ORIGINAL RESEARCH

Implementation of an intelligent clustering methodology for classification of terrorist acts

Akinyokun Oluwole C*[1], George Uduak D[2]

[1] *Department of Software Engineering, Federal University of Technology, Akure, Nigeria*
[2] *Department of Computer Science, University of Uyo, Nigeria*

Received: February 11, 2019 **Accepted:** March 28, 2019 **Online Published:** June 5, 2019
DOI: 10.5430/air.v8n1p61 **URL:** https://doi.org/10.5430/air.v8n1p61

ABSTRACT

Terrorist acts have elevated the level of violence, intimidation and pose a threat to life/property, peace and security in the world today. Deployed solutions to curb the occurrence of terrorism prove to be of insignificant value, hence there is the need for more solutions. The research aims at implementing an intelligent clustering methodology for classification of the acts of terrorism in Nigeria. Three experiments were carried out. In the first experiment, the qualitative terrorists data attributes were converted to quantitative attributes using an existing One-of-N (OoN) method and the processed data supplied to Adaptive Neuro-Fuzzy Inference System (ANFIS) (OoN-ANFIS) for training. The second experiment converted the qualitative data attributes to quantitative attributes using the formulated Rank-Frequency-Based (RFB) model before the data was supplied to ANFIS (RFB-ANFIS) for training. In the third experiment, which constitutes the current study, the RFB-processed data was used by Fuzzy C Means (FCM) to generate initial membership values for each point in the data set and then supplied to ANFIS (RFB-FCMANFIS). The results show that RFB-FCMANFIS model generated the least Root Mean Square Error (RMSE), Mean Absolute Error (MAE), training error and checking error of 0.002887, 0.004598, 0.0000713 and 0.0056155 respectively with the highest correlation coefficient of 0.99954, therefore indicating a superior classification capability using the RFB-FCMANFIS.

Key Words: Classification, Rank-Frequency-Based (RFB) model, OoN-ANFIS, RFB-ANFIS, RFB-FCMANFIS, Terrorist acts, Terrorism, Clustering, Neural network

1. INTRODUCTION

Terrorism is an act of violence, which instills fear, causes serious injuries and destroys lives and properties of the citizens. The main aim of terrorism is to force the government or organizations to follow a particular line of action dictated by the terrorists. It is observed that adequate use of scientific approaches is yet to be fully deployed towards the fight against terrorism. The lack of comprehensive database of various terrorism incidents accounts for the lack of knowledge on terrorism. It is also observed that the prevention of acts of terrorism is very challenging since the use of force to fight terrorism has proved to be of very insignificant value as terrorists are ready to die for what they believe. Furthermore, investigation, analysis and classification of terrorist acts yield little or no success because of the vagueness of its attributes in addition to uncertainty, confusion and varying tactics that often characterize the acts.[1]

Governments have employed preemptive, offensive and defensive measures as well as national, regional and international collaborations to further the fight. Law enforcement

*Correspondence: Akinyokun Oluwole C; Email: admin@akinyokun.com; Address: Department of Software Engineering, Federal University of Technology, Akure, Nigeria.

agencies devised the use of science, technology, intelligence gathering and other scientific and technological equipment. Some combating hardware devices such as Ultra Wide Band (UWB), Electromagnetic Pulse (EMP), active denial system, tear gas, sleep gas, psycho-chemicals and directed energy weapon which are very effective if the location of the terrorists and government agents are the same in.[2–4] The classification of terrorist acts is based on discovery of patterns which are useful information that was previously hidden from the owner or the public. The generated patterns drive decision support systems to assist humans in effective decision making.

Artificial Neural Network (ANN) and Fuzzy Logic (FL) are machine learning tools that rely on numeric data type to function and can learn from historical and operational data. ANNs are fault-tolerant and are basically used in situations where known information is used to infer some unknown information. FL on the other hand is very good at handling vague, ambiguous or imprecise information as well as being a powerful modeling tool for complex systems with high level of uncertainties and partial truths, which are characteristics of terrorist behaviors and activities.[5] The application of machine learning techniques in prediction, classification and clustering has been widely reported. In,[6] fuzzy ontology was used to extract terrorism events, while classification of terrorism events using Fuzzy Inference System (FIS) and Adaptive Neurofuzzy Inference System (ANFIS) was carried out in.[7] In order to evaluate supply chain management, the ANFIS is used in.[8] In,[9] a hybrid knowledge discovery system for oil spillage risks pattern classification using ANFIS is proposed. In,[10] a co-clustering approach to extract patterns from the global terrorism dataset is used. In,[11] clustering is identified as a potent tool in counter-terrorism and an enhanced clustering algorithm for use in the fight against terrorism in the society is developed. In,[12] FCM to analyze forensic data is proposed. In,[13] an approach for countering terrorism using soft computing model known as Competitive Neural Tree (CNeT) was adopted. In,[14] fuzzy logic driven expert system for the diagnosis and classification of heart failures was presented.

This work aims at implementing an intelligent clustering methodology for classifying terrorist acts designed in.[1] The specific objective is to implement a model that assigns numerical weights to qualitative terrorists attributes and also carry out a comparative analysis of the model with other models for classification of terrorist acts. Section 2.0 carries out a review of related works while the implementation of the clustering technique for classification of terrorist acts is presented in Section 3.0. Some conclusions are drawn in Section 4.0.

2. Related works

In,[15] hybrid classification algorithms for terrorism prediction in Middle East and North Africa was presented. The research was conducted in three major steps namely data reduction, data removal and experimentation using Waikato Environment for Knowledge Analysis (WEKA) software for the prediction and classification of terrorist groups. The research yielded unsatisfactory predictive accuracy and is not suited for uncertain, inaccurate, incomplete and complex situations like terrorism. In,[16] an experimental study of classification algorithms for prediction of terrorism was presented. Although the work offered approaches for handling missing data, compared and evaluated different classification algorithms implemented in WEKA such as Naïve Bayes, K-Nearest Neighbour, Tree Induction, Iterative Dichotomiser and Support Vector Machine, it was limited by its inability to incorporate methods that can enhance classification of terrorist groups. In,[7] FISs to classify terrorism events in four major steps; feature selection, information extraction, transformation of text data to list of vectors and construction of classification model is adopted. The research indicated that ANFIS is efficient in the prediction of terrorism events but could not provide accurate classification of tactics of terrorists.

In,[17] a classification technique for bioinformatics problems using ANFIS and FCM clustering in an attempt to memorize the data pattern supplied as input to the developed system is formulated. In,[9] a hybrid knowledge discovery system for classification of oil spillage risks pattern is developed and applied. The research adopted NN, FL and GA to develop an intelligent hybrid system that assisted in identification, extraction and classification of oil spillage risk patterns. Its inability to perform a comparative analysis of ANFIS performance with GA as the learning algorithm was a major limitation. In,[18] an adaptive neuro-fuzzy model with fuzzy clustering was proposed for nonlinear prediction and control where the FCM clustering was used to initialize the ANFIS. The research contributed an ANFIS model for nonlinear prediction and control but was unable to model the uncertainties and ambiguity inherent in engine operations and could not learn or recall the initial states of the system. In,[19] segmentation and classification of calcification and hemorrhage in the brain using FCM and ANFIS was proposed. The research developed a method that segmented and classified brain hemorrhage using contrast, correlations, energy and homogeneity as input features. Although, it provided a means of significantly classifying brain abnormalities, it lacked the ability to deal with the uncertainty and vague measurements of brain abnormalities inputs.

In,[20] a classification model for adaptive neuro-fuzzy inference system for prediction of heart diseases and classification of the degree of heart diseases using ANFIS was developed. The research was limited by the inadequate input parameters that can be used for different databases and inability to initially partition the dataset in order to present a preliminary pattern for the classifier. A subjective weighting method based on group decision making for ranking and measuring criteria values was presented in.[21] That work contributes a subjective weighting method used to generate relative weights to different criteria. The research suffers from the fact that the number of decision makers that respond in the ranking and scoring of the criteria is inadequate to present a reasonable and dependable judgment. It also lacks standardized values for the weights of the criteria as it used group total which is not a good representation of a set of data. In,[22] NN as a classification method in the behavioral sciences is presented. Data for that research was gathered from the groans of four fallow deer bucks which were tape-recorded during the rut. That work contributes a classification method for behavioral sciences using NN but lacks the facility to deal with the problem of imprecision in the vocalization patterns of the animal under study.

3. Implementation technique of the proposed system

The data for this research was collected from several sources such as electronic media, print media and most comprehensively from the Global Terrorism Database (GTB) 2017 of the National Consortium for the Study of Terrorism and Responses to Terrorism (START) hosted by the University of Maryland Center of Excellence, United States of America. The collected data spanned the period from 1983 to 2017. Three thousand, nine hundred and five (3,905) records with nine (9) attributes concerning Nigeria were selected with six of the attributes being quantitative (numeric) and three being qualitative (non-numeric).

3.1 Formulation of weight assignment model for qualitative data attributes

Fuzzy logic and neural networks are best suited for handling problems that are quantitative in nature. Solutions offered are often inhibited by qualitative data. However, there are several phenomena in nature that are not possibly expressed numerically. For instance, in terrorism, several qualitative data such as tactics used by terrorists, weapons used for attacks, the type of target/victims of attacks characterize the data set in addition to a few other quantitative data. Therefore, numerical representation of such qualitative data, taking into consideration their relative significance in the overall system, poses some challenges. A case that needs attention involves taking decisions in terms of prioritization and assignment of weights to several qualitative criteria and sub-criteria. There is the need to advance the application of neuro-fuzzy models to operate on qualitative data by applying some transformation techniques. In,[21] a subjective weighting method for ranking and measuring criteria values based on group decision making was presented and the criteria value was calculated as follows:

Criterion value = (frequency of the same rank for each criterion * sum of scores of the criterion in the same rank) + scores of other rank

It was observed that obtaining criteria weights from several expert decision makers was very difficult. The value of a criterion was obtained in terms of the total scores obtained from the same rank. However, in this work, we propose the use of the frequency of assigning the same rank to a criterion and the mean of scores awarded to a criterion in the same rank, to evaluate the value of the criteria, instead of the frequency of the same rank and sum of scores of the criterion in the same rank. The mean value is preferred because it is statistically established that a group of scores can best be represented by the measures of central tendency and not the total. The mean also possesses the minimum variance property and hence its preference in this work.

It must be understood that if the number of decision makers in a multi-criteria decision making process is large, there are possibilities that two or more respondents give the same ranking to a criterion although they may assign different scores to the criterion. Hence, the frequency of rankings must be taken into consideration. Suppose there are n qualitative criteria for a data attribute available for decision making and suppose a number of decision experts are given the assignment to rank each criterion in order of priority (importance) by giving scores based on a given scale. Based on these rankings and frequencies, a model for assigning numerical weights to the different criteria is needed. Taking all the variables into account, a Rank-Frequency-Based (RFB) model for obtaining the numerical value of a criterion is formulated as:

$$V_i = \sum_{k=1}^{K} f_k \bar{x}_k , \quad i = 1,2,\dots,n \tag{1}$$

where V_i is the value of the ith criterion, k is the assigned rank, f_k is the frequency of respondents assigning rank k, $\bar{x_k}$ is the mean of all the awarded scores under rank k, n is the total number of criteria and K is the total number of ranks. In other words, the value of the ith criterion is the sum of the products of the frequencies of respondents assigning the same rank and mean of scores in that rank under the ith criterion. Hence, the weight of the ith criteria (W_i) is the normalized value given as:

$$W_i = \frac{v_i}{\sum_{i=1}^{n} v_i} \tag{2}$$

3.2 Weight assignment to qualitative terrorists data attributes

The formulated model is used to assign numerical weights to qualitative data attributes. Questionnaires were administered to the officers of the Anti-Terrorism Squad (ATS), a unit of the Nigeria Police Force, through the Akwa Ibom State Police Headquarters. Every Nigeria Policeman, where ever he/she is posted, had passed through the same training and as such is well equipped with the necessary know how about crimes as well as terrorism matters in the case of the ATS officers. Therefore, since the variables and the constructs of the questionnaire were easy to understand by any officer of the ATS, a convenience sampling[23,24] technique was adopted in the administration of the questionnaire. The questionnaire was constructed with two sections. The first section had information on the respondent personal data. The second section has three tables representing the tactic deployed by the terrorists, weapon type used and victim type. In the first table, there are eight variables for tactics namely assassination, hijacking, kidnapping, hostage taking, bombing, armed assault, unarmed assault and infrastructure attack. The second table has eleven variables for weapon type namely biological, chemical, radiological, nuclear, firearms, explosives/bombs/dynamite, fake weapons, incendiary, melee, vehicle and sabotage equipment. The third table has eighteen variables for the victim type namely business, government, police, military, aircraft, diplomatic, educational institution, food or water supply, journalists and media, maritime, NGO, private citizens and property, religious figures/institutions, telecommunication, tourists, transportation (non-aviation), utilities and political party. A well-known scoring scale of 1 to 100 was chosen for the respondents to use in scoring each of the variables. This scale makes it easy for the respondents to rank the variables based on severity rating of the terrorist act through numbers. The severity increases from 1 (least severe) to 100 (most severe). The questionnaire is validated by experts, affirming that the variables and construct of the survey is adequate in solving the intended problem.

To commence the research process, one hundred questionnaires were administered. Fifty were returned properly completed, sixteen were not returned and thirty four were wrongly filled. The respondents, scores and ranks assigned to the tactic called Assassination and color-coded are depicted in Figure 1. The frequencies of respondents that ranked Assassination first, second, third, fourth, fifth, sixth, seventh and eighth are 30, 2, 5, 3, 3, 5, 1 and 1 respectively depicted in Figure 2. Their mean scores are 65.37, 73, 49.80, 44.33, 59.33, 47.40, 50 and 40 as depicted in Figure 3. Other tactics of terrorist acts were similarly scored and ranked.

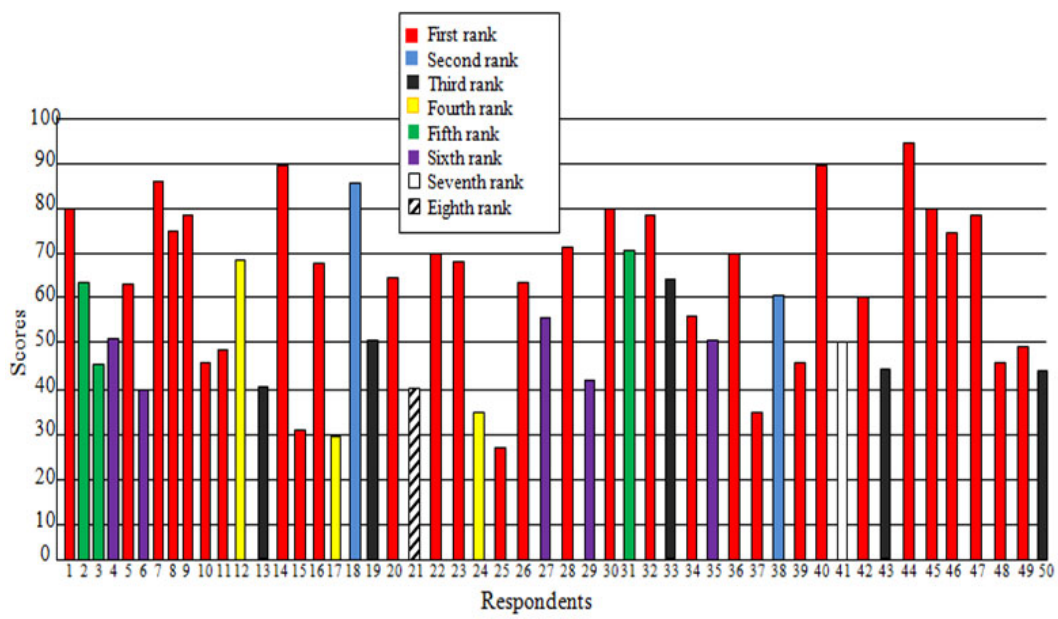

Figure 1. Respondents with Scores and Ranks for Assassination Tactic

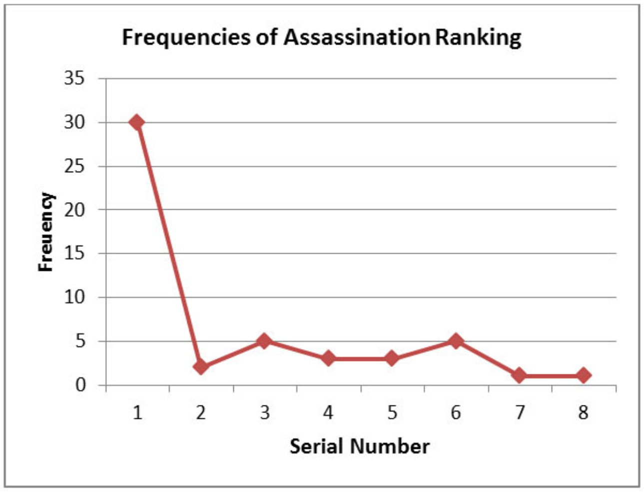

Figure 2. Frequency of Respondents Ranking Assassination

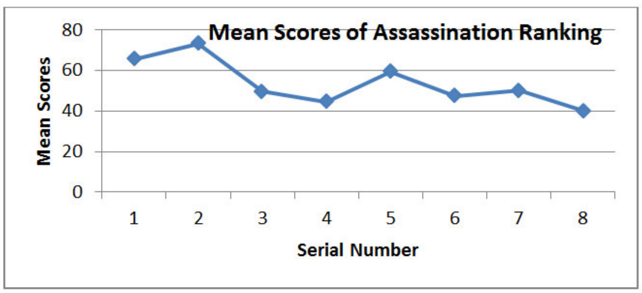

Figure 3. Mean Scores for Each Assassination Rank

Table 1. Values and Weights of Tactics Attribute

S/N	Attribute	Value	Weight
1	Assassination	2994	0.1468
2	Bombing	2720	0.1333
3	Kidnapping	2743	0.1345
4	Hostage taking	2682	0.1315
5	Hijacking	2813	0.1379
6	Armed Assault	2525	0.1238
7	Unarmed Assault	1995	0.0978
8	Infrastructure Attack	1928	0.0945
	Total	20400	

The values of the attributes are computed using Equation (1). For example, the value of Assassination tactic is computed as 30(65.3667) + 2(73) + 5(49.8) + 3(44.3333) + 3(59.3333) + 5(47.4) + 1(50) + 1(40) summing up to 2994. The total of all the tactic values is 20400 and the weight is calculated using Equation 2. Therefore, the weight of Assassination tactic, computed as 2994/20400 is 0.1468. The values for bombing, kidnapping, hostage taking, hijacking, armed assault, unarmed assault and infrastructure attack tactics are similarly computed and their weights generated as presented in Table 1. The computation for the weapons type and victim type attributes were similarly carried out and presented in Tables 2 and 3.

Table 2. Values and Weights of Weapons Attribute

S/N	Attribute	Value	Weight
1	Biological	3401	0.1149
2	Chemical	3328	0.1124
3	Radiological	3371	0.1139
4	Nuclear	3327	0.1124
5	Firearms	3165	0.1069
6	Explosives/Bombs/Dynamite	3115	0.1052
7	Fake Weapons	1507	0.0509
8	Incendiary	2106	0.0712
9	Melee	1680	0.0568
10	Vehicle	2528	0.0854
11	Sabotage Equipment	2071	0.0700
	Total	29599	

Table 3. Values and Weights of Victims Attribute

S/N	Attribute	Value	Weight
1	Business	3650	0.0734
2	Government	3285	0.0661
3	Police	3094	0.0622
4	Military	3060	0.0615
5	Aircraft	3668	0.0738
6	Diplomatic	2985	0.0600
7	Educational Institution	3615	0.0727
8	Food or Water Supply	2442	0.0491
9	Journalists and Media	2021	0.0407
10	Maritime	2371	0.0477
11	NGO	1953	0.0393
12	Private Citizens and Property	3897	0.0784
13	Religious Figures/Institutions	3829	0.0770
14	Telecommunication	2080	0.0418
15	Tourists	1796	0.0361
16	Transportation (non Aviation)	2697	0.0542
17	Utilities	1870	0.0376
18	Violent Political Party	1404	0.0282
	Total	49717	

3.3 Neural network implementation

Alyuda NeuroIntelligence[22] was used for the Neural Network (NN) training. Neural Network implementation starts with loading of the training data set by opening it with the File menu. The NN consists of one input layer, one hidden layer and one output layer. The nine input layer nodes consists of Tactics Type (TCT), Weapons Type (WPT), Victims Type (VTT), number of Victims Killed (VTK), number of Victims Wounded (VTW), number of Terrorists Killed (TRK), number of Terrorists Wounded (TRW), number of Terrorists Involved (TRI) and number of Terrorists Captured (TRC). The single output layer node of the network is the fatality. The logistic function was used for input and output activation function while the sum of squares is used for output error function. Correlation was the fitness criteria of the NN while the heuristic search is selected as the search method. The connection weights were randomized between +3 and -3. The network was trained using the quick propagation algorithm. To determine the best NN architecture, twenty NN training sessions with 500 iterations were carried out. It was discovered that the best architecture was the one with 9 input layer nodes, 9 hidden layer nodes, 1 output layer node, training error of 0.07880, validation error of 0.08048, testing error of 0.08135 and the highest correlation value of 0.63636 as presented in Table 4.

NN assists in input importance determination which was carried out using some error metrics and number of iterations as the stopping conditions. Consequently, Absolute Error (AE) was set at 0.028783 while the Mean Square Error (MSE) was set at 0.01. Also 500, 1000 and 1500 iterations were performed. The average values of the various input importance based on the selected stopping conditions were taken as presented in Table 5. Three of the inputs (WPT, VTK and TRK) with average values of 37.26%, 35.93% and 14.46% respectively contributed approximately 88% of the variability in the data set and therefore constituted the most important attributes in the data set.

Table 4. Best Neural Network Architecture Search

ID	Architecture	Training Error	Validation Error	Testing Error	Correlation
1	[9-1-1]	0.08335	0.08522	0.08622	0.63481
2	[9-2-1]	0.08270	0.08472	0.08592	-0.10565
3	[9-3-1]	0.07588	0.07811	0.07901	0.08893
4	[9-4-1]	0.08402	0.08572	0.08648	0.19715
5	[9-5-1]	0.07939	0.08105	0.08201	0.32748
6	[9-6-1]	0.08187	0.08365	0.08466	-0.04648
7	[9-7-1]	0.07898	0.08102	0.08197	-0.15817
8	[9-8-1]	0.08391	0.08558	0.08669	-0.12358
9	[9-9-1]	0.07880	0.08048	0.08135	0.63636
10	[9-10-1]	0.07825	0.08013	0.08114	0.10752
11	[9-11-1]	0.08068	0.08197	0.08290	0.44344
12	[9-12-1]	0.07948	0.08022	0.08154	0.17098
13	[9-13-1]	0.07365	0.07516	0.07608	0.08030
14	[9-14-1]	0.08064	0.08196	0.08285	0.31517
15	[9-15-1]	0.07795	0.07892	0.08015	0.00362
16	[9-16-1]	0.07591	0.07752	0.07860	0.16898
17	[9-17-1]	0.07671	0.07778	0.07900	-0.08633
18	[9-18-1]	0.07478	0.07586	0.07677	0.00758
19	[9-19-1]	0.07315	0.07432	0.07510	0.26016
20	[9-20-1]	0.07609	0.07742	0.07826	0.53761

Table 5. Input Importance Values

Input Name	Input Importance (%)					
	AE (0.028783)	MSE (0.01)	500 Iterations	1000 Iteration	1500 Iteration	Average Input Importance (%)
TCT	1.293035	0.587048	1.266304	0.555242	0.645411	0.8694
TRC	2.004044	1.250398	1.908148	2.227554	0.003719	1.4788
TRI	2.73408	2.190405	1.521139	2.465604	0.084454	1.7991
TRK	25.693464	14.096389	11.347536	12.54582	8.597631	14.4562
TRW	0.202302	3.12909	1.999376	1.386693	1.201134	1.5837
VTK	56.729719	22.141431	51.059602	13.2887	36.41153	35.9262
VTT	5.331001	1.872507	1.698336	7.529512	4.064483	4.0992
VTW	2.179441	3.880152	3.854691	2.652668	0.074084	2.5282
WPT	3.832914	50.852579	25.34487	57.34821	48.91756	37.2592

Fuzzy clustering was incorporated into the research to generate the initial membership values between 0 and 1 for each of the data points in the data set which serve as input to ANFIS. The command, “findcluster”, entered at the MATLAB command prompt loads the Fuzzy C-Means (FCM) clustering interface but to retrieve the membership functions for each of the data points, a MATLAB code was written and the result presented in Table 6.

Table 6. Sample FCM-Generated Membership Functions for Each Data Point

S/N	Membership Values for Training Data Set				Membership Values for Checking Data Set				Membership Values for Testing Data Set			
	Cluster1	Cluster2	Cluster3	Cluster4	Cluster1	Cluster2	Cluster3	Cluster4	Cluster1	Cluster2	Cluster3	Cluster4
1	0.99783	0.00016	0.00115	0.00086	0.98873	0.00027	0.00169	0.00931	0.76514	0.00150	0.22150	0.01186
2	0.99783	0.00016	0.00115	0.00086	0.99492	0.00011	0.00068	0.00429	0.95857	0.00044	0.03774	0.00326
3	0.99783	0.00016	0.00115	0.00086	0.95984	0.00065	0.00425	0.03527	0.27803	0.00110	0.71197	0.00891
4	0.99850	0.00011	0.00081	0.00058	0.34576	0.00508	0.03060	0.61857	0.98672	0.00009	0.01248	0.00071
5	0.99850	0.00011	0.00081	0.00058	0.98002	0.00035	0.00230	0.01733	0.05135	0.00037	0.94551	0.00277
6	0.99850	0.00011	0.00081	0.00058	0.98949	0.00021	0.00138	0.00893	0.98594	0.00010	0.01320	0.00077
7	0.00103	0.99553	0.00256	0.00088	0.96633	0.00057	0.00355	0.02955	0.95857	0.00044	0.03774	0.00326
8	0.91019	0.00592	0.04315	0.04075	0.99030	0.00020	0.00128	0.00823	0.36930	0.00100	0.62202	0.00768
9	0.94146	0.00404	0.02989	0.02461	0.98789	0.00029	0.00193	0.00990	0.95752	0.00044	0.03879	0.00326
10	0.95150	0.00342	0.02543	0.01966	0.97406	0.00068	0.00428	0.02098	0.95364	0.00050	0.04209	0.00376
11	0.98826	0.00090	0.00682	0.00402	0.97406	0.00068	0.00428	0.02098	0.02884	0.00702	0.02669	0.93746
12	0.62326	0.02394	0.14527	0.20754	0.99125	0.00020	0.00128	0.00728	0.66446	0.00107	0.32631	0.00816
13	0.00071	0.99692	0.00179	0.00059	0.97406	0.00068	0.00428	0.02098	0.94808	0.00058	0.04677	0.00457
14	0.22652	0.14563	0.21069	0.41717	0.97406	0.00068	0.00428	0.02098	0.95074	0.00028	0.04687	0.00211
15	0.97921	0.00144	0.01082	0.00854	0.99492	0.00011	0.00068	0.00429	0.04851	0.03718	0.04795	0.86637
16	0.98826	0.00090	0.00682	0.00402	0.97406	0.00068	0.00428	0.02098	0.01764	0.92222	0.01806	0.04208
17	0.97138	0.00195	0.01460	0.01208	0.95897	0.00067	0.00418	0.03618	0.00095	0.99580	0.00095	0.00230
18	0.99786	0.00016	0.00120	0.00079	0.97152	0.00053	0.00333	0.02461	0.33338	0.03023	0.48163	0.15476
19	0.99346	0.00047	0.00339	0.00268	0.03893	0.00948	0.92128	0.03031	0.95364	0.00050	0.04209	0.00376
20	0.98826	0.00090	0.00682	0.00402	0.00031	0.99854	0.00085	0.00030	0.05200	0.04067	0.05375	0.85358

Based on the input importance analysis carried out, weapon type, number of victims killed and number of terrorists killed was selected as the input variables for processing. The inputs are represented by the first three columns in each of the training, checking and testing data sets. The fourth column represents the output (fatality). The fuzzy data thus generated for each data point constitute the initial training data for ANFIS. Four cluster mean (cluster centers) for the training data are -0.96418, -0.73979, 0.309656 and 0.068801 respectively. The cluster centers for checking data are -0.96335, -0.89125, 0.293613 and 0.036498 respectively. The cluster centers for testing data are -0.98123, -0.8833, 0.314286 and 0.033181 respectively.

3.4 ANFIS implementation and results

The data set was pre-processed using the formulated Rank-Frequency-Based (RFB) model in order to convert the qualitative data attributes such as weapons used by terrorists, number of victims killed in an attack and number of terrorists killed in an attack to numeric attributes which were used by ANFIS. The 1000 records selected for training was segmented into 70% (700 records) for training, 15% (150 records) for checking and 15% (150 records) for testing data sets. In the research, three experiments were conducted. In the first experiment, the qualitative data attributes were processed using an existing OoN method adopted in Alyuda NeuroIntelligence and the data then supplied to ANFIS (OoN-ANFIS). The RMSE of OoN-ANFIS training and checking data are shown in Figure 4. The checking error, which were used to prevent over-fitting of the model, reduced from epoch 1 to epoch 6 and thereafter started increasing till epoch 33 where it reduced again erratically. This indicated that the best training parameters were optimally adjusted and selected at the sixth epoch. The average training and checking errors are 0.0000216 and 0.0007099 respectively.

In the second experiment, the qualitative data attributes were processed by the RFB model before the terrorist acts model was trained by ANFIS (RFB-ANFIS). The RMSE values for training and checking data are shown in Figure 5 with average errors of 0.006735 and 0.009383 respectively. Here the checking error decreases down to epoch 20, after which it started increasing sharply and behaving erratically, indicating that at the twentieth epoch, the optimum ANFIS model parameters were adjusted and selected.

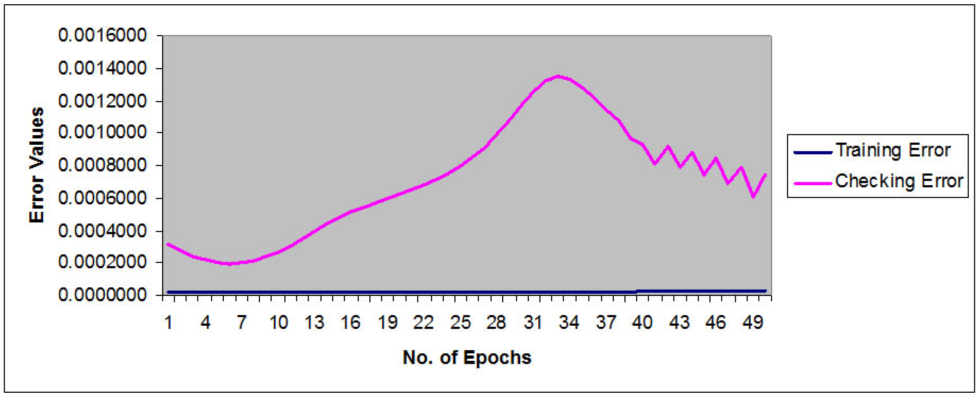

Figure 4. A Plot of Training and Checking Errors of OoN-ANFIS

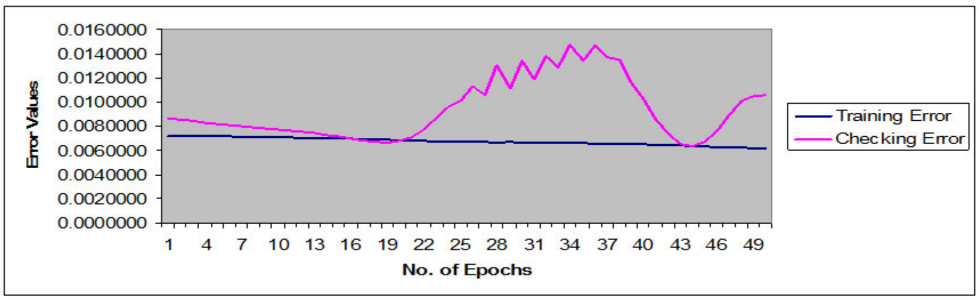

Figure 5. Plot of Training and Checking Errors

In the third experiment, the RFB-processed qualitative data attributes were used by FCM to generate initial membership values for each point in the data set and thereafter supplied to ANFIS (RFB-FCMANFIS) for training. The RMSE of training and checking data of the RFB-FCMANFIS training activity is shown in Figure 6. After the first epoch, the checking error started increasing sharply and then fluctuating at subsequent epochs therefore indicating a quicker adjustment and selection of optimum ANFIS parameters. The average training and checking errors are 0.0000713 and 0.0056155 respectively. The plot of error values involving OoN-ANFIS, RFB-ANFIS and RFB-FCMANFIS in this research is shown in Figure 7.

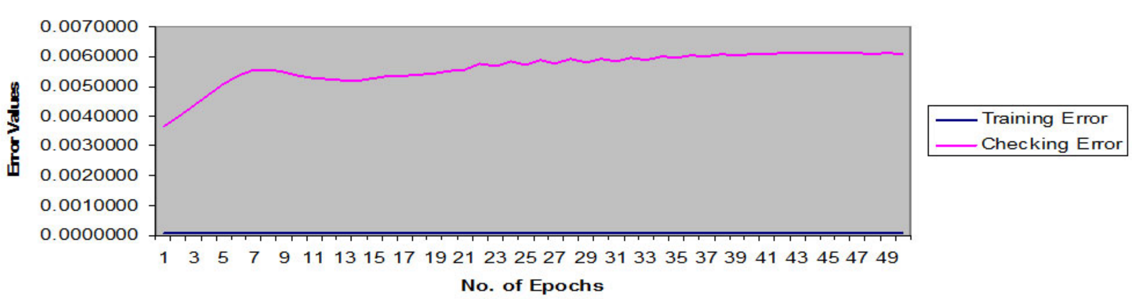

Figure 6. Training and Checking Error Plot

The metrics used to check the performance of the classification models include the Root Mean Squared Error (RMSE), the Mean Absolute Error (MAE), training error, checking error and correlation coefficient presented in Table 7. The RMSE values for OoN-ANFIS, RFB-ANFIS and RFB-FCMANFIS are 0.008594, 0.00517 and 0.002887 respectively while that of the research in7 is 0.08. The MAE values for OoN-ANFIS, RFB-ANFIS and RFB-FCMANFIS are 0.008662, 0.005267 and 0.004598 respectively.

Table 7. Research Summary

S/N	Parameter	Classification Model			
		1st Experiment (OoN-ANFIS)	2nd Experiment (RFB-ANFIS)	3rd Experiment (RFB-FCMANFIS)	Inyaem *et al.*(2010b)
1	Point of Best Model Generation	Epoch 6	Epoch 20	Epoch 1	
2	Average Training Error	0.0000216	0.006735	0.0000713	
3	Average Checking Error	0.00071	0.009383	0.0056155	
4	Average Execution Time for ANFIS Model Training (seconds)	67.9539	67.5002	68.0109	
5	Correlation	0.98004	0.97464	0.99954	
6	RMSE	0.008594	0.000517	0.002887	0.08
7	MAE	0.008662	0.005267	0.004598	

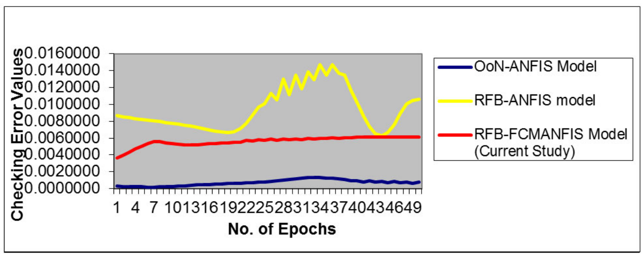

Figure 7. Plot of Different Error Values

The RFB-FCMANFIS model generated the least RMSE and MAE of 0.002887 and 0.004598 respectively. It also has a highest correlation of 0.99954. Therefore an improved classification model for terrorist acts is generated using RFB-FCMANFIS of this research. There is no significant variation in the average execution time of the three experiments as they all use approximately 68 seconds to execute showing the current study does not cause an increase in the time of execution compared to other models. There is a great improvement in the RMSE in the three experiments compared to what was obtained in the related work in.[7] The RFB-FCMANFIS in the third experiment converged faster than others and generated the classification model at the first epoch. Also there is a significant improvement on the classification accuracy for terrorist acts when the qualitative data is subjected to the RFB model. The classification performance of ANFIS is enhanced when the qualitative data are subjected to the RFB model and FCM is used to generate membership values to initialize the input for ANFIS training.

4. CONCLUSION

The acts of terrorism have elevated the level of violence, intimidation and threat to peace and security at large in the world today. Several approaches adopted to tackle the menace seem to be of insignificant impact against its occurrence. This research has implemented an intelligent clustering model used for the classification of the acts of terrorism to assist in prevention of occurrence or mitigation of effects of terrorism if it occurs. The formulated model used for the conversion from qualitative to quantitative data attributes proves to be a veritable tool in training terrorist classification model. This work serves as a dependable analytic and decision support tool in the terrorism domain. For further research, techniques such as simulation and the use of surrogate data to increase the sample size in order to improve the result should be studied.

REFERENCES

[1] George UD, Inyang UG, Akinyokun OC. Conceptualization of Intelligent Clustering Methodology for Terrorist Acts Classification in Nigeria. International Journal of Current Advanced Research. 2018; 7(7): 14200-14208. `http://dx.doi.org/10.24327/ijcar.2018.14208.2567`

[2] Perry WJ. Technology and National Security: Risks and Responsibilities. Proceedings of the Conference on Risk and Responsibility in Contemporary Engineering and Science: French and U.S. Perspective. 2003: 1-12.

[3] Chakraborty T. Educative Value of Information Technology in Tackling Terrorism with Community Policing. Bureau of Police Research and Development, New Delhi. 2009.

[4] Saeed F, Rahid M, Rehman HZ, et al. Tackling Terrorism in Pakistan. International Journal of Peace and Development Studies. 2012; 3(1): 1-5.

[5] Akinyokun OC. Neuro-Fuzzy Expert System for Evaluation of Human Resource Performance. First Bank of Nigeria Endowment Fund Lecture Series 1. Delivered at the Federal University of Technology, Akure. 2002.

[6] Inyaem U, Haruechaiyasak C, Meesad P, et al. Construction of Fuzzy Ontology-Based Terrorism Event Extraction. Proceedings of IEEE Computer Society's Third International Conference on Knowledge Discovery and Data Mining. 2010a: 391-394.

[7] Inyaem U, Horuechaiyasak C, Meesad P, et al. Terrorism Event Classification using Fuzzy Inference Systems. International Journal of Computer Science and Information Security. 2010b; 7(3): 247-256.

[8] Thipparat T. Application of Adaptive Neuro Fuzzy Inference System in Supply Chain Management Evaluation. Faculty of Management Sciences, Prince of Songkla Universit, Kohong Hatyai, Songkhla, Thailand. 2012.

[9] Inyang UG, Akinyokun OC. A hybrid Knowledge Discovery System for Oil Spillage Risks Pattern Classification. Artificial Intelligence Research. 2014; 3(4): 77-86. https://doi.org/10.5430/air.v3n4p77

[10] Adnan M, Rafi M. Extracting Patterns from Global Terrorist Dataset (GTD) using Co-Clustering Approach. Journal of Independent Studies and Research – Computing. 2015; 13(1): 7-13. https://doi.org/10.31645/jisrc/(2015).13.1.0002

[11] Dwivedi S, Pandey P, Tiwari MS, et al. Comparative Study of Clustering Algorithms used in Counter Terrorism. Journal of Computer Engineering (IOSR-JCE). 2014; 16(6): 13-17. https://doi.org/10.9790/0661-16681317

[12] Stoffel K, Cotofrei P, Han D. Fuzzy Methods for Forensic Data Analysis. Proceedings of the International Conference on Integration Techniques in Computational Forensics. 2010: 1-6.

[13] Marappan K, Nallaperumal K, Kannan KS, et al. A soft Computing Model to Counter Terrorism. International Journal of Computer Science and Network Security. 2008; 8(5): 141-147.

[14] Akinyokun OC, Iwasokun GB, Arekete SA, et al. Fuzzy Logic-Driven Expert System for Diagnosis of Heart Failure Disease. International Journal of Artificial Intelligence Research. 2015; 4(1): 12-21. https://doi.org/10.5430/air.v4n1p12

[15] Motaz M, Khorshid H, Tarek H, et al. Hybrid Classification Algorithms for Terrorism Prediction in Middle East and North Africa. International Journal of Emerging Trendsand Technology in Computer Science (IJETTCS). 2015; 4(3): 23-29.

[16] Ghada MT, Omar SS. An Experimental Study of Classification Algorithms for Terrorism Prediction. International Journal of Knowledge Engineering. 2015; 1(2): 107-112. https://doi.org/10.7763/IJKE.2015.V1.18

[17] Othman MF, Yau TMS. Neuro-Fuzzy Classification and Detection Technique for Bioinformatics Problems. Proceedings of the First Asian International Conference on Modelling and Simulation. IEEE Computer Society. 2007: 1-6.

[18] Al-Himyari BA, Yasin A, Gitano H. Adaptive Neuro-Fuzzy Model with Fuzzy Clustering for Nonlinear Prediction and Control. Asian Journal of Applied Sciences. 2014; 2(3): 343-348.

[19] Javiad Q, Arif M, Talpur S. Segmentation and Classification of Classification and Hemorrhage in the Brain using Fuzzy C-Mean and Adaptive Neuro-Fuzzy Inference System. QUAID-E-AWAM University Research Journal of Engineering, Science & Technology. 2016; 15(1): 50-63.

[20] Ziasabounchi N, Askerzade I. ANFIS-Based Classification Model for Heart Disease Prediction. International Journal of Electrical and Computer Sciences. 2014; 14(2): 7-12.

[21] Abbas T, Mahdi H, Masoumeh A, et al. A Subjective Weighting Method Based on Group Decision Making for Ranking and Measuring Criteria Values. Australian Journal of Basic and Applied Sciences. 2011; 5(12): 2034-2040.

[22] Alyuda NeuroIntelligence. Available from: https://www.alyuda.com

[23] Farahman F, Asgar MH. Rethinking Convenience Sampling: Defining Quality Criteria. Theory and practice in language studies. 2012; 2(4): 784-792. https://doi.org/10.4304/tpls.2.4.784-792

[24] Mohsin A. A Manual for Selecting Sampling Techniques in Research. MPRA Paper No. 70218. 2016. Available from: https://mpra.ub.uni-muenchen.de/70218/

[25] Salazar A, Safont G, Vergara L. Surrogate Techniques for Testing Fraud Detection Algorithms in Credit Card Operations, Proceedings of 2014 International Carnahan Conference on Security Technology (ICCST). 2014. p. 124-129.