International Research Journal of Engineering and Technology (IRJET) e-ISSN: 2395-0056
Volume: 06 Issue: 09 | Sep 2019 www.irjet.net p-ISSN: 2395-0072

An Evaluation of Hybrid Machine Learning Classifier Models for Identification of Terrorist Groups in the aftermath of an Attack

Peter Opiyo Oketch[1], Muhambe Titus Mukisa[2], Makiya Cyprian Ratemo[3]

[1]Msc. Student, Dept. of IT, Maseno University, Kenya
[2]Lecturer, Dept. of IT, Maseno University, Kenya
[3]Director, e-Learning, Kisii University, Kenya

Abstract - *The urgency of responding to a terrorist attack and the subsequent nature of analysis required to identify the terrorist group involved in an attack demands that the performance of the machine learning classifiers yield highly accurate outcomes. In order to improve the performance of machine learning classifiers, hybrid machine learning algorithms are used with the goal of improving the accuracy. The aim of the study was to build and evaluate hybrid classifier models for identification of terrorist groups. The research specifically sought to: build base classifiers (Naïve Bayes, K-Nearest Neighbor, Decision Trees, Support Vector Machines and Multi-Layer Perceptron); build hybrid classifier models from a combination of the base classifiers; and compare the performance of the hybrid and base classifiers. The study adopted an experimental research method using WEKA tool for data mining and real-world terrorist datasets for the period 1999-2017 for Sub-Saharan Africa region from the Global Terrorism Database. WEKA supervised filter Ranker was used to select 6 attributes of data and 784 records. The classifiers were evaluated using 10-fold cross validation. The study established that the optimal performance for all the classifiers was realized with a more balanced class at a resample rate of 1000%. The study concludes that hybrid classifiers perform better than base classifiers, and the best performing model was a hybrid combination of KNN, and DT. The study provides insights on the performance of hybrid machine learning classifiers and lays a foundation for further research in hybrid machine learning approaches.*

Key Words: Data mining, classification, Ensemble, Hybrid, Resample filter, class imbalance, WEKA, Naïve Bayes, Decision tree(J48), Majority voting, Support Vector Machine(smo), K-Nearest Neighbor (IBK), Multilayer perceptron

1. INTRODUCTION

According to the Global Terrorism Index (2016) [12], there has been a continued rise of terrorism which is a serious concern. Terrorist acts are planned and perpetrated by collections of loosely organized people operating in shadowy networks that are difficult to define and identify. Terrorism is considered a low-intensity form of warfare; however, terrorist plots and activities will leave an information signature, albeit not one that is easily detected (Popp, Armour , Senator, & Numrych, 2004) [25]. Data mining techniques like association, classification, clustering and prediction are widely used in counter terrorism. Data mining helps in discovering links, patterns, flaws in huge set of data which are very hard to find using human assistance (Ahsan & Shah , 2008) [1]. Classification is a supervised learning technique, it is used to predict the categorical class label of a given data instance, so as to classify it into one of the predetermined classes. It is a two-step process, in first step classification algorithm uses training dataset to build a classifier, and then in second step this classifier is used to predict the class label of a given unlabeled data instance (Verma, 2019)[32] The resulting classifier is then used to assign class labels to the testing instances where the values of the predictor features are known, and the value class label is unknown. The input data for the classification is a set of instances. Each instance is a record of data in the form of (x, y) where x is the features set and is y the target variable (class label). Classification model is a tool that is used to describe data (Descriptive Model) or a tool to predict the target variable for a new instance (Predictive Model) (Kotsiantis, 2007)[19]. The key question when dealing with classification is not whether a learning algorithm is superior to others, but under which conditions a particular method can significantly outperform others on a given application problem(Kotsiantis,Zaharakis,&Pinteals,2006)[20].

Traditional learning algorithms experience great challenge because of the high dimensionality of microarray data may bring some disadvantages, such as over-fitting, poor performance and low efficiency (Song, et al., 2015)[28]. The challenge of responding to a terrorist attack and the subsequent nature of analysis required to identify the group involved in an attack demands that the performance of classifiers yield highly accurate results. The technique of combining classifiers into an ensemble is a new approach for the improvement of the performance of traditional machine learning algorithms (Kuncheva & Whitaker, 2003)[22]. Hybrid classifier systems are ensemble classifiers that combine, and integrate different standard machine learning algorithms, resulting in improved performance, and more adaptivity (Kainulainen, 2010) [16]. Dietterch (2000) [7] has suggested numerous methods for the creation of ensemble of classifiers, although many methods of ensemble creation have been proposed, there is yet no clear picture of which method is best. Thus, active area of research in supervised learning is the study methods for the construction of good ensembles of classifiers (Villada & Drissi, 2002)[33]. Mechanisms that are used to build ensembles of classifiers include: using different subsets of training data with a single

learning method, using different and using different learning methods, the application of ensemble methods is suggested only if we are interested in the best possible classification accuracy(Kotsiantis, Zaharakis, & Pinteals, 2006)[20]. The aim of the study was to evaluate the performance of hybrid classifier models and compare it to the performance of traditional machine learning classifiers with an aim of establishing the best hybrid model for identification of terrorist groups. The study used terrorism dataset from the Global Terrorism Database (GTD).

1.1 Related Work

Various machine learning methods like K-Nearest Neighbour, Naïve Bayes, Decision Trees, Support Vector Machines, Multilayer Perceptron and Hybrid approach VOTE are built, evaluated and tested using terrorism data. The previous related studies primarily involved research on various application of machine learning in counterterrorism. Ozgul, Erdem, & Bowerman (2009) [23] developed crime prediction model (CPM) with group detection algorithms and CPM performed well on attributes of crime information to predict terrorist activities. In a study by Ivan (2009)[15], time-series methods were used to investigate the relationship between the number of global strategic armed forces-related incidents and the frequency of transnational terrorist attacks, the type of attacks, and the type of victims of terrorist attacks with data from transnational terrorism incidents from 1993 to 2004. Sachan and Roy (2012) [27] in their study realized more than 80% accuracy from a Terrorist Group Prediction Model developed to predict the terrorist group involved in a given attack in India from the year 1998 to 2008. In a study by Rizwan, Masrah, Aida, Payam, & Nasim (2013) [26] Two different classification algorithms namely, Naïve Bayes and Decision Tree for predicting "Crime Category" for different states in USA were compared. 10-fold cross validation was applied to the input dataset in the experiment, separately for both NB and DT to test the accuracy of the classifiers which showed that DT algorithm outperformed NB algorithm and achieved 83.951% accuracy in predicting "crime Category". Farysal, Wasi, & Usman (2014) [9] proposed a novel ensemble framework for the classification and prediction of terrorist groups in Pakistan that consisted of four base classifiers namely; NB, KNN, ID3, and Decision Stump (DS). Majority vote-based ensemble technique was used to combine these classifiers. The results of individual base classifiers were compared with the majority vote classifier and it was determined through experiments that the new approach achieved a considerably better level of accuracy and less classification error rate as compared to the individual classifiers. Khorsid, Abou, & Soliman (2015) [17] conducted an experimental study on 43335 terrorist events by applying hybrid supervised machine learning classifiers which proved SVM and RF gave better accuracy during classification. SVM classifier outperformed ANN and NB in classification accuracy for both attributes weapon type (89%) and attack type (81%) with leading nature of NB. In a study by Tolan & Soliman (2015) [31] SVM was more accurate than other classifiers especially NB, and KNN, the overall performance of NB and KNN was almost the same (Tolan & Soliman, 2015)[31]. According to Swanson (2016) [29], in the year 1970-1998, Hawkes Process was used to predict terrorist attacks in Northern Ireland which considered 5000 explosions. Random forest classifier (RF) yielded 79% accuracy for attack types and for weapon type the accuracy of classification is 86% as compared to other classifiers. Coffman and Marcus (2004) [6] in a social network analysis and pattern classification used to predict whether a person is terrorist or not resulted in 86% accuracy. A terrorist attack prediction model used a neural network to successfully predict the conflict in Liberia in 2010; the accuracy was between 0.65 and 0.74 (Blair, Blattman, & Hartman, 2017)[3]. Dong (2017) [8] used the 2010–2016 terrorist attacks data in India to empirically examine the effectiveness of machine learning based on back propagation (BP) neural networks in real-life terrorist attacks. The study found out that machine learning-based terrorist attack prediction paradigms, have a certain ability to anticipate terrorist attacks and can discover new knowledge regarding conflicts. Gundabathula & Vaidhehi (2018) [13] conducted a study on efficient modeling of behavior of terrorist groups in India, DT, NB, KNN, and ensemble approach were used. DT and ensemble classifiers gave highest accuracy, while there has been previous work on evaluation of hybrid classifiers in prediction of terrorist groups, none has considered to find out the best combination of classifiers among KNN, DT, MLP, SVM and NB in improving identification of terrorist groups.

Finally, complete content and organizational editing before formatting. Please take note of the following items when proofreading spelling and grammar:

1.2 Performance Evaluation

In order to determine the effectiveness of the classification algorithm used, a measurement is needed. Commonly used measurements include classification accuracy and error, F-Measure, precision, recall, Receiver Operating Characteristic (ROC) curves and Area under the Curve (AUC) (Freitas, deCarvalho, & Oliveira, 2007)[10]. These measurements can be calculated by the classification results commonly tabulated in a matrix format called a Confusion Matrix. Better differentiation of algorithms can be obtained by examining computational performance metrics such as build time and classification speed (Williams, Zander, & Armitage, 2006)[35]. Although training time varies according to the nature of the application task and dataset, specialists generally agree on a partial ordering of the major classes of learning algorithms.

2. MATERIALS AND METHODS

The study adopted an experimental research design. The term "experimental research design" is centrally concerned with constructing research that is high in causal (or internal) validity. Causal validity concerns the accuracy of statements regarding cause and effect relationships (Alexandrie, 2017)[3]. The data mining framework methodology followed in this study is the Cross-Industry Standard Process for Data Mining (CRISP-DM), a non-proprietary hierarchical process model designed by practitioners from different domains. It has proven to be the most commonly preferred framework

(Piatestsky, 2014)[24]. The framework breaks down the data mining process into six phases: understanding the business process and determining the ultimate data mining goals; identifying, collecting, and understanding key data sources; preparing data for data mining; selecting modelling techniques to use; evaluating and comparing results of different models against the initial goals and deploying the model. WEKA is used as a data mining tool for pre-processing, building classification models and evaluation because of its support for experimental data mining tasks and virtually all algorithms (Witten, Frank, & Hall, 2011)[36].

2.1 Data Collection

The Data for this study was collected from Global Terrorism Database (GTD). The data base is obtained from the National Consortium for the Study of Terrorism and Responses to Terrorism (START) initiative at University of Maryland, from their online interface at http://www.start.umd.edu/gtd/ as an open source database. GTD includes information about terrorist activities that happened throughout the world from 1970 to 2017.There are more than 170,000 cases recorded in GTD. It includes information on more than 14,400 assassinations, 52,000 bombings, and 5,600 kidnappings activities since 1970. Complete information of an incident is given in GTD, for example date of the event, location of the event, weapon used, number of causalities, attack type, group that is responsible for the event etc. There were a total 134 attributes and 113114 records in GTD. It is the most comprehensive open source database on terrorism activities and research throughout the world (START, 2018). The dataset for the study was 1999-2017 dataset for sub-Saharan Africa region from the GTD. According to the Global Terrorism Index (2016) [12] there were heightened terrorist activities in the sub-Saharan region between 1999 to 2016.

2.2 Data Pre-processing

Before running any classification algorithms on the data, the data must first be cleaned and transformed in what is called a pre-processing stage. During this pre-processing stage, several processes take place, including attribute selection, evaluating missing values, eliminating noisy data such as outliers, normalizing, and balancing unbalanced data. Data pre-processing improves data quality (Han & Kamber, 2006)[14]. WEKA supervised filter method ranked six (6) attributes as the most relevant in predicting of the terrorist groups class. Data obtained after reduction was both in numerical as well as textual nominal, only nominal numbers were considered. Attributes selected for the identification of terrorist groups are listed below:

i. Country (This attribute represents the country or location where the incident has happened)
ii. Region (It is a categorical variable and it represents the region in which the incident occurred)
iii. Attack type (It is a categorical variable and shows which kind of attack types executed e.g. assassination, bombing, kidnapping etc. there are total of 9 kinds of attack types recorded)
iv. Target type (This attribute represents the target category)
v. Gname (This attribute represents the group that is responsible for attack)
vi. Weapon type (This attribute shows the type of weapon used in attack)

The pre-processed dataset was converted to. ARFF to be used by WEKA. The attribute Gname was the class label which was intended to be predicted based on the values of other attributes. All the instances having class value as unknown were removed. Class label that had less than ten frequency of occurrence was also eliminated. This meant that the instances in which a terrorist group had participated in an incident less than ten times for the whole time period of 1999-2017 were removed. The column Gname which indicates the perpetrator group responsible for an incident consisted of apostrophes which were not accepted by WEKA for constructing the classification models. Therefore, these apostrophes from the group names were replaced by spaces. In this study the WEKA remove duplicates filter was used to remove duplicate instances. This resulted in the final dataset of 6 attributes and 784 records. The major pre-processing to be done was to solve the class imbalance problem in the data. Class imbalance problem is a situation in which the observations that belong to one class are significantly lower when compared to the observations of other classes. This causes biasness towards the classes with higher observations. In the final dataset WEKA resampling filter was applied to solve the class imbalance problem. This reduces class skew before applying a base classifier. It also ensures that the sub-sample is stratified so that the original class distribution is preserved in the sub-sample. WEKA Resampling filter produces a random sub-sample of dataset using either sampling with replacement or without replacement for imbalanced dataset. To achieve oversampling of the minority class, rather than under sampling of the majority class, so that both classes have the same numbers of instances (Gundabathula & Vaidhehi, 2018). Based on the data pre-processed there are 25 distinct terrorist groups, domestic and foreign that were involved in terror activities recorded between 1999 to 2017 for Sub-Saharan Africa region. Resampling is done at different rates and checked against accuracy rate, and build time.

2.3 Models Creation

Development of the hybrid classifier is two phased. The first phase is the construction and evaluation of base classifiers (selection of the modelling algorithms). The second phase is the combination of the selected base classifiers through bagging ensemble technique. Bagging combines them by majority voting and the most voted class is predicted (Breiman, 2001)[5]. Fig.1 shows the methodology followed for the identification system in this study. The experiments were conducted on Lenovo Ideapad 320 machine running an Intel processor core i5 third generation, 4GB ram, 1Terabyte

hard drive, MS Windows 10 Pro and WEKA 3.8.3. WEKA was used for modeling learning algorithms to build base classifiers, apply Majority voting combiner for integrating base classifiers into a hybrid and to analyze the performance of the classifiers using 10-fold cross validation.

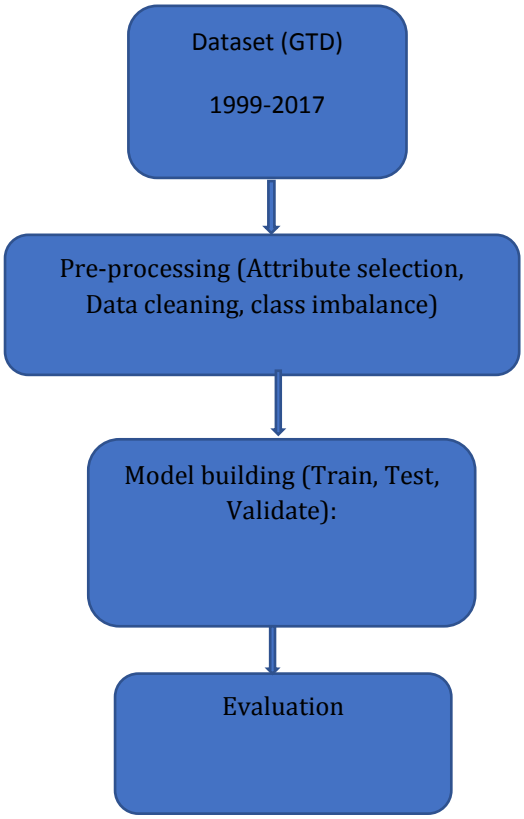

Fig -1: Identification flow

The classification models were constructed for the dataset using famous machine learning classification algorithms: Decision tree(J48), Naive Bayes, K-Nearest Neighbor (IBK), Support vector machine (SMO) and MLP and evaluated using 10-fold cross validation. In 10 cross validation method data set is divided randomly into 10 parts. 9 of those parts are used for training and reserved one tenth for testing. The procedure is repeated 10 times each time reserving a different tenth for testing. The aim is to overcome the problem of overfitting and make prediction more general. The advantage of this method is that all observations are used for both training and validation, and each observation is used for validation exactly once (Krogh & Vedelshy, 1995)[21]. J48 algorithm belongs to the Decision trees family which is used to generate decision trees. Naive Bayes is a probabilistic classifier which belongs to the Bayes family. IBK is a lazy learning algorithm which is the implementation of K-nearest neighbor's algorithm, SMO and MLP are function algorithms for the implementation of SVM and MLP respectively. 5 experiments were set up, one for each different machine learning classifier model. The hybrid Classifier models were built by combining the base classifiers: KNN, Naive Bayes, Decision Tree, SVM and MLP classifiers for better performance and results. The set of classifiers were trained in parallel, and their outputs combined afterwards to give the final decision. The hybrid classifier was evaluated using 10-fold cross validation. Fig. 2 demonstrates the flow of the hybrid model development.

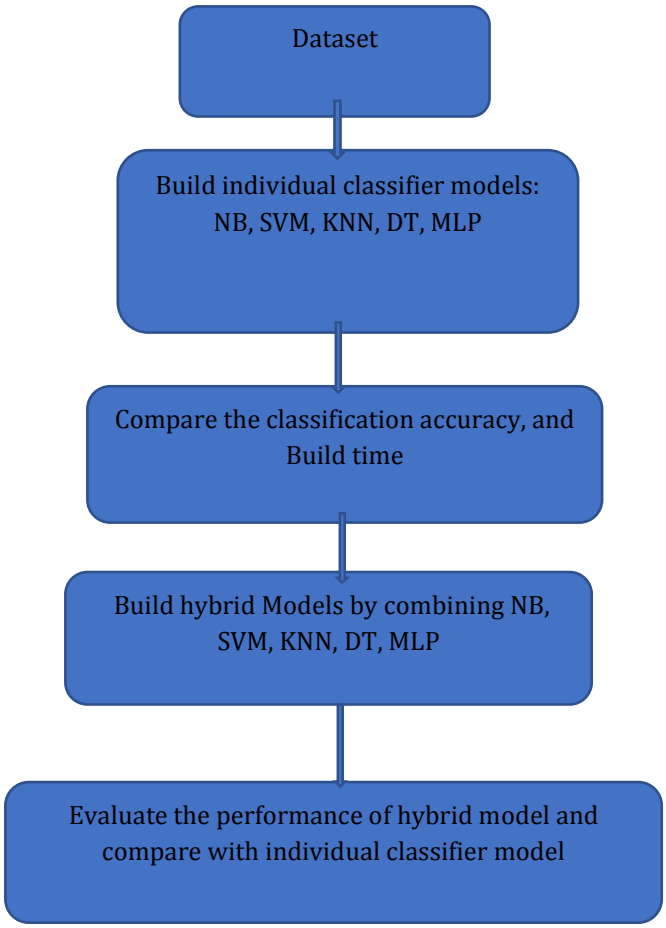

Fig -2: Classifier Model Creation Flow

The set of classifiers were trained in parallel, and their output combined afterwards using Vote Operator in WEKA, to give the final decision. The Vote Operator was applied on training dataset to create a model. Sub processes of vote operator KNN, Naive Bayes, Decision Tree, SVM and MLP classifiers were used. Each of these classifiers would get the dataset and generate a classification model, and then the vote operator applied all the classification models from its sub processes and assigned the predicted class with maximum votes to the unknown example set. For prediction of an unknown example set, the Vote operator applies all the classification models included in its sub process and assigns the predicted class with maximum votes to the unknown example (Bouziane, Messabih, & Chouarfia, 2011)[4]. The

Vote operator has sub processes, called base learners. This operator builds a classification model or regression model depending upon the data Set and learners. This operator uses a majority vote for predictions of the base learners provided in its sub process. While doing classification, all the operators in the sub process of the Vote operator accept the given training data set and generate a model for classification. For prediction of an unknown example set, the Vote operator applies all the classification models included in its sub process and assigns the predicted class with maximum votes to the unknown example. 4 experiments were set up for various optimum combinations of base classifiers based on the classifier model accuracy and build time. Accuracy of the classifier refers to the ability of a given classifier to correctly predict the class label for new or unseen data. The overall effectiveness of the algorithm is calculated by dividing the correct labeling against all classifications.

Accuracy=(TP+TN)/(TP+FP+TN+FN)

Build time is the time taken to model a classification algorithm

3. RESULT AND DISCUSSION

The main aim of the study was to build and evaluate hybrid machine learning classifier models for identification of terrorist groups in the aftermath of an attack. Several experiments were conducted, and the purpose was to establish if the models were optimally designed for classification and the performance of the hybrid classifier was better than individual base classifiers. Data mining and prediction tool WEKA used in this study and the model building has proved to be very efficient in prediction from the data available in GTD. The results of the study are organised in sub-topics focusing on specific objectives namely: Build base classifier models, Build hybrid classifier models, and comparison of performance between hybrid classifier models and base classifier models.

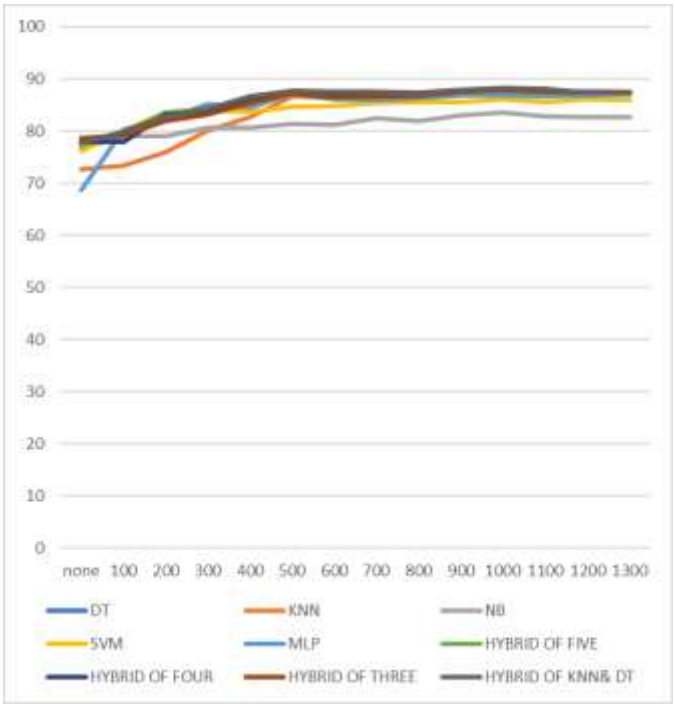

Chart -1: Classification accuracy

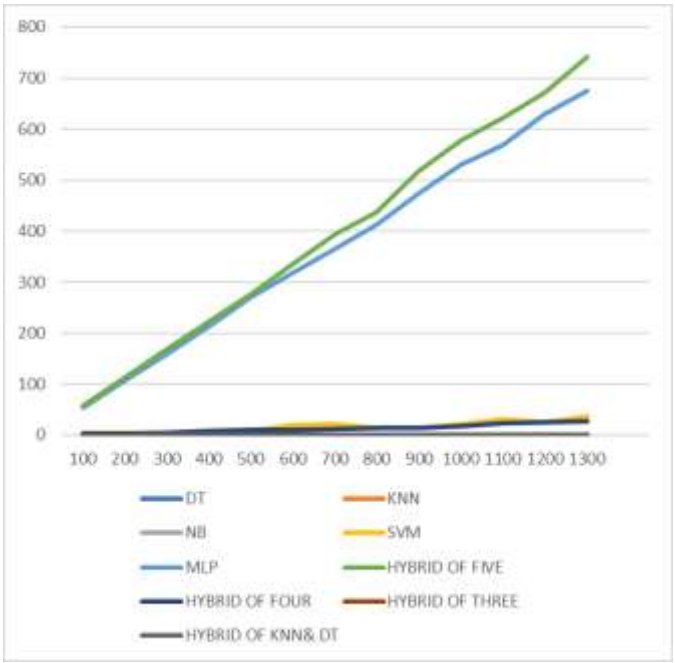

Chart -2: Build time

Table -1: Base classifiers and hybrid of DT, KNN, NB, SVM, MLP

Classifier models rate of accuracy (%)						
Resample (%)	DT	KNN	NB	SVM	MLP	Hybrid
None	77.42	72.7	76.66	76.15	68.62	77.17
100	79.08	73.34	79.08	80.1	80.23	79.6
200	82.46	75.89	78.95	83.61	82.27	83.48

300	83.72	80.1	80.57	83.93	85.25	84.23
400	85.59	82.62	80.68	83.65	84.53	85.78
500	86.89	86.66	81.3	84.55	87.09	87.4
600	86.31	86.44	81.19	84.89	86.12	86.82
700	86.42	86.37	82.45	85.42	85.9	86.6
800	86.4	86.24	81.9	85.48	86.34	87.04
900	87.34	87.56	82.96	85.63	86.73	87.54
1000	87.53	87.73	83.61	86.17	86.79	87.83
1100	87.45	87.63	82.75	85.6	86.55	87.44
1200	86.93	87.54	82.66	86.01	86.78	87.47
1300	87.13	87.52	82.61	85.96	87.06	87.25

Table -2: Base classifiers and hybrid of DT, KNN, NB, SVM

Classifier models rate of accuracy (%)						
Resample (%)	DT	KNN	NB	SVM	MLP	Hybrid2
None	77.42	72.7	76.66	76.15	68.62	77.93
100	79.08	73.34	79.08	80.1	80.23	77.93
200	82.46	75.89	78.95	83.61	82.27	83.16
300	83.72	80.1	80.57	83.93	85.25	83.8
400	85.59	82.62	80.68	83.65	84.53	85.71
500	86.89	86.66	81.3	84.55	87.09	87.27
600	86.31	86.44	81.19	84.89	86.12	86.93
700	86.42	86.37	82.45	85.42	85.9	86.59
800	86.4	86.24	81.9	85.48	86.34	86.86
900	87.34	87.56	82.96	85.63	86.73	87.6
1000	87.53	87.73	83.61	86.17	86.79	87.98
1100	87.45	87.63	82.75	85.6	86.55	87.62
1200	86.93	87.54	82.66	86.01	86.78	87.5
1300	87.13	87.52	82.61	85.96	87.06	87.33

Table -3: Base classifiers and hybrid of DT, KNN.NB

Classifier models rate of accuracy (%)						
Resample (%)	DT	KNN	NB	SVM	MLP	Hybrid3
None	77.42	72.7	76.66	76.15	68.62	78.57
100	79.08	73.34	79.08	80.1	80.23	79.21
200	82.46	75.89	78.95	83.61	82.27	81.95
300	83.72	80.1	80.57	83.93	85.25	83.21
400	85.59	82.62	80.68	83.65	84.53	85.33
500	86.89	86.66	81.3	84.55	87.09	87.17
600	86.31	86.44	81.19	84.89	86.12	86.5
700	86.42	86.37	82.45	85.42	85.9	86.66
800	86.4	86.24	81.9	85.48	86.34	86.8
900	87.34	87.56	82.96	85.63	86.73	87.6
1000	87.53	87.73	83.61	86.17	86.79	87.93
1100	87.45	87.63	82.75	85.6	86.55	87.5
1200	86.93	87.54	82.66	86.01	86.78	87.39
1300	87.13	87.52	82.61	85.96	87.06	87.37

Table -4: Base classifiers and hybrid of DT, KNN

Classifier models rate of accuracy (%)						
Resample (%)	DT	KNN	NB	SVM	MLP	Hybrid
None	77.42	72.7	76.66	76.15	68.62	77.81
100	79.08	73.34	79.08	80.1	80.23	79.85
200	82.46	75.89	78.95	83.61	82.27	82.78
300	83.72	80.1	80.57	83.93	85.25	84.23
400	85.59	82.62	80.68	83.65	84.53	86.7
500	86.89	86.66	81.3	84.55	87.09	87.68
600	86.31	86.44	81.19	84.89	86.12	87.54
700	86.42	86.37	82.45	85.42	85.9	87.5
800	86.4	86.24	81.9	85.48	86.34	87.45
900	87.34	87.56	82.96	85.63	86.73	87.98
1000	87.53	87.73	83.61	86.17	86.79	88.18
1100	87.45	87.63	82.75	85.6	86.55	88.08
1200	86.93	87.54	82.66	86.01	86.78	87.35
1300	87.13	87.52	82.61	85.96	87.06	87.44

Table -5: All classifiers and hybrid of DT, KNN, NB, SVM, MLP

Classifier models Build time(sec)						
Resample (%)	DT	KNN	NB	SVM	MLP	H5
None	0.06	0.01	0.01	4.27	54.34	66.91
100	0.01	0.01	0.01	4.05	53.42	56.67
200	0.02	0.01	0.02	3.76	107.4	113.2
300	0.02	0.01	0.01	4.77	157.8	168.2
400	0.01	0.01	0.01	7.33	212.9	222.6
500	0.03	0.01	0.01	8.74	270.6	276.8
600	0.03	0.01	0.01	18.82	318.4	335.1
700	0.02	0.01	0.01	21.22	365	393.8
800	0.04	0.01	0.01	13.32	413.1	437.5
900	0.04	0.01	0.01	14.16	473.8	516.7
1000	0.02	0.01	0.01	21.03	530.4	577.9
1100	0.02	0.01	0.01	30.97	569.1	622.2
1200	0.03	0.01	0.01	25.66	630.4	672.1
1300	0.04	0.01	0.01	36.58	674.8	741.8

Table -6: All classifiers and hybrid of DT, KNN, NB, SVM,

Classifier models Build time(sec)						
Resample (%)	DT	KNN	NB	SVM	MLP	H4
None	0.06	0.01	0.01	4.27	54.34	1.97
100	0.01	0.01	0.01	4.05	53.42	1.94
200	0.02	0.01	0.02	3.76	107.4	3.37
300	0.02	0.01	0.01	4.77	157.8	4.85
400	0.01	0.01	0.01	7.33	212.9	7.58
500	0.03	0.01	0.01	8.74	270.6	9.63
600	0.03	0.01	0.01	18.82	318.4	10.42
700	0.02	0.01	0.01	21.22	365	11.37
800	0.04	0.01	0.01	13.32	413.1	13.34
900	0.04	0.01	0.01	14.16	473.8	13.34
1000	0.02	0.01	0.01	21.03	530.4	17.83
1100	0.02	0.01	0.01	30.97	569.1	22.71

1200	0.03	0.01	0.01	25.66	630.4	24.8
1300	0.04	0.01	0.01	36.58	674.8	26.27

Table -7: All classifiers and hybrid of DT, KNN, NB

Classifier models Build time(sec)						
Resample (%)	DT	KNN	NB	SVM	MLP	H3
None	0.06	0.01	0.01	4.27	54.34	0.01
100	0.01	0.01	0.01	4.05	53.42	0.01
200	0.02	0.01	0.02	3.76	107.4	0.01
300	0.02	0.01	0.01	4.77	157.8	0.02
400	0.01	0.01	0.01	7.33	212.9	0.02
500	0.03	0.01	0.01	8.74	270.6	0.03
600	0.03	0.01	0.01	18.82	318.4	0.03
700	0.02	0.01	0.01	21.22	365	0.01
800	0.04	0.01	0.01	13.32	413.1	0.01
900	0.04	0.01	0.01	14.16	473.8	0.02
1000	0.02	0.01	0.01	21.03	530.4	0.03
1100	0.02	0.01	0.01	30.97	569.1	0.03
1200	0.03	0.01	0.01	25.66	630.4	0.05
1300	0.04	0.01	0.01	36.58	674.8	0.05

Table -8: All classifiers and hybrid of DT, KNN

Classifier models Build time(sec)						
Resample (%)	DT	KNN	NB	SVM	MLP	H2
None	0.06	0.01	0.01	4.27	54.34	0.01
100	0.01	0.01	0.01	4.05	53.42	0.01
200	0.02	0.01	0.02	3.76	107.4	0.01
300	0.02	0.01	0.01	4.77	157.8	0.02
400	0.01	0.01	0.01	7.33	212.9	0.03
500	0.03	0.01	0.01	8.74	270.6	0.03
600	0.03	0.01	0.01	18.82	318.4	0.03
700	0.02	0.01	0.01	21.22	365	0.03
800	0.04	0.01	0.01	13.32	413.1	0.03
900	0.04	0.01	0.01	14.16	473.8	0.02
1000	0.02	0.01	0.01	21.03	530.4	0.03
1100	0.02	0.01	0.01	30.97	569.1	0.03
1200	0.03	0.01	0.01	25.66	630.4	0.03
1300	0.04	0.01	0.01	36.58	674.8	0.03

From table 1, 2, 3, 4, 5, 6, 7 & 8, the study results demonstrate that optimal resample rate for all base classifiers was at resample rate of 1000%. An insight we can draw from the results is that class imbalance affects accuracy of classifiers. The results are consistent with studies which shows that imbalanced data set reduces performance and demonstrating the gains of using resampling in imbalanced data set (Garcia, Marques, & Sanchez, 2012)[11], NB performed worst with the lowest accuracy rate of 83.61%, and Build time of 0.01 seconds, followed by MLP with accuracy rate of 86.79%, and Build time of 530.4 seconds, SVM with accuracy rate of 86.17%, and Build time of 17.08 seconds, DT with accuracy of 87.53%, and Build time of 0.02 seconds the best was KNN with accuracy rate of 87.78%, and Build time of 0.01 seconds. The study answers the research question by determining the optimal design parameters for respective base classifiers and fit with the theory that different machine learning algorithms make different assumptions about the shape and structure of the function and how best to optimize a representation to approximate it, and why it is important to try a suite of different algorithms on a classification problem. The key question when dealing with classification is not whether a learning algorithm is superior to others, but under which conditions a particular method can significantly outperform others on a given application problem, and that KNN, NB, & DT take shortest Build time, SVM and MLP take longer Build time (Kotsiantis, Zaharakis, & Pinteals, 2006)[20]. Generalizability of the results is limited by resampling as a data approach to solving class imbalance. Hybrid of all five(H5): KNN, NB, DT, SVM and MLP performed worst with the lowest accuracy rate of 87.83%, and longest Build time of 577.9 seconds as shown in tables 1& 5, followed by hybrid of three(H3): KNN,DT,NB with accuracy rate of 87.93%, error rate of 12.07% and Build time of 0.03 seconds as shown in tables 3 & 7, hybrid of four(H4): KNN,DT,NB and SVM with accuracy rate of 87.98%, and Build time of 17.83 seconds as shown in tables 2 & 6, and the best combination was a hybrid of two with 88.18% accuracy rate, and 0.03 seconds Build time as shown in tables 4 & 8. The results fit with the theory that for good ensembles, base learners should be as more accurate as possible (Krogh & Vedelshy, 1995)[21] and reinforces the belief "many could be better than all" theorem may not be the fact (Zhou, Wu, & Tang, 2002)[37]. The study demonstrates that hybrid methods combine both selection and fusion techniques. The main idea is to use selection only and only if the best classifier is good enough to classify, otherwise a combination method is used (Vladislav, 2014)[34]. The results show that errors made by classifiers are independent and, and therefore the reason for majority vote hybrid outperform the best single classifier (Kim, Kim, Moon, & Ahn, 2011)[18]. The study determined the optimal design for a hybrid classifier and thereby answering the question, what was the best combination of base classifiers for an optimal hybrid classifier? The study is constrained by bagging as an ensemble combination technique. Generally, all classifiers yield highest accuracy rates at resample rate of 1000% and all combinations of hybrid classifier perform better than base classifiers. The accuracy rates improve with the resample rate up to 1000% and starts to decline as shown in in chart 1. The results confirm that better differentiation of classifiers can be done by examining the Build time (Williams, Zander, & Armitage, 2006)[35]. The inclusion of either an either MLP or SVM increases the overall Build time of a hybrid classifier as shown in chart 2. Increased computation is a weakness of ensembles, because in order to classify an input query, all component classifiers must be processed (Dietterich, 2000)[7]. Performance of various combinations of the hybrid perform differently in varying situations, confirming that generally, there is no ensemble method which outperforms other ensembles consistently (Ting & Witten, 1999)[30]. The study confirms that hybrid classifiers perform better than individual classifiers (Gundabathula &

Vaidhehi, 2018)[13]. DT performs better than NB confirming the results of a classification study between NB and DT (Rizwan, Masrah, Aida, Payam, & Nasim, 2013)[26]. The research answers the question, how does the performance levels of hybrid classifier compare with base classifiers?

3. CONCLUSION

In this study the problem of terrorism is addressed by using machine learning techniques. The GTD dataset is used to build machine learning algorithms. The dataset has been processed by using various pre-processing techniques and the class imbalance problem solved by using WEKA Resampling Filter. The pre-processed data is used to develop and evaluate models. The models are evaluated on various parameters like classification accuracy, classification error and Build time. The hybrid classifier is constructed by first constructing the base classifiers: NB, SVM, MLP, DT and KNN, and then the individual base classifiers are combined in parallel through bagging by use of majority voting rule technique. The classifiers were evaluated using 10 cross validation test option and Test split test option. Feature selecting was done in WEKA to remove features that had little relevance in determining terrorist groups. The study concludes hybrid classifier models perform better than traditional classifier models and that the best performing model was a hybrid combination of KNN, and DT with an accuracy rate of 88.18% and Build time of 0.03 seconds. Data imbalance affects the performance of machine learning classifiers and that a resample rate of 1000% yields optimum performance for classifiers.

ACKNOWLEDGEMENT

We wish to thank the National Consortium for the study of terrorism and response to terrorism (START) at the University of Maryland for granting access to their database for the purpose of this research.

REFERENCES

[1] Ahsan, S., & Shah , A. (2008). Data Mining, semantic Web and Advanced Information Technologies for fighting Terrorism. International symposium on Biometrics and Security Technologies, 1-5.

[2] Alexandrie, G. (2017). Surveillance cameras and crime: a review of randomized and natural experiments. Journal of Scandinavian Studies in Criminology and Crime Prevention, 210.

[3] Blair, R., Blattman, C., & Hartman, A. (2017). predicting local violence:Evidence from a panel survey in Liberia. Journal of Peace Research, 54,298-312.

[4] Bouziane, H., Messabih, B., & Chouarfia, A. (2011). profiles and majority voting based ensemble method for protein secondary structure prediction. Evolutionary Bioinformatics, 171-189.

[5] Breiman. (2001). Random forests. Machine Learning, 5-32.

[6] Coffman, T., & Marcus, S. (2004). Pattern Classification in Social network analysis: A case study, IEEE proceedings. Aerospace conference 5, pp. 3162-3175.

[7] Dietterich, T. G. (2000). An experimental comparison of three methods for constructing ensembles of Decision tress: bagging, boosting, and randomization. Mach Learn, 139-159.

[8] Dong, Q. (2017). Machine learning and conflict prediction: across disciplinary approach. world economics and politics, 7, 100-118.

[9] Farysal, G., Wasi, B. H., & Usman, Q. (2014). Terrorist group prediction using data classification. International Conferences of Artificial Intelligence and Pattern Recognition (pp. 17-19). Malaysia: Researchgate.

[10] Freitas, C. O., deCarvalho, J. M., & Oliveira, J. J. (2007). Confusion matrix disagreement for multiple classifiers. Progress in Pattern Recognition. Image Analysis and Application.

[11] Garcia, V., Marques, A., & Sanchez, J. (2012). Improving Risk Prediction by Pre processing imbalanced credit data.

[12] GTI. (2016). Global Terrorism Index. Institute for Economics and Peace.

[13] Gundabathula, V. T., & Vaidhehi, V. (2018). An Efficient Modelling of Terrorist Groups in India using Machine Learning Algorithms. India Journal of Science and Technology, 1-3.

[14] Han, J., & Kamber, M. (2006). Data Mining: Concepts and techniques. San Francisco: Morgan Kaufmann.

[15] Ivan, S. (2009). Has the global war on terror changed the terroridt threat? A time series intervention analysis. Studies in Conflict & Terrorism, 743-761.

[16] Kainulainen, L. (2010). Ensemble of locally linear models: Application to bankcruptcy prediction. Data Mining, 280-286.

[17] Khorsid, M. M., Abou, T. H., & Soliman, G. M. (2015). Hybrid classification Algorithms for Terrorism prediction in Middle East and North Africa,. International Journal of Emerging Trends & technology in Computer Science, 23-29.

[18] Kim, H., Kim, H., Moon, H., & Ahn, H. (2011). A weight adjusted voting algorithm for ensemble of classifiers. KoreanStatistical Society, 437-449.

[19] Kotsiantis, S. (2007). Supervised Machine Learning: a Review of Classification techniques. Informatica, 249-268.

[20] Kotsiantis, S., Zaharakis, I. D., & Pinteals, P. E. (2006). Machine learning: A review of classification and combining techniqies . springer, 23-32.

[21] Krogh, A., & Vedelshy, J. (1995). Neural network ensembles, cross validation and active learning. (G. Tesauro, D. S. Touretzky, & T. K. Leen, Eds.) Cambridge: MIT Press.

[22] Kuncheva, L. I., & Whitaker, C. J. (2003). Measures of diversity in classifier ensembles and their relationship with ensemble accuracy. Machine learning, 181-207.

[23] Ozgul, F., Erdem, Z., & Bowerman, C. (2009). Prediction of unsolved terrorist attacks using group detection algorithm. Pacific-Asia Workshop on Intelligence and Security Informatics (pp. 25-30). Berlin, Heidelberg: Springer.

[24] Piatestsky, G. (2014, 12 14). CRISP-DM, still the top methodology for analytics, data mining, or data science projects. Retrieved from Retrieved from KDnuggets: https://www.kdnuggets.com/2014/10/crisp-dm-top-methodology-analytics-data-mining-data-science-

projects.html): Retrieved from KDnuggets: https://www.kdnuggets.com/2014/10/crisp-dm-top-methodology-analytics-data-mining-data-science-projects.html)

[25] Popp, R., Armour , T., Senator, T., & Numrych. (2004). "Countering Terrorism through Information Technology". Communications of the ACM, 36-43.

[26] Rizwan, I., Masrah, A., Aida, A. M., Payam, H., & Nasim, K. (2013). An experimental study of classification algorithms for crime prediction. Indian Journal of Science and Technology.

[27] Sachan, A., & Roy, D. (2012). TGPM: Terrorist group prediction model for counterterrorism. International journal of Computer Applications, 44(10), 49-52.

[28] Song, N., Wang, K., Xu, M., Xie, X., Chen, G., & Wang, Y. (2015). Design and Analysis of Ensemble Classifier for Gene Expresion Data of Cancer. Advancement in Genetic Engineering, 1-2.

[29] Swanson, W. (2016). The eerie math that could predict Terrorist Attacks. Wahington: The Washington Post.

[30] Ting, K. M., & Witten, I. H. (1999). Issues in stacked generalization. Journal of Artificial Intelligence, 271-289.

[31] Tolan, G., & Soliman, O. (2015). An Experimental Study of Classification Algorithms for Terrorism Prediction,. International Journal of Knowledge engineering, 107-112.

[32] Verma, A. (2019). Evaluation of Classification Algorithms with Solutions to Class Imbalance Problem on Bank Marketing Dataset using WEKA. International Research Journal of Engineering and Technology, 54-60.

[33] Villada, R., & Drissi, Y. (2002). A perspective view and survey of meta-learning. Artificial intelligence, 77-95.

[34] Vladislav, M. (2014). Machine leearning of hybrid classification models for decision support, the use of the internet and development perspectives.

[35] Williams, N., Zander, S., & Armitage, G. (2006). A preliminary perofrmance comparison of five Machine learning algorithms for practical IP traffic flow classification. ACM SIGCOMM Communication Review, 5-16.

[36] Witten, I. H., Frank, E., & Hall, M. A. (2011). Data mining practical machine learning tools and techniques. Burlington: Morgan Kaufmann.

[37] Zhou, Z. H., Wu, J., & Tang, W. (2002). Ensembling neural networks: many could be better than all. Artificial Intelligence, 239-263.