VISVESVARAYA TECHNOLOGICAL UNIVERSITY

BELAGAVI

A PROJECT REPORT ON

"A Data Mining Approach to Predict User's Next Question in Q&A System"

Submitted in partial fulfillment for the requirements for the award of degree

BACHELOR OF ENGINEERING

IN

INFORMATION SCIENCE & ENGINEERING

Submitted by

Amulya R	**4VV12IS005**
Ashrith C	**4VV12IS008**
H V Tejas Taunad	**4VV12IS019**
Kavya Shree S	**4VV12IS024**

Under the guidance of

Prof. VIDYASHREE K P
Assistant Professor
Department of Information Science & Engineering
V.V.C.E, Mysuru

DEPARTMENT OF INFORMATION SCIENCE AND ENGINEERING

VIDYAVARDHAKA COLLEGE OF ENGINEERING

MYSURU-570002

2015-2016

VIDYAVARDHAKA COLLEGE OF ENGINEERING

Gokulam III Stage,

Mysuru-570002

Department of Information Science and Engineering

CERTIFICATE

This is to certify that the Project entitled "**A Data Mining Approach to Predict User's Next Question in Q&A System"** is a bonfire work carried out by **Amulya R (4VV12IS005), Ashrith C (4VV12IS008), H V Tejas Taunad (4VV12IS019), Kavya Shree S (4VV12IS024)** of 8th semester Information Science and Engineering as prescribed by **Visvesvaraya Technological University** Belgaum, during the academic year 2015-2016.It is certified that all the suggestions and corrections indicated for the internal assessment have been incorporated in the report .The project report has been approved as it satisfies the requirements in respect of project work prescribed for the said degree.

Signature of Guide	**Signature of HOD**	**Signature of principal**
Mrs.Vidyashree K P	**Mr.Ravikumar V**	**Dr.B Sadashive Gowda**
Asst. Prof, IS&E	HOD,IS&E	Principal

Name of the Examiners **Signature with Date**

1)

2)

ACKNOWLEDGEMENT

If words are considered as the token of acknowledgement, then the words play the heralding role of expressing our gratitude.

We owe our gratitude to The Principal, **Dr. B. Sadashive Gowda** for his whole hearted support.

We wish to express our deepest gratitude to **Dr. Ravi Kumar V,** Head of Department, Information Science and Engineering, VVCE, for his profound alacrity in our progress and his valuable suggestions.

We hereby like to thank **Prof. Vidyashree K P**, Assistant Professor, Information Science and Engineering for her constant guidance and encouragement

We also thank the project coordinator **Prof. Vinutha D C**, Associate Professor, Information Science and Engineering for her support and encouragement.

At the end, we are anxious to offer our sincere thanks to our family members and friends for their valuable suggestions, encouragement and unwavering support.

AMULYA R

ASHRITH C

H V TEJAS TAUNAD

KAVYA SHREE S

ABSTRACT

In a Question Answering system, the user submits a question and waits for the answer as the response. If the system is capable of predicting the user's future interest as the next question, its performance will improve greatly. This project predicts users' future questions based on the current interaction records of the user with the system. Their current interactions with the system show what they are interested in. These interaction records are maintained in the form of Questions log from which the user sessions are extracted. Based on the user sessions, the system predicts the next question for which the user may become interested in near future. A sample Questions Log is selected for the purpose of performing experiments. The model of Association rule mining is applied to predict the future question of the user. The online community consists of several discussion groups, help communities and forums. The Question Answering system is a useful tool in such communities. The proposed system is a Question Answering system with an extended functionality of predicting the users' next question. By prediction, we eliminate the work required for posting each question and improve the efficiency and user friendliness of the system. The prediction is done using the data mining technique of Association rule mining. In particular, we make use of the Apriori algorithm to generate association rules among the questions posted by the user. We analyze the question log of the users to obtain their history with the system and derive their area of interest. The predictions are obtained using the question log as an input to the algorithm

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES