National Forensic Sciences University

Knowledge | Wisdom | Fulfilment

An Institution of National Importance
(Ministry of Home Affairs, Government of India)

PROJECT REPORT

ON

"Forensic Analysis of Social Media Related to Terrorism and Gun Violence Using AI: An Analysis-Based Study"

Submitted To

School of Forensic Sciences

National Forensic Sciences University

For partial fulfilment for the award of degree

MASTER OF SCIENCE

In

FORENSIC SCIENCES

Submitted By

SHIRISH PANDITA

(022300100001025)

Under the Supervision of

Dr. G. Deepak Raj Rao

School of Forensic Science

National Forensic Sciences University,

Delhi Campus, New Delhi – 110085, India

May 2025

DECLARATION

I hereby declare that the thesis entitled "Forensic Analysis of Social Media Related to Terrorism and Gun Violence Using AI: An Analysis-Based Study" is a bona fide research work done by me, and no part of the thesis has been presented earlier for any degree, diploma or similar title at any other university

SHIRISH PANDITA

022300100001025

M.Sc. Forensic Science

2025

Date:

Place:

ORIGNALITY REPORT CERTIFICATE

I certify that

a. The work contained in the dissertation is original and has been done by myself under the supervision of my supervisor.

b. The work has not been submitted to any other Institute for any degree or diploma.

c. I have conformed to the norms and guidelines given in the Ethical Code of Conduct of the Institute.

d. Whenever I have used materials (data, theoretical analysis, and text) from other sources, I have given due credit to them by citing them in the text of the dissertation and giving their details in the references.

e. Whenever I have quoted written materials from other sources and due credit is given to the sources by citing them.

f. From the plagiarism test, it is found that the similarity index of whole dissertation within 10% and single paper is less than 10 % as per the university guidelines.

Shirish Pandita
Enrollment No.: 022300100001025

Date:
Place:

Forwarded by

Dr G. Deepak Raj Rao

Date_____

CERTIFICATE

This is to certify that the work contained in the dissertation entitled **"Forensic Analysis of Social Media Related to Terrorism and Gun Violence Using AI: An Analysis-Based Study"**, submitted by **SHIRISH PANDITA (Enrollment No.:022300100001025)** for the award of the degree of **Master of Science in Forensic Science (Cyber Specialisation)** to the **National Forensic Sciences University**, **Delhi Campus**, is a record of bona fide research works carried out by him under my supervision and guidance.

Dr. G. Deepak Raj Rao

Professor

School of Forensic Science

National Forensic Sciences University

Delhi Campus, Delhi, India

Date:

Place:

Forensic Analysis of Social Media Using AI:

An Analysis-Based Study

ACKNOWLEDGEMENT

I would like to express my sincere gratitude to Dr. G. Deepak Raj Rao, my supervisor, for his valuable guidance, constructive feedback, and continuous support throughout the course of this dissertation. His expertise and encouragement played a pivotal role in shaping the direction and depth of my research.

I am also thankful to the National Forensic Sciences University, Delhi Campus, for providing an enriching academic environment and the necessary resources that enabled me to carry out this research effectively.

I extend my heartfelt thanks to my mentors at the Forensic Science Laboratory (FSL), Jammu-Dr. Kumar Sourabh (Deputy Director-FSL Jammu), Mr. Kaiser Ali Bhat (ASO-Cyber), and Mr. Abul Roaf (ASO-Cyber) for their insightful discussions, practical knowledge, and kind support throughout the research process. Their mentorship has significantly contributed to my understanding of forensic applications in the field of cyber investigations.

I would also like to sincerely thank Ms. Leena Kamran for her invaluable assistance in the development of the NLP model and the AI-driven analysis of the data. Her contribution played a key role in strengthening the technical foundation and analytical rigor of this research.

With Sincere Regards,

SHIRISH PANDITA

022300100001025

M.Sc. Forensic Science (Cyber Specialization)

2025

ABSTRACT

With the increasing role of social media in the radicalization of individuals involved in terrorism and gun violence, forensic analysis of digital communication has become a vital field of inquiry. This study investigates the use of Artificial Intelligence (AI), particularly Natural Language Processing (NLP) and sentiment analysis, to trace the ideological and emotional trajectories of individuals prior to committing acts of violence. By collecting and analyzing publicly available social media posts from perpetrators of mass shootings and terrorist attacks, the research constructs detailed behavioral timelines, identifies linguistic markers of radicalization, and examines the progression from online expression to offline action. The methodology integrates AI-driven analysis with human validation to enhance accuracy and contextual understanding. Through comparative case studies, the study aims to establish patterns in digital behavior that can be used to flag potential threats. This research contributes to the development of predictive forensic tools and highlights the critical role of language-based AI in preempting ideologically motivated violence on social media platforms.

TABLE OF CONTENT

LIST OF FIGURES

LIST OF TABLES

GLOSSARY

No.	Keyword	Definition
1.	**Radicalisation**:	The process through which individuals adopt extreme political, social, or religious ideologies.
2.	**Twitter**:	A social media platform used for microblogging and sharing brief text-based posts.
3.	**Reddit**:	A forum-based platform where users share content and participate in discussions.
4.	**4chan**:	An anonymous image board used for posting and discussing controversial or fringe topics.
5.	**Tokenisation**:	The process of breaking text into individual elements, such as words or phrases.
6.	**Normalisation**:	Text preprocessing that converts data into a standard format (e.g., lowercasing).
7.	**Lemmatisation**:	Reducing a word to its base or dictionary form (lemma).
8.	**Sentiment**:	The emotional tone conveyed in a piece of text, such as positive, negative, or neutral.
9.	**Fear**:	An emotion reflecting perceived danger or threat.
10.	**Anger**:	An emotional response to perceived provocation or injustice.
11.	**Sadness**:	A feeling of sorrow or unhappiness.
12.	**Disgust**:	A feeling of revulsion or strong disapproval.
13.	**Joy**:	A positive emotion of happiness or pleasure.
14.	**Trust**:	Confidence in the reliability or truth of something or someone.
15.	**Anticipation**:	Expectation or prediction of a future event.
16.	**Surprise**:	A reaction to an unexpected event or outcome.
17.	**Ideological Hate**:	Hostility rooted in extreme ideological beliefs.
18.	**Personal Revenge**:	Retaliatory behaviour based on personal grievances.
19.	**Religious Extremism**:	Radical religious ideologies that justify violence.
20.	**Psychological Collapse**:	A breakdown of mental stability, often leading to irrational behaviour.
21.	**Obsessive Isolation**:	Withdrawal from society due to intense internal focus or paranoia.
22.	**Anomalies**:	Deviations in data patterns that may signal threats.
23.	**Algorithmic**:	Pertaining to a step-by-step procedure or set of rules used in data processing.
24.	**Sextortion**:	Blackmail involving sexual images or acts.

25.	**Extortion**:	The practice of obtaining something through force or threats.
26.	**Defamation**:.	Communication of a false statement that harms a person's reputation
27.	**Intimidation**:	Using fear or threats to influence others.
28.	**Bullying**:	Repeated aggressive behaviour intended to hurt or intimidate.
29.	**Terrorism**:	The unlawful use of violence and intimidation, especially against civilians, for political aims.

ABBREVATIONS

Acronym	Meaning
AI	Artificial Intelligence
NLP	Natural Language Processing
i.e.	That is
FBI	Federal Bureau of Investigation
DARPA	Defence Advanced Research Projects Agency
NSA	National Security Agency
CISA	Cybersecurity and Infrastructure Security Agency
HTML	HyperText Markup Language
CSV	Comma-Separated Values
ML	Machine Learning
API	Application Programming Interface
GUI	Graphical User Interface
PDF	Portable Document Formal

Introduction

Over the last two decades, the rapid growth of social media platforms has radically transformed the way individuals communicate with each other, share data, and interact on a worldwide scale. Through these platforms, over a billion users generate vast numbers of data daily. These platforms have emerged not only as hotspots for legitimate communication but also as a fertile ground for criminals for illicit activities. These digital interactions predominantly leave behind a rich trail of evidence; hence, the forensic analysis of social media has become a critical component in modern investigations, majorly those related to cybercrime, terrorism, fraud, and the spread of misinformation. As a response to the complications constituted by the complexity of social media, artificial intelligence has emerged as a metamorphic compulsion in the field of digital and cyber forensics. Advanced AI techniques, bridging machine learning Natural Language Processing (NLP), allow forensic experts to deskill the collection, processing, and analysis of vast numbers of data sets. AI as a tool not only paces the accuracy and progression of the investigation but also allows the extraction of detailed patterns and hidden links that might perhaps skip the human eye.

The integration of AI into social media applications/platforms provides investigators with a wide and real-time insight into digital communications. AI-moderated models can monitor live social media streams, flag potential threats, and even predict emerging criminal trends through predictive analytical data sets. Such advancements shift the investigative approach from purely reactive to a control-based strategy mode, allowing investigative authorities to identify and mitigate the risk before it escalates into a significant security risk. In a time when exponential technological advancements are constantly changing the face of digital communication, this change is essential. Despite these promising advancements, a number of issues complicate the seamless application of AI in social media forensics. As the internet data is heterogeneous and unstructured, vigorous methods for data integration and noise reduction are needed to ensure the accurate extraction of important insights. Moral dilemmas, evolving data protection regulations, and privacy concerns also provide significant challenges. In order to preserve both individual privacy and the integrity of digital evidence, AI systems must be developed to operate under strict legal guidelines. Furthermore, issues like algorithmic bias and the dynamic nature of online behaviour call for constant evaluation and refinement in order to ensure the effectiveness and equity of AI systems. The aim of this work is to provide a microscopic explanation of how AI and social media forensic tools interact. It will examine current methods, discuss the challenges and limitations they provide, and offer potential solutions to increase the reliability and ethical application of AI in this critical field. In order to help create rigorous, effective, and morally driven forensic procedures that can be adjusted to the quickly changing digital world, this initiative aims to integrate viewpoints from a variety of academic disciplines, including computer science, criminology, and legal studies. As society grows more interconnected, artificial intelligence plays a crucial role in digital and cyber forensics to guarantee the safety, reliability, and integrity of our online spaces.

Building on these ideas, it becomes certain that the integration of artificial intelligence for forensic analysis of social media pushes for a re-examination of both technical capabilities and ethical frameworks. One of the biggest challenges is striking a balance between equalising AI's powerful analytical capabilities and safeguarding individual privacy. With an AI system capable of

transitioning through large database sets to detect patterns and anomalies, there is a subsidiary risk of overreach, where innocent communication might be misinterpreted or someone's individuality is inadvertently targeted. Thus, establishing a linear oversight and clear regulatory guidelines is crucial in preventing the exploitation and ensuring that the investigative procedure remains just and transparent. Moreover, the application of artificial intelligence in the field of forensic sciences demands an interdisciplinary approach, intertwining insights from computer science, legal studies, and ethics. Forensic analysts must work closely with legal professionals to ensure that AI-driven evidence is technically sound along with having admissibility and relevance within the judiciary. This collaboration is necessary to bridge the gap between the rapidly evolving technologies and traditional legal approaches, ensuring that the civil and fundamental rights of an individual are upheld while enhancing public safety. In the modern age, artificial intelligence has served as a utility for progress as well as a tool for committing crime. The integration of AI into everyday life has changed how we interact with everything in our daily lives. On a brighter note, AI advancement has not only mitigated the daily hassle but also provided us with practical solutions for our problems. Along with these advancements, criminals ensured that they could also catch up with the new trends and use them for their benefit.

Alone in the year 2024, 17470 cases were registered under fraud for cybercrimes, comprising a total number of 1665 cases of credit/debit card fraud, 1690 ATM frauds, 6491 online banking frauds, and 2910 OTP frauds, and 4714 were categorised under “Other” for cybercrimes. (Press Information Bureau, 2023)

The rapid expansion of the AI domain has caused an eruption of modern crimes like deepfakes. This is not just limited to an individual but affects society, as manipulation of video-graphic or photographic data can and has led to the disruption of communal harmony. These video-graphic messages, which include speech from political leaders or people who hold an influential value in society, are altered using AI to spread a specific narrative, mostly hate speech or defamation of individuals. Alongside these video-graphic alterations, the most common criminal use of deepfakes is the generation of explicit images of victims to blackmail them for money or favours. These deepfakes are generated by AI-assisted face-swapping techniques as the face of the victim is swapped over an already explicit image. These images are then sold on various websites and marketed as original images, thus putting the social and personal life of the victim at stake. Major initiatives have been taken by different governments all around the world to address this issue to safeguard the rights of the individuals.

In the year 2021, under “The Information Technology (Intermediary Guidelines and Digital Media Ethics Code) Rules, 2021,” the Government of India issued guidelines and rules to safeguard the fundamental rights of an individual. These modifications ensured the removal of unauthorised publications of deepfakes in order to defame the individual or cause communal violence. They also ensured that artificially generated explicit images of individuals are flagged the same as unaltered explicit images of an individual to safeguard the right to privacy of the person, and prosecution of the convict is done in an orderly and fair manner. (Press Information Bureau, 2024)

In the year 2023, the Government of India issued an advisory to social media intermediaries to identify misinformation and deepfakes and remove any such content when reported within 36

hours of reporting. This ensured expeditious actions, well within the timeframe stipulated under IT Rules 2021, and disabled access to the content/information. (Collins and Ebrahimi, 2021)

Another emerging issue is the integration of deepfakes in live video conferencing, which alters real-time imagery; hence, it is more advanced than pre-moderated deepfake videos and images. This technique is orchestrated to spread real-time misinformation and cause chaos on a larger scale. Moreover, individuals from all classes of society are victims of this, especially women, as they are seen as a vulnerable target for sextortion.

Table 1-Overview of different type of risk associated with Cyber Crimes (National Security Agency et al., 2023)

Psychological harm	Financial harm	Societal harm
• **(S)extortion**	• Extortion	• News media manipulation
• **Defamation**	• Identity theft	• Damage to economic stability
• **Intimidation**	• Fraud (e.g. insurance/payment)	• Damage to the justice system
• **Bullying**	• Stock-price manipulation	• Damage to the scientific system
• **Undermining trust**	• Brand damage	• Erosion of trust
	• Reputational damage	• Damage to democracy

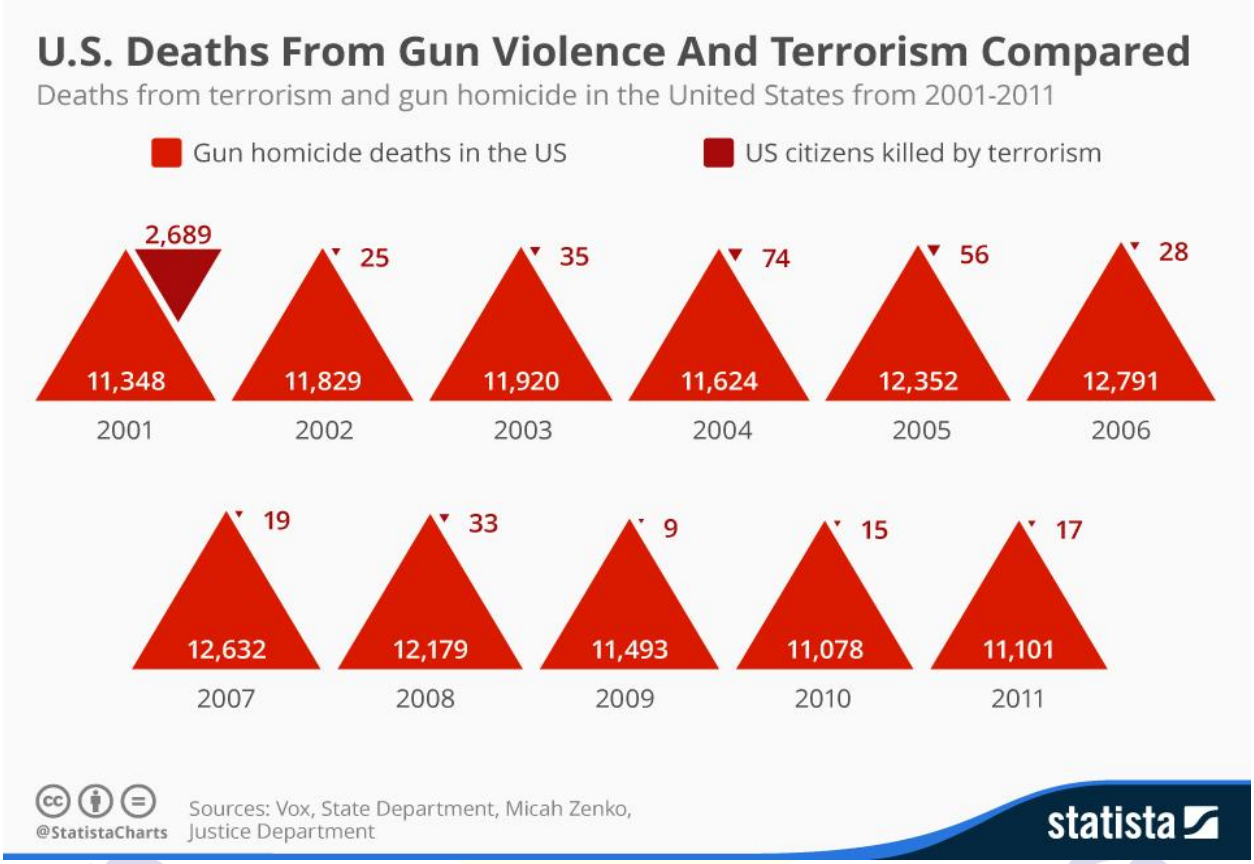
The above table provides a difference between 3 key potential impacts—reputational, financial, and manipulations of decision-making at three different levels: individual, organisational, and societal. In this table, the types of risks and levels of impact are presented separately.

In recent investor increasing digitalisation of personal expression has fundamentally transformed how individuals communicate their thoughts, ideologies and emotional studies. Wireless digital shift has opened up new avenues for connectivity and discourse it has also provided a fertile ground for the expression and evolution of extremist ideologies and violent intentions. The rise of lone wolf attacks on many ideologically motivated mass shootings and hate driven violence has drawn attention to the role that digital platform particularly social media play in both facilitating and concealing behavioural progression towards violence. This development has led to a growing interest within forensic science and behavioural analysis in the use of Artificial Intelligence (AI) for interpreting the digital trace left by such individuals. In particular Natural Language Processing (NLP) is a subfield of ai which focuses on computational interpretation of human language and has proven to be useful in detecting patterns in text that may reveal psychological state emotional instability or ideological fixations. Sentiment analysis emotional tagging and linguistic profiling once confined to marketing and customer feedback are now being applied to the domain of digital forensics and threat assessment.

This dissertation investigates the application of AI specifically NLP techniques to the social media post manifestos and online communication of known violent offenders by the analysis of emotional tone sentiment score and behavioural evolution across different time phrases such as early mid and final this study aims to identify reoccurring emotional trajectories that may serve as digital red

flags. The research also seeks to categorise offenders based on their emotional and psychological behaviour types such as ideological hatred personal revenge religious extremism or psychological collapse derived from their emotional data. The approach used in the study contains automated sentiment and emotional scoring with manual behavioural phase classification, resulting in a hybrid analytical mode this model allows for both large scale pattern recognition and contextual human-based judgment. It further addresses one of the core questions in modern forensic behaviour analysis which is "*Can emotional and linguistic patterns in digital content serve as early indicators of violent behaviour?*"

Figure1: Comparative Chart of U.S. Deaths from Gun Violence and Terrorism (Statista, n.d.)



Description: This chart compares the number of U.S. deaths resulting from gun violence to those from terrorism between 2001 and 2011, highlighting the relative impact of each.

By focusing on emotional aggression rather than ideological content alone this research expands the scope of digital science to induce the effective dimension of online radicalisation and violence. It offers a novel contribution to the field of cyber forensics, chronology and psychological profiling while proposing practical applications for law enforcement agencies, digital threat monitoring systems and early investigation strategies.

1.1 Forensic Applications

In this modern age, these social media platforms have turned a simple networking site into a vast repository of user-generated data, shaping public discourse, influencing political opinions, and even facilitating criminal activities. As the volume of online data increases, so do the challenges faced by forensic investigation agencies who navigate This data is filled with misinformation, defects, fraud, and cyberbullying. The traditional investment negative techniques are becoming less reliable for progress and ongoing investigation and analysing the immense amount of data generated on a daily basis. To address these challenges, artificial intelligence has emerged as a revolutionary tool in forensic investigation by offering enhanced capabilities in data extraction, threat detection, behavioural analysis, and predictive analytics.

AI-assisted forensic mythologies have not only streamlined the evidence collection process but have also empowered law enforcement agencies to identify patterns, predict criminal behaviour, and respond to the emerging cyber threats in real time, and this comprehensive analysis examines the forensic application of AI during a social media investigation. It is deep-rooted in the technological advancements that have provided automated evidence collection, discussing methods for detecting complex digital manipulations like defects and evaluation of ethical challenges posed by artificial intelligence. Moreover, the paper explores how predictive analytics and behavioural profiling contribute to and help in monitoring terrorism. Throughout this discussion, the integration of artificial intelligence with traditional forensics It is highlighted as a necessity in the fight against cybercrime.

The corresponding sections provide an in-depth analysis of each major application area supported by recent research findings and case studies. By minimising the gap between AI technologies and forensic science, the document offers an insight into how these advancements can be cultivated responsibly to protect the privacy of an individual and uphold justice in the digital era. In this modern age, these social media platforms have turned a simple networking site into a vast repository of user-generated data, shaping public discourse, influencing political opinions, and even facilitating criminal activities. As the volume of online data increases, so do the challenges faced by forensic investigation agencies who navigate This data is filled with misinformation, defects, fraud, and cyberbullying. The traditional investment negative techniques are becoming less reliable for progress and ongoing investigation and analysing the immense amount of data generated on a daily basis. To address these challenges, artificial intelligence has emerged as a revolutionising tool in forensic investigation by offering enhanced capabilities in data extraction, threat detection, behavioural analysis, and predictive analytics.

.

1.1.1 AI-Assisted Social Media Evidence Collection

(i) Automated Data Extraction

In the modern forensic investigation 1 of the most significant challenge is the vast volume of data available on social media platforms. The revolutionization of evidence collection has been done through AI power tools by automating the extraction of necessary and relevant information from these platforms. Sophisticated web scraping techniques and Natural Language Processing (NLP) algorithms are used by these tools to gather data such as posts, comments, images, videos and metadata without human interaction. Geolocation tags embedded within the photographs or timestamp information from post can be extracted from these automated systems which are crucial in establishing a suspect's timeline or verifying alibis (Casey, 2011)

Large data sets to isolate content that Mute specific criteria set by the investigator can be filtered by these tools. AI can quickly identify clusters of related posts and map out the interactions between different users across the platform such as those involving coordinated disinformation campaigns and pushed narratives. These techniques not only speed up the process but also ensures that any critical piece of information is not overlooked. Automated data extraction remains an indispensable element of forensic science as these digital landscapes evolve.

(ii) Sentiment Analysis and Contextual Interpretation

The textual content posted on social media is analysed by a branch of NLP to interpret the emotional tone behind the text, and thus this analysis is known as sentiment analysis. The detection of aggressive language, hate speech, or signs of radicalisation is done by the use of sentiment analysis by forensic analysts. A vast amount of textual data is processed, and AI can flag content that deviates from normal language patterns, which signals potential threats or criminal intent (Zellers et al., 2019). Contextual cues such as the frequency of certain keywords or the occurrence of slang and idiomatic expressions help build a comprehensive profile of an individual's online behaviour, which is analysed by these AI tools. These interpretations are invaluable in cases where subtitles may indicate crimes regarding grooming, cyberbullying, or promotion of violence. Search analytical techniques help us to differentiate between personal opinions and statements that could promote criminal activities.

(iii) Network Mapping and Relationship Analysis

Social media networks are inherently based on relations, meaning they contain an abundant amount of information about how individuals interact with each other on these platforms. Owned network mapping tools utilise graph theory and machine learning to understand these interactions, and by analysing these relationships between different users across the platform, these tools can identify key influencers, detect clusters of similar behaviour based on coordination, and reveal hidden links between unrelated accounts (Ferrara et al., 2016).

Network mapping can expose how radical groups operate online by highlighting nodes of high connectivity and influence in investigations that concern organised cybercrimes or terrorism.

1.1.2 AI-Driven Threat Detection and Cybercrime Prevention

(i) Deepfake Detection and Media Forensics

The threat of deepfake technology represents one of the most challenging sections in contemporary forensic science. Deepfakes are artificially generated videos or images that convincingly alter a person's appearance or speech, posing a significant risk to public trust and security. Criminals can use deepfakes to create false evidence, manipulate public opinion, or promote blackmailing (Chesney & Citron, 2019).

To counter these challenges, researchers and various agencies have developed artificially driven deepfake detection tools that analyse these digital media contents for cues of manipulation. The examination of these deep fakes is very simple and is based on inconsistencies in pixelated data, anomalies in facial moments, and the presence of inconsistent frequency in audio signals. Projects like DARPA's Media Forensics and Semantic Forensics have made progressive studies in this area. By developing algorithms that can accurately detect deepfakes and provide forensic expertise with reliable evidence (DARPA, 2020). Ongoing research and algorithm refinement are essential to stay ahead in the race of the evolving deepfake technology to counter malicious actors.

(ii) Identification of Fake Accounts and Bot Networks

A ripe environment has been created by the presence of fake accounts and automated bots on social media, which has contributed to the nourishment of cybercrime and disinformation on the Internet. These accounts are often employed to amplify false narratives, manipulate public sentiments, and even facilitate frauds. These accounts can be analyzed by a highly effective AI system in detecting patterns such as posting frequency, account creation dates, and linguistic characteristics of their content (Ferrar et al., 2016).

With a simple training method on large database bot behavior and human automated activities can be distinguished by these machine learning models. Once identified, these fake accounts and automated bots on social media can be removed or flagged in a systematic way, therefore curbing the spread of misinformation. The protection of the overall integrity of social media platforms and the protection of individual users are crucial capabilities that these machine learning models possess.

(iii) Anomaly detection in cyber-criminal behavior

In an ai driven forensic analysis and normally detection is a critical component which enables the identification of deviations from a typical online behavior of an account on these platforms. Establishing ground rules for normal activity such as typical login times, content engagement patterns and interaction frequencies these ai systems can flag anomalies that may indicate criminal activities on these platforms. For instance, a sudden abruption of activity from an account or unexpected changes and posting behaviors can sometimes signal a compromised account or a coordinated cyber-attack (Kaplan & Haenlein,2019)

A continuous monitoring of digital environment is achieved by these systems through anchoring models and unsupervised learning techniques. After the detection of anomalies alerts are generated and are further investigated by human analyst. For promptly addressing fraudulent and malicious behavior these proactive approaches are essential for preventing cybercrime before it escalates.

1.1.4 Behavioral Analysis and Predictive Analytics

(i) Predictive Analytics in Fraud Prevention

Predictive analytics employs historical data and machine learning algorithms to forecast future events. Fraudulent activities, which encompass online interaction transaction histories and payment modes, can be analyzed by AI-driven models. In the realm of financial fraud on social media, AI-driven predictive analytics models assess patterns in transaction histories, online interactions, and user behavior to identify potentially fraudulent activities. By comparing current behavior against established patterns, these models can provide early warnings of credit card fraud, phishing attempts, or identity theft (Ferguson, 2020).

New data continuously refine these systems and allow them to adapt to emerging fraud schemes. This ultimately results in financial institutions and law enforcement agencies getting equipped in a better way to mitigate loss and respond to these threats swiftly. It not only enhances accuracy, but the integration of predictive analytics in forensic investigation also contributes to a more secure online financial ecosystem.

(ii) Monitoring extremism and terrorism through social media

With the growth of social media, the threat posed by online extremism has increased significantly. These days, AI technologies are essential for keeping an eye on extremist content and stopping acts of terrorism. Forensic specialists can identify early indicators of radicalization in online forums by utilizing Natural Language Processing (NLP) and sentiment analysis. Potential terrorist networks can be identified by using AI algorithms that search for hashtags, phrases, and picture content that support extremist ideology (Zellers et al., 2019).

The analysis of user interaction has evolved from content analysis to behavioral tracking. Law enforcement agencies Use these dynamic monitoring tools to identify individuals transitioning from online radicalization to offline activities. Timely interventions are achieved by the proactive nature of these AI applications and help in the prevention of acts of terrorism before they occur in real time.

(iii) Psychological Profiling and Criminal Intent Detection

Forensic psychology has traditionally played an important role in understanding criminal behavior, and AI is increasingly supplementing these efforts with data-driven insights. Advanced artificial intelligence systems analyze linguistic patterns, social media activity, and interaction histories to create detailed psychological profiles of individuals. These profiles assist investigators to assess the likelihood of criminal intent and identify behavioral characteristics associated with violent or high-risk behavior (Casey, 2011)

Languages and human sentiment that might indicate a change in individual state of mind can be detected through psychological profiling using these ai tools for example the increase in aggressive language or the frequent use of certain terminologies that might trigger conspiracies and may ultimately be seen as warning signs. Law enforcement can develop more effective strategies by integrating these insights and other forensic data which will help them to develop more effective strategies and tailor their responses to specific threats and situations

1.1.4 Ethical and Practical Considerations in AI-Driven Forensic Investigations

(i) Balancing Privacy with Public Safety

While AI provides strong tools for forensic investigations, its use raises serious ethical problems, particularly around privacy. If automated gathering and analysis of personal data from social media is not adequately regulated, it has the potential to violate individual privacy rights. Legal frameworks such as Europe's General Data Protection Regulation (GDPR) and India's Information Technology (Intermediary Guidelines and Digital Media Ethics Code) Rules provide guidelines to ensure that AI is used ethically in forensic investigations (European Commission, 2024; Ministry of Electronics and IT, 2021).

A delicate balance between collecting Sufficient evidence to protect public safety and safeguard the privacy of the individual must be navigated by forensic experts. Practices like transparent data handling strict action control and rigorous oversight mechanism are therefore essential to maintain public trust and ensure that forensic investigations are conducted in an ethical manner

(ii) Mitigating Algorithmic Bias and Ensuring Fairness

Algorithmic prejudice is another significant issue in forensic investigations aided by AI. Artificial intelligence models are only as objective as the training data. The resulting algorithms may reinforce prejudices in their evaluations, resulting in unfair profiling and false allegations, if training datasets contain biases, whether based on socioeconomic position, gender, or race (Chesney & Citron, 2019).

Continuous efforts are required to diversify training data sets and implement fairness in these machine learning models for mitigating these issues. Regular audits and human oversight are necessary to ensure that ai system provides an unbiased and equitable outcome by addressing the algorithmic bias forensic applications of ai can enhance both their reliability and their legitimacy in the legal proceedings

(iii) The Arms Race: Criminal Innovation vs. Forensic Advancement

The rapid advancement of AI technologies has resulted in a continuous arms race between cybercriminals and forensic professionals. As forensic tools develop, crooks employ increasingly complex methods to escape detection. Fraudsters, for example, can use adversarial machine learning to subtly manipulate digital evidence, making it more difficult for automated systems to detect deviations (DARPA, 2020).

This continuous innovation in ai based forensic methods are necessitated by the constant push and pull. Adaptive Systems which are capable of learning from new types of criminal behavior in real

time should be made into development by the collaboration of researchers and law enforcement agencies. Only then through this sustained investment in research and development forensic tools will remain effective against the ever-changing tactics which are employed by cybercriminals.

1.2 Epilogue: Synthesis of Findings and Future Implications

The shift in how digital evidence is collected analyzed and utilized in the pursuit of criminal justice system the integration of ai in social media forensic investigation has played a significant role. AI has provided forensic experts with powerful tool to combat cybercrime which includes extraction of automated data sets and sentiment analysis to deep fake detection and predictive analytics. Technological advancements are rapidly increasing and are also bringing challenges which include piracy, algorithm bias and ethical oversight. We need to ensure that these innovations are used responsibly by providing a balanced approach which combines the technical powers of artificial intelligence with rigorous ethical standards and human oversight. The integrity of the forensic investigation can be maintained by the collaboration of policy makers law enforcement agencies and researchers to develop methodologies for the betterment of the future. Finally, the integration of AI will only be successful if it holds the promise of not only enhancing investigation techniques in an efficient way but also helps in protecting individual rights and upholding the rule of law in the digital world.

As the study focuses on understanding the emotional linguistic evolution of the violent offenders through AI driven analysis it is essential to position the investigation within the existing body of scholarly and forensic literature. The integration of artificial intelligence particularly Natural Language Processing (NLP) into the field of digital forensics is a relatively recent development. However, its applications are rapidly expanding from detecting spam and misinformation to more complex tasks such as emotion recognition, thread detection and behavior profiling. By understanding the shift in digital crime patterns and increasing reliance on online platforms by individuals for expressing extremist ideologies or emotional instability it has become important and critical to examine how NLP and sentiment analysis tool have previously been used to interpret human intent, emotional State and psychological indicators in digital communication.

It Is important to first understand what work has been done in this area as the use of artificial intelligence especially Natural Language Processing (NLP) is still fairly new in forensic size. It has shown a lot of promise as earlier studies have used these tools to analyze online content for things like hate speech misinformation and extremist language. However, many of them tend to skip how the offender's emotional state changes as they move closer to committing violence. While there is research on online radicalization and the influence of social media fewer studies have focused on tracking emotional changes over time to which this study calls the offender's "emotional progression". This is where the current research tries to contribute something new As this study Focuses on how ai has been used in forensic settings especially to study social media posts and a written statements by these offenders. It will also explore how past research has approached emotional analysis and behavioral predictions and what challenges exist in applying AI to human language and psychology. Around will help explain why this study is important and how it fits into a larger spectrum of forensic analysis and behavioral science

Literature Review

2.1 Historical overview

2.1.1 1990-Early 2000s

The field of digital forensics started rising in the 1990s when the law enforcement agencies validated the increasing need to collect, preserve, and analyse digital evidence. At the initial level, the efforts were done with a focus on recovering data from local storage devices like hard drives and floppy disks (Palmer, 2001). As cybercrimes such as hacking, identity theft, and digitally automated financial frauds became more and more prevalent, forensic specialists adapted to the new emerging techniques and advancements to investigate these growing threats.

During the same period. The Internet was evolving, and for the first time, the concept of social media emerged. These social media platforms started adapting to the modern age, and in fact, just like cybercrimes, these platforms were indeed a product of the modern age. Platforms like Six Degrees, Friendster, and MySpace paved the way for modern social networking platforms (Boyd & Ellison). These platforms not only allowed people to connect in new ways but also introduced digital risks like online identity fraud, cyberstalking, cyberbullying, and defamation.

At the same time, early researchers of this modern age were developing methods to analyse digital images for authenticity checks. Researchers like Hany Fraid Contributed significantly to image forensics by working on techniques like error level analysis, JPEG comparison, and noise pattern detection (Farid, 2009). These early methods provided the foundation for later achievements in AI-powered forensic tools.

2.1.2 Mid-2000s-2010s

As the 21st century reached the mid-2000s, platforms such as Facebook, Twitter, and YouTube revolutionised communication. The increasing popularity of social media also led to a sharp rise in cybercrimes, the spread of misinformation, and online scams (Casey, 2001).

Line four cement agencies recognised the importance of social media in the investigation and began incorporating it into their forensic practices. In 2012, the FBI launched the National Domestic Communication Assistance Centre (NDCAC) in support of digital investigations by providing expert assistance in handling data from social media, encrypted messaging apps, and online communication networks (FBI, 2012).

Forensic techniques developed during this period. Included:

- The extraction of metadata from social media posts to verify authenticity
- Analysing online behavioural patterns to detect criminal activities
- Geolocation and IP address tracking to identify suspects location
- Usage of automated keyword tracking to monitor hate speech, propaganda, and threats

With the subsidy rise of machine learning, law enforcement agencies also began using AI-powered algorithms to analyse large amounts of data from social media to help identify fraudulent accounts, misinformation campaigns, and cyber threats. In a more efficient way (Kaplan & Haenlein, 2019)

2.1.3 2010s–Mid-2010s

By 2010, artificial intelligence and machine learning were becoming essential in automating forensic investigations. These tools were AI-driven and improved the speed and accuracy of the digital investigation with some notable achievements, including

- Natural Language Processing (NLP): AI-powered NLP models helped forensic investigators analyse text from social media posts, identifying threats and misinformation (Zellers et al., 2019).
- Facial Recognition Technology: AI tools such as Clearview AI allowed law enforcement to match social media images with existing databases to identify individuals (Hill, 2020).
- Social Network Analysis (SNA): AI-assisted mapping of online relationships helped uncover cybercriminal networks (Ferrara et al., 2016).

The shift from manual forensic investigations to AI-powered automation marked a turning point, enabling real-time monitoring, threat detection, and pattern recognition in social media forensics.

2.1.4 Late 2010s–Present

In the late 2010s, deepfake technology—AI-generated manipulated videos and images—emerged as a major challenge in social media forensics. Initially created for entertainment and virtual reality, deepfakes quickly became a tool for fraud, misinformation, and blackmail (Chesney & Citron, 2019).

To counter this growing threat, forensic experts and governments developed AI-driven detection tools. The Defence Advanced Research Projects Agency (DARPA) launched projects such as Media Forensics (MediFor) and Semantic Forensics (SemaFor) to detect and analyse AI-generated media (DARPA, 2020).

Legal frameworks also evolved to address the misuse of AI-generated content:

- United States: The DEEPFAKES Accountability Act (2019) introduced laws against the malicious use of deepfakes (US Congress, 2019).
- • European Union: The AI Act (2024) established regulations for AI-generated misinformation (European Commission, 2024).
- • India: The IT Rules (2021) required deepfake content to be removed within 36 hours of being reported (Ministry of Electronics and IT, 2021).

Despite these efforts, the challenge remains ongoing, with AI-generated forgeries and forensic AI tools evolving side by side.

2.2 AI in Social Media Forensic Investigations

- **AI-Powered Evidence Collection:**

AI has transformed the way forensic investigators collect and analyse digital evidence. Some key applications include:

- Automated Data Extraction: AI tools scan social media platforms for posts, messages, and metadata relevant to criminal investigations.
- Sentiment Analysis: AI-powered NLP models help assess tone, emotion, and intent in online conversations to identify potential threats.
- Network Mapping: AI visualises connections between users engaged in criminal activities, making it easier to track offenders.

- **AI for Threat Detection and Cybercrime Prevention:**

AI plays a vital role in identifying and preventing cybercrimes through

- Deepfake Detection: AI models analyse visual and audio cues to distinguish real content from manipulated media.
- Fake Account Detection: AI algorithms help social media platforms identify and remove bots and fraudulent profiles.
- Anomaly Detection: AI detects unusual online activity, flagging potential security threats before they escalate.

- **Behavioural Analysis and Predictive Forensics:**

AI has also improved forensic investigations by predicting cybercrime trends and understanding criminal behaviour.

- Fraud Detection: AI examines transaction patterns to identify cases of identity theft, credit card fraud, and financial scams.
- Terrorism Prevention: AI monitors social media for extremist content and radicalisation indicators.
- • Psychological Profiling: AI analyses linguistic patterns to assess user behaviour and predict criminal intent.

AI has revolutionised forensic investigations, especially in the realm of social media forensics. By improving the speed and accuracy of investigations, AI helps law enforcement agencies combat cybercrimes, misinformation, and digital fraud. However, ethical concerns—such as privacy, bias, and misinformation risks—must be carefully managed. Moving forward, there is a need for a balanced approach that leverages AI's potential while maintaining ethical standards in forensic investigations.

Methodology

This study wraps around a comparative and exploratory design which combines elements of forensic content analysis with computational linguistic modeling and aims to identify and understand the behavioral progress of an individual who is involved Motivated through ideological or psychological means. The core issue is that violent offenders do not act in a vacuum their thoughts emotions and justifications are often expressed through a public or a semipublic digital space long before the act is committed (Weimann 2015). The study is rooted in the belief that digital content such as tweets manifestos or even chat logs is not just passive writing but often a reflection of internal psychological state and shifting emotions along with evolving intent. These texts when examined carefully can reveal subtle changes in language tone and emotional weight of the words that are otherwise overlooked.

Another factor shaping this methodological framework is the nature of the crimes being analyzed. Mass shootings, terrorist attacks, and other ideologically motivated violent acts frequently involve a complex interplay between extreme ideology, personal suffering, and a need for attention. Extreme ideology, personal suffering, and a craving for attention often interact in complex ways in mass shootings, terrorist attacks, and other violent acts driven by ideology. These objectives often find a voice in digital expression, notwithstanding their diversity. Most of the time, the offenders themselves are well aware of the possible consequences of their actions on the internet. They produce manifestos, stream information, or leave digital footprints in order to be seen, followed, and remembered. Therefore, rather than being accidental, their digital communication is crucial to the act itself.

Digital expression becomes, in a way, a staged performance of purpose, and the methodology is intended to analyze this performance alone. Instead of being seen as static or distinct, these texts are seen as part of a progressive narrative—a psychological accumulation conveyed through words. Consequently, the approach is both narrative-forensic and computational, recognizing that the offender's writing is a form of behavioral leakage or emotional projection (Meloy et al., 2012). Additionally, it is intentional to prioritize written language over visual or auditory content. Language remains the most direct and quantifiable form of self-expression. While films or memes may symbolically express ideology, words offer internal reasoning and cognitive justification. Many of these individuals verbally explain their moral stance, justify their violent acts, and articulate their objectives. Writing is therefore a crucial artifact for understanding the psycholinguistic evolution of these offenders. Because robots might not always be able to sense feelings and emotions this approach also analyze subjectively. Even the most advanced Natural Language Processing (NLP) algorithm can sometimes become confused by sarcasm coded language or ideological references. By striking a balance between automatic sentiment analysis and manual cross checking the work avoids this issue and enables more accurate interpretations in situations where AI might not alone be sufficient. This hybrid strategy preserves the research base in both technological and human viewpoint while strengthening the findings validity

The primary focus of this methodology is to track the progression of behavior over time. The study observes how the communication changes from one phase to another by collecting the written content of multiple individuals who later engaged in act of violence the data set have been compiled in a chronological order. It includes identifying when frustration begins to turn into anger when ideology starts to take more extreme form and when the final message carries a sense of resolution or moral justice and justification. Another crucial element of this approach is the focus on timeline-based development. Phases of violence are common, particularly when ideology is the driving force. The objective is not only to analyse the content but also to organise it along a chronological axis and monitor changes in language, emotion, and emotional markers as the individual approaches the action moment. This sequential method may reveal escalation patterns that conventional static text analysis typically overlooks.

The approach here is evidence based and centered on what these individuals have actually written rather than relying on a speculative psychological profiling emotional under current sentiment shifts and thematic patterns have been analyzed And it includes tracking the use of emotionally charged words repeated ideological keywords and the tone of the statements as the individual move closer to the act of violence.Offenders may complain about social isolation, display signs of identity ambiguity, or voice nebulous ideological grievances in the early stages. Usually, these words sound passive or rhetorical. However, as their viewpoint hardens, their tone shifts; their messages may become more focused, emotionally abusive, or self-righteous. In the later phases, there may be obvious signs of closure, such as direct statements of intent, purpose affirmations, or cryptic farewells. Capturing these subtle linguistic and emotional turns requires both technology and a forensic lens grounded in behavioural understanding.

To help with this Natural Language Processing (NLP) Tools are used to quantify and detect the changes that may not be immediately visible through manual reading And might fluctuate due to human error. A sentiment analysis helps to identify whether a message lean towards a negative neutral or a positive tone while the emotional detection highlights much deeper feelings such as fair, anger, or anticipation. These tools help to support the human reading of the text. The interpretations of these texts remain critical especially when dealing with complex ideologies and personal perspectives. Interpretability and utility are also taken into consideration when choosing Natural Language Processing (NLP) methods for the study. Because they offer transparency in language grading and are known to perform well on shorter, more casual texts, like tweets or online comments, TextBlob and VADER, for instance, are well-liked sentiment analysis tools (Hutto & Gilbert, 2014). This study favours models that provide forensic clarity over computational power alone, even in the phase of more advanced transformer-based models. Analysing the results statistically, behaviourally, and logically is crucial.

The study is designed in such a way that it is both structured and flexible. While it applies a clear timeline model by dividing content into early middle and final phase it also allows space for subjective judgment when a text doesn't fall into a specific category. The aim of this study is not to generalize or simplify the ideologies but to capture patterns of escalation and internal conflict

when they appear in a digital format. All the data used in the research is publicly available. No private communication or Confidential sources were involved. The intention of this research is not to label or speculate about the psychology of the individual but to examine how violence communicates itself before it happens in the real world and whether those communications carry warning signs that can be studied in a forensic setting.

What sets this method apart is the decision to let the offender's own language explain their growth. Instead of just employing psychological labels or pre-existing profiles, the study aims to understand them on their own terms through what they wrote, how they expressed it, and when they chose to say it. This achieves a balance between the computer-based process of an AI and the human intuition required for forensic work. While most online monitoring tools today are designed to flag content in real time, they often miss the gradual build-up that results in violent intent. This study, however, emphasizes the value of rational behavioral mapping not to predict specific outcomes, but to find emotional and linguistic cues shared by a number of offenders. We may be able to better understand how language turns into ideology and how ideology turns into action by comparing these patterns.

Ultimately, the goal of this methodology is not to replace existing approaches for risk assessment or investigative profiling, but to offer a different path for forensic analysis based on language, emotion, and behavior. It focuses on the complex nature of the subject, the ethical boundaries of academic research, and the potential uses of such endeavor in spotting and preventing dangers if done with caution.

In Forensic academia this type of methodology has not been used much particularly when seen from a timeline focused case comparative lens. The majority of the research either examines hate speech in general population or examines specific manifestos. With language serving as the main thread this research attempts to link Emotions ideology and behavior across time. Essentially it is an investigation into how individuals write themselves towards violence and how those changes in meaning and tone can be monitored, analyzed, comprehended and ultimately identified.

In forensic analysis the role of digital evidence has grown exponentially in recent years. Social media in particular has become a primary tool for individuals involved in mass shootings terrorism and extremist activities to express their grievances share ideologies and, in some cases, justify their violent actions. Social media posts tweets and manifestos offer direct insight into the thought process and emotions of individuals making them invaluable resource for forensic psychologist and criminal investigators aiming to understand the evolution of radicalization.

The study is based on the understanding that language as a powerful mode of expression holds the potential to reveal not only the cognitive state of individuals but also the psychological journey they undergo in transitioning from ideological expression to violent actions. The methodology adopts in this research is primarily driven by the hypothesis that radicalization and eventful violence are processes that can be traced through language from early expressions of dissatisfaction to explicit justification of violence. Through this lens social media posts become more than just expression of opinion or emotions they are vital artifacts of Radicalization that can help track behavioral changes and flag potential threats before they materialize into violent acts.

By focusing on written language and sentiment analysis this research aims to understand how individuals use social media to build, process and act upon violent ideologies. These specific individuals selected for the study are perpetrators of mass shooting terrorist attacks and those involved in ideologically motivated violence whose social media posts are central to understanding their path to violence. Over time these individuals express increasingly hostile behavior frustration and rage often culminating in manifestos or videos that articulate justifications for their violent actions. These digital traces are therefore crucial not only for understanding past events but also for creating predictive models that can help authorities and organizations identify and intervene in potential future incidents of violence.

This study seeks to know how people use social media to construct reinforced and act upon violent ideas by means of written language and sentiment analysis. The particular people chosen for the study are those engaged in ideologically driven violence whose social media posts are crucial for knowing their path to violence as well as perpetrators of mass shooting terrorist attacks. These people show more hostile behavior over time, often in frustration and rage, which leads to manifestos or films that explain rationales for their violent deeds. These digital footprints are therefore vital not only for analyzing past occurrences but also for developing predictive models that enable organizations and authorities spot and act in possible future violence. The research method uses sentiment analysis, Natural Language Processing (NLP), and human validation to examine the linguistic trajectory of these individuals' posts. It focusses on linguistic patterns, emotional intensity, and ideological development to identify important shifts in their discourse that may be early indicators of violence. By compiling a thorough history of online activity, this method aims to track the development of radicalization from quietly voiced discontent to overt threats of violence. Combining automated Natural Language Processing (NLP) algorithms with manual human validation allows for early detection of potential threats and offers a thorough understanding of the ideological, emotional, and psychological changes that take place over time.

Violent offenders really act in isolation. Thoughts frustration and ideological justification are often broadcasted to digital spaces both public and semipublic long before any real word action is taken (Weimann, 2015). These communications serve as early warning signs. When analyzed closely, they can reveal subtle but meaningful shifts in tone, emotional intensity, and lexical choice that are typically overlooked in conventional assessments. This study hypothesizes that language itself becomes a conduit for internal turmoil, and tracking these changes over time can provide valuable

forensic insights. In particular, the study pays attention to digital expressions from platforms historically linked with extremist discourse such as Twitter, Reddit, 4Chan, 8Kun and personal blogs, where many perpetrators have previously left behind a digital trail. Furthermore, the methodology draws from documented real-world incidents to ground its framework, For instance, the 2019 Christchurch shooter released a detailed manifesto and streamed his attack while the Buffalo supermarket shooter in 2022 had a timeline of radicalization documented online. These cases studies underscore how online content often precedes violent acts and can be critical for understanding the transition from thought to action. Through this comparative analysis of individual cases this research will help to identify common linguistic and emotional markers that appear across different platforms for different individual's radicalization journey. By keeping a track of these markers over time this study seeks to establish a pattern which can be used to flag emerging risks in the future. The ultimate goal of this research is to provide a methodological framework for the identification of early indicators of violent radicalization which can contribute to preemptive intervention strategies and aim at reducing The likelihood of future violent incidents

3.1 Data collection and preprocessing

For data collection and preprocessing publicly data of individuals who later become perpetrators of mass shooting terrorism or ideologically driven violence was systematically gathered. The primary source for this study includes Twitter Reddit and 4 Chan which are all significant platforms for expression of radical and extremist ideologies. These platforms provide an unfiltered real-time account of individuals thoughts grievances and plans as they often promote free speech hence revealing emotional and psychological upbringings of the person and underpinnings for their action

The data set for this study comprises a total number of ten individuals who were known for perpetrating mass violence and ideologically driven attacks. These individuals were selected based on the public availability of the return and digital communications prior the violent act. The textual data was extracted from a combination of manifestos, social media posts, interviews, video transcripts and online forums depending on the availability for each individual. Every effort was made to preserve the integrity of the original while ensuring the contextual accuracy of the data. Textual samples were selected to represent different behavioral phases which include early mid and final phase for mapping emotional progress and sentiment escalation in the form of a progressive timeline. Where necessary, offensive language was censored for ethical and academic sensitivity

The offenders analyzed in this study include:

1. **Brenton Tarrant** (New Zealand mosque shooter) – *Manifesto and online forums*

2. **Elliot Rodger** (Isla Vista killings) – *Manifesto, YouTube transcripts*

3. **Nikolas Cruz** (Parkland shooter) – *Instagram comments, forum posts*

4. **Omar Mateen** (Orlando nightclub shooter) – *Social media posts, 911 call transcripts*

5. **Robert Bowers** (Tree of Life Synagogue shooter) – *Gab posts*

6. **Stephen Paddock** (Las Vegas shooter) – *Investigation reports, indirect notes*

7. **Seung-Hui Cho** (Virginia Tech shooter) – *Manifesto, video monologue*

8. **Dylann Roof** (Charleston church shooter) – *Manifesto, jailhouse writings*

9. **Anders Breivik** (Norway attacks) – *1,500-page manifesto*

10. **Payton Gendron** (Buffalo supermarket shooter) – *Manifesto, Discord logs*

After the collection of posts the data undergoes a cleaning and preprocessing phase which ensures that the data set is both usable and efficient for analysis. This part is crucial because social media data is inherently noisy and unstructured As it often contains irrelevant and misleading information such as unrelated hashtags user mentions and links to external content the preprocessing steps involved the following:

Tokenization- The given text is broken down into individual tokens that is words or phrases which can be analyzed by NLP algorithms. This process allows the model to focus on core elements of the text and strip away unnecessary symbols and punctuations (Manning et al., 2008)

Normalization- The social media language often contains slangs, emojis and abbreviations Which might not be standard across all the platforms hence the text is normalized to convert all the characters into a uniform format. This Includes expanding abbreviations (e.g., “u” to “you”),

Converting slangs into standardized formats and removing unnecessary characters such as repeated punctuation marks.

Stop word removal- The language used on social media contain common words such as "the", "a" and "is" that carry little semantic meaning hence, they are removed from the data set to improve the processing in an efficient way. These words are termed stop words and are excluded from the analysis to allow the focus to remain on more meaningful content (Jurafsky & Martin, 2020).

Lemmatization- Words are reduced to their base or root form through this process so that the variation of a word is treated as a single item. For instance, "running" and "ran" Would be both reduce to the base "form run" (Bird et al., 2009). This helps in standardizing the data and ensures that different forms of the same word are not treated as a separate data set

Filtering irrelevant content- Across all the social media platforms many posts may contain links to external content unrelated hashtags or promotional materials such as advertisements which are often personalized. Therefore, it is necessary to remove them from the data set to ensure that the focus remains only on the content of the post. Only posts that reflect the individual's personal thoughts ideologies and actions are retained.

Once this data is preprocessed it is then organized and stored in a structured format which is suitable for analysis. This data set is then ready for the next part of the analysis where it will undergo sentiment analysis and NLP based categorization which helps in identifying emotional shifts ideological trends and escalations towards violence. At this point the data is ready to be processed through both manual and automated Methods for meaningful insights that will inform the comparative analysis and the timeline reconstruction of the data.

Automated models alone are insufficient especially when we are analyzing sarcasm, coded language or culturally contextualized ideologies. Thus, there was a need to maintain a hybrid approach in which automated analysis was followed by manual cross checking to maintain accuracy especially in case where AI misreads subtext. The method ensured that complex ideological narratives were not reduced to just sentiment scores but were understood within their context. Ethically speaking the study maintained a strict adherence to research standards. All the data used is publicly accessible and does not include private messages leaked content or any material gathered without the consent of the individual. While the goal of this study is to analyze pre-incident language for early warning signs, the research avoids speculating or racially profiling the individuals. Purpose is to investigate how language reflects behavioral transformations not to diagnose individuals or draw baseless conclusions.

3.2 Sentiment analysis and natural language processing (NLP)

When doing a forensic analysis understanding the psychological under-prints of violent radicalization requires more than just identifying the key themes in online communication. It requires examining the emotional weight of the words used the shifts in sentiments and the

evolving language pattern that indicates an individual's trajectory towards violent actions. This is where Sentiment Analysis and Natural Language Processing (NLP) Comes into play providing an essential tool for the identification of emotional and ideological shifts in the language of an individual.

Sentiment analysis is a subset NLP Which involves the computational detection of subjective information in text format including attitudes emotions and opinions (Pang & Lee,2008). When talking about social media post sentiment analysis solely focused on identifying whether the tone of the post is positive negative or neutral. Given the often-volatile nature of the content associated with mass shooting terrorist attacks and ideologically motivated violence the sentiment behind a post can reveal critical information about an individual's mental state, their frustrations and change in the ideology. Through sentiment analysis the emotions embedded in a written communication can be tracked and quantified for example early post from an individual may express frustration or dissatisfaction which may overtime evolve into anger hatred and ultimately call for violence. These shifts of language are not always immediately noticeable to human readers especially in the context of online communication where emotional context can be hindered by shortened language sarcasm or even coded speech. NLP Tools help extract these subtle emotional shifts by processing large volume of text data set and identify trends in sentiment that may not be apparent on the surface level and might be left behind due to human error. This allows forensic analyst to detect early signs of radicalization even in cases where the post seem unalarming at the first glance.

Figure2.1 Screenshot 1: Initial NLP Data Extraction Code

```
[ ] from vaderSentiment.vaderSentiment import SentimentIntensityAnalyzer
    from nrclex import NRCLex
    import pandas as pd
```

```
    df = pd.read_excel('/content/Violent Offender Real Text Analysis Dataset.xlsx', sheet_name='Sheet2')
    df.head()
```

	ID	Name	Date	Source Type	Phase	Text	Notes	Compound Score	NEG	NEU	POS	Emotion (AI)
0	1	Brenton Tarrant	2019-03-15	Manifesto	Early	"I am just a regular White man, from a regular...	Normalization of self before ideology	NaN	NaN	NaN	NaN	NaN
1	2	Brenton Tarrant	2019-03-15	Manifesto	Early	"The origins of my hate began with the death o...	Emotional trigger mentioned as turning point	NaN	NaN	NaN	NaN	NaN
2	3	Brenton Tarrant	2019-03-15	Manifesto	Early	"We must crush immigration and deport those in...	Expression of ideology, us-vs-them framing	NaN	NaN	NaN	NaN	NaN
3	4	Brenton Tarrant	2019-03-15	Manifesto	Mid	"There is no where left to go. No where to run."	Hopeless tone, isolation narrative	NaN	NaN	NaN	NaN	NaN
4	5	Brenton Tarrant	2019-03-15	Manifesto	Mid	"True change and victory only comes through fo...	Acceptance of violence as necessary	NaN	NaN	NaN	NaN	NaN

```
[ ] text_list = df['Text'].tolist() # String have unnecessary quotes
    cleaned = [s.strip('"') for s in text_list]
    print(cleaned)
```

```
['I am just a regular White man, from a regular family.', 'The origins of my hate began with the death of Ebba Akerlund.', 'We must crush immigr
```

Figure 2.2 Screenshot 2 : Interpretion of Scores

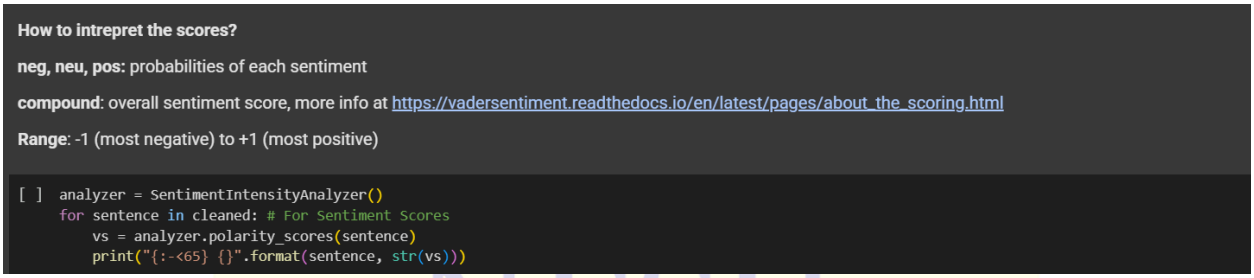


Figure 2.3 Screenshot 3 : Emotion Analysis

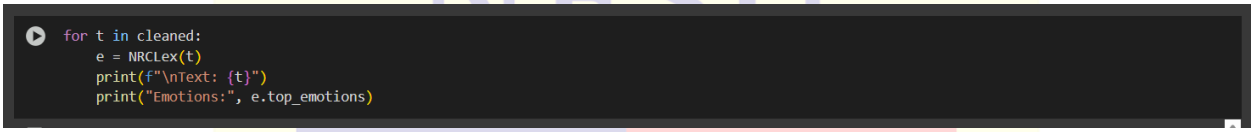


Figure 2.4 Screenshot 4 : Saving analysed data to Excel File

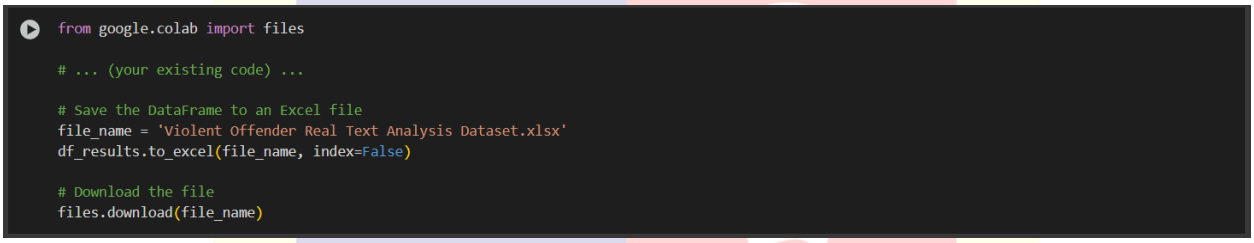
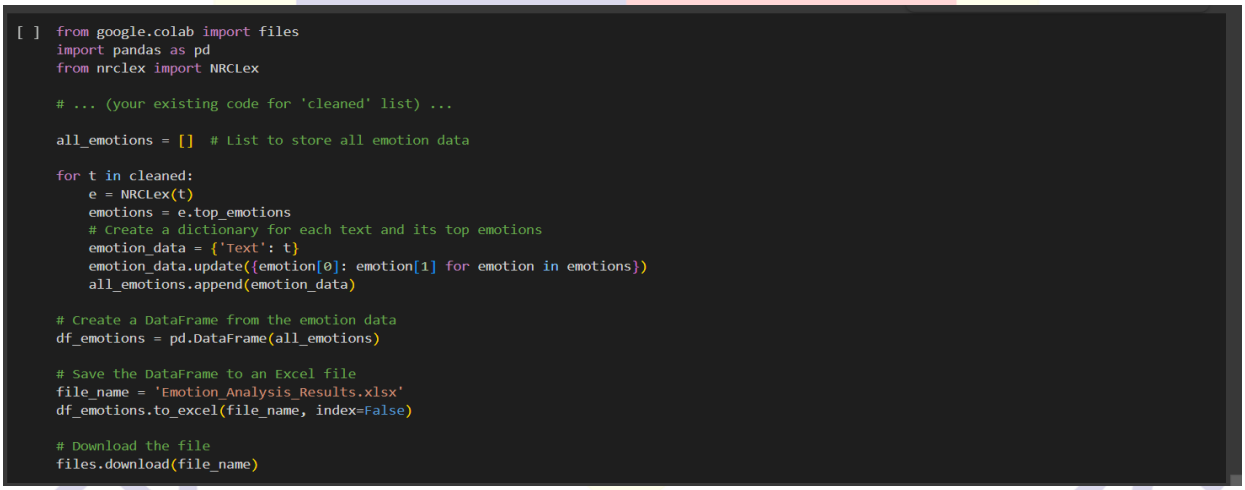


Figure 2.5 Screenshot 5 : Creation of Directory



On the other hand, Natural Language Processing (NLP) is a broader field which consists of computational linguistics that enables machines to understand interpret and generate human language (Jurafsky & Martin, 2020). It is a powerful tool for analyzing large dataset of textual content particularly when it comes to unstructured and noisy data as is the case with social media post. NLP techniques such as tokenization lemmatization and part of speech tagging Allows for the discretion of text into meaningful components hence making it possible to identify key phrases repeated keywords and ideological markers that may signal the onset of radicalization. For instance, if we take the process of tokenization, it breaks down text into smaller manageable pieces often into individual words or phrases that can be later analyzed for further patterns of linguistic

change. Lemmatization ensures that the variation of a word (e.g., “run”, “running” and “ran”) Are reduced to a single root form, which allows for more accurate analysis or the frequency and signifies the specific term (Bird et al., 2009). Hence my focusing on these fundamental linguistic elements NLP Algorithms are able to track the evolution of key ideological terms such as “Justice”, “Oppression” or “Revenge” across time. In the context of radicalization, these tools are particularly important for the identification of semantic shifts that occur as an individual moves from expressing general dissatisfaction with societal issues to advocating for violence as a solution for these problems. For example, an individual who begins by discussing “Corruption in the system” might gradually shift to more violent language such as “revolution”, “war” or “attack”. NLP algorithms can track these shifts by focusing on the context in which these words are used and allow for the detection of emergent threats before they manifest in physical violence.

The integration of sentiment analysis and NLP into forensic social media analysis allows researchers to build a more comprehensive understanding of how individuals express their ideologies and emotions overtime across these platforms. These tools help s in the identification of psychological markers of radicalization by highlighting the emotional weight of certain phrases and tracking how an individual stone become more aggressive or defensive as they approach the point of violent action. For instance, FN individual begins by expressing anger or frustration with political or social system but overtime as their view become more rigid their post may escalate into hostile rhetoric targeting specific ethnic groups or individuals.

In this research sentiment analysis and NLP tools are used in tandem to create a nuanced understanding of these emotional and ideological trajectories of an individual who is involved in ideologically motivated violence. By processing these large datasets of social media content these tools allow for the identification of significant emotional and linguistic shifts that might otherwise be overlooked by human analysis alone therefore, methodology does not rely on computational tools alone but also includes a crucial element of human validation which ensures the accuracy and contextual understanding of the analysis.

As the study progress N LP and sentiment analysis become essential in tracking the timeline of Radicalization. These tools not only help in the identification of patterns within an individual’s language but also allows the comparison of these patterns across multiple cases. By tracking the linguistic and emotional evolution of multiple individuals over time researchers can begin to identify common pathways of radicalization which leads to a deeper understanding of the factors that contribute to violent extremism hence provides a valuable insight into how language and sentiment evolves as individuals move closer to the point of action offering potential intervention for preventing violent outcomes. Through this sentiment analysis and NLP this research contributes to a growing body of forensic work that seeks to understand the link between language ideology and behavior. By automating the process of analysis for social media data these tools enable forensic analyst to detect early warning signs of radicalization and identify individuals who may be at risk of engaging in such violent acts. The combination of ai assisted analysis and human

expertise ensure that the emotional and ideological shift that define the path of radicalization is not only captured but also understood within the broader context of social and psychological factors.

3.3 Hybrid Analysis Model: AI and Human Validation

While Artificial intelligence (AI) has significantly transformed the landscape of digital forensics and social media analysis it is important for us to recognize that no automated system can entirely replace human judgment especially when it involves different ideologies, human emotions, and intent. This research adopts a hybrid analysis mode and integrates both ai driven rules and expert human validation to ensure a more accurate context aware and ethically responsible examination of radical trajectories. Base tools such as NLP algorithms and sentiment classifiers are essential for managing and analyzing large scale data set which are derived from social media platforms. These tools help in rapid processing of text identification of patterns and extraction of sentiment trends over time. These tools are particularly effective in the initial stage of analysis where the sheer volume of the data makes manual assessment totally impractical. By automating es like keyword extraction thematic clustering, sentiment scoring, and frequency analysis ai tools help in highlighting anomalies and key features that warrant closer investigation.

However, while these automated models are excellent for the reorganization of patterns, they often fall short when it comes to interpreting nuance, sarcasm, cultural references or coded languages all of which are frequently found in extremist discourse. For example, phrases that may appear isolated could carry significant ideological weight within specific subcultures or online communities. This is especially true in forums like 4Chan where language used is slightly stylized, ironic or purposely ambiguous. Therefore, relying solely on artificial intelligence can result in misinterpretation false positives or missed indicators. To mitigate these limitations this study incorporates a layer of human validation to review and interpret AI flagged data. Human analysis is responsible for assessing context verifying meaning and providing cultural and psychological interpretations that a machine cannot offer. This step is crucial for understanding the true intent behind statements especially in ambiguous or borderline cases where language might be misleading or indirect and have high chances of being taken out of context. Moreover, human validation serves a vital ethical foundation. Ai systems especially those trained on biased or limited data will eventually perpetrate stereotypes or disproportionately flag certain groups. By the incorporation of human oversight, the methodology ensures that findings are not only accurate but also sensitive to social and ethical concerns. A human review evaluates AI generated output to correct for bias, confirm relevance and maintain a balanced, objective perspective throughout the analysis.

The Iterative refinement of the AI tools is also supported by this hybrid approach. As human analysis valid ate the AI outputs the data can be fed back into the system to improve further performance and give better results. This form of machine learning feedback helps the model

evolve over time by making it more tuned to a specific pattern of a radicalization language and sentiment relevant to the study. It represents a dynamic collaboration between technology and human insight where each helps in strengthening the other. In the context of forensic research on ideologically motivated violence this hybrid model offers the best of both worlds which include the scalability and efficiency of AI combined with the in depth and ethical awareness of human analysis. This model ensures that the conclusions drawn from the data are not only computed but also psychologically and contextually valid hence providing a more comprehensive understanding of how online behavior translate into real time threats. Ultimately the integration of human and machine intelligence enhances the credibility and applicability of the research findings. Allows the investigator, policymaker and law enforcement to rely on a framework that is both data driven and deeply informed by human judgment hence an essential combination for addressing the complex evolving nature of online radicalization and violence.

3.4 Timeline Construction and Behavioral Evolution

One of the most critical components of this research involves the reconstruction of behavioral timeline for individuals who later on engaged in act of violence. The goal at this stage is to identify and document how the ideologies and emotional expression of these individuals evolved over time particularly across different social media platforms. By aligning post in a chronological manner, the study seeks to uncover temporal patterns in language sentiment and ideology that may signify a progression from passive disconnect to active radicalization and eventually violent intent. The temporal mapping process begins with the earliest available digital trace which often include mundane or ambiguous post and tracks them up to the final public communication before the incident happened which may include manifestos farewell message or direct threads. These timelines serve as a structured framework for observing major shifts in tone content intensity of language. They allow researchers to detect moments of transformation such as a turn from frustration to hatred or from wake political commentary to explicit calls for violence. By constructing these behavioral timelines, it involves not only the aggregation of content but segmenting the data into development phase. These segments may include an initial grievance phase a phase of ideological exploration and reinforcement and escalation phase marked by increasing hostile behavior and finally justification or declaration phase. By applying both NLP based trend analysis and human contextual interpretations the methodology identifies turning points that may correspond with external events social influence or psychological triggers. Moreover, these timelines provide valuable insight into the pace and rhythm of radicalization. In some cases, these individuals undergo a slow and cumulative shift with the ideology and anger building gradually over months or years And for some the radicalization process is rapid often catalyzing by triggering events such as loss of job political development or exposure to extremist content within a short period of time. This volatility is crucial for building effective intervention models as it underscores the need for flexible risk assessment tools that accounts for both long term and accelerated radicalization trajectories.

Another important aspect of this timeline construction is the integration of multi-platform data. Many individuals use different social media platforms for different purpose for example Twitter for broadcasting opinions, Reddit for engaging in discussions and 4Chan for venting anonymously. Mapping the activity of an individual across these platforms provide a more comprehensive picture of their ideological and emotional development. It also helps in contextualizing the functional values of each platform and the radicalization process, revealing, for example whether a user sort affirmation from echo chambers, tested out violent ideas in fringe forums or escalated language when met with encouragement or validation. These timelines also allow for correlation between online behavior and offline actions. Cases post expressing intent or rehearsing ideological narratives were made shortly before the actual event. By identifying such high-risk text the methodology, the importance of real-time monitoring tool that can flag behavioral red flags. This supports the argument for enhancing digital forensic capabilities among law enforcement and intelligence agencies particularly for the detachment of individuals in the final stage of violent planning. Ultimately the timed and construction offers forensic map of an individual's descent into violence. Binds computational precision with human interpretation to elevate the psychological and ideological progression behind their public expression. By charting the evolution of thought and emotions over time these timelines not only explain how radicalization occurred but also helps in the development of predictive framework for identifying and intervening in similar trajectories before they cumulate violence.

3.5 Comparative case study and cross analysis

The final phase of this research involves the competitive analysis of multiple individuals' timelines forming a cross-case study that seeks to identify common linguistic, emotional, and ideological patterns across diverse incident of violence. This comparative framework is essential in establishing whether certain features of Radicalization are universal or if they vary significantly depending on personal background environment lifestyle or platform usage.

Each case study begins with an individual profile which includes a compilation of their social media activity behavioral timeline sentiment trends and ideological markers. These profiles serve as the basis for a deeper comparative investigation. By examining the nuances of each individual's progression starting from initial expression of grievance to the justification of violence this research helps in the identification of reoccurring patterns and anomalies across different cases. For instance, while some individuals may use dehumanizing language towards specific groups while other may focus on self-victimization or moral righteousness as a means of justifying their action.

One core area of analysis is the linguistic commonality in expression of hatred, alienation and entitlement. By using NLP tools to analyze vocabulary usage phrase repetition and rhetorical structures this study isolates specific terms and sentences structured frequently associated with radical behavior and belief system. This is then supported by sentiment analysis which reveals a shared trajectory among many individuals hence a gradual intensification of negative sentiments

particularly anger, fear and contempt towards perceived enemies or systematic structure. Such emotional markers become critical in identifying psychological inflection points where passive beliefs turn into readiness to act. The cross analysis also incorporates ideological frameworks. In the study may be motivated by religious extremism white supremacy, misogyny or antigovernment sentiments yet their narratives often share structural similarities such as the portrayal of the self as a misunderstood hero or a martyr or the framing of violence as a last resort in response to injustice. These ideologies overlap each other and are more important for understanding how extremist rhetoric transcends specific doctrine and how online spaces facilitate and breed the adaptation of shared narratives across different ideological communities. Additionally, the study examines platform-based influence on Radicalization Trajectories for example individuals using fortune may exhibit more explicit and humanizing rhetoric due to platform's anonymity and lack of moderation whereas those on Twitter might adopt more coded language to bypass detection hence understanding how platform culture shapes expression helps contextualizing behavioral signals and Taylor platform specific intervention strategies. It also highlights the importance of analyzing not just what is said but where and how it is said.

A significant outcome of this competitive process is the identification of convergence zone which are moments where different individuals despite their diverse ideologies and personal histories demonstrate a similar behavior or posting style. These convergence zone can be used to develop a taxonomy of warning signs that could inform AI driven monitoring system or human led investigation protocols. For example, repeated references to perceived betrayal expression of divine or moral justification or sudden increase in the volume and frequency of post are among the signals found to be consistent across multiple cases regardless of the individual's identity and personal background. Furthermore, this cross-case study allows for contextual interpretation of anomalies which include instances where individuals deviate from expected patterns. Outliners are equally valuable as they help refine predictive models and caution against over generalizing the data. Not every individual who expresses anger or shares extreme content proceeds to violence and understanding the threshold of action is what separates rhetoric from reality which is crucial in developing responsible forensic frameworks.

In conclusion this competitive analysis not only reinforces the individual findings of each case but also contributes to a broader understanding of radicalization as a pattern yet personalized process. By identifying both commonalities and differences this phase supports the development of a holistic intervention strategy that is essential and sensitive to context flexible in application and grounded in empirical linguistic and psychological evidence

Results & Discussion

1. Introduction to analytical findings

The following chapter represents a detailed analysis of sentiments and emotions derived from a curated data set of text-based expression attributed to 10 violent offenders. The Dataset has been compiled manually and verified for authenticity including a total of 150 statements categorized according to their behavioral timeline which includes early-stage mid stage and final stage eventually leading up to acts of mass violence or terrorism. The core objective of this analysis is to uncover the emotional and psychological progression in these offenders' communications. Applying natural language processing NLP technique specifically sentiment polarity scoring and

emotion clarification this study aims to highlight how linguistic patterns evolve in proximity to the act of violence.

Each statement was processed using an AI model that evaluated:

- Sentiment Polarity: A compound score ranging from -1(strongly negative) to +1(Strongly positive)
- Sentiment distribution: Probability weights across negative (NEG), neutral (NEU) and positive (POS) dimensions
- Emotional categories: Classification into primary emotions such as anger fear trust sadness anticipation extra using lexicon-based tagging.

This mixed method approach combines AI generated outputs with manual labeling to construct behavioral timeline. These timelines were then crossed compared across offenders to detect shared progression patterns, peak emotional movements and shifts in intent expression.

2. Overall Sentiments Trends:

The sentiment analysis conducted on this data set comprises of 150 textual entries from 10 providing valuable insight into the progression of their emotional and psychological states across three behavioural phase which include early mid and final phase. Natural language processing NLP techniques each statement was evaluated for its sentiment polarity and intensity through lexicon-based models such as Vader this allowed for the computation of a compound sentiment score which ranges from -1 (strongly negative) to +1 (strongly positive) along with distribution probabilities across three core sentiment categories which include negative neutral and positive.

These findings revealed a measurable fluctuation in sentiment polarity over time. In most of the cases in the early phase sentiment scores were typically neutral or mildly negative. Offenders often expressed personal frustration ideological curiosity or social annihilation. Majority of the entries were filled with emotional fluctuation over the course of a timeline. The Max sentiment score of

all entries was 0.8126 and Min sentiment scored turns out to be -0.8481 whereas the Avg. sentiment

score was -0.12329. These sentiments were categorized into three basic types, NEG(Negative), NEU(Neutral) & POS(Positive) Having average score of 0.143827, 0.779513 & 0.07666 respectively. Furthermore, they were classified on a broader spectrum which included sentiments like Positive, Negative, Fear, Anger, Trust, Surprise, Sadness, Disgust, Joy & Anticipation.

On averaging the sentiment score of all 150 entries following results were obtained-

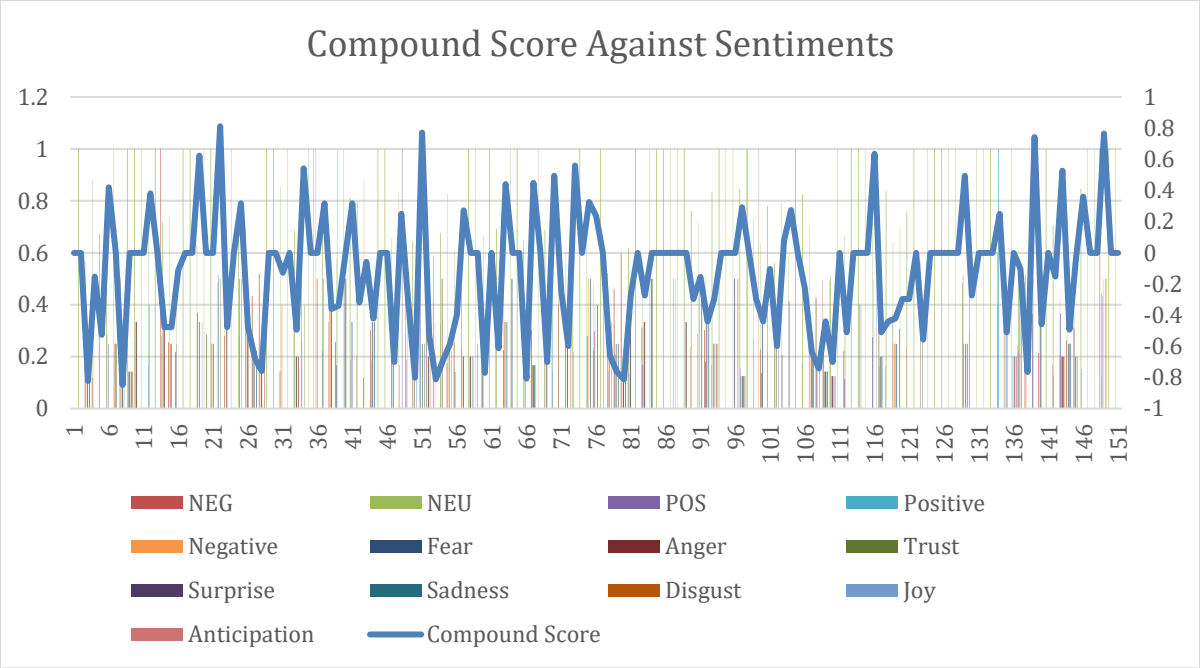
Table 2: Average Compound Score for 10 Sentiments

Positive	**0.132278245**
Negative	0.154465148
Fear	0.130948526
Anger	0.085100151
Trust	0.093098252
Surprise	0.031561584
Sadness	0.806717
Disgust	0.38106909
Joy	0.25283447
Anticipation	0.445386905

Table 3: Max/Min/Avg. Scores: NEU, POS, NEG

MAX Score	MIN Score	Average NEG	Average NEU	Average POS
0.8126	-0.8481	0.143827	0.779513	07666

Table 4: Compound Score Correlation with Sentiment Distribution



3. Emotion Distribution by Behavioral Phase

The emotional role of an offender can often reflect the underlying psychological escalations that leads to violent action. In this section the study analyzes emotional patterns associated with 150 manually curated text entries from 10 violent offenders. Using NLP based emotional classification models each text was Labeled with possibilities across 8 emotional categories **Anger, Fear, Sadness, Disgust, Joy, Trust, Anticipation**, and **Surprise**. The most dominating emotion for each entry was determined by the highest probability score of each offender. For better understanding of the temporal and behavioral dynamics of these emotions the data set was segmented into three phases early mid and final. Each phase represented a stage in the ideological or emotional evolution of the offender. The phase wise analysis below provides insight into the transformation of emotional tone as each individual progressed toward violent action.

3.1 Early Phase

The early phase signifies the initial part of the timeline which led to the domino effect. These four emotions largely dominate the early phase of these individuals and summarize the emotional capacity the evolved from. By the below table we can easily conclude that during the early-stage most individual show neutral emotion along with fear, sadness, disgust and anger with fear being the most dominant after a neutral emotional response

Table 5: Early Phase Behavioral Data

Phase	Fear	Anger	Trust	Surprise	Sadness	Disgust	Joy	Anticipation
Early	0	0	0	0	0	0	0	
Early	0.166667	0.166667			0.166667	0.166667		
Early								
Early	0	0	0	0	0	0	0	
Early	0	0	0	0	0	0	0	
Early			0.333333					0.333333333
Early	0	0	0	0	0	0	0	
Early	0.2	0.2			0.2	0.2		
Early			0.4					
Early								1
Early	0.285714							
Early				1				
Early								
Early	0	0	0	0	0	0	0	
Early	0.333333				0.333333			
Early	0.5							
Early			0.5					
Early	0.333333				0.333333			
Early	0.25				0.25			
Early	0.2	0.2			0.2	0.2		
Early	0.333333				0.333333			
Early	0.25					0.25		0.25
Early	0	0	0	0	0	0	0	
Early		0.2		0.2	0.2	0.2		
Early				0.181818				
Early	0.142857	0.142857	0.142857		0.142857	0.142857		
Early	0	0	0	0	0	0	0	
Early	0.285714							
Early	0	0	0	0	0	0	0	
Early	0	0	0	0	0	0	0	
Early	0	0	0	0	0	0	0	
Early			0.5					
Early								
Early								
Early	0.333333							

.

3.2 Mid Phase

The middle phase serves as the connecting bridge between the early and final stage. In this phase it is observed that individuals are heavily influenced by fear and anger almost on a similar scale with a slight sense of trust and anticipation. This mid phase directs the trajectory of emotions and acts as a catalyst For the final act of violence

Table 6: Mid Phase Behavioral Data

Phase	Fear	Anger	Trust	Surprise	Sadness	Disgust	Joy	Anticipation
Mid	0.25							
Mid		0.25	0.25					
Mid	0.25	0.25			0.25			
Mid	0.142857		0.142857			0.142857	0.142857	0.142857143
Mid			0.285714					
Mid		0.25	0.25					
Mid			0.5					
Mid	0.333333	0.333333						
Mid			0.5					
Mid		0.5						
Mid								
Mid		0.5						
Mid								
Mid	0.5		0.5					
Mid	0.2				0.2			
Mid			0.25				0.25	
Mid	0.2	0.2				0.2		
Mid								
Mid								
Mid	0.166667	0.166667	0.166667		0.166667	0.166667		
Mid	0.25							
Mid	0.333333	0.333333						
Mid	0.333333	0.333333						
Mid	0.5		0.5					
Mid	0.5							0.5
Mid		0.125	0.125	0.125	0.125		0.125	0.125
Mid								
Mid								1
Mid	0.125	0.125	0.125		0.125		0.125	0.125
Mid								
Mid				0.25			0.25	0.25
Mid	0.2	0.2			0.2	0.2		
Mid	0.25	0.25			0.25			
Mid	0.2				0.2			

3.3 Final Phase

The final stage concludes the domino effect as it is the last stage right before the individual commit's act of violence. Almost identical to the Mid Phase emotion of Fear and Anger were the most highlighting emotions with sadness and anticipation having significant influence. The final phase projects least number of emotions almost surfacing nonchalant behavior when compared to middle and early phase.

Table 7: Final Phase Behavioral Data

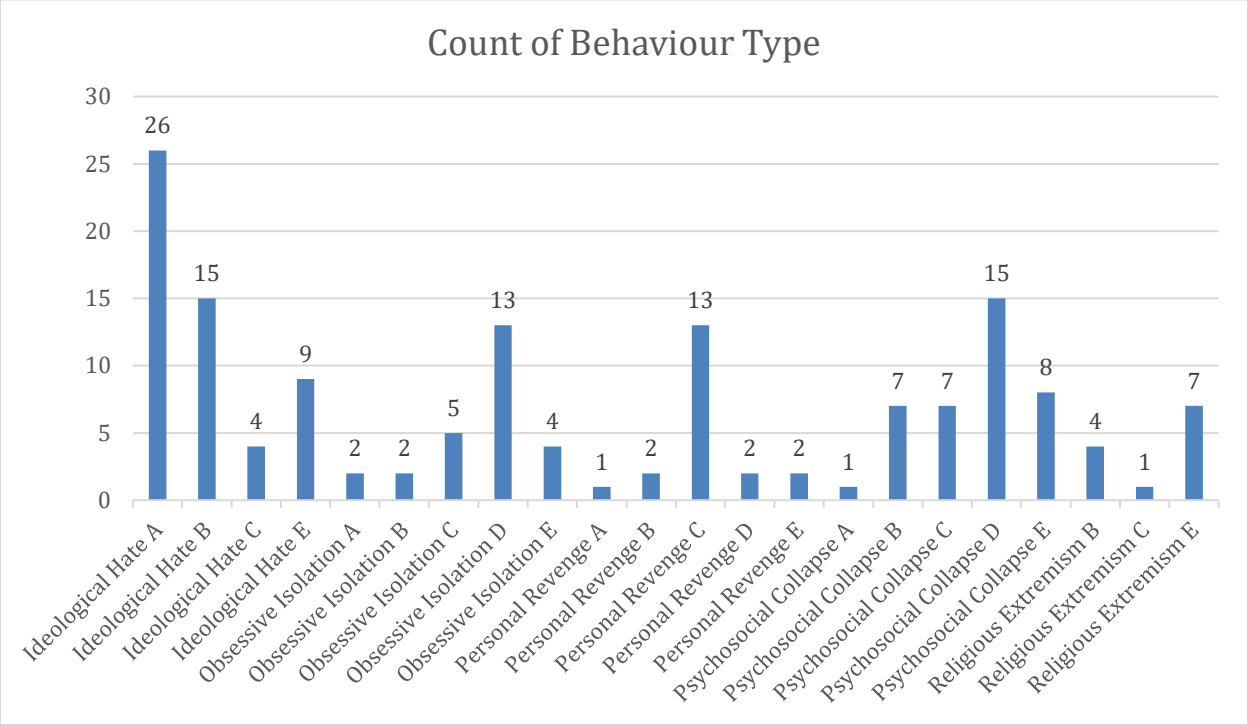
Phase	Fear	Anger	Trust	Surprise	Sadness	Disgust	Joy	Anticipation
Final	0.333333	0.333333						
Final								
Final								1
Final	0.333333	0.333333						
Final	0.25	0.25						
Final	0.333333	0.333333						
Final	0.222222				0.222222			
Final	0.166667	0.166667						0.166666667
Final								
Final	0.333333							
Final	0.333333				0.333333			
Final	0.5							
Final		0.222222				0.222222		
Final	0.2							
Final	0.2	0.2			0.2			0.2
Final		0.25						
Final								
Final								
Final	0.333333							0.333333333
Final					0.5			
Final					0.4			
Final			0.5					0.5
Final	0.333333	0.333333						
Final					0.5			
Final	0.4							
Final	0.333333				0.333333			
Final			0.4					
Final							0.5	
Final	0.2		0.2		0.2			
Final			0.25		0.25			
Final								
Final	0.333333		0.333333					
Final			0.2	0.2			0.2	0.2
Final								1
Final						0.5		

4. Comparative Cross-Case Analysis based on Behavioral Type:

Alongside emotional analysis behavior type analysis was also done for cross comparison and to identify similarities between individuals and their psychological state behavior type for based on 5 categories which were **Ideological Hate**, **Personal Revenge**, **Religious Extremism**, **Psychological Collapse** and **Obsessive Isolation**. These five behavioral types we are allocated

based on the intensity of the emotion and the tone of the Text/Statement from the individuals. This comparative cross-case analysis was done using NLP (Natural Language Processing) and human bias. Hence it serves as a hybrid method for accurately predicting and analyzing the data eliminating both machine and human error. These individuals were later assigned group IDs ranging from A to E based of intensity of behavioral type From early to final stage of each individual. Individuals who showed similar pattern of behavioral type were assigned same group IDs whereas those individuals who showed variation were assigned different group IDs. The group IDs were entirely based on human judgment. This maintained a balance between the hybrid analysis encompassing both AI driven approach and human judgment.

Table 8: Comparative Cross-Case Results by Ideological Motivation



5. Anomalies and Interpretive Insights

Apart from systematic and pattern oriented results certain individuals tend to show hostile behavior which often result into anomalies and hence need to be categorized separately. Similar is the case with Stephen Paddock, He tends to show little to no emotion from early to final stage comprising mostly of obsessive isolation behavior type with a slight psychological collapse in the middle and final stage. Individuals like Stephen Paddock are almost impossible to identify as they show no signs of violent behavior like other individuals. Hence this rules out as an exception in the analysis. This anomaly was identified using NLP (Natural Language Processing) along with human judgment to maintain the credibility.

6. Practical Implications and Theoretical Reflections:

The application of Artificial Intelligence specifically Natural Language Processing (NLP), In the forensic analysis of violent offenders for behavior presents a significant opportunity for both academic inquiry and real world implementation. This study integrates sentiment analysis and emotional profiling across temporal phases which include early mid and final of 10 high profile offenders offering a unique insight into the effective progression that may precede violent behavior. These findings carry profound implication for both digital forensic practices and behavioral threat assessment models.

One of the most prominent practical outcome of this research is the identification of emotion convergence patterns in final phase communication. Despite differing ideological religious or personal motives most offenders demonstrated a sharp emotional shift towards anger and fear in the final stage of their communicational timeline. Pattern supports the viability of emotion-based monitoring system wherein elevated markers of hostility resentment and dehumanization are detected through social media or manifesto analysis and may serve as early warning signs. Systems when integrated into forensic or law enforcement infrastructure can aid in the trailing of digital threats before escalation occurs. Moreover the utility of this emotional mapping extends to the development of the risk assessment. By categorization of offenders into behavioral type such as ideological hate, personal revenge, or psychological collapse which are based on emotional and sentiment trajectories digital forensic analyst can construct emotionally informed behavioral profiles. These profiles can guide further investigation support pattern matching for unknown subject analysis and inform psychological evaluations post incident. Theoretically speaking the study resonates with Alfred Adler's framework of individual psychology which portrays that individuals strive to overcome perceived inferiority through a compensation mechanism. Offenders like Elliot Rodger and Nicholas Cruz exemplify this progression from emotional state of sadness and rejection in early phase to anger and dehumanization contempt in the final declarations. This transition in emotional change reflects Adlerian compensation Which resembles a movement from psychological helplessness to aggressive overcompensation manifested in the form of both language and emotional expression. The emotional trajectory observed in such cases support the argument that radical behavior transformation is often emotionally rooted and not merely based on ideologies.

In addition the emotional evolution Which was uncovered in this research aligns with established radicalization framework such as the NYPD's 4 stage radicalisation model. This model outlines the progression of reradicalization to self-identification indoctrination and finally action. This progression of emotional tone in this study closely mirrors these stages by moving from emotional ambiguity such as fear and sadness to fix hostility like anger and fear in final communication. Such alignments lends the empirical support Two psychological and law enforcement models.

Nevertheless this study also uncovers the inherent limitations of AI based analysis when applied in isolation. Not all offenders display detachable emotional patterns in the case of Stephen Paddock

who demonstrated minimal emotional expression across all phases reveals a critical blind spot. His emotional sterility in communication defied both algorithmic classification and traditional behavioral analysis. This anomaly highlighted the need for contextual, human led interpretations by emphasizing that emotionally flat digital expression may still mask operative intent. Therefore AI models must function within a hybrid analysis environment by levering computational efficiency while preserving the nuance and interpretive depth offered by trained human analyst. Furthermore the presence of anomalies challenges the assumption that all violent behavior is preceded by emotional escalation. Some individuals may support emotional output deliberately or may be their communication style tend fall outside Standard lexicon based emotion classifiers. In such instances NLP must be supported with qualitative content analysis, behavioral cues and psychological background to ensure accuracy in thread detection. Finally this research reinforces the growing need for interdisciplinary collaboration. The convergence of forensic science, computer science behavioral psychology and technology is not just advantages but necessary in digital era where pre incident indicators are increasingly embedded in online behavior. Emotional and sentiment analysis when ethically implemented can serve as a critical component in prevention of violence, digital forensics and profiling of the offender.

Conclusion & Recommendations

1. Summary of key findings

This research aimed to explore the emotional and investigative evolution of violent offenders through the lens of artificial intelligence by focusing on sentiment analysis and emotion profiling across multiple behavioral phases. By applying natural language processing NLP tools to social media posts manifestos and public statements of 10 high profile mass violence perpetrator the study revealed significant emotional change of communication most notably a convergence towards anger and fear in final phase of communication.

The findings demonstrated that regardless of motive be ideological, personal or religious most offenders follow a consistent emotional trajectory early phase texts often contain signs of Fear and Sadness which later on transitioned into Anger Trust and Anticipation as the individuals approached the bridging point between the early stage and point of action. At the final stage these emotions converged to a sense of sadness and anticipation with fear and anger being the primary source of motive. Offenders like Elliott Rodgers and Nicholas Cruz reflected Adlerian model behavior i.e., transforming personal inferiority and rejection into acts of aggressive assertion. Meanwhile ideological actions such as Brenton Tarrant and Peyton Gendron display emotionally hardened narratives with final phase communication characterized by emotional detachment and justification for violence.

One of the most unique contributions of the study was the application of hybrid analysis by merging AI generated outputs with human forensic judgment. Natural language Processing (NLP) enabled the rapid detection of emotional markers it was the manual classification of phases and

behavior types that allowed for meaningful pattern recognition. This approach validates AI's rule as a tool of augmentation rather than replacement in forensic behavioral analysis. Furthermore, the identification of anomalies such as Stephen Paddock who exhibited minimal emotional signals has highlighted the need for caution in relying solely on linguistic profiling. Emotional neutrality does not equate to have no threat at all, As some individuals may operate outside recognizable radicalization trajectories. This finding reinforces the importance of human in the system in forensic AI framework.

2. Limitations of the study

While this research presents a meaningful insight into the emotional and linguistic progression of violent offenders using AI driven analysis it is important to acknowledge that the inherent limitation which shapes the scope and interpretation of findings stays alongside. Such limitations are discussed below:

1. **Data set and Diversity:** The study analyzed a total of 10 high profile offenders which while were sufficient for a focused case study approach limits the generalization of the result. The selection was based on the availability of public digital content and inherently excludes individuals whose communication was not archived or made public. Additionally, all subjects were male and the majority were from Western context hence it might have not reflected emotional trajectories in different socio-culture or gender specific environments.
2. **Variation in Source Type and Content Volume:** Not all offenders had uniform types or qualities of data. Individuals such as Brendan Tarrant had extensive manifesto with thousands of words while others had fragments or minimal digital footprints. This imbalance may have influenced the depth of emotional and sentiment analysis for each case. The reliance on different platforms introduced a variation in tone structure and linguistic intent which may affect the accuracy of comparative analysis. As different social media platforms have different policies individual might prefer different linguistic approach for different platforms.
3. **Manual Phase Classification:** Although behavioral phases like early mid and final were defined based on temporal arithmetic progression but in the end this categorization was conducted manually and relied heavily on subjective judgment. Despite efforts to standardize criteria human interpretation may introduce bias particularly in ambiguous or contextless entries. Such a scenario can be observed in the case of Stephen Paddock.
4. **Emotion Detection in Neutral or Deceptive Language:** Cases like Stephen Paddock revealed a critical limitation as an offender who exhibit emotionally neutral or linguistically sterile communication which Might have evaded detection by sentiment and emotional classifiers. Spaces meaningful behavioral indicators may be entirely absent from language highlighting the limits of digital forensics in profiling non expressive individuals.
5. **Ethical and Interpretive Risk:** The application of AI in behavioral analysis carries ethical risk particularly when used in predictive or preemptive policing contexts. Misclassification

or over reliance on AI scores without human judgment could lead to false positive, privacy violation or stigmatization of individuals based on ambiguous or out of context statement.

3. Recommendations for future research

Based on the findings and limitations of this study the following recommendations are proposed for further research and practical applications:

1. **Integration of NLP into threat assessment system:** Emotion and sentiment tracking tools should be incorporated into digital forensic monitoring system used by law enforcement and intelligence agencies. Boom can aid in prioritizing digital threats especially when patterns of emotional escalation are detected
2. **Development of emotional risk typologies:** Further studies should refine the standardized behavior types based on emotional trajectories such as ideological hate personal revenge and psychological collapse hence assisting in threat categorization profile.
3. **Expansion of data set Diversity:** Further research should involve a broader data set including lesser-known cases female offenders an ethnically diversified dataset to validate emotional pattern across different cultures and demographic environments.
4. **Cross validation with psychological reports:** He had detected emotional patterns should be cross referred with clinically evaluated and psychological autopsies to validate emotional indicators and improve predictive accuracy.
5. **Ethical oversight and data sensitivity:** Institutions applying such technologies must ensure ethical standards are upheld particularly concerning privacy, data consent and potential misuse of behavioral profiling system.

   The study lays the foundation for an emotional centric forensic profiling by offering bridge of computational linguistics, behavioral science and criminology. As digital communication continues to dominate the landscape of pre-incident behavior the importance and demand of tools that can interpret such emotional signals will grow. The integration of such tools into real world forensic processes represent not just a technological upgrade but a necessary evolution in the field of behavioral threat assessment.

References

1. Agarwal, A. and Sureka, A. (2015). Using Sentiment Analysis for Detecting Radicalization on Social Media. *Proceedings of the IEEE International Conference on Big Data*.

2. Bird, S., Klein, E. and Loper, E. (2009). *Natural Language Processing with Python*. O'Reilly Media.

3. Boyd, D.M. and Ellison, N.B. (2007). Social Network Sites: Definition, History, and Scholarship. *Journal of Computer-Mediated Communication*, 13(1), pp.210–230.

4. Cambria, E., Schuller, B., Xia, Y. and Havasi, C. (2013). New Avenues in Opinion Mining and Sentiment Analysis. *IEEE Intelligent Systems*, 28(2), pp.15–21.

5. Casey, E. (2001). *Digital Evidence and Computer Crime: Forensic Science, Computers, and the Internet*. Academic Press.

6. Casey, E. (2011). *Digital Evidence and Computer Crime: Forensic Science, Computers, and the Internet*. Academic Press.

7. Chakraborty, M., Gupta, D. and Sharma, D. (2021). Natural Language Processing in Forensics. *Journal of Digital Forensics*, 8(2), pp.34–48.

8. Chesney, R. and Citron, D.K. (2019). Deep Fakes: A Looming Challenge for Privacy, Democracy, and National Security. *California Law Review*, 107(6), pp.1753–1820.

9. Collins, A. and Ebrahimi, T. (2021). Forged Authenticity: Governing Deepfake Risks. *International Risk Governance Center, EPFL*. Available at: https://www.epfl.ch/research/domains/irgc/specific-risk-domains/projects-cybersecurity/forging-authenticity-governing-deepfake-risks/ [Accessed 24 Apr. 2025].

10. DARPA. (2020). *Media Forensics (MediFor)*. Defense Advanced Research Projects Agency. Available at: https://www.darpa.mil/program/media-forensics [Accessed 24 Apr. 2025].

11. European Commission. (2024). *Tackling Disinformation: Europe's Digital Future*. Available at: https://digital-strategy.ec.europa.eu/en [Accessed 24 Apr. 2025].

12. Farid, H. (2009). Exposing Digital Forgeries in Scientific Images. *Science*, 324(5928), pp.366–367.

13. Federal Bureau of Investigation. (2012). *Cyber Crime: A Growing Threat*. FBI.gov. Available at: https://www.fbi.gov/news/stories/cyber-crime-a-growing-threat [Accessed 24 Apr. 2025].

14. Ferguson, R. (2020). *Predictive Analytics: Crime Forecasting in the Digital Age*. Oxford University Press.

15. Ferrara, E., Varol, O., Davis, C., Menczer, F. and Flammini, A. (2016). The Rise of Social Bots. *Communications of the ACM*, 59(7), pp.96–104.

16. Hill, K. (2020). Facial Recognition Technology and its Impact on Forensic Science. *Journal of Digital Forensics*.

17. Hutto, C.J. and Gilbert, E. (2014). VADER: A Parsimonious Rule-based Model for Sentiment Analysis of Social Media Text. *Proceedings of the Eighth International AAAI Conference on Weblogs and Social Media*.

18. Jurafsky, D. and Martin, J.H. (2020). *Speech and Language Processing*. 3rd ed. Pearson.

19. Kaplan, A. and Haenlein, M. (2019). Siri, Alexa, and Other Digital Assistants: The Future of AI in Society. *Business Horizons*, 62(1), pp.15–25.

20. Manning, C.D., Raghavan, P. and Schütze, H. (2008). *Introduction to Information Retrieval*. Cambridge University Press.

21. Meloy, J.R., Hoffman, J., Guldimann, A. and James, D. (2012). The Role of Warning Behaviors in Threat Assessment: An Exploration and Suggested Typology. *Behavioral Sciences & the Law*, 30(3), pp.256–279.

22. Ministry of Electronics and IT. (2021). *The Information Technology (Intermediary Guidelines and Digital Media Ethics Code) Rules*. Government of India. Available at: https://www.meity.gov.in/content/notification-dated-25th-february-2021-gsr-139e-information-technology-intermediary/ [Accessed 24 Apr. 2025].

23. Monroe, B.L., Colaresi, M. and Quinn, K.M. (2008). Fightin' Words: Lexical Feature Selection and Evaluation for Identifying the Content of Political Conflict. *Political Analysis*, 16(4), pp.372–403.

24. National Security Agency, Federal Bureau of Investigation and Cybersecurity and Infrastructure Security Agency. (2023). *Contextualizing Deepfake Threats to Organizations*. Available at: https://media.defense.gov/2023/Sep/12/2003298925/-1/-1/0/CSI-DEEPFAKE-THREATS.PDF [Accessed 24 Apr. 2025].

25. Pang, B. and Lee, L. (2008). Opinion Mining and Sentiment Analysis. *Foundations and Trends in Information Retrieval*, 2(1–2), pp.1–135.

26. Palmer, G. (2001). *A Road Map for Digital Forensic Research*. Technical Report DFRWS-01-01. Digital Forensic Research Workshop.

27. Press Information Bureau. (2023). *Union Government issues advisory to social media intermediaries to identify misinformation and deepfakes*. Ministry of Electronics & IT, Government of India. Available at: https://pib.gov.in/Pressreleaseshare.aspx?PRID=1975445 [Accessed 24 Apr. 2025].

28. Press Information Bureau. (2024). *Cyber fraud and digital harassment*. Ministry of Home Affairs, Government of India. Available at: https://pib.gov.in/PressReleaseIframePage.aspx?PRID=2080186 [Accessed 24 Apr. 2025].

29. Statista (n.d.). U.S. Deaths From Gun Violence and Terrorism Compared. [online] Available at: https://www.statista.com/chart/3845/us-deaths-from-gun-violence-and-terrorism-in-comparison/ [Accessed 24 Apr. 2025].

30. US Congress. (2019). *DEEPFAKES Accountability Act: Hearing before the Subcommittee on Communications and Technology*. House Committee on Energy and Commerce. Washington, DC: Government Publishing Office.

31. Weimann, G. (2015). *Terrorism in Cyberspace: The Next Generation*. Columbia University Press.

32. Zellers, R., Holtzman, A., Rashkin, H., Bisk, Y., Farhadi, A., Roesner, F. and Choi, Y. (2019). Defending Against Neural Fake News. *NeurIPS Proceedings*.

33. Pang, B. and Lee, L. (2008). Opinion Mining and Sentiment Analysis. Foundations and Trends in Information Retrieval, 2(1–2), pp.1–135.

34. Chakraborty, M., Gupta, D. and Sharma, D. (2021). Natural Language Processing in Forensics. Journal of Digital Forensics, 8(2), pp.34–48.

35. Agarwal, A. and Sureka, A. (2015). Using Sentiment Analysis for Detecting Radicalization on Social Media. Proceedings of the IEEE International Conference on Big Data.

36. Cambria, E., Schuller, B., Xia, Y. and Havasi, C. (2013). New Avenues in Opinion Mining and Sentiment Analysis. IEEE Intelligent Systems, 28(2), pp.15–21.

37. Monroe, B.L., Colaresi, M. and Quinn, K.M. (2008). Fightin' Words: Lexical Feature Selection and Evaluation for Identifying the Content of Political Conflict. Political Analysis, 16(4), pp.372–403.

Appendices

Appendix A: Violent Offender Dataset Overview

This appendix includes the full dataset used in the study, comprising 151 entries across 10 offenders. Each entry contains text excerpts, sentiment scores, emotion values, behavioral phase labels, and source type. Due to formatting limitations, the dataset is provided as a separate Excel file Titled:

"Violent Offender Real Text Analysis Dataset.xlsx"

Violent Offender Real Text Analysis Dataset.

Appendix B: NLP Code for Sentiment & Emotion Analysis

The following Python code was used in Google Colab to perform sentiment scoring, emotion extraction, and semantic similarity analysis. The code utilizes libraries including pandas, transformers, nltk, and sklearn.

Sentiment_Analysis_NLP.ipynb

Forensic Analysis of Social Media Related to Terrorism and Gun Violence Using AI.pdf

ORIGINALITY REPORT

6%	5%	3%	2%
SIMILARITY INDEX	INTERNET SOURCES	PUBLICATIONS	STUDENT PAPERS

PRIMARY SOURCES

1	www.europarl.europa.eu Internet Source	<1%
2	Submitted to National Forensic Sciences University Student Paper	<1%
3	listens.online Internet Source	<1%
4	Karwan Mustafa Kareem. "The Intelligence Technology and Big Eye Secrets: Navigating the Complex World of Cybersecurity and Espionage", PsyArXiv, 2024 Publication	<1%
5	v.vibdoc.com Internet Source	<1%
6	papers.academic-conferences.org Internet Source	<1%
7	"Challenges and Opportunities in the Artificial Intelligence Era", Springer Science and Business Media LLC, 2025 Publication	<1%
8	ijlsi.com Internet Source	<1%
9	Submitted to CSU, Fullerton Student Paper	<1%
10	ai-magazine.com Internet Source	<1%