An Energy Consumption Prediction LSTM Model of Metallurgy Enterprises

Xueying Wang¹, Zhuchao Yu¹*, Pengfei Xi¹, Gaixia Chu²,³, Shiyang Lai¹, Jing Li¹, and Yuanzheng Zhang⁴

¹School of Business Administration, Northeastern University, Shenyang, China
²State Key Laboratory of Process Automation in Mining & Metallurgy, Beijing, China
³Qinghai KXXX Information Technology Co., LTD, Xining, China
⁴College of Medicine and Biological Information Engineering, Northeastern University, Shenyang, China

*Corresponding author: zcyu@mail.neu.edu.cn

Abstract. Aiming at the characteristics of multi-dimensional production data, complicated sources and diverse data structures of metallurgy enterprises, it is of great significance to study how to use energy management-related data to predict metallurgy enterprises’ energy consumption. Using an accurate measurement method, the enterprises can not only reduce the cost of energy consumption but also develop economic efficiency in producing for metallurgy enterprises, which results in low energy use efficiency, high enterprise cost, and weak scalability. In this paper, we establish an LSTM model to achieve energy consumption prediction for metallurgy enterprises by optimizing the model’s parameters using a grid search algorithm. We also compare our model's prediction results with other mainstream machine learning algorithms, i.e., MARS and SVM through indexes such as MSE, RMSE, MAPE, and MRAE to evaluate the prediction effect of the learning algorithm. According to our simulation, LSTM performs best in the task of energy consumption prediction.

1. Introduction
In recent years, metallurgy enterprises have been committed to solving the problem of pollution prevention, promoting energy conservation, emission reduction and green development. The harmonious development and ecological civilization have become a consensus of metallurgy enterprises [1]. In this regard, predicting energy consumption and optimizing energy allocation by applying historical data are of great significance for metallurgy enterprises to realize the transformation and upgrade of “green mines”. An accurate measurement method can not only reduce the cost of energy consumption but also develop economic efficiency for enterprises. Therefore, based on our practical experience of metallurgy workers, we introduce how to use an LSTM model with wavelet analysis to predict metallurgy enterprises’ energy consumption, which results in low energy use efficiency, high enterprise cost, and weak scalability.

2. Literature review
Energy consumption prediction is both the core part of energy management for enterprises and the important direction for energy conservation. In past study, Auto Regressive Moving Average model (ARMA) and Auto Regressive Integrated Moving Average model (ARIMA) [2] based on regression
method are the most typical method for all kinds of time series forecasting. In recent years, due to the high complexity and irregularity of actual production data, more and more researchers have begun to use machine learning algorithms to build statistical energy consumption models. SVR is used by Kim, K to build a solar power generation forecasting model [3]. Multivariate adaptive regression spline (MARS) algorithm is used to build the model to predict the relationship between load demands and several environmental variables [4]. Niu, DX constructed a general regression neural network (GRNN) forecasting model to forecasting total carbon emissions (TCE) and carbon emissions intensity (CEI) in 2016-2040 [5]. Besides, LSTM is considered to be the one of the effective methods for dealing with time series prediction problem. For example, in a short-term load forecasting task for a single resident household, Kong, WC [6] proved that the LSTM method is superior to other listed competitors' algorithms. Sahoo [7] proved the applicability of long short-term memory recurrent neural network (LSTM-RNN) in low-flow time series prediction. In order to maintain the high temperature required by the production, metallurgy enterprises will consume a large amount of energy during the production process. So, the energy consumption of the enterprise will show a significant difference in different months when the air temperature varies widely. The memory unit in the LSTM model can effectively solve the difficulty of timing-related sequence prediction.

3. Method

3.1. Energy consumption prediction model

Since there are many unstable fluctuations in the energy consumption time series data, we constructed our LSTM model as shown in Figure 1.

In this LSTM model, \(x_t\) is the input vector at the \(t\)-th time step, \(a_t\) is the activation value output at the \(t\)-th time step. \(i_t\), \(f_t\), and \(o_t\), respectively, represent the input gate, forget gate, and the output gate. \(C_t\) represents the stored value of the LSTM memory unit and \(c_t\) indicates the stored value of the candidate unit to be updated obtained according to the current \(x_t\), and \(w_t\), \(b\) are parameters of the LSTM unit. In addition, the memory cell value \(C_t\) and the activation value \(a_t\) of the next time step can be trained and obtained by entering the time series vector \(\{x_t = x_1, x_2, \ldots, x_n\}\) into the LSTM network.

The LSTM calculates the mapping from the input data set \(x_t\) to an output sequence \(y_t\), with three different gates: the input gate, the forget gate and the output gate. The calculation process of LSTM unit is as follows:
\[ c_i = \tanh(W_c \cdot [a_{i-1}, x_i] + b_c) \]  
\[ C_i = i_t \cdot c_i + f_t \cdot C_{i-1} \]  
\[ a_i = i_t \cdot \tanh(C_i) \]  

3.2. Model training and solving
The training parameters include the weight matrices \((W_{jv}, W_{fx}, W_{xv}, W_{xv}, W_{vx}, W_{vx}, W_{vx}, W_{vx})\) and the bias vectors \((b_j, b_i, b_c, b_t, b_u)\). The total error \(E\) from 1 to \(t\) is expressed as \[ \sum_{t=1}^{T} \frac{1}{2} || y_i - d_i ||^2, \]
where \(y_i - d_i\) is the error between the real value and the estimated value at time \(t\). We define the gradient of \(E\) to \(n_y\), \(t\) as
\[ \nabla_{n_y} E = \left( \frac{\partial y_i}{\partial n_y} \right)^T \nabla_{y_i} E = (y_i - d_i) \sigma(1 - y_i^2) \]  

4. Dataset and Experiments
The dataset consists of two parts, one is the 2017-2018 energy report provided by Western Copper Company, and the another is the local weather, temperature and wind conditions of the mine area.

4.1. Dataset
This paper has statistics on the total energy consumption of Western Copper Corporation. At the same time, we sort out the energy consumption of the four products of copper, electrolytic copper, lead and zinc produced by Western Copper, and choose the product output, comprehensive power consumption, overall energy consumption, and comprehensive water consumption data as variables to be placed in the model. The comprehensive energy consumption is energy consumables including electricity, general bituminous coal, coke, natural gas, gasoline, diesel, kerosene, and lubricants. According to the industry-specified discount coefficients, the corresponding conversion coefficients converted into tons of standard coal are 0.1229, 0.7143, 0.9714, 12.3, 1.4714, 1.4571, 1.4714 and 1.4143. The statistic results are shown in Table 1.

| PRODUCT          | VARIABLE | MAX  | MIN  | MEAN  | MEDIAN | SD      |
|------------------|----------|------|------|-------|--------|---------|
| Copper           | Pro output | 2643.02 | 1381.00 | 2107.00 | 2120.07 | 333.47  |
|                  | Power    | 776.15 | 426.30 | 608.30 | 613.90 | 66.83   |
|                  | Energy   | 953.90 | 523.90 | 747.60 | 754.51 | 82.13   |
|                  | Water    | 88694.04 | 42682.0 | 66575.00 | 66996.00 | 12870.71 |
| Lead and zinc    | Pro output | 2688.90 | 529.46 | 1937.00 | 2065.92 | 509.81  |
|                  | Power    | 404.30 | 83.40  | 291.80 | 304.50 | 63.01   |
|                  | Energy   | 496.90 | 102.40 | 358.60 | 374.20 | 77.43   |
|                  | Water    | 160762.00 | 21450.0 | 61389.03 | 46747.00 | 29923.64 |
| Electrolytic     | Pro output | 4791.06 | 1784.01 | 4216.05 | 4466.00 | 710.03  |
| copper           | power    | 228.01 | 99.40  | 185.80 | 191.80 | 31.98   |
|                  | energy   | 928.50 | 450.70 | 613.50 | 590.48 | 139.29  |
|                  | Water    | 5170.01 | 470.00 | 2116.02 | 1852.00 | 1379.88 |

The production of metallurgy enterprises will also be affected by factors such as weather and temperature. Daily weather-related dataset is obtained from the national historical weather query
platform (http://www.tianqihoubao.com). And the daily numbers of sunny, rainy, snowy, cloudy and sandy days in each month are treated as variables.

4.2. Data preprocessing
The observations were divided into 18 months as the training set and every 3 months was a prediction set, which means that the LSTM step size was set to 3.

4.3. Evaluation metrics
We selected MSE, RMSE [8], MAPE and MAPE as the evaluation metrics. They are calculated through

\[
MSE = \frac{1}{N} \sum_{i=1}^{N} \left| \hat{y}_i - y_i \right|^2
\]

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

\[
MAPE = \frac{1}{N} \sum_{i=1}^{N} \left| \frac{\hat{y}_i - y_i}{y_i} \right| \times 100\%
\]

\[
MARE = \frac{1}{N} \sum_{i=1}^{N} \left| \frac{\hat{y}_i - y_i}{\max(\hat{y}_i, y_i)} \right| \times 100\%
\]

where \( \hat{y}_i \) represents prediction, \( y_i \) is the real value and \( N \) is the number of testing samples. MSE and RMSE indicate the error between predicted result and true value. MAPE and MARE are employed to show the quality of different models. The smaller the MAPE and MARE values are, the better the model will be.

4.4. Experiment analysis
We set the parameters as epochs = 235, batch size = 18 and trained the LSTM. We also select the MARS and SVR models currently widely used in forecasting for experimental comparison. The copper prediction results of different models are visible in Figure 2.

![Figure 2](image)

**Figure 2.** The prediction results of copper.

The values of evaluation metrics based on different models are shown in Table 2. In the case of using the LSTM model, MSE, RMSE, MAPE, MARE are smaller than other models. So, the LSTM
model has better adaptability to time series data and can achieve better results in various production scenarios of mining enterprises.

Table 2. The values of evaluation metrics based on different models.

| PRODUCT      | INDEX  | MARS      | SVR        | LSTM       |
|--------------|--------|-----------|------------|------------|
| Copper       | MSE    | 5351.53   | 4468.83    | 2736.96    |
|              | RMSE   | 73.15     | 66.85      | 52.32      |
|              | MAPE   | 0.09      | 0.08       | 0.06       |
|              | MARE   | 0.09      | 0.09       | 0.06       |
| Lead and Zinc| MSE    | 394.29    | 636.21     | 77.20      |
|              | RMSE   | 19.86     | 25.22      | 7.89       |
|              | MAPE   | 0.04      | 0.06       | 0.02       |
|              | MARE   | 0.04      | 0.07       | 0.02       |
| Electrolytic Copper | MSE | 13147.76  | 9265.83    | 7333.48    |
|                | RMSE   | 114.66    | 96.26      | 85.64      |
|                | MAPE   | 0.19      | 0.14       | 0.13       |
|                | MARE   | 0.15      | 0.11       | 0.10       |

5. Conclusion
The LSTM model proposed in this article is used to predict the energy consumption of metallurgy enterprises. The experimental results show that the LSTM model has more stable performance and higher accuracy in multi-dimensional time-series data than the SVR and MARS model. And we will evaluate the future results more formally, in order to improve the accuracy of prediction continuously. After these processes, the energy consumption prediction model based on LSTM will help enterprises improve production efficiency and save costs.

Acknowledgments
Our thanks to the 2019 National College Students Innovation and Entrepreneurship Training Program at Northeastern University, China. And, funded by Open Foundation of State Key Laboratory of Process Automation in Mining & Metallurgy.

References
[1] Blam I and Kovalev S 2013 Ecological management in ferrous and non-ferrous metals industrial enterprises Voprosy statistiki 6 22-30.
[2] G.E.P. Box and Davida Pierce 1968 Distribution of residual autocorrelations in autoregressive-integrated moving average time series models Journal of the American Statistical Association 65 (332) 1509-1526.
[3] Kim K and Hur J 2019 Weighting factor selection of the ensemble model for improving forecast accuracy of photovoltaic generating resources Energies 12 3315.
[4] Nalcaci G, et al 2019 Long-term load forecasting: models based on MARS, ANN and LR methods Central European Journal of Operations Research 274 1033-49.
[5] Niu D X, Wang K K, Wu J, Sun L J, Liang Y, Xu X M and Yang X L 2020 Can China achieve its 2030 carbon emissions commitment? Scenario analysis based on an improved general regression neural network Journal of Cleaner Production 243 118558.
[6] Kong W C, et al 2019 Short-term residential load forecasting based on LSTM recurrent neural network IEEE Transactions on Smart Grid 10 841-51.
[7] Sahoo B B, Jha R, Singh A and Kumar D 2019 Long short-term memory (LSTM) Recurrent neural network for low-flow hydrological time series forecasting Acta Geophysica 67 1471-81.
[8] Mark Plutowski, Garrison W. Cottrell and Halbert White 1996 Experience with selecting exemplars from clean data Neural Multi-order 9 (2) 273-294.