Predicting Mine Dam Levels and Energy Consumption Using Artificial Intelligence Methods

Ali N Hasan, Bhekisipho Twala and Tshilidzi Marwala
Department of Electric and Electronic Engineering Science, University of Johannesburg, South Africa

Abstract**—** Four machine learning algorithms (artificial neural networks, a naïve Bayes' classifier, a support vector machines and decision trees) were applied for a single pump station mine to monitor and predict the dam levels and energy consumption. This work was undertaken to investigate the feasibility of using artificial intelligence in certain aspects of the mining industry. If successful, artificial intelligence systems could lead to improved safety and reduced electrical energy consumption. The results show neural networks to be more efficient when compared with support vector machines, a naïve Bayes' classifier and in particular, decision trees in terms of predicting underground dam levels. Artificial neural networks showed 60% accuracy, out-performing support vector machine, naïve Bayes' classifier and decision trees. For the prediction of water pump energy consumption, an artificial neural network and a naïve Bayes' classifier had the same accuracy of 99.0%, whereas a support vector machine and decision trees achieved a lower accuracy.

Index Terms**—** machine learning algorithms; deep gold mines, de-watering system, underground pump stations, energy consumption.

I. INTRODUCTION

There has been a growing concern about the total mining energy consumption in the South African mines, a substantial user of energy. In fact, the mining sector is a major electricity consumer in South Africa. It consumes approximately 23% of the total power generated [1]. In spite of this, very few studies have been carried out on monitoring, analyzing, and predicting energy usage [3].

In deep gold mines, de-watering systems are very important aspect of mining operations especially for the cooling of different mining levels and during mining. It is necessary to monitor underground dam levels to ensure the safety of miners and monitor pumps, as well as the pump power consumption to lower electricity costs [4].

A de-watering system is also a vital part of the mine system [4]. It consists mainly of underground dams, pumps, water pipes and surface dams, and in some cases refrigeration plants. Water is pumped from the underground dams to the next (higher) level or to a dam on the surface. Some mines have more than one underground pump station and others have only one.

In deep mines, underground dam levels must be monitored to ensure the dam's water level stays within safe limits, to prevent flooding or damage. These critical maximum and minimum levels are determined by mine personnel. This water cycling system is a relatively large energy consumer, and may therefore represent an opportunity for energy optimization and hence cost savings [5].

Several systems and methods have been developed to monitor the underground de-watering systems [5; 6], but none of them uses state-of-the-art Machine Learning (ML) or artificial intelligence algorithms. Recently there have been several applications for ML, the most significant being data mining. ML has also been successfully applied to improving the efficiency of systems and the design of advanced machines [7]. Other ML applications include classification and prediction tasks, for example to monitor and predict how a given system will behave in terms of energy demand according to the present inputs and factors [8].

In this paper a comparison has been made between Artificial Neural Networks (ANN), Support Vector Machines (SVMs), naïve Bayes' classifier (NBC) and Decision Trees (DT) in terms of their ability to predict underground dam levels and pump power consumption. The comparison between these algorithms was applied to a single pump station in a deep mine to determine which of these systems (ANN, SVM, NBC and DT) was the best at monitoring and predicting the underground dam levels and pump power consumption.

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

II. MINE LAYOUT

Mine ―A" is in a deep gold mine situated in South Africa. The mine de-watering system (Fig. 1) consists of one underground pump station, surface water dam and four underground dams. This pump station is comprised of four pumps. Two of these pumps have an installed power of 2.76 MW each; the remaining two are rated at 3.1 MW each.

978-1-4673-5851-4/13/$31.00

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.

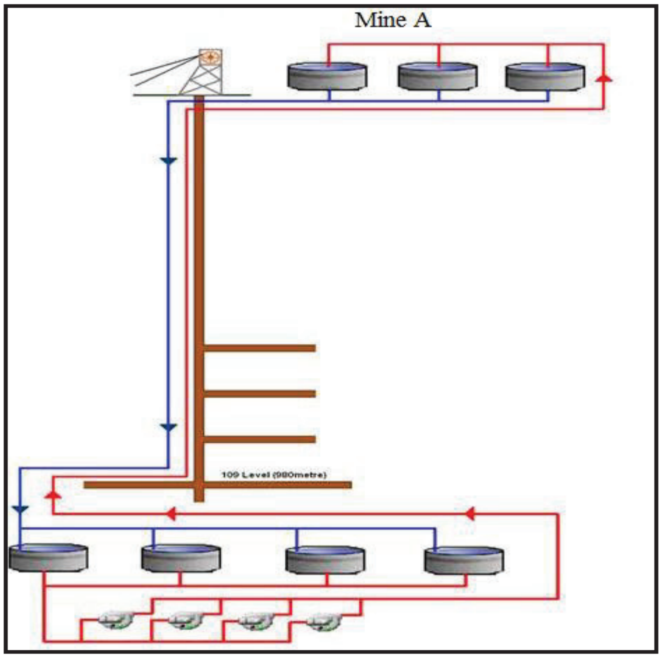

Fig. 1: Mine ―A” de-watering system [5]

III. MACHINE LEARNING ALGORITHMS

A. *Artificial Neural Network (ANN)*

The first algorithm tested was the ANN. The Neural network is one of the important components in Artificial Intelligence (AI). It has been studied for many years with the objective of achieving human-like performance in several fields, such as speech and image recognition, as well as in information retrieval. [16] Basically, an artificial neural network is a system on its own that receives an input, processes the data, and provides an output [11].

The Multilayer Perceptron (MLP) is a network of perceptrons. A perceptron is the simplest neural network representing a linear hyper-plane within instance space [12]. MLPs can be used to solve complex problems. Each MLP contains an input layer that includes 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].

An MLP is typically built as a back-propagation neural network. In a back-propagation neural network, the error is fed back to the same neuron. 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 calculate the output, which is composed of the weighted sums of the outputs from the previous layer [13].

B. *Support Vector Machine (SVM)*

A Support Vector Learning Machine (SVM) is finding application in pattern recognition, regression estimation, and operator inversion for ill-posed problems [12]. A Support Vector Machine Classifier (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. Instances that are closest to the maximum margin linear hyper-plane from the support vectors are correctly classified [13].

Among the possible hyper-planes, SVMs select the one where the distance of the hyper-plane from the closest data points (the ―margin”) is as large as possible [18]. Once instances have been identified from the support vector, the maximum margin linear hyper-plane can be constructed [13].

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 [17].
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) [17].

Decision trees are commonly used in operations research, specifically in decision analysis, to help identify a strategy most likely to reach a goal. If in practice, decisions have to be made online with no recall under incomplete knowledge, a decision tree should be paralleled by a probability model as a best choice model or by an online selection model algorithm. Another use of decision trees is as a descriptive means for calculating conditional probabilities [15].

D. Naïve Bayes' Classifier (NBC)

The naïve Bayes' classifier gives a simple approach, with clear semantics, to representing, using, and learning probabilistic knowledge [19]. 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 [20]. It is based on applying Bayes' theorem with strong (naïve) independence assumptions, or more specifically, independent feature model [21]. A naïve Bayes' classifier is a famous and popular technique because it is very fast approach and gives a high accuracy [21].

There are many datasets for which naïve Bayes' classifier does not do so well. That is because attributes are treated as if they are completely independent; the addition of redundant ones slants the learning process [20]-[21].

IV. EXPERIMENTAL SET-UP

Data for dam levels was collected over two months by using a pressure transmitter fitted onto the dams. This pressure transmitter was connected to a Programmable Logic Controller (PLC) fixed on the pump station, then connected via a optic fibre to a Supervisory Control And Data Acquisition (SCADA) system to log the data into spreadsheet. It logged a value every two seconds.

Data on the pump's energy consumption was logged for two months using a power logger mounted on the pump control panels, and connected to the underground PLCs and then to the SCADA on the surface. This data represents the run time status for each pump, where the value 1 denotes pump is on status, and a 0 denotes the off status. Each pump had a specified power capacity which determines the power consumption.

In this experiment the WEKA software was used to classify the mine data (power and dam level) using the ANN, SVM, and DT algorithms. WEKA is software which was developed at the Waikato University in New Zealand. It is a collection of open source programs with many data mining and machine learning algorithms [9; 10].

The dam level and pump's data was tested on a neural network of 5 hidden layers using WEKA. Each layer used the sigmoid function as the activation function. The data was split into 80% for training, and 20% for testing for both dam levels and energy consumption. The number of instances was 456 (sum of weights) and the attributes were five (pump 1, pump 2, pump 3, pump 4 and class). As mentioned before, the pumps were directly linked to the dam level, so the attributes here represented the pump's running status (on, off) and the class was the dam levels.

The SVM model includes three classifiers. The first one was the classifier for classes 0, 1 with 556 kernel evaluations. The second classifier was for classes 0, 2 and had 682 kernel evaluations. The third and last classifier was for classes 1, 2 and had 430 kernel evaluations.

A decision stump tree and a naïve Bayes' classifier were used to train the data. Once again, data was split to 80% to train and 20% to test for both algorithms, as in the previous methods.

According to the mine shaft engineer, the maximum and minimum dam levels for this mine were 65% and 30% respectively. The data was categorized in classes for simulation as shown in Table 1.

Table 1: 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%

Energy data was categorized for the classes as shown in Table 2. The energy data represents the energy consumed by the underground pumps.

Table 2: Energy consumption classes

Class	Energy consumption description	Energy level
1	No Energy Consumption	0 kW
2	Medium Energy Consumption	2760 kW
3	High Energy Consumption	5860 kW

V. EXPERIMENTAL RESULTS

Table 3 shows the number of instances (counts) and weight for each class.

Table 3: Class labels and weights

Class label	Count	weight
1	113	113.0
2	217	217.0
3	126	126.0

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

Table 4 shows ANN and SVM achieved higher accuracy for predicting dam levels compared to NBC and DT. In fact, in terms of all but one, performance measures, there was no significant difference in performance between SVM, NBC and ANN at the 5 % level of significance.

Table 4: Dam level prediction

Description	ANN	SVM	DT	NBC
Misclassification error	**40.00%**	44.00%	54.22%	45.00%
Kappa statistic	0.227	0.239	**0.119**	0.230
Mean absolute error	0.311	**0.344**	0.423	0.402
Root mean squared error	**0.411**	0.442	0.473	0.444
Relative absolute error	83.75%	**80.36%**	97.53%	93.89%
Root relative squared error	**90.43%**	95.00%	99.10%	95.21%

For the energy consumption data, the accuracy was higher than the dam level learner as shown in Table 5. There was a tie between ANN and NBC as they both achieved the highest accuracy rates in predicting energy consumption (99.0%), followed by the SVM (97.5%). However, the root mean square error for ANN was higher than for the NBC. The worst performance was by DT, with an error rate of 10.5%. Similar patterns were observed for the other performance measures with DT however, showing great similarities between the classes compared to the other three methods.

Table 5: Energy consumption

Description	ANN	SVM	DT	NBC
Misclassificatio n error	1.00%	2.50%	10.50%	1.00%
Kappa statistic	0.990	0.970	0.770	0.99
Mean absolute error	0.0054	0.222	0.108	0. 0009
Root mean squared error	0.0076	0.280	0.221	0.0024
Relative absolute error	1.65%	68.40%	30.91%	0.29%
Root relative squared error	1.86%	67.28%	54.81%	0.60%

VI. CONCLUSIONS

It can be seen from the results that the Artificial Neural Network (ANN) technique was the most accurate predictor among the four methods used for dam level prediction, with a 60% accuracy which is low accuracy for such an application, thus and to improve the accuracy an ensemble is recommended to be constructed out of the used classifiers. For the energy consumption, the ANN and NBC prediction accuracy was equal as both achieved a 99.0% accuracy, but the root mean squared error for ANN was higher than NBC and that favors NBC to be best performer. Support vector machine (SVM) had 56% prediction accuracy for dam levels and a 97.5% accuracy for energy consumption. A naïve Bayes' classifier accuracy for dam level gave 55%, while Decision Trees (DTs) had the lowest accuracy among the three methods as it had a 45.7% prediction accuracy for dam levels and 89.5% accuracy for energy consumption prediction. The four methods (ANN, NBC, DT and SVM) are known to give good results as classifiers. However, in this experiment and for the dam level part, ANN had slightly outperformed SVM, NBC and DT. This is most likely due to the nature of the dam level data and its continuous fluctuations as well as system complexity which tend to favor the ANN. For the energy prediction part, NBC and ANN had the same accuracy but a higher root mean squared error for ANN which made NBC the considered winner. That is most likely from the simplicity of the energy data.

VII. REFERENCES

[1] Statistics South Africa, ―Mining: Production and sales", Statistical release P2041, Available at: www.statssa.gov.za, Postal address: Private BagX44, Pretoria, 0001.

[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 at 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 Ouhada, T Marwala, B Twala, ―Optimization of the Compressed Air-Usage in South African Mines", Africon IEEE, Zambia, 2011.

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

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

[9] 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.

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

[11] ―Neuro AI - Intelligent systems and Neural Networks”, http://www.learnartificialneuralnetworks.com/, last accessed in March 2012.

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

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

[14] J Moody, C Darken, ―Fast Learning in Networks of Locally Tuned Processing Units”, Neural Computation, 1(2): 281-294, 1989

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

[16] H. Qin, ―Neural Network and Its Application in IR”, University of Illinois at Urbana-Champaign, 1999.

[17] X. Li, ―ATutorial on Induction of Decision Trees”, School of Information Technology and Electrical Engineering, The University of Queensland.

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

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

[20] T. Mitchell, H. McGraw, ―Machine Learning”, second edition, chapter one, January 2010.

[21] A. W. Moore, ―Naïve Bayes' Classifier”, School of Computer Science, Carnegie Mellon University, 2004.