World Academy of Science, Engineering and Technology
International Journal of Economics and Management Engineering
Vol:8, No:8, 2014

A Study of Classification Models to Predict Drill-Bit Breakage Using Degradation Signals

Bharatendra Rai

***Abstract*—**Cutting tools are widely used in manufacturing processes and drilling is the most commonly used machining process. Although drill-bits used in drilling may not be expensive, their breakage can cause damage to expensive work piece being drilled and at the same time has major impact on productivity. Predicting drill-bit breakage, therefore, is important in reducing cost and improving productivity. This study uses twenty features extracted from two degradation signals viz., thrust force and torque. The methodology used involves developing and comparing decision tree, random forest, and multinomial logistic regression models for classifying and predicting drill-bit breakage using degradation signals.

***Keywords*—**Degradation signal, drill-bit breakage, random forest, multinomial logistic regression.

I. Introduction

AMONG various machining processes used in manufacturing companies, drilling is a widely using cutting process. A reliable and non-intrusive method for predicting drill-bit breakage is highly desirable to reduce cost and improve productivity. Researchers have proposed several prediction models based on non-intrusive degradation signals extracted from the drilling process using appropriate sensors. These models are helpful for making suitable decisions about whether or not to continue drilling considering the severity level of the tool-breakage in different applications. When impact of tool-breakage is not severe on overall cost and productivity, the decision about the tool replacement can take into account extraction of maximum life from the cutting tool used. However, in situations where the impact of tool-breakage is severe on the overall cost and productivity, tool-life utilization may have a lower importance.

Reference [1] applied Mahalanobis-Taguchi System to monitor and predict drill-bit breakage using degradation signals. Reference [2] found monitoring of the kurtosis value obtained from the traverse and thrust vibrations to be effective for online detection of the drill-bit breakage. Reference [3] applied multi-layer perceptron neural network for tool-state classification using online data on the cutting forces and vibration, and reported achieving approximately 90% accuracy in tool-state classification. References [4]-[9] also developed models based on neural networks for online tool-condition monitoring. References [10], [11] proposed use of hidden Markov models for tool wear condition monitoring in drilling operations. Reference [12] provides a summary of various signal analysis methods for tool condition monitoring used in literature and reports statistical parameters to be one of the most frequently used with varying level of success.

This paper provides a study of the three methods viz., decision trees, random forest, and multinomial logistic regression for classifying degradation signals and predicting drill-bit breakage.

II. Two Drill-Bit Degradation Signals and Twenty Extracted Features

A. Data for the Study

The data used for the study is based on two degradation signals viz., thrust force and torque that were collected using an experimental setup consisting of a HAAS VF-1 CNC milling machine, a workstation with LabVIEW software for signal processing, a Kistler 9257B piezo-dynamometer and a National Instruments PCI-MIO-16XE-10 card for data acquisition [13]. Data on twelve drill-bits were collected at 250 Hz until their breakage. The recorded data consist of 380 to 460 data points per hole that were condensed to 24 root-mean-square (RMS) values per hole. The degradation signals in the form of RMS values provide a means for non-intrusive online tool condition monitoring with the help of suitably extracted features and aid in timely tool replacement before its breakage.

B. Features Extracted from the Degradation Signals

The RMS values of the two degradation signals, thrust force and torque are numerically summarized into ten features each viz., maximum, average, standard deviation, coefficient of variation, third quartile, kurtosis, skewness, mean-1 (25% low values trimmed), mean-2 (50% low values trimmed) and mean-3 (75% low values trimmed). The twenty features considered in this study are shown in Table I.

The dataset consists of degradation signals obtained from holes drilled using a total of twelve drill-bits. In all, data consists of 161 holes drilled by the twelve drill-bits under study. The number of holes successfully drilled using each drill-bit before their breakage are shown in Fig. 1.

B. Rai is with the Charlton College of Business, University of Massachusetts-Dartmouth, North Dartmouth, MA 02747 USA (phone: 508-910-6434; e-mail: brai@umassd.edu).

TABLE I
TWENTY FEATURES FROM THE DRILL-BIT DEGRADATION SIGNALS

No.	Degradation Signal	Feature	Notation
1	thrust force	maximum	A
2	thrust force	average	B
3	thrust force	standard deviation	C
4	thrust force	coefficient of variation	D
5	thrust force	third quartile	E
6	thrust force	kurtosis	F
7	thrust force	skewness	G
8	thrust force	mean-1 (25% low trim)	H
9	thrust force	mean-2 (50% low trim)	J
10	thrust force	mean-3 (75% low trim)	K
11	torque	maximum	L
12	torque	average	M
13	torque	standard deviation	N
14	torque	coefficient of variation	O
15	torque	third quartile	P
16	torque	kurtosis	Q
17	torque	skewness	R
18	torque	mean-1 (25% low trim)	S
19	torque	mean-2 (50% low trim)	T
20	torque	mean-3 (75% low trim)	U

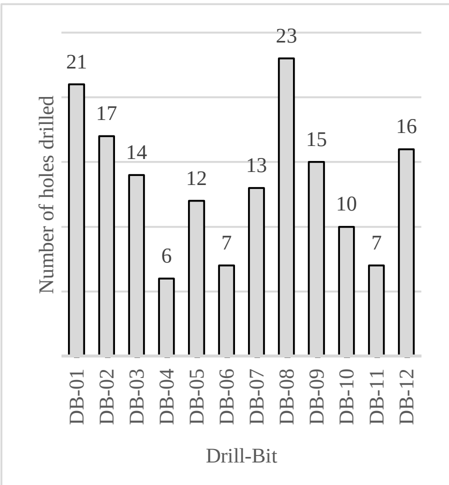

Fig. 1 Number of holes drilled per drill-bit

C. Coding of the Response

The response used for analyzing the data is that grouped into three categories viz., green, yellow, and red. A prediction involving response code red would indicate that the drill-bit is very likely to break and should be replaced immediately. Similarly, a prediction involving response code yellow would indicate that the drill-bit may break after drilling of two or three holes. And a prediction involving response code green would indicate that the drill-bit can continue to be used for drilling holes. For this study, response for data representing the last hole drilled before drill-bit breakage is coded as red. Data representing two drilled holes prior to the last hole drilled before drill-bit breakage are represented by response code yellow. Reponses for all remaining data are coded as green.

III. THREE CLASSIFICATION AND PREDICTION MODELS

In this study, three classification and prediction models are developed viz., decision trees, random forest, and multinomial logistic regression. The models are developed using R software. For developing the models, data are divided into training dataset (75%) and testing dataset (25%).Out of data on 161 drilled holes, 126 are used in the training dataset and remaining 35 are used in the testing dataset. Number of responses with green, yellow, and red in the training dataset are 98, 20, and 8 respectively. Similarly, number of responses with green, yellow, and red in the testing dataset are 27, 4, and 4 respectively. Performances of the three models are compared by calculating classification error percentage.

A. Decision Tree

A decision tree is constructed using *party* package in R software [14]. Level of significance for a variable to be added is specified at 0.05 and a minimum sample of five is specified for a split to take place. The decision tree obtained based on the training dataset is shown in Fig. 2.

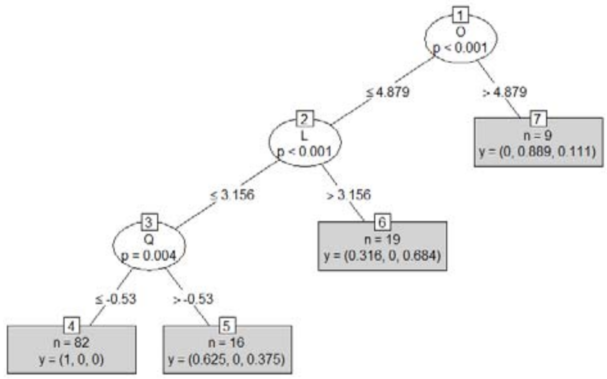

Fig. 2 Decision Tree

The decision tree includes features coded as ‘O’, ‘L’, and ‘Q’ representing coefficient of variation for torque, maximum torque, and kurtosis torque respectively. One of the advantages of using decision trees lies in the ease of interpretation of the if-then decision rules. For example, if coefficient of variation for torque is more than 4.879, then response code yellow is predicted with probability 0.889. This indicates that high amount of caution is needed to continue drilling holes as the drill-bit may not last more than 2 or 3 holes. Similarly, if coefficient of variation for torque is less than 4.879, and maximum torque is more than 3.156, then predicted response code is red with probability 0.684. This indicates that the drill-bit should be replaced immediately as the drill-bit is likely to break.

To assess the performance of the decision tree, classification errors are calculated for both training and testing datasets by comparing actual responses with the predicted responses. Table II provides the classification error percentages.

TABLE II
CLASSIFICATION ERROR RATES FOR DECISION TREE

Dataset	Predicted Response Code	Actual			Classification Error %
		Green	Yellow	Red	
training	green	92	6	0	6.1%
	yellow	6	13	0	31.6%
	red	0	1	8	11.1%
testing	green	23	1	0	4.2%
	yellow	3	3	0	50.0%
	red	1	0	4	20.0%

It is observed that the classification errors are lowest for green response codes and highest for yellow response codes. Lower classification error for green suggests that the decision tree model would support decisions to continue drilling holes with high confidence. For yellow response code, it should be noted that the misclassification for both training and testing datasets mostly occurs in predicting yellow response code when in reality the response should be green. In other words, when model predicts yellow, the actual response code is unlikely to be red. This could help in fine-tuning the decision making process for situations that have different levels of tool-breakage severity.

B. Random Forest

Random forests are extension of the idea of decision trees [15]. Unlike a single tree that is constructed in decision tree, multiple decision trees are constructed leading to a random forest. The output from all trees is combined to obtain a better model than what could be obtained from a single tree. The model is developed using *randomForest* package available from R software. A random forest model was constructed based on the training dataset by specifying number of trees grown to 5000 and number of predictors sampled for splitting at each node to 5. To assess the performance of the random forest, classification errors are calculated for both training and testing datasets by comparing actual responses with the predicted responses. Table III summarizes the classification error percentages for the random forest.

TABLE III
CLASSIFICATION ERROR RATES FOR RANDOM FOREST

Dataset	Predicted Response Code	Actual			Classification Error %
		Green	Yellow	Red	
training	green	92	8	0	8.0%
	yellow	6	9	2	47.1%
	red	0	3	6	33.3%
testing	green	24	0	0	0.0%
	yellow	3	4	0	42.9%
	red	0	0	4	0.0%

The classification errors for response code yellow are the highest for both training and testing datasets. Similarly, classification errors for response code green are the lowest for both training and testing datasets.

C. Multinomial Logistic Regression

In this study the response has three ordinal classes viz., green, yellow, and red indicating worsening levels of degradation signal. This study, with three ordinal classes, is well suited for a multinomial logistic regression model that is constructed using *multinom* function from package *nnet* in R software. Before running the model, response code red is specified as the baseline. The final multinomial logistic regression model is based on features coded as *A*, *K*, *L*, *O*, and *R*. This selection of features is based on two-tailed z-test with 0.05 level of significance. The three noncumulative probabilities of class membership are used for predicting the response code and subsequently arriving at classification errors. Table IV provides the classification error percentages for the multinomial logistic regression model.

TABLE IV
CLASSIFICATION ERROR RATES FOR MULTINOMIAL LOGISTIC REGRESSION

Dataset	Predicted Response Code	Actual			Classification Error %
		Green	Yellow	Red	
training	green	95	5	0	5.0%
	yellow	3	15	0	16.7%
	red	0	0	8	0.0%
testing	green	24	0	0	0.0%
	yellow	2	4	0	33.3%
	red	1	0	4	20.0%

The classification errors for response code yellow are the highest for both training and testing datasets which is similar to what was observed for earlier two models. It can also be observed that the response code yellow is often misclassified as green. However, it is interesting to note that there are no situations where a red response code is misclassified as green or yellow for both training and testing datasets.

IV. DISCUSSION OF THE THREE MODELS

Decision tree, random forest, and multinomial logistic regression models developed using the training dataset yielded overall classification errors of 10.3%, 15.1%, and 6.3% respectively. Similarly, overall classification errors for the three models using the testing dataset were 14.3%, 8.6%, and 8.6% respectively. Random forest and multinomial logistic regression models show lower overall classification errors for testing dataset compared to the decision tree model. However, all three models show high classification errors for response code yellow. The reason for such high error could lie in the way responses are coded. Number of data points before response code red that can be labeled as yellow can have a significant impact on the classification error. This choice can be influenced by the severity of the impact of a tool breakage.

Three models that were used for classification and prediction of drill-bit performance have certain advantages and weaknesses. Decision tree provides ease of interpretation of the model and prediction of results using if-then rules. Importance of the predictor variables can be judged by their higher position in the tree with the most important predictor variable placed at the top of the tree. For example, for the data used in this study coefficient of variation for torque is found to be the most important predictor variable and is shown in a node at the top of the decision tree. Decision tree models are nonlinear and non-parametric. Use of specified levels of significance and minimum sample size for the split to take place to grow a tree, help avoid over fitting that causes the tree to grow too large without improving predictive performance of the model. A weakness of decision tree approach is that it requires large sample size.

An extension of decision tree modeling is random forest. Random forest models also help to obtain variable importance. However, a large number of trees are needed to obtain a stable estimate of variable importance. In this study we have used 5000 trees where each tree is a based on bootstrap sample from the training dataset.

Compared to decision tree and random forest models, multinomial logistic regression models are more challenging to interpret. In addition, when dealing with several continuous predictors, correlation between variables can lead to an inappropriate model for application. In situations where number of classes are large, usually more than five, the response may be treated as continuous and multiple linear regression model may be used.

V. Conclusions

In this study classification and predictive models were developed to assist in making decision about whether or not to replace a tool in a drilling process. Twenty features were extracted from two degradation signals thrust force and torque. Data were divided into training (75%) and testing (25%) datasets. Three classification and prediction models viz., decision tree, random forest, and multinomial logistic regression were developed using the training dataset yielding overall classification errors 10.3%, 15.1%, and 6.3% respectively. Overall classification errors for the three models using testing dataset were 14.3%, 8.6%, and 8.6% respectively. Out of the three models studies, multinomial logistic regression yielded lower classification errors for both training and testing datasets.

References

[1] B. K. Rai, R. Chinnam, and N. Singh, "Prediction of drill-bit breakage from degradation signals using Mahalanobis-Taguchi system analysis,"*International Journal of Industrial and Systems Engineering*, Vol. 3(2), pp.134-148, 2008.

[2] T. I. El-Wardany, D. Gao, and M. A. Elbestawi, "Tool condition monitoring in drilling using vibration signature analysis,"*International Journal of Machine Tools Manufacturing*, Vol. 36, pp.687-711, 1996.

[3] D. E. Dimla, and P. M. Lister, "Online metal cutting tool condition monitoring. II: tool-state classification using multi-layer perceptron neural networks," *International Journal of Machine Tools and Manufacture Design, Research and Applications*, Vol. 40, pp.769-781, 2000.

[4] L. Fu,S. Ling, and C. Tseng, "On-line breakage monitoring of small drills with input impedance of driving motor,"*Mechanical Systems and Signal Processing*, Vol. 21, pp.457-465, 2007.

[5] N. Ghosh, Y. B. Ravi,A. Patra, S. Mukhopadhyay, S. Paul, A. R. Mohanty, and A. B. Chattopadhyay, "Estimation of tool wear during CNC milling using neural network-based sensor fusion,"*Mechanical Systems and Signal Processing*, Vol. 21, pp.466-479, 2007.

[6] J. Sun, G. S. Hong, Y. S. Wong, M. Rahman, and Z. G. Wang, "Effective training data selection in tool condition monitoring system," *International Journal of Machine Tools and Manufacture*, Vol. 46, pp.218-224, 2006.

[7] U. Zuperl, and F. Cus, "Tool cutting force modeling in ball-end milling using multilevel perceptron," *Journal of Materials Processing Technology*, Vol. 153, pp.268-275, 2004.

[8] I. Abu-Mahfouz, "Drilling wear detection and classification using vibration signals and artificial neural network," *International Journal of Machine Tools and Manufacture*, Vol. 43, pp.707-720, 2003.

[9] R. B. Chinnam, and P. Mohan, "Online Reliability Estimation of Physical Systems Using Neural Networks and Wavelets,"*Smart Engineering System Design*, Vol. 4, pp.253-264, 2002.

[10] H. M. Ertunc, K. A. Loparo, and H. Ocak, "Tool wear condition monitoring in drilling operations using hidden Markov models," *International Journal of Machine Tools and Manufacture*, Vol. 41, pp.1363-1384, 2001.

[11] H. M. Ertunc, and C. Oysu, "Drill wear monitoring using cutting force signals," *Mechatronics*, Vol. 14, pp.533-548, 2004.

[12] E. Jantunen, "A summary of methods applied to tool condition monitoring in drilling," *International Journal of Machine Tools and Manufacture*, Vol. 42, pp.997-1010, 2002.

[13] B. K. Rai, "Classification, Feature Selection and Prediction with Neural-Network Taguchi System," *International Journal of Industrial and Systems Engineering*, Vol. 4, No. 6, 645-664, 2009.

[14] D. T. Larose, and C. D. Larose, *Discovering Knowledge in Data: An Introduction to Data Mining*. Hoboken, New Jersey: John Wiley & Sons, 2014, ch 8.

[15] G. Shmueli, N. R. Patel, and P. C. Bruce, *Data Mining for Business Intelligence: Concepts, Techniques, and Applications*. Hoboken, New Jersey: John Wiley & Sons, 2010, ch 9.

Bharatendra Rai is PhD in industrial engineering from Wayne State University, Detroit in 2004. He has master of technology degree in quality, reliability, and operations research from Indian Statistical Institute, India in 1993, and another master's degree in statistics from Meerut University, India in 1991.

He is currently chair and associate professor in the department of decision and information sciences at Charlton College of Business, University of Massachusetts-Dartmouth. He has earlier worked as a quality and reliability engineer at Ford Motor Company in Detroit. He has co-authored a book titled 'Reliability Analysis and prediction from Warranty Data: Issues, Strategies, and Methods' (New York, NY: CRC Press Taylor & Francis Company, 2009). He is author of over 30 journal and conference articles including two publications in IEEE Transactions on Reliability.

Dr. Rai is a certified six-sigma black belt from American Society for Quality. Dr. Rai is also awarded certificates as ISO 9000 and ISO 14000 lead assessor by British Standards Institute and Marsden Environmental International respectively.