An Edge Intelligence Framework for Elegant Power Management in IoT-enabled Power Grids

Irina V. Pustokhina[*1], **Denis A. Pustokhin**[2]

[1] Department of Entrepreneurship and Logistics, Plekhanov Russian University of Economics, 117997, Moscow, Russia

[2] Department of Logistics, State University of Management, 109542, Moscow, Russia

Emails: pustohina.IV@rea.ru; da_pustohin@guu.ru

Abstract

The Internet of Things (IoT) is a concept that has the potential to attract new audiences in fields as diverse as manufacturing, healthcare, and more. IoT devices included into the sensor were the primary drivers of the massive data collection. To successfully combine, assess, and comprehend all programme objects, thus, self-adaptive algorithms based on AI are necessary. The proliferation of both massive datasets and resource-intensive IoT devices makes stringent power management essential. The proliferation of both massive datasets and resource-intensive Internet of Things devices makes stringent energy management essential. Combining IoT with AI-based techniques is crucial for equitable power distribution to compact mobile devices. To this end, we offer an efficient way to communicate between power utilities and end users by forecasting future power usage over short periods of time. Innovations include a revolutionary convolutional recurrent model for lightweight prediction method with low duration intricacy and minimum margins of error, as well as massive energy administration for edge devices via a centralised cloud-based data supervisory server. To maintain the power consumption and supply paradox efficiently, the suggested scheme has mobile nodes interact with a central remote server via an IoT network and then on to the corresponding power grid. We use a number of preparation methods to accommodate the varied electrical data, and then we construct a powerful decision-making engine for quick prediction on devices with limited resources.

Keywords: Internet of Things (IoT); Energy; Power Grids; Deep Learning

1. Introduction

Research into the power management at the power grid through the use of automated methods for demand forecasting in the future is an attractive field of study. Power systems are safe and reliable places for the distribution of electrical power to a wide range of users, including smart homes and industrial facilities, among other types of clients. The creation of electrical power takes place in power stations, its distribution takes place via smart networks, and its intake might take place in the household [1] or commercial structures, as well as in the industrial sector [2]. The quantity of energy that is consumed by individuals is the single most important factor that determines how much power is generated in power plants and transmitted across grids. The vast majority of customers are not knowledgeable about the energy requirements placed on electric networks, which results in monetary loss and the spending of unnecessary

DOI: https://doi.org/10.54216/JISIoT.060204
Received: July 26, 2022 Accepted: October 18, 2022

energy. In a similar vein, the farmers want to reduce costs as much as possible while simultaneously increasing the amount of energy they generate. This necessitates the use of effective planning and administration tactics.

These days, Internet of Things (IoT) and mobile devices are frequently outfitted with sophisticated sensors and high computation power, which enables them to gather and handle massive amounts of data that are made at the edge networks. Moreover, it is anticipated that there will be more than 10 billion intelligent things in the IoT linked to the Internet and that the total amount of mobile data would exceed 49 exabytes each month. The functions of processing and data storage could be offered by a system that combines cloud computing with edge computing. In this type of system, the tasks performed by users are automatically transmitted to either a cloud data center layer or an edge of the network. A great number of IoT applications call for the preparation and classification of data, which are subsequently put into practice to make predictions about future events through the application of machine learning (ML) techniques. Because the vast volume of data produced by IoT devices is typically acquired in personal settings, its sensitivity to information privacy is warranted by its very nature. Sending all of the data to a single server or cloud center that trains a deep learning model is therefore often not viable. Because of the restricted capabilities of mobile networks, sending a large volume of information through a wireless connection result in costly communication expenses as well as substantial network congestion. In addition to this drawback, the transfer of a large amount of data might take an extremely long time.

A well-thought-out plan for the production and utilization of power assures that it will be put to productive use in enterprises and households while also allowing power stations to generate the appropriate quantity of power. The microgrid is a way to communicate that maintains the integrity of the power transmission between the power supplier and the power user, and it is accountable for the steady state that exists for both sides. Techniques of power prediction are quite beneficial in this respect since they estimate the future power consumption of a client and the strain that is placed on the grids following that consumption. A faulty projection of power levels will result in higher costs as well as the waste of that power. For a residential house in the United Kingdom in 1984, it was estimated that there was a loss of 10 million pounds per year along with an increase of 1% in predicting inaccuracy [4]. Thus, accurate methods of projecting power requirements are necessary to make the best judgments for the future. Numerous energy forecasting systems are available, and their applications can be found in both residential and commercial settings. In the next paragraphs, we will talk about several representative methodologies that are especially relevant to the work that will be given in next Section.

The individualized predictive control systems can be deployed in a variety of practical scenarios, like day-ahead predictions, which help microgrids meet the required power needs [5]. The technologically intelligent systems that involve capacity planning play a crucial part in alleviating the current energy shortage and making a contribution to the greening of the climate. The majority of these techniques are instructional techniques that use deep learning, such as long short-term memory (LSTM), as it is the most widely used technique in the field of methodologies connected to energy forecasting. LSTM is a form of recurrent neural network (RNN) that is utilised extensively in a variety of fields, like video surveillance, for the purpose of completing sequencing and serial learning tasks [6]. In spite of the prevalence of LSTMs, hybrid techniques that combine fuzzy neural inference systems and evolutionary computation can be found in the relevant body of scholarly research concerning power forecasting. It has been demonstrated that accurate forecasting of power consumption may be achieved by the utilisation of spatial and temporal data in conjunction with one another. The researchers demonstrated that convolutional neural networks (CNNs) are superior to other methods for obtaining the representational characteristics of various variables that have an impact on the forecasting of power consumption. In addition, the incorporation of these representational characteristics with CNN's results in a reduction in the failure rates throughout the power grid data.

An in-depth analysis of the literature on topics connected to utilized demand forecasting reveals a number of unanswered questions that require more investigation. Obtaining preciseness in the predictive performance is the most important and difficult task when introducing an enhanced power prediction technique because it is the one that draws the most attention. The effective implementation of the different algorithms over the end devices, which is necessary for a productive interaction between pervasive computing in an Internet of Things (IoT) network for the purpose of power production, is another significant obstacle that is insufficiently addressed in the relevant research literature. Lately, resource-constrained devices in IoT contexts have demonstrated a significant degree of potential in machine vision [8, 9], medical [10], and a wide variety of other fields [11, 12]. In continuing with these issues, a significant concern is the decreased temporal difficulty of a power prediction model, especially when dealing with the issue of

DOI: https://doi.org/10.54216/JISIoT.060204
Received: July 26, 2022 Accepted: October 18, 2022

predictive modeling of power. In addition, the cloud [11] and fog computing [12], [13] technologies are seldom ever used in power modeling tasks, despite the fact that they offer reliable tools for effective big data analysis and fast decision making, including the prediction of abnormal power consumption. In an attempt to address these gaps, this work proposes to take the advantage of both DL and IT technologies to afford an intelligent system for power management in smart cities. The main contributions of the proposed solution can be summarized across the following points

1. This study introduces an elegant edge intelligence framework to promote scalable and efficient management of power distribution based on IoT-driven predictive modeling of power flows in power grids.
2. We provide a framework for deploying configurable devices with limited resources in a variety of consumer settings (smart homes or industries), where they can upload their immediate needs to a cloud-based monitoring server and provide insight into their anticipated future needs. Power control is optimized by the power grid, which process demands from homes and businesses, then sends out only that quantity of power. When reporting on consumers' out-of-the-ordinary energy consumption, a cloud server first screens out all normal requests. It stores energy predictions that can be utilized for more thorough examination later.
3. According to rigorous testing, we've established that our framework can serve as a standard for cutting-edge artificial intelligence applications in the field of power prediction. In the first set of tests, we show how well our framework performs in comparison to several models by selecting a normalization technique and a sequential model. Here, we investigate the relationship between a model's implementation duration and its accuracy by comparing various series-learning algorithms.

This remaining portion of the article is broken up into five key parts. In Section II, the techniques that are now considered state-of-the-art are discussed. In section III, we will discuss the proposed approach as well as the functionality of our framework. Section IV provides an explanation of the experimental specifics. Section V discusses the numerical outcome. The entire study is brought to a close in Section V, which also provides some suggestions for further study.

2. Literature review

This part is broken up into two primary subsections, the first of which is devoted to statistical techniques, and the second to deep learning-based solutions. The power prediction models that were in use during the time period in question did not place enough emphasis on the utilization of IoT devices with limited resources, which are becoming increasingly prevalent because of the computing skills and quick decision systems that they offer. In the sections that follow, these techniques are broken down into their respective categories, such as statistical ones and deep learning-based ones [1].

2.1. Statistical Power Analysis

The work [2] provided a way that ensures the application of low-performance machine learning algorithms on equipment in the IoT paradigm. Doing so will generate means for IoT consciousness, in which the KNN method was applied for the execution of the findings of the study, and the reliability of results is presented with a confusion matrix. The work [3] used the volume of the gas and wind velocity as input variables for a fuzzy system modeling wind energy prediction for microgrid applications, in which a set of models based on fuzzy logic and a modeling control scheme has been shown to be successful. For the purpose of constructing a regression model for electrical load forecasting from previous power data, the research [4] offered a hybrid model that incorporates the Wavelet Transform (WT) and Support Vector Machine (SVM) features. Information on electrical loads over time is broken down by the WT into a number of distinct series. The work [5] developed an ensemble predictive algorithm for wind energy that is based on modified naive Bayes classifiers is what we propose here. Multiple linear regression (MLR) and a naive Bayes classification model were used to anticipate the potential for wind is necessary to pick the integral that impacts the wind power outcomes from among the different climatic variables, such as heating rate, wind direction, and wind patterns. This was accomplished by performing a prediction of renewable sources using MLR and the naive Bayes classification algorithm based on the chosen climatic variables. In order to make an accurate forecast of wind power, we suggested combining the analog ensemble (AnEn) method with the composite learning method. in [6], following the screening of the coarse collection of characteristics with the Maximal Information Coefficient (MIC), the screening of the fine set of critical features affecting modeling is carried out employing the LightGBM feature extraction technique and the XGboost feature selection method, correspondingly. LightGBM and XGboost systems, both of

DOI: https://doi.org/10.54216/JISIoT.060204
Received: July 26, 2022 Accepted: October 18, 2022

which have strong prediction functions and receive the appropriate precise set of attributes and past workloads as inputs for the purpose of forecasting. The work [7], introduces the idea of crucial border matrices to offer a hardware-friendly template decline (TR) technique for the nearest neighbor (NN) classifiers. In addition, a physical system is built to illustrate the practicability of speeding up the suggested method by employing a field-programmable gate array (FPGA). In the beginning, the k-means centers are utilised in place of the complete template set as a replacement. Then, in order to improve the performance of the categorization, key border vectors are chosen via a training algorithm, which only requires one cycle to be finished. In addition, a global classifying strategy has been investigated and included in the algorithm in order to eliminate noisy boundary vectors. These vectors have the potential to mislead the classification in a generalized manner.

2.2. Deep Learning-based Power Analysis

The work [8] developed an advanced multi-channel bidirectional nested LSTM framework (MC-BiNLSTM) paired with discontinuous stationary wavelet transform for predicting energy usage in a manner that is both very precise and very effective. The primary contributions made by this research were the decomposition carried out with SWT to enhance correctness and the cooperative BiNLSTM framework to enhance performance. the work [9] developed a new deep learning (DL) model to accurately anticipate short-term power usage while simultaneously preserving effective interaction between power users and suppliers. The Model stack that has been proposed is made up of a number of layered spatiotemporal components, each of which is composed of a spatial transformer (ST) submodule and a temporal transformer (TT) submodule. The TT module represents the temporal dependencies present in the power data, whereas the ST module unearths concealed spatial features by merging convolutional and possesses an enhanced self-attention technique. The work [10] presented a PV-Net, a DL network for day-ahead Pv power prediction with a short-term horizon. To facilitate the effective recovery of spatial and temporal properties from PV power sequences, PV-Net reworks the gates of the gated recurrent unit (GRU) with convolution operation. For the purpose of modeling both ahead and back in time, the Conv-GRU units are layered in a bilateral manner to evade gradient vanishment across layers. In the work [11] the issues associated with STLF are addressed by employing a unique two-stream model known as STLF-Net. The first channel is built with Gated Recurrent Units (GRUs), which are intended to learn and record the long-term periodic representations of the information on power use. A stack of temporal convolutional (TC) modules is used to model the brief description and spatial interpretations that are being represented concurrently in the next channel. To allow the successful extraction of features while resolving concerns with gradient vanishing, the TC module has been designed with dilated causal convolutions and residual connection. This combination of mathematical operations has been given its name. The work [12] presented a detailed assessment of the many DL-based methodologies that are now in use and have been devised for power prediction of solar arrays and wind turbines in addition to electricity generation forecasting models. In addition to this, it covers the datasets that were used to train and test the various DL-based prediction models. This provides the research community with the ability to pick acceptable datasets to employ in their work.

3. Proposed Solution

The technical advancements made by our system include the forecasting of future power utilising a device with limited resources, which also features a lower error rate and computation that has been improved. The completion of the training process, which is required for a model to be useful in practical applications, entails a number of distinct processes. In the first step, raw data from an already existing dataset are preprocessed. This is accompanied by our innovative consecutive learning method, which is used to obtain the best-trained model, as will be discussed further down.

3.1. Data Preparation

The recording of data using smart meters involves a number of characteristics, such as date, time, active and reactive power, voltage, and others. The data pertaining to electrical power include these parameters. The smart meter serves as a hub that connects the wires of various home devices and other types of machinery to a single entire board. In most cases, the data is gathered over a period of one month or one year, during which time it develops a number of problems, including duplication, missing information, long-ranged parameters, and others. These inaccuracies are brought about by issues with the measurement apparatus itself, changes in the climate, issues with the metering, and mistakes made by personnel. Consequently, the methods of information purification and data normalization are required for the grid power data in order to achieve greater refining and more suitable outcomes.

DOI: https://doi.org/10.54216/JISIoT.060204
Received: July 26, 2022 Accepted: October 18, 2022

For the goal of training, our system makes use of a number of different preprocessing methods, which allow us to clean the data. First, we get rid of the values that aren't there and then we extract the useful data. The second step, which comes before the normalizing process, is the discovery of outliers. Its primary benefit is that it disregards any extraordinary odd digits that may have an effect on the spectrum of normalization values and pull the parameters closer to the maximum or minimum range. This is a massive benefit. The subsequent crucial step in the initialization step is normalizing, during which we utilized a variety of methods before moving on to the "standard transform choice" that is best suited for the final studies.

3.2. Predictive Model

Recurrent neural networks (RNN) are a type of deep model that belongs to a family called recurrent neural networks. These models take in as inputs not just the most recent information but also the most recent outputs. In order to accomplish this goal, the idea of a memory cell was implemented into the model to make it capable of preserving the concealed state throughout all time steps. Vanilla RNN is the simplest variant of RNNs, which calculate the, and this may be expressed as follows: In its most basic form, vanilla RNN is as follows:

$$H_t = tanh(X_t \cdot W_x + H_{t-1} \cdot W_h + b), \tag{1}$$

where W_x represent the weight parameters for the input sequence at time step t. W_h represent the weight parameters for the hidden state at the previous time step $t-1$. B is the bias parameter. Moreover, if the output is required at any time step, the hidden state can be passed to the linear layer or simply multiplied by the output parameter matrix as follow:

$$\tilde{Y}_t = H_t \cdot W_y \tag{2}$$

The structural design of RNNs, in contrast to that of convolutional networks, enables the preservation of sequential representations in text. However, the memory call in the previously mentioned vanilla RNN only performs well for short-term sequences, and it is unable to model long-term dependencies effectively due to issues with exploding gradients and/or vanishing gradients

LSTM was proposed to improve the vanilla RNN by adding a *memory cell*, C_t, store long-term sequential dependency, whereby three gating mechanisms were proposed to manage the information in the memory cell. Specifically, in particular, the *output gate*, O_t, decides the part of the information to be read out of the memory cell as an output, the *input gate*,I_t, control the adding of new inputs to the cell, and *forget gate*, F_t, was designed to control the removal of information from the cell. These gates seek to jointly control whether to recall or ignore the information in a hidden state H_t. The above calculation can be mathematically expressed as follow:

$$I_t = \sigma(X_t \cdot W_{xi} + H_{t-1} \cdot W_{hi} + b_i) \tag{3}$$

$$F_t = \sigma\left(X_t \cdot W_{xf} + H_{t-1} \cdot W_{hf} + b_f\right) \tag{4}$$

$$O_t = \sigma(X_t \cdot W_{xo} + H_{t-1} \cdot W_{ho} + b_o) \tag{5}$$

$$\tilde{C}_t = tanh(X_t \cdot W_{xc} + H_{t-1} \cdot W_{hc} + b_c) \tag{6}$$

$$C_t = F_t \odot C_{t-1} + I_t \odot \tilde{C}_t \tag{7}$$

$$H_t = O_t \odot tanh(C_t) \tag{8}$$

where σ represents the sigmoid activation, W represent the weight parameters of the connections of each gate. b_i, b_f, b_o, and b_c denote the bias vector of the input gate, forget gate, output gate, and cell respectively.

ConvLSTM [13] is a special edition of LSTM that was proposed for the predictive modeling of spatial-temporal video frames. distinctly, the input-to-state and state-to-state connectivity are designed with convolutional operation with the main aim to take the advantage of sequential learning of LSTM and spatial modeling of CNNs. the internal gats of ConvLSTM applied the convolution operation, instead of FCLs in standard LSTM. This can be mathematically represented as:

$$I_t = \sigma(X_t * W_{xi} + H_{t-1} * W_{hi} + C_{t-1} \odot W_{cI} + b_i), \tag{9}$$

$$F_t = \sigma\left(X_t * W_{xf} + H_{t-1} * W_{hf} + C_{t-1} \odot W_{cf} + b_f\right), \tag{10}$$

$$O_t = \sigma(X_t * W_{xo} + H_{t-1} * W_{ho} + C_{t-1} \odot W_{co} + b_o), \tag{11}$$

$$\tilde{C}_t = tanh(X_t * W_{xc} + H_{t-1} * W_{hc} + b_c), \tag{12}$$

Gated Recurrent Units (GRUs) were proposed by Kyunghyun Cho et al. in 2014 [14] as an extension to LSTM whereby only two gates control the memorization of the hidden state. Specifically, the reset gate R_t is accountable for regulating the volume of information that the GRU cell may have to consider from the earlier hidden state, while the update gate Z_t controls the amount of information that should be recalled from the hidden state H_t at previous time step. The internal operation of GRU can be expressed as follow:

$$R_t = \sigma(X_t \cdot W_{xr} + H_{t-1} \cdot W_{hr} + b_r) \tag{13}$$

$$Z_t = \sigma(X_t \cdot W_{xz} + H_{t-1} \cdot W_{hz} + b_z) \tag{14}$$

$$\widetilde{H}_t = tanh(X_t \cdot W_{xh} + (R_t \odot H_{t-1}) \cdot W_{hh} + b_h) \tag{15}$$

$$H_t = Z_t \odot H_{t-1} + (1 - Z_t) \odot \widetilde{H}_t \tag{16}$$

where σ represents the sigmoid activation, W represent the weight parameters of the connections of each gate. b_r, b_z, and b_h denote the bias vector of the reset gate, update gate, and hidden state respectively. The symbol $\odot$ is the Hamdard product.

3.3. System Design

In this part of the article, the implementation of the suggested Model is presented in a Smart city by using an IoT network. This section discuss the conceptual structure of a simulation of power management for a smart city. The structure is made up of three important levels, which are the grid management layer, the consumption layer, and the IoT-Enabled predictive control layer.

The process of handling energy demands and deliveries for the various smart city sectors is represented by the first layer of the architecture. In order to supply the grid stations with the necessary amounts of electrical energy, assets such as solar farms and turbines are utilized. This energy is then sold to a wide variety of end users. This layer is responsible for prediction models and other management duties that are pertinent to the situation. In addition to that, it has a cloud-fog server built in so that it may connect with entities that are consuming its services. To be more specific, the cloud-fog servers are responsible for saving and analyzing requests coming from a wide variety of industries before sending them on to the smart grid so they can answer to the respective consuming entity. The material that pertains to the sources of energy generation is not covered in this investigation because it is beyond the purview of this research.

The grid is a secure location from which the electrical energy can be distributed to a wide variety of consuming entities, each of which possesses unique characteristics such as location, responsiveness, and usage. Improving energy management typically results in a smarter grid because it makes it possible to devise a distribution strategy that is both efficient and effective, hence minimizing the amount of power that is wasted. Traditional energy networks hand out electricity to anyone who asks for it, but they don't ask any questions about how much they use, how the weather affects it, or any of the other myriad ecological variables that could lead to inefficient energy use. A smart grid, on the other hand, monitors the demands for electricity and then distributes it in the right amounts. Nevertheless, the grids frequently have poor performance because they are overburdened by excessive demands, they fail to pay attention to information that is pertinent to energy needs, and they do not have a strategy to accommodate anomalous energy requirements from various industries. The suggested framework provides a solution to this issue by incorporating an in-between cloud analysis operation. During this procedure, queries from ingesting parties are processed thru a series of particular analysis phases before being redirected to smart grids.

the process of energy management begins with the prediction of the usage of electricity, which comes prior to the scheduled transmission or delivery of needs that are specific to a particular entity. The proposed model receives, as an input, a time series of minute-level power usage from a particular entity. This time series is then used to make predictions for power use for the subsequent 60-minute timeframe. After receiving input data for three hours measured in kilowatts, the trained model is subsequently deployed through the resource-constrained device. It then makes a prediction regarding the forthcoming only one-hour usage. The amount of energy that is required is sent to cloud-fog servers, where it will be stored and analyzed with the request's histories and previous records in order to detect any anomalies. The abnormality may be seen as an unanticipated variation in the volume of requests coming from various smart urban industries. The smart grid responds to the demand and delivers the requested amount of electricity to the customer. This cycle repeats itself for all entities, and because the cloud-fog space is capable of such lightning-fast computing, the transitions between states are seamless.

DOI: https://doi.org/10.54216/JISIoT.060204
Received: July 26, 2022 Accepted: October 18, 2022

The layer of absorption. In a modern city, this surface is mostly made up of various consumption industries and entities, such as smart industry, smart transportation, smart housing, health, etc. Those entities might desire to investigate its load either overall, at the basis of a sub-entity, or at the degree of devices. The amount of energy that these various sectors and entities consume differs according to the ecological factors that are associated with them, such as their size, geography, climate, and so on. These units connected to the Internet of Things network through edge devices where the intended Model was installed to forecast the corresponding load at varying horizons, as was described before.

4. Setting up Experimentations

For this investigation, we make use of two publicly available datasets to conduct model experiments with residential and business data. To begin, nine different factors make up the Individual Household Electric Power Consumption (IHEPC) dataset. The time-series data were collected at a resolution of one minute in a residential residence in France between the years 2006 and 2010, and it was collected between those two years. The Independent System Operator New England (ISO-NE) is the subject of the second dataset. The zone covered by ISO-NE could be divided up into eight distinct regions. There are a total of seven years' worth of hourly time-series data that are utilized, starting in 2012 and continuing through 2018, for the goal of the model's training. In addition, one year's worth of data obtained (2019) is utilized for the goal of evaluating the model.

For evaluation purposes, four common metrics are adopted to estimate the predictive ability of the proposed edge intelligence model. The mathematical definition of these metrics is given as follows:

$$\text{Mean Square Error (MSE)}:=\frac{1}{N}\sum_{i=1}^{N}(y_i-\tilde{y}_i)^2 \tag{17}$$

$$\text{Root Mean Square Error } (RMSE)=\sqrt{\frac{1}{N}\sum_{i=1}^{N}(y_i-\tilde{y}_i)^2} \tag{18}$$

$$\text{Mean Absolute Error (MAE)}=\frac{1}{N}\sum_{i=1}^{N}|y_i-\tilde{y}_i| \tag{19}$$

$$\text{Mean Absolute Percentage Error (MAPE)}=\frac{1}{N}\sum_{i=1}^{N}\left|\frac{y_i-\tilde{y}_i}{y_i}\right|\times 100\% \tag{20}$$

Setup: In every single trial, the MAE is used in place of the model's loss function. The Adam optimizer is used because of the widespread recognition of its computational efficiency. An initial learning rate of 0.001 is used to initialize the optimization. This rate is then reduced by 50 after every 300 iterations. The number of participants in the training batch will be 196. To mitigate the impact of uncontrolled initialization, a 0 has been provided to each of the randomized seeds. All of the models' implementations are carried out in an environment based on Python 3.9, making use of the Keras 2.4.3 API and utilizing TensorFlow-GPU 2.2.0 as the backend. Training is carried out on a Laptop that is equipped with 48GB memory, and an NVIDIA RTX as a GPU accelerator.

5. Experimental Analysis

Through using two different datasets, we evaluate how well the suggested technique performs in comparison to other available benchmarks. In the following sections, we will explain the comparative analysis of popular technologies over residential and industrial datasets. In these sections, we will also clarify in depth how our suggested model is superior to the other model.

For the first dataset, the numerical results obtained from the proposed model are compared against the results of competing methods across different performance indicators. Since our focus s short-term predictive modeling, the comparative experimentations are performed under three scenarios namely the hour-ahead scenario (See Table 1), half-day-ahead scenario (See Table 2), and day-ahead scenario (See Table 3). The suggested model demonstrated its superiority and efficacy in modeling the spatial and temporal properties of the data as it was being learned, as evidenced by the fact that it obtained robust performance across all measures, exceeding all rival models.

DOI: https://doi.org/10.54216/JISIoT.060204
Received: July 26, 2022 Accepted: October 18, 2022

Table 1: Numerical results hour-ahead predictions for the competing methods on IHEPC data.

Models	**MSE**	**RMSE**	**MAE**	**MAPE (%)**
CNN[15]	0.3206	0.5224	0.4372	6996
EMD-LSTM [16]	0.3322	0.5988	0.3769	6615
LSTM [17]	0.4182	0.8025	0.4261	6810
GRU [18]	0.4701	0.8154	0.4648	7233
PV-Net [10]	0.4576	0.7364	0.4501	5768
Energy-Net [9]	0.3073	0.5677	0.4233	5541
Proposed	0.1827	0.5238	0.3430	3802

Table 2: Numerical results half-day-ahead predictions for the competing methods on IHEPC data.

Models	**MSE**	**RMSE**	**MAE**	**MAPE (%)**
CNN[15]	0.4327	0.6578	0.6289	6981
EMD-LSTM [16]	0.2761	0.5255	0.5737	8751
LSTM [17]	0.4265	0.6531	0.6913	8717
GRU [18]	0.5864	0.7658	0.6210	6518
PV-Net [10]	0.6567	0.8104	0.6127	6980
Energy-Net [9]	0.3391	0.5823	0.4619	6947
Proposed	0.1025	0.3202	0.4080	5708

Table 3: Numerical results day-ahead predictions for the competing methods on IHEPC data.

Models	**MSE**	**RMSE**	**MAE**	**MAPE (%)**
CNN[15]	0.4709	0.6862	0.3956	6827
EMD-LSTM [16]	0.5324	0.7297	0.4081	8406
LSTM [17]	0.6335	0.7959	0.4714	9544
GRU [18]	0.4608	0.6788	0.4490	9081
PV-Net [10]	0.6993	0.8362	0.6811	6703
Energy-Net [9]	0.4614	0.6793	0.4323	6758
Proposed	0.3085	0.5554	0.4542	0.5453

For the ISO dataset, The numerical results that are produced from the proposed model are compared with the results that are obtained from competing methods across a variety of performance indicators. Since we are primarily concerned with short-term predictive modelling, the comparative experiments are carried out under three different scenarios: the hour-ahead scenario (which can be found in Table 4), the half-day-ahead scenario (which can be found in Table 5), and the day-ahead scenario (which can be found in Table 6). The proposed model showed its supremacy and effectiveness in analysing the temporal and spatial characteristics of the data while it was being learned, as indicated by the fact that it obtained high efficiency throughout all indicators, surpassing all competing models in the process. This was reflected in the fact that it procured the highest score possible for each of the metrics.

DOI: https://doi.org/10.54216/JISIoT.060204
Received: July 26, 2022 Accepted: October 18, 2022

Table 4: Numerical results hour-ahead predictions for the competing methods on ISO-NE data.

Models	**MSE**	**RMSE**	**MAE**	**MAPE**
CNN[15]	0.5958	0.7719	0.4187	42.40
EMD-LSTM [16]	0.4667	0.6832	0.4285	57.50
LSTM [17]	0.7603	0.8720	0.4528	61.74
GRU [18]	0.4845	0.6961	0.4630	55.20
PV-Net [10]	0.3739	0.6115	0.3319	31.85
Energy-Net [9]	0.3083	0.5552	0.4335	26.98
Proposed	0.1518	0.3896	0.2725	23.75

Table 5: Numerical results half-day-ahead predictions for the competing methods on ISO-NE data.

Models	**MSE**	**RMSE**	**MAE**	**MAPE**
CNN[15]	0.3997	0.6322	0.4459	41.12
EMD-LSTM [16]	0.2956	0.5437	0.3912	52.67
LSTM [17]	0.6115	0.7820	0.2679	57.34
GRU [18]	0.4221	0.6497	0.4388	56.44
PV-Net [10]	0.3939	0.6276	0.1674	15.11
Energy-Net [9]	0.1000	0.3162	0.1993	29.40
Proposed	0.1840	0.4290	0.3017	25.71

Table 6: Numerical results day-ahead predictions for the competing methods on ISO-NE data.

Models	**MSE**	**RMSE**	**MAE**	**MAPE**
CNN[15]	0.6237	0.7897	0.6226	67.67
EMD-LSTM [16]	0.7126	0.8442	0.5288	78.02
LSTM [17]	1.0800	1.0392	0.7199	67.16
GRU [18]	0.8588	0.9267	0.6908	70.89
PV-Net [10]	0.5357	0.7319	0.4460	51.12
Energy-Net [9]	0.5097	0.7139	0.6036	45.39
Proposed	0.2453	0.4953	0.5807	51.24

6. Summary and Conclusions

IoT devices are becoming increasingly important in the resolution of a wide variety of issues on a regular basis, and there are a great many difficulties that they can solve. The majority of applications for these devices are in pattern recognition and machine learning, particularly for smart tracking and the identification of activities. It is uncommon for researchers to look into predicting the future of energy and managing it appropriately with the help of IoT devices. In particular, DL and its related concepts are not inferred to the edge of the network. For the purpose of our research, we employed light cognitively smart strategies that are operational over resource-constrained devices to predict the future consumption of energy and make its administration more efficient.

DOI: https://doi.org/10.54216/JISIoT.060204
Received: July 26, 2022 Accepted: October 18, 2022

Many different avenues can be investigated in the future to make this operate better. In the first place, as we were training and assessing our model, we planned to take into account more complicated characteristics and noisy data (such as environmental factors, user behavioral information, and so on). Second, there is the possibility of investigating decentralized, distributed, or federated training as a means of overcoming the ever-increasing volume of data and the growing physical separation between the various sources of information. Third, to make up for the scarcity of data, one could investigate the possibility of semi-supervised or self-supervised training. Fourth, the use of deep learning which is allowed by blockchain technology is the most effective method for resolving concerns over the confidentiality of the demand prediction model. Fifth, we want to expand our work by leveraging online learning to encourage capacity planning as a utility. In this scenario, the Energy-parameters Net's would need to be updated following the current power demands.

References

[1] A. Heidari, N. J. Navimipour, and M. Unal, “Applications of ML/DL in the management of smart cities and societies based on new trends in information technologies: A systematic literature review,” *Sustain. Cities Soc.*, vol. 85, p. 104089, Oct. 2022, doi: 10.1016/j.scs.2022.104089.

[2] M. O. Arowolo, R. O. Ogundokun, S. Misra, J. Oluranti, and A. F. Kadri, “K-Nearest Neighbour Algorithm for Classification of IoT-Based Edge Computing Device,” in *Internet of Things*, 2022.

[3] I. Akhtar, S. Kirmani, M. Ahmad, and S. Ahmad, “Average Monthly Wind Power Forecasting Using Fuzzy Approach,” *IEEE Access*, 2021, doi: 10.1109/ACCESS.2021.3056562.

[4] Aasim, S. N. Singh, and A. Mohapatra, “Data driven day-ahead electrical load forecasting through repeated wavelet transform assisted SVM model,” *Appl. Soft Comput.*, 2021, doi: 10.1016/j.asoc.2021.107730.

[5] G. Kim and J. Hur, “A short-term power output forecasting based on augmented naïve bayes classifiers for high wind power penetrations,” *Sustain.*, 2021, doi: 10.3390/su132212723.

[6] X. Yao, X. Fu, and C. Zong, “Short-Term Load Forecasting Method Based on Feature Preference Strategy and LightGBM-XGboost,” *IEEE Access*, vol. 10, pp. 75257–75268, 2022, doi: 10.1109/ACCESS.2022.3192011.

[7] W. Xia, Y. Mita, and T. Shibata, “A Nearest Neighbor Classifier Employing Critical Boundary Vectors for Efficient On-Chip Template Reduction,” *IEEE Trans. Neural Networks Learn. Syst.*, 2016, doi: 10.1109/TNNLS.2015.2437901.

[8] K. Yan, X. Zhou, and J. Chen, “Collaborative deep learning framework on IoT data with bidirectional NLSTM neural networks for energy consumption forecasting,” *J. Parallel Distrib. Comput.*, 2022, doi: 10.1016/j.jpdc.2022.01.012.

[9] M. Abdel-Basset, H. Hawash, R. K. Chakrabortty, and M. Ryan, “Energy-Net: A Deep Learning Approach for Smart Energy Management in IoT-Based Smart Cities,” *IEEE Internet Things J.*, 2021, doi: 10.1109/JIOT.2021.3063677.

[10] M. Abdel-Basset, H. Hawash, R. K. Chakrabortty, and M. Ryan, “PV-Net: An innovative deep learning approach for efficient forecasting of short-term photovoltaic energy production,” *J. Clean. Prod.*, 2021, doi: 10.1016/j.jclepro.2021.127037.

[11] M. Abdel-Basset, H. Hawash, K. Sallam, S. S. Askar, and M. Abouhawwash, “STLF-Net: Two-stream deep network for short-term load forecasting in residential buildings,” *J. King Saud Univ. - Comput. Inf. Sci.*, vol. 34, no. 7, pp. 4296–4311, Jul. 2022, doi: 10.1016/j.jksuci.2022.04.016.

[12] S. Aslam, H. Herodotou, S. M. Mohsin, N. Javaid, N. Ashraf, and S. Aslam, “A survey on deep learning methods for power load and renewable energy forecasting in smart microgrids,” *Renewable and Sustainable Energy Reviews*. 2021, doi: 10.1016/j.rser.2021.110992.

[13] X. Shi, Z. Chen, H. Wang, D. Y. Yeung, W. K. Wong, and W. C. Woo, “Convolutional LSTM network: A machine learning approach for precipitation nowcasting,” 2015.

DOI: https://doi.org/10.54216/JISIoT.060204
Received: July 26, 2022 Accepted: October 18, 2022

[14] K. Cho *et al.*, "Learning phrase representations using RNN encoder-decoder for statistical machine translation," 2014, doi: 10.3115/v1/d14-1179.

[15] C. Tian, J. Ma, C. Zhang, and P. Zhan, "A deep neural network model for short-term load forecast based on long short-term memory network and convolutional neural network," *Energies,* 2018, doi: 10.3390/en11123493.

[16] H. Zheng, J. Yuan, and L. Chen, "Short-Term Load Forecasting Using EMD-LSTM neural networks with a xgboost algorithm for feature importance evaluation," *Energies,* 2017, doi: 10.3390/en10081168.

[17] W. Kong, Z. Y. Dong, Y. Jia, D. J. Hill, Y. Xu, and Y. Zhang, "Short-Term Residential Load Forecasting Based on LSTM Recurrent Neural Network," *IEEE Trans. Smart Grid,* 2019, doi: 10.1109/TSG.2017.2753802.

[18] T. Han, K. Muhammad, T. Hussain, J. Lloret, and S. W. Baik, "An Efficient Deep Learning Framework for Intelligent Energy Management in IoT Networks," *IEEE Internet Things J.*, 2021, doi: 10.1109/JIOT.2020.3013306.

DOI: https://doi.org/10.54216/JISIoT.060204
Received: July 26, 2022 Accepted: October 18, 2022