A

Major Project

On

TWEETS CATEGORIZATION AND COMPARISON OF RESULTS USING MACHINE LEARNING MODELS

(Submitted in partial fulfillment of the requirements for the award of Degree)

BACHELOR OF TECHNOLOGY

In

COMPUTER SCIENCE AND ENGINEERING

By

MUDASSAR AHMED KHAN (197R1A0593)

MD ABDUL RAHMAN (207R5A0511)

MD.VAHEED (197R1A0591)

Under the Guidance of

K.PRAVEEN KUMAR

(Assistant Professor)

DEPARTMENT OF COMPUTER SCIENCE AND ENGINEERING

CMR TECHNICAL CAMPUS

UGC AUTONOMOUS

(Accredited by NAAC, NBA, Permanently Affiliated to JNTUH, Approved by AICTE, New Delhi) Recognized Under Section 2(f) & 12(B) of th e UGCAct.1956,

Kandlakoya (V),

Medchal Road, Hyderabad-501401.

2019-2023

DEPARTMENT OF COMPUTER SCIENCE AND ENGINEERING

CERTIFICATE

This is to certify that the project entitled **"TWEETS CATEGORIZATION AND COMPARISON OF RESULTS USING MACHINE LEARNING MODELS"** being submitted by **MUDASSAR AHMED KHAN(197R1A0593),MD ABDUL RAHMAN(207R5A0511) & MD.VAHEED(197R1A0591)** in partial fulfillment of the requirements for the award of the degree of B.Tech in Computer Science and Engineering to the Jawaharlal Nehru Technological University Hyderabad, is a record of bonafide work carried out by them under our guidance and supervision during the year 2022-23.

The results embodied in this thesis have not been submitted to any other University or Institute for the award of any degree or diploma.

K.PRAVEEN KUMAR
(Assistant Professor)
INTERNAL GUIDE

Dr.A. Raji Reddy
DIRECTOR

Dr. K. Srujan Raju
HOD

EXTERNAL EXAMINER

Submitted for viva voce Examination held on_____

ACKNOWLEDGEMENT

Apart from the efforts of us, the success of any project depends largely on the encouragement and guidelines of many others. We take this opportunity to express our gratitude to the people who have been instrumental in the successful completion of this project.

We take this opportunity to express my profound gratitude and deep regard to my guide K**. Praveen Kumar,** Associate Professor for his exemplary guidance, monitoring, and constant encouragement throughout the project work. The blessing, help, and advice given by him shall carry us a long way in the journey of life on which we are about to embark.

We also take this opportunity to express a deep sense of gratitude to the Project Review Committee (PRC) **Dr. Punyaban Patel, Ms. K. Shilpa, Dr. T . Subha Mastan Rao & J. Narasimharao** for their cordial support, valuable information and guidance, which helped us in completing this task through various stages.

We are also thankful to **Dr. K. Srujan Raju,** Head of the Department of Computer Science and Engineering, **Dr. Ashuthosh Saxena,** Dean R&D, and **Dr. D T V Dharmajee Rao,** Dean Academics for providing encouragement and support for completing this project successfully.

We are obliged to **Dr. A. Raji Reddy,** Director for being cooperative throughout the course of this project. We also express our sincere gratitude to Sri. **Ch. Gopal Reddy,** Chairman for providing excellent infrastructure and a nice atmosphere throughout the course of this project.

The guidance and support received from all the members of **CMR Technical Campus** contributed to the completion of the project. We are grateful for their constant support and help.

Finally, we would like to take this opportunity to thank our family for their constant encouragement, without which this assignment would not be completed. We sincerely acknowledge and thank all those who gave support directly and indirectly in the completion of this project.

MUDASSAR AHMED KHAN (197R1A0593)

MD ABDUL RAHMAN (207R5A0511)

MD.VAHEED (197R1A0591)

ABSTRACT

In recent years, research on Twitter sentiment analysis, which analyzes Twitter data (tweets) to extract user sentiments about a topic, has grown rapidly. Many researchers prefer the use of machine learning algorithms for such analysis. This study aims to perform a detailed sentiment analysis of tweets based on ordinal regression using machine learning techniques. The proposed approach consists of first pre-processing tweets and using a feature extraction method that creates an efficient feature. Then, under several classes, these features scoring and balancing. Multinomial logistic regression (SoftMax), Support Vector Regression (SVR), Decision Trees (DTs), and Random Forest (RF) algorithms are used for sentiment analysis classification in the proposed framework. For the actual implementation of this system, a twitter dataset publicly made available by the NLTK corpora resources is used. Experimental findings reveal that the proposed approach can detect ordinal regression using machine learning methods with good accuracy. Moreover, results indicate that Decision Trees obtain the best results outperforming all the other algorithms.

LIST OF FIGURES/TABLES

LIST OF SCREENSHOTS

TABLE OF CONTENTS

1. INTRODUCTION

/

2. LITERATURE SURVEY

3. SYSTEM ANALYSIS

4. ARCHITECTURE

5.IMPLEMENTATION

6.RESULTS

7.TESTING

8.CONCLUSION

9.BIBLIOGRAPHY

/

10.PAPER PUBLICATION

11.CERTIFICATES