FAKE DATA ANALYSIS AND DETECTION USING ENSEMBLED HYBRID ALGORITHM

A PROJECT REPORT

Submitted by

PALAGATI BHANU PRAKASH REDDY (0015113135)

MANDI PAVAN KUMAR REDDY (0015113134)

GANJIKUNTA MANASWINI REDDY (0015113144)

Under the guidance of

Dr. K. M. Mehata, Senior Professor, CSE Department

in partial fulfillment for the award of the degree of

BACHELOR OF TECHNOLOGY

IN

COMPUTER SCIENCE AND ENGINEERING

HINDUSTAN INSTITUTE OF TECHNOLOGY AND SCIENCE, CHENNAI - 603 103

APRIL 2019

HINDUSTAN
INSTITUTE OF TECHNOLOGY & SCIENCE
(DEEMED TO BE UNIVERSITY)

BONAFIDE CERTIFICATE

Certified that this project report **"FAKE DATA ANALYSIS AND DETECTION USING ENSEMBLED HYBRID ALGORITHM"** is the bonafide work of **"PALAGATI BHANU PRAKASH REDDY (0015113135) MANDI PAVAN KUMAR REDDY (0015113134) GANJIKUNTA MANASWINI REDDY (0015113144)"** who carried out the project work under my supervision during the academic year **2018-2019.**

SIGNATURE

HEAD OF THE DEPARTMENT

Dr. Rajeswari Mukesh

Professor

Department of Computer Science and Engineering

Hindustan Institute of Technology and Science

SIGNATURE

SUPERVISOR

Dr. K. M. Mehata

Senior Professor

Department of Computer Science and Engineering

Hindustan Institute of Technology and Science

INTERNAL EXAMINER

Name: _____

Designation: _____

EXTERNAL EXAMINER

Name: _____

Designation: _____

Institution Name: _____

Project Viva - voce conducted on _____

ACKNOWLEDGEMENT

It's our extreme pleasure to thank our Chancellor **Dr. (Mrs.) Elizabeth Verghese** and Vice Chancellor **Dr. K. P. Issac**, Hindustan Institute of Technology and Science for providing conducive environment which helped us to pursue our project in a diligent and an efficient manner.

We wish to express our sincere gratitude to Dean (Academics) **Dr. N. Vasudevan**, Hindustan Institute of Technology and Science for his valuable directions, suggestions and support.

We are thankful to **Dr. M. Rajeswari Mukesh,** Head of the Department, Electronics and Communication Engineering for having evinced keen interest in our project and for his continued support.

We are indebted to our project guide **Dr. K. M. Mehata,** Senior Professor, Electronics and Communication Engineering for his/her valuable guidance and technical support in the accomplishment of our project.

We also thank our Project Coordinators **Mr. S. K. Shankar,** Associate Professor, Computer Science and Engineering for their support and coordination throughout the accomplishment of the project.

Our sincere thanks to all the teaching and non-teaching staff and family members who have been constantly supporting us throughout the accomplishment of this project.

BHANU PRAKASH REDDY. P

PAVAN KUMAR REDDY. M

MANASWINI REDDY. G

DEDICATION

This project **“FAKE DATA ANALYSIS AND DETECTION USING ENSEMBLED HYBRID ALGORITHM”**, dedicated to our beloved parents and friends. This project is also dedicated to the staffs of Department of Computer Science and Engineering, Hindustan Institute of Technology and Science.

ABSTRACT

Fake data detection is the most important problem to be addressed in the recent years, there is lot of research going on in this field. Because of its serious impacts on the readers. Researchers, government and private agencies working together to solve the issue. This project represents a hybrid approach for fake data detection using the multinomial voting algorithm. This algorithm was tested with multiple fake news dataset which resulted in an accuracy score of 94 percent which is a benchmark in the field of machine learning where the other algorithms are at a range of 82 to 88 percent. The list of algorithms that have been used here is as follows Naïve Bayes, Random Forest, Decision Tree, Support Vector Machine, K Nearest Neighbours. All these algorithms use training data as the bag of words model which was created using Count Vectorizer. Experimental data has collected from the Kaggle data world. Python is used as a language to verify and validate the results. Tableau is used as a visualization tool. Implementation is carried out using default algorithm values. A fake news detection website is created to validate and visualize the real time use cases of the algorithm. This will create a change if it is used properly.

TABLE OF CONTENTS

LIST OF FIGURES