International Journal For Advanced Research In Science & Technology
A peer reviewed international journal
www.ijarst.in
ISSN: 2457-0362

Cyberbullying Prediction on Social Media using Big Data by implementing Machine Learning Algorithms

P.SUBHAN BASHA [1], DANDU RAJASEKHARREDDY[2]

[1] Associate Professor, Department of Computer science and Engineering ,**SRI SAI INSTITUTE OF TECHNOLOGY AND SCIENCE**, **ANNAMAYA**, AP, India.

[2, s] UG Scholars, Department of Computer science and Engineering, **SRI SAI INSTITUTE OF TECHNOLOGY AND SCIENCE**, **ANNAMAYA**, AP, India.

ABSTRACT:

Prior to the innovation of Information Communication Technologies (ICT), social interactions developed within small cultural boundaries, consisting of geo spatial locations. The latest tendencies of verbal exchange technology have drastically transcended the temporal and spatial barriers of conventional communications. These social technologies have created a revolution in user-generated information, on-line human networks, and wealthy human conduct-associated information. However, the misuse of social technology, consisting of social media (SM) platforms, has brought a brand new shape of aggression and violence that happens solely online. A new method of demonstrating competitive conduct on S web sites is highlighted in this work. The motivations for the development of prediction fashions to combat competitive conduct are also outlined. We comprehensively evaluate cyberbullying prediction fashions and discover the primary problems associated with the production of cyberbullying prediction fashions in SM. This paper gives insights into the general system for cyberbullying detection and, most significantly, overviews the methodology. Though information series and function engineering systems have been elaborated, maximum of the emphasis is on function choice algorithms. After which, the use of diverse gadgets gains knowledge of algorithms for prediction of cyberbullying behaviours. Finally, problems and demanding situations have been highlighted as well, which gave new study instructions for researchers to explore.

INTRODUCTION

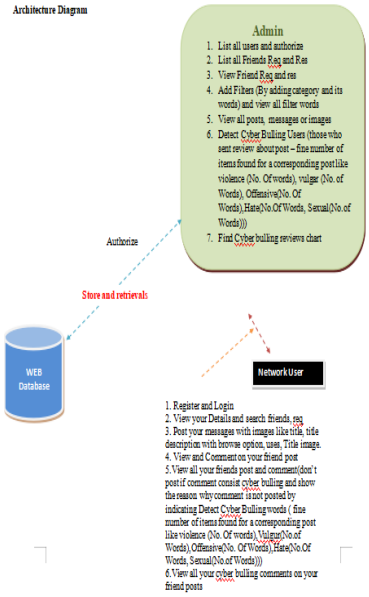
Machine or deep getting to know algorithms assist researchers apprehend massive information. Abundant statistics on human beings and their societies may be acquired in this massive information era. This acquisition became formerly impossible. By making use of devices getting to know algorithms to SM information, we can take advantage of historic information to expect the destiny of a huge variety of applications.

Machines getting to know algorithms offer an opportunity to effectively expect and come across terrible variety of human conduct, including cyberbullying. Big information evaluation can discover hidden information via deep getting to know from uncooked information. Big information analytics has stepped forward in numerous applications, and forecasting destiny has even emerged as

International Journal For Advanced Research In Science & Technology
A peer reviewed international journal
www.ijarst.in
ISSN: 2457-0362

viable via the mixture of massive information and device
getting to know algorithms. An insightful evaluation of information on human conduct and interplay to come across and restrain competitive conduct involves multifaceted angles and components and the merging of theorems and strategies from multidisciplinary and interdisciplinary fields. The accessibility of huge-scale information produces new studies questions, novel computational strategies, interdisciplinary approaches, and extraordinary possibilities to find out numerous essential inquiries quantitatively. However, using
_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.

ARCHITECTURE:



LITERATURE SURVEY:

Cybercrime detection in on-line communications: The experimental case of cyberbullying detection inside the Twitter community

The reputation of on-line social networks has created huge social communiques amongst their customers and this ends in a large quantity of user-generated communiques. In recent years, cyberbullying has grown into a prime problem with the increase of on-online communique and social media. Cyberbullying has been identified

International Journal For Advanced Research In Science & Technology
A peer reviewed international journal www.ijarst.in
ISSN: 2457-0362

currently as a severe country-wide fitness problem amongst on-online social community customers and growing green detection version holds top notch realistic. In this paper, we've proposed set of specific functions derived from Twitter; community, activity, user, and tweet content. Primarily based totally on those features, we advanced a supervised gadget gaining knowledge of answer for detecting cyberbullying inside Twitter. These outcomes imply that the proposed version primarily based totally on those functions gives a viable approach to detecting cyberbullying in on-line communique. Finally, we evaluate end results of the use of our proposed functions with the end result acquired from baseline functions. The assessment results display the importance of the proposed functions.

Improving cyberbullying detection with person's context, in Advances in Information Retrieval

The bad results of cyberbullying have become alarming each day and technical answers that permit suitable movement with the aid of computerized detection are nonetheless very limited. Up till now, research on cyberbullying detection has centred on person's remarks only, dismissing context consisting of users' traits and profile information. In this paper we display that taking a person's context under consideration improves the detection of cyberbullying.

Detecting Offensive Language in Social Media to Protect Adolescents' Online Safety.

A person-stage offensiveness detection appears to be the most possible method However, it's far beyond the researched area. To bridge this gap, we advocate the Lexical Syntactic Feature (LSF) structure to hit upon offensive content materials and pick out capacity offensive customers in social media. We distinguish the contribution of pejorative / profanities and obscenities in figuring out offensive content material, and introduce hand-authoring syntactic regulations in figuring out name-calling harassment. In particular, we contain a person's writing style, shape and particular cyber bullying content material as capabilities to expect the person's potential to ship out offensive content material. Results from experiments confirmed that our LSF framework finished drastically higher than current strategies in offensive content material detection.

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

International Journal For Advanced Research In Science & Technology
A peer reviewed international journal www.ijarst.in
ISSN: 2457-0362

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

International Journal For Advanced Research In Science & Technology
A peer reviewed international journal www.ijarst.in
ISSN: 2457-0362

have a set of features that are extracted and engineered.

Advantages

- The system is more effective due to LOGISTIC REGRESSION CLASSIFICATION and UNSUPERVISED MACHINE LEARNING.

An effective cyberbullying prediction models is to use a text classification approach that involves the construction of machine learning classifiers from labeled text instance and also 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

IMPLEMENTATION

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).

International Journal For Advanced Research In Science & Technology
A peer reviewed international journal www.ijarst.in
ISSN: 2457-0362

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

RESULT

Predicting Cyberbullying on Social Media in the Big Data Era Using

International Journal For Advanced Research In Science & Technology
A peer reviewed international journal www.ijarst.in
ISSN: 2457-0362

Machine Learning Algorithms: Review of Literature and Open Challenges

In this paper author is using various machine learning algorithms such as SVM, Random Forest, Naïve Bayes, KNearest Neighbours, and Decision Tree to predict cyberbullying posts from social networks. As additional algorithm we are using 'Extreme Machine Learning' algorithm which is an advance algorithm in machine learning area.

Using all algorithms we will build train model with normal and bullying messages and this train model will applied on new posts from users to predict whether new post is normal or contain bullying stuff.

This project consists of following modules

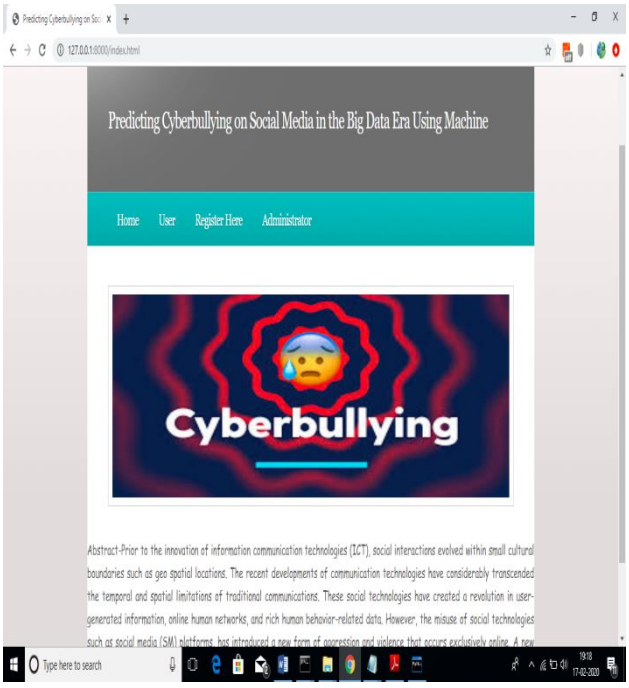
User module: using this module users can create an account. Using account details they can login to application and then send and view posts.

Admin Module: Admin can view all registered user account and then accept or reject new user account. Admin responsible to add new bullying messages to machine learning train dataset. Admin has to run all or at-least one SVM algorithm to perform bullying message detection from user side. Admin can view or monitor all posts send by all users.

Note: machine learning algorithms will predict bullying or non-bullying message if only some related data available in train model. So you can predict all those messages correctly which are entered by admin. All sample bullying and non-bullying messages are available inside 'Cyber/dataset.txt' file.

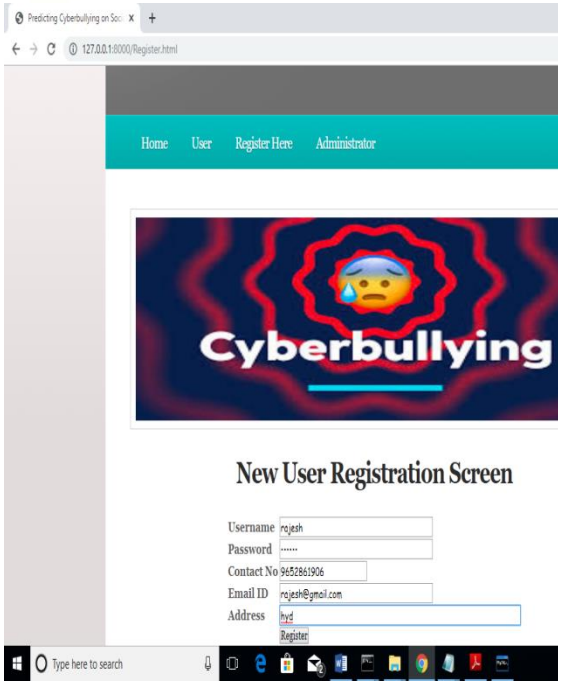
First create database by copying content from 'DB.txt' file and paste in MYSQL console

To run this project install DJANGO and then deploy 'Cyber' folder in DJANGO and start server and run in browser by entering URL as 'http://127.0.0.1:8000/index.html'.
After running above URL will get below screen

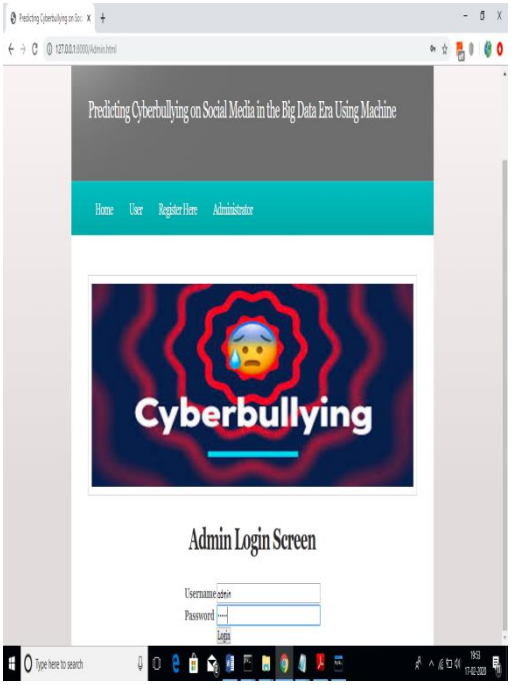


In above screen click on 'Register Here' link and add new user to create account

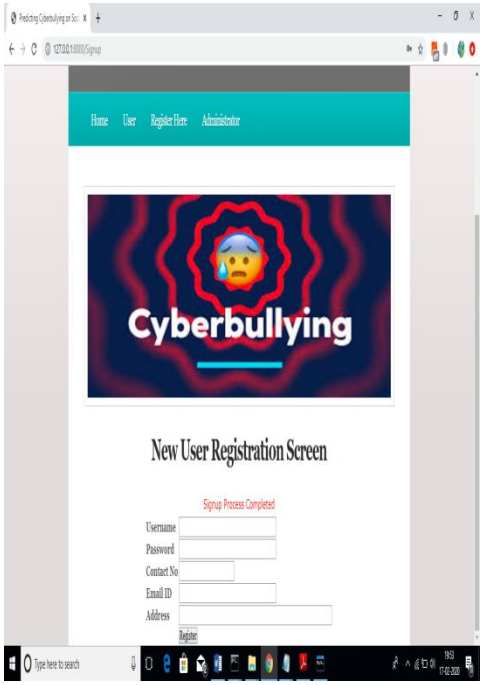
International Journal For Advanced Research In Science & Technology
A peer reviewed international journal
www.ijarst.in
ISSN: 2457-0362



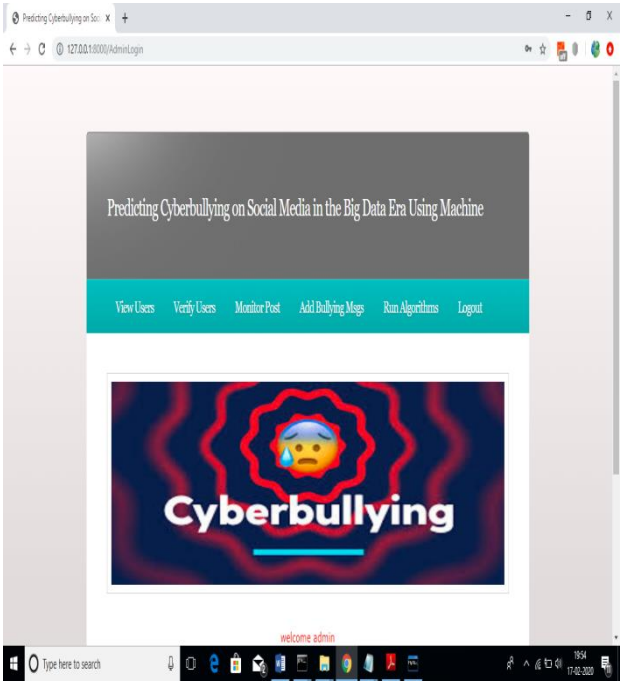
In above screen now click on 'Register' button to add details



In above screen login as 'admin' by giving username as 'admin' and password as 'admin'. After login will get below screen

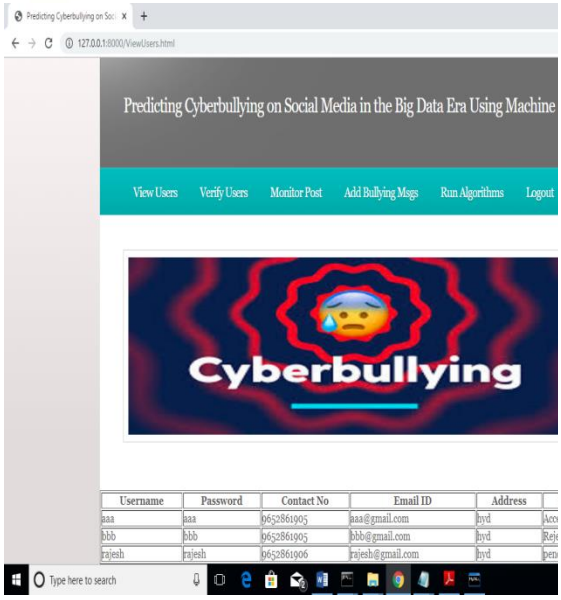


In above screen sign up process completed. Now click on 'Administrator' link to login as admin and give permission to new user

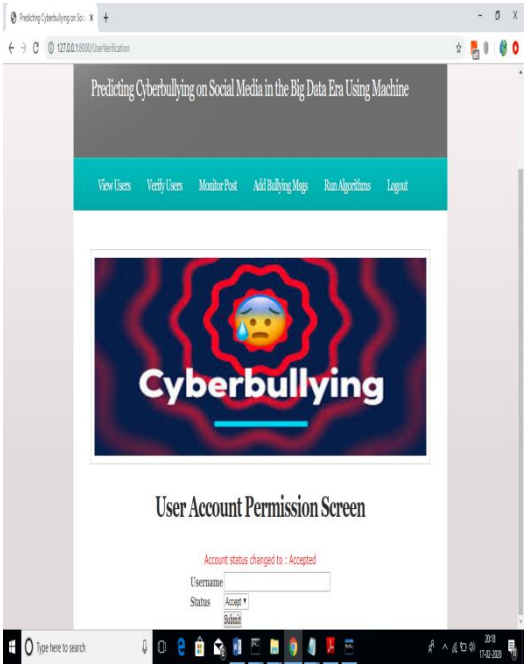


Now admin can click on 'View Users' link to view all users list

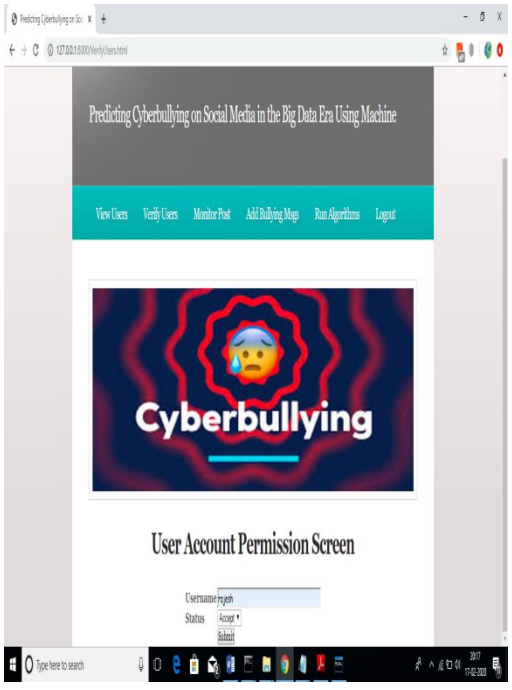
International Journal For Advanced Research In Science & Technology
A peer reviewed international journal
www.ijarst.in
ISSN: 2457-0362



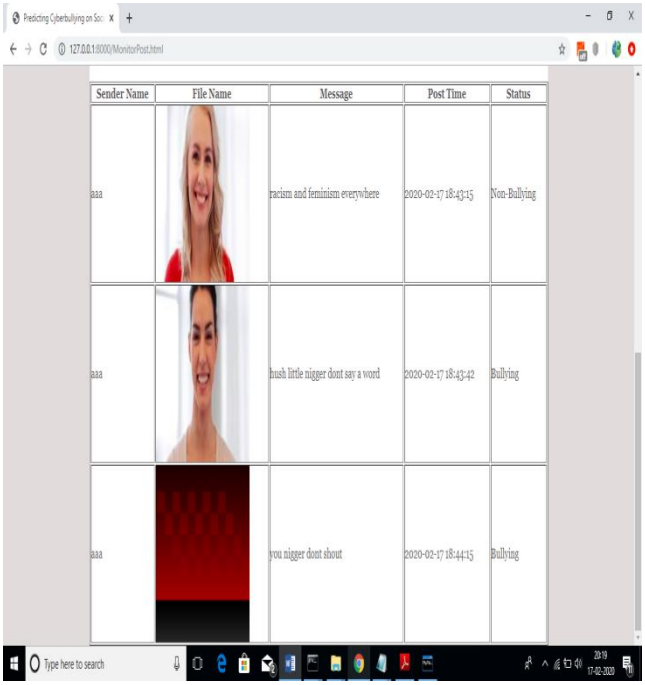
In above screen we can see 'rajesh' account is in pending state and to give permission to rajesh. Now admin will click on 'Verify Users' link to get below screen and to give permission



In above screen account state changed to 'Accepted'. Now admin can click on 'Monitor Post' to view all post from past users



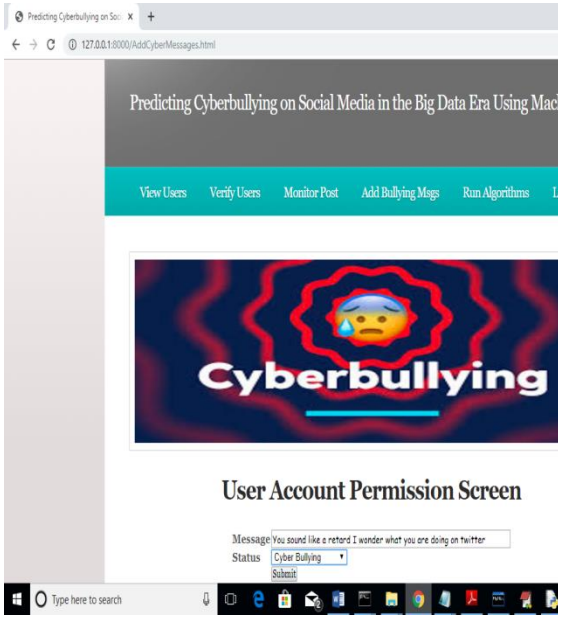
In above screen admin will enter username and then select 'Accept' or 'Reject' option to give permission



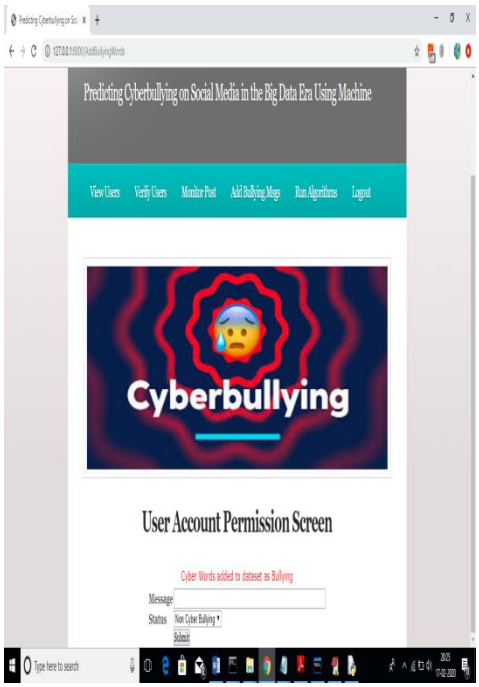
In above screen application will automatically detect whether message is non-bullying or bullying from machine learning algorithms. Now admin can

International Journal For Advanced Research In Science & Technology
A peer reviewed international journal
www.ijarst.in
ISSN: 2457-0362

click on 'Add Bullying Msgs' link to add words

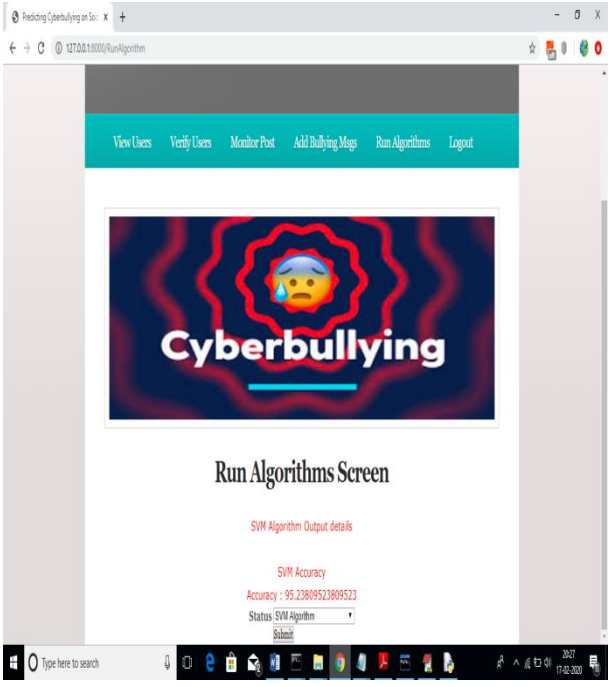


In above screen admin adding one sentence as 'Cyber Bullying' and similarly he can add all possible bullying and non-bullying messages. After adding messages will get below screen



Now admin can click on 'Run Algorithms' link to generate train model using entire dataset to predict user posts as normal or bullying

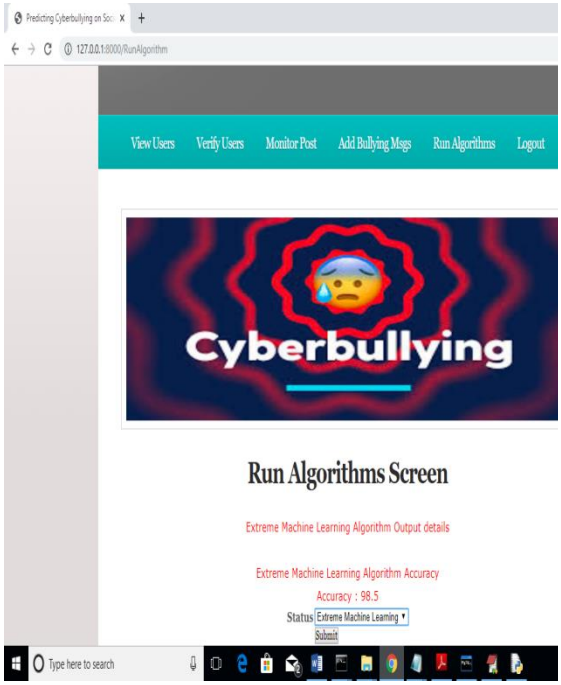
In above screen admin has to select each algorithm and click on 'Submit' button to train model and we will get accuracy also for each algorithm. Admin has to repeat this step whenever first time he starts the server or upon adding new bulling messages.



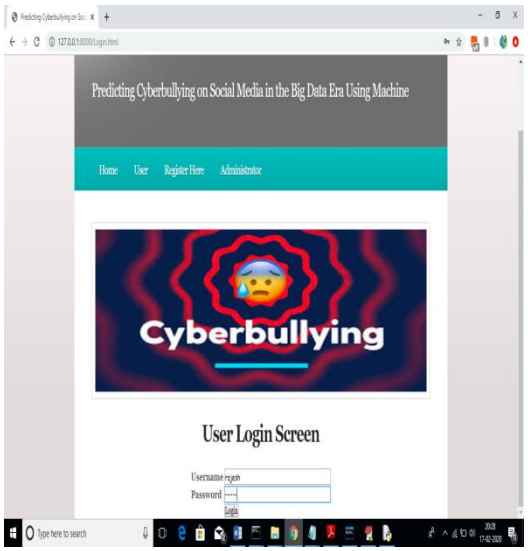
In above screen I ran SVM and got accuracy as 95. Similarly u need to

International Journal For Advanced Research In Science & Technology
A peer reviewed international journal
www.ijarst.in
ISSN: 2457-0362

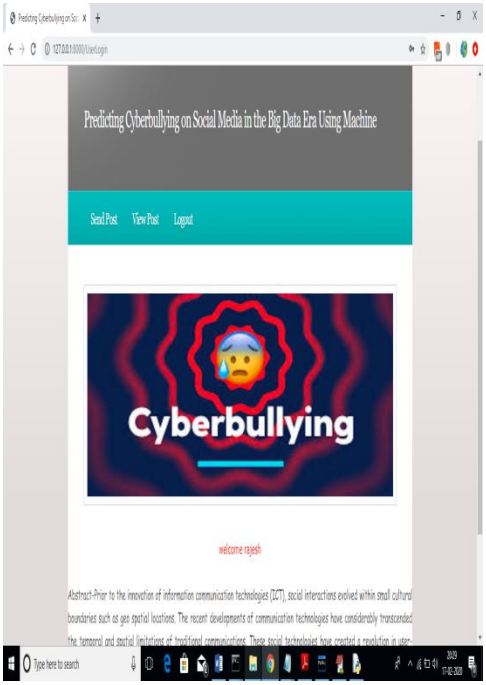
select all algorithms one by one and run it.



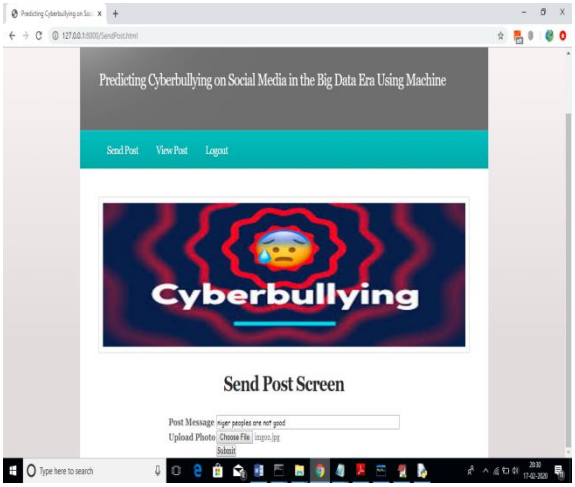
In above screen advance 'Extreme Machine Learning' algorithm gave 98% accuracy. Now admin logout and login as user to send posts.



In above screen rajesh user is login and after login will get below screen

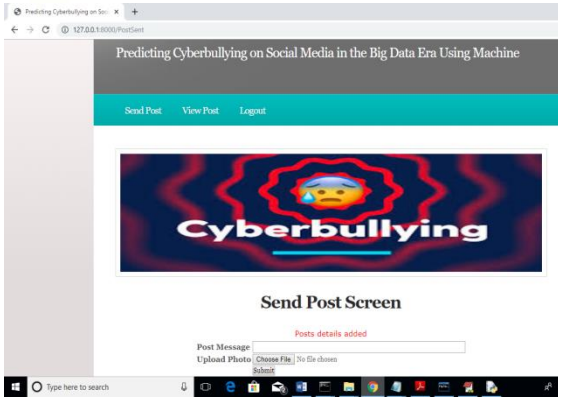


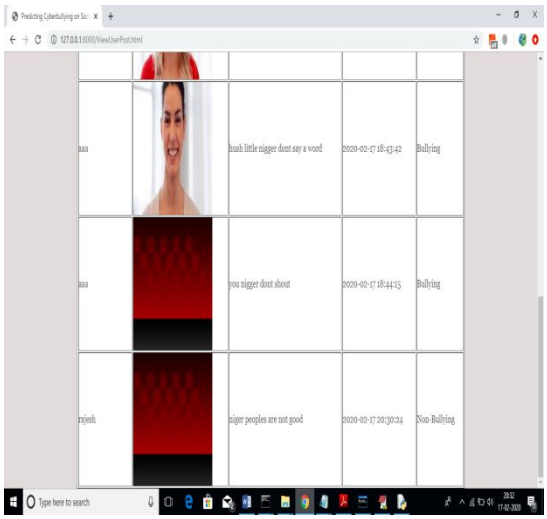
In above screen click on 'Send Post' link to get below screen



In above screen as post I added some messages and uploaded a photo also. After posting message will get below screen

International Journal For Advanced Research In Science & Technology
A peer reviewed international journal
www.ijarst.in
ISSN: 2457-0362





In above screen we are seeing posts from all users and rajesh post predicted as 'Non-Bullying'. Here based on words given in dataset will get prediction as bullying on non-bullying

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.

[2] 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.

[3] 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.

[4] 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
[5] 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. [6] 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. [7] BBC. (2012). Huge Rise in Social Media. [Online]. Available: http://www.bbc.com/news/uk-20851797
[8] 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. [9] 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. [10] 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