Logo

Description automatically generated**PERFORMANCE COMPARISON OF RECOMMENDER SYSTEMS**

**PROJECT – I REPORT**

Submitted by

|  |  |
| --- | --- |
| **REG NO** | **NAME** |
| **61072011130** | **NITHISHKUMAR S** |
| **61072011125** | **MAHAKASILINGAM A** |
|  |  |

Submitted in partial fulfilment for award of the degree

of

BACHELOR OF ENGINEERING

In

COMPUTER SCIENCE AND ENGINEERING

**Under the guidance of**

**Dr. S.Selvi M.E., Ph.D.,**

Assistant professor (Sr. Gr)

**GOVERNMENT COLLEGE OF ENGINEERING BARGUR,**

**KRISNAGIRI-635 104.**

**(Affiliated to Anna University, Accredited by NAAC with ‘B’ Grade)**

**ANNA UNIVERSITY:CHENNAI 600 025**

**DECEMBER 2022**

**ANNA UNIVERSITY: CHENNAI 600 025**

**BONAFIDE CERTIFICATE**

Certified that this project REPORT titled

“**PERFORMANCE COMPARISON OF RECOMMENDER SYSTEMS**”

is the bonafide work of NITHISHKUMAR S(61072011130), MAHAKASILINGAM A(61072011125) who carried out the mini project work under my supervision.



**Dr. S.Selvi M.E., Ph.D., Dr.J.NAFEESA BEGUM, M.E.,Ph.D.,**

**SUPERVISOR HEAD OF THE DEPARTMENT**

Assistant Professor (Sr. Gr), Professor,

Department of CSE, Department of CSE,

Govt. College of Engineering, Govt. College of Engineering,

Bargur-635 104. Bargur-635 104.

Submitted for the Mini Project Viva Voce Examination held on \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ at

GOVERNMENT COLLEGE OF ENGINEERING BARGUR, KRISHNAGIRI- 635 104.

**INTERNAL EXAMINER EXTERNAL EXAMINER**

**ACKNOWLEDGEMENT**

We express our sincere of gratitude to our Principal **Dr. R.Vijayan, M.E., PhD**, for his timely advice and his encouragement.

We express our regards to **Dr. J.NAFEESA BEGUM M.E.,Ph.D.,** **Head of Department** for her support, coordinate and for providing necessary facilities whenever needed to carry out the project.

We consider it as a great privilege to place a record or our sense of gratitude to our Internal Guide **Dr.S.SELVI PhD.,** Assistant professor, Department of Computer Science and Engineering for beginning the great inspiration and motivation to us to do the project, consistently guided us in completing the project.

At the very outset, we wish to express our sincere thanks to all those who were more involved in the completion of this project.

Our most sincere salutations go to Anna University, Chennai that gave us an opportunity to have sound base of Computer Science and Engineering.

We also express our thanks to our department faculty members, our parents and our beloved friends.

**ABSTRACT**

**Performance Comparison of Recommender Algorithms**

The recommender systems are used in many product-based companies for profiling users based on the content they generate. The data that gets generated include personalised information and user interaction data. These data are then evaluated to generate the personalised content for a user. The user may get a nice user experience when using the software application implemented with recommendation system.

The recommender system model implements four various algorithms for performance comparison. It requires large dataset for training and evaluation. After acquiring the datasets pre-processing begins. The pre-processing includes cleaning, instance selection, normalization, transformation, feature extraction. Missing values are nullified in cleaning process, instance selection involves selection of required data items. The data items are normalized produce a new instance which are transformed and a new feature is extracted. After pre-processing training of dataset begins. In training process, all features are analysed statistically and the model is evaluated. One key aspect of evaluation is to ensure that the trained model generalizes for data it was not trained on, using Cross-validation techniques. As there are different instance types, we associate them with a weight or strength, assuming that, for example, a comment in an article indicates a higher interest of the user on the item than a like, or than a simple view. This process is called data munging. Data munging of four different algorithms produces four different weights or strength values. The four different values are compared in a pictorial representation for better visualization.

The common way to assess the performance of a recommender system would be through standard metrics such as Accuracy, Precision or Recall. However, these metrics require ground truth knowledge about which recommendations are correct. The best recommendation algorithm is then chosen and is tested for appropriate user information.

|  |  |  |
| --- | --- | --- |
| **FIG**  **NO** | **FIGURES** | **PAGE**  **NO** |
| 1.1 | Traditional approach | 3 |
| 1.2 | Machine learning approach | 3 |
| 1.3 | Automatically adapting to change | 4 |
| 1.4 | Training set for supervised learning | 5 |
| 1.5 | Training set for unsupervised learning | 5 |
| 1.6 | Training set for reinforcement learning | 6 |
| 1.7 | Hybrid recommender system | 7 |
| 4.1 | System architecture | 20 |
| 4.2 | Dataflow diagram level 0 | 21 |
| 4.3 | Dataflow diagram level 1 | 22 |
| 4.4 | Dataflow diagram level 2 | 24 |
| 4.5 | Use case diagram for recommendation system | 27 |
| 4.6 | Activity diagram for recommendation system | 28 |
| 4.7 | Sequence diagram for recommendation system | 29 |
| 5.1 | Class diagram for hybrid recommender system | 32 |

**LIST OF FIGURES**

**TABLE OF CONTENTS**

**CHAPTER TITLE PAGE**

**NO NO**

**ABSTRACT** iv

**LIST OF FIGURES** v

**1 INTRODUCTION** 1

* 1. WHAT IS MACHINE LEARNING? 2
  2. WHY USE MACHINE LEARNING? 2
  3. TYPES OF MACHINE LEARNING 4
  4. RECOMMENDATION ALGORITHMS 6
  5. PROJECT DESCRIPTION 8

**2** **SYSTEM ANALYSIS** 9

2.1 LITERATURE SURVEY 10

2.1.1 Content based recommendation system 10

2.1.2 Collaborative filtering recommendation system 10

2.1.3 Hybrid recommender system: A Systematic 11

literature review

2.2 EXISTING SYSTEM 11

2.2.1 Disadvantages 11

2.3 PROPOSED SYSTEM 12

2.3.1 Advantages 12

**3** **SYSTEM REQUIREMENTS** 13

3.1 HARDWARE REQUIREMENTS 14

3.2 SOFTWARE REQUIREMENTS 14

3.3 SOFTWARE DESCRIPTION 15

3.3.1 Jupyter notebook 15

3.3.2 Pandas 15

3.3.3 Numpy 16

3.3.4 Sklearn 16

3.3.5 NLTK 17

3.3.6 SciPy 18

**4** **SYSTEM DESIGN** 19

4.1 SYSTEM ARCHITECTURE 20

4.2 DATAFLOW DIAGRAM 21

4.2.1 Level 0 – Dataflow diagram for hybrid 21

recommender system

4.2.2 Level 1 – Dataflow diagram for hybrid 22

recommender system

4.2.3 Level 2 – Dataflow diagram for hybrid 24

recommender system

4.3 UML DIAGRAM 26

4.3.1 Use case diagram 26

4.3.2 Activity diagram 27

4.3.3 Sequence diagram 28

**5** **MODULES DESCRIPTION**  30

5.1 MODULES 31

5.1.1 Data pre-processing 33

5.1.2 Training data 34

5.1.3 Evaluate for Content based filtering 35

5.1.4 Evaluate for Collaborative based filtering 36

5.1.5 Evaluate for popularity-based filtering 38

5.1.6 Evaluate for hybrid recommender system 39

**6** **SYSTEM TESTING**  41

6.1 OBJECTIVES OF TESTING 42

6.2 TESTING STRATEGIES 42

6.2.1 White box testing 42

6.2.2 Black box testing 42

6.2.3 Top-Down testing 42

6.2.4 Bottom-Up testing 42

6.2.5 Unit testing 43

6.2.6 Integration testing 43

6.3 TESTING RESULTS 44

**7** **CONCLUSION**  49

7.1 CONCLUSION 50

7.2 FUTURE WORK 50

**8** **APPENDICES**  51

APPENDIX I – SOURCE CODE 52

APPENDIX II – SCREENSHOTS 72

**9** **REFERENCES** 76