

Early detection of mining truck failure by modelling its operation with neural networks classification algorithms

H. HU and T.S. GOLOSINSKI

Mining Engineering, University of Missouri-Rolla, Missouri, USA

The paper reports on use of neural network algorithms to model performance and condition of mining trucks. The model inputs are selected digital data streams collected by truck Vital Information Management System, VIMS, during its operation in a mine. The models built as the result of this work allow for projection of truck condition and performance into future. As such they facilitate more efficient management of mine truck fleets.

Keywords: Modelling, modeling, neural networks, mining equipment, Vital Information Management System, Intelligent Miner, mining truck

Introduction

Modern mining equipment is equipped with numerous sensors that monitor its condition and performance. Data collected by these sensors are used to alert the operator to existence of abnormal operating conditions and to perform emergency shutdown if the pre-set upper or lower limits of the monitoring parameters are reached. These data are also used for post-failure diagnostics and for reporting and analysis of equipment performance. Availability of these voluminous data, together with availability of sophisticated data processing methods and tools, allow for extraction of additional information contained in the data. One set of methods that permit this is data mining^{1,2}.

In previous research the authors attempted to identify and extract this information by modelling mining truck operations using Decision-Tree algorithms. These models used VIMS @ collected data as inputs and were built to predict VIMS events. While these models have shown reasonable average prediction error rate (21%), the high standard deviation (26%) of error rate for test data sets indicates instability of the resulting predictions. As an example a typical model implies 14% probability of false alarms related to high engine speed on one hand, and 50% probability of missing high engine speed alarms on the other. Thus, these models are not sufficiently reliable and can not be used for prediction of events that may take place^{3,4}. Research presented in this paper uses neural network based modelling to improve prediction accuracy. This approach has allowed for significant reduction in errors of predictions for both the truck performance and its condition into the future. The IBM Intelligent Miner for Data was used to build the relevant models.

Data preparation

The data analysed in this research is of two types; it includes snapshot (event recorder) and datalogger VIMS records, each containing values of 70 truck parameters recorded over a period of time. These data were collected at six Caterpillar 789B trucks during their operation in a surface mine.

One snapshot stores a segment of truck history that

contains values of all 70 monitored parameters recorded during the period of six minutes. Value of each parameter is recorded once per second. The snapshot recording is triggered by occurrence of an abnormal situation or event, such as when a monitored parameter reaches a pre-set critical value. A snapshot record describes truck conditions during five minutes before the event and one minute after the event takes place. In this paper, every snapshot record is called 'event' for simplicity.

Unlike snapshot, the data logger records values of all truck parameters that are monitored by VIMS over varying periods of time, also at one-second intervals⁵. The recording needs to be triggered and stopped manually, with individual records covering periods of up to 30 minutes of truck operation. Datalogger records do not have to be associated with any events.

Of the 70 truck parameters monitored in the field, values of 26 were recorded as categorical and the remaining 44 as numeric values. The example of basic statistical description of the numerical parameter values is presented in Table I. The actual model inputs are statistics of recorded truck parameter values calculated for one-to three-minute time intervals of truck operation. The statistics include:

- Minimum
- Maximum
- Range
- Average
- Standard deviation
- Variance
- Regression slope
- Regression intercept
- Regression sum of square.

Neural networks

Research described in this paper used neural network algorithms to model truck operation. Neural networks are computer implementations of sophisticated pattern detection and machine learning algorithms used to build predictive models from large historical databases. They allow for construction of highly accurate predictive models that serve to solve a large number of different problems.

Table I
Example of numerical parameter values

Parameter name	Minimum value	Maximum value	Mean value	Standard deviation
AFTRCLR_TEMP_110	0	95	41.8	12.8
AMB_AIR_TEMP_791	0	38.5	21.9	7.0
ATMOS_PRES_790	0	93	89.4	9.1
BOOST-PRES_105	0	164	31.0	50.1

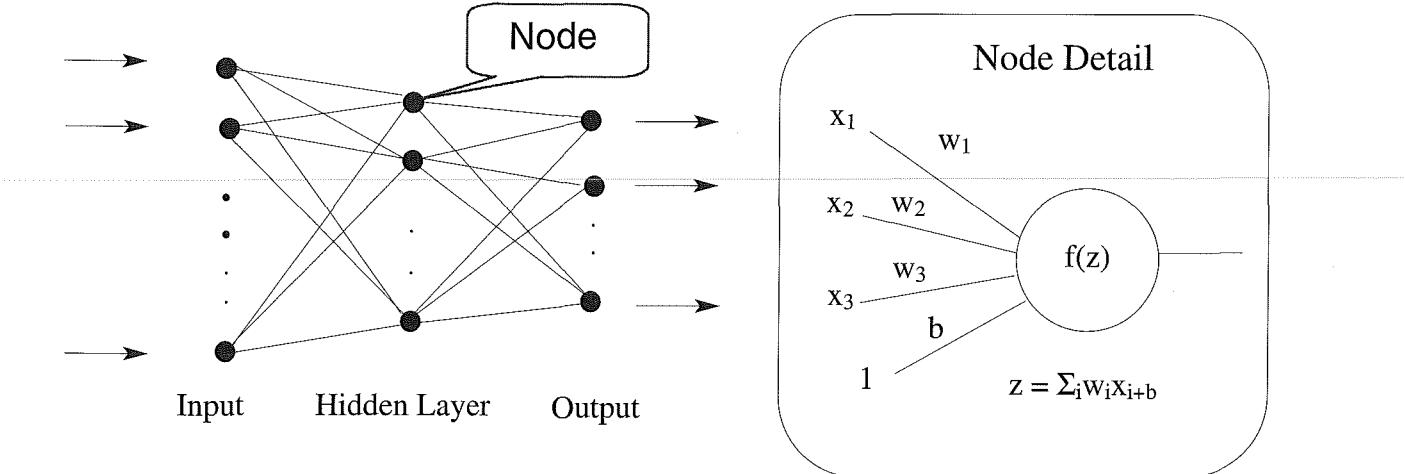


Figure 1. Back-propagation neural network architecture⁸

Various visualization techniques were used in conjunction with neural models to help explain and control the model, and to assure its clarity and transparency. IBM Intelligent Miner implementation of relevant algorithms was used to construct the models the tool. The ‘back propagation’ of a specific algorithm was used to train the neural network.

Discovery of an effective method of training a multiplayer neural network led to the re-emergence of neural networks as a tool for solving a wide variety of problems. This training method is called *back propagation* (of errors) or the generalized delta rule. It is simply a gradient descent method that minimizes the total squared error of the output computed by the network⁶.

A multilayer neural network with one layer of hidden units (Z units) is shown in Figure 1. The input layer contains the statistical parameters used to predict VIMS events. The output are the predicted VIMS events. The left side of Figure 1 is a basic neuron, output of which is calculated by the activation function $f(z)$. One of the most typical activation functions is the binary sigmoid function, which has the range of (0, 1) and is defined as Equations [1], [2].

$$f_1(x) = \frac{1}{1 + e^{-x}} \quad [1]$$

with the following property

$$f'_1(x) = f_1(x)[1 - f_1(x)] \quad [2]$$

The Z is calculated as the sum of the products of the inputs and related weights. The output units (Y units) and the hidden units may have biases, while the input value is

always one. Only the direction of information flow for the feed-forward phase of operation is shown in Figure 1. During the back-propagation phase of learning, signals are sent in the reverse direction.

VIMS event modelling

To facilitate modelling of truck operations the pattern of changes in VIMS parameter values was analysed, as associated with various events. The objective was to identify such patterns in changes of parameter values that, if repeated, may provide early indication of an impending failure. The other VIMS data were then screened for presence of these patterns. Initial analysis and modelling were done by building a decision tree classification model of the truck. The model was to allow for prediction of an occurrence of a selected event, based on the pattern of changes in values of other parameters.

Event prediction

The ‘high engine speed’ events were most numerous in analysed data. Consequently these were chosen to be the main targets of modelling and analysis. Caterpillar defines the ‘engine speed’ as the actual rotational speed of the crankshaft. For the modelled truck, this event is activated when the engine speed reaches 2250 rpm and deactivated when the speed drops to 1900 rpm. For comparative analysis VIMS data collected during normal operation was selected, grouped in ‘other’ class.

In this work the truck operation was divided in three-, two-, and one-minute intervals. For each interval the patterns of interest were sought. Figure 2 illustrates VIMS

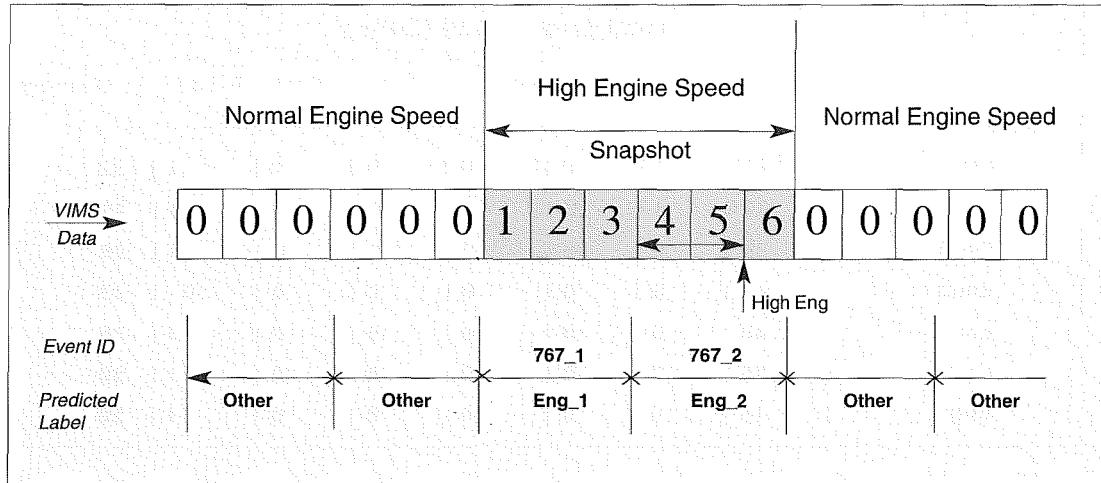


Figure 2. VIMS event prediction model

event prediction model of one VIMS event, ‘high engine speed’ based on three-minute truck operation intervals. To predict occurrence of this event statistics are defined for each three-minute interval of VIMS records. If the analysed three-minute VIMS records show similar statistics as do the first three minutes of the ‘high engine speed’ snapshot, the probability exists that the high engine speed will occur after another two minutes of truck operation.

As shown in Figure 2, the one-minute interval model can predict events that will occur within the next 4 minutes of truck operation. Similarly, the two-minute interval model can provide predictions of events that may occur within the following 3 minutes. The three-minute model can only predict events that will occur within the following two minutes³.

Modeling

The two main procedures of data mining are training, called also model construction, and testing called also model validation. In training mode, the function builds a model based on the selected input data. This model is later used as a classifier. In the test mode, the function uses a set of data to verify that the model created in the training mode produces results in satisfactory precision of predictions.

In this work all available data were split into two parts. Bulk of the data, 86.4%, were used for model training. There are remainder, 13.6% of available data, were used for model testing. The test data include dataset #1 (random selection) and dataset #2 (the whole set of all snapshot and datalogger data).

Three models were built based on the one-, two-, and three minute interval statistics. The error rate was defined for each and used for evaluation of the training and the testing processes. This was done by analysing confusion matrix determined for each model.

Results

During modeling, the one-minute and two-minute based models were unable to converge to predefined error rate after as many as 2000 passes in their training mode. Only three-minute model was able to satisfactorily converge. To quantify the distribution of the misclassifications of the neural network model runs the confusion matrix was used.

In every matrix, the number on the diagonal indicates the correct classification; other numbers indicate misclassification. The confusion matrix obtained for one-minute and two-minute interval statistics show that the network has not learned enough to be able to differentiate ‘high engine speed’ from ‘Other’. Only the model built from the three-minute statistical data is able to do so (Figures 3 to 7).

Discussion

Figure 8 compares results of Neural Network Classification modeling and that of Decision Tree Classification for three-minute truck operation intervals. The Neural Network Classification model run has lower error rates of training and test for dataset #1.

The error rates of the individual classes are presented in Table II. While Decision Tree model has 6% and 1% lower average error on training and test data of dataset #1, the Neural Network has 20% lower error rate (standard deviation) as evident in Figure 9. While the Neural Network model has somewhat higher overall error rate, it shows much higher stability. The performance prediction is quantified by accuracy of model runs on two datasets, ‘test #1’ and ‘test #2’. This compares to 14% of alarms missing and 50% false alarm probability in Decision Tree model. While runs of the Neural Network model show 17% of alarms missing, the probability of false alarm is only 29%. Thus Neural Network models allows for making more accurate prediction of the ‘high engine speed’ VIMS.

The Intelligent Miner incorporates the sensitivity analysis function, which is represented by a list of input fields ranked according to their respective importance to the classification function. The results are normalized so that they total 100%. A parameter that is listed as having 20% score is twice as important in making the desired classification as a parameter with a 10% score. The parameters with highest sensitivity, defined for the Neural Network model and not lower than 0.5 are shown in Table III. This Table also gives parameter descriptions.

Conclusions

The modelling approach presented in this paper compresses the VIMS data into statistical table and facilitates prediction

Total Errors = 24 (24%)										
Predicted Class -->		Other	Eng1	Eng3	Eng2	Eng4	Eng6	Eng5	Unknown	
Other		1353		0		0		0		L 32
Eng1		37		1		0		0) 28
Eng3		39		0		0		0) 28
Eng2		42		0		0		0		0 23
Eng4		40		0		0		0) 28
Eng6		36		0		0		0) 30
Eng5		15		0		0		0		3) 10

Figure 3. Training with one-minute statistical data

Total Errors = 15 (15.31%)						
Predicted Class -->		OTHER	ENG1	ENG2	ENG3	Unknown
OTHER		643		17		0 15
ENG1		14		44		0 10
ENG2		39		14		0 15
ENG3		58		2		0 5

Figure 4. Training with two-minute statistical data

Total Errors = 72 (12.7%)						
Predicted Class -->		OTHER	ENG1	ENG2	Unknown	
OTHER		373		47		3 15
ENG1		3		60		1 2
ENG2		1		0		68 1

Figure 5. Training with three-minute statistical data

of events with certain accuracy. The time horizon for predictions is two minutes.

Performance of Neural Network Classification modeling on the analysed data differs from that of Decision Tree Classification modeling. While the Neural Network model

built on the three-minute statistics of VIMS data sets shows an error rate close to that of the Decision Tree model, the error rate of standard deviation is only 6%. It follows that the first shows more robustness. The balanced probability that an alarm will be missed and that a false alarm will be

Total Errors = 15 (23.8%)					
Predicted Class -->		OTHER	ENG1	ENG2	Unknown
OTHER		41	8	1	1
ENG1		3	4	0	1
ENG2		1	0	3	0

Figure 6. Test #1 with three-minute statistical data

Total Errors = 9 (14%)					
Predicted Class -->		OTHER	ENG1	ENG2	Unknown
OTHER		44	6	0	2
ENG1		0	6	0	0
ENG2		1	0	5	0

Figure 7. Test #2 with three-minute statistical data

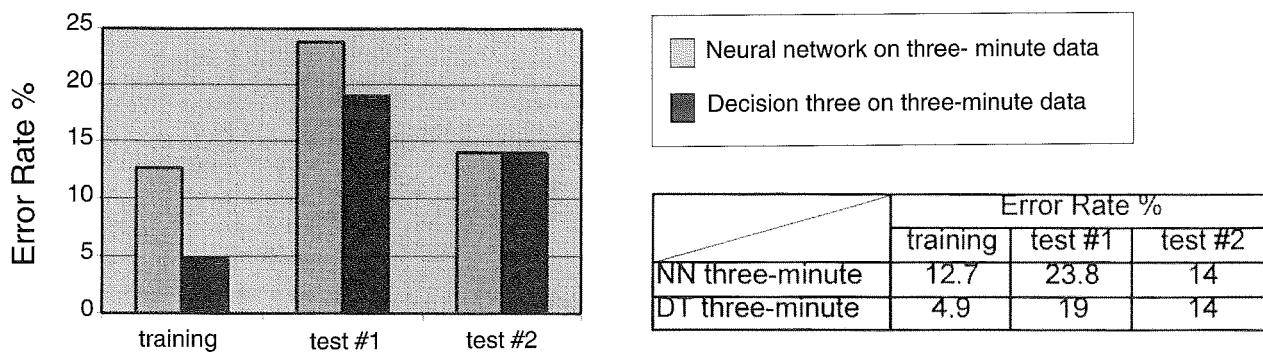


Figure 8. Total error rate comparison between Neural Network and Decision Tree classification

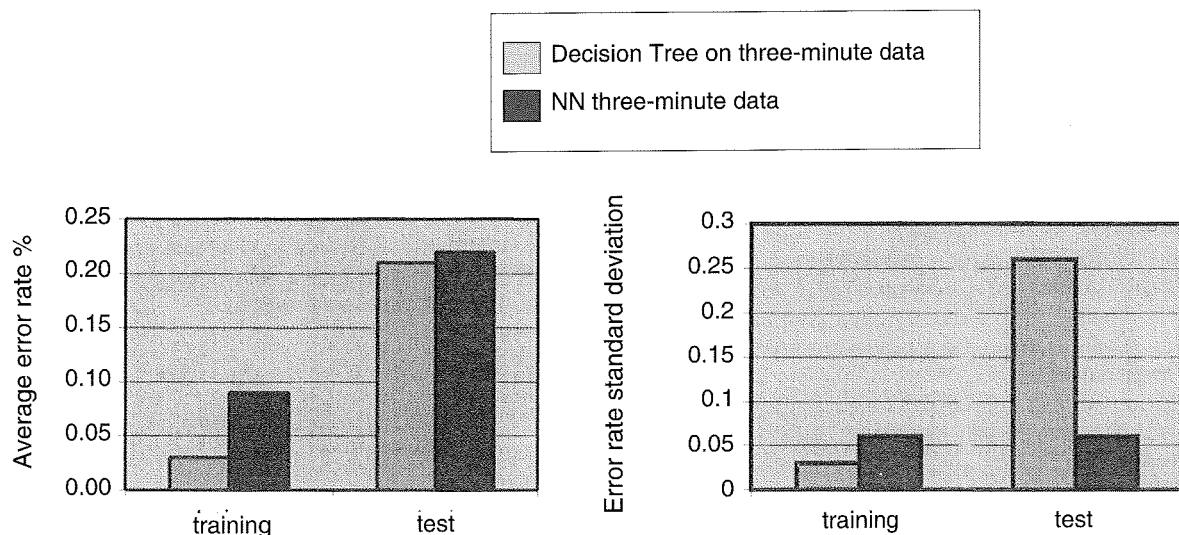


Figure 9. Error rate statistics comparison between Neural Network and Decision Tree classification

Table II
Neural network model performance calculation

Decision Tree—three minute		Training			Test #1			Test #2			Total test		
		Class	Correct	Total	Error rate %	Correct	Total	Error rate %	Correct	Total	Error rate %	Correct	Total
	Other	411	438	0.06	42	51	0.18	47	52	0.10	89	103	0.14
	Eng1	65	66	0.02	5	8	0.38	2	6	0.67	7	14	0.50
	Eng2	70	70	0.00	4	4	0.00	6	6	0.00	10	10	0.00
Average % (error rate)				0.03			0.18			0.25			0.21
Standard deviation (error rate)				0.03			0.19			0.36			0.26
N—three minute		Training			Test #1			Test #2			Total test		
	Class	Correct	Total	Error rate %	Correct	Total	Error rate %	Correct	Total	Error rate %	Correct	Total	Error rate %
	Other	373	438	0.15	41	51	0.20	44	52	0.15	85	103	0.17
	Eng1	60	66	0.09	4	8	0.50	6	6	0.00	10	14	0.29
	Eng2	68	70	0.03	3	4	0.25	5	6	0.17	8	10	0.20
Average % (error rate)				0.09			0.32			0.11			0.22
Standard deviation (error rate)				0.06			0.16			0.09			0.06

Table III
VIMS parameters with high sensitivity

Parameter	Description	Sensitivity
DIFF_FLTR_SW_AVG	The status of the differential axle oil filter (plugged or OK)	0.5
ENG_LOAD_MIN	This is calculated by the engine ECM (Electrical Control Module) after considering the engine speed, throttle switch position, throttle position, boost pressure, and atmospheric pressure and is shown as a per cent of a full load	0.5
ENG_SPD_REGR_SLOPE	The actual rotational speed of the crankshaft	0.5
GROUND_SPD_MIN	The speed of the machine relative to the ground	0.5
MACHINE_RACK_MAX	This is calculated from the four machine suspension cylinder pressures. VIMS takes the sum of the two diagonal suspension cylinder pressure minus the sum of the two other diagonal suspension cylinders pressures	0.5
PAYOUT_MIN	The payload is calculated by VIMS based on pressure of the four suspension cylinders	
RETARDER_AVERAGE	The status of the retarder system (On or Off)	0.5
RT_R_BRK_TEMP_MIN	The cooling oil temperature from the right rear brake	0.5
TC_OUT_TEMP_MAX	The oil temperature on the outlet of the torque converter	0.5

recorded is 17% and 29% respectively, a significant improvement over the results of Decision Tree Classification modeling.

Neural Network Classification needs more time for model training. Furthermore, the results of model runs are not as easy to interpret as that of Decision Tree Classification based models. There is also a risk that that Neural Network Classification model training runs may not converge in the absence of large databases.

All conclusions drawn here are valid for the investigated dataset of VIMS. The model prediction accuracy, while improved, remains relatively low. To improve the accuracy of the model, more VIMS data are needed, including both more values for each parameter and values of more VIMS parameters. The model prediction accuracy, while improved, remains relatively low.

Acknowledgements

Financial support for the research presented in this paper was provided by Caterpillar, Inc. of Peoria, Illinois. The IBM (International Business Machines) provided a copy of the IBM Intelligent Miner at no charge to the researchers under the IBM Scholars program.

References

1. GOLOSINSKI, T.S. Data mining uses in mining. *Proceedings, Computer Applications in the Minerals Industries* (APCOM), Beijing, China, 2001, pp. 763–766.
2. GOLOSINSKI, T.S. and HU, H. Data mining of mine

- equipment databases. *Proceedings of the Artificial Neural Networks in Engineering Conference* (ANNIE 2001), St. Louis, Missouri, U.S.A. 2001. pp. 387–396.
- 3. GOLOSINSKI, T.S. and HU, H. Modeling condition and performance of mining equipment. Proceedings, Mine Planning and Equipment Selection 2002, Czech (in print). 2002.
 - 4. HU, H. and GOLOSINSKI, T.S. Failure pattern recognition of a mining truck with a decision tree algorithm. Mineral Resources Engineering (in print). 2002.
 - 5. Caterpillar, Inc. Vital information management system (VIMS): system operation testing and adjusting. Company publication. 1999. pp. 10–13.
- 6. FAUSETT, V.L. Fundamentals of neural networks—architectures, algorithms, and applications' Florida Institute of Technology. 1994. 289 pp.
 - 7. IBM (International Business Machines Corporation). Manual: using the intelligent miner for data. Company publication. 2000. 298 pp.
 - 8. CLAIR, ST. D.C. CS404: Data mining and knowledge discovery. Lecture Note. Computer Science Department of University, Missouri-Rolla. 2001.

