**Mini Project Report On**

**PREDICTING SOCIAL MEDIA CYBERBULLYING IN THE BIG DATA ERA WITH MACHINE LEARNING ALGORITHMS**

Submitted to

**JAWAHARLAL NEHRU TECHNOLOGICAL UNIVERSITY**

in partial fulfillment of the requirement

for the award of the degree of

#### BACHELOR OF TECHNOLOGY IN

**ARTIFICIAL INTELLIGENCE AND MACHINE LEARNING**



###### SUBMITTED BY:

**MANDA MANASA (HTNO: 22TR1A7333) MAMIDI SWITHIKA (HTNO: 22TR1A7332) VANKAYALA SAIVARUNTEJA (HTNO: 22TR17359) MOHAMMED ISMAIL ANAS (HTNO:22TR1A7340)**

Under the Guidance of

**Dr. K. CHANDRASENA CHARY**

**TPO & PROFESSOR**

**DEPARTMENT OF ARTIFICIAL INTELLIGENCE AND MACHINE LEARNING**

**SREE CHAITANYA INSTITUTE OF TECHNOLOGICAL SCIENCES**

**LMD COLONY, KARIMNAGAR-505527**

**(Approved by AICTE, Affiliated to JNTUH, Hyderabad) JUNE-2025**

**SREE CHAITANYA INSTITUTE OF TECHNOLOGICAL SCIENCES**

###### LMD COLONY, KARIMNAGAR-505527

**(Approved by AICTE, Affiliated to JNTUH)**

Certificate

****

Certified that this Mini Project Report entitled, **“PREDICTING SOCIAL MEDIA CYBERBULLYING IN THE BIG DATA ERA WITH MACHINE LEARNING ALGORITHMS”** is the bonafide work of **M.MANASA (H.T.No.22TR1A7333), M.SWITHIKA (H.T.No.22TR1A7332), V.SAI VARUNTEJA (H.T.No. 22TR1A7359)** and **MD.ISMAIL ANAS (H.T.No. 22TR1A7340) of III Year, AIML** in the year 2025 in partial fulfillment of the requirements to award the Degree of Bachelor of Technology in **ARTIFICIAL INTELLIGENCE AND MACHINE LEARNING** of Sree Chaitanya Institute of Technological Sciences, Karimnagar.

**Dr. K. CHANDRASENA CHARY Dr. CHADA SAMPATH REDDY**

**TPO & PROFESSOR PROFESSOR**

**Project Guide Head of the Department**

**External Examiner**

# ACKNOWLEDGEMENT

We would like to express our sincere gratitude to our Project Guide,

**Dr. K. CHANDRASENA CHARY*, Professor*** whose knowledge and guidance has motivated us to achieve goals we never thought possible. The time we have spent working under his supervision has truly been a pleasure.

We are thankful to **Mrs. SHAGUFTHA BASHEER, Associate Professor & Project Coordinator** of AIML Department for her effort, guidance and all faculty members of AIML Department for their help during my course. Thanks to programmers and non-teaching staff of AIML Department, Sree Chaitanya Institute of Technological Sciences.

We also thank **Dr. CHADA SAMPATH REDDY, HOD & Professor of AIML Department** for providing seamless support and knowledge for the entire project work and also for providing right suggestions at every phase of the development of the project. He has consistently been a source of motivation, encouragement, and inspiration.

It is a great pleasure to convey our thanks to our principal **Dr.A.PRASAD RAJU, Principal**, Sree Chaitanya Institute of Technological Sciences and the College Management for permitting us to undertake this project and providing excellent facilities to carry out our project work.

Finally Special thanks to our parents, sisters and brothers for their support and encouragement throughout my life and this course. Thanks to all our friends and well wishers for their constant support.

# DECLARATION

We hereby declare that the work which is being presented in this report entitled, “**PREDICTING SOCIAL MEDIA CYBERBULLYING IN THE BIG DATA ERA WITH MACHINE LEARNING ALGORITHMS**”, submitted towards the partial fulfillment of the requirements for the award of the degree of **Bachelor of Technology** in **AIML**, Sree Chaitanya Institute of Technological Sciences, Karimnagar is an authentic record of our own work carried out under the supervision of **DR.K. CHANDRASENA CHARY, Professor, Department of AIML,** Sree Chaitanya Institute of Technological Sciences, Karimnagar.

To the best of our knowledge and belief, this project bears no resemblance with any report submitted to Sree Chaitanya Institute of Technological Sciences or any other University for the award of any degree or diploma.

**M.MANASA**

**H.T.No: 22TR1A7333 M.SWITHIKA**

**H.T.No:22TR1A733**

**V.SAI VARUNTEJA**

**H.T.No: 22TR1A7259**

**MD.ISMAIL ANAS**

**H.T.No: 22TR1A7340**

# ABSTRACT

Prior to the innovation of information communication technologies (ICT), social interactions evolved within small cultural boundaries such as geo spatial locations. The recent developments of communication technologies have considerably transcended the temporal and spatial limitations of traditional communications. These social technologies have created a revolution in user- generated information, online human networks, and rich human behavior- related data. However, the misuse of social technologies such as social media (SM) platforms, has introduced a new form of aggression and violence that occurs exclusively online. A new means of demonstrating aggressive behavior in SM websites are highlighted in this paper. The motivations for the construction of prediction models to fight aggressive behavior in SM are also outlined. We comprehensively review cyberbullying prediction models and identify the main issues related to the construction of cyberbullying prediction models in SM. This paper provides insights on the overall process for cyberbullying detection and most importantly overviews the methodology. Though data collection and feature engineering process has been elaborated, yet most of the emphasis is on feature selection algorithms and then using various machine learning algorithms for prediction of cyberbullying behaviors. Finally, the issues and challenges have been highlighted as well, which present new research directions for researchers to explore.

# LIST OF FIGURES

|  |  |  |
| --- | --- | --- |
| **S.No** | **Figure Description** | **Page No** |
| 1 | Fig:3.3: System Architecture | 19 |
| 2 | Fig:3.4: Data Flow Diagram | 20 |
| 3 | Fig:3.5: Flow Chart1: User | 21 |
| 4 | Fig:3.5: Flow Chart2: Admin | 22 |
| 5 | Fig:3.6: Use Case Diagram | 23 |
| 6 | Fig:3.7: Class Diagram | 24 |
| 7 | Fig:3.8: Sequence Diagram | 25 |
| 8 | Fig:5.1: Home Page | 29 |
| 9 | Fig:5.2: Server Login | 29 |
| 10 | Fig:5.3: User Login | 30 |
| 11 | Fig:5.4: User Registration | 30 |
| 12 | Fig:5.5: Registration Successful | 31 |
| 13 | Fig:5.6: Server Authorization | 31 |
| 14 | Fig:5.7: Uploading Post | 32 |
| 15 | Fig:5.8: Uploading a Normal Comment | 32 |
| 16 | Fig:5.9: Uploading a Cyberbullying | 33 |
| 17 | Comments  Fig:5.10: Comment details | 33 |
| 18 | Fig:5.11: Cyberbullying Comments | 34 |

#### TABLE OF CONTENTS

Abstract V

[List of Figures VI](#_TOC_250019)

Page Nos Start – End

[Chapter 1: INTRODUCTION 9-10](#_TOC_250018)

[Chapter 2: PROJECT ANALYSIS 10-16](#_TOC_250017)

* 1. [: EXISTING SYSTEM 11-12](#_TOC_250016)
  2. [: PROPOSED SYSTEM 12-13](#_TOC_250015)
  3. [: SYSTEM REQUIREMENTS 13](#_TOC_250014)
  4. [: PRELIMINARY INVESTIGATION 14](#_TOC_250013)

2.5: FEASIBILITY ANALYSIS………………………………………………………...15-16

[Chapter 3: PROJECT DESIGN… 17-25](#_TOC_250011)

* 1. [: INPUT DESIGN 17](#_TOC_250010)
  2. [: OUTPUT DESIGN 18](#_TOC_250009)
  3. [: SYSTEM ARCHITECTURE 19](#_TOC_250008)
  4. [: DATA FLOW DIAGRAM 20](#_TOC_250007)
  5. [: FLOW CHART1:USER 21](#_TOC_250006)
  6. : FLOW CHART2: ADMIN 22
  7. [: USE CASE DIAGRAM 23](#_TOC_250005)
  8. [: CLASS DIAGRAM 24](#_TOC_250004)
  9. [: SEQUENCE DIAGRAM 25](#_TOC_250003)

[Chapter 4: IMPLEMENTATION… …..26-28](#_TOC_250002)

[Chapter 5: RESULTS… …..29-34](#_TOC_250001)

[Chapter 6: TESTING …..35-41](#_TOC_250000)

* 1. : SYSTEM TESTING 35
  2. : TYPES OF TESTING 35
     1. : UNIT TESTING 35
     2. : INTEGRATION TESTING 35-36
     3. : FUNCTIONAL TESTING ..36
     4. : SYSTEM TESTING .36
     5. : WHITEBOX TESTING 37
     6. : BLACKBOX TESTING …37
  3. : TESTING METHODOLOGIES ….37

Chapter 7: CONCLUSION… 43

REFERENCES… 44-47

#### CHAPTER 1 INTRODUCTION

Machine or deep learning algorithms help researchers understand big data [1]. Abundant information on humans and their societies can be obtained in this big data era, but this acquisition was previously impossible [2]. One of the main sources of human-related data is social media (SM). By applying machine learning algorithms to SM data, we can exploit historical data to predict the future of a wide range of applications. Machine learning algorithms provide an opportunity to effectively predict and detect negative forms of human behavior, such as cyberbullying [3]. Big data analysis can uncover hidden knowledge through deep learning from raw data [1]. Big data analytics has improved several applications,

and forecasting the future has even become possible through the combination of big data and machine learning algorithms [4].

An insightful analysis of data on human behavior and interaction to detect and restrain aggressive behavior involves multifaceted angles and aspects and the merging of theorems and techniques from multidisciplinary and interdisciplinary

\_elds. The accessibility of large-scale data produces new research questions, novel computational methods, interdisciplinary approaches, and outstanding opportunities to discover several vital inquiries quantitatively. However, using traditional methods (statistical methods) in this context is challenging in terms of scale and accuracy. These methods are commonly based on organized data on human behavior and small-scale human networks (traditional social networks). Applying these methods to large online social networks (OSNs) in terms of scale and extent causes several issues. On the one hand, the explosive growth of OSNs enhances and disseminates aggressive forms of behavior by providing platforms and networks to commit and propagate such behavior. On the other hand, OSNs

offer important data for exploring human behavior and interaction at a large scale, and these data can be used by researchers to develop effective methods of detecting and restraining misbehavior and/or aggressive behavior. OSNs provide criminals with tools to perform aggressive actions and networks to commit misconduct. Therefore, methods that address both aspects (content and network) should be optimized to detect and restrain aggressive behavior in complex systems.

#### CHAPTER 2 PROJECT ANALYSIS

#### EXISTING SYSTEM

* + - State-of-the-art research has developed features to improve the performance of cyberbullying prediction. For example, a lexical syntactic feature has been proposed to deal with the prediction of offensive language; this method is better than traditional learning-based approaches in terms of precision . Dadvar *et al.* examined gender information from profile information and developed a gender-based approach for cyberbullying prediction by using datasets from Myspace as a basis. The gender feature was selected to improve the discrimination capability of a classifier. Age and gender were included as features in other studies, but these features are limited to the information provided by users in their online profiles.
    - Several studies focused on cyberbullying prediction based on profane words as a feature. Similarly, a lexicon of profane words was constructed to indicate bullying, and these words were used as features for input to machine learning algorithms. Using profane words as features demonstrates a significant improvement in model performance. For example, the number of ``bad'' words and the density of ``bad'' words were proposed as features for input to machine learning in a previous work. The study concluded that the percentage of ``bad'' words in a text is indicative of cyberbullying. Another research expanded a list of pre-defined profane words and allocated different weights to create bullying features. These features were concatenated with bag-of-words and latent semantic features and used as a feature input for a machine learning algorithm.

##### Disadvantages

* + - * The System is not much affective due to Semi supervised machine learning techniques.
        + The system doesn’t have sentiment classification for cyberbullying.

#### PROPOSED SYSTEM

* The proposed system is constructing cyberbullying prediction models is to use a text classification approach that involves the construction of machine learning classifiers from labeled text instances. Another means is to use a lexicon-based model that involves computing orientation for a document from the semantic orientation of words or phrases in the document. Generally, the lexicon in lexicon-based models can be constructed manually or automatically by using seed words to expand the list of words. However, cyberbullying prediction using the lexicon-based approach is rare in literature.
* The primary reason is that the texts on SM websites are written in an unstructured manner, thus making it difficult for the lexicon-based approach to detect cyberbullying based only on lexicons. However, lexicons are used to extract features, which are often utilized as inputs to machine learning algorithms. For example, lexicon based approaches, such as using a profane- based dictionary to detect the number of profane words in a post, are adopted as profane features to machine learning models. The key to effective cyberbullying prediction is to have a set of features that are extracted and engineered.

#### SYSTEM REQUIREMENTS

**Hardware Requirements:-**

* Processor - Pentium –IV
* RAM - 16GB
* Hard Disk - 512 GB
* Key Board - Standard Windows Keyboard
* Mouse - Two or Three Button Mouse
* Monitor - SVGA

**Software Requirements:**

* + Operating System - Windows 10
  + Coding Language - Java/J2EE(JSP,Servlet)
  + Front End - J2EE
  + Back End - MySQL

#### PRELIMINARY INVESTIGATION

The first and foremost strategy for development of a project starts from the thought of designing a mail enabled platform for a small firm in which it is easy and convenient of sending and receiving messages, there is a search engine ,address book and also including some entertaining games. When it is approved by the organization and our project guide the first activity, ie. preliminary investigation begins. The activity has three parts:

* **Request Clarification**
* **Feasibility Study**
* **Request Approval**

#### REQUEST CLARIFICATION

After the approval of the request to the organization and project guide, with an investigation being considered, the project request must be examined to determine precisely what the system requires

Here our project is basically meant for users within the company whose systems can be interconnected by the Local Area Network(LAN). In today’s busy schedule man need everything should be provided in a readymade manner. So taking into consideration of the vastly use of the net in day to day life, the corresponding development of the portal came into existence.

#### 2.5 FEASIBILITY ANALYSIS

An important outcome of preliminary investigation is the determination that the system request is feasible. This is possible only if it is feasible within limited resource and time. The different feasibilities that have to be analyzed are

* **Operational Feasibility**
* **Economic Feasibility**
* **Technical Feasibility**

##### Operational Feasibility

Operational Feasibility deals with the study of prospects of the system to be developed. This system operationally eliminates all the tensions of the Admin and helps him in effectively tracking the project progress. This kind of automation will surely reduce the time and energy, which previously consumed in manual work. Based on the study, the system is proved to be operationally feasible.

##### Economic Feasibility

Economic Feasibility or Cost-benefit is an assessment of the economic justification for a computer based project. As hardware was installed from the beginning & for lots of purposes thus the cost on project of hardware is low. Since the system is a network based, any number of employees connected to the LAN within that organization can use this tool from at anytime. The Virtual Private Network is to be developed using the existing resources of the organization. So the project is economically feasible.

##### Technical Feasibility

According to Roger S. Pressman, Technical Feasibility is the assessment of the technical resources of the organization. The organization needs IBM compatible machines with a graphical web browser connected to the Internet and Intranet. The system is developed for platform Independent environment.

Java Server Pages, JavaScript, HTML, SQL server and WebLogic Server are used to develop the system. The technical feasibility has been carried out. The system is technically feasible for development and can be developed with the existing facility.

#### REQUEST APPROVAL

Not all request projects are desirable or feasible. Some organization receives so many project requests from client users that only few of them are pursued. However, those projects that are both feasible and desirable should be put into schedule. After a project request is approved, it cost, priority, completion time and personnel requirement is estimated and used to determine where to add it to any project list. Truly speaking, the approval of those above factors, development works can be launched.

#### CHAPTER 3 PROJECT DESIGN

#### INPUT DESIGN

Input Design plays a vital role in the life cycle of software development, it requires very careful attention of developers. The input design is to feed data to the application as accurate as possible. So inputs are supposed to be designed effectively so that the errors occurring while feeding are minimized. According to Software Engineering Concepts, the input forms or screens are designed to provide to have a validation control over the input limit, range and other related validations.

This system has input screens in almost all the modules. Error messages are developed to alert the user whenever he commits some mistakes and guides him in the right way so that invalid entries are not made. Let us see deeply about this under module design.

Input design is the process of converting the user created input into a computer-based format. The goal of the input design is to make the data entry logical and free from errors. The error is in the input are controlled by the input design. The application has been developed in user-friendly manner. The forms have been designed in such a way during the processing the cursor is placed in the position where must be entered. The user is also provided with in an option to select an appropriate input from various alternatives related to the field in certain cases.

Validations are required for each data entered. Whenever a user enters an erroneous data, error message is displayed and the user can move on to the subsequent pages after completing all the entries in the current page.

#### OUTPUT DESIGN

The Output from the computer is required to mainly create an efficient method of communication within the company primarily among the project leader and his team members, in other words, the administrator and the clients. The output of VPN is the system which allows the project leader to manage his clients in terms of creating new clients and assigning new projects to them, maintaining a record of the project validity and providing folder level access to each client on the user side depending on the projects allotted to him. After completion of a project, a new project may be assigned to the client. User authentication procedures are maintained at the initial stages itself. A new user may be created by the administrator himself or a user can himself register as a new user but the task of assigning projects and validating a new user rests with the administrator only.

The application starts running when it is executed for the first time. The server has to be started and then the internet explorer in used as the browser. The project will run on the local area network so the server machine will serve as the administrator while the other connected systems can act as the clients. The developed system is highly user friendly and can be easily understood by anyone using it even for the first time.

#### SYSTEM ARCHITECTURE



Admin

1. List all users and authorize
2. List all Friends Req and Res
3. View Friend Req and res
4. Add Filters( By adding category and its words) and view all filter words
5. View all posts ie messages or images
6. Detect Cyber Bulling Users (those who sent review about post – fine number of items found for a corresponding post like violence (No. Of words),Vulgur(No.of Words),Offensive(No. Of Words),Hate(No.Of Words, Sexual(No.of Words)))
7. Find Cyber bulling reviews chart

**Network User**

Authorize

**Store and retrievals**

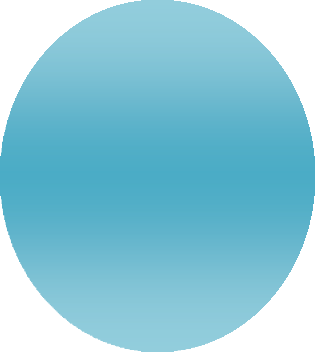
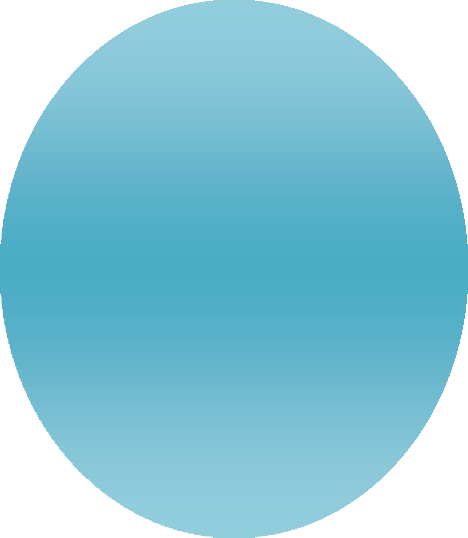
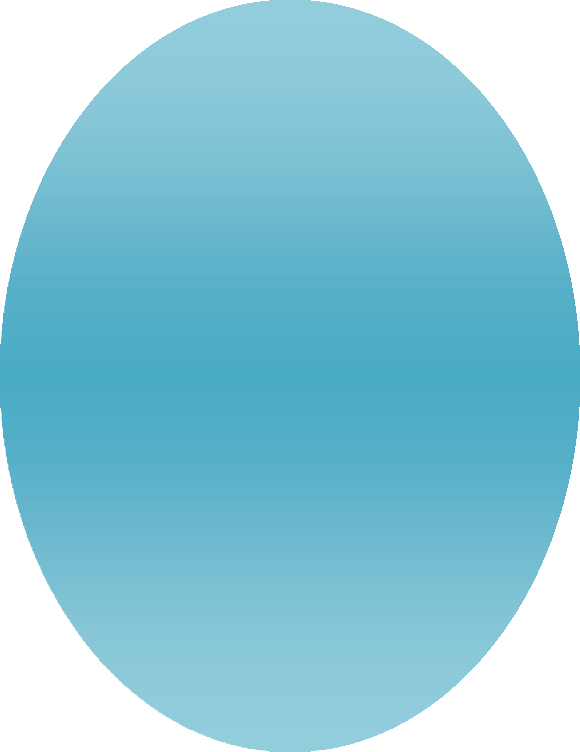
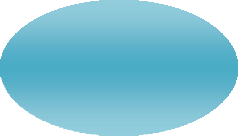
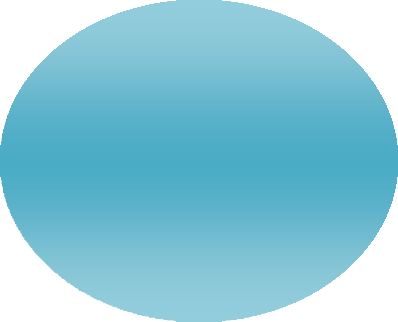
* + 1. Register and Login

**WEB**

**Database**

* + 1. View your Details and search friends, req
    2. Post your messages with images like title, title description with browse option, uses, Title image.
    3. View and Comment on your friend post
    4. View all your friends post and comment(don’t post if comment consist cyber bulling and show the reason why comment is not posted by indicating Detect Cyber Bulling words ( fine number of items found for a corresponding post like violence (No. Of words),Vulgur(No.of Words),Offensive(No. Of Words),Hate(No.Of Words, Sexual(No.of Words)))
    5. View all your cyber bulling comments on your friend posts

##### Data Flow Diagram



Web Server

Login, Add Filters ( By adding category and its words) and view all filter words

Admin

View all users and authorize

Request Response

Register and Login

Post your messages with images like title, title description with browse

Request

View your Details and search friends

List all users and authorize, List all Friends Req and Res,View Friend Req and res, View all posts ie messages or images, Detect Cyber Bulling Users (those who sent review about post – fine number of items found for a corresponding post like violence (No. Of words),Vulgur(No.of Words),Offensive(No. Of

Words),Hate(No.Of Words,



End User

req ,View and Comment on your friend post ,View all your friends post and comment(don’t post if comment consist cyber bulling and show the reason why comment

Sexual(No.of Words))),Find Response

##### Flow Chart1: User

**Start**

**User Register**

**Login lLogin**

##### 

**NO**

**YES**

##### 

Search friends, req Friends

**Username & Password Wrong**

##### 

##### YES

**NO**

##### 

**View all friends and posts**

**Log Out**

##### 

##### 

Post your messages

**Request for Keys**

View all your cyber bulling comments on your friend posts

No

**Flow Chart 2: Admin**

**Start**

**Admin**

**YES**

**Login**

**NO**



**Username & Password Wrong**



**YES**

**NO**

**Detect Cyber Bulling Users**

**List all users and authorize**

**View all posts ie messages or images**

Add Filters on Messages

**List Attackers**

**Log Out**

**List all Friends Req and Res**

#### USE CASE DIAGRAM

**Social Network**



**User**

Register And Login

**Admin**

List all users and

Search friend and Req Friend

List all Friends Req and

View Friend’s Post and make comments

Add Filters

View all your cyber bulling comments on your friend posts

View all posts ie messages or images

Detect Cyber Bulling Users

#### CLASS DIAGRAM

## Web Server -- Admin

Method

Name, Password, DOB, Gender, Address, City, Country, Email, Mobile, Image, Pincode, title, title description with browse option, uses, Title image.

List all users and authorize

,List all Friends Req and Res

,View Friend Req and res

,Add Filters ,View all posts ie messages or images ,Detect cyber bulling attackers

Users, contents, Search\_ Details,

Members

## Login

Method

User Name, Password

Login (), Register (), Reset ()

Name, Password, DOB,

Gender, Address, City, Country, Email, Mobile, Image, Pincode

Register (), Reset ()

**Register**

Method

Members

Members

## End User

Method

title, title description with browse option, uses, Title image.

Register and Login,View your Details, search friends, Post your messages,View Comments on posts View all your cyber bulling attacks,Logout()

Members

#### SEQUENCE DIAGRAM



Network User

Admin

### Check Login Status Login confirmation

**Add Posts**

### Receive Status

Added Post confirmation

View and Comment on your friend post& Make Comment

View all your cyber bulling comments on your friend posts

List all users and authorize

List all Friends Req and Res

Add filters

Detect Cyber Bullying users

View all posts ie

messages or

images

#### CHAPTER 4 IMPLEMENTATION

###### MODULES:

* **Admin**

In this module, the Admin has to login by using valid user name and password. After login successful he can perform some operations such as view and authorize users, view all friends request and responses, Add and View Filters, View all posts, Detect Cyber Bullying Users, Find Cyber Bullying Reviews Chart.

**Viewing and Authorizing Users**

In this module, the admin views all users details and authorize them for login permission. User Details such as User Name, Address, Email Id, Mobile Number.

**Viewing all Friends Request and Response**

In this module, the admin can see all the friends’ requests and response history. Details such as Requested User Name and Image, and Requested to User Name and Image, status and date.

**Add and View Filters**

In this module, the admin can add filters (like Violence, Vulgar, Offensive, Hate, and Sexual) as Categories with the words those related to corresponding filters.

**View all posts**

In this module, the admin can see all the posts added by the users with post details like post name, description and post image.

**Detect Cyber Bullying Users**

In this module, the admin can see all the Cyber Bullying Users (The users who had posted a comment on posts using cyber bullying words which are all listed by the admin to detect and filter). In this, the results shown as, Number of items found for a corresponding post like Violence (no. of words belongs to Violence Filter used in comments by the users), Vulgar (no. of words belongs to Vulgar Filter used in comments by the users), Offensive (no. of words belongs to Offensive Filter used in comments by the users), Hate (no. of words belongs to Hate Filter used in comments by the users), Sexual (no. of words belongs to Sexual Filter used in comments by the users).

**Find Cyber Bullying Reviews Chart**

In this module, the admin can see all the posts with number of cyber bullying comments posted by users for particular post.

* **User**

In this module, there are n numbers of users are present. User should register before performing any operations. Once user registers, their details will be stored to the database. After registration successful, he has to login by using authorized user name and password. Once Login is successful user can perform some operations like viewing their profile details, searching for friends and sending friend requests, Posting Your Messages as Posts by giving details, View and Comment on Friend Posts, viewing all friends posts and comment, view all your cyber bullying comments on your friend posts.

**Viewing Profile Details, Search and Request Friends**

In this module, the user can see their own profile details, such as their address, email, mobile number, profile Image.

The user can search for friends and can send friend requests or can accept friend requests.

**Add Posts**

In this, the user can add their own posts by giving post details such as, post title, description, uses, and image of post.

**View and Comment on Your Friends Post**

In this, the user can see his entire friend’s post details (post title, description, uses, creator and image of post) and can comment on posts.

**View all Friends Posts and Comment (Cyber bullying Related)**

In this, the user can see his all friend’s post details (post title, description, uses, creator and image of post) and can comment on posts.

Don’t Post If the comment consists of Cyber bullying words and Shows the reason why comment is not posted by indicating Detected Cyber Bullying Words like Numbers of Cyber Bullying words Related to Filter Violence found in comment, Numbers of Cyber Bullying words Related to Filter Vulgar found in comment, Numbers of Cyber Bullying words Related to Offensive found in comment, Numbers of Cyber Bullying words Related to Hate found in comment, Numbers of Cyber Bullying words Related to Sexual found in comment,

**View all Your Cyber bullying comments on your friend posts**

The user can see all his posted cyber bullying comments on their friend created posts.

#### CHAPTER 5 RESULTS

****

Fig 1: Home Page

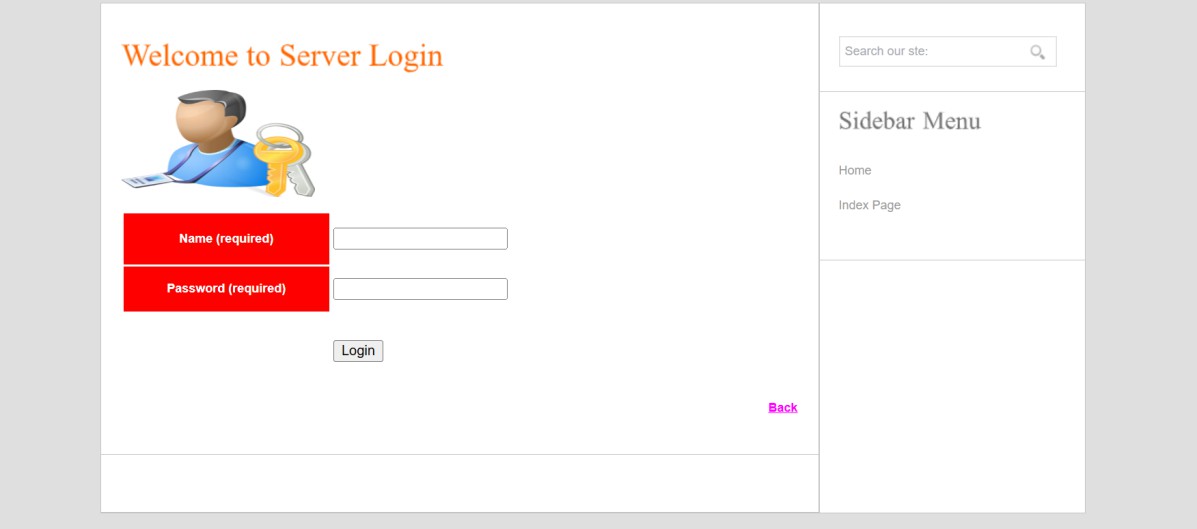


Fig 2: Server Login

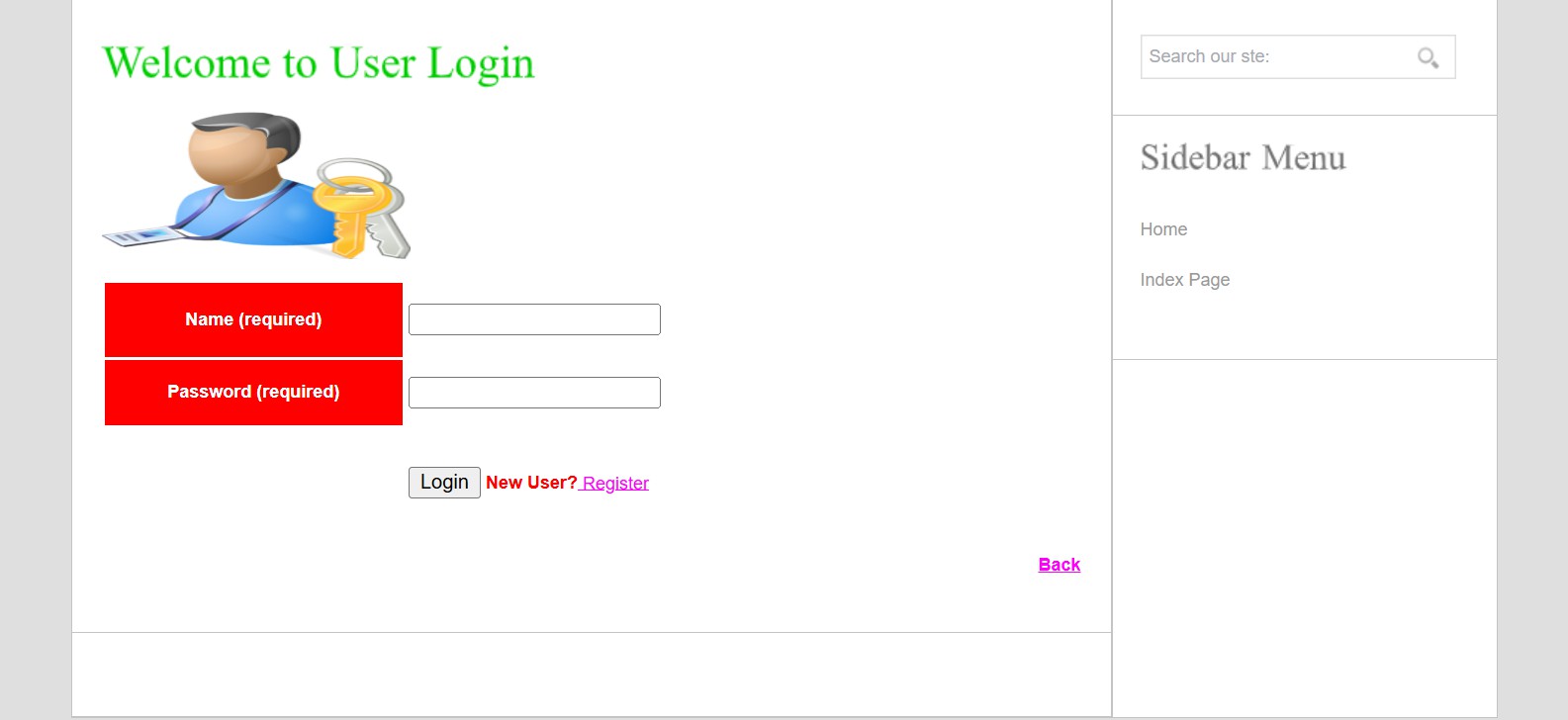


Fig 3: User Login

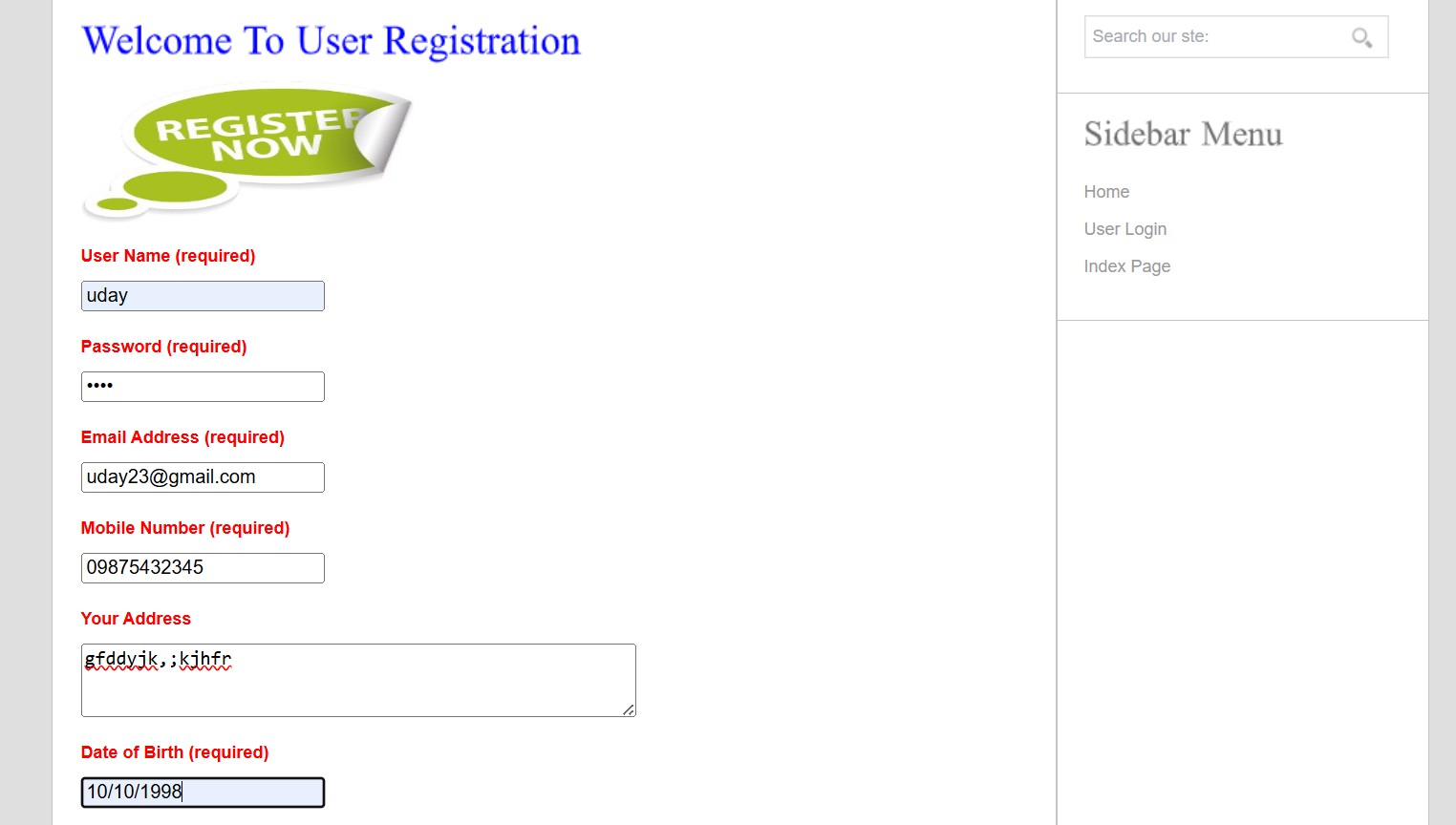


Fig 4: User Registration



Fig 5: Registration Successful



Fig 6: Server Authorization

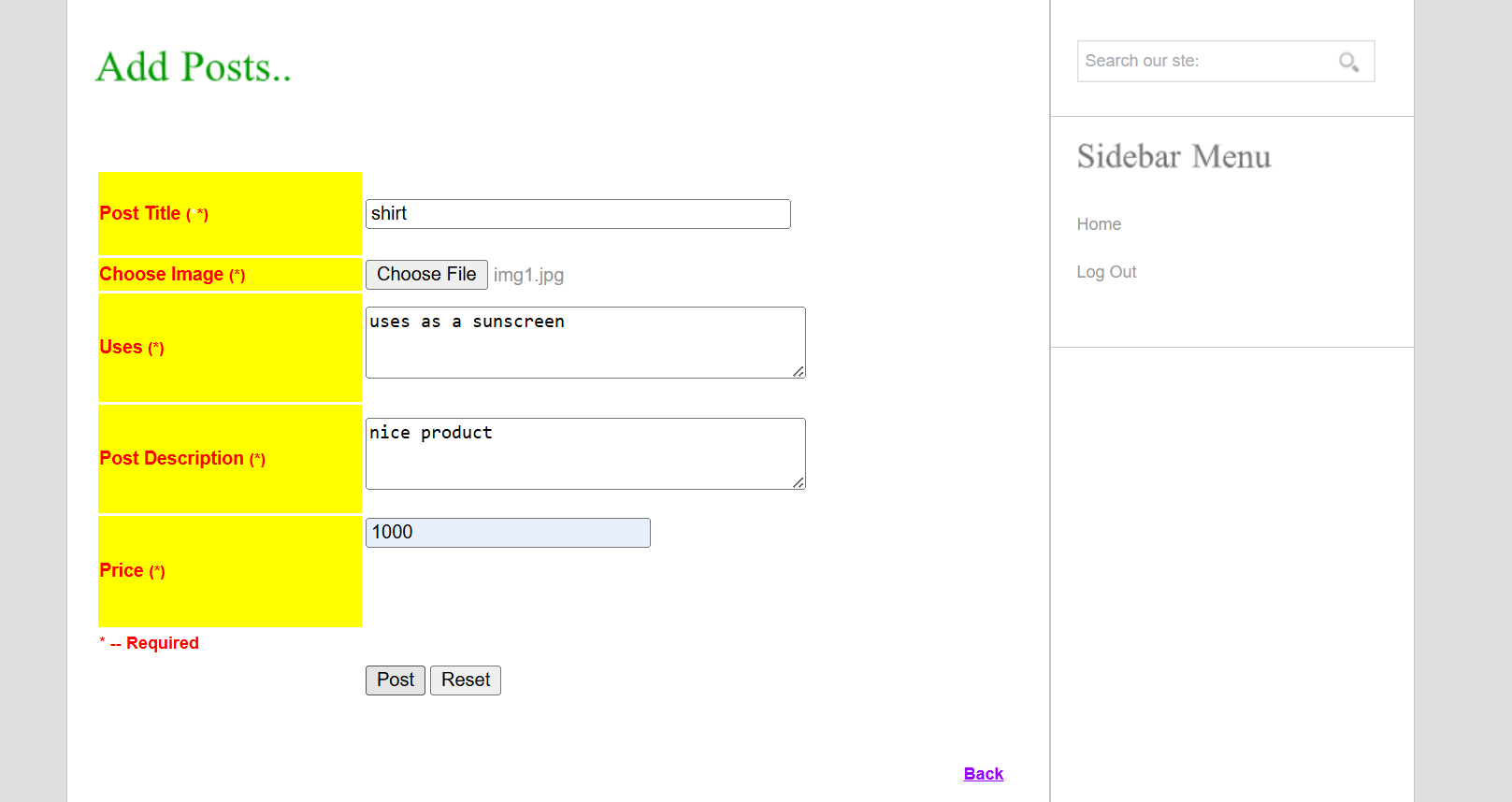


Fig 7: Uploading Post

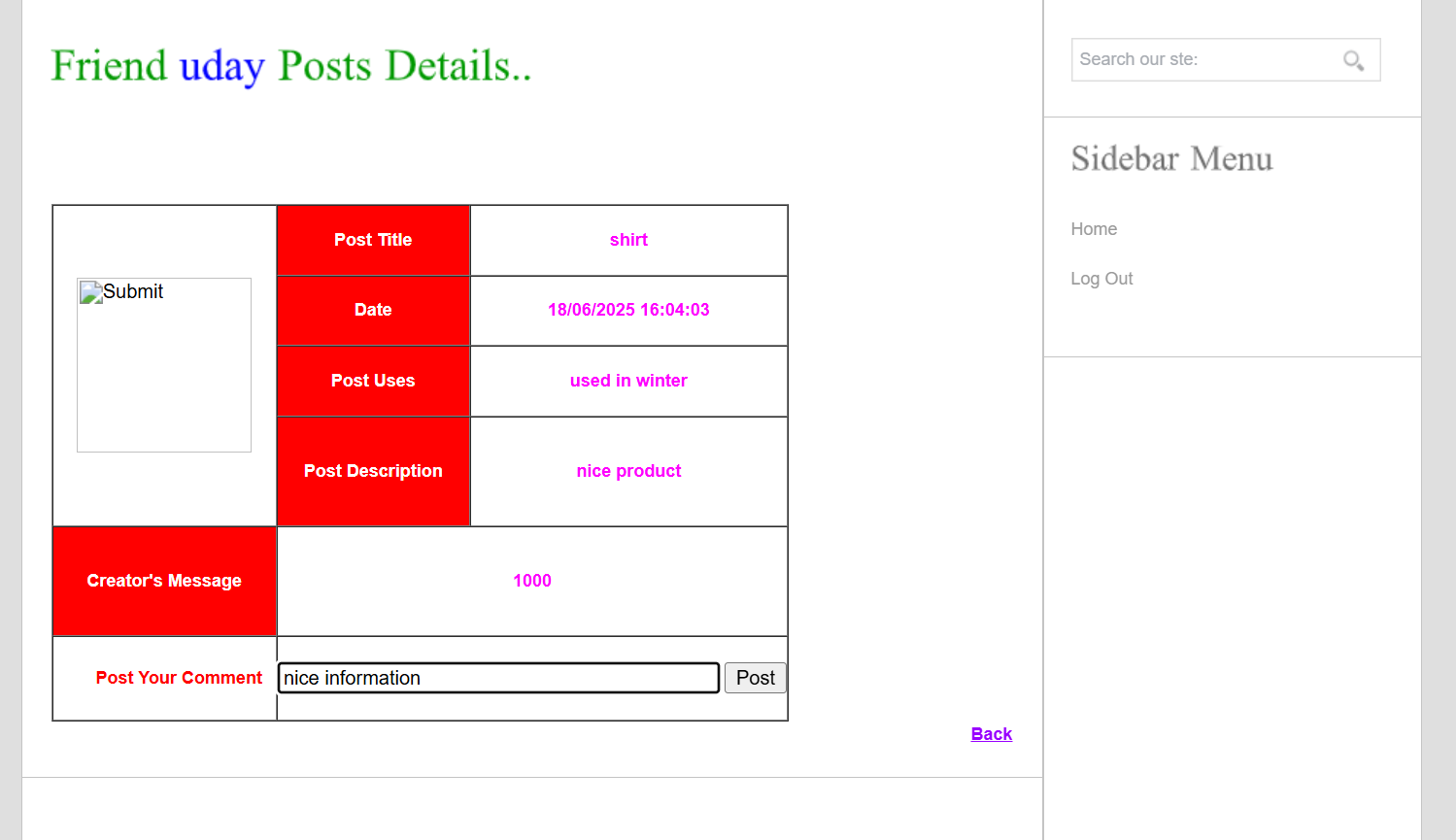


Fig 8: Uploading a Normal Comment

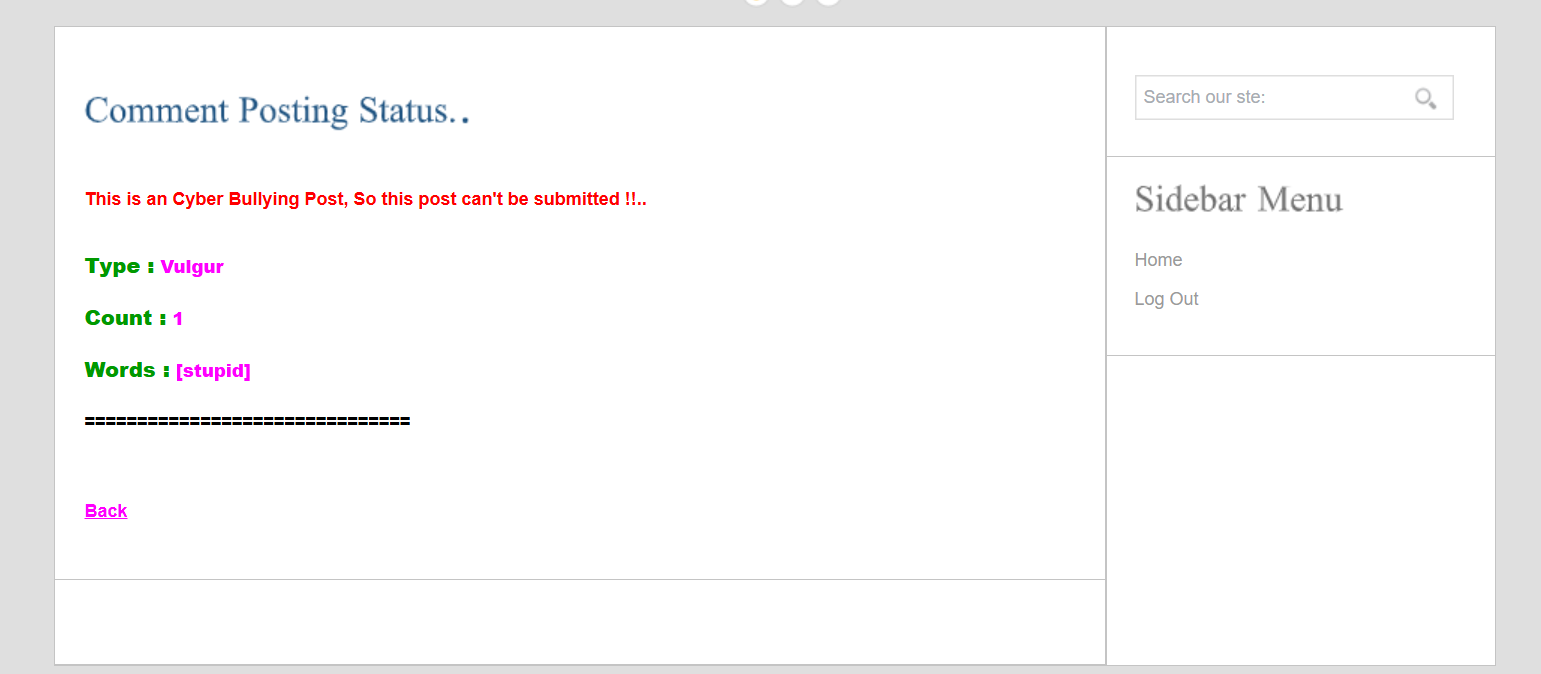


Fig 9: Uploading Cyberbullying Comment



Fig 10: Comment Details



Fig 11: Cyberbullying Comments

* 1. **: SYSTEM TESTING**

#### CHAPTER 6 TESTING

The purpose of testing is to discover errors. Testing is the process of trying to discover every conceivable fault or weakness in a work product. It provides a way to check the functionality of components,sub assemblies, assemblies and/or a finished product It is the process of exercising software with the intent of ensuring that the Software system meets its requirements and user expectations and does not fail in an unacceptable manner.There are various types of test. Each test type addresses a specific testing requirement.

###### : TYPES OF TESTING

* + 1. **: Unit testing**

Unit testing involves the design of test cases that validate that the internal program logic is functioning properly, and that program inputs produce valid outputs. All decision branches and internal code flow should be validated. It is the testing of individual software units of the application .it is done after the completion of an individual unit before integration. This is a structural testing, that relies on knowledge of its construction and is invasive. Unit tests perform basic tests at component level and test a specific business process, application, and/or system configuration. Unit tests ensure that each unique path of a business process performs accurately to the documented specifications and contains clearly defined inputs and expected results.

* + 1. **: Integration testing**

Integration tests are designed to test integrated software components to determine if they actually run as one program. Testing is event driven and is more concerned with the basic outcome of screens or fields. Integration tests demonstrate that although the components were individually satisfaction, as

combination of components is correct and consistent. Integration testing is specifically aimed at exposing the problems that arise from the combination of components.

* + 1. **: Functional Testing**

Functional tests provide systematic demonstrations that functions tested are available as specified by the business and technical requirements, system documentation, and user manuals.

Functional testing is centered on the following items:

Valid Input : identified classes of valid input must be accepted. Invalid Input : identified classes of invalid input must be rejected. Functions : identified functions must be exercised.

Output : identified classes of application outputs must be exercised.

Systems/Procedures: interfacing systems or procedures must be invoked.

0rganization and preparation of functional tests is focused on requirements, key functions, or special test cases. In addition, systematic coverage pertaining to identify Business process flows; data fields, predefined processes, and successive processes must be considered for testing. Before functional testing is complete, additional tests are identified and the effective value of current tests is determined.

* + 1. **: System Testing**

System testing ensures that the entire integrated software system meets requirements. It tests a configuration to ensure known and predictable results. An example of system testing is the configuration oriented system integration test. System testing is based on process descriptions and flows, emphasizing pre-driven process links and integration points.

* + 1. **: White Box Testing**

White Box Testing is a testing in which in which the software tester has knowledge of the inner workings, structure and language of the software, or at least its purpose. It is purpose. It is used to test areas that cannot be reached from a black box level.

* + 1. **: Black Box Testing**

Black Box Testing is testing the software without any knowledge of the inner workings, structure or language of the module being tested. Black box tests, as most other kinds of tests, must be written from a definitive source document, such as specification or requirements document, such as specification or requirements document. It is a testing in which the software under test is treated, as a black box .you cannot “see” into it. The test provides inputs and responds to outputs without considering how the software works.

###### : TESTING METHODOLOGIES

The following are the Testing Methodologies:

* **Unit Testing.**
* **Integration Testing.**
* **User Acceptance Testing.**
* **Output Testing.**
* **Validation Testing. Unit Testing**

Unit testing focuses verification effort on the smallest unit of Software design that is the module. Unit testing exercises specific paths in a module’s control structure to

ensure complete coverage and maximum error detection. This test focuses on each module individually, ensuring that it functions properly as a unit. Hence, the naming is Unit Testing.

During this testing, each module is tested individually and the module interfaces are verified for the consistency with design

specification. All important processing path are tested for the expected results. All error handling paths are also tested.

**Integration Testing**

Integration testing addresses the issues associated with the dual problems of verification and program construction. After the software has been integrated a set of high order tests are conducted. The main objective in this testing process is to take unit tested modules and builds a program structure that has been dictated by design.

**The following are the types of Integration Testing:**

1. **Top Down Integration**

This method is an incremental approach to the construction of program structure. Modules are integrated by moving downward through the control hierarchy, beginning with the main program module. The module subordinates to the main program module are incorporated into the structure in either a depth first or breadth first manner.

In this method, the software is tested from main module and individual stubs are replaced when the test proceeds downwards.

1. **Bottom-up Integration**

This method begins the construction and testing with the modules at the lowest level in the program structure. Since the modules are integrated from the bottom up, processing required for modules subordinate to a given level is always available and the need for stubs is eliminated. The bottom up integration strategy may be implemented with the following steps:

* + The low-level modules are combined into clusters into clusters that perform a specific Software sub-function.
  + A driver (i.e.) the control program for testing is written to coordinate test case input and output.
  + The cluster is tested.
  + Drivers are removed and clusters are combined moving upward in the program structure

The bottom up approaches tests each module individually and then each module is module is integrated with a main module and tested for functionality.

**User Acceptance Testing**

User Acceptance of a system is the key factor for the success of any system. The system under consideration is tested for user acceptance by constantly keeping in touch with the prospective system users at the time of developing and making changes wherever required. The system developed provides a friendly user interface that can easily be understood even by a person who is new to the system.

**Output Testing**

After performing the validation testing, the next step is output testing of the proposed system, since no system could be useful if it does not produce the required output in the specified format. Asking the users about the format required by them tests the outputs generated or displayed by the system under consideration. Hence the output format is considered in 2 ways – one is on screen and another in printed format.

**Validation Checking**

Validation checks are performed on the following fields.

**Text Field:**

The text field can contain only the number of characters lesser than or equal to its size. The text fields are alphanumeric in some tables and alphabetic in other tables. Incorrect entry always flashes and error message.

**Numeric Field:**

The numeric field can contain only numbers from 0 to 9. An entry of any character flashes an error messages. The individual modules are checked for accuracy and what it has to perform. Each module is

subjected to test run along with sample data. The individually tested modules are integrated into a single system. Testing involves executing the real data information is used in the program the existence of any program defect is inferred from the output. The testing should be planned so that all the requirements are individually tested.

A successful test is one that gives out the defects for the inappropriate data and produces and output revealing the errors in the system.

**Preparation of Test Data**

Taking various kinds of test data does the above testing. Preparation of test data plays a vital role in the system testing. After preparing the test data the system under study is tested using that test data. While testing the system by using test data errors are again uncovered and corrected by using above testing steps and corrections are also noted for future use.

**Using Live Test Data:**

Live test data are those that are actually extracted from organization files. After a system is partially constructed, programmers or analysts often ask users to key in a set of data from their normal activities. Then, the systems person uses this data as a way to partially test the system. In other instances, programmers or analysts extract a set of live data from the files and have them entered themselves.

It is difficult to obtain live data in sufficient amounts to conduct extensive testing. And, although it is realistic data that will show how the system will perform for the typical processing requirement, assuming that the live data entered are in fact typical, such data generally will not test all combinations or formats that can enter the system. This bias toward typical values then does not provide a true systems test and in fact ignores the cases most likely to cause system failure.

**Using Artificial Test Data:**

Artificial test data are created solely for test purposes, since they can be generated to test all combinations of formats and values. In other words, the artificial data, which can quickly be prepared by a data generating utility program in the information systems department, make possible the testing of all login and control paths through the program.

The most effective test programs use artificial test data generated by persons other than those who wrote the programs. Often, an independent team of testers formulates a testing plan, using the systems specifications.

The package “Virtual Private Network” has satisfied all the requirements specified as per software requirement specification and was accepted.

###### USER TRAINING

Whenever a new system is developed, user training is required to educate them about the working of the system so that it can be put to efficient use by those for whom the system has been primarily designed. For this purpose the normal working of the project was demonstrated to the prospective users. Its working is easily understandable and since the expected users are people who have good knowledge of computers, the use of this system is very easy.

###### MAINTAINENCE

This covers a wide range of activities including correcting code and design errors. To reduce the need for maintenance in the long run, we have more accurately defined the user’s requirements during the process of system development. Depending on the requirements, this system has been developed to satisfy the needs to the largest possible extent. With development in technology, it may be possible to add many more features based on the requirements in future. The coding and designing is simple and easy to understand which will make maintenance easier.

###### TESTING STRATEGY :

A strategy for system testing integrates system test cases and design techniques into a well planned series of steps that results in the successful construction of software. The testing strategy must co- operate test planning, test case design, test execution, and the resultant data collection and evaluation .A strategy for software testing must accommodate low-level tests that are necessary to verify that a small source code segment has been correctly implemented as well as high level tests that validate major system functions against user requirements.

Software testing is a critical element of software quality assurance and represents the ultimate review of specification design and coding. Testing represents an interesting anomaly for the software. Thus, a series of testing are performed for the proposed system before the system is ready for user acceptance testing.

#### CHAPTER 7 CONCLUSION

This study reviewed existing literature to detect aggressive behavior on SM websites by using machine learning approaches. We specifically reviewed four aspects of detecting cyberbullying messages by using machine learning approaches, namely, data collection, feature engineering, construction of cyberbullying detection model, and evaluation of constructed cyberbullying detection models. Several types of discriminative features that were used to detect cyberbullying in online social networking sites were also summarized. In addition, the most effective supervised machine learning classifiers for classifying cyberbullying messages in online social networking sites were identified. One of the main contributions of current paper is the definition of evaluation metrics to successfully identify the significant parameter so the various machine learning algorithms can be evaluated against each other. Most importantly we summarized and identified the important factors for detecting cyberbullying through machine learning techniques specially supervised learning. For this purpose, we have used accuracy, precision recall and f-measure which gives us the area under the curve function for modeling the behaviors in cyberbullying. Finally, the main issues and open research challenges were described and discussed.

#### REFERENCES

1. V. Subrahmanian and S. Kumar, ``Predicting human behavior: The next

frontiers,'' *Science*, vol. 355, no. 6324, p. 489, 2017.

1. H. Lauw, J. C. Shafer, R. Agrawal, and A. Ntoulas, ``Homophily in the digital world: A LiveJournal case study,'' *IEEE Internet Comput.*, vol. 14, no. 2, pp. 15\_23, Mar./Apr. 2010.
2. M. A. Al-Garadi, K. D. Varathan, and S. D. Ravana, ``Cybercrime detection

in online communications: The experimental case of cyberbullying detection in the Twitter network,'' *Comput. Hum. Behav.*, vol. 63,

pp. 433\_443, Oct. 2016.

1. L. Phillips, C. Dowling, K. Shaffer, N. Hodas, and S. Volkova, ``Using social media to predict the future: A systematic literature review,'' 2017, *arXiv:1706.06134*.[Online]. Available: <https://arxiv.org/abs/1706.06134>
2. H. Quan, J. Wu, and Y. Shi, ``Online social networks & social network services: A technical survey,'' in *Pervasive Communication Handbook*. Boca Raton, FL, USA: CRC Press, 2011, p. 4.
3. J. K. Peterson and J. Densley, ``Is social media a gang? Toward a selection,

facilitation, or enhancement explanation of cyber violence,'' *Aggression Violent Behav.*, 2016.

1. BBC. (2012). *Huge Rise in Social Media*. [Online]. Available: <http://www.bbc.com/news/uk-20851797>
2. P. A.Watters and N. Phair, ``Detecting illicit drugs on social media using

automated social media intelligence analysis (ASMIA),'' in *Cyberspace Safety and Security*. Berlin, Germany: Springer, 2012, pp. 66\_76.

1. M. Fire, R. Goldschmidt, and Y. Elovici, ``Online social networks: Threats and solutions,'' *IEEE Commun. Surveys Tuts.*, vol. 16, no. 4,

pp. 2019\_2036, 4th Quart., 2014.

1. N. M. Shekokar and K. B. Kansara, ``Security against sybil attack in social

network,'' in *Proc. Int. Conf. Inf. Commun. Embedded Syst. (ICICES)*, 2016, pp. 1\_5.

1. J. Ratkiewicz, M. Conover, M. Meiss, B. Gonçalves, A. Flammini, and

F. Menczer, ``Detecting and tracking political abuse in social media,'' in

*Proc. 5th Int. AAAI Conf. Weblogs Social Media*, 2011, pp. 297\_304.

1. A. Aggarwal, A. Rajadesingan, and P. Kumaraguru, ``PhishAri:

Automatic

realtime phishing detection on Twitter,'' in *Proc. eCrime Res. Summit (eCrime)*, Oct. 2012, pp. 1\_12.

1. S. Yardi *et al.*, ``Detecting spam in a Twitter network,'' *First Monday*, Jan. 2009. [Online]. Available: https://\_rstmonday.org/ article/view/2793/2431
2. C. Yang, R. Harkreader, J. Zhang, S. Shin, and G. Gu, ``Analyzing spammers' social networks for fun and pro\_t: A case study of cyber criminal ecosystem on twitter,'' in *Proc. 21st Int. Conf. World Wide Web*, 2012, pp. 71\_80.
3. G. R. S.Weir, F. Toolan, and D. Smeed, ``The threats of social networking:

Old wine in new bottles?'' *Inf. Secur. Tech. Rep.*, vol. 16, no. 2, pp. 38\_43, 2011.

1. M. J. Magro, ``A review of social media use in e-government,''

*Administ.*

*Sci.*, vol. 2, no. 2, pp. 148\_161, 2012.

1. M. Dadvar, D. Trieschnigg, R. Ordelman, and F. de Jong, ``Improving cyberbullying detection with user context,'' in *Advances in Information*

*Retrieval*. Berlin, Germany: Springer, 2013, pp. 693\_696.

1. Y. Chen, Y. Zhou, S. Zhu, and H. Xu, ``Detecting offensive language in social media to protect adolescent online safety,'' in *Proc. Int. Conf. Privacy, Secur., Risk Trust (PASSAT)*, Sep. 2012, pp. 71\_80.
2. V. S. Chavan and S. S. Shylaja, ``Machine learning approach for detection

of cyber-aggressive comments by peers on social media network,'' in

*Proc. Int. Conf. Adv. Comput., Commun. Inform. (ICACCI)*, Aug. 2015,

pp. 2354\_2358.

1. W. Dong, S. S. Liao, Y. Xu, and X. Feng, ``Leading effect of social media

for \_nancial fraud disclosure: A text mining based analytics,'' in *Proc. AMCIS*, San Diego, CA, USA, 2016.

1. M. S. Rahman, T.-K. Huang, H. V. Madhyastha, and M. Faloutsos,

``FRAppE: Detecting malicious Facebook applications,'' in *Proc. 8th Int. Conf. Emerg. Netw. Exp. Technol.*, 2012, pp. 313\_324.

1. S. Abu-Nimeh, T. Chen, and O. Alzubi, ``Malicious and spam posts in

online social networks,'' *Computer*, vol. 44, no. 9, pp. 23\_28, Sep. 2011.

1. B. Doerr, M. Fouz, and T. Friedrich, ``Why rumors spread so quickly in

social networks,'' *Commun. ACM*, vol. 55, no. 6, pp. 70\_75, Jun. 2012.

1. J. W. Patchin and S. Hinduja, *Words Wound: Delete Cyberbul- lying and Make Kindness Go Viral*. Golden Valley, MN, USA:

Free Spirit Publishing, 2013.

1. J. Cheng, C. Danescu-Niculescu-Mizil, and J. Leskovec, ``Antisocial behavior in online discussion communities,'' in *Proc. 9th Int. AAAI Conf. Web Social Media*, Apr. 2015.
2. S. Liu, J. Zhang, and Y. Xiang, ``Statistical detection of online drifting

Twitter spam: Invited paper,'' in *Proc. 11th ACM Asia Conf. Comput. Commun. Secur.*, 2016, pp. 1\_10.

1. Z. Miller, B. Dickinson, W. Deitrick, W. Hu, and A. H. Wang,

``Twitter

spammer detection using data stream clustering,'' *Inf. Sci.*, vol. 260,

pp. 64\_73, Mar. 2014.

1. M. Jiang, S. Kumar, V. S. Subrahmanian, and C. Faloutsos, ``KDD 2017

tutorial: Data-driven approaches towards malicious behavior modeling,''

*Dimensions*, vol. 19, p. 42, 2017.

1. S. Y. Jeong, Y. S. Koh, and G. Dobbie, ``Phishing detection on Twitter

streams,'' in *Proc. Paci\_c\_Asia Conf. Knowl. Discovery Data Mining*. Cham, Switzerland: Springer, 2016, pp. 141\_153.

1. I. Frommholz, H. M. Al-Khateeb, M. Potthast, Z. Ghasem, M. Shukla, and E. Short, ``On textual analysis and machine learning for cyberstalking

detection,'' *Datenbank-Spektrum*, vol. 16, no. 2, pp. 127\_135, 2016.