Evaluation of weather forecast with DL

Pritish Samant
Deep Learning
Hochschule Hamm-Lippstadt
Lippstadt, Germany
pritish-sanjay.samant@stud.hshl.de

***Abstract*—Weather forecasts play a very important role in today's world in almost all sectors of the world economy. The accuracy of making decisions related to weather evaluation using deep learning is extremely important to ensure the smooth functioning of society. Earlier, it used to be just numerical evaluation, but recent developments in deep learning technologies have ensured that the new technologies can be effectively used for weather evaluation. In this paper, we will explore weather evaluation using deep learning techniques as well as the implementation of different techniques such as ANN and CNN.**

***Index Terms*—weather evaluation, Deep learning , ANN, CNN**

I. Introduction

The weather in this world constantly keeps changing, and this affects all the human beings on earth. Changes in weather are constant and have severe effects on this planet. The effects of weather change are sometimes drastic, which can lead to the devastation of agriculture, thus directly affecting humans. It can also lead to the end of civilization. Sometimes even a small change in weather can affect how we eat and what we eat, as well as how we live. For example, reliable weather forecasts allow us to provide early warning of natural disasters like cyclones, tsunamis, cloud bursts, etc. that can seriously inflict damage to both lives and property [1]. To avoid all of these, there was the necessity of developing a technique that we could use to avoid destruction. We, as humans, had started predicting or evaluating weather. Since earlier times, we have been using statistics or mathematics to evaluate weather. Although this was nowhere near perfect, we were still using it. Weather being very unpredictable, Building statistical or mathematical models for evaluating weather where only a few factors can be considered for evaluating weather was really difficult and time-consuming. This created a gap in the research on weather evaluation. There was always this need for another way of building models for weather evaluation. The prediction application in science and technology that makes use of the atmospheric conditions at a specific place and time is called weather forecasting [2].

With recent developments in the fields of machine learning and deep learning, research has begun to fulfill the need for weather evaluation using these methods. It was later found that machine learning techniques and deep learning algorithms can be implemented to evaluate or predict weather in real time, in less time, and with greater accuracy. Many researchers have been inspired to investigate hidden hierarchical patterns in the vast amount of weather datasets for weather forecasting by the emergence of deep learning techniques in the last ten years, the widespread availability of massive weather observation data, and the development of information and computer technology [3]. Also, the algorithms can be trained to learn weather patterns and improve their accuracy in their predictions. All this can be done using the computational power of sophisticated cybersystems. Also, more layers or factors can be added or trained to predict more accurate weather and learn weather patterns. The use of deep learning in weather modeling and representation has been spurred by the numerous researchers who have documented the technology's successful applications in a variety of fields. [4].

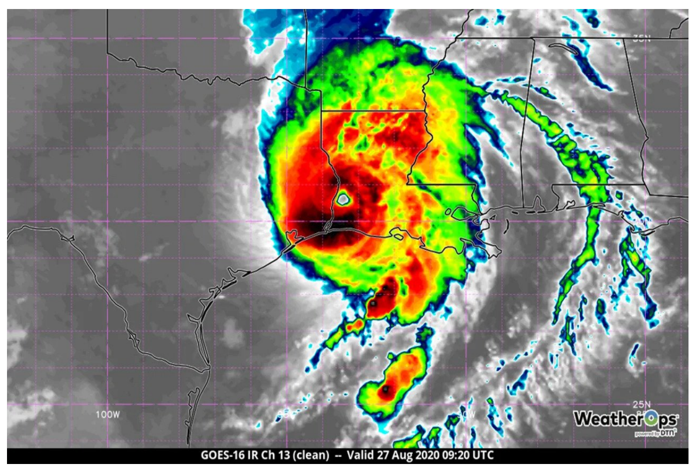

Fig. 1. Image of Stormy weather [9]

In this paper, we are going to see the weather evaluation using a deep learning technique called convolutional neural network (CNN), which is a network architecture of deep learning. Deep learning structures algorithms in layers to create an "artificial neural network" that can learn and make intelligent decisions on its own, and this is how deep learning is used in this experiment to make intelligent decisions [5]. It learns the weather patterns and trains itself on the available weather data to make accurate weather predictions. Also, in this paper, we will be using a weather dataset to analyze, learn from weather data, and predict weather using a deep learning technique.

II. WEATHER EVALUATION

Predicting weather is extremely challenging, and in this challenging situation, one of the toughest parts is predicting snow, rain, hail, etc. Different parameters are used to predict the weather. These can be called features for predicting weather using machine learning or deep learning. These parameters are temperature, pressure, humidity, and wind speed. These features are somehow connected to each other and can be used to predict weather at any location, provided we use the above features for the same location. These parameters will change for different locations. Companies use temperature and precipitation forecasts to estimate their needs for the next few days, and these forecasts are essential to agriculture. [6].

For instance, we can say that one of the factors depends on the other factors in a given location or area. All these features can be expressed in a form that a machine can understand, which is why instead of just mentioning the features, we use the values of the features as input. Then we train this data to predict the unknown or future weather. But since we are using this in a small area, the accuracy will be affected. The more you train, predict, and increase the area, the more accurate it will be. The used features, or unique attributes, use machine learning techniques, including deep learning, where weights are also assigned and are an important part of the prediction. In essence, a model is an approach that produces an objective value based on the unique attributes and weights assigned to each training variable. [7].

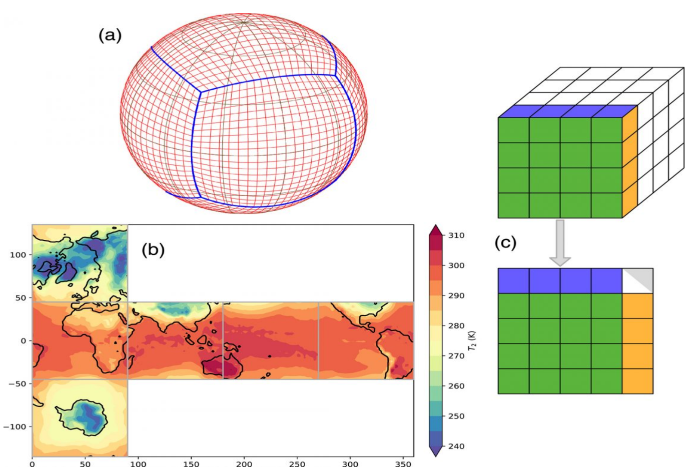

Fig. 2. First the authors divide the planet's surface into a grid with a six-sided cube (top left) and then flatten out the six sides into a 2-D shape, like in a paper model (bottom left). This new technique let the authors use standard machine learning techniques, developed for 2-D images, for weather forecasting [11]. [10]

Practically, we can have as many features as our machine can handle. The model has to be made in such a way that it has an error margin in case the forecasts go wrong. If the model assumes a cloud is in a slightly different location, then even with all the correct values of the features, we cannot accurately forecast the weather. The weather is very dynamic. It changes every minute. So we have to be careful with the forecasts, as there are constant changes in the weather, and one incorrect value of temperature, position of the cloud, or any other feature will result in the predictions or forecasts being inaccurate, and the inaccuracy will grow exponentially. There are some ways to tackle this. One can use stochastic models for the same. We can add a small amount of noise to the model, see the output, and compare the deviation between the noise and the output. If it is more than the added noise, we use the probabilistic prediction as the output. And this is how weather forecasts are done. The unstable nature of the atmosphere, the enormous processing power needed to solve intricate equations that represent the state of the atmosphere, errors in the initial condition measurement, and a lack of knowledge about the workings of the atmosphere can all contribute to inaccurate forecasts at times. [8]. A little difference in inputs will generate different outputs, even if it is slight. The advancements in technology and prediction models are really fast, and it can be assumed that new weather models can replace the current ones in no time. These advancements are a boon in helping us predict unwanted events or natural disasters. There's a great future in this sector, specifically in creating machine learning and deep learning models for weather forecasting.

III. DEEP LEARNING

Deep learning is a part of machine learning. Whereas machine learning is a subset of the domain of data science, Machine learning is nothing but the use of algorithms to train the input data, compute or calculate it, and return a possible output. Deep learning is just a technique that works similar to a human brain but is part of machine learning. This is inspired by the human brain. It is based on how a human brain could think and analyze the data in a logical way, and for this, deep learning uses a structure with multiple layers called a neural network. These neural networks analyze the information or data in a similar way to what a human brain does. Neural networks can be trained by being taught on a completely new set of data. Imagine we want to identify between a cat and a dog using data science techniques. If we use machine learning, we need to feed the machine the features of dogs and cats, and using these features and input data, the machine can differentiate between a dog and a cat. If we use deep learning for this, we don't need to do anything except feed the input data, which would be images of dogs and cats. The machine on its own, like a human brain, will identify features and train itself, and then it will differentiate between a dog and a cat.

This means that a deep learning algorithm in computer vision will first build representations based on its own low-level representations, which it will learn from a raw image, and then repeat the process for higher levels. citeb13. These neural network structures have layers, and on these layers are nodes called neurons. Every layer in the network has these

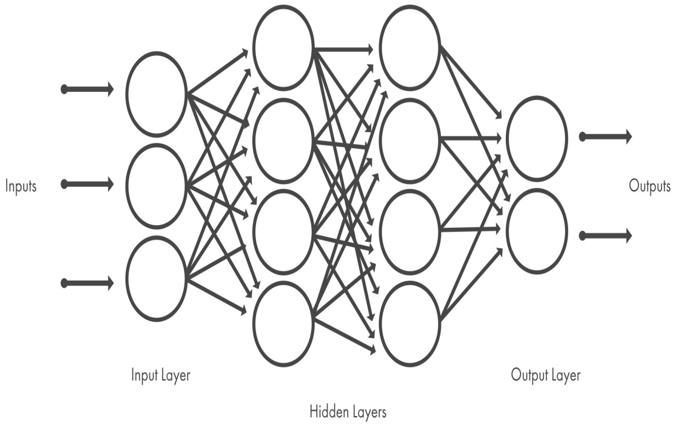

Fig. 3. Deep Learning Network [12]

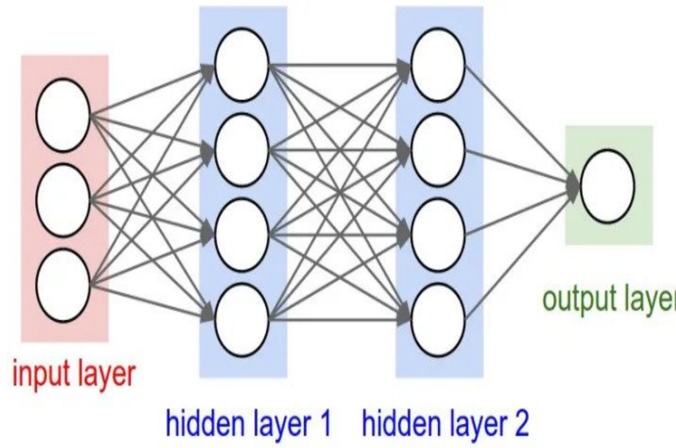

Fig. 4. Artificial Neural Network [14]

neurons. In a neural network, the data or information travels from one neuron to another through the layers along with its weight. Deep learning has its own challenges. To train the neural network and classify, differentiate, or predict to be accurate, the model requires a huge data set. A smaller data set would affect the accuracy or ability. It also requires high computational power. To overcome this, graphical processing units are used instead of central processing units. This, in turn, increases the cost. Training time, accuracy, or the optimization of the model depends on the size of the data and the number of hidden layers inside the neural network, as these hidden layers can range from one hidden layer to multiple layers.

A. Artificial Neural Network (ANN)

Consider a neural network as a box with two ends on opposite sides, where one can put all the inputs through one end and the output at the other end. Inside are all the hidden layers that we cannot see, but the computation happens inside the box. This neural network includes many units called neurons. Every layer has these units, called neurons. These neurons of every layer are connected to the neurons of other layers, and every neuron has its own weight. In the figure below, we can see that the circles or dots are units called neurons, and they are present in each and every layer of the network. This network looks like the network inside our brain. And thus the name Neural Network is derived. The hidden layers can either be one or multiple layers, depending on the model and the input data we have to process. Generally, we have been provided with the weights of the neurons, but if we have to calculate them, then it becomes very important to accurately calculate the weights of the network. This accuracy in turn drives the accuracy of processing the data. The aim of using neural networks to simulate the human brain is to create machine learning models that can perform tasks that are difficult for computers to perform using traditional methods [15].

In other words, artificial neural networks can be explained. Although we have talked about how the neural network looks like the human brain, Most of the functions of the human brain can be performed by an artificial neural network. While it does most of the functions, it cannot do all that the brain does, as the human brain is a very complex nervous system. Regarding inputs, neurons can be divided into two categories: single input neurons and multiple input neurons [16]. Similar to the human brain, the artificial neural network does all the computations but not all. These networks do imitate the process by which the brain functions. We have to feed the data into the input. Then the input goes through the input layer and then to the hidden layers, where all the computing is done. This computing is nothing, but the network learns the input again and again and tries to memorize the data, and then it is passed to the output layer. This is where we test this model on sample data to predict or classify the data. This is part of a deep learning technique called an artificial neural network. A survey of a few scientific studies on neural networks demonstrates how effectively and in many areas—including stability, fault detection, prediction, tracking maximum power points, load component and power analysis, energy efficiency, power fluctuations, frequency regulation, and storage energy—neural networks can optimize power systems [23]. After researching artificial neural networks, we can say that, due to only parallel computing, artificial intelligence is becoming more and more necessary as technology advances daily [22].

In an artificial neural network, each neuron has its own weight, and as stated before, these neurons are connected with each other. When the data from the input layer gets transferred to the next layer, then the weight of the initial neuron also gets transferred. The learning, recall, and generalization capabilities of artificial neural networks are what define them. Now we know that the artificial neural network works similarly to our human brain. It also trains similar to our brain. These are also called processing units since they are capable of imitating the brain, and the brain is considered a human's processing unit. The neurons have their own internal state called the activation level. This activation level works similarly to the weights. They are transferred when the connections between neurons are made within different layers.

B. Convolutional Neural Network (CNN)

A convolutional neural network (CNN) is a neural network mostly used in the image processing domain. Which is basically predicting, identifying, or doing anything related to images. It is also used in NLP as well as in deep learning and artificial intelligence. A neural network consists of three layers: an input layer, hidden layers, and an output layer [24]. There can be a single hidden layer or multiple hidden layers. The input layer processes inputs into the model, whereas the hidden layers compute or do the actions inside the model, and the output layer shows the output of the model, which has input from the input layer and calculations from the hidden layers. All these layers have neurons, and every neuron is connected to each other. Every neuron has a weight. Convolutional neural networks operate in a distinct manner because they view data as spatial rather than connected to each other. A convolutional neural network (CNN) is a deep learning algorithm that can process huge data points with less cost. They work by taking an input (2D image), convolving it using filters or a kernel, and then producing output volumes [17].

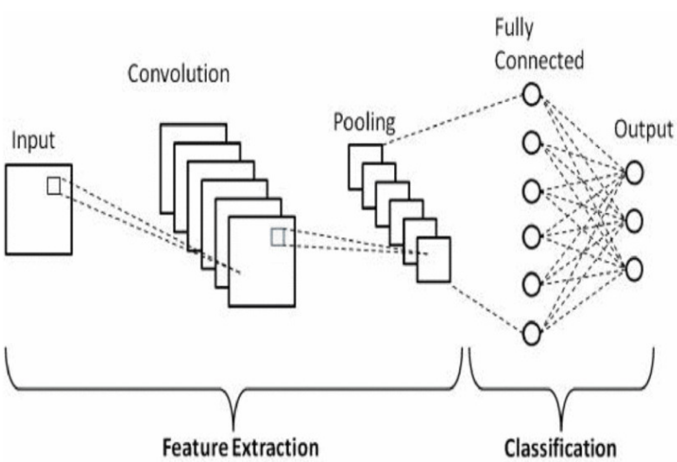

Fig. 5. Convolutional Neural Network [19]

The difference between a simple neural network and a convolutional neural network (CNN) is the number of layers. Although a convolutional neural network is similar to a simple neural network, it has some unique features or unique layers. It has a convolutional layer and a pooling layer. But the rest of the parameters are similar to those of a simple neural network. The relu layer is an activation function. It is important for the data we want to use as input to maintain its originality. convolutional layer and pooling layers are very important. This reduces parameters and processes the output faster. There are two methods for achieving this: average pooling, which just takes the average, and max pooling, which takes the maximum input of a certain convolved feature. In this way, a convolutional neural network builds the image or the output. Certain convolutional neural networks use more complex convolution techniques, like tiled convolution, deconvolution, and dilated convolution, which are based on linear convolution [18]. There are different ways to train such a neural network. For unlabeled data, unsupervised methods are preferred, and for this, autoencoders are used, which helps with the data. Another way is to use GAN. GAN is the Generative Adversarial Network. For this, you train twice. The first training gives you some samples similar to the data. And after the second training, the actual output can be identified between the samples and the model. Recurrent neural networks, as their name suggests, feed data back into themselves, whereas convolutional neural networks filter spatial data as they are fed forward. Sequential data is more suitable for recurrent neural networks.

IV. USE CASE

In this section, we are using deep learning for weather evaluation. For the same, we will be using artificial neural networks (ANN) and convolutional neural networks (CNN) as architectures for evaluation. This dataset [20] consists of different weather metrics for 18 different cities across Europe. This dataset showcases data from the period 2000 to 2010. The task is to evaluate the weather conditions for barbeques in the open. And the data is classified into two categories: true and false. For this, we will consider weather data for Dresden city and try to figure out favorable conditions to barbeque outside in the open in Dresden. This dataset can be found on Kaggle under the name "Weather Prediction." For this, we will compare ANN and CNN and also figure out their accuracy in their predictions and decide which algorithm performs better with respect to this example. Here are some of the snippets of the Python code [21] used for performing deep learning techniques.

```
import numpy as np
import pandas as pd
import seaborn as sns
import matplotlib.pyplot as plt
%matplotlib inline
.
.
.
.
df = pd.read_csv(r"dataset.csv")
df_bbq = pd.read_csv(r"dataset-bbq.csv")
```

With data preprocessing, we first imported all the required libraries, like Numpy, Pandas, Seaborn, and Matplotlib. After this, we imported both the datasets—weather and barbecue—using Pandas. Since it contains 18 cities in Europe, we decided to focus on Dresden. For this, we tailored the dataset to the data from Dresden by trimming all the unnecessary columns. Similarly, we can tailor another dataset for Dresden, which has only true or false classifications.

After this, using Matplotlib and Seaborn, we prepared plots for different weather parameters like wind gust, wind speed, cloud cover, humidity, pressure, global radiation, precipitation, sunshine, minimum temperature, and maximum temperature. This leads us further to the creation of models for artificial neural networks (ANN) and convolutional neural networks (CNN).

Firstly, we will start with the creation of a simple ANN model. For the model building of ANN, we use a sequential model. In total, we use three layers for this model. All three are the dense layers. The first dense layer is using Relu as its activation function. The second dense layer is also using Relu as its activation function. The third and final layer of this model is the dense layer, and its activation function is sigmoid. The loss function we are using here is binary_crossentropy, and the optimizer is set to Adam. Verbose in this model is set to 1, and we are using 200 epochs to train the model. After training, the accuracy of the ANN model comes around at 97.94%.

```
#A very simple ANN model with 2 layers
model = Sequential([
        Dense(X_train.shape[1], activation="relu"
        Dense(X_train.shape[1]/2, activation="rel
        Dense(1, activation = 'sigmoid'),
    ])

model.compile(loss='binary_crossentropy',
optimizer='adam', metrics = ['accuracy'])

model.fit(x=X_train,
          y=y_train,
          epochs=200,
          validation_data=(X_test, y_test),
          verbose=1)
```

Similarly, we constructed a CNN model. For the model building of CNN, we use a sequential model. In total, we have 4 layers for this model. The first layer is the reshape layer, which reshapes the input. The second layer is the 1D convolutional layer, which has Relu as its activation function. Then the third layer is the flatten layer, which flattens the output. And the fourth and final layer is the dense layer, which has sigmoid as its activation function. Similar to the ANN model, we are using binary_crossentropy as the loss function, Adam as the optimizer, Verbose as 1, and we are using 200 epochs to train the CN model. After training, the accuracy of the CNN model comes around at 96.71%.

```
model_2 = Sequential(
    [Reshape((1,X_train.shape[1],1)),
     Conv1D(filters=10, kernel_size=2,
    activation='relu',
    input_shape = (1,X_train.shape[1],1)),
     Flatten(),
     Dense(1, activation = 'sigmoid')
                  ])

model_2.compile(loss='binary_crossentropy',
optimizer='adam',metrics=['accuracy'])

model_2.fit(x=X_train,
          y=y_train,
          epochs=200,
          validation_data=(X_test, y_test),
          verbose=1)
```

After comparing ANN and CNN, we can say that in this specific example, ANN is more accurate than CNN. This also establishes the fact that CNN can be more productive and effective in terms of image processing, while ANN is better suited for other forms of data. Although this cannot be a general statement, Rather, it is specific to this example of weather evaluation using deep learning.

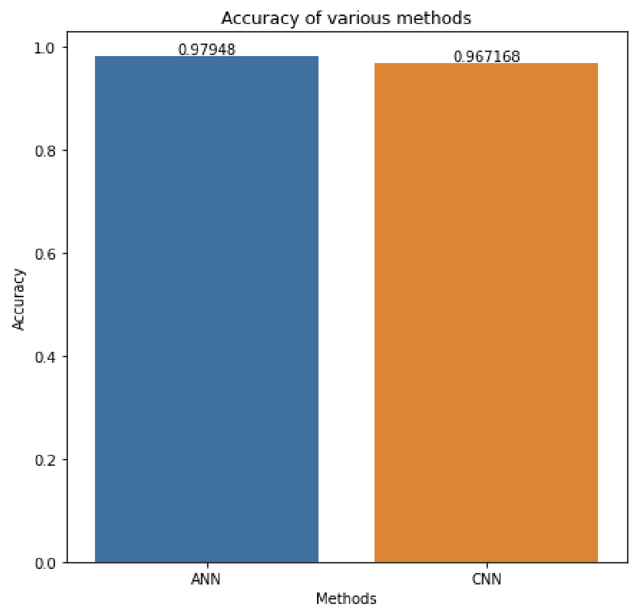

Fig. 6. Accuracy of various methods

V. CONCLUSION

In this paper, we have seen how the weather evaluation is done using deep learning. For this, we used two different algorithms for predicting or evaluating weather: ANN and CNN. In the use case, we tried to figure out whether the weather is favorable to barbeque outdoors. We used Dresden's data for such an evaluation. We compared the accuracies of these two different techniques and figured out the more accurate technique for the above use case. More and more models like this need to be developed to improve their accuracy and predict weather more accurately. As weather impacts most of our lives, we need to develop and improve the existing models as technology advances.

VI. FUTURE WORK

From the above example, it is evident that deep learning can be used to predict or evaluate weather. With respect to the above use case, the model can be optimized further with more data points. The current model is trained on a few data points compared to the scale of data used for weather evaluations in real time. This affects the accuracy and predictions as well. Therefore, more data points may solve the issue and will be comparable with real-time weather evaluations. With respect to research in the deep learning domain, it is increasing day by day as time and technology progress. Newer, more complex, and more accurate models for such evaluation in real-time weather are being optimized.

References

[1] F. Ahmad, M. Tarik, M. Ahmad and M. Z. Ansari, "Weather Forecasting Using Deep Learning Algorithms," 2023 International Conference on Recent Advances in Electrical, Electronics Digital Healthcare Technologies (REEDCON), New Delhi, India, 2023, pp. 498-502, doi: 10.1109/REEDCON57544.2023.10150439.

[2] P. S. Mung and S. Phyu, "Time Series Weather Data Forecasting Using Deep Learning," 2023 IEEE Conference on Computer Applications (ICCA), Yangon, Myanmar, 2023, pp. 254-259, doi: 10.1109/ICCA51723.2023.10182058.

[3] A. G. Salman, B. Kanigoro and Y. Heryadi, "Weather forecasting using deep learning techniques," 2015 International Conference on Advanced Computer Science and Information Systems (ICACSIS), Depok, Indonesia, 2015, pp. 281-285, doi: 10.1109/ICACSIS.2015.7415154.

[4] A. G. Salman, B. Kanigoro and Y. Heryadi, "Weather forecasting using deep learning techniques," 2015 International Conference on Advanced Computer Science and Information Systems (ICACSIS), Depok, Indonesia, 2015, pp. 281-285, doi: 10.1109/ICACSIS.2015.7415154.

[5] Amaratunga, T. (2021). What Is Deep Learning?. In: Deep Learning on Windows. Apress, Berkeley, CA. https://doi.org/10.1007/978-1-4842-6431-7_1.

[6] P. S. Mung and S. Phyu, "Time Series Weather Data Forecasting Using Deep Learning," 2023 IEEE Conference on Computer Applications (ICCA), Yangon, Myanmar, 2023, pp. 254-259, doi: 10.1109/ICCA51723.2023.10182058.

[7] 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), Noida, India, 2022, pp. 422-427, doi: 10.1109/Confluence52989.2022.9734133.

[8] P. S. Mung and S. Phyu, "Time Series Weather Data Forecasting Using Deep Learning," 2023 IEEE Conference on Computer Applications (ICCA), Yangon, Myanmar, 2023, pp. 254-259, doi: 10.1109/ICCA51723.2023.10182058.

[9] https://www.zdnet.com/article/fair-forecast-how-180-meteorologists-are-delivering-good-enough-weather-data/.

[10] https://environment.uw.edu/news/2020/12/a-i-model-shows-promise-to-generate-faster-more-accurate-weather-forecasts/.

[11] https://environment.uw.edu/news/2020/12/a-i-model-shows-promise-to-generate-faster-more-accurate-weather-forecasts/.

[12] https://de.mathworks.com/discovery/deep-learning.html

[13] F. Q. Lauzon, "An introduction to deep learning," 2012 11th International Conference on Information Science, Signal Processing and their Applications (ISSPA), Montreal, QC, Canada, 2012, pp. 1438-1439, doi: 10.1109/ISSPA.2012.6310529.

[14] https://www.digitaltrends.com/computing/what-is-an-artificial-neural-network/.

[15] N. Saha, A. Swetapadma and M. Mondal, "A Brief Review on Artificial Neural Network: Network Structures and Applications," 2023 9th International Conference on Advanced Computing and Communication Systems (ICACCS), Coimbatore, India, 2023, pp. 1974-1979, doi: 10.1109/ICACCS57279.2023.10112753.

[16] N. Saha, A. Swetapadma and M. Mondal, "A Brief Review on Artificial Neural Network: Network Structures and Applications," 2023 9th International Conference on Advanced Computing and Communication Systems (ICACCS), Coimbatore, India, 2023, pp. 1974-1979, doi: 10.1109/ICACCS57279.2023.10112753.

[17] R. Chauhan, K. K. Ghanshala and R. C. Joshi, "Convolutional Neural Network (CNN) for Image Detection and Recognition," 2018 First International Conference on Secure Cyber Computing and Communication (ICSCCC), Jalandhar, India, 2018, pp. 278-282, doi: 10.1109/ICSCCC.2018.8703316.

[18] D. Dai, "An Introduction of CNN: Models and Training on Neural Network Models," 2021 International Conference on Big Data, Artificial Intelligence and Risk Management (ICBAR), Shanghai, China, 2021, pp. 135-138, doi: 10.1109/ICBAR55169.2021.00037.

[19] D. Dai, "An Introduction of CNN: Models and Training on Neural Network Models," 2021 International Conference on Big Data, Artificial Intelligence and Risk Management (ICBAR), Shanghai, China, 2021, pp. 135-138, doi: 10.1109/ICBAR55169.2021.00037.

[20] https://www.kaggle.com/datasets/thedevastator/weather-prediction

[21] https://www.kaggle.com/code/lorresprz/ann-cnn-randomforest-predicting-nice-weather/notebook

[22] M. Mishra and M. Srivastava, "A view of Artificial Neural Network," 2014 International Conference on Advances in Engineering Technology Research (ICAETR - 2014), Unnao, India, 2014, pp. 1-3, doi: 10.1109/ICAETR.2014.7012785.

[23] M. A. Obeidat, A. M. Mansour, B. Al Omaireen, J. Abdallah, F. Khazalah and M. Alaqtash, "A Deep Review and Analysis of Artificial Neural Network Use in Power Application with Further Recommendation and Future Direction," 2021 12th International Renewable Engineering Conference (IREC), Amman, Jordan, 2021, pp. 1-5, doi: 10.1109/IREC51415.2021.9427846.

[24] Rahul Ranjan, Richard Partl, Ricarda Erhart, Nithin Kurup, Harald Schnidar medRxiv 2021.09.24.21264011; doi: https://doi.org/10.1101/2021.09.24.21264011