International Journal of Innovative Technology and Exploring Engineering (IJITEE)
ISSN: 2278-3075 (Online), Volume-9 Issue-6, April 2020

Chennai Hydrological Drought Prediction

Priyansha T Jat, Nikhil Nautiyal, M. Pushpalatha

***Abstract*: Chennai also referred to as Madras is located on the Coromandel Coast off the Bay of Bengal. It's the largest cultural, economic and academic center of south India. The city faced an acute water shortage in June 2019. Chennai was entirely captivated with spring water resources to fulfill its water needs. There are four reservoirs within the city, namely, Red Hills, Cholavaram, Poondi and Chembarambakkam, with a combined capacity of 11,057 mcft. These are the main sources of water for the town.**
Apart from the reservoirs, alternate sources of fresh water are Minjur and Nemelli desalination plants; Cauvery water from Veeranam lake; aquifers in Minjur, Panchetty and Neyveli; Here is an initiative to place together a dataset that has the data about the varied water sources available within the city using LSTM (Long Short Term Memory), TensorFlow, NumPy, Keras and ANN (Artificial Neural Network).
The idea is to figure out whether we can use this dataset to:
1. Visualize the water need / usage of the town
2. Identify whether the water sources availability is going to be able to meet the wants till the following monsoon?
3. How bad is the current water crisis compared to previous years?*

Keywords* : *ANN, Keras, LSTM, NumPy, TensorFlow.

I. INTRODUCTION

With 140 cm of annual rainfall, Chennai has no business to ever get into a drought situation. But it did; for over 170 straight days without a drop of rain, the town has slipped into such a horrendous water crisis that even the international media, notably CNN (Cable News Network) and also the New York Times, ran stories on that. Hollywood star Leonardo Di Caprio highlighted the matter to his 32 million Instagram followers through a post in June.

Making weather forecast is domineering and demands operational responsibilities administrated by meteorological services globally the world. The goal is complicated because within the meteorology arena, all verdicts are supposed to be made within the visage of uncertainty. Different scientists over the world have developed stochastic weather models which are supported by random number of generators whose output resembles the weather data to which they have been fit. Rainfall is a non-linear phenomenon and requires a non-statistical and complex method for its prediction. One such method is using an artificial neural network.

Revised Manuscript Received on April 30, 2020.
***** Correspondence Author
Priyansha T Jat*, Computer Science and Engineering, SRM IST KTR, Tamil Nadu India. Email: priyanshatjat@gmail.com
Nikhil Nautiyal, Computer Science and Engineering, SRM IST KTR, Tamil Nadu India. Email: nnautiyal9@gmail.com
Dr. M. Pushpalatha, Computer Science and Engineering, SRM IST KTR, Tamil Nadu India. Email: pushpalatha.m@ktr.srmuniv.ac.in

This is because ANN (Artificial Neural Network) model relies on 'prediction' by smartly 'analyzing' the pattern from a persisting historical set of information with myriad number of data entries. the remaining ANN, other than ANN the remaining models are either statistical or mathematical. These models are proved to be highly accurate in calculating, but not as much in predicting as they can't adapt to the irregularly varying trends of information which might have neither been deduced from a formula nor written in type of a function. Such real-life scenarios are well understood by 'artificial neurons' as they acquire their learnings from practice; back-propagating errors in next deduction then on, to be precise. This might compromise the accuracy but give us a stronger advantage in 'understanding the problem', concluding from it or duplicating it.

Amongst various weather changes, rainfall has foremost significant role in our day to day life. Human civilization amply depends upon its frequency and amount to varied scales. Several stochastic models are attempted to predict the rainfall timing, to analyze its seasonal inconsistency, to forecast monthly/yearly rainfall over a particular geographic region. This paper aims to build an ANN model which can predict an average rainfall in a month within Chennai, Tamil Nadu.

Economy of India highly depends on summer monsoon. Therefore, predicting rainfall can prove to be a challenging task for atmospheric scientists of India. Back-propagating ANN used to predict typical rainfall in monsoon in summer over Chennai city. The catch here lies within the incontrovertible detail that several MANN (My Artificial Neural Network) models are endeavored to seek for most effective mapping.

II. LITERATURE SURVEY

Various assessments on rate of rainfall estimate with use of a fake neural-framework have fundamentally been established on without fail, month to month and yearly precipitation desires with higher importance on water and meteorology, the officials French et al., made use of a fake neural-framework to check precipitation in two dimensions, an hour in advance. Despite the way this was not extraordinarily productive, the work extended an appreciation in using fake neural-frameworks (ANNs) in inspecting composite geophysical methods.

In year 1995, Michaelides et al utilized a neural system for estimating missing precipitation information over the Cyprus. Expectations were made by means of day by day precipitation perceptions in the surrounding localities. Christodoulou et al. applied a counterfeit neural system to further foresee precipitation rate, from climate radar information, in some other investigation.

Two AI classifiers, the factual K-Nearest Neighbor (KNN) classifier with radar information and a self-organizing map (SOM) were utilized as contribution to neural system and downpour check estimates as yields while preparing. Normal mistake rate for anticipated precipitation rate come out to be 23 %.

From their examinations, Nayak et al., affirmed that the back-propagation arrange, notwithstanding different sorts of neural systems, can be depended upon in precipitation forecast with better outcomes contrasted with measurable and numerical techniques.

III. NEED FOR THIS MODEL AND CURRENT SCENARIO

As indicated by the review report of the Comptroller and Auditor General of India, 2017, the developed zone in the Chennai metropolitan territory has developed from 90.88 sq.km in 1979, to 541.14 sq.km in 2016. Then again, the water spread territory of lakes has contracted to 91.31 sq.km from 100.98 sq.km in the previous 40 years.

In each satellite picture beneath, the zone set apart in red focuses to the degree of that lake in 1973 and the water body underneath delineates the real lake degree starting at 2018. The streets that confuse and the structures worked over the red region were once a part of the lake which were later infringed upon.

Fig. 1 Satellite Image 1

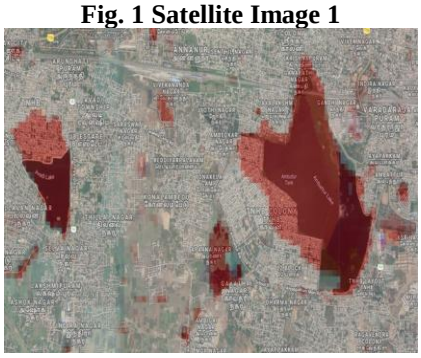

Fig. 2 Satellite Image 2

IV. LSTM

No one will begin their speculation without consistent preparation. In this report, we see every word depending on our comprehension of previously used words. Everything can't be discarded and thought with no preparation yet again. Ones musings has fidelity. Conventional neural systems cannot try this, and it proves to be a big insufficiency. Let us picture a scenario, envision if we would like to rearrange what kind of occasion goes on at every point of a motion-picture. It is unconventional, the way a neural system could utilize its thought process over previous occasions within the film to teach ones that appear later.

Intermittent neural systems have the ability of addressing this kind of issues. These systems have loops within them which enables the data to continue.

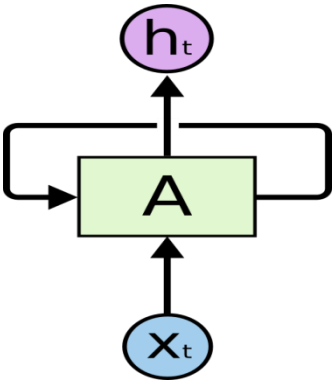

Fig. 3 Recurrent Neural Networks have loops.

Long short-term Memory System; also known as LSTM's are a sole sort of RNN, suitable to adapt ordeal conditions. These were first presented by Schmidhuber and Hochreiter (1997) and were polished and encouraged by several individuals in following work. They work enormously well on a large assortment of issues and are generally utilized currently.

Ideally, LSTMs should be far away from long-haul reliance issue. To recollect data for extensive stretches of our time is their default conduct for all purposes, not something they battle to learns.

All repetitive neural systems have the type of a sequence of neural systems with rehashing modules. For standard RNNs, (Recurrent Neural Networks) rehashing module will have a elementary structure, as an example, a sole tanh layer.

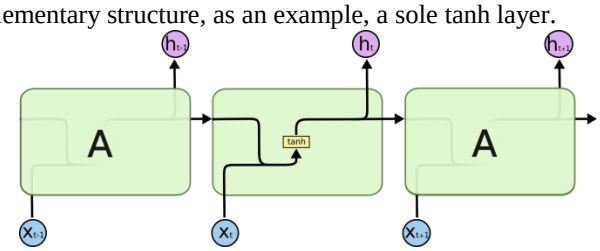

Fig. 4 There is a single layer in Repeating module in a standard RNN

LSTMs even have a chain structure, but there will be a different structure for repeating module rather than having one neural network layer, there are four, interacting in a very special way. The repeating module in an LSTM contains four interacting layers.

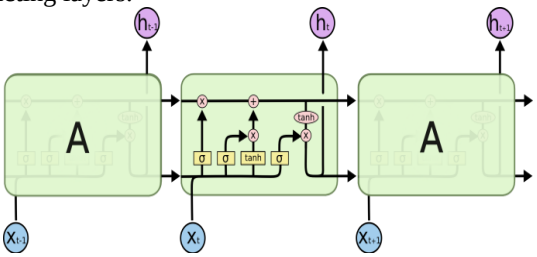

Fig. 5 Four interacting layers in repeating module in an LSTM

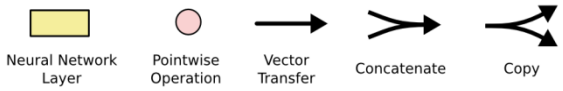

Fig. 6 Block Diagram

In the above diagram, there is a complete vector in every line, from one node's output to others' inputs. Pink circle represents pointwise operations, such as vector addition, whereas learned neural network layers are represented by yellow colored boxes. Merging lines denote replication of its content and its copies visiting several different locations.

V. TENSORFLOW

Google's TensorFlow is the most acclaimed profound learning library on the planet right now. Google item uses AI in most of its items to improve the web crawler, interpretation, picture subtitling or proposals.

To deliver a solid model, Google clients can encounter a quicker and progressively refined hunt with use of AI (Artificial Intelligence). On the off chance that the client types a watchword the pursuit bar, Google gives a proposal as to what the following word could be.

The plan was to build it such that it could run on multiple GPUs (Graphic Processing Units) or CPUs (Central Processing Units) and mobile operating systems. It also has numerous wrappers in various languages such as Java, Python or C++.

Table-I: List of Algorithms

	TensorFlow supports and has a built-in API for the following algorithms:
1.	Linear Regression: tf.estimator.LinearRegressor
2.	Classification: tf.estimator.LinearClassifier
3.	Deep Learning Classification: tf.estimator.DNNLinearClassfier
4.	Deep Learning wipe and deep: tf.estimator.DNNLinearCombinedClassifier
5.	Booster Tree Regression: tf.estimator.BoostedTreesRegressor
6.	Boosted Tree Classification: tf.estimator.BoostedTreesClassifier

VI. NUMPY

CPython reference usage of Python is a non-streamlining bytecode mediator and NumPy mainly focuses on it. Numerical calculations made for this rendition of Python recurrently run rather slowly as compared to accumulated reciprocals.

NumPy gives multidimensional clusters and effectively working administrators that work on exhibits and requires reworking some code, which is mostly inward circles using NumPy. This is how it can address gradualness issue halfway through.

Python uses NumPy exhibits to work on information and store data. Images that have several different channels are essentially referred to as 3D clusters and that is why veiling, ordering or cutting using dissimilar exhibits are extremely skillful approaches to obtain explicit pixels of an image. The NumPy cluster unravels the process of programming and troubleshooting as it is a comprehensive structure of OpenCV.

VII. KERAS

Keras is designed to empower fast experimenting with profound neural systems. It is an open-source neural-organize library which is written in Python. It is suitable for running on Theano, TensorFlow PlaidML, Microsoft, or Cognitive Toolkit. It centers around being user friendly, extensible, and measured.

Keras simplifies the process of working with image and content data. It contains several uses of normally utilized neural-arrange building squares, for example, layers, destinations, actuation capacities, and a huge collection of devices.

Keras does not withstand typical neural systems, it supports convolutional and intermittent neural systems. It reinforces other basic utility layers like pooling, dropout, and cluster standardization.

Clients use Keras to convert deep models into products for cell phones, both iOS and android and on web or JVM (Java Virtual Machine). Similarly, it enables utilization of disseminated preparing deep learning models on groups of GPU and THU (Tensor Handling Units).

VIII. NEURAL-NETWORKS

Neural system is a series of iterations and calculations that attempts to connect basic information through a procedure that copies the way, a human mind works. Right now, systems suggest to frameworks of neurons, either counterfeit or natural. Neural systems have the ability of adjusting themselves according to evolving data input; this is how the system comes up with the most suitable result without having to upgrade the criteria of yield.

Artificial Neural networks work on the rule of gaining from a preparation set. There are an assortment of neural system models and learning methodology. Two classes of neural systems that are generally utilized for forecast applications are recurrent networks and feed-forward networks. Both systems are trained using back-propagation algorithm. At the point when this algorithm is utilized for weight change, the condition of the framework is doing gradient descent; moving toward the path opposite to the largest local slope on the execution surface. As it were, the loads are being refreshed the descending way.

Back propagation algorithm, is one of the most predominant of all the supervised learning models in ANN. This algorithm addresses the weight of the interconnectivity neuron by utilizing steepest gradient descent strategy. Gradient descent strategy is used to calculate the network's weight and regulate the weight of interconnections in order to lessen the error in output. Error function at the output neuron is as follows:

$$E = \frac{1}{2}\sum_k (T_k - A_k)^2 \quad (1)$$

Here, A_k and T_k represent the predicted and actual values of output neuron, where k is the output neuron.

Resultant weight vector of the trained network represents its learning regarding the matter and is employed to use to a brand-new set of data to gauge the performance of the model. During this research, for all applied models back-propagation algorithm term has been used.

Network design Prediction networks generally take measured historical data, and future condition is simulated after some processing stages. During this research after rigorous testing of various ANN structures, RN (Recurrent network) and TLRN networks were selected in accordance with their higher performance, so amongst these two, TLRN proved to better as it had higher abilities. Hence, TLRN was preferred ANN type for drought prediction during this project. These two networks are briefly introduced within the below. - Recurrent Networks (RN) this sort of network may be categorized into partially and fully recurrent. This is different from an MLP as it has a memory element.

Recurrent networks are generally stronger but they're also tougher to train, and it is comparatively more complex to understand their properties. To build the simplest architecture of a network for this study, several different structures were tested. Several parameters like number of hidden layers and processing elements in them, output functions and kind of transfer along with sort of learning rule and its parameters are evaluated and considered. After several iterations of differing types of output and transfer functions for output and hidden layers, it could be concluded that a tangent hyperbolic function will be foremost suitable one for hidden layer. On the other hand, for output layer, the sigmoid function may be a more compatible.

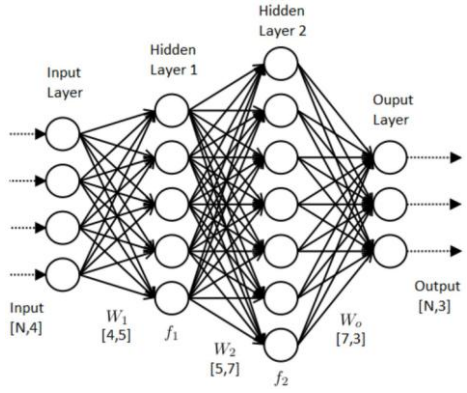

Fig. 7 Typical Recurrent Network with unit memory in hidden layers.

IX. RESULT AND DISCUSSION

LSTM turns out to be most accurate and efficient algorithm with an accuracy of about 94%. It took a total of 167 second to train the model in first epoch with initial accuracy of 86%. As the model was trained with repeated epochs, accuracy gradually increased. Total loss turned out to be 0.0160.

Efficiency and the performance of the device can be increased if a rather powerful GPU unit is used instead of current default GPU.

A fan can be used as a cooling solution for the device as the thermal throttling of the device leads to performance loss.

The novelty of this attempt lies in it the size of dataset and speed as there is limited work in past which uses LSTM for such bulky data. Several additional variables were included and factored in predicting the results.

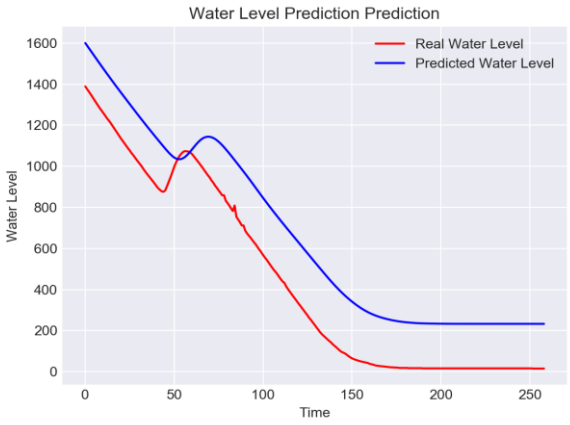

Fig. 8 Resultant graph showing actual and predicted values

X axis: Time (in days)

Y axis: Water level (mcft)

X. CONCLUSION

This study is aimed to make people more aware of the crisis and help them prepare for such event in the future.

This study will create awareness among citizen and government bodies about scarcity of water in future. Currently it only considers 4 major sources of water in Chennai. It is planned to add more sources and gather more in-depth knowledge about factors affecting water shortage and how much they contribute to it.

REFERENCES

1. Chennai Metropolitan Water Supply & Sewage Board
2. Mary N. Ahuna, Thomas J. Afullo, Akintunde A. Alonge, "Rainfall rate prediction based on artificial neural networks for rain fade mitigation over earth-satellite link", AFRICON 2017 IEEE, pp. 579-584, 2017.
3. Hassanuddin Mohamed Noor, David Ndzi, Guangguang Yang, Noor Zuraidin Mohd Safar, "Rainfall-based river flow prediction using NARX in Malaysia", Signal Processing & its Applications (CSPA) 2017 IEEE 13th International Colloquium on, pp. 67-72, 2017.
4. Lourenco Bueno, Pyramo Costa, Israel Mendes, Enderson Cruz, Daniel Leite, "Evolving ensemble of fuzzy models for multivariate time series prediction", Fuzzy Systems (FUZZ-IEEE) 2015 IEEE International Conference on, pp. 1-6, 2015.
5. S. Renuga Devi, C. Venkatesh, Pranay Agarwal, P. Arulmozhivarman, "Daily rainfall forecasting using artificial neural networks for early warning of landslides", Advances in Computing Communications and Informatics (ICACCI 2014 International Conference on, pp. 2218-2224, 2014.
6. Syeiva Nurul Desylvia, Taufik Djatna, Agus Buono, "A monsoon onset and offset prediction model using back-propagation and moron method: A case in drought region", Advanced Computer Science and Information Systems (ICACSIS) 2015 International Conference on, pp. 201-206, 2015.
7. Zaher Mundher Yaseen, Mazen Ismaeel Ghareb, Isa Ebtehaj, Hossein Bonakdari, Ridwan Siddique, Salim Heddam, Ali A. Yusif, Ravinesh Deo, "Rainfall Pattern Forecasting Using Novel Hybrid Intelligent Model Based ANFIS-FFA", Water Resources Management, 2017.

8. Omid Seyedashraf, Abbas Rezaei, Ali Akbar Akhtari, "Application of computational intelligence methods for complex two-phase flow pattern recognition", Journal of the Brazilian Society of Mechanical Sciences and Engineering, vol. 40, 2018.
9. Siti Amely Jumaat, Flora Crocker, Mohd Helmy Abd Wahab, Nur Hanis Mohammad Radzi, Muhammad Fakri Othman, "Prediction of Photovoltaic (PV) Output Using Artificial Neutral Network (ANN)Based on Ambient Factors", Journal of Physics: Conference Series, vol. 1049, pp. 012088, 2018.
10. Deepak Kumar, K. Vatsala, Sushmitha Pattanashetty, S. Sandhya, Die Anästhesiologie, pp. 413, 2019.

AUTHORS PROFILE

Priyansha T Jat is a final semester student currently pursuing B-tech in Computer Science and Engineering from SRM IST KTR, Tamil Nadu with an aggregate of 88%. Her interest lies in Machine learning, App Development and Software Development. She has worked on several machine learning projects during graduation.

Nikhil Nautiyal is currently pursuing final year of B-Tech in Computer Science Engineering at SRM IST, Tamil Nadu. He will be completing his undergraduate degree by June 2020 and has keen interest in Machine learning and artificial intelligence. He has done works on several fields of embedded systems in his academic career.

M Pushpalatha is currently working as Professor in Department of Computer Science and Engineering, SRM Institute of Science and Technology. She was sponsored as a short term scholar to University of Southern California in the year 2011. She has received M.E degree from Madras University in 2001 and Ph.D degree in Computer Science and Engineering from SRM Institute of Science and Technology in 2014. On the research perspective, she continues to work as an investigator for many DRDO and DST projects and her research interests include Wireless ad hoc networks, Sensor Networks and Internet of Things.