Journal of Mines, Metals and Fuels, 70(9) : 476-483; 2022. DOI: 10.18311/jmmf/2022/32098

Print ISSN : 0022-2755

Journal of Mines, Metals and Fuels

Contents available at: www.informaticsjournals.com/index.php/jmmf

Breakdown and Productivity Prediction of Dragline using Machine Learning Algorithms

Vikram Seervi[1*], Nilesh Pratap Singh[1], Nawal Kishore[2] and Rajeev Verma[3]

[1]*PhD Research Scholar, Department of Mining Engineering, IIT-BHU, Varanasi - 221005, Uttar Pradesh, India; Vikramseervi.rs.min17@itbhu.ac.in*

[2]*Assistant Professor, Department of Mining Engineering, IIT-BHU, Varanasi - 221005, Uttar Pradesh, India*

[3]*Assistant Professor, Department of Mining Engineering, College of Technology and Engineering, Udaipur - 313001, Rajasthan, India*

Abstract

Dragline operations play a major role in the overall production of coal in open cast mining. Hence, it becomes necessary to maximize the working hours and minimize the idle and breakdown hours as it affects the overall production of a mine. There is also a shortage of skilled labour for dragline operations and combined with the time-to-time breakdown of dragline, it results in a production deficit. In this study, extensive research is carried out using machine learning algorithms on data obtained from one of the largest opencast mines in Singrauli. The data consists of the parameters that were maintained by the staff on a regular basis, and the algorithm tried to learn the underlying patterns between the independent and dependent variables and find the correlation between the parameters that have a significant impact on productivity and breakdown, which were the dependent variables. The results obtained from the algorithms are encouraging and, with certain improvements in data collection procedures, can improve the prediction accuracy to an effective level. An increase in the frequency of data collection and expanding the data recording using sensors to the electrical and mechanical parameters along with the specific type of failure in the dragline machine will further improve the accuracy of the model and can provide beforehand information so that the machine could be handed over to maintenance department for the change of faulty parts and necessary precautions that can be taken to prevent the breakdown which will result in an overall reduction of idle and breakdown hours and increase in overall production.

Keywords: *Artificial Neural Network, Breakdown Hours, Dragline, Machine Learning, Productivity*

1.0 Introduction

1.1 Dragline Overview and Operating Mechanism

Dragline is made up of a major frame structure that spins around its base. A boom that covers the excavation site is attached to the main structure. A bucket suspended beneath the boom is connected to a hoist and drag ropes. While the drag ropes pull the bucket across the material being moved, the hoist ropes elevate and lower the bucket. The bucket is open where the drag ropes are attached at the leading edge, and as it is pulled across the ground, it fills up. The upper frame rotates as the bucket is hoisted once it has been filled, allowing the material to be discharged away from the diggings.

*Author for correspondence

Electricity powers the large draglines used in the coal industry (Chaoji & Dey, 2000). They are connected to the high voltage grid of the state's power system due to their significant power usage; their supply voltage is 66 kV. Internal transformers then lower this power supply to power the electrical driving systems. Each motion circuit's drive system consists of a generator and a motor. A huge dragline has four motion circuits: hoist, drag, rotation (swing), and propel. The hoist and drag motions are powered by a system of geared machinery inside the main house structure, which also houses the hoist and drag ropes in massive winch drums (Dayawansa *et al.*, 2008). The geared machinery is directly linked to electric motors. The operator, who is seated in a cab above the bucket, controls whether the drum pays the ropes in the forward or reverse motion. The digging action is created by this technique. The operator controls the hoist and drag motions with hand levers; the swing motion of the upper frame to the bottom is controlled by foot pedals (Rzhevsky, 1987). Electric signals are sent from the levers and pedals to the main electric drive control panels. Most draglines in use today feature Programmable Logic Control (PLC) units to track and translate operator signals into commands for the drive units. The PLC receives additional inputs from various motion transducers to help control the motions. However, as businesses modernise, fewer draglines are still using outdated techniques to control the drives instead of PLC. Therefore, a dragline is a piece of equipment made up of numerous mechanical and electrical components which must be integrated to perform cycles of digging and dumping. To increase machine reliability, new system control technologies are being deployed. Draglines are frequently used, which creates a heavy workload on its mechanical and electrical

Figure 1. Dragline working in open cast mine.

primary components. The maintenance strategy includes a milestone where substantial component repairing and replacement are necessary to sustain reliability in service. This work is structured so that the biggest amount of maintenance may be done in the shortest period.

The incidence of breakdown malfunctions of draglines immediately affects production and other simultaneous operations (Arunraj & Maiti, 2007), and hence it is the objective of mine management staff to optimize the use of draglines and minimize the breakdown and idle time to ramp up the production. Unwanted breakdown of the machine significantly impacts the machine performance and efficiency and shoots up the maintenance (Vidyasagar and Kishorilal, 2016) cost of mining machinery accounts for an increase of around 50% cost in the overall operational cost. The incidence of failure due to inadequate maintenance actions leads to downtime damages, and it lowers the availability, reliability, and efficiency of the operation. To bring down the cost and increase the operational efficiency of machine, forehand failure prediction becomes more important.

1.2 Prediction of Dragline Breakdown

Failure prediction has been an area of extensive research and analysis, and there are certain evaluations and analyses done to study them. One of them makes use of potential Failure, Impacts, and Criticality Assessment (FMECA) to pinpoint the dragline system's crucial failure points. It also uses FMECA to systematically assess the likelihood of failures occurring and the detectability of possible failure modes to fully comprehend the reasons behind failures and how they affect the system's performance. A criticality study of dragline parts and an evaluation of RPN were used to highlight the other risk estimating method, which uses the Risk Priority Number (RPN), which considers four aspects, primarily failure incidence, production loss, performance degradation, and detectability (Sahu & Palei, 2000). The study calculates RPN utilizing the status of existence, damages, defect, and detectability of malfunction on dragline failure data. RPN varies on a scale of 1 to 250. The higher RPN suggests the more critical component of the dragline system is at risk.

$\text{RPN} = \alpha \times \beta \times \boldsymbol{\mu} \times \Phi,$

α denotes each component's risk and the possibility that it may fail

β represents a reduction in production measured in breakdown time losses.
μ indicates a failure effect that reduces the system's performance in terms of the occurrence of a component fault.
Φ Indicates detection performance in terms of root cause and failure identification.

1.3 Prediction of Productivity (m^3/hr)

In recent years theoretical analysis has been used for the calculation of annual production, though good for approximation for the annual production (Seervi *et al.*, 2022) analysis, due to the specific nature of a particular mine with its own conditions and its features, a general theoretical model may not encompass the intricacies of a particular mine. Artificial intelligence, with its robust tools, can learn features in relation to productivity (output variable), especially with the help of historical data; machine learning models can solve the above problem. (Rai *et al.*, 2022). Using machine learning algorithms to predict cast blasting performance in surface mining predicted the performance of dragline in production (Rai *et al.* 2011), which involved removal of overburden which has a significant role in overall production. Hyper-parameters tuning of Random Forest Regressor with GridSearch CV gave the parameters required for optimum performance. The model required an R^2 value of 69.16% and MSE of 6.532 on the training set, and the performance of the model on the test data yielded an R^2 of 67.37% and MSE of 12.366, respectively.

2.0 Objectives of the Study

(i) Breakdown prediction using machine learning algorithms.
(ii) Productivity analysis and finding correlation between the parameters i.e., independent variables and the dependent variables.
(iii) Effect of input variables, namely
- Height to burden (H/b) ratio
- Height to width (H/W) ratio
- length to width (L/W) ratio
- Effective in-hole explosive density (de – te/m^3)
- Powder factor (PF) (m^3 /kg – volume of rock broken per kg of explosive)
- Average delay per unit width of burden (ms/m),
- On the dependent variables mainly productivity and the casting percentage.

3.0 Field study

The study was conducted in one of the major opencast coal mines of Northern Coalfields Limited (NCL), Singrauli Coalfields. The Singrauli Coalfield lies between latitudes $23^0 47'$& $24^0 12'$ N, longitudes $81^0 48'$ & $82^0 52'$ E. It occupies an area of over 2,200 km^2. There are 9,121 million tonnes of proven coal reserves in the Singrauli coalfield's northeast, spread across an extent of around 220 km^2. The remaining inferred or indicated reserves total 2,724 million tonnes. Important coal seams include Jhingurda (130-162 m thick), Purewa (8-25 m thick), and Turra (12-22 m thick) in this area of the Singrauli coalfield (Singh, 2004). The overburden above the Turra coal seam is removed by dragline for full exposure of coal. The site map of the study zone is shown in Figure 4. The area has been chosen because it is the only coalfield in India, where opencast mining is used for all coal production. Another distinctive feature of this region is that the large volume of excavation is carried out by deploying large walking draglines operating in tandem to meet the desired rate of coal exposure and coal extraction. This coalfield has the highest number of draglines in India, mostly working in tandem. The various sizes of draglines ranging from 10 to 24 cubic meter bucket size with boom lengths of 72 to 96 m are being deployed in this region. At present, 23 draglines are in operation in this region.

4.0 Machine Learning Models

4.1 Selecting Algorithms

Algorithms for Machine Learning (ML) are thought of as information processing elements that can remember, generalise, and learn from training data. Since ML algorithms have a solid computational foundation, they could be trained to simulate complex physical processes. Numerous ML models have been applied to dragline operations and mining engineering applications. However, all machine learning models share a fundamental framework that includes retraining the model for the least error and then validating it through testing.

Though, the model architecture differs from one another. We have decided to various types of models few of them mentioned here are Ridge Regression, Neural Network Regressor, and XG Boost algorithms, to cover all the possibilities and to find which kind of algorithm is successfully able to encompass the given features for

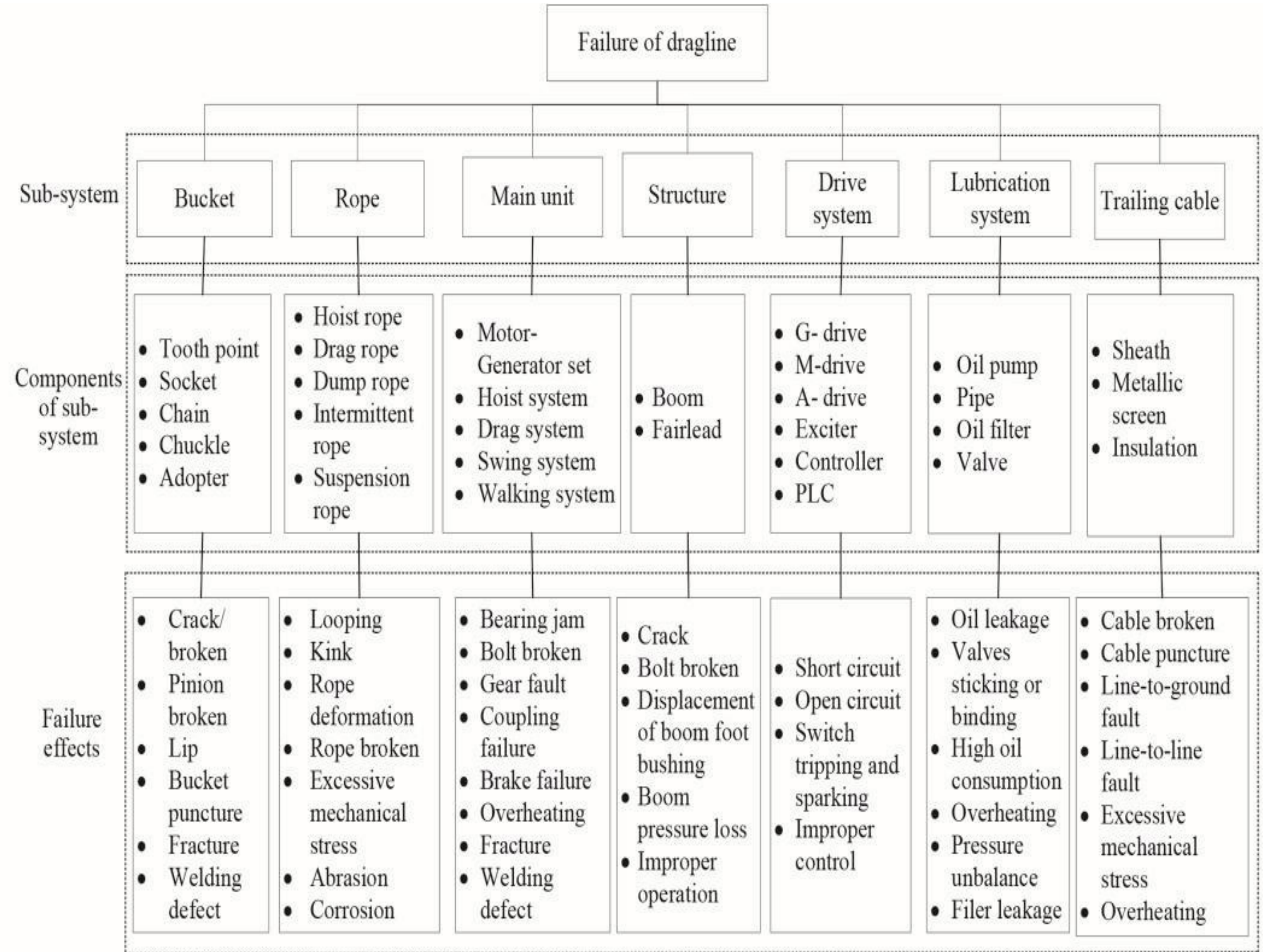

Figure 2. Chart showing various dragline sub-system and their effect of failure.

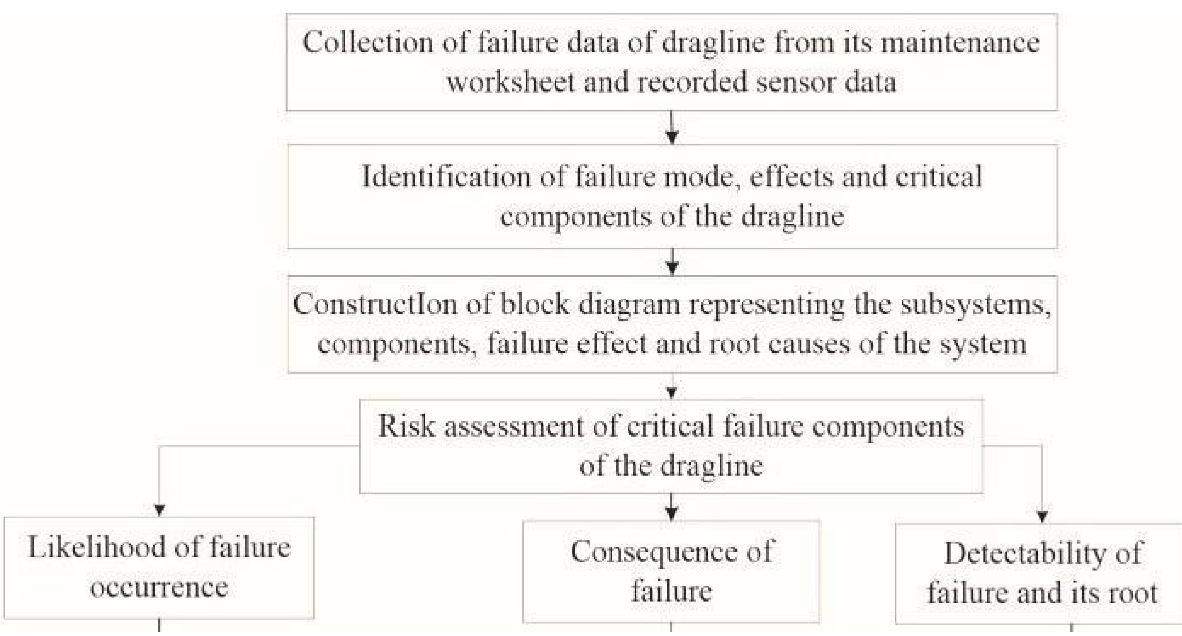

Figure 3. Dragline failure data collection methodology and failure assessment.

output prediction. Since previous research Regression models give the best performance, so we have tried to find which algorithms gives better performance.

4.2 Ridge Regression Model

Ridge regression is a technique for calculating the coefficients of multiple-regression models in cases when linearly independent variables have strong correlation. When linear 22 regression models have a certain multicollinear (highly correlated) independent variables, ridge regression was devised as a potential remedy to the imprecision of least square estimators. This was done by developing a Ridge Regression Estimator (RR). Given that its variance, as well as mean square estimator, are frequently lower than just the least square estimators previously computed, this provides a more accurate ridge parameters estimate. In Ridge Regression, the regression coefficients are computed using the formula:

$$\hat{\boldsymbol{\beta}}_{ridge} = (\boldsymbol{X}^T\boldsymbol{X} + k\boldsymbol{I}_p)^{-1}\boldsymbol{X}^T\boldsymbol{y}$$

I_p is the p x p Identity Matrix
k > 0 and is small

4.3 Neural Network Regressor

In neural networks, several neural layers unite to form a network, or we might say that certain layers have outputs that act as inputs for those other layers. The most typical sort of layer used to build a basic neural network is the fully connected layer, wherein neurons in a single layer are not coupled to one another, and adjacent layers are fully connected pairwise. The first layer act as an input for the second layer and the second layer for the third and hence according to the architecture. There are many activation functions that, along with the combination of weights, decide whether to activate or not a particular neuron.

4.4 XG Boost Regressor

Among various tree-based sequential models, the gradient boosting method known as XG Boost or Xgboost Boosting (XGB) is well-known for its excellent accuracy and speed. By addressing the shortcomings of the Gradient Boosting methodology, XG Boost speeds up computation by focusing on the allocation of characteristics across all data points rather than evaluating the loss of all potential splits and forming a new branch. This narrows the search space for all potential splits.

This method replaces the traditional Gini index with a new parameter defined as the similarity measure for node selection & breaking in decision trees during the optimization phase.

Similarity Score=Gradient^2/(Hessian+λ)

Where "Gradient^2" is the squared sum of the residuals, "Hessian" is the number of residuals, and " λ " is a regularisation hyperparameter. This similarity score is used to determine the correct node. The knowledge obtained reveals the distinction between old and new similarities, indicating the degree of homogeneity attained by dividing the node at a certain place.

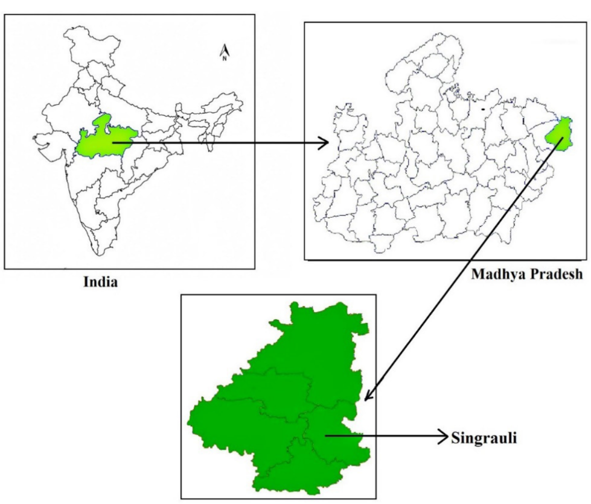

Figure 4. Location map of the study area.

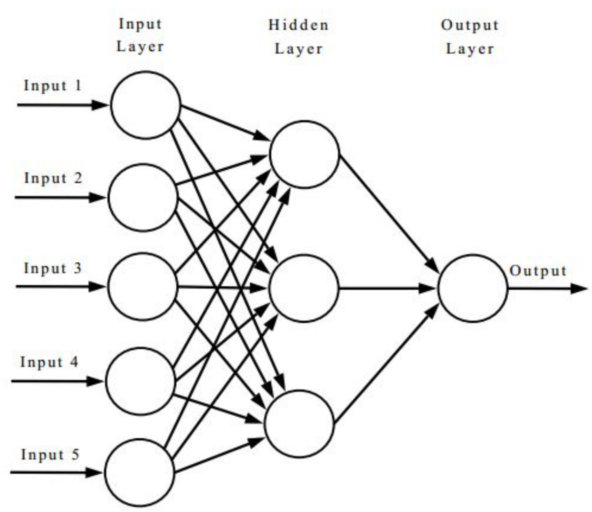

Figure 5. Schematic diagram of a neural network.

Information_Gain = LeftSimilarity + RightSimilarity - SimilarityforRoot

The residuals are first obtained using the chosen loss function and a single leaf tree that has been built. The residuals for succeeding trees originate from the preceding tree's forecasts. The following formula is used to determine the new set of residuals:

NewResidual = OldResiduals + ρ ∑PredictedResidual

5.0 Statistics of Dataset

5.1 Breakdown Hours Dataset Statistics

Output Variable: Breakdown hours

Input Variable: Solid (ton), Rehandling (ton), Working Hours, Idle Hours, Maintenance Hours

5.2 Productivity Data Statistics

Output Variables: Productivity (tons/hour)

Input Variables: Height to width ratio, length to width ratio, bench height to burden ratio, powder factor(m^3/kg), average delay(ms), drop-V(m), cast% and cast volume (m^3)

6.0 Results and Discussions

6.1 Model Evaluation Metrics

For evaluating the effectiveness of various ML models, various performance or validation metrics, such as coefficient of determination (R^2), residual error, and Root Mean Square Error (RMSE), have been frequently used. A statistical technique for evaluating the accuracy of the created models for predicting the actual data points is the R^2 value, which describes the convergent validity of an ML model. The R^2 value ranges from 0 to 1, where a value of 0 indicates that the produced model does not match the presented dataset, and a value of 1 suggests that it does.

6.2 Training Data Result

$$R^2 = 1 - \frac{SS_{RES}}{SS_{TOT}} = 1 - \frac{\sum_i (y_i - \hat{y}_i)^2}{\sum_i (y_i - \overline{y})^2}$$

6.2.1 Breakdown Hours Result

In this simulation, we used various models to fit our dataset of total Breakdown hours in a month. Null values are replaced with mean values and all the volume, and time variables were considered as Input Variables.

Table 1. Overburden handling, working and breakdown hours dataset statistics

Feature	Mean	Std. Deviation	Minimum	Maximum
Solid	192091.96	82827.95	0.00	377760.00
Rehandling	53738.72	34620.73	0.00	216075.00
Working Hours	476.83	144.93	0.00	634.50
Idle Hours	80.80	63.10	0.00	566.00
Maintenance Hours	100.19	135.37	0.00	744.00

Table 2. Dragline Blasting Bench Parameters Data Sets

Feature	Mean	Std. Deviation	Minimum	Maximum
Height/Width	0.64	0.037	0.58	0.82
Width/Length	2.50	1.18	0.78	6.14
$\frac{\text{Bench Height}}{\text{Burden}}$	4.83	0.135	4.60	5.02
Powder Factor	1.50	0.094	1.32	1.73
Average Delay	13.08	1.94	9.13	16.19
Drop - V	14.97	2.85	10.00	19.00
Cast%	30.18	4.91	22.00	38.00
Cast Volume(m^3)	173228.60	45557.56	106657.32	255682.86

Model Name	R^2 score
Ridge Regressor	0.97
Neural Network Regressor	0.07
XG Boost	0.98

It is evident from the results that both Ridge Regression and XG Boost fits the training dataset very well as compared to Neural Network Regressor.

6.2.2 Productivity Dataset Result

In this, we used three regression models to fit our dataset for productivity prediction (m^3/hr). Null values are replaced with mean values due to the shortage of data and all the relevant variables were considered for Input.

Model Name	R^2 score
Ridge Regressor	0.92
Neural Network Regressor	0.80
XG Boost	0.93

Here also, it is evident from the results that both Ridge Regression and XG Boost fit the training dataset very well as compared to Neural Network Regressor, but here Neural Network also fits the dataset, well as compared to the earlier dataset where it was a setback.

6.3 Testing Data Result

6.3.1 Breakdown Hours Test Result

In ths simulation we used the above models to predict total breakdown hours in a month.

Model Name	R^2 score
Ridge Regressor	0.93
Neural Network Regressor	0.02
XG Boost	0.79

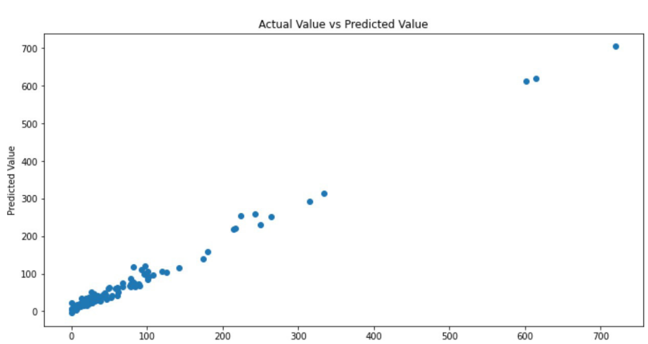

Figure 6. Values predicted by XG Boost Model vs. actual values.

It is evident from the results that Ridge Regression fits the testing dataset well as compared to Neural Network Regressor and XG Boost model.

6.3.2 Productivity Dataset Test Result

In this we used three regression models for prediction of productivity (m^3/hr).

Model Name	R^2 score
Ridge Regressor	0.98
Neural Network Regressor	0.51
XG Boost	0.90

Here also it is evident from the results that Ridge Regression outperforms XG Boost and neural network in prediction.

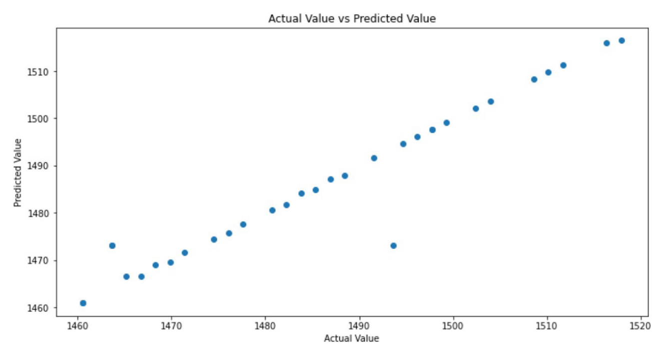

Figure 7. Values predicted by XG Boost Model vs. actual values.

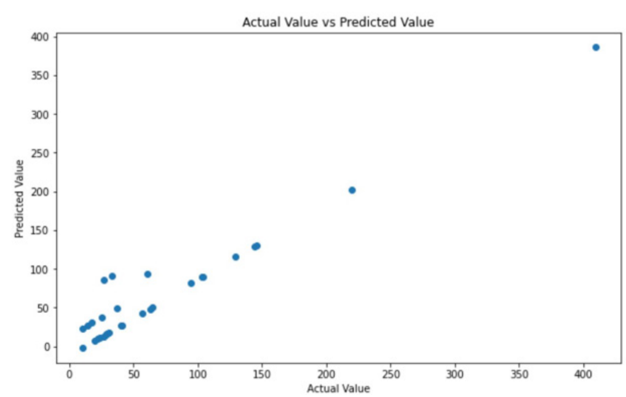

Figure 8. Values predicted by Ridge Regression model vs. actual values on test data.

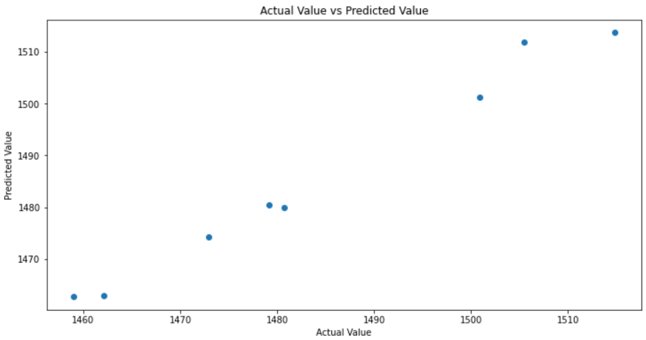

Figure 9. Values predicted by Ridge Regression model vs. actual values on test data.

7.0 Conclusion

Though the neural network is a sophisticated algorithm that can learn more complex patterns and works better than normal supervised algorithms like Ridge Regression and XG Boost (Bagging-based algorithm) in which boosting all the models of the ensemble are weighed based on their performance, on the contrary, neural networks are weighed according to forward and backward propagation algorithm.

But it is evident from the results that in both the cases Ridge Regression and in the second dataset XG Boost gives better results, this could be attributed to-

- Less amount of data, as neural network requires lots of data to learn complex features.
- Less complexity in the features required to predict the output.
- High Variance in the dataset.
- In the first dataset, where XG Boost having R square value of 0.98 on training data and 0.79 on testing data indicates that the model is prone to overfitting because it fits the training data quite well, but it performs poorly on unseen data, and as tree-based models are prone to overfitting that certainly seems to be the case.
- Similarly, in the second dataset, both Ridge Regression and XG Boost gave commendable results, so any of them can be used for prediction analysis.

8.0 References

1. Arunraj, N. S. & Maiti, J. (2007). Risk-based maintenance techniques and applications. *Journal of Hazardous Materials,* **142**(3), 653-661. https://doi.org/10.1016/j.jhazmat.2006.06.069 PMid:16887261
2. Dayawansa, D., Kuruppu, M. & Mashiri, F., (2008). Deterioration mechanisms in dragline wire ropes. Advanced Materials Research. *Trans Tech Publications Ltd.,* **41**, 199-204. https://doi.org/10.4028/www.scientific.net/AMR.41-42.199
3. Sahu, A.R. & Palei, S.K. (2020). Real-time fault diagnosis of HEMM using Bayesian Network: A case study on drag system of dragline. *Engineering Failure Analysis,* **118**, 104917. https://doi.org/10.1016/j.engfailanal.2020.104917
4. Vidyasagar, D. & Kishorilal, D. B. (2016). Maintenance and performance analysis of draglines used in mines. *Int J Comput Eng Res,* **6,** 24-27.
5. Rai, P., Yadav, U. & Kumar, A. (2011). Productivity analysis of draglines operating in horizontal and vertical tandem mode of operation in a coal mine- A case study. *Geotechnical and Geological Engineering* https://doi.org/10.1007/s10706-011-9398-9
6. Rai, S.S., Murthy, V. M. S. R., Kumar, R., Maniteja, M. & Singh, A.K. (2022). Using machine learning algorithms to predict cast blasting performance in surface mining. *Mining Technology.* https://doi.org/10.1080/25726668.2022.2078090
7. Singh, R.D. (2004). Principles and practices of modern coal mining. p. 54, New Age International (P) Limited 1997/2004.
8. Seervi, V., Kishore, N., & Verma, A. (2022). Selection of mode of tandem dragline operations by utilizing 3-dimensional computer graphics balancing diagram: *A case study. Journal of Mines, Metals and Fuels,* **70**(3), 112-123
9. Chaoji, S.V., & Dey, B.C. (2000). Dragline operation in mines - An overview. *Jl of Mines Metals & Fuels,* **XLVIII** (5), 84-93.
10. Rzhevsky, V. V. (1987). Opencast Mining Technology and Integrated Mechanization, Mir Publishers, Moscow.