ISSN: 2229-6948(ONLINE)
DOI: 10.21917/ijct.2021.0374

ENSEMBLE NEURO-FUZZY BASED SYSTEM FOR VEHICLE THEFT PREDICTION AND RECOVERY

Femi Johnson[1] and Akintunde Saminu[2]
[1]*Department of Computer Science, Federal University of Agriculture, Abeokuta, Nigeria*
[2]*Department of Information Technology, National Open University of Nigeria, Nigeria*

Abstract

Vehicle theft is continuously being reported as a global prevalent crime. It often aids the perpetuation of other related crimes such as kidnapping, armed robbery, terrorism and human trafficking. The traditional mode of combating vehicle theft crime is faced with abnormallies hindering accurate, timely prediction and recovery of stolen vehicles from criminals. This paper presents a computational Artificial Intelligence (A.I) technique known as Ensemble Neuro-Fuzzy modeled system with the aim of minimizing investigation time and number of deployed security operatives towards achieving a high successful rate in the prediction, detection and recovery of stolen cars. A collection of data collected from the Criminal Investigation Department of the Nigeria Police Force, were further analyzed through Dimensionality Reduction formula and Routine Activity Approach (RTA) to extract the most significant features. Dataset were sub-divided into 60, 20 and 20% for training, testing and validating the model respectively. A significant result of 92.91% obtained with this model showed that it is most efficient in predicting, detecting and recovering of stolen vehicles as compared with other machine learning algorithms such as Random Tree, Naïve Bayes, J48 and Decision Rule of prediction accuracy of 86.51%, 71.24%, 67.68% and 55.73% respectively.

Keywords:

Machine learning, Neuro-fuzzy, Prediction, Recovery, Selection, Significant features, Vehicle theft

1. INTRODUCTION

Crime is a prevalence criminal activity which poses threat to the peaceful co-existence of people, communities, nations, countries and the world at large. It has spanned through the existence of mankind from little or unorganized form of theft [12] to a well complex and organized form requiring sophisticated gadgets combined with special intelligence for it to be curbed [38] [10]. Perpetuators of crimes are often referred to as criminals [24] and are categorized based on the type of crime committed [26] [17] ranging from murder, man slaughter, rape, human trafficking, drug addictions, armed robbery, theft, cybercrime, property theft, amongst others [37] [12]. Property theft especially car theft or vehicle theft as defined in [5] [11] [34] is the successful and unsuccessful attempts by unauthorized persons to take a car without the consent of the rightful owners. It constitutes a sizeable portion of the crime problem [16] not only in Nigeria but in other nations such as the Great Britain and America. The terms car theft and vehicle theft are often used interchangeably.

Vehicle thefts including cars, busses and lorries on roads constitute a major constraint to commuters [4] [11] and transporters in Nigeria. It possess a disturbing effect not only to security personnel [22] but also to businesses as goods and services transported from one location to another are easily hijacked by criminals thereby causing great financial loss to industries [9] [22] and in the long term affects the nation though decrease in her Gross Domestic Product (GDP). The financial loss [31] emanating from vehicle theft annually is estimated to be about 51% of the amount loss by victims to property crime [16] [43], combined with the cost of getting vehicles insured most especially in advance countries it is very obvious that vehicle theft have great financial and psychological effects on victims and the society [13].

Recently, statistics revealed that the recent increase in car theft is due to the dwindling purchasing power of the Nigerian naira against the dollar which makes importation of vehicles very expensive. About 85-90% of stolen cars in neighboring countries are shipped to Nigeria as fairly used (TOKUNBO) vehicles for purchase by un-noticing citizens [1] thereby causing some well-meaning Nigerians often arraigned and charged to court for car theft, occasionally with granted bail to the tune of large sum of money. There also exist various techniques for identifying vehicles [6] [31] [36] [40].

There is a high risk of vehicle theft in urban and densely populated areas [39] with high concentrations of industries and residential houses [3] [15]. Repair shops, religious centres, stadiums and large business outfits with car parks are also locations prone to car theft [27] [2] [21].

However, the persistent increase in the theft of cars [22] has made manufacturers equip vehicles with advance security features [33] [29] such as strong ignition and steering locks, installing of alarms, tracking devices, and object recognition devices. Researchers are also driven to carry out detailed examination into the factors necessitating and aiding the theft of cars [14] [5], predict the occurrence of crime and mode of attack as well as determine the various tracking modes to apprehend criminals who engage in such act [30] [35]. Several intelligent machine learning algorithms are available for prediction [19] [25] [32] and classification of data. They also provide varied percentage and degree of accuracy [7] [43] depending on the nature and type of data provided.

In this paper, the occurrence and pattern of car theft is exploited by analysing data from previous reported cases using computational artificial intelligence and machine learning techniques looking into the characteristics of criminals, time and mode of operation, locations of incidents and other associated factors in order to help citizens determine the safety of their vehicles at in a given location and also aid security personnel predict recovery locations for stolen vehicles by criminals, thus informing them of the best approach and route to be adopted in recovering stolen vehicles at a particular time.

This paper is structured as follows: section 1 is an introduction into vehicle theft, section 2 provides a detailed literature review of related works on vehicle theft as a criminal act and the techniques adopted for tracking and recovering of stolen vehicles.

Section 3 presents the research methodology, the architectural framework, Algorithms and proposed Neuro-Fuzzy model for vehicle theft prediction and recovery with section 4 detailing and depicting the results, training datasets, results analysis, evaluation and discussion. Finally, section 5 concludes the paper with contribution to knowledge and recommendation of the research methodology.

2. RELATED WORKS

In [34] a proposed method for analyzing car thefts and recoveries was presented with connections to modelling origin-destination point patterns. Two major datasets were collected for this research modelling. The first being a dataset consisting of about 4016 records of car theft's locations with a recovery percentage of 10% in the year 2015 from Neva, Mexico and the other from Belo Horizonte, Brazil comprising of 5250 pairs of stolen car data and location of recovery from August 2000 to July, 2001. The first set of data were categorized into two subsets and varied with variables including population of citizens, number of apartments, health conditions of citizens and count of gainfully employed people. The other category was varied with different crime types linked to car theft to include murder, burglary, robbery and kidnapping. Furthermore, the second dataset provides a joint of pair information regarding recovery location and areas of car theft. A number of car theft event modelling techniques were analyzed with two performing better than the others. These two modelling techniques are the Non-Homogeneous Poisson Process (NHPP) and the Log-Gaussian Cox Processes (LGCP) model. A p-thinning cross validation was adopted to validate test data combined with a Continuous Rank Probability (CRK) function for the accuracy evaluation of predicted areas of car theft recovery. Results generated indicate that the recovery locations are dependent on the theft location.

Initiator prediction and Near Repeat analysis was used by [14] to significantly detect the factors responsible for car thefts and residential burglar in the city of Indianapolis, Indiana for a period of a year precisely in 2013 from geocoded X-Y coordinates crime data provided electronically by the Information and Intelligence department of the Indiana Metropolitan Police Division (IMPD). A group of four categories (social disorganization, crime generators, geographic locations and date of occurrences) were created for nineteen variables used in the prediction. In addition, a dual reference data table was adopted to maximize hit rate of the geocoded information by recording incident address as a pair of street corners rather than the general address listing of street names, a total of 8,075 residential burglaries and 3,149 motor vehicle theft incidents were re- geocoded with the aid of the point distance tool in ArcGIS. The distance between the two identified points was calculated to aid accurate precision. Crime generators informed by criminal environmental Specialist included in this research were liquor stores, parks, shops, bars, ATMs and banks. Physical features of locations (trails, rail roads, rivers, police patrol zones) aiding or hindering car theft were also considered as some facilitated criminal movements between crime scenes and escape routes. Further analysis was performed on the data similar to that performed in [8] to extract relevant social disorganization variables including racial heterogeneity, geographic mobility and population density from the dataset to predict the number of targeted cars at risks. In addition, the Near Repeat Calculator (NRC) software of version 1.3 which adopts the use Knox test by comparing each event in a data-set with every other event with records of both spatial and temporal distances between the two points was used to identify significant spatiotemporal clusters. Other utility functions of the software were used to identify the number of times an incident acted as an initiator in the cluster.

The preliminary study involving integration of environmental factors researched by [39] predicted and classified crime rates into four levels ranging from low to high. Street level images of captured areas where the crime events took place in the year 2014 and 2015 were collected in Chicago city region of the United States, America. These images were classified into equal sizes to aid accurate prediction of the 4-Cardinal Siamese Convolutional Neural Network (4CSCNY) adopted. Four cardinal points images collected with respect to a given reference positions serves as input to the model with pre-trained frozen weights adopted from the famous Alexnet Architecture. A single output generated was attached to a descriptor which is then finally classified by a Multi-layer Perceptron (MLP), into one of the earlier labeled levels. Although attention was not focused on the quality of images nor special attributes but on the whole use of captured images, yet an overall accuracy of 54.3% was recorded.

Research in [41] exploited spatial-temporal data comprising of data related to public security, Meteorology, Human mobility, Public Service Complain and Point of Interests (POI) in ne borough urban city in New York with machine learning technique leveraged to predict crime occurrence in other urban centres. Extracted features and patterns from collected dataset were analysed and classified into two groups; intra and inter-patterns based on the proximity of the regions in relations to the collected data. A total of 30, 11 and 311 datasets were extracted from Meteorology, Point of Interest and Public Service Complaint (PSC) data respectively. These data became the base for the integration of their proposed Transfer learning Model used to train parameters for the prediction of crime in other cities in New York.

Crime prediction model deployed by [20] in their research based on the application of Multi-Modal data combined with an environmental theory and Break Window Theory (BWT) using Deep Neural Network (DNN). Dataset for their research include crime related statistical record obtained from various online sources including the City of Chicago Data portal, Demographics obtained from America Fact-Finder, weather data from Weather Underground and captured Images from Google Street view. Statistical Analysis on data helped to determine the correlation between collected date and crime occurrence. The DNN embedded with feature extraction level and four layers (Spatial, Temporal, Joint Feature Representation and Environmental context layer) was used in predicting was also used for predicting crime.

A comparison of two machine learning algorithms (K-Nearest Neighbourhood and Naives Bayes classifier) implemented on python by [28] for crime analysis in multistate is one of the applications of artificial intelligence adopted in crime prediction and classification. The analysis was performed to ascertain, compare the validity and determine the appropriateness of the best algorithm for multistate crime data in India. In the K-NN, a marked value of K which determines the success of classification by extracting its associated nearest neighbours from a group of data is used in classification and could be trained continuously until the best result is achieved. However, a similar technique with a conditional model known as Naïve Bayes classifier was compared using the same dataset grouped into training and testing data respectively. An accuracy of 87% was recorded against the former classifier with 77% accuracy level. Similarly, the execution time comparison revealed that the Naïve Bayes utilizes less time than the others with time duration of 0.03 seconds.

The effectiveness and accuracy of machine and soft computing technique were also utilized in [32] [36] [44] was also noticed when [45] proposed exploring the influence of truck proportion on freeway traffic safety using adaptive network-based fuzzy inference system. The study sourced used data from VISSIM. Simulation and orthogonal experimental was outlined for standardization of the data used, together with the combination of SSAM to evaluate the influence of truck proportion on traffic flow parameters and traffic conflicts. It was later proposed that the critical and conventional conflict prediction model built on the Adaptive Network-based Fuzzy Inference System (ANFIS) in establishing the influence of truck proportion on freeway traffic safety could be adopted. Although, the study showed an increasing traffic conflict and number of serious conflicts than general conflicts for truck proportion ranging from 0.4 under 3200 veh/h to 0.6 under 2600 veh/h. below 3000 veh/h, there was an upsurge in travel time and average delay. On the other hand, there is a reduction in the mean speed and speed of small car as the truck ratio rises. Results also showed that ANFIS model can correctly ascertain the influence of truck proportion on traffic conflicts under diverse traffic volume, and also substantiate the learning capacity of ANFIS. The authors in [18] developed a multi-ANFIS model based synchronous tracking control of high-speed electric multiple units. The study presented a developed modelling and operation procedure in actualizing the simultaneous monitoring of vehicle body with its corresponding motion. The study built an expanded ANFIS model via learning from data gathered from vehicle motions in real-time to illustrate the high-speed electric multiple unit (HSEMU), while implementing the simultaneous monitoring using a model predictive controller. Results from the study revealed that the modelling and operational procedure drastically enhanced the effective running of HSEMU with respect to vehicle protection, promptness, ease and parking precision.

In [23], the authors proposed speed prediction for triggering vehicle activated signs. The study used Adaptive Neuro fuzzy (ANFIS), regression tree (CART) and random forest (RF) model to determine a precise predictive model built on historical traffic speed data to derive an ideal trigger speed throughout each period. The models built were tested in contrast to findings acquired from artificial neural network (ANN), multiple linear regressions (MLR) and naïve prediction using traffic speed data retrieved at various locations in Sweden. The study revealed RF as an effective technique for predicting mean speed for both the short and extended period. Similarly, the paper showed an upturn in response time with regards to computational complexity, functioning and other features to the predictive model, and at the same time offering a low estimation error.

The research conducted by [1] [5] adopted the use of Charnov's Prey Selection model, Routine Activity and other associated variables including flow of cars, guardianship level and social disorganizing factor to determine the most common car model stolen by criminals. Lagos state, Nigeria with a population of over seventeen million people and known to be the centre of the nation commercial, financial and economic capital was used for the research. Secondary data comprising of information relating to number of car thefts, model of cars involved, place of theft, and number of thefts between 2009 and 2013 collected from the Lagos State Police command was used. Areas of theft were classified into three based on population density as high, medium and low. Three major places identified with high risks of car theft are residential homes, parking lots and streets. Findings revealed that unemployed male between the ages of 20-39 years were likely to be engaged in this criminal act and most common targeted areas are parking lots in densely populated areas of the state.

3. METHODOLOGY

This research approach for predicting vehicle theft recovery locations is based on an ensemble of Adaptive Neuro-fuzzy Inference System and significant feature selection of attributes (Neuro-fuzzy) model. A given collected dataset from the Criminal Investigation Department of the Nigeria Police Force; Nigeria serves as the input with feature selection technique to reduce the number of factors from the dataset in order to retain only the subset of relevant factors to be simulated in the Neuro-fuzzy model for prediction. The model learns from the extracted dataset and provides appropriate predictions of vehicle safety and recovery locations by tuning a set of membership function parameters combining the back propagation algorithm and least squares method. The collated dataset consisting of investigations with 394 instances and 14 attributes is divided into proportions for use in training and testing of the model.

3.1 DATA REPRESENTATION

Given that there exist a single dataset $\{A\}$ with n attributes combined to form an n-dimensional vector obtained from the collection of datasets (b, c, d, e, f and g) from six different years (2015-2020) with multiple similar attributes (n_1, $n_2,\ldots,n_6$) respectively. This single dataset $\{A\}$ referred to as Unified-Multi Data (UMDat) is the basis for our research which is used for training and testing.

$$\{A\} = (\{b\}\cup\{c\}\cup\{d\}\cup\{e\}\cup\{f\}\cup\{g\}). \tag{1}$$

where b, c, d, e, f, and g are dataset provided from 2015, 2016, 2017, 2018, 2019 and 2020 respectively by the Criminal Investigation Department of the Nigeria Police Force, Abeokuta.

3.2 CLUSTER CLASSIFICATION OF RESEARCH AREA

The research area for the study is Ogun state, Nigeria. It shares boundaries with Oyo, Lagos, Republic of Benin and Ondo states towards the north, south, west and east respectively. The state is located within latitudes 6°N-8°N and longitudes 3°E -5°E with twenty local government areas. The towns and villages located in each local government area of the state serve as a cluster for use in this research.

3.3 MATHEMATICAL MODELLING - PRINCIPAL COMPONENT ANALYSIS

From the derived Unified-Multi Data (UMDat) dataset $\{A\}$, a feature extraction approach referred to the Principal Component Analysis (PCA) is used to create a less complex dataset which is easily interpretable to enhance the prediction of the Model.

Definition 1: *n* - dimensional vector

A set of independent Definition 1: *n*-dimensional vector variables α associated with each dataset collected in the set (A) with *n*-attributes collected to form set A as shown in Eq.(2).

$$\alpha = (\alpha_1, \alpha_2, \alpha_3, \ldots, \alpha_n) \tag{2}$$

where α - independent variable attributes

Definition 2: Class Vector

Suppose that a class vector denoted by β is assigned to the attributes generated from the α - dimensional vectors collected in the dataset in Eq.(2).

Hence, the class vector is a dependent variable used to determine the initial safety level for the training set. It is represented in Eq.(3) below.

$$\{(\beta_1, \beta_1), (\beta_2, \beta_2) \ldots (\beta_n, \beta_n)\} \tag{3}$$

Definition 3: Dimensionality reduction

Given that the Unified-Multi Data (UMDat) dataset with n-dimensional random variables α (Eq.(2)) then a lower dimensional representation of α called the dimensionality reduction can be derived using the Eq.(4)

$$R = \{r_1, r_2, r_3, \ldots, r_k\}, \text{ with } k \leq n \tag{4}$$

where R captures the data attributes (α) in the Unified-Multi Data (UMDat) $\{A\}$, according to variance maximization criterion. Elements in R are the hidden components.

In this, the PCA maps data from the space of p-variables to a new space of un-correlated k variables over the dataset.

$$R_k = W^T\alpha_1 + \ldots + W^T\alpha_n \tag{5}$$

Eq.(5) depicts the dimension reduction linear transformation function used to derive the new set r from the set (A) as a result of the linear transformation of attributes α such that $k \leq n$ and W^T is the linear transformation weight matrix. The elements in R_k are the extracted factors of k-size and α_n is the original factors of n-size

3.4 ARCHITECTURAL FRAMEWORK

The architectural framework for the Neuro-fuzzy model employs 263 numbers of nodes, 560 linear parameters, 32 Non-linear parameters, total numbers of parameters is 592, training data pairs are 263, and checking data pairs are 79 and 112 fizzy rules to predict recovery location of stolen cars. The architecture in Fig.1 consists of two major phases: Training and Testing phases. In the training phase feature extraction is first applied to the raw criminal data to remove noise and then the resultant data is stored in memory as the training dataset.

The ANFIS model then uses the training dataset for learning. In the testing phase the ANFIS algorithm is presented with a new instance for classification; at run- time, the ANFIS performs an alignment with the memory where the training dataset are stored using the Euclidean distance function and finally classify the new instance into the closest neighbor in the training dataset.

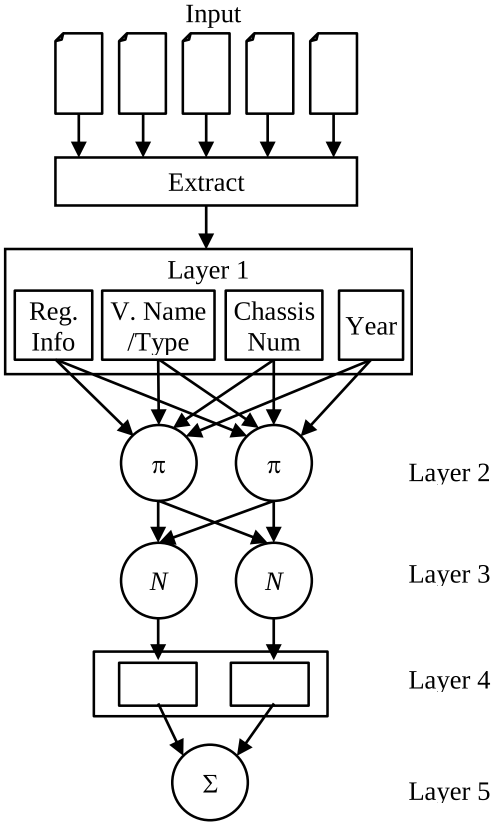

Fig.1. Ensemble Neuro-Fuzzy Model for vehicle theft prediction and recovery

3.5 PSEUDO CODES: FEATURE SELECTION AND PREDICTION ALGORITHMS

The pseudocode relying on the Principal Components Analysis (PCA) used in extracting important factors and attributes for the prediction of car theft safety is indicated in Algorithm 1. A reduced training data model adopted by this algorithm maps four input variables to their respective membership functions, membership functions to associated rules. These rules are linked to a set of output variables with associated membership functions. The membership functions to a single –valued output for decision making. This determines the predicted recovery location for vehicles thus enabling security agencies channel a greater percentage of resources, personnel and effort towards predicted location in attempt to recover stolen vehicles while an appreciable portion is also directed towards other location in a bid to hasten the recovery process.

A combination of Least Square and the Back propagation algorithm is used to train the Adaptive Neuro-Fuzzy Inference system at vary levels of epoch until the best result is achieved.

Algorithm 1: Feature Selection

Input: Dataset; $\alpha_1, \ldots, \alpha_n$ *n*-dimension vector, α

Output: $R_1, \ldots, R_k$; *k*-dimension vector, $R \ni k \leq n$

Begin

{

$\alpha \leftarrow n \times k$ data matrix with α in each row

$\alpha \leftarrow \frac{1}{n}\sum_{i=1}^{n}\alpha_i$ -α_i in α //from each row

$$COV \leftarrow \frac{1}{n-1}\alpha \times \alpha$$

Compute eigenvalue $e_1,\ldots, e_n$ of *COV* and sort them

Compute matrix *V* which satisfies $V^{-1}\times COV\times V=D$ //*D* representing the diagonal matrix of eigenvalue of COV

Return *k*-dimension// the first *k* column of *V*

}

End

Algorithm 2: Model Training

Input: *r*(*k*)

Output: *Y*(*n*)

Begin

{ ϶

Initialization. Set

$W_k(1) = 0$ for $i = 1,2,\ldots,m$

Filtering for time $n = 1,2, \ldots$, compute

$Y(n) = \sum_{i=1}^{p} w_1(n) r_1(k)$

$e(n) = d(n)-y(n)$

$w_k(n+1) = w_k(n)+e(n)r_k(n)$ for $i=1,2,\ldots,m$

while (stopping criteria=true) do

//forward pass

For i = length of r_j in *N*

For j = length of nodes in *N*

F(train): $i\rightarrow j$

$a_j = \sum_{i=1}^{p} r_j w_{ji}$ workout the net inputs

$f(a_j)=(1+\exp^{-aj})^{-1}$

}

{

// backward pass of output nodes

Filtering

// backward pass of hidden nodes

$d_j = f^1 j(a_j)\sum_{k=1}^{K}(d_k - W_{kj})$ // error of hidden nodes

$\Delta W_{jh}(n+1)=x_h+d_h\cdot\Delta W_{jh}$

}

Return *Y*(*n*)

End

Algorithm 3: ANFIS Prediction

Input: $r_1,.., r_k$, *k*-dimension vector, α ϶ $k\leq n$

Output: *Y*, Prediction Value

Begin

{϶

$R \leftarrow r_1,.., r_k$ // load reduced data set

Generate FIS

#For each parameter $q\in\alpha$ do

Weight = $w_1z_1+(w_nz_n/w_1z_1)$// hybrid Train network

RMSE = *MSE* //test the FIS model

$MSE = \frac{1}{n}\sum_{t=1}^{n}(F_t - y_t)^2$

Return *Y*

}

END

4. IMPLEMENTATION, RESULTS AND DISCUSSION

The training dataset, being a derived set from the collection of a Unified-Multi Data (UMDat) contains four factors deployed as input into the Neuro-Fuzzy vehicle safety prediction model in relation to recovery location as the output. This dataset is divided into three subgroups, pre-processed in matrix form with four input columns and a single output column. The factors as contained in the dataset include name, address, vehicle name and type, engine number, chassis number, theft location, month and recorded year of theft incidence.

4.1 NEURO-FUZZY INFERENCE GENERATION, MODEL TRAINING AND TESTING

Fuzzy Inference System (FIS) is generated with the specified number of input parameters sets to include the required number of membership functions and types. The model is trained until the desired result is obtained with the least minimum error.

In this paper, zero (0) was selected for the error tolerance and an epoch of hundred (100) for a more accurate and prediction. Testing data are also loaded into the model

4.2 ENSEMBLE NEURO-FUZZY MODEL BASED SYSTEM VALIDATION

Dataset used for validating the model is also loaded from the MATLAB workspace into the system. It is depicted as balls super imposed on both training and testing dataset with the aim of testing the efficiency of the system at each epoch. A list of 112 different rules generated by the Neuro-Fuzzy model for the prediction of vehicle theft recovery level and location is obtained. An automatic generated rule editor allows for modification of the input factors combined with if- then (construct) rules for accurate prediction. An ensemble Neuro-Fuzzy model structure with assigned membership function and with rule values logically represented in tabular form.

4.3 MACHINE LEARNING PREDICTIONS, RESULT ANALYSIS AND EVALUATION

It was established that time of the day, place and security precautions attached to vehicles are keys factors which could help determine the safety of cars The comparison of accuracy prediction of the ensemble Neuro- fuzzy model with other machine learning algorithms (such as J48, Naïve Bayes, and

Random Forest) embedded in WEKA JPI on the transformed range of data from the Unified –Multi

Data (UMDat) set predicted varying degree of correctness and errors. Evaluations based on well-known evaluation metrics including the Relative Absolute Error (RAE), Root Mean Square Error (RMSE), Mean Absolute Error (MAE) and Root Relative Square Error (RRSE) was also analyzed. The results of the accuracy comparison are depicted as shown in Fig.2 with error values recorded by other commonly used machine learning algorithms as given in Table.1.

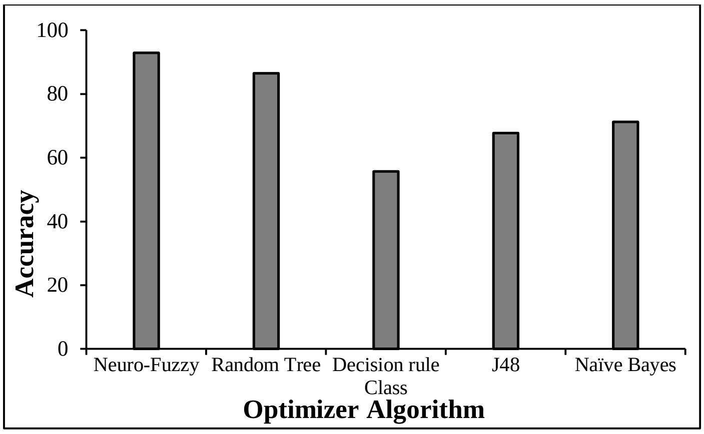

Fig.2. Model Evaluation Comparison

Table.2. Error Metric Values

Algorithm	RAE (%)	RMSE (%)	MAE (%)	RRSE (%)
J48	37.37	0.7757	0.036	80.05
Random tree	39.40	0.1158	0.038	52.80
Naïve Bayes	38.10	0.1476	0.037	67.25

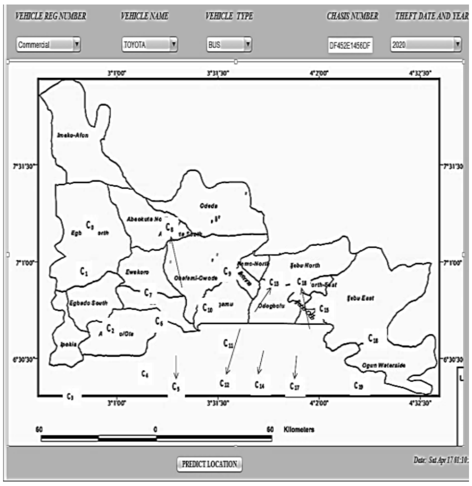

Fig.3. Vehicle theft Prediction input data and attribute interface

Furthermore, a user-friendly decision system program with a graphical interface is developed using Java programming language to aid users, most especially security personnels concerned with the prediction and recovery of stolen vehicles in their operations.

The Fig.3 and Fig.4 are interfaces of the Java developed Ensemble Neuro-Fuzzy based decision support system for vehicle theft prediction and recovery. Users can input attributes (registration number, vehicle name, type, chassis number, year of theft) associated with the stolen vehicles in Fig.5 and generate sample predicted values as shown in Fig.6 for prompt actions towards recovery of the stolen vehicle.

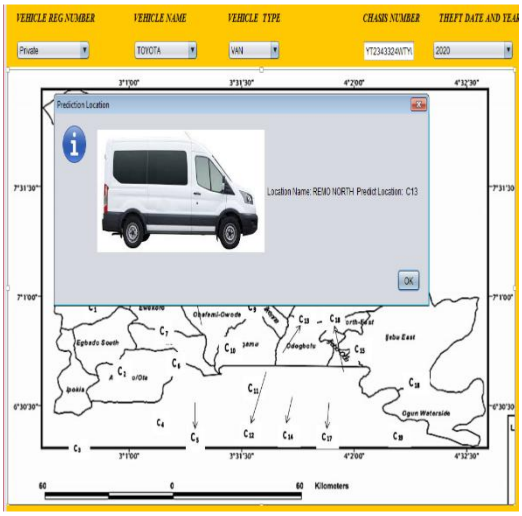

Fig.4. Predicted Recovery location for stolen vehicles

5. CONCLUSION

The report of vehicle theft is no longer strange news; it has become a major constituent (headline) of broadcast news on timely basis not only in Nigeria but in the world at large. Criminals who engage in this activity have varied means and manner for performing their operations. They do not only target some specific periods and times of the day but also prefer quiet locations to enable them go un-noticeable after breaking security measures fixed in vehicles. However, despite the attempt and efforts made by security personnel, only about 32% of stolen cars are tracked, recovered and returned to the rightful owners.

Hence, the need for the adopting a more accurate technique for predicting recovery locations of stolen vehicles, increase detection rate and reducing the incidents of car thefts especially when criminals are aware that there are intelligent systems which can be deployed to easily track them.

Well known classifiers including Naïve Bayes, J-48 and others used in this research showed some level of accuracy but the prediction rate of 92.91 % accuracy depicted by the novel ensemble Neuro-Fuzzy model surpassed others. Thus, indicating that the Neuro- fuzzy model will be a better choice for the accurate predictions when compared with other classifiers.

Furthermore, it is therefore recommended that adequate security precautions be put in place for vehicles regardless of the time or places where they are kept or parked. Adoption of this Neuro-Fuzzy based system will aid security personnels take appropriate decision at given times in order to reduce car theft incidents, minimize investigation time, man power and energy in predicting cases associated with vehicle theft and recoveries.

REFERENCES

[1] Adewumi Isreal Badiora, “Ecological Theories and Spatial Decision Making of Motor Vehicle Theft (MVT) Offenders

in Nigeria", *Journal of Applied Security Research*, Vol. 12, No. 3.pp. 374-391, 2017.

[2] M.A. Andresen, A.S. Curman and S.J. Linning, "The Trajectories of Crime at Places: Understanding the Patterns of Disaggregated Crime Types", *Journal of Quantitative Criminology*, Vol. 33, pp. 427-449, 2017.

[3] Andrew P. Wheeler and Wouter Steenbeek, "Mapping the Risk Terrain for Crime Using Machine Learning", *Journal of Quantitative Criminology*, Vol. 13, pp. 1-36, 2020.

[4] Assessment of Auto Vehicle Theft, "Auto Theft Intelligence Coordination Center (ATICC) Report", Available at http://www.rmiia.org/downloads/ATICC-Annual-Report-2019.pdf, Accessed at 2020.

[5] A. Badiora, "Motor Vehicle Theft: An Examination of Offenders' Characteristics and Targeted Locations in Lagos, Nigeria", *Kriminoloji Dergisi: Turkish Journal of Criminology and Criminal Justice*, Vol. 4, No. 2, pp. 59-70, 2012.

[6] S. Bai, Z. Liu and C. Yao, "Classify Vehicles in Traffic Scene Images with Deformable Part-Based Models", *Machine Vision and Applications*, Vol. 29, No. 3, pp. 393-403, 2018.

[7] Y. Bengio, A. Courville and P. Vincent, "Representation Learning: A Review and New Perspectives", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 8, pp. 1798-1828, 2013.

[8] J. Bergstra and Y. Bengio, "Random Search for Hyper Parameter Optimization", *The Journal of Machine Learning Research*, Vol. 13, No. 1, pp. 281-305, 2012.

[9] A. Braga and R. Clarke, "Explaining High-Risk Concentrations of Crime in the City: Social Disorganization, Crime Opportunities, and Important Next Steps", *Journal of Research in Crime and Delinquency*, Vol. 51, pp. 480-498, 2014.

[10] S. Das and N. Roy, "Applications of Artificial Intelligence in Machine Learning: Review and Prospect", *International Journal of Computer Applications*, Vol. 9, pp. 115-129, 2015.

[11] Deborah Lamm Weisel, William R. Smith, G. David Garson, Alexi Pavliche and Julie Wartell, "Motor Vehicle Theft: Crime and Spatial Analysis in a Non-Urban Region", Available at https://nij.ojp.gov/library/publications/motor-vehicle-theft-crime-and-spatial-analysis-non-urban-region, Accessed at 2006.

[12] J. Douglas, A.W. Burgess and R.K. Ressler, "Crime Classification Manual: A Standard System for Investigating and Classifying Violent Crime", John Wiley and Sons, 2013.

[13] G. Drawve, S.A. Thomas and J.T. Walker, "Bringing the Physical Environment Back into Neighborhood Research: The Utility of RTM for Developing an Aggregate Neighborhood Risk of Crime Measure", *Journal of Criminal Justice*, Vol. 44, pp. 21-29, 2016.

[14] Eric L. Piza and Jeremy G. Carter, "Predicting Initiator and Near Repeat Events in Spatiotemporal Crime Patterns: An Analysis of Residential Burglary and Motor Vehicle Theft", *Justice Quarterly*, Vol. 4, pp. 1-30, 2017.

[15] P. Gera and R. Vohra, "City Crime Profiling using Cluster Analysis", *International Journal of Computer Science and Information Technologies*, Vol. 5, No. 4, pp. 5145-5148, 2014.

[16] Gohar Petrossian and Ronald V. Clarke, "Export of Stolen Vehicles Across Land Borders," Problem-Oriented Guides for Police Problem-Specific Guides", Available at https://popcenter.asu.edu/sites/default/files/problems/pdfs/export_stolen_vehicles.pdf, Accessed at 2018.

[17] M. Goyal, V. Bhatnagar and A. Jain, "A Classification Framework for Data Mining Applications in Criminal Science and Investigation", *Proceedings of International Conference on Data Mining Trends and Applications in Criminal Science and Investigations*, pp. 32-51, 2016.

[18] Hui Yang, Yia-Tang Fuya and Diavi Yang, "Multi-ANFIS Model Based Synchronous Tracking Control of High-Speed Electric Multiple Unit", *IEEE Transactions on Fuzzy Systems*, Vol. 26, No. 3, pp. 1472-1484, 2019.

[19] H. Huttunen, F.S. Yancheshmeh and K. Chen, "Car Type Recognition with Deep Neural Networks", *Proceedings of International Conference on Intelligent Vehicles*, pp. 1115-1120, 2016.

[20] Hyeon-Woo Kang and Hang-Bong Kang, "Prediction of Crime Occurrence from Multi-Modal Data using Deep Learning", *PLoS ONE*, Vol. 12, No. 4, pp. 200-215, 2017.

[21] O.E. Isafiade, A. Bagula and S. Berman, "On the Advancement of using Data Mining for Crime Situation Recognition: A Comparative Review", *Proceedings of International Conference on Data Mining Trends and Applications in Criminal Science and Investigations*, pp. 1-31, 2016.

[22] N. Jain, P. Sharma and D. Kalbande, "Computerized Forensic Approach using Data Mining Techniques", *Proceedings of the ACM Symposium on Women in Research*, pp. 55-60, 2016.

[23] D. Jomaa and M. Dougherty, "Speed Prediction for Triggering Vehicle Activated Signs", *Corpus Engineering Journal*, Vol. 6, No. 2, pp. 1-16, 2016.

[24] L.W. Kennedy and H. Buccine Schraeder, "Vulnerability and Exposure to Crime: Applying Risk Terrain Modelling to the Study of Assault in Chicago". *Applied Spatial Analysis and Policy*, Vol. 9, pp. 529-548, 2016.

[25] P.K. Kim and K.T. Lim, "Vehicle Type Classification using Bagging and Convolutional Neural Network on Multi View Surveillance Image", *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition Workshops*, pp. 41-46, 2017.

[26] Y.A. Kim and J.R. Hipp, "Street Egohood: An Alternative Perspective of Measuring Neighborhood and Spatial Patterns of Crime", *Quant Criminology*, Vol. 36, pp. 29-66, 2020.

[27] D. Kuang, P.J. Brantingham and A.L. Bertozzi, "Crime Topic Modelling", *Crime Science*, Vol. 6, pp. 1-20, 2017.

[28] Mrinalini Jangra and Shaveta Kalsi, "Crime Analysis for Multistate Network using Naive Bayes Classifier", *International Journal of Computer Science and Mobile Computing*, Vol. 8, No. 6, pp. 134-143, 2019.

[29] N. Morgan, O. Shaw, A. Feist and C. Byron, "Reducing Criminal Opportunity: Vehicle Security and Vehicle Crime", Available at https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/489097/horr87.pdf, Accessed at 2016.

[30] Nesquivel, O. Nicolis and B.P. Márquez, "Predicting Motor Vehicle Theft in Santiago De Chile using Graph-Convolutional LSTM", *Proceedings of International Conference on Chilean Computer Science Society*, pp. 1-7, 2020.

[31] Y. Peng, J.S. Jin and Y. Cui, "Vehicle Type Classification using Data Mining Techniques", *Proceedings of International Conference on Era of Interactive Media*, pp. 325-335, 2013.

[32] Y. Peng, J.S. Jin and S. Luo, "Vehicle Type Classification using PCA with Self-Clustering", *Proceedings of International Conference on Multimedia and Expo Workshops*, pp. 384-389, 2012.

[33] B.M. Prakoso, "The Efforts of the Sabhara Unit Patrol Unit in Preventing Crime of Motor Vehicle Theft in Sumedang Police Jurisdiction: Upaya Unit Patroli Satuan Sabhara dalam Mencegah Tindak Pidana Pencurian Kendaraan Bermotor di Wilayah Hukum Polres Sumedang", *Indonesian Journal of Police Studies*, Vol. 5, No. 1, pp. 17-34, 2021.

[34] Shirota Shinichiro, Alan. E. Gelfand and Jorge Mateu, "Analyzing Car Thefts and Recoveries with Connections to Modelling Origin", *Proceedings of International Conference on Destination Point Patterns*, Vol. 3, pp. 1-28, 2020.

[35] G. Song, W. Bernasco, L. Liu and W. Liao, "Crime Feeds on Legal Activities: Daily Mobility Flows help to Explain Thieves' Target Location Choices", *Journal of Quantitative Criminology*, Vol. 35, pp. 831-854, 2019.

[36] W. Sun, X. Zhang and Y. Jin, "Vehicle Type Recognition Combining Global and Local Features Via Two-Stage Classification", *Mathematical Problems in Engineering*, Vol. 2017, pp. 134- 145, 2017.

[37] N.V. Kousik, "Analyses on Artificial Intelligence Framework to Detect Crime Pattern", *Proceedings of International Conference on* Intelligent Data Analytics for Terror Threat Prediction: Architectures, Methodologies, Techniques and Applications, pp. 119-132, 2021.

[38] Tayal, D. K., Jain, A., Arora, S., Agarwal, S., Gupta, T., and Tyagi, N. "Crime detection and criminal identification in India using data mining techniques", AI and SOCIETY, vol. 30(1), pp. 117-127, 2015.

[39] K. Praghash and T. Karthikeyan, "An Investigation of Garbage Disposal Electric Vehicles (GDEVs) Integrated with Deep Neural Networking (DNN) and Intelligent Transportation System (ITS) in Smart City Management System (SCMS)", *Wireless Personal Communications*, Vol. 23, pp. 1-20, 2021.

[40] Xingyang Ni and Heikki Huttunen, "Vehicle Attribute Recognition by Appearance: Computer Vision Methods for Vehicle Type, Make and Model Classification", *Journal of Signal Processing Systems*, Vol. 34, pp. 1-12, 2020.

[41] Xiangyu Zhao and Jiliang Tang "Exploring Transfer Learning for Crime Prediction", *Proceedings of International Conference on Data Mining*, pp. 1-3, 2017.

[42] Z. Xu, W. Yang and L. Huang, "Towards End-to-End License Plate Detection and Recognition: A Large Dataset and Baseline", *Proceedings of European Conference on Computer Vision*, pp. 255- 271, 2018.

[43] L. Yang, P. Luo and X. Tang, "A Large-Scale Car Dataset for Fine-Grained Categorization and Verification", *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition*, pp. 3973-3981, 2015.

[44] Y. Zhang and L. Hoover, "Space-Time Clustering of Crime Events and Neighborhood Characteristics in Houston", *Criminal Justice Review*, Vol. 40, pp. 340-360, 2015.

[45] Shiwen Zhang, Yingying Xing, Jian Lu and H. Michael Zhang, "Exploring the Influence of Truck Proportion on Freeway Traffic Safety using Adaptive Network-Based Fuzzy Inference System", *Journal of Advanced Transportation*, Vol. 2019, pp.1-14, 2019.