Journal Homepage: - www.journalijar.com

INTERNATIONAL JOURNAL OF ADVANCED RESEARCH (IJAR)

Article DOI: 10.21474/IJAR01/18173
DOI URL: http://dx.doi.org/10.21474/IJAR01/18173

RESEARCH ARTICLE

A COMPREHENSIVE REVIEW ON CYBERBULLYING PREVENTION

Dr. Rajesh L.[1], Ajay V.S[2], Basannagouda Biradar[2], Bhanav V.H[2] and Parwez Ahmed[2]

1. Associate Professor, Department of Information Science and Engineering, Dayananda Sagar Academy of Technology and Management, Bengaluru, India.
2. B.E. Students, Department of Information Science and Engineering, Dayananda Sagar Academy of Technology and Management, Bengaluru, India.

Manuscript Info

Manuscript History
Received: 15 November 2023
Final Accepted: 19 December 2023
Published: January 2024

Key words:-
Cyberbullying, Naive Bayes, Support Vector Machine (SVM), Classification, Machine Learning, Online Social Network, Text Classification in Social Media

Abstract

The rise of cyberbullying in the digital age necessitates innovative approaches for prevention and intervention. This project proposes a probabilistic framework for Cyber Bullying Prevention utilizing machine learning techniques, support Vector Machines (SVM) and Naïve Bayes classifiers, in particular, are the focus.. The objective is to develop a robust system capable of identifying and mitigating instances of cyberbullying in online environments. The proposed model integrates advanced natural language processing and sentiment analysis to effectively analyze textual content and contextual cues within digital communications. By employing SVM and Naïve Bayes algorithms, the system aims to discern patterns indicative of cyberbullying behavior, achieving a probabilistic assessment of the likelihood of such occurrences. Through a training dataset enriched with diverse cyberbullying instances, the model learns to generalize its understanding and adapt to evolving online communication dynamics. The project further explores feature engineering and optimization techniques to enhance the classifiers' performance. The ultimate goal is to provide a proactive and accurate cyberbullying prevention tool that empowers users, platform administrators, and law enforcement agencies to intervene and mitigate the impact of cyberbullying effectively. This research contributes to the broader discourse on utilizing machine learning in social contexts, emphasizing the importance of collaborative efforts to create a safer digital space for all users.

Introduction:-

In the ever-evolving landscape of the digital age, the escalation of cyberbullying represents a substantial risk to the welfare and security of individuals within online communities. Addressing this issue requires a proactive and sophisticated approach that leverages the capabilities of machine learning. This project introduces a comprehensive framework for Cyber Bullying Prevention, utilizing two key machine learning algorithms include Naïve Bayes and Support Vector Machines (SVM).The primary focus is on developing robust classifiers capable of discerning patterns indicative of cyberbullying behavior within the vast realm of online social networks and text-based communication platforms. Cyberbullying, characterized by the malicious use of digital technologies to harm individuals, is a

Corresponding Author:- Dr. Rajesh L.
Address:- Associate Professor, Department of Information Science and Engineering, Dayananda Sagar Academy of Technology and Management, Bengaluru, India.
Mail-id: rajesh-ise@dsatm.edu.in

pervasive issue demanding novel and effective preventive measures. By harnessing the power of machine learning, particularly Naïve Bayes and SVM, this research seeks to create a system that can analyze textual content, recognize contextual cues, and probabilistically assess the likelihood of cyberbullying instances. Key components of the project include advanced natural language processing, sentiment analysis, and the incorporation of diverse cyberbullying instances within the training dataset. This research contributes to the broader field of machine learning in social contexts, emphasizing the significance of text classification for creating a safer online environment. The amalgamation of these technologies aims to empower users, platform administrators, and law enforcement agencies to proactively identify and address cyberbullying, thereby fostering a more secure and inclusive digital space.

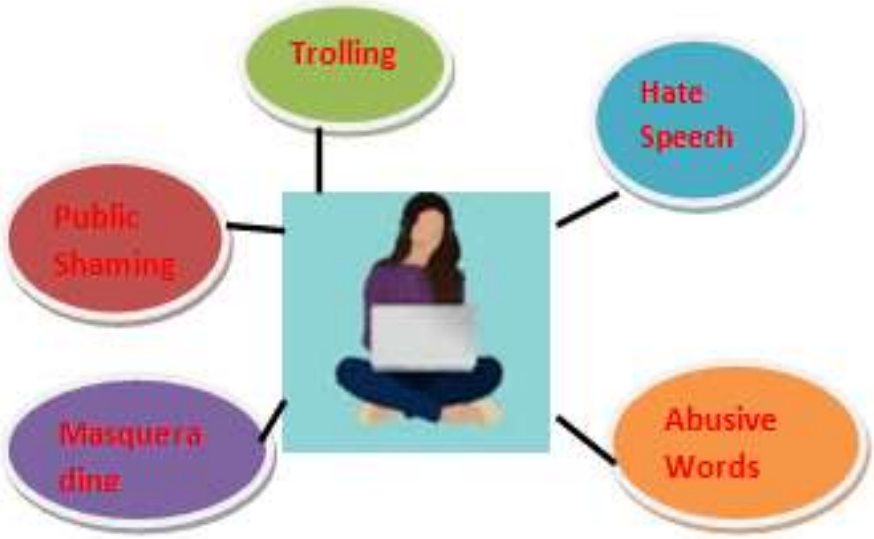

Fig 1:- Cyberbullying.

Literature Survey:-

The proposition put forth by the authors R. R. Dalvi et al, in the paper [1] is that, the paper tackles the growing concern of cyberbullying on social media platforms, particularly Twitter, by proposing a machine learning-based method for detection and prevention. With the increasing use of social networking sites, the study focuses on identifying commonalities in language used by bullies in tweets. Two classifiers, SVM (Support Vector Machine) and Naïve Bayes, are employed for training and testing on a dataset collected from various sources. Notably, SVM surpasses Naïve Bayes, achieving an accuracy of 71.25%, highlighting its effectiveness in discerning cyberbullying content. The model utilizes the Twitter API to retrieve real-time tweets for evaluation, demonstrating its capability to recognize and log instances of bullying. The proposed solution involves preprocessing, feature extraction, and algorithm selection, presenting a comprehensive framework for addressing cyberbullying on social media. The study underscores the importance of proactive measures in identifying and preventing cyberbullying, making a valuable contribution to the ongoing discourse on the integration of machine learning and societal challenges.

The contention made by the authors S. Mestry et al, in the paper [2] is that escalating issue of cyberbullying on social media platforms. The paper introduces a deep learning system that utilizes fastText word embedding and Convolutional Neural Network (CNN) for the automated identification and classification of toxic comments. The model aims to categorize discussions into six classes: Toxic, Severe Toxic, Obscene, Threat, Insult, and Identity-hate, providing a comprehensive solution to the challenges posed by negative content. The proposed approach leverages fastText's efficiency in word embedding, outperforming Word2Vec and GLOVE models. The authors emphasize the importance of automating the comment classification process to save time and manual effort in platform moderation. The study utilizes Kaggle's Toxic Comment Classification Dataset and focuses on creating a word embedding mechanism, efficient pre-processing, and CNN-based classification. The proposed model demonstrates promising results in enhancing social media experiences by effectively identifying and addressing toxic comments, contributing to the ongoing discourse on mitigating online negativity.

In research paper [3] the authors M. A. Al-Garadi et al, explore the challenges and methodologies of predicting and combating cyberbullying in the realm of social media (SM). They emphasize the transformative impact of Information Communication Technologies (ICT) on interpersonal communication, extending beyond traditional boundaries. The paper delves into the development of predictive models to address the rising online aggression facilitated by social technologies. The focus is on cyberbullying prediction models, covering data collection, feature engineering, and the application of various machine learning algorithms. The authors highlight the subjective nature of analyzing big data, incorporating human bias and subjectivity in defining cyberbullying. Challenges include the dynamic nature of language, cultural variations, and the need for unsupervised machine learning to discern patterns effectively. The study identifies the necessity of predicting cyberbullying severity, considering cultural shifts and language dynamics. The paper concludes by advocating for interdisciplinary collaboration and the exploration of advanced deep learning techniques to process and analyze the rapidly generated big data from social media, especially in the context of cyberbullying.

M. Yao et al. expressed that [4], a novel algorithm named CONcISE is introduced for detecting cyberbullying on Instagram, incorporating optimal online feature selection. The algorithm addresses the need for accurate methods to detect cyberbullying and the scalability of existing approaches. It frames the task of cyberbullying detection as a sequential hypothesis testing problem and introduces an innovative algorithm with the goal of reducing the number of features employed in classification. The approach was demonstrated using real-world Instagram data and showed significant improvements in recall while diminishing the average feature count in contrast to prevailing methods is a central objective. The investigation concentrates on Instagram, given its highest percentage of users reporting incidents of cyberbullying. The platform's features, such as pictures, captions, hashtags, and comments, are used to detect cyberbullying behavior, which includes hateful comments, humiliating images, and impersonation of victims. The study also highlights the staggering statistics of cyberbullying and its consequences, emphasizing the need for effective detection methods. The proposed algorithm uses a step-by-step assessment process aimed at minimizing the utilization of features for cyberbullying detection without compromising accuracy.

It fine-tunes the cost function to reduce the number of features, taking into account the effort and cost associated with feature evaluation. The methodology was assessed against baseline approaches, encompassing both supervised and weakly supervised learning methods, and outperformed them in terms of recall, precision, F-measure, and MCC. Additionally, the algorithm was compared to a non-sequential testing method, showing comparable performance while using significantly fewer features.

The authors Hinduja et al., assert in the paper [5] that, This exploratory study investigates cyberbullying through an online survey, collecting data from over 6,800 respondents between December 2004 and January 2005. The survey targeted adolescents by linking to various websites. Despite efforts, a gender bias emerged, with 82% of respondents being female. The study challenges traditional bullying research, revealing that cyberbullying does not discriminate based on gender or race. The findings establish a link between cyberbullying and traditional schoolyard bullying, indicating that those bullied at school are more likely to experience cyberbullying. Qualitative details from victims highlight the severity of cyberbullying, including physical threats. The study acknowledges limitations, such as the convenience sample and the cross-sectional design, emphasizing the need for future research to replicate the study with a more controlled sample. The paper concludes by proposing policy solutions, emphasizing awareness and ongoing discussions among parents, caregivers, teachers, and law enforcement to address the complex issue of cyberbullying.

The authors in the paper [6] puts forward the author Yun-yin Huang et al suggestion that the study investigates cyberbullying in Taiwan junior high school students of among 545, addressing the gap in limited research on the subject in the East Asian context. The study explores the frequency and correlates of cyberbullying, including gender, academic achievement, technology types, and anonymity. Findings reveal that male students are more likely to engage in cyberbullying, academic achievement does not significantly affect cyberbullying, and instant messenger users experience more cyberbullying than users of other technologies. Indifference is identified as the dominant attitude toward cyberbullying, raising concerns about the lack of prevention efforts. Peers, often the first confidants for teenagers facing cyberbullying, tend to take no action due to a preference for conflict avoidance and group harmony. The study emphasizes Taiwan's cultural context, including Confucian philosophy. The prevalence of cyberbullying in Taiwan, driven by the popularity of instant messengers, underscores the need for increased awareness and prevention efforts. Limitations include the self-report method, and future research is recommended to

explore causal relationships and additional factors, such as school climate and psychological conditions of cyberbullying roles.

The contention made by the authors Barbara A et al in the paper [7] discusses the use of online social marketing approaches for cyberbullying prevention and intervention targeted at young people. Traditionally, interventions have focused on individual, class, and whole school community levels, but with mixed results. The document explores the potential of online social marketing campaigns to go beyond traditional school-based interventions and influence cyberbullying prevention and reduction. It presents the Safe and Well Online Study, which involved the development and evaluation of four online social marketing campaigns aimed at promoting respect, positive attitudes, and help-seeking behaviors among young people. The document emphasizes the shift from solely focusing on the "message" to considering causal pathways and contributing conditions of ill/health, and how social marketing approaches can promote protective well-being factors and address risk factors related to cyberbullying and aggression. It highlights the need to move beyond school settings and consider the digital environment in which young people socialize. The methodology section outlines the research design, including thematic literature reviews, participatory design studies, age-cohort studies, digital data collection, and post-campaign qualitative studies. The study involved a large number of participants and aimed to understand the impact and reach of the online campaigns.

According to the authors Li et al thesis, in paper [8], the research article examines the trends and sex disparities in school bullying victimization among U.S. youth from 2011 to 2019. The study used data from the national Youth Risk Behavior Survey (YRBS) and included 72,605 high school students [1]. The findings revealed that the prevalence of traditional victimization and cybervictimization among high school students was 19.74% and 15.38% respectively, with female students reporting higher rates of both forms of victimization. The study did not observe a decline in the prevalence of traditional victimization and cybervictimization over the years. The overlap between traditional victimization and cybervictimization was about 60%. The prevalence of being bullied through electronic means was found to be higher than previously reported, indicating that cyberbullying is not a low-frequency phenomenon. The study suggests that more work is needed to address bullying and achieve the Healthy People 2020 goal of reducing bullying among U.S. high school students, especially among female students. The findings highlight the need for evidence-based interventions to address sex disparities in traditional and cyberbullying.

The authors Pornpongtechavanich et al [9] advocate for the application of a flipped classroom with a challenge-based learning model on an online streaming ecosystem to develop coping skills in cyberbullying. The study involved synthesizing international research papers, selecting experts, and evaluating the suitability of the model. The research found that cyberbullying is a significant concern, with factors such as gender, age, anxiety, depression, and social support contributing to cyberbullying. The study also highlighted the increasing use of online streaming platforms and the factors that affect their effectiveness, such as ease of use, media options, credibility, and content. The designed flipped classroom model, combined with challenge-based learning, was assessed by experts and found to be highly suitable for developing coping skills in cyberbullying. The model aimed to develop both cognitive and practical skills to prevent cyberbullying, and the evaluation results confirmed its effectiveness. The study concluded that the designed model is appropriate for skill development to prevent cyberbullying and that it holds significant potential for promoting cyberbullying skills development. The authors also acknowledged the importance of future research to put the model into practice and further validate its effectiveness.

According to the authors Hinduja S et al paper “A Preliminary Look at Cyberbullying" [12] discusses the emergence of cyberbullying and its impact on adolescents. The authors, Justin W. Patchin and Sameer Hinduja, explore the nature and extent of online bullying, as well as its potential negative repercussions on both victims and instigators. They conducted a pilot study to assess the prevalence of cyberbullying and its characteristics. The document also discusses the effects of traditional bullying and the transition to cyberbullying, highlighting the potential long-term implications of mistreatment during adolescence. Key findings from the study show that cyberbullying is prevalent among adolescents, with over 29% reporting being victims of online bullying and 11% admitting to bullying others online. The study also reveals the negative effects of cyberbullying on victims, including frustration, anger, sadness, and impacts on school, home, and friendships. Additionally, the document highlights the lack of significant associations between age, race, or gender and the likelihood of being a victim of online bullying. The authors emphasize the need for further research to better understand the scope, prevalence, and nuances of cyberbullying. They also suggest the importance of parental supervision, teacher vigilance, and police intervention to address and mitigate the harms of cyberbullying. In summary, the document provides valuable insights into the prevalence and

impact of cyberbullying on adolescents, highlighting the urgency of further research and the need for proactive measures to address this growing social concern.

The authors Huang H forward the idea in the paper [11] the reviews research on bullying and peer victimization in Chinese schools. The study seeks to understand the factors contributing to these issues and their implications for practice and policy. The authors highlight the prevalence and definition of bullying in Chinese schools, emphasizing cultural differences in the understanding of bullying. They also discuss the socio-demographic factors associated with bullying, such as age, gender, and behavioral/mental health problems. Additionally, the authors explore factors at the microsystem level, including the role of parents, peers, and teachers, as well as the mesosystem level, focusing on the interactions between these microsystems. They also examine the ecosystem level, discussing the impact of mass media on bullying behaviors. Finally, the study delves into the macrosystem level, addressing the influence of cultural values, academic achievement, and collectivism versus individualism on bullying and peer victimization. The authors propose implications for practice and research, emphasizing the need for ecologically-based prevention and intervention strategies. They suggest the importance of assessing socio-demographic characteristics, such as gender, age and behavioral/mental health problems, when addressing bullying. Additionally, they recommend involving parents and teachers in bullying prevention efforts and promoting a whole-school approach to address these issues effectively. Overall, the study provides a comprehensive overview of the factors associated with bullying and peer victimization in Chinese schools, highlighting the need for culturally relevant intervention measures and policy considerations.

According to the authors Tangen et al viewpoint in the paper [14], it highlights the prevalence and consequences of cyberbullying, emphasizing the need for intervention programs. The study compares bullying incidences and students' perceptions in a school with a Philosophy for Children (P4C) approach and other schools. Surprisingly, while the P4C School reported more face-to-face bullying, there were no significant differences in cyberbullying incidences. Both groups of pupils felt that teachers were more likely to prevent face-to-face bullying than cyberbullying, indicating a need for explicit teaching strategies about cyberbullying. The P4C approach focuses on developing critical thinking and social skills through philosophical discussions, aiming to address the broader context of social interactions and problem-solving. However, it was observed that despite participating in the P4C program, students did not show a significant difference in cyberbullying incidences. This suggests a potential disconnect between the program's objectives and students' abilities to address bullying situations on their own.

Author's name	**Data set used**	**Methodology used**	**Efficiency**	**Year of publication**
Phisit Pornpongtechavanich et al [9]	Flipped Classroom with Challenge-Based Learning Management Model	With a challenge-based Learning management model	90%	11 NOV 2021
Li et al [8]	Youth Risk Behavior Surveillance System (YRBSS)	The YRBS is an ongoing, biennial, cross-sectional, school-based survey of a representative sample of high school students from across the U.S. that monitors the prevalence of health-related behaviors	-	21 OCT 2020
R. R. Dalvi et al[1]	Kaggle, GitHub, and social media platforms	SVM (Support vector machine) and Naïve Bayes algorithm	SVM-71.25% Naïve bayes-52.70%	2020
S. Mestry et al [2]	Kaggle's Toxic Comment Classification Dataset on Wikipedia's talk	Word2Vec and GLOVE model.	10-70%	2019

	page edits			
Barbara A et al [7]	Participatory Design (PD) study, Digital Data.	Participatory design, age-cohort study, digital data collection, and post-campaign qualitative study.	-	2018
M. Yao et al [4]	Instagram data set , Noswearing.com	Optimization Algorithm DRFS (Dimensionality Reduction and Feature Selection) RF (Random Forest)	~90%	2018
Yun-yin Huang et al [6]	Self-report and peer-report surveys	Self-report and peer-report surveys are the most frequently used methods of collecting data on school bullying	-	NOV 2010
Hinduja et al [5]	an on-line survey methodology from over 6800 respondents	Data Analysis Software: Tools like SPSS tools used	-	1 FEB 2008
M. A. Al-Garadi et al [3]	Academic databases and search engines	SVM, Naive Bayes Algorithm, Random Forest, Decision Tree	30-70%	2nd April 2006

Table 1:- Comparative study of existing approaches.

The study also revealed that students perceived adults to be less aware of cyberbullying and less likely to prevent it compared to face-to-face bullying. There was a clear indication that teachers and guidance counsellors need to extend prevention and intervention programs to address cyberbullying at a much younger age. The findings emphasize the importance of explicit teaching about cyberbullying and the need for teachers to engage in the digital world of students to build their confidence in adults' ability to address cyberbullying. In conclusion, the study sheds light on the necessity of comprehensive anti-cyberbullying programs in primary schools and the importance of teacher training to effectively address and prevent cyberbullying. The P4C approach, while valuable in developing critical thinking skills, may need to be supplemented with explicit anti-cyberbullying strategies. Further research is explored to the effectiveness of such programs and warranted in preventing cyberbullying.

Problem Identified:-

The research papers collectively address the pervasive issue of cyberbullying, emphasizing its escalating prevalence and detrimental consequences in contemporary society. Focus areas include machine learning and deep learning methods for detection and prevention on platforms like Twitter and Instagram. Challenges in predicting and combating cyberbullying in the big data era are explored, emphasizing interdisciplinary collaboration. Unique features of Instagram are highlighted in a paper presenting a novel algorithm for detection. Gender biases, the link between cyberbullying and traditional bullying, and online threat severity are revealed in an exploratory analysis. Cultural nuances and gender-related patterns in cyberbullying among junior high school students in Taiwan are examined. Online social marketing approaches are advocated in another paper, urging campaigns beyond traditional interventions. Trends and sex disparities in school bullying victimization among U.S. youth persist, requiring evidence-based interventions.

A flipped classroom model is proposed for developing coping skills, considering factors like gender, age, anxiety, depression, and social support. The emergence of online bullying and its negative repercussions on adolescents are underscored, calling for further research and proactive measures. Factors associated with bullying and peer victimization in Chinese schools are reviewed, proposing ecologically-based prevention strategies. A primary school's cyberbullying prevention approach reveals a potential gap, emphasizing the necessity of explicit teaching strategies and teacher engagement. In summary, the multifaceted nature of cyberbullying necessitates comprehensive, interdisciplinary, and culturally relevant prevention and intervention approaches.

Conclusion:-

The amalgamation of diverse research on cyberbullying emphasizes its pervasive and evolving nature, transcending cultural boundaries. The studies advocate for proactive, multifaceted interventions, including technological solutions and educational strategies. Machine learning-based approaches demonstrate promise in identifying and preventing cyberbullying on social media platforms. Cultural nuances, evidenced in studies from specific regions, highlight the

importance of tailoring prevention strategies. Additionally, research underscores the need for holistic approaches, such as online social marketing campaigns and flipped classroom models, leveraging digital platforms to instill coping skills among the younger generation. In conclusion, addressing cyberbullying requires interdisciplinary efforts, integrating technology with education, and emphasizing ongoing research and awareness campaigns for a safer digital space

References:-

[1] R. R. Dalvi, S. Baliram Chavan and A. Halbe, "Detecting A Twitter Cyberbullying Using Machine Learning," 4th International Conference on Intelligent Computing and Control Systems (ICICCS), Madurai, India, 2020, pp. 297-301, doi: 10.1109/ICICCS48265.2020.9120893.
[2] S. Mestry, H. Singh, R. Chauhan, V. Bisht and K. Tiwari, "Automation in Social Networking Comments With the Help of Robust fastText and CNN," 1st International Conference on Innovations in Information and Communication Technology (ICIICT), Chennai, India, 2019, pp. 1-4, doi: 10.1109/ICIICT1.2019.8741503.
[3] M. A. Al-Garadi et al., "Predicting Cyberbullying on Social Media in the Big Data Era Using Machine Learning Algorithms: Review of Literature and Open Challenges," in IEEE Access, vol. 7, pp. 70701-70718, 2019, doi: 10.1109/ACCESS.2019.2918354.
[4] M. Yao, C. Chelmis and D. -S. Zois, "Cyberbullying Detection on Instagram with Optimal Online Feature Selection," IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining (ASONAM), Barcelona, Spain, 2018, pp. 401-408, doi: 10.1109/ASONAM.2018.8508329.
[5] Hinduja, Sameer & Patchin, Justin. Cyberbullying: An Exploratory Analysis of Factors Related to Offending and Victimization. Deviant Behavior – DEVIANT BEHAV. 29. 129-156. 10.1080/01639620701457816. 2008.
[6] Yun-yin Huang, Chien Chou, An analysis of multiple factors of cyberbullying among junior high school students in Taiwan,Computers in Human Behavior, Volume 26, Issue 6, 2010, Pages 1581-1590, ISSN 0747-5632, https://doi.org/10.1016/j.chb.2010.06.005.
[7] Barbara A. Spears, Carmel Taddeo, Alan Barnes, 6 - Online social marketing approaches to inform cyber/bullying prevention and intervention: What have we learnt?, Academic Press,2018,Pages 75-94,ISBN 9780128114230,https://doi.org/10.1016/B978-0-12-811423-0.00006-7.
[8] Li, R., Lian, Q., Su, Q. et al. Trends and sex disparities in school bullying victimization among U.S. youth, 2011–2019. BMC Public Health 20, 1583, 2020. https://doi.org/10.1186/s12889-020-09677-3
[9] Phisit Pornpongtechavanich, Kawitsara Eumbunnapong, and Pallop Piriyasurawong, "Flipped Classroom with Challenge-Based Learning Model on an Online Streaming Ecosystem to Develop Coping Skills in Cyberbullying," International Journal of Information and Education Technology vol. 11, no. 11, pp. 523-531, 2021.
[10] Rajesh L & Dr. Mohan H S, An multilevel efficient energy clustering protocol with secure routing (MEECSR) in WSNs. International Journal of Applied Science and Engineering, volume-18, issue 6, 2021. https://doi.org/10.6703/IJASE.202106_18(2).008.
[11] Huang, H., Hong, J.S., & Espelage, D.L, Understanding factors associated with bullying and peer victimization in Chinese schools within ecological contexts. Journal of Child and Family Studies, 22, 2013, 881–892.
[12] Hinduja, S., & Patchin, J.W. Bullying beyond the schoolyard: Preventing and responding to cyberbullying. Thousand Oaks, CA: Corwin Press, 2008.
[13] L., Rajesh and Mohan H. S. Shuffled Shepherd Squirrel Optimization and Fractional LMS Model for In-Network Aggregation in Wireless Sensor Network. International Journal of Business Data Communications and Networking, vol.18, no.1 2022: pp.1-20. http://doi.org/10.4018/IJBDCN.309412
[14] Tangen, Donna & Campbell, Marilyn, Cyberbullying prevention: one primary school's approach. Australian Journal of Guidance and Counselling, 20(2),2010: pp. 225-23.
[15] Hokoda, A., Lu, H. -H.A., & Angeles, M. School bullying in Taiwanese adolescents. Journal of Emotional Abuse, 6(4),2006, 69–90.