

# **Visvesvaraya Technological University**



**A Project Report  
on**

## **A Weight based Personalized Recommendation using Idiocentric and Collaborative Filtering**

**submitted in partial fulfillment of the requirements for the award of the degree of**

**BACHELOR OF ENGINEERING  
in  
COMPUTER SCIENCE & ENGINEERING**

**by**

**Sanjay G Huilgol – 1PI09CS083  
Vijay Mahantesh SM – 1PI09CS118  
Vijesh M – 1PI09CS119**

**under the guidance of  
Dr. Kavi Mahesh**

**January 2013 – May 2013**



**Department of Computer Science & Engineering  
PES Institute of Technology**

**(An autonomous institute under VTU, Belgaum and UGC, New Delhi)  
100 Feet Ring Road, BSK- III Stage, Bangalore – 560085**



**PES INSTITUTE OF TECHNOLOGY**  
(Autonomous institute under VTU, Belgaum)  
**100 Feet Ring Road, BSK- III Stage, Bangalore – 560 085**

**DEPARTMENT OF COMPUTER SCIENCE & ENGINEERING**

**CERTIFICATE**

Certified that the project work entitled “**A Weight based Personalized using Idiocentric and Collaborative Filtering**” is bonafide work carried out by

**Sanjay G Huilgol, 1PI09CS083,**

**Vijay Mahantesh SM, 1PI09CS118,**

**Vijesh M, 1PI09CS119,**

in partial fulfillment for the award of degree of Bachelor of Engineering in Computer Science and Engineering of the Visvesvaraya Technological University, Belgaum during the academic semester January 2013 - May 2013. It is certified that all corrections and suggestions indicated for internal assessment have been incorporated in the report deposited in the departmental library. The project report has been approved as it satisfies the academic requirements in respect of the project work prescribed for the said Bachelor of Engineering degree.

Signature of the Guide  
Dr. Kavi Mahesh

Signature of the HOD  
Prof. Nitin V. Pujari

Signature of the Principal  
Dr. K N B Murthy

External Viva  
Name of the Examiners

Signature with date

1. ....
2. ....

.....  
.....

## **Acknowledgement:**

The thrill and excitement of finishing a project can be felt in the last stage when the report is being written. When we look back from where it all started we recollect all the people who brought this project from infancy to its completion.

At the outset, we would like to express our heartfelt thanks towards our Principal **Dr. K N Balasubramanya Murthy** for providing us with a congenial environment to carry out our academic projects.

We express our deep sense of gratitude to our HOD, **Prof. Nitin V Pujari** for his invaluable support and guidance and providing all the necessary resources, which always encourages us to excel.

We consider it a privilege to acknowledge and express gratitude to our guide, **Dr. Kavi Mahesh** for his continuous support and motivational guidance throughout.

We also express our sincere thanks to all the teaching and non-teaching staff of the **Department of Computer Science, PESIT** for their kind support and ever helping nature. We would also like to thank our project coordinator, **Mr. Badri Prasad**, for all his cordial support during the project timeline.

A friend in need is a friend indeed is rightly said. We take this opportunity to thank all our friends for their constant and invaluable feedback which encouraged us to a great extent. In particular, we would like to thank **Sandeep Raju P** and **Rakesh Kumar R** for their invaluable effort in helping us with designing the approach and building the user interface.

Lastly, we are very grateful to several web resources like Google, Wikipedia and other useful technical articles without which this project would not have seen the light of day.

- **The Team**

## **Abstract:**

Recommending items to users is one of the most frequently stumbled upon requirements in the e-commerce, or the internet in general. Many techniques have been developed for recommending items. The techniques can be broadly classified into three categories: Collaborative Filtering, Content Based Filtering and Hybrid Recommendations. In this project, we have developed a Graph-based recommendation algorithm that uses both the idiocentric and collaborative behavior of the user to recommend the items back to the users. We have developed methods to carry out unsupervised dimensionality reduction on the dataset. We have deduced some of the characteristic qualities of the users and have constructed a custom built profile for each of them. We have used the movielens\_1m dataset to test the accuracy of our algorithm. A survey paper on movielens dataset by Bao and Xia, from autumn 2012 Machine Learning Course at Stanford University, compares four parameterized approaches for rating prediction. The least RMS error from all the approaches was found to be 0.917. Using the non-parameterized approach that we have developed, we were able to achieve an RMS error of 0.903.

## **Table of Contents:**

i.	Certificate	
ii.	Acknowledgement	
iii.	Abstract	
1.	Introduction .....	1
2.	Problem Definition.....	4
3.	Literature Survey.....	6
4.	Project Requirement Definition.....	9
4.1.	Target Users.....	10
4.2.	Software Requirements.....	10
4.3.	Hardware Requirements.....	11
4.4.	Assumptions and Dependencies.....	12
4.5.	Languages Used.....	13
5.	System Requirements Specification.....	14
6.	Gantt Chart.....	16
7.	System Design.....	18
7.1.	Component Diagram.....	19
7.2.	Activity Diagram.....	20
8.	Detailed Design.....	21
8.1.	Building Item Graph.....	22
8.2.	Building User Profiles.....	22
8.3.	Unsupervised Dimensionality Reduction.....	23
8.4.	User Similarity.....	24
8.5.	Clustering.....	25
8.6.	Idiocentric Recommendation.....	26
8.7.	Collaborative Filtering.....	26
8.8.	A weighted combination for recommendation.....	27
8.8.1.	Rank based weighted average.....	27
8.8.2.	Score based weighted average.....	28
9.	Hurdles: Design Decisions and Optimizations.....	29
9.1.	Git Versioning System.....	30
9.2.	Mockup UI.....	30
9.3.	Choosing Language.....	30
9.4.	Databases.....	31
9.5.	Genericity.....	31

9.4.	32-bit Addressable Woes.....	31
9.5.	Graph.....	32
9.6.	Categorization of Non-Categorical Data.....	33
9.7.	Parallelization of Algorithms.....	33
10.	Implementation.....	36
10.1.	Item Graph.....	37
10.2.	User Profile.....	39
10.3.	Dimensionality Reduction.....	40
10.4.	User similarity.....	41
10.5.	Idiocentric Recommendation.....	42
10.6.	Clustering.....	43
10.7.	Collaborative Filtering.....	44
11.	Project Highlights.....	46
11.1.	In-homogeneity in Item Properties.....	47
11.2.	Unsupervised Dimensionality Reduction.....	47
11.3.	User Profiling - deducing the weights of attributes just by using ratings.....	48
11.4.	User Profiling - deducing the weights of values for each attributes just by using ratings.....	49
11.5.	Rating prediction algorithm.....	50
11.6.	Extendibility of the approach to items with no ratings.....	50
11.7.	Possibility of modifying this algorithm to the Link Prediction problem if the user set and the item set are the same.....	51
11.8.	Idiosyncrasy factor for each user.....	51
12.	Integration.....	52
12.1.	Algorithmic Integration.....	53
12.2.	Front-end Back-end Integration.....	53
12.2.1.	Communication from UI to backend.....	53
12.2.2.	Communication from backend to UI.....	54
13.	Testing.....	55
14.	Screenshots.....	60
15.	Conclusion.....	66
16.	Future Enhancements.....	70
17.	References.....	73