UNIVERSITY PARTNER

UNIVERSITY OF WOLVERHAMPTON

HERALD COLLEGE KATHMANDU

Complex Systems (6CS014)

AI and Complex Systems in Criminal Justice

Student Name : Nayan Raj Khanal

Student Number : 2227486

Module Leader : Mr. Deepson Shrestha

Tutor : Mr. Akash Adhikari

Submission Date : 20th January, 2024

Word Count : 2332

ACKNOWLEDGEMNT

I extended my sincere gratitude to all those involved for this report. I would like to thank my instructor Mr. Akash Adhikari for their guidance and support throughout this task. I would also like to thank Mr. Deepson Shrestha our module leader and the team of Complex System for their insights and assistance. I would like to extend my thanks to the researchers and authors whose valuable work helped me finish writing this report. Lastly, I would like to thank my friends and family for pushing me to complete this task.

ABSTRACT

In this report, we investigate the current state of complex and artificial intelligence (AI) implementation in the Criminal Justice system, with a specific focus on predictive policing. After an extensive literature review, this report covers how complex systems and AI have the potential to improve law enforcement's crime prediction capabilities and aid in suspect identification. While advanced AI models have shown positive results, there are still certain challenges that need to be addressed, such as data quality issues, heavy system resources, and transparency concerns. It is important to continue the research to strike a balance between innovation and responsible implementation, which is vital for the effective and ethical use of AI in law enforcement. Addressing these limitations is essential to maximize AI benefits while ensuring ethical practices in the Criminal Justice system.

TABLE OF CONTENTS

TABLE OF FIGURES

1. INTRODUCTION

In today's world, several fields have seen massive technological advancement. Likewise, the use of AI in the field of Criminal Justice has also created a significant buzz of attention and debate in recent years. More and more law enforcement agencies are increasingly turning to predictive policing to enhance their arsenal when it comes to crime prevention and resource allocation. The integration of these technologies raises eyebrows for the general masses with concerns about morals, privacy, and the overall effectiveness of such approaches. By exploring the current papers and analyzing findings, this report aims to shed light on the current landscape and implications of AI in criminal justice.

2. AIM AND OBJECTIVES

2.1. AIM

The main aim of this report is to thoroughly investigate the implementation of AI and complex systems within the context of Criminal Justice, specifically in predictive policing, and then share the findings with the general masses so that they can draw their conclusions.

2.2. OBJECTIVES

- To investigate the current state of AI implementation in law enforcement.
- To assess the impact of predictive policing on crime prevention and resource allocation.
- To explore ethical challenges associated with AI in Criminal Justice.
- To analyze the effectiveness of AI in enhancing overall criminal justice outcomes.

3. LITERATURE REVIEW

The paper “Research on Anti-terrorism Intelligence Mining Method Based on Attention Neural Networks” follows a comprehensive methodology by utilizing two powerful models: an autoencoder (AE) network and a Long Short-Term Memory (LSTM) network. Multiple stages are involved in this research: data restoration, Word2Vec encoding, and feature extraction using an AE network. Data reconstruction and feature extraction are handled by the autoencoder network while the analysis of sequential data is performed by the LSTM network. The unique combination of the AE and LSTM network is the core of this methodology and allows for precise prediction of sequences related to terrorist organizations and casualties within the dataset. The results show that the AE-LSTM network model outshines traditional models like SVM, MLP, and logistic regression when it comes to predicting terrorist organizations and casualties. The reason for the improvement is its enhanced training iterations which refine its predictive capabilities. The reliability of the model fosters confidence in its accuracy. While this model is quite skilled, its main problems come from needing lots of information to learn and needing a powerful computer. These difficulties could make it tricky to use in real- life situations where there isn't much data or the computers aren't strong enough. (Bai, et al., 2020)

Algorithm	Accuracy of casualty prediction	Accuracy in predicting potential terrorist groups
The SVM	0.75	0.80
Multilayer perceptron	0.56	0.76
Logistic regression	0.78	0.64
The ATTENtion-free LSTM network	0.89	0.85
AE-LSTM network	0.90	0.91

Figure 1. Comparison of Traditional and Proposed Algorithms (Bai, et al., 2020)

The paper “A Machine Learning and Deep Learning Integrated Model to Detect Criminal Activities” utilizes a combination of machine learning and deep learning techniques, including Long-term Recurrent Convolutional Network (LRCN) for criminal activity detection, YOLOv7 for weapon detection, dlib's HOG+Linear SVM for criminal prediction, and TextBlob for sentiment analysis. Each module has its specific task and uses different datasets and algorithms. To evaluate the system's performance, the datasets are preprocessed, the models are trained, and metrics like accuracy, F1-score, and confusion matrices are used. The results of the study show that the system is effective. The criminalactivity detection module has an accuracy of 84%, the weapon detection module has an accuracy of 87%, and the criminal recognitionmodule has an impressive accuracy rate of 99.38%. The sentiment analysis from social media provides important information on potential criminal activity. These results demonstrate the system's capability to detect crime and identify suspects in real-time. The results were great but the study also identified some limitations. While some modules have high accuracies, others can be improved and it needs to improve the depth of social media analysis and use geolocation-based hotspot predictions to enhance its predictive capabilities. (Arpa, et al., 2023)

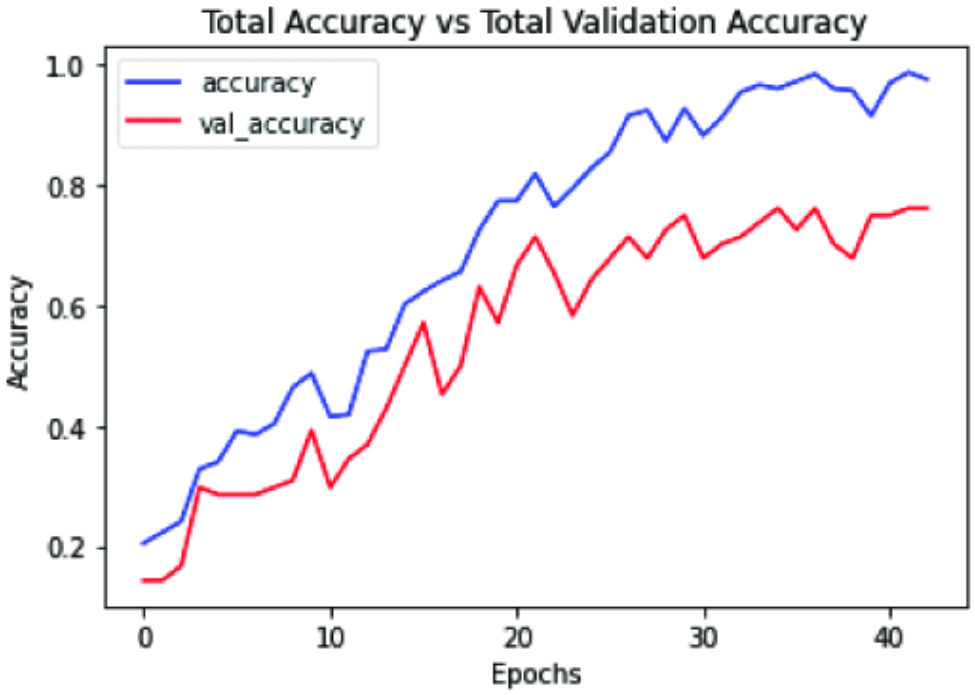

Figure 2. Total Accuracy VS Total Validation Accuracy (Criminal Activity Detection) (Arpa, et al., 2023)

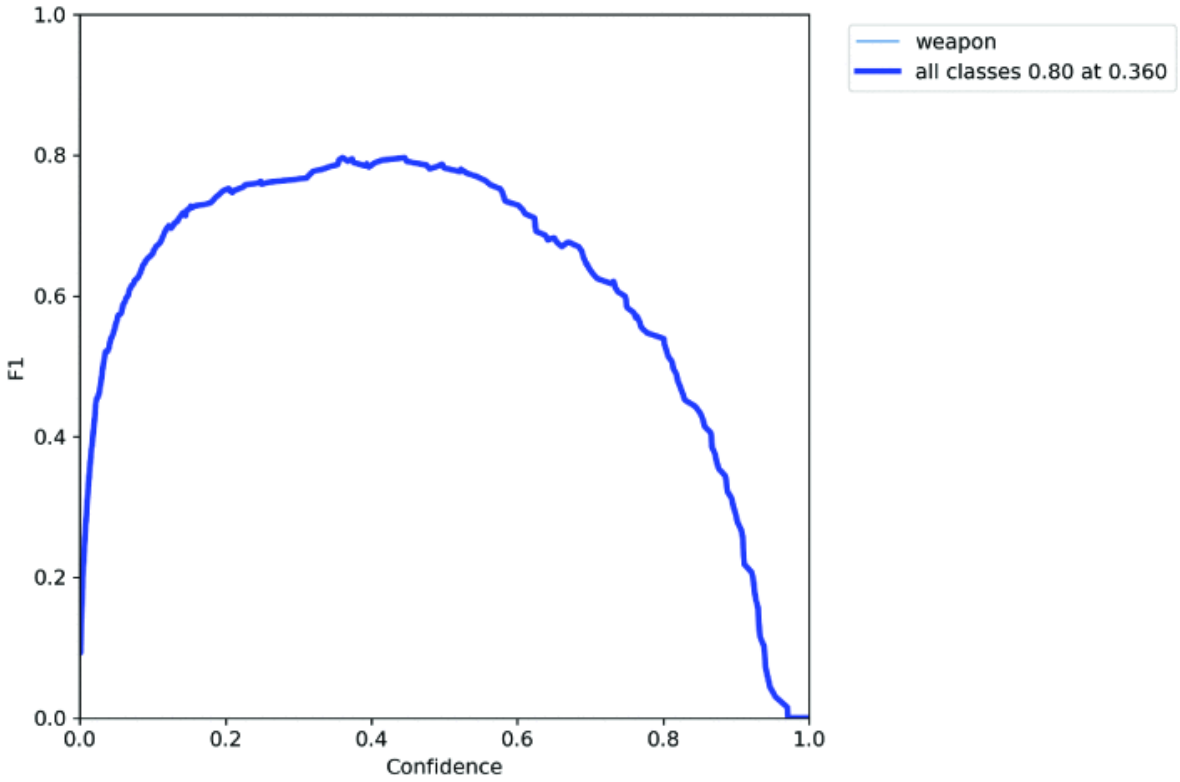

Figure 3. F1 Score Curve for a Custom YOLOv7 Weapon Detection (Arpa, et al., 2023)

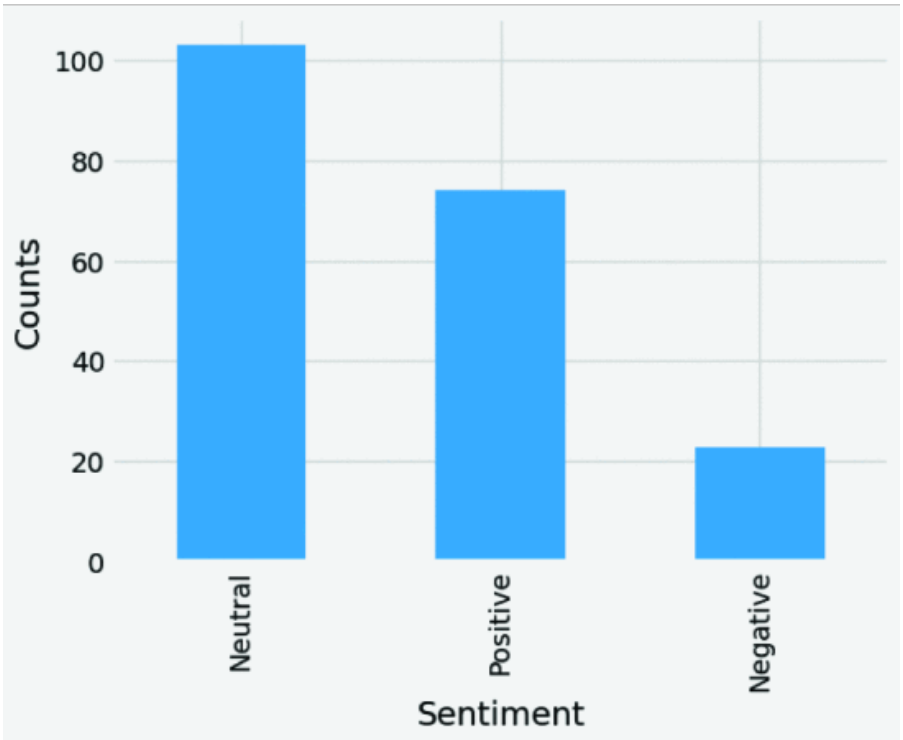

Figure 4. Bar Chart of Sentiment (Arpa, et al., 2023)

The paper “The Power of Predictive Analytics: Forecasting Crime Trends in High-Risk Areas for Crime Prevention using Machine Learning” analyzed crime data from Chicago to create predictive models for crime count and type that may occur in specific areas within a city. The study used various machine learning algorithms, such as Random Forest, Naive Bayes, Decision Tree, K Nearest Neighbor, Logistic Regression, and Neural Integrated Learning. The model produced positive results, with high accuracy in predicting both crime count and type. Initially, traditional machine learning algorithms were used, with the Random Forest algorithm performing very well in predicting both crime count and type achieving the highest accuracy of 99.70% and 90.11% respectively. However, after integrating Neural Networks into the models, there were notable improvements in accuracy across all algorithms, especially in random forest producing the best accuracy of 99.78% and 91.01% respectively. The paper does point out one significant defect, which was missing data in the dataset, especially in the early years of records. Even though this missing data was handled by careful procedure during the preparation phase, the paper's accuracy in predicting crime count and type may have been influenced by the limitations within the dataset. This affects the applicability of the findings for further work. (Saini & Kaur, 2023)

S. No.	Accuracy	Precision	Recall	F1 Score
Naïve Bayes	0.9952	0.9934	0.9952	0.9943
NINB	0.9977	0.9968	0.9977	0.9968
Logistic Regression	0.9956	0.9933	0.9956	0.9945
NILR	0.9977	0.9968	0.9977	0.9968
Decision Tree	0.9956	0.9933	0.9956	0.9945
NIDT	0.9977	0.9968	0.9977	0.9968
Random Forest	0.9970	0.9969	0.9970	0.9969
NIRF	0.9978	0.9979	0.9977	0.9979
K Nearest Neighbor	0.9955	0.9942	0.9955	0.9946
NIKNN	0.9977	0.9968	0.9977	0.9968

Figure 5. Comparative analysis for Predicting Number of Crimes (Saini & Kaur, 2023)

S. No.	Accuracy	Precision	Recall	F1 Score
Naïve Bayes	0.8331	0.8098	0.8331	0.8097
NINB	0.8844	0.8829	0.8715	0.8770
Logistic Regression	0.4736	0.3645	0.4736	0.3938
NILR	0.8881	0.8881	0.8878	0.8879
Decision Tree	0.7174	0.5898	0.7174	0.6344
NIDT	0.8125	0.8459	0.7518	0.7932
Random Forest	0.9011	0.8881	0.9011	0.8882
NIRF	0.9101	0.9003	0.9134	0.8990
K Nearest Neighbor	0.6220	0.6084	0.6220	0.6074
NIKNN	0.8848	0.8864	0.8838	0.8851

Figure 6. Comparative analysis for Predicting Crime Type (Saini & Kaur, 2023)

The paper “Predicting High-Risk Areas for Crime Hotspot Using Hybrid KNN Machine Learning Framework” involves training the KNN algorithm with historical crime data from cities, such as Boston. The study considers social, environmental, and geographic factors that impact crime frequency and examines spatial and temporal patterns. To evaluate different crime mapping and machine learning techniques, the study utilizes Friedman's and Nemenyi's tests. The study found that the KNN algorithm is effective in identifying crime hotspots and can detect distinct temporal patterns in crime scenarios. The study also found that the accuracy of the model is influenced by the number of samples used and that the algorithm is better than traditional methods. The study used feature importance analysis to identify key factors that contribute to accurate predictions. Despite the accuracy of the KNN algorithm for predicting crime hotspots, the paper points out some limitations. Its accuracy is heavily dependent on the quality of training data and the algorithm's ability to identify relevant patterns. Another issue is the curse of dimensionality, which means that the more features present in the data, the less effective the algorithm becomes. (K, et al., 2023)

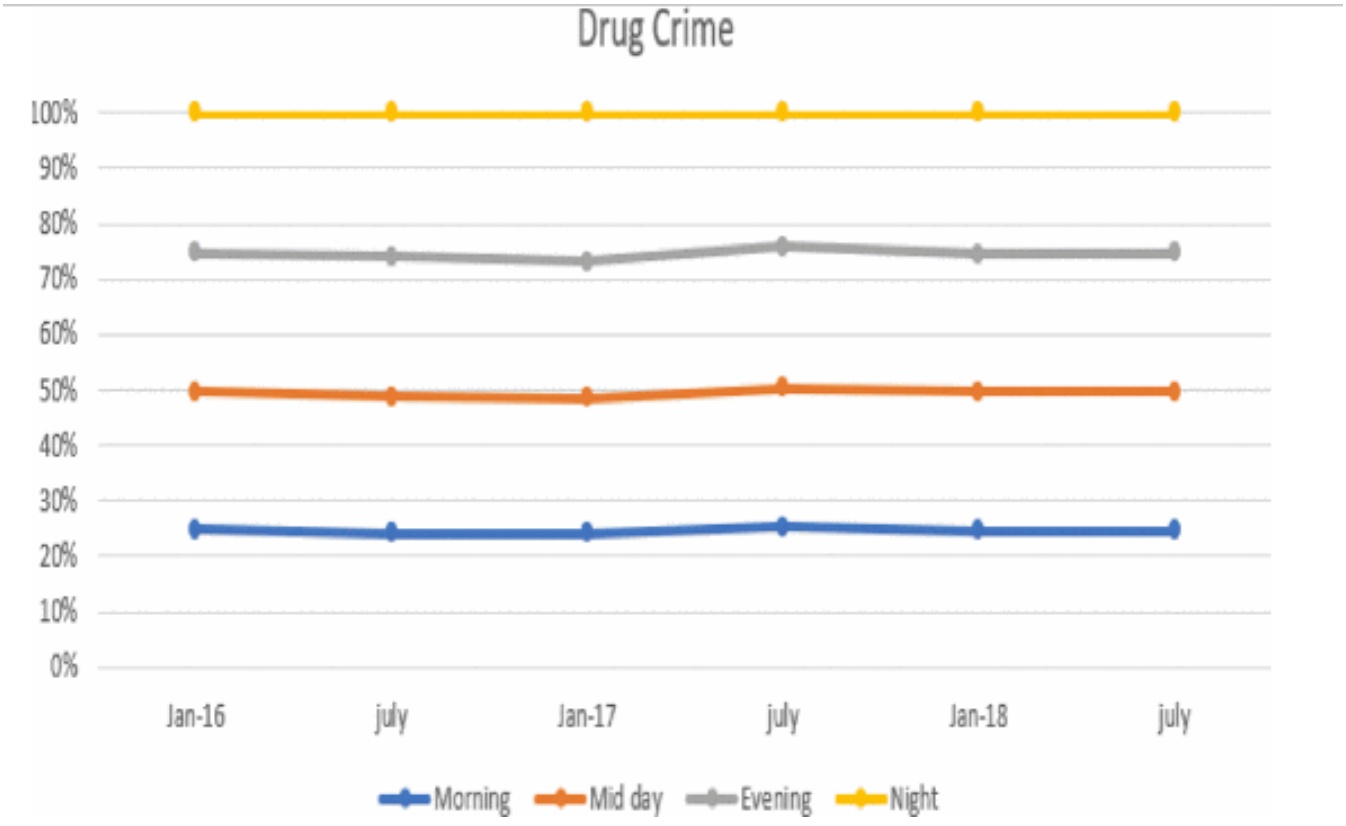

Figure 7. Comparison of time series using KNN algorithm (K, et al., 2023)

The paper “Crime Classification Using Machine Learning and Data Analytic” aimed to predict crime by analyzing data. The methodology used included data visualization, outlier handling, and feature extraction. The main goal was to solve class imbalance and dependency issues. The study used the Chicago Crime dataset to identify spatial and temporal trends. To handle outliers, mean substitution was used. Feature extraction involved clustering, temporal analysis, and district division. Machine learning algorithms like Naive Bayes, Decision Trees, and Random Forests were applied to predict the type of crime. The results showed that the Random Forest Classifier was the most accurate, with a 55.03% accuracy rate in predicting the top three crime categories. The Naive Bayes Classifier achieved an accuracy of 53.33%, while the Decision Tree Classifier reached an accuracy of 51.35%. However, predicting the top 10 crime categories was less accurate due to an imbalance in the number of cases. The study found that limited features and high dependence among them negatively affected the algorithms' performance in predicting crime types. The study indicates the main cause of low prediction accuracy resulting in limitations is due to the limited feature set. To improve the predictions, more features need to be extracted, including education standards, poverty index, traffic, weather, and financial data that contribute to crime in the community. (Haider, et al., 2022)

	Naive Bayes Classifier			**Decision Tree Classifier**			**Random Forest Classifier**		
	BATTERY	CRIMINAL DAMAGE	THEFT	BATTERY	CRIMINAL DAMAGE	THEFT	BATTERY	CRIMINAL DAMAGE	THEFT
Sensitivity	0.6746	0.0885	0.5664	0.6458	0.0036	0.6699	0.5874	0.0244	0.6786
Specificity	0.5562	0.9432	0.7165	0.5950	0.9987	0.6261	0.6255	0.9852	0.5866
Precision	0.4607	0.2701	0.6187	0.4726	0.3929	0.5927	0.4685	0.2817	0.5714
F1-Score	0.5475	0.1333	0.5914	0.5458	0.0071	0.6289	0.5213	0.0448	0.6204

Figure 8. Detailed results of top 3 crime categories (Haider, et al., 2022)

The paper “An Approach to Criminal Suspect Prediction Software using Machine Learning Classifiers” used both supervised and unsupervised learning techniques. Supervised learning used labeled data to train the model, while unsupervised learning found patterns in unlabeled data. The process involved preparing the data, training the model, grouping similar data points, and predicting outcomes using machine learning techniques like Regression, K-means, and KNN algorithms. By analyzing recorded data attributes, this approach can identify potential suspects and crime patterns. The study showed promising outcomes, with a 90% accuracy rate in identifying suspects. This high accuracy suggests significant progress in predictive policing, demonstrating the potential effectiveness of AI-based models in identifying suspects and predicting crime hotspots. The results highlight the models' strong predictive power, indicating their potential practical application in law enforcement for crime prediction and prevention. The paper highlights issues with its crime prediction. It fails to explain the complex machine-learning techniques employed and provide details about the data used for predictions. The lack of information about the model's performance in real-life scenarios and its accuracy testing is another limitation. Fixing these issues could make the model more reliable and useful in a variety of settings. (Arya, et al., 2023)

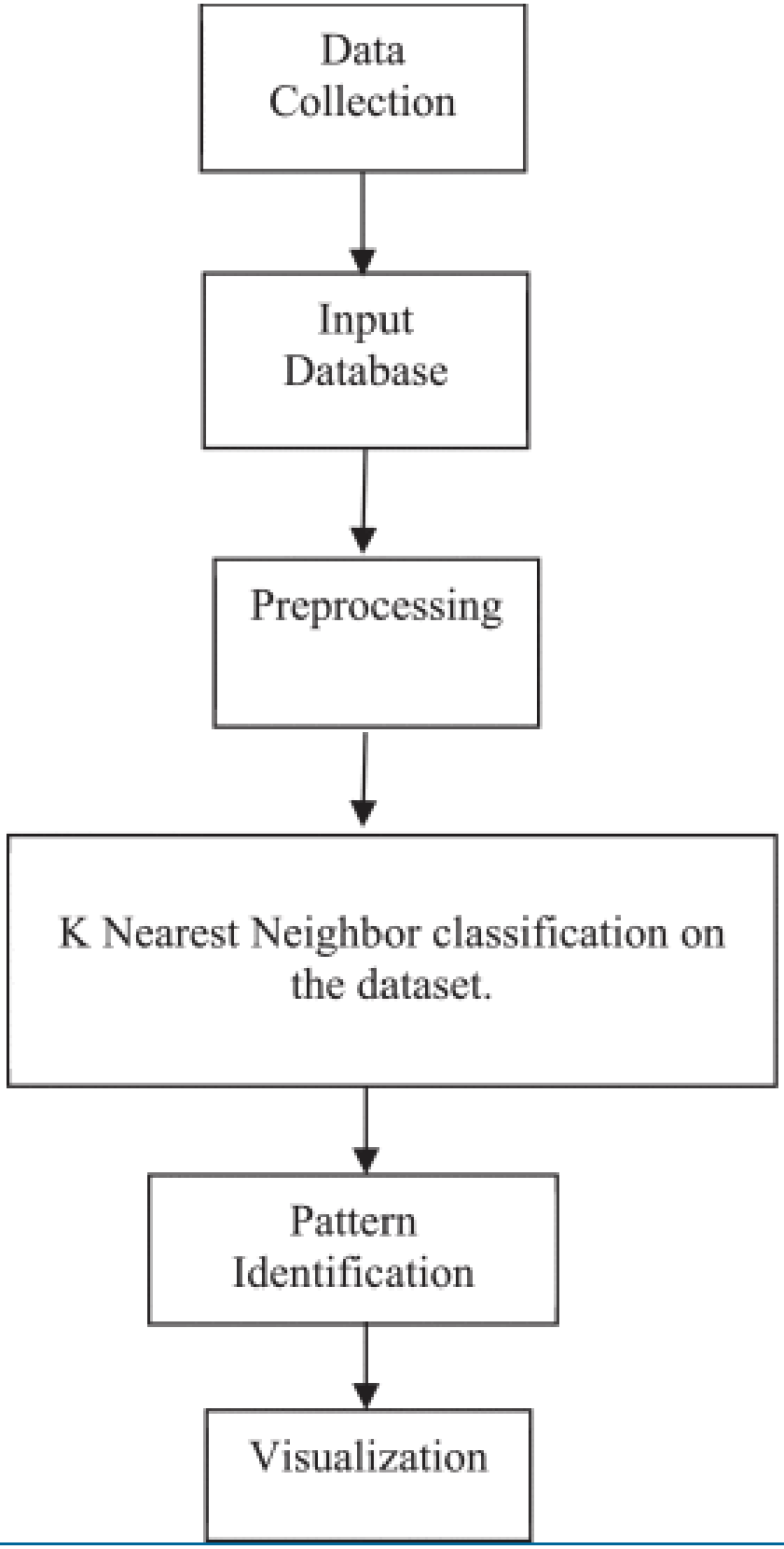

Figure 9. Flow Chart of the Prediction of Suspects (Arya, et al., 2023)

The paper “Crime Forecasting using Interpretable Regression Techniques” includes gathering data from diverse communities in the US including demographic data, crime rates, and unique features of each locality. The data for burglary rates were carefully analyzed by handling missing values and selecting significant factors. Different regression models are used to ensure accurate predictions and advanced techniques such as SHAP XAI, and LIME XAI are applied to explain the decision-making process. The study successfully predicted burglary rates using different regression models, with an average predictive accuracy of 94.0%. The Linear Regression model proved to be the most efficient. The analysis provided insights into the main factors contributing to burglary, improving understanding of crime dynamics. The report mentions some limitations. It only focuses on burglary patterns and doesn't consider other crime patterns. It excluded some community features from the analysis due to their abundance. Time-based crime analysis was not done, despite its importance and the report does not utilize Benford's law for better predictions. (David, et al., 2023)

Algorithm	Accuracy	Cross Validation 10-Fold:	Running time
Linear Regression	94.0%	97.96%	0:00:00:048737
LASSO Regression	94.0%	88.47%	0:00:00:181487
Ridge Regression	94.0%	97.96%	0:00:00:053866

Figure 10. Table of performance for the different models (David, et al., 2023)

4. ANALYSIS

Upon analyzing the paper “Research on Anti-terrorism Intelligence Mining Method Based on Attention Neural Networks”, the AE-LSTM network, which uses attention mechanisms, is promising for improving anti-terrorism intelligence mining. It is better than traditional models and could be useful in real-world situations. However, it needs a lot of data and computing power, which means it needs to be optimized to work well in places where resources are limited. (Bai, et al., 2020)

Upon analyzing the paper “A Machine Learning and Deep Learning Integrated Model to Detect Criminal Activities”, it highlights the potential for artificial intelligence to transform crime prevention. The system's accuracy in detecting criminal activities, weapons, and suspects is highlighted and it is high, potentially ready for real-world testing. The study also acknowledges the limitations and outlines future directions for improvements. (Arpa, et al., 2023)

Upon analyzing the paper “The Power of Predictive Analytics: Forecasting Crime Trends in High-Risk Areas for Crime Prevention using Machine Learning”, it introduces a promising method to improve crime predictive models by utilizing both traditional algorithms and Neural Networks. By combining these approaches, the paper has put forth a model that can predict crime count and type accurately. This can have a significant impact on other cities and datasets that aim to improve their predictive models. (Saini & Kaur, 2023)

Upon analyzing the paper “Predicting High-Risk Areas for Crime Hotspot Using Hybrid KNN Machine Learning Framework”, it shows that the hybrid KNN algorithm has potential in predicting crime hotspots, but also has limitations. It can help law enforcement understand crime patterns and create crime prevention strategies. However, there are concerns about the quality of the data, therefore, other machine learning models should be considered before real-world use for crime prediction methods (K, et al., 2023)

Upon analyzing the paper “Crime Classification Using Machine Learning and Data Analytic”, it discusses how machine learning algorithms have the potential to predict crimes, but face challenges such as limited features, class imbalance, and feature interdependence negatively affecting accuracy. To improve, it is recommended to extract more features, refine models, and have better crime reporting systems to collect accurate data. (Haider, et al., 2022)

Upon analyzing the paper “An Approach to Criminal Suspect Prediction Software using Machine Learning Classifiers”, it demonstrates that an AI-powered model can accurately identify suspects. However, due to a lack of details about the data on which it was tested, more research and real-world testing are needed to make sure these models work well in different situations before deploying them to law enforcement agencies. (Arya, et al., 2023)

Upon analyzing the paper “Crime Forecasting using Interpretable Regression Techniques”, it used regression models and interpretability techniques to predict crime. It was comprehensive but looking into the limitations the paper only focused on predicting burglaries and excluded numerous community-related features. This means it may not be useful for broader crime prevention strategies. To improve predictions, more data and diverse crime patterns need to be analyzed. (David, et al., 2023)

5. CONCLUSION

The main aim of the report is to show how complex systems and AI can benefit law enforcement in predictive policing. The literature review supports this aim by demonstrating the positive outcomes in crime prediction and suspect identification. However, the review also identifies some limitations such as data quality issues, optimization, and transparency concerns. Therefore, from the analysis, we can conclude that for an effective and ethical use of AI in law enforcement it is crucial to have a balance between innovation and responsible implementation. In summary, by addressing the limitations we can maximize the benefits of AI in the Criminal Justice system, therefore enhancing the workforce of law enforcement.

6. REFERENCES

Arpa, K. F. et al., 2023. *A Machine Learning and Deep Learning Integrated Model to Detect Criminal Activities.* Bangkok, IEEE.

Arya, A. et al., 2023. *An Approach to Criminal Suspect Prediction Software using Machine Learning Classifiers.* Greater Noida, IEEE.

Bai, C., Li, A., Gao, Z. & Cui, X., 2020. *Research on Anti-terrorism Intelligence Mining Method Based on Attention Neural Networks.* Weihai, IEEE.

David, M., Mbabazi, E. S., Nabende, J. N. & Marvin, G., 2023. *Crime Forecasting using Interpretable Regression Techniques.* Tirunelveli, IEEE.

Haider, M. N., Mumtaz, R. & Hassan Zaidi, S. M., 2022. *Crime Classification Using Machine Learning and Data Analytic.* Marietta, IEEE.

K, V. et al., 2023. *Predicting High-Risk Areas for Crime Hotspot Using Hybrid KNN Machine Learning Framework.* Coimbatore, IEEE.

Saini, I. S. & Kaur, N., 2023. *The Power of Predictive Analytics: Forecasting Crime Trends in High-Risk Areas for Crime Prevention using Machine Learning.* Delhi, IEEE.