A Study on Implementation of Multivariate LSTM in Estimation of Gross Calorific Value of Coals

Makarand M Upkare, Archis A Mahore, Ayush S Kohade

Department of Chemical Engineering
Vishwakarma Institute of Technology, Pune, Maharashtra, India
Email: makarand.upkare@gmail.com

Abstract — This paper aims to report the Gross Calorific Value of coal using Multivariate LSTM. The total of ten parameters are taken into the consideration which are non-linear models. These are developed using the LSTM methodology. The first LSTM model developed here uses the major constituents of the proximate and ultimate analyses as inputs while the remaining four sub-models make the different combinations of the constituents of the stated analyses. It was reported that more accurate results are obtained in ANN based models than RNN based models. It was also found that some features of coal samples like moisture content, ash content may dominate other features like volatile matter, He density, etc. in estimation of GCV of coals.

Keywords — Gross Calorific Value, Long Short-Term Memory, RNN, Root Mean Square Error.

I. INTRODUCTION

Coal is a combustible black or brownish-black sedimentary rock formed from traces of ferns and other primitive animals which were covered and buried millions of years ago. Coal is mostly the carbon with small amounts of other elements like hydrogen, sulphur, oxygen, and nitrogen. It is an important source of energy because of its abundance and versatility. There are very many industrial products which are made from coal, and it is majorly used in thermal power plants, metals, and chemicals. The reason that coal is called primary fossil fuel is because at present times, the demand for energy is drastically increasing around the globe due to acceleration in industrial development.

The chemical composition of coals is characterized in terms of their proximate as well as the ultimate analyses. The difference in both is that the proximate analysis investigates the contents of volatile matter, contents of moisture, contents of fixed carbon and contents of ash. On the other hand, the ultimate analysis studies the content of different elements like carbon, hydrogen, nitrogen, sulphur, and oxygen. The heat of combustion which is also called the calorific value is the amount of heat evolved when a unit weight of the fuel is burnt completely and the combustion products cooled to a standard temperature of 298 K. In general, it is termed as gross calorific value (GCV) and sometimes as higher heating value, HHV.

Depending on the moisture content, the ash content and the type of coal, the level of gross calorific value (GCV) varies significantly and it is decided on the basis of dry mineral matter free which is generally termed as DMF, as received, and dry basis. Among these, the DMF basis is useful for scientific evaluation and classification of coals while in commercial applications calorific values are commonly determined using a received or dry basis.

In this paper, various non-linear correlations are proposed for GCV estimation of coals with the LSTM method. Models have been predicted on the basis of an Artificial Intelligence (AI) formalism namely Long Short-Term Memory (LSTM) that has been introduced in order to estimate GCV with a special focus on Indian coals.

II. LITERATURE REVIEW

Various methods have been adopted in finding gross calorific value (GCV).

Hoang Nguyen et al [1] reported a paper on the gross calorific value (GCV) of coal. In this paper, there were 8 AI models and it was determined very accurately and rapidly using them based on big data of 2583 observations of coal samples.

Zhimin Lu [2] proposed a rapid GCV determination method in his thesis which combined laser-induced breakdown spectroscopy (LIBS) technique with two parts, one was artificial neural networks (ANN) and the another was Genetic Algorithm (GA).

Shagufta U. Patel [3] developed a total of seven nonlinear models using the artificial neural networks (ANN) methodology for the estimation of GCV with a special focus on Indian coals.

Priya Kumari [4] reported a paper on the application of non-unitary exponents to the moisture content and ash yield. She also introduced some non-linear terms for the equation to provide empirical correlations with improved prediction accuracy especially for local and national application which is further evaluated. The evaluation was achieved in Excel using its built-in optimizer called Solver which was used to evaluate a generic equation and optimize its coefficients when applied to a dataset of 756 coals from three Indian coal basins of that dataset.

Li Jing et al. [5] in his thesis, analyzed the relationship between industry analysis data and calorific value of fire coal into the furnace. The original independent variables were selected as five kinds which were 1. moisture, 2. ash, 3. sulphur content, 4. volatile matter and 5. fixed carbon in fire coal.

Roselito de et al. [6] Albuquerque Teixeira presented a new report for improving generalization of multilayer perceptrons. This algorithm used a multi-objective optimization (MOO) approach to balance between the error of the norm of network weight vectors and the training data to avoid overfitting.

S.A. Channiwala [7] reported a unified correlation in the paper for making the computation of higher heating values (HHV) from elemental analysis of the fuels.

Willy Wojsznis [8] developed a technique of multi-objective optimization for Model Predictive Control (MPC). In this technique, the optimizations were having three levels of the objective function in priority: 1. handling constraints, 2. maximizing economics, and 3. maintaining control.

GP Rangaiah [9] presented MOO and compared it with single objective optimization, and then MOO studies on design of energy efficient processes which were published from January 2013 to February 2015, that are reviewed.

Hui Liu [10] reported a comprehensive review on the multi-objective optimization technologies in the subtopic of wind energy forecasting where he introduces the basic theories in brief and the certain methods which are related to the multi-objective optimization.

III. METHODOLOGY

A. Dataset

In order to find the gross calorific value (GCV) of the coals, there are certainly two types of correlations where one is used only for coals and the other type includes solid, liquid and gaseous fuels. The dataset contains several parameters, which are mentioned below:

Q: Calorific value
C_M: percentage of moisture at 60% relative humidity at 40ºC
C_A: percentage of ash at 60% relative humidity at 40ºC
V_m: percentage of fixed volatiles on dry ash free basis
F_C: percentage of fixed carbon on dry ash free basis
C_C: percentage of carbon
C_H: percentage of hydrogen
C_S: percentage of sulphur
C_N: percentage of nitrogen
C_O: oxygen content in coal
P_{He}: Helium density

These parameters have certain correlations with each other. The ranges of the mass percentage value (on dry basis) over which the correlation is valid are:

$0\% \leq CC \leq 92.29\%$
$0.43\% \leq CH \leq 25.15\%$
$0\% \leq CO \leq 50\%$
$0\% \leq CN \leq 5.6\%$
$0\% \leq CS \leq 94.08\%$
$0\% \leq CA \leq 71.49\%$
$4.475 \text{ MJ/kg} \leq Q \leq 55.345 \text{ MJ/kg}$.

But the correlations that were proposed in the past have certain drawbacks and difficulties where it required the certain values from the elemental analysis of coals which needed costly equipment. To overcome this difficulty, S.A. Channiwala [7] presented and developed a correlation which was totally based upon the proximate analysis-based correlation to predict the Gross Calorific value. This correlation was used for all solid carbonaceous materials such as coals, lignite, all types of biomass materials, and char to residue-derived fuels.

The correlation is as follows:

$Q = 0.3536 * F_C + 0.1559 * V_m - 0.0078 * C_A (MJ/Kg)$

To understand the proper fittings of the linear relationships, the suitable data is considered from proximate and ultimate analyses of a large number of coals mined from different regions in India.

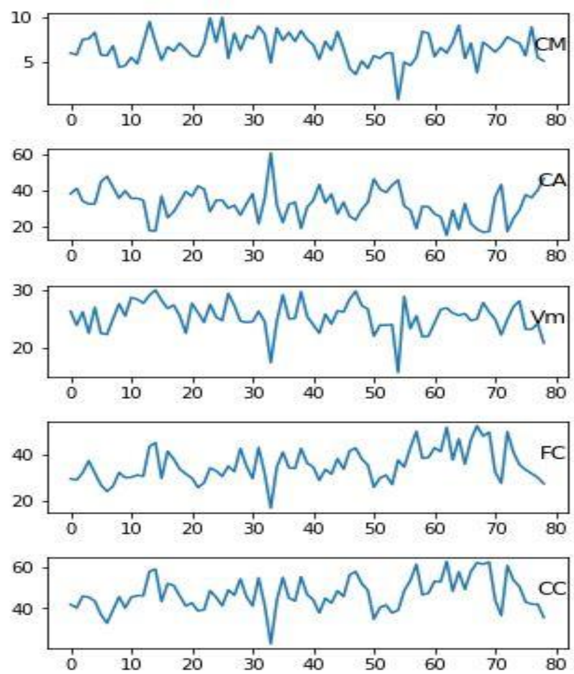

Figure 1a

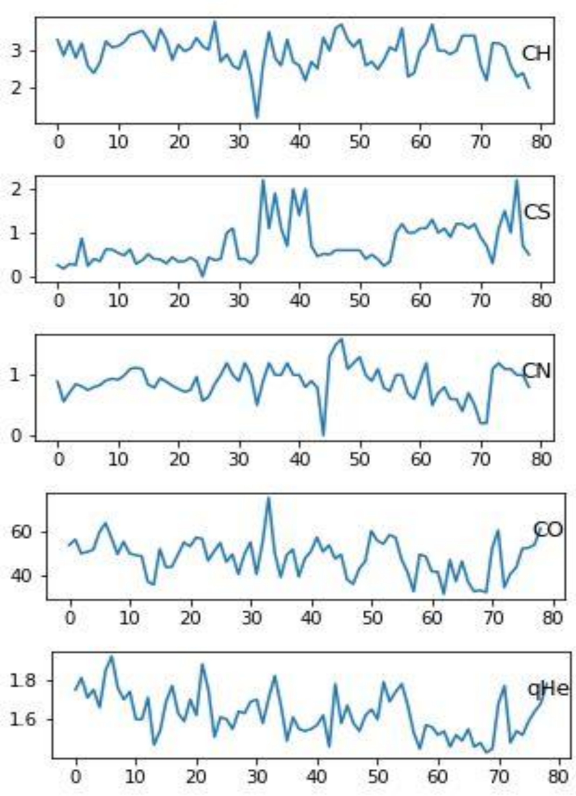

Figure 1b

Figure 1a and 1b depicts the plots of different features of multiple analyses which shows the various trends of various features of multiple coal samples.

B. LSTM

Long Short-Term Memory (LSTM) is an artificial architecture of Recurrent Neural Network (RNN) which is generally used in the field of Deep Learning, and is efficient enough to learn the long-term dependencies. In LSTM, there are certain feedback connections where it processes the entire list of the data (eg. Video). By utilizing experience in modeling core engineering problems [13, 14], various techniques for mathematical optimization of data were shortlisted. From the techniques available, RNN based LSTM provided a fresher scope to the study of the current available dataset.

LSTM is designed in such a way that it avoids the long-term dependency problem. A common LSTM unit comprises four parts: a cell, an input gate, an output gate and a forget gate. The function of a cell is that it remembers the values over arbitrary time intervals and the three gates function in such a way that they regulate the flow of information into and out of the cell.

The major advantage of RNN is that it can model the number of records or the time collection data where every pattern will be assumed to be dependent on previous one. These RNN with LSTM are also implemented with convolutional layers in order to extend the powerful pixel neighborhood.

C. Working

Different types of RNN-LSTM models were implemented to study the relationship between GCV of coals and features of coals as well as the internal dependence on their past and future values. Multiple preprocessing and postprocessing dataset manipulation steps were implemented to get the best possible trained and predicted fit for the models.

Step 1: Raw Data

The raw dataset was referred from the study [3]. 79 different coal samples were analyzed in this dataset for all the stated features from different coal mines of India.

Step 2: Data Preprocessing

The data was screened for any anomalies like missing values, NAN values or data labels which cannot be interpreted by LSTM. After cleaning, the final data set had 12 columns, 1 and 12 being Sr. No. and GCV (experimental value MJ/kg) and the rest 10 being various features of coals which affect GCV values.

After the preprocessing of the data, it was divided into training and testing [3]. 64 coal samples were reserved for training the model while the remaining 15 were used to test the trained model.

Step 3: Normalization, Scaling of data and Reshaping of dataset

As LSTM models are sensitive to numerical data, it is advisable to normalize and scale the complete data in minimum and maximum value bounds. In these models, the data is scaled from 0 to 1. The models implemented multivariate LSTM technique, thus the dimensions of inputs to dimension of output outnumbered. Thus, reshaping of data with no. of features to number of coal samples used was adjusted correspondingly to output i.e., GCV of coal.

Step 4: Training neural network

LSTM works on the principle of RNN. It requires a certain number of sets of previous input and output parameters to be trained. All of the trained models mentioned in Models 1 to 5, uses 5 is to 1 timesteps i.e., 5 previous inputs of features and their respective output GCV to predict next 1 GCV output. After this the data is fed to the neural network and trained for prediction assigning random biases and weights. Our LSTM model is composed of a sequential input layer followed by 2 LSTM layers with ReLU activation and a dropout layer and then finally a dense output layer.

```
model = Sequential()
model.add(LSTM(64, activation='relu', input_shape=(trainX.shape[1], trainX.shape[2]), return_sequences=Tru
model.add(LSTM(32, activation='relu', return_sequences=False))
model.add(Dropout(0.2))
model.add(Dense(trainY.shape[1]))
model.compile(optimizer='adam', loss='mse')
```

Figure 2: Code snippet of trained LSTM model

```
Model: "sequential"
_________________________________________________________________
Layer (type)                 Output Shape              Param #
=================================================================
lstm (LSTM)                  (None, 5, 64)             17152
_________________________________________________________________
lstm_1 (LSTM)                (None, 32)                12416
_________________________________________________________________
dropout (Dropout)            (None, 32)                0
_________________________________________________________________
dense (Dense)                (None, 1)                 33
=================================================================
Total params: 29,601
Trainable params: 29,601
Non-trainable params: 0
_________________________________________________________________
```

Figure 3: MLSTM Model Summary

Step 5: Data post processing and result generation

The output generated by the trained model has to be in the original scale and shape to get a relatable predicted out form. Thus, the trained model is again rescaled and reshaped to original form to get a single column of GCV as an output. Then the trained data is tested and a prediction plot is generated. This prediction vs original plot is evaluated with Root mean squared error.

D. Algorithm

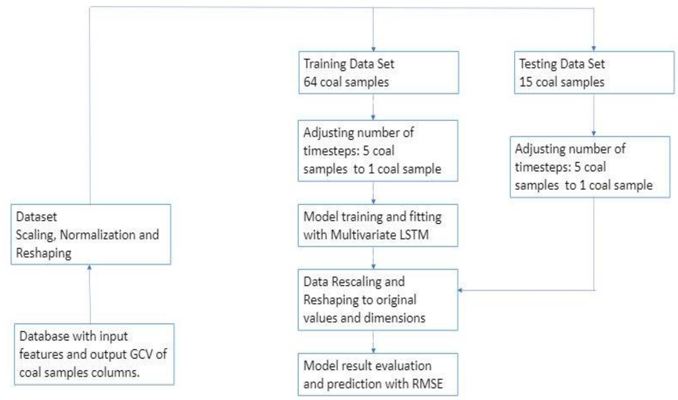

Figure 4: Design algorithm of novel deep learning study of use of MLSTM in estimation of coal gross calorific value

Following model design algorithm was implemented to study the use of multivariate LSTM in estimation of coal gross calorific. The algorithm comprises all the steps that are required for fitting a general data into the RNN based LSTM model and predicting the nature of data.

IV. RESULTS AND DISCUSSIONS

Model 1-Model 5(both trained and predicted): Comparison of trained vs predicted GCV on the basis of input features using multivariate LSTM modelling.

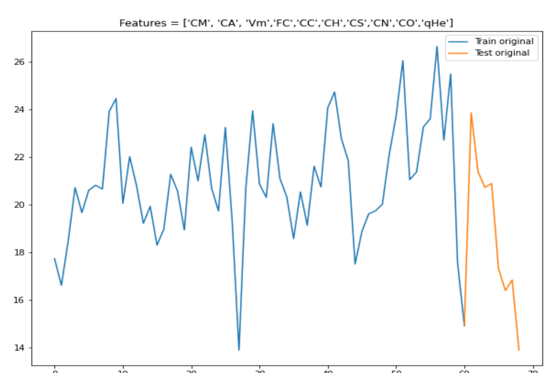

Figure 5(a): Model 1(trained)

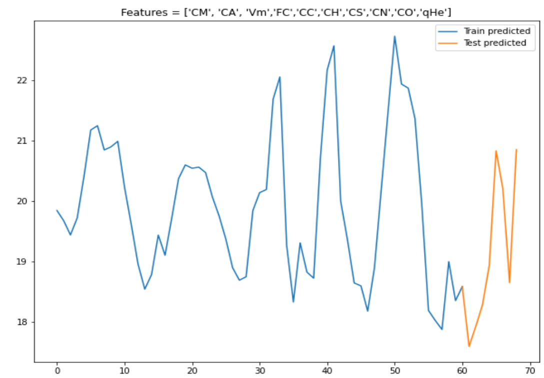

Figure 5(b): Model 1(predicted)

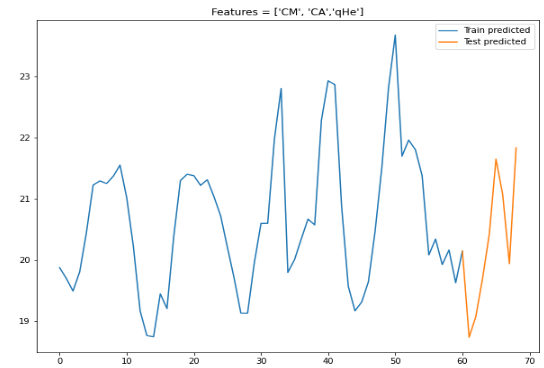

Figure 7(b): Model 3(predicted)

Figure 6(a): Model 2(trained)

Figure 8(a): Model 4(trained)

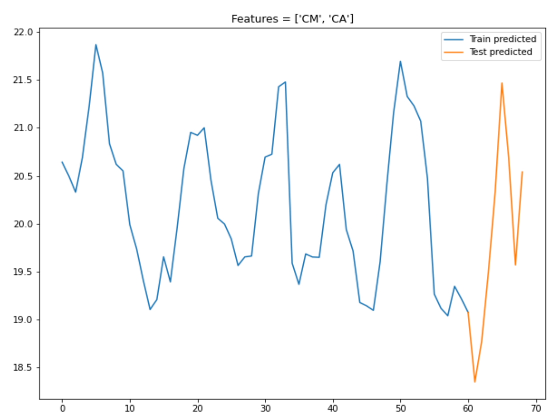

Figure 6(b): Model 2(predicted)

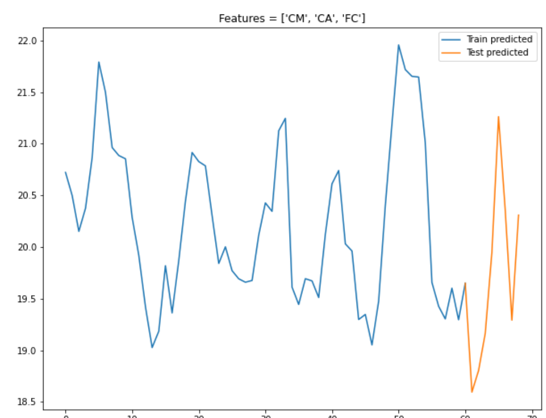

Figure 8(b): Model 4(predicted)

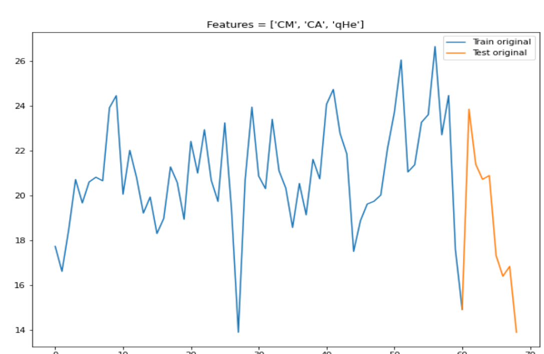

Figure 7(a): Model 3(trained)

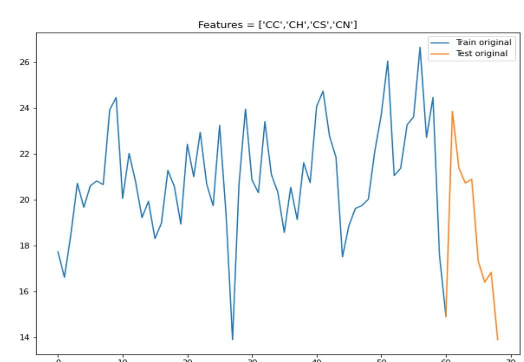

Figure 9(a): Model 5(trained)

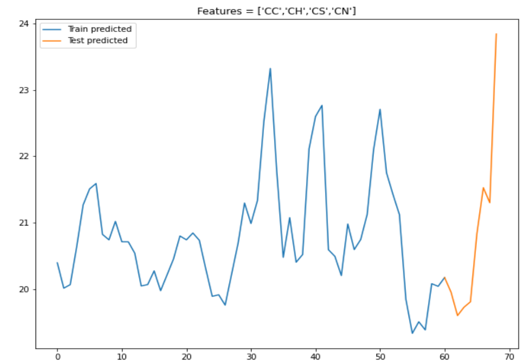

Figure 9(b): Model 5(predicted)

With the help of RNN (Recurrent Neural Network) based architecture, multivariate LSTM models were created with respective combinations of features that influence the generic properties of coal. Models were evaluated in terms of root mean square error (RMSE) for train and test dataset separately as mentioned in Table 1.

As observed in the comparison figures, the plots have a significant variation in trained and test data graphs which indicates that with the help of multivariate LSTM modelling, GCV of coals may not have a recursive nature between features and their experimental values and thus may not be predicted on the basis past GCV of coal samples.

$$\mathrm{RMSE} = \sqrt{\frac{1}{N}\sum\left(\hat{Y}_i - Y_i\right)^2}$$

Formula 1: RMSE calculation

Even Though there are anomalies in the prediction of GCV in MLSTM models, they can be utilized and studied further as generic results correspond precisely to more accurate models in the paper [11].

LSTM Model 1 which includes all features has the least RMSE suggesting the best estimation of GCV of coal samples. The RMSE for Training and Test data set with respective sizes of (63,16) is (19.84, 19.64) respectively. This result suggests that all of the following features can be used to provide the most accurate estimation of GCV of coal samples.

Model 2 takes into account features with more physical relevance like ash and moisture content of coals and gives an average RMSE of (20.3, 20.038) for Training and Test dataset respectively. Considering all the models, this combination of features ranks at the second last possible estimation of GCV of coals.

Table 1: Statistical Analysis of GCV prediction and performance description of multivariate LSTM model

	Model Description	Sizes		Root Mean Square Error (RSME)	
		Training set size	Test set size	Training Set	Test Set
Model 1	Moisture, ash, volatile matter, fixed carbon, carbon, hydrogen, sulphur, nitrogen, oxygen and He-density	63	16	19.84714457	19.64347501
Model 2	Ash, moisture	64	15	20.30648382	20.03876926
Model 3	Ash, moisture, He-density	64	15	20.47069978	20.50457071
Model 4	Ash, moisture, fixed carbon	64	15	19.89864931	20.08845489
Model 5	Carbon, hydrogen, sulphur, nitrogen	64	15	19.89388349	20.04766717

Model 3 and 4 have basic physical features of coal samples constant i.e., Ash and Moisture content differing in chemical features i.e. He-density and fixed carbon content respectively. RMSE of model 3 suggests the worst possible combination of features to predict the estimation of GCV of coal. Model 4 in comparison performs fairly with ranking 3rd in all models. This indicates that the fixed carbon content in the coal dominates GCV more than the relative He-density in the coal samples.

Model 5 has been developed with chemical features of coal samples and performs better in the category of specific selection of combination of various features excluding model 1 which has all features of coal. Model 5 has RMSE (19.8938, 20.047) respectively for training and test datasets which on lags model 1.

V. FUTURE SCOPE

Majority of LSTM literature and studies had been conducted on time series datasets. Usually, these datasets have a large number of sample data points enabling better training and prediction. Our dataset had a limited sample size giving a fair fit to LSTM modelling.

Also, the fact that GCV of coals are not interdependent on previous and future values caused the RNN based structures to underperform. For better understanding, a larger dataset with initial optimization of training and testing datasets with their chronological time entries can be added as a feature such that they complement RNN and MLSTM in the future.

Furthermore, a study on individual features of coal samples in relation with GCV of coal samples can be studied with the help of LSTM to understand the trend of GCV in more detail.

VI. CONCLUSION

Various modelling architectures, optimization techniques and Artificial Neural Network algorithms have been designed to estimate the Gross Calorific Value of coal and to predict its nature. As stated, and observed in various found literature on particular topics, GCV has a dependence on various analyzed features and may have a nonlinear relationship.

In this paper, RNN based Multivariate LSTM models were designed to observe the nature between past and future values of GCV of coals and its relative quantified feature values. 5 models with different combinations of features of coals were curated and modelled.

The results of MLSTM models are particularly ambitious in nature and have not completely disregarded the chances of inter-relation between GCV values and their features. Although it can be concluded that more accurate results are obtained in ANN based models than RNN based models.

The results also show that some features of coal samples like moisture content, ash content may dominate other features like volatile matter, He density, etc in estimation of GCV of coals.

This paper majorly attempted to use LSTM as a helpful tool in relating past and future data points but it can be concluded that estimation of GCV of coals is dependent only on its features and to a larger extent is independent of its past and future values.

REFERENCES

[1] Nguyen, H., Bui, HB. & Bui, XN. Rapid Determination of Gross Calorific Value of Coal Using Artificial Neural Network and Particle Swarm Optimization. Nat Resour Res 30, 621–638 (2021). https://doi.org/10.1007/s11053-020-09727-y.

[2] Zhimin Lu "Rapid Determination of the Gross Calorific Value of Coal Using Laser-Induced Breakdown Spectroscopy Coupled with Artificial Neural Networks (ANN) and Genetic Algorithm (GA)" Energy & Fuels 2017 31 (4), 3849-3855, DOI: 10.1021/acs.energyfuels.7b00025.

[3] Shagufta U. Patel, "Estimation of gross calorific value of coals using artificial neural networks" Fuel, Volume 86, Issue 3, 2007, Pages 334-344, ISSN 0016-2361, https://doi.org/10.1016/j.fuel.2006.07.036.

[4] Kumari, P. et al., "Predictions of Gross Calorific Value of Indian Coals from their Moisture and Ash Content" J Geol Soc India 93, 437–442 (2019), https://doi.org/10.1007/s12594-019-1198-5.

[5] Li Jing, "Predicting the gross calorific value of coal based on support vector machine and partial least squares algorithm," 2016 IEEE International Conference on Knowledge Engineering and Applications (ICKEA), 2016, pp. 221-225, doi: 10.1109/ICKEA.2016.7803023.

[6] Roselito de Albuquerque Teixeira, "Improving generalization of MLPs with multi-objective optimization, Neurocomputing, Volume 35, Issues 1–4, 2000, Pages 189-194, ISSN 0925-2312, https://doi.org/10.1016/S0925-2312(00)00327-1.

[7] S.A. Channiwala, "A unified correlation for estimating HHV of solid, liquid and gaseous fuels", Fuel, Volume 81, Issue 8, 2002, Pages 1051-1063, ISSN 0016-2361, https://doi.org/10.1016/S0016-2361(01)00131-4.

[8] Wojsznis, Willy, "Multi-objective optimization for model predictive control" ISA transactions, 2007, 46.351-61.10.1016/j.isatra.2006.10.002.

[9] GP Rangaiah, "Multi-objective optimization for the design and operation of energy efficient chemical processes and power generation", Current Opinion in Chemical Engineering, Volume 10, 2015, Pages 49-62, ISSN 2211-3398, https://doi.org/10.1016/j.coche.2015.08.006.

[10] Hui Liu, "A review on multi-objective optimization framework in wind energy forecasting techniques and applications", Energy Conversion and Management, Volume 224, 2020, 113324, ISSN 0196-8904, https://doi.org/10.1016/j.enconman.2020.113324.

[11] Nandi, Somnath & Ghosh, Soumitra & Tambe, Sanjeev & Kulkarni, Bhaskar. (2001). Artificial neural-network-assisted stochastic process optimization strategies. AIChE Journal. 47. 126 - 141. 10.1002/aic.690470113.

[12] Makarand M. Upkare, Kalpendra B. Rajurkar, Pankaj R. Gupta, and R. Jaganathan, "Mathematical Modeling and Simulation of Bubble Column Reactor for Aerial Liquid Phase Cyclohexane Oxidation", AIP Conference Proceedings 1298, 151 (2010); https://doi.org/10.1063/1.3516293.

[13] Makarand M. Upkare, Kalpendra B. Rajurkar, Samir K. Das, and R. Jaganathan, "Dynamic Modeling of Hydro- Formylation of 1-Decene on Rh/C Catalyst in Bubble Column Slurry Reactor", AIP Conference Proceedings 1298, 143 (2010); https://doi.org/10.1063/1.3516292.