Journal of Theoretical and Applied Information Technology
20th November 2015. Vol.81. No.2

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

CURRENT ISSUES IN ENSEMBLE METHODS AND ITS APPLICATIONS

NURA MUHAMMAD BABA[1], MOKHAIRI MAKHTAR[2], SYED ABDULLAH FADZLI[3] , MOHD KHALID AWANG[4]

Faculty of Informatics and Computing, University of Sultan Zainal Abidin, Tembila Campus, 22000 Besut, Malaysia[1,2,3,4].

E-mail: nabnurtrn@gmail.com[1], mokhairi@unisza.edu.my[2], fadzlihasan@unisza.edu.my[3], khalid@unisza.edu.my[4]

ABSTRACT

This paper reviewed the current state-of-the-art of optimization of ensemble methods so as to provide us with a better direction of how we will conduct our research in the future. The primary aim of ensemble method is to integrate a set of models that are used for solving different tasks so as to come up with enhanced composite global model, which produces higher accuracy and reliable estimate than what can be achieved through a single model. Diversity, combination strategies, number of based classifiers, types of ensemble, and performance measures are the key factors to be considered in the build of committees. When the numbers of base classifiers become huge, ensemble methods incurred high storage space and computational time, selective ensemble is proposed by most literatures to solve these problems. In terms of optimization techniques, multi-objectives techniques have become the better ones to use due to their efficiency in terms of optimization process and they provide a set of near optimal solution instead of just a single solution. When comparing the performance of ensemble methods, most of the time, accuracy alone cannot differentiate which classifiers perform best; for this reason, other performance measures such as AUC, F-measure, TPR, TNR, FPR, FNR, RMSE were used. Based on the reviewed literatures, we concluded that in our proposed methodology we would come up with a new method for comparing and searching for relevant classifiers from a collection of models that would be used as a model for predicting the quality of water to achieve higher performance rate than other previous work.

Keywords: *Ensemble Method, Classification Model, Diversity Measures, Performance Measures.*

1.0 INTRODUCTION

Ensemble learning is a machine learning discipline in which many base classifiers are trained on given datasets in order to provide a solution to given problems [1,2,3,4]. An ensemble consists of a group of base classifiers that are trained (such as neural networks or decision trees), whose decisions are integrated for classifying new instances [4] (see Figure 1). It is sometimes referred to as a mixture of experts [5, 6], committees [7], multi-classifier systems, fusion of experts [8], selection or thinning [9]. The main aim of ensemble method is to integrate a set of models that are used for solving different tasks so as to come up with enhanced composite global model which produces higher accuracy and reliable estimate than what can be achieved through a single model [10,11,12].

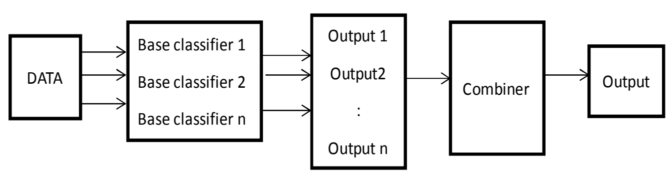

Figure1: General Concept Of Ensemble Classifier

In Figure 1, data is fed into various classifiers, different outputs were obtained, and they are then combined into a single output by the combiner.

The method falls into two categories, namely homogeneous and heterogeneous ensemble; if the ensemble is made up of the same type of learning algorithm, say neural network, then it is called homogeneous, but if it is made up of more than one different learning algorithm, for example, neural

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

network and decision trees, then it is referred to as heterogeneous [12,7]. However, when ensemble is built with models that are homogeneous, neither high accuracy nor diversity would make the ensemble obtain higher accuracy than the individual classifiers [13]. For this reason, we are going to make use of heterogeneous ensemble.

Ensemble method is chosen because it has been proven that it produces more accurate results than when a single model is used to solve the same problem [4]. Ensemble technology was introduced to the area of data classification and has since obtained great success. In order to achieve great success with the ensemble method, two criteria are taken into consideration, which are that the ensemble should have enough diversity introduced into it, and secondly, a suitable integrated method must be chosen in order to combine the decision of the base classifiers to a single output. The term diversity refers to the fact that indicates that the base classifiers errors are uncorrelated. Diversity is typically considered as a quantified estimation of the distinction of making the same errors among models in an ensemble [14] or it can simply be put as the difference between base classifiers in the ensemble [15]. It is grouped into two: pairwise and non-pairwise [16, 12]. Their examples can be found in literatures [12, 17] such as entropy, double default measure and Q-statistic to generalize diversity measures. However, the pairwise has the drawback of not having effectiveness in measuring diversity and shows no or little relation with the accuracy of the ensemble because it only considers the difference of two models and hence, they are not valuable [14]. Others split diversity into destructive and constructive or negative and positive diversity [14]. Combination or integration method is used to combine the output of the base classifiers in the ensemble. They are categorized differently in the literature, such as fusion, selection, and hybrid [7], static and adaptive [5], utility-based and evidence-based [18], evidence, fusion, genetic algorithm, and voting based aggregation techniques [20]. Fusion based methods are the ones in which all classifiers are assumed to be of equal experience in the whole feature space, and all classifier's decision are taken into account for any given input pattern. Examples are sum, majority voting, naïve Bayesian, neural networks, fuzzy neural networks, and fuzzy connectives among others. For selection based methods, only one classifier is needed to classify the input pattern correctly, for example, dynamic classifier selection, as suggested in [21], is one of the main methods. Hybrid Methods combine both selection and fusion techniques to provide the most suitable output to classify the input pattern. The main idea is to use selection only, and only if the best classifier is good enough to classify the test pattern, otherwise, a combination method is used. Examples are: dynamic classifier selection based on multiple classifier behavior and dynamic classifier selection based on decision templates [21]. Static combiners are independent of the feature vector. They are further subdivided into Trainable and Non trainable. The trainable combiner undergoes individual training phase to increase the ensemble performance, e.g. weighted averaging and stacked generalization. Non-trainable performs voting independently of the performance. Examples are: Borda count, averaging and voting. Adaptive means individual experts only need to perform well in their region of expertise and not on all inputs, e.g. a mixture of experts and Hierarchical mixture of experts. Utility-based is the type that does not make use of prior knowledge or evidence to make decisions, e.g. simple averaging, voting techniques while decision-based are the ones that use previous evidence to make decision, for example, Dempster-Shafer theory of evidence [22].

Other researchers suggest that the performance of the ensembles depends on two properties, which are the individual success of the base classifiers of the ensemble and the independence of the base classifier's results from each other [23]. Another researcher suggests that the accuracy of individual models, diversity among the individual models, decision making strategy, and number of base classifiers used for constructing an ensemble [24, 14] are among the factors responsible for the success of an ensemble.

There are two types of methods in machine learning, namely supervised and unsupervised learning. In supervised learning, there are defined rules, and the outcomes are known while in unsupervised learning, the algorithm follows certain rules to learn by itself and comes up with the result.

Ensemble methods had been widely applied in many fields such as web ranking algorithm, classification and clustering, time series and regression problems, and water quality application, among others.

Throughout the process of the research, four main steps are followed: choose your data, pre-process the data, apply classifiers to the pre-process data and select the combiner to output the result. To get the best out of the chosen classifier, the data is split into training and test set [25]. Training data set

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

is used to build and train the model while testing set is used for validating and testing the accuracy of the model [26].

Various researchers apply optimization algorithm to the ensemble methods in order to optimize their models. The use of optimization techniques becomes necessary because certain classification problems demand a high computation or are unfeasible to solve [7]. Optimization is defined as a branch of applied mathematics that deals with the minimization or maximization of a certain function, possibly under constraints. It has evolved towards the study and application of algorithms to solve mathematical problems on computers. Today, the field adds to numerous disciplines, extending from dynamical systems and control, statistics, calculations and complexity theory. It is used in different areas such as engineering design, machine learning and information retrieval, management, finance, and economics [29]. There are various optimization techniques that are used by researchers, among which are teaching learning-based algorithm (TLBO), Ant Colony Optimization (ACO), Genetic algorithm (GA), Particle Swarm Optimization, Harmony Search and Differential Evolution, etc.

We are going to make use of water dataset in our research because the prediction of water quality has become necessary due to the fact that polluted or unhealthy water had been extremely affecting the life of human's beings, plants and animals. This condition also leads to the outbreak of diseases. According to World Health Organization (WHO) in 1983, about 80% of all disease in human beings is water-borne [27]. Water is an essential requirement of human life and activities associated with industry, agriculture, and others, and it is considered as one of the most delicate parts of the environment [28].

2. CURRENT STUDY ON ENSEMBLE METHODS

The proposed paper by Santana et al. [7] compared the use of two optimization techniques in heterogeneous ensembles to find out whether an ensemble built with 3, 6 and 12 individual classifiers together with or without feature selection algorithm will perform better. Results showed that an ACO optimization outperformed GA in ensembles built with fewer individual classifiers while GA outperformed ACO in ensembles constructed with more individual classifiers.

In paper [8], the use of Meta-learning technique and multi-objective optimization was proposed to make the ensemble system to investigate how the initial configuration of an ensemble would affect the outcome of NSGA II optimization algorithm. Results indicate that when Meta-learning is used, a more accurate ensemble system is obtained in more than 50% of the cases analyzed. An empirical investigation of Rand, Meta, and Equal was conducted, in which the lowest error rate and statistically significant result was obtained with Meta.

The author in [13] develops a heterogeneous ensemble and a framework for constructing different kinds of ensemble for classifying spam emails. Results indicate that the heterogeneous ensemble can increase diversity as well as performance when compared to individual classifiers and other ensemble models.

Meanwhile, the authors in [14] focus on examining how the ensemble accuracy was impacted by what components and the degree of their effects. Three results were found, where they found that as the number of base classifiers increases, so does the diversity. Secondly, in terms of the accuracy, as the diversity increases, the accuracy also increases when voting strategy is used. Their final finding indicates that as the number of members increase, so does accuracy and diversity but the accuracy increases even further when odd numbers are used in building the ensembles instead of even numbers.

In [30], they compared two evolutionary algorithms based on stacking ensembles optimization techniques. Result showed that ACO algorithm is more flexible in terms of meta-classifier selection, has a larger search and GA ensemble can only find the best classifier for some dataset if it is either majority voting scheme or model tree. They concluded that GA is superior in terms of accuracy than ACO while ACO is more proficient than GA.

In [31], an ensemble approach was used to develop a global optimization method for classification models that aim to improve the accuracy of many classifiers on any given dataset. Results showed an overall improvement in all the classifiers, some high and others are very low, varying between 1% to 3% depending upon the complexity of the algorithm and how it handles bias and variance.

A study by [32] predicted the Water Quality Index of Thai Chin River using an ensemble built with support vector regression with radial basis and neural network created based on truth and falsity approach using four parameters. Genetic Algorithm

Journal of Theoretical and Applied Information Technology
20th November 2015. Vol.81. No.2

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

was used to assign weights to the selected parameters. Experimental results indicate that the ensemble provides better performance accuracy when compared to the result of individual base classifiers.

In [33], the author applied ensemble clustering to address three problems, which are inconsistency of result, knowledge beyond data and grouping with multiple objects. Inconsistency of result is solved through calibration while multiple objectives are solved by partitioning the data selectively prior to clustering.

[34] proposes a multi-view evolutionary algorithm that used multi-objective optimization approach for improving document clustering. Three contributions were achieved; the first contribution is to use multiple views to generate an initial set of candidate clustering solution. Secondly, to combine the individual clusters from different clustering solutions and the third contribution is to use the multi-objective ranking system from NSGA II to guide the optimization.

[35] introduced selective ensembles based on TLBO algorithm that comprises of these phases. The first phase is to reduce the training set through relief f algorithm to produce multiple training subsets through bootstrap technology. The training sets are trained with multiple base classifiers. RBFSVM is chosen as the classifier. Results showed an improved classification accuracy of 7.79% and a reduction of computational and storage space compared to other previous methods.

The research in [36] used an ensemble composed of seven ANN-based classifiers to predict porosity and permeability of petroleum reservoirs. Diversity was created by randomly assigning the number of hidden neurons and random selection of input data. R-Square, RMSE, and MAE statistical measures were used to compare the results. Results indicate that their model outperformed other ensemble techniques.

The author in [37] introduced a new method to calculate approximate gradient-based stochastic perturbation. Finding suggests that the constraints production optimization problem is solved successfully when the proposed method was combined with projected gradient methods.

An ensemble composed of eight features ranked algorithm was aimed to determine the cause of shellfish farm closure in various locations as carried out by [38]. Rain and salinity are the two main factors that indicate farm closure. Results further reveal that rain is the main cause of closure in the Southern and coastal region where land use is low whereas salinity in Northern and oceanic region where land use is high.

The approach in [39] employed a method for constructing ensemble based on SMBO. Extra computational cost was not generated at learning time by the method. Experimental result showed a generalized performance and convergence speed on 22 regressional and 39 classification data sets.

Research by [40] developed a new method for building dynamic ensemble from a collection of classifiers to predict sea water quality from spectrum channel data. GA search algorithm was used to optimize the ensemble. Experimental results were compared with SVM algorithm. Results showed that their method outperformed SVM, but the performance of the ensemble is critically affected by the quality of the population in the ensemble.

The use of the confusion matrix to compare the performance of multi-class predictive models of toxic and non-toxic class was studied by [41]. Results indicate that by converting multi classifiers confusion matrix to binary class, a simple solution was obtained that helped to analyze and categorize the performance of multi-class classifiers from a collection of models.

The research produced in [42] designed an ensemble of ANN network that determines the number of hidden neurons using randomization methods and applied it to the petroleum reservoir data set. Results indicate that their methods are more efficient than using other methods such as trial and error method for determining the number of hidden neurons, and also their methods outperform the accuracy of each base classifier.

The researchers in [43] aimed to improve random forest accuracy by optimizing large number of decision trees within the forest by choosing only uncorrelated and good trees. They improve the accuracy of the enhanced random forest through maximization of individual trees strength and minimize the correlation between the trees in the forest. An increased difference from 1% to 6% was achieved with the four experimental datasets.

3. DISCUSSION OF CURRENT STUDY

3.1 Reason for Ensembles

Among the advantages that ensemble method offers are: it reduces the probability of over fitting [44] and bias or variance error [45], it exploits the

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

idea that different classifiers can provide complementary information about patterns to be classified, thereby improving the effectiveness of the overall recognition process [16]. However, the ensemble method has a limitation of increased processing time of a system since they are more complex than single classifiers [7]. According to [46], ensembles models are constructed through six different sets: those that are established by (i) manipulation of the output labels, (ii) clustering, (iii) feature space, (iv) training patterns, (v) training parameters and (vi) error function.

When constructing an ensemble, the ensemble size affects the accuracy of the ensemble. If we use a smaller number of individual classifiers, the ensemble will not perform properly, whereas if we use a large number of individual classifiers, the ensemble accuracy improves but will lead to increase of storage space and computational time [8,48]. Due to the above problem, selective or pruning ensemble was proposed by many literatures such as [23, 49, 35, and 9] that determine the optimal number of base classifiers for the construction of the ensemble system. It also reduces the complexity, storage requirements and enhances the performance of the system [50].

Some authors reported that smaller size ensemble performs better than larger size ensemble for imbalanced dataset because the individual classifiers within the ensemble are so similar to each other, that adding too many such classifiers only makes the combined classifier over-complex [51]. Some researchers made use of homogeneous ensemble while others used heterogeneous ensemble.

3.2 Difference among Famous Ensemble Methods

Bagging and Boosting are among the most popular ensemble methods, and one of the most frequent ways that researchers used to compare their proposed ensemble method. A number of differences exist between bagging and boosting. Bagging uses majority voting scheme while boosting uses weighted majority voting during decision-making [30]. Classifiers generated by boosting are independent of each other while that of bagging are dependent on each other [8]. In bagging, instances selected to train base classifiers were bootstrapped duplicate copies of the training data, which implies that each instance has the same chance of being in each training dataset. While in boosting, the training dataset of each subsequent classifier gradually concentrates on instances misclassified by earlier trained classifiers [6]. But Bagging has the advantage over Boosting in that it reduces variance and minimizes error [31].

3.3 How are Performance of Ensembles Models Evaluated or Tested against other Methods?

Generally, when researchers construct their ensemble models, the next step is to evaluate its performance with other known famous ensemble methods such as bagging, boosting and Random Forest, so as to validate its effectiveness. These models are evaluated in terms of performance measures such as Accuracy, AUC, G-mean, F-measure, TPR, TNR, FPR, FNR, Mann-Whitney rank sum test, correlation coefficient (R2), root-mean squared error (RMSE), and mean absolute error (MAE), Hypothesis Test (t – test), Wilcoxon and Friedman test, among others. AUC was used by [43, 52, 53], as it had been shown that it is statistically consistent and more discriminant than accuracy [52] and also offers better performance measure when evaluating how well a classifier can stabilize the performance between classes [15]. AUC and G-mean are widely used especially in class imbalance learning. They have the advantage over recall, precision and F-measure in that they are better indicators to show the performance trade-off between classes than overall accuracy for their insensitivity to class distributions [54]. F-Measure is used by [55] as it had been shown by [56] to be a favorable measure. R square, RMSE and MAE as well as Algebraic rules can also be used as combination strategies.

3.4 Ways to Create or Increase Diversity

When an ensemble is built with identical base classifiers, there will not be any improvement in terms of accuracy than when a single classification model is used [21, 46]. For an ensemble to be useful, the base classifiers to be utilized in the construction needs to have features that are termed as diversity. Diversity is created in many different ways according to the literatures studied. In bagging, diversity was created through resampling, where each training subset is kept separate from each other. In random forests, it was created by choosing alternative branches to be put in decision trees while in boosting, diversity was achieved by concentrating on where the current model makes errors; and in stacking results, a bundle of various types of models are consolidated using an alternate model, rather than just voting [15, 57]. Diversity in ANN network can be created by assigning the number of hidden neurons randomly using a randomized algorithm [42] and through Random selection of input data for each base classifier [36].

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

Diversity is also achieved through manipulation of training parameters by assigning different values for each of the parameters, for example in [58], different weights are assigned to the network that are used to train the base neural network. Another method to achieve diversity is by manipulation of feature space, i.e. when the base classifiers were trained with different feature subsets such as in [59, 60]. In [6], diversity in ensemble was attained by using different subsets of training data, different subsets of the available features to train each classifier (i.e. random subspace method), using different parameters for the base classifier, and using different base classifiers as the ensemble members. However, some literatures such as [61, 62] reported that measuring diversity and using it explicitly does not necessarily relate to the success of constructing ensemble, also high diversity alone does not necessarily relate to the ensemble attaining higher accuracy especially when the base classifiers have little accuracy among them [13].

Usually, the combination method determines the based model to be used, for example, if using a combination method that returns probability values, say average combiner, then one cannot use SVM or linear discriminant analysis as the base model.

3.5 Why Use Optimization Techniques?

In terms of optimization techniques, many papers used different methods. Genetic algorithm is one of the most popular algorithms and was used by [7, 31, 48,] as it can deal with bigger search spaces and to look for near-optimal solutions without an exact depiction of the problem. In addition, they can undoubtedly fuse prior knowledge into the system [63]. In terms of feature selection, genetic algorithm (GA) has been shown to produce better results than other feature reduction techniques [65], as it deals with a population of solution rather than a single solution [66]. [35] used TLBO as it does not require any specific parameters to be set, has fast convergent rate, simple principle and globe optimization, [36]. PSO was used by [67], as it has the merit of not having an overlapped and mutation calculation [67] but cannot solve the problem of non-coordinated system [67]. It also suffers from high risk of getting trapped and not being able to be enhanced any further when the region explored by the particles happens to be of low quality than the earlier particle best option [68]. ACO was applied in [69] and [30], but according to [30], ACO wastes time in combining individual and meta-level classifiers and when making priority assigning as well as selection between strong and weak learners. But the evolutionary algorithms had the limitation of inability to converge on global optima [70], possibility of over-fitting [71], occasional inefficiency [72], as well as time and computational complexity [73]. In order to overcome the above problems, some authors opted to use multi-objective optimization approach such as NSGA II [48] as it enforces the search/optimization process and provides a set of near optimal solutions instead of just a single solution. In addition, when using multi-objective optimization, we can analyze the combination of the two most important objectives in the performance of an ensemble system, which are: accuracy and diversity, using two recently proposed measures [48].

3.6 Ways of Minimizing Attributes and Their Merits

To enhance the performance of models for any given dataset, selections of attributes are carried out by using a different method such as the use of feature selection or sometimes by the use of optimization techniques. Majority of literatures use feature selection as it aims to improve the quality of the obtained results in the sense that it provides different subsets of attributes for the individual classifiers, aiming to reduce redundancy, dimensionality among the attributes in the dataset and to increase diversity. Some literatures such as [7, 74, 75], used genetic algorithm and [76], ant colony optimization was used by [77] to select attributes for the base classifier. In utilizing feature selection techniques, its target is to enhance the quality of the obtained results. In the setting of ensembles, for example, feature selection techniques give diverse subsets of attributes to the base classifiers, aiming to decrease excess among the attributes of a pattern and to improve the diversity in such models [7].

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

4. PROPOSED ENSEMBLE METHOD

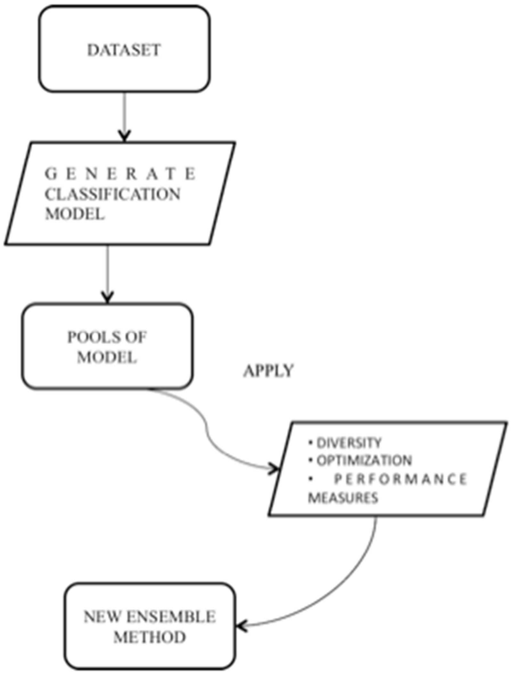

Figure 2: Proposed Ensemble Method

In Figure 2, a given dataset, i.e. water dataset is chosen. A set of models is then generated from pools of models that are compatible with the selected dataset. Diversity measures, for example pair wise diversity of which different types such as correlation coefficient, disagreement measure were applied so as to compute the diversity of the ensembles. Optimization techniques such as genetic algorithm and performance measure such as AUC, are applied to the dataset. Different results are then combined and compared from which a new or enhanced ensemble method would be chosen as the optimal method.

5. CONCLUSIONS

This paper reviewed several studies carried out by other researchers on ensemble models. The review work presented here provides us with the idea of how we will conduct our research using optimization techniques of ensemble methods for water quality application. Most of the researchers used heterogeneous ensemble approach in their work because it provides better performance. However, sometimes the ensemble method tends to increase storage space and computational time due to the presence of a vast number of base classifiers in the ensemble. For this reason, selective ensemble had been proposed by most literatures to tackle the drawback of the ensemble and so we choose to use it in our proposed methodology. Diversity as one of the most important factors used in the construction of ensemble is achieved through four ways, which are: using different subsets of training data, different subsets of the available features to train each classifier, different parameter values for the base classifier, and using different base classifiers as the ensemble members. In terms of optimization techniques, multi-objective techniques have become the better ones to use due to their efficiency in terms of search or optimization process and they provide a set of near optimal solution instead of just a single solution. When comparing the performance of ensemble methods, most of the time, accuracy alone cannot differentiate which classifiers performed best; for this reason, other performance measures such as AUC, FMeasure, TPR, TNR, FPR, FNR, RMSE were used. Based on the literatures reviewed, we concluded that in our proposed methodology, we would come up with a new method for comparing and searching for relevant classifiers from a collection of models that would be used as a model for predicting the quality of water to achieve higher performance rate than other previous researches.

In summary, the survey reviews a number of current issues in ensemble learning that include the identification of all the four ways of attaining diversity, current ensemble methods and their drawbacks. Among the optimization techniques, the multi optimization technique was found to be the best method. The work also discovered that the performance of the ensemble is not only measured by accuracy alone but by other techniques such as AUC, F-measure, TPR, TNR, FPR, FNR, Mann-Whitney rank sum test, correlation coefficient (R2), root-mean-squared error (RMSE), and mean absolute error (MAE), Hypothesis Test (t – test), among others.

6.0 ACKNOWLEDGEMENT

This research is partially supported by UniSZA (Grant No. UNISZA/13/GU/029) and (FRGS/2/ 2013/07/UniSZA/02/2). The first author acknowledges the financial support received from the Kano State Government, under the Dr. Rabi'u Musa Kwankwaso Administration.

REFERENCES

[1] Zhao Hui. (2013). *"Intrusion Detection Ensemble Algorithm Based on Bagging and Neighborhood Rough Set"*, International Journal of Security and Its Applications (IJSIA), Vol.7, No.5, pp. 193-204, SERSC.

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

[2] Chen Tao. (2011), *"Selective SVM Ensemble Based on Accelerating Genetic Algorithm"*, Application Research of Computers, Issue 1, pp. 139-141, Ori Probe Information Service.

[3] Zhou, Z. H., & Tang, W. (2003). *"Selective Ensemble of Decision Trees"*, In Rough Sets, Fuzzy Sets, Data Mining, and Granular Computing, Lecture Notes in Computer Science Volume 2639, pp. 476-483, Springer Berlin Heidelberg.

[4] Dietterich T., (2000). "*Ensemble Methods in Machine Learning"*, In J. Kittler and F.Roll, editors, First International Workshop on Multiple Classifier Systems, Lecture Notes in Computer Science, pp. 1-15. Springer-Verlag.

[5] Ricardo Gutierrez-Osuna () "*L25: Ensemble Learning"*, CSCE 666 Pattern Analysis CSE@TAMU" Lecture Notes pp. 1 – 15, Texas A&M University. Available online at: http://search.kedirijaya.com/detail/research.cs.tamu.edu/prism/lectures/pr/pr_l25.pdf. Accessed on December, 15,2014.

[6] Polikar, R. (2012). "*Ensemble learning"*, In C. Zhang and Y. Ma (eds.), *Ensemble Machine Learning: Methods and Applications*, Springer Science + Business Media, LLC 2012, pp. 1-34, Springer US.

[7] Santana, L. E. A., Silva, L., Canuto, A. M., Pintro, F., & Vale, K. O. (2010). "*A Comparative Analysis of Genetic Algorithm and Ant Colony Optimization to Select Attributes for an Heterogeneous Ensemble of Classifiers"*, In Evolutionary Computation (CEC), 2010 IEEE Congress on pp. 1-8, IEEE.

[8] Neto, A. A. F., & Canuto, A. M. (2014). "*Meta-Learning and Multi-Objective Optimization to Design Ensemble of Classifiers"*, 2014 Brazilian Conference on Intelligent Systems. pp. 91 – 96, IEEE.

[9] R. E. Banfield, L. O. Hall, K. W. Bowyer and W. P. Kegelmeyer (2002). *"Ensemble Diversity Measures and Their Application to Thinning"*, Information Fusion, vol. 6, no. 1, pp. 49-62, Elsevier B.V.

[10] Quinlan, J. R., (1996). "*Bagging, Boosting, and C4.5"*, In Proceedings of the Thirteenth National Conference on Artificial Intelligence, Vol. 1. AAAI Press, pp. 725-730, ACM Digital Library.

[11] Opitz, D. and Maclin, R., (1999). "*Popular Ensemble Methods: An Empirical Study"*, Journal Of Artificial Intelligence Research, Volume 11, pages 169-198, Cornel University Library.

[12] Kuncheva, L. I., & Whitaker, C. J. (2003). "*Measures of Diversity in Classifier Ensembles and Their Relationships with the Ensemble Accuracy"*, Machine learning, Volume 51, Issue 2, pp.181-207, Kluwer Academic Publisher.

[13] Wang, W. (2010). "*Heterogeneous Bayesian Ensembles for Classifying Spam Emails"*, The 2010 International Joint Conference on Neural Networks (IJCNN), pp. 1-8, IEEE.

[14] Wang, W. (2008). "*Some Fundamental Issues in Ensemble Methods"*, In Neural Networks, IJCNN. IEEE World Congress on Computational Intelligence. IEEE International Joint Conference on pp. 2243-2250. IEEE.

[15] Wang, S., & Yao, X. (2013). "*Relationships Between Diversity of Classification Ensembles and Single-Class Performance Measures"*, Knowledge and Data Engineering, IEEE Transactions on, vol. 25, No. 1, pp. 206-219. IEEE.

[16] L. I. Kuncheva, (2005). *"Combining Pattern Classifiers: Methods and Algorithms"*, New York: Wiley.

[17] K. Tang, P.N. Suganthan, and X. Yao, (2006). *"An Analysis of Diversity Measures,"* Machine Learning, vol. 65, pp. 247-271, Springer.

[18] K. Ghosh, Y.S. Ng, and R. Srinivasan, (2011). *"Evaluation of Decision Fusion Strategies for Effective Collaboration among Heterogeneous Fault Diagnostic Methods"*, Computers and Chemical Engineering, Volume 35, Issue 2, 9 February 2011, Pages 342–355, Elsevier.

[19] Woods, K., Kegelmeyer, W. P., & Bowyer, K. (1997). *"Combination of Multiple Classifiers Using Local Accuracy Estimates"*, IEEE Transactions on Pattern Analysis and Machine Intelligence, 19, pp. 405-410. IEEE.

[20] Shruti Asmita and K k Shukla. (2014). *"Review on the Architecture, Algorithm and Fusion Strategies in Ensemble Learning"*, International Journal of Computer Applications, Volume 108, Number 8, pp. 21-28.

[21] Kuncheva, L. I. (2004). "*Combining Pattern Classifier"*, Methods and Algorithms. John Wiley & Sons.

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

[22] Shafer, G. (1976). "*A Mathematical Theory of Evidence"*, Vol. 1, Princeton: Princeton University Press.

[23] G. Brown, J.L. Wyatt, and P. Tino, (2005). *"Managing Diversity in Regression Ensembles,"* The Journal of Machine Learning Research, vol. 6, pp. 1621-1650, ACM Digital Library.

[24] Makhtar, M., Yang, L., Neagu, D., & Ridley, M. (2012). "*Optimisation of Classifier Ensemble for Predictive Toxicology Applications",* In Computer Modelling and Simulation (UKSim), 2012 UKSim 14th International Conference on pp. 236-241. IEEE.

[25] J. Han, M. Kamber, and J. Pei, (2006). "*Data Mining: Concepts and Techniques"*, 2nd edition, Morgan Kaufmann.

[26] S. Kotsiantis, D. Kanellopoulos, and P. Pintelas, (2006) *"Data Preprocessing for Supervised Leaning",* International Journal of Computer Science, vol.1, no.2, pp.111–117, World Enformatika Society.

[27] Tebbutt TH (1983). "*Principles of Water Quality Control"*. 3rd Edn. pp. 42. Pergaman Press Oxford.

[28] J. Das and B. C. Acharya, (2003). *"Hydrology and Assessment of Lotic Water Quality in Cuttack City, India,"* Water, Air, and Soil Pollution, vol. 150, no. 1–4, pp. 163–175, Springer.

[29] Optimization Models https://inst.eecs.berkeley.edu/~ee127a/book/login/l_intro_main.html , (visited on 10 January, 2015)

[30] Gupta, A., & Thakkar, A. R. (2014). *Optimization of Stacking Ensemble Configuration Based on Various Meta heuristic Algorithms*. In *Advance Computing Conference (IACC), IEEE International* pp. 444-451, IEEE.

[31] Anwar, H., Qamar, U., & Muzaffar Qureshi, A. W. (2014). *Global Optimization Ensemble Model for Classification Methods.* The Scientific World Journal, Hindawi Publishing Corporation

[32] P. Kraipeerapun, S. Amornsamankul, *"Prediction of WQI for Tha Chin River Using an Ensemble of Support Vector Regression and Complementary Neural Networks",* Recent Advances in Information Science, Proceedings of the 7th European Computing Conference (ECC '13), pp. 36 – 41.

[33] Charkhabi, M., Dhot, T., & Mojarad, S. A. (2014*). "Cluster Ensembles, Majority Vote, Voter Eligibility and Privileged Voters",* International Journal of Machine Learning & Computing, Vol. 4, No.3, pp. 275 – 278.

[34] Wahid, A., Gao, X., & Andreae, P. (2014). "*Multi-View Clustering of Web Documents using Multi-Objective Genetic Algorithm"*, In Evolutionary Computation (CEC), 2014 IEEE Congress, pp. 2625-2632. IEEE.

[35] Tao Chen A, (2014). *"Selective Ensemble Classification Method on Microarray Data",*Journal of Chemical and Pharmaceutical Research, JCPRC5, vol. 6(6), pp. 2860-2866, CODEN(USA).

[36] Anifowose, F., Labadin, J., & Abdulraheem, A. (2013). "*Predicting Petroleum Reservoir Properties from Downhole Sensor Data using an Ensemble Model of Neural Networks",* In Proceedings of Workshop on Machine Learning for Sensory Data Analysis, pp. 27 - 34. ACM.

[37] Wei, S., Cheng, L., Huang, W., & Gu, H. (2014). "*A New Approximate Gradient Algorithm Applied in Constrained Reservoir Production Optimization",* Journal of Industrial and Intelligent Information Vol. 2, No.3, pp. 194 – 199, Engineering and Technology Publishing.

[38] Rahman, A., D'Este, C., & McCulloch, J. (2013*). "Ensemble Feature Ranking for Shellfish Farm Closure Cause Identification",* In Proceedings of Workshop on Machine Learning for Sensory Data Analysis pp. 13 – 18, ACM Digital Library.

[39] Lacoste, A., Larochelle, H., Laviolette, F., & Marchand, M. (2014). *Sequential Model-Based Ensemble Optimization*. *arXiv preprint arXiv:1402.0796*.

[40] Zeng, B., Luo, Z., & Wei, J. (2008). *"Sea Water Pollution Assessment Based on Ensemble of Classifiers",* In Natural Computation, 2008. ICNC'08. Fourth International Conference on Vol. 1, pp. 241-245. IEEE.

[41] Makhtar, M., Neagu, D. C., & Ridley, M. J. (2011). "*Comparing Multi-Class Classifiers: on the Similarity of Confusion Matrices for*

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

Predictive Toxicology Applications", In Intelligent Data Engineering and Automated Learning Ideal 2011, pp. 252-261. Springer Berlin Heidelberg.

[42] Anifowose, F., Labadin, J., & Abdulraheem, A. (2013). *"Ensemble Model of Artificial Neural Networks with Randomized Number of Hidden Neurons". In Information Technology in Asia (CITA), 2013 8th International Conference on* pp. 1-5. IEEE.

[43] Bharathidason, S., & Jothi Venkataeswaran, C. (2014). "*Improving Classification Accuracy based on Random Forest Model with Uncorrelated High Performing Trees",* International Journal of Computer Applications, vol. 101, issue 13, pp. 26-30, CROSSREF.

[44] M.P. Perrone and L.N. Cooper, (1993) *"When Networks Disagree: Ensemble Methods for Hybrid Neural Networks,"* Neural Networks for Speech and Image Processing, Chapman-Hill.

[45] Y. Sun, M.S. Kamel, A.K. Wong, and Y. Wang, (2007). *"Cost-Sensitive Boosting for Classification of Imbalanced Data,"* Pattern Recognition, vol. 40, no. 12, pp. 3358-3378, Elsevier.

[46] Hansen, L., & Salamon, P. (1990). *"Neural Network Ensembles",* IEEE Trans Pattern Analysis and Machine Intelligence , 12, pp. 993-1001. IEEE.

[47][49] A. Rahman and B. Verma, (2013). *"Ensemble Classifier Generation using Non–Uniform Layered Clustering and Genetic Algorithm*", Elsevier Knowledge Based Systems, vol. 43, pp. 30 – 42, Elsevier.

[48] Antonino A. Feitosa Neto, Anne M. P. Canuto and Teresa B Ludermir, (2013). *"Using Good and Bad Diversity Measures in the design of Ensemble Systems: A Genetic Algorithm Approach",* IEEE Congress on Evolutionary Computation, pp. 789 – 796. IEEE.

[49] Sylvester, J., Chawla, N. (2006). *"Evolutionary Ensemble Creation and Thinning*", In: IJCNN 06 International Joint Conference on Neural Networks, pp. 5148-5155, IEEE.

[50] Windeatt, T., & Zor, C. (2013). *"Ensemble Pruning using Spectral Coefficients",* IEEE Transactions on Neural Networks and Learning Systems, volume 24, Issue 4, pp. 673-678, IEEE.

[51] R.E. Schapire, Y. Freund, P. Bartlett, and W.S. Lee,(1998). *"Boosting the Margin: A New Explanation for the Effectiveness of Voting Methods",* The Ann. Statistics, vol. 26, no. 5, pp. 1651-1686, JSTOR.

[52] C.X. Ling, J. Huang, and H. Zhang, (2003). *"AUC: A Statistically Consistent and More Discriminating Measure Than Accuracy,"* Proceedings of the 18th international joint conference on Artificial intelligence (IJCAI '03), pp. 329-341, Morgan Kaufmann Publishers Inc. San Francisco, CA, US.

[53] A.P. Bradley, (1997). *"The Use of the Area under the Roc Curve in the Evaluation of Machine Learning Algorithms",* Pattern Recognition, vol. 30, no. 7, pp. 1145-1159, Elsevier Science Inc. New York, NY, USA.

[54] H. He and E.A. Garcia, (2009). *"Learning from Imbalanced Data,"* IEEE Trans. Knowledge and Data Eng., vol. 21, no. 9, pp. 1263-1284, IEEE.

[55] N.V. Chawla and J. Sylvester, (2007). *"Exploiting Diversity in Ensembles: Improving the Performance on Unbalanced Datasets,"* Proceedings of the 7th international conference on Multiple classifier systems, vol. 4472, pp. 397-406, Springer-Verlag Berlin, Heidelberg.

[56] M.V. Joshi, (2002). *"On Evaluating Performance of Classifiers for Rare Classes,"* Proc. IEEE Int'l Conf. Data Mining, pp. 641-661, IEEE.

[57] Data Mining with Weka MOOC Material http://www.cs.waikato.ac.nz/ml/weka/mooc/dataminingwithweka/transcripts/Transcript4-6.txt (3 January, 2015).

[58] T. Yamaguchi, K.J. Mackin, E. Nunohiro, J.G. Park, K. Hara, K. Matsushita, M. Ohshiro, K. Yamasaki, (2009). *"Artificial Neural Network Ensemble-Based Land-Cover Classifiers using Modis Data"*, Artificial Life and Robotics, vol. 13, issue 2, pp. 570–574.

[59] L.I. Kuncheva, J.J. Rodriguez, C.O. Plumpton, D.E. Linden, S.J. Johnston, (2010). *"Random Subspace Ensembles for FMRI Classification"*, IEEE Transaction on Medical Imaging, vol. 29, issue 2, pp. 531–542, IEEE.

[60] T. K. Ho, (1998). *"Random Subspace Method for Constructing Decision Forests,"* IEEE

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

Transactions on Pattern Analysis and Machine Intelligence, vol. 20, no. 8, pp. 832–844, IEEE.

[61] D. de Oliveira, A. Canuto, and M. de Souto, (2009)., *"Use Of Multi-Objective Genetic Algorithms To Investigate The Diversity/Accuracy Dilemma in Heterogeneous Ensembles,"* in International Joint Conference on Neural Networks (IJCNN), pp. 2339 –2346. IEEE.

[62] T. Windeatt, (2005). *"Diversity Measures for Multiple Classifier System Analysis and Design",* Information Fusion, Volume 6, Issue 1, Pages 21–36, Elsevier B.V.

[63] David E. Goldberg. (1989). *"Genetic Algorithms in Search, Optimization and Machine Learning",* (1st ed.). Addison-Wesley Longman Publishing Co., Inc., Boston, MA, USA.

[64] John H. Holland. (1992) "*Adaptation in Natural and Artificial Systems: An Introductory Analysis with Applications to Biology, Control and Artificial Intelligence",* MIT Press, Cambridge, MA, USA.

[65] M. L. Raymer, W. F. Punch, E. D. Goodman, L. A. Kuhn, and A. K. Jain, (2000). *"Dimensionality Reduction Using Genetic Algorithms",* IEEE Transactions on Evolutionary Computation, vol. 4, no. 2, pp. 164–171, IEEE.

[66] Santos, E., Sabourin, R., Maupin, P. (2006). *"Single and Multi-Objective Genetic Algorithms for the Selection of Ensemble of Classifiers",* In: International Joint Conference on Neural Networks, pp. 3070-3077, IEEE.

[67] Bai, Q. (2010). "*Analysis of Particle Swarm Optimization Algorithm"*. Computer and Information Science, Vol 3, No 1, pp. 180 – 184, Canadian Center of Science and Education.

[68] Shailendra S. Aote et al. (2013) *"A Brief Review on Particle Swarm Optimization: Limitations & Future Directions"*, International Journal of Computer Science Engineering (IJCSE), Volume 14– No.1, pp. 196-200.

[69] Chen, YiJun, and Man-Leung Wong. (2012). *"Applying Ant Colony Optimization in Configuring Stacking Ensemble",* Soft Computing and Intelligent Systems (SCIS) and 13th International Symposium on Advanced Intelligent Systems (ISIS), 2012 Joint 6th International Conference on. IEEE.

[70] Davoian, K., Reichel, A., Wolfram-Manfred, L. (2006). *"Comparison and Analysis of Mutation-Based Evolutionary Algorithms for Ann Parameters Optimization*". In Crone, S.F. Lessmann, S., Stahlbock, R. (eds.) International Conference on Data Mining (Las Vegas, Nevada, USA). pp. 51-56, CSREA Press.

[71] Batchis, P. (2013). "*An Evolutionary Algorithm for Neural Network Learning using Direct Encoding"*. Resource 53, Chinese Digital Library, Available online: www.cs.rutgers.edu/~mlittman/courses/ml03/iCML03/.../batc his.pdf. Accessed January 16

[72] Azzini, A. (2006). "*A New Genetic Approach for Neural Network Design and Optimization"*. PhD Dissertation, Universita Degli Studi Di Milano, pp. 38.

[73] Gao, W. (2012). "*Study on New Improved Hybrid Genetic Algorithm*", In: Zeng, D. (ed.) Advances in Information Technology and Industry Applications. Lecture Notes in Electrical Engineering 136, Springer, Volume 136, pp. 505-512. Springer Berlin Heidelberg.

[74] Lee, M. et al. (2008). "*A Two-Step Approach for Feature Selection and Classifier Ensemble Construction in Computer-Aided Diagnosis",* IEEE International Symposium on Computer-Based Medical System, pp. 548-553, Albuquerque: IEEE Computer Society.

[75] L Oliveira, M Morita and R Sabourin, (2006) "*Feature Selection for Ensembles Applied to Handwriting Recognition*", International Journal of Document Analysis and Recognition (IJDAR), Volume 8, Number 4, pp. 262-279, Springer-Verlag

[76] K Robbins, W Zhang and J Bertrand, (2007). *"The Ant Colony Algorithm for Feature Selection in High-Dimension Gene Expression Data for Disease Classification*". Mathematical Medicine and Biology pp. 413-426. Oxford University Press.

[77] H Kanan and K Faez. (2008). *"An Improved Feature Selection Method Based on Ant Colony Optimization (ACO) Evaluated on Face Recognition System*", Applied Mathematics and Computation, Volume 205, Issue 2, pp. 716-725, Elsevier.