Drought prediction using artificial neural network

Pardha Saradhi Metta
Department of computer science and engineering Vellore institute of technology
Vellore, Tamil Nadu, India
saradhi1996@gmail.com

Anish Chintamaneni
Department of computer science and engineering Karunya institute of technology and Sciences Coimbatore,
Tamil Nadu, India
canish99@gmail.com

Ashish Kumar
Department of Information Technology Netaji Subhas University of Technology
Dwarka, Delhi, India
ksndelash@gmail.com

Jatin Yadav
Department of Instrumentation and Technology Karunya institute of technology and Sciences Coimbatore,
Tamil Nadu, India
yadavjatin77@gmail.com

Vishesh Kumar
Department of computer science and Engineering SRM Institute of Science and Technology Kattankulathur, Tamil Nadu, India
mail4vishesh@gmail.com

Bharath Bhaskar
Department of Electronics and Communications PES University
Bangalore, Karnataka, India
bharathbr12@gmail.com

2022 2nd International Conference on Advance Computing and Innovative Technologies in Engineering (ICACITE) | 978-1-6654-3789-9/22/$31.00 ©2022 IEEE | DOI: 10.1109/ICACITE53722.2022.9823838

Abstract—Drought has been one of the most frightening natural calamity that is being faced by the humans, and it is generally faced by the farmers all around the world. It can occur anywhere in the world and is also known as one of the slow disaster as it stays for a very longer time period and maybe even more if it wants to be harsher. Drought not only impacts the human life but also the agriculture, world economies and energy that had been infiltrated by the farmers. Everything goes on stakes during this calamity and the basic needs such as food are not easy to get during these times and the inflammation is at peak due to supply and demand instability. There are many ways which can be done to avoid the drought, which include drip irrigation to save water during the times of drought, rainwater harvesting can also be a saviour, crop engineering, desalination of water and many more things. The main solution for this would be to analyse the weather and the future outcomes of it which would help in preparing for the worse situation and also the soil prediction can help a lot when it comes to predicting this situation. The paper aims to take the data of weather and soil together to predict how droughts can be avoided. Deep learning techniques will help to get a great idea if the drought is going to happen or not.

Index Terms—Drought prediction, artificial neural network, Deep Learning, predictive modelling, neural networks

I. Introduction

It is observed as challenges of agricultural systems that how to alleviate the effects of drought. Drought impacts farming systems both economically and environmentally[3]. In terms of economic impact, drought damages effect the production of agriculture and due to this, economic damage is observed in the field of industries elated to agriculture, in addition, it effect the employment of the people and also reduce the production. From an environmental perspective, drought reduce the crop production and also the quality of the soil to introduce the slat content in large quantity. It also affects the irrigation system of the soil. To see this type of environment, prevention of water resources should be first step and planning of water resources will be the part of it. This will help to reduce the drought effect in the future in specific location of the country [2]. The use of drought indicators allows for the evaluation and forecasting of drought. The Palmer Index, the Crop Moisture Index, and the Standardized Precipitation Index are some of the most widely used drought indexes (SPI). Because of its uniformity, the Palmer Index and the SPI have long been the most used indexes for drought predictions[1].

II. Related Work

Before applying any algorithm our first objective is to work on the datasets to collect and to work on them.When deep learning is used it is generally used with the unsupervised techniques for the neural networks and it can solve the problems that are faced very easily, outperforming traditional machine learning techniques by avoiding becoming caught in the poor solutions. Using numerous hierarchical layers, the Deep Learning technique may extract data characteristics from raw data[9]. Its capacity which is mostly underestimated and is generally used in the learning process of the experience, helping in the development of models at multiple sample locations. DL-based techniques have been shown in a few studies to extract more usable information that can be utilised in the problem that is being faced and also included in the model that has been used by the author in finding the solution for this thing[6]. The results show that DL-based algorithms outperform multi-layered perceptron and SVR in areas of prediction accuracy. While working on any dataset, it is very important to have relevant information about that particular topic on which with are working. Here in terms of drought with should have relevant information on the surface area, the unit of area, etc. Based on the Conv1D model, the new DL model is proposed[4]. The Keras library, which is built on top of TensorFlow, is used to create. An input layer, an output layer, and an infinite number of hidden layers are often found in a CNN design. The described DL model is of the sequential kind, and it may be used to solve modelling issues with one or more time series[7]. The input layer's job is to take in the signal (input data) and send it to the hidden layer (s). The model's computation engine is hidden layers. Depending on the situation, they may have one or more Conv1D layers, max-pooling layers, dropout layers, and flatten layers. CNN's core part is the Conv1D layer. It is made up of one-dimensional filters/kernels that extract features from the input signal, a kernel size that specifies the length of the filter, and an activation function that determines the neurons' threshold limit. The model is trained on the causal and target elements and extracts the sequence's hidden information[8].The dropout layer's purpose is to assign zero values to the network's neurons at random, making the network less sensitive to tiny variations

and thus enhancing the model's accuracy on unknown input. The data from the convolutional/pooling/dropout layers is transferred to the output layer through a flatten layer, which turns the output into one dimension. The output layer is a densely linked layer that is connected to all of the neurons in the previous layer and is in charge of producing the model's output. To correct the model design, numerous hyper-parameters have to be defined[10]. [1]According to a study SVM (Support Vector Machine), the method can also be used in the classification of drought problems. An SVM model is a description of the cases as factors in space, drawn so that the cases of the different categories are divided by as broad a gap as possible. New samples are then mapped into the same region and allocated to a category solely depending on which part of the space they fall into. According to research by Anteneh Belayneh and Jan Adamowski algorithms like Support Vector Regression (SVR), Artificial Neural Network (ANN), and coupled wavelet-ANN can also be used based on the Standard Precipitation Index (SPI). [11][3]

III. Methodology

A. Data preparation

The dataset contains 20000 images of different surroundings, the data has been converted into the csv format for the model to read and perform on it. The dataset has been acquired from Nasa sat4 images which consist of 28 * 28 * 4 and were taken from space. There are different classes that can be seen in the figure 1 which are water, barren land and tree. A smaller test set has been used as the dataset which has been selected is very large[14].

B. Data Preprocessing:

Data is very important in this research problem, without data prediction, approximation or any operations are not possible. The categorization of the data is also play an important role. Therefore, the preparation of dataset is little bit difficult task specially in this problem. The classes of the pictures are of water, barren land and tree which can be shown in figure 1[12]. To make the images clear one has to take care of the background as the features would be detected on the main image or the main part which is focused in the image. Consider as example, the background removal is undertaken for all the images of water, barren land and tree class. This is undertaken by the authors to make the dataset clean and without any error. Each row in the file looks like this [1,0,0,0], where only one of the 4 integers is 1. If it is one, then it is that class related to the order I mentioned above. If it had the aforementioned values, the image is a picture of desolate terrain. If it was [0,1,0,0], then it would be forest land. If it was [0,0,1,0], then it would be grassland and so forth[13]. In the last step which is very critical step and it divide the dataset into train test split dataset. This step is important because of dependency in the performance of t Finally, the last step would be the most critical step to divide the dataset into train test split dataset which is important of the model because of depnedancy on the modal performance. It is necessary to provide a unknown dataset to the model and see that it perform the same level as it performed on the training dataset [15].

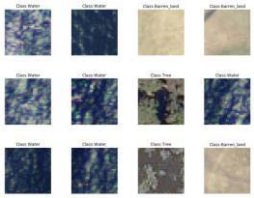

Fig. 1. Dataset overview

C. Feature Visualization

Feature visualization is very important in the terms of dataset. Because if the feature is not collected or extracted well then the feature will not perform well. For the performance of the model, it is very important that features are visualize. It is responsible for the output of the model or prediction of the model. The correlation matrix is the matrix which shows the relationship between sets and variables. This correlation matrix allows us to visualize the relation heights pair which having the highest correlation. The correlation matrix can be visualizing in fig.3. In the next, seaborn pairplot plotted which is basically a python library to visualize data which is based on Matplotlib[16]. This function shows a very attractive and informative picturization of statistical graphs. The function pairplot represent the relationship in database in graphical way using X-axis and Y-axis. This method complete the requirement of matplotlib such as Default parameters and while working with data frames. In the last, the importance of feature is performed to get to know that which feature is important or play a vital role to determine the output of the model [17].

D. Model Architecture

The main motive or the function of the model is do a categorization which is generally based on the target of the characteristics which is potrayed by the techniques such as image processing and extraction of the features. After that it is followed by the classification of the images which a very useful technique. Based on the application, score is compared to either merely a threshold that determines whether an item exists or not (for example, the incidence of a plant or animal in the shot), or it is compared to additional scores to separate groups of things.The artificial neural network (ANN) technique converts the human brain's NN (neural network) structure into a mathematical model[18]. It creates a simulated NN (neural network) framework by arranging a large number of neurons and connecting them. It derives the weight parameters of each neuron in the network via data set training. ANN may be divided into two phases: training learning and recognition categorization, depending on the technique used to create the network model . First and foremost, our method improves and adjusts network parameters to make picture identification and categorization faster and more accurate. The most extensively utilized ANN models these days are back propagation (BP) networks and radial basis function (RBF) neural networks.In the area of AI,ANN is generally used to create the existing network of the neurons that are an important component of the human brain and because of that the computers can utilize it and choose just like the humans. Just like us the computers are trained to operate like they are the brain cells of any operation like the human brain is the main part of any operation, the artificial neural network is worked in the similar manner. Just as the information is stored in the human brain the information is also stored in the neural network and which is utilized to do computations and give a proper output with the help of the neurons.

IV. Experimental Results

The model finally succeeded in generating the best results needed by the authors in properly executing the detection strategy and striking a big blow to the agricultural industry. After running the model with appropriate parameters of the artificial neural network that the accuracy received by the authors is 99 percent and the precision and recall can be seen in the figure 4.

V. Conclusion

The agricultural sector is one of the main industries on which people depend on and if it doesn't perform properly due to causes like drought then it may bring many difficulties on the human lives. To overcome this issue, the authors have devised a drought prediction system seeking to anticipate whether the drought is likely to occur and if it may cause the problem or not. They have achieved an accuracy of 99 percent which is utilized for the identification of the drought.

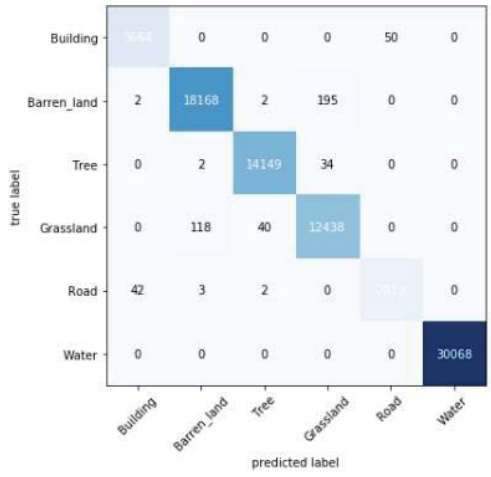

Fig. 2. Confusion Matrix

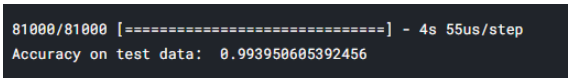

Fig. 3. Accuracy

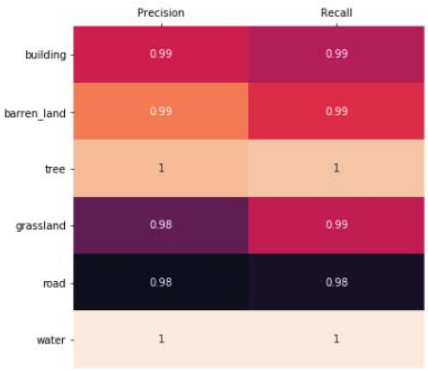

Fig. 4. Precision and recall

Future Work

There is a lot of work to be done further which includes using or integration of iot with the current model, so it can capture the images and also provide other necessary details which can be included to improve the performance of the model and the issue can be solved very easily. The next thing that can be done is to use a different dataset and test out the current model and also use different ensemble techniques to predict and get a better accuracy.

References

[1] M. I. K. S. S. R. D. S. S. M. M. P. a. K. C. Rajib Maity, "Potential of Deep Learning in drought assessment by extracting information from hydrometeorological precursors".

[2] D. K. K. B. AISHWARYA M IYENGAR, "DROUGHT PREDICTION USING MACHINE LEARNING ALGORITHM".

[3] J. F. A. Anteneh Belayneh, "Drought forecasting using new machine learning methods".

[4] G. Berhan, S. Hill, T. Tadesse and S. Atnafu, "Drought Prediction System for Improved Climate Change Mitigation," in IEEE Transactions on Geoscience and Remote Sensing, vol. 52, no. 7, pp. 4032-4037, July 2014, doi: 10.1109/TGRS.2013.2279020.

[5] Z. Gong, R. Sun and Lianshui Li, "Risk prediction of agricultural drought in China," 2010 Seventh International Conference on Fuzzy Systems and Knowledge Discovery, 2010, pp. 927-930, doi: 10.1109/FSKD.2010.5569129.

[6] Y. Liu and L. Wang, "Drought Prediction Method Based on an Improved CEEMDAN-QR-BL Model," in IEEE Access, vol. 9, pp. 6050-6062, 2021, doi: 10.1109/ACCESS.2020.3048745.

[7] J. Wang, Z. Xia, L. Guo and D. Liang, "Recognizing and forecasting the hydrologic drought in the upper of Weihe basin," 2011 International Conference on Electrical and Control Engineering, 2011, pp. 1677-1680, doi: 10.1109/ICECENG.2011.6058279.

[8] N. A. Agana and A. Homaifar, "A deep learning based approach for long-term drought prediction," SoutheastCon 2017, 2017, pp. 1-8, doi: 10.1109/SECON.2017.7925314.

[9] W. Wang, L. Sun, Z. Pei, Y. Chen and X. Zhang, "Analysis of Temporal and Spatial Variation of Growing Season Drought in Jiling Province Based on Standardized Precipitation Evapotranspiration Index," 2019 8th International Conference on Agro-Geoinformatics (Agro-Geoinformatics), 2019, pp. 1-5, doi: 10.1109/Agro-Geoinformatics.2019.8820436.

[10] M. Kansara, P. Maity, H. Malgaonkar and A. Save, "A Novel Approach for Early Prediction of Drought," 2020 6th International Conference on Advanced Computing and Communication Systems (ICACCS), 2020, pp. 768-773, doi: 10.1109/ICACCS48705.2020.9074405.

[11] R. Inoubli, A. B. Abbes, I. R. Farah, V. Singh, T. Tadesse and M. T. Sattari, "A review of drought monitoring using remote sensing and data mining methods," 2020 5th International Conference on Advanced Technologies for Signal and Image Processing (ATSIP), 2020, pp. 1-6, doi: 10.1109/ATSIP49331.2020.9231697.

[12] X. Liu, L. Ren, F. Yuan and B. Yang, "Meteorological Drought Forecasting Using Markov Chain Model," 2009 International Conference on Environmental Science and Information Application Technology, 2009, pp. 23-26, doi: 10.1109/ESIAT.2009.19.

[13] A. Mohammadinezhad and M. Jalili, "Optimization of echo state networks for drought prediction based on remote sensing data," 2013 IEEE 8th Conference on Industrial Electronics and Applications (ICIEA), 2013, pp. 126-130, doi: 10.1109/ICIEA.2013.6566352.

[14] G. Berhan, S. Atnafu, T. Tadesse and S. Hill, "Drought information mining from satellite images for improved climate change mitigation," 2012 World Congress on Information and Communication Technologies, 2012, pp. 197-202, doi: 10.1109/WICT.2012.6409074.

[15] V. Shah and S. Modi, "Comparative Analysis of Psychometric Prediction System," 2021 Smart Technologies, Communication and Robotics (STCR), 2021, pp. 1-5, doi: 10.1109/STCR51658.2021.9588950.

[16] S. Chaudhari, V. Sardar, D. S. Rahul, M. Chandan, M. S. Shivakale and K. R. Harini, "Performance Analysis of CNN, AlexNet and VGGNet Models for Drought Prediction using Satellite Images," 2021 Asian Conference on Innovation in Technology (ASIANCON), 2021, pp. 1- 6, doi: 10.1109/ASIANCON51346.2021.9545068.

[17] A. Kaur and S. K. Sood, "Cloud-Centric IoT-Based Green Framework for Smart Drought Prediction," in IEEE Internet of Things Journal, vol. 7, no. 2, pp. 1111-1121, Feb. 2020, doi: 10.1109/JIOT.2019.2951610. bibitemb18S. Modi and M. H. Bohara, "Facial Emotion Recognition using Convolution Neural Network," 2021 5th International Conference on Intelligent Computing and Control Systems (ICICCS), 2021, pp. 1339-1344, doi: 10.1109/ICICCS51141.2021.9432156.