

### A Project Report On

**“NAMED ENTITY RECOGNITION IN SOCIAL MEDIA USING MACHINE LEARNING”**

***Submitted in the partial fulfillment for the award of the Bachelor of Engineering degree in Computer Science and Engineering***

**Submitted By**

|  |  |
| --- | --- |
| **AISHWARYA M S** | **4AD21CS004** |
| **ESHWAR A M** | **4AD21CS021** |
| **KUSHAL R** | **4AD21CS039** |
| **LAKSHMI PRIYA H P** | **4AD21CS040** |

**Under the guidance of**

## Mrs. Sushma V

Assistant Professor

Department of Computer Science & Engineering ATME College of Engineering

**Department of Computer Science & Engineering ATME College of Engineering**

**13th Kilometer, Mysore – Kanakapura - Bangalore Road, Mysore-570028**

**2024-2025**



**Department of Computer Science and Engineering**

**CERTIFICATE**

This is to certify that the Project Work entitled **“NAMED ENTITY RECOGNITION IN SOCIAL MEDIA USING MACHINE LEARNING”** is the Bonafide work carried out by **Aishwarya M S (4AD21CS004), Eshwar A M (4AD21CS021), Kushal R (4AD21CS039)** and **Lakshmi Priya H P (4AD21CS040)** in partial fulfillment for the award of degree of Bachelor of Engineering in Computer Science and Engineering from Visvesvaraya Technological University, Belagavi during the year 2024-2025. The report has been approved and satisfies the academic requirement with respect to Project Work prescribed for Bachelor of Engineering degree.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Signature of Guide |  | Signature of HOD |  | Signature of Principal |
| **Mrs. Sushma V**  Assistant Professor |  | **Dr. Puttegowda D**  Professor & Head |  | **Dr. L Basavaraj**  Principal |
|  |  | **External Viva** |  |  |
| **Name of Examiners** |  |  |  | **Signature with Date** |
| **1…………………….** |  |  |  | **……………………..** |
| **2…………………….** |  |  |  | **……………………..** |

The successful completion of our work would be incomplete without the mention of the names of the people who have made it possible. We are indebted to several individuals who have helped us to complete our project.

We are thankful to **Dr. L Basavaraj**, **Principal**, ATME College of Engineering for having granted us permission and extended full use of the college facilities to carry out our project successfully.

We express our profound gratitude to **Dr. Puttegowda D,** Professor & Head, Department of Computer Science and Engineering for his consistent co-operation and support.

At the outset we express our profound gratitude to our guide **Mrs.** **Sushma V**, Assistant Professor, Department of Computer Science and Engineering for her consistent co-operation and support.

We are greatly indebted to our project coordinators **Dr. Sunitha Patel M S** and **Dr. Shilpa B L**, Associate Professors, Department of Computer Science and Engineering for her timely inquiries into the progress of the project

We also thank all teaching and non-teaching staff members of the department for their valuable assistance throughout our academic tenure.

Lastly, we would like to thank our family and friends for their cooperation and support for successful completion of our project.

|  |  |
| --- | --- |
| **Aishwarya M S** | **(4AD21CS004)** |
| **Eshwar A M** | **(4AD21CS021)** |
| **Kushal R** | **(4AD21CS039)** |
| **Lakshmi Priya H P** | **(4AD21CS040)** |

I



The growing prevalence of social media platforms has led to an overwhelming amount of unstructured textual data, making it challenging to extract meaningful information. This project introduces a system for Named Entity Recognition (NER) on social media text using machine learning techniques, specifically leveraging a fine-tuned DistilBERT model. The system fetches posts from Reddit based on user-specified keywords and subreddits, processes the content, and identifies key entities such as names, organizations, and locations.

Using token classification with an aggregation strategy, the model analyzes noisy and informal text often prevalent in social media. Extracted entities are displayed in interactive visualizations, including frequency charts and entity type distributions, for better interpretability. The project highlights the challenges of working with noisy social media data and provides a robust, scalable solution for text analytics. By automating entity extraction, this system demonstrates the potential of machine learning in streamlining information retrieval from vast online datasets, with applications in market research, sentiment analysis, and digital investigations.



|  |  |  |
| --- | --- | --- |
| **Chapter No.** | **Chapter Name** | **Page No.** |
|  | ACKNOWLEDGEMENT | i |
|  | ABSTRACT  TABLE OF CONTENTS | ii  iii-iv |
|  | LIST OF FIGURES | v |
| **Chapter 1** | **Introduction** | **1-2** |
| 1.1 | Overview | 1 |
| 1.2 | Existing System | 1 |
| 1.3 | Problem Statement | 1 |
| 1.4 | Proposed System | 1 |
| 1.5 | Advantages Over Current System | 2 |
| **Chapter 2** | **Literature Survey** | **3-9** |
| 2.1 | Survey Papers | 3-8 |
| 2.2 | Related Work | 8-9 |

**Chapter 3 System Requirements and Specification 10-15**

|  |  |  |
| --- | --- | --- |
| 3.1 | Functional Requirements | 10 |
| 3.2 | Non-Functional Requirements | 11-15 |
| **Chapter 4** | **System Analysis** | **16-17** |
| **Chapter 5** | **Methodology** | **18-25** |
| 5.1 | System Architecture | 18-22 |
| 5.2 | Proposed System | 23-25 |
| **Chapter 6** | **Implementation** | **26-34** |
| 6.1 | Stages of Implementation | 27-30 |
| 6.2 | Code Snippets | 31-36 |



|  |  |  |
| --- | --- | --- |
| **Chapter 7** | **System Testing** | **37-37** |
| 7.1 | Testing Strategy | 37 |
| 7.2 | Unit Testing | 38 |
| 7.3 | Integration Testing | 39 |
| 7.4 | Validation Testing | 39 |
| 7.5 | Output Testing | 40 |
| 7.6 | System Testing | 40 |
| 7.7 | Performance and Scalability Testing | 41 |
| **Chapter 8**  8.1  8.2 | **Results**  Comparative Study  Snapshots | **42-45**  42-45  46-47 |
|  | **Conclusion** | 48 |
|  | **Future Scope** | 49 |
|  | **References** | 50-51 |
|  |  |  |



**LIST OF FIGURES**

|  |  |  |
| --- | --- | --- |
| **Figure No.** | **Figure Name** | **Page No.** |
| 5.1 | System Architecture | 22 |
| 5.2 | Proposed System | 24 |
| 8.2.1 | Home Page | 46 |
| 8.2.2 | Result Page | 46 |
| 8.2.3.1 | Combined Entity Analysis | 47 |
| 8.2.3.2 | Combined Entity Type Analysis | 47 |
|  |  |  |
|  |  |  |
|  |  |  |
|  |  |  |
|  |  |  |
|  |  |  |
|  |  |  |
|  |  |  |
|  |  |  |
|  |  |  |
|  |  |  |
|  |  |  |