VISVESVARAYA TECHNOLOGICAL UNIVERSITY, BELAGAVI

A PROJECT REPORT ON

"REAL-TIME SPAM DETECTION ON TWITTER USING MACHINE LEARNING"

Submitted in partial fulfilment for the award of Degree of,

BACHELOR OF ENGINEERING

IN

COMPUTER SCIENCE &ENGINEERING

By

AISHWARYA HOSAMANI	**4AL14CS001**
HANISHREE M S	**4AL16CS036**
HARSHITHA M	**4AL16CS038**
POOJASHREE T	**4AL16CS064**

Under the Guidance of

Miss. Shilpa

Assistant Professor

DEPARTMENT OF COMPUTER SCIENCE & ENGINEERING

ALVA'S INSTITUTE OF ENGINEERING AND TECHNOLOGY

MOODBIDRI-574225, KARNATAKA, 2019 – 2020

ALVA'S INSTITUTE OF ENGINEERING AND TECHNOLOGY

MIJAR, MOODBIDRI D.K. -574225, KARNATAKA

DEPARTMENT OF COMPUTER SCIENCE & ENGINEERING

CERTIFICATE

This is to certify that the project entitled **"REAL-TIME SPAM DETECTION ON TWITTER USING MACHINE LEARNING"** has been successfully completed by

AISHWARYA HOSAMANI	**4AL14CS001**
HANISHREE M S	**4AL16CS036**
HARSHITHA M	**4AL16CS038**
POOJASHREE T	**4AL16CS064**

the bonafide students of **DEPARTMENT OF COMPUTER SCIENCE & ENGINEERING, ALVA'S INSTITUTE OF ENGINEERING AND TECHNOLOGY** of the **VISVESVARAYA TECHNOLOGICAL UNIVERSITY, BELAGAVI** during the year 2019–2020. It is certified that all corrections/suggestions indicated for Internal Assessment have been incorporated in the report deposited in the departmental library. The project report has been approved as it satisfies the academic requirements in respect of Project work prescribed for the Bachelor of Engineering Degree.

Miss. Shilpa	**Dr. Manjunath Kotari**	**Dr. Peter Fernandes**
Project Guide	**Head of the department**	**Principal**

External Viva

Name of the Examiners	**Signature with Date**
1. Dr. Manjunath Kotari	
2. Vivek Sharma. S	Vivek Sharma

ABSTRACT

With the increased popularity of online social networks, spammers find these platforms easily accessible to trap users in malicious activities by posting spam messages. In this work, we have taken Twitter platform and performed spam tweets detection. To stop spammers, Google Safe Browsing and Twitter's Bootmaker tools detect and block spam tweets. These tools can block malicious links; however, they cannot protect the user in real-time as early as possible. Thus, industries and researchers have applied different approaches to make spam free social network platform. Some of them are only based on user-based features while others are based on tweet-based features only.

However, there is no comprehensive solution that can consolidate tweet's text information along with the user-based features. To solve this issue, we propose a modelled system which takes the user and tweet-based features along with the tweet text feature to classify the tweets. The benefit of using tweet text feature is that we can identify the spam tweets even if the spammer creates a new account which was not possible only with the user and tweet-based features.