IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received November 4, 2020, accepted November 14, 2020, date of publication November 24, 2020, date of current version December 24, 2020.

Digital Object Identifier 10.1109/ACCESS.2020.3040083

HSIC Bottleneck Based Distributed Deep Learning Model for Load Forecasting in Smart Grid With a Comprehensive Survey

MD. AKHTARUZZAMAN[1,2], MOHAMMAD KAMRUL HASAN[1], (Senior Member, IEEE), S. RAYHAN KABIR[2], SITI NORUL HUDA SHEIKH ABDULLAH[1], MUHAMMAD JAFAR SADEQ[2], AND EKLAS HOSSAIN[3], (Senior Member, IEEE)

[1]Center for Cyber Security, Faculty of Information Science and Technology, Universiti Kebangsaan Malaysia, Bangi 43600, Malaysia
[2]Department of Computer Science and Engineering, Asian University of Bangladesh, Dhaka 1230, Bangladesh
[3]Department of Electrical Engineering and Renewable Energy, Oregon Institute of Technology, Klamath Falls, OR 97601, USA

Corresponding author: Mohammad Kamrul Hasan (mkhasan@ukm.edu.my)

This work was supported in part by the National University of Malaysia under Research Grant GGPM-2020-028, and FRGS-1-2019-ICT-UKM-02.

ABSTRACT Load forecasting is a vital part of smart grids for predicting the required electrical power using artificial intelligence (AI). Deep learning is broadly used for load forecasting in the smart grid using the artificial neural network (ANN). Generally, computing the deep learning in the smart grid requires massive data aggregation or centralization and significant computational time. This paper presents a survey of deep learning-based load forecasting techniques from 2015 to 2020. This survey discusses the studies based on their deep learning techniques, Distributed Deep Learning (DDL) techniques, Back Propagation (BP) based works, and non-BP based works in the load forecasting process. Consequent to the survey, it was determined that data aggregation dependency would be beneficial for reducing computational time in load forecasting. Therefore, a conceptual model of DDL for smart grids has been presented, where the HSIC (Hilbert-Schmidt Independence Criterion) Bottleneck technique has been incorporated to provide higher accuracy.

INDEX TERMS Energy management, smart grid application, load forecasting, IoT, smart metering, distributed neural network, distributed machine learning, distributed deep learning.

NOMENCLATURE

ABBREVIATIONS

AAL	Ambient Assistive Living
AE	Auto-encoder
AGI	Artificial General Intelligence
AI	Artificial Intelligence
AIFIS	Adaptive Neuro-Fuzzy Inference System
ANN	Artificial Neural Network
ARIMA	Auto-regressive Integrated Moving Average
ART	Adaptive Resonance Theory
ARTMAP	Predictive Adaptive Resonance Theory (ART) Mapping
BLSTM	Bi-directional Long Short-term Memory
BP	Back-Propagation
BPNN	Back-Propagation Neural Network
BRNN	Bi-directional Recurrent Neural Network
Cascade NN	Cascade Neural Network
CCHP	Combined Cooling, Heating, and Power
CNN	Convolutional Neural Network
CoT	Cloud of Things
CSRM	Clear Sky Solar Radiation Model
CVNN	Complex Valued Neural Network
DAI	Decentralized Artificial Intelligence
DAE	Denoising Auto-encoder
DE-RELM	Differential Evolution Recurrent Extreme Learning Machine
DBN	Deep Belief Network
DC	Direct Current
DDL	Distributed Deep Learning
DML	Distributed Machine Learning
DNN	Deep Neural Network
DQN	Deep Q-Network
DSG	Distributed Smart Grid
EEMD	Ensemble Empirical Mode Decomposition
ResNet	Residual Neural Network

The associate editor coordinating the review of this manuscript and approving it for publication was Sanjeevikumar Padmanaban.

ELM	Extreme Learning Machine
Elman ANN	Elman Artificial Neural Network
EMD	Empirical Mode Decomposition
ENN	Ensemble Neural Network
ESAENARX	Efficient Sparse Auto-encoder Nonlinear Autoregressive Network with Exogenous
ESN	Echo State Network
F-MODM	Fuzzy Multi-Objective Decision Making
FDD	Frequency Domain Decomposition
FFNN	Feed-forward Neural Network
GMDL	Group Method of Data Handling
GPU	Graphics Processing Unit
Grey-NN	Grey / Gray Neural Network
GRNN	General Regression Neural Network
GRU	Gated Recurrent Unit
GWDO	Genetic Wind-Driven Optimization
HCI	Human Computer Interaction
HSIC	Hilbert-Schmidt Independence Criterion
HTM	Hierarchical Temporal Memory
IMF	Intrinsic Mode Function
IoT	Internet of Things
IRB	Iterative Res-Blocks
KF	Kalman Filtering
KNN	K-nearest Neighbor
KNN-ANN	K-nearest Neighbor based Artificial Neural Network
KPCA	Kernel Principal Component Analysis
LDPC	Low-density parity-check code
LRD	Load Range Discretization
LSM	Liquid State Machine
LSTF	Long Short-term Memory
LTLF	Long-term Load Forecasting
MI	Mutual Information
MI-ANN	Mutual Information based Artificial Neural Network
MLP	Multi-layer Perceptron
MLR	Multivariable Linear Regression
MTLF	Mid-term Load Forecasting
NAR	Non-linear Auto-regressive
NAR-ANN	Non-linear autoregressive Neural Network
NARX	Nonlinear Auto-regressive Network with Exogenous inputs
Neuro-evolution	Evolutionary Neural Network
Neuro-fuzzy	Fuzzy Neural Network
PCA	Principal Component Analysis
Physical-NN	Physical Neural Network
PNN	Probabilistic Neural Network
PSR	Phase Space Reconstruction
PV	Photo-voltaic
RBF	Radial Basis Functions
RBM	Restricted Boltzmann Machine
RES	Renewable Energy Source
RFNN	Recurrent Fuzzy Neural Network
RNN	Recurrent Neural Network
S2S	Sequence to Sequence
SAE	Sparse Auto-encoder
SDPSO	Switching Delayed Particle Swarm Optimization
SGD	Stochastic Gradient Descent
SLP	Single-layer Perceptron
SRWNN	Self-Recurrent Wavelet Neural Network
SSA	Singular Spectrum Analysis
STLF	Short-term Load Forecasting
SVM	Support Vector Machine
SVR	Support Vector Regression
SWEMD	Sliding Window Empirical Mode Decomposition (EMD)
UKF	Unscented Kalman Filter
UQR	Unified Quantile Regression
VSTLF	Very Short-term Load Forecasting
WNN	Wavelet Neural Network
WPT	Wavelet Packet Transform

I. INTRODUCTION

Artificial Intelligence (AI) is a fundamental theme of future technology research and development. In many nations, smart grids are being developed to be an intelligent layer to improve power distribution, control, and generation [1]. The smart grids are being established with intelligent devices and sensors to computerize and improve various applications' productivity, including metering distribution. The machine learning-based smart meter system contributes effectively to the Ambient Assistive Living (AAL) area for detecting daily living activities [2]. Machine learning has been used with smart meters for improving end-user load modeling machine learning [3]. AI also combined with edge computing and edge analytics in smart power meters [4]. At present, the smart grid, smart home, and smart meter success depend on AI and communication security [5].

Machine learning, deep learning, and swarm intelligence are the common mechanisms of AI that are broadly used in smart grids for different purposes. Machine learning is being used for analyzing big data in the smart grid and for security aspects of the Internet of Things (IoT) in the smart grid [6]. Deep learning (Figure 1) is a family of machine learning based on artificial neural networks (ANN) that are used in the smart grid for predicting load forecasting, price forecasting, solar forecasting, power quality, and other purposes [7]–[9]. This article focuses on deep learning for load forecasting from the distributed computing perspective. Figure 1 shows a section of deep learning where ANN is used.

Load forecasting is an essential outcome of a smart grid system. It predicts the short-term, medium-term, and long-term demand for electrical power to the users. Deep learning is widely used for load forecasting in smart grids [10]. Although a subset of machine learning, deep learning is more useful than other traditional machine learning algorithms. It also facilitates the use of other machine learning algorithms

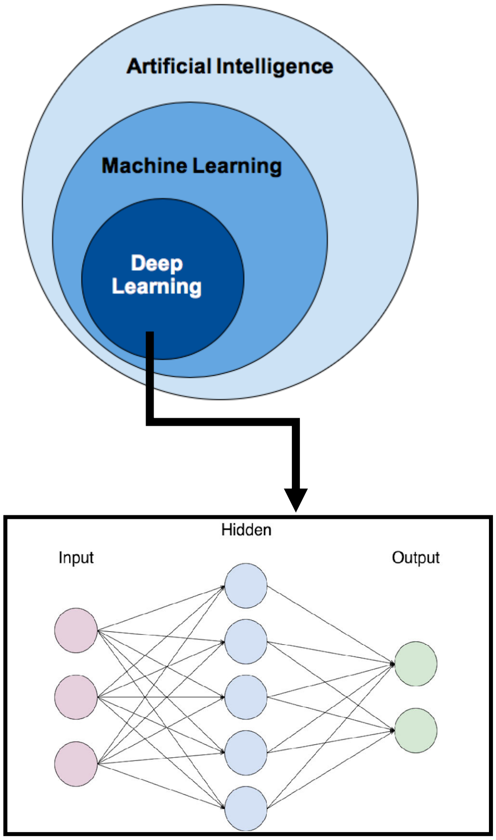

FIGURE 1. Deep learning is a subfield of machine learning and AI and it is based on ANN computation. The key feature of deep learning is that ANN works in any deep learning process along with other machine learning techniques.

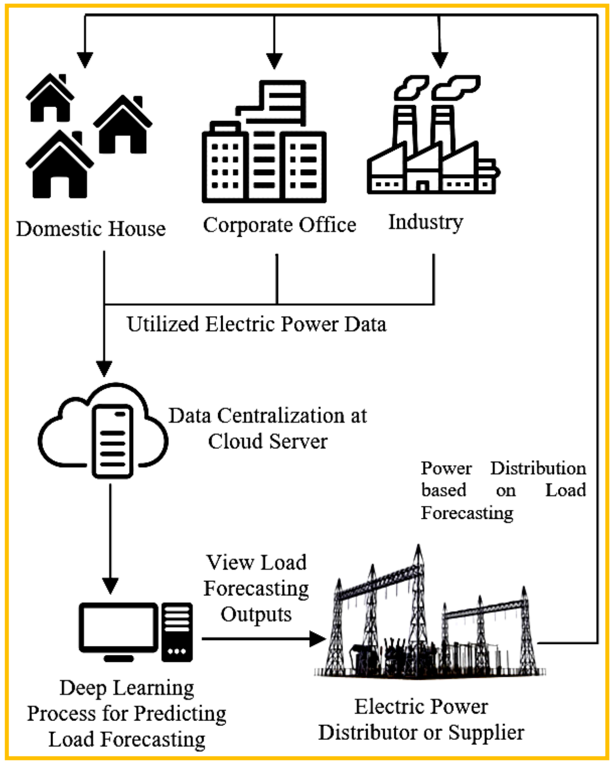

FIGURE 2. Load forecasting and electrical power distribution in a smart grid system. Here, the cloud server contains different utilized electric power data from different areas. Based on the data and load forecasting process, a power distributor can predict the future load demand of these areas.

in conjunction with ANN for better results. For that reason, it is well known that it provides precise accuracy. Various techniques have been combined with ANN for smart grids, including Decision Tree, Random Forrest, Support Vector Machine, K-nearest Neighbor (KNN), and other machine learning algorithms, obtaining better outcomes for power management, load forecasting, and other purposes [11], [12].

Electrical power distribution companies will be particularly benefited from smart grids by determining better power distribution among the consumers or clients (home, industry, and corporate office). Currently, different electrical power utilization data are aggregated centrally. Deep learning performs load forecasting from this huge data, and hence the power distributors can predict the demand for electrical power distribution and provide power to different consumers accordingly. The process is shown in Figure 2.

In smart grids, different electrical power utilization data are aggregated in centralized cloud servers or cloud storage. Due to this large amount of data, the computation time for deep learning is high [65], [77], [96]. Therefore, various machine learning techniques are being applied to deep learning to enhance computation time and efficiency [13]. Distributed Machine Learning (DML) [18] is one technique that is already being used to enhance computation time [63], so Distributed Deep Learning (DDL) [14], a subset of DML, may also be applied to reduce the data being processed and may obtain useful and valid outputs compared to deep central learning. Smart grids are large-scale network systems consisting of physical power and an information network. When considering distributed storage and distribute power distribution control, Distributed Smart Grid (DSG) [15] appears to be a reasonable alternative to investigate in such systems. Distributed Artificial Intelligence or Decentralized Artificial Intelligence (DAI) [16], [17] can thus be employed in smart grids for obtaining effectual output.

DML is normally used in decentralized and distributed systems such as IoT and wireless communication [18]. DDL has become common in computing the training processes of multiple devices by using ANN and other machine learning algorithms [19]. Therefore, DDL has potential applications in future smart grids for reducing ANN computation time and decreasing huge data centralization dependency.

At the same time, IoT is a very essential network mechanism in the smart grid system [20]. Cloud of Things [21] is a cloud layer of IoT used for monitoring, managing, and analyzing data and information provided by using a cloud server [22], [23]. A deep learning-based CoT model has

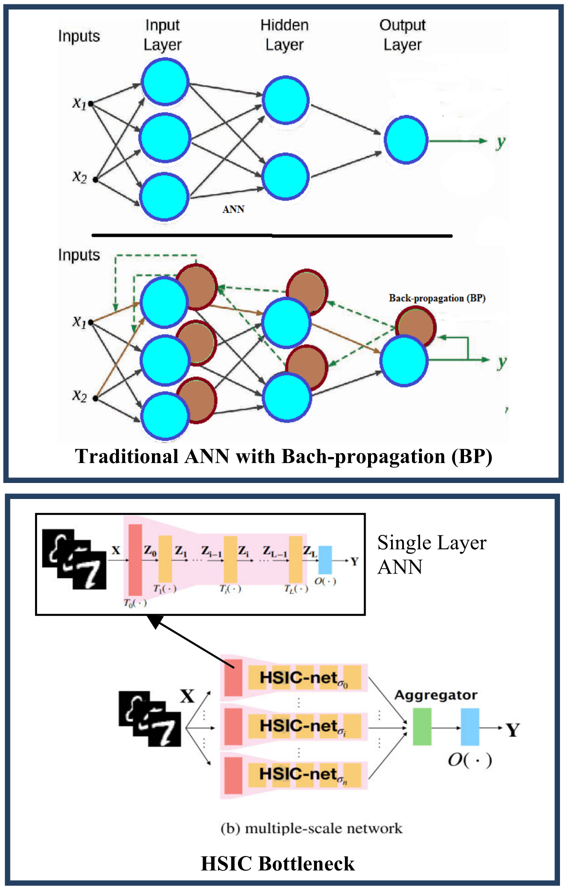

FIGURE 3. **Traditional ANN contains BP process where every neuron is connected. HSIC Bottleneck works without BP that contains multiple ANN layers where one neuron is not connected to another layer's neuron.**

been implemented to identify the traffic in a heterogeneous network [24]. Therefore, DDL may be integrated with CoT for better performance in a smart grid.

Typically, Back-Propagation (BP) method is used in deep learning processes to find results with good accuracy. BP has some disadvantages such as weight transport, update locking, vanishing gradients, and exploding gradients [25]–[28]. In smart grids, BP takes additional computation time on top of that caused by huge centralized data aggregates [96]. So, it would be beneficial to reduce the computation times for load forecasting in the smart grid. Due to the disadvantages of BP, non-BP based works have been proposed for load forecasting [44], [46], [49]. In other application areas, alternative approaches to BP are also being investigated [28], [151], [152], [153]. The BP process with traditional ANN has been illustrated in Figure 3 [162].

HSIC (Hilbert-Schmidt Independence Criterion) Bottleneck [28] is one such novel deep learning method that provides good accuracy results without BP. It provides single-layer ANN with multiple scales ANN network with no connection among the single-layer ANNs. The outputs of the single-layer ANNs are aggregated. This is potentially a distributed training process and solves some major disadvantages of BP [28]. The difference between HSIC Bottleneck and traditional ANN is shown in Figure 3. HSIC Bottleneck has been developed based on Hilbert-Schmidt Independence Criterion (HSIC) [29], Information Bottleneck [30], and Dependency Bottleneck [31] theorems. The information bottleneck method has been used in 5G networks for decoding LDPC (Low-density parity-check code) codes [32] and wireless communication for broadcasting data [33]. In conjunction with HSIC, DDL can reduce the overall deep learning computation time and decrease data centralization dependency if multiple ANN processes were to occur at the distributed smart meters. Combining HSIC bottleneck and DDL would appear to be an extremely suitable technique for load forecasting in smart grids. From a C2C (Consumer to consumer) information quality [166] perspective, this manuscript creates a better impact.

This manuscript presents a comprehensive survey of deep learning techniques of load forecasting from 2015 to 2020. After presenting the survey, we have discussed a model that is based on current major challenges in load forecasting. The contributions of this contribute in six ways:

- This article provides a comprehensive survey on deep learning-based load forecasting techniques including BP and non-BP-based works. Hence, some current issues and possible research scopes are highlighted. The outcome of the survey presented the major focused areas and major ANN techniques of the load forecasting processes.
- The outcome of the survey highlighted the major used ANN strategies, the most important load forecasting type, major focused areas, and current major challenges.
- Based on the survey, a conceptual model has been proposed that may ameliorate some present challenges of load forecasting. Such as high computational time, large data, Data requirement or limited data, and Additional BP process.
- Because of huge data aggregation at the cloud server, load forecasting takes high computational time. The model presents a distributed deep learning-based distributed load forecasting approach for reducing the high computational time.
- Mitigating huge data aggregation or centralization dependency of the cloud server and reducing overall ANN computation time in the load forecasting process.
- Avoiding additional BP processes and BP complexity at ANN in load forecasting by using the HSIC Bottleneck-based distributed neural network.

The remaining of the paper covers a survey of various deep learning techniques for load forecasting in the smart grid (Section II), including major focus areas of load forecasting research. Section III contains current research scopes and problem statements. Hence, a conceptual model for ameliorating the current issues is presented in Section IV. The outcomes of this research are presented in Section V Finally, in Section VI, the conclusions of the paper and future works are presented.

II. DEEP LEARNING TECHNIQUES IN LOAD FORECASTING

A. CLASSIFICATION OF LOAD FORECASTING

Before delving into the deep learning techniques of load forecasting, it is necessary to classify load forecasting. Generally, there are four types of load forecasting: long-term load forecasting (LTLF), mid-term or medium-term load forecasting (MTLF), short-term load forecasting (STLF), and very short term load forecasting (VSTLF) [34] (Figure 4). VSTLF helps to calculate the electric loads of half-hour or a few minutes [35]. The purpose of the STLF is to measure the load from one hour to the next few weeks [36]. MTLF predicts the load from one month to one or more years [37]. LTLF is a vital load forecasting approach in electric power generation, transmission, and distribution. Normally, LTLF predicts the electricity load demand for up to ten years and it also helps in decision-making for more than ten years [38].

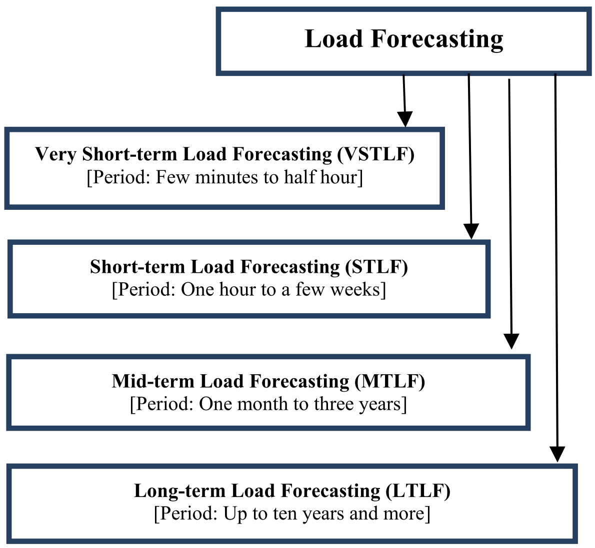

FIGURE 4. Classification of load forecasting.

B. DIFFERENT ANN TECHNIQUES IN DEEP LEARNING BASED LOAD FORECASTING

ANN is the primary and fundamental portion of deep learning. For developing any deep learning process, researchers or developers need to integrate different kinds of machine learning algorithms with ANN structure. Table 1 demonstrates different types of ANN that have been used in proposed load forecasting systems, along with their characteristics. Table 2 displays some relevant works from 2015 to 2020 related to deep learning-based load forecasting in smart grids.

Feed-forward Neural Network (FFNN) is the basic concept of ANN. FFNN is a forward propagation technique without BP. FFNN has been used with Grasshopper Optimization Algorithm [42], Fuzzy Multi-Objective Decision Making (F-MODM) [44], Particle Swarm Optimization [100], Copula [123], and other machine learning mechanisms in various load forecasting models. KNN-ANN is the combination of

TABLE 1. Different types of ANN in load forecasting.

ANN Type	Characteristics
FFNN	ANN without BP
SLP	ANN without hidden layer
BPNN	Combination of Multi-layer Perceptron (MLP) based FFNN with BP
DNN	ANN with multiple hidden layers
Distributed Neural Network	DDL based ANN where multiple parts of an ANN are distributed
GMDH	A type of inductive algorithm with neurons
PNN	ANN using Probability Distribution Function (PDF)
CNN	Convolutional layer with DNN
Cascade NN	It can learn any finite input-output reference.
CVNN	It uses complex arithmetic rules
ELM	No need for parameters tuning of hidden nodes
MI-ANN	Combination of ANN and Mutual Information
KNN-ANN	Combination of KNN and ANN
NAR-NN	It predicts time-series from the past values
AE	Representation learning based unsupervised ANN
SAE	A type of AE which contains a small number of active hidden unit
DAE	Stochastic version of AE
WNN	Combination of ANN and wavelet analysis
WaveNet	Raw audio waveform based DNN
DQN	Combination of DNN and reinforcement learning
ENN	Ensemble learning based ANN
RNN	Relationship between nodes like a graph
LSTM	RNN architecture based ANN which provides the feedback connections
BRNN	Two opposite direction based hidden layers are connected at RNN
BLSTM	Two opposite direction based hidden layers are connected at LSTM
GRU	Look like LSTM but it works without an output gate
ESN	A type of RNN with sparsely connected hidden layers
RBM	A generative and stochastic ANN

TABLE 1. *(Continued.)* **Different types of ANN in load forecasting.**

Elman ANN	A type of RNN which contains input neurons more than hidden neurons
	RNN based architecture with a machine translation
Neuro-fuzzy	Combinations of fuzzy logic and ANN
RFNN	It is based on the RNN structure to connect fuzzy logic.
ANFIS	Takagi–Sugeno fuzzy-based Neuro-fuzzy ANN
LSM	A kind of Spiking Neural Networks (SNN) and contains a large number of neurons
Physical-NN	ANN with physical hardware
HTM	Tree-shaped hierarchy represented in the ANN
ART network	Several ANN using for both supervised and unsupervised learning
Grey-NN	Grey model-based ANN
Neuro-evolution	Producing ANN by using evolutionary algorithms
RBF network	Radial basis functions (RBF) based ANN
DBN	Generates pre-training process
ResNet	It can skip one or more layer connections
GRNN	Single-pass based learning process where no need BP

K-nearest Neighbor (KNN) and ANN which is used with FFNN in the hydro-thermal unit generation-based VSTLF process [75]. BPNN is a version of FFNN where BP is connected to a Multi-layer Perceptron (MLP) based FFNN. A low and medium voltage smart grid [144] is an important matter for predicting accurate load forecasting. BPNN is broadly used in low voltage grids [43], load shedding [45], and PV (Photovoltaic) generation [104]. FFNN and BPNN based hybrid ANN have been proposed for predicting MTLF and STLF in low voltage smart grid [43]. BPNN with fuzzy logic has also been proposed for the medium voltage at self-optimized STLF [136]. Deep Neural Network (DNN) is a mechanism built upon FFNN where there are multiple hidden layers. Iterative ResBlocks (IRB) based DNN has been developed for individual residential loads [58]. BP algorithm has been utilized with DNN [116]. Convolutional Neural Network (CNN) is an enhancement on DNN with the additional convolutional layer and it is another popular ANN technique in load forecasting. Although CNN is broadly used in image-related work, it is also used for probabilistic load forecasting [47], over-fitting issues [65], feature redundancy, and environment-friendly smart grid [77]. CNN-based encoded images have been applied for load forecasting [62].

From the works surveyed, most have combined different types of ANN (hybrid ANN) [145] for solving various complex issues in load forecasting. A recent review study of load forecasting has been indicated the vital parts of deep learning models with single and hybrid models [9]. Single-layer Perceptron (SLP), which has no hidden layer, was used for increasing SLTF speed in a hybrid SLP with Recurrent Neural Network (RNN), Gated Recurrent Unit (GRU), Encoder-Decoder Architecture, and Auto-encoder (AE) [98]. RNN is a directed graph-oriented ANN and hybrid RNN is broadly used in various load forecasting processes for data cleaning [59], big data [71], computational time, high dimensional data [133], and other purposes. LSTM and GRU are techniques built upon RNN. LSTM was the most used ANN approach in load forecasting, demonstrating better performance in different complex situations such as low-high-frequency components [53], probabilistic load forecasting [55], PV generation [66], CCHP (Combined Cooling, Heating, and Power) system [81], Hybrid energy systems, [87], etc. A combination of three ANNs, namely LSTM, BPNN, and DQN (Deep Q-Network), have been used for establishing a similar day selection based STLF model [40]. WNN (Wavelet Neural Network) with BPNN has been used in MTLF to determine the reduction of the unavoidable stochastic part [41]. Neuro-fuzzy is a fuzzy logic-based ANN technique and has been applied to RNN, LSTM, ELM (Extreme Learning Machine) in STLF [91], and other ANNs for improving some issues in load forecasting. Table 1 presents various other types of ANN, such as Ensemble Neural Network (ENN), Deep Belief Network (DBN), and ART network, which have been applied for the various load forecasting model. Based on Table 1, Table 2 demonstrates the applied techniques, deep learning model, focus area, forecasting type, and the major contribution of different research works from 2015 to 2020.

After reviewing 105 relevant deep learning-based load forecasting research works, it was found that STLF was the most important forecasting area, with 88 works on STLF. Accordingly, MTLF and VSTLF are considered in 24 and 17 works respectively. LTLF has fewest the research works with a total of 10. Some works focused on multiple load forecasting processes. In one, FFNN, DNN, Grasshopper Optimization algorithm-based load forecasting has been developed for MTLF and STLF [42]. In another work, Neuro-evolution and RNN based load forecasting was applied for both VSTLF and STLF [133].

It was also found that LSTM is the most widely used ANN technique for load forecasting. For purposes such as improving feature selection [49], feature extraction [78], computational time [84], etc., LSTM has been found to provide more impactful outputs than other ANN techniques [40], [50], [53] (Table 4). BPNN was the next most used technique for load forecasting. Although CNN is widely used for visual

TABLE 2. Most relevant deep learning based load forecasting works from 2015-2020.

Ref. No.	Year	Applied Techniques	Developed Model	Forecasting Type	Focus Area	Major Contribution
[39]	2020	Takagi–Sugeno model, RFNN, Fuzzy rules, Nonlinear System, BP	Takagi–Sugeno based RFNN model for LTLF	LTLF	1) Week period and holidays 2) Violation of Gaussian Distribution 3) Meteorological data (load, temperature, humidity, etc.) 4) Weather factor	Retaining the temperature data among the weather stations in the LTLF process and managing the holiday features and violation of Gaussian distribution.
[40]	2020	LSTM, BPNN, DQN, Reinforcement Learning	Reinforcement Learning based similar day selection model in STLF algorithm	STLF	1) Similar day selection	Establishes a similar day selection based STLF model by using the reinforcement learning on the BPNN for Korea.
[41]	2020	WNN, BPNN, Cuckoo Search Algorithm, Singular Spectrum Analysis (SSA)	WNN and cuckoo search based load forecasting for two buildings paradigm	MTLF	1) Wavelet Disintegration 2) Unavoidable Stochastic Part	Exploring the wavelet disintegration in load forecasting to know the reduction of the unavoidable stochastic part.
[42]	2020	FFNN, DNN, Grasshopper Optimization Algorithm (GOA), Regressive model, BP	Multi-layer ANN and GOA based load forecasting	MTLF and STLF	1) Different hours and different days 2) Weather factors	The model can useable to forecast loads at the daily and hourly over a month using GOA and Multi-layer ANN.
[43]	2020	FFNN, BPNN	FFNN and BPNN for LTLF of total electric power demand	LTLF	1) Low or medium voltage 2) Learning error	Reduction of mean square error (MSE) at load forecasting process in the low voltage grid.
[44]	2020	FFNN, Neuro-fuzzy, Fuzzy Multi-Objective Decision Making (F-MODM)	Combined F-MODM and FFNN model to predict electric load for 1 hour ahead.	VSTLF	1) Weather factors 2) Thermal unit generation	Develops a weather data based load forecasting that combines F-MODM and FFNN for 1 hour ahead.
[45]	2020	Aaptation Mechanism, BPNN	Advance BP based load forecasting	STLF and MTLF	1) Load shedding 2) Cost saving	Predicts advance BP based load forecasting on to reduce the load shedding, power loss and cost generation.
[46]	2020	Net-load strategy, Wavelet Transform, Cascade NN, DNN	Integration of wavelet transform and DNN for net-load forecasting	STLF and MTLF	1) PV generation and PV penetration 2) Wind generation 3) Energy price 4) Softmax layer	Integrating wavelet transform and DNN for measuring photovoltaic (PV) generation, wind generation, and load demand.
[47]	2020	CNN, Load Range Discretization (LRD), BP	CNN with LRD for predicting load forecasting	STLF and MTLF	1) Probability distribution 2) Probabilistic load forecasting	Generates the probability distribution in load forecasting using CNN and LRD.
[48]	2020	RNN, GRU, BP	RNN based electric demand forecasting for residential buildings	VSTLF, STLF and MTLF	1) Load demand of residential buildings	Combining RNN and GRU for predicting the load demand of the residential buildings over the short to the medium period.
[49]	2020	MI-ANN, LSTM, RBM, Genetic Wind-Driven Optimization (GWDO)	GWDO heuristic algorithm based load forecasting process	STLF	1) Feature selection 2) Wind generation	Combines MI-ANN, GWDO and RBM for STLF which systematically works on both linear and non-linear systems.
[50]	2020	ENN, Bagging method, Bosting method, Ensemble Learning, BP	STLF method by using Bagged-boosted ANN	STLF	1) Reduction of bias 2) Variance of load forecasting	Applies bagged-boosted ANN in load forecasting to reduce the bias and variance.

TABLE 2. ***(Continued.)*** **Most relevant deep learning based load forecasting works from 2015-2020.**

Ref. No.	Year	Applied Techniques	Developed Model	Forecasting Type	Focus Area	Major Contribution
[51]	2020	Neuro-fuzzy, Singular Spectrum Analysis (SSA), ART networks, Fuzzy ARTMAP (Predictive ART), BP	Fuzzy ARTMAP ANN and SSA based load forecasting process in disaggregated levels	STLF	1) Disaggregated levels 2) Cost saving (Computation cost) 3) Data requirement or limited data	Combining Fuzzy logic, ARTMAP and SSA in disaggregated levels of load forecasting for reducing computational cost and data requirements.
[52]	2020	MI-ANN, RBM, Jaya algorithm, BP	Entropy MI-ANN based feature selection in MTLF	MTLF	1) Data and feature redundancy 2) Consumption behaviors 3) Feature selection	Removes the relevance and redundancy of feature selection in MTLF and improves the forecasting validity.
[53]	2020	LSTM, Ensemble Empirical Mode Decomposition (EEMD), Multivariable Linear Regression (MLR)	MLR and LSTM based hybrid STLF	STLF	1) Low and high frequency components 2) Big data	Analyzing electric load from the large data and low to high-frequency components.
[54]	2020	BPNN, WNN, Kalman Filtering (KF), Clustering techniques	Clustering techniques based STLF model	STLF	1) Data clustering 2) Data discrimination	Using six different models based on BPNN, WNN, and KF for clustering data in load forecasting
[55]	2020	CNN, LSTM, Unified Quantile Regression (UQR), Time-cognition, BP	UQR based ANN for residential load forecasting	STLF	1) Probabilistic load forecasting	Managing the probabilistic residential load forecasting by using UQR based deep learning
[56]	2020	ELM, Support Vector Machine (SVM), XGboost, Decision tree, Genetic Algorithm	ELM, SVM and deep learning based STLF	STLF	1) Data cleaning 2) Data and feature redundancy 3) Hyper parameter tuning 4) Feature extraction	Removing the data cleaning and redundancy, tuning the hyper parameter and extracting the features at load forecasting.
[57]	2020	LSTM, ANN	Load forecasting for predicting system marginal price	STLF, MTLF and LTLF	1) System marginal price (SMP) 2) Energy price	Utilizing the load forecasting to enhance SMP by using LSTM and ANN
[58]	2020	Iterative ResBlocks (IRB), DNN, Spatio -temporal correlation	IRB-DNN based residential STLF model	STLF	1) Individual residential load	IRB-DNN based STLF for individual residential users by using the Spatio-temporal correlation.
[59]	2020	RNN , LSTM, Frequency Domain Decomposition (FDD), iForest, Mallat	STLF based on LSTM and FDD	STLF	1) Data cleaning 2) Low and high frequency components	Developing LSTM bsed STLF to clean the data and to obtain the low-high frequency signal.
[60]	2020	RNN, LSTM, GRU, BP, Sequence to Sequence (S2S), Attention Mechanism	S2S RNN for load forecasting	STLF	1) Attention mechanisms 2) Time dependencies	Proposes S2S RNN for load forecasting and predicts different attention mechanisms
[61]	2020	LSTM, Cascade NN, ENN, BP, Levenberg-Marquardt algorithm, Ensemble Learning	Ensemble learning based load forecasting model by using Cascade neural network	STLF and MTLF	1) Learning error 2) Data clustering	Integrates LSTM, Cascade neural network, and clustering for testing load forecasting and decreases the mean absolute percentage error.
[62]	2020	CNN, Image encoding, BP, Recurrence Plot, Gramian Angular field, Markov Transition field.	Image encoding and CNN based load forecasting model	STLF and MTLF	1) Demand response 2) Individual residential load	Applying CNN on time series data based encoded image to predicting load forecasting for the single residential loads.

TABLE 2. ***(Continued.)*** **Most relevant deep learning based load forecasting works from 2015-2020.**

Ref. No.	Year	Applied Techniques	Developed Model	Forecasting Type	Focus Area	Major Contribution
[63]	2020	DML, Apache Spark, Apache Hadoop, Linear Regression, Generalized Linear Regression, Decision Tree, Random Forest, Gradient-boosted Trees, Distributed computing	Spark and Hadoop based DML for load forecasting	VSTLF, STLF, MTLF and LTLF	1) Big data 2) DML and DDL 3) Training and computational time 4) Parallel computing 5) Feature selection 6) Real-time load	Develops DML for reducing testing time and training time of load forecasting with high accuracy for the big data. Here, ANN is unavailable but researchers stated that ANN can be used in the future work.
[64]	2020	CNN, BP	Electric vehicle charging oriented and CNN based probabilistic load forecasting	STLF	1) Probabilistic load forecasting 2) Feature extraction 3) Electric vehicle	Establishing CNN to predict traffic flow feature based load forecasting for charging the electric vehicle.
[65]	2020	CNN, Mutual Information (MI), MI-ANN, ReliefF, Kernel Principal Component Analysis (KPCA), BP	Big data and enhanced CNN based electric price forecasting	VSTLF, STLF, MTLF and LTLF	1) Big data 2) Over-fitting issue 3) Electric price 4) Feature selection 5) Data and feature redundancy 6) Training and computational time	Reducing over-fitting issue, overall computational time and ignoring feature redundancy with high performance in big data based electric forecasting by using CNN, MI, KPCA, etc. techniques
[66]	2020	LSTM, Bayesian deep learning, Bayesian theory	Bayesian deep learning based residential load forecasting	VSTLF	1) PV generation and PV penetration 2) Probabilistic load forecasting 3) Day–ahead 4) Data clustering 5) Individual residential load	Establishing LSTM based Bayesian deep learning for predicting probabilistic-residential load forecasting at PV system
[67]	2019	ENN, Ensemble averaging method, BP	Ensemble ANN based day–ahead load forecasting	STLF	1) Day–ahead 2) Environment friendly smart grid	Developing advance multi-load forecasting based distribution management systems for green energy policies
[68]	2019	CNN, Wavelet Packet Transform (WPT), BP	Big data analysis based wind power load forecasting	STLF	1) Environment friendly smart grid 2) Day–ahead 3) Electric price 4) Big data 5) Cost saving 6) Feature extraction 7) Wind generation	Combines CNN and WPT for using decomposed input at big data based environment friendly load forecasting
[69]	2019	KNN-ANN, BPNN, Spark,	Multivariable weighted KNN-ANN based load forecasting	STLF and MTLF	1) Big data 2) Weather factor 3) large time series 4) Training and computational time 5) Learning error	Dealing with multiple time series and multivariate data simultaneously and provides multiple prediction outputs of load forecasting
[70]	2019	LSTM, BPNN, Adaptive Moment Estimation	Deep LSTM based electric price and load forecasting	STLF and MTLF	1) Big data 2) Electric price	Predicting load forecasting with electric price by using LSTM from the big data
[71]	2019	RNN, AE, SAE, ELM, BP, Nonlinear Auto-regressive Network with Exogenous inputs (NARX), Differential Evolution, NAR-NN	DE-RELM and ESAENARX for big data based load forecasting	STLF	1) Big data 2) Electric price 3) Feature extraction 4) Learning error	Two ANN model named DE-RELM and ESAENARX have been developed for extracting quality features, reducing learning error and smoothing training inputs.

TABLE 2. ***(Continued.)*** **Most relevant deep learning based load forecasting works from 2015-2020.**

Ref. No.	Year	Applied Techniques	Developed Model	Forecasting Type	Focus Area	Major Contribution
[72]	2019	Parallel deep learning, DC-DC converter	Parallel deep learning based load forecasting for energy storage	STLF, MTLF and LTLF	1) Recovery of waste heat resource 2) Energy storage 3) Parallel computing	Using parallel deep learning for controlling hybrid energy storage in distributed system
[73]	2019	BPNN, Bayesian Regularization, Levenberg-Marquardt algorithm	Bayesian Regularization and Levenberg-Marquardt based ANN for VSTLF and STLF	VSTLF and STLF	1) Time horizon 2) Single building load forecasting	Integrating the ANN with Bayesian Regularization and Levenberg-Marquardt for the single district buildings' load forecasting
[74]	2019	DBN, BP, Phase Space Reconstruction (PSR), Levenberg-Marquardt algorithm	PSR-DBN method for VSTLF	VSTLF	1) High distributed power penetration 2) Bus load forecasting	Predicting the PSR-DBN based VSTLF from the perspectives of bus-load forecasting and distributed power penetration.
[75]	2019	KNN-ANN, FFNN, Euclidean theory	Euclidean KNN-ANN model for hourly VSTLF	VSTLF	1) Meteorological data (load, humidity, etc.) 2) Thermal unit generation	Establishing VSTLF for the hydro thermal unit generation by using KNN-ANN and meteorological data.
[76]	2019	CNN, LSTM, SAE, BP, Terminal coding, Fusion method, Affinity Propagation	Terminal coding and fusion ANN based day-ahead load forecasting	STLF and MTLF	1) Classification of similar load forecasting. 2) Day-ahead load	Classifying same group load forecasting and aggregating the load with various scales.
[77]	2019	CNN, BP	CNN based day-ahead load forecasting for individual resident	STLF	1) Individual resident load (day-ahead) 3) Training and computational time 3) Environment friendly smart grid	Developing LoadCNN for reducing the ANN training time, CO_2 (Carbon dioxide) emission, and environment friendly smart grid.
[78]	2019	LSTM	LSTM based real-time load forecasting for the electricity market	STLF	1) Meteorological data (load, temperature, humidity, etc.) 2) Feature extraction 3) Real-time load	Extracting the features from meteorological data and other information to predict load forecasting for the real-time electricity spot market.
[78]	2019	WaveNet, WNN, ENN, Ensemble learning, Bootstrapping, Stacked Generalization	WaveNet, WNN, ENN based residential STLF	STLF	1) Nonlinear STLF design. 2) Feature selection	Modeling load time series in nonlinear STLF by using WaveNet, WNN, and ensemble learning techniques.
[80]	2018	ResNet, ENN, BP, Monte Carlo Dropout, Ensemble learning	ResNet based load forecasting	STLF	1) Feature extraction 2) Meteorological data (load, temperature,etc.)	Developing ResNet based STLF where no need additional feature extraction process
[81]	2019	LSTM, CNN, BP	Load forecasting for CCHP (Combined Cooling, Heating, and Power) system	STLF	1) Energy management in CCHP systems.	Coupling the electric load, heating, and gas for STLF by using LSTM and CNN at the CCHP system.
[82]	2019	CNN, BP, XG-Boost, Decision Tree, Recursive Feature Elimination, Random Forest, Support Vector Regression (SVR)	CNN, SVR and data mining based load forecasting model	STLF	1) Training and computational time 2) Electric price 3) Generation loss prevention	Predicting STLF within the electric prices by using CNN, SVR, and data mining algorithms and also reducing the ANN computation time.
[83]	2019	GRNN, Empirical Mode Decomposition, Intrinsic mode function (IMF), Fruit Fly Optimization,	GRNN and feature extraction based STLF	STLF	1) Input reduction 3) Feature extraction 4) Data and feature redundancy	Reducing input, decomposing load series and extracting features for improving accuracy at STLF process.

TABLE 2. ***(Continued.)*** **Most relevant deep learning based load forecasting works from 2015-2020.**

Ref. No.	Year	Applied Techniques	Developed Model	Forecasting Type	Focus Area	Major Contribution
[84]	2019	BRNN, LSTM, GRU, BP	LSTM and GRU based BRNN model for STLF	STLF	1) Training and computational time 2) Over-fitting issue	Improving the accuracy and system quickness with preventing the over-fitting issue.
[85]	2019	BLSTM, Attention Mechanism, Rolling Update	BLSTM based STLF model	STLF	1) Week period and holidays 2) Learning error	Establishing BLSTM at STLF for obtaining higher accuracy and smaller error.
[86]	2019	WaveNet, CNN, LSTM, BP	CNN-LSTM-WaveNet based hybrid STLF	STLF	1) Learning error 3) Unknown and Identical distribution 4) Hybrid energy systems	Obtaining better performance in the Root Mean Squared Error, Mean Absolute Error, and Mean Absolute Percentage Error of STLF process.
[87]	2019	ENN, Ensemble learning, LSTM, Quantile method, Quantile forecasting, Distributed computing, Parallel computing	Ensemble learning and Quantile method based deep learning model for probabilistic load forecasting	STLF	1) Probabilistic load forecasting 2) Feature extraction 3) Feature selection 4) Individual load forecasting 5) Day-ahead 6) Parallel computing 7) DML and DDL	Developing load forecasting which does not require feature selection and extraction and establishing Quantile method for individual load forecasting. Moreover, this study has been prepared for future use of DDL based load forecasting.
[88]	2019	AE, DAE, Unsupervised learning, BP	Stacked DAE based unsupervised STLF	STLF	1) Random initialization of ANN 2) Gradient vanishing 3) Learning error	Ignoring random initialization, reducing error, mitigating gradient vanishing at unsupervised STLF.
[89]	2019	DBN, BP, Copula, Tail dependence	DBN and Copula based STLF	STLF	1) Unnecessary hidden neurons 2) High dimensional data 3) Over-fitting issue 4) Meteorological data 5) Peak load 6) Energy price	Adopts DBN for less hidden neuron based STLF process which ignores over-fitting issue and improves the peak load forecasting accuracy with the electric prices and temperature factor.
[90]	2019	LSTM, GRU	LSTM and GRU based single and multiple time sequence load forecasting	STLF and MTLF	1) Single and multiple time scale sequences 2) Time variations	Training LSTM and GRU with single and multiple time scale sequences for predicting load forecasting.
[91]	2019	RNN, LSTM, ELM, BP, Fuzzy logic, Neuro-fuzzy, ANFIS	Hybrid AI-based load forecasting model for the Power Distribution Network	STLF	1) Data cleaning 2) Meteorological data (load, temperature, etc.)	Developing STLF for South African Distribution Network where used temperature data and data cleaning method.
[92]	2019	AE, ELM	AE and ELM based STLF process for End-user transformer	STLF	1) Feature extraction 2) Low pass filter 3) Training and computational time 4) Highly volatile period	Improving training time, extracting deep features, and decomposing time-series data from AE and ELM based STLF process.
[93]	2019	LSTM , One-Hot Encoding	LSTM and one-hot encoding based STLF for virtual power plant	STLF	1) Virtual power plant	Establishing STLF for cloud-based power plant or virtual power plant
[94]	2019	CNN, Fuzzy logic, BP, Dropout technique	CNN and fuzzy time series based STLF	STLF	1) Over-fitting issue 2) Feature extraction	Improving feature extraction with high accuracy and resolving the over-fitting issue by using CNN, dropout technique and fuzzy logic.

TABLE 2. ***(Continued.)*** **Most relevant deep learning based load forecasting works from 2015-2020.**

Ref. No.	Year	Applied Techniques	Developed Model	Forecasting Type	Focus Area	Major Contribution
[95]	2018	BPNN, Data mining, Sub-network (Subnet)	Subnet oriented distributed load forecasting using BP	STLF	1) Distributed load forecasting 2) Weather factor 3) Meteorological Data	Dividing load forecasting into the few subnets for obtaining better accuracy and performance.
[96]	2018	Neuro-fuzzy, ANFIS, Genetic algorithm, Particle Swarm Optimization	Neuro-fuzzy,ANFIS and Swarm Optimization based VSTLF	VSTLF	1) Training and computational time 2) BP alternative 3) Feature selection	Decreasing training or execution time, reducing the complexity of feature selection at VSTLF process without BP
[97]	2018	HTM, Sparse Distributed Representations, HTM Spatial Pooler	HTM Spatial Pooler based STLF process	STLF	1) Handling class overlapping	Improving load forecasts by using HTM with overlapping temporal classification.
[98]	2018	RNN, GRU, Encoder-Decoder Architecture, AE, SLP, Sequence to Sequence (S2S), Neural Machine Translation, BP	SLP, RNN, GRU, AE and Encoder-Decoder based hybrid deep learning model for STLF	STLF	1) Training and computational time 2) Learning error 3) Number of epochs	Reducing the training time and computation error at STLT by using SLP, RNN, GRU, Encoder-Decoder and AE.
[99]	2018	Neuro-fuzzy, ANFIS, BPNN, Levenberg-Marquardt algorithm	ANFIS and BPNN based LTLF model	LTLF	1) Weather factor 2) Meteorological data (load, temperature, etc.)	Combining Levenberg-Marquardt BP based ANN and ANFIS for predicting LTLF effectively.
[100]	2018	FFNN, Particle Swarm Optimization, MLP	ANN and Particle Swarm Optimization based MTLF process	MTLF	1) Environment emission 2) Peak load	Combining FFNN and Particle Swarm Optimization for MTLF from the environment emission and peak load points of views.
[101]	2018	Distributed Neural Network, DML, DDL	DML load forecasting process	STLF	1) Feature extraction 2) Training and computational time 3) Cost saving 4) Peak load 5) Parallel computing 6) DML and DDL	Applying Distributed Neural Network at load forecasting and speed-up the training time with high accuracy. Moreover, this study has been represented DDL based STLF.
[102]	2018	ART network, ARTMAP (Predictive ART), Fuzzy logic	Fuzzy-ARTMAP ANN based multi-nodal load forecasting	STLF	1) Training and computational time 2) BP alternative 3) Multi-modal load forecasting	Establishing Fuzzy-ARTMAP based ANN for load forecasting which is faster than BPNN and reducing additional computational time.
[103]	2018	AE, DAE, Support Vector Regression (SVR)	Stacked DAE based day-ahead load forecasting process	STLF	1) Day-ahead 2) Meteorological data 3) Weather factor 4) Feature extraction	Obtaining higher-level features at load forecasting process from the lower-level information with lower error.
[104]	2018	BPNN	BPNN based load forecasting model for PV (Photovoltaic) generation	MTLF and LTLF	1) PV generation and PV penetration 2) Peak load 3) Maximum power load identification	Identifying the installed capacity and maximum power load of PV generation at MTLF and LTLF
[105]	2018	WNN, ELM, Friedman test, Post-hoc test, Bootstrapping	Improved WNN and Generalized ELM for probabilistic load forecasting	STLF	1) Probabilistic load forecasting 2) Noisy data 3) Feature extraction	Mitigating the problem of probabilistic load forecasting which considers nosy data and non-iterative training process.
[106]	2018	ELM, RBF network, Unscented Kalman Filter (UKF) kernel	RBF and UKF kernel functions based ELM for STLF	STLF	1) Noisy data 2) Capturing characteristics	Capturing characteristics by using RBF and UKF and forecasting from the de-noised electric data by using ELM.

TABLE 2. ***(Continued.)*** **Most relevant deep learning based load forecasting works from 2015-2020.**

Ref. No.	Year	Applied Techniques	Developed Model	Forecasting Type	Focus Area	Major Contribution
[107]	2017	GMDH, Wavelet transform, Data mining	Wavelet transform and GMDH based load forecasting	STLF	1) Day-ahead	Utilizing GMDH and Wavelet transform for improving the accuracy of STLF
[108]	2017	BPNN, Regression Analysis	Regression Analysis and ANN based MTLF	MTLF	1) Learning error	It reduces the learning error with batter accuracy at MTLF
[109]	2017	CVNN, BP	CVNN based VSTLF with high Photovoltaic (PV) penetration	VSTLF	1) Energy storage 2) Parallel computing 3) PV generation and PV penetration	Establishing parallel load forecasting with PV penetration for energy storage.
[110]	2017	ELM, Switching Delayed Particle Swarm Optimization (SDPSO)	SDPSO-EML based STLF	STLF	1) Unnecessary hidden neurons 2) Over training	Reducing unnecessary hidden neurons and over training problem with high accuracy
[111]	2017	DBN, Copula, BP	DBN-Copula based day-ahead and week-ahead STLF	STLF	1) Learning error 2) Day-ahead 3) Meteorological data 4) Energy price	Improving DBN and Copula based STLF where have electric price and ahead meteorological data
[112]	2017	CNN, BP, K-means clustering	CNN and K-means clustering based STLF for big data	STLF	1) Big data 2) Data clustering	Improving big data based STLF by using CNN and K-means clustering techniques
[113]	2017	LSTM, Quantile method, BP	Quantile LSTM based residential STLF	STLF	1) Individual residential load 2) Probabilistic load forecasting	Improving STLF for individual resident users and probabilistic forecasting by using LSTM and Quantile method
[114]	2017	Grey-NN, Augmented Dickey–Fuller test, Wavelet decomposition	Grey-NN and Wavelet decomposition based STLF	STLF	1) Low and high frequency components 2) High dimensional data	Developing Grey-NN for maintaining high-frequency components at STLF
[115]	2017	ELM, DDL, BP, Parallel deep learning, MapReduce, Cloud Computing	Distributed ELM and cloud computing based load forecasting approach	STLF	1) Data requirement 2) Training and computational time 3) Cost saving 4) Data re-train 5) DML and DDL 6) Parallel computing	Reducing data requirements, computational time, cost with high accuracy. Researchers established DDL based on ELM for STLF where no need data re-train process.
[116]	2017	DNN, BP, Levenberge-Marquardt algorithm,	Dynamic ANN based STLF	STLF	1) Low and high frequency components 2) Cost saving	Enhancing STLF and reducing the high cost by using dynamic DNN.
[117]	2017	ENN, Ensemble learning, Boosting method, BP	Boosting ANN based STLF process	STLF	1) Learning error 2) Training and computational time	Reducing the forecasting error and computation time by using Boosting method based ENN.
[118]	2017	LSTM, Empirical Mode Decomposition (EMD), Xgboost, K-means algorithm, BP	LSTM-EMD- Xgboost based STLF	STLF	1) Intrinsic mode function (IMF) 2) Similar days 3) Meteorological data	Combines LSTM, EMD, Xgboost and K-means algorithm for enhancing similar days STLF performance
[119]	2017	Elman ANN, Empirical Mode Decomposition (EMD), Sliding Window EMD (SWEMD), BP	SWEMD- Elman ANN based building level load forecasting	STLF	1) Data and feature redundancy 2) Feature Selection 3) Over-fitting issue	Combines SWEMD with Elman ANN for reducing over-fitting problem and data redundancy at load forecasting.
[120]	2017	Neuro-fuzzy, Fuzzy Logic, Levenberg-Marquardt, BP	Fuzzy logic and ANN based LTLF process	LTLF	1) Learning error 2) Training and computational time 3) Econometric data	Training econometric data (population) and reducing error at LTLF by using ANN and fuzzy logic.

TABLE 2. ***(Continued.)*** **Most relevant deep learning based load forecasting works from 2015-2020.**

Ref. No.	Year	Applied Techniques	Developed Model	Forecasting Type	Focus Area	Major Contribution
[121]	2016	ELM, DML, Smart Gateway Platform, Online Machine Learning , Shared Machine Learning,	EML and smart gateway based DML model for STLF	STLF	1) Training and computational time 2) Peak load 3) Data requirement or limited data 4) Cost saving 5) DML and DDL	Developing EML based DML in STLF process for reducing computational time with limited data. Here, several EML works at several gateways in smart grid.
[122]	2016	BPNN, Hadoop, MapReduce	BPNN and Machine learning algorithm s based load forecasting for Big data	MTLF	1) Big data 2) Noisy data	Predicting load forecasting from big data using BPNN and other machine learning algorithms.
[123]	2016	FFNN, Copula, Correlation coefficient	Correlation coefficients based multi-area STLF	STLF	1) Learning error 2) Multi-area STLF	Developing the correlation coefficients among the forecasting errors from different areas.
[124]	2016	BPNN, Support Vector Regression (SVR)	Sensor-based load forecasting for event venues	VSTLF and STLF	1) Big Data 2) Sensor-based load forecasting	Establishes sensor-based load forecasting for the event venues.
[125]	2016	LSTM, BP, Sequence to Sequence (S2S) architecture	S2S based LSTM for building load forecasting	STLF	1) Building load forecasting 2) Individual residential load	Comparing between LSTM and S2S based LSTM at STLF and S2S based LSTM presents better performances.
[126]	2016	LSM, PNN	LSM-PNN based load forecasting	STLF	1) Data requirement or limited data 2) Power scheduling	Providing performance better than RNN at STLF by using LSM based PNN and
[127]	2016	FFNN, NAR-NN, Auto-regressive Integrated Moving Average (ARIMA	Multi-model ANN for MTLF	MTLF	1) Data requirement or limited data	Combines three NAR-NN models with an FFNN to use small training data at MTLF.
[128]	2016	Grey-NN, BPNN, Regression analysis,	Grey-NN and Regression analysis based load forecasting	MTLF and LTLF	1) Data requirement or limited data	Improving MTLF and LTLF by using Grey-NN with low data requirements
[129]	2016	Neuro-fuzzy, Fuzzy Logic	Neuro-fuzzy based VSTLF	VSTLF	1) Classification of similar load forecasting.	Improving VSTLF for five-minute intervals by using Neuro-fuzzy.
[130]	2016	DBN, RBM, BP	DBN, RBM and BP based multi ANN for STLF	STLF	1) Training and computational time 2) Learning error 3) Day-ahead	Developing 24 hour ahead STLF using DBN-RBM to reduce the computational time and learning error.
[131]	2016	Neuro-fuzzy, Graphics Processing Unit (GPU), Fuzzy Logic, Evolutionary computation	GPU based load forecasting	VSTLF and STLF	1) Training and computational time 2) Parallel computing	Reducing the forecasting execution time more than CPU by using GPU based deep learning process.
[132]	2016	DNN, FFNN	DNN based STLF model	STLF	1) Meteorological data 2) Econometric data 3) Feature Selection 4) Learning error 5) Weather factor	Improving STLF with learning error reduction based on meteorological data, economic data and weather factors.
[133]	2016	Neuro-evolution, RNN, Cartesian genetic programming	Recurrent neuro-evolution based highly fluctuating electrical load forecasting	VSTLF and STLF	1) High dimensional data 2) Highly fluctuating electrical loads	Developing RNN and neuro-evolution based load forecasting for highly fluctuating electrical loads

TABLE 2. ***(Continued.)*** **Most relevant deep learning based load forecasting works from 2015-2020.**

Ref. No.	Year	Applied Techniques	Developed Model	Forecasting Type	Focus Area	Major Contribution
[134]	2015	ART network, ARTMAP (Predictive ART), Online machine learning, Hyper-sphere method, Distributed training	Distributed training based Hyper-sphere ARTMAP load forecasting	VSTLF and STLF	1) Training and computational time 2) Data proliferation problem 3) Noisy data	Predicts load forecasting quickly with high accuracy and decreases the data proliferation problem at ARTMAP based load forecasting model Moreover, this process is based on distributed training.
[135]	2015	BPNN, Soft computing, Fuzzy logic, Random Forest, Auto-regressive Integrated Moving Average (ARIMA)	Soft computing with machine learning based hourly load forecasting	STLF	1) Feature selection 2) Cost saving 3) Training and computational time	Using soft computing and entropy based feature selection at STLF for reducing the computational cost with quickly prediction.
[136]	2015	BPNN, Fuzzy logic	Self-optimized STLF	STLF	1) Low or medium voltage 2) High dimensional data	Developing self-optimized STLF for medium voltage and high uncertainty
[137]	2015	ESN, Genetic algorithm, Principal Component Analysis (PCA), Auto-regressive Integrated Moving Average (ARIMA)	ESN- PCA decomposition based load forecasting	VSTLF and STLF	1) High dimensional data 2) Training and computational time	Predicts load forecasting using ESN, PCA decomposition, ARIMA and genetic algorithm when forecasting period is higher or very long
[138]	2015	Physical-NN, Clear Sky solar Radiation Model (CSRM), Soft computing, BP	Hybrid physical-NN based load forecasting	STLF and MTLF	1) Day-ahead 2) PV generation and PV penetration 3) Meteorological data (load, humidity, temperature, etc.) 4) Weather factor	Developing CSRM and soft computing based Physical-NN in load forecasting for ignoring renewable energy source (RES) and PV related problems.
[139]	2015	BPNN, MI-ANN, Correlation-based Feature Selection, Mutual Information (MI), RReliefF, Levenberg-Marquardt algorithm, Model Tree Rules	Correlation and instance feature selection techniques based load forecasting process	VSTLF and STLF	1) Feature selection 2) High dimensional data	Establishing correlation and instance based feature selection at load forecasting process for identifying highly related features
[140]	2015	WNN, RNN, BP, Levenberg-Marquardt algorithm	Self-Recurrent Wavelet Neural Network (SRWNN), based STLF	STLF	1) Highly volatile period	Applying SRWNN (Combination of WNN and RNN) at load forecasting for non-smooth and highly volatile load time series.
[141]	2015	RBF network, Error Correction algorithm	RBF based STLF	STLF	1) Learning error	Applying RBF with Error Correction algorithm at STLF for increasing the accuracy.
[142]	2015	BPNN, Adaptive Differential Evolution Algorithm	BPNN and Adaptive Differential Evolution based time series forecasting model	STLF	1) Learning error 2) Training and computational time 3) Cost saving	Combining BPNN and Adaptive Differential Evolution Algorithm for improving time searies based load forcasting
[143]	2015	MI-ANN, Mutual Information (MI)	MI-ANN and modified feature selection based day-ahead load forecasting	STLF	1) Feature selection 2) Data and feature redundancy 3) Day-ahead load 4) Training and computational time	Reducing data redundancy and execution time with high accuracy in the data ahead load forecasting process

TABLE 3. Research works and load forecasting types.

Forecasting Type	Total Number of Works	Research Works	Major Reasons of Utilization	Technical Problems Facing
VSTLF	17	[44], [48], [63], [65], [66], [73], [74], [75], [96], [109], [124], [129], [131], [133], [134], [137], [139]	1) Load demand of residential buildings 2) PV generation and PV penetration 3) Single building load forecasting 4) Bus load forecasting 5) Thermal unit generation 6) Classification of similar load forecasting.	1) Over-fitting issue, 2) Big data 3) High training and computational time, 4) Feature extraction and selection related issue, 5) Data requirement or limited data, 6) High dimensional data related ,problem, 7) Noisy data, 8) Data proliferation problem
STLF	88	[40], [42], [45] – [65], [67] – [73], [76] – [95], [97], [98], [101] – [103], [105] – [107], [110] – [119], [121], [123] – [126], [130] – [143]	1) Day-ahead load forecasting 2) Hourly load forecasting 3) Cost saving 4) Similar day load forecasting 5) PV generation and PV penetration 6) Wind generation 7) Load demand of residential buildings 8) Load demand of different hours and on different days 9) Individual residential electric load 10) Single building load forecasting 11) Peak load 12) Load demand of different hours and different days 13) Load shedding 14) Probabilistic load forecasting	1) High training and computational time, 2) Variance of load forecasting 3) Data requirement or limited data, 4) Data and feature redundancy, 5) High computational cost, 6) Data discrimination, 7) Low and high frequency components related issues, 8) Time dependencies, 9) Learning error and system accuracy, 10) Big data, 11) Over-fitting issue 12) Large time series, 13) Non environment friendly smart grid issue, 14) Unknown and Identical distribution, 15) Unnecessary hidden neurons, 16) High dimensional data related issue, 17) Single and multiple time scale sequences, 18) Highly volatile period, 19) Feature extraction and selection related issue, 20) Noisy data, 21) Data clustering related issue, 22) Data proliferation problem
MTLF	24	[41], [42], [45] – [48], [52], [57], [61] – [63], [65], [69], [70], [72], [76], [90], [100], [104], [108], [122], [127], [128], [138]	1) Load demand of different hours and different days 2) Load demand of residential buildings 3) Individual residential electric load 4) Peak load 5) Maximum power load identification 6) PV generation and PV penetration 7) Probabilistic load forecasting	1) Over-fitting issue, 2) Learning error and system accuracy 3) Big data, 4) High training and computational time, 5) Large time series, 6) Single and multiple time scale sequences, 7) Feature extraction and selection related issue

imagery analysis, CNN was the next most widely used. ELM, FFNN, DNN, RNN, Neuro-fuzzy, ENN, and AE techniques have also seen significant use. In a few cases, MI-ANN, WNN, GRU, DBN, RBM, ANFIS, and ART network approaches were applied to obtain better performances. Cascade NN, KNN-ANN, SAE, DAE, WaveNet, RBF network, and Grey-NN techniques were used much in load forecasting. Other ANN techniques, such as CVNN, NAR-NN, DQN, BRNN, BLSTM, etc, have been used rarely.

The various VSTLF, STLF, MTLF, and LTLF based works are summarized in Table 3. Table 3 also presents the major reasons for utilizing different forecasting types and technical problems. The breakdown of ANN technique usage is shown in Table 4.

C. BP AND NON-BP BASED LOAD FORECASTING

From Table 4, we see that BPNN, which is FFNN with BP, is the second-most used ANN for load forecasting among the papers surveyed. BPNN is the updated class of FFNN which contains an additional BP algorithm. Because of its accurate prediction, BP has also been used in RFNN [39], WNN [40], RNN [48], ENN [50], ART network [51], LSTM [61], and other techniques. Additionally, CNN is trained with BP and BP is a necessary part of CNN [146]. Sometimes, BP was modified by using the Levenberg-Marquardt algorithm [61], [73], [99]. The majority of the works surveyed used BP with different ANN techniques. Others tried to avoid BP to reduce BP's disadvantages [25], [26], [27] and ameliorate issues related to loading forecasting. The total number of BP and non-BP based works are displayed in Table 5.

At the non-BP-based works, researchers have tried to develop a hybrid ANN model for ignoring BP decency and reducing computational time. Most of the non-Bp based works have been focused on computational time and learning error. On the other hand, we found that LSTM is the most usable ANN at load forecasting. Most of the studies on LSTM

TABLE 4. **ANN and load forecasting works.**

ANN Type	Total Number of Works	Research Works	Major Focused Areas
LSTM	22	[40], [49], [53], [55], [57], [59], [60], [61], [66], [70], [76], [78], [81], [84], [86], [87], [90], [91], [93], [113], [118], [125]	1) Training and computational time 2) Over-fitting issue 3) Similar day selection 4) Feature selection and extraction 5) Wind generation 6) Learning error 7) Big data 8) Probabilistic load forecasting 9) Energy price 10) Data cleaning 11) Low and high frequency components 12) Data clustering 13) PV generation and PV penetration 14) Day–ahead load forecasting 15) Individual residential load 16) Meteorological data (load, temperature, humidity, etc.) 17) DML and DDL 18) Single and multiple time scale sequences 19) Week period and holidays 20) Hybrid energy systems
BPNN	19	[40], [41], [43], [45], [54], [69], [70], [73], [95], [99], [104], [108], [122], [124], [128], [135], [136], [139], [142]	1) Training and computational time 2) Learning error 3) Similar day selection 4) Unavoidable Stochastic Part 5) Low or medium voltage 6) Cost saving 7) Data clustering 8) Data discrimination 9) Big Data 10) Weather factor 11) Electric price 12) Single building load forecasting 13) Distributed load forecasting 14) Meteorological data (load, temperature, humidity, etc.) 15) PV generation and PV penetration 16) Peak load 17) Noisy data 18) Data requirement or limited data 19) Feature selection and extraction 20) High dimensional data
CNN	13	[43], [55], [62], [64], [65], [68], [76], [77], [81], [82], [86], [94], [112]	1) Training and computational time 2) Cost saving 3) Learning error 4) Big data 5) Probabilistic load forecasting 6) Individual residential load 7) Feature selection and extraction 8) Electric price 9) Over-fitting issue 10) Data and feature redundancy 11) Environment friendly smart grid 12) Day–ahead load forecasting 13) Electric price 14) Wind generation 15) Individual resident load 16) Data clustering 17) Probability distribution 18) Generation loss prevention 19) Hybrid energy systems

TABLE 4. ***(Continued.)*** **ANN and load forecasting works.**

ANN Type	Total Number of Works	Research Works	Major Focused Areas
Neuro-fuzzy	8	[44], [81], [91], [96], [99], [120], [129], [131]	1) Training and computational time 2) Learning error 3) Weather factors 4) Thermal unit generation 5) Cost saving 6) Data requirement or limited data 7) Data cleaning 8) Meteorological data (load, temperature, etc.) 9) Feature selection and extraction 10) Weather factor 11) Econometric data 12) Similar load forecasting 13) Parallel computing 14) BP alternative
ENN	7	[50], [61], [67], [79], [80], [87], [117]	1) Training and computational time 2) Reduction of bias 3) Variance of load forecasting 4) Learning error 5) Data clustering 6) Day-ahead load forecasting 7) Environment friendly smart grid 8) Feature selection and extraction 9) Meteorological data (load, temperature, humidity, etc.) 10) Input reduction 11) Data and feature redundancy 12) Probabilistic load forecasting 13) Individual load forecasting 14) Parallel computing 15) DML and DDL
AE	6	[71], [76], [88], [92], [98], [103]	1) Training and computational time 2) Learning error 3) Big data 4) Electric price 5) Feature extraction 6) Random initialization of ANN 7) Gradient vanishing 8) Day-ahead load forecasting 9) Meteorological data (load, temperature, humidity, etc.) 10) Weather factor
MI-ANN	5	[49], [52], [65], [139], [143]	1) Training and computational time 2) Feature selection 3) Wind generation 4) Data and feature redundancy 5) Consumption behavior 6) Big data 7) Over-fitting issue 8) Electric price 9) High dimensional data 10) Day-ahead load forecasting
WNN	5	[41], [54], [79], [105], [140]	1) Wavelet disintegration 2) Data clustering 3) Data discrimination 4) Feature selection and extraction 5) Probabilistic load forecasting 6) Noisy data 7) Highly volatile period

TABLE 4. *(Continued.)* **ANN and load forecasting works.**

ANN Type	Total Number of Works	Research Works	Major Focused Areas
GRU	5	[48], [60], [84], [90], [98]	1) Training and computational time 2) Over-fitting issue 3) Learning error 4) Load demand of residential buildings 5) Time dependencies 6) Single and multiple time scale sequences 7) Time variations
DBN	4	[74], [89], [111], [130]	1) Training and computational time 2) Learning error 3) High distributed power penetration 4) Bus load forecasting 5) Unnecessary hidden neurons 6) High dimensional data 7) Over-fitting issue 8) Meteorological data (load, temperature, humidity, etc.) 9) Peak load 10) Energy price 11) Day-ahead load forecasting
RBM	3	[49], [52], [130]	1) Training and computational time 2) Learning error 3) Day-ahead load forecasting 4) Feature selection 5) Wind generation 6) Data and feature redundancy 7) Consumption behaviors
ANFIS	3	[91], [96], [99]	1) Training and computational time 2) BP alternative 3) Feature selection 4) Data cleaning 5) Meteorological data (load, temperature, humidity, etc.) 6) Weather factor
ART network	3	[51], [102], [134]	1) Training and computational time 2) Data proliferation problem 3) Noisy data 4) Cost saving 5) Data requirement or limited data 6) BP alternative
Cascade NN	2	[46], [61],	1) Learning error 3) Data clustering 4) PV generation and PV penetration 5) Wind generation 6) Energy price 7) Softmax layer
KNN-ANN	2	[69], [75]	1) Training and computational time 2) Learning error 3) Big data 4) Weather factor 5) large time series 6) Meteorological data (load, temperature, humidity, etc.) 7) Thermal unit generation

TABLE 4. *(Continued.)* **ANN and load forecasting works.**

ANN Type	Total Number of Works	Research Works	Major Focused Areas
KNN-ANN	2	[69], [75]	1) Training and computational time 2) Learning error 3) Big data 4) Weather factor 5) large time series 6) Meteorological data (load, temperature, humidity, etc.) 7) Thermal unit generation
DAE	2	[88], [103]	1) Random initialization of ANN 2) Gradient vanishing 3) Learning error 4) Day-ahead load forecasting 5) Meteorological data (load, temperature, humidity, etc.) 6) Weather factor 7) Feature extraction
WaveNet	2	[79], [86]	1) Learning error 2) Feature selection 3) Unknown and Identical distribution 4) Hybrid energy systems
RBF network	2	[106], [141]	1) Learning error 2) Noisy data 3) Capturing characteristics
Grey-NN	2	[114], [128]	1) Low and high frequency components 2) High dimensional data 3) Data requirement or limited data
NAR-NN	2	[71] [127]	1) Data requirement or limited data 2) Big Data 3) Feature extraction 4) Learning error
SLP	1	[98]	1) Training and computational time 2) Learning error 3) Number of epochs
Distributed Neural Network	1	[101]	1) Feature extraction 2) Training and computational time 3) Cost saving 4) Peak load 5) Parallel computing 6) DML and DDL
GMDH	1	[107]	1) Day-ahead load forecasting
PNN	1	[126]	1) Data requirement or limited data 2) Power scheduling
CVNN	1	[109]	1) Energy storage 2) Parallel computing 3) PV generation and PV penetration
DQN	1	[40]	1) Similar day selection

TABLE 4. ***(Continued.)*** **ANN and load forecasting works.**

ANN Type	Total Number of Works	Research Works	Major Focused Areas
BRNN	1	[84]	1) Training and computational time 2) Over-fitting issue
BLSTM	1	[85]	1) Week period and holidays 2) Learning error
ESN	1	[137]	1) Training and computational time 2) High dimensional data
Elman ANN	1	[119]	1) Data and feature redundancy 2) Feature Selection 3) Over-fitting issue
Encoder-Decoder Architecture	1	[98]	1) Training and computational time 2) Learning error 3) Number of epochs
RFNN	1	[39]	1) Week period and holidays 2) Violation of Gaussian Distribution 3) Meteorological data (load, temperature, humidity, etc.) 4) Weather factor
LSM	1	[126]	1) Data requirement or limited data 2) Power scheduling
Physical-NN	1	[138]	1) Day-ahead load forecasting 2) PV generation and PV penetration 3) Meteorological data (load, humidity, temperature, etc.) 4) Weather factor
HTM	1	[97]	1) Handling class overlapping
Neuro-evolution	1	[133]	1) High dimensional data 2) Highly fluctuating electrical loads
ResNet	1	[80]	1) Feature extraction 2) Meteorological data (load, temperature, humidity, etc.)
GRNN	1	[83]	1) Input reduction 2) Feature extraction 3) Data and feature redundancy

are based on the load forecasting process, researchers have avoided the BP process.

D. FOCUS AREAS

For load forecasting, different papers focused on different issues. Of these issues, much of the research has been devoted to improving the speed and accuracy of load forecasting, and so, as can be seen from Table 6, training and computational time, learning error and feature extraction have received a large share of the attention from researchers in smart grids.

Others have focused on improving accuracy by adding data that are expected to be correlated with power consumption. Datasets used for load forecasting included smart meter electric data, power loss data, temperature, weather, etc., depending on different power grid requirements. For example, the weather has an impact on power usage by consumers an important part of the smart grid used

TABLE 5. Different types of ANN in load forecasting.

BP Status	Total Number of Works	Research Works
BP based works	62	[39] – [43], [45], [47], [49], [50] – [52], [54], [55], [60] – [62], [64], [65], [67] – [71], [73], [74], [76], [77], [80] – [82], [84], [86], [88], [89], [91], [94], [95], [98], [99], [104], [108], [109], [111] – [113], [115] – [120], [122], [124], [125], [128], [130], [135], [136], [138] – [140], [142]
Non-BP based works	43	[44], [46], [49], [53], [56] – [59], [63], [66], [72], [75], [78], [79], [83], [85], [87], [90], [92], [93], [96], [97], [100] – [103], [105] – [107], [110], [114], [121], [123], [126], [127], [129], [131] – [134], [137], [141], [143]

meteorological data (temperature and humidity) Geographical, cultural, and income variances in different places. Where it mean that solutions for one area may not work for another area of the world, so load forecasting methods need to be reviewed for every locality.

Day-ahead load forecasting is used to predict day-to-day power consumption, which is conceptually so many works are focused on Day-ahead load forecasting or STLF. Table 6 shows other focus areas, including feature selection, big data, and feature selection. Parallel computing has been attempted for load forecasting, and it was found that it improved the computational time [63], waste heat resource recovery [72], PV generation related issues [109].

E. DML AND DDL IN LOAD FORECASTING

There are a small number of works found on DML and DDL based load forecasting. DML has been used for reducing machine learning training and computational time and improving feature selection in a big data-based load forecasting process [63]. Researchers are not used ANN in this work because their used libraries or tools do not directly support ANN but researchers stated that they will apply ANN in their future work. ENN and LSTM based hybrid ANN has been applied as a DDL for probabilistic load forecasting and improving feature extraction-selection [87]. In this work, the BP process ignores researchers. A distributed neural network has been used to reduce the computational time [98]. Here, researchers also ignored an additional BP process for reducing computational time.

ELM-based DDL process has been used in a model that ignores additional data re-train and reduced the training time [115]. In this work, researchers used the BP process. After reviewing the DML and DDL based load-forecasting work, we stated that maximum DML and DDL based works

TABLE 6. Top focus areas in load forecasting.

Focus Area	Total Number of Works	Research Works
Training and computational time	22	[63], [65], [69], [77], [82], [84], [92], [96], [98], [101], [102], [115], [117], [120], [121], [130], [131], [134], [135], [137], [142], [143],
Learning error	17	[43], [61], [69], [71], [85], [86], [88], [98], [108], [111], [117], [120], [123], [130], [132], [141], [142]
Feature extraction	13	[56], [64], [68], [71], [78], [80], [83], [87], [92], [94], [101], [103], [105]
Meteorological data (load, temperature, humidity, etc.)	13	[39], [75], [78], [80], [89], [91], [95], [99], [103], [111], [118], [132], [138]
Day-ahead	12	[66], [67], [68], [76], [77], [87], [103], [107], [111], [130], [138], [143]
Feature selection	10	[52], [63], [65], [79], [87], [96], [119], [132], [135], [139], [143]
Big data	10	[53], [63], [65], [68], [69], [70], [71], [112], [122], [124]
Weather factor	9	[39], [42], [44], [69], [95], [99], [103], [132], [138]
Cost-saving	9	[45], [51], [68], [101], [115], [116], [121], [135], [142]
Energy price	9	[46], [57], [67], [68], [70], [71], [82], [89], [111]
Probabilistic load forecasting	7	[47], [55], [64], [66], [87], [105], [113]
Parallel computing	7	[63], [72], [87], [101], [109], [115], [131]
Individual residential load	7	[58], [62], [66], [77], [87], [113], [125]
Data requirement or limited data	6	[51], [115], [121], [126], [127], [128]
Data and feature redundancy	6	[52], [56], [65], [83], [119], [143]
High dimensional data	6	[89], [114], [133], [136], [137], [139],
PV generation and PV penetration	5	[46], [66], [104], [109], [138]
Over-fitting issue	5	[65], [84], [89], [94], [119]
Peak Load	5	[89], [100], [101], [104], [121],
DML and DDL	5	[63], [87], [102], [115], [121]

are not based on the BP mechanism. Another ELM-based DML has been used at STLF where several EML works at several smart gateways [121]. This research reduces the

TABLE 7. DML and DDL based load forecasting techniques.

Ref. No.	Year	Applied ANN Technique	Major Solution / Problem Area	BP Availability
[63]	2020	Here, ANN is not available but at the DML process but DDL will apply in their future work.	1) Training and computational time 2) Big data 3) Feature selection	No
[87]	2019	ENN and LSTM	1) Probabilistic load forecasting 2) Feature extraction 3) Feature selection	No
[101]	2018	Distributed Neural Network	1) Training and computational time 2) Cost saving 3) Feature extraction	No
[115]	2017	ELM	1) Training and computational time 2) Data requirement or limited data 3) Cost saving 4) Data re-train	Yes
[121]	2016	ELM	1) Training and computational time 2) Data requirement or limited data 3) Cost saving 4) Peak Load	No

computational time and costs with limited data. Table 7 have shown the DML and DDL based research works.

III. CURRENT RESEARCH SCOPES

From the 105 works surveyed, it was found that although many have focused on deep learning for load forecasting in smart grids, few have tried to apply distributed computing using DML and DDL. Therefore, the reduction of data aggregation dependency is a barely-explored scope for possible research scopes in load forecasting to reduce computational time and learning error. Such work would be extremely beneficial as there are not many works on mitigation of data centralization and computation load on central servers, problems which decentralization and distribution can help resolve. Some works have tried to demonstrate a reduction in training and computation time using limited database load forecasting [51], [115], [121], but the use of DML and DDL is still open for study. The current challenges and possible research scopes have shown in Table 8.

Another possible avenue of future work is reducing learning error. Many have used BP for this purpose because it provides high accuracy. However, BP takes additional training time, so some work was devoted to improving BP, and some others avoided it altogether. Therefore, the new HSIC bottleneck method may be attempted as it is expected to provide high accuracy without BP and it supports distributed ANN training. Hence, a combination of HSIC bottleneck with DDL-based load forecasting may be found to reduce computational time, computational cost, and learning error. Specific works show scope for development, including applying ANN for evaluating DML-based load forecasting with big data, where MLib, and Apache Spark were used, but ANN was not applied [63]. In another case, ENN and LSTM based DDL processes were used, but the focus may be given in the future on big data, limited data, cost-saving, and computational time [87]. The distributed neural network has been used at load forecasting for reducing computational time and cost [101], but this method may be supported by techniques to mitigate dependency on data centralization.

ART network-based distributed training process was applied for load forecasting to reduce computational time [134]. HSIC bottleneck can be added to this technique for better accuracy. Each work focuses on one or two long-term, short-term, very short-term, and medium-term demands. How the proposed technique may apply to the other demands is open for investigation. Another problem is the impact of meteorological data on power usage, which varies from one land to another due to income, geography, and cultural differences. Therefore, each location needs specific study for that place to ensure that the deep learning methods used are suitable.

Potential research scopes have been identified which are displayed at Table 8, but there may be some more such scopes, which are unrecognized from our point of views where fruitful research may be done.

IV. CONCEPTUAL FRAMEWORK FOR DDL BASED LOAD FORECASTING

A. OBJECTIVES

Based on the research scopes a conceptual model may be established to reduce the data centralization dependency and computational time. The focus of this model is on HSIC Bottleneck and DDL for load forecasting in the smart grid. Its objectives are as follows:

1. Use DDL for load forecasting from multiple devices in the smart grid.
2. Reduce data centralization or aggregation dependency of cloud servers for load forecasting.
3. To reduce the overall load forecasting computational time by reducing data aggregation.
4. To perform load forecasting quickly and accurately.
5. Performing load forecasting for different areas with low learning error.

TABLE 8. Current challenges and future research scopes.

No.	Targeted Problem Areas	Current Challenges	Future Research Scopes
1.	Training and computational time	Reducing the overall load forecasting computational time by reducing data aggregation or centralization at the cloud server.	Establishing DDL based load forecasting for reducing data centralization or aggregation dependency at cloud server.
2.	Learning Error	Performing load forecasting accurately with quickly.	DDL provides a distributed training process and to do research becomes essential to enhance the accuracy with low computational time.
3.	Additional BP process	BP provides good accuracy but takes additional training time. Moreover, the overall load forecasting takes much time. Load forecasting without BP with high accuracy is a big challenge.	Some ANN mechanisms have been revealed which works without BP and provides high accuracy. It becomes essential to do the experiment on load forecasting without BP based ANN techniques.
4.	Big Data	Big data takes a lot of time. ANN is not available at DML based big data load forecasting model [60] because the applied tools and libraries are not supported ANN directly.	If somehow the Distributed Neural Network applies to the big data-based load forecasting, it would provide better improvement.
5.	Data requirement or limited data	Huge data are aggregated from the distributed area for load forecasting. As a result, the computational time increases day by day at any system. So, load forecasting with limited data and good accuracy become challenges.	If load forecasting occurs from the distributed area and the forecasting outputs aggregates at the cloud server, it may possible to obtain batter accuracy by using limited data. DDL is a very valuable technology for implementing this process in the future smart grid.
6	Feature selection and extraction	The state of the dataset in load forecasting in one region differs greatly from the state of the dataset in load forecasting in another region. So, Variations are found in feature selection and extraction work.	Feature selection and extraction become an important research scope because in many cases the accuracy of load forecasting depends on feature extraction and selection.
7	Meteorological data (load, temperature, humidity)	Sometimes, the variable or parameter sections and Input section of ANN create issues and changes to the ANN program become necessary.	The same load forecasting process, which would work for both meteorological data and non-meteorological data, becomes necessary to do research.

6. Ignoring the limited data-oriented load forecasting process and reducing huge data aggregation (Big Data) of the cloud server.
7. To avoid additional BP process during load forecasting.
8. To establish deep learning-based load forecasting without BP and also with high accuracy.
9. Develop the DDL mechanism using distributed neural network-based HSIC Bottleneck.
10. To measure the performance of HSIC Bottleneck for load forecasting.
11. To manage the meteorological data from distributed training perspectives and to enhance the feature extraction and selection performances.

B. ADVANTAGES OF HSIC BOTTLENECK

HSIC Bottleneck is a bottleneck layer-based multi-scale single neural network [28]. The bottleneck layer contains fewer neurons than the forward and backward. This ANN concept is based on Information Bottleneck [30], Dependency Bottleneck [31], Hilbert Space, Hilbert-Schmidt, HSIC [29] theories. Some advantages of HSIC bottleneck are given below:

- No need additional BP process and it reduces some disadvantages of BP such as vanishing gradients, exploding gradients, updates locking, etc.
- BP takes additional processing time at ANN. So, HSIC Bottleneck can reduce the additional computational time.

- For the large data and large training error, the Exploding Gradient issue occurs at ANN. Moreover, the Vanishing Gradient issue happened when ANN working with BP and gradient-based learning methods. A group of researchers has tried to reduce the DAE based Vanishing Gradient issue at DAE and BP based STLF [88]. HSIC Bottleneck does not sustain the exploding gradients and vanishing gradients issues.
- It is based on a single layer-based distributed training process were used the Stochastic Gradient Descent (SGD) algorithm. So, this ANN method can be enabled to use as a Distributed Neural Network in IoT and distribute load forecasting system. For that reason, it can provide better accuracy from different distributed perspectives
- Softmax is typically the final output layer which is an additional layer of ANN. Wavelet transform and DNN based load forecasting process have been mitigated Softmax layers based issue. At HSIC Bottleneck no required additional Softmax layer. For that reason, it reduces more computational time.
- At BP-based ANN techniques, generally, it is not possible to update the information of the first layer until the last layer value update. This issue is called Update Locking Problem. This ANN can ignore this issue.

Recently this ANN model has been used in the computer vision research area for recognizing facial expression [158]. Here, researchers used this model with CNN and got better performances. Computer vision is a popular research field of AI where the system can detect an object such as face detection from the dark image [159], Era identification from old heritage image [160], [161], etc. Because of the better performance of HSIC Bottleneck in the computer vision field, we can also hope that this model may enable us to reduce some severe challenges of load forecasting.

C. DDL-BASED DISTRIBUTED LOAD FORECASTING ARCHITECTURE

HSIC Bottleneck with DDL and distributed neural network structures where multiple Raspberry Pi micro-computers may be connected to the smart grid. Raspberry Pi is a micro-computer that is broadly used in IoT, cloud computing, wireless communication, and big data. Recently, Raspberry Pi has been used in energy management [147], [148] and in big data-based IoT solutions [149], [150], [155]–[157]. This architecture has been developed from the CoT, virtual power plant, and parallel computing point of view. Figure 5 presents the conceptual model.

In this model, two smart meters are connected with two different Raspberry Pi devices through the RS485 converter. Raspberry Pi is a Single-Board Computer (SBC) that can be easily set up with a smart meter and requires little space. Every Raspberry Pi contains data storage. Raspberry Pi can enable HSIC bottleneck-based single layer ANN computation using Python, Keras, and TensorFlow libraries. The devices will acquire and store electric energy data and meteorological data (temperature, etc.).

After computing load forecasting from different distributed devices using a distributed neural network-based HSIC Bottleneck model, the system would provide multiple ANN training data to the central server. Node.js, a server-side scripting framework based on JavaScript, may be used for aggregating multiple ANN training data. Keras and TensorFlow libraries can be used for computing deep learning-based aggregation. It is expected that the central server will not require raw data from the distributed devices, reducing communication load, and thus lends itself to be applied to resource-constrained wireless sensor networks. MongoDB, a NoSQL database, may support the server-side computing platforms such as Node.js. After computing the aggregation function of HSIC Bottleneck the overall predicted values would be obtained. The model architecture, required components, and fully distributed load forecasting process have been illustrated in Figure 5.

D. ADVANTAGES OF THE DDL-BASED LOAD FORECASTING MODEL

Based on the major challenges (Table 8) and HSIC Bottleneck advantages, this manuscript has been presented a DDL based distributed load forecasting architecture. Some major advantages of the model are given below:

- Here, no required additional BP process at load forecasting. So, it may reduce the computational or training time and BP disadvantages during load forecasting.
- Huge data aggregation takes much computational time and cost at load forecasting. At the model, the load forecasting process occurs from distributed devices of multiple areas. For that reason cloud servers not required huge data aggregation.
- On the other hand, obtaining better accuracy from limited data is another challenge in load forecasting. In this DDL based load forecasting model, the cloud server only aggregates the outputs of distributed ANN training data. So, it may be reduced more computational time without huge data, big data, and limited data-related issues.
- Here, load forecasting works from multiple areas by using the Distributed Neural Network concept. If multiple ANN processes happen from different distributed areas, it may possible to obtain different load demands from different points of view. So, it indicates that this model can provide better accuracy with law learning error.
- Generally, Exploding Gradient and Vanishing Gradient issues happened at the BP process when ANN working with large data and gradient-based learning methods. These issues take much computational time and cost. HSIC Bottleneck reduces these issues by ignoring BP. So, it can enhance the quickness of the load forecasting speed.

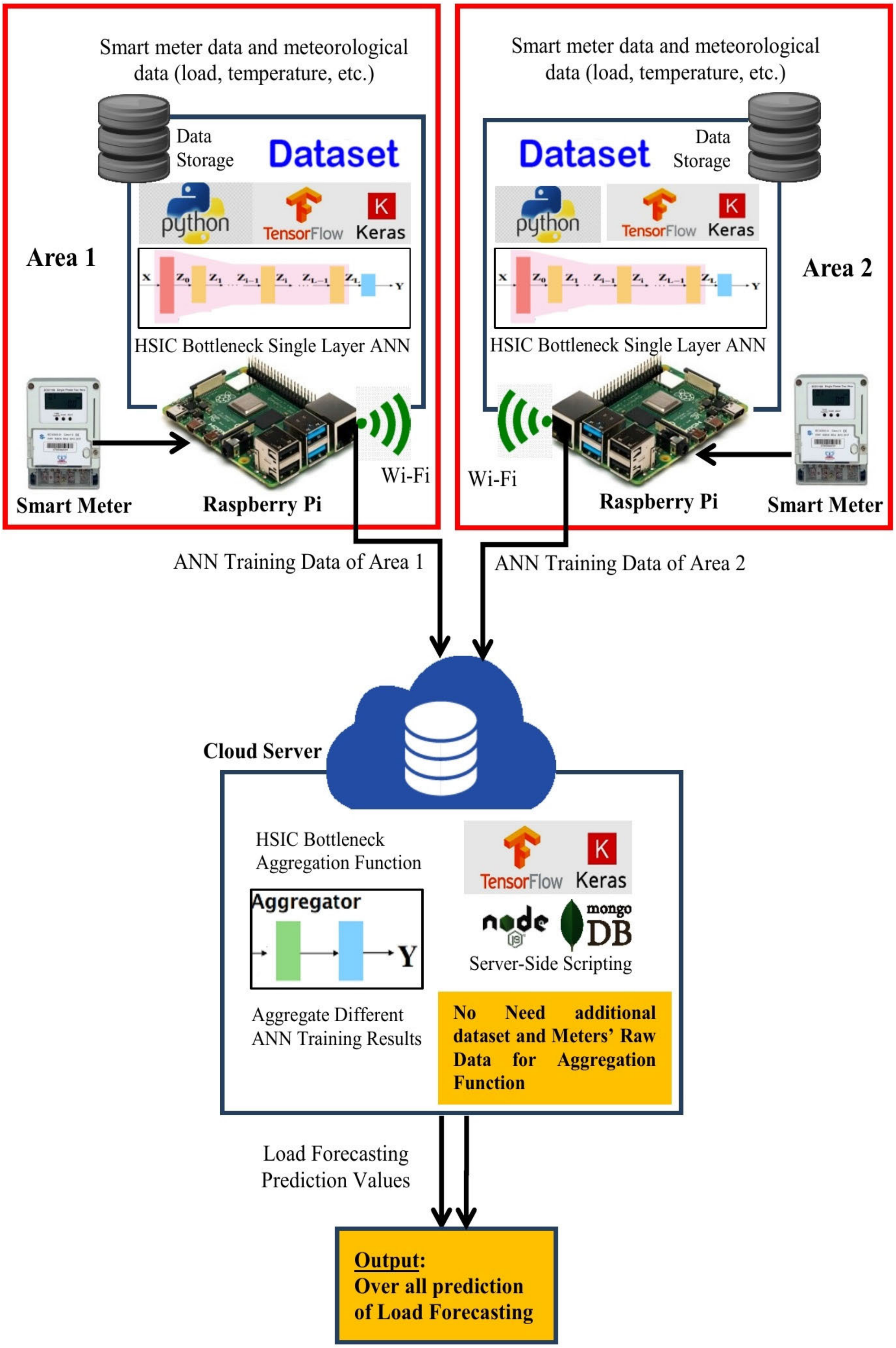

FIGURE 5. HSIC Bottleneck based DDL model for load forecasting. Here, the HSIC Bottleneck played as distributed neural network. In this process cloud server would not aggregate huge data from power consumer. Here, the cloud server aggregate multiple ANN training data or learning data.

- Because of HSIC Bottleneck, this load forecasting model not required an additional Softmax layer. So, reduces more computational time.

E. IMPLEMENTATION AND BARRIERS

Implementing the DDL-based load forecasting model can be possible for all types of load forecasting processes (VSTLF, STLF, MTLF, and LTLF). The primary two barriers to implementing this model are distributed to data storage and load forecasting devices. The distribution contains the data of regular smart meter data. It is essential to set up data storage in multiple distributed areas and maintain this storage from the cloud server. It may create many difficulties, but it can be maintained from the Virtual Power Plant and CoT point of view. If one device (Raspberry Pi) damage or shut down from an area, it may affect the load forecasting process. So, regular maintenance and supervision become necessary for implementing this model.

Moreover, we also assume that controlling the deep learning output data flow of the distribution area is another barrier during the system implementation. Relative Direction Learning algorithm [163]–[165] is a recent approach that is based on Human-Computer Interaction and Artificial General Intelligence (AGI). The relative direction learning algorithm controls the direction (Left, Right, Forward, and Backward) and it can help this model control the output data flow direction from distributed area perspectives. Information security may be needed at this process because data stored at distributed storage. Moreover, Blockchain is an innovative technology that is used in the decentralized and distributed network. This technology has been created as an important role in price intelligence [167], electronic medical records [168], and remote database access protocol-based database [169]. So, it is possible reducing the security issues by using Blockchain at the distributed load forecasting system.

V. OUTCOMES

The major outcomes after performing the comprehensive survey and proposed model are given as follows:

- STLF is the most important load forecasting type than VSTLF, MTLF, and LTLF. STLF has become extremely necessary in regular life due to individual buildings, residential areas, day-ahead load forecasting, hourly load forecasting, etc. for forecasting loads. For that reason, researchers have deeply focused on STLF related issues and also faced more problems at STLF than VSTLF, MTLF, and LTLF (Table 3).
- Among ANN strategies, LSTM is the most widely used ANN strategy. Although most researchers have used the hybrid ANN model, LSTM, along with other ANNs, has been overlooked in most major focus areas and problem-solving. In addition to LSTM, the BPNN, CNN, ELM, RNN, ENN, etc. ANN models also showed well contribution in a variety of key areas.
- Because of better accuracy, BP has been used at maximum researchers. Moreover, BP takes additional time and the load forecasting process also takes a huge time. Some researchers have tried to avoid additional BP process and developed hybrid ANN models for reducing computational time. Moreover, some researchers have not used BP at some LSTM based load forecasting techniques.
- Because of huge data aggregation, load forecasting takes much training and computational time. For that reason, the reduction of training and computational time is the major challenge in load forecasting. Moreover, high accuracy or learning error is another major concern of load forecasting. Alongside computational time reduction, researchers have also focused on high accuracy by reducing the learning error.
- Few pieces of research have been through about DDL and DML based load forecasting process [63], [87], [101], [115], [121]. Most researchers have tried to ignore the BP process for reducing computational time among these works. These works also indicate that DML and DDL can be created a potential impact on the load forecasting for reducing computational time, big data, and limited data-related issues.
- After reviewing the works, we have found that training or computational time, learning error, additional BP process, Big Data, data requirement, or limited data are major challenges at today's load forecasting processes. We have also found that DDL-based load forecasting could be a potential future research scope to mitigate the challenges.
- After studying the research works, current challenges, and possible research scopes, we proposed the HSIC Bottleneck based DDL model for load forecasting. Here, HSIC Bottleneck is a deep learning process that provides high accuracy without BP. The revealed conceptual model is based on distributed neural network architecture. Here, multiple load forecasting process occurs from multiple areas and aggregate the prediction data at the cloud server. For that reason, no need for huge data aggregation at the cloud server and no need limited data-oriented load forecasting process. Therefore, the model would enable us to reduce the challenges.

VI. CONCLUSION

Nowadays, VSTLF, STLF, MTLF, and LTLF approaches are necessary to estimate power suppliers' required electrical power demand. Within these approaches, STLF had been concerned for reducing day-ahead and hourly ahead load forecasting related problems to many researchers, while very few researchers had been interested in LTLF and VSTLF. After the STLF, many researchers had concerned about MTLF. Despite all these identified studies, this research is still open to apply and adapt many novel combined models for electricity and power prediction. Moreover, individual magnified attention to study VSTLF and MTLF should have been additionally dedicated to fulfilling the detected gap in the field. Deep learning from a smart grid is being intensively studied to support this load forecasting need, and there are many ANN techniques and models in this regard. Although the maximum study has been focused on hybrid ANN, LSTM is the most used among ANN techniques because it can solve many complex issues.

Maximum researchers used the BP process at load forecasting to obtain better accuracy, but it takes additional computational time. The load forecasting process takes much time because there is a huge data aggregate at the cloud server in the smart grid system. On the other hand, the BP process takes more additional processing time. For that reason, few groups of researchers had developed the non-BP-based hybrid ANN. Moreover, we have learned from this survey that researchers had not used BP at maximum LSTM based load forecasting. We have also learned from the survey that huge training time or computational time is the major issue in load forecasting because the cloud server contains enormous data. Alongside computational time, learning error, big data, limited data, etc., are becoming a concern in the forecasting process. For reducing Big data related issue in demand forecasting, researchers have tried to developed limited data based load forecasting with high accuracy where no need for massive data.

Our survey found that computational time, learning error, additional BP process, Big Data, and data requirement are the significant challenges in load forecasting. After studying the current challenges and positive aspects of the HSIC bottleneck, we have realized that the Distributed Neural Network-based DDL process becomes essential for reducing the challenges. The proposed model illustrates a novel DDL-based load forecasting model, and it also illustrates how to reduce the current challenges from this model. However, many current load forecasting proposals may suffer from data centralization, training time, and computational time issues. There is comparatively very little work on ameliorating these issues through decentralization. Therefore, a DDL model for load forecasting may be proposed that could solve some of these issues. It is expected that the findings of this survey and the proposed model will be of benefit to researchers, policymakers, and practitioners in this field.

Moreover, this research's contribution would significantly impact developing DDL-based load forecasting and understanding the current challenges. This paper's first significant contribution is to acknowledge the current challenges of load forecasting such as computational time, learning error, other BP processes, big data, data requirement or limited data, Feature selection and extraction, and meteorological data (load, temperature, humidity). This paper's second significant contribution is to demonstrate the DDL-based load forecasting, where we have discussed how a Distributed Neural Network can enable us to reduce the current challenges. It needs to experiment on this model for obtaining real visibility and mitigating the current issues and challenges. In this study's future work, we will try to develop the proposed model for a real-life smart grid perspective. The future study could think of the real-time implementation to improve the model and measure the HSIC bottleneck, DDL, and Distributed Neural Network from the load forecasting point of view. Besides, Blockchain is an innovative technology that is used in the decentralized and distributed system. Alongside AI, we will also try to apply blockchain technology in future studies.

ACKNOWLEDGMENT

The authors would like to acknowledge the support given by the Universiti Kebangsaan Malaysia (UKM) under the Grant GGPM-2020-028, FRGS-1-2019-ICT-UKM-02, and FRGS/1/2020/ICT03/UKM/02/6. The author also acknowledges to Oregon Institute of Technology, USA, and the Asian University of Bangladesh for the collaborative work.

CONFLICTS OF INTEREST

The authors declare no conflict of interest.

REFERENCES

[1] A. H. Al-Badi, R. Ahshan, N. Hosseinzadeh, R. Ghorbani, and E. Hossain, "Survey of smart grid concepts and technological demonstrations worldwide emphasizing on the oman perspective," *Appl. Syst. Innov.*, vol. 3, no. 1, 2020, Art. no. 5.

[2] C. Chalmers, P. Fergus, C. A. C. Montanez, S. Sikdar, F. Ball, and B. Kendall, "Detecting activities of daily living and routine behaviours in dementia patients living alone using smart meter load disaggregation," *IEEE Trans. Emerg. Topics Comput.*, early access, May 11, 2020, doi: 10.1109/TETC.2020.2993177.

[3] B. Thayer, D. Engel, I. Chakraborty, K. Schneider, L. Ponder, and K. Fox, "Improving end-use load modeling using machine learning and smart meter data," in *Proc. 53rd Hawaii Int. Conf. Syst. Sci.*, Maui, HI, USA, 2020, pp. 3055–3064.

[4] Y.-Y. Chen, Y.-H. Lin, C.-C. Kung, M.-H. Chung, and I.-H. Yen, "Design and implementation of cloud analytics-assisted smart power meters considering advanced artificial intelligence as edge analytics in demand-side management for smart homes," *Sensors*, vol. 19, no. 9, p. 2047, May 2019.

[5] R. T. Kreutzer and M. Sirrenberg, "Fields of application of artificial intelligence—Energy sector, smart home, mobility and transport," in *Understanding Artificial Intelligence, Management for Professionals*. Singapore: Springer, Sep. 2019, pp. 195–210.

[6] E. Hossain, I. Khan, F. Un-Noor, S. S. Sikander, and M. S. H. Sunny, "Application of big data and machine learning in smart grid, and associated security concerns: A review," *IEEE Access*, vol. 7, pp. 13960–13988, 2019.

[7] A. K. Ozcanli, F. Yaprakdal, and M. Baysal, "Deep learning methods and applications for electrical power systems: A comprehensive review," *Int. J. Energy Res.*, vol. 40, no. 9, pp. 7136–7157, 2020.

[8] L. Hernandez, C. Baladron, J. M. Aguiar, B. Carro, A. J. Sanchez-Esguevillas, J. Lloret, and J. Massana, "A survey on electric power demand forecasting: Future trends in smart grids, microgrids and smart buildings," *IEEE Commun. Surveys Tuts.*, vol. 16, no. 3, pp. 1460–1495, 3rd Quart., 2014.

[9] A. A. Mir *et al.*, "A review of electricity demand forecasting in low and middle income countries: The demand determinants and horizons," *Sustainability*, vol. 12, no. 15, 2020, Art. no. 5931.

[10] A. A. Mamun, M. Hoq, E. Hossain, and R. Bayindir, "A hybrid deep learning model with evolutionary algorithm for short-term load forecasting," in *Proc. 8th Int. Conf. Renew. Energy Res. Appl. (ICRERA)*, Brasov, Romania, Nov. 2019, pp. 886–891, doi: 10.1109/ICRERA47325.2019.8996550.

[11] S. M. Miraftabzadeh, F. Foiadelli, M. Longo, and M. Pasetti, "A survey of machine learning applications for power system analytics," in *Proc. IEEE Int. Conf. Environ. Electr. Eng., IEEE Ind. Commercial Power Syst. Eur. (EEEIC I CPS Europe)*, Genova, Italy, Jun. 2019, pp. 1–5.

[12] A. Almalaq and G. Edwards, "A review of deep learning methods applied on load forecasting," in *Proc. 16th IEEE Int. Conf. Mach. Learn. Appl. (ICMLA)*, Cancun, Mexico, Dec. 2017, pp. 511–516.

[13] A. Arif, N. Javaid, M. Anwar, A. Naeem, H. Gul, and S. Fareed, ''Electricity load and price forecasting using machine learning algorithms in smart grid: A survey,'' in *Web, Artificial Intelligence and Network Application—WAINA* (Advances in Intelligent Systems and Computing), vol. 1150, L. Barolli, Ed. Cham, Switzerland: Springer, 2020, pp. 471–483.

[14] T. Ben-Nun and T. Hoefler, ''Demystifying parallel and distributed deep learning: An in-depth concurrency analysis,'' *ACM Comput. Surveys*, vol. 52, no. 4, Sep. 2019, Art. no. 65.

[15] J. Zhang, A. Hasandka, J. Wei, S. Alam, T. Elgindy, A. Florita, and B.-M. Hodge, ''Hybrid communication architectures for distributed smart grid applications,'' *Energies*, vol. 11, no. 4, Apr. 2018, Art. no. 871.

[16] I. Ioannou, V. Vassiliou, C. Christophorou, and A. Pitsillides, ''Distributed artificial intelligence solution for D2D communication in 5G networks,'' *IEEE Syst. J.*, vol. 14, no. 3, pp. 4232–4241, Sep. 2020.

[17] G. A. Montes and B. Goertzel, ''Distributed, decentralized, and democratized artificial intelligence,'' *Technol. Forecasting Social Change*, vol. 141, pp. 354–358, Apr. 2019.

[18] J. Verbraeken, M. Wolting, J. Katzy, J. Kloppenburg, T. Verbelen, and J. S. Rellermeyer, ''A survey on distributed machine learning,'' *ACM Comput. Surveys*, vol. 53, no. 2, 2020, Art. no. 30.

[19] Z. Tang, S. Shi, X. Chu, W. Wang, and B. Li, ''Communication-efficient distributed deep learning: A comprehensive survey,'' 2020, *arXiv:2003.06307*. [Online]. Available: http://arxiv.org/abs/2003.06307

[20] N. Sakib, E. Hossain, and S. I. Ahamed, ''A qualitative study on the united states Internet of energy: A step towards computational sustainability,'' *IEEE Access*, vol. 8, pp. 69003–69037, 2020, doi: 10.1109/ACCESS.2020.2986317.

[21] D. Dias, F. C. Delicato, P. F. Pires, A. R. Rocha, and E. Y. Nakagawa, ''An overview of reference architectures for cloud of things,'' in *Proc. 35th Annu. ACM Symp. Appl. Comput.*, Brno, Czech Republic, Mar. 2020, pp. 1498–1505.

[22] B. Alohali and V. G. Vassilakis, ''Protecting data confidentiality in the cloud of things,'' *Int. J. Hyperconnectivity Internet Things*, vol. 1, no. 1, pp. 29–46, Jan. 2017.

[23] M. J. Sadeq, S. R. Kabir, R. Haque, J. Ferdaws, M. Akhtaruzzaman, R. Forhat, and S. M. Allayear, ''A cloud of things (CoT) approach for monitoring product purchase and price hike,'' in *Intelligent Computing and Innovation on Data Science* (Lecture Notes in Networks and Systems), vol. 118, S. L. Peng, Ed. Singapore: Springer, 2020, pp. 359–368.

[24] B. He, J. Wang, Q. Qi, H. Sun, and J. Liao, ''Towards intelligent provisioning of virtualized network functions in cloud of things: A deep reinforcement learning based approach,'' *IEEE Trans. Cloud Comput.*, early access, Apr. 6, 2020, doi: 10.1109/TCC.2020.2985651.

[25] J.-H. Yi, W.-H. Xu, and Y.-T. Chen, ''Novel back propagation optimization by cuckoo search algorithm,'' *Sci. World J.*, vol. 2014, Mar. 2014, Art. no. 878262.

[26] L. T. Nghia, A. H. Quyen, H. H. Pham, and A. T. Nguyen, ''A hybrid artificial neural network-genetic algorithm for load shedding,'' *Int. J. Electr. Comput. Eng.*, vol. 10, no. 3, pp. 2250–2258, Jun. 2020.

[27] A. Repetto. (Aug. 2017). *Toward Data Science, The Problem with Back-Propagation*. Accessed: Jun. 2, 2020. [Online]. Available: https://towardsdatascience.com/the-problem-with-back-propagation-13aa84aabd71

[28] W. K. Ma, J. P. Lewis, and W. B. Kleijn, ''The HSIC bottleneck: Deep learning without back-propagation,'' in *The 34th AAAI Conf. Artif. Intell. (AAAI), AAAI Tech. Track, Machine Learn.*, New York, NY, USA, 2020, vol. 34, no. 4, pp. 5085–5092.

[29] A. Gretton, O. Bousquet, A. Smola, and B. Schölkopf, ''Measuring statistical dependence with Hilbert-Schmidt norms,'' in *Algorithmic Learning Theory* (Lecture Notes in Computer Science), vol. 3734, S. Jainet, Ed. Berlin, Germany: Springer, 2005, pp. 63–77.

[30] N. Tishby, F. Pereira, and W. Bialek, ''The information bottleneck method,'' in *Proc. 37th Annu. Allerton Conf. Commun., Control Comput.*, 1999, pp. 368–377.

[31] D. Wu, Y. Zhao, Y.-H. H. Tsai, M. Yamada, and R. Salakhutdinov, '''Dependency Bottleneck' in auto-encoding architectures: An empirical study,'' 2018, *arXiv:1802.05408*. [Online]. Available: http://arxiv.org/abs/1802.05408

[32] M. Stark, L. Wang, G. Bauch, and R. D. Wesel, ''Decoding rate-compatible 5G-LDPC codes with coarse quantization using the information bottleneck method,'' *IEEE Open J. Commun. Soc.*, vol. 1, pp. 646–660, 2020.

[33] A. Steiner and S. S. Shitz, ''Broadcast approach for the information bottleneck channel,'' in *Proc. IEEE Int. Conf. Microw., Antennas, Commun. Electron. Syst. (COMCAS)*, Tel-Aviv, Israel, Nov. 2019, pp. 1–5.

[34] M. Q. Raza and A. Khosravi, ''A review on artificial intelligence based load demand forecasting techniques for smart grid and buildings,'' *Renew. Sustain. Energy Rev.*, vol. 50, pp. 1352–1372, Oct. 2015.

[35] C. Guan, P. B. Luh, L. D. Michel, Y. Wang, and P. B. Friedland, ''Very short-term load forecasting: Wavelet neural networks with data pre-filtering,'' *IEEE Trans. Power Syst.*, vol. 28, no. 1, pp. 30–41, Feb. 2013.

[36] M. Jacob, C. Neves, and D. V. Greetham, ''Short term load forecasting,'' in *Forecasting and Assessing Risk of Individual Electricity Peaks, Mathematics of Planet Earth*. Cham, Switzerland: Springer, Sep. 2019, pp. 15–37.

[37] P. Bunnoon, K. Chalermyanont, and C. Limsakul, ''Mid term load forecasting of the country using statistical methodology: Case study in thailand,'' in *Proc. Int. Conf. Signal Process. Syst.*, 2009, pp. 324–928.

[38] K. B. Lindberg, P. Seljom, H. Madsen, D. Fischer, and M. Korpås, ''Long-term electricity load forecasting: Current and future trends,'' *Utilities Policy*, vol. 58, pp. 102–119, Jun. 2019.

[39] Z. Wen, L. Xie, Q. Fan, and H. Feng, ''Long term electric load forecasting based on TS-type recurrent fuzzy neural network model,'' *Electr. Power Syst. Res.*, vol. 179, Feb. 2020, Art. no. 106106.

[40] R.-J. Park, K.-B. Song, and B.-S. Kwon, ''Short-term load forecasting algorithm using a similar day selection method based on reinforcement learning,'' *Energies*, vol. 13, no. 10, May 2020, Art. no. 2640.

[41] Z. Yuan, W. Wang, H. Wang, and S. Mizzi, ''Combination of cuckoo search and wavelet neural network for midterm building energy forecast,'' *Energy*, vol. 202, Jul. 2020, Art. no. 117728.

[42] M. Talaat, M. A. Farahat, N. Mansour, and A. Y. Hatata, ''Load forecasting based on grasshopper optimization and a multilayer feed-forward neural network using regressive approach,'' *Energy*, vol. 196, Apr. 2020, Art. no. 117087.

[43] A. Masoumi, F. Jabari, S. G. Zadeh, and B. Mohammadi-Ivatloo, ''Long-term load forecasting approach using dynamic feed-forward back-propagation artificial neural network,'' in *Optimization of Power System Problems* (Studies in Systems, Decision and Control), vol. 262, M. P. Hajiabbas and B. Mohammadi-Ivatloo, Eds. Cham, Switzerland: Springer, 2020, pp. 233–257.

[44] E. Yundra, U. Negeri Surabaya, U. Kartini, L. Wardani, D. Ardianto, U. N. Surabaya, U. N. Surabaya, and U. N. Surabaya, ''Hybrid model combined fuzzy multi-objective decision making with feed forward neural network (F-MODM-FFNN) for very short-term load forecasting based on weather data,'' *Int. J. Intell. Eng. Syst.*, vol. 13, no. 4, pp. 182–195, Aug. 2020.

[45] M. J. A. Soeb *et al.*, ''Application of advanced back propagation algorithm in electric load forecasting,'' *J. Math. Stat. Sci.*, vol. 6, no. 3, pp. 95–102, 2020.

[46] M. Alipour, J. Aghaei, M. Norouzi, T. Niknam, S. Hashemi, and M. Lehtonen, ''A novel electrical net-load forecasting model based on deep neural networks and wavelet transform integration,'' *Energy*, vol. 205, Aug. 2020, Art. no. 118106.

[47] Q. Huang, J. Li, and M. Zhu, ''An improved convolutional neural network with load range discretization for probabilistic load forecasting,'' *Energy*, vol. 203, Jul. 2020, Art. no. 117902.

[48] L. Wen, K. Zhou, and S. Yang, ''Load demand forecasting of residential buildings using a deep learning model,'' *Electr. Power Syst. Res.*, vol. 179, Feb. 2020, Art. no. 106073.

[49] G. Hafeez, K. S. Alimgeer, and I. Khan, ''Electric load forecasting based on deep learning and optimized by heuristic algorithm in smart grid,'' *Appl. Energy*, vol. 269, Jul. 2020, Art. no. 114915.

[50] A. S. Khwaja, A. Anpalagan, M. Naeem, and B. Venkatesh, ''Joint bagged-boosted artificial neural networks: Using ensemble machine learning to improve short-term electricity load forecasting,'' *Electr. Power Syst. Res.*, vol. 179, Feb. 2020, Art. no. 106080.

[51] M. R. Müller, G. Gaio, E. M. Carreno, A. D. P. Lotufo, and L. A. Teixeira, ''Electrical load forecasting in disaggregated levels using fuzzy ARTMAP artificial neural network and noise removal by singular spectrum analysis,'' *Social Netw. Appl. Sci.*, vol. 2, no. 7, Jul. 2020, Art. no. 1218.

[52] O. Samuel, F. A. Alzahrani, R. J. U. Hussen Khan, H. Farooq, M. Shafiq, M. K. Afzal, and N. Javaid, ''Towards modified entropy mutual information feature selection to forecast medium-term load using a deep learning model in smart homes,'' *Entropy*, vol. 22, no. 1, Jan. 2020, Art. no. 68.

[53] J. Li, D. Deng, J. Zhao, D. Cai, W. Hu, M. Zhang, and Q. Huang, "A novel hybrid short-term load forecasting method of smart grid using MLR and LSTM neural network," *IEEE Trans. Ind. Informat.*, early access, Jun. 5, 2020, doi: 10.1109/TII.2020.3000184.

[54] H. H. H. Aly, "A proposed intelligent short-term load forecasting hybrid models of ANN, WNN and KF based on clustering techniques for smart grid," *Electr. Power Syst. Res.*, vol. 182, May 2020, Art. no. 106191.

[55] Z. Deng, B. Wang, H. Guo, C. Chai, Y. Wang, and Z. Zhu, "Unified quantile regression deep neural network with time-cognition for probabilistic residential load forecasting," *Complexity*, vol. 2020, Jan. 2020, Art. no. 9147545.

[56] W. Ahmad, N. Ayub, T. Ali, M. Irfan, M. Awais, M. Shiraz, and A. Glowacz, "Towards short term electricity load forecasting using improved support vector machine and extreme learning machine," *Energies*, vol. 13, no. 11, Jun. 2020, Art. no. 2907.

[57] K. Yudantaka, J.-S. Kim, and H. Song, "Dual deep learning networks based load forecasting with partial real-time information and its application to system marginal price prediction," *Energies*, vol. 13, no. 1, Dec. 2019, Art. no. 148.

[58] Y. Hong, Y. Zhou, Q. Li, W. Xu, and X. Zheng, "A deep learning method for short-term residential load forecasting in smart grid," *IEEE Access*, vol. 8, pp. 55785–55797, 2020.

[59] Q. Zhang, Y. Ma, G. Li, J. Ma, and J. Ding, "Short-term load forecasting based on frequency domain decomposition and deep learning," *Math. Problems Eng.*, vol. 2020, Feb. 2020, Art. no. 7240320.

[60] L. Sehovac and K. Grolinger, "Deep learning for load forecasting: Sequence to sequence recurrent neural networks with attention," *IEEE Access*, vol. 8, pp. 36411–36426, 2020.

[61] L. Wang, S. Mao, B. M. Wilamowski, and R. M. Nelms, "Ensemble learning for load forecasting," *IEEE Trans. Green Commun. Netw.*, vol. 4, no. 2, pp. 616–628, Jun. 2020.

[62] A. Estebsari and R. Rajabi, "Single residential load forecasting using deep learning and image encoding techniques," *Electronics*, vol. 9, no. 1, Jan. 2020, Art. no. 68.

[63] D. Syed, S. S. Refaat, and H. Abu-Rub, "Performance evaluation of distributed machine learning for load forecasting in smart grids," in *Proc. Cybern. Informat. (K I)*, Velke Karlovice, Czech Republic, Jan. 2020, pp. 1–6.

[64] X. Zhang *et al.*, "Deep-learning-based probabilistic forecasting of electric vehicle charging load with a novel queuing model," *IEEE Trans. Cybern.*, 2020.

[65] M. Adil, N. Javaid, N. Daood, M. Asim, I. Ullah, and M. Bilal, "Big data based electricity price forecasting using enhanced convolutional neural network in the smart grid," in *Web, Artificial Intelligence and Network Applications* (Advances in Intelligent Systems and Computing), vol. 1150, L. Barolli, Ed. Cham, Switzerland: Springer, 2020, pp. 1189–1201.

[66] M. Sun, T. Zhang, Y. Wang, G. Strbac, and C. Kang, "Using Bayesian deep learning to capture uncertainty for residential net load forecasting," *IEEE Trans. Power Syst.*, vol. 35, no. 1, pp. 188–201, Jan. 2020.

[67] M. Saviozzi, S. Massucco, and F. Silvestro, "Implementation of advanced functionalities for distribution management systems: Load forecasting and modeling through artificial neural networks ensembles," *Electr. Power Syst. Res.*, vol. 167, pp. 230–239, Feb. 2019.

[68] S. Mujeeb, T. A. Alghamdi, S. Ullah, A. Fatima, N. Javaid, and T. Saba, "Exploiting deep learning for wind power forecasting based on big data analytics," *Appl. Sci.*, vol. 9, no. 20, Oct. 2019, Art. no. 4417.

[69] R. Talavera-Llames, R. Pérez-Chacón, A. Troncoso, and F. Martínez-Álvarez, "MV-kWNN: A novel multivariate and multi-output weighted nearest neighbours algorithm for big data time series forecasting," *Neurocomputing*, vol. 353, pp. 56–73, Aug. 2019.

[70] S. Mujeeb, N. Javaid, M. Ilahi, Z. Wadud, F. Ishmanov, and M. Afzal, "Deep long short-term memory: A new price and load forecasting scheme for big data in smart cities," *Sustainability*, vol. 11, no. 4, Feb. 2019, Art. no. 987.

[71] S. Mujeeb and N. Javaid, "ESAENARX and DE-RELM: Novel schemes for big data predictive analytics of electricity load and price," *Sustain. Cities Soc.*, vol. 51, Nov. 2019, Art. no. 101642.

[72] S. Yang, J. Wu, H. Qin, Q. Xie, Z. Xu, and Y. Hua, "Distributed buildings energy storage charging load forecasting method considering parallel deep learning model," *Concurrency Comput., Pract. Exper.*, Nov. 2019, Art. no. e5580.

[73] H. Dagdougui, F. Bagheri, H. Le, and L. Dessaint, "Neural network model for short-term and very-short-term load forecasting in district buildings," *Energy Buildings*, vol. 203, Nov. 2019, Art. no. 109408.

[74] T. Shi, F. Mei, J. Lu, J. Lu, Y. Pan, C. Zhou, J. Wu, and J. Zheng, "Phase space reconstruction algorithm and deep learning-based very short-term bus load forecasting," *Energies*, vol. 12, no. 22, Nov. 2019, Art. no. 4349.

[75] U. T. Kartini, D. Ardianto, and L. Wardani, "Very short term load forecasting based on meteorological with modelling k-NN-feed forward neural network," *J. Electr. Syst.*, vol. 15, no. 1, pp. 1–16, Mar. 2019.

[76] H. Chen, S. Wang, S. Wang, and Y. Li, "Day-ahead aggregated load forecasting based on two-terminal sparse coding and deep neural network fusion," *Electr. Power Syst. Res.*, vol. 177, Dec. 2019, Art. no. 105987.

[77] Y. Huang, N. Wang, W. Gao, X. Guo, C. Huang, T. Hao, and J. Zhan, "LoadCNN: A low training cost deep learning model for day-ahead individual residential load forecasting," Dec. 2019, *arXiv:1908.00298*. [Online]. Available: http://arxiv.org/abs/1908.00298

[78] Q. Zhang, J. Lu, Z. Yang, and M. Tu, "A deep learning based real-time load forecasting method in electricity spot market," *J. Phys., Conf. Ser.*, vol. 1176, no. 6, Mar. 2019, Art. no. 062068.

[79] G. T. Ribeiro, V. C. Mariani, and L. D. S. Coelho, "Enhanced ensemble structures using wavelet neural networks applied to short-term load forecasting," *Eng. Appl. Artif. Intell.*, vol. 82, pp. 272–281, Jun. 2019.

[80] K. Chen, K. Chen, Q. Wang, Z. He, J. Hu, and J. He, "Short-term load forecasting with deep residual networks," *IEEE Trans. Smart Grid*, vol. 10, no. 4, pp. 3943–3952, Jul. 2019.

[81] R. Zhu, W. Guo, and X. Gong, "Short-term load forecasting for CCHP systems considering the correlation between heating, gas and electrical loads based on deep learning," *Energies*, vol. 12, no. 17, Aug. 2019, Art. no. 3308.

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

[83] Y. Liang, D. Niu, and W.-C. Hong, "Short term load forecasting based on feature extraction and improved general regression neural network model," *Energy*, vol. 166, pp. 653–663, Jan. 2019.

[84] X. Tang, Y. Dai, T. Wang, and Y. Chen, "Short-term power load forecasting based on multi-layer bidirectional recurrent neural network," *IET Gener., Transmiss. Distrib.*, vol. 13, no. 17, pp. 3847–3854, Sep. 2019.

[85] S. Wang, X. Wang, S. Wang, and D. Wang, "Bi-directional long short-term memory method based on attention mechanism and rolling update for short-term load forecasting," *Int. J. Electr. Power Energy Syst.*, vol. 109, pp. 470–479, Jul. 2019.

[86] S. H. Pramono, M. Rohmatillah, E. Maulana, R. N. Hasanah, and F. Hario, "Deep learning-based short-term load forecasting for supporting demand response program in hybrid energy system," *Energies*, vol. 12, no. 17, p. 3359, Aug. 2019.

[87] Y. Yang, W. Hong, and S. Li, "Deep ensemble learning based probabilistic load forecasting in smart grids," *Energy*, vol. 189, Dec. 2019, Art. no. 116324.

[88] P. Liu, P. Zheng, and Z. Chen, "Deep learning with stacked denoising auto-encoder for short-term electric load forecasting," *Energies*, vol. 12, no. 12, p. 2445, Jun. 2019.

[89] T. Ouyang, Y. He, H. Li, Z. Sun, and S. Baek, "Modeling and forecasting short-term power load with copula model and deep belief network," *IEEE Trans. Emerg. Topics Comput. Intell.*, vol. 3, no. 2, pp. 127–136, Apr. 2019.

[90] S. Bouktif, A. Fiaz, A. Ouni, and M. A. Serhani, "Single and multi-sequence deep learning models for short and medium term electric load forecasting," *Energy*, vol. 12, no. 1, Jan. 2019, Art. no. 149.

[91] S. Motepe, A. N. Hasan, and R. Stopforth, "Improving load forecasting process for a power distribution network using hybrid AI and deep learning algorithms," *IEEE Access*, vol. 7, pp. 82584–82598, 2019.

[92] Q. Chen, M. Xia, T. Lu, X. Jiang, W. Liu, and Q. Sun, "Short-term load forecasting based on deep learning for end-user transformer subject to volatile electric heating loads," *IEEE Access*, vol. 7, pp. 162697–162707, 2019.

[93] K. H. Kim, B. Chang, and H. K. Choi, "Deep learning based short-term electric load forecasting models using one-hot encoding," *J. Inst. Korean Electr. Electron. Eng.*, vol. 23, no. 3, pp. 852–857, Sep. 2019.

[94] H. J. Sadaei, P. C. de Lima e Silva, F. G. Guimarães, and M. H. Lee, "Short-term load forecasting by using a combined method of convolutional neural networks and fuzzy time series," *Energy*, vol. 175, pp. 365–377, May 2019.

[95] D. Liu, L. Zeng, C. Li, K. Ma, Y. Chen, and Y. Cao, "A distributed short-term load forecasting method based on local weather information," *IEEE Syst. J.*, vol. 12, no. 1, pp. 208–215, Mar. 2018.

[96] Y. K. Semero, J. Zhang, D. Zheng, and D. Wei, "An accurate very short-term electric load forecasting model with binary genetic algorithm based feature selection for microgrid applications," *Electr. Power Compon. Syst.*, vol. 46, nos. 14–15, pp. 1570–1579, 2018.

[97] E. N. Osegi, "Using the hierarchical temporal memory spatial pooler for short-term forecasting of electrical load time series," *Appl. Comput. Inform.*, Jul. 2018.

[98] M. M. Tripathi, "Zero initialization of modified gated recurrent encoder-decoder network for short term load forecasting," Dec. 2018, *arXiv:1812.03425*. [Online]. Available: http://arxiv.org/abs/1812.03425

[99] N. Ammar, M. Sulaiman, and A. F. M. Nor, "Long-term load forecasting of power systems using artificial neural network and ANFIS," *ARPN J. Eng. Appl. Sci.*, vol. 13, no. 3, pp. 828–834, Feb. 2018.

[100] A. Heydari, F. Keynia, D. A. Garcia, and L. De Santoli, "Mid-term load power forecasting considering environment emission using a hybrid intelligent approach," in *Proc. 5th Int. Symp. Environ.-Friendly Energies Appl. (EFEA)*, Rome, Italy, Sep. 2018, pp. 1–5.

[101] H. Huang, H. Xu, Y. Cai, R. S. Khalid, and H. Yu, "Distributed machine learning on smart-gateway network toward real-time smart-grid energy management with behavior cognition," *ACM Trans. Design Autom. Electron. Syst.*, vol. 23, no. 5, Oct. 2018, Art. no. 56.

[102] T. Abreu, A. J. Amorim, C. R. Santos-Junior, A. D. P. Lotufo, and C. R. Minussi, "Multinodal load forecasting for distribution systems using a fuzzy-artmap neural network," *Appl. Soft Comput.*, vol. 71, pp. 307–316, Oct. 2018.

[103] C. Tong, J. Li, C. Lang, F. Kong, J. Niu, and J. J. P. C. Rodrigues, "An efficient deep model for day-ahead electricity load forecasting with stacked denoising auto-encoders," *J. Parallel Distrib. Comput.*, vol. 117, pp. 267–273, Jul. 2018.

[104] Y. Song, H. Chen, K. Yuan, C. Sun, Z. Xue, X. Jin, W. Liu, and J. Han, "Medium and long term load forecasting considering the uncertainty of distributed installed capacity of photovoltaic generation," in *Proc. 13th IEEE Conf. Ind. Electron. Appl. (ICIEA)*, Wuhan, China, May 2018, pp. 1691–1696.

[105] M. Rafiei, T. Niknam, J. Aghaei, M. Shafie-Khah, and J. P. S. Catalao, "Probabilistic load forecasting using an improved wavelet neural network trained by generalized extreme learning machine," *IEEE Trans. Smart Grid*, vol. 9, no. 6, pp. 6961–6971, Nov. 2018.

[106] Y. Chen, M. Kloft, Y. Yang, C. Li, and L. Li, "Mixed kernel based extreme learning machine for electric load forecasting," *Neurocomputing*, vol. 312, pp. 90–106, Oct. 2018.

[107] T. Yuniarti, I. Surjandari, E. Muslim, and E. Laoh, "Data mining approach for short term load forecasting by combining wavelet transform and group method of data handling (WGMDH)," in *Proc. 3rd Int. Conf. Sci. Inf. Technol. (ICSITech)*, Bandung, Indonesia, Oct. 2017, pp. 53–58.

[108] I. A. Samuel, E. Adetiba, I. A. Odigwe, and F. C. Felly-Njoku, "A comparative study of regression analysis and artificial neural network methods for medium-term load forecasting," *Indian J. Sci. Technol.*, vol. 10, no. 10, pp. 1–7, Mar. 2017.

[109] S. Sepasi, E. Reihani, A. M. Howlader, L. R. Roose, and M. M. Matsuura, "Very short term load forecasting of a distribution system with high PV penetration," *Renew. Energy*, vol. 106, pp. 142–148, Jun. 2017.

[110] N. Zeng, H. Zhang, W. Liu, J. Liang, and F. E. Alsaadi, "A switching delayed PSO optimized extreme learning machine for short-term load forecasting," *Neurocomputing*, vol. 240, pp. 175–182, May 2017.

[111] Y. He, J. Deng, and H. Li, "Short-term power load forecasting with deep belief network and copula models," in *Proc. 9th Int. Conf. Intell. Hum.-Mach. Syst. Cybern. (IHMSC)*, Hangzhou, China, Aug. 2017, pp. 191–194.

[112] X. Dong, L. Qian, and L. Huang, "Short-term load forecasting in smart grid: A combined CNN and K-means clustering approach," in *Proc. IEEE Int. Conf. Big Data Smart Comput. (BigComp)*, Jeju, South Korea, Feb. 2017, pp. 119–125.

[113] D. Gan, Y. Wang, N. Zhang, and W. Zhu, "Enhancing short-term probabilistic residential load forecasting with quantile long-short-term memory," *J. Eng.*, vol. 2017, no. 14, pp. 2622–2627, 2017.

[114] B. Li, J. Zhang, Y. He, and Y. Wang, "Short-term load-forecasting method based on wavelet decomposition with second-order gray neural network model combined with ADF test," *IEEE Access*, vol. 5, pp. 16324–16331, 2017.

[115] R.-D. Wang, X.-S. Sun, X. Yang, and H. Hu, "Cloud computing and extreme learning machine for a distributed energy consumption forecasting in equipment-manufacturing enterprises," *Cybern. Inf. Technol.*, vol. 16, no. 6, pp. 83–97, Dec. 2016.

[116] M. Mordjaoui, S. Haddad, A. Medoued, and A. Laouafi, "Electric load forecasting by using dynamic neural network," *Int. J. Hydrogen Energy*, vol. 42, no. 28, pp. 17655–17663, Jul. 2017.

[117] A. S. Khwaja, X. Zhang, A. Anpalagan, and B. Venkatesh, "Boosted neural networks for improved short-term electric load forecasting," *Electr. Power Syst. Res.*, vol. 143, pp. 431–437, Feb. 2017.

[118] 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*, vol. 10, no. 8, p. 1168, Aug. 2017.

[119] Y. Liu, W. Wang, and N. Ghadimi, "Electricity load forecasting by an improved forecast engine for building level consumers," *Energy*, vol. 139, pp. 18–30, Nov. 2017.

[120] A. O. Melodi, S. T. Adeniyi, and R. H. Oluwaniyi, "Long term load forecasting for Nigeria's electric power grid using ANN and fuzzy logic models," in *Proc. IEEE 3rd Int. Conf. Electro-Technol. Nat. Develop. (NIGERCON)*, Owerri, Nigeria, Nov. 2017, pp. 962–968.

[121] H. Xu, H. Huang, R. S. Khalid, and H. Yu, "Distributed machine learning based smart-grid energy management with occupant cognition," in *Proc. IEEE Int. Conf. Smart Grid Commun. (SmartGridComm)*, Sydney, NSW, Australia, Nov. 2016, pp. 491–496.

[122] M. N. Rahman, A. Esmailpour, and J. Zhao, "Machine learning with big data an efficient electricity generation forecasting system," *Big Data Res.*, vol. 5, pp. 9–15, Sep. 2016.

[123] J. Cheng, N. Zhang, Y. Wang, C. Kang, W. Zhu, M. Luo, and H. Que, "Evaluating the spatial correlations of multi-area load forecasting errors," in *Proc. Int. Conf. Probabilistic Methods Appl. Power Syst. (PMAPS)*, Beijing, China, Oct. 2016, pp. 1–6.

[124] K. Grolinger, A. L'Heureux, M. A. M. Capretz, and L. Seewald, "Energy forecasting for event venues: Big data and prediction accuracy," *Energy Buildings*, vol. 112, pp. 222–233, Jan. 2016.

[125] D. L. Marino, K. Amarasinghe, and M. Manic, "Building energy load forecasting using deep neural networks," in *Proc. 42nd Annu. Conf. IEEE Ind. Electron. Soc. (IECON)*, Florence, Italy, Oct. 2016, pp. 7046–7051.

[126] Y. Yunsong, "Research of power load forecasting model based on improved liquid state machine," *Autom. Instrum.*, vol. 2016, no. 11, p. 71, 2016.

[127] R. M. Nezzar, N. Farah, M. T. Khadir, and L. Chouireb, "Mid-long term load forecasting using multi-model artificial neural networks," *Int. J. Electr. Eng. Informat.*, vol. 8, no. 2, pp. 389–401, Jun. 2016.

[128] Y. Feng, "Study on medium and long term power load forecasting based on combination forecasting model," *Chem. Eng. Trans.*, vol. 51, pp. 859–864, Aug. 2016.

[129] I. N. da Silva and L. C. M. de Andrade, "Efficient neurofuzzy model to very short-term load forecasting," *IEEE Latin Amer. Trans.*, vol. 14, no. 2, pp. 721–728, Feb. 2016.

[130] A. Dedinec, S. Filiposka, A. Dedinec, and L. Kocarev, "Deep belief network based electricity load forecasting: An analysis of macedonian case," *Energy*, vol. 115, pp. 1688–1700, Nov. 2016.

[131] V. N. Coelho, I. M. Coelho, E. Rios, A. S. T. Filho, A. J. R. Reis, B. N. Coelho, A. Alves, G. G. Netto, M. J. F. Souza, and F. G. Guimarães, "A hybrid deep learning forecasting model using GPU disaggregated function evaluations applied for household electricity demand forecasting," *Energy Procedia*, vol. 103, pp. 280–285, Dec. 2016.

[132] M. M. Eljazzar and E. E. Hemayed, "Feature selection and optimization of artificial neural network for short term load forecasting," in *Proc. 18th Int. Middle East Power Syst. Conf. (MEPCON)*, Cairo, Egypt, Dec. 2016, pp. 827–831.

[133] G. M. Khan and F. Zafari, "Dynamic feedback neuro-evolutionary networks for forecasting the highly fluctuating electrical loads," *Genetic Program. Evolvable Mach.*, vol. 17, no. 4, pp. 391–408, May 2016.

[134] X. Lu, J. Wang, Y. Cai, and J. Zhao, "Distributed HS-ARTMAP and its forecasting model for electricity load," *Appl. Soft Comput.*, vol. 32, pp. 13–22, Jul. 2015.

[135] S. Jurado, À. Nebot, F. Mugica, and N. Avellana, "Hybrid methodologies for electricity load forecasting: Entropy-based feature selection with machine learning and soft computing techniques," *Energy*, vol. 86, pp. 276–291, Jun. 2015.

[136] T. S. Mahmoud, D. Habibi, M. Y. Hassan, and O. Bass, "Modelling self-optimised short term load forecasting for medium voltage loads using tunning fuzzy systems and artificial neural networks," *Energy Convers. Manage.*, vol. 106, pp. 1396–1408, Dec. 2015.
[137] F. M. Bianchi, E. De Santis, A. Rizzi, and A. Sadeghian, "Short-term electric load forecasting using echo state networks and PCA decomposition," *IEEE Access*, vol. 3, pp. 1931–1943, 2015.
[138] A. Dolara, F. Grimaccia, S. Leva, M. Mussetta, and E. Ogliari, "A physical hybrid artificial neural network for short term forecasting of PV plant power output," *Energies*, vol. 8, no. 2, pp. 1138–1153, Feb. 2015.
[139] I. Koprinska, M. Rana, and V. G. Agelidis, "Correlation and instance based feature selection for electricity load forecasting," *Knowl.-Based Syst.*, vol. 82, pp. 29–40, Jul. 2015.
[140] H. Chitsaz, H. Shaker, H. Zareipour, D. Wood, and N. Amjady, "Short-term electricity load forecasting of buildings in microgrid," *Energy Buildings*, vol. 99, pp. 50–60, Jul. 2015.
[141] C. Cecati, J. Kolbusz, P. Rozycki, P. Siano, and B. M. Wilamowski, "A novel RBF training algorithm for short-term electric load forecasting and comparative studies," *IEEE Trans. Ind. Electron.*, vol. 62, no. 10, pp. 6519–6529, Oct. 2015.
[142] L. Wang, Y. Zeng, and T. Chen, "Back propagation neural network with adaptive differential evolution algorithm for time series forecasting," *Expert Syst. Appl.*, vol. 42, no. 2, pp. 855–863, Feb. 2015.
[143] Y. W. Foo, C. Goh, H. C. Lim, Z.-H. Zhan, and Y. Li, "Evolutionary neural network based energy consumption forecast for cloud computing," in *Proc. Int. Conf. Cloud Comput. Res. Innov. (ICCCRI)*, Singapore, Oct. 2015, pp. 53–64.
[144] A. K. M. A. Habib, S. M. A. Motakabber, M. I. Ibrahimy, and M. K. Hasan, "Active voltage balancing circuit using single switched-capacitor and series LC resonant energy carrier," *Electron. Lett.*, vol. 56, no. 20, pp. 1036–1039, Jul. 2020.
[145] M. K. Hasan, M. M. Ahmed, A. H. A. Hashim, A. Razzaque, S. Islam, and B. Pandey, "A novel artificial intelligence based timing synchronization scheme for smart grid applications," *Wireless Pers. Commun.*, vol. 23, pp. 1–8, Apr. 2020.
[146] N. Aloysius and M. Geetha, "A review on deep convolutional neural networks," in *Proc. Int. Conf. Commun. Signal Process. (ICCSP)*, Chennai, India, 2017, pp. 0588–0592.
[147] L. Khichadi, S. G. Kumar, and K. Nagamani, "Building a cloud solution for energy management using raspberry pi," in *Proc. Int. Conf. Commun. Electron. Syst. (ICCES)*, Coimbatore, India, Jul. 2019, pp. 422–426.
[148] D. I. Brandao, R. Patric dos Santos, W. Silva, T. R. De Oliveira, and P. F. Donoso-Garcia, "Model-free energy management system for hybrid AC/DC microgrids," *IEEE Trans. Ind. Electron.*, early access, Apr. 7, 2020, doi: 10.1109/TIE.2020.2984993.
[149] W. Hajji and F. Tso, "Understanding the performance of low power raspberry pi cloud for big data," *Electronics*, vol. 5, no. 4, Jun. 2016, Art. no. 29.
[150] B. Mohebali, A. Tahmassebi, A. H. Gandomi, and A. Meyer-Baese, "A big data inspired preprocessing scheme for bandwidth use optimization in smart cities applications using Raspberry Pi," in *Proc. SPIE, Big Data, Learn., Anal., Appl.*, Baltimore, MD, USA, vol. 10989, 2019, Art. no. 1098902.
[151] S. Kim and B. C. Ko, "Building deep random ferns without backpropagation," *IEEE Access*, vol. 8, pp. 8533–8542, 2020.
[152] A. Nøkland, "Direct feedback alignment provides learning in deep neural networks," in *Proc. 30th Conf. Neural Inf. Process. Syst. (NIPS)*, Barcelona, Spain, Dec. 2016, pp. 1045–1053.
[153] A. Ghazvini, S. N. H. S. Abdullah, M. Kamrul Hasan, and D. Z. A. Bin Kasim, "Crime spatiotemporal prediction with fused objective function in time delay neural network," *IEEE Access*, vol. 8, pp. 115167–115183, 2020.
[154] A. A. Mamun *et al.*, "A comprcomprehensive review of the lehensive review of the load forecasting techniques using single and hybrid predictive models," *IEEE Access*, vol. 8, pp. 134911–134939, 2020.
[155] Z. May, M. K. Alam, K. Husain, and M. K. Hasan, "An enhanced dynamic transmission opportunity scheme to support varying traffic load over wireless campus networks," *PLoS ONE*, vol. 15, no. 8, Aug. 2020, Art. no. e0238073.
[156] A. K. Raeespour and A. M. Patel, "Design and evaluation of a virtual private network architecture for collaborating specialist users," *Asia-Pacific J. Inf. Technol. Multimedia*, vol. 5, no. 1, 2016.
[157] Z. E. Ahmed, M. K. Hasan, R. A. Saeed, R. Hassan, S. Islam, R. A. Mokhtar, S. Khan, and M. Akhtaruzzaman, "Optimizing energy consumption for cloud Internet of Things," *Frontiers Phys.*, vol. 8, Oct. 2020.
[158] W. Yang, H. Gao, Y. Jiang, J. Yu, J. Sun, J. Liu, and Z. Ju, "A cascaded feature pyramid network with non-backward propagation for facial expression recognition," *IEEE Sensors J.*, early access, May 25, 2020, doi: 10.1109/JSEN.2020.2997182.
[159] S. M. Allayear, M. F. A. Bhuiyan, M. M. Alam, S. R. Kabir, M. T. A. Munna, and M. S. Hasan, "Human face detection in excessive dark image by using contrast stretching, histogram equalization and adaptive equalization," *Int. J. Eng. Technol.*, vol. 7, no. 4, pp. 3984–3989, 2018.
[160] M. S. Hasan, S. R. Kabir, M. Akhtaruzzaman, M. J. Sadeq, M. M. Alam, S. M. Allayear, M. S. Uddin, M. Rahman, R. Forhat, R. Haque, and H. A. Arju, "Identification of construction era for indian subcontinent ancient and heritage buildings by using deep learning," in *Proc. 5th Int. Congr. Inf. Commun. Technol. (ICICT)* in Advances in Intelligent Systems and Computing, vol. 1183, X. S. Yang, Ed. Singapore: Springer, 2020, pp. 631–640.
[161] S. R. Kabir, M. Akhtaruzzaman, and R. Haque, "Performance analysis of different feature detection techniques for modern and old buildings," in *Proc. CEUR Workshop*, vol. 2280, Dec. 2018, pp. 120–127.
[162] M. A. Norrie, "An introduction to machine learning," in *Data Analytics: Concepts, Techniques, and Applications*, M. Ahmed and A. S. K. Pathan, Eds. Boca Raton, FL, USA: CRC Press, 2018, doi: 10.1201/9780429446177.
[163] S. R. Kabir, S. M. Allayear, M. M. Alam, and M. T. A. Munna, "A computational technique for intelligent computers to learn and identify the human's relative directions," in *Proc. Int. Conf. Intell. Sustain. Syst. (ICISS)*, Palladam, India, Dec. 2017, pp. 1037–1040.
[164] S. R. Kabir, M. M. Alam, S. M. Allayear, M. T. A. Munna, S. S. Hossain, and S. S. M. M. Rahman, "Relative direction: Location path providing method for allied intelligent agent," in *Proc. Adv. Comput. Data Sci. (ICACDS)* in Communications in Computer and Information Science, vol. 905, M. Singh, Ed. Singapore: Springer, 2018, pp. 381–391.
[165] S. R. Kabir, "Computation of multi-agent based relative direction learning specification," M.S. thesis, Daffodil Int. Univ., Dhaka, Bangladesh, 2017.
[166] R. Haque, I. Mahmud, M. H. Sharif, S. R. Kabir, A. Chowdhury, F. Akter, and A. B. Akhi, "Modeling the role of C2C information quality on purchase decision in Facebook," in *Challenges and Opportunities in the Digital Era* (Lecture Notes in Computer Science), vol. 11195, S. Al-Sharhan, Ed. Cham, Switzerland: Springer, 2018, pp. 244–254.
[167] M. Akhtaruzzaman, "A combined model of blockchain, price intelligence and IoT for reducing the corruption and poverty," in *Proc. 6th Int. Conf. Poverty Sustain. Develop.*, Colombo, Sri Lanka, vol. 6, 2019, pp. 13–24.
[168] R. Haque, H. Sarwar, S. R. Kabir, R. Forhat, M. J. Sadeq, M. Akhtaruzzaman, and N. Haque, "Blockchain-based information security of electronic medical records (EMR) in A healthcare communication system," in *Intelligent Computing and Innovation on Data Science* (Lecture Notes in Networks and Systems), vol. 118, S. L. Peng, Ed. Singapore: Springer, 2020, p. 641.
[169] M. J. Sadeq, S. R. Kabir, M. Akter, R. Forhat, R. Haque, and M. Akhtaruzzaman, "Integration of Blockchain and Remote Database Access Protocol-based Database," in *Proc. 5th Int. Congr. Inf. Commun. Technol.* in Advances in Intelligent Systems and Computing, vol. 1184, X. S. Yang, Ed. Singapore: Springer, 2020, pp. 533–539.