SAPTHAGIRI COLLEGE OF ENGINEERING
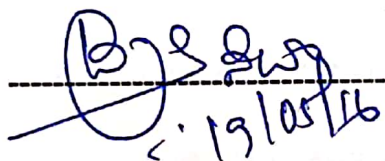
14/5, Chikkasandra, Hesaraghatta Main Road, Bangalore-560057

Department of Computer Science and Engineering

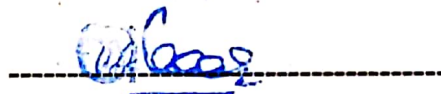# Certificate

SAPTHAGIRI
College of Engineering

Certified that the project work entitled **"An Approach to Enhance User Based Recommendation Engine for Big Data"** carried out by **NARAHARI P RAO (1SG12CS060), S NITIN SRINIVAS (1SG12CS068),** bonafide students of **this institute**, in partial fulfillment for the award of **Bachelor of Engineering** in **Computer Science and Engineering** of Visvesvaraya Technological University, Belgaum during the academic year **2015-16.** It is certified that all corrections/suggestions indicated for Internal Assessment have been incorporated in the report deposited in the library. The project report has been approved as it satisfies the academic requirements in respect of Project Work (10CS85) prescribed for the said degree.

19/05/16

Signature of the Guide

Dr. Prashanth C.M

Professor & HOD

Dr. Aswatha Kumar. M
Principal

Signature of the Principal

Sapthagiri College of Engineering
No. 14/5, Chikkasandra,
Hesaraghatta Main Road,
Bangalore-560 057

Dr. Aswatha Kumar M

Principal

Name of the Examiners

1........................

2........................

Signature with date

....

....

ABSTRACT

With the vast amount of data available today, organizations are looking for more accurate ways of using this data for improving productivity and user experience. Recommender system is one such technology that proactively suggests items of interest to users based on their objective behavior on their explicitly stated preferences. It is no longer a website add-on, but a necessary component. According to ChoiceStream survey, 45% of users are more likely to shop at a website that employs recommender technology.

Recommendation Engine is one of the most important parts of all commercial and social websites. Whenever a user searches for a book, music, movies or any other product, recommender systems plays a huge role in suggesting items that are similar. Such recommendations not only ensure active engagement of customers in business but also help the user make quicker decisions. Recommendations in general are of two types content based and user based. The project explores User based Recommendation Engines built using collaborative filtering approach by making use of Alternating Least Squares technique. Also an in-depth research is made on Recommendation engine running on a distributed architecture like Apache Hadoop, Apache Spark. Comparision and analysis of the results obtained from both Hadoop and Spark implementations of the Recommendation Engine is also done to understand how Spark is advantageous over hadoop.

This project aims at understanding what regularization constant is and inter-relationship of regularization constant and other factors that affect the accuracy of Recommendation Engine. Best practices are formulated as a result of the research to select optimal regularization constant value that helps us improve the system by reducing Root Mean Square Error. Here, the optimal value of regularization constant is determined as 10, number of iterations as 20 and rank as 8, resulting in the Root Mean Square Error value of 0.8256 which is an improvement of 3.46% compared to the NetFlix prize winning accuracy.