Annals of Mathematics and Computer Science
ISSN: 2789-7206
Vol 17 (2023) 56-68

MACHINE LEARNING METHODS IN POWER FORECASTING: A SYSTEMATIC REVIEW

SAIDJON KAMOLOV[1]*

ABSTRACT. Forecasting power consumption is a crucial aspect of managing cities and regions. Accurate forecasts ensure smooth and uninterrupted operation of consumer and industrial units. While the traditional methods of forecasting have been useful, the advent of big data has enabled the use machine learning techniques. In this paper, we discuss the applications of machine learning to power forecasting. We describe the technical details of the existing methods and highlight their strengths and weaknesses. We find that there is no single method that fits all scenarios. The optimal method depends on various factors including the characteristics of the data, the size of the data, and the forecasting horizon.

1. INTRODUCTION

Electricity is the bedrock of the modern society. It sustains almost every part of our daily lives. Therefore, power consumption management is a key responsibility of the public and private utilities. Effective power management requires accurate forecasting of energy power consumption. It is a complex issue that involves several dozen variables with a high degree of stochasticity. While the traditional autoregressive and statistical models have been useful in forecasting power demand, they do not leverage the availability of big data which has recently emerged. The increased capacity to gather and store data allows for better analysis and deeper insight. Enter machine learning and deep learning algorithms. Machine learning leverages big data to develop accurate models that learn from data [24, 25, 26, 52, 53]. Intelligent systems have been increasingly used for power demand forecasting. Indeed, machine learning based forecasting algorithms have become more popular than the traditional models in the forecasting literature [12, 14, 27, 29]. In addition, hybrid methods that combine statistical and deep learning methods have also shown promising results.

Given the popularity and importance of power demand forecasting, the amount of literature dedicated to the topic has significantly increased over the last decade. To navigate the existing sea of literature, we propose a survey of the current research. To this end, our goal in this paper is to analyze the existing literature related to power demand forecasting using machine learning techniques. We identify the major machine

Date: Received: Jul 8, 2023; Accepted: Aug 24, 2023.
* Corresponding author.
Key words and phrases. machine learning, deep learning, forecasting, power demand, photovoltaic energy.

learning algorithms employed in power demand forecasting and discuss the technical details as well as present the achievements of the methods. In addition, we investigate the advantages and disadvantages of each approach. In the end, we provide guidelines to compare and select the optimal machine learning algorithms.

Our paper is structured as follows. In Section 2, we provide an overview of the existing machine learning methods employed for power demand forecasting. In Section 3, we delve into the details of each method. Section 4 concludes the paper.

2. Machine Learning and Power Forecasting

There exists a range of machine learning approaches used in power demand forecasting including regression, support vector machines, artificial neural networks, deep learning, and ensemble learning models. Each model has its own strengths and weaknesses. The performance of the models are not uniform and depend on the characteristics of a particular problem. While regression models perform better in scenarios with limited data availability, deep learning models excel in problems where a large amount of data is available. Therefore, a proper understanding of each method is required in order to identify the optimal approach for a given problem.

Regression models are some of the most commonly used tools in power demand forecasting. They offer a quantitative measure of the relationship between the dependent variable (power demand) and one or more independent variables (such as temperature, time of day, and so on). The two key types are linear and non-linear regression models. Linear regression models are the simplest, assuming a straight-line relationship between the variables. Non-linear regression models, on the other hand, accommodate more complex relationships and are thus generally more accurate but also computationally more intensive. Multiple regression models, which involve more than one independent variable, are often used for power demand forecasting due to the multiple factors influencing power demand.

Support Vector Machines (SVM) have been widely used in power demand forecasting due to their unique capabilities in addressing non-linear, high-dimensional, and sparse data issues. SVMs work by mapping input vectors to a higher-dimensional space where a hyperplane can be used to perform linear regression or classification. This characteristic allows SVMs to capture complex relationships in the data, making them a powerful tool in this domain.

Artificial Neural Networks (ANN) are a subset of machine learning models loosely inspired by the human brain. They are capable of learning from non-linear and complex data, making them ideal for power demand forecasting. Feed-forward neural networks (FFNNs) and radial basis function networks (RBFNs) are the two most popular types used in power demand forecasting. FFNNs are well-suited to model complex non-linear relationships, while RBFNs offer advantages in terms of speed and simplicity of training. The major drawback of ANNs is their ”black box” nature, which can make the interpretation of results challenging.

Deep learning is a subfield of machine learning that focuses on algorithms inspired by the structure and function of the brain called artificial neural networks. Deep learning was originally conceived as an extension of ANNs with multiple layers of neurons for better feature extraction. However, it has evolved into many different architectures. Specifically, Recurrent Neural Networks (RNNs) and Long Short-Term Memory (LSTM) networks have been particularly effective in power demand forecasting. RNNs have loops that allow information to be carried across nodes while processing sequential data. This gives them a crucial advantage in power demand forecasting, as they can use the historical data (previous power demands) to influence the network's forecasts. However, RNNs have difficulties handling long sequences due to the ”vanishing gradient” problem. To address the issue of the vanishing gradient, researchers developed LSTM, a type of RNN, overcomes this problem by introducing a mechanism called ”gates”. These allow the network to selectively remember or forget information, thereby effectively handling longer sequences. LSTM networks have shown exceptional performance in predicting power demand due to their ability to learn from long-term dependencies, which is a common scenario in power demand data.

Ensemble methods combine predictions from multiple machine learning algorithms to deliver superior predictive performance. The basic principle behind the ensemble approach is to leverage the different strengths of various algorithms, thereby mitigating their individual weaknesses. Bagging and boosting are two popular ensemble methods used in power demand forecasting. Bagging improves prediction accuracy by reducing the variance of the estimate, while boosting reduces bias.

Each of these machine learning techniques offers its unique advantages and challenges. The choice of method typically depends on the specifics of the power demand forecasting problem at hand, such as the nature of the data, the forecasting horizon (short, medium, or long-term), and the computational resources available. It's also worth noting that hybrid models, which combine two or more of these methods, are becoming increasingly popular, as they offer the potential to capture the strengths of multiple methods while mitigating their weaknesses. Machine learning, with its capability to handle large, complex datasets and uncover hidden patterns, offers a powerful set of tools for power demand forecasting and is set to play an increasingly important role in this area.

3. Survey of Machine Learning in Power Forecasting

3.1. **Regression.** In the field of machine learning, regression-based models are frequently used for forecasting tasks [49, 19, 20] , including power demand forecasting. These models are designed to predict a continuous outcome variable (in this case, power demand) based on one or more input features. In this subsection, we present the details of some of the commonly used regression-based models: linear regression, polynomial regression, and ridge regression.

The linear regression model predicts the dependent variable as a linear combination of the independent variables. The key equation of a simple linear regression model is:

$$Y = \beta_0 + \beta_1 X + \epsilon,$$

where: Y is the dependent variable, X is the independent variable, β_0 is the y-intercept, β_1 is the slope (effect of X on Y), and ϵ is the error term.

Linear regression was evaluated against neural networks in photovoltaic power forecasting models based on different input methods [4] A study by the researchers in [15] used linear regression to predict power demand in Spain. They used time series data of previous electricity consumption and other variables such as weather data. Their model achieved competitive results in short-term forecasting.

The polynomial regression is an extension of linear regression where the model is not only a linear function of the parameters but also a polynomial function of the independent variables. A second degree polynomial regression model can be represented as:

$$Y = \beta_0 + \beta_1 X + \beta_2 X^2 + \epsilon.$$

Polynomial regression was considered in [42], where the authors forecasted power demand in Alice Springs, one of the geographically solar energy-rich areas in Australia, with various environmental parameters. In a study by researchers in [47], polynomial regression was used for long-term load forecasting in Taiwan's electricity system. Factors such as economic indicators, population growth, and technological advancement were included in the model. The polynomial regression provided better accuracy than linear regression for long-term forecasting .

The ridge regression method used to deal with multicollinearity in multiple linear regression data. It uses a penalty term to the sum of squared coefficients in the cost function. The key equation for Ridge Regression is:

$$\min ||Y - X\beta||^2 + \lambda||\beta||^2,$$

where: λ is the tuning parameter controlling the strength of the penalty term, $||.||$ denotes the Euclidean norm, and β are the parameters to be learned.

The authors in [41] attempted to forecast intermittency of Photovoltaic (PV) energy in smart grid using ridge regression. A study by researchers in [3] applied Ridge Regression to short-term load forecasting in the Saudi electricity system. The model incorporated weather data, previous load patterns, and calendar variables. Ridge regression improved the model's performance in the presence of highly correlated input variables.

3.2. **Support Vector Machines.** Support Vector Machines (SVMs) are a class of machine learning models used for both classification and regression tasks. Their primary aim is to determine the optimal hyperplane that best separates the data in a high-dimensional feature space, making them particularly useful for complex non-linear problems.

When used for regression (Support Vector Regression, or SVR), SVM aims to fit the best line within a predefined or specified boundary called an ϵ-tube to approximate the actual data points. The key equation for Support Vector Regression can be expressed as follows:

$$\min_{\mathbf{w},b,\xi,\xi^*} \frac{1}{2}\|\mathbf{w}\|^2 + C\sum_{i=1}^{n}(\xi_i + \xi_i^*)$$

$$y_i - \mathbf{w}^\top \phi(x_i) - b \leq \varepsilon + \xi_i$$

$$\mathbf{w}^\top \phi(x_i) + b - y_i \leq \varepsilon + \xi_i^*$$

$$\xi_i, \xi_i^* \geq 0,$$

where: y_i are the target values, x_i are the input feature vectors, $\mathbf{w}$ is the normal vector to the hyperplane, b is the bias, $\phi(x_i)$ is the function mapping x_i to a higher dimensional space, ε defines the margin of tolerance where no penalty is given to errors, ξ_i and ξ_i^* are slack variables that allow for deviations outside the ϵ-tube, C is a regularization parameter which determines the trade-off between the flatness of the hyperplane and tolerance to deviations larger than ϵ.

SVMs have been used extensively in the literature. Photovoltaic power forecasting based on a support vector machine with improved ant colony optimization was proposed in [46] which showed enhanced results. The authors in [58] used a single SVM model to predict daily power demand in China. Their model considered historical power consumption data, weather information, and day-of-week information. Similarly, the researchers in [7] used a standard SVM model to predict hourly electricity load in Taiwan. Their model took in input data on weather conditions, day of the week, and time of day. Their SVM model performed better than several competing forecasting models. In [39] the authors used a combination of SVM, WLS, and PSO for short-term load forecasting in Brazil. The PSO was used to optimize the parameters of the SVM. The combined model outperformed a standalone SVM model. In the study by researchers in Iran used a hybrid of wavelet transform (WT) and SVM for short-term load forecasting in Iran [5]. The WT was used to decompose the load time series into several sub-series, which were then forecasted using SVMs. The authors in [9] used Least Squares SVM to construct a time series model for short-term wind power forecasting achieving accurate results. Least Squares SVM was used in the study by [**?**]. In particular, an LS-SVM model was used for day-ahead electricity load forecasting in Belgium. Their model was fed with past load data, temperature data, and calendar variables. LS-SVM outperformed traditional ARIMA models in this study. On the other hand, Kernel SVM was used in [54], in benchmark analysis of day-ahead solar power forecasting techniques using weather predictions. Similarly, Kernel SVM was used in [18]. The researchers implemented a Kernel SVM model for forecasting half-hourly electricity demand in Australia. The predictors included past demand data, temperature data, and time variables. The Kernel SVM model was able to capture complex non-linear relationships in the data.

3.3. **Artificial Neural Networks.** Artificial Neural Networks (ANNs) are a type of machine learning model that is inspired by the biological neural networks of the human brain. ANNs consist of layers of interconnected nodes or ”neurons” that transmit and process information [22, 23]. They are capable of learning complex patterns and relationships in data, making them widely used for tasks like power demand forecasting.

The simplest form of an ANN is the Feed-Forward Neural Network (FFNN), also known as the Multi-layer Perceptron (MLP). It consists of an input layer, one or more hidden layers, and an output layer. Each node in a layer is connected to all nodes in the subsequent layer, with each connection having an associated weight. A node takes the weighted sum of its inputs, applies a bias, and then passes the result through an activation function to produce its output.

The equation for a neuron's output in an FFNN can be represented as follows:

$$y = f\left(\sum_{j=1}^{n} w_j x_j + b\right),$$

where: y is the neuron's output, f is the activation function, w_j are the weights associated with the connections to the neuron, x_j are the inputs to the neuron, and b is the bias.

Training an ANN involves adjusting the weights and biases of the network to minimize the difference between the network's predictions and the actual values. This is typically done using a method such as backpropagation and an optimization algorithm like stochastic gradient descent.

ANNs have been a popular tool for power demand forecasting. One of the active areas of research has been forecasting solar power generation [48]. Another source of clean energy - namely wind power - has also attracted considerable amount attention from forecasting reseachers [45]. MLP was used as the main model for short-term load forecasting in Brazil [16]. They utilized historical load data, temperature data, and calendar variables as inputs to the model. The authors successfully developed probabilistic wind power forecasting using radial basis function neural networks [50]. The results showed that the proposed approach significantly outperformed the benchmark models

3.4. **Deep Learning.** Deep learning is a branch of machine learning that is based on artificial neural networks (ANN) with multiple layers. These layers are designed to simulate the human brain's structure and function, making deep learning models adept at recognizing patterns in large, complex datasets. The number of layers has grown over time reaching thousands in recent models. The key feature of deep learning models is the ability to learn features that are used for classification and regression tasks [13, 28]. In power demand forecasting, deep learning models can handle the complexity and volatility of power demand data, making them effective in producing accurate forecasts [6].

One of the most popular deep learning models used in forecasting tasks is the Long Short-Term Memory (LSTM) network [21]. LSTMs are a type of Recurrent Neural Network (RNN) designed to avoid the long-term dependency problem, allowing them to remember their inputs over a long period. The traditional RNN architecture allows for past inputs to be used to calculate future inputs using the hidden state mechanism. However, due to the repeated matrix multiplication during the back-propagation process the gradient signal tends to zero resulting in the loss of the signal to update the upstream weights. LSTM was proposed to resolve the issue of vanishing gradients.

The LSTM architecture was originally developed in 1998 [17]. Each LSTM unit consists of a cell state and three gates: input, forget, and output gate. The cell state runs straight down the entire chain, with some minor linear interactions. The gates are responsible for the removal or addition of information to the cell state.

The key equations of the LSTM cell are the following:

Forget gate:

$$f_t = \sigma(W_f \cdot [h_{t-1}, x_t] + b_f)$$

Input gate:

$$i_t = \sigma(W_i \cdot [h_{t-1}, x_t] + b_i)$$

Cell state:

$$C_t = f_t * C_{t-1} + i_t \cdot \tanh(W_C \cdot [h_{t-1}, x_t] + b_C)$$

Output gate:

$$o_t = \sigma(W_o \cdot [h_{t-1}, x_t] + b_o)$$
$$h_t = o_t \cdot \tanh(C_t),$$

where: x_t is the input at time t, h_t is the output at time t, C_t is the cell state at time t, f_t, i_t, and o_t are the forget gate, input gate, and output gate, respectively, W and b are weights and biases, respectively, σ is the sigmoid function, and $\tanh$ is the hyperbolic tangent function.

A number of authors have used LSTM to forecast power demand. The researchers in [57] proposed a day-ahead PV power forecasting method based on LSTM-RNN model and time correlation modification under partial daily pattern prediction framework. A novel long term solar photovoltaic power forecasting approach was proposed in [51], where the authors proposed using LSTM with Nadam optimizer. The approach was applied to the case of India with encouraging results. A study in [37] applied LSTM for short-term residential power demand forecasting. The authors used historical power demand data, weather data, and time-related features. The LSTM model outperformed traditional machine learning models. In [55], a CNN-LSTM model was used to forecast the power demand in China. Empirical results showed that the proposed prediction model is effective, and it proves that the organic fusion of time series data and text data can effectively improve forecasting performance. Another LSTM-CNN model was proposed in [56], where the authors used the temporal features of the data for extraction first by the long-short term memory network, and then the spatial features of the data are extracted by the convolutional neural network model. The results showed that the hybrid prediction model had better prediction effect than the single prediction model.

3.5. **Convolutional Neural Network.** Convolutional neural networks (CNNs) are a type of deep learning model primarily used for image processing but also have applications in time series analysis, including power demand forecasting. CNNs are composed of an input layer, multiple hidden layers (convolutional layers, pooling layers, and fully connected layers), and an output layer.

The convolution operation in the first layer of a CNN is defined as follows:

$$z^{[l]} = w^{[l]} * a^{[l-1]} + b^{[l]}$$

$$a^{[l]} = \sigma(z^{[l]}),$$

where: $w^{[l]}$ is the kernel or filter in the l^{th} layer, $a^{[l-1]}$ is the output activations from the previous layer, $b^{[l]}$ is the bias in the l^{th} layer, $*$ represents the convolution operation, σ is the activation function, $z^{[l]}$ and $a^{[l]}$ are the pre-activation and post-activation values, respectively, in the l^{th} layer.

After convolutional layers, pooling layers are often applied which reduce the spatial size of the representation to reduce the amount of parameters and computation in the network, and hence to also control overfitting. The pooling layer operates independently on each depth slice of the input and resizes it spatially [33, 35].

CNN is popular approach for forecasting power demand especially when used in conjunction with LSTM. A hybrid method based VMD-CNN-GRU combination was used to model short-term forecasting of wind power considering spatio-temporal features [59]. Another hybrid approach based on the combination of CNN and LSTM was proposed in [40] to forecast wind power generation. The researchers in [38] proposed a hybrid model consisting of a CNN and LSTM for stable power generation forecasting. The CNN is used to classify weather conditions, while the LSTM is utilized to learn power generation patterns based on the weather conditions. Similarly, the authors in [2] utilized the CNN-LSTM architecture to forecast power demand in Rabat, Morocco achieving promising results.

3.6. **Ensemble Learning.** Ensemble learning methods involve combining multiple machine learning models to make predictions. The idea behind ensemble methods is that a group of weak learners can come together to form a strong learner, thereby increasing the accuracy of the model.

There are several types of ensemble methods, but three of the most common are bagging, boosting, and stacking. Bagging is an ensemble method that involves training the same algorithm many times by using different subsets sampled from the training data. The most common bagging algorithm is the Random Forest algorithm. In bagging, the final output decision is made by averaging the output (for regression) or by majority voting (for classification). Boosting is an ensemble method that trains many models sequentially. Each new model is adjusted according to the errors of the previous model. The final decision is determined by the weighted sum of the decisions of all models. One of the most popular boosting algorithms is the Gradient Boosting Machine (GBM). Stacking is an ensemble learning technique that combines multiple classification or regression models via a meta-classifier or a meta-regressor. The base level models are trained based on a complete training set, then the meta-model is fitted based on the outputs of the base level models as inputs. Several ensemble methods also involve feature selection [30, 31, 32, 34]

Ensemble learning methods have recently gained popularity in forecasting power demand. A calibrated ensemble of model chains was proposed by the authors in [43] for probabilistic photovoltaic power forecasting. A short-term photovoltaic power forecasting method with adaptive stochastic configuration network ensemble was proposed in [11]. Random forest was used in [10] for multivariable response surface in short-term load forecasting. The proposed method hybridized random forest model and the mean generating function model to achieve promising results. An interpretable probabilistic model for short-term solar power forecasting using natural gradient boosting was proposed in [44]. In particular, a two stage probabilistic forecasting framework able to generate highly accurate, reliable, and sharp forecasts yet offering full transparency on both the point forecasts and the prediction intervals A stacking approach was proposed in [1] for forecasting photovoltaic (PV) power generation. Concretely, a stacked ensemble algorithm (Stack-ETR) to forecast PV output power one day ahead, utilizing three machine learning (ML) algorithms, namely, random forest regressor (RFR), extreme gradient boosting (XGBoost), and adaptive boosting (AdaBoost), as base models.

4. Comparative Analysis

The choice of a machine learning algorithm for power forecasting is determined by various factors, including data characteristics, computational resources, model interpretability, and accuracy requirements. Each algorithm has its strengths and weaknesses.

TABLE 1. Sampling methods used in the study.

Algorithm	Strength	Weakness
Linear Regression	Simple to understand and explain, fast to model and predict, works well with numerical and well-prepared data.	Assumes a linear relationship between inputs and output, sensitive to outliers and may perform poorly with irrelevant features.
SVM	Effective in high-dimensional spaces, capable of modeling non-linear relationships, robust against overfitting, especially in high-dimensional space.	Can be sensitive to the choice of the kernel parameters, computationally intensive, hence less suitable for larger datasets, not easy to interpret.
Deep learning	Able to model and learn complex non-linear relationships, can automatically extract features in deep learning, scalable, and works well with large-scale datasets.	Requires large amounts of data, computationally expensive, model parameters are not easily interpretable, can overfit if not carefully regularized and the network architecture is not properly designed.
Ensemble Methods	Can achieve higher accuracy and stability by combining several models, reduces overfitting, and works well with unbalanced and missing data.	Lack of model interpretability, can be computationally expensive, careful tuning required.

It is crucial to keep in mind that there is no one-size-fits-all algorithm for power forecasting. A suitable model often depends on the specifics of the problem, including the volume and quality of the data, the computational resources available, and the specific forecasting objectives. Hence, the best practice is often to try multiple models and select the one that works best for the specific problem.

5. Conclusion

In this paper, we discussed various approaches to forecasting power consumption using machine learning methods. We described the technical details and identified strengths and weaknesses of each approach. There is no single method that consistently outperforms the rest. The optimal method depends on the specifics of the data. In case when a large amount of data is available, deep learning models may be the optimal choice, while in situations where the data is scarce, it may be more effective to apply linear regression. Ensemble approaches have recently gained popularity as they combine several methods to reduce the weaknesses on the constituent models while highlighting their strengths.

References

[1] Abdellatif, A., Mubarak, H., Ahmad, S., Ahmed, T., Shafiullah, G. M., Hammoudeh, A., ... & Gheni, H. M. (2022). Forecasting photovoltaic power generation with a stacking ensemble model. Sustainability, 14(17), 11083.

[2] Agga, A., Abbou, A., Labbadi, M., El Houm, Y., & Ali, I. H. O. (2022). CNN-LSTM: An efficient hybrid deep learning architecture for predicting short-term photovoltaic power production. Electric Power Systems Research, 208, 107908.

[3] Alfares, H. K., & Nazeeruddin, M. (2002). Electric load forecasting: literature survey and classification of methods. International journal of systems science, 33(1), 23-34.

[4] AlShafeey, Mutaz, & Csaba Csáki. ”Evaluating neural network and linear regression photovoltaic power forecasting models based on different input methods.” Energy Reports 7 (2021): 7601-7614.

[5] Amjady, N., Keynia, F., & Zareipour, H. (2010). Short-term load forecast of microgrids by a new bilevel prediction strategy. IEEE Transactions on smart grid, 1(3), 286-294.

[6] Aslam, S., Herodotou, H., Mohsin, S. M., Javaid, N., Ashraf, N., & Aslam, S. (2021). A survey on deep learning methods for power load and renewable energy forecasting in smart microgrids. Renewable and Sustainable Energy Reviews, 144, 110992.

[7] Chen, B. J., & Chang, M. W. (2004). Load forecasting using support vector machines: A study on EUNITE competition 2001. IEEE transactions on power systems, 19(4), 1821-1830.

[8] De Brabanter, K., Karsmakers, P., Ojeda, F., Alzate, C., De Brabanter, J., Pelckmans, K., ... & Suykens, J. A. (2010). LS-SVMlab toolbox user's guide: version 1.7. Katholieke Universiteit Leuven.

[9] Ding, M., Zhou, H., Xie, H., Wu, M., Liu, K. Z., Nakanishi, Y., & Yokoyama, R. (2021). A time series model based on hybrid-kernel least-squares support vector machine for short-term wind power forecasting. ISA transactions, 108, 58-68.

[10] Fan, G. F., Zhang, L. Z., Yu, M., Hong, W. C., & Dong, S. Q. (2022). Applications of random forest in multivariable response surface for short-term load forecasting. International Journal of Electrical Power & Energy Systems, 139, 108073.

[11] Guo, X., Wang, X., Ao, Y., Dai, W., & Gao, Y. (2022). Short-term photovoltaic power forecasting with adaptive stochastic configuration network ensemble. Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery, 12(6), e1477.

[12] Gurrib, I., & Kamalov, F. (2019). The implementation of an adjusted relative strength index model in foreign currency and energy markets of emerging and developed economies. Macroeconomics and Finance in Emerging Market Economies, 12(2), 105-123.

[13] Gurrib, I., & Kamalov, F. (2022). Predicting bitcoin price movements using sentiment analysis: a machine learning approach. Studies in Economics and Finance, 39(3), 347-364.

[14] Gurrib, I., Kamalov, F., & Smail, L. (2021, August). Bitcoin price forecasting: Linear discriminant analysis with sentiment evaluation. In The 7th Annual International Conference on Arab Women in Computing in Conjunction with the 2nd Forum of Women in Research (pp. 1-5).
[15] Hernandez, L., Baladron, C., Aguiar, J. M., Carro, B., Sanchez-Esguevillas, A. J., Lloret, J., & Massana, J. (2014). A survey on electric power demand forecasting: future trends in smart grids, microgrids and smart buildings. IEEE Communications Surveys & Tutorials, 16(3), 1460-1495.
[16] Hippert, H. S., Pedreira, C. E., & Souza, R. C. (2001). Neural networks for short-term load forecasting: A review and evaluation. IEEE Transactions on power systems, 16(1), 44-55.
[17] Hochreiter, S., & Schmidhuber, J. (1997). Long short-term memory. Neural computation, 9(8), 1735-1780.
[18] Hyndman, R. J., & Fan, S. (2009). Density forecasting for long-term peak electricity demand. IEEE Transactions on Power Systems, 25(2), 1142-1153.
[19] Kamalov, F., Gurrib, I., & Thabtah, F. (2021, September). Autoregressive and neural network models: a comparative study with linearly lagged series. In 2021 International Conference on Innovation and Intelligence for Informatics, Computing, and Technologies (3ICT) (pp. 175-180). IEEE.
[20] Kamalov, F., Gurrib, I., Moussa, S., & Nazir, A. (2022, August). A Comparative Study of Autoregressive and Neural Network Models: Forecasting the GARCH Process. In International Conference on Intelligent Computing (pp. 589-603). Cham: Springer International Publishing.
[21] Kamalov, F. (2020). Forecasting significant stock price changes using neural networks. Neural Computing and Applications, 32, 17655-17667.
[22] Kamalov, F., Nazir, A., Safaraliev, M., Cherukuri, A. K., & Zgheib, R. (2021, November). Comparative analysis of activation functions in neural networks. In 2021 28th IEEE International Conference on Electronics, Circuits, and Systems (ICECS) (pp. 1-6). IEEE.
[23] Kamalov, F., & Elnagar, A. (2021, October). Kernel density estimation-based sampling for neural network classification. In 2021 International Symposium on Networks, Computers and Communications (ISNCC) (pp. 1-4). IEEE.
[24] Kamalov, F. (2020). Kernel density estimation based sampling for imbalanced class distribution. Information Sciences, 512, 1192-1201.
[25] Kamalov, F., Sulieman, H., & Santandreu Calonge, D. (2021). Machine learning based approach to exam cheating detection. Plos one, 16(8), e0254340.
[26] Kamalov, F., & Denisov, D. (2020). Gamma distribution-based sampling for imbalanced data. Knowledge-Based Systems, 207, 106368.
[27] Kamalov, F., Gurrib, I., & Rajab, K. (2021). Financial Forecasting with Machine Learning: Price Vs Return. Journal of Computer Science, 17(3), 251-264.
[28] Kamalov, F., Smail, L., & Gurrib, I. (2020, November). Stock price forecast with deep learning. In 2020 International Conference on Decision Aid Sciences and Application (DASA) (pp. 1098-1102). IEEE.
[29] Kamalov, F., Smail, L., & Gurrib, I. (2020, December). Forecasting with deep learning: S&P 500 index. In 2020 13th International Symposium on Computational Intelligence and Design (ISCID) (pp. 422-425). IEEE.
[30] Kamalov, F., Moussa, S., Zgheib, R., & Mashaal, O. (2020, December). Feature selection for intrusion detection systems. In 2020 13th international symposium on computational intelligence and design (ISCID) (pp. 265-269). IEEE.
[31] Kamalov, F., Thabtah, F., & Leung, H. H. (2022). Feature selection in imbalanced data. Annals of Data Science, 1-15.
[32] Kamalov, F. (2021). Orthogonal variance decomposition based feature selection. Expert Systems with Applications, 182, 115191.
[33] Kamalov, F., Rajab, K., Cherukuri, A., Elnagar, A., & Safaraliev, M. (2022). Deep Learning for Covid-19 Forecasting: state-of-the-art review. Neurocomputing.
[34] Kamalov, F., Moussa, S., El Khatib, Z., & Mnaouer, A. B. (2021, October). Orthogonal variance-based feature selection for intrusion detection systems. In 2021 International Symposium on Networks, Computers and Communications (ISNCC) (pp. 1-5). IEEE.

[35] Kamalov, F., & Gurrib, I. (2022). Machine learning-based forecasting of significant daily returns in foreign exchange markets. International Journal of Business Intelligence and Data Mining, 21(4), 465-483.

[36] Kamalov, F., & Leung, H. H. (2020). Outlier detection in high dimensional data. Journal of Information & Knowledge Management, 19(01), 2040013.

[37] Kong, W., Dong, Z. Y., Jia, Y., Hill, D. J., Xu, Y., & Zhang, Y. (2017). Short-term residential load forecasting based on LSTM recurrent neural network. IEEE transactions on smart grid, 10(1), 841-851.

[38] Lim, S. C., Huh, J. H., Hong, S. H., Park, C. Y., & Kim, J. C. (2022). Solar Power Forecasting Using CNN-LSTM Hybrid Model. Energies, 15(21), 8233.

[39] Lu, H., Azimi, M., & Iseley, T. (2019). Short-term load forecasting of urban gas using a hybrid model based on improved fruit fly optimization algorithm and support vector machine. Energy Reports, 5, 666-677.

[40] Malakouti, S. M., Ghiasi, A. R., Ghavifekr, A. A., & Emami, P. (2022). Predicting wind power generation using machine learning and CNN-LSTM approaches. Wind Engineering, 46(6), 1853-1869. Chicago

[41] Massaoudi, M., Refaat, S. S., Abu-Rub, H., Chihi, I., & Wesleti, F. S. (2020, July). A hybrid Bayesian ridge regression-CWT-catboost model for PV power forecasting. In 2020 IEEE kansas power and energy conference (KPEC) (pp. 1-5). IEEE.

[42] Mahmud, K., Azam, S., Karim, A., Zobaed, S., Shanmugam, B., & Mathur, D. (2021). Machine learning based PV power generation forecasting in alice springs. IEEE Access, 9, 46117-46128.

[43] Mayer, M. J., & Yang, D. (2022). Probabilistic photovoltaic power forecasting using a calibrated ensemble of model chains. Renewable and Sustainable Energy Reviews, 168, 112821.

[44] Mitrentsis, G., & Lens, H. (2022). An interpretable probabilistic model for short-term solar power forecasting using natural gradient boosting. Applied Energy, 309, 118473.

[45] Nazir, M. S., Alturise, F., Alshmrany, S., Nazir, H. M. J., Bilal, M., Abdalla, A. N., ... & M. Ali, Z. (2020). Wind generation forecasting methods and proliferation of artificial neural network: A review of five years research trend. Sustainability, 12(9), 3778.

[46] Pan, M., Li, C., Gao, R., Huang, Y., You, H., Gu, T., & Qin, F. (2020). Photovoltaic power forecasting based on a support vector machine with improved ant colony optimization. Journal of Cleaner Production, 277, 123948.

[47] Pao, H. T., & Fu, H. C. (2013). Renewable energy, non-renewable energy and economic growth in Brazil. Renewable and Sustainable Energy Reviews, 25, 381-392.

[48] Pazikadin, A. R., Rifai, D., Ali, K., Malik, M. Z., Abdalla, A. N., & Faraj, M. A. (2020). Solar irradiance measurement instrumentation and power solar generation forecasting based on Artificial Neural Networks (ANN): A review of five years research trend. Science of The Total Environment, 715, 136848.

[49] Rajab, K., Kamalov, F., & Cherukuri, A. K. (2022). Forecasting COVID-19: vector autoregression-based model. Arabian journal for science and engineering, 1-10.

[50] Sideratos, G., & Hatziargyriou, N. D. (2012). Probabilistic wind power forecasting using radial basis function neural networks. IEEE Transactions on Power Systems, 27(4), 1788-1796.

[51] Sharma, J., Soni, S., Paliwal, P., Saboor, S., Chaurasiya, P. K., Sharifpur, M., ... & Afzal, A. (2022). A novel long term solar photovoltaic power forecasting approach using LSTM with Nadam optimizer: A case study of India. Energy Science & Engineering, 10(8), 2909-2929.

[52] Thabtah, F., Kamalov, F., Hammoud, S., & Shahamiri, S. R. (2020). Least Loss: A simplified filter method for feature selection. Information Sciences, 534, 1-15.

[53] Thabtah, F., & Kamalov, F. (2017). Phishing detection: a case analysis on classifiers with rules using machine learning. Journal of Information & Knowledge Management, 16(04), 1750034.

[54] Visser, L., AlSkaif, T., & Van Sark, W. (2019, June). Benchmark analysis of day-ahead solar power forecasting techniques using weather predictions. In 2019 IEEE 46th Photovoltaic Specialists Conference (PVSC) (pp. 2111-2116). IEEE.

[55] Wang, D., Gan, J., Mao, J., Chen, F., & Yu, L. (2023). Forecasting power demand in China with a CNN-LSTM model including multimodal information. Energy, 263, 126012.

[56] Wang, K., Qi, X., & Liu, H. (2019). Photovoltaic power forecasting based LSTM-Convolutional Network. Energy, 189, 116225.

[57] Wang, F., Xuan, Z., Zhen, Z., Li, K., Wang, T., & Shi, M. (2020). A day-ahead PV power forecasting method based on LSTM-RNN model and time correlation modification under partial daily pattern prediction framework. Energy Conversion and Management, 212, 112766.

[58] Zhang, W., Liu, F., Zheng, X., & Li, Y. (2015, November). A hybrid EMD-SVM based short-term wind power forecasting model. In 2015 IEEE PES Asia-Pacific Power and Energy Engineering Conference (APPEEC) (pp. 1-5). IEEE. Chicago

[59] Zhao, Z., Yun, S., Jia, L., Guo, J., Meng, Y., He, N., ... & Yang, L. (2023). Hybrid VMD-CNN-GRU-based model for short-term forecasting of wind power considering spatio-temporal features. Engineering Applications of Artificial Intelligence, 121, 105982.

[1]FACULTY OF ENGINEERING, TAJIK TECHNICAL UNIVERSITY, DUSHANBE, TAJIKISTAN.
Email address: said.kamolov@ttu.tj