Human Arenas
https://doi.org/10.1007/s42087-022-00312-8

ARENA OF TECHNOLOGIES

Artificial Intelligence Implementation to Counteract Cybercrimes Against Children in Pakistan

Saadia Anwar Pasha[1] · Sana Ali[2] · Riadh Jeljeli[3]

Received: 30 June 2022 / Revised: 14 September 2022 / Accepted: 20 September 2022

Abstract

Increased internet usage also enhances cyberattacks, particularly in a developing country like Pakistan. These cybercrimes are common against children, demanding the implementation of automated systems and models to detect and counteract these crimes. By keeping in view the growing importance of AI-enabled cybersecurity systems, this article provides insights regarding implementing the relevant systems and models to counteract cybercrimes against children in Pakistan. The researchers reviewed current studies witnessing the performance, reliability, and results of AI implementation in different countries to relate their possible potential further to detect, takedown, and trace online violence against children effectively. Findings showed that AI-enabled software, i.e., Spotlight, SomeBuddy, Google AI Tool, etc., and models such as DAPHNE, iCOP toolkit, and PrevBOT can identify and takedown any indecent activity against children. Besides, detecting the perpetrators' URLs, domains, and personal emails can further help the children resume internet usage in a healthy online environment. Thus, it is concluded that the internet technology is also creating vectors for abuse and exploitation against children. Harnessing powerful technology, i.e., AI, to analyze and manage the data can also enrich investigative functions. Towards this, local government and law enforcement agencies should resort to suggested tools that may identify even keywords and images. Further, the researchers have provided policy recommendations and discussed the limitations accordingly.

Keywords Cybercrimes · Children · Artificial intelligence · Internet · Online child sexual abuse · Cyberbullying · Extortion

✉ Sana Ali
sana_leo1990@hotmail.com

Saadia Anwar Pasha
saadia.pasha@aiou.edu.pk

Riadh Jeljeli
riadh.jeljeli@aau.ac.ae

[1] Department of Mass Communication, Allama Iqbal Open University, Islamabad, Pakistan

[2] Allama Iqbal Open University, Islamabad, Pakistan

[3] College of Communication and Media, Al Ain University, Al Ain, United Arab Emirates

Published online: 10 October 2022

Introduction

There is a wealth of opportunities for the users on the internet. However, today, when from financial assets to one's personal information is available on the internet, data preservation, privacy, and security are some of the major concerns, increasing the need for cybersecurity as comparatively more effective and strong (Bhatele et al., 2019). For instance, individuals saving their bank account details and residential information in their online cloud storage account are more vulnerable to cyber threats indicating an inevitable need to increase their cyber security. Cyber security creates and sustains the integrity, availability, and confidentiality of the users' data without compromising any private information. In this regard, the contemporary trends in internet usage and adoption raise several concerns regarding a sudden confrontation with cybercrimes. However, the role of artificial intelligence in counteracting these cyber crimes is prominent, providing several instant solutions to the law enforcement agencies (Newman, 2019). As noted by Coole et al. (2021), using artificial intelligence in the cyber system provides several benefits for operational security systems. For example, artificial intelligence helps increase the speed and probability of cybercrimes' detection, reduces operator fatigue, and assists security personnel for providing more attention where needed. Besides, Artificial Intelligence also helps reduce the cost at the management level, supports the decision-making process, provides effective interventions to counteract the internal threats, and directs resource allocation.

Similarly, an ever-increasing number of cybercrimes against children is a thought-provoking phenomenon. As noted by Kumar (2021), children are the most vulnerable component of our society that can be easily manipulated and exploited. Especially today, when the internet offers greater accessibility and ease of use, targeting a child has become common. Especially after 2019, the number of cyber violence against children has increased 400 times more, raising questions regarding the credibility of the internet as a safe place for children (Durkin, 2017). Consequently, increasing internet usage among children without knowing the potential threats is considerable. We cannot deny malicious activities on different online platforms that require one single click, leading to adverse outcomes for the users and children (Alhumaid et al., 2021).

Particularly, cybercrimes against children in developing countries are another major concern that requires serious consideration. For example, Federal Intelligence Agency (FIA) Pakistan reported daily 266 cybercrimes against children, mainly including child pornography and harassment, in the country. Still, most cybercrimes against children remained unreported due to children's lack of understanding about the incident and parents' neglect (Child Welfare Information Gateway, 2006; Global Human Rights Defence, 2021a). Yet, these increased cybercrimes against children can be counteracted by cyber security systems accompanied by artificial intelligence. Emerging technologies such as cloud computing, internet banking, and others can use artificial intelligence, which can also be applied to control cybercrimes against children (Chandra, 2019). Vilks (2019) further argued that AI-based cyber security not only identifies any criminal activity but also helps determine crimes against children. Countries, where cyber security is adopted and accompanied by Artificial Intelligence, are comparatively more vigilant to identify and counteract the relevant activities.

Thus, by keeping in view the number of cybercrimes against children in Pakistan, this article focuses on highlighting and addressing the online incidents against children and potential ways to avoid them. In the first section, the researchers have discussed the importance of Artificial Intelligence in dealing with cybercrimes in general and cybercrimes against children in particular. The second section discusses the most common cybercrimes against children worldwide and in Pakistan. The following section will also involve literature discussing how and to what extent Artificial Intelligence can help identify relevant crimes. The third section involves a general discussion and suggestions for implementing AI-based systems to identify and deal with the cybercrimes against Pakistani children. Finally, conclusions are made in the fourth chapter, and study limitations are discussed accordingly.

Cybercrimes Against Children

Online emotional abuse and cyberbullying happen when someone resorts to online technology to target, harass, threaten, or cause psychological harm. As today, smart devices are adopted and owned even by children, they often confront this harassment besides psychological trauma. Several cases of cyberbullying involve posting personal information and photos of a child, sharing other content that may harm the child and their reputation. Fake accounts pretending to be someone are also found to harm and target a child, causing emotional damage and bullying (Kidshealth.org, 2020). A report represented by Ipos Global Advisor conducted a study in 28 countries to investigate cyberbullying against children. Results revealed that the reports of cyberbullying have tremendously increased after 2011. Among the selected countries, 54% of South African children reported cyber-bullying against them. However, 65% of children from Latin America also reported cyberbullying, indicating increased cyber violence against them (Jarno & November, 2020).

However, despite online bullying and emotional abuse being common against children, sexual crimes are comparatively more prevalent (Katz, 2020). An investigative team by Europol led by the Te Tari Taiwhenua Department of Internal Affairs started interrogation against cyber-crimes against children in Croatia, Hungary, Austria, Czechia, Canada, Spain, USA, Greece, Australia, UK, and Slovenia. In 2019, the team found many users utilizing different online platforms to distribute and exchange images and videos containing explicit sexual and sadistic activities with children and infants. The team acquired 32 GB of files, leading to the international opening of approximately 836 cases (EUROPOL, 2022). Notably, the team encountered almost 90,000 international accounts, arrested 46 suspects in New Zealand, and more than 100 individuals were also arrested across the Europe. In two of the cases in Hungary and Austria, the suspects were molesting their own children, who were 6 and 8 years old, which were later safeguarded. Another case in Spain involved a suspect possessing child pornographic material and sharing sexual content with adults without their knowledge and consent (EUROPOL, 2022).

Similarly, a mixed-method study in Sri Lanka revealed that 28% of children had experienced some kind of online violence against them. More specifically, irrespective of gender, 27% of children reported receiving indecent images. However, the number of female children (29%) receiving the relevant images remained higher than that of male children (27%), while 27% of children also reported facing cyberbullying and extortion, and 20% reported that their indecent images were being shared on the internet (SavetheChildren.net, 2020) (Table 1).

Cybercrimes Against Children in Pakistan

Children are the most vulnerable sections of society and are easily exploited in the cyber world due to a lack of maturity level. It is observed that besides bullying and emotional abuse, sexual exploitation of children is prevailing at many online platforms. The offenders

Table 1 Reports concerning cybercrimes against children worldwide and in Pakistan

Source	Region	Type of online violence	Summary
Jalil (2018)	Pakistan	Pornography and sexual violence	Local authorities found a wing regulating online child pornography in Hussain Khanwala village, Kasur, Pakistan, for commercial purposes. Further reports revealed more than 285 children were sexually abused for pornography and silenced by their parents due to stigmatization. Police also recovered more than 400 pornographic videos of young boys engaging in on-camera sexual acts with the adults
SavetheChildren.net (2020)	Sri Lanka	Cyberbullying, extortion, harassment, revenge porn	Irrespective of gender, 27% of children reported receiving indecent images. However, the number of female children (29%) receiving the relevant images remained higher than that of male children (27%), while 27% of children also reported facing cyberbullying and extortion, and 20% reported that their indecent images were being shared on the internet
Jarno and November (2020)	Europe, America, Africa, Arab States	Cyberbullying, sexual crimes	The reports of cyber cullying have tremendously increased after 2011. Among the selected countries, 54% of South African children reported cyber-bullying against them. However, 65% of children from Latin America also reported cyberbullying, indicating increased cyber violence against them
AIN (2021)	Pakistan	Cyberbullying, child pornography, sextortion, harassment, revenge porn	In 2020, a total of 2960 cases were recorded, indicating a 4% increase compared to 2019 only in Punjab province. These cases involved physical and sexual violence, leading to even life-threatening situations for the victims. Of these cases, 51% of victims were females, and 49% were male children

Table 1 (continued)

Source	Region	Type of online violence	Summary
EUROPOL (2022)	Croatia, Hungary, Austria, Czech, Canada, Spain, USA, Greece, Australia, UK, and Slovenia	Sexual and sadistic content	The team found many users utilizing different online platforms to distribute and exchange images and videos containing explicit sexual and sadistic activities with children and infants. The team acquired 32 GB files, leading to the opening of approximately 836 cases internationally

chat online with young children by wrongly stating/representing their age and lure them towards sex. With the latest technology, it has become easy for criminals to contact children (Lewczuk et al., 2021).

Violence against children is a burning issue in Pakistan like other countries. For example, in 2020, a total of 2,960 cases were recorded, indicating a 4% increase compared to 2019 only in Punjab province. These cases involved physical and sexual violence, leading to even life-threatening situations for the victims. 51% of victims were females, and 49% were male children in these cases (AIN, 2021). However, these cases only indicate violence in the non-virtual environment, indicating a lack of research and consideration towards online violence against children in the country. Notably, today there are 61.34 million internet users in Pakistan. The number of internet users significantly increased from 2020 to 2021 by 11 million due to the COVID-19 outbreak. These internet users comprise almost 27% of the total population, indicating a significant number of youngsters below 16 years old. As a result, children often confront several cybercrimes, including cyberbullying, online violence, and harassment (DigitalPakistan, 2021). In 2019, Pakistan's Human Rights Minister, Dr. Shireen Mazari declared that today Pakistan is one of the most prominent countries to produce and disseminate child pornography. Other crimes such as identity theft, stalking, and online harassment also prevail against children in Pakistan (Global Human Rights Defence, 2021b).

Cybercrimes against Pakistani children are primarily seen in terms of child sexual abuse and pornography. During the past few years, young internet users in Pakistan have widely experienced cyber-bullying, sextortion, and revenge porn (Global Human Rights Defence, 2021). One of the prominent incidents of online violence against children in Pakistan can be traced back to 2015 when the local authorities found a wing regulating online child pornography for commercial purposes in Hussain Khanwala village, Kasur, Pakistan. Further reports revealed more than 285 children were sexually abused for pornography and silenced by their parents through fear and threats (Jalil, 2018). Police also recovered over 400 pornographic videos of young boys engaging in on-camera sexual acts with the adults (Zehra Abid, 2019).

In 2020, Federal Intelligence Agency (FIA) witnessed a potential increase in online harassment against children, including child pornography. According to the records, Federal Intelligence Agency (FIA) registered 260 complaints daily, reaching 94,500 complaints by the end of 2020. These crimes involve identity theft, online blackmailing, defamation, and child pornography (Global Human Rights Defence, 2021). According to Kasim Abbasi (2021), the accumulative ratio of cyber violence against children and in general indicates several types of crimes causing physical and psychological harm to the young users. Fake profiles, online blackmailing, defamation, sharing, receiving, and recording child pornography and other have enormously increased in Pakistan.

Artificial Intelligence in Crime Detection in General Context

Artificial intelligence is an important field of computer sciences that, e.g., bots, imitate human intellect. In other words, robots mimic human intellect in a digital environment as they contain smart algorithms making an evaluation based on the provided information. Artificial intelligence plays a significant role in almost every industry, including communication, cybersecurity, and forensics (Dupont et al., 2021). Today, when internet usage is increasing globally, users are more vulnerable to cyber-attacks. Cyber-criminals are always searching for new ways to harm people, steal their data, and put their lives at stake. As a result, it is impossible to overlook these cyber-crimes leading to strong cyber-security adoption among them (Thuy & Hieu, 2020). According to Caldwell et al. (2020), the internet provides us with various entertainment, information, communication, and educational opportunities; threats posed by cyber-crimes further endanger these new opportunities.

Consequently, cyber-security, including new approaches, i.e., Artificial Intelligence, has become an integral consideration for everyone. Here Thuy and Hieu (2020) further argued that cyber-criminals could contact and invade from any geographical region, identify personal information, defraud us financially, or harm our reputation. In such a situation, Rehnström (2021) suggested implementing Machine Learning and Artificial Intelligence as critical approaches to combat the cyber-crimes. For instance, Artificial Intelligence analyzed the trends in cybercrimes and prevention in the past. The importance of Artificial Intelligence can be determined by the fact that it can also work with the conventional security systems hand in hand (Rigano, 2019). Today, emerging security systems across the globe are attaining different ideas from several cyber events and utilizing them to identify threats, i.e., phishing and malware attacks. Podoletz (2022) noted that artificial intelligence automatically detects all the cyber threats or data breaches, which further alerts the systems to strengthen the cyber-security and counteract any potential invasion. As Rouhollahi (2021) argued, whatever form the Artificial Intelligence takes, technology will provide in-depth details of an event and ensure security against such incidents.

AI-Enabled Approaches to Counteract Cybercrimes Against Children in Pakistan

As system models have a potential to counteract online crimes against children, still, many ask how AI can help mitigate online violence against children. As a result, software developers and providers are providing influential AI applications that can thwart online crimes against children (Coole et al., 2021) in Pakistan. It is notable that, the intelligent agent system represents a small part of an entire AI system, known as computational intelligence (CI). Computational intelligence (CI) involves some strong yet different nature inspired capabilities, i.e., Fuzzy Logic, Swarm Intelligence, Artificial Immune System, Machine Learning, and Neural Networks (Dilek, 2017). These approaches facilitate the decision-making particularly, when the cybersecurity of internet users is required. When we say "Nature Inspired Immune System," we mean that the AI technology that can imitate the natural immune system, focused on counteracting the crimes as having abilities such as detection, memorization, process, and classify the information (Dilek, 2017). Thus, the following are some relevant AI-enabled systems that not only detect the cybercrimes against children but also provide a pathway to counteract them:

1. Child Safe AI

Child Safe AI is one of the pioneering AI-based platforms that monitor web content, particularly child abuse-based material, ensuring reduced child abuse in the online environment. The US law enforcement agencies also deploy Child Safe AI, which actively gathers signals regarding exploitative activities or material from the online environment. The relevant system also assists many organizations by continuously monitoring and evaluating the online material, i.e., images, videos, chats, and other content (Nolan & Brodowski, 2019).

2. Spotlight

Developed by Thorn, Spotlight is a digital identifier of online crimes against children, especially child trafficking and child sexual abuse. This technology uses predictive analytics to detect the relevant activities and victims. According to Oriel (2022), Spotlight identifies the activities and victims of child sexual abuse and online trafficking by data obtained from escort websites and sex advertisements. Like Child Safe AI, Spotlight is also used by the US Federal Department to detect child trafficking activities. It is also notable that, Spotlight has helped to detect and solve more than 14,874 cases of online child trafficking during the past 4 years.

3. AI technology by UNICRI

According to the United Nations Secretary-General Antonio Guterres, the combined efforts to counteract online crimes can protect children and ensure peace for all. AI technology by the United Nations Interregional Crime and Justice and Research Institute (UNICRI) uses Robotics and Artificial Intelligence for the law enforcement (particularly to identify and locate the long-missing children) (UNICRI, 2020a). Besides, it also helps to detect child and human trafficking sites and activities and identify illicit online pornographic material (UNICRI, 2020b). However, despite UNICRI is using Robotics and AI, still, the relevant technology is not much used (Oriel, 2022). According to UNICRI, employing Artificial Intelligence to ensure online safety for children also requires regular monitoring and updating the relevant system. Besides, incorporating AI technology is important due to the fact that, during the after the COVID-19 outbreak, internet usage among children increased further leading to an increased cases of online crimes against children. As a part of definitive emerging technologies today, artificial intelligence not only detects the crimes, but also monitor what a human eye, sometimes, cannot detect (UNICRI, 2021).

4. Google's AI Tool

Technology giant Google introduced an AI toolkit to counteract online crimes against children. This AI toolkit involved image processing through Deep Neural Networks that further helped the non-governmental organizations and investigators to detect the audio and video content based on child sexual abuse. The relevant AI toolkit also assists the classifiers in monitoring the offenders by detecting the content not identified by child sexual abuse material (Hunt et al., 2020). According to Detrick, the Artificial Intelligence-enabled tool by Google can help the reviewers to scan and identify more than 700% indecent material of children available online. Notably, Google's AI Tool is available free of cost; however, it needs human moderators to carefully evaluate the indecent images and other types of content by hand, needing human efficiency and effort as aiding intelligent systems for the crime detection (Detrick, 2018).

5. Safer

Thorn developed Safer as one of the leading AI companies, able to detect 99% of cybercrimes against children. With the help of Safer, an online platform can detect, remove, and block the sources and platforms used for cybercrimes against children. Safer had previously taken down more than 100,000 relevant files in its beta phase, and more improvements are yet to come (Gray et al., 2016). Safer provides the following services:

(a) Image Hash Matching creates perpetual and cryptographic hashes to match them with the previous child sexual abuse containing material, mainly to detect the indecent images of children.
(b) CSAM Image Classifier helps classify whether an image is consistent with the content that involves child sexual abuse.
(c) Video Hash Matching, like Image Hash Matching, also produces perpetual and cryptographic hashes to match them with the previous child sexual abuse containing material, mainly to detect the indecent images of children.

6. Griffeye

According to UK Home Office (2015), Griffeye uses different computer vision and recognition tools such as facial recognition to scan and detect the images based on the parameters of age and nudity. Griffeye is also adopted and implemented by US federal agencies to counteract any online activities against children (GRIFFEYE, 2018). According to the official Griffeye platform, online crimes against children can be traced through three resources: Analyze DI Pro software that can be used by individuals to import, process, and review the complex information regarding indecent images and videos of children available online. Analyze CS Operations involves group efforts facilitating crime detection. More specifically, teams using Analyze CS Operations are provided with digital base for analyzing and reviewing the relevant material. Finally, Analyze CS Enterprise facilitates the organizations in resuming ongoing interrogations regarding cybercrimes against children. The Analyze CS Enterprise is based on a dynamic repository to examine and identify the suspicious information over time (Digital Forensics, 2021).

7. SomeBuddy

Another important AI-enabled software, “SomeBuddy,” was created and designed by UNICEF to protect children using the internet for socialization, education, and entertainment. SomeBuddy as a “First-aid” platform helps and facilitates children report cyberbullying and harassment (UNICEF, 2019). It involves a strong combination of human supervision and Artificial Intelligence (AI) to detect and categorize the type of harassment, providing them with the most suitable recommendations concerning the current step. According to UNICEF (2019), the primary objective of SomeBuddy is to provide psychological and legal support to the children facing any challenging situation in cyberspace. Besides, it also spreads awareness among children to effectively identify and counteract any potential crime. According to CrimeDetector, initially, SomeBuddy adopted a virtual assistant interface; however, the developers found the chatbots as ineffective due to their broader conversational nature. Thus, SomeBuddy adopted simple yet automated approach to provide the users with

Table 2 AI-enabled approaches and models to counteract cybercrimes against children

Source	Approach/model	Approach/target	Summary
Cano et al. (2014)	DAPHNE	Analysis of bigram and trigram features, sentiment features, psycho-linguistic features, sentiment polarity features etc	Using a stage classifier involving a unigram of words approach and a tenfold cross-validation five-trial setting, in their work, Cano and their colleagues emphasized the characterization of the predators' online behavior, including sentiment polarity, syntactical content, discourse patterns, psycholinguistics, and a bag of words (BOW). Further, for the characterization process, the researchers provided a framework
Aiello and McFarland (2014)	iCOP	N-grams and specialized vocabulary detection	The P2P Engine helps to monitor the online traffic on the P2P networks. This P2P Engine supports traffic monitoring on Gnutella. However, other monitors can also plug in to increase their monitoring ability. This P2P Engine gathers information such as URLs, IP addresses, hash values, metadata, and filenames when the user is seen sharing the file
Gray et al. (2016)	Safer	Image Hash Matching, CSAM Image Classifier, Video Hash Matching	With the help of Safer, an online platform can detect, remove, and block the sources and platforms used for cybercrimes against children. Safer had previously taken down more than 100,000 relevant files in its beta phase, and more improvements are yet to come
GRIFFEYE (2018)	Griffeye	Facial recognition system (visual data review)	Griffeye uses different computer vision and recognition tools such as facial recognition to scan and detect the images based on the parameters of age and nudity. Griffeye is also adopted and implemented by US federal agencies to counteract online activities against children
UNICEF (2019)	SomeBuddy	Human supervision and artificial intelligence (AI) to detect and categorize the type of harassment	AI-enabled software "SomeBuddy" was created and designed by UNICEF to protect children using the internet for socialization, education, and entertainment. SomeBuddy, as a "First-aid" platform, helps and facilitates children to report cyberbullying and harassment
Nolan and Brodowski (2019)	Child Safe AI	Child sex trafficking ads detection on the commercial platforms	Child Safe AI is one of the pioneering AI-based platforms that monitor web content, particularly child abuse-based material, ensuring reduced child abuse in the online environment. The US law enforcement agencies also deploy Child Safe AI, which actively gathers signals regarding exploitative activities or material from the online environment

Table 2 (continued)

Source	Approach/model	Approach/target	Summary
UNICRI (2020b)	Robotics and artificial intelligence	AI and machine learning approaches for the prosecution if CSA material	AI technology by the United Nations Interregional Crime and Justice and Research Institute uses robotics and artificial intelligence to identify and locate the long-missing children. Besides, it also helps to detect child and human trafficking sites and activities and identify illicit online pornographic material
Hunt et al. (2020)	Google's AI Tool	Image processing through deep neural networks	AI toolkit involves image processing through deep neural networks that further helped the non-governmental organizations and reviewers to detect and sort out audio and video content based on child sexual abuse. The relevant AI tool also assists the classifiers in monitoring the offenders by detecting the content not identified by child sexual abuse material
Sunde and Sunde (2021)	PrevBOT	Problem Persons (PP) and Problematic Spaces (PS) classification	Besides, monitoring and interacting through chatrooms, PrevBOT contains the characteristics of machine learning and forensic linguistics influenced by AiBA technology. PrevBOT also enabled to predict and compute the details important for highlighting the preventive measures against online child sexual abuse in particular. Figure 3 illustrates the graphical model regarding PrevBOT to classify Problem Persons (PP) and Problematic Spaces (PS)
Oriel (2022)	Spotlight	Predictive analytics gathered from escort websites and sex advertisements	Spotlight is a digital identifier of online crimes against children, especially child trafficking and sexual abuse. This technology uses predictive analytics to detect the relevant activities and victims. Spotlight identifies the activities and victims of child sexual abuse and online trafficking by data obtained from escort websites and sex advertisements

more time to elaborate the obtained data. The relevant design output aided improving and streamlining the intake functions (CrimeDetector, 2019). Besides, SomeBuddy also stresses the legal experts across the globe to thoroughly review the obtained information, especially "false positives" and "false negatives" (CrimeDetector, 2019) (Table 2).

DAPHNE: Detecting Grooming in Online Social Media

Online crimes against children, especially sharing and possessing indecent images of children, have become common today. Besides, interacting with online child abusers and sextortion is also prevailing. Exposure to online criminals, especially pedophiles, is of greater concern, demanding strong consideration from the technology developers and online lawmaking authorities. In this regard, classified supervisors using independent feature types such as content, polarity, semantic frames, n-gram, and psycholinguistics can help counteract cybercrimes, particularly online child grooming (Peersman, 2017). Rigano (2019) proposed that the merged features can also help to counteract the online child grooming.

The study conducted by Cano et al. (2014) further proposed generating binary classifiers to detect the child grooming attempts on social media as the researchers used stage-labeled sentences (i.e., approach stages section, grooming, trust development) that were labeled as the positive set while the sentences labeled as "other" were the negative set. The researchers further suggested using a stage classifier with a unigram of words approach and a tenfold cross-validation five-trial setting. In their work, Cano and their colleagues emphasized upon the characterization of the predators' online behavior, including sentiment polarity, syntactical content, discourse patterns, psycholinguistics, and a bag of words (BOW). Further, for the characterization process, the researchers provided a framework (see Fig. 1). In the framework proposed by Cano et al. (2014), the first step involves extracting predators' communication patterns from the PJ chat log. In the second phase, the identifier pre-processes these conversation lines. Then, each conversation is represented in the feature extraction, following the information-gaining approach. Further, the researchers also designed a vector machine for the experiment as a supervised discriminative model.

iCOP: Identifying New CSAM on P2P Networks

The severity of online crimes against children demands a proliferation of monitoring and detection technology. A study by Aiello and McFarland (2014) proposed a toolkit, iCOP, to identify online child abuse material on P2P networks. However, before discussing and proposing the iCOP toolkit, the researchers proposed a classification approach that involves

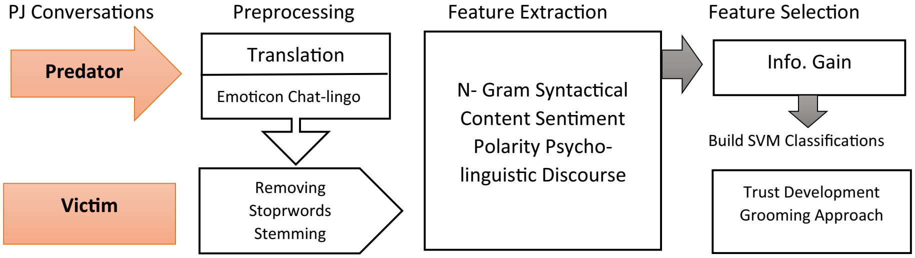

Fig. 1 Identification and characterization of the child grooming phases (Cano et al., 2014)

specialized vocabulary and n-grams to share child abuse material on P2P networks. Both specialized vocabulary and n-grams can automatically identify the child abuse material from the millions of files shared online. Further, a potential video and image classification approach using multiple and multimodal feature descriptions leads to a strong identification of child abuse material on online platforms.

Similarly, the above-discussed approaches are merged into an iCOP toolkit to counteract the child abuse-based online material further. As shown in Fig. 2, the proposed toolkit has two primary components, including the P2P Engine and iCOP Analysis Engine (Pupillo & Fantin, 2021). The P2P Engine helps to monitor the online traffic on the P2P networks. This P2P Engine supports traffic monitoring on Gnutella. However, other monitors can also plug in to increase their monitoring ability. This P2P Engine gathers information such as URLs, IP addresses, hash values, metadata, and filenames when the user is seen sharing the file. The latter is thus important to detect the originator and receiver of the child abuse-based content file.

The PrevBOT Concept

According to Kumar (2021), online crimes against children are proliferating today. To counteract these concerns, there is a need to establish an intelligence crime detection system. Further, Sunde and Sunde proposed PrevBOT concept commonly used by law enforcement authorities. PrevBOT was first conceptualized after the Sweetie 2.0 that can monitor open conversations and communicate automatically through chat rooms. Besides, monitoring and interacting through chatrooms, PrevBOT contains the characteristics of machine learning and forensic linguistics influenced by AiBA technology (Schermer et al., 2019). PrevBOT also enabled to predict and compute the details important for highlighting the preventive measures against online child sexual abuse in particular. Figure 3 illustrates the graphical model regarding PrevBOT to classify Problem Persons (PP) and Problematic Spaces (PS).

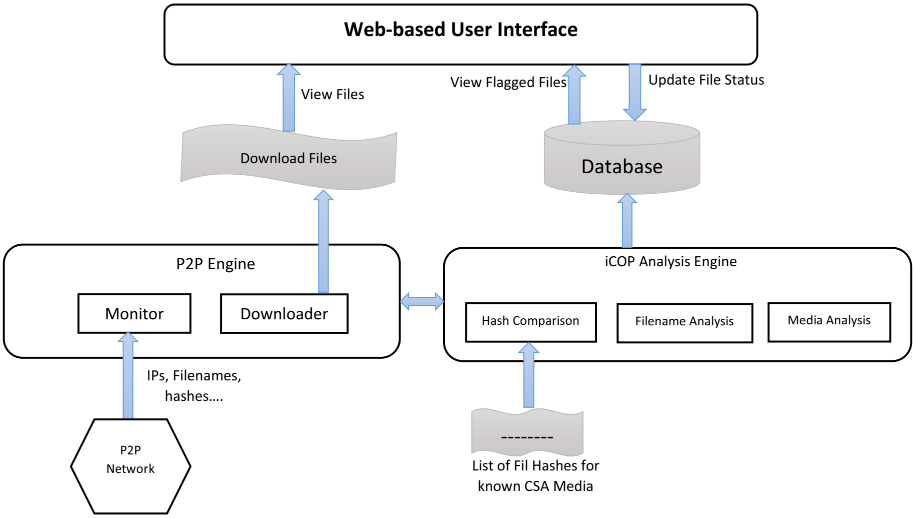

Fig. 2 iCOP toolkit (Aiello & McFarland, 2014)

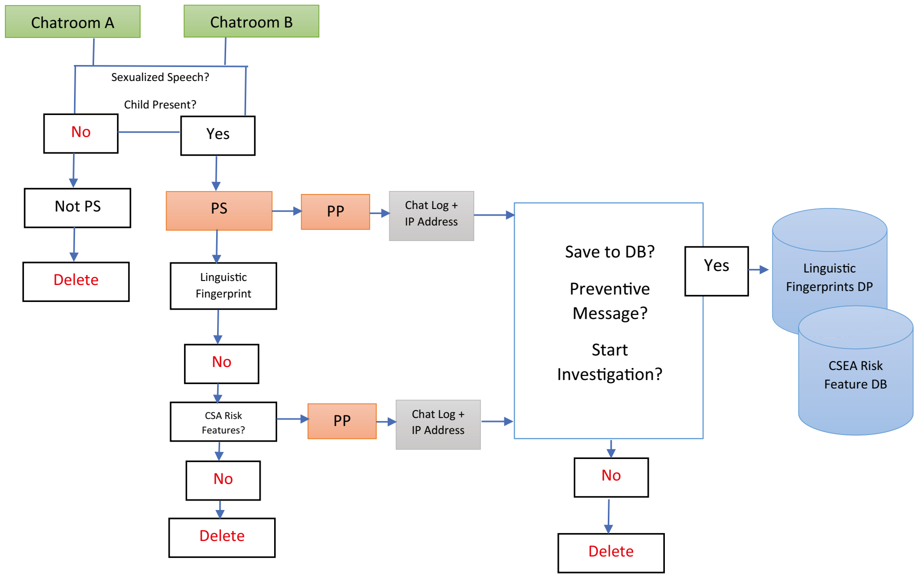

Fig. 3 Model regarding PrevBOT to classify Problem Persons (PP) and Problematic Spaces (PS)

According to Sunde and Sunde, PrevBOT is based on the Machine Learning algorithms and generates computations and predictions regarding the information consistent to online crimes against children. The output is statements, i.e., “carry a degree of uncertainty.” Notably, the PrevBOT has two primary objectives: to classify the Problematic Persons and to classify the Problematic Spaces (Sunde & Sunde, 2021). For example, when a chatroom is flagged as Problematic Space, PrevBOT predicts whether the behavior of individual(s) is suspicious (Problematic Persons (PS)). Further, PrevBOT uses a series of questions to classify and detect the problematic chat and behaviors through other different tactics. For example, at one stage, it uses “Linguistics Fingerprints” to compare pre-existing conversations with the new conversations. By using different techniques, PrevBOT facilitates reliable investigations for identifying the suspicious behaviors in chatrooms even if the persons are known to the police from previous cases of cybercrimes against children (Finch & Ryckman, 2020).

Discussion and Policy Recommendations

Fundamentally, the role and integration of information communication technology in our society limit access to counteract every single cyberattack. Talking particularly about the cybercrimes against children, the magnitude of these crimes is increasing in Pakistan. An algorithm-based system can help not only to counteract the crime but also aid in determining both previous and recent patterns in cybercrimes. The data used to modify and upgrade these systems is based on factual reports, defining the most considerable patterns of crimes especially against children (Global Information Security, 2019).

The empirical resources, available today, indicate that there are several approaches to identify and counteract cybercrimes against children. For instance, Neural Networks as an integral part of Intrusion Detection System can be used to detect the spam material and also to conduct the forensic investigations (EPCAT, 2021). According to Yadav and their colleagues, the ability of Artificial Intelligence to process huge amount of information cannot be denied. For example, a law enforcement team identifies suspects and has to go through the files saved into their computer systems. In such a situation, the investigation team will need human force to scrutinize all the content one by one that will be time-consuming, further threatening the investigators interest in the case. However, using Artificial Intelligence-enabled tools and approaches will not only help them to improve and speedup their investigations but also help to counteract the potential crimes in the future (Yadav & Husain, 2018). According to Pupillo and Fantin (2021), major platforms such as Facebook, YouTube, Tiktok, and others use Artificial Intelligence to identify and takedown the relevant activities. In many cases, tracing the criminals by recognizing their URLs also remained prominent as the relevant content was directly reported to the law enforcement agencies, leading to an instant detention and punishments.

Thus, to mitigate these cybercrimes, the Federal Intelligence Agency (FIA) Pakistan has created a special wing named “Cyber Crime Wing” (CCW) (Zia UL Islam et al., 2019). This Cyber Crime Wing was established under the laws that were designed under the Prevention of Electronic Crimes Act (PPECA) 2016, which criminalizes online violence against children with a punishment of a fine of up to five million rupees or 7 years of imprisonment or both. Victims can directly access the Cyber Crime Wing (CCW) through email or phone calls to lodge their complaints, ensuring an effective strategy to deal with the cyberattacks against children (AIN, 2021). However, these Cyber Crime Wing (CCW) and other bodies lack any prominent technological approach (Coole et al., 2021; Zia UL Islam et al., 2019) to automatically detect and takedown content and activities indicating a loophole in the cybersecurity systems. For this purpose, this article tends to provide some recommendations to the government, law enforcement authorities, and software providers in Pakistan, including:

1. The trend of using Artificial Intelligence to counteract cybercrimes against children is increasing, and Pakistan’s criminal justice authorities should also consider them (Velasco, 2022). The local government should grant sufficient funds and training to the law enforcement agencies, especially Cyber Crime Wing.
2. It is notable that, although cybercrimes against online users is a punishable crime in Pakistan, yet no practical policy is given to specify these crimes against children. In this regard, it is important to create a national Artificial Intelligence strategy to create a primary framework to implement it for both public and private sector law enforcement organizations with a potential focus on children welfare and their rights (Global Information Security, 2019).
3. The suggested softwares and models should be deployed and also identify suspicious activities through the suggested software models (See Gray et al., 2016; Podoletz, 2022; UNESCO, 2019). For this purpose, professionally trained individuals should be recruited.
4. Creating additional task forces to counteract cybercrimes against children may also coordinate with the Artificial Intelligence-enabled system. The relevant coordination may help identify the suspicious activities and content and help the task force to take instant action (Erokhina & Letuta, 2020).
5. Deploying models like PrevBOT, DAPHNE and iCOP to identify and counteract the stages of grooming in a cyber environment require skilled agents (Aiello & McFarland,

2014; Charalambous et al., 2016; Peersman, 2017). Software providers should recruit skilled and experienced individuals for deployment purposes so that they may also keep the systems updated.

6. Efficiently deploying network tools already used by UNICEF (Walker, 2019) and other concerned bodies can help investigators gather information about the host servers and URLs. Investigators with vigilance and prior experience in detecting and takedown any indecent activity can also help gather information about the user, website owner, and host country. If the concerned authorities use an AI-enabled system, improved forensic tools can further assist in identifying the email and URL information leading to easy detention of the cybercriminals (Charalambous et al., 2016).
7. The government of Pakistan can take an example of incorporating AI in counteracting online crimes against children by the Australian Police Department. Introducing “AiLecs” is trained to identify the suspicious activities, particularly, the content showing any indecent activity against children. It is notable that, the AI cannot be fully influential yet when the algorithms work as expected, cybercrimes against children can be counteracted (Panda Media Center, 2022).

Conclusion

The rise of cybercrimes against children in Pakistan further questions the credibility and use of online platforms among children. These crimes also threaten children’s life-long well-being as the relevant crimes are also evolving. Anonymity and accessibility are two prominent features that further halt a direct identification of the criminals and their origins. However, where technology has facilitated criminals, it also enables forensics and principal investigators to identify and takedown cyberattacks in the better possible way.

Deploying technology, updating existing cybersecurity systems, and using Artificial Intelligence (AI) besides human force can further help to counteract these crimes. As technology evolves, communication and interactivity are enhanced from conventional platforms to computers and now smart devices. Children are becoming increasingly vulnerable to sexual abuse, bullying, and emotional violence in online environments. Today, the rise of internet technology is also creating vectors for abuse and exploitation against children. Harnessing powerful technology, i.e., AI, to analyze and manage the data can also enrich investigative functions. Towards this, local government and law enforcement agencies should resort to suggested tools that may identify even keywords and images.

Limitations and Contributions

This study highlights an important area of interest which is formally underrepresented in the Pakistani scenario. Despite increasing cybercrimes against children in the country, deploying technology-assisted defense and identification systems is still under debate. The current study will not only provide an idea regarding AI’s significance in counteracting cybercrimes against children but also proposes potential technological systems and models with practical evidence from the other regions. However, certain limitations are considered in the current research. First, this study lacks any primary data that narrow down its scope. Second, the researchers have highlighted some selective softwares and models. Finally, the third limitation involves the lack of data sufficiently witnessing cyberattacks in the Pakistani scenario. Yet, this study is a heads-up call for the government, law enforcement agencies, and policymakers, further filling the gap in the existing literature.

Declarations

Conflict of Interest The authors declare no competing interests.

References

Aiello, L. M., & McFarland, D. (2014). Detecting child grooming behaviour patterns on social media. *Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics), 8851*, 16. https://doi.org/10.1007/978-3-319-13734-6

Alhumaid, K., Habes, M., & Salloum, S. A. (2021). Examining the factors influencing the mobile learning usage during COVID-19 pandemic: An integrated SEM-ANN method. *IEEE Access, 9*(July), 102567–102578. https://doi.org/10.1109/ACCESS.2021.3097753

AIN. (2021). *Violence against children continues unabated in Pakistan's Punjab province*. https://www.aninews.in/news/world/asia/violence-against-children-continues-unabated-in-pakistans-punjab-province20210621041438/#:~:text=Atotalnumberof2%2C960,49percentwereboys

Bhatele, K. R., Shrivastava, H., & Kumari, N. (2019). *The role of artificial intelligence in cyber security*. *January*, 170–192. https://doi.org/10.4018/978-1-5225-8241-0.ch009

Caldwell, M., Andrews, J. T. A., Tanay, T., & Griffin, L. D. (2020). AI-enabled future crime. *Crime Science, 9*(1), 1–13. https://doi.org/10.1186/s40163-020-00123-8

Cano, A. E., Fernandez, M., & Alani, H. (2014). Detecting child grooming behaviour patterns on social media. *Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics), 8851*, 412–427. https://doi.org/10.1007/978-3-319-13734-6_30

Chandra, M. (2019). Reduction of cyber crimes by effective use of artificial intelligence techniques. *International Journal of Recent Technology and Engineering (IJRTE), 4*, 8643–8645. https://doi.org/10.35940/ijrte.D8566.118419

Charalambous, E., Kavallieros, D., Brewster, B., Leventakis, G., Koutras, N., & Papalexandratos, G. (2016). Combatting cybercrime and sexual exploitation of children: An open source toolkit. *Open Source Intelligence Investigation*. https://doi.org/10.1007/978-3-319-47671-1

Child Welfare Information Gateway. (2006). *Child abuse and neglect what is child abuse and neglect?* https://tedibear.ecu.edu/wp-content/pv-uploads/sites/189/2019/07/Child-Abuse-and-Neglect.pdf

Coole, M., Evans, D., & Medbury, J. (2021). *Artificial intelligence in security: Opportunities and implications*.

CrimeDetector. (2019). *SomeBuddy*. 1–14.

Detrick, H. (2018). *Google's new AI tool to fight child sexual abuse will help reviewers scan 700% more material*. https://fortune.com/2018/09/04/google-child-sexual-abuse-prevention-ai-tool/

Digital Forensics. (2021). *Griffeye*. https://www.dataexpert.eu/products/digital-forensics-griffeye/

DigitalPakistan. (2021). *Digital 2021: Pakistan*. https://datareportal.com/reports/digital-2021-pakistan

Dilek, S. (2017). *Applications of Artificial Intelligence Techniques to Combating Cyber Crimes: A Review.*, *6*(1), 21–39.

Dupont, B., Stevens, Y., Westermann, H., & Joyce, M. (2021). Artificial intelligence in the context of crime and criminal justice. *SSRN Electronic Journal*. https://doi.org/10.2139/ssrn.3857367

Durkin, K. F. (2017). *Encyclopedia of cyber behavior*. *January 2012*. https://doi.org/10.4018/978-1-4666-0315-8.ch066

EPCAT. (2021). *The role of artificial intelligence in protecting children in the digital space*. https://ecpat.org/ai-digitalspace/

Erokhina, E. V, & Letuta, T. V. (2020). *Juvenile cybersecurity and artificial intelligence system*. *156*(Iscde), 607–611.

EUROPOL. (2022). *146 children worldwide saved in an operation targeting child abuse online*. https://www.europol.europa.eu/media-press/newsroom/news/146-children-worldwide-saved-in-operation-targeting-child-abuse-online

Finch, A., & Ryckman, L. (2020). *Strategies to prevent online sexual abuse of children: A systematic review of the literature protocol*. *March*, 1–7.

Global Human Rights Defence. (2021a). *Cybercrime towards children in Pakistan*. https://ghrd.org/cybercrime-towards-children-in-pakistan/

Global Human Rights Defence. (2021b). *Cybercrime towards children in Pakistan*. https://ghrd.org/cybercrime-towards-children-in-pakistan/#:~:text=Young people between the ages,from over 160 countries worldwide.

Global Information Security. (2019). *Global information society watch 2019 artificial intelligence:*

Gray, J., Pesevska, D. J., Sethi, D., Gonzalez, M. D. R., & Yon, Y. (2016). Handbook on developing national action plans to prevent child maltreatment. *World Health Organization*. http://www.euro.who.int/__data/assets/pdf_file/0019/329500/Child-maltreatment-PAP-handbook.pdf

GRIFFEYE. (2018). *Analyze DI Pro: Powerful software for your digital media investigations*. 30–31.

Hunt, X., Tomlinson, M., Sikander, S., Skeen, S., Marlow, M., du Toit, S., & Eisner, M. (2020). Artificial intelligence, big data, and mHealth: The frontiers of the prevention of violence against children. *Frontiers in Artificial Intelligence*, *3*(October). https://doi.org/10.3389/frai.2020.543305

Jalil, X. (2018). *Is something wrong with Kasur?* https://www.dawn.com/news/1384248

Jarno, K., & November, T. (2020). *Cyberbullying, an overlooked and ever growing danger to the development of children*. *November*, 1–23.

Kasim Abbasi. (2021). *Cybercrime increases by 83pc in three years*. https://www.thenews.com.pk/print/884453-cybercrime-increases-by-83pc-in-three-years

Katz, A. (2020). *Vulnerable children in a digital world*.

Kidshealth.org. (2020). *Cyberbullying*. https://kidshealth.org/en/parents/cyberbullying.html

Kumar, S. (2021). *Crime against children in cyber world*. *January*.

Lewczuk, K., Wójcik, A., & Gola, M. (2021). Increase in the prevalence of online pornography use: Objective data analysis from the period between 2004 and 2016 in Poland. *Archives of Sexual Behavior*. https://doi.org/10.1007/s10508-021-02090-w

Newman, J. C. (2019). *Toward AI security: Global aspirations for a more resilient future*. 1–89. https://cltc.berkeley.edu/wp-content/uploads/2019/02/CLTC_Cussins_Toward_AI_Security.pdf

Nolan, C., & Brodowski, M. L. (2019). *Emerging practices in the prevention of child abuse and neglect*.

Oriel, A. (2022). *Top 8 AI technologies to mitigate child abuse and child trafficking*. https://www.analyticsinsight.net/top-8-ai-technologies-mitigate-child-abuse-child-trafficking/

Panda Media Center. (2022). *Australian police develop a new AI tool to combat child abuse*. https://www.pandasecurity.com/en/mediacenter/technology/police-ai-abuse/

Peersman, C. (2017). *Using AI for supporting cybercrime investigations*.

Podoletz, L. (2022). We have to talk about emotional AI and crime. *AI and Society, 0123456789*. https://doi.org/10.1007/s00146-022-01435-w

Pupillo, L., & Fantin, S. (2021). *Artificial intelligence and cybersecurity*.

Rehnström, F. (2021). *How capable is artificial intelligence (AI) in crime prediction and prevention?* https://www.diva-portal.org/smash/get/diva2:1581808/FULLTEXT01.pdf

Rigano, C. (2019). Intelligence to address criminal. *National Institute of Justice, 3*(280), 1–10.

Rouhollahi, Z. (2021). *Towards artificial intelligence enabled financial crime detection*.

SavetheChildren.net. (2020). *Online violence against children in Sri Lanka: A national research on incidence, nature and scope contents*. https://srilanka.savethechildren.net/sites/srilanka.savethechildren.net/files/OnlineViolenceAgainstChildreninSriLanka.pdf

Schermer, B. W., Georgieva, I., Hof, S. V. D., & Koops, B.-J. (2019). *Legal aspects of Sweetie 2.0*. https://research.tilburguniversity.edu/en/publications/legal-aspects-of-sweetie-20-2

Sunde, N., & Sunde, I. M. (2021). *Article fagfellevurdert conceptualizing an AI-based police robot for preventing online child sexual exploitation and abuse: July*. https://doi.org/10.18261/issn.2703-7045-2021-02-01

Thuy, N. X., & Hieu, N. D. (2020). *Developing artificial intelligence in fighting, preventing and combating the digital business crimes*. *115*(Insyma), 447–450. https://doi.org/10.2991/aebmr.k.200127.090

UK Home Office. (2015). *The Child Abuse Image Database (CAID)*. *November*, 1–5. https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/478006/4817_CAID_Booklet_v13_online4.pdf

UNESCO. (2019). *Child online safety: Minimizing the risk of violence, abuse and exploitation online*.

UNICEF. (2019). *SomeBuddy*. 1–14.

UNICRI. (2020a). *Artificial intelligence and robotics for law enforcement*. https://unicri.it/artificial-intelligence-and-robotics-law-enforcement

UNICRI. (2020b). *The new age of technology*. https://unicri.it/topics/ai_robotics

UNICRI. (2021). *A high-level side-event to the 76th session of the United Nations general assembly*. https://unicri.it/index.php/News/Child-Sexual-Exploitation-Online-Side-event-UNGA76

Velasco, C. (2022). Cybercrime and artificial intelligence. An overview of the work of international organizations on criminal justice and the international applicable instruments. *ERA Forum*, *23*(1), 109–126. https://doi.org/10.1007/s12027-022-00702-z

Vilks, A. (2019). *Cybercrime and sexual exploitation of children in e-environment in the context of strengthening urban and rural security. 01010.*

Walker, S. (2019). *A guide to the UN cybercrime debate*. *March*.

Yadav, S., & Husain, M. S. (2018). *Application of artificial intelligence in fighting against cyber crimes: A review Available Online at* www.ijarcs.info *Application of artificial intelligence in fighting against cyber crimes: A review*. *April 2020*.

Zehra Abid. (2019). *In Pakistan's Kasur, child rapes and killings continue unabated | Child rights | Al Jazeera*. https://www.aljazeera.com/features/2019/10/28/in-pakistans-kasur-child-rapes-and-killings-continue-unabated

Zia UL Islam, Khan, M. A., & Zubair, M. (2019). Cybercrime and Pakistan. *Global Political Review*, *IV*(II), 12–19. https://doi.org/10.31703/gpr.2019(iv-ii).02

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.