

A
PROJECT REPORT
ON
“RUMOR DETECTION ON TWITTER DATASET”

SUBMITTED TO
SHIVAJI UNIVERSITY, KOLHAPUR
IN THE PARTIAL FULFILLMENT OF REQUIREMENT FOR THE AWARD OF DEGREE IN
BACHELOR OF TECHNOLOGY IN COMPUTER SCIENCE AND ENGINEERING

SUBMITTED BY

- | | |
|--------------------|--|
| 1. 19UCS107 | Miss. Mrunali Sanjay Prabhavalkar |
| 2. 20UCS306 | Mr. Santosh Ashok Powar |
| 3. 20UCS310 | Miss. Mayuri Suryakant Shete |
| 4. 20UCS311 | Miss. Aishwarya Anand Sutar |
| 5. 20UCS313 | Miss. Disha Manoj Teware |

UNDER THE GUIDANCE OF

Prof. S. S. Sangewar



DEPARTMENT OF COMPUTER SCIENCE AND ENGINEERINGDKTE
SOCIETY'S TEXTILE AND ENGINEERING INSTITUTE, ICHALKARANJI
2022-2023

D.K.T.E. SOCIETY'S TEXTILE AND ENGINEERING INSTITUTE, ICHALKARANJI
(AN AUTONOMOUS INSTITUTE)

DEPARTMENT OF COMPUTER SCIENCE & ENGINEERING



Promoting Excellence in
Teaching, Learning & Research

CERTIFICATE

This is to certify that, project work entitled

“RUMOR DETECTION ON TWITTER DATASET”

is a Bonafide record of project work carried out in this college by

- | | |
|--------------------|--|
| 1. 19UCS107 | Miss. Mrunali Sanjay Prabhavalkar |
| 2. 20UCS306 | Mr. Santosh Ashok Powar |
| 3. 20UCS310 | Miss. Mayuri Suryakant Shete |
| 4. 20UCS311 | Miss. Aishwarya Anand Sutar |
| 5. 20UCS313 | Miss. Disha Manoj Teware |

is in the partial fulfillment of the award of degree Bachelor in Technology in Computer Science & Engineering prescribed by Shivaji University, Kolhapur for the academic year 2022-2023.

Prof. S. S. Sangewar

(PROJECT GUIDE)

PROF.(DR.) D.V. KODAVADE

(HOD CSE DEPT.)

PROF.(DR.)L. S. ADMUTHE

(DIRECTOR)

EXAMINER: _____

DECLARATION

We hereby declare that, the project work report entitled “**RUMOR DETECTION ON TWITTER DATASET**” which is being submitted to D.K.T.E. Society’s Textile and Engineering Institute Ichalkaranji, affiliated to Shivaji University, Kolhapur is in partial fulfillment of degree B.Tech.(CSE). It is a Bonafide report of the work carried out by us. The material contained in this report has not been submitted to any university or institution for the award of any degree. Further, we declare that we have not violated any of the provisions under the Copyright and Piracy / Cyber / IPR Act Amended from time to time.

- | | |
|--------------------|--|
| 1. 19UCS107 | Miss. Mrunali Sanjay Prabhavalkar |
| 2. 20UCS306 | Mr. Santosh Ashok Powar |
| 3. 20UCS310 | Miss. Mayuri Suryakant Shete |
| 4. 20UCS311 | Miss. Aishwarya Anand Sutar |
| 5. 20UCS313 | Miss. Disha Manoj Teware |

ACKNOWLEDGEMENT

With great pleasure we wish to express our deep sense of gratitude to **Prof. S. S. Sangewar** for his valuable guidance, support and encouragement in completion of this project report. Also, we would like to take the opportunity to thank our head of department **Dr. V. Kodavade** for his co-operation in preparing this project report.

We feel gratified to record our cordial thanks to other staff members of the Computer Science and Engineering Department for their support, help and assistance which they extended as and when required.

Thank you,

1. 19UCS107	Miss. Mrunali Sanjay Prabhavalkar
2. 20UCS306	Mr. Santosh Ashok Powar
3. 20UCS310	Miss. Mayuri Suryakant Shete
4. 20UCS311	Miss. Aishwarya Anand Sutar
5. 20UCS313	Miss. Disha Manoj Teware

ABSTRACT

Emergence in the social network leads to the extensive and faster diffusion of news than conventional news channels. Verification of data is challenging due to massive information on a social network. Unverified information can be a rumor or fake news that causes damage to an individuals and organizations, revealing the harmful impact on humanity. Therefore, it is vital to combat rumor diffusion to minimize the adverse effects on society. Despite vigorous efforts to deal with this issue, researchers mainly focused on temporal dynamics of posts and other features like a user, network, content-based, which demonstrate a moderate accuracy. The time series features are associated with an event that suppresses the other quality features related to each post. There is a scope for improvement in the accuracy, so this paper focuses on post-wise features such as user-based, content-based and lexical-based features along with post sequences. We proposed a framework that uses various essential features and combines two deep learning models. Word embedding is utilized with bidirectional long short-term memory (BiLSTM) and combined with post-wise features using a multilayer perceptron (MLP), which improves accuracy. The experiments on the real-world dataset of Twitter demonstrate a notable improvement in accuracy compared to state-of-the-art approaches.

Keywords: Deep learning . Lexical features . Rumor . Rumor detection . Social network

INDEX

Title	Page.no
1. INTRODUCTION	1
1.1 Problem definition	2
1.2 Aim and objective of the project	2
1.3 Scope and limitation of the project	2
1.4 Timeline of the project	3
1.5 Project Cost	4
2. BACKGROUND STUDY AND LITERATURE OVERVIEW	5-7
2.1 Literature overview	8
2.2 Critical appraisal of other people's work	8
2.3 Investigation of current project and related work	8
3. REQUIREMENT ANALYSIS	9
3.1 Software and hardware Requirement	9
3.2 Requirement Specification	10
3.3 Use case Diagram	11
4. SYSTEM DESIGN	12
4.1 Architectural Design	12
4.2 Modules	12-13
4.3 Algorithmic description of each module	14-15
4.4 System Modeling	16
4.4.1. Class Diagram	16
4.4.2. Data flow Diagram	17-18
4.4.3. Activity Diagram	19
5. IMPLEMENTATION	20
5.1 Environmental Setting for Running the Project	20
5.2 Detailed Description of Methods	20-25
5.3 Implementation Details	26-27
5.4 Code for Bi-LSTM_UCL	28
6. INTEGRATION AND TESTING	29
6.1 Unit Testing	29
6.2 Integration Testing	30

7. PERFORMANCE ANALYSIS	31
8. FUTURE SCOPE	32-33
9. APPLICATIONS	34
10. OUTPUTS	35-36
11. PLAGIARISM REPORT	37-39
12. REFERENCES	40