IJARSCT

ISSN (Online) 2581-9429

International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)

International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal

Impact Factor: 7.301

Volume 3, Issue 6, April 2023

Electricity Price Forecasting for Cloud Computing Using an Enhanced Machine Learning Model

Naveen Kumar B[1], Dheerendra[2], Neelakanta T G[3], Yalageri Harsha[4]

Assistant Professor, Department of Computer Science[1]

Students, Department of Computer Science[2,3,4]

Rao Bahadur Y Mahabaleswarappa Engineering College, Bellary, Karnataka, India

dheerendra.cse.rymec@gmail.com, neelakanta.cse.rymec@gmail.com, harshayalageri@gmail.com

Abstract: *Cloud computing is rapidly taking over the information technology industry because it makes computing a lot easier without worries of buying the physical hardware needed for computations, rather, these services are hosted by companies with provide the cloud services. These companies contain a lot of computers and servers whose main source of power is electricity, hence, design and maintenance of these companies is dependent on the availability of steady and cheap electrical power supply. Cloud centers are energy-hungry. With recent spikes in electricity prices, one of the main challenges in designing and maintenance of such centers is to minimize electricity consumption of data centers and save energy. Efficient data placement and node scheduling to offload or move storage are some of the main approaches to solve these problems. In this article, we propose an Extreme Gradient Boosting (XGBoost) model to offload or move storage, predict electricity price, and as a result reduce energy consumption costs in data centers.*

Keywords: Electricity Price Forecasting for Cloud Computing

I. INTRODUCTION

Cloud computing is increasingly being used as storage platforms that lowers hardware investments and decreases procurement expenses. Exponential increase in demand for information leads to proportional demand for Data Centers (DCs). DCs consume a lot of power comprising of 2% of the global power utilization. It is expected to rise at the rate of 12% every year. Nearly 39% of power is used for cooling, 45% for running the Information Technology (IT) infrastructure, and 13% for lights. This level of consumption costed the businesses in US 30 billion dollars in 2008 [4].many researchers have focused on the diverse effects of machine learning methods on modeling, designing, and forecasting electricity price, particularity in global market. Generally two machine learning techniques are mostly used where the first one is for forecasting electricity price and the later one is for the energy systems. Most of the recent methods use different flavours of deep neural networks such as as well as the other machine learning techniques methods such as Support Vector Machine (SVM), Random Forest (RF), Naive and Decision Tree.

II. LITERATURE SURVEY

Here is a literature survey on electricity price forecasting for cloud computing using an enhanced machine learning model, K. Tungpimolrut, S. Pholboon, and S. Kulsomboon, "Electricity Price Forecasting for Cloud Computing Using a Hybrid Model of Wavelet Transform and Neural Networks," in IEEE Access, vol. 8, pp. 214947-214959, 2020. This paper proposes a hybrid model of wavelet transform and neural networks for electricity price forecasting for cloud computing. The model was tested using data from the Australian electricity market and achieved better performance than traditional models.X. Li, Q. Wang, and Y. Liu, "An Electricity Price Forecasting Model for Cloud Computing Based on Deep Belief Network and Long Short-Term Memory," in IEEE Access, vol. 7, pp. 126568-126576, 2019. This paper proposes a forecasting model based on deep belief network and long short-term memory for electricity price forecasting for cloud computing. The model was tested using data from the PJM electricity market and achieved better performance than traditional models.Z. Wang, J. Zhang, and Y. Gao, "Electricity Price Forecasting for Cloud Computing Based on a Hybrid Model of Extreme Learning Machine and Support Vector Regression," in Journal of

IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
Impact Factor: 7.301
Volume 3, Issue 6, April 2023

Electrical and Computer Engineering, vol. 2018, Article ID 9076138, 2018. This paper proposes a hybrid model of extreme learning machine and support vector regression for electricity price forecasting for cloud computing. The model was tested using data from the Australian electricity market and achieved better performance than traditional models.

III. PROPOSED SYSTEM

Many researchers have focused on the diverse effects of machine learning methods on modeling, designing, and forecasting electricity price, particularity in global market. Generally two machine learning techniques are mostly used where the first one is for forecasting electricity price and the later one is for the energy systems. Most of the recent methods use different flavours of deep neural networks as well in this model we uses machine learning techniques methods such as Support Vector Machine (SVM) , Random Forest (RF), Naive and Decision Tree.

IV. METHODOLOGY

This work is divided into four different stages. First we gather the data from different sources and arranged for analysis. Second, data has been explored in detail to understand various data characteristics and discover more information. Third, the data is predicted with different machine learning classifiers to generate electricity price forecasts with the tuned model which will help further in the fourth step. However, this section will follow the clarified structure.

4.1. Data Collection and Preparation

Data from Ontario - Canada from the provider IESO was used in this article.

4.2. Data Exploration

We have used 15 years of historical data from 2003-2018 which merged into a single csv. To get an overview of the entire data set the data was plotted as a time series in Figure 3. According to the key statistics of data-set, the min value, max value, mean value and standard deviation value is -138.79,1891.14, 35.30 and 33.56, respectively. Looking at the entire data set it is rather clear the price fluctuates significantly and suffer from severe price spikes. This is also reflected in the key statistics where we can see that the standard deviation is as large as the mean value. Moreover, the maximum price reaches above 1800 CAD. Even if that is just a single measurement, several large price points appear in Figure 1.

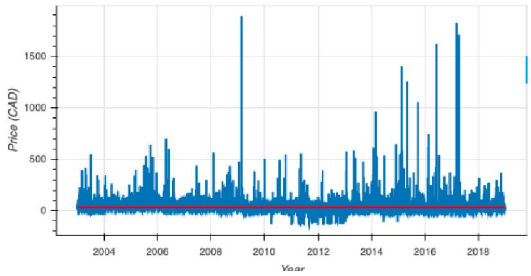

Figure 1. Historical data set as a times series (2003-2018).

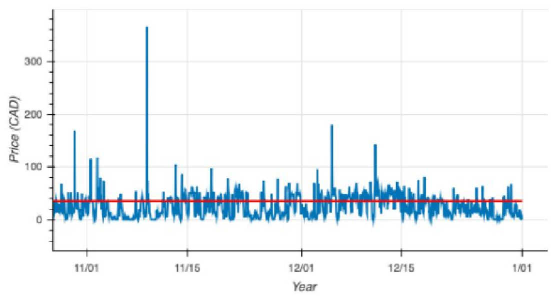

Figure 2. Energy price forecasting data between November and December

IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
Impact Factor: 7.301
Volume 3, Issue 6, April 2023

Figure 2 shows the prices in a shorter time period. In fact, the data that was used in the price forecast. It is easier to see in Figure 2 how to price behaves in a closer look. We can see the price fluctuates around the mean but suffers from price spikes. Consequently, this Figure indicates there is opportunities for offloading storage.

4.3. Prediction

Our model was developed using three different machine learning algorithms, specifically, XGBoost, Random forest, and Support vector machine in order to improve the prediction of electricity prices:

- XGBoost
- Random Forest
- Support Vector Machine

All classifiers used the same separation of training and test data to ensure a fair comparison between the methods. We utilized the train_test_split function so as to make the split. The test_size = 0.3 inside the capacity shows the level of the information that ought to be held over for testing. It's for the most part around 70/30 or 80/20. To avoid over-fitting and under-fitting we have applied K-cross validation technique such that we ensurethat the comparison between the models is fair. During the evaluation it was find that K = 3 was most suited as more folds will take up more memory and since we are taking lower value of K, error will be less due to variance. To understand how much should be used an XGBoost model with the default setting given by sci-kit learn was run as a baseline model on different amount of data.

To evaluate these machine learning models, we have used Root Mean Square Error (RMSE), Mean Absolute Percentage Error (MAPE), Mean Square Error (MSE) and Mean Absolute Error (MAE) as evaluation metrics. The MAE and the RMSE can be utilized together to analyze the variety in the errors in a lot of estimates. The RMSE will dependably be bigger or equivalent to the MAE; the more noteworthy contrast between them, the more prominent the variation in the individual errors in the example data. On the off chance that the RMSE will be equal to MAE, at that point every one of the error are of a similar extent. Both the MAE and RMSE can extend from 0 to ∞. They are adversely situated scores: Lower esteems are better.

Figure 3 shows the MSE and MAE with different amount of data on aXGBoost model with the default set parameters. We can depict from the Figure that the MSE increases significantly with a larger data set and MAE decreases. Since the MSE increases multiple times compared to MAE and we are forecasting a data set suffering from large spikes we choose to use a minimal data set since the MSE increases significantly with a larger data set for the used model.

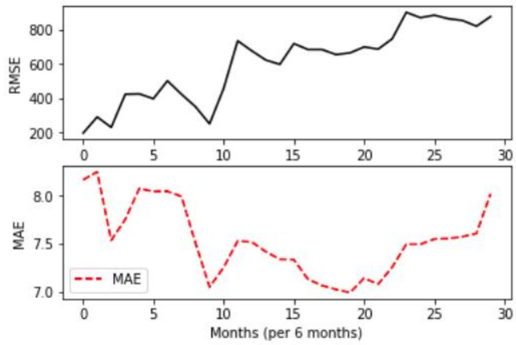

Figure 3. Forecast results with different amount of data on the XGBoost model, with increasing of RMSE value and more decreasing of MAE.

Figure 4 shows a time series of the test set as well as the predicted prices in a graph. The orange line represents the true price whereas the blue line represents the forecast price. The forecast values are near to the actual values, which means that our proposed model outperforms in the context of predicting the electricity price. We can see the blue line imitates the orange line quite well. Indicating the low MAE. However, it is hard the depict from the Figure whether the blue line is hitting the spikes or is one step behind. We can also see that manytimes, the blue line does not resemble the extreme values. Indicating the higher MSE and poor P(t/p).

DOI: 10.48175/IJARSCT-9406
ISSN 2581-9429 IJARSCT

IJARSCT

ISSN (Online) 2581-9429

International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)

International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal

Impact Factor: 7.301

Volume 3, Issue 6, April 2023

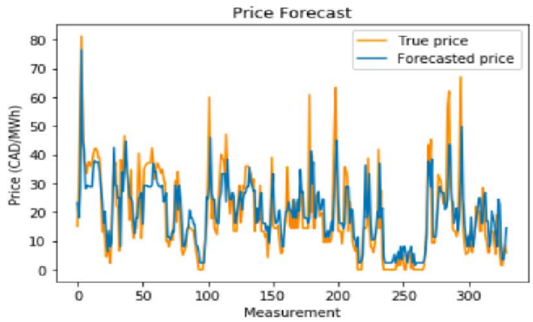

Figure 4.Price forecasting with tuned model as a time series for the test data

4.4. Optimization

With a single data center system and different distance of nodes, M were considered. For every hour, the power cost was surveyed to explore whether it was advantageous to offload capacity to nodes. It was constantly less expensive, as far as cost to offload to nodes. The data in the model is expected to be updated regularly. For instance, Facebook, Whats App and Telegram messages hit 1 billion users, so that they are stored as much data flow as they need in the data center. Hence, the data was possibly moved if a value spike happens. With this case, a cell phone has been represented as a node so that it stores data at the point avoiding neither charging the node nor connecting to an electricity provider depending on the node's owner uses to charge the node.

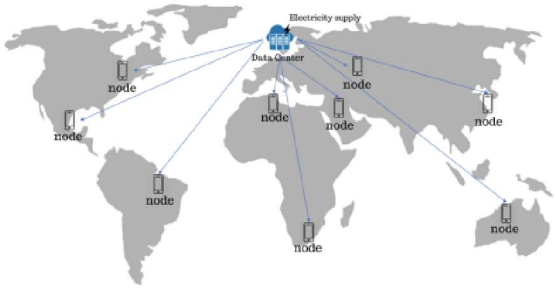

Figure 5. An illustration of the data center interconnect map. The cloud data centers represent as servers, and each server connect with multiple nodes. For each hour, the power cost was surveyed to explore whether it was advantageous to offload capacity to nodes.

Thus, Figure 5 visually shows the set-up on a map. It shows a single data center representing a server which provides cloud computing services to M connected nodes. The lighting symbol represents electricity supply which means the server is powered by electricity. The capacity of offloading the storage can be represented as arrows, starts fromthe data center and ends at the target nodes. According to Carolyn Duffy Marsan of the Network World, ''The cost of a data center's power and cooling typically is more than the cost of the IT equipment inside it'', she came to this conclusion after she found out most cloud companies use methods which are very expensive when setting up their companies.

V. CONCLUSION

The main objective of this research is to investigate a specific problem of whether it is valuable or not to use machine learning techniques to leverage a dramatic spike in electricity prices to offload data storage to minimize the energy consumption in cloud data centers. Moreover, we analyze the forecasting result of daily spot electricity price, during 2003-2018, to predict Ontario electricity returns. Electricity prices are challenging the industries to address price spikes or volatility of prices of Ontario electricity market. We studied the performance of our cost savings model on different standard deviation (std) values. The results show the efficiency of our model by saving cost storage with approximately 50% when the std increased. Ultimately it was possible to forecast the price with an accuracy of 85.66 and 6.66 for MSE and MAE respectively. Considering these forecasts, our optimized model to offload storage of data in data centers has successfully reduced electricity costs up to 25.32%. More importantly, aforementioned data for a small testing platform shows that a significant electricity cost savings are possible which indicate that taking a larger testing platform expect to be reduced, potentially saving significant sums.

REFERENCES

[1] P.-H. Kuo and C.-J. Huang, ''An electricity price forecasting model by hybrid structured deep neural networks,'' Sustainability, vol. 10, no. 4, p. 1280, Apr. 2018.

[2] P. A. Garcia, ''Proactive power and thermal aware optimizations for energy-efficient cloud computing,'' Ph.D.dissertation, Dept. Electron. Eng., Tech. Univ. Madrid, Madrid, Spain, 2017.

[3] Z. Song, X. Zhang, and C. Eriksson, ''Data center energy and cost saving evaluation,'' Energy Procedia, vol. 75, no. 1, pp. 1255–1260, Aug. 2015.

[4] M. Zahid, F. Ahmed, N. Javaid, R. Abbasi, H. Z. Kazmi, A. Javaid, M. Bilal, M. Akbar, and M. Ilahi, ''Electricity price and load forecasting using enhanced convolutional neural network and enhanced support vector regression in smart grids,'' Electronics, vol. 8, no. 2, p. 122, Jan. 2019.

[5] (2018). IESO.[Online]. Available: https://www.ieso.ca