Prediction of breakdown hours of load haul dumper by long short term memory network

Vaibhav Bisht¹, Sarthak Kumar¹, Anil Kumar Agrawal¹, Mohd A H Siddiqui¹,², Somnath Chattopadhyaya¹

¹Indian Institute of Technology (Indian School of Mines), Dhanbad, Jharkhand 834001, India.
²Coal India Limited, Western Coal Field Limited, Nagpur, 440001,

anil.102024@gmail.com

Abstract. The profitability and feasibility of a system depends on the optimisation of the work of the subsystem as a unit. Here, the breakdown hours of an LHD machine is predicted by modelling a time series on a Neural Network. By accurately predicting the breakdown hours one can plan for efficient mining operations, maintenance and parts supply chain for replacements.

1. Introduction

Mine productions are sophisticated systems comprising of many subsystems. The profitability and feasibility of the system depend on the optimization of the work of the subsystem simultaneously as a unit. LHD machines are employed to haul ore or waste rock from the site of digging and unload it into ore passes or trucks. Being able to predict the time for which the LHD will breakdown help be able to plan for the mining operations and preparing for maintenance activity of the LHD. The breakdown hours of the LHD depend on complex factors and can be formulated as a time series. Traditionally time series have been modelled using Autoregressive Moving Average (ARMA) Model. However, with the advances in neural networks and computation power time series have been modelled using various deep neural network. In this paper, try to model the breakdown data for four LHD machines using Long Short Term Memory networks.

2. Literature survey

Effectiveness of an LHD is mainly affected by the availability, reliability and maintainability of the machine, and its capability to perform as expected. Traditionally, the breakdown of the LHD machines is studied as a stochastic process and modelled using Markov models. Several case studies have been done to perform the reliability analysis of LHD machines [1,2,3]. These studies use graphical methods and statistical tests to find the best fit distribution for characterisation of the failure data. These methods are based on some a priori assumptions which may not be easy to validate for real-life problems.

Treating the reliability analysis as a time series problem helps to deal with these biases. Ho and Xie (1998) [4] have used the Autoregressive Integrated Moving Average (ARMA) model [5] for reliability forecasting and analysis. However, comparable or even better results have been achieved using artificial neural networks [6,7]. Neural networks’ ability to approximate any distribution has provided exceptional results. This study aims to take the application of neural networks to reliability analysis further by using a long short term memory (LSTM) network. LSTMs are a type of (recurrent neural network) RNN that have been shown to be very useful for training temporal and sequential data. LSTMs have solved the problems of vanishing gradient, and long term dependencies faced with standard RNNs.
and have found use in varied problems. Many studies have been done using them on financial market data [8,9]. In this analysis, we are going to fit an LSTM on breakdown data of four LHD machines and analyse its performance.

3. Machine details
Load Haul Dumper (LHD) machines are loaders which are similar to conventional front end loaders but developed for the toughest of hard rock mining applications, with overall production economy, safety and reliability in mind. They are rugged, highly manoeuvrable and exceptionally productive. More than 75% of the world's underground metal mines use LHD for handling the muck of their excavations.

LHD has powerful prime movers, advanced drive train technology, heavy planetary axles, four-wheel drive, articulated steering and ergonomic controls. Their narrower, longer and lower profile make them most suitable for underground application where height and width are limited. As the length is not a limitation in a tunnel and decline LHD are designed with sufficient length. The range improves axial weight distribution and bucket capacity can be enhanced. The two-part construction with central articulation helps in tracking and manoeuvrability. In mining, there is a limitation for shifting heavy equipment. Sometimes, an LHD has to be moved through a shaft while dismantled.

Western coalfields limited, Nagpur uses LHD machines manufactured by EIMCO ELECON. Data of four LHD machines were collected, the serial numbers of which were 811-682, 811-784, 811-785 and 811-685.

4. ANN model
The neural network is a learning model that achieves efficient results in a wide range of both supervised and unsupervised machine learning tasks. Machine perception tasks, where the raw underlying features are not individually interpretable, is where they are uniquely suited. Unlike traditional methods that rely upon hand-engineered features, their success is attributed to their ability to learn hierarchical representations. Despite their ability to understand complex underlying relations, standard neural networks have limitations. Most notably, they rely on the assumption of independence among the training and test examples. The entire state of the system is lost after each data point is processed. If each model is generated independently, this presents no problem. But in case of a time series where data points are related, this is unacceptable. Additionally, standard networks take fixed-length vectors as input. Thus new network structures have been proposed to extend these powerful learning tools to model data with temporal or sequential structure and varying length inputs and outputs.

RNNs are connectionist models with the ability to selectively pass information across sequence steps while processing sequential data one element at a time. Thus they can be used to model data where input and/or output consists of sequences of elements that are autocorrelated. Further, RNNs can simultaneously model sequential and time dependencies on multiple scales. Traditionally, recurrent neural networks have been difficult to train, and often contain millions of parameters. However, recent advances in network architectures, optimization techniques, and parallel computation have enabled solving many machine learning problems with them. In particular, systems based on long short-term memory (LSTM) architectures have demonstrated exceptional results on tasks as varied as language translation, image captioning, time series modelling, and handwriting recognition.

To overcome the problem of vanishing gradients, Hochreiter and Schmidhuber [1997] introduced the LSTM model. Standard RNN has a chain-like structure with a repeating neural network module, as shown in figure 1.
The module takes in previous cell state $h_{t-1}$ and input $x_t$ and the tanh neural layer performs the computation. The updated cell state $h_t$ is a function of previous cell state $h_{t-1}$ and input $x_t$.

To overcome the problem of long term dependencies the repeating module of a standard RNN is modified to produce the repeating module of an LSTM, as shown in figure 2.

The different interacting layers in an LSTM selectively control the flow of information. The key building block is a structure called gate which enables the LSTM to selectively add or remove information from the cell state. And barriers consist of a neural net layer like a sigmoid and a pointwise multiplication. Enabled by the interacting neural net layers and pointwise operations LSTM process information in four simple steps:

- Forget – It forgets the irrelevant parts of the previous state.
\cdot Store – It stores relevant new information into the cell state.
\cdot Update – It selectively updates the current cell state values.
\cdot Output – The output gate controls what information is sent to the next time step.

For this analysis, the LSTM code has been implemented using Keras, as shown in figure 3.

```
In [48]: regressor = Sequential()
    | regressor.add(Dense(units=10))
    | regressor.add(LSTM(units=100, return_sequences=True, input_shape=(X_train_min.shape[1],1)))
    | regressor.add(Dense(units=100))
    | regressor.add(Dropout(0.2))
    | regressor.add(Dense(units=1))
    | regressor.compile(optimizer='adam', loss='mean_squared_error')
```

Fig 3: Keras implementation of the LSTM network

5. Data collection and modelling
The data for this experiment was provided by the Western Coalfields Limited, Saoner Mine. The breakdown data of each of the four LHD machines was recorded, starting from July 2017 up to December 2019.

To prepare the data to be fed into the neural network, the following steps were taken:
\* The breakdown data were scaled to the range of [0,1] using the MinMaxScaler() function provided by the sci-kit learn library.
\* The sequential scaled data were then arranged into an input dataset and an output dataset. For each time step, the corresponding output data is the next time step in the series.
\* The input and output dataset was then split into training, validation and test sets with an approximate split of 70/15/15 per cent.

Finally, the training, validation and test sets were reshaped to be fed into the long short term memory network.

6. Analysis
We have used the Mean Squared Error(MSE) as the cost function to be optimised by the adam optimiser. MSE is the mean of the squares of the error calculated for each data point and unlike Mean Absolute Error(MAE) punishes harshly output values that stray far away from the actual value. The derivatives calculated on the cost function are backpropagated in the network to update the weights. Root Mean Squared Error(RMSE) has been used to report the model error for the four LHD machines. RMSE is the square root of MSE and hence makes it easy to interpret the result in comparison to the input values.

\[
RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^{n} (\hat{y}_i - y_i)^2}
\]

7. Result
The sequential breakdown data for the four LHD machines was fit into the LSTM model. Table 1 and figures 4, 5, 6, and 7 show the comparison of the actual vs predicted breakdown hours for the four machines using a line plot.

The average RMSE on the validation set for the four machines comes out to be 27.461075 and the average RMSE on the test set for the four machines comes out to be 27.7685, showing the model does well on both validation and testing set across all machines.

**Table 1. Comparison of the actual vs predicted breakdown hours**
| LHD No    | Validation RMSE | Test RMSE   |
|-----------|-----------------|-------------|
| 811-682   | 14.59941        | 13.98681    |
| 811-784   | 33.42955        | 32.85165    |
| 811-785   | 39.53129        | 28.09883    |
| 811-607   | 22.28421        | 36.13755    |
| Average   | 27.461075       | 27.7685     |

8. Conclusion
A priori knowledge of the breakdown hours of an LHD machine in an underground mining operation allows for effective resource management, forecasting production and is also helpful for the scheduling preventive maintenance of the machine. With the proposed model, we are able to predict the breakdown hours for the next month given the data for previous months with an acceptable margin of error. The LSTM model has been successful in capturing the complexity and dependence of the series on time with good accuracy.

Acknowledgement
Authors are thankful to Coal India Limited for field investigation as well as technical supports.

References

1. Barabady J and Kumar U 2008. Reliability analysis of mining equipment: A case study of a crushing plant at Jajarm Bauxite Mine in Iran. *Reliability engineering & system safety*, 93(4), pp.647-653.
2. Barabady J 2005 January. Reliability and maintainability analysis of crushing plants in Jajarm Bauxite Mine of Iran. In *Annual Reliability and Maintainability Symposium, 2005. Proceedings*. (pp. 109-115). IEEE.
3. Vashistha S, Agrawal A K, Siddiqui M A and Chattopadhyaya S 2019, November. Reliability and Maintainability Analysis of LHD Loader at Saoner Mines, Nagpur, India. In *IOP Conference Series: Materials Science and Engineering* (Vol. 691, No. 1, p. 012013). IOP Publishing
4. Ho, S L and Xie M 1998. The use of ARIMA models for reliability forecasting and analysis. *Computers & industrial engineering*, 35(1-2), pp.213-216.
5. Box G E, Jenkins G M and Reinsel G C 2011. *Time series analysis: forecasting and control* (Vol. 734). John Wiley & Sons.
6. Chatterjee S and Bandopadhyay S 2012. Reliability estimation using a genetic algorithm-based artificial neural network: An application to a load-haul-dump machine. *Expert Systems with Applications*, 39(12), pp.10943-10951.
7. Ho S L, Xie M and Goh T N 2002. A comparative study of neural network and Box-Jenkins ARIMA modeling in time series prediction. *Computers & Industrial Engineering*, 42(2-4), pp.371-375.
8. Roondiwala M, Patel H and Varma S 2017. Predicting stock prices using LSTM. *International Journal of Science and Research (IJSR)*, 6(4), pp.1754-1756.
9. Tsang G, Deng J and Xie X 2018, August. Recurrent Neural Networks for Financial Time-Series Modelling. In 2018 24th International Conference on Pattern Recognition (ICPR) (pp. 892-897). IEEE.