Cyber Bullying Detection And User Based Recommendation on OSN

MOHAMMED LIJAS C

MES20MCA-2028

PRODUCT OWNER-HYDERALI.K

TABLE OF CONTENTS

1. Introduction
2. Modules
3. Methodology
4. Developing Environment
5. Data Flow Diagram
6. User story
7. Product backlog
8. Project plan
9. Sprint plan
10. Sprint1 actual

INTRODUCTION

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 project 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. We are using Machine learning algorithms for filtering bulling images from the OSN. SIFT is a common technique to perform object recognition on images. By using SIFT, we explore the similarity of objects contained in bullied images and the similarity of objects contained in non-bullied images relatively.

MODULES

- Admin:

1. View Users
2. Add bullying words
3. Add and manage bullying and non-bullying image dataset
4. Add good words
5. View bullying words
6. View good words
7. View report
8. Add bullying and normal images

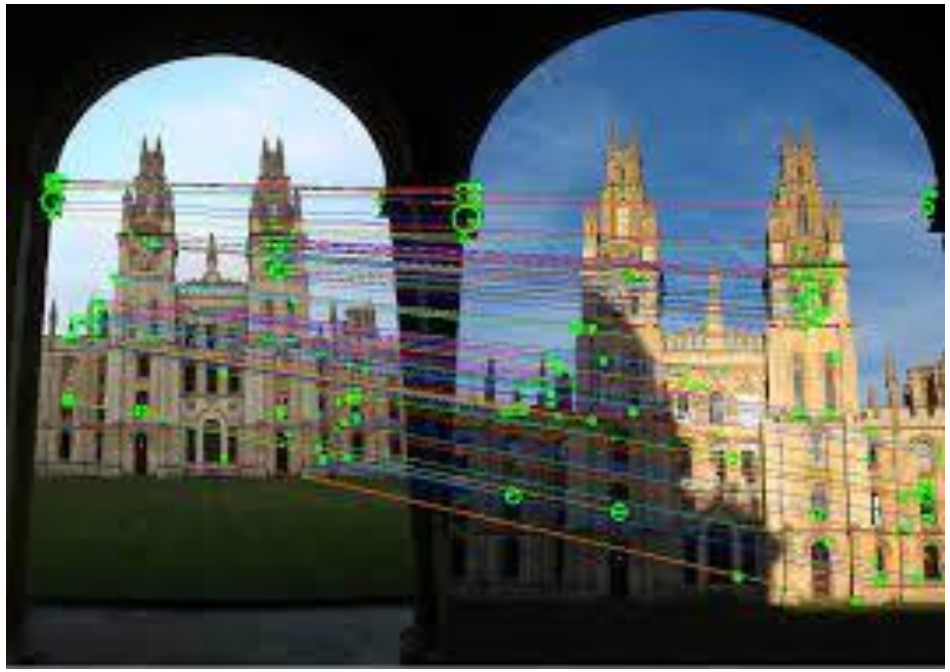
- User:

1. Registration
2. Login
3. Add post
4. View my post
5. Chat
6. Add bullying words
7. Send friend request
8. View friend request
9. View recommendation

METHODLOGY

1.SIFT(Scale-Invariant Feature Transform)

The SIFT features are **local and based on the appearance of the object at particular interest points, and are invariant to image scale and rotation**. They are also robust to changes in illumination, noise, and minor changes in viewpoint.



METHODLOGY

2. KNN matching

The Nearest-Neighbor Matching is an alternative way to stratification to match treated and comparison units. It takes each treated unit and search for the comparison unit(s)with the closest p-score. After the match, we compute the difference in the outcomes between each treated and comparison pair and average these differences across number of treated units.

3. Term Frequency Inverse Document Frequency (TF-IDF)

TF-IDF is frequently used in machine learning algorithms in various capacities, including stopword removal. These are common words like “a, the, an, it” that occur frequently but hold little informational value. TF-IDF consists of two components, term frequency, and inverse document frequency Term frequency can be determined by counting the number of occurrences of a term in a document. IDF is calculated by dividing the total number of documents by the number of documents in the collection containing the term. It’s useful for reducing the weight of terms that are common within a collection of documents

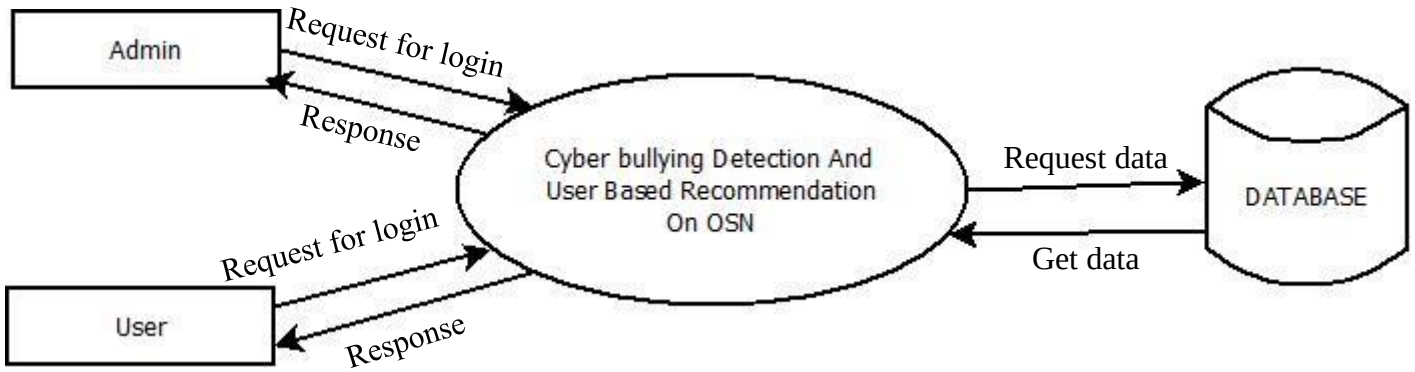
DEVELOPING ENVIRONMENT

- **Hardware Requirements**
  - Processor : i3 and above
  - RAM : 4GB
  - Storage : 500GB Hard disk

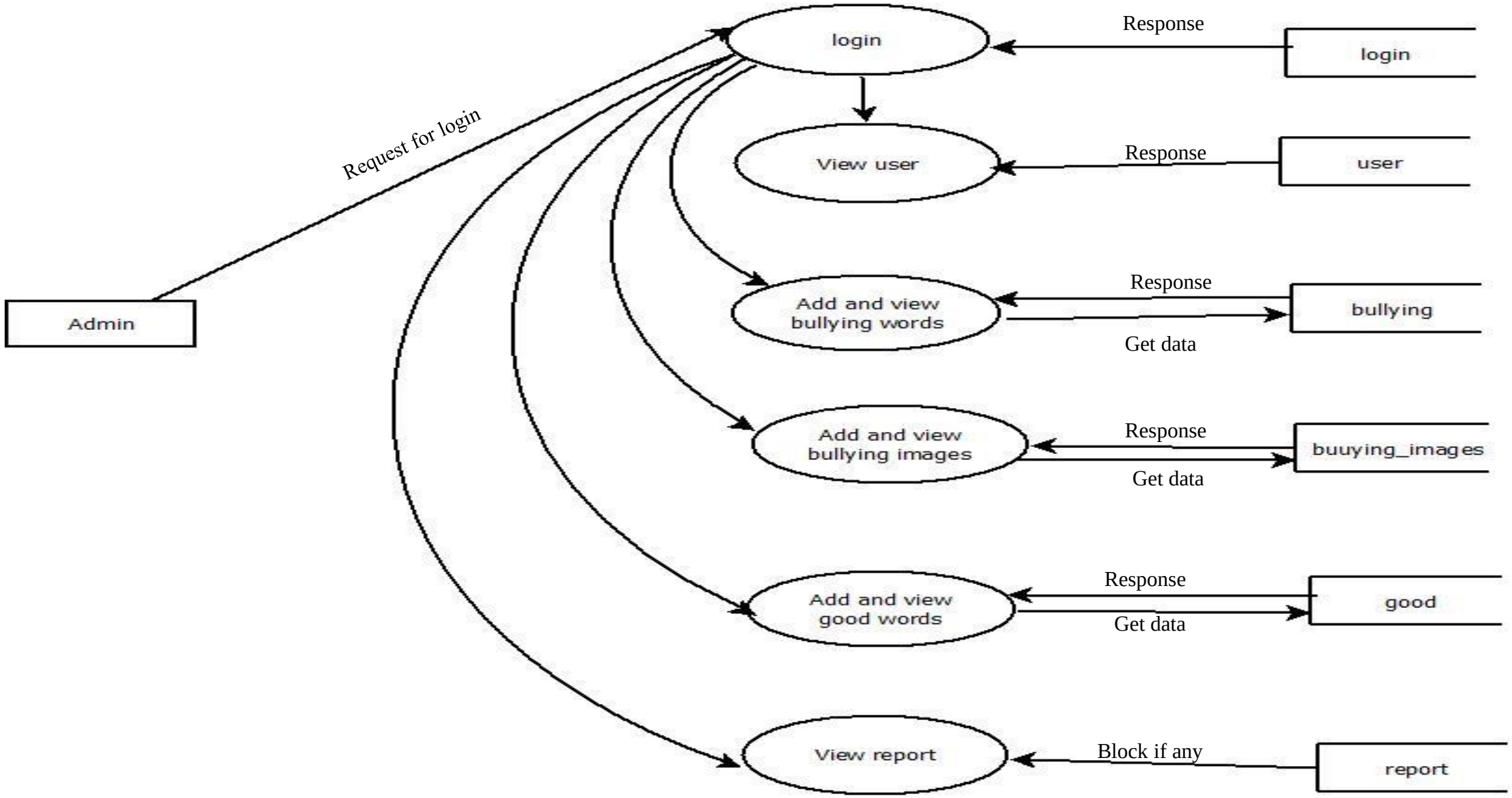
- **Software Requirements**
  - Front End: HTML, CSS, JAVASCRIPT
  - Back End: Python, MYSQL
  - IDE: : JetBrains PyCharm
  - FRAME WORK USED: Flask

DATA FLOW DIAGRAM

LEVEL 0



LEVEL 1.1


Admin
Request for login
login
Response
login
View user
Response
user
Add and view
bullying words
Response
bullying
Get data
Add and view
bullying images
Response
buuying_images
Get data
Add and view
good words
Response
good
Get data
View report
Block if any
report

LEVEL 0

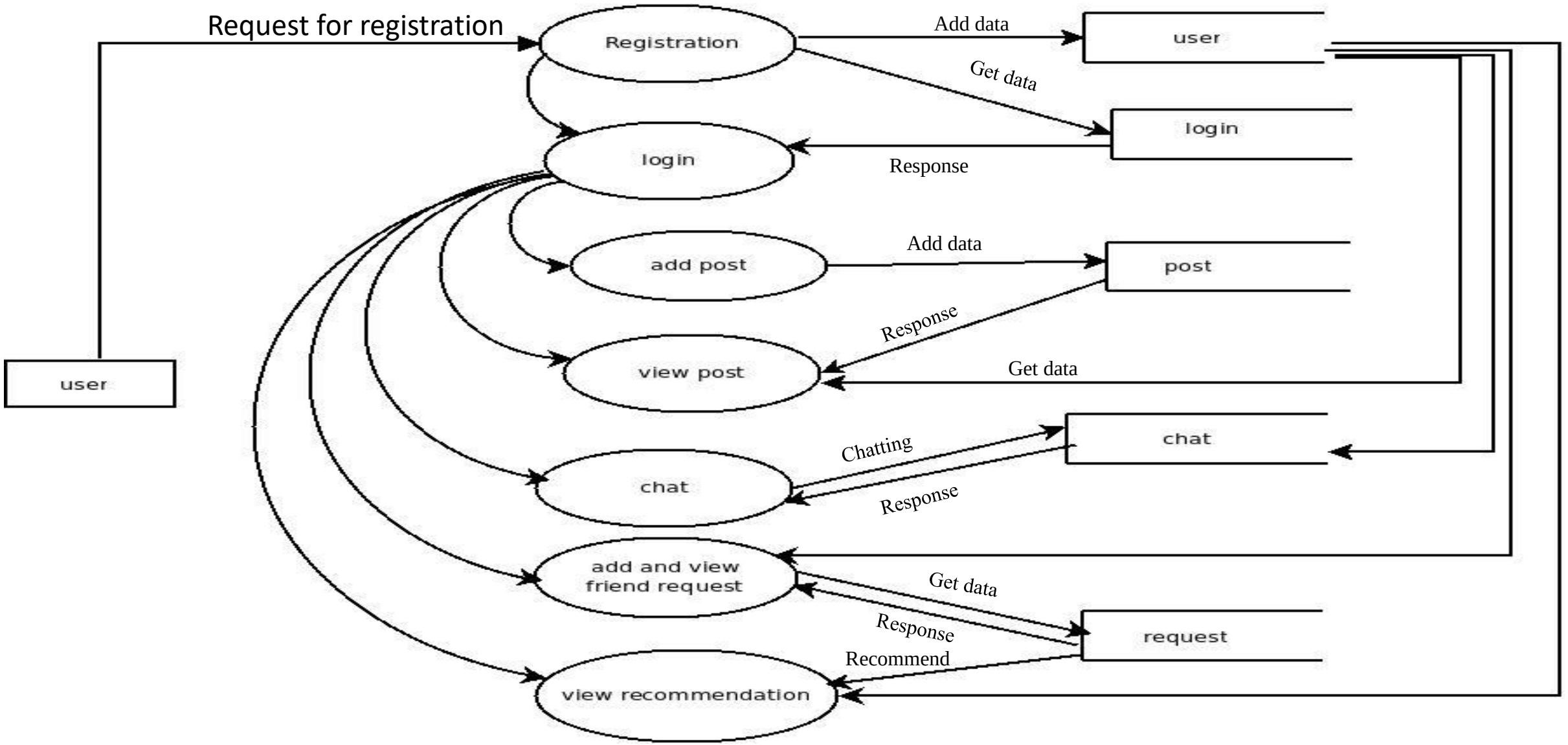

Request for registration
Registration
Add data
user
Get data
login
login
Response
add post
Add data
post
Response
view post
Get data
user
chat
Chatting
chat
Response
add and view
friend request
Get data
Response
request
Recommend
view recommendation

TABLE DESIGN

	Field	Type	Comment
	l_id	int(11) NOT NULL	contains the login id
	user_name	varchar(100) NULL	contains the name of users
	password	varchar(50) NULL	contains the password
	usertype	varchar(50) NOT NULL	contains the usertype

login

	Field	Type	Comment
	l_id	int(11) NULL	contains the id of login
	user_id	int(11) NOT NULL	contains the id of user
	f_name	varchar(100) NULL	contains the first name
	l_name	varchar(100) NULL	contains the last name
	gender	varchar(100) NULL	contains the gender
	d_o_b	varchar(100) NULL	contains the date of birth
	phone_number	bigint(100) NULL	contains the phone number
	photo	varchar(100) NULL	contains photo
	bio	varchar(100) NULL	contains the bio

user

	Field	Type	Comment
	l_id	int(11) NULL	contains the login id
🔑	bull_id	int(11) NOT NULL	contains the bullying id
	bull_word	varchar(100) NULL	contains the bullying word
	date	date NULL	contains the date

bully word

	Field	Type	Comment
🔑	id	int(11) NOT NULL	
	lid	int(11) NOT NULL	
	word	varchar(500) NOT NULL	
	date	varchar(100) NOT NULL	

good word

	Field	Type	Comment
🔑	id	int(11) NOT NULL	
	img	varchar(20) NULL	
	type	varchar(20) NULL	

bully image

	Field	Type	Comment
🔑	post_id	int(11) NOT NULL	contains post id
	l_id	int(11) NULL	contains login id
	post	varchar(100) NULL	contains post
	caption	varchar(100) NULL	contains caption
	date	date NULL	contains

post

	Field	Type	Comment
🔑	report_id	int(11) NOT NULL	
	l_id	int(11) NULL	
	post_id	varchar(11) NULL	
	comment	text NULL	
	date	date NULL	
	status	varchar(100) NULL	

report

	Field	Type	Comment
🔑	request_id	int(11) NOT NULL	
	from_id	int(11) NULL	
	to_id	int(11) NULL	
	date	date NULL	
	status	varchar(100) NULL	

request

USER STORY

User Story ID	As a type of User	I want to <perform some task>	So that I can < Achieve Some Goal>
1	Admin	Login	Login successful with correct username and password
2	Admin	View users	View all user details
3	Admin	Add and view bullying words	Add and manage bullying words
4	Admin	Add and view bullying images	Add and manage bullying images
5	Admin	Add and view good words	Add and manage good words for dataset
6	Admin	View report	View reports about bullying posts
7	User	Registration	Register to the OSN
8	User	Login	Login successful with correct username and password
9	User	Add post	Add a new post to OSN

10	User	View post	View all post OSN
11	User	Post comment	Post comment as either image or text
12	User	Add and manage friend request	Accept or reject friend request
13	User	Chat	Chat with friends
14	User	View recommendation	View recommendation

PRDUCT BACKLOG

User story ID	Priority <High/Medium/Low>	Size (Hours)	Sprint <#>	Status <Planned/In progress/Completed>	Release Date	Release Goal
1	Medium	8	1	Completed	1/05/2022	Table and form Designs
2	High	10		Completed	15/05/2022	Coding
4	Medium	6	2	Completed	28/05/2022	Comment section divided into a) Image comment b) Text Comment
6	High	5	3	Completed	5/06/2022	Comment do actions as a) Bullying detection b) View text comment c) Block d) View image comment as image
8	High		4	Completed	20/06/2022	Feature extraction and training

9	High		5	Completed	24/06/2022	Prediction of bullying image
10	High			Completed	29/06/2022	Block users

PROJECT PLAN

User Story ID	Task Name	Start Date	End Date	Hours	Status
1	Sprint 1	20/04/2022	1/05/2022	18	Completed
2		4/05/2022	15/05/2022		Completed
4	Sprint 2	17/05/2022	28/05/2022	11	Completed
5		29/05/2022	1/06/2022		Completed
6	Sprint 3	2/06/2022	5/06/2022	5	Completed
8	Sprint 4	6/06/2022	20/06/2022	11	Completed
9	Sprint 5	20/06/2022	24/06/2022	8	Completed
10		25/06/2022	29/06/2022		Completed

SPRINT PLAN

Backlog Item	Status And Completion Date	Original Estimation in Hours	Day 1	Day 2	Day 3	Day 4	Day 5	Day 6	Day 7	Day 8	Day 9	Day 10	Day 11	Day 12	Day 13	Day 14
UserStory#1,#2,#3			hrs	hrs	hrs	hrs	hrs	hrs	hrs	hrs	hrs	hrs	hrs	hrs	hrs	hrs
Form Designing	1/05/2022	8	1	1	1	2	0	1	1	0	1	0	0	0	0	0
Coding	15/05/2022	10	1	3	0	1	1	0	1	0	0	3	0	0	0	0
UserStory#4, #5																
	28/05/2022	6	1	1	0	1	1	2	0	0	0	0	0	0	0	0
Add and manage bullying words	1/06/2022	5	0	4	1	0	0	0	0	0	0	0	0	0	0	0
UserStory#4,#5																
Add and manage bullying images	5/06/2022	3	0	0	0	0	1	1	1	0	0	0	0	0	0	0
Feature extraction	20/06/2022	11	1	2	1	2	0	0	0	0	0	0	2	0	2	1

UserStory#8,#9																
Prediction	24/06/2022	4	0	0	0	1	2	1	0	0	0	0	0	0	0	0
Block Users	29/06/2022	4	0	0	2	0	0	0	0	0	0	2	0	0	0	0
Total		53	4	11	5	5	6	4	8	4	0	1	5	2	2	2

SPRINT ACTUAL

Backlog Item	Status And Completion Date	Original Estimation in Hours	Day 1	Day 2	Day 3	Day 4	Day 5	Day 6	Day 7	Day 8	Day 9	Day 10	Day 11	Day 12	Day 13	Day 14
UserStory#1,#2,#3			hrs	hrs	hrs	hrs	hrs	hrs	hrs	hrs	hrs	hrs	hrs	hrs	hrs	hrs
Form Designing	1/05/2022	8	1	1	1	2	0	1	1	0	1	0	0	0	0	0
Coding	15/05/2022	10	1	3	0	1	1	0	1	0	0	3	0	0	0	0
UserStory#4, #5																
	28/05/2022	6	1	1	0	1	1	2	0	0	0	0	0	0	0	0
Add and manage bullying words	1/06/2022	5	0	4	1	0	0	0	0	0	0	0	0	0	0	0
UserStory#4,#5																
Add and manage bullying images	5/06/2022	3	0	0	0	0	1	1	1	0	0	0	0	0	0	0
Feature extraction	20/06/2022	11	1	2	1	2	0	0	0	0	0	0	2	0	2	1

UserStory#8,#9																
Prediction	24/06/2022	4	0	0	0	1	2	1	0	0	0	0	0	0	0	0
Block Users	29/06/2022	4	0	0	2	0	0	0	0	0	0	2	0	0	0	0
Total		53	4	11	5	5	6	4	8	4	0	1	5	2	2	2

THANK YOU