Advances and Evaluation of Intelligent Techniques in Short-Term Load Forecasting

Jamshaid Iqbal Janjua[1, 2, *], Tahir Abbas[3]
[1] Al-Khawarizmi Institute of Computer Science, University of Engineering & Technology (UET), Lahore, Pakistan
[2] School of Computer Sciences, National College of Business Administration & Economics, Lahore, Pakistan 3Department of
[3] Computer Sciences, TIMES Institute, Multan, Pakistan jamshaid.janjua@kics.edu.pk
*Corresponding Author:

***Abstract*—In order to maintain the stability of the power grid, various load forecasting methods have emerged in endlessly. However, due to different characteristics such as algorithm generalization capabilities and model complexity, their applicability to load forecasting varies. This article discusses short-term power load forecasting in the past five years. In summary, this paper summarizes the various dimensions such as experimental data sets, data preprocessing, prediction algorithms, optimization models, and evaluation methods of current research status in power load forecasting. In addition to the advantages, disadvantages, and applicability of various prediction algorithms, this paper also sums up and prospects the development trends of the short-term power load forecasting system, so as to provide a reference for the selection of power system load forecasting models in the future.**

Keywords—Short-Term Load Forecasting, Deep Learning Combination Model, Long Short-Term Memory Network, Machine Learning.

I. INTRODUCTION

The contemporary electrical grid is an intricate and ever-changing system that necessitates sophisticated management systems to guarantee stability and efficiency. With the increasing worldwide population and surging demand for electricity, there is a growing necessity for better energy management mechanisms. Machine Learning (ML) is a highly effective method for tackling these difficulties, especially in the field of Short-Term Load Forecasting (STLF), which is crucial for the day-to-day functioning of power systems [1]. Short-term load forecasting (STLF) is a critical undertaking that entails projecting the power consumption within a limited timeframe, usually spanning from a few hours to a week. Precise predictions allow power firms to optimize energy management, scheduling, and assure a consistent energy supply by efficiently aligning load and supply [2]. Precise load forecasting is crucial in incorporating renewable energy sources into the power grid. This is because it enables demand response applications and can result in decreased energy costs in manufacturing [3].

Current research has concentrated on evaluating different machine learning algorithms to ascertain their efficacy in short-term load forecasting. Various algorithms, including logistic regression, support vector machines, naive Bayes, decision tree classifiers, K-nearest neighbor, and neural networks, have been utilized and evaluated for their accuracy and forecast error [4]. Advanced iterations of these algorithms, such as the Enhanced Decision Tree Classifier, have been created to optimize control variables and attain superior forecast outcomes.

Furthermore, there have been proposals for innovative methods that integrate various machine learning techniques, such as the parallel LSTM-CNN Network (PLCNet). These hybrid models utilize the advantages of extended short-term memory networks and convolutional neural networks to enhance the accuracy of predicting [5].Additional cutting-edge techniques involve the utilization of multimodal evolutionary algorithms and ensemble learning models to optimize the forecasting process and improve accuracy [6].

Researchers have investigated the use of clustering approaches, such as K-Means and Fuzzy C-Means, with neural networks to enhance the efficiency and precision of Short-Term Load Forecasting (STLF) [1]. In addition, ensemble approaches that integrate wavelet-based schemes, hybrid learning algorithms, and feature selection have demonstrated promising outcomes in improving forecasting accuracy [7]. However, the task of precisely predicting high demand periods and effectively handling the uncertainties related to external causes still persists, notwithstanding these progressions. In response to this issue, scholars have suggested fusion forecasting methodologies and data preprocessing techniques that combine the strengths of different machine learning algorithms [8]. In addition, techniques for managing missing data, a prevalent problem in power load data analysis, have been created to enhance the precision of ultra-short-term load prediction for industrial power consumers [9].

The field of STLF is witnessing rapid advancements with the application of ML techniques. This paper aims to evaluate the efficiency and accuracy of various ML models in STLF, providing insights into their comparative performance and potential for practical implementation in smart grid systems.

Load forecasting not only occupies a dominant position in comprehensive energy management and energy dispatch, but is also particularly important for electricity price setting and stable operation of the power grid [10]. As early as the 1990s, experts and scholars have conducted research on load

Received: 05- 09- 2024
Revised: 30-12- 2024
Published: 31- 12- 2024

This article is freely accessible under the Creative Commons Attribution License, allowing for its use, distribution, and reproduction in any format, as long as the original work is correctly cited. © 2024 The Authors.

forecasting, such as regression analysis traditional load forecasting algorithms such as method and time series method [11] , but the prediction ability of such models only performs well for linear models, while user electricity load has typical nonlinear characteristics. Especially in recent years, with the widespread popularization of intelligent equipment, electricity consumption has increased significantly. More and more factors affecting power load, such as weather, economy, population, geographical location, etc., are captured, and the difficulty of load forecasting has also increased. In order to comprehensively consider multiple influencing factors, intelligent algorithms such as ANN [12], SVM , and ensemble learning [13] have begun to be widely used in load forecasting. For example respectively adopted automatic load forecasting. The power load prediction is enhanced by utilizing a fuzzy inference system that adjusts to the network, a multi-layer perceptron neural network, and the seasonal autoregressive comprehensive moving average approach. The intelligent prediction algorithm offers a broader variety of applications and significantly higher accuracy compared to the classic prediction algorithm.

II. ELECTRIC POWER LOAD FORECASTING

Electrical Load to refer to the electric power borne by the power supply area or power grid at a certain moment. According to the different user structures, electric load can be divided into industrial load, commercial load, urban civil load, rural load and other loads. Among them, urban civil load It mainly refers to the electricity load of urban residents' household appliances, which is closely related to residents' daily life and is the current main research field.

Load forecasting involves comprehensive consideration of natural elements, societal effects, system operation, and other situations. It utilizes various mathematical approaches or intelligent models to anticipate the load value of users within a specific future time period. Load forecasting can be categorized into different time frames based on the study objective and prediction duration. These categories include ultra-short-term load forecasting, short-term load forecasting, medium-term load forecasting, and long-term load forecasting [14]. Precise short-term load forecasting is highly important for developing power generation strategies, scheduling unit operations such as start-up, shutdown, and maintenance, and determining electricity pricing.

Factors Affecting Load Forecasting Electricity is the basic energy source that maintains users' daily lives and is easily affected by many factors such as weather, physical characteristics of buildings, user behavior, population density, etc. This article divides the influencing factors of users' electricity load into regional weather factors, economic factors, building. There are four major categories: structure and population density.

(a) Weather factors mainly include indoor and outdoor temperature, air humidity, sun exposure angle, wind speed, etc.; weather is one of the main factors affecting user load, especially when the temperature rises in summer, the power consumption of high- power appliances such as air conditioning and refrigeration raised dramatically.

(b) Humanistic characteristics refer to the density of residents, living customs and characteristics of residents in a certain area. Regional customs and habits, residents' occupational characteristics and the age composition of the family population will have a certain impact on users' electricity consumption habits and load levels.

(c) The building structure includes factors such as building materials, greening planning, and location of residents; factors such as building height and building location affect the user's electricity load by affecting solar radiation and ventilation effects, while building materials will affect the cooling and heating of the room. In addition, indoor equipment conditions will also have a greater impact on users' electricity load.

Economic factors refer to household income, urban energy ratio and price, regional living standards and education levels, etc.; household economic level plays a decisive role in residents' electricity consumption behavior and equipment ownership, thereby affecting users' electricity load.

III. LOAD FORECASTING ALGORITHMS

Currently, load forecasting can be categorized into standard forecasting and intelligent forecasting models that rely on deep learning. This article provides a concise introduction to regression analysis, artificial neural networks (ANN), support vector machines (SVM), convolutional neural networks (CNN), recurrent neural networks (RNN), and other forecasting methods. It also describes the advantages, weaknesses, and scope of application for each forecasting method. The goal is to offer a reference for future load forecasting.

3.1 Regression Analysis Method Regression analysis involves creating a regression equation to forecast the future changes in the dependent variable by analyzing both the dependent and independent variables. The construction of the model is straightforward and the forecast speed is rapid. Nevertheless, due to the swift advancement of science and technology, there is an increasing amount of valuable information that may be gathered. The amount of data on factors that influence electrical load is growing steadily, and regression analysis algorithms frequently overlook the inherent patterns in load fluctuations, leading to a decrease in prediction accuracy. Currently, regression analysis is increasingly being employed as a fundamental paradigm for addressing load forecasting issues. For example, based on the basic regression model, assumed variables are introduced to describe the periodic change pattern of time, and instantaneous temperature, lagging temperature and 24-hour average temperature are also considered to achieve user load prediction.

3.2 Time Series Method

The time series method was initially suggested by American researchers [15]. This approach formulates a mathematical framework to depict the correlation between time and load levels using historical time series data. [16] Employed auto regressive and moving average models to replicate various seasonal cycles (daily, weekly, quarterly, and yearly) offline data. These models accurately captured both seasonal and non-seasonal load cycles, resulting in significant reduction. The margin of error in the prediction is decreased. The time series analysis algorithm exclusively takes into account time variables, necessitates a minimal amount of data, and exhibits a rapid prediction speed. Nevertheless, the model is theoretically intricate, demands a high level of stationarity in the original data, and fails to account for other sources of uncertainty. The ultimate forecast indicates a significant discrepancy in accuracy.

Statistical load forecasting models encompass many techniques such as regression analysis, time series approaches, gray model comparable day method, and load derivation method. The gray model is more adept at managing factors of uncertainty. The method is primarily applicable for long-term load forecasting. The similar day approach partially addresses the issue of "accumulated temperature effect" in load forecasting, but it heavily depends on historical data. On the other hand, load derivation demonstrates superior performance for ultra-short-term load forecasting, but the calculation involved is more complex. The level of complexity is greater.

3.3 Artificial Neural Network ANN has gained popularity as a result of its strong capacity for nonlinear mapping and its adaptable network structure [17]. Currently, there is no universally accepted standard for determining the number of neurons and network layers in an artificial neural network (ANN). Additionally, ANN still faces challenges such as sluggish learning convergence speed and the tendency to choose local optimal solutions. Numerous professionals and scholars have created enhancements to these issues. The literature use the bat method for optimizing the selection of model parameters. It also devises a gray correlation model based on load curve form and distance similarity in order to make the selection. On days with similar load, the root mean square error (RMSE) is decreased when comparing it to the single artificial neural network (ANN) and other models.

Users exhibit significant variations in power demand as a result of diverse geographical locations, regional economies, and other factors. Methods such as principal component analysis and Pearson correlation coefficient exhibit a certain level of subjectivity. In order to tackle this issue, the literature employs the least squares approach and the minimum. Models such as variance and genetic algorithm assess the weight of each factor and employ a weighted approach to determine the importance ranking of the influencing factors. This ranking is then combined with the radial basis function model to make predictions. The study in literature [18] introduces a novel framework for generating random embedded distributions and employs the kernel density estimation method for estimating purposes. The prediction results of multiple neural networks are aggregated to create a probability density function. The final prediction value is then calculated using the expectation estimation method, which is appropriate for situations with limited data. The determination of the BPNN structure and the number of hidden layer nodes requires iterative experimentation. Literature [19] Employ the resilient CIM measurement technique as the loss function of ANN in order to mitigate the influence of noise and outliers.

3.4 Support Vector Machine

SVM [20] was first mainly used for data classification. Due to its good nonlinear data processing capabilities, it can also be used to handle forecast loading problems. Compared with BPNN, SVM has a fast convergence speed and no network layers and local optimality. However, for complex structural data, the calculation is difficult and the model is difficult to implement. Used decision trees and weighted average methods to divide seasonal attributes, and used time series and SVM support vector machines to perform analysis on different variables showing better prediction effect. Literature [21] uses the least squares support vector machine for load forecasting, and combines it with the improved parallel particle swarm optimization algorithm to optimize parameter selection, which effectively improves the data processing efficiency and algorithm prediction accuracy. proposed a linear asymmetric loss function oriented to prediction cost, which implements different punishment measures for over- or under-prediction, and at the same time fully optimizes the insensitive parameters of SVM based, which reduces greatly cost of load forecasting. Literature [22, 23] combine empirical wavelet transform, variational mode decomposition and SVM based model respectively, decompose the data in the data preprocessing stage, and merge the prediction results of all sub-sequences to obtain more satisfactory results.

The pattern search algorithm exhibits superior performance in local search, whereas the global swarm optimization algorithm possesses the ability to conduct global search. The literature provides a comprehensive algorithm that combines the advantages of these two techniques and utilizes multilayer granularity to enhance the training efficiency of SVM. The model has great training efficiency, but, its construction is characterized by complexity and difficulty. The literature [24] presents a method that incorporates basic load forecasting to determine real-time power costs. This method utilizes weighted gray correlation projection. The program forecasts the demand during holidays, while the PSO algorithm fine-tunes the parameters of the SVM model to optimize the prediction of the base loads. This paradigm holds greater practical value.

Many conventional load forecasting models are significantly influenced by previous data and do not take into account complete parameters, leading to a low level of accuracy in their predictions. The emergence of deep learning forecasting models has been facilitated by the advancement of big data. Deep learning models have arisen as a result of their potent ability to generalize. Features such as the capability for unsupervised learning are extensively utilized in domains like image recognition and natural language processing [25]. This research employs convolutional neural networks (CNN), long short-term memory (LSTM) networks, deep belief networks, and ensemble learning and combination models to anticipate short-term load. Taking the method as an illustration, we provide a concise overview of the application of deep learning models in short-term load forecasting.

3.5 Convolutional Neural Network CNN utilizes convolutional layers to extract features from input data and use fully connected layers for regression prediction and classification. Originally, CNN was mostly utilized for image processing. However, if the data is presented in two or three dimensions, similar to an image, CNN can also be employed for data prediction. The literature utilized a hybrid prediction technique by combining the optimal characteristics of CNN and K-means clustering. This method was developed utilizing extensive data sets acquired from the power grid and implemented the Kmeans algorithm. Group them into clusters, and the resulting clusters are utilized to train a Convolutional Neural Network (CNN). Ultimately, the predictions from each cluster subset are combined to yield the final load prediction results. [26] The approach involves transforming multi-variable time series data into multi-channel pictures and feeding them into a CNN model. Additionally, logical models like fuzzy space and spectrum are utilized to represent the onedimensional structure of the time series. This helps effectively address the over-fitting issue of the CNN model. In the literature, historical load is utilized as the input feature, with a focus on reducing the load data by examining the distribution of the data time series [27].

In order to minimize noise errors and improve the time series properties, we employ a multi-temporalspatial-scale temporal convolutional network to analyze load data. This approach allows us to simultaneously capture the nonlinear and time series properties of the load data. [28] In their study, the researchers utilized a Convolutional Neural Network (CNN) to forecast residential load. They employed a technique called nonlinear relationship extraction (NRE) to extract and combine load and temperature data from the previous two weeks, encompassing three days before, during, and after the forecasted period. This data was organized into a loadtemperature cube. The CNN model utilized a twodimensional convolution operator to extract local features from load values in the surrounding area.

Additionally, the load-temperature cube was used to learn hidden nonlinear load temperature features to train the CNN model. To further enhance the accuracy of the predictions, the researchers combined the CNN model with Gaussian kernel Support Vector Regression (SVR) to minimize prediction errors. The results of the study demonstrated that this method outperformed other approaches in terms of prediction performance.

3.6 Recurrent Neural Network Load forecasting has benefited from RNNs. The literature addresses volatility and uncertainty in residential electrical load forecasts. It uses residential consumer load profiles to train a deep recurrent neural network. This network directly handles and learns uncertainty, eliminating model overfitting. Recurrent Neural Networks (RNNs) are used for offline learning, since each training session cannot immediately learn from new data. Fekri et al. [29] use RNN to capture temporal correlation to solve this problem. The RNN weight is updated using new data, and the hyperparameters are automatically adjusted. In literature, a recursive inspection neural network model for short-term power load forecasting is developed using RNN and a one-dimensional CNN [30]. CNN evaluates high and low points in power load time series data, while RNN learns long-term and short-term temporal correlations. Compared to MLP, the enhancements improve Mean Absolute Percentage Error.

RNN acquires the pattern representation of previous time steps by utilizing shared parameters across all time steps. However, as the sequence length increases, the memory of past patterns rapidly diminishes. LSTM employs gated memory units to effectively preserve data information over an extended duration, hence enhancing prediction accuracy. The field of literature extensively examines the cyclical nature of energy consumption data, use autocorrelation to understand the attributes of affecting factors, and develops an LSTM network to model and forecast sequential data. This approach aligns with conventional models like ARMA and ANN. RMSE is substantially decreased in comparison. The study in literature LSTM for predicting the energy consumption of non-residential users. It use K-means classification to identify patterns in user energy consumption behavior and analyzes the performance of non-residential consumers in different time series using the Spearman correlation coefficient. The proposed architecture presents a framework for forecasting non-residential load using a multisequence node for each layer. They fed 371 daily electrical load data points into a fully connected layer consisting of 53 neurons. The output of the fully connected layer serves as the input for the LSTM loop layer, substantially reducing the duration of the input load sequence and mitigating the vanishing gradient problem. Confronted with In the case of highdimensional huge feature data, the literature [32] creates a feature matrix of influencing elements, applies density clustering to group users into 2 categories, and employs a 3-layer BiLSTM to forecast the load of a specific location in the upcoming week.

Experts and academics highly value the gate recurrent unit (GRU) for its exceptional memory learning capability, in addition to LSTM. Du et al.developed a load matrix with historical data that includes both spatial and temporal correlation. They used a 3-layer Convolutional Neural Network (CNN) to extract features from the load data, and a Gated Recurrent Unit (GRU) to map the sequence to the final predicted value. The ongoing combination of influencing elements on electricity load results in a progressively increasing non-stationarity of the load sequence. To enhance the stability of the load sequence, employed Empirical Mode Decomposition (EMD) to isolate the non-stationary factors of the load sequence. They then utilized Pearson correlation analysis to examine the serial correlation and fed the sequence with higher correlation into the Gated Recurrent Unit (GRU) for prediction. On the other hand, the study [33] employed differential decomposition and error compensation smoothing. The GRU is utilized for prediction following the load sequence. Both approaches significantly enhance the accuracy of the model. Employ a two-dimensional convolution layer to extract the latent characteristics of the power load matrix, effectively circumventing the issue of excessive parameters in the convolution layer.

Table-1 CNN Models in Short-Term Load Forecasting

Predictive Model	Activation Function	Dataset	Factors	Comparative Models	Assessment Methods	Advantages and Disadvantages
CNN [26]	ReLU	Hourly load data generated by a Malaysian power supply company	Weather factors	ARIMA, FTS, LSTM	APE, MAPE, RMSE	Use images to represent load data to improve prediction accuracy
MTCN [27]	Linear	15-minute-level power load data for china	Historical load, weather factors, other factors	prediction models including BPNN, LSTM, SVM, RBF, etc.	MAPE, RMSE, MAE, R2	Effectively resolves the issue of gradient vanishing and amplification.
NRE_CNN [28]	ReLU	Load data for a location in Canada and for apartments in UMass	Historical load, weather factors	CNN, FFNN, LSTM, SVR	MAPE, MAE, MSE, NRMSE	Model theory is complex and only applicable to a small number of data sets

The size of the convolution kernel is determined by the autocorrelation coefficient of the electric load. For prediction purposes, the model leverages the strengths of LSTM and GRU networks, utilizing the LSTM unit for forward propagation and the GRU unit for back propagation, thereby achieving enhanced forecast accuracy.

3.7 Deep Learning Models, Aside from frequently utilized models like CNN, LSTM, and GRU, additional deep learning models have also been employed in short-term load forecasting. For instance, in their study, researchers employed a deep belief network consisting of multilayer restricted Boltzmann machines for load forecasting purposes. Following the unsupervised training conducted layer by layer, the model parameters are further refined using the supervised back propagation training method. Additionally, the electricity price variable is incorporated, resulting in a significant enhancement in prediction accuracy.

In order to reduce the computational cost problems caused by the complex network structure and limited deterministic point prediction, [34] used the K nearest neighbor algorithm to find historical power load time series characteristics similar to future values, and used DBN for prediction. The literature uses the optimization (CSO) algorithm to optimize DBN and the mutual information algorithm to screen influencing factors. Both models show good performance.

The article [35] used autoencoders and cascade neural networks to build a DNN model for load forecasting. Reference [36] considered historical load and meteorological data such as temperature and sunshine at the same time, and used gradient boosting regression tree to carry out probabilistic net load forecasting. Literature studies the impact of photovoltaic penetration, seasonality and load demand increase on net load forecast intervals, and uses quantile regression and dynamic Gaussian models to carry out probabilistic forecasting of daily load intervals. Literature a combination of phase space reconstruction and BNN is used for short-term net load forecasting.

Deep learning models can mine load data set features at a deeper level and improve prediction accuracy, but model framework and parameter selection are still difficult problems that need to be solved.

3.8 Integrated Learning and Combination Models

High stability and accuracy have always been the goals pursued in model training, but the results of single model training are often unsatisfactory. Therefore, prominent professionals and scholars amalgamate numerous feeble learners to create a robust learner, aiming to achieve superior outcomes. Ensemble learning, often known as [37], refers to the process of combining many models to make predictions. introduced a multi-modal evolutionary algorithm that utilizes comprehensive weighted vector angle and shift-based density estimation. They also developed a method that includes prediction performance evaluation, model attribute analysis, intelligent decision support scheme for structural and fusion strategy optimization, and optimal model preference selection. The goal of this method is to enhance ensemble learning models based on random vector function chain networks.

time-consuming. To address this problem, the literature [42] Decision trees are used as weak learners to construct random forest (RF) models for loading forecasting. Rough sets, gray projection technology, and fruit fly optimization algorithms are used to optimize the models and good results have been achieved. The XGBoost algorithm uses second-order derivatives, which improves calculation efficiency

and effectively reduces over-fitting problems. However, each iteration needs to traverse the entire data set, which is not suitable for processing larger data sets and has higher memory requirements. In order To solve this problem, the literature the stacking technology of integrated learning to combine XGBoost with three models: light gradient boosting machine and multi-layer perceptron MLP for load forecasting, which shows good performance advantages. [39] constructed a two-layer XGBoos model structure of a data processing layer that screens feature sets and a load prediction layer that optimizes parameter selection. This model uses the feature sets and loads screened in the first layer as the second layer input, effectively avoiding the problem of interdependence of feature data. In addition to the above integrated learning models, some experts also integrate neural network, ensemble learning and other algorithms for combined prediction, as shown in Table 2. [42 power system into several small areas and used fuzzy neural Multiple algorithms such as fuzzy neural network and SVM are used to predict the electricity load of a single area respectively, and the locally linear model tree method is used to combine the prediction results of small areas, combined with fuzzy inference to flexibly set Network topology can more effectively extract load distribution trends. Literature [43] combines CNN with RF, which effectively reduces MAPE. To address the problem of low prediction accuracy caused by few data features, use CNN to extract users the local trend of electricity load was combined with the LSTM model for prediction. [40, 41] used the reciprocal error method to select the best weight of the model combination, and combined Figure 1: Comparison of Actual Load Data with Predictions from Various Machine Learning Models To ensure a precise comparison of the pros and of each model, the dataset was evenly partitioned. The initial 70% was designated as the training dataset, while the remaining 30% was allocated as the test fitting problems. However, each iteration needs to traverse the entire data set, which is not suitable for processing larger data sets and has higher memory requirements. In order To solve this problem, the literature [38] uses the stacking technology of integrated learning to combine XGBoost with three models: light gradient layer perceptron MLP for load forecasting, which shows good performance layer XGBoost model structure of a data processing layer that screens feature sets and a load prediction layer that optimizes parameter selection. This model uses the feature sets and loads screened in the first layer as the second e problem of In addition to the above integrated learning models, some experts also integrate neural network, ensemble learning and other algorithms for combined 2] divided the m into several small areas and used fuzzy neural Multiple algorithms such as fuzzy neural network and SVM are used to predict the electricity load of a single area respectively, and the locally linear model tree method is used to combine the ults of small areas, combined with fuzzy inference to flexibly set Network topology can more effectively extract load distribution trends. Literature] combines CNN with RF, which effectively reduces MAPE. To address the problem of low cy caused by few data features, use CNN to extract users the local trend of electricity load was combined with the LSTM model for prediction. used the reciprocal error method to select the best weight of the model combination, and combined the time series predicted by the XGBoost and LSTM models. The data is weighted and combined, effectively reducing the prediction error of a single model.Used CNN to extract data Seq2Seq to predict user energy loads. At the same time, attention mechanisms and multi methods were introduced to improve the accuracy of load prediction. The article [preprocessing, multi objective optimization, model prediction and model evaluation into a new forecasting system, used EMD to decompose the load data, and combined three prediction models of RBF, ELM and GRNN for prediction. The results are weighted to form the final predicted value. Studies have discovered that deep learning prediction algorithms have effectively majority of load forecasting analysis issues. The combined model surpasses a single model in terms of performance and accuracy. Exploring additional permutations of deep learning models is a prominent area of future research.

IV. Experimental Analyses

In order to further analyze the advantages and disadvantages of each prediction model, this paper uses electricity load data for four datasets to conduct experimental comparisons, and selects the more typical SVM [20], Bagging [42], and XGBoost [38] and other models are used for experiments. Datasets include foreign building energy consumption data (buildata) [45, 46] and Belgian load data set (beldata) domestic data sets include the load of the National Electrician Mathematical Modeling Competition data set (nemmdata) and the real power consumption data set of PRC (prcdata).

Dataset. This experiment was carried out using Python. The projected values of five frequently employed load forecasting techniques and the customer's electricity usage were compared. Figure 1 displays the comparison of actual electrical load data with predictions from several machine learning models for the month of March 2018. The AdaBoost model is represented by a dashed orange line, while the Bagging model is shown with a dashed purple line. The figure suggests that the Bagging model has a higher degree of numerical fitting to the actual load data, indicating better forecasting accuracy. On the other hand, the AdaBoost model's predictions are less aligned with the actual data, pointing to a need for improvement in its forecasting performance. The XGBoost model, represented by a dashed red line in Figure 1, is described as reasonably precise. While it may not closely align with the actual load as much as other models, it nevertheless offers a decent amount of predictive accuracy.

The predictions of the SVM model are represented by a dashed line in the color brown. The graphic displays the performance of the Support Vector Machine (SVM) model using three distinct kernel functions. Nevertheless, the SVM model's capacity to accurately fit the data is shown to be suboptimal, possibly due to its difficulties in handling the intricate variations included in the load data. Ultimately, the LSTM model, indicated by a dotted green line, demonstrates a significant level of numerical conformity. The predictions of this model closely align with the actual load data, demonstrating its exceptional ability to capture temporal trends. As a result, it is considered a very accurate model for load forecasting. To summarize, Figure 1 illustrates that the Bagging and LSTM models exhibit strong performance in accurately fitting the real load data, however the AdaBoost model shows potential for enhancement. The XGBoost model provides a reasonable level of accuracy, while the SVM model has the least favorable fitting performance among the provided models because to its limits in data processing.

Table-2 Combinational Models in Short-Term Load Forecasting

Predictive Model	Combination Method	Dataset	Influencing Factors	Comparative Models	Assessment Methods
XGB+LGBN+MLP [38]	Stacked stack sets nonlinear weights	Hourly temperature and load data	Temperature, historical load	Single models such as XGB, MLP, etc.	MAPE, MAE, R2, etc.
Dual-layer XGBoost [39]	Serial	15-minute daily load data	Meteorological factors	XGBoost, BPNN, ARIMA	MAPE, RMSE
XGBoost_LSTM [40]	MAPE-RW search weights	15-minute load data of a certain place	Historical load	XGBoost_GRU	MAPE
CNN+LSTM+XGBoost [41]	MAPE-RW searches for optimal weights	Spain hourly load data	Temperature, historical load	SVR, LSTM, CNN_LSTM	MAPE, RMSE

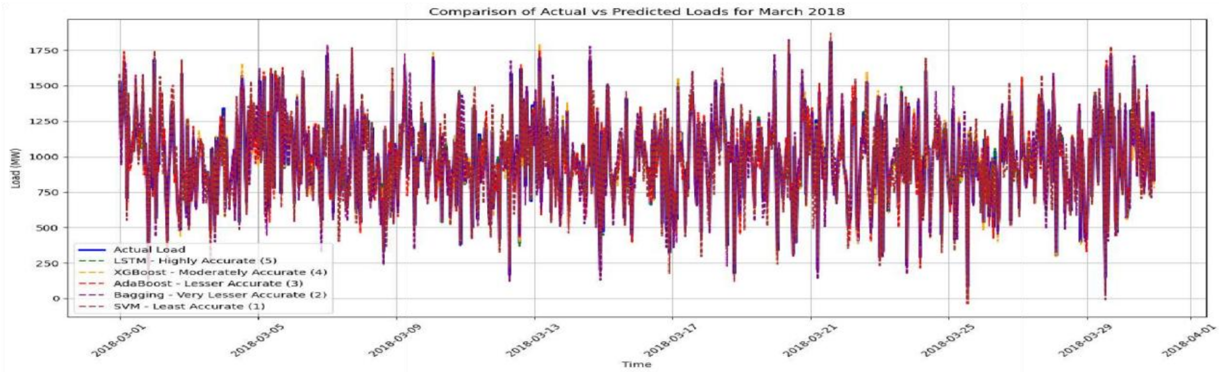

Figure 1: Comparison of Actual Load Data with Predictions from Various Machine Learning Models

The Mean Absolute Percentage Error (MAPE), calculated as the average of 15 iterations, is presented in Table 3 for the running outcomes of 5 frequently employed models. The experimental results reveal that the Bagging model exhibits robust generalization capability and is particularly suitable for datasets of diverse sizes, as evidenced by its consistently low MAPE values throughout multiple runs. The prediction accuracy of this model is really outstanding, giving it a dependable option for tasks involving forecasting. On the other hand, AdaBoost, albeit displaying promise, necessitates lengthier execution durations and enhancements in precision as compared to the Bagging model. Enhancements to AdaBoost may improve its performance and lead to more accurate forecasting results. Integrated learning models regularly demonstrate high performance on small datasets, exhibiting steady performance levels. Nevertheless, when dealing with high-dimensional sparse data such as the buildenergy dataset, it may be necessary to fine-tune the prediction accuracy of XGBoost models in order to enhance the outcomes. Overall, both the Bagging and AdaBoost models exhibit strong generalization capabilities and are versatile in handling diverse datasets. XGBoost stands out for its superior prediction accuracy, although it comes with higher memory requirements and may not be optimal for large-scale datasets. AdaBoost is highlighted for its fast training speed, but it may face challenges when processing performance, requiring careful consideration for optimal results. High dimensional sparse data due to inherent limitations. The LSTM model shows promising prediction results, but the selection of model parameters significantly influences its forecasting

Table-3 Comparative MAPE for Models

Models	buildata	beldata	nemmdata	prcdata
SVM	0.35	0.14	0.23	0.14
Bagging	0.29	0.09	0.16	0.09
AdaBoost	0.36	0.17	0.23	0.15
XGBoost	0.28	0.18	0.72	0.18
LSTM	0.68	0.14	0.40	0.21

V. Conclusion And Outlook

Load forecasting has become an essential component of the power system industry due to the rapid advancement of the information age [5]. This article provides a concise overview of classic forecasting methods and advanced forecasting algorithms that utilize deep learning techniques. Traditional load forecasting algorithms are known for their speed and simplicity. However, they have some limitations when it comes to handling nonlinear data [16, 17]. Intelligent prediction algorithms based on deep learning are favored by major experts because of their strong generalization ability and high self-learning ability. However, there are also problems such as model complexity, over-fitting, and local optimal solutions [25]. Model combination strategies and weight parameter setting issues have become obstacles to the development of integrated learning models. In practical applications, short-term power load forecasting still needs to be developed in the following aspects:

(a) The advancement of intelligent equipment allows for the acquisition of an increasing number of influencing elements on load data. Nevertheless, the scarcity of high-quality labelled data persists due to service providers' practice of maintaining the confidentiality of real-time and historical data for security purposes. The data obtained by sensors is plagued by issues such as data duplication, incorrect labeling, and occasional interruptions in data streams. Consequently, the need for effective and dependable data preparation has gained significant significance in load forecasting [14].

(b) Conventional models and intelligent models possess distinct advantages. Current research is focused on maximizing the benefits of each model, leveraging their strengths, mitigating shortcomings, and enhancing model accuracy [33, 35, 36].

(c) The current process of predicting power demand entails dealing with many challenges related to quantifying uncertainty. These challenges include ensuring the correctness and completeness of the training data, addressing the constraints of deep learning models, and evaluating their performance. The resolution of uncertainty quantification problems is crucial for enhancing the dependability and precision of deep learning models.

REFERENCES

[1] Athanasios Ioannis Arvanitidis, D. Bargiotas, Aspassia Daskalopulu, Dimitrios Kontogiannis, I. Panapakidis and L. Tsoukalas. "Clustering Informed MLP Models for Fast and Accurate Short-Term Load Forecasting." Energies (2022). https://doi.org/10.3390/en15041295.

[2] Behnam Farsi, Manar Amayri, N. Bouguila and U. Eicker. "On Short-Term Load Forecasting Using Machine Learning Techniques and a Novel Parallel Deep LSTM-CNN Approach." IEEE Access, 9 (2021): 31191-31212.

[3] Bastian Dietrich, J. Walther, M. Weigold and E. Abele. "Machine learning based very short term load forecasting of machine tools." Applied Energy, 276 (2020): 115440. https://doi.org/10.1016/j.apenergy.2020.115440.

[4] T. Alquthami, M. Zulfiqar, Muhammad Kamran, A. H. Milyani and M. B. Rasheed. "A Performance Comparison of Machine Learning Algorithms for Load Forecasting in Smart Grid." IEEE Access, PP (2022): 1-1. https://doi.org/10.1109/access.2022.3171270.

[5] International Renewable Energy Agency. Renewable energy statistics 2020.
https://www.irena.org/publications/2020/Jul/Renewable-energystatistics-2020. (2020-07-01) [2021-11-11]

[6] Yi Hu, B. Qu, Jie Wang, Jing J. Liang, Yanli Wang, Kunjie Yu, Yaxin Li and Kangjia Qiao. "Short-term load forecasting using multimodal evolutionary algorithm and random vector functional link network based ensemble learning." Applied Energy, 285 (2021): 116415. https://doi.org/10.1016/J.APENERGY.2020.116415

[7] Songjiang Li, Peng Wang and L. Goel. "A Novel Wavelet-Based Ensemble Method for Short-Term Load Forecasting with Hybrid Neural Networks and Feature Selection." IEEE Transactions on Power Systems, 31 (2016): 1788-1798. https://doi.org/10.1109/TPWRS.2015.2438322.

[8] Weilin Guo, Liang Che, M. Shahidehpour and Xin Wan. "Machine-Learning based methods in short-term load forecasting." The Electricity Journal, 34 (2021): 106884. https://doi.org/10.1016/j.tej.2020.106884

[9] Haoyu Jiang, A. Wu, Bo Wang, P. Xu and Gang Yao. "Industrial Ultra-Short-Term Load Forecasting With Data Completion." IEEE Access, 8 (2020): 158928-158940. https://doi.org/10.1109/ACCESS.2020.3017655

[10] Rangelov, D., Boerger, M., Tcholtchev, N., Lämmel, P., & Hauswirth, M. (2023). Design and Development of a Short-Term Photovoltaic Power Output Forecasting Method Based on Random Forest, Deep Neural Network and LSTM Using Readily Available Weather Features. IEEE Access, 11, 41578-41595. https://doi.org/10.1109/ACCESS.2023.3270714.

[11] N. Giamarelos, Myron Papadimitrakis, M. Stogiannos, E. Zois, N. Livanos and A. Alexandridis. "A Machine Learning Model Ensemble for Mixed Power Load Forecasting across Multiple Time Horizons." Sensors (Basel, Switzerland), 23 (2023). https://doi.org/10.3390/s23125436.

[12] Wang, C., & Li, W. (2023). A hybrid model of modal decomposition and gated recurrent units for short-term load forecasting. PeerJ Computer Science, 9. https://doi.org/10.7717/peerj-cs.1514.

[13] Nespoli, A., Leva, S., Mussetta, M., & Ogliari, E. (2022). A Selective Ensemble Approach for Accuracy Improvement and Computational Load Reduction in ANN-based PV power forecasting. IEEE Access, PP, 1-1. https://doi.org/10.1109/ACCESS.2022.3158364.

[14] Qin, C., Srivastava, A., Saber, A., Matthews, D., & Davies, K. (2023). Geometric Deep-Learning-Based Spatiotemporal Forecasting for Inverter-Based Solar Power. IEEE Systems Journal, 17, 3425-3435. https://doi.org/10.1109/JSYST.2023.3250403.

[15] Box GEP, Jenkins GM. Time Series Analysis: Forecasting and Control. San Francisco: Holden-Day, 1976. 238–242.

[16] Zhou, W., Tao, H., Ding, S., & Li, Y. (2022). Electricity consumption and production forecasting considering seasonal patterns: An investigation based on a novel seasonal discrete grey model. Journal of the Operational Research Society, 74, 1346 - 1361. https://doi.org/10.1080/01605682.2022.2085065.

[17] Yang, J., Tang, Y., & Duan, H. (2022). Application of Fuzzy Support Vector Machine in Short-Term Power Load Forecasting. J. Cases Inf. Technol., 24, 1-10. https://doi.org/10.4018/jcit.295248.

[18] Zulfiqar, M., Kamran, M., Rasheed, M., Alquthami, T., & Milyani, A. (2022). A Short-Term Load Forecasting Model Based on Self-Adaptive Momentum Factor and Wavelet Neural Network in Smart Grid. IEEE Access, 10, 77587-77602. https://doi.org/10.1109/access.2022.3192433.

[19] Cai Qiuna, Chao Zhu, Su Binghong, et al. Short-term load forecasting method of neural network based on a new type of robust loss. Power Grid Technology, 2020, 44(11): 4132-4139.

[20] Han, M., Tan, A., & Zhong, J. (2022). Improved particle swarm optimization combined with least squares support vector machines for short-term load forecasting. , 12176, 1217612 - 1217612-11. https://doi.org/10.1117/12.2636473.

[21] Xu, J., Liu, M., Xiang, W., Chen, H., Mao, Y., & Jia, T. (2023). Short-term Load Forecasting Based on Multiple PSO-LSSVM under Electricity Market Environment. Journal of Physics: Conference Series, 2427. https://doi.org/10.1088/17426596/2427/1/012024.

[22] Zhang ZC, Hong WC. Application of variational mode decomposition and chaotic gray wolf optimizer with support vector regression for forecasting electric loads. Knowledge-Based Systems, 2021, 228: 107297.DOI:10.1016/j.knosys.2021.107297

[23] Qin, B., Huang, X., Wang, X., & Guo, L. (2023). Ultra-shortterm wind power prediction based on double decomposition and LSSVM. Transactions of the Institute of Measurement and Control, 45, 2627 - 2636. https://doi.org/10.1177/01423312231153258 .

[24] Dai YM, Zhao P. A hybrid load forecasting model based on support vector machine with intelligent methods for feature selection and parameter optimization. Applied Energy, 2020, 279: 115332.DOI:10.1016/j.apenergy.2020.115332

[25] Jahangir, H., Tayarani, H., Gougheri, S., Golkar, M., Ahmadian, A., & Elkamel, A. (2021). Deep Learning-Based Forecasting Approach in Smart Grids With Microclustering and Bidirectional LSTM Network. IEEE Transactions on Industrial Electronics, 68, 8298-8309. https://doi.org/10.1109/TIE.2020.3009604.

[26] . Sadaei HJ, de Lima e Silva PC, Guimarães FG, et al. Short-term load forecasting by using a combined method of convolutional neural networks and fuzzy time series. Energy, 2019, 175: 365-377.DOI:10.1016/j. energy.2019.03.081

[27] Yin LF, Xie JX. Multi-temporal-spatial-scale temporal convolution network for short-term load forecasting of power systems. Applied Energy, 2021, 283: 116328.DOI:10.1016/j.apenergy.2020.116328

[28] Imani M. Electrical load-temperature CNN for residential load forecasting. Energy, 2021, 227: 120480.DOI:10.1016/j.energy.2021.120480

[29] Fekri MN, Patel H, Grolinger K, et al. Deep learning for load forecasting with smart meter data: Online adaptive recurrent neural network. Applied Energy, 2021, 282: 116177.DOI:10.1016/j.apenergy.2020.116177

[30] Mahjoub, S., Chrifi-Alaoui, L., Marhic, B., & Delahoche, L. (2022). Predicting Energy Consumption Using LSTM, MultiLayer GRU and Drop-GRU Neural Networks. Sensors (Basel, Switzerland), 22. https://doi.org/10.3390/s22114062.

[31] Zang HX, Xu RQ, Cheng LL, et al. Residential load forecasting based on LSTM fusing self-attention mechanism with pooling. Energy, 2021, 229: 120682.DOI:10.1016/j.energy.2021.120682

[32] Ijaz, K., Hussain, Z., Ahmad, J., Ali, S., Adnan, M., & Khosa, I. (2022). A Novel Temporal Feature Selection Based LSTM Model for Electrical Short-Term Load Forecasting. IEEE Access, 10, 82596-82613. https://doi.org/10.1109/ACCESS.2022.3196476.

[33] Feng, C., Xu, K., & Ma, H. (2022). Research on Impulse Power Load Forecasting Based on Improved Recurrent Neural Networks. Computational Intelligence and Neuroscience, 2022. https://doi.org/10.1155/2022/2784563.

[34] Zhang, Y., & Li, G. (2023). Power load forecasting based on deep neural network. 12719, 127190V - 127190V-5. https://doi.org/10.1117/12.2685798.

[35] Alipour M, Aghaei J, Norouzi M, et al. A novel electrical netload forecasting model based on deep neural networks and wavelet transform integration. Energy, 2020, 205: 118106.DOI:10.1016/j.energy.2020.118106

[36] Qin, C., Srivastava, A., Saber, A., Matthews, D., & Davies, K. (2023). Geometric Deep-Learning-Based Spatiotemporal Forecasting for Inverter-Based Solar Power. IEEE Systems Journal, 17, 3425-3435. https://doi.org/10.1109/JSYST.2023.3250403.

[37] Aduama, P., Zhang, Z., & Sumaiti, A. (2023). Multi Feature Data Fusion-Based Load Forecasting of Electric Vehicle Charging Stations Using a Deep Learning Model. ArXiv, abs/2301.13774. https://doi.org/10.3390/en16031309.

[38] Massaoudi M, Refaat SS, Chihi I, et al. A novel stacked generalization ensemble-based hybrid LGBM-XGB-MLP model for short-term load forecasting. Energy, 2021, 214: 118874.DOI:10.1016/j.energy.2020.118874

[39] Sun Chao, Lu Qi, Zhu Sixian, et al. Ultra-short-term power load forecasting based on the double-layer XGBoost algorithm considering the influence of multiple features. High Voltage Technology, 2021, 47(8): 2885-2895.

[40] Chen Zhenyu, Liu Jinbo, Li Chen, et al. Ultra-short-term power load forecasting based on the combined model of LSTM and XGBoost. Power Grid Technology, 2020, 44(2): 614-620.

[41] Zhuang Jiayi, Yang Guohua, Zheng Haofeng, et al. CNN-LSTMXGBoost short-term power loading forecasting method based on multi-model fusion. China Electric Power, 2021, 54(5): 46-55.

[42] Zhang, J., Zhao, Z., Yan, J., & Cheng, P. (2023). Ultra-ShortTerm Wind Power Forecasting Based on CGAN-CNN-LSTM Model Supported by Lidar. Sensors (Basel, Switzerland), 23. https://doi.org/10.3390/s23094369.

[43] Liu, F., Dong, T., Hou, T., & Liu, Y. (2021). A Hybrid ShortTerm Load Forecasting Model Based on Improved Fuzzy CMeans Clustering, Random Forest and Deep Neural Networks. IEEE Access, 9, 59754-59765. https://doi.org/10.1109/ACCESS.2021.3063123.

[44] Nie Y, Jiang P, Zhang HP. A novel hybrid model based on combined preprocessing method and advanced optimization algorithm for power load forecasting. Applied Soft Computing, 2020, 97: 106809.DOI:10.1016/j.asoc.2020.106809

[45] Lazzeroni, P., Lorenti, G., & Repetto, M. (2023). A Data-Driven Approach to Predict Hourly Load Profiles From Time-of-Use Electricity Bills. IEEE Access, 11, 60501-60515. https://doi.org/10.1109/ACCESS.2023.3286020.

[46] Zeng Ming, Zhang Shuo. Analysis of comprehensive energy development in the "14th Five-Year Plan" power plan. China Electric Power Enterprise Management, 2020(13): 2628.DOI:10.3969/j.issn.1007-3361.2020.13.006.