JOURNAL OF CRITICAL REVIEWS

ISSN- 2394-5125 VOL 10, ISSUE 04, 2023

A NEW DETECTION OF CHILD PREDATOR'S CYBER HARASSERS ON SOCIAL MEDIA

K. Ramya Sri[1], K. Geethika[2], L. Chandana[2], M. Sruthi[2], N. Ramya[2]

[1]Assistant Professor, [2]UG Scholar, [1,2]Department of Computer Science and Engineering

[1,2]Malla Reddy Engineering College for Women (A), Maisammaguda, Medchal, Telangana.

chowdarygeethika42@gmail.com, lasmaichandana@gmail.com, madinenisruthi@gmail.com, nunavathramya123@gmail.com

Abstract

Professional psychologists need to understand the dangers of online sexual harassment and how to protect young people from sex predators using the internet. Although the net has several positive aspects, one in all the foremost pernicious aspects is its potential use for on-line sexual postulation. The internet shows a medium that allows sex predators to enter numerous children in a relatively anonymous environment. The main objective of our project is to detect child predator base on comments and post of social media account and send predator record to cyber cell admin. a recent national survey indicated that about one in five youth are solicited for sex over the internet annually (finkelhor, mitchell, & wolak, 2000; mitchell, finkelhor, & wolak, 2001). This project report presents our current development to enable the creation of the system. As a result, with the developed system, child predator accounts detection any report to admin for further action.

Keywords: Child predators, Cyber harassers, social media.

1. Introduction

Child predator detection system on social media is a web based application. This project aims to detect child predator comments and post on social media like fb, insta etc. and send report to cyber cell admin. To develop an well-designed database to store all comments and post of social online contact of children in pedophiles is a rapidly growing problem on social media. As of March 2014, the national society for the hindrance of cruelty to kids (nspcc), reported that i) 12-tone system of 11-16 year olds within the kingdom have received unwanted sexual messages; and ii) 8% of 11-16 year olds in the uk have received requests to send or respond to a sexual message. The detection of kids cybersexual offenders is so a crucial issue that must be addressed. Kids in their teens have begun to use social media as their main means of communication. Moreover, a recent study of cognition, adolescents and mobile phones (scamp) has revealed that 70% of 11–12-year-olds in the uk now own a mobile phone rising to 90% by age 14. A common attack of pedophiles is the so-called online child grooming, where adults eventually exchange sexually explicit content through social media outlets. Such grooming consists of building a trust-relationship with a minor, which finally leads into convincing a child to meet them in person. Previous research on detecting cyber pedophilia online, including the efforts of the first international sexual predator identification competition.

2. Literature Survey

Muhamad Ali, Partick Bours. "Ensemble technique for sexual predator identification". They study a good technique for sexual predator identification. Cyber grooming may be a compelling drawback worldwide today and plenty of reports powerfully instructed that it becomes terribly imperative to tackle this drawback to safeguard the kids from sexual exploitation. during this study, we have a tendency to propose a good technique for sexual predator identification in on-line chats supported two-stage classification. the aim of the primary stage is to tell apart predatory languages from the traditional ones whereas the second stage aims to inform apart between the predator user and therefore

the victim at intervals one predatory conversation. Finally, some distinctive predators square measure derived from the second stage result. we have a tendency to investigate many machine learning classifiers as well as Naive Bayes, Support Vector Machine, Neural Network, provision Regression, Random Forest, K-Nearest Neighbours, and call Tree with Bag of Words options victimization many totally different term weight strategies for this task. we have a tendency to additionally projected 2 ensemble techniques to enhance the classification task. The experiment results on PAN12 dataset show that our greatest technique victimization soft vote primarily based} ensemble for initial stage And Naive Bayes based technique for the second stage obtained an F zero.5 -score of zero.9348, which might place as favourite within the PAN12 competition ranking

Michael Ashcroft; Lisa Kaati; Maxime Meyer "A Step Towards sleuthing on-line Grooming -- characteristic Adults simulation to be Children" They enforced machine-controlled analysis of chat area language to discover and attainable tries of grooming nline grooming may be a major drawback in today's society wherever additional and longer is spent on-line. To become friends and establish a relationship with their young victims in on-line communities, groomers typically faux to be kids. during this paper, we have a tendency to describe AN approach that may be wont to discover if AN adult is simulation to be a baby during a chat area language. The approach involves a 2-step method whereby authors square measure initial classified as being kids or adults, so every kid is being examined and false kids distinguished from real kids. Our results show that notwithstanding it's arduous to separate standard adults from kids in chat logs it's attainable to tell apart real kids from adult's simulation to be kids with a high accuracy. during this paper, we are going to discuss the accuracy of the strategies projected, additionally because the options that were vital in their success. we have a tendency to believe that this work is a vital step towards machine-controlled analysis of chat area language to discover and attainable tries of grooming. Our approach wherever we have a tendency to use text analysis to tell apart adults World Health Organization square measure simulation to be kids from actual kids may be wont to inform kids regarding verity age of the person who they're human activity. this might be a step towards creating the web safer for young kids and eliminate grooming.

Patrick Bours, Halvor Kulsrud Detection of Cyber Grooming in on-line Conversation They enforced system to discover on-line cyber grooming. during this paper, we are going to specialise in the detection of sexual predators in on-line chat conversations. we have a tendency to use three totally different approaches (message-based, author-based and conversation-based) combined with five {different|totally totally different|completely different} classification algorithms and a pair of different options sets. the simplest results were obtained victimization either the author-based approach with the Neural Network classifier on the TF-IDF feature set, or the conversation-based approach victimization the Ridge or the Naïve Bayes classifier on the TF-IDF feature set. during this paper, for the primary time, we have a tendency to checked out however fast a predator may be detected, and located that in most cases 26-161 messages of a language were comfortable. This constitutes solely alittle fraction of the complete conversations, showing that we will have AN early detection system of sexual predators rather than knowing looking back that a baby was the victim of a sexual predator.

Stefan C. Dombrowski, John W. LeMasney, and Claude Elwood Shannon A. Dickson they study regarding skilled psychologist's ought to additional absolutely perceive the risks of on-line sexual solicitation and ways that during which to safeguard youth from sexual predators World Health Organization use the web. though the web has several positive aspects, one in every of the foremost pernicious aspects is its potential use for on-line sexual predation. the web represents a medium that permits sexual predators access to innumerable kids during a comparatively anonymous setting. this text reviews the overall ways of sexual perpetrators and their characteristics, additionally because the

on-line ways and characteristics of the cyber sexual predator. data on a way to shield kids from this crime through a review of technological, psych instructional, and legal issues is provided. an outline of the relevant laws as they relate to on-line solicitation and active psychologists is additionally provided

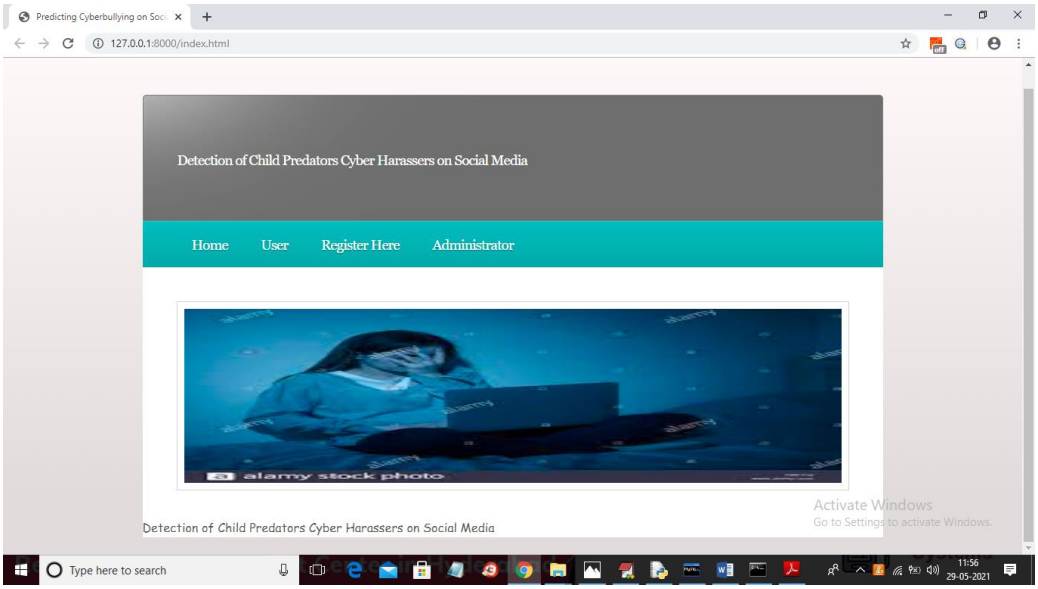
Hee-Eun Lee Tatiana Ermakova Vasilis Ververis Benjamin Fabian This gift analysis provides a comprehensive synthesis and an interpretation of the present analysis accomplishments and challenges within the CSAM detection domain, expressly considering the size of policy and legal framework, distribution channels, and detection applications and implementations. Among alternative aspects, it reveals and aggregates data associated with image hash info, keywords, web-crawler, detection supported filenames and information, and visual detection. The findings recommend that CSAM detection applications yield the simplest results if multiple approaches square measure utilized in combination, such as deep-learning algorithms with multi-modal image or video descriptors incorporate along. Deep-learning techniques were shown to surpass alternative detection strategies for unknown CSAM.

Muhammad Ali Fauzi Apostle Bours during this study, we have a tendency to propose a good technique for sexual predator identification in on-line chats supported two-stage classification. the aim of the primary stage is to tell apart predatory languages from the traditional ones whereas the second stage aims to inform apart between the predator user and therefore the victim at intervals one predatory conversation. Finally, some distinctive predators square measure derived from the second stage result. we have a tendency to investigate many machine learning classifiers as well as Naive Bayes, Support Vector Machine, Neural Network, provision Regression, Random Forest, K-Nearest Neighbours, and call Tree with Bag of Words options victimization many totally different term weight strategies for this task.

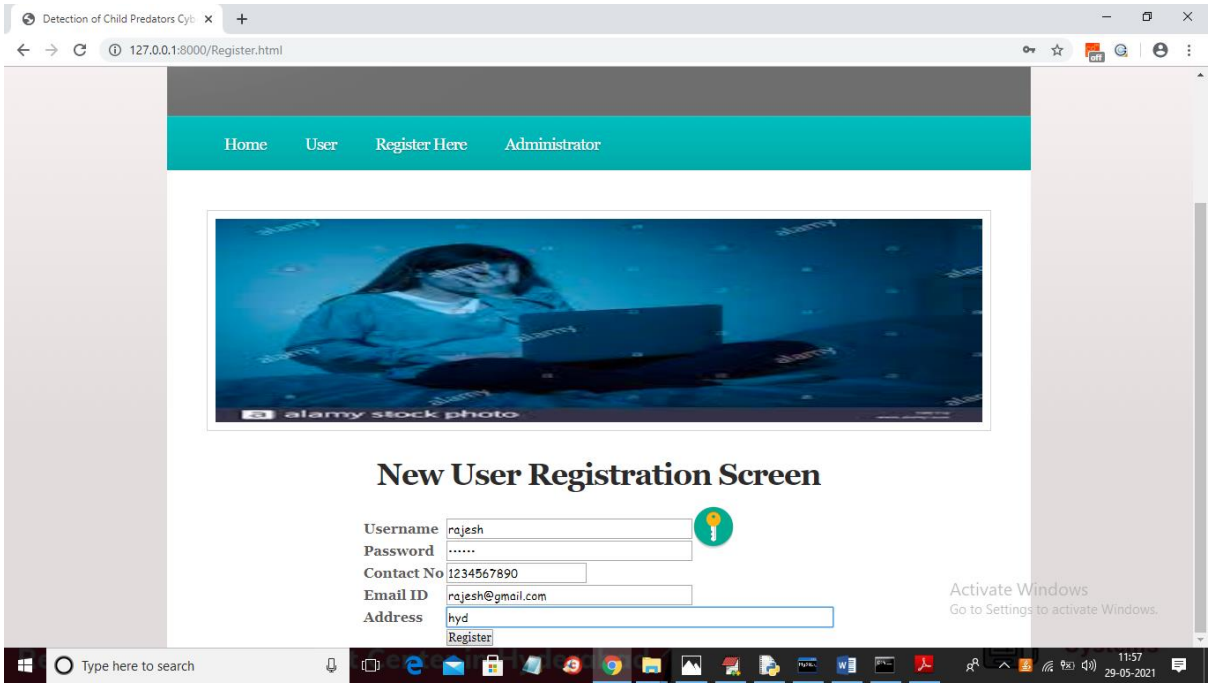
3. Proposed System

In this project we are using various machine learning algorithms such as SVM, Random Forest, Naïve Bayes, KNearest Neighbours, and Decision Tree to predict child harasser's posts from social networks. Using all algorithms, we will build train model with normal and harasser's word and messages and this train model will applied on new posts from users to predict whether new post is normal or contain harasser's stuff.

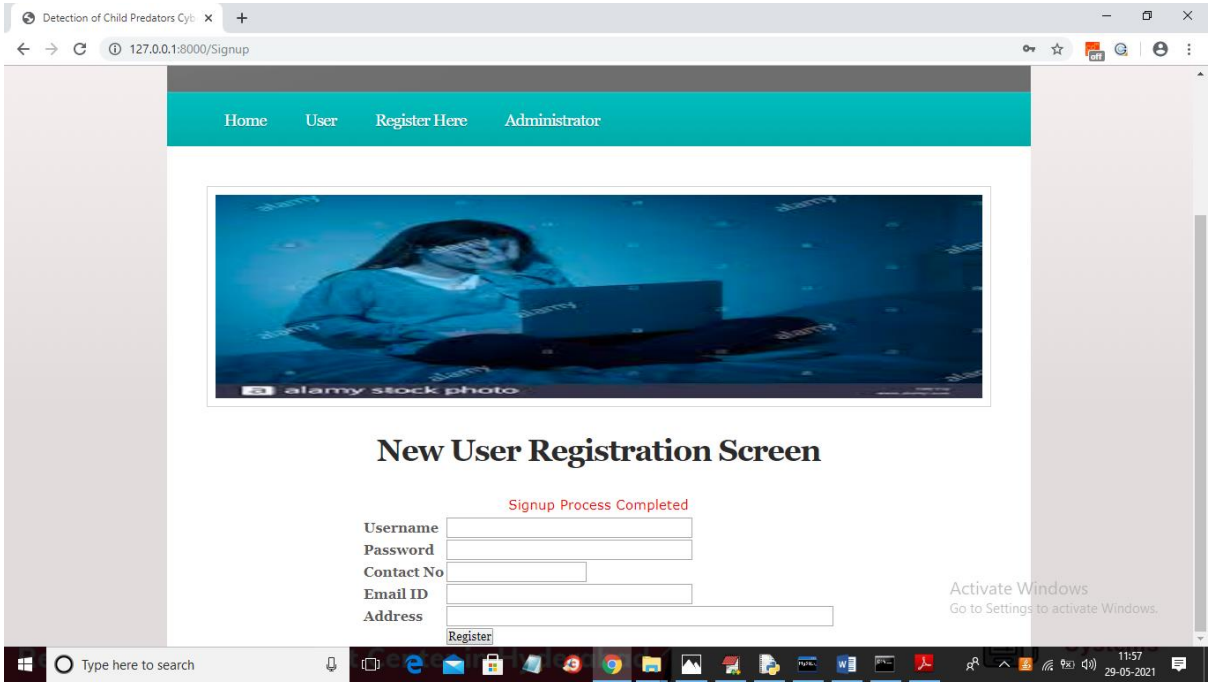
4. Results



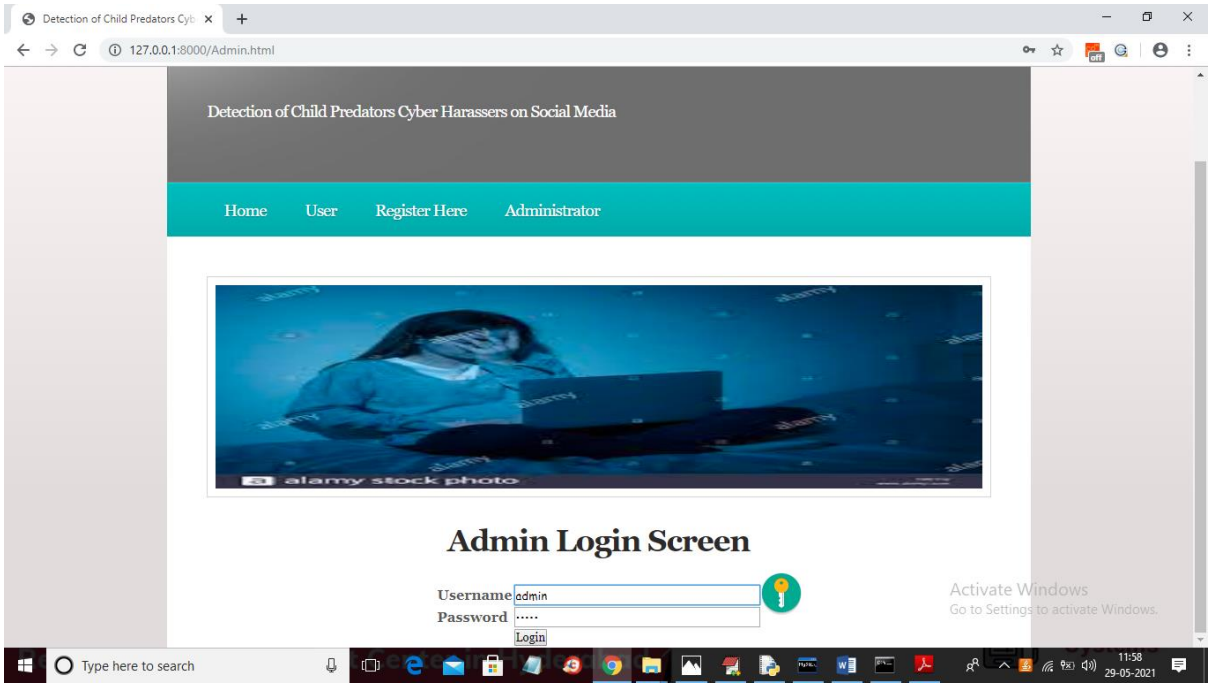
In above screen click on ‘Register Here’ link and add new user to create account



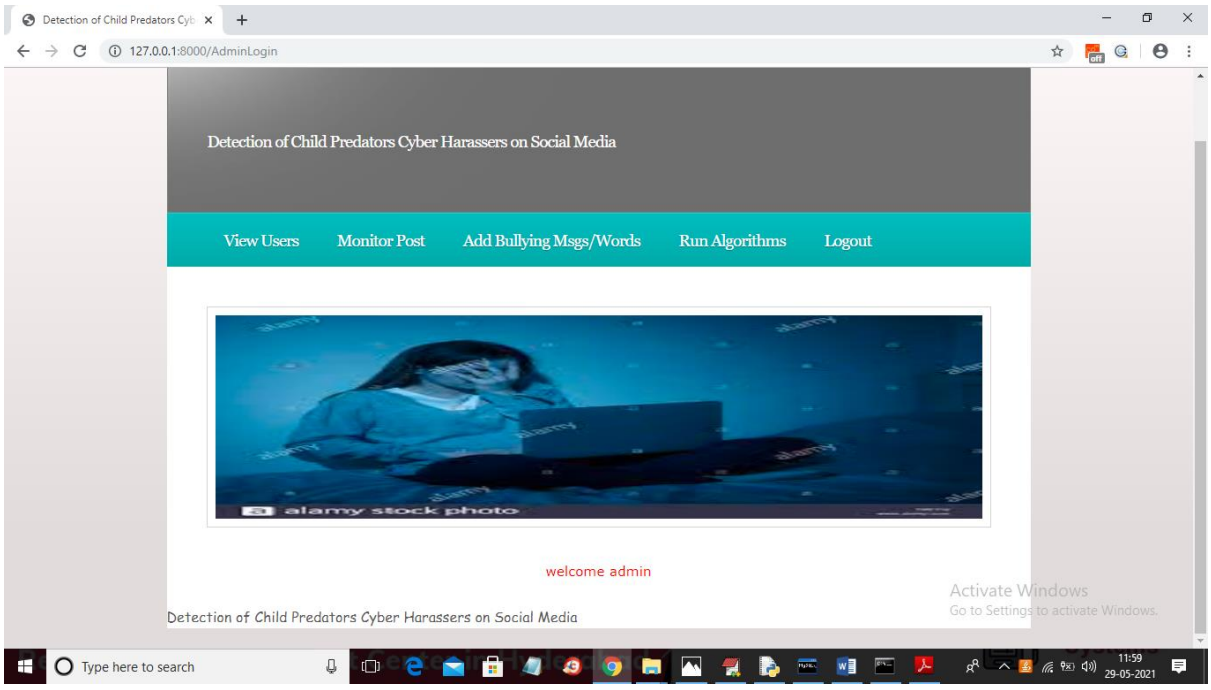
In above screen now click on ‘Register’ button to add details



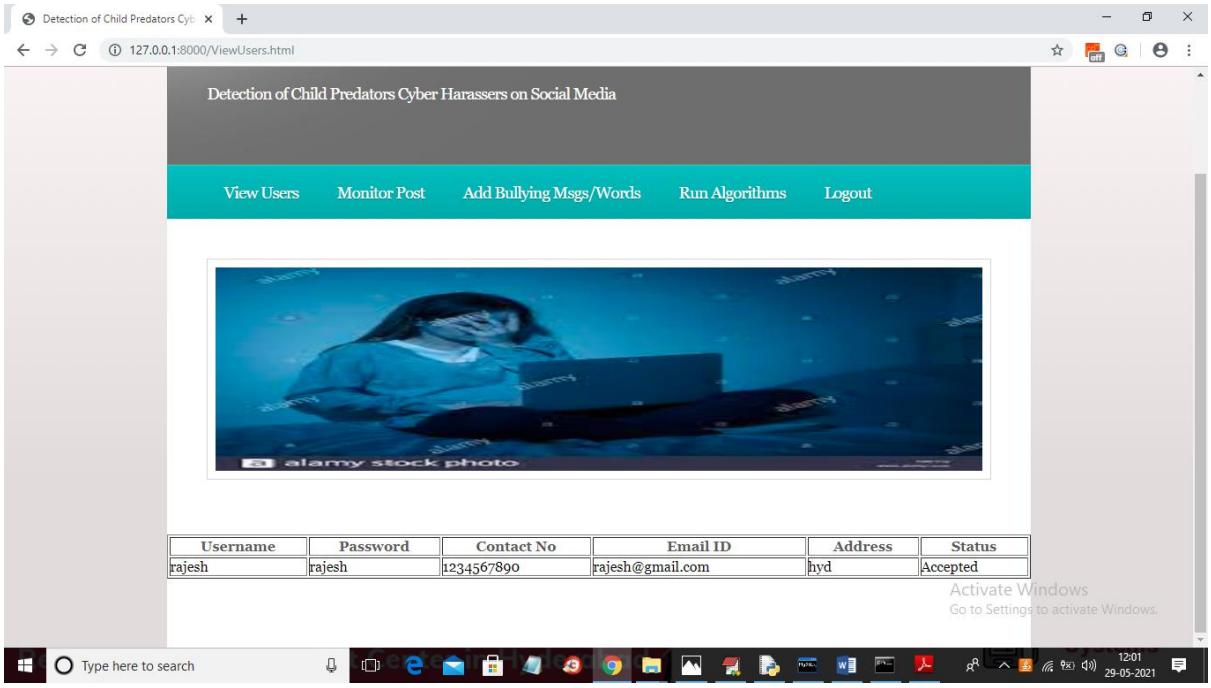
In above screen sign up process completed. Now click on ‘Administrator’ link to login as admin view new user details



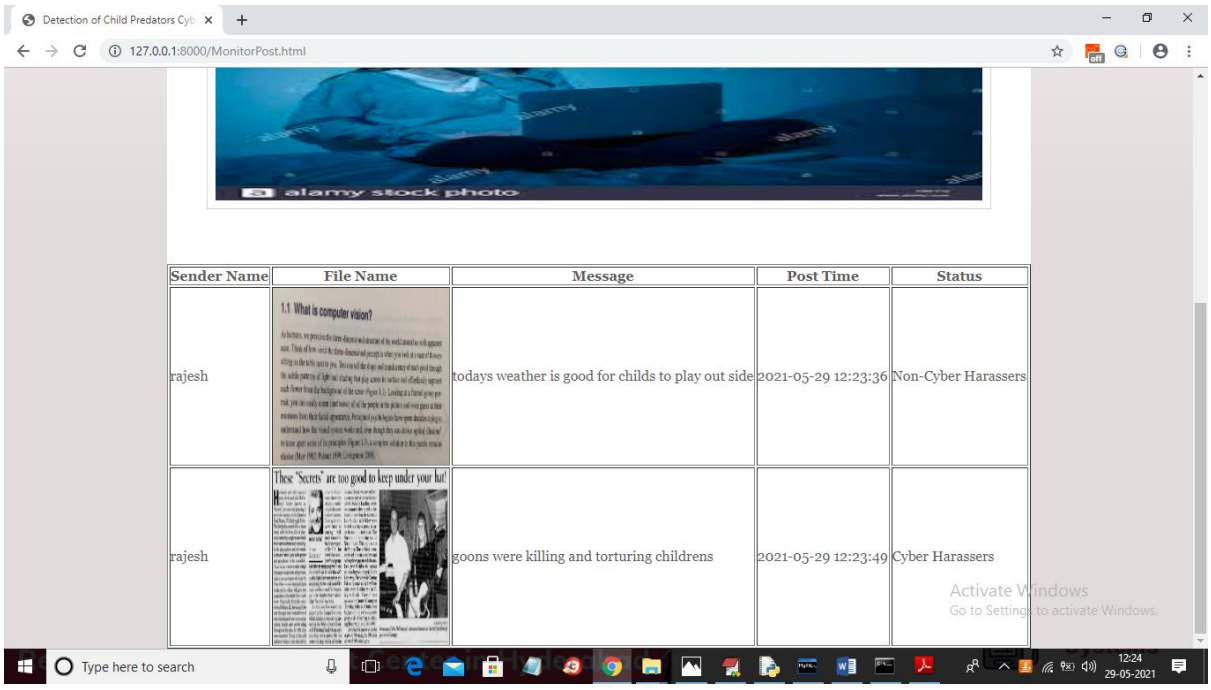
In above screen login as ‘admin’ by giving username as ‘admin’ and password as ‘admin’. After login will get below screen



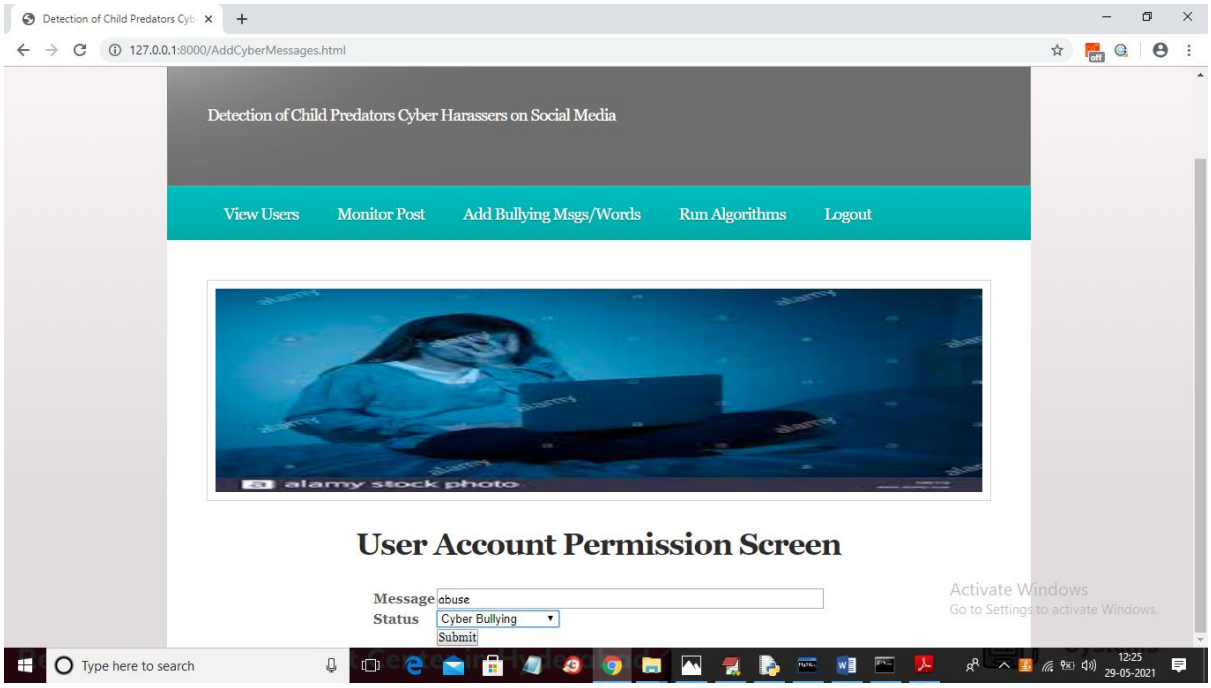
Now admin can click on ‘View Users’ link to view all users list



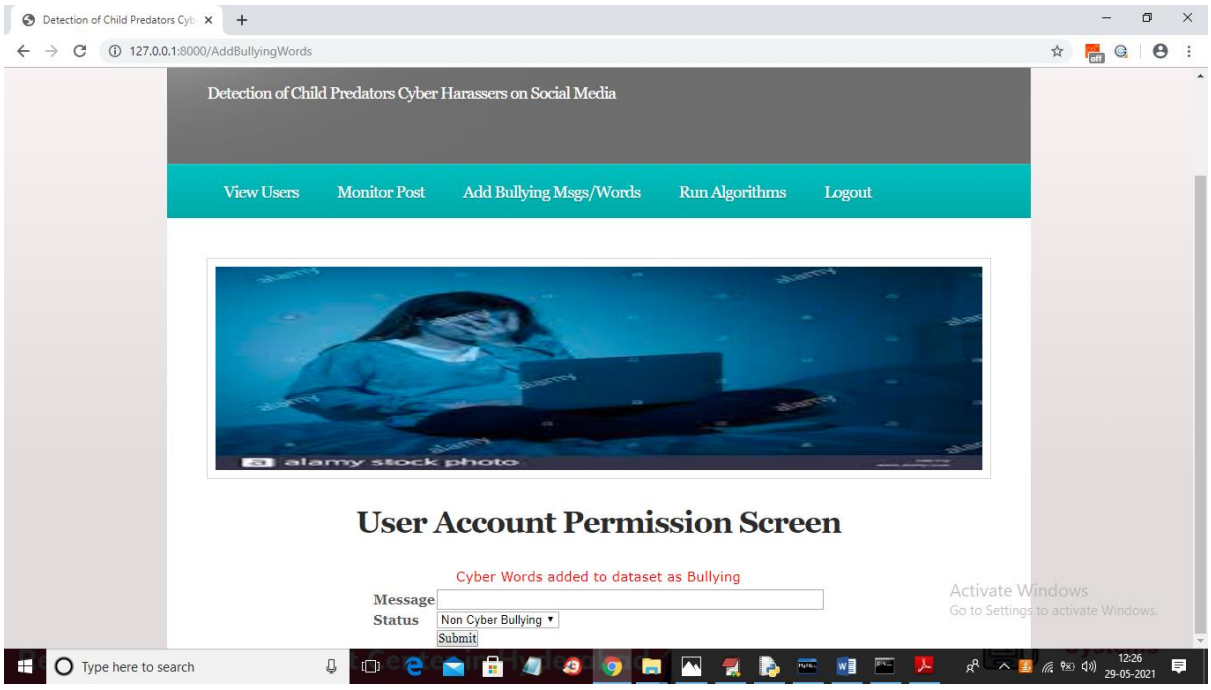
In above screen we can see ‘rajesh’ account created. . Now admin can click on ‘Monitor Post’ to view all post from past users.



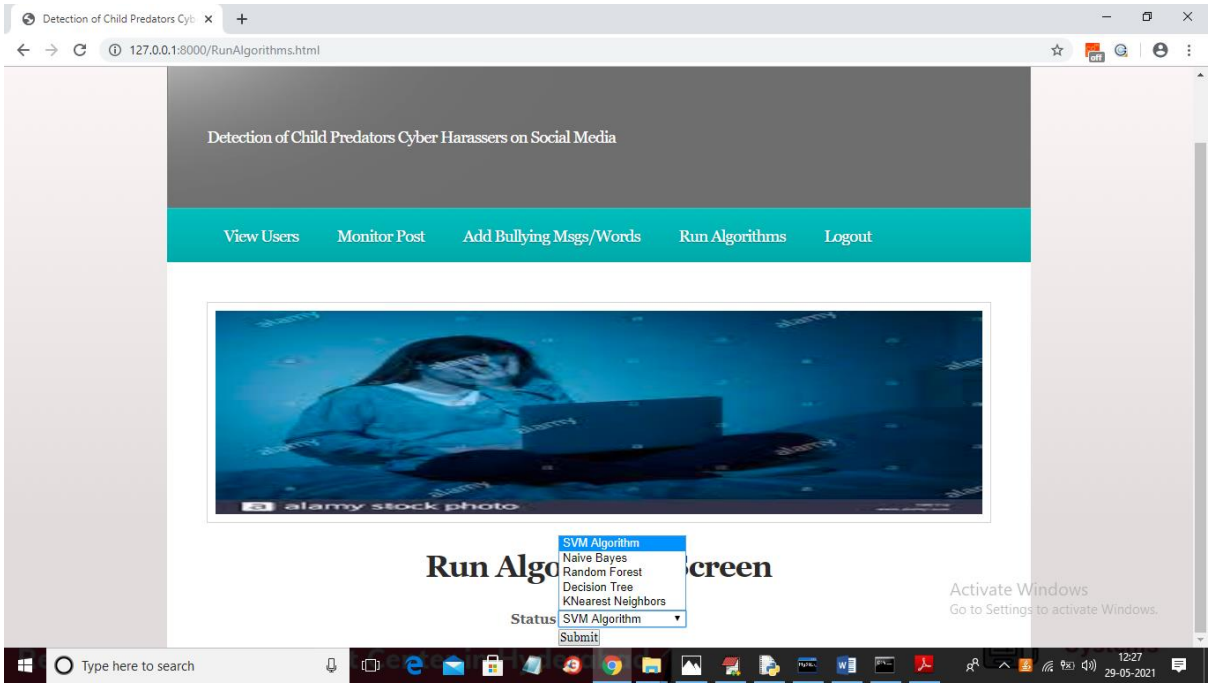
In above screen application will automatically detect whether message is non-cyber harasser’s or harassers from machine learning algorithms. Now admin can click on ‘Add Bullying Msgs/words’ link to add words.



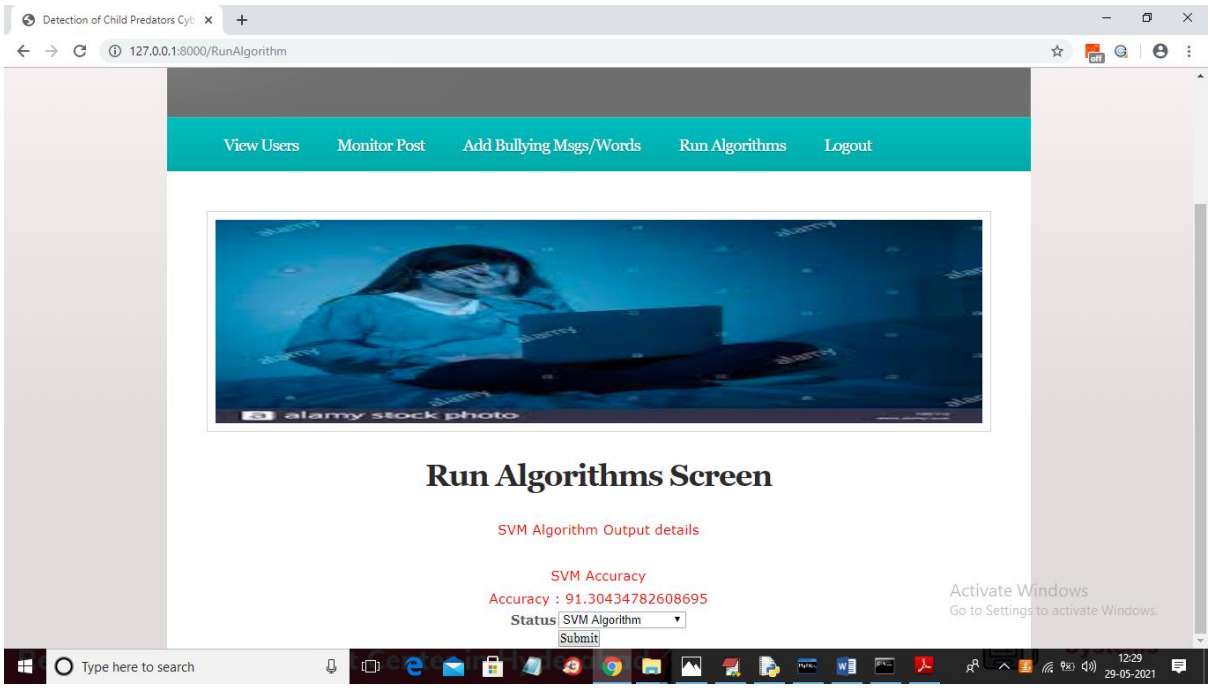
In above screen admin adding one word as 'Cyber Bullying' and similarly he can add all possible bullying and non-bullying messages. After adding messages will get below screen



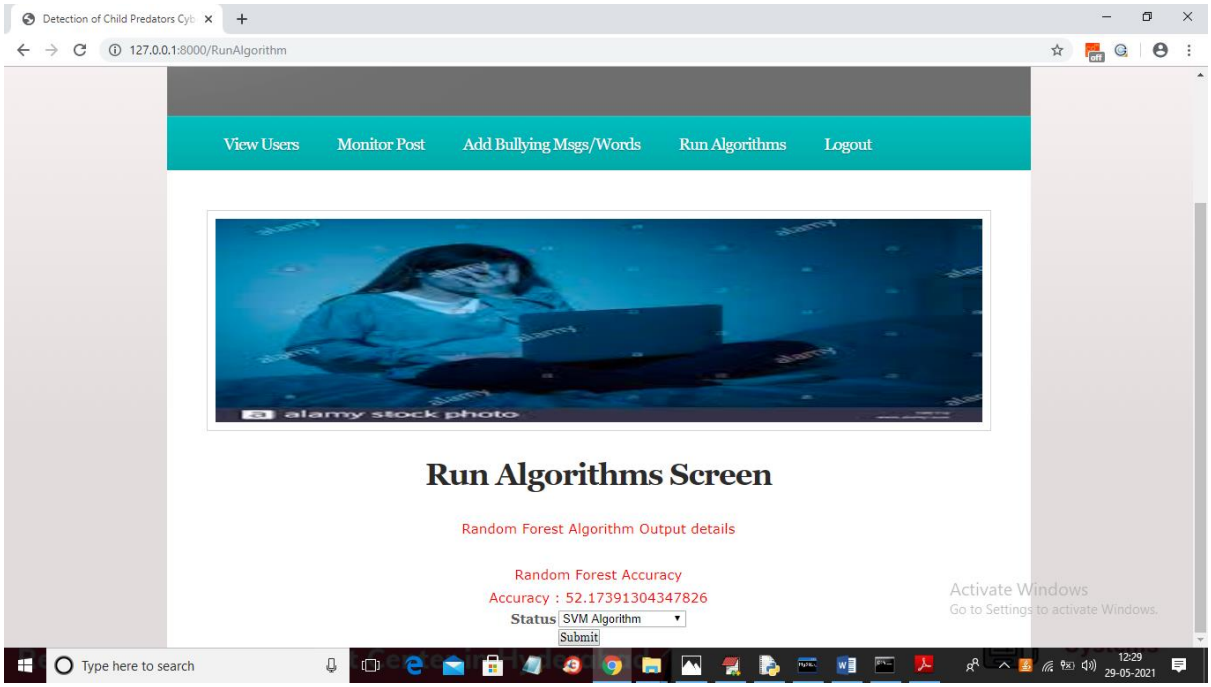
Now admin can click on 'Run Algorithms' link to generate train model using entire dataset to predict user posts as normal or bullying/harasser's.



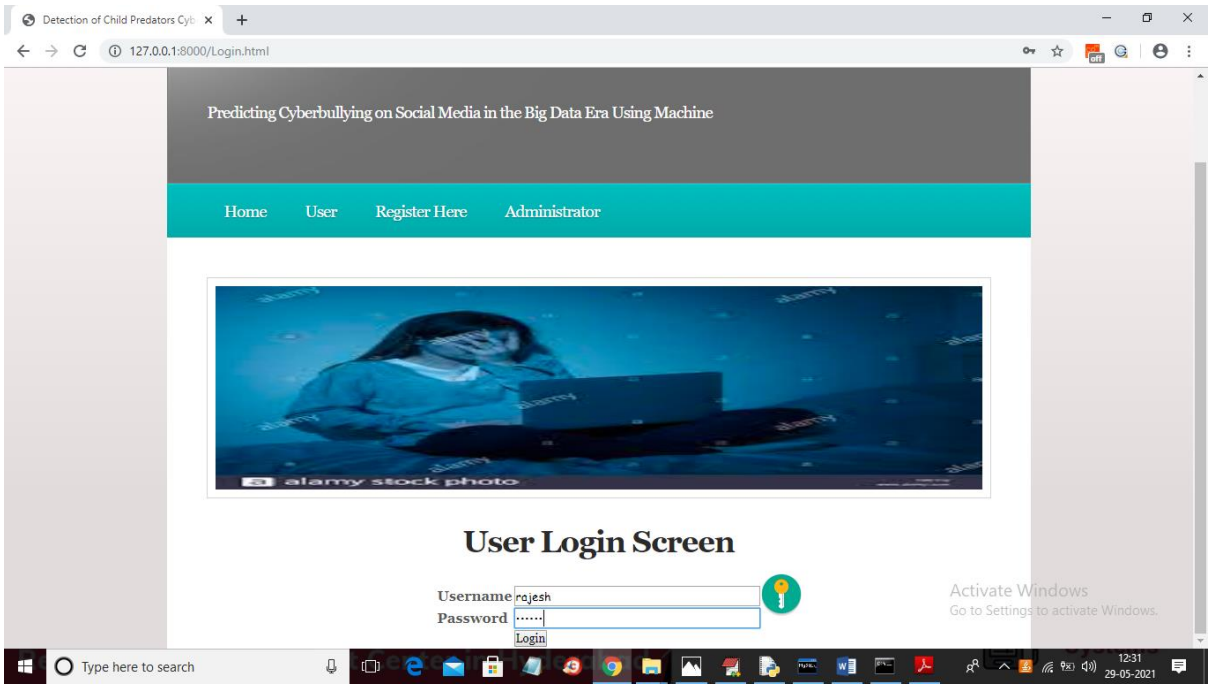
In above screen admin has to select each algorithm and click on 'Submit' button to train model and we will get accuracy also for each algorithm. Admin has to repeat this step whenever first time he starts the server or upon adding new bulling messages. Admin has to run at least SVM algorithm to perform automatic detection of harasser's or non-harasser's.



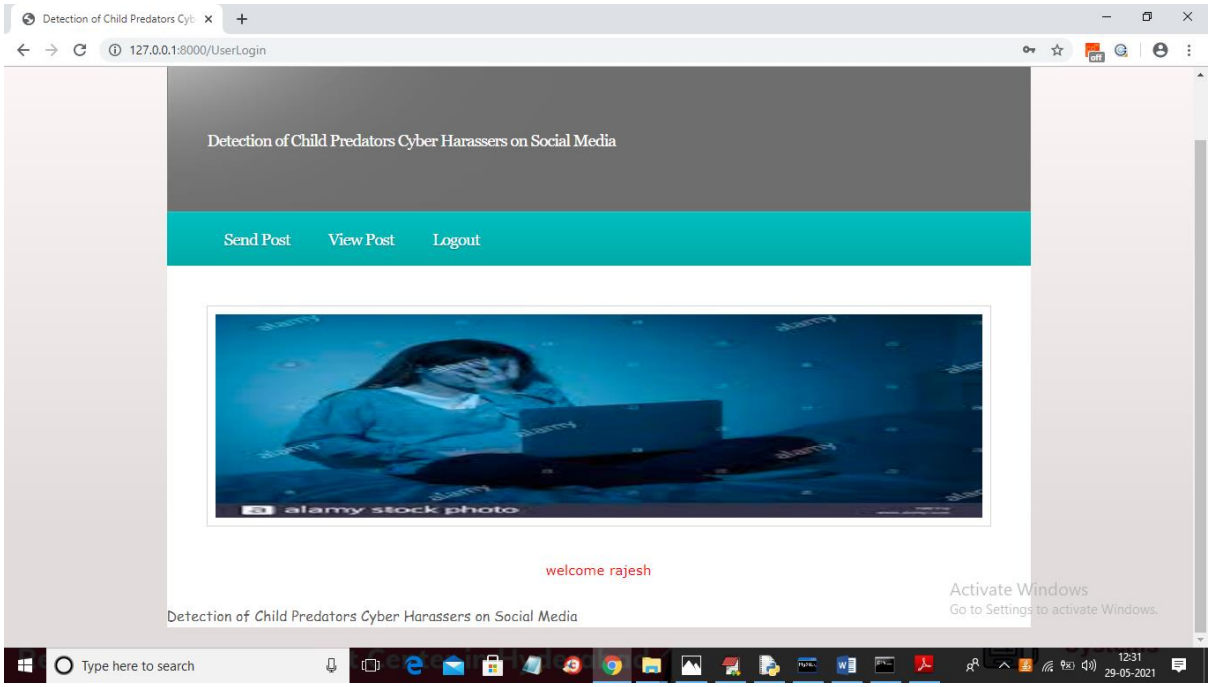
In above screen I ran SVM and got accuracy as 91. Similarly u need to select all algorithms one by one and run it.



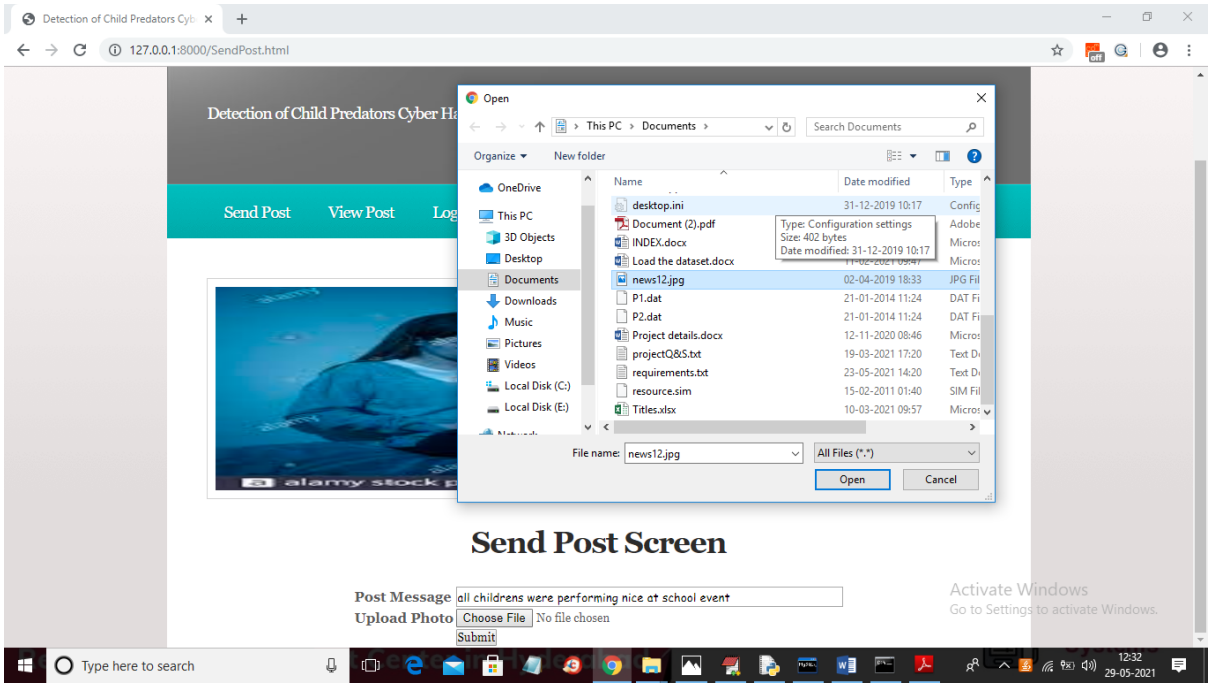
In above screen advance 'Random Forest' algorithm gave 52% accuracy. Now admin logout and login as user to send posts.



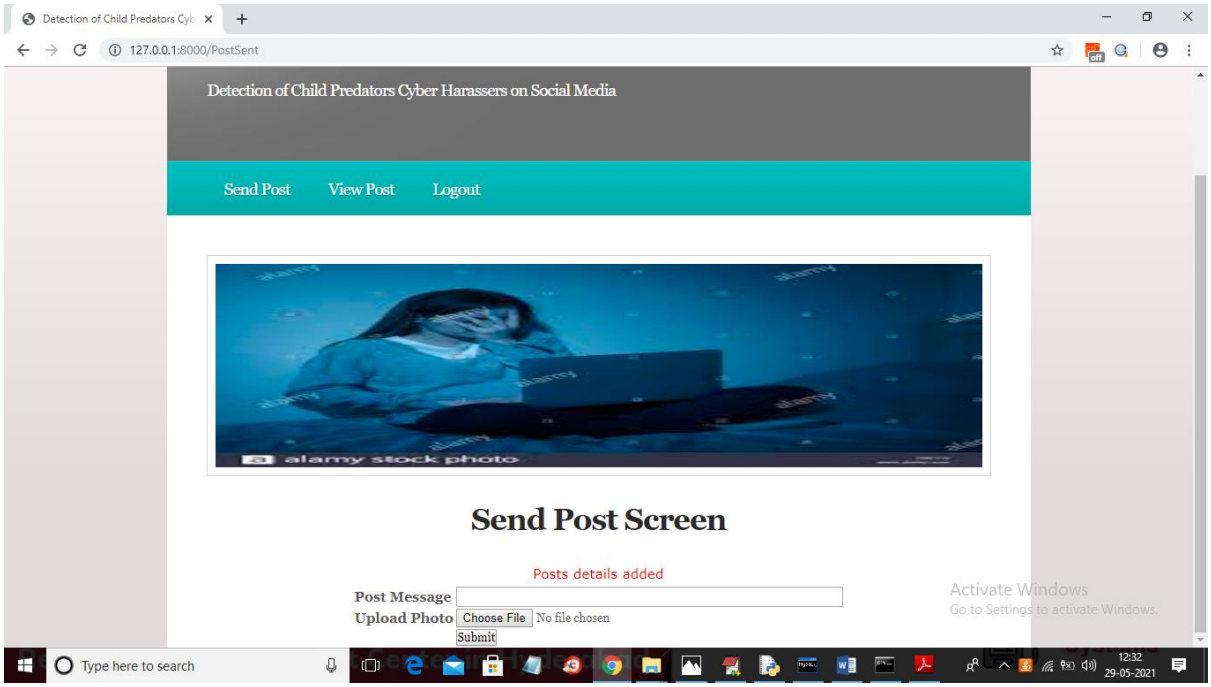
In above screen rajesh user is login and after login will get below screen



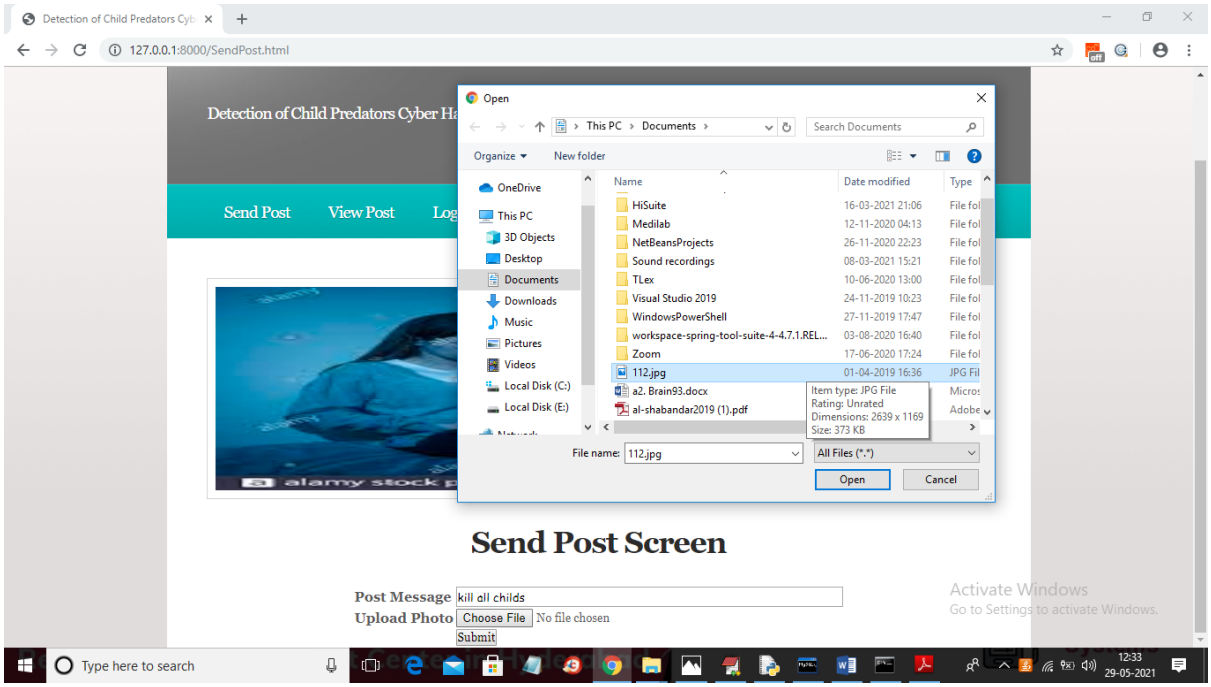
In above screen click on 'Send Post' link to get below screen



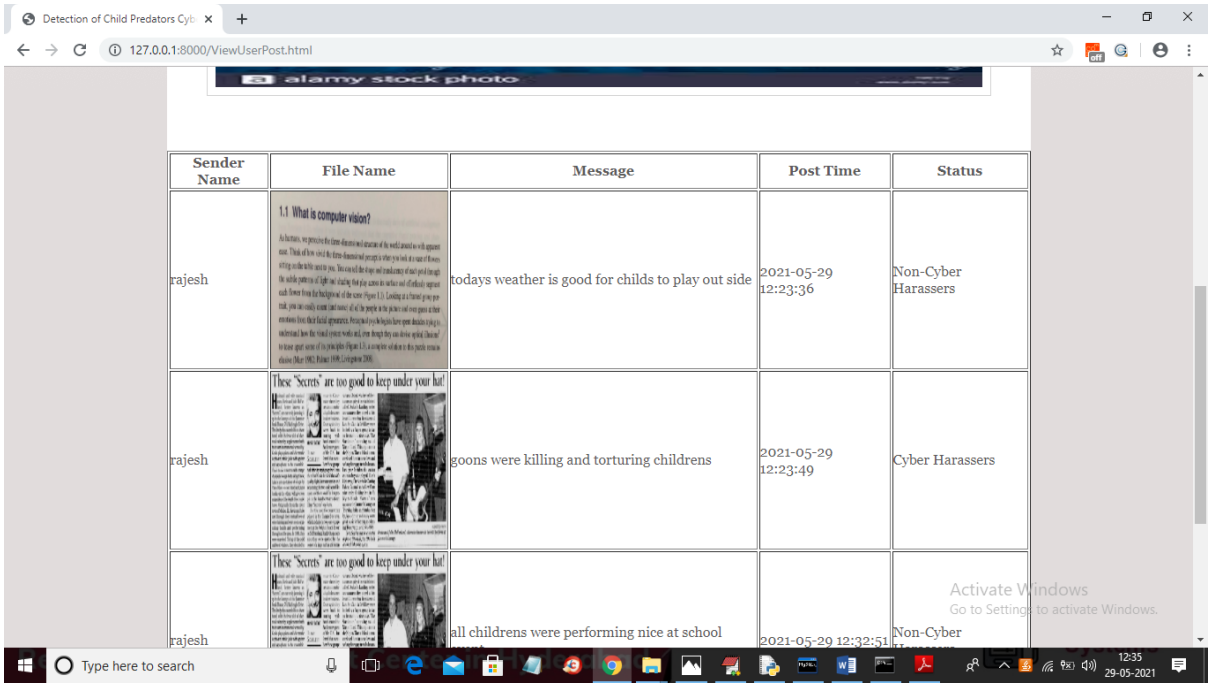
In above screen as post I added some messages and uploaded a photo also. After posting message will get below screen

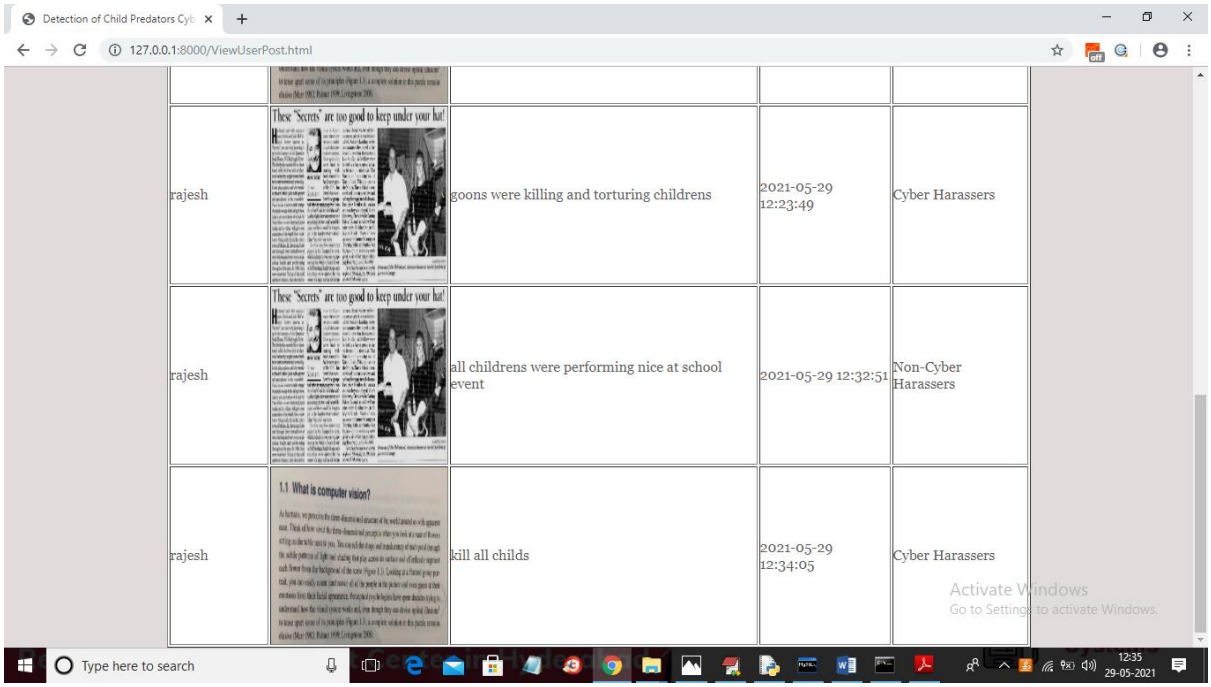


Now try another message



In above screen added another message and now click on ‘View Post’ link to get all posts sent by users





In above screen we are seeing posts from all users and we can see with the help of machine learning SVM algorithms we are automatically able predict message as cyber or non-cyber harasser's. Similarly you can signup new users and then post new messages.

Machine learning will predict harasser or non-harasser word based on dataset records so you add all possible harasser and non-harasser words to dataset by using 'add words' module from admin. After adding words then run algorithms link to train model and then you can see application will predict harasser or non-harasser automatically.

5. CONCLUSION

The cost to youngsters and society of sexual commission is simply too nice to overlook the hazards of on-line solicitation. The aim of the groomer is to build a relationship with a child inorder to gain access to that child. When grooming takes place it is common that an adult groomer is pretending to be a child with common hobbies or interests to build a relationship that includes trust with the child. In this project, we detect predators of children for child safety. And send report to cyber admin for action.

REFERENCES

[1] C. H. Ngejane, G. Mabuza-Hocquet, J. H. P. Eloff, and S. Lefophane, “Mitigating online sexual grooming cybercrime on social media using machine learning: A desktop survey,” in 2018 International Conference on Advances in Big Data, Computing and Data Communication Systems (icABCD), Aug 2018, pp. 1–6.

[2] N. Pendar, “Toward spotting the pedophile telling victim from predator in text chats,” in International Conference on Semantic Computing (ICSC 2007), Sep. 2007, pp. 235–241.

[3] I. McGhee, J. Bayzick, A. Kontostathis, L. Edwards, A. McBride, and E. Jakubowski, “Learning to identify internet sexual predation,” International Journal of Electronic Commerce, vol. 15, no. 3, pp. 103– 122, 2011.

[4] G. Inches and F. Crestani, “Overview of the international sexual predator identification competition at PAN-2012,” in CLEF 2012 Evaluation Labs and Workshop, Online Working Notes, Rome, Italy, September 17-20, 2012, 2012.

[5] E. Villatoro-Tello, A. Juarez-Gonz ´ alez, H. J. Escalante, M. Montes-y- ´ Gomez, and L. V. Pineda, “A two-step approach for effective detection ´ of misbehaving users in chats,” in CLEF 2012 Evaluation Labs and Workshop, Online Working Notes, Rome, Italy, September 17-20, 2012, 2012.

[6] G. Eriksson and J. Karlgren, “Features for modelling characteristics of conversations,” in CLEF 2012 Evaluation Labs and Workshop, Online Working Notes, Rome, Italy, September 17-20, 2012, 2012

[7] Muhammad Ali Fauzi, Patric Bours, “Ensemble Method for Sexual Predator Identification”. IEEE, 2020,25 June 2020

[8] Michael Ashcroft, Lisa Katti, Maxime Meyer “A Step Towards Detecting Online Grooming-Identifying Adults Pretending to be Children”. European Intelligence and Security Informatics Conference,2019

[9] Elif Varol Altay, Bilal Altas “Detection of Cyber Grooming in Online Conversation”. International Conference on Big Data Deep Learning and Fighting cyber–Terrorism Ankara, Turkey IEEE 03/12/2018

[10] Amparo Elizabeth Cano, Miriam Fernandes, Harith Alani “Detecting Child Grooming Behavior Patterns on Social Media” IEEE 2014.

[11] Batoul Haipar, Maroun Phamoun, Fadi yamout, “Cyber Bullying detection-A Multilingual Technique,” IEEE, DOI 10.1109/EMS 2016.

[12] John Ibanez Rodriguez, Santiago Rocha Duran, Daniel Diaz-Lopez, Javie Pastor-Galindo, Felix Gomez Marmol, “A Conversational Agent to Detect Online Sex Offenders.” IEEE 27 october 2020.

[13] Stefan C. Dombrowski, John W. LeMasney, and C. Emmanuel Ahia. “Protecting Children from online sexual PredatorsTechnological psychoeducational and legal considerations.”2004. www.ijcrt.org © 2021 IJCRT | Volume 9, Issue 7 July 2021 | ISSN:

2320-2882 IJCRT2107068 International Journal of Creative Research Thoughts (IJCRT) www.ijcrt.org a491

[14] Janis Wolak, David Finkelhor. "Online Predators and their Victims myths, realities and implications for prevention and treatment" February-March 2008.

[15] Amparo Elizabeth Cano, Miriam Fernandez, and Harith Alani, "Detecting Child Grooming Behaviour Patterns on Social Media."23 November 2015, Cambridge University Pressed.

[16] M. W. R. Miah, J. Yearwood, and S. Kulkarni, "Detection of child exploiting chats from a mixed chat dataset as a text classification task," in Proceedings of the Australasian Language Technology Association Workshop 2011, Canberra, Australia, December 2011, pp. 157–165.

[17] E. Villatoro-Tello, A. Jurez-Gonzlez, H. J. Escalante, M. M. y Gmez, and L. V. Pineda, "A two-step approach for effective detection of misbehaving users in chats." in CLEF (Online Working Notes/Labs/Workshop), P. Forner, J. Karlgren, and C. WomserHacker, Eds., 2012.

[18] H. J. Escalante, E. Villatoro-Tello, A. Juarez, M. Montes-y G´omez,´and L. Villasenor, "Sexual predator detection in chats with chained˜ classifiers," in Proceedings of the 4th Workshop on Computational Approaches to Subjectivity, Sentiment and Social Media Analysis. Atlanta, Georgia: Association for Computational Linguistics, June 2013, pp. 46–54.

[19] C. Peersman, F. Vaassen, V. Van Asch, and W. Daelemans, "Conversation level constraints on pedophile detection in chat rooms," in CLEF 2012 Conference and Labs of the Evaluation Forum – Uncovering Plagiarism, Authorship, and Social Software Misuse (PAN), P. Forner, J. Karlgren, and C. Womser-Hacker, Eds., Rome, Italy, 2012.

[20] C. Laorden, P. Galn-Garca, I. Santos, B. Sanz, J. M. G. Hidalgo, and P. G. Bringas, "Negobot: A conversational agent based on game theory for the detection of paedophile behaviour," in CISIS/ICEUTE/SOCO Special Sessions'12. Springer Berlin Heidelberg, 2012, pp. 261–270.

[21] F. Rangel, P. Rosso, I. Chugur, M. Potthast, M. Trenkmann, B. Stein, B. Verhoeven, and W. Daelemans, "Overview of the 2nd author profiling task at pan 2014," in CLEF 2014 Evaluation Labs and Workshop – Working Notes Papers, Sheffield, UK, 2014/09/18 2014.

[22] R. O'Connell, "A typology of child cyber sexploitation and online grooming practices," Cyberspace Research Unit, Tech. Rep., 2003.

[23] A. Kontostathis, "Chatcoder: Toward the tracking and categorization of internet predators," in Proc. of text mining workshop 2009 held in conjunction with the the 9th SIAM international conference on data mining., 2009.

[24] "Spaghetti book club - book reviews by kids for kids!" http://www.spaghettibookclub.org/.

[25] Snap: Web data: Amazon reviews," http://snap.stanford.edu/data/webAmazon-links.html.

[26] J. McAuley and J. Leskovec, "Hidden factors and hidden topics: Understanding rating dimensions with review text," in Proceedings of the 7th ACM Conference on Recommender Systems, ser. RecSys '13. New York, NY, USA: ACM, 2013, pp. 165–172.

[27] "The blog authorship corpus," http://u.cs.biu.ac.il/ koppel/BlogCorpus.htm.

[28] Y. Freund and R. E. Schapire, "Experiments with a new boosting algorithm," in In Proceedings of the thirteenth International Conference on Machine Learning., 1996, pp. 148–156.

ISSN- 2394-5125 VOL 10, ISSUE 04, 2023