ENIN 834 – Fall 2018

Project Report

Title:

Using the ANN for Forecasting the Precipitation and Temperature for 6-hours ahead in Regina Airport

Students:

Mahnaz Olfatimalamiri (200379893)

Shubham Sharma (200398667)

Mohammad Parhamian (200386457)

Professor:

Dr. Rene V. Mayorga

Table of Contents

Introduction

Weather forecasts are often viewed as the most important services provided by the meteorological profession (Leipper 1995). These forecasts however possess no intrinsic economic value unless they influence the behavior of individuals and organizations whose activities are sensitive to weather conditions (Murphy 1994). The Australian Bureau of Meteorology (BOM) provides Terminal Aerodrome Forecasts (TAFs) for all major airports in Australia with each capital city TAF being valid for 24 h. These forecasts are used by the Commercial airlines and Air Services Australia (formerly the Civil Aviation Safety Authority) for flight planning, in-flight decision making, and optimization of airport operations. BOM staff routinely review their forecasting performance to seek ongoing improvement, and this study is part of a coordinated effort between Macquarie University and the BoM in relation to the Latter's National Fog Project. The main use of TAFs by airlines and aircraft operators is for flight planning, both pre- and intraflight. The core component of this planning is the pilot's "alternative fuel" decision (pretakeoff). Such decisions are based on TAFs issued for the intended destination.

When certain criteria are exceeded, legislation requires that additional fuel be uplifted should extra flying time or diversion to an alternate landing site be necessary. When these criteria are not exceeded, depending upon the fuel policy of the airline, the pilot must decide whether to carry the additional fuel. The profit margins of most airlines are relatively low; they have been previously estimated at approximately 2% (Bonné 2005). Unexpected costs can therefore shift a flight from making money for the airline, to incurring a substantial loss.

Daily Weather forecasting is used for multiple reasons in multiple areas like agriculture, energy supply, transportations, etc. Accuracy of weather conditions shown in forecast reports is very necessary. In this project, the review is conducted to investigate a better approach for forecasting which compares many techniques such as ANFIS, Artificial Neural Network, Ensemble Neural Network, Backpropagation Network, Radial Basis Function Network, and General Regression Neural Network which are used for different types of forecasting. Among which ANN with the accurate algorithm performs prediction with minimal error. Artificial neural network is simply understood as a nonlinear statistical data modeling tool that presents complex relationships between predictors input layer and predict output layer through a synapse system hidden layers connecting predictors with predict.

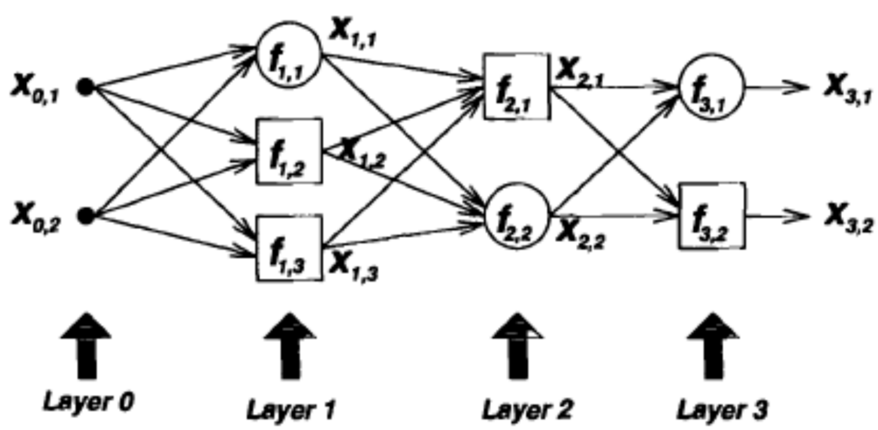


Figure 1 ANN model

This project reviews various techniques and focuses mainly on ANN for the Precipitation and Temperature forecasting in the Regina airport. The technique uses 10 inputs parameters to forecast the daily weather in terms of Humidity, dew point, temperature, cloud cover, Time, Wind Direction, Sea Pressure, and pressure station in order to forecast the precipitation and Temperature in the airport.

Weather forecasting is the application of science and technology to predict the state of the atmosphere for a given location. Weather forecasts are made by collecting quantitative data about the current state of the atmosphere and using scientific understanding of atmospheric processes to project how the atmosphere will evolve. The chaotic nature of the atmosphere, the massive computational power required to solve the equations that describe the atmosphere, error involved in measuring the initial conditions, and an incomplete understanding of atmospheric processes mean that forecasts become less accurate as the difference in current time and the time for which the forecast is being made increases. There are a variety of end uses to weather forecasts. Weather warnings are important forecasts because they are used to protect life and property. Forecasts based on temperature and precipitations are important to agriculture, and therefore to traders within commodity markets. Temperature forecasts are used by utility companies to estimate demand over coming days. On an everyday basis, people use weather forecasts to determine what to wear on a given day. Since outdoor activities are severely curtailed by heavy rain, snow and the wind chill, forecasts can be used to plan activities around these events, and to plan ahead and survive them. In order to predict weather in a very effective way and to help overcome all such problems we have proposed a weather forecasting model using ANN.

Problem Definition

The study region is City of Regina Saskatchewan, which is situated in the southeastern part of Canada, between Near the US border, 104.6553 western longitude and 50.4324 Northern latitude. There are several rivers and lakes in this region.

Figure 2 Topologic map of the region

It is a uniform surface. The average annual temperature high is 9 °C and the average annual temperature low is -3 °C. Therefore, the average difference in high and low temperature is about 12 °C which shows the temperature variation is quiet large in this region.

The average annual precipitation is about 390 mm. (15.3 in) The annual precipitation totals are averages of weather data collected during 1981 to 2010.

Figure 3 shows a graph of our normalized data for hourly temperature. Figure 4 shows a graph of our normalized data for hourly Precipitation.

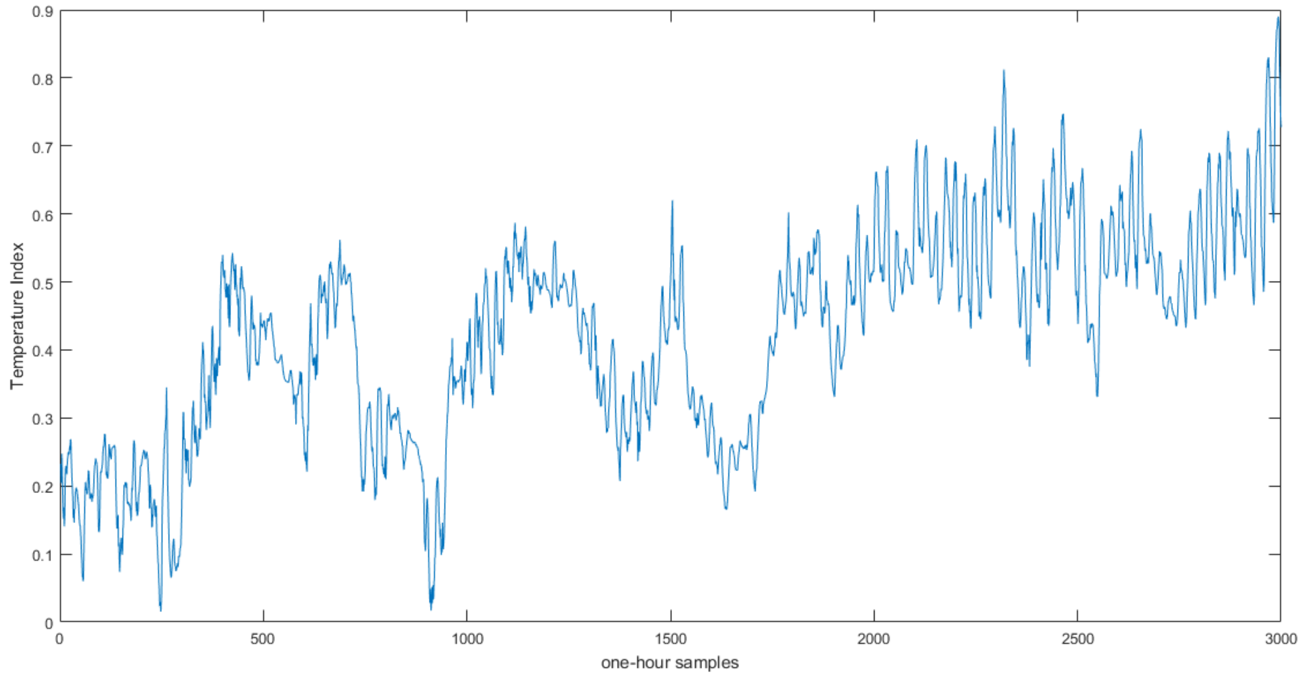


Figure 3 Normalized hourly temperature for first 3000 samples

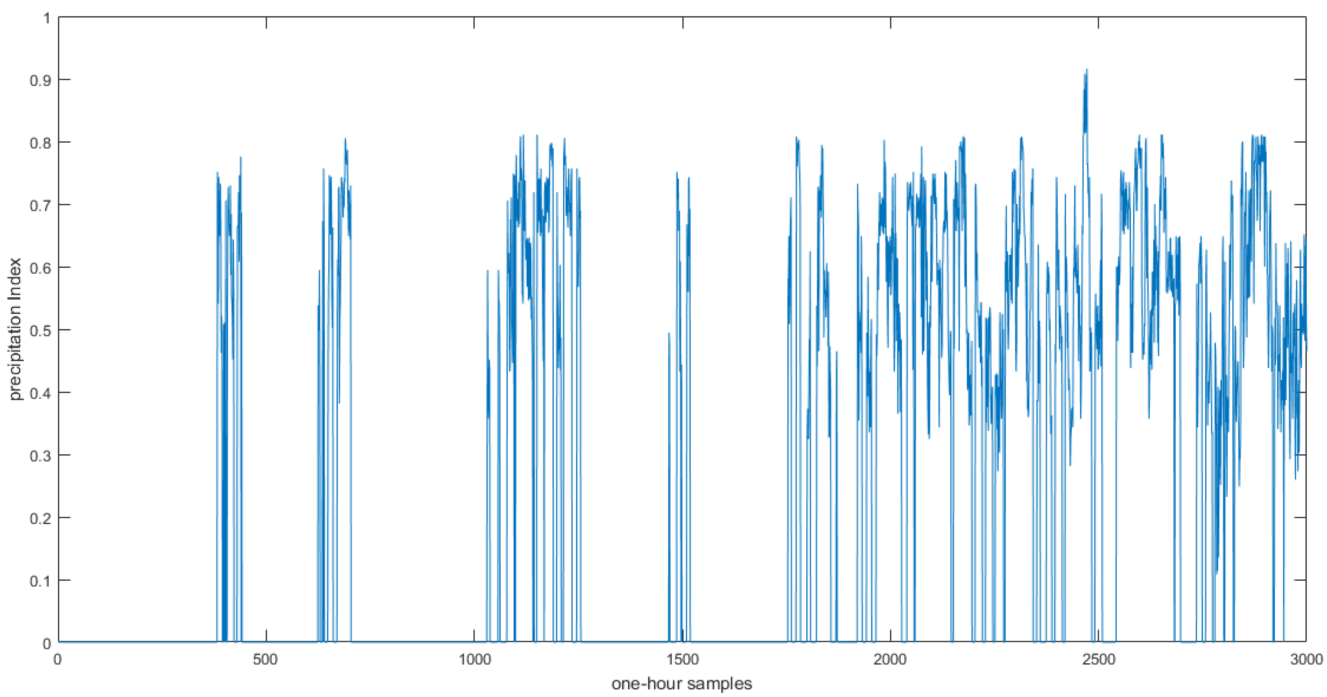


Figure 4 Normalized hourly Precipitation for first 3000 samples

Data Specification:

In this project, we got the Weather data from Regina Airport and Grassland Website. The main aim is to create and train, test, and check the data that can predict the individual weather components for e.g. temperature, humidity, and cloud cover, etc. We have used the hourly data from 2017. Our study has examined the temperature and precipitation because our main focus is on the design of the ANN which itself has numerous parameters to vary and optimize during experimentation. The model once created can then be feed with other weather factors in a similar fashion. This is a sample of our data from excel.

Data from 2017-Regina airport and Grassland Website

time	wind Direction	wind Speed	Humidity	dew Point	visibility	pressure Sea	cloud Cover	pressure station	temperature	Percipitation forecast
1	28	4	76	-32	24100	104.7	0	97.02	-29	0
2	23	4	75	-31.8	24100	104.8	0	97.05	-28.6	0
3	29	4	75	-31.9	24100	104.8	0	97.09	-28.8	0
4	27	14	74	-31	24100	104.8	0	97.09	-27.6	0
5	28	13	76	-29.5	24100	104.8	0	97.09	-26.5	0
6	29	11	78	-27.2	24100	104.7	0	97.07	-24.4	0
7	29	12	77	-25.9	24100	104.7	1	97.07	-22.9	0
8	29	15	79	-25.5	16100	104.7	7	97.07	-22.7	0
9	30	12	78	-26.3	16100	104.8	8	97.08	-23.4	0
10	30	13	76	-27.7	16100	104.8	8	97.09	-24.7	0
11	29	15	75	-29.1	19300	104.8	8	97.13	-26	0
12	29	16	73	-31.2	24100	104.9	7	97.15	-27.8	0
13	29	14	73	-31.9	24100	104.8	7	97.11	-28.6	0
14	28	8	73	-33.8	24100	104.8	7	97.08	-30.4	0
15	29	10	71	-35.2	24100	104.8	1	97.06	-31.6	0
16	27	10	71	-35.6	24100	104.8	1	97.06	-32	0
17	27	9	71	-35.4	24100	104.8	0	97.02	-31.8	0

Figure 5- Data from 2017-Regina Airport and Grassland Website

Input and Target Variables

The dataset consists of 8758 sample corresponding to the 1 year arranged column-wise in an Excel sheet which is later imported into the MATLAB workspace. The inputs and targets are listed in table1. 7000 samples were used in ANN system and other samples were reserved to be used for creditability of this research. Among these 7000 samples, neural network used 70% for training, 15% percent for checking and 15% for testing.

Table 1 Neural Network Variables

Variables	Distribution
Inputs: Time, Wind direction, Wind speed, Humidity, Dew point, Visibility, Sea Pressure, Cloud coverage, Station Pressure, Temperature	Hourly
Output: Temperature, precipitation	Hourly, 6 hours ahead

Methodology:

This project was done with help of MATLAB coding. When the data is prepared and organized in an excel datasheet, it needs to be imported to MATLAB and normalized. When the dataset is imported from an excel file, MATLAB recognized it as a table not matrix so we have to change it to matrix. Following code shows how it is implemented:

```
x=readtable('dataset.xlsx');
x=x(2:8759,1:9);
x=x{:,:};
y=size(x)
```

Normalization of data set is important because the different inputs has very different ranges. For example while the Pressure ranges between 92.2 KPa and 99.7 KPa (7.5 KPa difference), the temperature ranges between -35 to +41. Therefore the data set should be normalized using the following equation:

$$x_{new} = \frac{x - x_{min}}{x_{max} - x_{min}}$$

Which is done in Matlab with following 2 loops code:

```
%normalizing:

for j=1:y(1,2)
```

```
    for i=1: y(1,1)
    xnormal(i,j)=((x(i,j)-min(x(:,j))))/(max(x(:,j))-min(x(:,j)));
    end
 end
```

where y is the size matrix of the dataset matrix. (Here 8758 samples by 9 input and target variables).

It's also important to have a smooth dataset with no non-value or missing data, because usually large data sets consist of missing data. The following code is used to recover missing data by using interpolation. Since the dataset is arranged for every hour the average of previous and next element would be simply used for interpolation. The code is as follow:

```
%Eliminating NANS
mis_data=0;
 for i=1:y(1,1)
    for j=1:11
       if any(isnan(x(i,j)));
        x(i,j)=(x(i-1,j)+x(i+1,j))/2;
        mis_data=mis_data+1;
       end
    end
 end
```

The next step is to prepare the input and target data for training the ANN. Following code separate first 7000 samples to be used as training data in next step.

```
%Samples for ANN
   xinput_train= xnormal(1:8,1:7001);
   xoutput_train=xnormal(8:9,7:7007);
   inputs=xinput_train;
   targets=xoutput_train;
```

Now it's time to define the ANN system. It's consisted of two layers with 60 and 40 neurons in first and second hidden layer, respectively.

```
% Create a Fitting Network
hiddenLayerSize = [60 40];
net = fitnet(hiddenLayerSize);
```

70% of dataset is used for training, 15% for validation and 15% for testing. Dividing this portions is done randomly with ‘dividerand’ function, and dividing is do in every sampling step. ‘trainlm’ indicates that utilized training function is Levenberg-Marquardt. Mean square error is used as performance function. Number of epochs is 50, maximum validation fail is set to 50, time is infinite, and train parameter goal is 1e-4. Three plot functions are used: Plot histogram, Regression and Performance.

```
% Setup Division of Data for Training, Validation, Testing
net.divideFcn = 'dividerand';  % Divide data randomly
net.divideMode = 'sample';  % Divide up every sample
net.divideParam.trainRatio = 70/100;
net.divideParam.valRatio = 15/100;
net.divideParam.testRatio = 15/100;
% For help on training function 'trainlm' type: help trainlm
% For a list of all training functions type: help nntrain
net.trainFcn = 'trainlm';  % Levenberg-Marquardt

% Choose a Performance Function
net.performFcn = 'mse';  % Mean squared error

% Choose Plot Functions
net.plotFcns = {'plotperform','ploterrhist','plotregression', };

net.trainparam.epochs = 50;
net.trainParam.goal = 1e-4;
net.trainParam.show = 1;
net.trainParam.time=inf;
net.trainParam.max_fail=50;
```

Next part is training which is done by MATLAB command of ‘train’ as it is used in following:

```
% Train the Network
[net,tr] = train(net,inputs,targets);
```

It’s important to recalculate training, test and validate the network:

```
% Test the Network
outputs = net(inputs);
errors = gsubtract(targets,outputs);
performance = perform(net,targets,outputs)

% Recalculate Training, Validation and Test Performance
trainTargets = targets .* tr.trainMask{1};
valTargets = targets  .* tr.valMask{1};
```

```
testTargets = targets  .* tr.testMask{1};
trainPerformance = perform(net,trainTargets,outputs)
valPerformance = perform(net,valTargets,outputs)
testPerformance = perform(net,testTargets,outputs)

% View the Network
view(net)
```

Now the Training is done, ANN can be saved for future access. Normalized data is returned to their original units. A storing matrix is implemented with a loop to store inputs and predicted output for comparison between real data (which wasn't used for training ANN) and Predicted output of ANN.

```
begin=7119;
Real=[];
Predict=[];
for k=1:16
    xinput_test1= xnormal(1:8,begin);
    xtarget_test1=xnormal(8:9,begin+6);
    output_test1=net(xinput_test1);
Real(1,k)=xtarget_test1(1,:);
Predict(1,k)=output_test1(1,:);
Real(2,k)=xtarget_test1(2,:);
Predict(2,k)=output_test1(2,:);

begin=begin+53;
end

maxtemp=max(xinput_train(:,8))
mintemp=min(xinput_train(:,8))

maxprecip=max(xinput_train(:,9))
%minprecip=0

for l=2:16
    if Predict(2,l)<0
    Predict(2,l)=0;
    end
end

output1=(maxtemp-mintemp).*Predict(1,:)+mintemp;
real1=(maxtemp-mintemp).*Real(1,:)+mintemp;
output2= maxprecip.*Predict(2,:);
```

```
real2= maxprecip.*Real(2,:);

subplot(2,2,2)
plot(real1,'r-o');
hold on
plot(output1,'b-*');
legend('real','Predicted');
xlabel('number of predictions');
ylabel('Predicted Temprature');

subplot(2,2,1)
plot(real2,'r-o');
hold on
plot(output2,'b-*');
legend('real','Predicted');
xlabel('number of predictions');
ylabel('Predicted Percipitation');
```

Results and Discussions:

The ANN model code was run into MATLAB and Results described are as follows:

Fig. 1 represents the simulated model of the prediction model used for weather forcasting having 9 inputs , 2 outputs and 2 hidden layers having 60 and 30 neurons in layer 1 and layer 2 respectively.

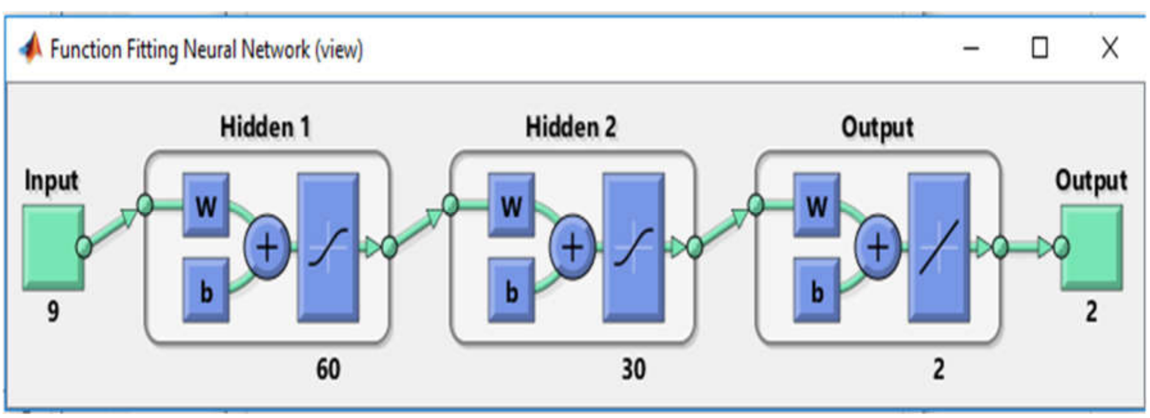


Fig. 6

Training of the proposed model is done and the final settled training parameters is shown in Fig. 7

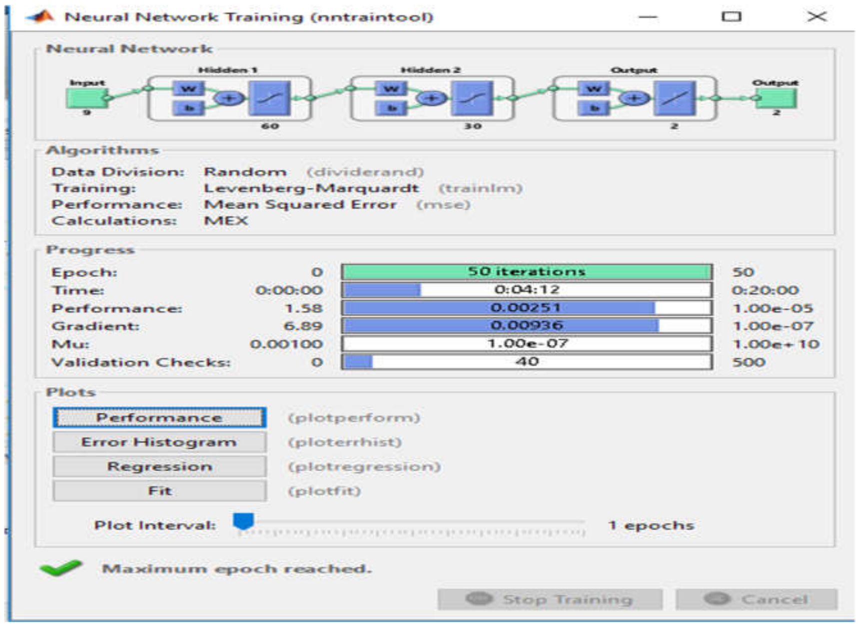
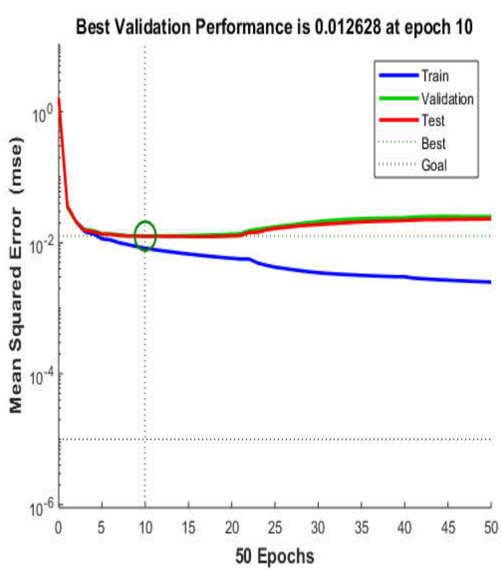


Fig.7

Training Plots are shown as follows



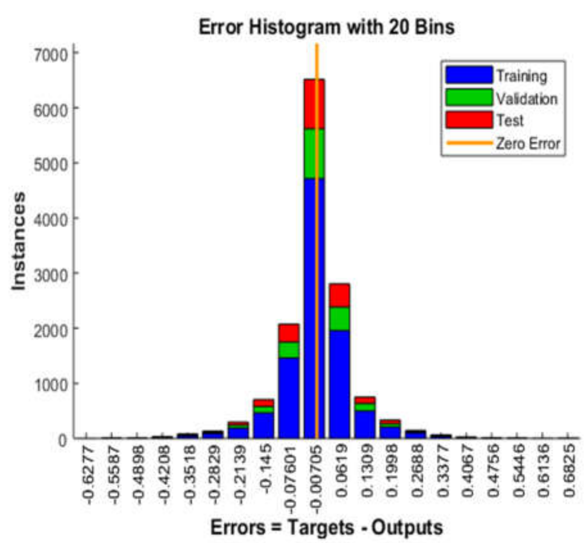
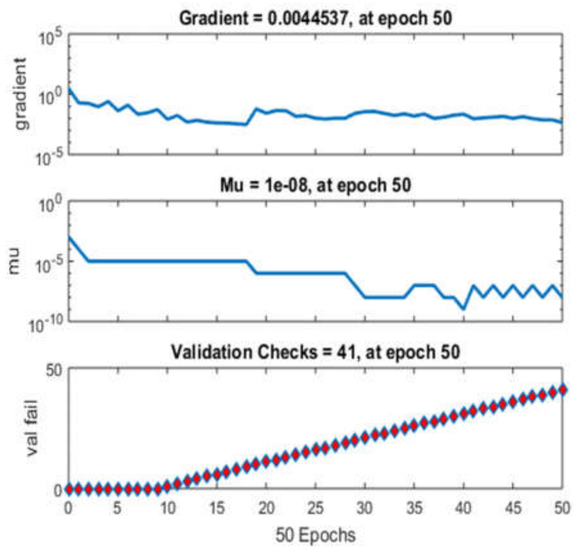


Fig. 8 System Performance

Fig. 9 Error Histogram



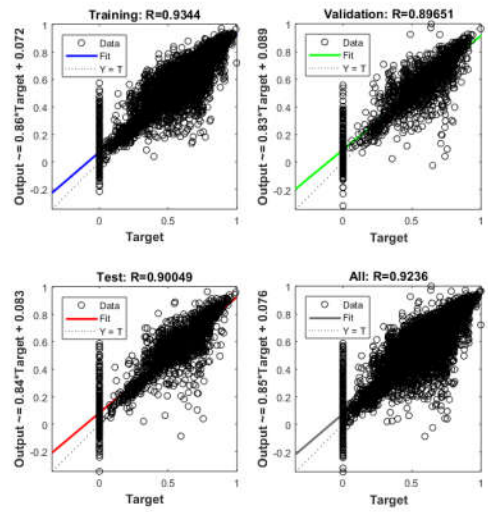


Fig. 10 Training State Plots

Fig. 11 Regression

After training the above model, system is fed input data after each 48 hours span and system output is generated and the generated out is being compared with the real data to verify the above model. The comparison graphs for precipitation and temperature parameters are shown in below curves.

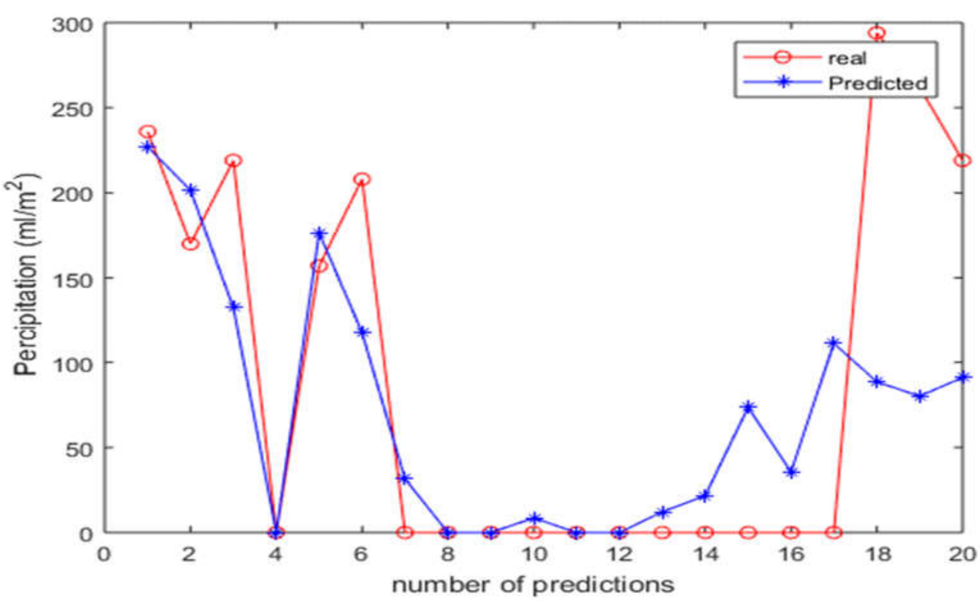


Fig. 12 Precipitation comparison(real vs predicted)

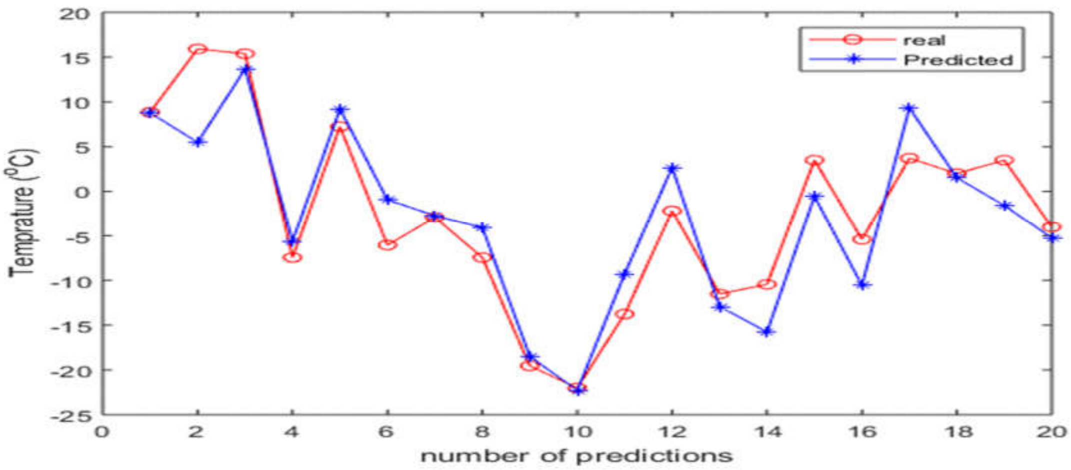


Fig. 13 Temperature comparison (real vs predicted)

Fig. 7 and 8 present the actual system performance by comparing the model output to the actual data. Hence results shows that proposed Weather forecasting model gives accurate results.

Conclusion and Future Works:

After observing the Output curves, it can be adhered that Temperature and Precipitation prediction can be done effectively with the proposed model. In most of the cases, predicted output curve is following the real data curve in both of the outputs i.e. Temperature as well as Precipitation.

Although System is showing promising results with less errors yet future work can be done to improve its effectiveness. Model can still be modifiable to get better accuracy by adding, subtracting hidden layers or adjusting hidden neurons. Model can be extended to used to predict other weather parameters like pressure, windspeed, cloud cover etc.

References:

1. Imran Maqsood, Muhammad Riaz Khan, and Ajith Abraham. An ensemble of neural networks for weather forecasting, Neural
Comput & Applic (2004) 13: 112–122.

2. Ball , R. , Tissot , P. (2006) Demonstration of Artificial Neural Network in Matlab

3. G. R. Gainieva, L. D. Nikitin, M. M. Naimark, N. N. Nazarov, and G. P. Tkachenko, Influence of Batch Composition and
Clinkering Properties on the Hot Strength of Coke and Blast-Furnace Operation, ISSN 1068-364X, Coke and Chemistry,
2008, Vol. 51, No. 10, pp. 390–393. © Allerton Press, Inc., 2008

4. *http://climate.weatheroffice.gc.ca/climateData/canada_e.html* .The website of the weather data of Canada

5. Haider, Adnan and Hanif, Nadeem, M., *Inflation Forecasting In Pakistan Using Artificial Neural Networks*

6. Dr. S. Santhosh Baboo and I.Kadar Shereef. An e_cient weather forecasting system using Artificial neural network. International Journal of Environmental Science and Development, 1(4):321-326, 2010.

7. Amanpreet Kaur, J K Sharma, and Sunil Agrawal. Artificial neural networks in forecasting maximum and minimum relative
humidity. International Journal of Computer Science and Network Security, 11(5):197-199, May 2011.

8. Sanjay Mathur, Avinash Kumar, and Mahesh Chandra. A feature based neural network model for weather forecasting, World
Academy of Science, Engineering and Technology 34 2007.

9. Mohsen Hayati and Zahra Mohebi. Application of artificial neural networks for temperatureforecasting. World Academy of
Science, Engineering and Technology, 28:275{279, 2007.

10. Rosmina Bustami, Nabil Bessaih, Charles Bong, and Suhaila Suhaili. Artificial neural network for precipitation and water
level predictions of bedup river. IAENG International Journal of Computer Science, 34(2):228-233, 2007.]

11. Sergio Caltagirone. Air temperature prediction using evolutionary arti_cial neural networks. Master's thesis, University of
Portland College of Engineering, 5000 N. Willamette Blvd. Portland, OR 97207, 12 2001.

https://gis.wsask.ca/Html5Viewer/index.html?viewer=WaterWells.WellsViewer
https://www.currentresults.com/Weather/Canada/Saskatchewan/Places/regina-temperatures-by-month-average.php
https://www.currentresults.com/Weather/Canada/Cities/precipitation-annual-average.php