VIT®

Vellore Institute of Technology

(Deemed to be University under section 3 of UGC Act, 1956)

School of Information Technology and Engineering

M.Tech Software Engineering

Title: Weather Prediction using Back Propagation

DESIGN PATTERNS – SWE2019

SLOT –C1

FINAL REPORT

SUBMITTED BY:

16MIS0196 - A BHANUCHANDRA

16MIS0417 – M.VAMSI KRISHNA

ASSIGNED BY:

Dr. SENTHIL KUMAR. M

Assistant Professor(Senior)

ABSTRACT:

Back Propagation algorithm is used for analysis of Weather Forecasting for the month of one month with three sets of data. Back Propagation Algorithm is the best one to give the output with less error. So here we implement our project using Back propagation Algorithm tool and explains the graphical output obtained in the process with three data sets and one target data set. Back Propagation Algorithm is the best one to give the output with less error**.** MATLAB is used for the simulation of process. Using ANN is an innovative approach to construct a computationally intelligent system that is able to process non-linear weather conditions within a specific domain, and make prediction. A number of researches have been done. The Neural Networks supports different types of training or learning algorithms. One such algorithm is **Back Propagation Neural Network (BPN) technique**. The main advantage of the BPN neural network method is that it can fairly approximate a large class of functions. This method is more efficient than numerical differentiation.

INTRODUCTION:

Weather forecasting is the most important of all fields. Weather forecasting plays vital role in fields like Agriculture, Transportation, Air Way Travels, and etc… Now a days it is easy to know the current weather report and it is important to predict and analyze the weather report for future. Agriculture is the main area mainly depended on weather and rainfalls now a days the agriculture has went in loss and the agriculture field is diminishing. And the weather tremendously increased hot in weather. By today there exists a number of ways to know the weather and we can also able to see the weather data for the upcoming days also. Weather report should be some knowledgeable to the users. So giving the weather data with most accurate or with less error will be more helpful for the user. There exists many ways to predicting the data for weather forecasting even for the next century. But applying the best method will give the best result. Neural network deals with many areas and also in weather forecasting. Neural network clearly talks with the inputs, outputs and the sample data that are used. Neural network has a lots of algorithms to solve the different kinds of problems in different ways. Fuzzy also used to handle the data in a very easy set of fuzzy set instead of the crisp data. Fuzzy values can be easily handled by labeling it with the corresponding labeling, for example the weather can be labeled as very hot, hot, moderate, cold, very cold. So handling

the fuzzy data for training will also easier for the work. MATLAB is programming software used to solve mathematical problems and used in image processing. MATLAB also analyze the data to give outputs.

RELATED WORKS:

In area of weather prediction system there are many researchers have tried by using some data mining methodologies. The K-means clustering is to divide the land and other areas to discover the interesting patterns. Dataset analysed on the basis of clustering and classification. The probability of weather prediction is discussed. Weather is predicted based on the uncertainties of initial conditions and model formulation. The accuracy of results is more important because on the weather forecasting because some peoples are relay on the site to manage their events. The reason depend on climate decision of events is taken. Work on the real time dataset normalized using min-max normalization. Quantitative forecast like temperature is the important in agriculture area, to plan for the future and planning to trip in the city, as well as to traders within commodity markets. Weather and climate disasters are on the rise in the India no national capacity to provide skillful long range severe weather outlooks. Many variables that may affect weather, considering all the current statistics on the climate, physical equations to describe particle interactions on the smallest scale are aggregated to model an entire weather system. To assess the current state of the science and to identify what is required to develop long range severe weather prediction system for the future prediction and the preventive measure to do. Forecasting helps to take necessary measures, to prevent damage to life and property to a large extent. Quantitative forecast like temperature is important in agriculture area, as well as to traders within commodity markets. Different categories of forecasting methods are BPN and Naive approach etc.,

LITERATURE REVIEW:

S.NO	TITLE	DESCRIPTION
1	Weather Prediction Using Data Mining	Weather forecasting is mainly concerned with the prediction of weather condition in the given future time. Weather forecasts provide critical information about future weather. There are various approaches available in

		weather forecasting, from relatively simple observation of the sky to highly complex computerized mathematical models. The prediction of weather condition is essential for various applications. Some of them are climate monitoring, drought detection, severe weather prediction, agriculture and production, planning in energy industry, aviation industry, communication, pollution dispersal, and so forth. In military operations, there is a considerable historical record of instances when weather conditions have altered the course of battles. Accurate prediction of weather conditions is a difficult task due to the dynamic nature of atmosphere.
2	Weather prediction using the data mining Techniques	Data mining techniques has been the subject of several research papers. Climate change may have a serious impact on the availability especially in countries of human resources. Knowledge discovery from temporal, spatial and spatiotemporal data is critical for climate change and climate effect on the environment. Climate change will significantly impact, public health. Historical climate statistics for future forecasting. The evolution of the tools and techniques available to gather data about weather, water, temperature, etc. In order to prevent these public health impacts, we can take actions to prevent climate change by analyzing historical data. Climate change refers to long-term shifts in weather conditions and patterns of extreme weather events
3	Application of Data Mining Techniques in Weather Prediction and Climate Change Studies	Weather forecasting is a vital application in meteorology and has been one of the most scientifically and technologically challenging problems around the world in the last century. In this paper, we investigate the use of data mining techniques in forecasting maximum temperature, rainfall, evaporation and wind speed. This was carried out using Artificial Neural Network and Decision Tree algorithms and meteorological data collected between 2000 and 2009 from the city of Ibadan, Nigeria. A data model for the meteorological data was developed and this was used to train the classifier algorithms.

4	Data Mining Techniques Using To Weather Prediction	Meteorological department worldwide is putting a great effort in the research areas of weather prediction. Since, India is an agriculture based country most of the people are dependent on the weather conditions. Large amount of India population depend on monsoon. It is always a matter of concern for the scientists to exactly forecast the weather conditions. Weather forecast includes prediction of rain, fog, winds, clouds, lightening, storm etc. One the biggest challenge in weather forecasting is its unpredictable and dynamic climate data sets, which can frequently change according to global climatic changes. Many techniques have been applied and suggested out of which data mining is considered as the most feasible approach towards Weather forecasting. Data Mining has the capability to mine hidden patterns, relations and provide verification/validation of data sets based on certain input conditions.
5	Analysis of Data Mining Techniques for Weather Prediction	Weather prediction has been a challenging problem in meteorological department since years. Even after the technological and scientific advancement, the accuracy in prediction of weather has never been sufficient. Even in current date this domain remains as a research topic in which scientists and mathematicians are working to produce a model or an algorithm that will accurately predict weather. There have been immense improvements in the sensors that are responsible for recording the data from the environment and cancel the noise present in them; Indian Journal of Science and Technology, Vol 9(38), DOI: 10.17485/ijst/2016/v9i38/101962, October 2016 ISSN (Print) : 0974-6846 ISSN (Online) : 0974-5645 along with this new models have been proposed which include different attributes related to weather to make accurate prediction. In present era weather forecasting and analysis has become a challenging problem around the world from the last century. The reason behind are the two main factors: Firstly, it is useful for many human activities like agriculture sector, tourism and natural disaster prevention. Secondly, due to various technological advances like the

6	Application of Data Mining Techniques in Weather Data Analysis	growth of computational power and ongoing improvements in measuring systems.
7	A study on prediction of rainfall using datamining technique	India is an agricultural country and its economy is largely based upon crop productivity and rainfall. For analyzing the crop productivity, rainfall prediction is require and necessary to all farmers. Rainfall Prediction is the application of science and technology to predict the state of the atmosphere. It is important to exactly determine the rainfall for effective use of water resources, crop productivity and pre planning of water structures. Using different data mining techniques it can predict rainfall. Data mining techniques are used to estimate the rainfall numerically. This paper focuses some of the popular data mining algorithms for rainfall prediction. Naive Bayes, K-Nearest Neighbour algorithm, Decision Tree, Neural Network and fuzzy logic are some of the algorithms compared in this paper. From that comparison, it can analyze which method gives better accuracy for rainfall prediction.
8	A Weather Forecasting Model using the Data Mining Technique	The weather conditions are changing continuously and the entire world is suffers from the changing Clemet and their side effects. Therefore pattern on changing weather conditions are required to observe. With this aim the proposed work is intended to investigate about the weather condition pattern and their forecasting model. On the other hand data mining technique enables us to analyse the data and extract the valuable patterns from the data. Therefore in order to understand fluctuating patterns of the weather conditions the data mining based predictive model is reported in this work. The proposed data model analyse the historical weather data and identify the significant on the data Rainfall prediction is one of the challenging tasks in weather forecasting. Accurate and timely rainfall prediction can be very helpful to take effective security measures in advance regarding: ongoing construction projects, transportation activities, agricultural tasks, flight

9	Rainfall Prediction using Data Mining Techniques: A Systematic Literature Review	operations and flood situation, etc. Data mining techniques can effectively predict the rainfall by extracting the hidden patterns among available features of past weather data.
10	Weather Forecasting Using Data Mining	Weather forecasting is the application of science and technology to predict the state of the atmosphere for a given location. Ancient weather forecasting methods usually relied on observed patterns of events, also termed pattern recognition. For example, it might be observed that if the sunset was particularly red, the following day often brought fair weather. However, not all of these predictions prove reliable. Here this system will predict weather based on parameters such as temperature, humidity and wind. This system is a web application with effective graphical user interface. User will login to the system using his user ID and password. User will enter current temperature; humidity and wind, System will take this parameter and will predict weather from previous data in database
11	Smart Weather Forecasting Using Machine Learning: A Case Study in Tennessee	Traditionally, weather predictions are performed with the help of large complex models of physics, which utilize different atmospheric conditions over a long period of time. These conditions are often unstable because of perturbations of the weather system, causing the models to provide inaccurate forecasts. The models are generally run on hundreds of nodes in a large High Performance Computing (HPC) environment which consumes a large amount of energy. In this paper, we present a weather prediction technique that utilizes historical data from multiple weather stations to train simple machine learning models, which can provide usable forecasts about certain weather conditions for the near future within a very short period of time

Existing system:

Smart Weather Forecasting Using Machine Learning

Machine Learning for Weather Forecasting Machine learning is a data science technique which creates a model from a training dataset. A model is basically a formula which outputs a target value based on individual weights and values for each training variable. In each record, corresponding weights (sometimes between 0 and 1) to each variable tells the model how that variable is related to the target value. There must be sufficient amount of training data to determine the best possible weights of all the variables. When the weights are learned as accurately as possible, a model can predict the correct output or the target value given a test data record. Utilizing simple machine learning techniques allow us be relieved from the complex and resource-hungry weather models of traditional weather stations. It has immense possibilities in the realm of weather forecasting. Such a forecasting model can be offered to the public as web services very easily.

Methodology:

In this case study, we aim to use ML techniques to predict the temperature of the next day at any particular hour of the city of Nashville, Tennessee, based on the weather data of the current day of this city and a couple of its surrounding cities.

Performance Measure:

They used root mean square error(RMSE) to evaluate their models

$$RMSE = \sqrt{\frac{\sum_{t=1}^{n} (\hat{y}_t - y_t)^2}{n}}$$

Performance Comparison of Models

At first, only one-week data is used as training set, that week is the immediate previous week of the test week. As we said earlier, the test data comprises of seven days of data starting from September 1, 2018. Thus, first week of training data represents the weather data starting from the 25th August and ending on the 31st August, 2018. Two weeks of training refers to the data of the previous two weeks of the test week, and so on. Using only one week, RMSE is noticeably high, nearly 3.3. As we increase the number of weeks, RMSE drops considerably. At 5-week situation, RMSE is nearly 3.03

Proposed System:

The back propagation learning algorithm is mostly used in neural networks. This learning algorithm is used for multilayer feed forward networks which use differentiable activation functions for updating weights. The network which is using back propagation learning algorithm is known as Back Propagation Networks (BPN). To classify the provided input this algorithm provides a procedure for updating the weights. The basic concept for this weight updating is gradient descent method. The main aim is to train the network so that it produces a balance between its ability to memorize and its ability to produce the output with the help of trained data. Here the generated output is compared with the target output and if error occurs it is propagated back to the hidden layer. Following are the methods which are analyzed and verified.

- The conventional Back Propagation (BP).
- The Back Propagation with Adaptive Gain (BP-AG).
- Back Propagation with Adaptive Gain, Adaptive Momentum and Adaptive Learning Rate (BPAGAMAL).
- Back propagation with Adaptive momentum (BP-AM).
- Back propagation with momentum and adaptive learning rate (BPAL).

In multilayer neural network the inputs need to be presented repeatedly, for which the weights has to be adjusted before the network is able to settle down into an optimal solution. Network paralysis occurs when the adjustment of weight during training are done in large values. Due to large change in weights can make most of the units to function at extreme values, in a region where the activation function of the derivatives becomes very small

Methodology:

We are using BACK-PROPAGATION-NETWORK for our project Temperature Prediction:

1. The network is first initialized by setting up all its weights to be small random numbers say between 0 and 1.

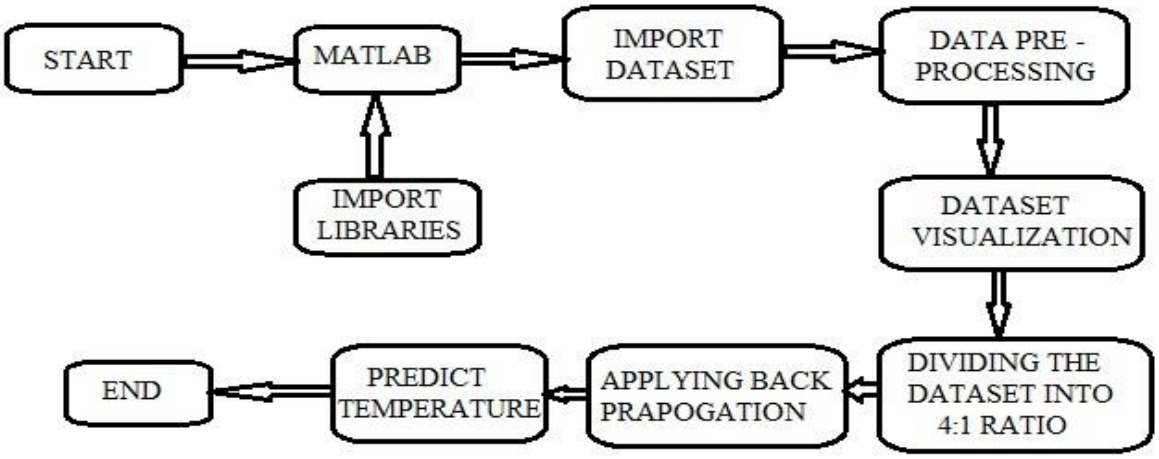
2. The input pattern is applied and the output calculated (this is called the forward pass). The calculation gives an output which is completely different to what actually needed (the Target), since all the weights are random.

3. We then calculate the Error of each neuron, which is essentially: Target – Actual Output. This error is then used mathematically to change the weights in such a way that the error will get smaller.

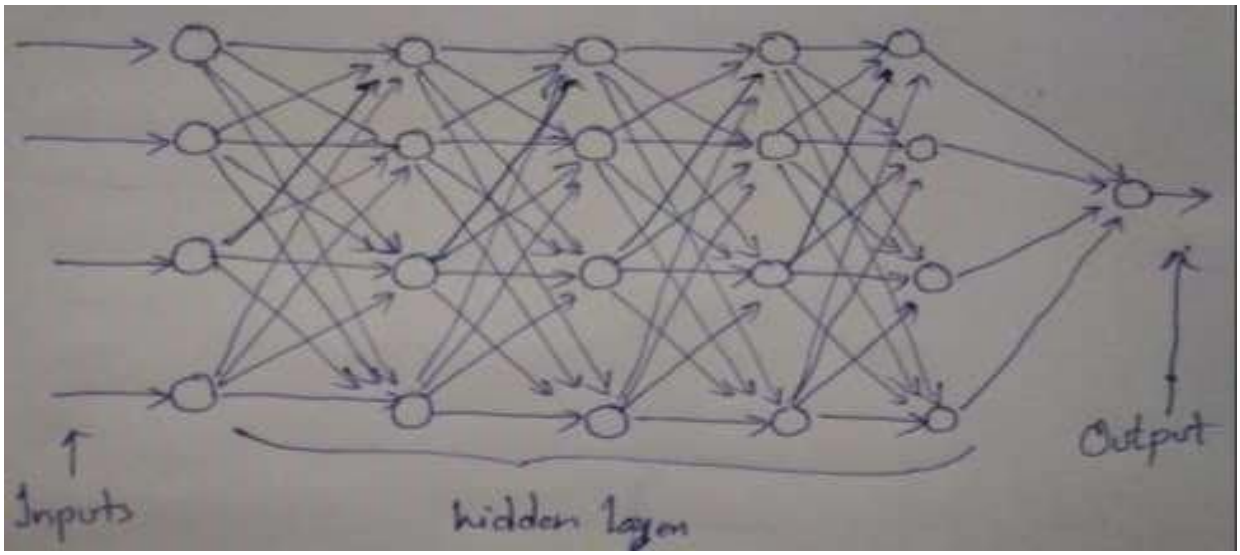
4. The Output of each neuron will try to get closer to its Target (this part is called the reverse pass).

5. This process is repeated again and again until the error is minimal.

DESIGN FRAMEWORK:



ARCHITECTURE DIAGRAM:



CHANLLENGES

The training of the BPN is done in three stages-the feedforward of the input training pattern, the calculation and back-propagation of the error and updating of weights. The testing of the BPN involves the computation of feed forward phase only. There can be more than one hidden layer but only one hidden layer is sufficient. Even though the training is very slow, once the network is trained it can produce its outputs rapidly.

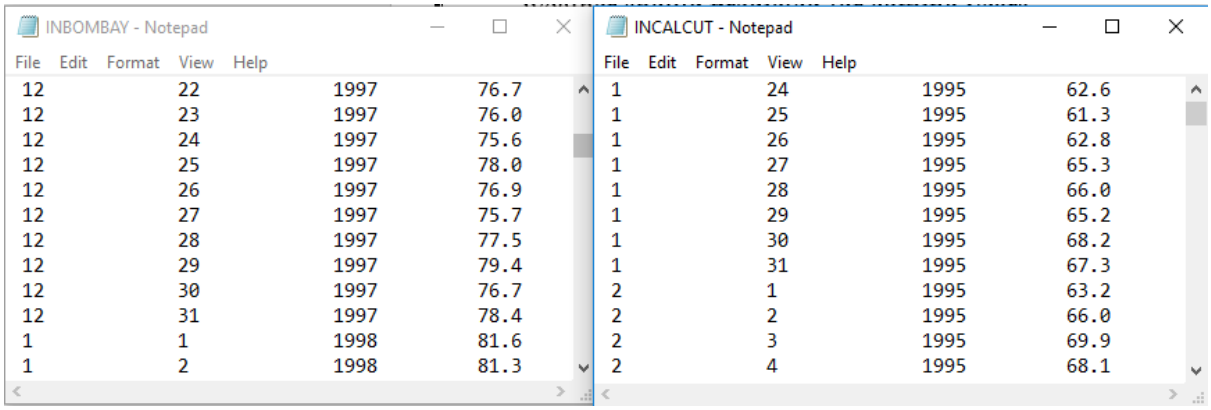
DATA SET, DATA SOURCE, CHARACTERIZATION, PREPROCESSING

Data collection is the process gathering of the required data for the processing purpose. We have collected from the locale city center and crawled from the open weather site for the particular region. The data has been collected of the

last few years. We have taken few parameters such as temperature, Prepare the training data set using pre-processing on the raw data. First we want to take the dataset that contains the information that is called raw data. After we want to check for missing values and outliers and noisy data in the given dataset there should not be a missing values, noisy data and outliers. By checking these values go to another step. Identify useful attributes for classification. In this all four attributes are useful for the classification. Learn a model using training examples in Training set. Decision tree model we will take 75 samples for the training dataset.

DATA PREPARATION

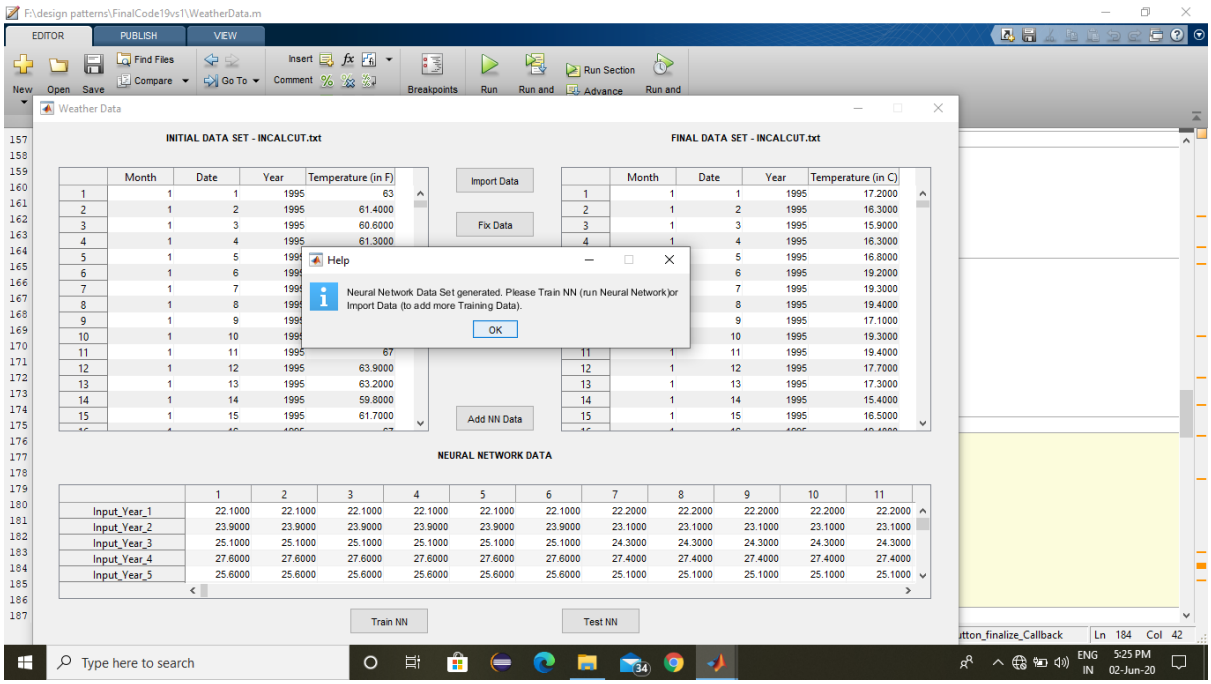
The dataset contains previous 20 years of data with daywise weather report. It has some missing values also. Based on the previous data the weather should predicted for coming years

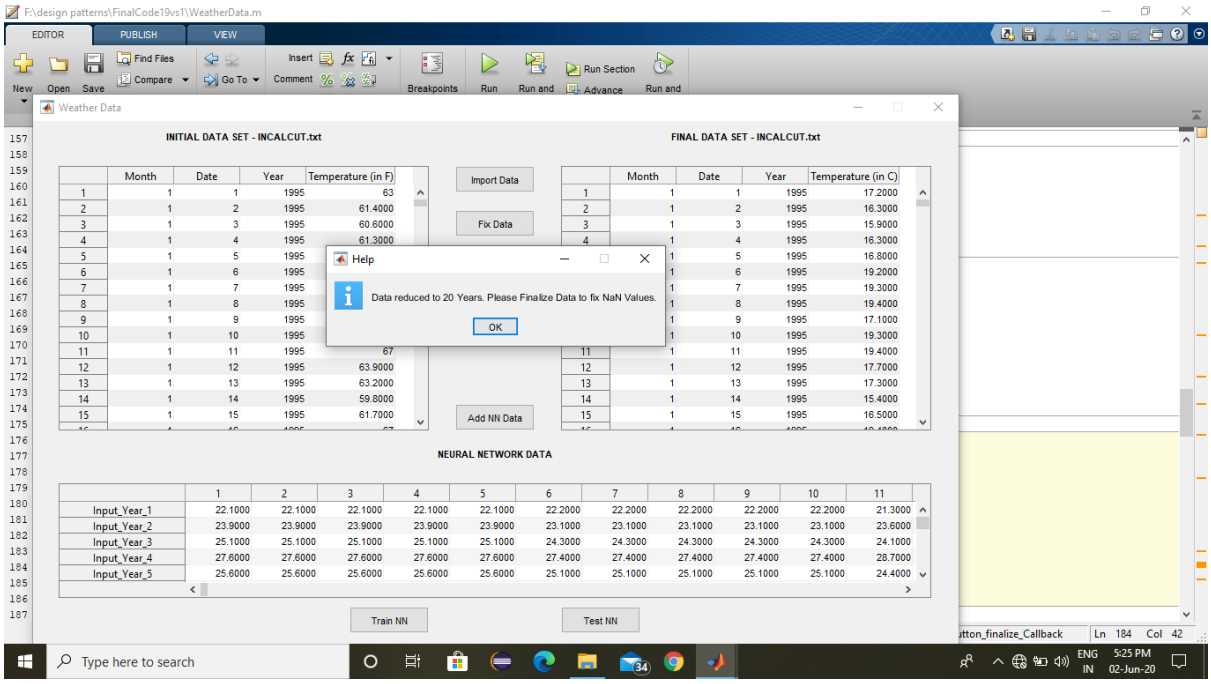


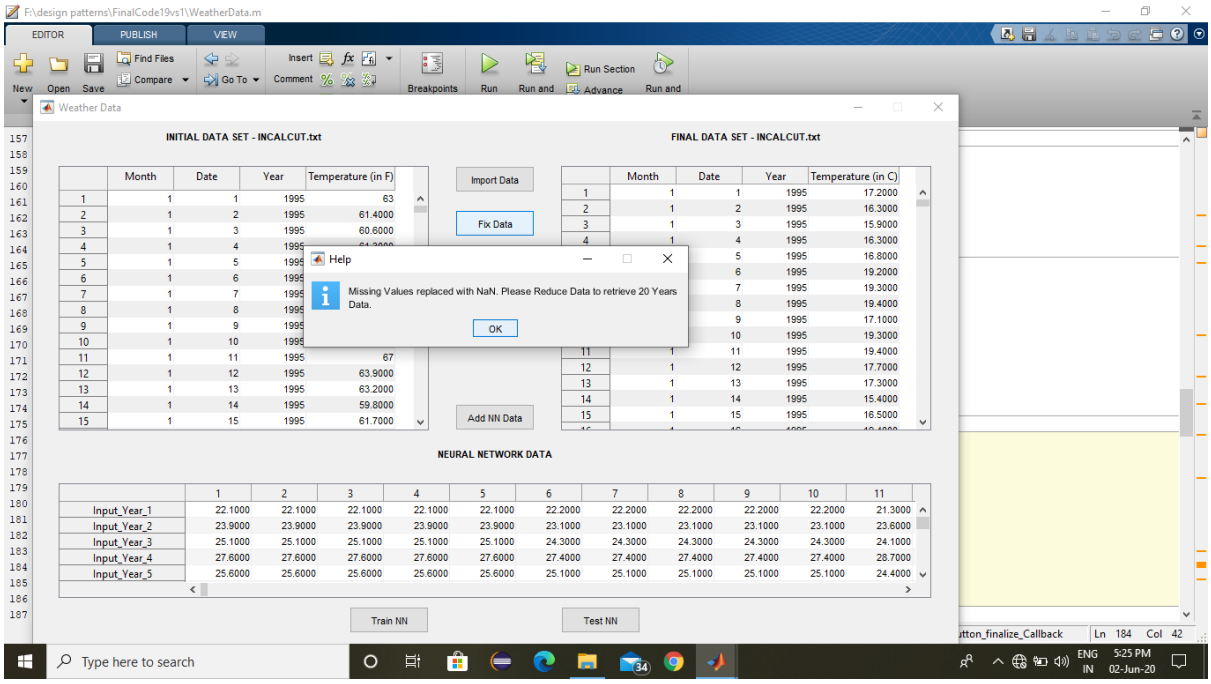
Dataset Source Link:

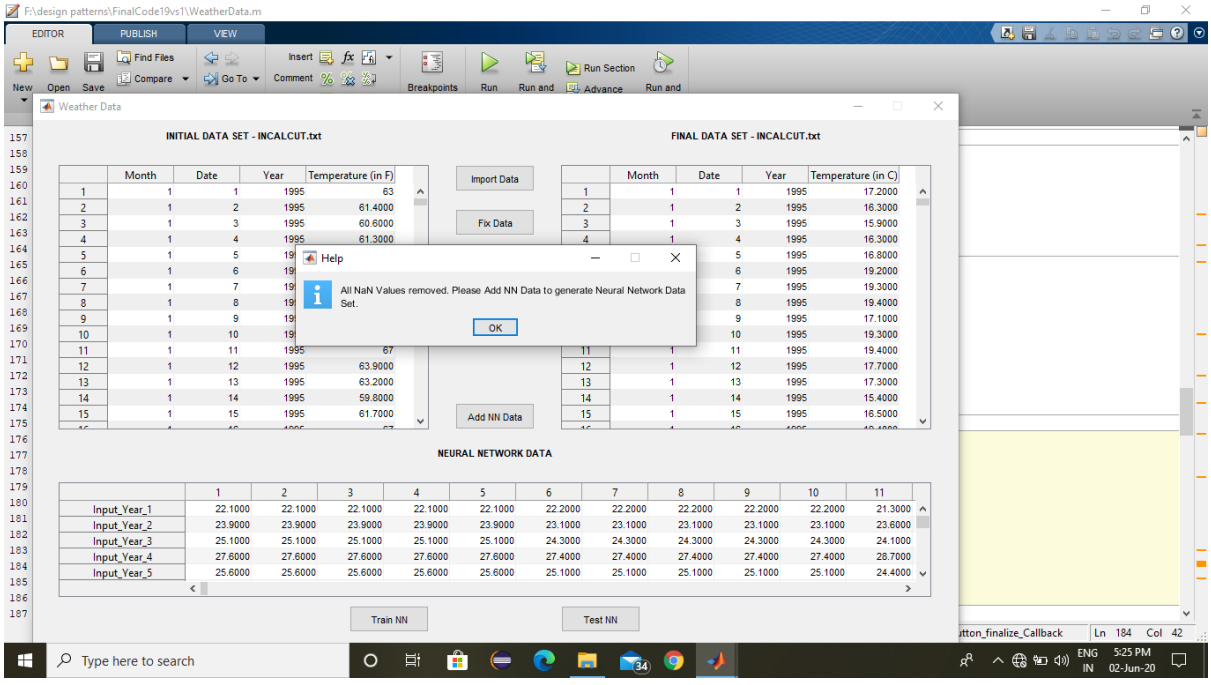
https://www.mathworks.com/matlabcentral/fileexchange/55884-temperature-pattern-prediction

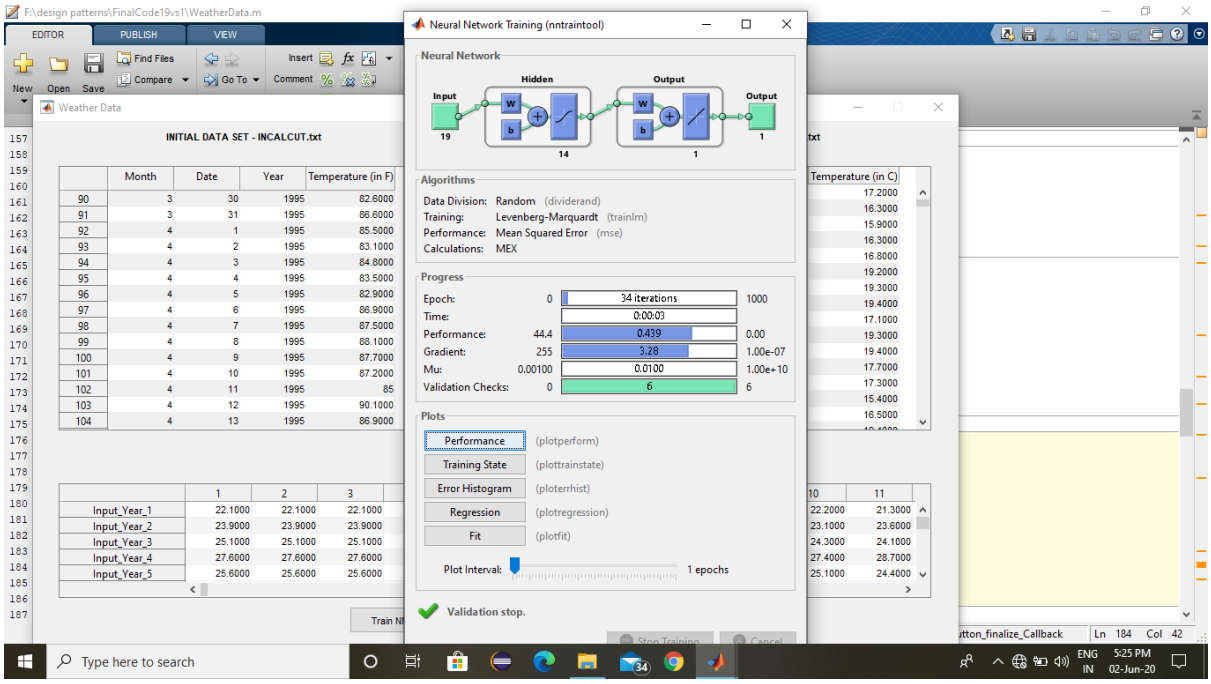
System Implementation

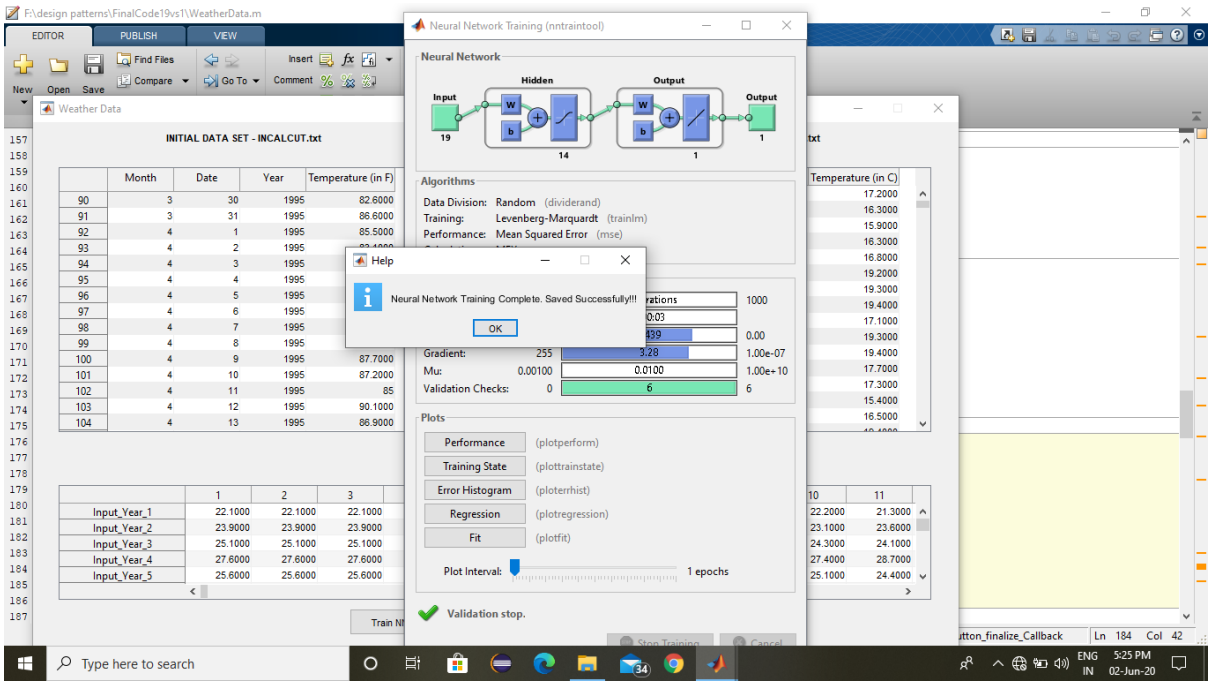





F:\design patterns\FinalCode19vs1\WeatherData.m
EDITOR
PUBLISH
VIEW
Weather Data
INITIAL DATA SET - INCALCUT.txt
FINAL DATA SET - INCALCUT.txt
Import Data
Fix Data
Add NN Data
Help
Missing Values replaced with NaN. Please Reduce Data to retrieve 20 Years Data.
OK
NEURAL NETWORK DATA
Train NN
Test NN
Type here to search


F:\design patterns\FinalCode19vs1\WeatherData.m
EDITOR
PUBLISH
VIEW
Weather Data
INITIAL DATA SET - INCALCUT.txt
FINAL DATA SET - INCALCUT.txt
Import Data
Fix Data
Add NN Data
Help
All NaN Values removed. Please Add NN Data to generate Neural Network Data Set.
OK
NEURAL NETWORK DATA
Train NN
Test NN
Type here to search





CODE:

```
function varargout = WeatherData(varargin)

gui_Singleton = 1; gui_State = struct('gui_Name',       mfilename, ...
```

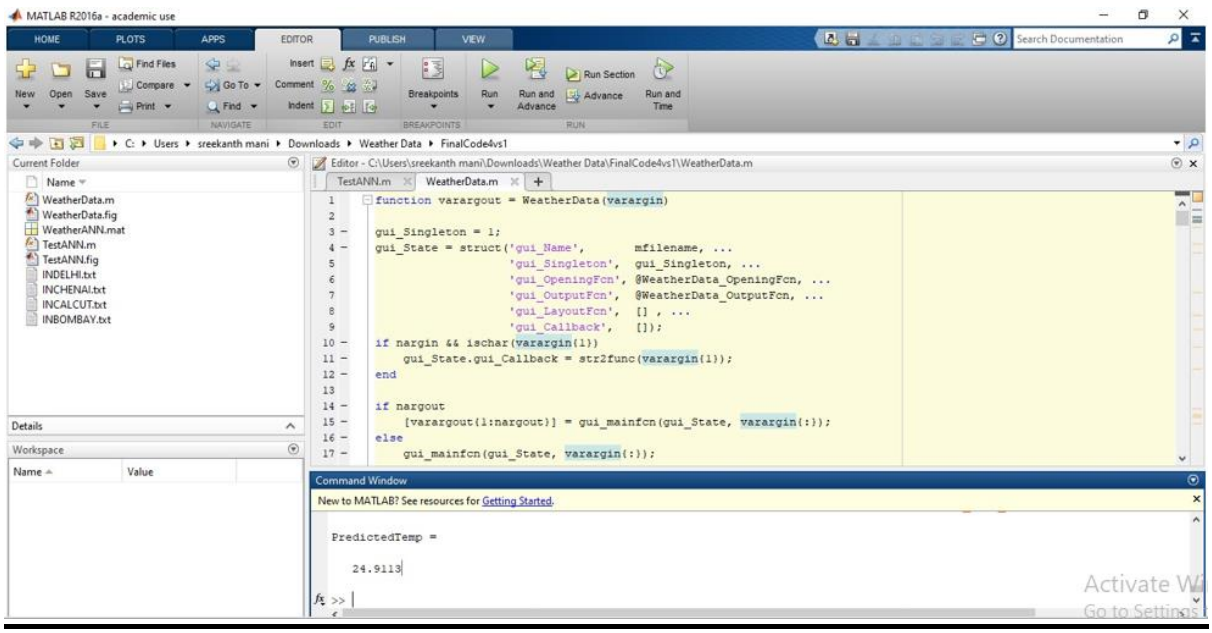
```
                'gui_Singleton',  gui_Singleton, ...
                'gui_OpeningFcn', @WeatherData_OpeningFcn, ...
                'gui_OutputFcn',  @WeatherData_OutputFcn, ...
                'gui_LayoutFcn',  [] , ...
                'gui_Callback',   []);
if nargin && ischar(varargin{1})
   gui_State.gui_Callback = str2func(varargin{1}); end
if nargout
   [varargout{1:nargout}] = gui_mainfcn(gui_State, varargin{:}); else     gui_mainfcn(gui_State, varargin{:}); end
function WeatherData_OpeningFcn(hObject, eventdata, handles, varargin)
NNDATA = double([]);
handles.NNDATA = NNDATA;
handles.output = hObject;
guidata(hObject, handles);
function varargout = WeatherData_OutputFcn(hObject, eventdata, handles)  varargout{1} = handles.output;
function button_import_Callback(hObject, eventdata, handles)
filename = uigetfile('.txt'); delimiter = ' ';
formatSpec = '%f%f%f%f%[^\n\r]';
fileID = fopen(filename,'r');
dataArray          =          textscan(fileID,     formatSpec,          'Delimiter',          delimiter,
'MultipleDelimsAsOne', true, 'EmptyValue' ,NaN, 'ReturnOnError', false);
fclose(fileID);
TEMPDATA = [dataArray{1:end-1}];
```

```
set(handles.uitable1, 'data', TEMPDATA);

set(handles.text1, 'string', ['INITIAL DATA SET - ' filename]);

handles.filename = filename;

handles.TEMPDATA = TEMPDATA;

guidata(hObject, handles);

msgbox(['Data imported from ' filename '. Please Fix Data to replace Missing Values with NaN.'], 'Help', 'help');

function button_fix_Callback(hObject, eventdata, handles) for i = 1: length(handles.TEMPDATA)     if
handles.TEMPDATA(i, 4) == -99         handles.TEMPDATA(i, 4) = NaN;     end

   if (handles.TEMPDATA(i, 1) == 2 && handles.TEMPDATA(i, 2) == 28) &&

(handles.TEMPDATA(i+1, 2) == 1)         for j = length(handles.TEMPDATA)-1:-1:i+1
handles.TEMPDATA(j+1, :) = handles.TEMPDATA(j, :);         end

      handles.TEMPDATA(i+1, 1) = 2;         handles.TEMPDATA(i+1, 2) = 29;         handles.TEMPDATA(i+1, 3) =
handles.TEMPDATA(i, 3);         handles.TEMPDATA(i+1, 4) = NaN;     end end

set(handles.uitable1, 'data', handles.TEMPDATA);

guidata(hObject, handles);

msgbox('Missing Values replaced with NaN. Please Reduce Data to retrieve 20

Years Data.', 'Help', 'help');

function button_reduce_Callback(hObject, eventdata, handles)

VARTEMP(7320, 4) = 0;

for i = 1: 7320

   VARTEMP(i, :) = handles.TEMPDATA(i, :);

end set(handles.uitable1, 'data', VARTEMP);

handles.VARTEMP = VARTEMP;

guidata(hObject, handles);

msgbox('Data reduced to 20 Years. Please Finalize Data to fix NaN Values.', 'Help', 'help');
```

```
function button_finalize_Callback(hObject, eventdata, handles)

TIME = 1: 7320;

VTEMP = (handles.VARTEMP(:, 4)-32)*5/9;

VTEMP = round(VTEMP*10)/10;

Xi = TIME;

Yi = interp1(TIME, VTEMP, Xi, 'spline'); warning('off','MATLAB:interp1:NaNstrip'); handles.VARTEMP(:, 4) = Yi;

set(handles.uitable2, 'data', handles.VARTEMP); set(handles.text2, 'string', ['FINAL DATA SET - '
handles.filename]);

guidata(hObject, handles);

msgbox('All NaN Values removed. Please Add NN Data to generate Neural

Network Data Set.', 'Help', 'help');

function button_add_Callback(hObject, eventdata, handles)

odatasize = length(handles.NNDATA); ndatasize = odatasize + 5856; handles.NNDATA = resizem
(handles.NNDATA, [5 ndatasize]); handles.NNDATA (1, odatasize+1:ndatasize) = handles.VARTEMP (1: 5856, 4);
handles.NNDATA (2, odatasize+1:ndatasize) = handles.VARTEMP (367: 6222, 4); handles.NNDATA (3,
odatasize+1:ndatasize) = handles.VARTEMP (733: 6588, 4); handles.NNDATA (4, odatasize+1:ndatasize) =
handles.VARTEMP (1099: 6954,

4);

handles.NNDATA (5, odatasize+1:ndatasize) = handles.VARTEMP (1465: 7320,

4);

nninput = handles.NNDATA(1:4, :); nnoutput = handles.NNDATA(5, :);

set(handles.uitable3, 'data', handles.NNDATA);

handles.nninput = nninput; handles.nnoutput = nnoutput; guidata(hObject, handles); msgbox('Neural Network
Data Set generated. Please Train NN (run Neural Network)or Import Data (to add more Training Data).', 'Help',
'help');

function button_train_Callback(hObject, eventdata, handles)

inputs = handles.nninput; targets = handles.nnoutput;
```

```
hiddenLayerSize = 4;

WeatherANN = fitnet(hiddenLayerSize);

WeatherANN.inputs{1}.processFcns = {'removeconstantrows','mapminmax'};

WeatherANN.outputs{2}.processFcns = {'removeconstantrows','mapminmax'};

WeatherANN.divideFcn = 'dividerand';

WeatherANN.divideMode = 'sample';

WeatherANN.divideParam.trainRatio = 80/100;

WeatherANN.divideParam.valRatio = 10/100;

WeatherANN.divideParam.testRatio = 10/100;

WeatherANN.trainFcn = 'trainlm';

WeatherANN.performFcn = 'mse';

WeatherANN.plotFcns = {'plotperform','plottrainstate','ploterrhist', ...   'plotregression', 'plotfit'};

[WeatherANN,tr] = train(WeatherANN,inputs,targets);

outputs = WeatherANN(inputs); errors = gsubtract(targets,outputs);

trainTargets = targets .* tr.trainMask{1}; valTargets = targets  .* tr.valMask{1}; testTargets = targets  .*
tr.testMask{1};

performance = perform(WeatherANN,targets,outputs) trainPerformance =
perform(WeatherANN,trainTargets,outputs) valPerformance = perform(WeatherANN,valTargets,outputs)
testPerformance = perform(WeatherANN,testTargets,outputs)

save WeatherANN; msgbox('Neural Network Training Complete. Saved Successfully!!!', 'Help', 'help');

function button_test_Callback(hObject, eventdata, handles)

TestANN;
```

Results:



Description of input and output:

20 years data is taken from a particular city or a palce. 4

inputs 1 output

- 4 temperatures will be taken as basic input and 5th years temperatures will be taken as expected output and check with the actual output . here we calculate the accuracy by (Actual – Predicted)/no of observations.

4 Years:

Starting year – final year

- ⇨ 1 - 16;
- ⇨ 2 - 17;
- ⇨ 3 - 18;
- ⇨ 4 - 19;

Expected output year

- ⇨ 5-20

Consider the values :

22.1000 23.9000 25.1000 27.6000

Above values are of 4 years

25.6000 - this value is of 5th years which will be taken as the expected output this is actual value

The predicted ie the output is **24.9737**

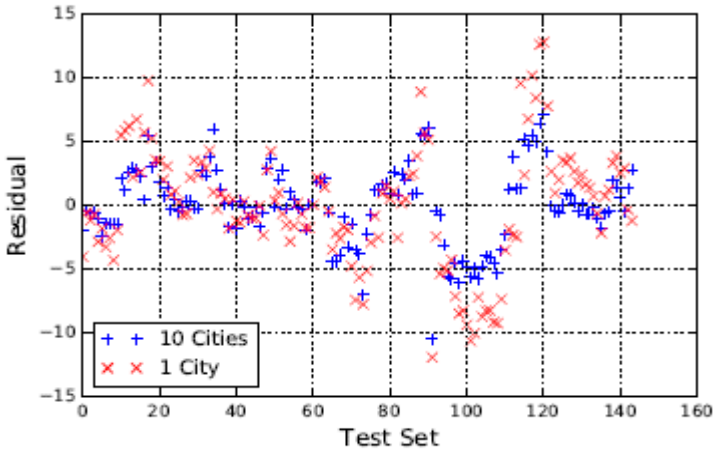
Accuracy: (25.6000-24.9737)/4=0.156575

After training many times the accuracy decreases and hence we can predict the approximate temperature

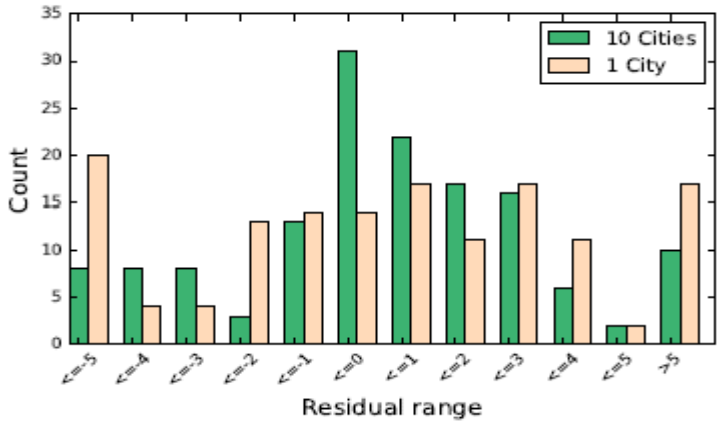
Comparison between both algorithms:

Existing system:

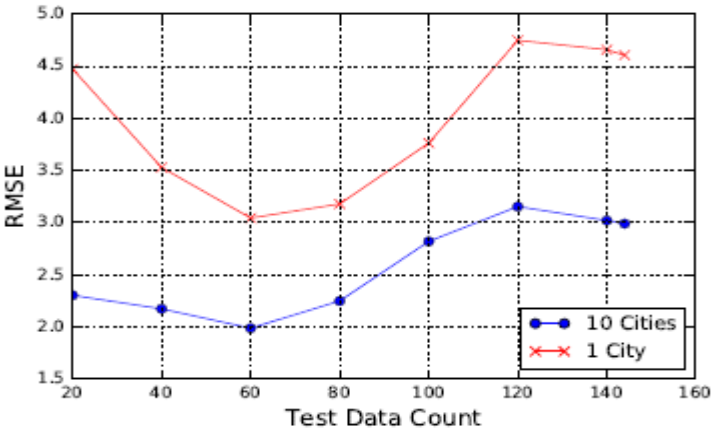
RMSE is nearly 3.03



a)Error in prediction of the test set



b)distribution of errors in prediction of the test set



c)RMSE for different number of test data.

Proposed System:

Accuracy: 0.156575

SUMMARY AND CONCLUSIONS

After going through all the above study & discussion we see that applying NN with BPN algorithm for forecasting the weather forecasting. The dataset selection, input variable selection, the relationships & inter-conditions is most feasible rather than any other algorithm for weather forecasting approach. The study also says that NN with BPN algorithms are the best combination for weather dependencies among the data, the proper training set and the proper ANN architecture are most vital for the best prediction results.

Finally the trained neural network can predict the future temperature with less error.

REFERENCES:

Base Paper:
https://www.ijcaonline.org/research/volume140/number3/tyagi-2016-ijca-909252.pdf

http://www.ijettjournal.org/volume-3/issue-1/IJETT-V3I1P204.pdf

http://www.ijsce.org/wp-content/uploads/papers/v2i5/E1053102512.pdf

https://www.researchgate.net/publication/330369173_Smart_Weather_Forecasting_Using_Machine_Learning_A_Case_Study_in_Tennessee

****THANK YOU****