063     3.0
2203  154953     5.0

Mean Squared Error: 13.2667
```

**Model Prediction of Content Base Filtering approach**

```
make_recommendation('Georgia (1995)', item_user_mat_sparse, best_model, movie_to_index, 10)
```

```
Working on Recommendation....

Viewer who watches this movie Georgia (1995) also watches the following movies.
Jumanji (1995)
Grumpier Old Men (1995)
Father of the Bride Part II (1995)
GoldenEye (1995)
American President, The (1995)
Nixon (1995)
Casino (1995)
Sense and Sensibility (1995)
Ace Ventura: When Nature Calls (1995)
Copycat (1995)
```

**Model Prediction of Collaborative filtering approach**