International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 10 Issue: 12
DOI: https://doi.org/10.17762/ijritcc.v10i12.5890
***Article Received**: 18 October 2022 **Revised**: 12 November 2022 **Accepted**: 18 December 2022*

A State-of-the-Art Review of Time Series Forecasting Using Deep Learning Approaches

Radhika Chandrasekaran , Senthil Kumar Paramasivan
School of Information Technology and Engineering
Vellore Institute of Technology
Vellore, India.
radhika.c2020@vitstudent.ac.in, psenthilkumar@vit.ac.in

Abstract— Time series forecasting has recently emerged as a crucial study area with a wide spectrum of real-world applications. The complexity of data processing originates from the amount of data processed in the digital world. Despite a long history of successful time-series research using classic statistical methodologies, there are some limits in dealing with an enormous amount of data and non-linearity. Deep learning techniques effectually handle the complicated nature of time series data. The effective analysis of deep learning approaches like Artificial Neural Networks (ANN), Convolutional Neural Networks (CNN), Recurrent Neural Networks (RNN), Long short-term memory (LSTM), Gated Recurrent Unit (GRU), Autoencoders, and other techniques like attention mechanism, transfer learning, and dimensionality reduction are discussed with their merits and limitations. The performance evaluation metrics used to validate the model's accuracy are discussed. This paper reviews various time series applications using deep learning approaches with their benefits, challenges, and opportunities.

Keywords- Time series forecasting; Deep Learning; Convolutional Neural Network; Long Short-Term Memory; Gated Recurrent Unit; Recurrent Neural Network.

I. INTRODUCTION

A time series is an observation of an ordered sequence that combines the explicit order dependence between the observations taken sequentially over time. Time series forecasting benefits numerous fields in preplanning by using historical data to forecast future results. The time horizon of predictions is mostly followed in time series like short-term, medium-term, and long-term forecasting and accurate predictions require frequent updation. Time series components are the reason which affects the values of observation including trend, seasonality, cyclic component, and noise. The categorization of time series can be deterministic or non-deterministic, stationary or non-stationary and its classification can be structured or unstructured, univariate or multivariate, single-step or multi-step, regression or classification, linear or non-linear, etc. [1-2].

Due to its ability to achieve greater accuracy when trained on enormous amounts of data with conventional methods, deep learning has recently become an exciting study that increases the researcher's attention to handling time series challenges like nonlinearities and generalization. Most models work well when the data is stationary, when it is non-stationary data, traditional approaches fail. Even though traditional time series forecasting has various advantages, including cleaning and analyzing data, filtering noise from signals in the dataset, and projecting the future with the greatest accuracy it has significant drawbacks, such as incompetent prediction and difficulty handling large datasets. The missing values in the dataset, outliers, difficulty identifying complicated patterns, and insufficient support for long-term forecasting all impair model performance. The deep learning approach is recommended for overcoming the aforementioned challenges by automatically learning features and tasks from unstructured and raw data. Deep learning avoids manual feature extraction and accepts multiple images as input, feature extraction and classification are carried out effectively and it uses GPUs to perform a massive parallel computation that is scalable for massive amounts of data. Deep learning is considered to be a high-end infrastructure to train the data in a reasonable amount of time. It works well with a large volume of image, text, and video data. It is adaptable to new future problems by easily updating new data using backpropagation.

Recent years have seen a dramatic increase in the number of documents about deep learning for time series forecasting. According to the Scopus database, between 2015 and 2022, 41,180 documents related to the search term ["time AND series AND forecasting AND deep AND learning"] were published. Figure 1 illustrates recent publications from the past seven years and demonstrates the exponential growth of time series forecasting research papers. There have been 7385 documents published as of 2022, and that number is still increasing.

Figure 2 below displays the distribution of publications based on documentation type in several fields of deep learning time series forecasting based on the Scopus database. This research examines the advantages, difficulties, and opportunities of deep learning techniques for time series forecasting.

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 10 Issue: 12
DOI: https://doi.org/10.17762/ijritcc.v10i12.5890
***Article Received**: 18 October 2022 **Revised**: 12 November 2022 **Accepted**: 18 December 2022*

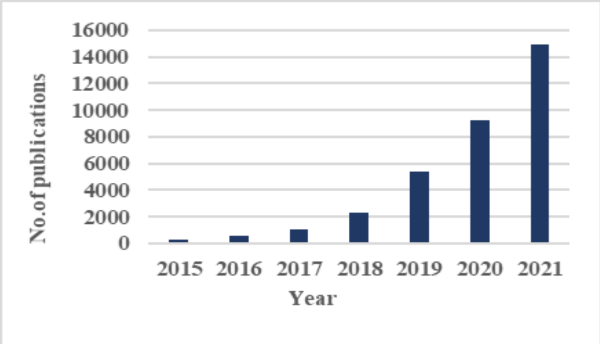

Figure 1. Global Research Trends in Time series Forecasting using Deep learning

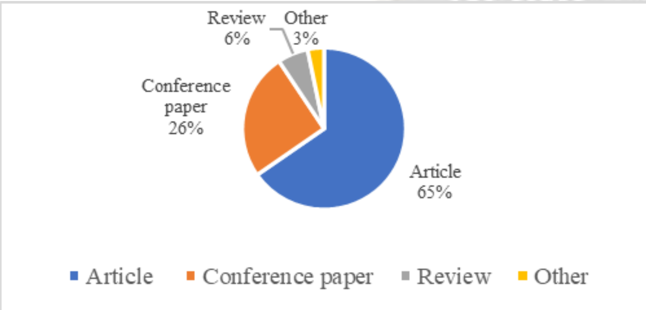

Figure 2. Distribution of publications based on the document type.

Time-series applications can be used in a variety of sectors, as shown in Figure 3, for efficient decision-making and planning. Research articles were gathered for this study from a variety of publishers, including IEEE, Elsevier, Springer, arXiv, MDPI, and others.

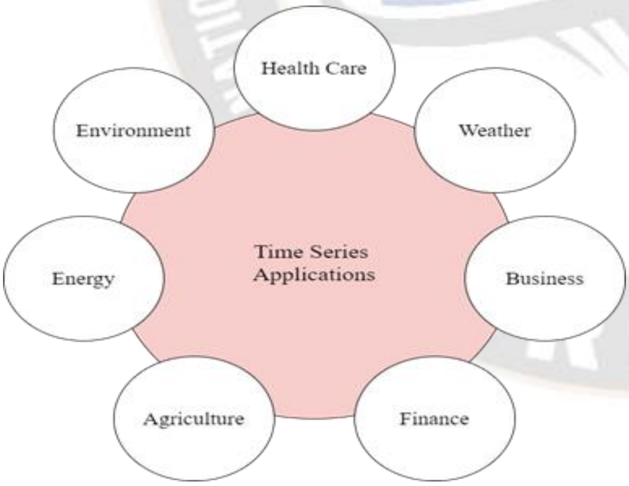

Figure 3. Different Applications of Time series forecasting

Deep learning offers a promising outcome for time series applications like Health care, Finance, Environment, Agriculture, Energy, Weather, Business, and others. Healthcare forecasting can be carried out on different sub-topics like patient care, health insurance, and research development. The prediction helps health care to improve the efficacies of operational management, accurate diagnosis, and treatment in personalized medicine, and cognizance to enhance cohort treatment [1][3-6].

Financial forecasting is the ability to effectively estimate future revenues, cash flow, and costs. It also decides the company's direction and makes lucrative, confident financial investment-related decisions. Financial forecasting may also include sub-topics like stock market prediction, gold price prediction, cryptocurrency price prediction, etc. Financial forecasting has the benefit of reducing financial risk, the future budget makes it easy, and can undergo contingency planning during challenging financial times [2][7-10]. Weather forecasting is effective in predicting future atmospheric properties, daily rainfall variations, and climatic changes. The major benefit of weather forecasting is to minimize the disasters, risks, and losses related to weather. It also improves societal benefits, public health, and safety, and supports economic prosperity and life quality [11-15].

Time series forecasting in agriculture can include sub-topics like monitoring crop and soil, crop yield prediction, price forecast, disease diagnosis, etc. Before harvesting, many agencies might plan decisions depending on crop yield to support farmers and improve production [16-20].

For proper power system planning and to identify the peak demand, electricity forecasting is applied. Forecasting peak demand is difficult due to the constant fluctuation and increases the energy usage. For efficient power system operation and affording power supply continuously for consumers. Developing smart grids can accurately forecast the electric load and it favors the power companies to efficiently schedule load and minimizes unnecessary production of electricity [21-30].

Environmental forecasting may include the sub-topics like air quality forecasting, water quality forecasting, etc. Air quality management depends on the time services data to identify the population's exposure to airborne pollutants and determines compliance of air standards with local ambiance. To manage and prevent pollution in the air, air quality prediction is required [31-35]. Business forecasting and other applications in time series have many benefits with the prediction of the future can be focused, on customer satisfaction, and progress toward goals [36-41].

1.1 *Research contribution*

The applications of deep learning in time series forecasting can be explored and this study provides key answers to the following research questions given below,

(a) What are the different applications of time series forecasting?

(b) What are the various deep learning approaches and techniques used in time series applications?

(c) How the performance of the deep learning models is compared with traditional approaches?

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 10 Issue: 12
DOI: https://doi.org/10.17762/ijritcc.v10i12.5890
***Article Received**: 18 October 2022 **Revised**: 12 November 2022 **Accepted**: 18 December 2022*

This study categorizes various deep learning algorithms for time series applications. Deep learning approaches in time series models are discussed in Section II. The attention mechanism and transfer learning are discussed in Sections III and IV. The necessity of dimensionality reduction for dealing with high dimensionality data is summarized in Section V and followed by time series applications that integrate domain-specific models with single and hybrid deep learning models are discussed in section VI. Section VII explores performance criteria for evaluating the predicted model. Finally, in section VIII, the conclusion and recommendations are provided.

II. DEEP LEARNING APPROACHES IN TIME SERIES FORECASTING

Artificial intelligence (AI) is an extensive area concerned with the construction of a smart machine that is capable of performing tasks with human intelligence and plays a vital role in multi-disciplinary areas like speech and image recognition, robotics, autopilot technology, stock trading, etc. Machine learning (ML) is a branch of artificial intelligence that enables computers to recognize patterns in data automatically, make decisions without explicit programming, and provide reliable results. Due to the increase in the availability of data, the complexity of handling data is enlarged. Machine learning has certain limitations it requires large stores of training data.

Training data labeling is a difficult task, time consumption is more and it requires high resources to process the data.ML performance degrades when the number of data increases. Deep learning, a subtype of machine learning that can handle massive volumes of data and produces highly accurate results, is used to get over the constraints of machine learning.

A deep neural network with numerous layers that is inspired by the functionality of the human brain is an artificial neural network [21]. It enables the intricate nonlinear relationship between the predictors and the response variable [42]. Information processing units in ANN are interconnected to identify the patterns known as neurons. The neural network has a basic processing unit that takes the input signal and the outcome is passed to other connected neurons, known as perceptron [10].

The ANN can be expressed as,

$$y = \sum_{k=1}^{n} w_k x_k + b \qquad (1)$$

Where ‘y’ represents output, ‘x’ and ‘w’ represent input and weight, and ‘b’ biases [43].

ANN has the benefit of accurate mapping between input and output, it doesn't require scaled and stationary data as input. Handling missing data, storing information, parallel processing and fault tolerance, etc. are some other advantages. Hardware dependence, unexplained network behavior, overfitting, data duration, and high computational cost are the challenges faced in ANN. Some of the applications of ANN are bitcoin price forecasting [10], financial prediction [16], and monthly electricity demand forecasting [21]. Multilayer perceptron (MLP) is a multilayer ANN and it is a fully connected feed-forward neural network (FNN) with no feedback. During the training, the weight and bias are estimated to network with a supervised algorithm [9][22-23]. In a neural network, it contains one or more hidden layers. As depicted in Figure 4, MLP is a feed-forward network with three layers, comprising input, hidden, and output layers. All other nodes except input use a non-linear activation function. The count and size of the layer decide the depth and complexity of the network.

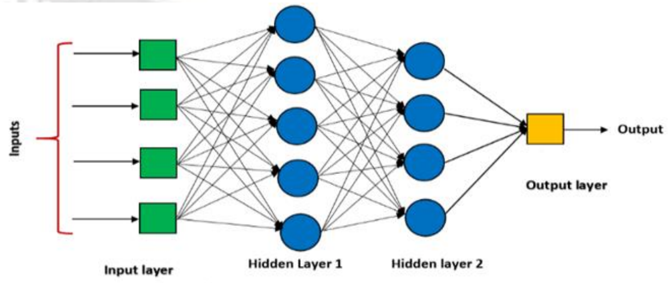

Figure 4. The Block diagram of Multilayer perceptron

When the amount of data increases the performance of deep learning simultaneously increases. Due to the availability of high-performance GPUs and massive of data, deep learning approaches provide better accuracy in various domains. Deep learning models can handle highly complex data with expensive resources. The drawback of the deep learning model is it requires enormous data to perform better than other models and requires expensive resources like GPUs to train complex data that may increase the computation cost.

2.1 Convolutional Neural Network (CNN)

A feed-forward network also known as ConvNet is a CNN that identifies, extracts high-level features, and classifies the image objects of input. Due to its excellent learning and extraction of features from raw input automatically this approach is preferred in time series forecasting problems. It can capture any application's dependencies of a temporal and spatial image through a relevant filter. To extract information from an image pixel, a ConvNet design comprises multiple hidden layers.

The architecture consists of a convolutional layer, an activation function like ReLU, a pooling layer, and a fully connected layer. The procedure begins by filtering and retrieving relevant information from the input images before performing convolutional operations in the convolution layer. In comparison to previous activation functions like sigmoid and sigmoid and tanh, ReLU increases the network's nonlinearity.

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 10 Issue: 12
DOI: https://doi.org/10.17762/ijritcc.v10i12.5890
***Article Received**: 18 October 2022 **Revised**: 12 November 2022 **Accepted**: 18 December 2022*

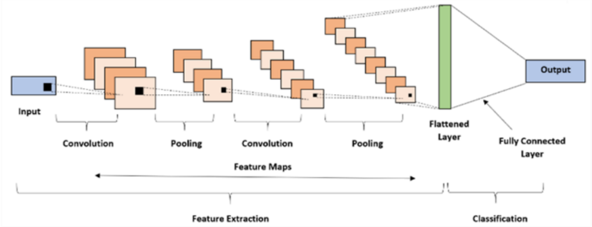

Figure 5. The Block Diagram of Convolutional Neural Network

ReLU also avoids vanishing gradients and has higher computational efficiency.

Pooling is the next step in the feature mapping dimensionality reduction process, which reduces the spatial size and lowers the computing cost. Pooling can be divided into two types: average and maximum. Max pooling returns the greatest value of the image, whereas Average pooling returns the average values of the image section covered by the kernel [24]. Finally, as shown in Figure 5, the fully connected layers connect all of the layers in the network. Compared to other classification methods, CNN requires less pre-processing and is scalable to large datasets which are considered to be its advantage. Due to the max pool, the need for a large dataset to compute and train the model, and the fact that CNN is spatially invariant to input data, CNN operates slowly and can be applied for energy consumption forecasting [31] and power demand forecasting [25].

2.2 Recurrent Neural Network (RNN)

RNN is a recursive method for working with sequential data, including text and movies, which are collections of images [44]. RNN loses control of the next context fragment to be sent. The network's backtracking training process naturally experiences the exploding and vanishing gradients problem, which makes it difficult to manage how much past data is to be forgotten and how much context should be maintained in memory with specifics when it can be discarded. The error is backtracked in each phase of the backtracking procedure from the most recent timestamp to the first. At each timestamp, the error that occurred is calculated and this will update the weights that cause difficulty in training and slow computation. As shown in Figure 6, ‘x’ and ‘y’ represent input and output, ‘h’ is the hidden layer, ‘t’ and ‘w’ are timestamps and weight.

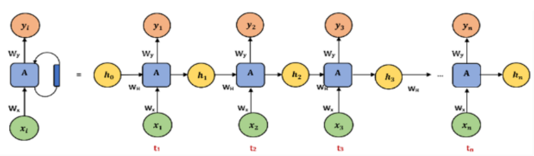

Figure 6 The block diagram of Recurrent Neural Network

The advantage of RNN is its capacity of handling complex data and missing values do not affect its performance. Due to their inability to manage long-term dependencies, RNNs cannot store information for an extended period. The major disadvantage of RNN is it suffers from a problem with exploding and vanishing gradients and the computational cost is high. Electric load forecasting [23], citation count prediction [36], wind energy forecasting [45], and air quality forecasting [32] are some of the applications of RNN.

2.3 Long Short-Term Memory (LSTM)

The long-term dependencies of the time series are learned using the LSTM, which is a successful sequential learning model. Input gate “i” output gate "o" and forget gate "f" are the three gates that make up the cell structure (c), as shown in Figure 7.

When the information enters the cell structure with specific rules it determines whether the entered information is valuable or not valuable. LSTM is used to overcome the shortcomings of RNN and handle the temporal situation of data. The LSTM network can extract and retain data in memory for a long period, and it can handle both long-term and short-term dependencies [2][22] [24][26][32][44].

LSTM can deal with the problem of exploding or vanishing gradients. But it is susceptible to overfitting and difficult to implement dropout. In LSTM, training becomes difficult if it demands a lot of computation. Stock market prediction [7], Weather prediction [11], and Monthly electricity demand forecasting [21] are some of the areas using LSTM. The variants of LSTM include Bidirectional LSTM (BiLSTM), Convolutional LSTM (ConvLSTM), and Stacked LSTM. A bidirectional LSTM uses two LSTMs, one of which receives input in the forward direction and the other which is used in a backward way.

The model is designed for signal processing where the signal moves across time both forward and backward. BiLSTM is employed to categorize text and predict diseases [37]. Before the LSTM process, the convolution layer extracts the flowing information in convolutional LSTM. One-dimensional LSTM input data is suitable for satellite, radar image, and video data. ConvLSTM is capable of handling 3-D data and is employed in applications such as intelligent financial forecasting [2]. Stacked LSTM consists of multiple LSTM layers and allows greater model complexity. Stacked LSTM is used in disease forecasting [4].

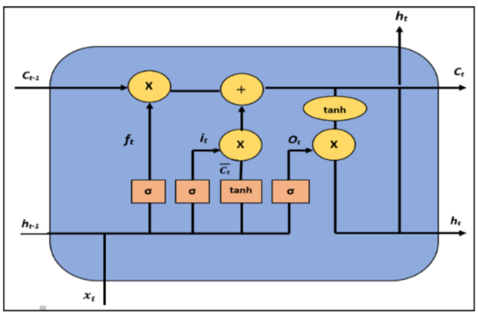

Figure 7. The block diagram of the Long Short-Term Memory

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 10 Issue: 12
DOI: https://doi.org/10.17762/ijritcc.v10i12.5890
***Article Received**: 18 October 2022 **Revised**: 12 November 2022 **Accepted**: 18 December 2022*

2.4 Gated Recurrent Unit (GRU)

The update gate (z_t) and reset gate (r_t) of the GRU as illustrated in Figure 8, which is an upgraded form of the LSTM. Due to lighter processing, it outperforms LSTM by sequentially learning both short- and long-term dependencies. The reset gate (r_t) controls the degree to which the preceding instant's data information was concealed, while the update gate (z_t) controls the status of the previous instant in the hidden state (h_t).

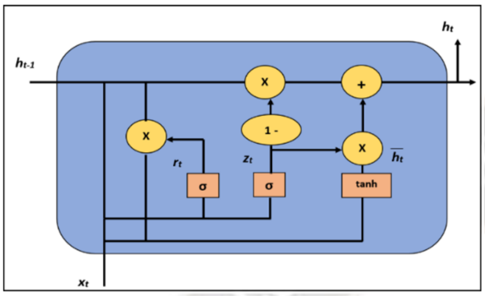

Figure 8. The block diagram of the Gated Recurrent Unit

The GRU can be expressed as,

$$z_t = \sigma(w_z = [h_{t-1}, x_t]) \quad (2)$$

$$r_t = \sigma(w_r = [h_{t-1}, x_t]) \quad (3)$$

$$\tilde{h}_t = tanh(W[r_t * h_{t-1}, x_t]) \quad (4)$$

$$h_t = (1 - z_t) * h_{t-1} + z_t * \tilde{h}_t \quad (5)$$

GRU uses less memory and it is quicker than LSTM. The disadvantages of GRU include slow convergence and low learning efficiency. Some of the applications of GRU include Intelligent financial forecasting, Electric load forecasting, and Petroleum production forecasting.

2.5 Deep Belief Network (DBN)

A generative model called DBN is based on specific probabilistic properties. Deep belief networks are recommended to get over the drawbacks of ordinary neural networks, such as their slow learning rate, poor parameter selection, and training procedure. On top of a deep belief, networks are numerous layers of RBM modules [46]. The network consists of two layers hidden and a visible layer. Some of the application of DBN includes wind power forecasting [27] and Ultra short-term industrial power demand forecasting [28].

2.6 Variational Autoencoders (VAE)

Variational Autoencoders is an effective deep learning model to handle complex data in terms of unsupervised learning. Dimensionality reduction is an attractive feature of VAE to compress the high dimensional data for flexibility and solves the overfitting problem in a better way and improves data sampling during training. VAE model composes of an encoder and decoder-like autoencoder for encoding and decoding via unsupervised learning. Some of the applications of VAE are face recognition, handwritten digits, and hospital resource management [3][6].

2.7 Generative Adversarial Networks (GAN)

The generative adversarial network is advanced generative modeling using deep neural network methods such as CNN. In general, generative modeling is an unsupervised machine learning task that finds and learns the regularities. The advantages of GAN are it normally generalizes and provides brief ideas with limited data. The generative model (G), which depicts the data distribution, is one of the two models in GAN. To estimate the training data probability sample from the generator, the discriminative model (D) is used (G) [42][47].

The cost objective function of GAN is,

$$\min_G \max_D E_{x\sim p_d}[\log D(x)] + E_{x}^{1}{\sim}p_d[\log(1 - D(\hat{x}))] \quad (6)$$

The input of G is '$\boldsymbol{z}$', which is usually sampled from a uniform or Gaussian distribution. where p_d - real data distribution, p_g - model distribution defined by,

$$\hat{x} = G(z) \quad (7)$$

'D' receives either generated sample $\hat{x}$ or real data '$\boldsymbol{x}$', and 'D' is trained to distinguish them by maximizing the cost function. While 'G' is trained to generate more and more realistic samples by minimizing the cost function. The training stops when 'D' and 'G' achieve the Nash equilibrium, where none of them can be further improved through training. This network can be used in financial forecasting applications [48].

2.8 Sequence to Sequence Model (seq2seq)

The encoder-decoder design of the seq2seq framework is based on language processing and machine translation. The mapping of the input to output sequence includes tags and attention values. The sequence includes items like words, letters and time series, etc. The input sequence was read by the first network. The second network will learn to produce an output sequence [23][28]. The seq2seq architecture solves the RNN problem to generate the output sequence of arbitrary length and the model was first used in machine translation and image captioning. The drawbacks of the model are the sequence of output relies more on the context in the final output of the encoder defined by the hidden layer. In case when the sequence is too long there is a high possibility of losing the initial context at the end stage. Attention mechanisms can overcome the problem by focusing on every stage of the sequence. This model is used in applications like citation count prediction [36].

III. ATTENTION MECHANISM

The attention mechanism is the extension of an encoder-decoder based neural machine translation system initially in natural language processing and other applications like speech recognition and computer vision. The encoder and decoder are considered as a stack of RNN and LSTM modules working in two steps. Initially, encoder LSTM process the input sequence entirely to encode into a context vector, and decoder units produce words in sequential order one after the other.

Transformer architecture solves the sequence transduction problems and the main task is transforming any input sequence into an output sequence with encoding, machine translation, and decoding stages. To detect similar sequence and image representations transformer architecture with an attention mechanism is applied this helps the model to learn faster and generalize better features [49].

IV. TRANSFER LEARNING

Transfer learning is the method of training the source dataset of a base network and transferring the features learned with weights to the target dataset of the second network. It reuses the pre-trained model and improves the generalization of a new problem suffering from the limited dataset. It effectively trains the limited amount of dataset to overcome time consumption and cost. In computer vision tasks like object localization and image identification, this model uses deep neural network generalization abilities [50-52].

V. DIMENSIONALITY REDUCTION IN TIME SERIES FORECASTING

Dimensionality reduction techniques are mostly used for the visualization of data when the features are correlated and redundant. Dimensionality reduction is divided into two components including feature selection and feature extraction [26][53-54].

Feature extraction transforms high dimensional space into low dimension space and is common in the field of speech recognition and bioinformatics etc. The filter, wrapper, and embedded-based techniques of feature selection are further separated into finding a subset of the original collection of variables [2]. Missing value, low variance, strong correlation, back feature deletion, principal component analysis (PCA), linear discriminant analysis (LDA), and generalized discriminant analysis are a few techniques for dimensionality reduction (GDA). In some applications, such as wind forecasting. High correlation can be removed among the values using PCA [7][46].

The feature-based method composes of two stages they are feature extraction and clustering. Feature extraction transforms the input into low-dimensional feature vectors first and using clustering algorithm features are extracted. Some of the advantages of dimensionality reduction include reducing computation time, removing redundant features, data compression, multicollinearity, and removing noise. The ranker-based feature selection can be applied in short-term load forecasting [55][57-58] and stock market prediction [56].

TABLE 1 SUMMARY OF TIME SERIES FORECASTING IN HEALTH CARE

Authors	Forecasting Applications	Methodology	Performance Evaluation Metrics	Result & Discussions
Ayoobi N.et.al. [1]	New cases and death rates for COVID-19 prediction.	LSTM ConvLSTM GRU Bi-LSTM Bi-GRU Bi-Conv-LSTM	Mean squared log error (MLSE) MAPE Root means squared log error (RMSLE) EV	Bi-directional approaches outperform compared with other models.
Harrou et al. [3]	Hospital resource Management	Variational autoencoder(VAE)	R^2 RMSE MAE Explained variance (EV)	Efficient hospital resource management. The suggested method solves the overfitting, removes outliers, and improves data quality compared to CNN, LSTM, ConvLSTM, BiLSTM, RNN, GRU, and RBM
Shastri et al. [4]	Covid -19 prediction	Bi- LSTM Conv LSTM Stacked LSTM	MAPE	The result proves that Convolutional LSTM outperforms and produces high accuracy in disease forecasting.
Bharati et.al. [5]	Lung disease prediction	Visual geometry group (VGG) + data augmentation + spatial transformer network (STN) + CNN - (VDSNet)	Recall Precision Fβ score	Lung disease detection becomes simple.
Li & Liu [6]	Stress detection	1D CNN + MLP	Accuracy F1 Score Precision Recall	The findings show that the two-deep neural networks greatly outperformed the conventional machine learning methods for both tasks.

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 10 Issue: 12
DOI: https://doi.org/10.17762/ijritcc.v10i12.5890
***Article Received**: 18 October 2022 **Revised**: 12 November 2022 **Accepted**: 18 December 2022*

VI. TIME SERIES FORECASTING APPLICATIONS USING DEEP LEARNING APPROACHES

The study was based on various single and hybrid approaches in the deep learning model. Hybridization is combining two or more deep learning, machine learning, and optimization approaches to progress the predictive model with better performance. Time series forecasting with deep learning approaches can be applied to different applications for attaining accuracy in prediction which helps fine decision-making and planning in prior. Out of many applications, a few applications like health care, finance, environment, weather, energy, agriculture, and business applications survey is discussed below with their methodology, performance metrics, and result.

6.1 Health care

Due to the increase of data, complexity in health care has augmented. In the medical industry, various forms of data are generated, such as electronic health records, text, sensor data, images, etc. Due to their unstructured, poorly annotated, and heterogeneous form, data handling becomes complex [59]. The advancement of current deep learning technologies has enabled the healthcare industry to diagnose the disease at an early stage to especially support elderly people [60]. Effective deep learning paradigms can learn from complex data from beginning to end, and with their computing capabilities, they can enable quick, precise, and efficient operations in the healthcare industry.

Some of the challenges of using deep learning in health care are the limitation of data in case of rare diseases can minimize the accuracy. Before handling the raw data, it needs to be pre-processed. Feature enrichment, temporal modeling, and model privacy can be focused on for future research in health care. The summary of Health care in time series forecasting is discussed in Table 1.

6.2 Finance

To increase profitability, firms utilize a technique known as financial forecasting to project future sales, costs, and cash flow. Financial forecasting, which has a variety of subdomains including stock price prediction and commodity price prediction, primarily forecasts prices (oil, gold, etc.) forex price prediction, cryptocurrency price predictions, etc as given in Table 2. The deep learning approach is required in financial forecasting to overcome certain drawbacks like non-stationary data, leakage of data, Overfitting, and data unavailability.

TABLE 2 SUMMARY OF TIME SERIES FORECASTING IN FINANCE

Authors	Forecasting Applications	Methodology	Performance Evaluation Metrics	Result & Discussions
Niu et al. [2]	Intelligent Finance Forecasting.	RReliefF algorithm. Binary grey wolf optimizer with cuckoo search ELM. LSTM. ConvLSTM. GRU. Error Correction model.	MAE MAPE RMSE Median of absolute percentage error (MdAPE) Theil U statistics (U) Pearson correlation coefficient (R) Index of argument (IA)	Improves the forecasting framework performance and accuracy of multivariate forecasting significantly.
Shen and Shafiq. [7]	Stock Market	FE RFE PCA LSTM	MSE MAE F1 Score TPR (recall) TNR (specificity) FPR (fall-out) FNR (miss rate) Binary accuracy	Better accuracy is obtained with this model.
Chen et.al. [8]	Stock price forecasting	BiLSTM + MLP Attention Mechanism PCA	MSE EVS MAE MSLE MedAE R^2	The proposed model provides better performance.
Hrasko et al. [9]	Energy prediction Dollar prediction Normalized pressure prediction	DBN based Gaussian-Bernoulli Restricted Boltzmann Machine (GBRBM) + FNN and Backpropagation (BP).	RMSE	GBRBM + FNN performance was promising and outperforms BP.

Mudassir et al. [10]	Bitcoin price Forecasting	ANN SANN SVM LSTM	MAE RMSE MAPE Accuracy F1- score AUC	Detects fraudulent activities and anomalous behavior..

6.3 Weather Forecasting

Meteorologists deploy computer models to forecast the weather and the processing power has increased dramatically. Despite this, meteorologists still struggle to forecast weather for a few days. Certain considerations, including the amount of data available, the length of time available to review it, and the complexity of weather events, limit their capacity to forecast the weather. Deep learning algorithms play a vital role in forecasting weather conditions accurately as illustrated in Table 3.

TABLE 3 SUMMARY OF TIME SERIES FORECASTING IN WEATHER

Authors	**Forecasting Applications**	**Methodology**	**Performance Evaluation Metrics**	**Result & Discussions**
Hewage et al. [11]	Weather Forecasting	LSTM Temporal Convolution network (TCN).	MSE	Overcome vanishing gradient. High memory capacity is used.
Khan & Maity [12]	Rainfall prediction	Conv1D + MLP GCM simulation	Co efficient of correlation (r) RMSE Nash–Sutcliffe Efficiency (NSE)	Performs better than other models when predicting rainfall
Bajpai.et.al [13]	Rainfall Prediction	A deep and wide rainfall prediction model (DWRPM)	RMSE MAE MSE	The results are promising and the model has generalization ability. The same model works well for forecasting rainfall in different atmospheric zones of Rajasthan.
Sun.et.al [14]	Short-term rainfall Prediction	Conv- 3D-GRU	Critical success index (CSI) Heidke skill score (HSS) MSE MAE B-MSE, B-MAE	Improves short-term rainfall Prediction accuracy.
Dabhade.et.al [15]	Extreme weather detection (Cyclone)	Convolutional long-short term memory (ConvLSTM)	peak signal-to-noise ratio (PSNR), structural index similarity (SSIM)	The suggested concept can significantly enhance cyclone detection and prediction.

6.4 Agriculture

Deep learning algorithms in the field of agriculture can assess and predict the effects of weather, seasonal sunlight, animal, bird, and insect migration patterns, crop pesticides, use of specialty fertilizers, planting cycles, and irrigation cycles on yield [61-62] as shown in Table 4.

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 10 Issue: 12
DOI: https://doi.org/10.17762/ijritcc.v10i12.5890
***Article Received**: 18 October 2022 **Revised**: 12 November 2022 **Accepted**: 18 December 2022*

TABLE 4 SUMMARY OF TIME SERIES FORECASTING IN AGRICULTURE

Authors	Forecasting Applications	Methodology	Performance Evaluation Metrics	Result & Discussions
Li et al. [16]	Agriculture market future price forecasting.	Differential privacy multimodal – VAE long and short-term prediction model (DP-MAELS)	MAPE Root relative squared error (RSE) Relative absolute error (RAE) Empirical correlation coefficient (R) Theil U	DP-MAELS outperforms other approaches with improved robustness and accuracy of prediction as an alternative to multivariate forecasting.
Sharma et.al [17]	Wheat crop yield prediction	CNN – LSTM-	RMSE	The model outperforms and incorporates additional information that will improve yield estimates.
Sagan et.al [18]	Field-scale crop yield prediction.	2D and 3D CNN	RMSE	Reliable yield predictions and improves prediction accuracy.
Jin.et.al [19]	Smart Agriculture Sensing	EMD CNN GRU IoT	RMSE	The suggested predictor can produce precise forecasts for the next 24 hours, giving producers access to enough climate data.
Khaki et.al [20]	Crop yield prediction	DNN	RMSE	The outcomes show that environmental factors affected crop production more so than genetics.

TABLE 5 SUMMARY OF TIME SERIES FORECASTING IN ENVIRONMENT

Authors	Forecasting Applications	Methodology	Performance Evaluation Metrics	Result & Discussions
Xie et al. [31]	PM2.5 pollution prediction. Energy consumption forecast. Human activity recognition (HAR).	Enhanced grey wolf optimizer (GWO). CNN-LSTM.	RMSE MAE	Gains high accuracy and avoids overfitting and underfitting with the efficient network.
Freeman et.al [32]	Air Quality Forecasting	RNN with LSTM	MAE RMSE	Reduces errors in both training and test data with limited and optimized features.
Bata et.al. [33]	Short-term water demand forecasting	Regression tree Self-organizing map	MAPE NRMSE	The fused model complexity is analyzed and the hybrid approach provides promising outcomes.
Chu et.al [34]	Water disposal Prediction	CNN + MLP	Accuracy Precision Recall	The model outperforms and achieves better accuracy in the classification of waste disposed of.
Du et.al. [35]	Air Quality Forecasting	1D-CNN + BiLSTM	RMSE MAE	The model handles PM2.5 air pollution with good accuracy.

6.5 Environment

Accurate forecasting should be able to identify impending environmental changes. Environmental forecasting may include subtopics like air quality prediction for identifying air quality levels, water quality prediction, climatic changes, water demand forecasting, etc. [63]. To identify future threats and opportunities and it helps to understand the transformation of the industries using deep learning techniques as given in Table 5.

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 10 Issue: 12
DOI: https://doi.org/10.17762/ijritcc.v10i12.5890
***Article Received**: 18 October 2022 **Revised**: 12 November 2022 **Accepted**: 18 December 2022*

6.6 Energy

Energy forecasting has certain challenges in predicting the future electricity requirement based on demand. To overcome uncertainty and blackout in forecasting, certain factors require essential focus. Forecasting the electricity load is a complex task, it depends on capital investment. Moreover, demand forecasting relies open to the effects of many human actions as well. Uncertainty is mainly caused due to the influence of changes in public perception, policies and weather data, etc as shown in Table 6.

TABLE 6 SUMMARY OF TIME SERIES FORECASTING IN ENERGY

Authors	**Forecasting Applications**	**Methodology**	**Performance Evaluation Metrics**	**Result & Discussions**
Son and Kim [21]	Monthly Electricity demand forecasting	LSTM	MAE RMSE MAPE MBE Unpaired peak accuracy (UPA)	Making accurate predictions for efficient power system planning using LSTM with weather and social variables.
Mehdipour Pirbazari et al. [22]	Short term load forecasting	SVR Gradient boosted regression tree (GBRT) FFNN LSTM	MAE RMSE MASE Daily Peak Mean Average Percentage Error (DpMAPE) Cumulative weighted error (CWE)	Appliances should be scheduled according to the price of electricity during peak hours to save energy and improve daily peak load performance.
Gasparin et.al [23]	Electric load forecasting	Elmann RNNs Seq to Seq model TCN LSTM GRU DFNN	RMSE MAE NRMSE R^2	Accurate load forecasting with nonlinearity provides efficient operation and management of smart grids with critical infrastructure.
Shao et al. [24]	Power Consumption forecasting.	CNN – LSTM with Discrete wavelet transform (DWT). Domain fusion (DF – CNNLSTM).	RMSE MAE MAPE	With unseen data, future predictions can be accurate. Random initialization is not considered to be sensitive.
Kim et al. [25]	Hybrid Power demand Forecasting	LSTM + CNN	MAPE RRMSE	LSTM + CNN outperforms with better accuracy.
Bouktif.et.al. [26]	Electric load forecasting	Genetic algorithm (GA) enhanced LSTM - RNN	MAE RMSE Coefficient of variation (CVRMSE)	Performance was good with high confidence.
Hu et al. [27]	Wind power Forecasting	Convolution based spatial - temporal wind power predictor (CSTWPP)	MAE RMSE MASE	Significantly outperforms univariate methods. GPU makes the training of CSTWPP fast.
Tan et al. [28]	Ultra-short-term industrial power demand forecasting	LSTM BRSB (Bagging, Random Subspace, and Boosting) ensemble algorithm.	MAE MAPE Normalized Root Mean Square Error (NRMSE) Peak Absolute Percentage Error (PAPE)	Improves loss function, result obtains low forecasting error, and the accuracy of the model is improved.
Yan et.al. [29]	Short term load forecasting – Individual households	CNN + LSTM	RMSE MAE MAPE	When compared to other models, the suggested model performs better.
Liu et.al. [30]	Wind power prediction	Wavelet decomposition + LSTM	MAPE	When compared to other models, the suggested model performs better.
Yan et.al. [64]	Energy consumption of individual households	Ensemble LSTM with SWT	RMSE MAPE MBE	The result produced outperforms other compared models.

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 10 Issue: 12
DOI: https://doi.org/10.17762/ijritcc.v10i12.5890
***Article Received**: 18 October 2022 **Revised**: 12 November 2022 **Accepted**: 18 December 2022*

Accurate forecasting is difficult and it is risky to bond the plans depending on load forecasting. By integrating uncertainty into analysis techniques, the prominence of planning will produce an accurate forecast and construct a model that adapts to the changes easily.

Statistical model suffers in handling raw data due to non-linearity. Deep learning algorithms provide better performance compared with other algorithms based on previous studies.

6.7 Business and other applications

Nowadays, in the business world, frequent changes occur and make managers find a new paths to survive in the competitive business environment. Businessmen have to figure out new ways to generate profit and expansion prospects.

The new laws and regulations put a further burden on businesses. Applications for business forecasting can be analyzed using deep learning techniques as shown in Table 7.

TABLE 7 SUMMARY OF TIME SERIES IN BUSINESS AND OTHER APPLICATIONS

Authors	**Forecasting Applications**	**Methodology**	**Performance Evaluation Metrics**	**Result & Discussions**
Abrishami and Aliakbary [36]	Citation Count Prediction	Simple RNN Seq2Seq model	RMSE R^2	Citation count prediction outperforms with high accuracy and a highly cited paper is predicted.
Sharfuddin et.al. [37]	Text classification	Deep RNN with BiLSTM	Confusion Matrix	This approach is used for sentiment classification and an accuracy of 85.67% is achieved using this hybrid approach.
Suhermi et.al. [38]	Roll motion prediction	ANN ARIMA	Root mean squared error prediction (RMSEP)	The suggested model results in a better result.
Ali Shah.et.al. [39]	Workplace Absenteeism Prediction	DNN Synthetic Minority Oversampling Technique (SMOTE)	Accuracy TPR, FPR, F1-Score, Recursive operating characteristic curve (ROC)	Higher accuracy for real-world problems with large datasets.
Kaneko .et.al [40]	Retail store sales prediction	Logistic Regression	Accuracy Precision Recall F-measure AUC	The current study reveals that deep learning is useful for examining the point-of-sale (POS) data of retail establishments.
Vallés-Pérez et.al. [41]	Sales Forecasting	Sequence to sequence RNN Transformers	Root Mean Squared Logarithmic Error (RMSLE) Root Mean Squared Weighted Logarithmic Error (RMSWLE) Mean Absolute Logarithmic Error (MALE)	The seq2seq trimmed model produced the best results for the least amount of processing effort.
Sagheer and Kotb [44]	Petroleum Production forecasting	Deep long-short term memory recurrent network (DLSTM).	RMSE Root means square percentage error (RMSPE) MAPE	Performs better than deep RNN and deep GRU models.

VII. PERFORMANCE EVALUATION METRICS

The performance of the models in time series forecasting is assessed by the assessment measures. The measurements offer feedback on forecasting accuracy that aids in model improvement for increased accuracy. Below, Table 8 lists the most common typical evaluation metrics for time series forecasting.

Table 8 Performance Evaluation Metrics

Performance Metrics	**Equations**
Mean square error (MSE)	$\frac{1}{n}\sum_{i=1}^{n}\left(y_{act} - \hat{y}_{pred}\right)^2$
Root Mean square error (RMSE)	$\sqrt{\frac{1}{N_{samples}}\sum_{i=1}^{N_{samples}}(y_{act} - \hat{y}_{pre})^2}$
Mean absolute error (MAE)	$\frac{1}{n}\sum_{i=1}^{n}\lvert y_{pred} - y_{act}\rvert$
Mean absolute percentage error (MAPE)	$\frac{1}{n}\sum_{i=1}^{n}\left\lvert\frac{y_{act} - \hat{y}_{pred}}{y_{act}}\right\rvert$

ISSN: 2321-8169 Volume: 10 Issue: 12
DOI: https://doi.org/10.17762/ijritcc.v10i12.5890
***Article Received**: 18 October 2022 **Revised**: 12 November 2022 **Accepted**: 18 December 2022*

Coefficient of determination (or) R-Squared (R^2)	$\frac{\sum_{i=1}^{n}[(y_i-\bar{y})\cdot(\hat{y}_i-\bar{y})]^2}{\sqrt{\sum_{i=1}^{n}(y_i-\bar{y})^2}\cdot\sqrt{\sum_{i=1}^{n}(\hat{y}_i-\bar{y})^2}}$

Where y_i, y_{act} = actual value, $\hat{y}_i$, y_{pred} = predicted value and $\bar{y}$ = average of y_i values.

VIII. CONCLUSION AND RECOMMENDATIONS

Time series forecasting has large attention in various research domains. The study provides a detailed analysis of certain applications using deep learning approaches to improve efficiency and preciseness. Different applications including health care, financial forecasting, environmental prediction, weather forecasting, energy forecasting, agriculture, business, and others are analysed with their methodology, performance evaluation metrics, and the result. The deep learning approaches and techniques utilized in many applications are discussed, and time series forecasting frequently employs analysis techniques including CNN, LSTM, ANN, GRU, and RNN. The transfer learning can be applied to an application with limited data to avoid overfitting during training. For precise prediction, hybrid models are preferred which is a combination of various machine learning, deep learning, and statistical algorithms. For assessing the model's error, the evaluation metrics RMSE, MAE, MAPE, MSE, and R^2 are most frequently employed in time series forecasting. In addition to these models, confusion matrices are employed to assess the model's accuracy and precision.

The ability to handle difficult time series problems in numerous industrial applications makes deep learning approaches the most significant subfield of machine learning. According to the examination, the majority of the data produced are from different time series. Nowadays, handling complicated data is becoming a difficult issue in many study disciplines because of the unavoidable enormous availability of data and its nonlinearity. Deep learning techniques can perform well and are scalable for enormous data with computational resources like GPUs. Unlike machine learning methods that use parallel processing for large amounts of data, deep learning approaches use less time.

The overview of different deep learning strategies, as well as techniques like dimensionality reduction to prevent multi-collinearity and to cut down on processing time and storage when working with high dimensional data, are discussed. The time-series data can be adjusted using attention mechanisms and transfer learning when the data is limited. The accuracy of the model can be improved through performance evaluation metrics. In comparison to machine learning techniques, the deep learning techniques used in this review's hybrid approaches for a variety of applications perform better at predicting, planning, and scheduling complex time series data. The deep learning algorithms can be integrated with big data and IoT in the future to explore better results in various time series applications.

REFERENCES

[1] Ayoobi, N., Sharifrazi, D., Alizadehsani, R., Shoeibi, A., Gorriz, J. M., Moosaei, H., ... & Mosavi, A. (2021). Time series forecasting of new cases and new deaths rate for COVID-19 using deep learning methods. Results in Physics, 27, 104495.

[2] Niu, T., Wang, J., Lu, H., Yang, W., & Du, P. (2020). Developing a deep learning framework with two-stage feature selection for multivariate financial time series forecasting. Expert Systems with Applications, 148, 113237.

[3] Harrou, F., Dairi, A., Kadri, F., & Sun, Y. (2020). Forecasting emergency department overcrowding: A deep learning framework. Chaos, Solitons & Fractals, 139, 110247.

[4] Shastri, S., Singh, K., Kumar, S., Kour, P., & Mansotra, V. (2020). Time series forecasting of Covid-19 using deep learning models: India-USA comparative case study. Chaos, Solitons & Fractals, 140, 110227.

[5] Bharati, S., Podder, P., & Mondal, M. R. H. (2020). Hybrid deep learning for detecting lung diseases from X-ray images. Informatics in Medicine Unlocked, 20, 100391.

[6] Li, R., & Liu, Z. (2020). Stress detection using deep neural networks. BMC Medical Informatics and Decision Making, 20(11), 1-10.

[7] Shen, J., & Shafiq, M. O. (2020). Short-term stock market price trend prediction using a comprehensive deep learning system. Journal of big Data, 7(1), 1-33.

[8] Chen, Q., Zhang, W., & Lou, Y. (2020). Forecasting stock prices using a hybrid deep learning model integrating attention mechanism, multi-layer perceptron, and bidirectional long-short term memory neural network. IEEE Access, 8, 117365-117376.

[9] Hrasko, R., Pacheco, A. G., & Krohling, R. A. (2015). Time series prediction using restricted boltzmann machines and backpropagation. Procedia Computer Science, 55, 990-999.

[10] Mudassir, M., Bennbaia, S., Unal, D., & Hammoudeh, M. (2020). Time-series forecasting of Bitcoin prices using high-dimensional features: a machine learning approach. Neural computing and applications, 1-15.

[11] Hewage, P., Trovati, M., Pereira, E., & Behera, A. (2021). Deep learning-based effective fine-grained weather forecasting model. Pattern Analysis and Applications, 24(1), 343-366.

[12] Khan, M. I., & Maity, R. (2020). Hybrid deep learning approach for multi-step-ahead daily rainfall prediction using GCM simulations. IEEE Access, 8, 52774-52784.

[13] Bajpai, V., Bansal, A., Verma, K., & Agarwal, S. (2020). Prediction of Rainfall in Rajasthan, India using Deep and Wide Neural Network. arXiv preprint arXiv:2010.11787.

[14] Sun, D., Wu, J., Huang, H., Wang, R., Liang, F., & Xinhua, H. (2021). Prediction of short-time rainfall based on deep learning. Mathematical Problems in Engineering, 2021.

[15] Dabhade, A., Roy, S., Moustafa, M. S., Mohamed, S. A., El Gendy, R., & Barma, S. (2021, December). Extreme Weather Event (Cyclone) Detection in India Using Advanced Deep Learning Techniques. In 2021 9th International Conference on Orange Technology (ICOT) (pp. 1-4). IEEE.

[16] Li, H., Cui, Y., Wang, S., Liu, J., Qin, J., & Yang, Y. (2020). Multivariate financial time-series prediction with certified robustness. IEEE Access, 8, 109133-109143.

[17] Sharma, S., Rai, S., & Krishnan, N. C. (2020). Wheat crop yield prediction using deep LSTM model. arXiv preprint arXiv:2011.01498.
[18] Sagan, V., Maimaitijiang, M., Bhadra, S., Maimaitiyiming, M., Brown, D. R., Sidike, P., & Fritschi, F. B. (2021). Field-scale crop yield prediction using multi-temporal WorldView-3 and PlanetScope satellite data and deep learning. ISPRS journal of photogrammetry and remote sensing, 174, 265-281.
[19] Jin, X. B., Yang, N. X., Wang, X. Y., Bai, Y. T., Su, T. L., & Kong, J. L. (2020). Hybrid deep learning predictor for smart agriculture sensing based on empirical mode decomposition and gated recurrent unit group model. Sensors, 20(5), 1334.
[20] Khaki, S., & Wang, L. (2019). Crop yield prediction using deep neural networks. Frontiers in plant science, 10, 621.
[21] Son, H., & Kim, C. (2020). A deep learning approach to forecasting monthly demand for residential–sector electricity. Sustainability, 12(8), 3103.
[22] Mehdipour Pirbazari, A., Farmanbar, M., Chakravorty, A., & Rong, C. (2020). Short-term load forecasting using smart meter data: A generalization analysis. Processes, 8(4), 484.
[23] Gasparin, A., Lukovic, S., & Alippi, C. (2022). Deep learning for time series forecasting: The electric load case. CAAI Transactions on Intelligence Technology, 7(1), 1-25.
[24] Shao, X., Pu, C., Zhang, Y., & Kim, C. S. (2020). Domain fusion CNN-LSTM for short-term power consumption forecasting. IEEE Access, 8, 188352-188362.
[25] Kim, M., Choi, W., Jeon, Y., & Liu, L. (2019). A hybrid neural network model for power demand forecasting. Energies, 12(5), 931.
[26] Bouktif, S., Fiaz, A., Ouni, A., & Serhani, M. A. (2018). Optimal deep learning lstm model for electric load forecasting using feature selection and genetic algorithm: Comparison with machine learning approaches. Energies, 11(7), 1636.
[27] Hu, T., Wu, W., Guo, Q., Sun, H., Shi, L., & Shen, X. (2019). Very short-term spatial and temporal wind power forecasting: A deep learning approach. CSEE Journal of Power and Energy Systems, 6(2), 434-443.
[28] Tan, M., Yuan, S., Li, S., Su, Y., Li, H., & He, F. (2019). Ultra-short-term industrial power demand forecasting using LSTM based hybrid ensemble learning. IEEE transactions on power systems, 35(4), 2937-2948.
[29] Yan, K., Wang, X., Du, Y., Jin, N., Huang, H., & Zhou, H. (2018). Multi-step short-term power consumption forecasting with a hybrid deep learning strategy. Energies, 11(11), 3089.
[30] Liu, B., Zhao, S., Yu, X., Zhang, L., & Wang, Q. (2020). A novel deep learning approach for wind power forecasting based on WD-LSTM model. Energies, 13(18), 4964.
[31] Xie, H., Zhang, L., & Lim, C. P. (2020). Evolving CNN-LSTM models for time series prediction using enhanced grey wolf optimizer. IEEE Access, 8, 161519-161541.
[32] Freeman, B. S., Taylor, G., Gharabaghi, B., & Thé, J. (2018). Forecasting air quality time series using deep learning. Journal of the Air & Waste Management Association, 68(8), 866-886.
[33] Bata, M. T., Carriveau, R., & Ting, D. S. K. (2020). Short-term water demand forecasting using hybrid supervised and unsupervised machine learning model. Smart Water, 5(1), 1-18.
[34] Chu, Y., Huang, C., Xie, X., Tan, B., Kamal, S., & Xiong, X. (2018). Multilayer hybrid deep-learning method for waste classification and recycling. Computational Intelligence and Neuroscience, 2018.
[35] Du, S., Li, T., Yang, Y., & Horng, S. J. (2019). Deep air quality forecasting using hybrid deep learning framework. IEEE Transactions on Knowledge and Data Engineering, 33(6), 2412-2424.
[36] Abrishami, A., & Aliakbary, S. (2019). Predicting citation counts based on deep neural network learning techniques. Journal of Informetrics, 13(2), 485-499.
[37] Sharfuddin, A. A., Tihami, M. N., & Islam, M. S. (2018, September). A deep recurrent neural network with bilstm model for sentiment classification. In 2018 International conference on Bangla speech and language processing (ICBSLP) (pp. 1-4). IEEE.
[38] Suhermi, N., Prastyo, D. D., & Ali, B. (2018). Roll motion prediction using a hybrid deep learning and ARIMA model. Procedia computer science, 144, 251-258.
[39] Ali Shah, S. A., Uddin, I., Aziz, F., Ahmad, S., Al-Khasawneh, M. A., & Sharaf, M. (2020). An enhanced deep neural network for predicting workplace absenteeism. Complexity, 2020.
[40] Kaneko, Y., & Yada, K. (2016). A deep learning approach for the prediction of retail store sales. In 2016 IEEE 16th International conference on data mining workshops (ICDMW) (pp. 531-537). IEEE.
[41] Vallés-Pérez, I., Soria-Olivas, E., Martínez-Sober, M., Serrano-López, A. J., Gómez-Sanchís, J., & Mateo, F. (2022). Approaching sales forecasting using recurrent neural networks and transformers. Expert Systems with Applications, 201, 116993.
[42] Tealab, A. (2018). Time series forecasting using artificial neural networks methodologies: A systematic review. Future Computing and Informatics Journal, 3(2), 334-340.
[43] Rajagukguk, R. A., Ramadhan, R. A., & Lee, H. J. (2020). A review on deep learning models for forecasting time series data of solar irradiance and photovoltaic power. Energies, 13(24), 6623.
[44] Sagheer, A., & Kotb, M. (2019). Time series forecasting of petroleum production using deep LSTM recurrent networks. Neurocomputing, 323, 203-213.
[45] Paramasivan, S. K. (2021). Deep Learning Based Recurrent Neural Networks to Enhance the Performance of Wind Energy Forecasting: A Review. Rev. d'Intelligence Artif., 35(1), 1-10.
[46] Dong, S., Wang, P., & Abbas, K. (2021). A survey on deep learning and its applications. Computer Science Review, 40, 100379.
[47] Goodfellow, I., Pouget-Abadie, J., Mirza, M., Xu, B., Warde-Farley, D., Ozair, S., ... & Bengio, Y. (2014). Generative adversarial nets. Advances in neural information processing systems, 27.

[48] Fu, R., Chen, J., Zeng, S., Zhuang, Y., & Sudjianto, A. (2019). Time series simulation by conditional generative adversarial net. arXiv preprint arXiv:1904.11419.
[49] Wu, N., Green, B., Ben, X., & O'Banion, S. (2020). Deep transformer models for time series forecasting: The influenza prevalence case. arXiv preprint arXiv:2001.08317.
[50] Fawaz, H. I., Forestier, G., Weber, J., Idoumghar, L., & Muller, P. A. (2018, December). Transfer learning for time series classification. In 2018 IEEE international conference on big data (Big Data) (pp. 1367-1376). IEEE.
[51] Ye, R., & Dai, Q. (2021). Implementing transfer learning across different datasets for time series forecasting. Pattern Recognition, 109, 107617.
[52] Ma, J., Cheng, J. C., Lin, C., Tan, Y., & Zhang, J. (2019). Improving air quality prediction accuracy at larger temporal resolutions using deep learning and transfer learning techniques. Atmospheric Environment, 214, 116885.
[53] Senthil, K. P. (2017). A review of soft computing techniques in short-term load forecasting. Int J Appl Eng Res, 12(18), 7202-7206.
[54] Senthil Kumar, P., & Lopez, D. (2016). A review on feature selection methods for high dimensional data. International Journal of Engineering and Technology, 8(2), 669-672.
[55] Subbiah, S. S., & Chinnappan, J. (2020). An improved short term load forecasting with ranker based feature selection technique. Journal of Intelligent & Fuzzy Systems, 39(5), 6783-6800.
[56] Swaroop, G., Senthil Kumar, P., & Muthamil Selvan, T. (2014). An efficient model for share market prediction using data mining techniques. Int J Appl Eng Res, 9(17), 3807-3812.
[57] Subbiah, S. S. (2022). Deep Learning Based Load Forecasting with Decomposition and Feature Selection Techniques. Journal of scientific & Industrial research, 81, 505-517.
[58] Subbiah, S. S., & Chinnappan, J. (2022). Deep learning based short term load forecasting with hybrid feature selection. Electric Power Systems Research, 210, 108065.
[59] Miotto, R., Wang, F., Wang, S., Jiang, X., & Dudley, J. T. (2018). Deep learning for healthcare: review, opportunities and challenges. Briefings in bioinformatics, 19(6), 1236-1246.
[60] Qin, F. Y., Lv, Z. Q., Wang, D. N., Hu, B., & Wu, C. (2020). Health status prediction for the elderly based on machine learning. Archives of gerontology and geriatrics, 90, 104121.
[61] Agila, N., & Senthil Kumar, P. (2020). An efficient crop identification using deep learning. International Journal of Scientific & Technology Research, 9(1), 2805-2808.
[62] Amani, M. A., & Marinello, F. (2022). A deep learning-based model to reduce costs and increase productivity in the case of small datasets: A case study in cotton cultivation. Agriculture, 12(2), 267.
[63] Li, J., Li, X., & Wang, K. (2019). Atmospheric PM2. 5 concentration prediction based on time series and interactive multiple model approach. Advances in Meteorology, 2019.
[64] Yan, K., Li, W., Ji, Z., Qi, M., & Du, Y. (2019). A hybrid LSTM neural network for energy consumption forecasting of individual households. Ieee Access, 7, 157633-157642.