

# **VISVESVARAYA TECHNOLOGICAL UNIVERSITY**



**H.K.E. SOCIETY'S S.L.N. COLLEGE OF ENGINEERING  
RAICHUR - 584 135**



**A PROJECT REPORT ON**

## **“USER INTEREST BASED SOCIAL MEDIA DATA RETRIEVAL SYSTEM”**

**Submitted By:**

<b>ANUDEEP BABU K</b>	<b>(3SL19CS006)</b>
<b>BASANAGOWDA</b>	<b>(3SL19CS011)</b>
<b>RAGHAVENDRA D</b>	<b>(3SL19CS035)</b>
<b>VINODKUMAR SANKRANTHI</b>	<b>(3SL19CS055)</b>

**Under the Guidance of  
PROF.SURESH PATEL**

**Head of the Department  
PROF.SUMANGALA. I**

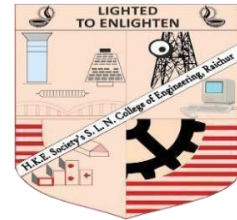
**Department of Computer Science & Engineering**  
(Affiliated to VTU - Belagavi, Affiliated to AICTE, Accredited by NAAC)  
Yeramarus Camp, Raichur-584135, Karnataka

**2022-2023**

# VISVESVARAYA TECHNOLOGICAL UNIVERSITY



**H.K.E. SOCIETY'S S.L.N. COLLEGE OF ENGINEERING  
RAICHUR - 584 135**



## **DEPARTMENT OF COMPUTER SCIENCE & ENGINEERING** **CERTIFICATE**

Certified that the Project work on topic “**User Interest Based Social Media Data Retrieval System**” is a bonafied work carried out by **Anudeep babu k (3SL19CS006), Basanagowda (3SL19CS011), Raghavendra D(3SL19CS035), Vinodkumar Sankranthi (3SL19CS055)** in partial fulfilment for the award of degree of **Bachelor of Engineering in Computer Science & Engineering of the Visvesvaraya Technological University, Belgaum** during the year **2022-2023**. It is certified that all corrections/suggestions indicated for internal assessments have been incorporated in the report deposited in the department. The project report has been approved as it satisfies the academic requirements in respect of project work prescribed for the **Bachelor of Engineering Degree**.

**Signature of the Guide**

**Prof.Suresh Patel**

**Signature of the Coordinator**

**Prof.Sujatha J**

**Name of the examiners**

- 1.
- 2.

**Signature of the HOD**

**Prof.Sumangala. I**

**Signature of the Principal**

**Dr.R. Basawaraja**

**Signature with date**

---

## **ACKNOWLEDGEMENT**

---

This satisfaction and euphoria that accompany the successful completion of any task would be but incomplete without the mention of the people who made it with constant guidance and encouragement and crowned our effort with success.

We would like to express our gratitude to our principal **DR. R. BASAWARAJA S.L.N. COLLEGE OF ENGINEERING**, for providing congenial environment and surroundings to work in.

A hearty thanks to our beloved HOD. **Prof. SUMANGALA I**, for her encouragement and support.

Our sincere thanks to our beloved coordinator **Prof.SUJATHA J**, Department of Computer Science & Engineering, for her kind support and cooperation.

We express our sincere thanks to project guide **Prof.SURESH PATEL** Dept. Of **COMPUTER SCIENCE AND ENGINEERING**, for her constant encouragement and support throughout our course especially for the useful suggestions given during the course of the project period.

We also thank all the staff members of department of Computer Science and Engineering and all those who have directly or helped us with valuable suggestions in the successful completion of this project.

Last but not the least we would like to thank our beloved parents for their blessing, love, and encouragement to successfully complete the task by meeting all the requirements

**ANUDEEP BABU K (3SL19CS006)**  
**BASANAGOWDA (3SL19CS011)**  
**RAGHAVENDRA D (3SL19CS035)**  
**VINODKUMAR SANKRANTHI (3SL19CS055)**

## **ABSTRACT**

Recently, there has been a significant rise in the ecommerce industry and more specifically in people buying products online. There has been a lot of research being done on figuring out the buying patterns of a user and more importantly the factors which determine whether the user will buy the product or not. In this study, we will be researching on whether it is possible to identify and predict the purchase intention of a user for a product and target that user towards the product with a personalized advertisement or a deal. Further, we wish to develop software that will help the businesses identify potential customers for their products by estimating their purchase intention in measurable terms from their tweets and user profile data on twitter. After applying various text analytical models to tweets data, we have found that it is indeed possible to predict if a user have shown purchase intention towards a product or not, and after doing some analysis we have found that people who had initially shown purchase intention towards the product have in most cases also bought the product.

Digital marketing is taken into account the well-liked method comparing to traditional marketing. It can be used by both researchers and academicians for social media marketing and to predict the customers purchase intention. The Proposed work revolves around some valuable information and processes in accordance to the behavior of customer during the online purchase. Business owners, scientists, researchers all post their ads, details on the Web so that they can be linked to owners quickly and easily by web scrap searching on searchable product websites to gain a lot of data from websites. Hence, customer price and rating of product evaluation and prediction has become an important research area. The analysis is done by Support Vector Machine (SVM-Linear) to gather several information and provide variation analysis. The major goal remains to investigate and analyze the extracted dataset using ML oriented algorithms with best accuracy possible. The analysis has a proper path to sentimental analysis of parameters in accordance to the ratings and price of the product to find proper accurate calculations.

## **LIST OF FIGURES**

**Fig 4.1 : Steps involved in User Interest Based Social Media Data Retrieval System**

**Fig 4.2 : Flow Diagram 1**

**Fig 4.3 : Flow Diagram 2**

**Fig 4.4 : Sequential Diagram**

**Fig 4.5 : SVM Algorithm**

**Fig 4.6 : Logistic Regression algorithm**

**Fig 4.7 : Logistic Regression Algorithm**

**Fig 4.8 : Neural Networks Algorithm**

**Fig 5.1 : Data set Processing**

**Fig 6.1 : home page 1**

**Fig 6.2 : home page 2**

**Fig 6.3 : upload of Data**

**Fig 6.4 : Test of Data**

**Fig 6.5 : Stats**

**Fig 6.6 : Count of Tweets**

**Fig 6.7 : Analysis**

**Fig 6.8 : Potential Customers**

**Fig 6.9 : Potential Data Results**

**Fig 6.10 : Select Machine Mode**

**Fig 6.11 : Select the Document Vector**

**Fig 6.12 : Select Level**

**Fig 6.13 : Positive and Negative Tweets**

## **TABLE OF CONTENTS**

<b>Chapter 1 Introduction .....</b>	<b>01</b>
1.1 Objectives	
1.2 Problem Statement	
<b>Chapter 2 Literature Review .....</b>	<b>04</b>
<b>Chapter 3 Methodology .....</b>	<b>07</b>
3.1 Existing System	
3.2 Proposed System	
3.2.1 Data Preprocessing	
3.2.2 Algorithms	
3.2.3 Document Vector	
3.2.4 Evaluation	
<b>Chapter 4 Design .....</b>	<b>14</b>
4.1 System Architecture	
4.2 Machine Learning	
4.3 Algorithms	
4.3.1 Support Vector Machine (Svm)	
4.3.2 Naive Bayes Algorithm	
4.3.3 Logistic Regression	
4.3.4 Decision Tree Algorithm	
4.3.5 Neural Network	

**Chapter 5 System Overview ..... 32**

5.1 Background And Motivation

5.2 Project Goal

5.3 Project Requirements

5.4 Errors

5.5 Trails And Tests

5.6 Advantages Of User Interest Based Social Media

Data Retrieval System

**Chapter 6 Experimental Analysis ..... 38**

6.1 Hardware Requirements

6.2 Software Requirements

6.3 Tools And Technology Details

6.3.1 Python

6.3.2 Visual Studio Code

6.3.3 Django Framework

6.3.4 Machine Learning Techniques

6.3.5 Data Analysis

6.4 Results

6.4.1 Formation Of Document Vector

6.4.2 Modelling

6.4.3 Experimentation And Results

6.5 Snapshots



<b>Chapter 7 Conclusion And Future Scope .....</b>	<b>55</b>
7.1 Conclusion	
7.2 Future Work	
<b>Reference.....</b>	<b>57</b>