ELECTRICITY BILL PRICE FORECASTING WITH ARIMA MODEL USING THE CONCEPT OF MACHINE LEARNING

Anjali Krishnan[1], Jinu P Sainudeen[2]
[1]MLMCE, India,anjalikrishnan2013@gmail.com
[2]MLMCE, India, jinu.sainudeen@mangalam.in

ABSTRACT

Price forecasting has always played a pivotal role in the success of every institution or life course. Electricity price forecasting is one of the key elements among price forecasting. Just like its importance, the difficulties in forecasting will always be imminent like difficulties to handle with huge price and redundancy from feature selection. Eventhough there are multiple existing systems currently available, the proposed system is far more advance as it can predict the future values in time series with ARIMA model.For feature extraction kernel principal component analysis is being used. For feature selection kernel principal component analysis is being used.

Key words: ARIMA Model, KPCA

1. INTRODUCTION

The main benefit of smart grid is to save money, manage energy consumption, energy reliability and protecting the environment. Extreme price volatility, which can be up to two orders of magnitude higher than that of any other commodity or financial asset, has forced market participants to hedge not only volume but also price risk. A power market company able to forecast the volatile wholesale prices with a reasonable level of accuracy can adjust its bidding strategy and its own production or consumption schedule in order to reduce the risk or maximize the profits in day- ahead trading. A ballpark estimate of savings from a 1% reduction in the mean absolute percentage error (MAPE) of short-term price forecasts is $300,000 per year for a utility with 1GW peak lo**ad.**

Clients are able to join in activities of smart grid, where the energy cost can be reduced by energy preservation and load shifting[1]. The electricity price is influenced by various factors, such as fuel price, electricity requirement, renewable energy supply, etc, and it varies hourly[4]. On the one hand, these unique and specific characteristics lead to price dynamics not observed in any other market, exhibiting seasonality at the daily, weekly and annual levels, and abrupt, short-lived and generally unanticipated price spikes. On the other hand, they have encouraged researchers to intensify their efforts in the development of better forecasting techniques.

Extreme price volatility, which can be up to two orders of magnitude higher than that of any other commodi**ty or** financial asset, has forced market participants to hedge not only against volume risk but also a against price movements. Price forecasts from a few hours to a few months ahead have become of particular interest to power portfolio managers. A generator, utility company or large industrial consumer who is able to forecast the volatile wholesale prices with a reasonable level of accuracy can adjust its bidding strategy and its own production or consumption schedule in order to reduce the risk or maximize the profits in day-ahead trading.

2. LITERATURE SURVEY

With respect to clients, they are really anxious to know whether the power cost surpasses the particular client characterized edges, which they used to choose to turn the heap on or on the other hand off. Under this situation, clients require the power value grouping. Henceforth, some particular edges in view of point value estimating calculations are utilized to order the power cost. Capacity estimate methods are the key of point value anticipating calculations, in which the fundamental procedure of value arrangement is imitated by a value display [2]. A hybrid feature selector based on Grey Correlation Analysis (GCA) to eliminate the feature redundancy. Second, an integration of Kernel function and Principle Component Analysis (KPCA) is used in feature extraction process to realize the dimensionality reduction[3]. The bootstrapping technology is utilized for studying uncertainty and a generalized extreme learning machine method is proposed for wavelet neural networks[5].

Selecting proper and useful features plays an essential role in improving the forecasting accuracy. multivariable mutual information is applied for feature selection to select the more relation between the output price and the candidate input features. Support vector regression (SVR) is applied as the regression system[6]. Load and price forecasts are necessary for optimal operation planning in competitive electricity markets. However, most of the load and price forecast methods suffer from lack of an efficient feature selection technique with the ability of modeling the nonlinearities and interacting features of the forecast processes[7]. Kernels are employed in Support Vector

Machines (SVM) to map the nonlinear model into a higher dimensional feature space where the linear learning is adopted. Every kernel has its advantages and disadvantages. Preferably, the 'good' characteristics of two or more kernels should be combined[8]. In order to overcome the problem that it is difficult for support vector machine to deal with uncertain information system, fuzzy theory and rough set are introduced to get two uncertain support vector machines, which are fuzzy support vector machine and fuzzy rough support vector machine respectively[9]. Real-time pricing (RTP) of electricity is one of the key technologies in smart grid. In RTP, the electricity price is set according to power consumption. Information about the price and power consumption is communicated through a communication network[10].

3. PROPOSED SYSTEM

The objective of our system is to do productive and exact gauging of power cost. To accomplish this, we have to process the crude information, make sense of the chose accompanying measurements are critical for the handling execution of our proposed structure.
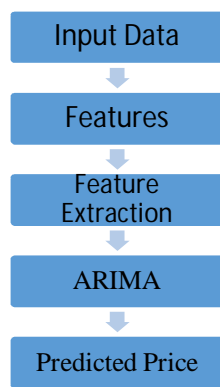

Figure 1: Proposed System

The primary issue in electricity price forecasting is accuracy. However, various factors influence the electricity, which makes the classifier training difficult. Feature selection is different from dimensionality reduction. Figure 1 is the proposed architecture. Both methods seek to reduce the number of attributes in the dataset, but a dimensionality reduction method do so by creating new combinations of attributes, where as feature selection methods include and exclude attributes present in the data without changing them. The main benefits of feature selection includes:

- **Reduces Overfitting**: Less redundant data means less opportunity to make decisions based on noise.
- **Improves Accuracy**: Less misleading data means modeling accuracy improves.
- **Reduces Training Time**: fewer data points reduce algorithm complexity and algorithms train faster.
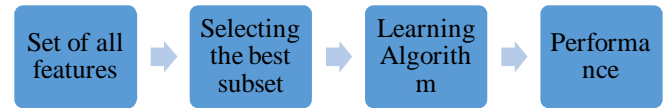

Figure 2: Feature Selection

Minimum-redundancy-maximum-relevance (mRMR) feature selection The relevance of a feature set S for the class c is defined by the average value of all mutual information values between the individual feature f_i and the class c as follows:

$$D(S,c) = \frac{1}{|S|} \sum_{f_i \in S} I(f_i, c)$$

The redundancy of all features in the set S is the average value of all mutual information values between the feature f_i and the feature f_j:

$$R(S) = \frac{1}{|S|^2} \sum_{f_i, f_j \in S} I(f_i, f_j)$$

In machine learning, pattern recognition and in image processing, feature extraction starts from an initial set of measured data and builds derived values (features) intended to be informative and non-redundant, facilitating the subsequent learning and generalization steps, and in some cases leading to better human interpretations. Feature extraction is a dimensionality reduction process, where an initial set of raw variables is reduced to more manageable groups (features) for processing, while still accurately and completely describing the original data set. Figure 2 shows the basic steps used in machine learning.

Kernel Principal Component Analysis (KPCA)
In the field of multivariate statistics, kernel principal component analysis (kernel PCA)[1] is an extension of principal component analysis (PCA) using techniques of kernel methods. Using a kernel, the originally linear operations of PCA are performed in a reproducing kernel Hilbert space.
Conventional PCA operates on zero-centered data; that is,

$$\frac{1}{N} \sum_{i=1}^{N} x_i = 0$$

where x_i is vector of one of the N multivariate observations. It operates by diagonalizing the covariance matrix.

ARIMA Model

In statistics and econometrics, and in particular in time series analysis, an autoregressive integrated moving average (ARIMA) model is a generalization of an autoregressive moving average (ARMA) model. Both of these models are fitted to time series data either to better understand the data or to predict future points in the series (forecasting). ARIMA models are applied in some cases where data show evidence of non-stationarity, where an

initial differencing step (corresponding to the "integrated" part of the model) can be applied one or more times to eliminate the non-stationarity.

The AR part of ARIMA indicates that the evolving variable of interest is regressed on its own lagged (i.e., prior) values. The MA part indicates that the regression error is actually a linear combination of error terms whose values occurred contemporaneously and at various times in the past. The I (for "integrated") indicates that the data values have been replaced with the difference between their values and the previous values (and this differencing process may have been performed more than once). The purpose of each of these features is to make the model fit the data as well as possible.

Non-seasonal ARIMA models are generally denoted ARIMA(p,d,q) where parameters p, d, and q are non-negative integers, p is the order (number of time lags) of the autoregressive model, d is the degree of differencing (the number of times the data have had past values subtracted), and q is the order of the moving-average model. Seasonal ARIMA models are usually denoted ARIMA(p,d,q)(P,D,Q)$_m$, where m refers to the number of periods in each season, and the uppercase P,D,Q refer to the autoregressive, differencing, and moving average terms for the seasonal part of the ARIMA model.

Parameters :-
endog (array-like) – The endogenous variable.

order (iterable) – The (p,d,q) order of the model for the number of AR parameters, differences, and MA parameters to use.

exog (array-like, optional) – An optional array of exogenous variables. This should not include a constant or trend. You can specify this in the fit method.

dates (array-like of datetime, optional) – An array-like object of datetime objects. If a pandas object is given for endog or exog, it is assumed to have a DateIndex.

freq (str, optional) – The frequency of the time-series. A Pandas offset or 'B', 'D', 'W', 'M', 'A', or 'Q'. This is optional if dates are given.

f exogenous variables are given, then the model that is fit is

$$\emptyset(L)\,(y_{t-}x_t\beta') \;=\; \theta(L)\varepsilon_t$$

where ϕ and θθare polynomials in the lag operator, L. This is the regression model with ARMA errors, or ARMAX model. This specification is used, whether or not the model is fit using conditional sum of square or maximum-likelihood, using the method argument in statsmodel.tsa.arime_model.ARIMA.fit. Therefore, for now, CSS and mle refer to estimation method only.

4. EXPERIMENT & RESULTS

In order to investigate the capability of our proposed framework, we develop a simulator with Python according to the system framework. In order to eliminate the redundant information within the features, KPCA is used to extract the principle components. With the consideration of the huge amounts of data, it is significant for our framework to adequately utilize computation resources and support the parallel computing. Different from traditional electricity price forecasting approaches where the date is processed sequentially, the proposed framework is easy to implement on a parallelized and distributed system.
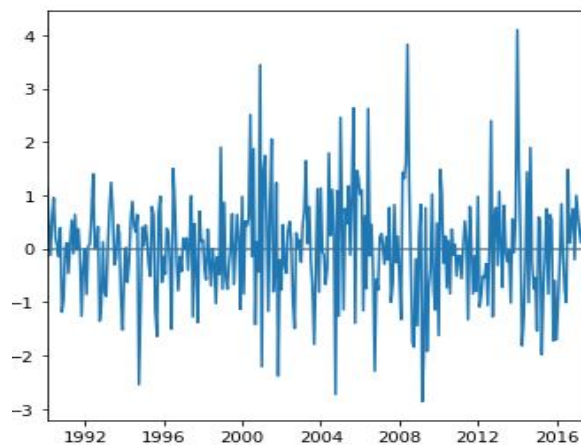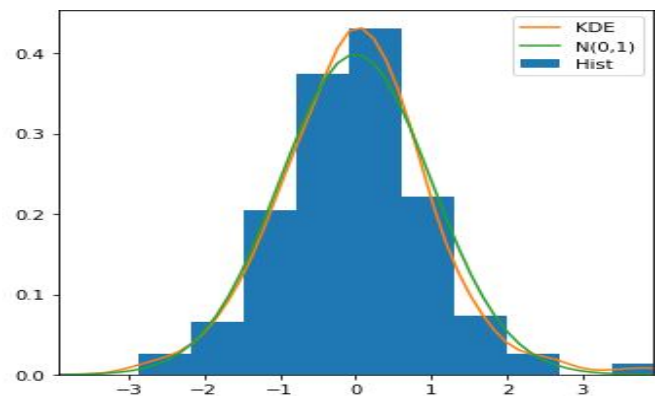

Figure 3: Standardized Residual

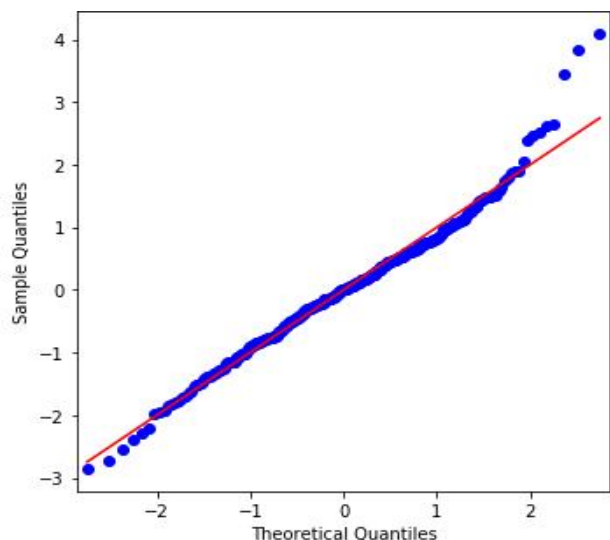**Figure 4:** Histogram plus estimated density
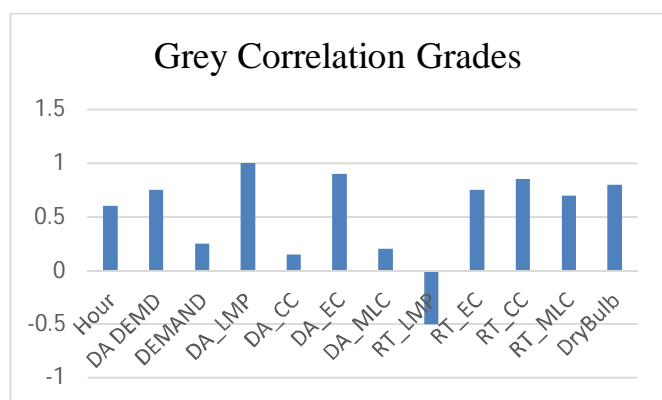
Figure 5: Normal Q-Q

Figure 6: Grey Correlation grade of each feature

Since the physical meaning of each feature in the framework is different, the dimension of data is not necessarily the same. Therefore, when the grey correlation grade analysis is carried out, the non-dimensional data processing will be executed accordingly. In this processing, each feature of data is non dimensionalized by dividing the data via their average value.

5. CONCLUSION

In this paper, we have researched the power cost estimating issue in shrewd matrix by means of joint thought of highlight building and classifier parameters modification. . Electricity price forecasting is one of the key elements among price forecasting. Just like its importance, the difficulties in forecasting will always be imminent like difficulties to handle with huge price and redundancy from feature selection. KPCA is applied to extract new features with less redundancy. The proposed system is far more advance as it can predict the future values in time series with ARIMA model. The numerical results have shown that our proposed framework is more accurate than other benchmark algorithms.

REFERENCES

1. Kun Wang, Chenhan Xu,Song Gou and Albert Y ZomayaP. **Robust bigdata analytics for electricity price forecasting in the smart grid**, *IEEE Trans. On Big Data*, Vol. 12,2017.

2. Hossam Mosbah and Mohamed El-Hawary. **Month using multilayer neural network prediction** , *Canadian Journal of Electrical & Compter Engineering*, vol.39, pp. 283-291, 2016.
https://doi.org/10.1109/CJECE.2016.2586939

3. Koichi Kobayashi. **Optimal Real-time pricing of electricity based on switched Markov chain models & welfare maximization**, *International conference on Cyber- physical systems.*, pp. 35-39, 2016.
https://doi.org/10.1109/CPSNA.2016.16

4. Abdussalam Mohamed and M.E El- Hawary. **Effective input features selection for electricity price forecasting**, *IEEE Canadian conference on electrical and computer engineering* pp. 1-5, 2016.
https://doi.org/10.1109/CCECE.2016.7726626

5. ZHI-WEIQIU.**Multivariable mutual information based feature selection for electricity price forecasting**, *International conference on machine learning and cybermetrix,* pp. 1-6,2012.
https://doi.org/10.1109/ICMLC.2012.6358906

6. Oveis Abedinia, Nima Amjady and Hamidreza Zareipour**. A new feature selection technique for load and price forecast of electrical power systems**, *IEEE transaction on power system*, pp.1-11,2016.
https://doi.org/10.1109/TPWRS.2016.2556620

7. Zhiyong Du, Xianfang wang, Liyuan Zheng and Zhulin Zheng. **Nonlinear system modeling based on KPCA and MKSVM**, *International colloquium computing, communication, control and management*, pp 1-4, 2009.

8. Hu Han, Jianwu Dang and Enen Ren**. Comparative study on two uncertain support vector machines**, *International conference on advanced computational intelligence* , pp. 1-4, 2012.
https://doi.org/10.1109/ICACI.2012.6463192

9. Harish Varshney, Avinash Sharma and Rajesh Kumar. **A hybrid approach to price forecasting incorporating exogenous variables for a day ahead electricity market**, *International conference on power electronics, intelligent control and energy system*s, pp.1-6,2016.
https://doi.org/10.1109/ICPEICES.2016.7853355

10. Mehdi Rafiei, Taher Niknam and Mohammad- Hassan Khooban. **Probabilistic Forecasting of hourly electricity price by generalization of ELM for usage in improved wavelet neural network**, *IEEE transactions on Industrial Informatics*, pp 1-8,2016.
https://doi.org/10.1109/TII.2016.2585378