Advanced Journal of Interdisciplinary Studies 9 (2026) 2440-2464

Journal homepage: www.ajids.dmu.edu.et

Volume 9(2), March 2026

Crime Detection and Identification using Hybrid Model to Enhance Crime Prevention in Amhara Region

Ayisheshim Almaw[1]*, Debalkew Germew[2], Banchalem Adugnaw[3]

[1]Department of Computer Science, Debre Markos University, Burie Campus, Burie, Ethiopia
[2]Department of Computer Science, Debre Markos University, Burie Campus, Burie, Ethiopia
[3]Department of Computer Science, Debre Markos University, Burie Campus, Burie, Ethiopia

Abstract

Crime prevention and investigation has become vital problem for law enforcement agencies due to poor crime data analysis and forecasting techniques. Knowing crime patterns for law enforcement agencies helps to prevent and reduce crime. Law enforcement agencies can perform actively and respond earlier if they have better information about crime patterns. The aim of crime analysis and forecasting is to get meaningful information from large crime dataset and to assist law enforcement bodies as well as to reduce crimes. This paper introduces to predict crime type and crime rate using hybrid of classification models and long short term memory model respectively. The experimentation was divided into four categories: statistical data analysis and visualization, base classification model, Hybrid model of base classification models for crime type and LSTM for crime rate forecast. For hybrid model, combining three different classification models. The experimentation result shows the hybrid model provides an accuracy of 84.2%. Predicting particular crime type and rate accurately helps to improve crime prevention and to optimize resource allocation in law enforcement agencies.

Keywords: hybrid model, LSTM, crime prediction, randomtree

*Corresponding author: Email: ayisheshim_almaw@dmu.edu.et; ayishu165@gmail.com
Phone: +251-953774446; ORCID: https://orcid.org/0000-0001-7212-0181

Article information: Received: 28 September 2025 Revised: 20 January 2026
Accepted: 25 February 2026 Available online: 31 March 2026

Doi: http://doi.org/10.20372/ajids.2026.2440

1: Introduction

Different scholars are studying and exploring on crime data analysis and prediction that has grown up quickly over the past. Most of the scholars are gathering data from official, acquired from the survey and uploaded to web sites. Due to the sensitivity of crime data most of the organization hide criminal data for safety and confidential issues. Researchers have introduced different data mining models for crime data analysis, prediction, criminal identifications, crime identification and crime hotspot area identification. Some of the papers are discussed here.

(Emmanuel A., et al.2017) applied used J48 model to predict the level of crime that will occur in a particular location in the future. The authors used only one classifier to predict the level of crime that will occur in the future. (ShijuS et.el., 2006) Introduced Naive Bayes to predict the location that has a high possibility of crime occurrence and performed better with an accuracy of 90% for crime prediction. The authors used priori algorithm to know the most frequently occurred, crime pattern in certain locations and visualized high crime prone locations on Graphed and store results on mangoDB. (Gupta A.,et al.,2017). Applied different algorithms, namely; Naïve Bayes, OneR, Decision Table, J48, JRip and BayesNet on crime data to know the best performed algorithm on crime data based execution time for predicting the crime that will occur on certain season. JRip performs better among nominated algorithms with accuracy of 73.71%.

Chung-Hsien Y.et el. (2011) applied both individual classifier (like SVM, artificial neural network, Naïve Bayes one nearest neighbor and J48) and ensemble model (voting) for crime hotspot prediction. Authors introduced forecasting technique that have high residential burglaries density and the possibility that will occur in forthcoming. The authors formulate two phases namely; a proper method of forming data is measured to store spatial data and number of crimes and crime associated activities by police.(Shahi A.,et el. 2015) introduced advantages of ensemble learning in decision support system that shows ensemble learning provides better

Several researchers studied various algorithms and combined them get better prediction called ensemble learning. This are combining more than two different algorithms and obtain final result by using combination rules that every classifier responses on the same data heterogeneous ensemble teaching (Breiman, 1996; Sesmero, 2015).homogenous ensemble combines the output of the same classifiers on the training set using bootstrap aggregating (bagging) techniques (Carson et.al., 2008). In Ethiopia, conducting a research to determine the nature of crime is very important. (Meti, (2016) conducted their study on the impact of socio-economic factors on committing crime or not. The result of the study implies demography have high with committing crimes. (GIRMAW . ,2016) studied by applying decision tree and Naïve Bayes to predict and identify levels of crime but crime type prediction is not conducted. (Robel ., 2013) studied on determinant factors of crime, criminal and victim factors for predicting crime severity level in Debre Markos Ketema administrative police Office.(Tsion, 2019) studied on crime which is time-based crime and can be forecasted using LSTM time-series forecasting model. Ethiopian crime record dataset (Leul, 2003), (Endalew, 2017), (Tsion, 2019),(Meron., 2020). The research done by (Leul, 2003) did not consider location-based data but location is a critical attribute for analysing crime data. Tsion (2019) have studied the feasibility of LSTM-RNN model on crime dataset. They have used time and location data to predict the next crime type to be committed.(Meron., 2020) studied to predict crime types using hybrid of deep learning algorithms called FFANN and LSTM-RNN

In general, in the literature there is no work that studies the crime prediction using ensemble or hybrid approach except (Bagheri. et al., 2013). and there is no paper that analyzes the impact of time and season on crime occurrences (Peng. et al.2009).

Recurrent neural networks (RNN) work best when data patterns change over time. A straightforward structured model with built-in feedback makes up this deep learning model. It can function as a forecasting engine thanks to a loop (Tayaba A. et al, 2019). In a feed-forwardneural network, signals move in only one direction, from input to output, through each layer in turn. In an RNN, however, the output from one layer is added to the next input and fed back into the same layer. The ability to operate in sequence opens up RNN to a wide range of applications (Hui H et al., 2019). In contrast to feed-forward neural nets, an RNN can accept a sequence of values as input and produce a sequence of values as output. The LSTM has been chosen to create temporal connections and define and maintain an internal memory cell state over the course of this model's entire life cycle. Additionally, they can handle multi-step forecasts and multivariate inputs and are easy to understand, approximate non-linear functions that are robust to noise (Ramirez-Alcocer et al.,2019) . Time-series data are predicted using LSTM-RNN. A study by Krishnan et al. (2018) was created to forecast the number of crimes that will occur in upcoming months and years. On the crime records from India, they used LSTM. In a study published in (Tsion ,2019) used crime data from the Addis Abeba, Bole sub-city to predict the type of crime that would occur at a given time and place.

1.1. Objective

1.1.1. General Objective

ThemainobjectiveofthisstudyisCrimeDetectionandIdentificationusingHybridModelto Enhance Crime Prevention in Amhara Region.

1.1.2. Specific objective

- Toanalyzeandinvestigatekindofcrimeoccurredmorefrequentlyincertainlocation
- To analyze the impact of Day, month and season on crime occurrence rate in certain location.
- To investigate the performance of Naïve Bayes,J48,JRip and Random Tree on crime prediction
- To analyze the performance of base models and their hybrid model on crime prediction

1.2. Significance of the Study

Crime data analysis and prediction help to prevent recurring crimes in certain location by recognizing the patterns of crimes occurred in the past or identifying the most common types of crime in certain location at a particular interval of day, month and season. Crime data analysis and prediction helps the police officer to predict future crimes that may occur in the future based on historical data in order to increase prevention efforts and distribute resources/human power wisely in more heavily affected locations to handle crime. It also benefits people to stay far from the places with certain interval of time of the day, month and seasonal on with preserving living style. Extracting new information for existing crime data called crime data analysis and prediction will be an important part of police officer tasks in the future to develop better training programs and prevent unnecessary use of resources through review of crime statistics occurred in different locations

2: Design and Methodology

This chapter deals on research methods and methodology that have been used for this study. Including data collection and Preprocessing, data analysis, model development and evaluation methods and tools that will be used to the research in this study will be discussed as follows.

2.1 Data Preprocessing

2.1.1 Data set Source and Preparation

The police commission's zonal cities of the Amhara region provided the data set for this study, which was originally in manual form but was later transformed to electronic form by the researcher. Using a technique document analysis, we acquired the dataset. Due to time constraints, we were only able to use crime records from September 2006 through 2013 E.C.

2.1.2 Data cleaning

Choosing an attribute and description: This step includes selecting an important attribute and diminishing an attribute which isn't required for crime forecast from crime data. Attributes which are used for the proposed work are crime type, crime committed date and city. Crime occurrence date is extricated to create three new attributes specifically; day, month and season of the year.

Detecting an outlier: It"s critical to recognize the extreme and lowest value which is out of the identified range. This makes a difference to eliminate the outliers from the record, Outlier is a kind of noise and they are any values of attribute which are not in identified range and the value is replaced by attribute mean. No outlier is detected in this crime dataset.

2.1.3 Data reduction and transformation

We have dropped kebele and time columns, because time attribute is almost not recorded in manual crime paper. From 80 with missing value all rows are dropped. Data transformation is applied on crime_occurrence_date. From crime_occurrence_date three new attributes are generated which contains day of the week, month of the year and season information, namely, named as crime_occurrence_date, crime_occurrence_month and offense_occurrence_season

2.2 Proposed Model

2.2.1 NaiveBayes Algorithms

Naïve Bayes algorithm works based on conditional probability, which is the likelihood that something will happen given a prior event. It is feasible to determine the likelihood of an event using the conditional probability as a basis. Naive Bayes operates under the presumption that a feature's existence or absence is unrelated to an attribute. Steps that Naïve Bayes works as follows;

- Create the frequency table for crime dataset.
- Create the likelihood tables by finding the probabilities
- Using the formula calculate the occurrence of event.

$$P(Cn|Y)=\frac{P(Cn)P(Y|Cn)}{P(Y)}$$

Where,Y is an instance which wishes to be classified ,Cn is a probable class crime type to be predicted and P (Cn|Y) is the likelihood of vector Y going to class Cn.

Prediction result is obtained: Based on the above formula the one which have greater value than the other will be the predictor.

2.2.2 J48 Algorithm

J48 is crucial to symbolically represent and depict decisions and decision-making. By developing decision rules, it aids in the prediction of the variable's class label and offers solutions through the use of tree visualization. A crime dataset consists of five attributes, and the J48 method chooses the best attribute as a root node; the choice of which attribute to be the root node will be based on information gain. Each attribute's value is calculated by information gain, and the sorted and ordered attributes are applied to the tree

2.2.3 Random Tree Algorithm

A random tree generates a large number of individual learners who build a collection of trees to address a certain problem. It builds a tree by amassing trees and producing random samples. The following steps gives the general technique that how random tree works.

- Select„X‟parametersamong„Y‟parametersrandomlyincrimedata.WhereX<Y;
- For„ A ‟parameters compute the node„ Z‟ better by split point.
- Divide the node into child nodes.
- Continue in this manner until the desired number of nodes is reached.

2.2.4 JRip Algorithm

William W. Cohen suggested Repeated Incremental Pruning to Generate Error Reduction, and this class implements a propositional rule learner. It is founded on reduced error pruning, a widely popular and useful method used in decision tree algorithms. Classes are evaluated as they grow in size, and an initial set of class rules is produced using incremental reduced error. JRip (RIPPER) moves forward by treating all of the training data's examples of a specific decision asa class and locating a set of rules that apply to every member of that class. The process then moves on to the following class and repeat itself, continuing until all classes have been covered.

2.2.5 Random Forest Algorithm

The dataset is partitioned into subsets and given to each decision tree. During the training stage, each decision tree produces a forecast result, and when a new data point happens, at that point based on the majority of results, the Random Forest model predicts the final decision.

Steps to build random forest are given below

Step-1: Select arbitrary K data points from the training set.

Step-2: Construct the decision trees related with the selected data points

(Subsets).

Step-3: Select the number N for decision trees that you need to construct.

Step-4:Repeat Step1 & 2.

Step-5: For new data points, find each decision tree prediction and assign the new data point to the category that acquired the most votes.

2.3 Proposed Hybrid Model

Hybrid model is a method by which diversified models are tactically generated and combined together to increase prediction of model (Bagheri, et al.,2013), (Rahman A. et al,2014). The models which are being combined are called base models and the models used to combine outputs are known as hybrid. This work introduces training the different models on the same data and combining these different models to improve the prediction performance called voting ensemble model.

Voting ensemble is a model that combines the forecasts from numerous other models which is trained separately. It is a strategy that will be used to improve model performance, ideally accomplishing better performance than any single model used in the hybrid. Voting ensemble Known as a meta-model, it is used with any collection of existing trained models and the existing models don't got to be aware that they are being used within the ensemble. Voting ensemble is a method for enhancing predicting performance by merging multiple models (Güneş F. et al,2017). Due to the various representation approaches and methods of hypothesis discovery used by each program. This facilitates the integration of algorithms to produce higher performance. The models are generated and then combined. In voting, many classifiers from entirely unrelated techniques, such as JRip, Naive Bayes, random trees, and J48, are combined.

Figure1 demonstrates how the same crime data is provided to three different models individually in order to create the model and determine each classifier's performance. Classifiers which are combined are known as base learners. It compares the model performance and combines the results after creating the basis models for four classifiers. An ensemble model is formed without including a base learner who performed poorly on the crime data. The output of the models is combined using the Meta learner. The outcome of better performing models is combined to provide the final prediction result.

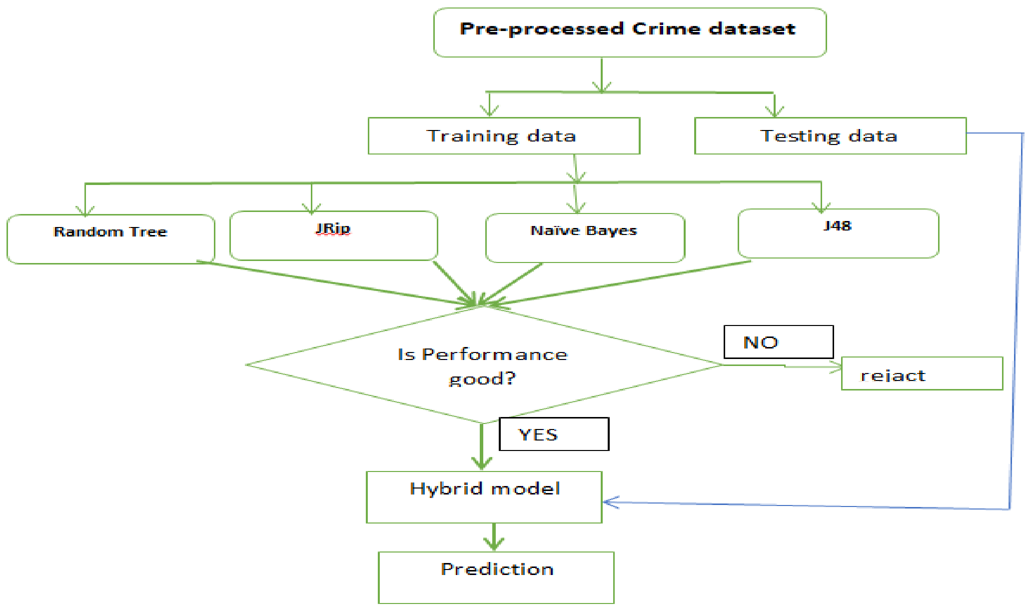

Figure1: Block diagram for proposed hybrid model once data preprocessing is completed,

Splitting dataset: for training the model and testing its performance the data need to be split into train and test data. 80% of the data is used for training and 20% for testing the model.

- **Traindata**: this is used to train the base model and hybrid model.
- **Testdata:**this data helps to evaluate the model which trained by using traindata.
- **Implementmodel:**hybrid model is made for prediction depending base model's better accuracy on the training data
- **Train the model:** in this module the train data will be feed to the implemented model and the model will train based on the given data.
- **Test the model:** in this module the test data will be feed to the model in order to measure the performance of the model.
- **Prediction result:** this module shows the result of the prediction model. The to be committed crime type will be predicted.

Hybrid model is formed according to the following technique and it works as the follow.

Step1.Train base model on training data separately.

Step 2. Evaluate the accuracy of each base model.

Step3.Reject weak performed model.

Step4.Combine well performed models into one.

Step5. Evaluate the accuracy of combined ensemble model.

2.4 Time Series Crime Rate Prediction Using LSTM

2.4.1 Time Series Data Preprocessing

A collection of data points that are regularly paced apart in time and taken at regular intervals is referred to as time series data and is gathered over a regular time period. Trend, seasonality, and errorareamongthemostimportantcomponentsofatime-seriesdataset.Usinghistoricaldatasets with temporal features, time series prediction attempts to predict future patterns and (Hochreiter et al.,1997). Crime rate for each month per year is calculated using excel formula so that we can generatet imeseries crimedata. Data set consists of daily crime rate(1/1/2006to30/12/2013E.C). In this case, the aim is to forecast the monthly crime rate, thus data are to be grouped by Month. Date attributed is grouped to month-year and crime_ rate attributes are used for crime rate forecasting. Crime rate attribute is derived from total count of crimes occurred for each month and categorize month-year pair.

2.4.2 Long Term Short Memory (LSTM)

When modeling sequential data, Keras' LSTM (Long Short-Term Memory) layer is crucial. It is made to deal with the difficulties of identifying and handling long-term dependencies in sequential input. The network can learn patterns and relationships in sequences, such as time series or natural language data, thanks to the layer's memory cells, which have long-term retention capabilities. LSTMs store data in a closed cell that is distinct from an RNN's typical flow. Similar to how data is stored in computer memory, a cell can be read from, written to, or both. By opening and closing ports, deciding when to permit reading and writing, deciding what to store and what to delete, the cell makes decisions

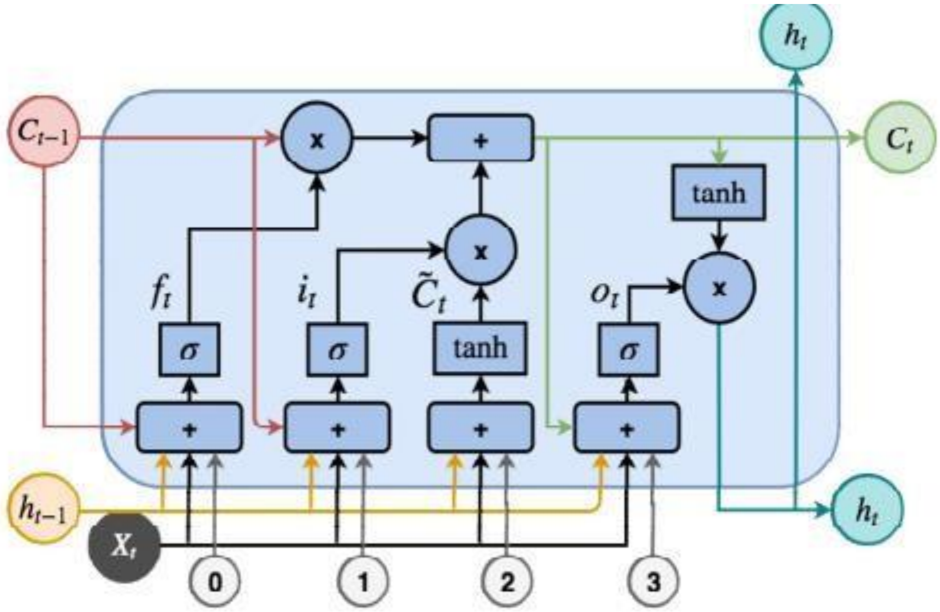

Figure2:Structure of an LSTMcell(Ramirez-Alcoceret al.,2019).

Where;

- xt:inputattimet
- ht:hiddenstateattimet
- ct:cellstateattimet
- ft:forgetgateattimet
- it:inputgateattimet
- ot:outputgateattimet

2.5 Performance Analysis of Base model and hybrid model

- **Supplied Test Set:** Supplied test set helps to know how well the algorithm is done on the unseen data. It can be used when it is vital to know or have to be testing the algorithm's knowledge against particular test set. The whole dataset is divided into two using another technique training set and test set. The classifier is ready on the complete training dataset and test set is used to assess the performance of the classifier on unseen data. Supplied test set is used for forecast after the show has been modeled**.**

- **Use Training Set:** The classifier is prepared on the total data test set, at that point the classifier is assessed on the same data set. in use training test set perfect algorithms can memorize all the training tests and accomplish better accuracy. It is used to make descriptive model and helps to form the model.

- **Confusion Matrix:** Confusion matrix shows in tabular format to visualize the execution of a demonstrate that shows up the adequacy of a model. Confusion matrix shows the test output of the prediction model. Columns in confusion matrix show the records inside the predicted class name and columns indicate the records inside the actual class. Confusion matrix can be X × Y matrix where „X‟ is the actual classes and „Y‟ is the predicted classes. The value of „a‟ and „b‟ are the same and represent the number of classes**.**

FN (False Negative): Demonstrates positive records that are mislabeled as negative. False negative of class 1 is all records that are not classified as class 1. In order to calculate for each class replace class 1 by another five individual class, at that point result is gotten for each class. The whole of FN‟s for a class is the whole value within the equivalent row without the TP.

TN (True Negative): Shows the negative tuples that are directly classified by the model. Genuine negative for class 1 is non-class 1 records that aren't classified as class 1. Order to calculate for each class replace class 1 by another five individual class, at that point result is gotten for each class.TheTN‟s for a class is the total wholeness of all columns and rows without that class‟s column and row.

Accuracy: is the complete of precisely classified occurrence divided by the overall number of inaccurately and precisely classified occurrence.

TP rate (Recallor Sensitivity)

TP (True Positive**)** demonstrates the positive records that are accurately classified by the model. TP is the total sum of diagonal records inside the matrix table. Among six classes 1, 2,3,4,5 and6 TP of class 1 is all class 1 instances that are classified as class 1. So TP rate indicates the positive records that are accurately classified by the model.

Precision:It is the measure of the having no error on condition that exact class has been predicted.

FP rate: indicates the negative tuples that are incorrectly classified as positive. FP of class 1 is non-class 1 records that are classified as class 1. In order to calculate for each class replace class 1 by another five individual class, then result is obtained for each class.

Based on above equation FPrate can be computed as= $\frac{P}{(FP+TN)}$

2.6 Tools

Data mining tools used for this implementation are Weka 3.9.6 for classification and Anaconda (Jupiter Notebook) python distribution 3.9 for data analysis and visualization.

2.6.1 Anaconda Python Distribution

Continuum Analytics created the sophisticated Python package repository known as Anaconda. Python distribution interpreter Anaconda comes with a variety of packages, including Matplotlib, Seaborn, and pandas. Jupiter Notebook used for data analysis and visualization (ScientificPython Development Environment). The following is a description of some libraries that areused for data processing and visualization.

Pandas: In order to handle the data dynamically, it stores the data in tabular form into rows and columns using data frames. The pandas object "pd" is used to load crime data as a csv file and create data frames.

Seaborn: a python library used for statically data visualization. It offers a sophisticated interface for creating statistics graphs. It can handle time series data and functions well with the Pandas data structure. Because it was created on top of matplotlib, it has pre-installed themes for customizing MatPlotlib graphics.

2.6.2 Weka

Weka is a free resource that includes a variety of algorithms and visualization tools for data analysis and predictive modeling, respectively. It is a collection of various data mining methods designed to address specific real- world data mining problems. Weka is a free data mining application that includes a variety of data preprocessing, classification, clustering, association, and visualization methods. It is platform independent and highly customizable. It is simple to use thanks to its graphical interface.

2.6.3 TensorFlow

It was developed and open sourced by Google and supports deployment across CPUs, GPUs,and mobile and edge devices too and popular and widely used deep learning frameworks. Efficiently works with mathematical expressions involving multi-dimensional arrays, Good support of deep neural networks and machine learning concepts and High scalability of computation across machines and huge data sets are basic features of Tens or Flow

2.6.4 Keras

User-friendly API (Application Program Interface) ,high-level neural network, which is written in Python and capable of running on top of Tensor Flow . It is designed to be modular, fast and easy to use (Moolayil, 2019), (Borr, 2018).

3. Experimental Analysis and Crime Prediction

3.1 Crime Data Analysis and Visualization

As a first step, statistical analysis is done to examine crime data from various angles.Theanswerstoquestionslikewhatdayoftheweekhasthe highest rate of crime, which city has the highest rate of crime during a particular period of day, month, and season are made possible through data visualization and analysis.

3.2 Analysis of Crime Occurrence Rate

3.2.1 Analysis of Crime rate Occurrence Frequency

From figure 3 it is observed that beating is the most occurred crime type with frequency of 2205, and Larceny next with frequency of 1836, Intimidation with 1668, public disorder crime with 1086.

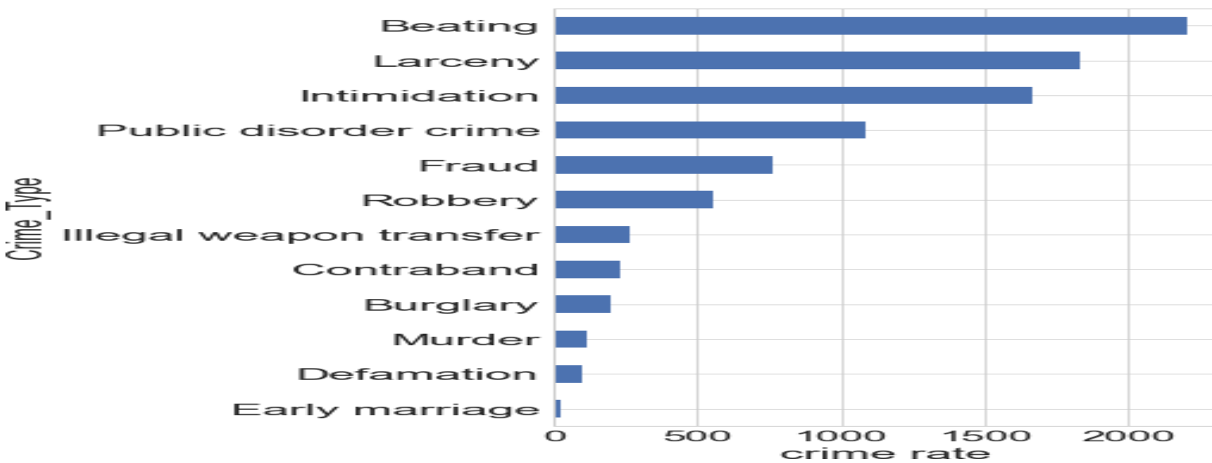

Figure 3: Crime type occurrence rate.

3.2.2 Analysis of Crime Occurrence Rate in Different Season of the Year

From figure 4 it is observed that season of autumn has the maximum number of crime occurrence rate and spring has less number of crime occurrences rate. As it is observed crime occurrence rate decrease while moving from autumn, winter, summer and spring respectively.

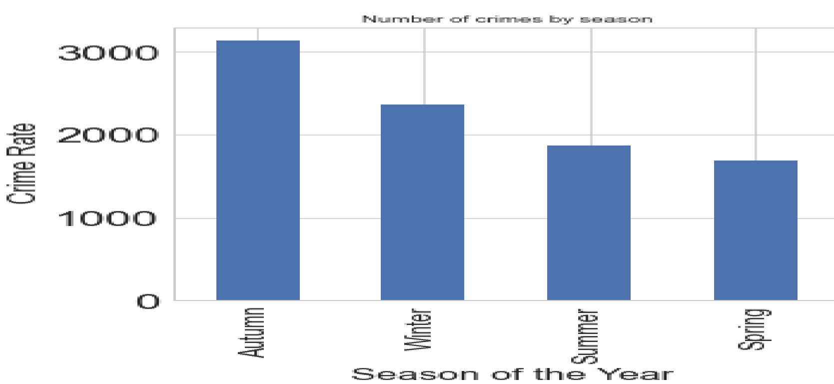

Figure 4: Crime occurrence in different seasons of the year.

3.2.3 Analysis of Crime Occurrence Rate Based on Month

Figure 5 show October has the maximum number of crime occurrences rate and April has less number of crime occurrences. As it is observed from the graph crime occurrence rate increases on October, November, December and September relatively compared to other months of the year.

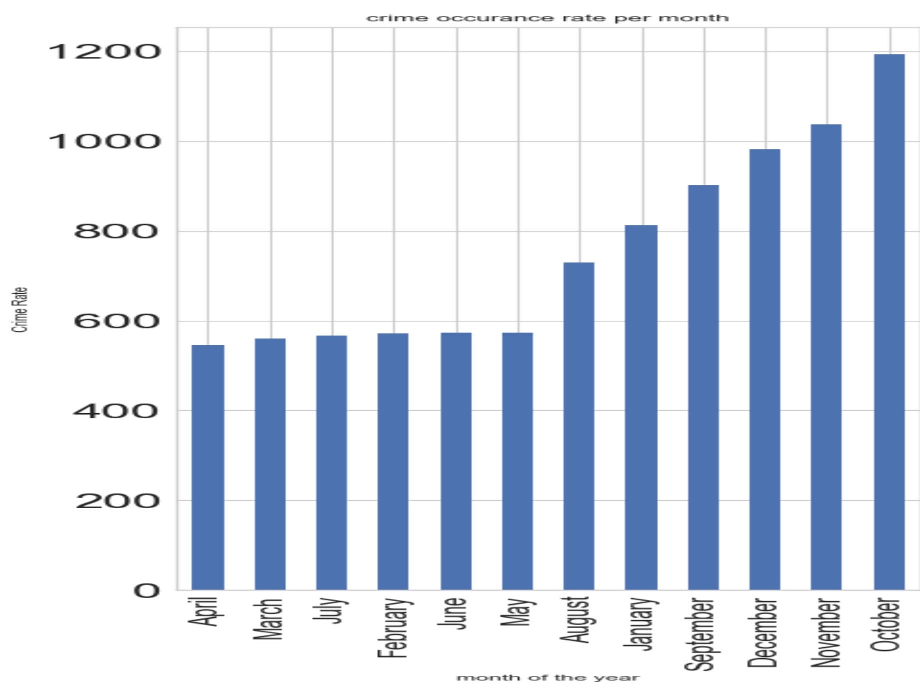

Figure 5: Crime occurrence in different month of the year

3.2.4 Analysis of Crime Occurrences Rate per Day of the Week

Figure 6 shows the number of crime occurrences in seven days of the week. It is observed that crime occurrence rate is high on Monday and less on Friday.

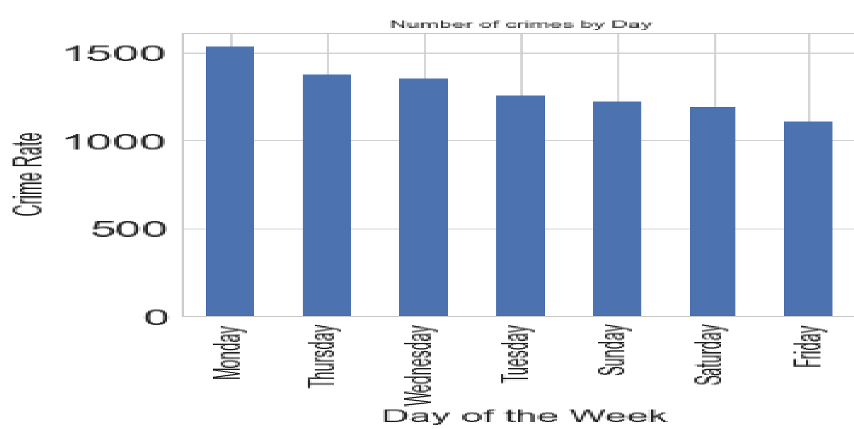

Figure 6: Crime occurrence rate per day of the week.

3.2.5 Analysis of Crime Occurrences Rate per City

Figure 7 shows the number of crime occurrences rate in nine police districts. As it is observed that district Bahir dar have highest crime rate compared to other zonal cities and the rate is described as figure below.

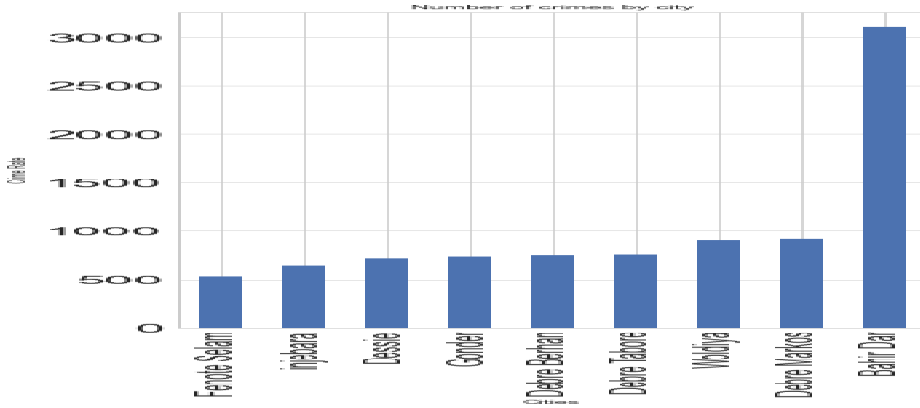

Figure7: Crime occurrence for each police districts.

3.3 Identifying Day of the Week that have High Crime Rate per City

Figure 8 shows days of the week which have the highest crime rate in nine zonal cities. As observed from Bahir dar on Monday and Wednesday have high crime rate compared to other days and less on Friday and thus day. For Woldiya crime occurrence rate is high on thusday,Saturday and less on Friday.

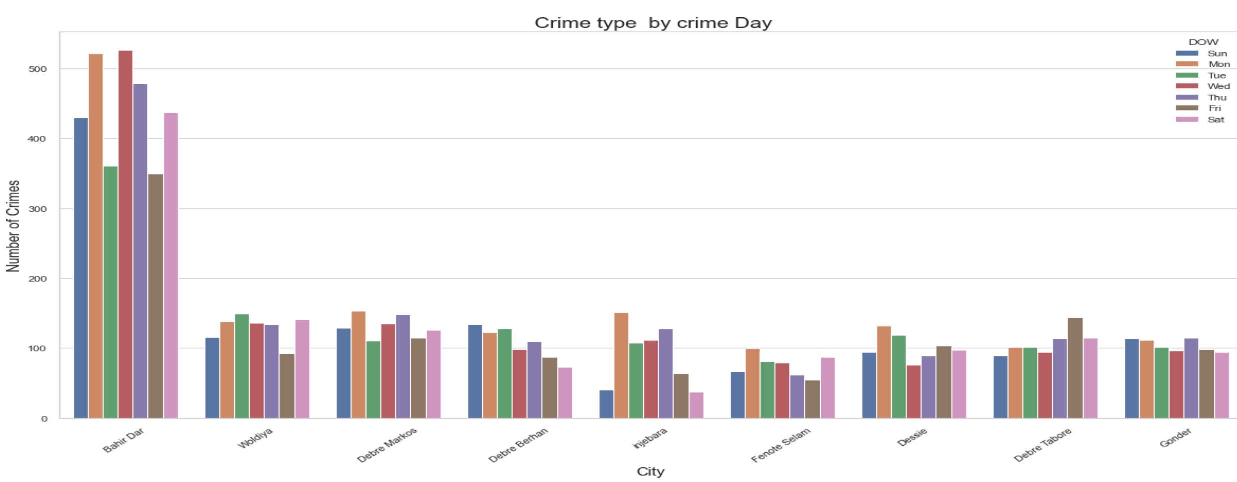

Figure 8: Identifying day of the week that have high crime rate on particular city

3.3.1 Type of Crimes Occurred in Different Days of the Week

From figure 9 Larceny is mostly occurred crime on Monday and Wednesday and less on Sunday and Friday. And the same across in most days of the week. Beating crime rate is the highest on Sunday and Wednesday and less Friday. Public disorder crime occurrence increase on the Monday, Sunday and Wednesday and less on Friday. The Intimidation crime type occurrence increase on Monday and Thursday and decreases from Wednesday.

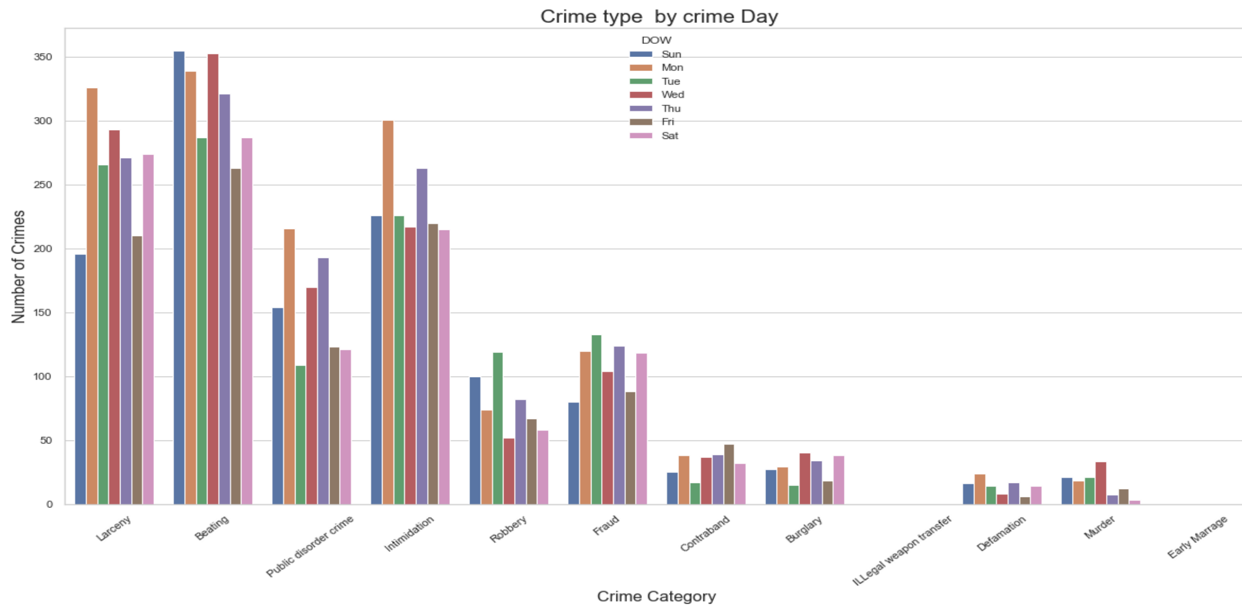

Figure 9: Crime type occurrence within days of the week.

3.4 Base Model Building and Performance Evaluation

In this step, the execution of J48, random tree, JRip and Naive Bayes classifiers are tested independently within the crime data. To conduct the experiment cleaned and preprocessed crime data are converted into an attribute relation file format (.ARRF file) which is the standard file format for Weka tool with nominal data type. In order to produce the model training data is used with known output values to build the show. Testing was done independently for each of the classifiers and precision measurements were calculated separately. The same number of training set and test sets are applied for all classifiers to build the model. The crime data is divided into two parts, namely training set 80% of crime data to build the model and 20% is used immediately after building the model to check the performance of the model. The accuracy of model presents the correctly classified records. Scientifically, a confusion matrix is used to contrast the classification accuracy for crime data with correctly and incorrectly predicted records by the model. In general the following steps are performed in order to build each base classifier using tool as follows;

1. Preprocessed crime data are loaded into the Wekatool.
2. Choose"classify"tab and click choose button to select base model.
3. After selecting a base classifier, it needs different parameters to specify. For this work default values are used for the experiment.
4. From the text options select" use training set" option as performance evaluation technique.

5. Click "start" to build the model. When model building is completed different evaluation metric results are displayed on the right panel.
6. View accuracy of each base classifier.
7. To use the model compare the accuracy of base classifiers with another classifier and ifperformed better save the model to classify new records.
8. To predict new instance based on selected model choose "supplied test set" and load the crime test data then click start.
9. Re-evaluate model on current test set and analyze the errors of the model.

3.4.1 Experimental Result of Naïve Bayes Algorithm

Naive Bayes considers two assumptions over the parameters in the dataset; these parameters are all equally significant and the value of one parameter is not related to another attribute value. Only 62. 30 % of instances are correctly classified and 37.69% Instances are incorrectly classified.

Table1:DetailedAccuracyper Classfor Naïve Bayes Model

	TP Rate	FPRate	Precision	Recall	F-Measure	Class
	0.752	0.148	0.629	0.752	0.685	Beating
	0.681	0.109	0.607	0.681	0.642	Larceny
	0.320	0.009	0.707	0.320	0.441	Robbery
	0.697	0.107	0.596	0.697	0.642	Intimidation
	0.410	0.015	0.711	0.410	0.520	Fraud
	0.655	0.053	0.625	0.655	0.640	Publicdisordercrime
	0.254	0.000	1.000	0.254	0.404	Defamation
	0.367	0.002	0.840	0.367	0.511	Contraband
	0.627	0.015	0.552	0.627	0.587	Illegalweapontransfer
	0.272	0.000	0.957	0.272	0.423	Murder
	0.588	0.002	0.370	0.588	0.455	Earlymarriage
	0.197	0.004	0.567	0.197	0.292	Burglary
WeightedAvg.	0.623	0.087	0.639	0.623	0.611	

3.4.2 Experimental Result forJRip Algorithm

JRipPerformsOnly35.79%ofinstancesarecorrectlyclassifiedand64.2%Instancesare incorrectly classified which is very week.

3.4.3 Experimental ResultforJ48Algorithm

J48 model correctly classified 79.135% (5727) instances and incorrectly classified 20.865 % (1510) instances of 80% training data. Table 6 and 7 describes detailed accuracy per class and confusion matrix generated from the tool for J48 model is given below respectively.

3.4.4 Experimental Result of Random Forest Algorithm

Among 80% training data random forest model correctly classified 83.8469 %(6068) instances and incorrectly classified 16.1531 % (1169) instances of training data. Table 8 and 9 describes confusion matrix and detailed accuracy per class.

3.4.5 Experimental Result of Random Tree Algorithm

Among 80% training data random tree model correctly classified 84.5102%(6116)instances and incorrectly classified 15.4898% (1121) instances of training data. Table 10 and 11 describes confusion matrix and detailed accuracy per class generated from the tool for random tree model is given below respectively.

3.5 Hybrid model and Performance Evaluation

All nominated algorithms which are to form hybrid are not suitable to combine and to give better prediction result. One model may provide better performance than the other nominated algorithms. There is no single algorithm that can homogeneously perform on the same dataset. So it requires algorithm selection technique to form an ensemble, for this classifier selection by accuracy technique is used to select classifiers which are to be combined. To build an ensemble model first it is important to evaluate the performance of base classifiers and classifiers which have better accuracy are used for combining.Therearedifferentstepstoformensemblemodel.Thefollowingstepsshowshowensemble model is constructed.

Step1: Crime data training set on different classifiers

Step2: Compare and evaluate the performance of algorithms

Step 3: Eliminate classifiers which are performed weak out of ensemble

Step4: Ensemble the remained algorithms which perform better accuracy

Step 5: Compare and evaluate the performance of ensemble models

3.5.1 Combining J48, Random forest and Random Tree

Combination of J48, Random forest and Random Tree model when J48 used as a combiner correctly classified 84.2062% (6094) instances and incorrectly classified 15.7938% (1143) instances of 80% training data. This shows that combination of J48, Random forest and Random Tree model shows better performance. Table 12 and 13 describes confusion matrix and detailed accuracy per class generated from the tool for combination of J48, Random forest and Random Tree used as a combiner model is given below respectively.

Table2: Detailed accuracy per class by combining J48 and random tree and random forest using J48.

	TPRate	FP Rate	Precision	Recall	F-Measure	Class
	0.860	0.063	0.820	0.860	0.839	Beating
	0.834	0.036	0.851	0.834	0.843	Larceny
	0.737	0.011	0.809	0.737	0.771	Robbery
	0.851	0.028	0.874	0.851	0.862	Intimidation
	0.847	0.012	0.864	0.847	0.855	Fraud
	0.811	0.027	0.804	0.811	0.808	Publicdisordercrime
	0.873	0.001	0.861	0.873	0.867	Defamation
	0.879	0.006	0.818	0.879	0.848	Contraband
	0.951	0.003	0.902	0.951	0.926	Illegalweapontransfer
	0.914	0.001	0.881	0.914	0.897	Murder
	1.000	0.001	0.810	1.000	0.895	Earlymarriage
	0.827	0.002	0.911	0.827	0.867	Burglary
WeightedAvg.	0.842	0.033	0.843	0.842	0.842	

Table3:ConfusionmatrixforcombinationofJ48andrandomtreeandrandomforestusing J48.

abcdefghijkl<-- classifiedas

1555701958176121553O4|a= Beating

991201213023581501 10|b= Larceny

402532218416165000 |c= Robbery

83471711341217356405 |d=Intimidation

29 2161950211 0 3 1 0 01|e= Fraud

78 3282113701 1 4 0 0 33|f= Publicdisordercrime

0005316200000|g= Defamation

511022501890001 |h= Contraband

01522000194000 |i=Illegalweapontransfer

2101020017400|j= Murder

000000 0000170|k=Earlymarriage

62083024320143 |l = Burglary

3.6 Crime Prediction

Random tree achieves better accuracy than other nominated algorithms for crime prediction. So it is used for crime prediction. Hybrid model is trained on 80% of crime data using use training set and 20% crime data are used in order to know the unseen class labels using supplied test set. The test in the experiment shows a particular crime is assumed to be unseen or unknown and needs to be predicted based on the given attributes.

3.7 Time Series Crime Prediction Model using Long Short-Term Memory

3.7.1 Time series forecasting

Sincedatapreprocessinghascometoanend, the’monthlycrimerateperyearplotispresentedas follows.

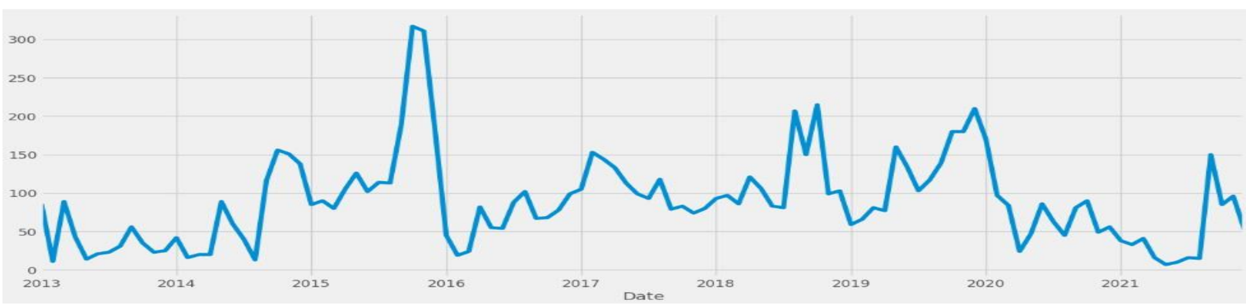

Figure11:' monthly crime rate per year

Plot below comprising all actual crime rate data, theL STM model's training outputs and the

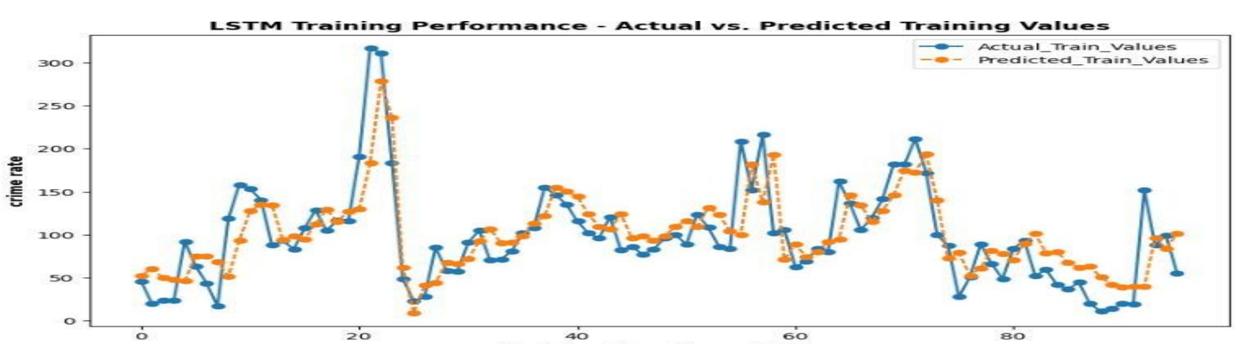

Figure12:L TSM actual vs trained value performance comparison

3.7.2 LSTM Prediction in to the future crime rate

atthelast,LSTMmodelpredictioncapabilitiesbytrainingthisdeeplearningneuralnetworkwith all the past monthly crime data from crime rate to predict the crime rate values of the next 12 months(from01-2014(2022)to12-2014(2022)).TheflowingPlotdepictstheoriginaltimeseries crime rate, and the 12-month LSTM forecasts.

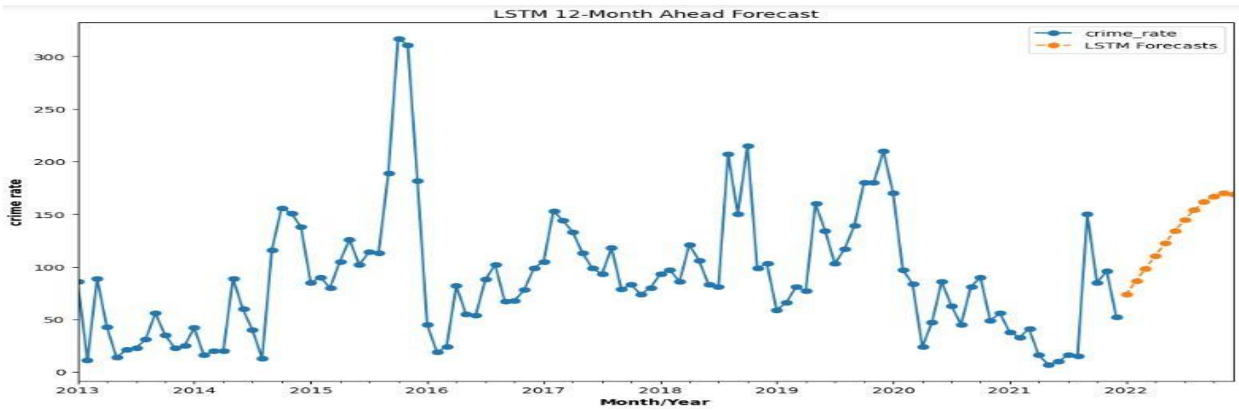

Figure13: LSTM12 month ahead prediction graph

4: Conclusion

Crime prediction and finding relevant information from a large amount of crime data is really important but challenging. Law enforcement agencies can perform actions and respond earlier if police enforcement agency has better information about crime patterns and trends in different places. To come up with this problem, this work introduced hybrid model for crime prediction to improve predictive performance. As the experiment shows the more improved model is obtained from the hybrid model with accuracy of 84.2062%.

In general, crime prediction and discovery of important information from huge crime data is very crucial. Having this kind of information helps to enforce law agencies to distribute resources wisely. Crime prediction is also important to the people to stay away from some particular locations within a certain time of the day, month and season of the year along with saving living style. In addition, Knowing this kind of information helps the people to choose the environment to live in and travelling places. It also reduces economic and life loss during crime occurrence in a country and creates economic growth and political stability in a country.

5. Acknowledgement

We are thankful to Debre Markos University giving us this opportunity to fund and complete this study. We also thankful to Debre Markos University Burie Campus research directorate office for their guidance, and support until the completion of this work. We would like to thank all police commission and police stations for their cooperation in gathering crime data that we need for this research

6. Declaration of Conflict of Interest

The authors declare that there is no conflict of interest regarding the publication of this paper.

7. Authors’ Contribution

Ayisheshim Almaw designed the study, conducted data analysis, and drafted the manuscript. Debalkew Germew supervised the research and contributed to methodology design. Banchalem Adugnaw assisted in data preprocessing and model validation. All authors reviewed and approved the final manuscript.

8. REFERENCES

1. Ahishakiye, E., Taremwa, D., Omulo, E. O., & Niyonzima, I. (2017). Crime prediction using decision tree (J48) classification algorithm. *International Journal of Computer and Information Technology*, *6*(3), 188-195.
2. Breiman, L. (1996). Bagging predictors. *Machine learning*, *24*(2), 123-140.
3. Calhoun, C. C., Stobbart, C. E., Thomas, D. M., Villarrubia, J. A., Brown, D. E., & Conklin, J. H. (2008, April). Improving crime data sharing and analysis tools for a web-based crime analysis toolkit: Webcat 2.2. In *2008 IEEE systems and information engineering design symposium* (pp. 40-45). IEEE.
4. Cesario, E., Catlett, C., & Talia, D. (2016, August). Forecasting crimes using autoregressive models. In *2016 IEEE 14th Intl Conf on Dependable, Autonomic and Secure Computing, 14th Intl Conf on Pervasive Intelligence and Computing, 2nd Intl Conf on Big Data Intelligence and Computing and Cyber Science and Technology Congress (DASC/PiCom/DataCom/CyberSciTech)* (pp. 795-802). IEEE.
5. Yu, C. H., Ward, M. W., Morabito, M., & Ding, W. (2011, December). Crime forecasting using data mining techniques. In *2011 IEEE 11th international conference on data mining workshops* (pp. 779-786). IEEE.
6. Czarnowski, I., & Jędrzejowicz, P. (2016, October). An approach to machine classification based on stacked generalization and instance selection. In *2016 IEEE International Conference on Systems, Man, and Cybernetics (SMC)* (pp. 004836-004841). IEEE.
7. Endalew, A. (2017). Crime Analysis For The Case of Bahir Dar City Administration Using Data Mining Approach. *Bahir Dar University*.(Msc dissertation)
8. Gashler, M., Giraud-Carrier, C., & Martinez, T. (2008, December). Decision tree ensemble: Small heterogeneous is better than large homogeneous. In *2008 Seventh international conference on machine learning and applications* (pp. 900-905). IEEE.
9. Gupta, A., Syed, A., Mohammad, A., & Halgamuge, M. N. (2016). A comparative study of classification algorithms using data mining: crime and accidents in Denver City the USA. *International Journal of Advanced Computer Science and Applications*, *7*(7), 374-381.
10. Güneş, F., Wolfinger, R., & Tan, P. Y. (2017, April). Stacked ensemble models for improved prediction accuracy. In *Proc. Static Anal. Symp* (pp. 1-19).
11. He, H., Hu, R., Zhang, Y., Jiao, R., & Zhu, H. (2018, June). Hourly Day-Ahead Power Forecasting for PV Plant Based on Bidirectional LSTM. In *International Workshop on High*

Performance Computing for Advanced Modeling and Simulation in Nuclear Energy and Environmental Science (pp. 208-222). Singapore: Springer Singapore.

12. Kilimci, Z. H., Akyokus, S., & Omurca, S. I. (2016, August). The effectiveness of homogenous ensemble classifiers for Turkish and English texts. In *2016 International Symposium on INnovations in Intelligent SysTems and Applications (INISTA)* (pp. 1-7). IEEE.
13. Krishnan, A., Sarguru, A., & Sheela, A. S. (2018). Predicitve analysis of crime data using deep learning. *Int. J. Pure Appl. Math*, *118*(20), 4023-4031.
14. Meti, K. (2016). An assessment of socio-economic factors on crime: a case study of Kaliti correctional administration, Ethiopia. *Master's Thesis, Addis Ababa University*.
15. Peng, W., Chen, J., & Zhou, H. (2009). An implementation of ID3-decision tree learning algorithm. *From web. arch. usyd. edu. au/wpeng/DecisionTree2. pdf Retrieved date: May, 13*.
16. Polikar, R. (2006). Ensemble based systems in decision making. *IEEE Circuits and systems magazine*, *6*(3), 21-45.
17. Rahman, A., & Tasnim, S. (2014). Ensemble classifiers and their applications: A review. *arXiv preprint arXiv:1404.4088*.
18. Rajesh, D. G., & Punithavalli, M. (2012). A study on the existing computer aided crime analysis tools to handle indian state of affairs. *International Journal of Computer Science Issues (IJCSI)*, *9*(2), 528
19. Ramirez-Alcocer, U. M., Tello-Leal, E., & Mata-Torres, J. A. (2019, July). Predicting incidents of crime through LSTM neural networks in smart city domain. In *The Eighth International Conference on Smart Cities, Systems, Devices and Technologies* (pp. 32-37).
20. Sathyadevan, S., Devan, M. S., & Gangadharan, S. S. (2014, August). Crime analysis and prediction using data mining. In *2014 First international conference on networks & soft computing (ICNSC2014)* (pp. 406-412). IEEE.
21. Sesmero, M. P., Ledezma, A. I., & Sanchis, A. (2015). Generating ensembles of heterogeneous classifiers using stacked generalization. *Wiley interdisciplinary reviews: data mining and knowledge discovery*, *5*(1), 21-34.
22. Abbasi, T., Lim, K. H., & Yam, K. S. (2019, April). Predictive maintenance of oil and gas equipment using recurrent neural network. In *Iop conference series: Materials science and engineering* (Vol. 495, No. 1, p. 012067). IOP Publishing.

23. Thongsatapornwatana, U. (2016, January). A survey of data mining techniques for analyzing crime patterns. In *2016 Second Asian Conference on Defence Technology (ACDT)* (pp. 123-128). IEEE..
24. Meskela, T. E., Afework, Y. K., Ayele, N. A., Teferi, M. W., & Mengist, T. B. (2020). Designing time series crime prediction model using long short-term memory recurrent neural network. *International Journal of Recent Technology and Engineering (IJRTE)*, *9*(4), 402-405.
25. https://www.cs.waikato.ac.nz/ml/weka/
26. https://conda.io/docs/user-guide/install/download.html.Lastaccessedon18/3/2022.
27. https://pandas.pydata.org/. Last accessed on 18/3/2022.
28. https://seaborn.pydata.org/.Lastaccessedon18/3/2022.