Journal of Information Systems Engineering and Management
2025, 10(9s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

Scientific Analysis of Various Computational Intelligence Methods used for Weather Forecasting

Umesh Kumar[1*], Neetu Sharma[2]
[1]*Government Mahila Engineering college Ajmer, Rajasthan, India*
[2]*Government Engineering college Ajmer, Rajasthan, India*
* **Corresponding Author**: umesh@gweca.ac.in

ARTICLE INFO

Received: 29 Oct 2024

Revised: 20 Dec 2024

Accepted: 30 Dec 2024

ABSTRACT

Numerous studies in the field of weather forecasting have been conducted since technology started to develop in order to better understand how to manage weather by employing the appropriate kind of forecasting. This work is an evaluation report of facts and figures from the literature on weather forecasting incorporate with machine learning(ML), and deep learning(DL) models. Meteorologists, scientists, and researchers have created a wide range of designs, models, simulation systems, and prototypes to increase prediction accuracy. The first portion examines the literature on previous weather forecasting work, the application of numerous ML and DL models for weather forecasting along with the associated challenges. The second Section is all about the analysis description and drawbacks of current DL weather forecasting models. Several flaws were discovered following the study of prior models. The most common concerns are that running several equations simultaneously which are non-linear in nature requires a significant amount of computer resources and takes a long time to process. computer using Data-driven modeling techniques can be used to reduce the complexity of earlier models. ML and DL, in particular, can more accurately reflect a physical process's nonlinear or intricate underlying features.

Keywords: Weather Forecasting, ML, DL, MISO,

I. INTRODUCTION

A. Machine Learning

Machine Learning is an significant topic of discussion among researchers pertaining to reviewing and validating various weather forecast models. It provides a family of algorithms which enables the computer with the ability to automatically learn from experience and perform better by adapting to new situations [1]. The ML algorithms may be supervised, semi-supervised, unsupervised or reinforcement. In supervised ML algorithms, models learn from the past labeled data and use it to generalize the new data [2]. In contrast, the unsupervised ML algorithms study the unlabeled data and try to draw inference rather than figuring out the correct output [3]. The semi-supervised ML algorithms use labeled as well as unlabeled data and fall somewhere in between the above two learning methods. The reinforcement ML algorithms interact with the environment to learn the best action. As supervised learning algorithms analyze the past-labeled data to predict the unseen patterns, it suits to the problem of Weather forecasting. Accuracy is one of the most vital factor in forecasting. Additionally, data related with past weather conditions consist of input data and output data, can be used to create a training data set. Therefore, machine learning models using supervised algorithms have been widely used in weather forecasting [4].

Numerous machine learning methodologies are currently used for weather forecasting [5] and these approaches can be combined with different approaches to improve the efficiency of models. Short-term weather forecasting is a strong suit for ML, but medium-to-long-term climate forecasting is difficult because of complicated variables and data constraints[6].

General Machine Learning Models

Neural network is a small computing network model in which each neuron generates a single output value from a vector of input values. Logistic regression is also called as model with minimal neural network. It is technically a neural network comprises of a single hidden layer. The logistic regression model can be defined in two words such as logistic and regression [7].

Another straightforward and popular supervised machine learning technique for classification and regression tasks is the DT, in which the nodes represent the choice features, the branches show the likely outcomes from the nodes, and the leaves describe the classes[8]. Every single parent node should have at least one child node in the DT. The DT algorithms are used to handle both classification (classification trees) and regression (regression trees) problems [9]. The source dataset, which serves as the DT's root node, is divided into its succeeding nodes to create the DT. A set of guidelines based on classification features is used to partition the dataset. [10]. The main limitation of DT is it is highly sensitive to over fitting problem and node overlapping is there. However, this is reduced in random forest classifier, which constitute ensemble of DTs for classification [11].

Random Forests (RF) is the most famous techniques in information mining. Multiple DTs are built in a random fashion and in vast manner during the training time, gives us the impression like a forest and the class is predicted by considering the mode of the classes. It's an algorithm for guided ensemble learning. The Random Forest Tree [12] combines a bunch of weak learners to create a more powerful categorisation prediction. To put it another way, it is a random forest composed of DTs (weak learners) and strong predictors.Using a learning Algorithm, the main purpose of random forest is to merge several base-level predictors into a single, reliable, and efficient predictor [13].The training data and the objective values for the testing data are the two arrays used to train the classifiers when creating a random forest[14]. Leo Breiman and Adele Cutler [15] created the Random Forests computation.

Support Vector Machine (SVM) is a popular statistical learning concept based supervised ML technique that is defined using an optimal separating hyper plane, which can be used for ML applications such as classification and regression [16]. The vectors (training data) that define the hyper plane are called support vectors. The SVM technique was used in various extents, such as image, numerical, and text classification and regression applications. The kernel, which is a set of mathematical operations, is the foundation of SVM approaches. Data conversion into the required format is the kernel's responsibility. Different SVM methods utilize various sorts of kernel functions. [17]. The most common kernel functions are linear, nonlinear, polynomial, and sigmoid. The SVM technique takes a long training time on massive datasets than other ML algorithms.

ANNs are self-adaptive data driven nonlinear models that try to simulate the structural components and underlying functionalities of the biological neural network[18]. In order to do so, the basics of ANN are kept same to that of biological neural network. The simulation of biological neural network starts with its basic processing element i.e., “neuron”. Equivalent to the components of biological neuron such as dendrite, soma, and axon, the artificial neuron has inputs, weights, bias, a summation unit, a transfer (activation) function and an output unit. In biological neuron, the information is input to the neuron through dendrites; soma processes it and produces the output through axon. Similarly, in artificial neuron, the information is input to the neuron through weighted links. Then, it computes the weighted summation of inputs and “processes” the computed sum with a transfer function to produce the output.

Limitations of Machine Learning Models

Nowadays, artificial intelligence and ML techniques are preferred for data analytics because they can effectively handle real-time, non-linear data. Statistical framework serves as the foundation for ML models, which are implemented as supervised and unsupervised algorithms. The following are problems that ML algorithms generally run into while attempting to predict events:

Time Series Issues: The majority of ML algorithms are created focused on classification issues, while forecasting requires a distinct algorithmic strategy. For time series analysis, models that perceive the capability of incorporating the aforementioned states for forecasting the ensuing outcomes are highly suited. Recurrent neural networks can be used for managing time series forecasting.

Lack of Quality Data: The input needs to be aligned [19] based on time and in defined time intervals. Errors in the training set and missing values will propagate into the predicting variable. Additionally, more detailed data are anticipated for better pattern learning, whereas aggregation favours learning at the macro-level. The forecast may deviate significantly from the final result if the data pattern changes suddenly.

Accuracy Concern: Another key issue is that the model may not yield the same accuracy for test data because it must be trained for a change in data values each time. Despite their effectiveness, ML models aim towards forecasting accuracy that is on par with statistical techniques. Future errors, heterogeneous series types, over-fitting in essential preprocessing, and unacceptable preprocessing are the justifications [20].

Lack of Interpretability: Lack of transparency in a model's underlying logic and workings can lead to severe issues because it makes it impossible for specialists to comprehend the rationale and validate the decisions made by the system [21]. Any artificial intelligence or ML system's success is mostly influenced by interpretability.

Stochastic Not Deterministic: For prediction, many ML models concentrate on probability measurements. They therefore do not adhere to any physical restrictions, such as the allowed ranges for the forecasting parameters. For instance, the value of rainfall cannot be negative.

B. Deep Learning

Deep learning (DL) is a promising research area as it is a subset of ML methods and further ML is a subset of artificial intelligence. DL approaches achieved good performances in various fields like speech recognition, automatic navigation systems, sentiment analysis, and pattern matching [22][23]. In addition to the input and output layers, DL approaches generate numerous hidden layers to perform feature selection and training. As a result, DL techniques outperform ML techniques in terms of results [24][25]. Research proposed by G.E. Hinton in 2006 [26], which attempts to model data by combining different non-linear system with complex architecture. It is the subfield of ML in which learning methods are performed based on the data representation or feature learning. In DL system, set of methods are used to allow the system to discover the representation automatically required for feature detection or classification process from the raw data. Fig. I shows the process of DL. Neural Networks are the set of algorithms used in ML to perform complex mapping from the input to the output. Neural networks have the ability to provide the training on supervised and unsupervised learning.

DL is basically an artificial neural network and between the input layers and output layers of a DL network are several hidden layers. It could even accept raw data as input and generate appropriate feature representation with different levels of abstraction at each hidden layer [27].

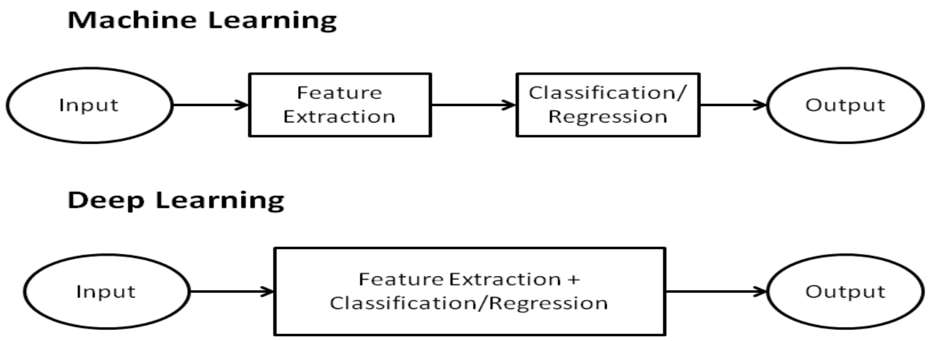

Fig.1 ML v/s DL

DL permits the computational frameworks that have multiprocessing layers to learn depictions of data with multiple abstraction levels. It is a sort of ML that facilitates computers to learn as of experience and comprehend the world grounded on a hierarchy of concepts. DL has been used in a variety of domains over the past ten years with remarkable results in terms of accuracy. The graph in Fig. II shows the rise of DL approach for the period of 2009 to 2019. The X-axis shows the years while Y-axis shows the interest of researchers in DL, which continuously increasing over the time.

One of the key advantages of DL, as shown in Fig. 1, is that algorithms can incrementally learn detailed features from the input data, eliminating the need for domain experts and hardcore extraction of features. However, in ML an expert person of the domain is required to extract features from data to reduce the complexity and by this way patterns can be made more clear for the learning algorithms. Consequently, DL solves complete problem at a time more precisely while ML divides the whole problem into different modules, and then solve each of them separately, and finally merge the results of various modules.

II. LITERATURE SURVEY

Recent days, DL methods are doing impressive performance in various applications [28] across multiple datasets including weather prediction which are not needed to provide with the pre-defined features, but it can learn the features from the dataset itself [29]. DL methods that contain many networks such as Artificial Intelligence, ML, DL, CNN, RNN and DBN used to perform the learning process effectively. The DL methods includes the enhanced software engineering, improved learning procedures and accessibility of computational power and training data [30] which are motivated by neuro science and achieved better impact in different areas such as Natural Language Processing and Computer Vision [31].

The authors have conducted a survey of weather forecasting techniques, covering their classifications, benefits, drawbacks, and limits also summarises, evaluates, and clarifies the work done on DL in the past [32]. Recent days, DL is one of the successful leaning approaches, because of the three important reasons such as improve the abilities of chip processing, lower cost of hardware and significant improvement in ML algorithms[33]. Some of the important state of the art models based on ML and DL for forecasting the weather are discussed along with their merits and demerits in Table 1.

Table 1 : ML and DL based weather forecasting techniques.

Weather forecasting based on ML & DL	**Input/output parameters**	**Merits**	**Demerits**
“Modelling & forecasting the daily maximum temperature using abductive ML (Abdel-Aal and Elhadidy, 1995)” [34]	"Inputs are Humidity Temperature, Wind Direction, Pressure, Wind Speed" "Output is Maximum Temperature"	"Reduced computation. Faster and more automated model. Improved prediction accuracy compared to statistical prediction models" "Better longterm forecasting"	"Used to predict individual output parameters" "Not suitable for region-specific forecasting."
“ANN based Snowfall and rainfall forecasting from weather radar images (Ochiai et al., 1995)” [35]	"Inputs are Rainfall and Snowfall" "Outputs are Rainfall and Snowfal"	"Greater accuracy then traditional cross-correlation methods" "Reduced computational training time compared to traditional ANN"	"Predicts one parameter at a time" "complex" "Short Term Forecasting only"
“Rainfall estimation using M-PHONN model (Qi and Zhang, 2001)”[36]	"Inputs are Cloud Top Temperature, Cloud Growth, and Rainfall." "Output is Rainfall"	"Improved speed and accuracy thentraditional ANN. " "Ideal for very short time rainfall forecasting."	"Requires satellite data to process Complex model"
“Intelligent weather monitoring systems using connectionist models (Maqsood et al., 2002)” [37]	"Inputs are Wind speed, Max. and Min. Temp. on daily basis." "Outputs are Wind Speed , Daily basis Max. and Min. Temperature. "	"Greater prediction accuracy " "Better forecasting for long term." "Can be used to predict all parameters simultaneously."	"Not suitable for region-specific weather forecasts" "More Complex"
“An efficient weather forecasting system using ANN (Baboo and Shereef,	"Inputs are Wind speed , Humidity, Pressure, Temperature.	"Short and long-term forecasting can be used. Accurate forecasting for	"Single parameter prediction only.

Weather forecasting based on ML & DL	Input/output parameters	Merits	Demerits
2010)" [38]	Output is Temperature"	Short term. Less complex."	Time-consuming"
"A rough set-based fuzzy neural network algorithm for weather prediction (Li and Liu, 2005)"[39]	"Input are Wind speed, Dew and Temperature. Outputs are Wind speed, Dew, visibility and Temperature."	"One model can be used to predict four parameters. Greater accuracy in short and medium-term forecasting. Good for region-specific weather forecasts"	"Complex Special equipment is required to measure the Dew, temperature."
"DNN based ultra-short-term wind forecasting (Dalto et al., 2015)" [40]	"Input is Wind Output is Wind"	"Outperforms shallow ones. It can be used for local wind forecasts."	"Short-term forecasting limitation. Complex model."
"Weather forecasting using DL techniques (Salman et al., 2015)" [41]	"Input is Rainfall Output is Rainfall"	"Accurate RNN-based prediction results when compared with CRBM (Conditional Restricted Boltzmann Machine) and CN (Convolutional Network) models."	"Limited to short-term forecasting. Vanishing gradient issue. Prediction of single output parameter only."
"Short-term local weather forecast using dense weather station by a DNN (Yonekura et al., 2018a)" [42]	"Inputs are Temperature, Humidity, Pressure, Wind, and Rain. Outputs Temperature or Rain"	"Compared to SVR and RF, yield the maximum accuracy for rain prediction (RF). Up to an hour of data prediction accuracy."	"Limitation of very short-term forecasting. limited to predict one weather parameter High model complexity"
"Convolutional LSTM Network: A ML Approach for Precipitation Nowcasting (Shi et al., 2015)"[43].	"Input and output is Precipitation"	"Appropriate for area-specific prediction. No issue of vanishing gradient".	"One output prediction only."
"Nevada weather forecasting: A DL approach (Hossain et al., 2015)" [44]	"Inputs are Wind speed, Humidity, Pressure and Temperature Output is Temperature"	"Accurately predict long-term temperatures. Used for area-specific temperature forecasts."	"Limited to a single output parameter. Not appropriate for very short-term temperature forecasting."
"LSTM Recurrent Neural Networks based Sequence to Sequence Weather Forecasting (Akram and El, 2016)" [45]	"Inputs and Outputs are Temperature, Humidity, and Wind Speed.	"Predict common weather variables with reasonable accuracy up to one day. Suitable for regional forecasting."	"suitably predicting one output parameter at a time. Model is complex."
"A DL Based on Bidirectional Gated Recurrent Unit for the prediction of Wind Power	"Inputs are Wind Speed and Wind Direction. Output is Wind Power"	"up to 6 hours wind power prediction. appropriate for both area-specific and regional wind	"Only Single parameter output. Considerable time consuming to train the network compared

Weather forecasting based on ML & DL	Input/output parameters	Merits	Demerits
(Deng et al., 2019)"[46]		power forecasts."	to LSTM."
Deep Learning Based Prediction Of Weather Using Hybrid_stacked Bi-Long Short Term Memory(U Sharma et al., 2022)"[47]	"Inputs are Railfall, Temperature ,Humidity ,Windspeed ,Visibility and pressure Outputs are Temp, Humidity, Rainfall and Wind speed"	"Accurately predict various weather parameters for long-term. Can be used for area-specific forecasts."	"Limited to predicting one output parameter at a time. Complex Model."

Researchers have used different ML and DL models, either separately or in combination, with input parameters such as wind speed, dew point, rainfall, pressure, snow, humidity and temperature as shown in Table I. The majority of methods for NN-based weather forecasting rely on five or fewer linked parameters, and each model has advantages and disadvantages of its own.

III. DEEP LEARNING MODELS: ANALYSIS

Table I in Section II lists the benefits and drawbacks of cutting-edge ML and DL techniques and Table II in section III emphasises the distinctions and contributions of current DL weather forecasting models. The regression difficulty in fine-grained weather forecasting for certain user groups inside a given domain is addressed by a model that is suggested based on these findings.

Table 2. DL forecasting models along with their Analysis, Description and drawbacks

DL Model	Contribution and Analysis	Drawbacks
"Ultrashort-term wind forecasting model (Dalto et al., 2015)" [40].	"Contributions to deep neural networks for weather forecasting better than shallow ones in MSE values. Very accurate wind forecast compared to flat networks. Regional or region-specific forecasts. "	"SISO model & Can predict wind parameters for less than 1 hour. Higher computational complexity compared to neural networks."
"Deep neural network rainfall prediction model (Hernández et al., 2016)" [48].	"Contributed an architecture based on a deep learning rainfall prediction model that outperforms ML models in multi-layer recognition. More accurate precipitation forecast based on MSE. Area-specific forecasts up to 24 hours."	"SISO(Single input Single output) model. Not suitable for local area forecasting. Inaccuracy in results due to non consideration of observations of the interrelated parameters. "
"RNN short-term forecasting model (Salman et al., 2015)" [41].	"Contributed to Deep Learning RNN for better weather forecasting than CRBM and CN networks." "Precipitation forecast with higher accuracy than CRBM and CN based on MSE values." Region-Specific & Short Term Forecast "	"Higher memory consumption compared to CRBM and CN based models. SISO model. Applies to 6-hour forecasts only. not considering related parameters"
"Deep neural network shortterm	"Contributed to deep models that outperform SVR and RF scores compared to MSE scores.	"MISO (Multiple input Single output) model

DL Model	Contribution and Analysis	Drawbacks
forecasting model (Yonekura et al., 2018b)" [42].	Rain and temperature forecast with higher accuracy than SVR and RF. Predicts single parameter at a time. Good for local forecasting Take five input parameters and resulting single parameter as output (Multi-Input Single-Output)."	Can forecast for less than 60 minutes. Not appropriate for regional forecasts. Large amount of computation compared to SVR and RF. The vanishing gradient problem persists during the training process."
"LSTM multilayered weather forecasting model (Salman et al., 2018)" [49].	" comparative study of single layer LSTM model and multi-layer LSTM model in weather forecasting. Result shows that multi-layer LSTMs provide accurate predictions compared to single-layer ones. Accurate humidity, dew point, temperature, and pressure predictions are based on MSE compared to ground truth. Regional or region-specific MISO-based forecasts."	"Computational complexity is higher for multi-layer LSTM models compared to single-layer models. Productive output requires special hardware such as a GPU memory unit. For long input sequences, storing partial results of various cell gates consumes a lot of memory."
"Precipitation nowcasting model (Shi and Dustdar, 2016)" [50].	"Author contributed convolutional LSTM networks that outperform fully connected LSTM approaches to precipitation nowcasting. Accurate precipitation forecast compared to ground truth based on MSE values. Very short-term Area-specific forecast."	"Predicts less than 1-hour forecast. SISO model. Inaccuracy in results due to non consideration of observations of the interrelated parameters. Productive output requires a GPU memory unit."
"DNN based feature representation model (Liu et al 2014)" [51].	"Contribution of DNN models that characterize raw weather data layer by layer and outperform traditional SVR. Higher accuracy temperature, dew, pressure and wind speed predictions compared to the SVR. Regional long-term forecast up to one day. Four input parameters and one output parameter at the same time (multi-input single output)."	"Not suitable for region-specific weather forecasts. Only one output can be predicted at a time. This model requires specific hardware such as a GPU memory unit for productive output. More computation than SVR."
"A DNN model for temperature forecasting (Hossain et al., 2015)" [44].	"Contributed a stacked auto-encoder DL model" "As compared to the ground truth forecasts temperature up to 97.97% accurate. Area-specific long-term forecasts up to 30 days based on MISO."	"At a time predicts only single output parameter. lot of memory is used to store partial results. Training process is time-consuming."
"Sequence to sequence weather forecasting model (Akram and El, 2016)" [45].	"The authors suggested using multi-stacked LSTMs to map sequences of weather values that are all the same length. Based on MSE values, reasonably accurate forecasts of temperature, humidity, and wind speed compared to the actual weather. Regional prediction up to 24 hours based on MIMO."	"Not appropriate for regional weather forecasting. needs a lot of memory and several cell gates because of the lengthy input sequence. For faster output, it is advised to use GPU memory units."
"Bi-LSTM wind	"Bidirectional gated recurrent networks were	"Parameters can be predicted for 6 hrs

DL Model	Contribution and Analysis	Drawbacks
power forecasting model (Deng et al., 2019)" [46].	suggested by the authors for forecasting weather power. Based on the MSE values, the model determined a precise link between the power, wind speed and direction in comparison to the actual situation. Local or regional prediction based on MISO."	Uses a single observational weather parameter without taking into account how those parameters connect to one another to make an accurate prediction. Advised to employ specialized hardware like GPU memory units."

Analyzing the Table.2 above, it is clear that both regional weather forecasting and area specific forecasting can use DL techniques. Each approach has its own disadvantages. Authors [41][42][51] outperformed traditional ML approaches using DL models. Furthermore, Authors[45][40][41][49] found that in forecasting of weather, DNN outperformed shallow ones. As a result, DNN were chosen for this investigation over shallow neural networks. In addition, Table.1 discusses all cutting-edge deep models that use fewer than five interrelated parameters. This research contributed to the development of a full weather forecasting model with a few additional input parameters.

In Table.2, there are three types of regression, some of which are multi-input and multi-output, some of which are multi-input and single-output, and some of which are single-input and single-output. All of these regressions can achieve cutting-edge results using DL techniques. The prediction accuracy of any given parameter is determined using two factors: previous observations and prior observations of connected parameters [52]. Furthermore, parameters associated with past observations can generate a more accurate forecast than parameters with a lower degree of connection [53]. MIMO produces lower MSE values when compared with MISO [54] also Jiancheng Qin et al. [55] states that "Although the MIMO model structure is more complicated compared to the other three models, there is no guarantee that the MIMO model will produce the best forecasting results", therefore MIMO is also not considered. As a consequence, MISO model is considered to overcome the regression problem in our proposed work for weather forecasting.

IV. RESULTS AND DISCUSSIONS

A comprehensive examination of weather forecasting models reveals that each model has advantages and disadvantages, yet certain concerns can be handled by upgrading the model without sacrificing the method's originality. A statistical method ARIMA is used for most forecasting applications that competes equally with ML models; however, due to its univariate forecasting nature, DNN outperforms it. The following difficulties have been noticed in existing techniques.

- Traditional ANN relies on a gradient descent approach for back propagation, which causes the vanishing gradient problem.
- Inadequate memory for storing past time series data states.
- Longer forecasting periods may experience local minimums, necessitating more epochs for accurate prediction. This raises computational complexity. Standard models like ARIMA and ANN have high prediction accuracy during regular seasons, but this decreases during harsh seasons.
- Prediction models require a minimum of 10 years of data to understand weather patterns, but most existing research (with a few exceptions) does not cover enough time.
- Existing neural network-based weather prediction models often use fewer than five or more inter-related input parameters, resulting in limited output parameters that are insufficient for accurate forecasting.
- SISO and MIMO regression models have reduced MSE values compared to MISO.

As a result, certain features in the weather forecast are regarded crucial, which contributes to the occurrence of the aforementioned issues in present models. This review provides insight into the implementation of a DL model to tackle some existing issues by taking into account several key elements for efficient weather forecasting.

V. CONCLUSION

DL model requires more computation power and tremendous amount of data for training and ML models having simpler architectures with fast training but When both are compared, DL produces better long-term forecasting results when combined with select cutting-edge algorithms. The DL models are utilised in conjunction with a variety of other climate models, and data is entered into the system from all around the world, resulting in a more precise and long-term prediction. In the meanwhile, it is admirable to investigate merging data of any region with the data of local weather station. This may improve the prediction accuracy of the intended models.

REFERENCES

[1] S. I. Diaz, E. F. Combarro and S.M. Mazza, "Machine Learning Applied to prediction of citrus production," Spanish journal of agricultural research, vol. 15, no. 2, p. 7, 2017.

[2] N. Singh, S. Chaturvedi, and S. Akhter, "Weather Forecasting Using Machine Learning Algorithm," 2019 Int. Conf. Signal Process. Commun. ICSC 2019, pp. 171–174, 2019.

[3] R. I. Rasel, N. Sultana, and P. Meesad, "An application of data mining and machine learning for weather forecasting," Adv. Intell. Syst. Comput., vol. 566, pp. 169–178, 2018.

[4] S. Murugan Bhagavathi et al., "Weather forecasting and prediction using hybrid C5.0 machine learning algorithm," Int. J. Commun. Syst., vol. 34, no. 10, pp. 1–14, 2021.

[5] S. Scher and G. Messori, "Predicting weather forecast uncertainty with machine learning," Q. J. R. Meteorol. Soc., vol. 144, no. 717, pp. 2830–2841, 2018.

[6] Chen, L., Han, B., Wang, X., Zhao, J., Yang, W., & Yang, Z. "Machine learning methods in weather and climate applications: A survey". Applied Sciences, 13(21), 12019, 2023

[7] M. S. Balamurugan and R. Manojkumar, "Study of short term rain forecasting using machine learning based approach," Wirel. Networks, vol. 3, 2019.

[8] W. Wei, Z. Yan, and P. D. Jones, "A decision-tree approach to seasonal prediction of extreme precipitation in eastern China," Int. J. Climatol., vol. 40, no. 1, pp. 255–272, 2020.

[9] Z. M. Fadlullah et al., "State-of-the-Art Deep Learning: Evolving Machine Intelligence Toward Tomorrow's Intelligent Network Traffic Control Systems," IEEE Commun. Surv. Tutorials, vol. 19, no. 4, pp. 2432–2455, 2017.

[10] N. Kalchbrenner, E. Grefenstette, and P. Blunsom, "A convolutional neural network for modelling sentences," 52nd Annu. Meet. Assoc. Comput. Linguist. ACL 2014 - Proc. Conf., vol. 1, pp. 655–665, 2014.

[11] P. Sermanet, S. Chintala, and Y. Lecun, "Convolutional neural networks applied to house numbers digit classification," Proc. - Int. Conf. Pattern Recognit., no. Icpr, pp. 3288–3291, 2012.

[12] W. Zaremba, I. Sutskever, and O. Vinyals, "Recurrent Neural Network Regularization," no. 2013, pp. 1–8, 2014.

[13] J. Chung, C. Gulcehre, K. Cho, and Y. Bengio, "Empirical Evaluation of Gated Recurrent Neural Networks on Sequence Modeling," arXiv preprint arXiv:1412.3555, pp. 1–9, 2014.

[14] R. Jozefowicz, W. Zaremba, and I. Sutskever, "An empirical exploration of Recurrent Network architectures," 32nd Int. Conf. Mach. Learn. ICML 2015, vol. 3, pp. 2332–2340, 2015.

[15] A. Cutler, D. R. Cutler, and J. R. Stevens, "Ensemble Machine Learning," Ensemble Mach. Learn., 2012.

[16] M. Sun, "Support Vector Machine Models for Classification," Encycl. Bus. Anal. Optim., pp. 2395–2409, 2014.

[17] A. N. Wang, Y. Zhao, Y. T. Hou, and Y. L. Li, "A novel construction of SVM compound kernel function," 2010 Int. Conf. Logist. Syst. Intell. Manag. ICLSIM 2010, vol. 3, no. 60843007, pp. 1462–1465, 2010.

[18] K. C. Luk, J. E. Ball, and A. Sharma, "A study of optimal model lag and spatial inputs to artificial neural network for rainfall forecasting," J. Hydrol., vol. 227, no. 1–4, pp. 56–65, 2000.

[19] P. Chronis, G. Giannopoulos, and S. Athanasiou, "Open issues and challenges on time series forecasting for water consumption," CEUR Workshop Proc., vol. 1558, 2016.

[20] S. Makridakis, E. Spiliotis, and V. Assimakopoulos, "Statistical and Machine Learning forecasting methods: Concerns and ways forward," PLoS One, vol. 13, no. 3, pp. 1–26, 2018.

[21] D. V. Carvalho, E. M. Pereira, and J. S. Cardoso, "Machine learning interpretability: A survey on methods and metrics," Electron., vol. 8, no. 8, pp. 1–34, 2019.

[22] C. Pradeepthi, V. J. Vijaya Geetha, S. Ramasubbareddy, and K. Govinda, Prediction of real estate price using clustering techniques, vol. 1054. 2020.

[23] S. Dargan, M. Kumar, M. R. Ayyagari, and G. Kumar, “A Survey of Deep Learning and Its Applications: A New Paradigm to Machine Learning,” Arch. Comput. Methods Eng., vol. 27, no. 4, pp. 1071–1092, 2020.

[24] Y. Xin et al., “Machine Learning and Deep Learning Methods for Cybersecurity,” IEEE Access, vol. 6, no. c, pp. 35365–35381, 2018.

[25] A. Basit, M. Zafar, X. Liu, A. R. Javed, Z. Jalil, and K. Kifayat, “A comprehensive survey of AI-enabled phishing attacks detection techniques,” Telecommun. Syst., vol. 76, no. 1, pp. 139–154, 2021.

[26] W. Liu, Z. Wang, X. Liu, N. Zeng, Y. Liu, and F. E. Alsaadi, “A survey of deep neural network architectures and their applications,” Neurocomputing, vol. 234, pp. 11–26, 2017.

[27] J. C. B. Gamboa, “Deep Learning for Time-Series Analysis,” arXiv preprint arXiv:1701.01887, 2017.

[28] S. Sengupta et al., “A review of deep learning with special emphasis on architectures, applications and recent trends,” Knowledge-Based Systems, vol. 194, p. 105596, 2020.

[29] S. Corbella, S. Srinivas, and F. Cabitza, “Applications of deep learning in dentistry,” Oral Surg. Oral Med. Oral Pathol. Oral Radiol., vol. 132, no. 2, pp. 225–238, 2021.

[30] J. Ahmad, H. Farman, and Z. Jan, “Deep Learning: Convergence to Big Data Analytics,” Springer Singapore, 2019.

[31] P. Patel and A. Thakkar, “The upsurge of deep learning for computer vision applications,” Int. J. Electr. Comput. Eng., vol. 10, no. 1, pp. 538–548, 2020.

[32] U. Kumar and N. Sharma, "A Scientific Analysis of Weather Forecasting Techniques and its Classification." Journal of Management World, 917-923, 2024.

[33] U. Sharma and C. Sharma. “Analysis and Literature Review of Weather Prediction and Forecasting Methods, TEST Engineering and Management, vol. 83, pp.12539-12549, May-June 2020

[34] R. E. Abdel-Aal and M. A. Elhadidy, “Modeling and forecasting the daily maximum temperature using abductive machine learning,” Weather & Forecasting, vol. 10, no. 2. pp. 310–325, 1995.

[35] K. Ochiai, H. Suzuki, S. Suzuki, N. Sonehara, and Y. Tokunaga, “Snowfall and rainfall forecasting from the images of weather radar with artificial neural networks,” Neural Networks Signal Process. - Proc. IEEE Workshop., pp. 473–481, 1996.

[36] M. Zhang and J. Crane, “Rainfall estimation using SPHONN model,” Proc. Int. Conf. Artif. Intell. IC-AI’04, vol. 2, pp. 695–701, 2004.

[37] I. Maqsood, M. R. Khan, and A. Abraham, “Intelligent Weather Monitoring Systems Using Connectionist Models,” Neural, Parallel Sci. Comput., vol. 10, pp. 157–178, 2002.

[38] S. S. Baboo and I. K. Shereef, An efficient weather forecasting system using artificial neural network,” International journal of environmental science and development, vol. 1, no. 4, pp. 321–326, 2010.

[39] K. Li and Y. S. Liu, “A rough set based fuzzy neural network algorithm for weather prediction,” 2005 Int. Conf. Mach. Learn. Cybern. ICMLC 2005, no. August, pp. 1888–1892, 2005.

[40] M. Dalto, J. Matusko, and M. Vasak, “Deep neural networks for ultra-short-term wind forecasting,” Proc. IEEE Int. Conf. Ind. Technol., vol. 2015-June, no. June, pp. 1657–1663, 2015.

[41] M. G. Salman, “Weather Forecasting using Deep Learning Techniques,” pp. 281–285, 2015.

[42] K. Yonekura, H. Hattori, and T. Suzuki, “Short-term local weather forecast using dense weather station by deep neural network,” IEEE International Conference on Big Data, no. 1, pp. 1683–1690, 2019.

[43] Z. Chao, F. Pu, Y. Yin, B. Han, and X. Chen, “Research on real-time local rainfall prediction based on MEMS sensors,” J. Sensors, vol. 2018, pp. 1–9, 2018.

[44] M. Hossain and S. J. Louis, “Forecasting the Weather of Nevada : A Deep Learning Approach,” vol. 2, pp. 2–7, 2015.

[45] M. Akram and C. El, “Sequence to Sequence Weather Forecasting with Long Short-Term Memory Recurrent Neural Networks,” Int. J. Comput. Appl., vol. 143, no. 11, pp. 7–11, 2016.

[46] Y. Deng, H. Jia, P. Li, X. Tong, X. Qiu, and F. Li, “A deep learning methodology based on bidirectional gated recurrent unit for wind power prediction,” Proc. 14th IEEE Conf. Ind. Electron. Appl. ICIEA 2019, pp. 591–595, 2019.

[47] U. Sharma and C. Sharma, "Deep Learning Based Prediction Of Weather Using Hybrid_stacked Bi-Long Short Term Memory," 2022 12th International Conference on Cloud Computing, Data Science & Engineering (Confluence), 2022, pp. 422-427, doi: 10.1109/Confluence52989.2022.9734133.

[48] Y. Echevarría, L. Sánchez, and C. Blanco, “Hybrid Artificial Intelligent Systems,” Lect. Notes Comput. Sci. (including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics), vol. 9648, pp. 250–260, 2016.

[49] A. G. Salman, Y. Heryadi, E. Abdurahman, and W. Suparta, "Single Layer & Multi-layer Long Short-Term Memory (LSTM) Model with Intermediate Variables for Weather Forecasting," Procedia Comput. Sci., vol. 135, pp. 89–98, 2018.

[50] W. Shi and S. Dustdar, "Cloud Cover Why Do We Need Edge Computing? The Promise of Edge Computing," Computer, no. 0018, 2016.

[51] J. N. K. Liu, Y. Hu, J. J. You, and P. W. Chan, "Deep neural network based feature representation for weather forecasting," Proc. 2014 Int. Conf. Artif. Intell. ICAI 2014 - WORLDCOMP 2014, pp. 105–110, 2014.

[52] J. Ding, J. Dai, Y. Chen, F. Hu, and X. Tang, "Helicity as a Method for Forecasting Severe Weather Events," Adv. Atmos. Sci., vol. 13, no. 4, pp. 533–538, 1996.

[53] Z. Di et al., "Assessing WRF model parameter sensitivity: A case study with 5 day summer precipitation forecasting in the Greater Beijing Area," Geophys. Res. Lett., vol. 42, no. 2, pp. 579–587, 2015.

[54] P. Hewage, M. Trovati, E. Pereira, and A. Behera, "Deep learning - based effective fine - grained weather forecasting model," Pattern Anal. Appl., vol. 24, no. 1, pp. 343–366, 2021.

[55] J. Qin, J. Yang, Y. Chen, Q. Ye, and H. Li, "Two-stage short-term wind power forecasting algorithm using different feature-learning models," Fundam. Res., vol. 1, no. 4, pp. 472–481, 2021.