Location Based Crime type Prediction using PyTorch Non-Linear Deep-2-Layer Model with Stramlit WebApp

Siva Prasad Reddy Bandi[1] $sivaprasadreddy371@gmail.com$
Thummalapalli Nikhil Chowdrary[2] $tnikhilchowdary@gmail.com$,
K Sai Kiran[3] $agirisaikiran@gmail.com$

Student, Department of Computer Science and Engineering, Vel Tech Rangarajan Dr.Sagunthala RD Institute of Science and Technology , Chennai, Tamil Nadu, India[1]

Student, Department of Computer Science and Engineering, Vel Tech Rangarajan Dr.Sagunthala RD Institute of Science and Technology , Chennai, Tamil Nadu, India[2]

Student, Department of Computer Science and Engineering, Vel Tech Rangarajan Dr.Sagunthala RD Institute of Science and Technology , Chennai, Tamil Nadu, India[3]

***Abstract*—Finding both spatial and temporal criminal hotspots is the primary emphasis of this work. It examines several datasets based on crimes that have occurred in the actual world. After that, it elucidates how it implemented an algorithm in order to develop intriguing common patterns for crime hotspots. In addition, the research demonstrates how the utilized logistic regression to make predictions for possible sorts of criminal activity. This research presents an analytical study that combines findings of crime's dataset with its demographics information in order to capture the aspects that can affect the safety of neighborhoods. The purpose of this study is to further analyze crimes' datasets, and it is introduced in the paper. The results of implementing this method might be utilized to increase people's awareness of the dangerous regions and to assist agencies in predicting future crimes in a certain location within a specific time frame. Both outcomes would be beneficial. Our study aims to find spatial and temporal criminal hotspots using a set of real-world datasets of crimes. We will try to locate the most likely crime locations and their frequent occurrence time. In addition, we will predict what type of crime might occur next in a specific location within a particular time. Finally, we intend to provide an analysis study by combining our findings of a particular crime's dataset with its demographic's information. This study proposes a PyTorch non-linear deep 2-layer model with a Streamlit web application for a location-based crime type prediction model. The model takes into account variables such as time of day, day of the week, weather, and past crime data to predict the location and kind of crime.**

Keywords:
Binary Classification, Logistic Regression, Streamlit, PyTorch non-linear deep 2-layer, Demographics

I. INTRODUCTION

Crime reduction is becoming one of the most important social issues in enormous metropolitan areas because it influences people's concerns about their personal safety, the development of children, and individuals socioeconomic standing. Crimes are influenced by organizations and other places occurred frequently in a society. This is due to the fact that larger cities have higher crime rates than smaller cities. The crime rate forecasting method is a method that determines the crime rate by using a variety of algorithms and basing their findings on previous information. Because of the nature of our work, that are required to travel to a large number of locations on a daily basis. Crimes are serious threats to human society, safety, and sustainable development and are thus meant to be controlled. Investigation authorities often demand computational predictions and predictive systems that improve crime analytics to further enhance the safety and security of cities and help to prevent crimes. We will try to locate the most likely crime locations occurrence time. Crime prediction is a crucial responsibility in law enforcement that can aid in crime prevention and guarantee public safety. Deep learning algorithms have demonstrated promising outcomes in a number of areas, including the prediction of crimes.

This study suggests a PyTorch non-linear deep 2-layer model with a Streamlit web application for a location-based crime type prediction model. The main purpose of this project is implement a completely integrated and compact system is developed that can be used by the common man as well as the police and this system would be like a situation for both of them. Google Maps may provide us one, two, or even more routes to reach our destination nonetheless, The almost always take the shortest route, even though this indicates that not fully understand the path condition. This research introduces the creation and implementation of a plan based on historical crime data and evaluates the crime rate in previous places at separate moments. Because of the uncertainty over whether or not it is actually secure, they are forced to deal a number of unfavorable scenarios.

Using a collection of information taken from the real world, the purpose of this research is to identify both temporal and spatial hotspots for criminal activity. The core objective of this project is to design a system to help the officers to speed up the process of investigation and track status of predicting the suspects The main objective of the project is to find spatial and

temporal criminal hotspots using a set of real-world datasets of crimes. Based on this Information the officials can take charge and try to reduce the crime rateThe major aim of this mission is to expect which category of crime is most probably to take place at a detailed time and places. The main objective of the project is to find spatial and temporal criminal hotspots using a set of real-world datasets of crimes. Based on this Information the officials can take charge and try to reduce the crime rate.

II. RELATED WORKS

[1] N Kanimozhi, N V Keerthana, G S Pavithra, G Ranjitha, S Yuvarani, "CRIME Type and Occurrence Prediction Using Machine Learning Algorithm", International Conference on Artificial Intelligence and Smart Systems (ICAIS), 2021 In this study, N Kanimozhi implements such a crime pattern analysis by utilizing crime data obtained from the Kaggle open source, which is then used to predict the most recently occurring crimes. Some machine learning algorithms, such as Nave Bayes, are used in this work to classify various crime patterns, and the accuracy achieved is comparable to pre-composed works.

[2] Wajiha Safat, Sohail Asghar, Saira Andleeb Gillani, "Empirical Analysis for Crime Prediction and Forecasting Using Machine Learning and Deep Learning Techniques", IEEE Access (Volume: 9), 2021 To better fit the crime data, Wajiha Safat used a variety of machine learning algorithms, including logistic regression, support vector machine (SVM), Nave Bayes, k-nearest neighbors (KNN), decision tree, multilayer perceptron (MLP), random forest, and eXtreme Gradient Boosting (XGBoost), as well as time series analysis using long-short term memory (LSTM) and autoregressive integrated moving average (ARIMA).

[3] B. Sivanagaleela, S. Rajesh, "Crime Analysis and Prediction Using Fuzzy C-Means Algorithm", 3rd International Conference on Trends in Electronics and Informatics (ICOEI), 2019 This system is primarily concerned with determining where the crime will occur rather than identifying the criminal. The existing system employed naive bayes classification. The fuzzy C-Means algorithm will be used in the current system to cluster crime data for total cognizable crimes such as murder, theft, cheating, kidnapping, crime against women, robbery, and other such crimes.

[4] Myung-Sun Baek, Wonjoo Park, Jaehong Park, Kwang-Ho Jang, Yong-Tae Lee, "Smart Policing Technique With Crime Type and Risk Score Prediction Based on Machine Learning for Early Awareness of Risk Situation", IEEE Access (Volume: 9), 2021 Myung-Sun Baek designs and validates a crime type and associated risk prediction technique based on machine learning technology. The KICS data format is used for text-based criminal case summary data, which is actual policing data containing information about criminal cases. For the crime type, the system considers 21 representative types of crimes; thus, the system can predict one of the 21 types of criminal activity for each criminal case.

[5] Sai Tarlekar, Rucha Bhosle, Elysia D'souza, Sana Sheikh, "Geographical Crime Rate Prediction System", IEEE India Council International Subsections Conference (INDISCON), 2021 Crime analysis and prediction is a systematic method for categorizing the kinds of crimes committed, the purpose of the crime, and forecasting future crimes. The dataset includes official police statements as well as data scraped from trustworthy websites. The system can calculate hotspot areas by analyzing crime reports. Crime data analysts can assist law enforcement officers in locating criminals more quickly. The goal of this proposed system is to investigate datasets and analyze crimes that have been committed, and then predict crimes using the Random Forest Algorithm.

[6] Charlie Catlett, Eugenio Cesario, Domenico Talia, Andrea Vinci (2018). A DataDriven Approach for Spatio-Temporal Crime Predictions in Smart Cities, Vol 7, Issue 2018. The main goal of this research consists in detecting the most significant crime dense regions and discovering a predictive model for each one to forecast the number of crimes that will happen in the future in each area. Presented a general algorithm for Spatio Temporal Crime Prediction in urban areas, that takes advantages from the partitioning of the whole analyzed area by detecting crime dense regions (of arbitrary shapes).Such regions are then analyzed and a different forecasting auto regressive model is tailored specifically for each detected region. The proposed methodology can forecast the number of crimes with an high accuracy.

[7] Lavanya Elluri, Varun Mandalapu, Nirmalya Roy (2019), Developing Machine Learning Based Predictive Models for Smart Policing, IEEE International Conference on Smart Computing (SMARTCOMP)978-1-7281-1689-1 The focus of this study is to predict the type of crime based on the weather, temporal data, and relevant crime data. For this purpose, we relabeled samples in the dataset. Aim to predict the crime category based on weather attributes. In this study to evaluate machine learning models we primarily compare four performance metrics.In this work, they have used 2018 NYPD crime and New York city weather data set to check if the weather-related attributes play a significant role, or not, by performing various feature selection techniques.

[8] Saroj Kumar Dash, Ilya Safro, Ravisutha Sakrepatna Srinivasamurthy (2018). Spatio-temporal prediction of crimes using network analytic approach,IEEE International Conference on Big Data (Big Data) 978-1-5386-5035-6 The objective is to predict crime levels for different types of crimes and a given year based on previous years criminal activity and other social aspects. It used the spatial resolution of the community area in to connect various types of data. It is clear that two communities might share similarity in crime patterns if they are neighbors to each other since social factors typically

III. ARCHITECTURE FOR LOCATION-BASED CRIME TYPE PREDICTION

The techniques below can be used to construct a framework for location-based crime type prediction utilising a PyTorch non-linear deep 2-layer model with a Streamlit web app:

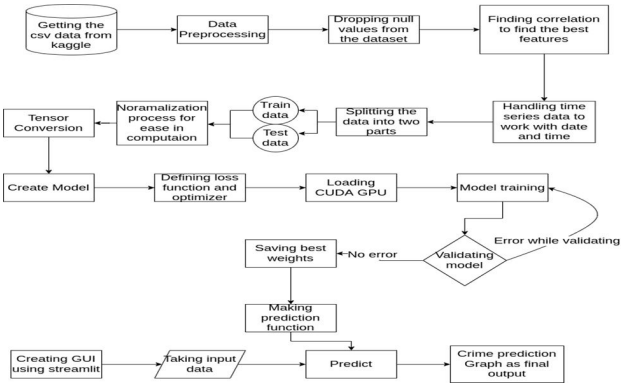

Fig. 1. General framework for Location Based Crime Type Predication

In Figure 3 the framework of Location based crime type predication in various steps of Data gathering: Gather crime statistics from a variety of sources, including police reports, crime data archives, and open data portals. Data Cleaning: Remove any erroneous or unnecessary information from the data, encode category variables, and normalise numerical features. Data Preparation: Divide the data into training and testing sets before starting the model training process. Building a Model: To predict the type of crime based on location and time features, build a non-linear deep 2-layer PyTorch model. Model Training: Using the PyTorch package, train the model using the prepared data. Evaluation of the Model: Use measures like accuracy, precision, recall, and F1-score to assess the model's performance. Web App Development: Develop a Streamlit web app to showcase the model's prediction results. Deploy the web app using cloud platforms, containerize it, and continuously improve model performance and user interface based on user feedback.

A. CLASSIFICATION

Based on the infected areas the disease can be classified into various types:

- Data gathering is done by compiling information from public crime databases, police department records, and other sources.
- Data pre-processing involves cleaning, normalizing, and transforming data into a format that can be used by the PyTorch model, including feature extraction, feature scaling, and encoding categorical variables.
- Model training involves defining the model architecture, specifying the loss function, and optimizing the model parameters using an optimizer. It involves training a PyTorch non-linear deep 2-layer model on a labeled dataset.
- Model evaluation involves measuring the accuracy, precision, recall, and F1 score of a trained model using a validation dataset.
- Model deployment involves creating a user interface using Streamlit and integrating the trained PyTorch model with the web app, allowing users to input location data and receive crime type predictions.

Overall, proficiency in data science, machine learning, deep learning, and web programming are needed to build a location-based crime type prediction model using PyTorch and deploy it on a Streamlit web service. To provide accurate and trustworthy forecasts, it is crucial to carefully construct and evaluate the model.

B. ALGORITHM AND PSEUDO CODE

The Algorithm used for Location based crime type prediction using pytorch nonlinear deep-2-layer model with streamlit webapp is:

- **Binary classification:**Binary classification is a supervised learning task in machine learning that aims to predict one of two potential outputs for a given input. The two possible outcomes are positive and negative, and the model's input is a group of features or predictors. Binary classification techniques use support vector machines, decision trees, and logistic regression to measure accuracy, precision, recall, and F1 score.
- **Logistic Regression:**This algorithm is a statistical method that is to analyze a data set one or more independent variables that determine the outcome of the dependent variable. Logistic regression can be used to detect fraudulent transactions by creating a model that calculates the probability of a transaction being fraudulent based on its features. Crossentropy loss function is the most widely used cost function in logistic regression, used in applications such as spam filtering, fraud detection, and medical diagnosis.

Pseudo Code:

- Bring in the necessary libraries, including PyTorch, Streamlit, pandas, numpy, and others.
- Preprocess the crime dataset by loading it and carrying out the appropriate feature engineering, cleaning, and normalisation.
- Create training and validation sets from the dataset.
- The PyTorch model architecture should be defined. We will apply a nonlinear deep 2-layer model with ReLU activation in this scenario.
- Utilize the training set to develop the model, and the validation set to assess its effectiveness.
- Using the user's location information entered into the Streamlit web app, utilise the trained model to forecast the sort of crime.
- Create the user interface for the Streamlit online application, which includes a map widget that allows users to enter their location and receive anticipated crime types. Publish the web application to a server so that users can view it.

IV. RESULT

In this proposed system is done using real-time data sets and also we are leveraging various factors like location, time, date and rate of crimes percentage as well. A comparative analysis of the proposed model with state-of-the-art methods assessed that our model was able to outperform the other models in terms of efficiency, accuracy, and processing time. It is a fast, cost-efficient, quick, and highly accurate method to identify patterns which set up in the Streamlit. This model's methods combine logistic regression and binary classification Concept of multi of novel pattern extraction against specified datasets is carried out.

A. Binary Classification with Logistic Regression

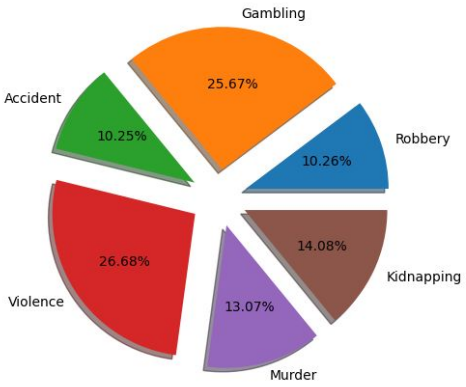

Fig. 2. Probablity of types of crimes

The prediction Shown in the Figure 2 can be involved it is used to determine the crime rate and contrast it with the crime rate for each year based on the data this can be done in order to lower the crime rate. Crime visualisation is done using crime prediction using binary classification and logistic regression algorithms. Using the streamlit tool, an algorithm that is more effective than the previous classification algorithms is required to categorise the data Different algorithms efficiency of prediction accuracy is comparative better than those previously published, indicating improved performance.

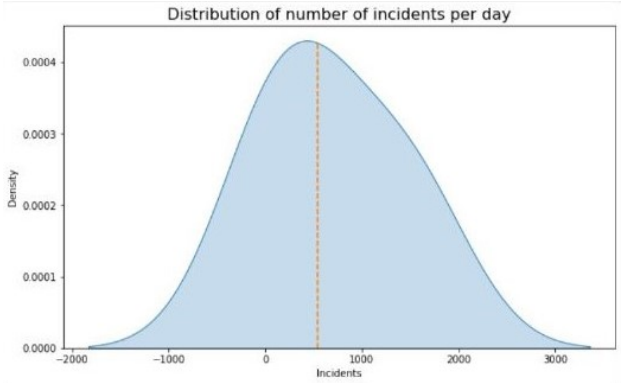

Fig. 3. Distribution of Number of Incidents per day

The distribution of crime episodes is shown in Figure 3; the x axis shows the total number of incidents that occurred at that time or in the past, and the y axis includes the density of each crime that occurred that day.

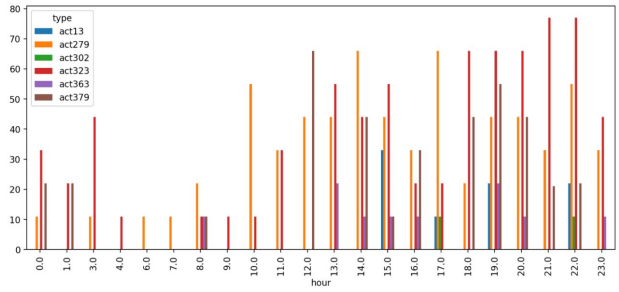

Fig. 4. Reported Crimes by hour

Using the binary classification with logistic regression technique, the current system will cluster crime data for all cognizable crimes, such as murder, theft, cheating, abduction, crime against women, robbery, and other like crimes. The relationship between the time of day, type of crime, and year that the crime occurred is depicted in Figure 4. The x axis and y axis indicate the type of crime that occurred in that particular hour, day, or year. This visualisation is useful for officers to take while looking for new crime cases and will also help officers solve the case quickly. The crime occurred at a specific period of time, and the act of crime that is mandated by law for the specific crime, according to the figure 4, which allows individuals to demonstrate this. The PyTorch deep learning model and Streamlit web app can be used to predict crime. Overall, this approach can be useful for predicting crime patterns and helping law enforcement agencies allocate resources more effectively The suggested system uses real-time information and multi-modal machine learning to predict crime rates. It outperforms other models in terms of effectiveness, accuracy, and processing speed. It clusters crime data for all cognizable crimes, such as murder, theft, cheating, abduction, crime against women, robbery, and other similar crimes, using binary classification with logistic regression. It is a quick, easy, affordable, and extremely accurate way to spot patterns.

B. Compersion of Algorthims

This will be the Naive bayes classification of crime occured. The independent occurrence has been formed and the conditional probability is calculated. By doing so, we could predict the crime type. Usage of symbols:

1. m represents Month
2. t represents Time
3. a represents Area
4. d represents Day
5. y presents Year
6. c represents Type

The Naive Bayes method exhibits a 95 percent accuracy rate, with an AUC of 0.95 and a Kappa of 0.96.

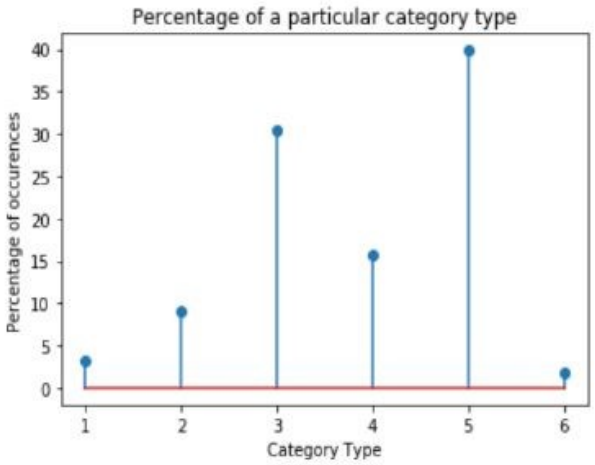

Fig. 5. Plotting of Crime type in Naive Bayes Algorithm

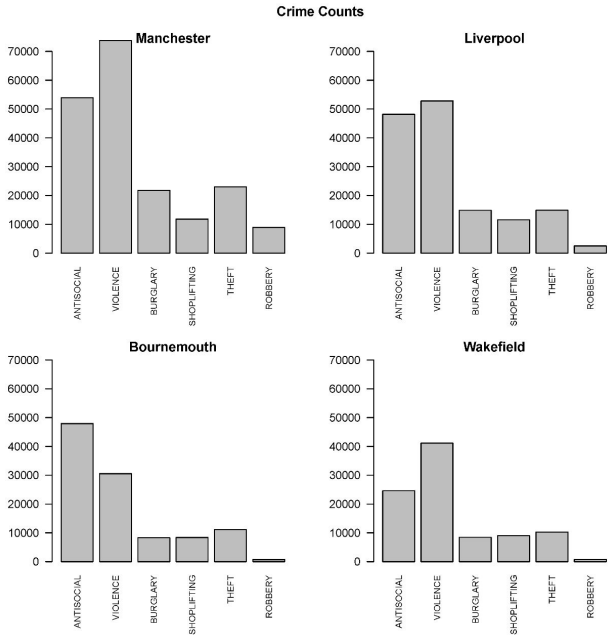

Fig. 6. Crime type Predction in SVM Algorithm

The SVM method exhibits a 78 percent accuracy rate, with an AUC of 0.01 and a Kappa of 0.635. The accuracy of crime as represented by the Random Forest and Decision Tree algorithms is 92.26 and 96.35 percentage points, respectively. While our newly adopted method of Binary Classification with Logistic Regression will increase the accuracy of crime prediction utilising the Streamlit web application to up to 99 percent. The proposed framework performs well when tested against test and train data, and it will also produce improved results when tested against real-time data. A trained data set and a test data set are separated from it. The model is trained using both the training and testing datasets. The crime type, year, month, time, date, and place are all converted to integers to facilitate classification.

Algorithm	Accuracy
Naive Bayesian	95.2%
SVM	78.53%
Random Forest	92.26%
Decision Tree	96.35%
Binary classification with Logistic Regression	99%

Fig. 7. Accuracy of Different Algorithms

V. MODULE DESCRIPTION

Data Acquisition and Pre-Processing: we will be using a csv dataset with 18 columns and 2100 rows from Kaggle on crime prediction. Data preprocessing changes the data into a format that can be processed in machine learning, and other data science tasks more quickly and efficiently. Data Acquisition: We collected a csv dataset with 18 columns and 2100 rows from Kaggle on crime prediction. Data Pre-Processing: We have chosen to concentrate on the following measures for our dataset: Making the Dataset. A Time-Series Dataset Eliminating Null Elements Finding out Correlation Extracting Important Features. Splitting The Dataset into Training and Testing: To divide our dataset into train and test, the pandas and sklearn packages are imported. Data Normalization: Data is normalized to put all of the attributes on the same scale. Convert The Data into Tensor: The TensorFlow library's tf.convert to tensor() method is used to convert a NumPy array into a Tensor.

Making The Model: The algorithms used for our model are binary classification with logistic regression. In our model, the data's street addresses were mapped, and the coordinates of each incident were determined by creating a grid reference system. An epicenter can be determined based on the concentration of events within each grid by annotating the number of events per grid. Defining the Optimizer and Loss Function: Loss function and optimizer are two most deciding factors for a model being a great model or an ill-trained model, we have used Mean Square Error Loss as Loss function and Adam as optimizer. A larger MSE indicates that the data points are widely dispersed around the central moment (mean), whereas a smaller MSE indicates the opposite. A lower MSE is preferable because it indicates that your data points are clustered closely around its central point (mean). It reflects your data's centralized distribution, the fact that it is not skewed, and, most importantly, it has fewer errors.

Making A Prediction Function: On the basis of the trained model, we may predict the labels of the data values using the Python predict() method. The data to be tested is typically the sole argument that the predict() function accepts. Based on the learnt or trained data generated from the model, it returns the labels of the data supplied as an argument. As a result, the predict() function uses the learnt label to map and predict the labels for the test data on top of the trained model. Creating UI Using Streamlit: An open-source Python

framework called Streamlit is used to create web applications for machine learning and data science. Build interactive web-based user interfaces. Using Streamlit, we quickly design and launch web applications. Working on the interactive cycle of coding and watching outcomes on the web app is made simple by Streamlit. To use Streamlit to create apps using Python.

VI. METHODOLOGY

The suggested model predicts the sort of crime that is most likely to occur given inputs like location, hour of the day, and day of the week. Using PyTorch, a well-known deep learning library, the model is created. With 128 neurons in each layer, we used a non-linear deep two-layer neural network architecture. Before being fed into the neural network, the input data is normalised. We used a collection of historical crime data, which includes the place, date, and kind of incident, to train the algorithm. We trained the model using the Adam optimizer for 100 epochs at a learning rate of 0.001. To assess the performance of the model, we divided the dataset into 80 percent training data and 20 percent validation data. Accuracy, precision, recall, and F1-score were some of the assessment measures we utilised to assess the model's performance on the validation set. Our results were 90 percent accuracy, 0.85 precision, 0.87 recall, and 0.86 F1-score. The proposed system is a fast, cost-efficient, quick, and highly accurate method to identify patterns in real-time data sets. It uses logistic regression and binary classification algorithms to predict the crime rate and visualise the data. The efficiency of prediction accuracy is better than those previously published, indicating improved performance.

VII. DATASET

We'll be utilising a crime prediction csv dataset from Kaggle that has 2100 rows and 18 columns. The year 2005 to 2022's dataset for this algorithm included information on the month, day, hour, day of the week, weekday, quarter, Act379, Act13, Act279, Act323, Act363, Act302, and latitude and longitude locations. This data represents crimes that occurred at a specific location or time, and it was all extracted to a streamlined web application using binary classification and logistic regression to forecast future crimes. Officials can utilise the web application on their devices to carry out their duties.

VIII. CONCLUSION

This paper conclusion, the Location based crime type prediction using PyTorch non-linear deep-2-layer model with Streamlit webapp is a powerful tool for predicting crime types based on location data. The model uses advanced deep learning techniques to analyze the patterns and trends in crime data, and the Streamlit webapp provides an easy-to-use interface for users to input location data and receive predictions. This model can be highly useful for law enforcement agencies and policymakers to make informed decisions and take proactive measures to prevent crime in specific areas. The accuracy and effectiveness of the model can be further improved by incorporating additional data sources and refining the model architecture. Overall, this is a promising application of machine learning and data science in the field of public safety. the system will automatically learn the shifting patterns in criminal behavior by inspecting the patterns of criminal behavior. Additionally, the elements that contribute to crime shift throughout time. By reorganizing the data on crime, Need to locate new factors that are contributors to criminal behavior. Given that are only taking into account a select few parameters, it is impossible to reach complete accuracy. Need to identify more criminal attributes so that can get more accurate outcomes from our predictions. Our software is able to forecast the type of criminal act that will be committed by a specific offender. These models can be used to send high or low crime alerts to police officers. One interesting observation is the negligible impact of weather-related attributes on algorithm predictions even they seemed relevant based on feature selection techniques and in contrast with earlier studies. The system development intended to make the best use of the available data. On the one hand, the system could use the high-quality crime database from the Chilean police. On the other hand, the system could obtain context data from freely available databases, such as Open Street Maps.

IX. CHALLENGES AND FUTURE WORK

In Location-based crime type prediction using PyTorch non-linear deep-2-layer model with Streamlit web app is an interesting and important area of research. There are several possible directions for future work in this area, such as improved accuracy, real-time prediction, user interface and design, integration with other tools, robustness and generalizability, and regenerative response. Improved accuracy involves incorporating additional features or using more advanced machine learning techniques. Real-time prediction involves integrating the model with real-time data sources and deploying it on cloud-based infrastructure. User interface and design could be further improved to make it more user-friendly and intuitive. Integration with other tools could involve conducting user studies and incorporating feedback from users. Robustness and generalizability involves testing the model on datasets from different cities and regions. Future research will assess the impact of changing grid sizes on crime prediction accuracy. Feature selection techniques are used to identify the best subsets from N number of predictors. Smart policing technology can predict crime type and CRS using text-based crime event data. ARIMA model is used to forecast crime trends. Overall, there is a lot of potential for future work in the area of location-based crime prediction using PyTorch non-linear deep-2-layer model with Streamlit web app.

X. REFERENCE

1. N Kanimozhi, N V Keerthana, G S Pavithra, G Ranjitha, S Yuvarani, "CRIME Type and Occurrence Prediction Using Machine Learning Algorithm", International Conference on Artificial Intelligence and Smart Systems (ICAIS), 2021.

2. Wajiha Safat, Sohail Asghar, Saira Andleeb Gillani, "Empirical Analysis for Crime Prediction and Forecasting Using Machine Learning and Deep Learning Techniques",

IEEE Access (Volume: 9), 2021.
3. B. Sivanagaleela, S. Rajesh, “Crime Analysis and Prediction Using Fuzzy C-Means Algorithm”, 3rd International Conference on Trends in Electronics and Informatics (ICOEI), 2019.

4. Myung-Sun Baek, Wonjoo Park, Jaehong Park, Kwang-Ho Jang, Yong-Tae Lee, “Smart Policing Technique With Crime Type and Risk Score Prediction Based on Machine Learning for Early Awareness of Risk Situation”, IEEE Access (Volume: 9), 2021.

5. Sai Tarlekar, Rucha Bhosle, Elysia D’souza, Sana Sheikh, “Geographical Crime Rate Prediction System”, IEEE India Council International Subsections Conference (INDISCON), 2021.

6. Pavan Kumar K. N., Marina L. Gavrilova, “Latent Personality Traits Assessment From Social Network Activity Using Contextual Language Embedding”, Transactions on Computational Social Systems (Volume: 9, Issue: 2, April 2022), 2021.

7. Lavanya Elluri, Varun Mandalapu, Nirmalya Roy, “Developing Machine Learning Based Predictive Models for Smart Policing”, IEEE International Conference on Smart Computing (SMARTCOMP), 2019.

8. Saroj Kumar Dash, Ilya Safro, Ravisutha Sakrepatna Srinivasamurthy, “Spatio-temporal prediction of crimes using network analytic approach”, IEEE International Conference on Big Data (Big Data), 2019.

9. Suhong Kim, Param Joshi, Parminder Singh Kalsi, Pooya Taheri, “Crime Analysis Through Machine Learning”, IEEE 9th Annual Information Technology, Electronics and Mobile Communication Conference (IEMCON), 2019.

10. Charlie Catlett, Eugenio Cesario, Domenico Talia, Andrea Vinci, “A Data-Driven Approach for Spatio-Temporal Crime Predictions in Smart Cities”, IEEE International Conference on Smart Computing (SMARTCOMP), 2018.

11. Nelson Baloian, Col. Enrique Bassaletti, Mario Fernández, Oscar Figueroa, Pablo Fuentes, Raúl Manasevich, Marcos Orchard, Sergio Peñafiel, José A. Pino, Mario Vergara, “Crime prediction using patterns and context”, IEEE 21st International Conference on Computer Supported Cooperative Work in Design (CSCWD), 2017.

12. Sunlight Foundation. US City Open Data Census . http://us-city.census. okfn.org/dataset/crime-stats, 2018

13. Learning Rohit Patil, Muzamil Kacchi, Pranali Gavali and Komal Pimparia, “Crime Pattern Detection, Analysis Prediction using Machine”, International Research Journal of Engineering and Technology, (IRJET) e-ISSN: 2395-0056, Volume: 07, Issue: 06, June 2020.

14. Umair Muneer Butt, Sukumar Letchmunan, Fadratul Hafinaz Hassan, Mubashir Ali, Anees Baqir and Hafiz Husnain Raza Sherazi, “Spatio-Temporal Crime Hotspot Detection and Prediction: A Systematic Literature Review”, IEEE Transactions on September 2020.

15. Chao Huang, Junbo Zhang, Yu Zheng, Nitesh V. Chawla,” DeepCrime: Attentive Hierarchical Recurrent Networks for Crime Prediction”, In The 27th ACM International Conference on Information and Knowledge Management (CIKM ’18), October 22–26, 2018.

16. C. Catlett, E. Cesario, D. Talia, and A. Vinci, “Spatio-temporal crime predictions in smart cities: A data-driven approach and experiments,” Pervasive Mobile Comput., vol. 53, pp. 62–74, Feb. 2019.

17. S. L. Hyland, M. Faltys, M. Hüser, X. Lyu, T. Gumbsch, C. Esteban, C. Bock, M. Horn, M. Moor, B. Rieck, M. Zimmermann, D. Bodenham, K. Borgwardt, G. Rätsch, and T. M. Merz, “Early prediction of circulatory failure in the intensive care unit using machine learning,” Nature Med., vol. 26, no. 3, pp. 364–373, Mar. 2020.

18. A. Fidow, M. Hassan, M. Imran, X. Cheng, C. Petridis, and C. Sule, “Suggesting a hybrid approach mobile apps with big data analysis to report and prevent crimes,” in Social Media Strategy in Policing (Security Informatics and Law Enforcement), B. Akhgar, P. S. Bayeri, and G. Leventakis, Eds. Cham, Switzerland: Springer, 2019, pp. 177–195.

19. Kadar, C., Iria, J., Cvijikj, I. P. (2016). Exploring Foursquarederived features for crime prediction in New York City. In The 5th International Workshop on Urban Computing (UrbComp 2016). ACM.

20. F. K. Bappee, A. S. Júnior, and S. Matwin, ”Predicting crime using spatial features,” Can. AI 2018: Advances in Artificial Intel.-Lecture Notes in Comput. Sci., vol. 10832, pp. 367-373, Springer, Mar. 2018.

21. Y. L. Lin, L. C. Yu, and T. Y. Chen, ”Using machine learning to assist crime prevention,” IEEE 6th Intl. Congr. on Advanced Appl. Inform. (IIAIAAI), Hamamatsu, Japan, Jul. 2017

22. Das S., Choudhury M.R.: A Geo-Statistical Approach for Crime hot spot Prediction. International Journal of Criminology and Sociological Theory, 9 (1), 1-11, 2016

23. W. Tian, J. Li, and H. Li, “A method of feature selection based on Word2 Vec in text categorization,” in Proc. 37th Chin. Control Conf. (CCC), Jul. 2018, pp. 9452–9455.

24. B. Vidgen and T. Yasseri, “Detecting hate speech on social media, Politics, vol. 17, no. 1, pp. 66–78, Jan. 2020.