# VARDHAMAN COLLEGE OF ENGINEERING

**(AUTONOMOUS)**

Shamshabad - 501 218, Hyderabad

## DEPARTMENT OF COMPUTER SCIENCE AND ENGINEERING

**B. TECH VIII SEMESTER CSE-A**

| BATCH<br>2018-22 | PROJECT WORK<br>2021-22 | ACADEMIC<br>YEAR 2021-22 |
|---|---|---|

**Batch Id :** 18CSPW-A19

**Title of the Project : Identifying and Categorizing Offensive language in Social Media**

**TEAM MEMBERS**

| S. No. | Roll Number | Student Name | Mail Id | Contact Number |
|---|---|---|---|---|
| 1 | 18881A0531 | Anurima Vaishnavi Kumar | kumaranurima_18cs@vardhaman.org | 7995047632 |
| 2 | 18881A0537 | M.Vaishnavi | mvaishnavi_18cs@vardhaman .org | 9014909390 |

**GUIDE DETAILS**

| | | |
|---|---|---|
| **Name of the Guide** | : | Mrs M Nikita |
| **Designation** | : | Assistant Professor |
| **Department** | : | Computer Science and Engineering |
| **Mail Id** | : | nikitha@vardhaman.org |
| **Contact Number** | : | 8790531091 |
| **Area of interest** | : | Computer Networks,Computer Organization,Cryptography and Network Security |

| | | |
|---|---|---|
| *Signature of the Projects Incharge* | *Signature of the Guide with Date* | *Signature of the Project Coordinator* |

**Batch Id :** 18CSPW-A19

**Title of the Project : Identifying and Categorizing Offensive language in Social Media**

## ABSTRACT

Offensive language is pervasive in social media. Individuals frequently take advantage of the perceived anonymity of computer-mediated communication, using this to engage in behavior that many of them would not consider in real life. Online communities, social media platforms, and technology companies have been investing heavily in ways to cope with offensive language to prevent abusive behavior in social media. One of the most effective strategies for tackling this problem is to use computational methods to identify offense, aggression, and hate speech in user generated content (e.g. posts, comments, microblogs, etc.).. This system classifies a tweet as either offensive or not offensive (Sub-task A) and further classifies offensive tweets into categories (Sub-tasks B – C). Some sort of grid search approach is taken where multiple techniques for preprocessing, feature extraction and classification are implemented and combinations of them all are tried to achieve the best model for the given dataset.

## Objectives

**1.** The main objective is to categorize the tweets into two main categories i.e, offensive or not offensive and further classify those offensive tweets into individual, group etc.

**2.** We use machine learning models and NLP models to determine the category/class a particular tweet/test data falls in.

## Scope of Work:

It can be used by Twitter, Instagram, and Facebook where there is a tremendous amount

of negativity, profanity, defamation. Users can enable offensive language detection in their comments section, say in Instagram, and the algorithm will automatically block/report the user if his language/comment is offensive in any manner.

**Commercializable: Yes**

**REFERENCES:**

[1]. Georgios K. Pitsilis, Heri Ramampiaro and Helge Langseth, 'Detecting Offensive Language in Tweets Using Deep Learning'

URL- https://arxiv.org/pdf/1801.04433.pdf

[2]. Myan Sheriff, Sherine Mamdouh and Wegdan Ghazi, CTSys at SemEval-2018 Task 3: Irony in Tweets

URL: http://aclweb.org/anthology/S18-1094

[3]. Shivam Bansal, Ultimate Guide to Understand & Implement Natural Language Processing

URL- https://www.analyticsvidhya.com/blog/2017/01/ultimate-guide-to-understand-implement-naturallanguage-processing-codes-in-python/

**Date of Submission: 20-10-2021**

***Signature of the Guide with Date*** ***Signature of the Project In-Charge***