

Application of Deep Learning in Power load Analysis

Xinhua Duan

Department of Mathematics and Information Engineering,
Puyang Vocational and Technical College, Puyang Henan
457000, China

*Corresponding Author Email: xinhuaduan4213@163.com

Received: April 16, 2020. Revised: October 19, 2020. Accepted: October 21, 2020. Published: October 26, 2020.

Abstract - Aiming at the problems of slow model training speed and poor prediction effect of traditional power load prediction algorithm, a parallel load prediction method based on deep learning is proposed. The method is based on the MapReduce parallel calculating framework, and the deep belief network model, which is used to parallel training the sample data with the historical load and the weather information, and the model of the training model to predict the load value. The experimental results show that the average root-mean-square error between the predicted power load value and the actual value of the prediction method in this paper is 2.86%. The prediction accuracy is higher than the traditional method, and the training and prediction time are effectively reduced, which can adapt to the prediction requirements of large-scale power data.

Keywords - Deep learning, Power load analysis, parallel computing

I. INTRODUCTION

Power load forecasting is a series of forecasting work which takes electric load as the object, including forecasting the future power demand (power), the future power consumption (energy) and the load curve. Power system load forecasting is an important part of power system planning, and it is also the basis of power grid dispatching, regulation, control and so on. Power load forecasting is an important basis for safe dispatching of electric power, which is of great significance in ensuring the stability, reliability and economic operation of power system. In recent years, Domestic and foreign experts have done a lot of research on forecasting theory, and put forward a variety of load forecasting models, such as BP neural network, wavelet analysis and supporting vector machine, to provide a strong support for power system

load forecasting. Among them, BP neural network has strong ability of nonlinear mapping and generalization, but it is easy to produce over-fitting and local optimum. Wavelet analysis has strong ability of approximation and fault tolerance, but the selection of wavelet basis function and parameter initialization have no certain criterion, supporting vector machine can eliminate a large number of redundant samples, which has good robustness. But it is difficult to implement large-scale sample training, so it is difficult to solve the multi-classification problem. The regression analysis method has the advantages of fast training speed and low error rate when the data is less, but it has poor effect on a large number of data processing. Because of the limitation of precision and range of application, the above power load forecasting algorithms can not only meet the more and more complex power system load forecasting requirements. In addition, with the construction of smart grid in China, a great deal of data will be produced, which will bring great challenge to the computing efficiency.

In order to improve the prediction accuracy and improve the training efficiency, this paper presents a deep learning load forecasting model based on parallel computing framework for big data on the power user side. Firstly, the deep belief network (DBN) is combined with MapReduce parallel computing framework, then the load forecasting model is trained by historical load data and weather data, and the power load forecasting model is used to forecast power load. After introducing big data and parallel operation, the method can make more reasonable use of computing resources and avoid the problem of dimensionality disaster in power system.

II. PARALLEL LOAD FORECASTING BASED ON DEPTH LEARNING

Deep learning training process

The training process of deep learning is as follows.

Step one, use bottom-up unsupervised learning, that is, to train from the bottom up to the top layer by layer. Using untagged data (or tagged data) to train each layer of parameters in a hierarchical manner. This step can be regarded as an unsupervised training process, because of the limitation of model capacity and the constraint of sparsity. So that the model can learn the structure of the data itself and get the features which are more expressive than the input. This process can also be regarded as the process of feature learning. This step is also the biggest difference between deep learning and traditional neural networks [1-2].

Step two, top-down supervised learning, that is, training through tagged data, error top-down transmission, fine-tuning the network. Based on the parameters obtained from the first step, the parameters of the whole multilayer model are further fine-tuned, which is a supervised training process. The first step of training is similar to the random initialization process of neural network, but because the initial value of the first step of deep learning is not random, it is obtained by learning the

structure of input data. Therefore, this initial value is closer to the global optimum than the initial value of neural network and can achieve better results. So, the deep learning algorithm is superior to the neural network to a great extent due to the feature learning process of the first step.

Power load forecasting system Architecture

Timely, accurate and reliable basic information is the basis of network load forecasting. In order to realize the timely and accurate load forecasting, this paper uses the traditional big data processing framework to construct the load forecasting framework, which includes all the links of big data technology, and the functions of each layer are clear and all the layers are coordinated and cooperate with each other, to meet the requirements of power grid load calculation. The system architecture is divided into resource layer, platform layer and application layer, as shown in figure 1. Among them, the resource layer generates, receives the basic information such as historical load data and weather data, and then transmits the data information to the platform layer. The platform layer provides data processing, data storage, cluster management and other functions, and provides interfaces to the application layer, which is the most important part of the whole system architecture [3-6]. The application layer outputs the final data processing results, and visualizes, query the prediction results and so on.

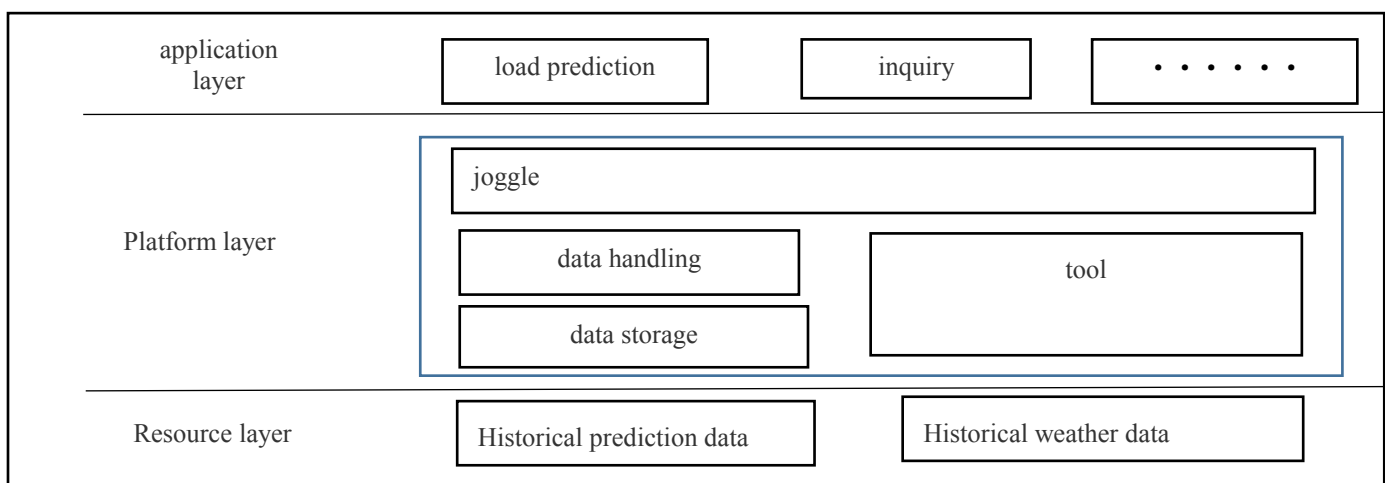


Fig. 1 Power load forecasting system architecture

Deep belief Network Model

The deep belief network is a common model of deep learning and an unsupervised learning model based on energy. The model defines the probability distribution of the visible variable v and the hidden variable h through the energy function. The energy function is defined as:

$$E(v, h | \theta) = -\sum_{i=1}^M \sum_{j=1}^N W_{ij} v_i h_j - \sum_{i=1}^M b_i v_i - \sum_{j=1}^N a_j h_j \quad (1)$$

Among them, W_{ij} is the link weight between the visible layer neuron v_i and the hidden layer neuron h_j , a and b are the bias between the visible unit and the hidden unit, θ represents the parameter group $\{W, a, b\}$ [7-8].

When the parameters are determined, the joint probability distribution of v and h can be obtained by the joint configuration energy function based on v and h .

$$p_{\theta}(v, h | \theta) = \frac{1}{Z_{\theta}} e^{-E(v, h | \theta)} \quad (2)$$

When v is given, the activation state of each hidden unit is conditional independent. In this case, the formula for calculating the activation probability of the j hidden unit is

$$p_{\theta}(h_j = 1 | \theta) = \sigma(a_j + \sum v_i W_{ij}) \quad (3)$$

Similarly, when h is given, the activation probability of the number i visible unit is

$$p_{\theta}(h_j = 1 | \theta) = \sigma(a_j + \sum v_i W_{ij}) \quad (4)$$

In order to obtain each parameter $\{W, a, b\}$, the training data need to be fitted. According to the principle of maximization of likelihood function, using Gibbs Sampling (Gibbs sampling method), the weight updating formula is derived.

$$W_{ij} \leftarrow W_{ij} + \eta [\langle v_i h_j \rangle_{data} - \langle v_i h_j \rangle_{model}] \quad (5)$$

$$b_i \leftarrow b_i + \eta [\langle v_i \rangle_{data} - \langle v_i \rangle_{model}] \quad (6)$$

$$a_j \leftarrow a_j + \eta [\langle h_j \rangle_{data} - \langle h_j \rangle_{model}] \quad (7)$$

Parallel load forecasting process based on DBN

Data processing in platform layer is the core of load forecasting. The data processing adopts the MapReduce parallel computing architecture, using three MapReduce job classes to execute the training process of the algorithm, and the output of the former MapReduce is taken as the input of the latter. The data processing flow is divided into three parts: data pre-processing, generating forecasting model and forming parallel DBN load forecasting model [9-11]:

1) The data pre-processing is accomplished by the first MapReduce, and the collected data is cleaned and divided into blocks, which can be used by the next MapReduce.

2) The prediction model is generated in parallel. The pre-processed data block is transferred into a prediction model by using Map, to realize the parallel algorithm. The MapReduce is only used in the Map process;

3) A parallel DBN load forecasting model can be generated by combining K models generated by Map, where K models correspond to k group model parameters,

and the predicted model parameters are the average values of K group parameters. This process is completed by the third MapReduce.

The parallel load forecasting flow chart based on DBN is shown in figure 2. The model is built in Hadoop distributed cluster, and the deep belief network prediction algorithm is parallelized by MapReduce to realize data mining and computational prediction. The parallel execution process improves the accuracy of load forecasting and the ability of load forecasting system to deal with power big data.

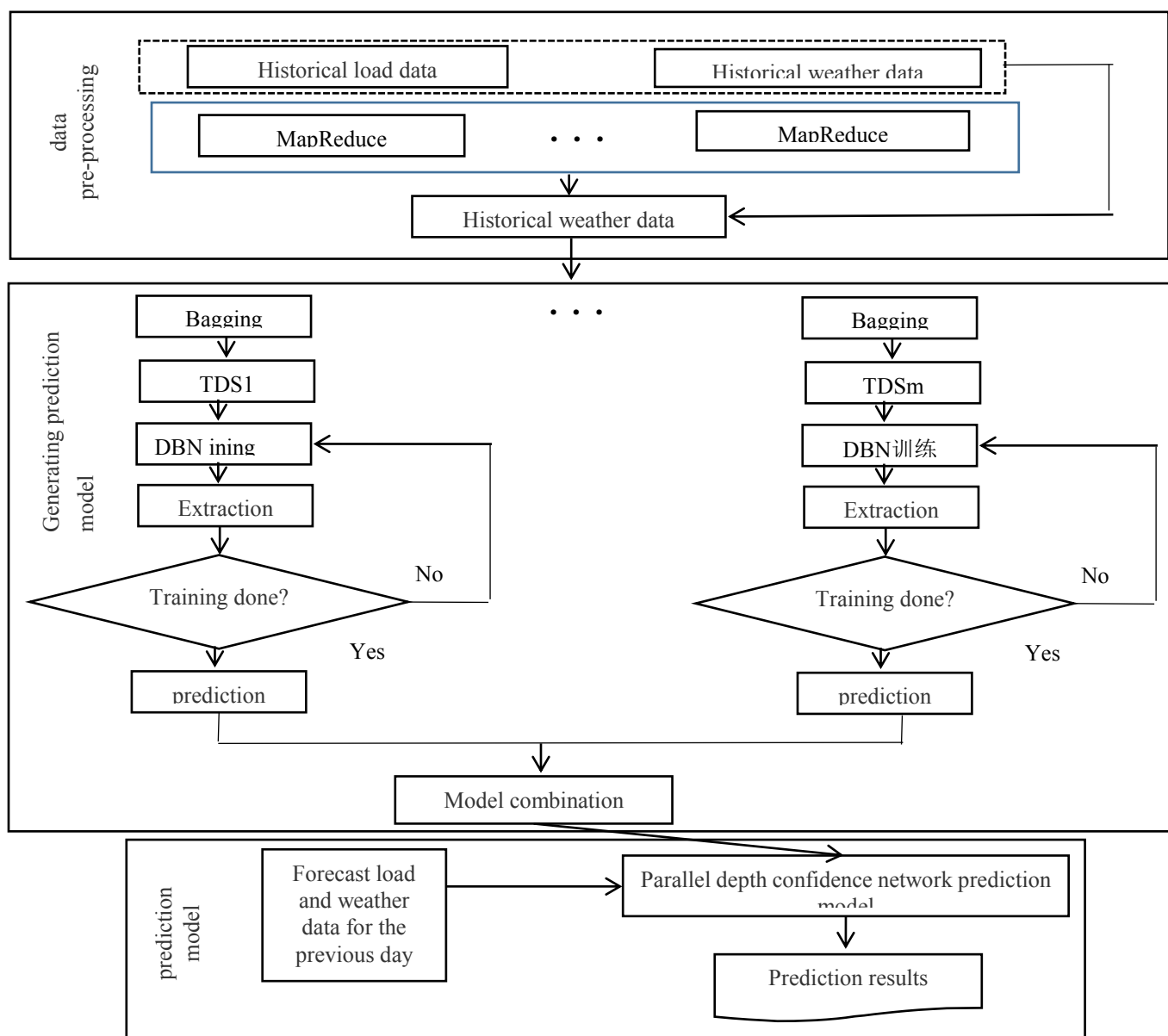


Fig. 2 parallel load forecasting flow chart based on DBN

Data pre-processing:

Data pre-processing is to remove incomplete, noisy and low quality data. For the collected data, the irrelevant and noisy data must be removed by data cleaning, and then the data from multiple data sources are integrated and stored in a consistent

data store. Finally, all kinds of input data are normalized into a form suitable for data mining. Use formula (8) to normalize the sample data [12-13]

$$r_i = \frac{x_i - \min(x_i)}{\max(x_i) - \min(x_i)} \quad (8)$$

The load data in this paper come from the data acquisition and monitoring control system, and the weather data from the weather forecast system. The load data acquisition period is 15 min. After normalization, the above data are in the range of [0, 1], and the weather data are divided into three kinds of weather conditions, which are defined as:

Weather condition={clear, cloudy, rain}

To better incorporate weather data into the DBN prediction model, define weather conditions as: clear=0, cloudy=0.5, rain=1.

Prediction model generation

Step one: Bagging algorithm is used to extract m samples randomly from the pre-processed data block, and the data set TDS1, TDS2, ... TDSm is composed by label, one TDS corresponds to one DBN training set, which can not only keep the diversity of the original data set, but also improve the training efficiency.

Step two: K prediction models are generated in parallel, one Map trains generates one prediction model, and K prediction models are generated to obtain K group parameters $\theta = \{W, a, b\}$, and the final values are taken as the average values of the parameters.

Step three: By combining K prediction models, DBN parallel load forecasting model is formed, and deployed in Hadoop distributed cluster, finally power load forecasting is realized.

III. EXPERIMENT AND RESULT ANALYSIS

Evaluation of performance indicators

There are many factors affecting power load forecasting, and losses will inevitably occur in the process of deep learning network training, which results in some errors between the predicted value and the actual value. In order to evaluate the performance of the load forecasting method proposed in this paper, the RMS error (RMSE) with high sensitivity is used as the evaluation index.

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n \left| \frac{f_i - \hat{f}_i}{f_i} \right|^2} \times 100\% \quad (9)$$

Among them, f_i is the actual load, \hat{f}_i is the forecasting load

Experimental results and analysis

The experiment environment is composed of 5 computers with the same configuration. Each computer CPU is Intel Core i5-4590, main frequency 3.30 GHz, memory 4GB and hard disk 500GB. The experimental data are derived from the historical load information and weather information of a certain area in Guangdong Province,

covering the partial load values from 20 July to 30 July 2018 and the weather data from July 21 to July 31, 2018. The collection period of load information is 15 min. The weather information includes the highest temperature, the lowest temperature and the weather characteristic, in which the weather characteristic includes the sunny day,

the cloudy day, the rainy day denoted with 0, 0.5, 1 respectively. The above experimental data are used as model training data to predict the power load on July 31, 2018. Partial training data set as shown in Table 1 [14].

TABLE I Partial training data sets

Date	2016			2017			2018			
	7/20	...	7/30	7/20	...	7/30	7/20	...	7/30	7/31
Load	0.2461	...	0.2217	0.2531	...	0.2305	0.2123	...	0.2119	-
	0.2243	...	0.2025	0.2316	...	0.2113	0.2012	...	0.1963	-
	0.2122	...	0.1968	0.2213	...	0.2012	0.2001	...	0.1923	-
	-
	0.8148	...	0.8632	0.8863	...	0.8906	0.8019	...	0.8179	-
	0.8226	...	0.8328	0.8654	...	0.8738	0.8136	...	0.8229	-
maximum temperature	-	...	0.2426	0.2385	...	0.2321	0.2316	...	0.2317	0.2362
minimum temperature	-	...	0.3032	0.3125	...	0.3162	0.2977	...	0.2936	0.3026
Weather characteristics	-	...	0	1	...	1	0	...	0	1

The results and analysis are as follows:

The first step is to analyse the performance of parallel computing.

Due to the limited amount of experimental data, the original data set is artificially expanded to 10.24 GB, for parallel computing performance testing through a distributed cluster of five computers. One of them acts as the primary node and the other computers as data nodes. In addition, an open source java persistence plug-in, Data Nucleus,

which is easy to integrate and develop, is used to integrate the experimental data sets into the Hadoop cluster. The comparison between the parallel algorithm and the traditional algorithm is shown in fig.3. When the sample data is small (< 1000), the difference of the prediction time of the two algorithms is very small, and the computational time of the traditional algorithm increases obviously with the increase of the amount of data. The time used in the parallel computing model grows slowly, and its computational efficiency is outstanding [15-16].

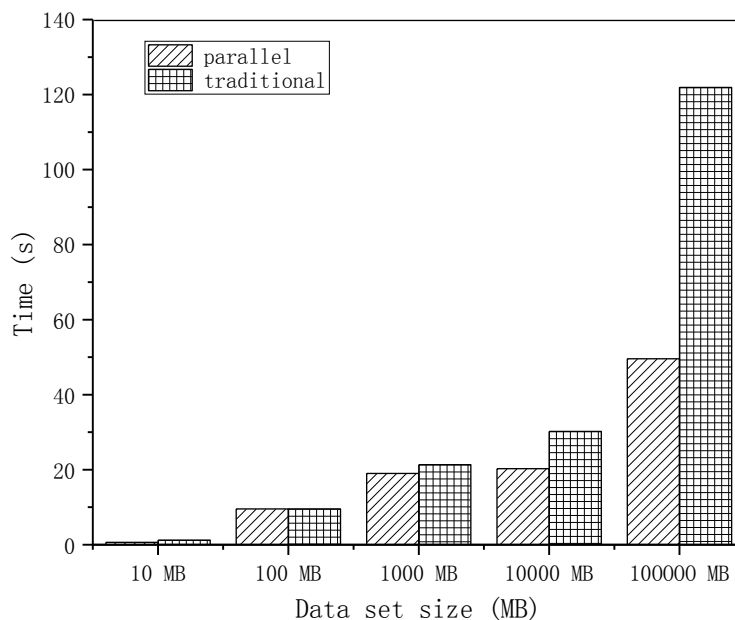


Fig. 3 Comparison of time consumption between traditional algorithm and parallel algorithm

Step two, load forecasting results

The experimental results are as shown in Table 2 and figure 4, where Table 2 lists time points, forecasted load values, actual load values and root mean square error (RMS) data. The average RMS

error between predicted load values and actual load values is only 2.86%. At the same time, it can be seen from the comparison curve between the forecasted load and the actual load in figure 4 that the predicted value is basically consistent with the actual value, and the prediction accuracy of this model is quite high [12].

TABLE II Load forecasting errors

time point	predicted value	actual value	root-mean-square error (RMSE)%
1	0.2351	0.2124	6.12
2	0.2038	0.1981	3.72
3	0.2009	0.1932	4.38
...
95	0.8292	0.8183	1.38
96	0.8155	0.8233	-0.92
average	0.5312	0.5236	2.91

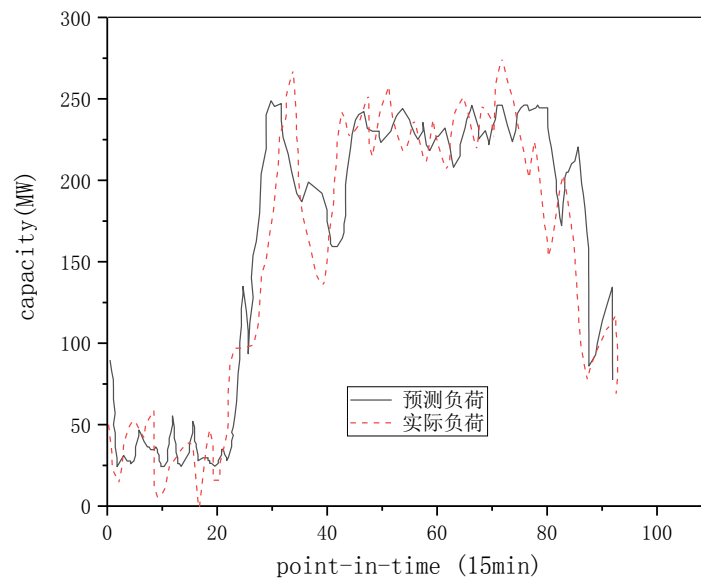


Fig. 4 Comparison of prediction results

IV. CONCLUSION

Aiming at the problems of long time consuming and poor precision in traditional load forecasting methods, this paper innovatively combines power big data technology, parallel computing framework and deep belief network algorithm, proposes a parallel load forecasting method. Based on the pre-processing data of historical load information and weather information, the DBN network model is trained in parallel, and the training model is built in Hadoop distributed cluster, which effectively improves the accuracy of load forecasting. At the same time, the prediction time is greatly reduced by parallel computation. With the development of micro-grid, the grid connection of renewable energy, such as photovoltaic power generation, wind power generation and so on, makes it more difficult to predict the load change, which puts forward higher requirements for load forecasting methods. For this research, the future development direction and prospects are to improve the forecasting methods with the prediction of PV generation, wind power generation and load forecasting after grid-connected and apply deep learning algorithm to other problems in power system.

ACKNOWLEDGMENT

Science Foundations of Nanjing Institute of Technology (No. YKJ201421, No. JCYJ201615, No. YKJ201420, No. YKJ201423)

REFERENCES

- [1] Chen, L., Wang, Z., Wang, G. "Application of LSTM Networks in Short-Term Power Load Forecasting Under the Deep Learning Framework". *Electric Power Information & Communication Technology*, 2017, 1(2), 20-24.
- [2] Liu, D., Rui, G., Qiang, D., et al. "The application of target tracking method in power load analysis based on computer vision in complex scenes". *Cluster Computing*, 2018, (11), 1-9.
- [3] Khalaf MM. "Algorithms And Optimal Choice For Power Plants Based On M-Polar Fuzzy Soft Set Decision Making Criteria". *Acta Electronica Malaysia*, 2020, 4(1), 11-23.
- [4] Litjens, G., Kooi, T., Bejnordi, B.E., et al. "A survey on deep learning in medical image analysis". *Medical Image Analysis*, 2017, 42(9):60-88.
- [5] Xue, Y., Chen, S., Qin, J., et al. "Application of Deep Learning in Automated Analysis of Molecular Images in Cancer: A Survey". *Contrast Media Mol Imaging*, 2017, 2017(6), 1-10.
- [6] Alireza M. H, Farhad S. S. "Effects Of Welding Parameters, Time Interval And Preheating On Residual Stress And Distortion Of Joining St52 Stiffener Ring In An Aisi 4130 Tubular Shell". *Acta Mechanica Malaysia*, 2020, 3(1), 11-15.
- [7] Aguilar JL. "Sustaining The Supply Chain Management System Of A Multi-Purpose Cooperative In Tiaong, Quezon". *Information Management and Computer Science*, 2019, 2(1), 04-09.

- [8] Yuan, R.C., Yan, H., Zhou, X.M., et al. "Application and Architecture of Power Dispatching & Distribution System Using Big Data Technology". *Advanced Materials Research*, 2015, 1070-1072, 1425-1429.
- [9] Khan, S., Yairi, T. "A review on the application of deep learning in system health management". *Mechanical Systems & Signal Processing*, 2018, 107(1), 241-265.
- [10] Li, M.H., Mestre, T.A., Fox, S.H., et al. "Automated vision-based analysis of levodopa-induced dyskinesia with deep learning". *Conf Proc IEEE Eng Med Biol Soc*, 2017, 3377-3380.
- [11] Gupta JK, Gupta SK. "IOT Based Statistical Approach For Human Crowd Density Estimation-Design And Analysis". *Acta Informatica Malaysia*, 2020, 4(1), 22-25.
- [12] Laghari, J.A., Mokhlis, H., Bakar, A.H.A., et al. "Application of computational intelligence techniques for load shedding in power systems: A review". *Energy Conversion & Management*, 2013, 75(11), 130-140.
- [13] Guha, D., Roy, P.K., Banerjee, S. "Application of backtracking search algorithm in load frequency control of multi-area interconnected power system". *Ain Shams Engineering Journal*, 2016, 9(2), 11-12.
- [14] Dong-Sheng, X.U., Yang, W., Wei, Z., et al. "Application of SPSS in characteristic of short load and its forecasting". *Power System Protection & Control*, 2009, 37(21), 147-151.
- [15] Erlano-De Torres JA. "Performance of Sophomore Information Technology Students In Assembly Language Subject Based On Different Teaching Methods". *Information Management and Computer Science*, 2019, 2(1), 10-13.
- [16] Madrazo, C.F., Cacha, I.H., Iglesias, L.L., et al. "Application of a Convolutional Neural Network for image classification to the analysis of collisions in High Energy Physics". 2017, 9(2), 11-12.
- [17] Zhang, D., Han, X., Deng, C. "Review on the research and practice of deep learning and reinforcement learning in smart grids". *CSEE Journal of Power and Energy Systems*, 2018, 4(3), 362-370.

**Creative Commons Attribution License 4.0
(Attribution 4.0 International, CC BY 4.0)**

This article is published under the terms of the Creative Commons Attribution License 4.0

https://creativecommons.org/licenses/by/4.0/deed.en_US