2014 International Joint Conference on Neural Networks (IJCNN)
July 6-11, 2014, Beijing, China

Moving Towards Accurate Monitoring and Prediction of Gold Mine Underground Dam Levels

Ali N Hasan, Bhekisipho Twala, and Tshilidzi Marwala
The Center for Intelligent System Modeling (CISM), Department of Electrical and Electronic Engineering University of Johannesburg, South Africa

***Abstract*— In this paper a comparison between an ensembles (multi-classifier) constructed of several machine learning methods (support vector machine, artificial neural network, naïve Bayesian classifier, decision trees, radial basis function and k nearest neighbors) versus each single classifiers of these methods in term of gold mine underground dam levels prediction is presented. The ensembles as well as the single classifiers are used to classify, thus monitoring and predicting the underground water dam levels on a single-pump station deep gold in South Africa. In order to improve the classification accuracy an ensemble was constructed based on each single classifier performance, therefore, five ensembles were built and tested. In terms of misclassification error, the results show the ensemble to be more efficient for classification of underground water dam levels compared to each of the single classifiers.**

***Index Terms*— Support vector machines, classification, ensembles, neural networks, naïve Bayesian, gold mines, de-watering system, and underground dam levels.**

I. INTRODUCTION

South Africa is the major economic nation on the African continent. Mining has been the pillar of the South African economy for many years and has certainly contributed significantly to the economy and welfare of the country [1].

In deep gold mines, de-watering system, or clear-water pumping system is vital for mining process especially for cooling different mining levels and mining purposes. In spite of this, few studies have been conducted on controlling monitoring, analyzing, and predicting the underground dam levels [2]. It is very essential to monitor and observe the underground dam levels for the safety of miners and pumps [3].

The mine water pumping system generally consists of pumping stations with dams on specific underground levels, and in several cases fridge plants. The water being pumped from underground is water already used for mining purposes [4].

In the deep mines, underground dam levels must be monitored and controlled in order to ensure the dam's water level stays within safe limits, in order to prevent flooding or damage. These critical maximum and minimum levels are determined by the mine personnel [5]-[6].

Recently, an excessive amount of interesting research work has been done in the area of machine learning (ML) and artificial intelligence for prediction, classification and optimization purposes, in fields such as robotics, management and statistical sciences. There are many systems and methods that have been established to monitor and control the underground water pumping systems [5]-[6], but none of them uses state-of-the-art ML or artificial intelligence methods. Currently, there have been several applications for ML, the most significant being data mining. ML has also been successfully applied to improving the efficiency and accuracy of systems and the design of sophisticated machines [7]. Other ML applications include classification and prediction tasks, for instance, to monitor and predict how a given system would behave according to the present inputs and factors in terms of energy demand [8]. This work was carried out to inspect the viability of using machine learning and artificial intelligence in certain aspects of the mining industry. If successful, artificial intelligence systems could lead to improved safety and reduced risks and accidents.

Ensembles or multi-classifier methods have recently become as a common learning method, not only because of their straightforward implementation, but also due to their outstanding predictive performance on practical and real-life problems [9]. An ensemble contains a set of individually trained classifiers (for example decision trees or neural networks) whose predictions are combined when classifying distinctive instances. Ensemble methods aim to improve the predictive performance of a given statistical learning or model fitting technique. The general principle of ensemble methods is to create a linear combination of specific model fitting method, instead of using a single fit of the method [10]. Earlier, researches have shown that an ensemble is often more accurate than any of the single classifiers in the ensemble. Two relatively new but famous methods for creating ensembles are Bagging and Boosting [25].

The major contribution of the paper is the comparison between six solid single classifier methods (artificial neural networks and support vector machines, k nearest neighbors, naïve Bayesian, decision trees and radial basis function), on the one hand, against a combination of the same classifiers (ensemble) in terms of their ability and accuracy to predicting underground dam levels in a South

African mine. The comparison between these algorithms is applied on a single pump station in a deep mine in order to determine the best method (in terms of predictive accuracy) that the mine could apply when monitoring underground dam levels.

The layout of the paper is as follows: section 2 gives a mine layout situated in South Africa. In Section 3 methods used in the current investigation in the paper are briefly described. Comparative experiments on dam levels databases are presented in Section 4 followed by the major results in Section 5. Section 6 contains concluding remarks.

II. MINE LAYOUT

Mine A is situated in the North West of South Africa. This is mine mainly produces gold and Uranium. Figure 1 shows the pumping system layout. The mine has one main pump stations.

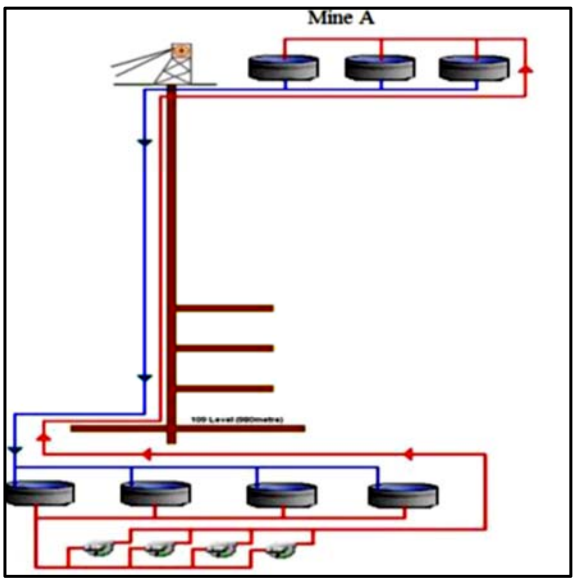

Fig. 1. Mine A clear-water pumping system [5]

Water is pumped directly from the underground dams to the surface dam. From the surface dam, some of the water is fed back to the various mine levels for mining purposes, while the rest goes to the gold plant. The average capacity of each underground dam was considered for monitoring its water level as all four dams are connected. On the other hand, the surface dam level was not taken into consideration as it has sufficient capacity to accommodate all the mine water without any risk of flooding [5] [28].

III. MACHINE LEARNING ALGORITHMS

A. *Artificial Neural Network (ANN)*

The first algorithm to test is ANN. Neural network is one of the significant components in Artificial Intelligence (AI) [11]. It has been studied for many years with the objective of achieving human-like performance in several fields, for instance speech and image recognition, as well as information retrieval. [12]. Basically, an artificial neural network is a system on its own that receives an input, process the data, and delivers an output [13]. Multilayer Perceptron (MLP) is a network of perceptrons. A perceptron is the simplest neural network representing a linear hyper-plane within instance space [12]. MLP's can be used to solve complex problems. Each MLP contains an input layer that contains at least one hidden layer and an output layer. A layer is an arrangement of neurons that include hidden ones which do not have any connection to the external sources [13].

The neuron output is the threshold weighted sum of all inputs from the previous layer. This process is continued iteratively until the error can be tolerated or reaches specific threshold. Activation functions use the input into the neurons to compute the output, which is comprised of weighted sums of the outputs from the previous layer [20].

B. *Support Vector Machine (SVM)*

Support Vector Machine method (SVM) is finding application in pattern recognition, regression estimation, and operator inversion for ill-posed problems [14]-[11]. SVM, or as it is called SMO in the Waikato Environment for Knowledge Analysis (WEKA), can be used to solve two-class (binary) classification problems. These classifiers find a maximum margin linear hyper-plane within the instance spaces that provides the greatest separation between the two classes [15]. Instances that are closest to the maximum margin linear hyper-plane from the support vectors are correctly classified [14].

Among the possible hyper-planes, SVMs choose the one where the distance of the hyper-plane from the nearest data points (the "margin") is as large as possible [15]. Once instances from the support vector have been recognized, the maximum margin linear hyper-plane can be created [15]-[20].

C. Decision Trees (DT)

A decision tree is a decision support tool that uses a tree-like graph or model of decisions and their possible consequences, including chance event outcomes, resource costs, and utility. It is one way to display an algorithm [16]. A decision tree builds an interpretable model that represents a set of rules. It is a popular tool for classification that is relatively fast to train and to use to make predictions. This decision tree has several advantages. Firstly, it naturally handles missing data. That is, when a decision is made on a missing value both sub-branches are traversed and a prediction is made using a weighted vote. Secondly, it naturally handles nominal attributes. For instance, the number of splits can be made equal to the number of nominal values. Moreover, a binary split can be made by grouping the nominal values into subsets (called sub-setting). While a decision tree is fast to train, one disadvantage is that it requires a large number of examples to make significant splits (to create a more general model) [16]. Decision trees are commonly used in operations research, specifically in decision analysis, to help identify a strategy most likely to reach a goal [17].

D. Naïve Bayes' Classifier (NBC)

The naïve Bayes' classifier gives a simple approach, with clear semantics, to representing, using, and learning probabilistic knowledge [18]. Basically, a naive Bayes' classifier assumes that the presence (or absence) of a particular feature of a class is unrelated to the presence (or absence) of any other feature, given the class variable. It is based on applying Bayes' theorem with strong (naïve) independence assumptions, or more specifically, independent feature model [18]. A naïve Bayes' classifier is a famous and popular technique because it is very fast approach and gives a high accuracy [17].

E. Radial Basis Function Classifier (RBF)

The radial basis function (RBF) network is a special type of neural networks with several distinctive features [19]. A RBF network consists of two layer feed-forward neural network. In between the input layer and the output layer there is a hidden one with hidden processing units which implement the radial basis function. The input layer broadcasts the coordinates of the input vector to each of the units in the hidden layer [20]. Each unit in the hidden layer then produces an activation based on the associated radial basis function. Finally, each unit in the output layer computes a linear combination of the activations of the hidden units [20].

F. k Nearest Neighbour Classifier (kNN)

The aim of the k Nearest Neighbours (kNN) method is to use a data-set wherein the data points are separated into few separate classes to predict the classification of a new sample point [21]. kNN is a solid classifier of nonparametric discrimination, or supervised learning [22]. Each instance, to be classified, is characterized by c values x_i , $i = 1 \ldots c$ and is therefore represented by a point in c-dimensional space . The distance between the two instances can be defined in different ways, the simplest one is the usual Euclidean metric [22].

G. Ensemble Classifier or multiple classifiers

An ensemble of classifiers is a set of classifiers whose individual decisions (weighted votes) are combined in some way to classify new examples [23]. Several techniques of combining the predictions of multiple classifiers have been investigated to produce a single classifier [24]. The resulting classifier (hereafter referred to as an *ensemble*) is usually more accurate than any of the single classifiers that are used to construct the ensemble. Ensemble has many other names such as, ensemble methods, committee, classifier fusion, combination, aggregation…etc [25]. Both theoretical and empirical research has showed that a good ensemble is one where the individual classifiers in the ensemble are both accurate and make their errors on different parts of the input space [25].

IV. EXPERIMENTAL SET-UP

1- Individual classifiers

Data for underground dam levels was gathered for a period of three months by using pressure transmitter fitted on the dams. This pressure transmitter is connected to a programmable logic controller's (PLC) fixed on the pump station, then via fibre optics to a supervisory control and data acquisition (SCADA) system to log the data on a spread sheet. It logs a value every two seconds.

In this experiment the WEKA software is used to classify the dam level data for the six single classifiers ANN, SVM, NBC, kNN, DT and RBF. For the simulation purposes the data was averaged over 30 second's intervals. For all the algorithms default parameters are used. WEKA is a software which was developed at Waikato University in New Zealand. It is a collection of open source of numerous data mining and machine learning algorithms [26]-[27]. The data split was, 80% to train and 20% to test for all algorithms, as in the previous methods. Data was trained and tested separately. The maximum and minimum dam levels in this mine were provided by the mine's shaft engineer. The maximum is 65% and the minimum is 25%. The data is categorized in classes for simulation. Table 1 illustrates the classes for the underground dam level.

TABLE 1: Underground dam level classes

Class	Dam Level	Numeric Dam Level
1	Low Level	0%- 40%
2	Medium Level	41%- 55%
3	High Level	More than 55%

After specifying the classes and the split-to-train percentage, the data was processed by WEKA and the most suitable classifier that achieved the maximum number of correctly classified instances was determined.

2- Ensemble classifier

For the ensemble experiment the split-to-train percentage was 60% to train and 40% to test. To construct the ensemble, MATLAB (matrix laboratory) simulator was created and programmed. The underground water dam level's data was organized in arrays as an ".m" file which can deal MATLAB function, script, or class to be suitable for the MATLAB simulator. This simulator included all the six classifiers and used majority vote algorithm to combine the classifiers output. Majority vote was used because it gives the best results among other combination algorithms, such as the sum, the maximum, the minimum, the average, products and the Bayes algorithms for this particular case.

The ensemble was constructed based on the most accurate classifier (i.e. started with the most two accurate classifiers, then add the less accurate one, until combining all the classifiers). The ensemble only used for the underground water dam's level as it will be seen later. The steps below summarize the ensemble construction.

1- Divide original dataset into K training datasets, TR1, TR2, ...,TRk;

2. Create k single models (N1, N2, N_k) with the different training datasets TR1, TR2, …, TRk to determine k single classifiers (ensemble members) generated by different algorithms.

3. Select the most accurate single classifiers from n classifiers.

4. Combine the most two accurate single classifiers, and then add the third less accurate one until all six classifiers have been combined, and find out which ensemble achieves the highest accuracy among the five possibilities.

7. Combine the multiple classifiers into aggregate output using majority voting algorithm.

6. Compare the classification accuracy between each ensemble starting with the first two single classifier's accuracy to the all classifiers accuracy.

The main performance measures used in this experiment is the classification and the misclassification accuracy. The secondary measures are the mean square error (MSE) and root mean square error (RMSE).

V. EXPERIMENTAL RESULTS

1- Individual classifier's results

For the underground dam level classification, classifiers performed as shown in Table 2.

Table 2: Underground dam level's classification results

Description	ANN	SVM	DT	NBC	kNN	RBF
Misclassification error	40.50%	41.06%	54.40%	41.56%	41.76%	40.50%
Mean absolute error	0.3406	0.3455	0.4124	0.4016	0.3455	0.3530
Root mean squared error	0.4212	0.4443	0.4614	0.4425	0.4443	0.4214

This performance needs to be improved in term of classification accuracy. However, ANN and RBF are the best classifiers with accuracy of 59.5% and the worst performance is DT with accuracy of 45.5%. SVM, kNN and NBC have all almost the same performance with a slight advantage for SVM.

2- Ensemble classifier's results

As mentioned in the experimental set-up section, five ensembles were built and tested with the data (ens1, ens2, ens3, ens4, and ens5) with a training percentage of 60% and 40% for testing. The main performance measure used for the ensemble is the classification accuracy. The five ensembles were constructed as follows based on the performance accuracy of each single classifier:

1- ens1: MLP + RBF

2- ens2: MLP+RBF+KNN

3- ens3: MLP+RBF+KNN+SVM

4- ens4: MLP+ RBF+SVM+KNN+NBC

5- ens5:MLP+RBF+SVM+KNN+NBC+DT

Ensemble prediction accuracy results are shown in figure 2. It can be seen that the accuracy for the underground dam level prediction was strongly improved. The highest accuracy was achieved by ens3 with 89%. ens1 achieved the lowest accuracy of 73%.

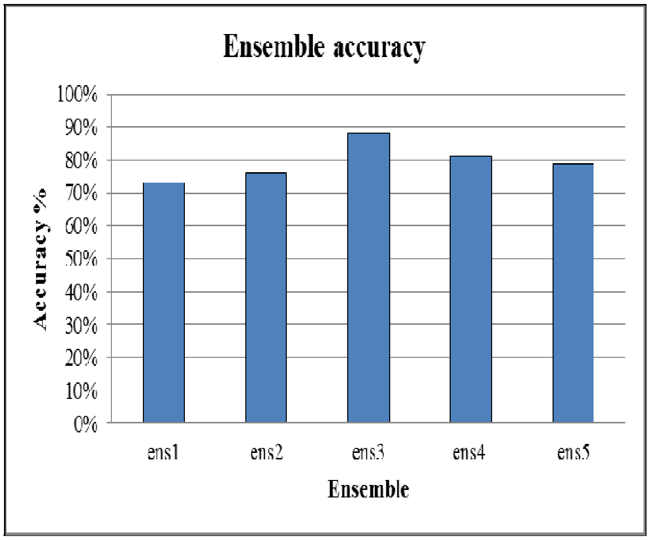

Fig. 2: Ensemble prediction accuracy

VI. CONCLUSIONS

For single classifiers, ANN and RBF were the leading methods with 59.5% accuracy, while DT had the highest misclassification error of 54.4%. This is most likely due to the dam level data inconsistency caused by the continuous dam water fluctuations, as well as system complexity, both tend to favor the ANN and RBF, however results needed to be improved to suit the critical nature of the application. Thus, an ensemble method was presented by combining the most accurate single classifiers in order to improve the prediction accuracy of the underground dam levels. Five different ensembles were built and tested. The ensemble method improved the accuracy and reached a classification accuracy of 89%. This means that the accuracy improved from 59.5% (ANN and RBF) to 89% (ens3). It can be noticed that all five ensembles performed better than any single classifier for dam level experiment. However it is also worth to mention that constructing an ensemble is more complex in terms of time and structure than using a single classifier. Prediction results suggest that using artificial intelligence in monitoring and controlling the mine de-watering system could be efficient and applicable in certain mining aspects. However each mine has to be treated as separate case, as these results may differ as each mine has its own structure and layout. This work could be investigated and applied on other mining aspects, such as compressed air network, smelters and so on.

REFERENCES

[1] Georg J. Coakly, “The mineral industry of South Africa”, U.S. Geological Survey Minerals. Yearbook-2000, 2000.

[2] Preliminary energy outlook for South Africa, Energy Research Institute, Dept of Mechanical Engineering, University of cape Town, Private Bag, Rondebosch 7701, October 2001.

[3] C Chen, B Das, D J Cook, “Energy Prediction Based on Residents Activity”, Washington State University, Pullman WA 99164, USA. Available on http://www.eecs.wsu.edu/~cook/pubs/kdd10p2.pdf.

[4] A N Hasan, “Maximising Load Shifting Results with Minimal Impact on Water Quality”, Final project report presented in partial fulfillment of the requirements for the degree Master of Engineering, Electrical Engineering, North West University, Potchefstroom campus, November 2009.

[5] E H Mathews, A N Hasan, R P Pelzer, “Case studies: The Environmental Impact Of DSM Projects”, ICUE, Cape Town, 2008.

[6] A N Hasan, K Ouahada, T Marwala, B Twala, “Optimization of the Compressed Air-Usage in South African Mines”, Africon IEEE, Livingstone, 2011.

[7] S. Piramuthu,” Evaluating feature selection methods for learning in data mining applications”, European Journal of Operational Research, 2004, available online at www.sciencedirect.com.

[8] D. Michie, D.J. Spiegelhalter, C.C. Taylor, “Machine Learning, Neural and Statistical Classification”, February 17, 1994.

[9] L. Breiman, “Bagging predictors”, Machine Learning 24(2), P 123-140, 1996.

[10] P. Buhlmann, “Bagging, Boosting and Ensemble Methods”, ETH Zurich, Seminar f¨ur Statistik, HG G17, CH-8092 Zurich, Switzerland, 2010.

[11] I. Witten, E. Frank, “Data Mining, Practical Machine Learning Tools and Techniques”, second edition, ISBN: 0-12-088407-0, 2005

[12] “Neuro AI - Intelligent systems and Neural Networks”,http://www.learnartificialneuralnetworks.com/, last accessed on March 2012.

[13] A Folleco, T Khoshgoftaar, J Van Hulse, A Napolitano, “ Identifying Learners Robust to Low Quality Data”, Florida Atlantic University, Florida, USA, October 12, 2008.

[14] C.J.C. Burges and B. Schölkopf, "Improving the Accuracy and Speed of Support Vector Machines", Advances in Neural Information Processing Systems, 1997.

[15] J. P. Lewis, “A Short SVM (Support vector Machine) Tutorial”, CGIT Lab/IMSC, version 0.zz, University of Southern California, December 2004.

[16] Y. Yuan and M.J. Shaw, “Induction of fuzzy decision trees”. Fuzzy Sets and Systems 69 (1995), pp. 125–139.

[17] T. Mitchell, H. McGraw, “Machine learning”, Second Edition, Chapter One, January 2010.

[18] A. W. Moore, “Naïve Bayes’ Classifier”, School of Computer Science, Carnegie Mellon University, 2004.

[19] M. J. L. Orr, "Introduction to radial basis function networks," Technical Report, Center for Cognitive Science, University of Edinburgh, 1996.

[20] V. Kecman, “Learning and Soft Computing: Support Vector Machines, Neural Networks, and Fuzzy Logic Models”, the MIT Press, Cambridge, Massachusetts, London, England, 2001.

[21] M. Tsypin, H. Röder, “On the Reliability of kNN Classification”, Proceedings of the World Congress on

Engineering and Computer Science, San Francisco, USA, October 2007.

[22] O. Sutton, “Introduction to k Nearest Neighbor Classification and Condensed Nearest Neighbour Data Reduction”, February, 2012.

[23] D. Optiz, R. Maclin, “Popular Ensemble Methods: An Empirical Study”, Journal of Artificial Intelligence Research, P 169-198, August 1999.

[24] L. Xu, A. Krzyzak, and C.Y. Suen, “Methods of Combining Multiple Classifiers and Their Applications to Handwriting recognition”, IEEE Transaction on Man and Cybernetics, Vol 22, No 3, May/June 1992.

[25] S. Vemulapalli, X Luo, J. Pitrelli, I, Zitouni, “Using Bagging and Boosting Techniques for Improving Coreference Resolution”, Informatica 34, P 111-118, 2010.

[26] R R. Bouckaert, E Frank, M Hall, R Kirkby, P Reutemann, A Seewald, D Scuse, “WEKA Manual for Version 3-7-3”, University of Waikato, December 6, 2010.

[27] W. Wang, “Tutorial for WEKA a Data Mining Tool”, University of East Anglia, February 19, 2010.

[28] A. Hasan, B. Twala, T, Marwala, “Predicting Mine Dam Levels and Energy Consumption Using Artificial Intelligence Methods”, IEEE Symposium Series on Computational Intelligence SSCI, Singapore, 2013.