JETIR.ORG ISSN: 2349-5162 | ESTD Year : 2014 | Monthly Issue
JOURNAL OF EMERGING TECHNOLOGIES AND INNOVATIVE RESEARCH (JETIR)
An International Scholarly Open Access, Peer-reviewed, Refereed Journal

DETECTION OF CHILD PREDATORS CYBER HARASSERS ON SOCIAL MEDIA

M.SUDHAKAR REDDY, MCA Student, Department of CSE, Amrita Sai institute of science and Technology, Andhra Pradesh, India
Mr.Dr P. CHIRANJEEVI, Department of CSE, Amrita Sai institute of science and Technology, Andhra Pradesh, India.

Abstract - Professional psychologists must comprehend the dangers of online sexual harassment and take steps to protect young people from sexual predators who utilize the internet. Although the internet has numerous advantageous features, one of its most sinister aspects is its potential for online sexual exploitation. The internet provides a medium for sex predators to target numerous children in a relatively anonymous environment. The prime aim of our project is to identify child predators based on their comments and posts on social media accounts and forward the predator records to the cyber cell administration. According to a recent national study, approximately 20% of youths are solicited for sexual purposes online every year (Finkelhor, Mitchell, & Wolak, 2000; Mitchell, Finkelhor, & Wolak, 2001).

This project report details our ongoing progress in the development of a system that addresses this issue. The result of this system will be the detection of child predator accounts and the ability to report these cases to the administration for appropriate action.

Psychologists must possess an apprehension of the hazards posed by cyberbullying and arm themselves with the techniques necessary to defend children from it. The prospect of online sexual exploitation constitutes one of the most pernicious facets of the internet, notwithstanding its advantageous features. The internet offers anonymity to predators, enabling them to reach numerous children with ease. Our project aspires to determine the social media accounts utilized by predators and furnish a report on the predator to the cyber cell administrator (Wolak, 2000; Mitchell, Finkelhor, & Wolak, 2001). This study report exemplifies our recent endeavors to fabricate a system that can assist the administrator in executing further measures subsequent to receiving a report from a victim of sexual assault.

1. INTRODUCTION

A web-based application has been designed to tackle the issue of child predator activity on social media platforms, such as Facebook and Instagram. The primary aim of the application is to identify and report instances of child predator behavior, in the form of comments and posts, to the Cyber Cell Administrator. For the effective detection and reporting of child predator activity, a comprehensive database is being established to store all comments and posts related to children's online social activities. The problem of child predator activity on social media is rapidly growing, and it is a matter of concern. A report by the National Society for the Prevention of Cruelty to Children in March 2014 highlights the extent of the problem, revealing that 12% of 11-16 year olds in the UK have received unwanted sexual messages, and 8% have received requests to send or respond to sexual messages.

The task of detecting cybersexual offenders who target children is a pressing matter that deserves utmost consideration. With the increasing usage of social media as the main mode of communication by teenagers, the threat of being exposed to such offenders becomes more pronounced. A study called SCAMP (Study of Cognition, Adolescents and Mobile Phones) showed that a significant proportion of children in the UK, 70% of 11-12 year olds and 90% of 14 year olds, possess a mobile phone. A tactic often employed by such offenders is referred to as online child grooming, where they engage in exchanging sexually explicit content through social media. This grooming involves establishing a relationship of trust with a minor, with the intention of eventually meeting them in person. Previous research has been conducted to detect cyber pedophilia online, including through the First International Sexual Predator Identification Competition.

Purpose

The aim of this effort is to provide a secure environment for minors who utilize social media platforms, for example, Facebook and Instagram. The endeavor intends to observe and report the online behavior of predators by tracking their comments and posts. The collected data will be submitted to the responsible administrator at the cyber cell. The heightened concern for the safety of minors utilizing social media has prompted the necessity for a robust system that can identify and report any sexually explicit content aimed towards them. A study conducted in March 2014 discovered that a significant percentage of 11-16 year olds, 12% in the state and 8% in the UK, had received unwanted sexual communications or invitations to participate in sexually explicit exchanges. It is imperative that this issue is addressed to ensure the protection of minors on the internet.

The utilization of social media for communication by minors has experienced a significant upsurge. A study called "Adolescents, Cell Phones, and Intelligence (SCAMP)" discovered that a considerable number of 11-12-year-old minors in the UK, around 70%, possess a cell phone, with this figure increasing to a substantial 90% when they reach the age of 14. This has resulted in the proliferation of an occurrence referred to as "online grooming," whereby adults use social media to foster trust with minors, thus enabling them to make personal contact. Prior research endeavors have concentrated on recognizing individuals engaging in cyber-pedophilia activities via the internet and have necessitated the initiation of international inquiries.

2. LITERATURE SURVEY

The realm of online entertainment, comprising games, audio chats, and numerous forums, has integrated the aspect of hunting into their designs. To ensure the safety of minors from the risk of online sexual abuse and exploitation, a system for detecting child predators has been devised. This system is equipped to identify potential dangers while minors are participating in games or utilizing online voice chat. The significance of having a system in place to recognize and counteract child predators becomes even more crucial as minors become more active on social media platforms. The present system uses either a dialogue-based approach with the Ridge or Naive Bayes classifier in the TF-IDF feature set, or a 5-level Neural Network algorithm.

With the purpose of preserving minors from the perils posed by cyberbullying and online sexual misconduct, it is imperative for psychologists to be well-informed regarding these hazards. The capability of perpetrators to reach a multitude of minors while concealed has been made possible through the internet. Thus, a project was instituted with the aim of establishing a system to distinguish social media accounts utilized by predators and to inform the cyber cell administrator of such accounts.

This project report offers an overview of the latest endeavors towards this objective and underscores the requirement for an improved system. The final objective is to grant the administrator the ability to take action after obtaining a report of sexual assault from a victim. By effectively recognizing and reporting these predators, it is hoped that the potential harm to minors can be curtailed and their safety on the internet can be enhanced. The outcome of the experiment conducted on the PAN12 dataset showcases the proficiency of our two-stage method, which involves the utilization of a soft vote ensemble for the preliminary stage and a Naive Bayes methodology for the subsequent stage. This technique obtained a remarkable F 0.5-score of 0.9348, thus ranking it among the premier performers in the PAN12 competition standings.

The authors, Michael Ashcroft, Lisa Kaati, and Maxime Meyer, in their scholarly composition "A Step towards Detecting Online Grooming - Characterizing Adults Pretending to be Children," tackle the matter of online grooming which is a pressing issue in the current society where a considerable amount of time is spent on the internet. Grooming perpetrators, to form connections and relationships with their youthful victims in online communities, often disguise themselves as minors. The authors present a two-fold method aimed at determining if an adult is pretending to be a child in chat room language. This approach commences by categorizing authors as minors or adults, followed by a scrutiny of each minor to distinguish between genuine minors and those impersonating minors. The outcomes of the study signify the feasibility of accurately determining the real children from adults masquerading as minors in chat logs, despite the challenge of separating typical adults from minors in such records. This scholarly article showcases the precision of the methodologies proposed and highlights the critical elements that played a role in their efficacy.

3. SYSTEM ANALYSIS:

The methodology for discovering malign content on a platform requires a combination of machine learning algorithms and various Python libraries, including Pandas.

The initial phase involves the examination of multiple posts to discern any malevolent conduct through statistical analysis. Subsequently, individuals whose level of suspicion surpasses a designated threshold are classified as suspects.

An in-depth analysis of the suspected user's posts, including multimedia elements such as images, audio recordings, and videos, is then performed. This analysis is executed by utilizing image and audio analysis tools, in conjunction with artificial intelligence, to conclude whether the suspect falls under the category of a predator.

The results obtained from this process aid in recognizing patterns of child grooming. Finally, potential predators are reported to the law enforcement agencies.

3.1 EXISTING SYSTEM

Existing methods for the discovery of online child predators are present in the realm of gaming, audio chat, and various online leisure sources. These methods serve to protect children from being subjected to sexual exploitation while they partake in online gaming or audio chatting activities. Nevertheless, in this age where the internet prevails, many children have begun to use social media platforms as their primary means of social interaction. Thus, the absence of a specific detection system to protect children on these platforms exposes them to a risk of harm from sexual predators.

The extant methodology encompasses the utilization of five algorithms for categorization: the Neural Network Classifier functioning on the TF-IDF feature set and the conversation-centered strategy utilizing the Ridge or Naive Bayes Classifier, which operates on the TF-IDF feature set.

Our intended system intends to adopt a unique technique for image and text classification with the intention of elevating accuracy as compared to pre-existing systems. This technique will be the Support Vector Machine (SVM), a machine learning method that is controlled and applied to two-category classification difficulties.

3.2 PROPOSED SYSTEM

In our project, the objective is to identify occurrences of child harassment on social media by utilizing several machine learning algorithms, specifically the Support Vector Machine (SVM), Random Forest, Naive Bayes, K-Nearest Neighbors, and Decision Tree. The training of the models will be conducted through the amalgamation of normal and harassing words and messages. The model, once trained, will be imposed upon new posts from users with the intent of determining if they consist of normal content or content that is harassing.

3.3 IMPLEMENTATION

Modules:

User module:

The User Component of the platform permits individuals to establish a profile by registering, whereby they gain access to the application upon successful login. Upon entering the system, users are equipped to send and inspect posts.

Admin Module:

The Administrator Module affords the administrator with the authority to inspect all recorded user accounts and to make judgements on the acceptance or rejection of new user accounts. The administrator has the duty of supplementing the machine learning training data set with new harassing or non-harassing messages. The administrator is obligated to execute one or more Support Vector Machine algorithms for the identification of harassing messages from the user side. Furthermore, the administrator is empowered to oversee and peruse all messages transmitted by all users.

4. SYSTEM DESIGN:

4.1 SYSTEM ARCHITECTURE

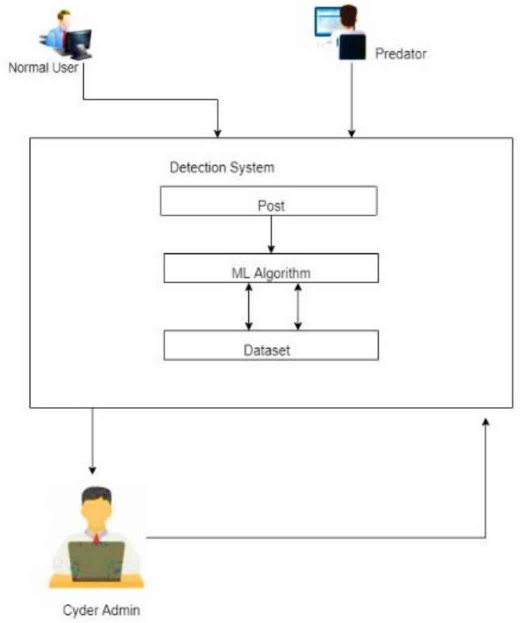


Figure 4.1: Architecture diagram

4.2 UML DIAGRAMS

4.2.1 USE CASE DIAGRAM

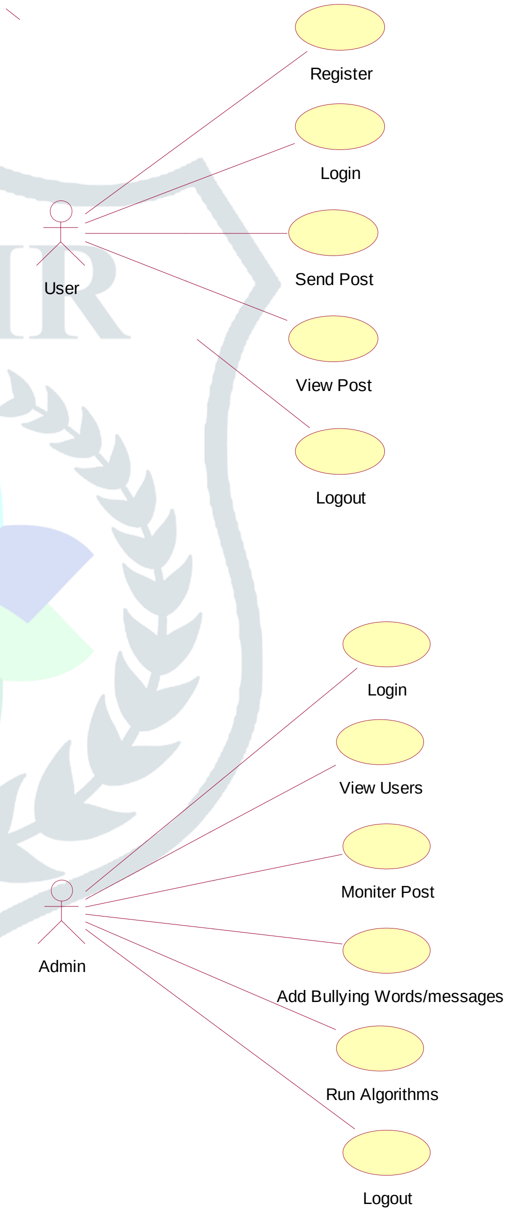


Figure 4.2.1 Case Diagram

4.2.2 SEQUENCEDIAGRAM

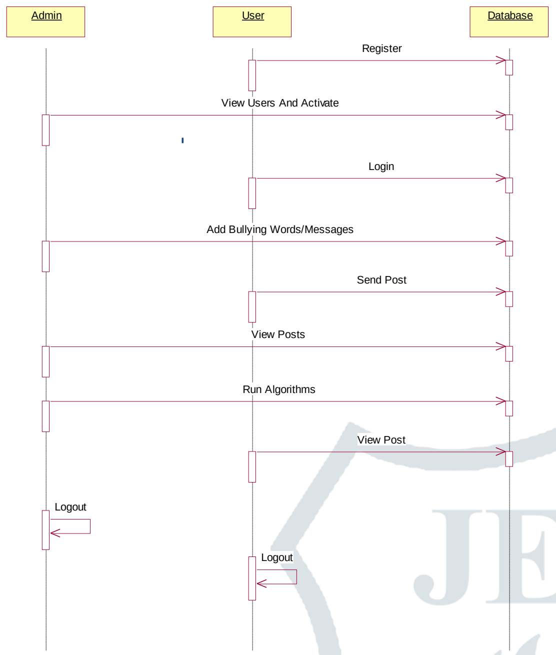


Figure 4.2.2: Sequence diagram

4.2.3 COLLABORATION DIAGRAM

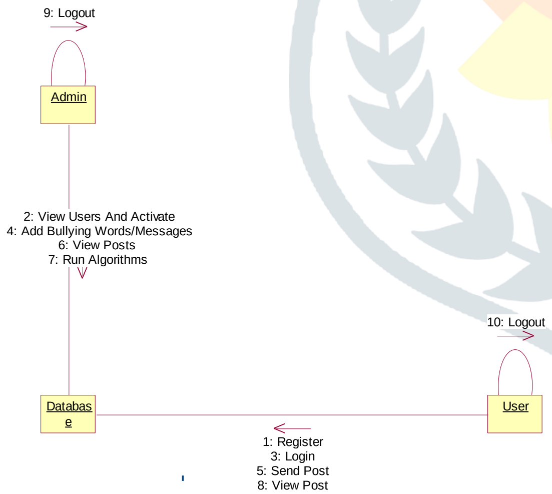


Figure 4.2.3: **COLLABORATION DIAGRAM**

CLASS DIAGRAM:

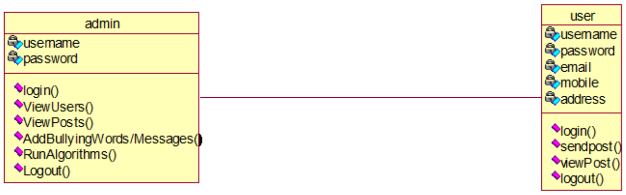
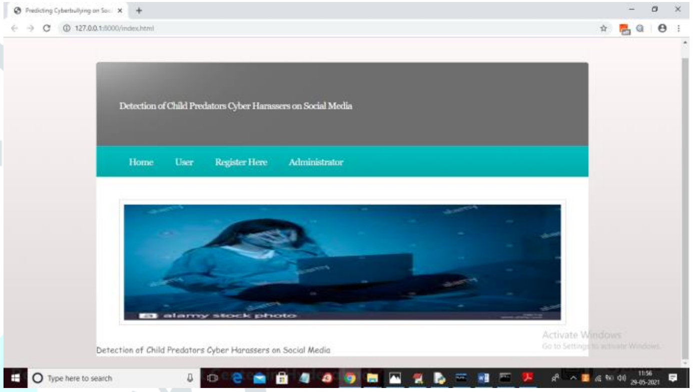


Figure 4.2.4: Class Diagram

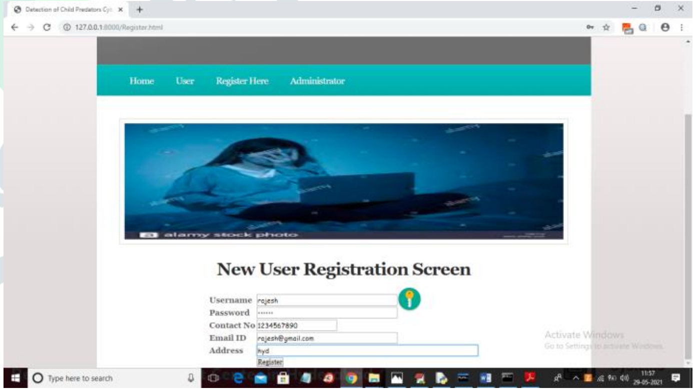
5. SCREEN SHOTS


Predicting Cyberbullying on Soc
127.0.0.1:8000/index.html
Home
User
Register Here
Administrator
alamy stock photo
Activate Windows
Go to Settings to activate Windows.
Type here to search

Detection of Child Predators Cyber Harassers on Social Media

Detection of Child Predators Cyber Harassers on Social Media

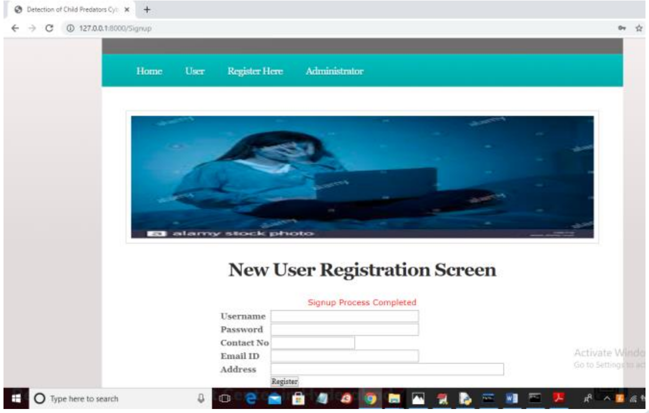
To establish a new user account, navigate to the aforementioned screen and activate the "Register Here" connection.


Detection of Child Predators Cy
127.0.0.1:8000/Register.html
Home
User
Register Here
Administrator
alamy stock photo
Activate Windows
Go to Settings to activate Windows.
Type here to search

New User Registration Screen

Username rajesh
Password ......
Contact No 1234567890
Email ID rajesh@gmail.com
Address hyd
Register

Please proceed to clicking the "Register" button displayed above, in order to input the relevant information.



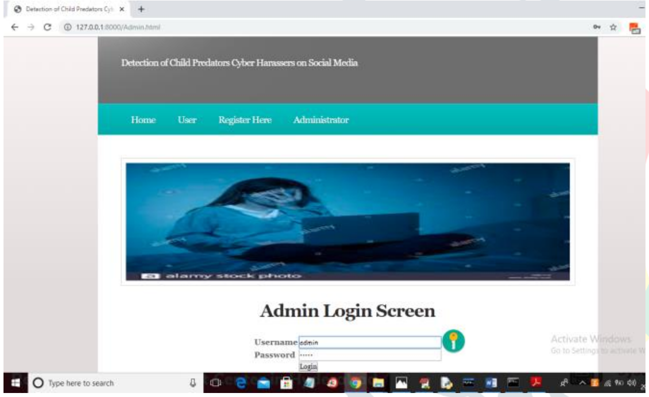
New User Registration Screen

Signup Process Completed

Username	
Password	
Contact No	
Email ID	
Address	

Register

Once the sign-up procedure on the screen mentioned above has been accomplished, access the new user information by clicking the "Administrator" link and logging in as the administrator.

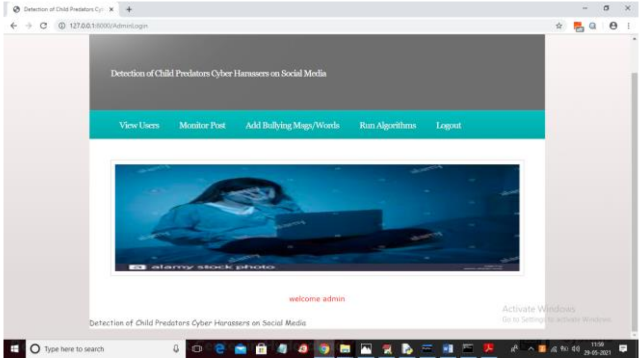

Detection of Child Predators Cy
127.0.0.1:8000/Admin.html
Detection of Child Predators Cyber Harassers on Social Media
Home
User
Register Here
Administrator
alamy stock photo
Type here to search

Admin Login Screen

Username	admin
Password	•••••

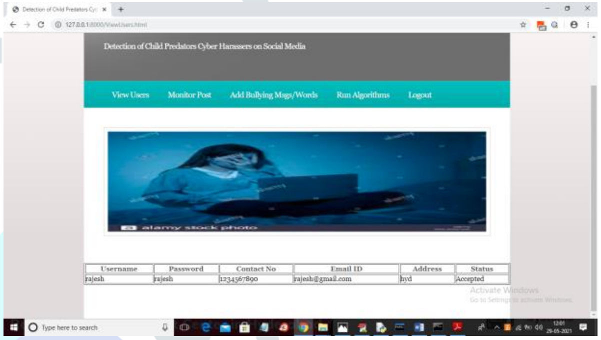
Login

In order to access the below screen, one must log in as the "admin" user on the aforementioned screen by providing "admin" as both the username and password. Upon successful login, the subsequent screen will become accessible.


Detection of Child Predators Cy
127.0.0.1:8000/AdminLogin
Detection of Child Predators Cyber Harassers on Social Media
View Users
Monitor Post
Add Bullying Msgs/Words
Run Algorithms
Logout
alamy stock photo
Detection of Child Predators Cyber Harassers on Social Media
Type here to search

welcome admin

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


Detection of Child Predators Cy
127.0.0.1:8000/ViewUsers.html
Detection of Child Predators Cyber Harassers on Social Media
View Users
Monitor Post
Add Bullying Msgs/Words
Run Algorithms
Logout
alamy stock photo
Type here to search

Username	Password	Contact No	Email ID	Address	Status
rajesh	rajesh	1234567890	rajesh@gmail.com	hyd	Accepted

In the screen above, the creation of the "Rajesh" account is demonstrated. The administrator can gain access to a history of posts made by users by clicking the "Monitor Posts" button.

6. CONCLUSION

The cost to youth and society from sexual exploitation is too severe to disregard the risks posed by online soliciting. Child groomers aim to establish connections with minors to gain access to them, frequently masquerading as a child sharing common interests and pastimes. The objective is to create a relationship built on trust with the child. This project endeavors to detect these predators for the sake of child protection and, upon detection, to promptly notify the cyber administrative authorities for appropriate action.

The method for scrutinizing suspect content on a platform involves the following successive steps:

- Procurement of data from the suspected user's posts, incorporating multimedia aspects such as images, audio, and videos.
- Examination of the procured data through the utilization of the IGPL Python package, Urllib, artificial intelligence, and the NSFW library.
- Determining the suspect's classification as either a suspect or predator, based on the outcome of the examination.
- Analysis of patterns of child grooming and statistical results for classifying the individual as a predator.
  Automated reporting of the predator classification to a Gmail address, which is stored on the server.

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.