MES COLLEGE OF ENGINEERING-KUTTIPPURAM
DEPARTMENT OF COMPUTER APPLICATIONS
20MCA245– MINI PROJECT

Mini Project Proposal (III Semester MCA)

Approval of the mini project proposal is mandatory to continue and submit the project work.

The mini project proposal should clearly state the project objectives and the environment of the proposed project to be undertaken.

The following documents are to be submitted for approval

1. Pro forma for approval of the mini project (Present in this document)
2. Synopsis/Abstract with following contents
    1. Title of the Mini Project.
    2. Introduction and Objectives of the Project.
    3. Tools / Platform, Hardware and Software Requirement
    4. Problem Definition and Initial Requirements
    5. Basic functionalities of the project

The abstract should be submitted in the format given in the 3rd page of this document.

The Abstract in the given format shall be uploaded on or before **01.12.21**

MES COLLEGE OF ENGINEERING, KUTTIPPURAM
DEPARTMENT OF COMPUTER APPLICATIONS
20MCA245 – MINI PROJECT

PRO FORMA FOR THE APPROVAL OF THE THIRD SEMESTER MINI PROJECT

(Note: All entries of the pro forma for approval should be filled up with appropriate and complete information. Incomplete Pro forma of approval in any respect will be rejected.)

Mini Project Proposal No : ___1_____ *(Filled by the Department)*	Academic Year : __2021-2022_____

1.Title of the Project : Cyber *bullying Detection on Social Networks Using Machine Learning Approaches*

2.Name of the Guide : Mr. VASUDEVAN T V

3.Number of the Student: MES20MCA-2028

4.Student Details (in BLOCK LETTERS)

Name MOHAMMED LIJAS C Roll Number 28

Signature

1.

Date: 1/11/2021

Approval Status : Approved / Not Approved__

Signature of Committee Members

Comments of The Mini Project Guide Dated Signature

Initial Submission :

First Review :

Second Review :

Comments of The Project Coordinator Dated Signature

Initial Submission:

First Review

Second Review

Final Comments :

Dated Signature of HOD

Cyber bullying Detection on Social Networks Using Machine Learning Approaches

MOHAMMED LIJAS C

Introduction & Objectives:

Cyberbullying can be defined as aggressive, intentional actions performed by an individual or a group of people via digital communication methods such as sending messages and posting comments against a victim. Different from traditional bullying that usually occurs at school during faceto-face communication, cyber bullying on social media can take place anywhere at any time. For bullies, they are free to hurt their peers' feelings because they do not need to face someone and can hide behind the Internet. For victims, they are easily exposed to harassment since all of us, especially youth, are constantly connected to Internet or social media.

Cyberbullying victimization rate ranges from 10% to 40%. In the United States, approximately 43% of teenagers were ever bullied on social media. The same as traditional bullying, cyberbullying has negative, insidious and sweeping impacts on children. The outcomes for victims under cyberbullying may even be tragic such as the occurrence of self-injurious behaviour or suicides. One way to address the cyberbullying problem is to automatically detect and promptly report bullying messages so that proper measures can be taken to prevent possible tragedies.

The use of social media has grown exponentially over time with the growth of the Internet and has become the most influential networking platform in the 21st century. However, the enhancement of social connectivity often creates negative impacts on society that contribute to a couple of bad phenomena such as online abuse, harassment cyberbullying, cybercrime and online trolling. Cyberbullying frequently leads to serious mental and physical distress, particularly for women and children, and even sometimes force them to attempt suicide. Online harassment attracts attention due to its strong negative social impact. Many incidents have recently occurred worldwide due to online harassment, such as sharing private chats, rumours, and sexual remarks. Therefore, the identification of bullying text or message on social media has gained a growing amount of attention among researchers. The purpose of this research is to design and develop an effective technique to detect online abusive and bullying messages by merging natural language processing and machine learning. Two distinct features, namely Bag-of - Words and term frequency-inverse text frequency (TF-IDF), are used to analyse the accuracy level of four distinct machine learning algorithms.

Millions of users creating the social revolution. Users' social behavior influences them to connect with others with same mentality. Social networks are constituted because of its user or organizations common interest in some social emerging issues. The popular social networking sites are Facebook, Twitter, MySpace, Orkut, LinkedIn, Google plus etc. which are actually online social networking (OSN) sites. However, the large amount of online users and their diverse and dynamic interests possess great challenges to support recommendation of friends on SNS's for each of the users. In this paper, we proposed a novel friend recommendation framework (FRF) based on the behavior of users on particular SNS's. The proposed method is consisted of the following stages: measuring

the frequency of the activities done by the users and updating the dataset according to the activities, applying FP-Growth algorithm to classify the user behavior with some criteria, then apply multilayer thresholding for friend recommendation.

Problem Definition:

Existing system

Social media is a platform that allows people to post anything like photos, videos, documents extensively and interact with society . People connect with social media using their computers or smartphones. The most popular social media includes Facebook, Twitter, Instagram, TikTok and so on. Nowadays, social media is involved in different sectors like education , business , and also for the noble cause . Social media is also enhancing the world's economy through creating many new job opportunities .Although social media has a lot of benefits, it also has some drawbacks.Using this media, malevolent users conduct unethical and fraudulent acts to hurt others feelings and damage their reputation. Recently, cyberbullying has been one of the major social media issues. Cyberbullying or cyber-harassment refers to an electronic method of bullying or harassment. Cyberbullying and cyber-harassment are also known as online bullying. As the digital realm has grown and technology has progressed, cyberbullying has become relatively common, particularly amongst adolescents. So it is very important to control these kind of harassment on online platforms. In existing system there is no effective methods to prevent Cyberbullying .

Proposed system

In this context, we suggest a cyberbullying detection model based on machine learning that can detect whether a text relates to cyberbullying or not. We conduct experiments with two data sets from twitter and Facebook's comments and posts. For performance analysis, we use two different feature vectors BoW and TF-IDF. The results indicate that TF-IDF feature provides better accuracy than BoW. Users behaviour could be defined in several approaches like association rules in perspective of mining, complex graph activities, sequence mining etc. Suppose for two different user we have same behaviour we can recommend them each other.

Basic functionalities:

Functional module description

- **Social Network**

A social network has to be created This social network will be maintained by the admin Users have to register in this social network. Registered users can socialize with other users using this .

- **Filtering**

All the content in this social network will be filtered and only after that it will reach the user. For filtering rules are kept. The Text filtering is a collection of words called bag of words is constructed and the words included in this are filtered. These words are filtered directly and are also extracted from their latent structure. Based upon the percentage of the bullying content in that word the behavior of the user is defined. Such users are enlisted in the blacklist and based upon their future activities they may be either added to the block list or may be removed from the black list.

TF-IDF stands for *term frequency-inverse document frequency*

In information retrieval, **tf–idf**, **TF*IDF**, or **TFIDF**, short for **term frequency–inverse document frequency**, is a numerical statistic that is intended to reflect how important a word is to a document in a collection or corpus. It is often used as a weighting factor in searches of information retrieval, text mining, and user modeling. The tf–idf

value increases proportionally to the number of times a word appears in the document and is offset by the number of documents in the corpus that contain the word, which helps to adjust for the fact that some words appear more frequently in general. tf–idf is one of the most popular term-weighting schemes today. A survey conducted in 2015 showed that 83% of text-based recommender systems in digital libraries use tf–idf.

Users behaviour could be defined in several approaches like association rules in perspective of mining, complex graph activities, sequence mining etc.

User Module Description

- ADMIN
  - View user
  - Add bullying words
  - Add good words
  - View bullying words
  - View good words
  - View report
- USER
  - Registration
  - Add post
  - View my post
  - Chat
  - Add bullying words
  - Add friend request
  - View friend request
  - View recommendation.

HARDWARE AND SOFTWARE REQUIREMENT

This specifies the hardware and the support software required to carry out the development

Hardware Requirements

The selection of hardware is very important in the existence and proper working of any software. Then selection hardware, the size and capacity requirements are also important.

1. Processor : 64 bit
2. RAM : Min 3 GB
3. Hard Disk : 10 GB

Software Requirements

One of the most difficult task is selecting software for the system, once the system requirements is found out then we have to determine whether a particular software package fits for those system requirements. The application requirement:

1 . OPERATING SYSTEM: WINDOWS 10

2 . FRONT END: HTML, CSS, JAVASCRIPT

3. BACK END: Mysql

4 . IDE: JetbrainsPycharm, Android studio

5. TECHNOLOGY USED: PYTHON,JAVA

6 .FRAME WORK USED: Flask