Implementation Procedure on Efficient Distribution of Water using Machine Learning Techniques

Feiroz Khan [1], Mayukh Dasgupta [2], Ankur Ankit [3], Alankrit Yadav [4]
[1] Assistant Professor, Compute Science and Engineering, SRM University, Chennai, Tamil Nadu, India.
[2, 3, 4] Compute Science and Engineering, SRM University, Chennai, Tamil Nadu, India.

Abstract – This paper mainly deals with the functionality prospective of implementing machine learning to water distribution in an urban or rural locality. To this date all the work has been done through human labors. This process although seems to be full proof but has quite a number of loopholes. Firstly, the above process is time consuming as humans can work only at a certain rate. Secondly, human error cannot be left out of the equation as well. When taken into considerations, all these problems cater largely to inadequate water distributions and loss of water in huge quantities. All these problem scenarios can be checked with the implementation of a simple Machine Learning algorithm into the system. Although much of the work still has to be taken care by human methods itself. This also renounces the possibility of job opportunities becoming less due to automation in the system.

Index Terms – Machine Learning, water distribution, LSTM, doppler sensors, time stamp prediction, neural network

1. INTRODUCTION

The above mentioned method uses an LSTM network. This network has the ability to store data of a previous time stamp for a longer period of time. This quality of the ANN(Artificial Neural Network) will help us to analyse past data and compare it with a data set of present time stamp. The LSTM neural net incorporates various activation functions that controls the flow of data within the network. These activation functions decide which data is to be given as output for further evaluation and analysis. Mainly, the LSTM network is divided into three sectors as per their working patterns.

The furnished output from this neural network using past and present dataset will be used to design a new water flowing architecture within a city. The output obtained will be a better way of distributing water in a society considering various parameters such as daily water usage, cooking water usage, bathing water usage, water used by hospitals, industries, etc. Once all these data sets are gathered and evaluated properly, the current problem of water distribution can be solved to some extent.

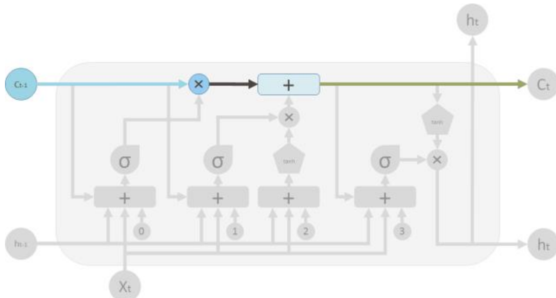
Currently during water crisis in a certain area the government cuts off the water distribution in a certain locality and redirects that water to the water deficient area. This process although takes care of the problem temporarily does not give a permanent solution. By only analyzing the data sets properly can and predicting an efficient usage can we solve this problem. If such a problem arises in the future after implementation of this process, the officials will know exactly how much water deficiency is being created and in which areas and how much water do we need to extract from another location to solve the problem.

2. LSTM NETWORK

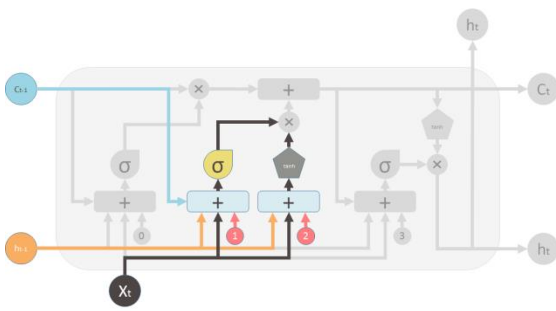
The LSTM(Long Short Term Memory) network stores previous time stamps as well as memory related will those time stamps. It passes these old memories through a main or boss line and continuously takes in present data at each time stamp for each parameter/attribute. Which time stamp and which attributes are to be delivered are decided by the algorithm and the design of the LSTM neural network. Various activation functions like sigmoid, hyperbolic tangent and ReLU are used for this validation purpose. The output of the LSTM is the predicted data along with the required time stamp. BPTT(Back Propagation Through Time) algorithm is an important aspect in this network. We carefully module the incorporation of the BPTT algorithm keeping in mind that when our present data set is being compared no gradient loss takes place. Significant loss of gradient can result in faulty output or less convolution within the hidden layers.

3. WORKING Of LSTM NETWORK

The working of LSTM can be understood by dividing it into several parts for better understanding. To visualize it generally the working of LSTM is that of water pipeline system. You open certain valves in the piping system to let the water flow from one pipe to the other. These valves are the activation functions and the pipes are the data flow lines. New memory will come in from a T-shaped junction and merge with the old memory that is flowing through the main line. There are certain validation points in the main line which are denoted by X. If you want to forget the old memory you multiply that memory with a vector close to 0 and if you want to let that memory pass then you multiply with a vector 1.

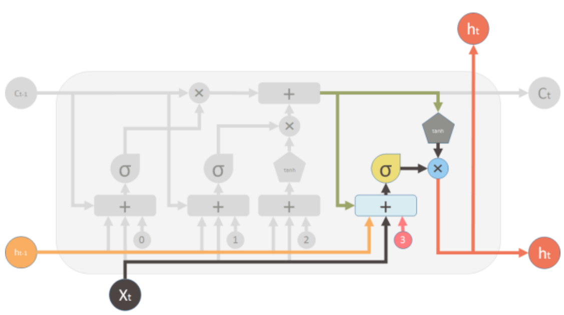


The second part is the memory organization part. Here we compare the old memory with the new memory. But before comparing we validate which memory we need to pass and which not. To do this a sigmoid function is generally used along with a 0 bias valve. The resultant is the summation of all these valves.



The new memory is generated by another half of this section which constitutes of the hyperbolic tan activation function.

The final part actually requires generating the output. This part comprises a valve which takes input from the old output , the current input and the new memory. The algebraic summation of all these inputs gives the output from this part. The bias decides how much of the output is to be generated. Finally, we have a output C_t with a time stamp h_t.



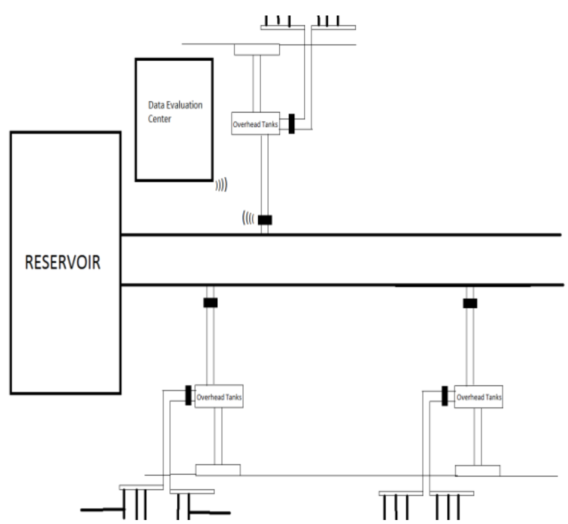
This is the general structure and working of an LSTM structure. One single modification needs to be done to make this structure more efficient and to get a better set of results. The addition of leaky ReLU activation function instead of sigmoid and tanh functions.

4. LEAKY RELU ACTIVATION FUNCTION

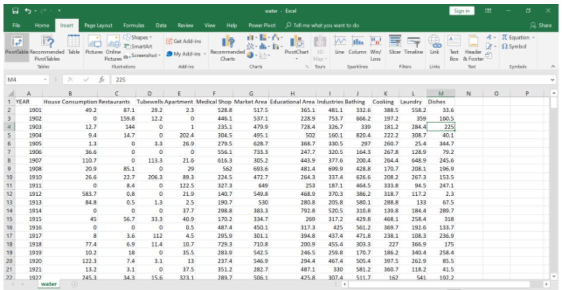
The reason for selecting this activation function is due to its capability of accepting values less than 0. Most of the activation functions often take values between 0 and 1. So as soon as a value is encountered less than 0 the graph just merges with the X axis and refuses to go any lower to show the actual readings. Leaky ReLU deals with this problem. It reads values beyond 0 and 1 and thus gives a more accurate output to our data sets.

5. TOTAL SYSTEM DESIGN

Once the algorithm and the entire neural network is trained and ready we can implement it in a real world scenario. Currently, the water distribution network is quite simple. The main pipelines follow the road patterns i.e. the pipelines run beneath the roadways. These main pipelines extract water from the reservoir. The main pipelines are further divided locality wise which in turn are further divided into neighborhoods. The locality pipes lead to a overhead tank from which vein pipes come out and distribute water to various households. The Doppler ultrasonic sensors can be attached to the outsourcing pipes from the overhead tanks to retrieve real time water distribution data. This data can then be sent to the data managing center where the data will be fed into the algorithm to predict new water distribution architecture. According to this new architecture water can then be transferred from a neighboring locality at the right amount so that none of the areas face water deficit problems.



6. DATASETS AND TEST RESULTS

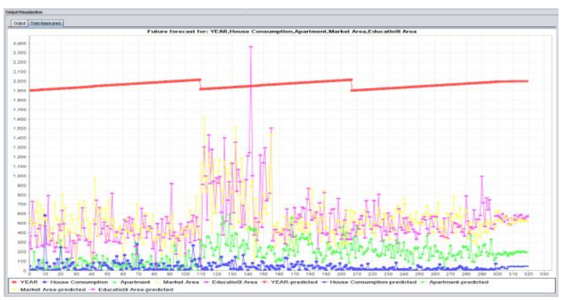


YEAR	House Consumption	Restaurants	Tubewells	Apartment	Medical Shop	Market Area	Educational Area	Industries	Bathing	Cooking	Laundry	Dishes
1901	49.2	87.1	29.2	2.3	528.8	517.5	365.1	481.1	332.6	388.5	558.2	33.6
1902	0	159.8	12.2	0	446.1	537.1	228.9	753.7	666.2	197.2	359	160.5
1903	12.7	144	0	1	235.1	479.9	728.4	326.7	339	181.2	284.4	225
1904	9.4	14.7	0	202.4	304.5	495.1	502	160.1	820.4	222.2	308.7	40.1
1905	1.3	0	3.3	26.9	279.5	628.7	368.7	330.5	297	260.7	25.4	344.7
1906	36.6	0	0	0	556.1	733.3	247.7	320.5	164.3	267.8	128.9	79.2
1907	110.7	0	113.3	21.6	616.3	305.2	443.9	377.6	200.4	264.4	648.9	245.6
1908	20.9	85.1	0	29	562	693.6	481.4	699.9	428.8	170.7	208.1	196.9
1910	26.6	22.7	206.3	89.3	224.5	472.7	264.3	337.4	626.6	208.2	267.3	153.5
1911	0	8.4	0	122.5	327.3	649	253	187.1	464.5	333.8	94.5	247.1
1912	583.7	0.8	0	21.9	140.7	549.8	468.9	370.3	386.2	318.7	117.2	2.3
1913	84.8	0.5	1.3	2.5	190.7	530	280.8	205.8	580.1	288.8	133	67.5
1914	0	0	0	37.7	298.8	383.3	792.8	520.5	310.8	139.8	184.4	289.7
1915	45	56.7	33.3	40.9	170.2	334.7	269	317.2	429.8	468.1	258.4	318
1916	0	0	0	0.5	487.4	450.1	317.3	425	561.2	369.7	192.6	133.7
1917	8	3.6	112	4.5	295.9	301.1	394.8	437.4	471.8	238.1	108.3	236.9
1918	77.4	6.9	11.4	10.7	729.3	710.8	200.9	455.4	303.3	227	366.9	175
1919	10.2	18	0	35.5	283.9	542.5	246.5	259.8	170.7	186.2	340.4	258.4
1920	122.3	7.4	3.1	13	237.4	546.9	294.4	467.4	505.4	397.5	262.9	85.5
1921	13.2	3.1	0	37.5	351.2	282.7	487.1	330	581.2	360.7	118.2	41.5
1922	245.3	34.3	15.6	323.1	289.7	506.1	425.8	307.4	511.7	162	541	192.2

The above dataset was used to train our model. This dataset consists of various water usage parameters like cooking, bathing, drinking, hospital usage, etc. All these data are acquired over a time period of 1901-2015. According to this data set our neural network has thirteen input neurons and five hidden layers.

7. RESULTS AND DISCUSSIONS

Below are the results acquired from WEKA tool which gives a time stamp prediction of two years.



The data plots are done using squared and circular notation whereas the predictions are depicted using jagged edges for two years in the future.

The classification of data was done by Random Forest Tree algorithm which gave an accuracy up to 92.7%. The above graph shows us plotting of parameters such as House consumption, Apartments, Market Areas, and Education Area along with the years. The data is plotted till 2015 and with two years of time stamp the prediction is done for 2016 and 2017.

8. CONCLUSION

We talked about the application of LSTM network in WDS. We also mentioned the application of leaky ReLU activation function in the neural network. The usage of ultrasonic flow meter will give us real time data that will increase the efficiency of the system. Since the entire system is based on observed data from the desired output is of great accuracy. It also reduces human effort and data is obtained at a very short notice. No major changes in the base foundation of th city pipelines is required as this process only focuses in distributing water through the existing pipelines. Excess of water loss can be reduced to a great extent as water will be delivered as per the needs of a household. This process is quite cost effective as the ultrasonic flow meters come at a very low cost.

REFERENCES

[1] A comparative study of artificial neural network architecture for time series prediction of water distribution using flow data S.R.Mounce

[2] Application of Machine Learning Techniques in water distribution networks assisted by domain experts Luis M. Camarinha-Matos, Fernando J.Martinelli

[3] Ultrasonic Flow Meter Basics John Flood

[4] Wall mount ultrasonic flow meter data sheet

Authors

First Author - Mayukh Dasgupta, B.Tech Third Year, SRM Instititue of Science and Technology, Chennai, Tamil Nadu, India.

Second Author - Ankur Ankit, B.Tech Third Year, SRM Institute of Science and Technology, Chennai, Tamil Nadu, India.

Third Author - Alankrit Yadav, B.Tech Third Year, SRM Institute of Science and Technology, Chennai, Tamil Nadu, India.

Correspondence Author - Mr.Feiroz Khan, Assistant Professor, , SRM Institute of Science and Technology, Chennai, Tamil Nadu, India.