iJRASET
International Journal For Research in
Applied Science and Engineering Technology

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 8 **Issue:** VII **Month of publication:** July 2020

DOI: https://doi.org/10.22214/ijraset.2020.30357

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.429
Volume 8 Issue VII July 2020- Available at www.ijraset.com

Rate of Penetration Prediction using K-means and Ensembles, a Machine Learning Approach

Hitesh Hinduja[1], Arishma Datta[2]

***Abstract:** Rate of Penetration (ROP) prediction is an important aspect of drilling in the Oil & Gas Industry. Several studies have been carried out to predict ROP. Primarily, Artificial Neural Networks (ANN) has been used. In this paper, the objective is to explore a new approach to predict ROP using K-means and Ensemble of Gradient Boosting Model (GBM) technique. Nine input parameters are used for ROP prediction- True vertical depth, weight on bit, standpipe pressure, flow-rate, torque, equivalent circulating density and RPM. The model is evaluated on the basis of accuracy, R^2 and Root-mean square error (RMSE)*
***Keywords:** Rate of Penetration, Prediction, Gradient Boosting Machine (GBM), K-means, Ensemble, ANN, Random Forest*

I. INTRODUCTION

These days in the Oil & Gas industry, cost efficiency is an important aspect. Prediction of drilling parameters and optimization cost has been extensively studied and researched upon.

The primary aim of these studies is to maximize the performance and decrease the probability of encountering problems and thus reducing the non-productive time during drilling. In most cases, the cost of drilling is reduced by increasing its drilling speed. This is mainly achieved by maximizing the Rate of Penetration(ROP).

ROP is dependent on many drilling parameters hence the key task is to derive a relation between the optimum drilling parameters that will maximize ROP thus minimizing cost.

Therefore, this research has been a focus area for many researchers and major oil & gas companies. This paper presents a technique which predicts the Rate of Penetration of drill bit with high accuracy. Various models have been tried and tested and finally the ensemble of GBM models gives the best results.

The scope of this paper is to present a technique apart from Artificial Neural Networks(ANN) and thus avoiding black-box methods to predict the ROP.

Prior to this research paper, the GBM model has not been used in ROP prediction in the oil and gas sector.. Moreover, this paper also provides effective choices of hyper-tuning parameters in the GBM model for better prediction.

II. METHODOLOGY

The methodology followed is to implement each regression model with different parameters and evaluate the highest accuracy model. Regression is widely used in ROP prediction and therefore it is important to determine whether it is justified to change the technique and avoid black-box methodology

III.IMPLEMENTATION

A. Input / Output Data

The data collected had 9 input parameters for each of 4 wells in one cluster (Oil field). Among these parameters, multicollinearity was identified and finally 7 input parameters were used viz: - True vertical depth(TVD), Standpipe pressure(SPP), Equivalent circulating density(ECD), Mud flow rate, Weight on Bit(WOB), Rotations per minute (RPM), Torque. The dataset was divided into a training set, a cross validation set and a test set. The accuracy was evaluated on the test set

B. Background

For every well, there are different lithologies (layers) on the inside. This helps companies identify the position of the drill-bit as it moves from one layer to another layer. Based on the properties of each layer, the speed of the drill bit reduces or increases. For example:- As the drill bit goes to lower depths, the speed of the drill bit increases due to high force inside the well. Identifying this depth is very crucial as the majority of drill bit breakdowns happen here hence it is important to predict the speed and adjust it accordingly. Below is the image which shows the lithology inside the well

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.429
Volume 8 Issue VII July 2020- Available at www.ijraset.com

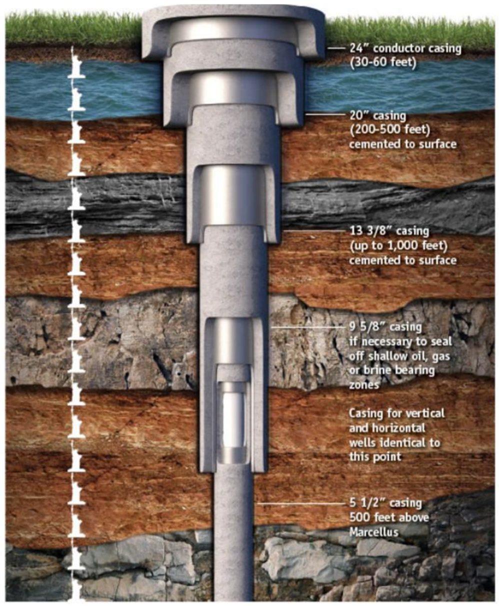

Fig. 1 Different oil well layers. Credits:- Google Images

C. Procedure

The process started by splitting the data into 3 parts viz: - 70% training set, 20% cross validation set and 10% test set. The key problem encountered in the data set was the unavailability of data on lithology(layers) inside the well which determines the exact change in Rate of penetration as the depth increases. Each well has a different number of lithologies with each lithology having its own properties and the ROP varies significantly when the drill goes from one layer of lithology to the other layer of lithology. The model accuracy and the statistics obtained before establishing lithology is as shown in the table below: -

ALGORITHM	R2	RMSE
Linear Regression	0.18	4.77
Support Vector Regression	0.21(without tuning), 0.27(with tuning)	4.64, 4.37
Random Forest	0.426	4.03
Lasso	0.4	4.31
Gradient Boosting Machine (GBM)	0.39	4.08
XGBoost Regressor	0.44	4.03
Ensemble (3 GBMs)	0.48	3.92
Neural Networks	0.47 (2 hidden layers)	3.96
Glm, log transformations	0.23	4.59

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.429
Volume 8 Issue VII July 2020- Available at www.ijraset.com

The major reason for the model failure is non-linearity in the data primarily due to missing lithologies in the data set. The machine learning models fail to predict the ROP when the drill moves from one layer to the other layer and thus leading to low accuracy. To overcome this drawback and to understand the layers inside a well, K-means clustering algorithm was implemented. By using k-means clustering algorithm on the full data set with all the 7 parameters, layers were established. At first, manually 5 clusters were taken and using elbow point method, 8 clusters emerged to be significant. Finally, the dataset was divided into 8 clusters with each row in the dataset attributed to a specific layer inside the well based on the properties of the cluster. Clustering algorithms especially K-means expect data to be scaled hence before the clustering process a major task was to center scale the data. Once the clustering was completed, the dataset was rescaled to its original form with an additional variable of layer number in the dataset which was the cluster number. This variable was then converted into a categorical variable using one-hot encoding technique. The graph for optimal number of clusters is shown below: -

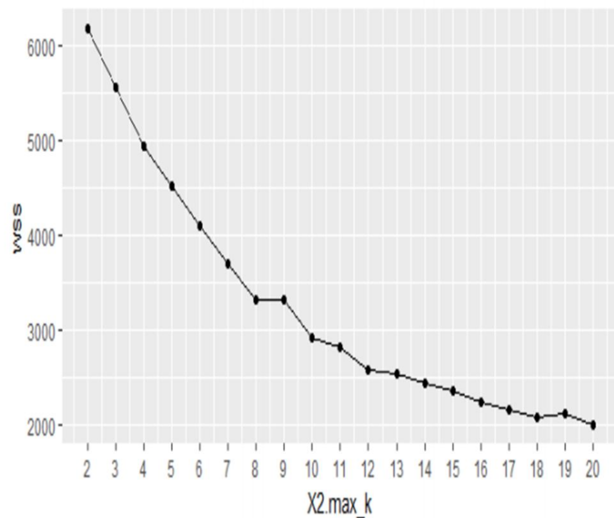

Fig. 2 An elbow curve for identifying ideal number of clusters (lithologies) inside a well

After k-means clustering, same algorithms were again performed and the results obtained are as below in the table:-

ALGORITHM	R2	RMSE
Linear Regression	0.29	4.4
Support vector regression	0.55(without tuning), 0.51(with tuning)	3.4990, 3.67
Lasso	0.4	4.0
Random Forest	0.81	2.05
Gradient Boosting Machine (GBM)	0.84	1.72886
Ensemble (3 GBMs)	0.87	1.699

Thus, the above table clearly states that the Gradient Boosting Machine model gives the best accuracy and lowest RMSE. To further improve the model's results, an ensemble of 3 GBM models was used by varying the learning rate and other hyperparameters. Using ensemble, the model successfully predicted the ROP with 87% accuracy and 1.699 RMSE.

IV. CONCLUSION

This paper has shown implementation of significant machine learning models for ROP Prediction. It also avoids the black-box methods such as neural networks and thus accurately predicts the Rate of penetration of a drill bit which optimizes cost and reduces the non-productive time in the Oil & Gas industry.

REFERENCES

[1] Abtahi A., Butt S., and Molgaard, J., and Arvani F., 2011. "Bit Wear Analysis and Optimization for Vibration Assisted Rotary Drilling) VARD (using Impregnated Diamond Bits", Memorial University of Newfoundland, St. John's, NL, Canada.

[2] Adrian G. Bors 1996. "Introduction to the Radial Basis Function (RBF) Networks", University of York. Available: wwwsers.cs.york.ac.uk/adrian/Papers/ Others/OSEE01.pdf>

[3] Awasthi, Ankur 2008. "Intelligent oilfield operations with application to drilling and production of hydrocarbon wells", University of Houston, 2008, 370p

[4] Bataee M., & Mohseni, S., 2011. "Application of artificial intelligent systems in ROP optimization: A case study in Shadegan oil field". SPE Middle East Unconventional Gas Conference and Exhibition 2011.

[5] Bourgoyne A.T. Jr., Young F.S., 1974 "A Multiple Regression Approach to Optimal Drilling and Abnormal Pressure Detection", SPE 4238, August 1974.Available: ttp://www.onepetro.org/mslib/servlet/ onepetropreview?id=00004238>.

[6] Eren Tuna, 2010. "Real-TimeOptimization Of Drilling Parameters During Drilling Operations Dissertation" PhD Technical University of the Middle East, Turkey. "Getting Started with Matlab 7" www.mathworks.com (2007).

[7] Akgun, F., Drilling Rate at the Technical Limit, International Journal of Petroleum Science and Technology, vol. 1, pp. 99-118, 2007.

[8] Moran David, Ibrahim Hani, Purwanto Arifin, Smith International and Jerry Osmond, 2010. "Sophisticated ROP Prediction Technologies Based on Neural Networks Delivers Accurate Drill Time Result", SPE Asia Pacific Drilling Technology Conference and Exhibition, 2010.

[9] Paiaman A. M., Al-Askari M. K. Ghassem, Salmani B, AlAnazi, B. D. and Masihi M. 2009. "Effect of Drilling Fluid Properties on Rate of Penetration", NAFTA 60 (3) 129-134 (2009).

[10] Rampersad, P.R., Hareland, G.; and Boonyapaluk, P., 1994. "Drilling Optimization Using Drilling Data and Available Technology", SPE Latin America/Caribbean Petroleum Engineering Conference, 27-29 April 1994, Buenos Aires, Argentina.

[11] Gidh Yashodhan, Ibrahim Hani, Purwanto Arifin, and Bits smith, 2011. "Real-time drilling parameter optimization system increases ROP by predicting/managing bit wear". SPE Digital Energy conference and Exhibition, 2011.