Prediction of Rainfall Using Backpropagation Neural Network Model

Enireddy.Vamsidhar
Lecturer
Dept. of I.T.,
Gitam University
Visakhapatnam

K.V.S.R.P.Varma
Asst.Professor
Dept. of C.S.E,
Gitam University
Visakhapatnam

P.Sankara Rao
Asst.Professor
Dept. of C.S.E,
Gitam University
Visakhapatnam

Ravikanth satapati
Assst.professor
Dept. of C.S.E,
Gitam University
Visakhapatnam

Abstract

Agriculture is the predominant occupation in India, accounting for about 52% of employment. The Irrigation facilities are inadequate, as revealed by the fact that only 52.6% of the land was irrigated in 2009–10 which result in farmers still being dependent on rainfall, specifically the Monsoon season. A good monsoon results in a robust growth for the economy as a whole, while a poor monsoon leads to a sluggish growth. . Artificial neural network is one of the most widely used supervised techniques of data mining. In this paper we used the back propagation neural network model for predicting the rainfall based on humidity, dew point and pressure in the country INDIA. Two-Third of the data was used for training and One-third for testing .The number of training and testing patterns are 250 training and 120 testing .In the training we obtained 99.79% of accuracy and in Testing we obtained 94.28% of accuracy. From these results we can predict the rainfall for the future.

Key Words: Agriculture, Neural Networks, Back Propagation, Prediction

Introduction

The economy of India is the twelfth largest economy in the world by nominal value [1] and the fourth largest by purchasing power parity (PPP)[2].In the 1990s, following economic reform from the socialist-inspired economy of post-independence India, the country began to experience rapid economic growth, as markets opened for international competition and investment. In the 21st century, India is an emerging economic power with vast human and natural resources, and a huge knowledge base. Economists predict that by 2020 [3] India will be among the leading economies of the world.

Agriculture is the predominant occupation in India, accounting for about 52% of employment. The service sector makes up a further 34%, and industrial sector around 14% [4] Agricultural and allied sectors accounted for about 60% of the total workforce in 2009 same as in 1993–94.Most of the sub-continent depends on the rainfall for the agriculture needs. The rainfall data is available for the data mining techniques which can be useful for predicting the rainfall which can very useful for taking decisions over crop planting in the areas.

Artificial neural networks have been extensively used in these days in various aspects of science and engineering because of its ability to model both linear and non-linear systems without the need to make assumptions as are implicit in most traditional statistical approaches. ANN has been a aggressive model over the simple linear regression model.[5]

Backpropagation, or propagation of error, is a common method of teaching artificial neural networks how to perform a given task. It was first described by Arthur E. Bryson and Yu-Chi Ho in 1969,[6],[7]but it wasn't until 1986, through the work of David E. Rumelhart, Geoffrey E. Hinton and Ronald J. Williams, that it gained recognition, and it led to a “renaissance” in the field of artificial neural network research.

It is a supervised learning method, and is an implementation of the Delta rule. It requires a teacher that knows, or can calculate, the desired output for any given input. It is most useful for feed-forward networks (networks that have no feedback, or simply,

that have no connections that loop). The term is an abbreviation for "backwards propagation of errors". Backpropagation requires that the activation function used by the artificial neurons (or "nodes") is differentiable

In this paper the rainfall data has been taken from the period 1901-2000 and the data was taken from the website(www.tyndall.ac.uk).The quantative prediction of the monthly rainfall by backpropagation neural network is examined.

Methodology

Predictive mining is a task that it performs inference on the current data in order to make a prediction. Here the rainfall data can be grouped as a time series set because it consists of sequence of values in time. A time series data can be denoted as

Y=f(t)

Where y can be any single valued variable which develops in time t, in this work, y is a monthly rainfall values to forecast time-series data, it involves knowing the past history of f and extrapolating it to the future. The characteristic of the forecasting model is non-linear system, so that the backpropagation neural network can be applied in time-series prediction areas. Min-max normalization performs a linear transformation on the original data. Suppose that min_A and max_A are the minimum and maximum values of an attribute, A. Min-max normalization maps a value, *v*, of A to *v'* in the range [new min_A; new max_A] by computing

$$v' = \frac{v - min_A}{max_A - min_A}(new\_max_A - new\_min_A) + new\_min_A$$

Min-max normalization preserves the relationships among the original data values.
It will encounter an “out-of-bounds” error if a future input case for normalization falls
Outside of the original data range for A. In this work, we set the new_min_A to 0.0 and new_max_A to 1.0.

Neural network architecture

A three layer feed-forward neural network architecture was created by initializing the weights of the neural network by random values in between -1.0 to 1.0.The data were processed into 11 variable: SUM(RAIN)[t],SUM(RAIN)[t-1 to t-10].

The input nodes correspond to summary of monthly rainfall from the year 1901 to the year 2000.In this work, the architecture of neural network in this research is 3:7:1(input node: hidden node: output node).If n is the number of input nodes then the number of hidden layer nodes are 2n+1.

Data preprocessing

A monthly rainfall data in the period of 1901-2000 were collected and some data preprocessing steps on raw set of monthly rainfall data as shown below:

1) Firstly, a monthly rainfall data were cleaned by filling in missing values with mean values.
2) Secondly, a monthly rainfall data were normalized by min-max normalization into a specified range 0.0 to 1.0.

Actual Data

Attribute/type of data	Pressure	Humidity	Dew Point
Actual Data	992.1632	95.6122	6.6571
Normalized values	0.956122	0.949216	0.500078

Dividing 2/3 rd as training and 1/3rd for testing.

No. of inputs(n)	Attribute1	Attribute2	Attribute3	No of layers in hidden layer(2n+1)	Class label (out put node)
3	0.956122	0.949216	0.500078	7	0.500078

By using neural network we construct the model for training set and consider test set by removing the actual class label or placing the class label=0 and apply it on training model by that we predict the class label value for test data and compare with the actual class label to get mean square error.

No. of inputs(n)	Attribute 1	Attribute 2	Attribute 3	No of layers in hidden layer(2n+1)	Class label (output node)
3	0.73449	0.962571	0.469931	7	0

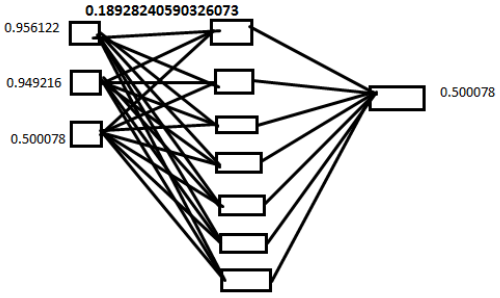

Fig 1. Architecture of network

Estimation of Accuracy

To estimate accuracy [8] of a prediction, the accuracy is defined as

$$mse = [\frac{\frac{1}{2}\sum_{i=1}^{N}(t_i - o_i)^2}{N}]$$

Accuracy =100-*mse.*

Data type	Mean Square Error	Accuracy
Training	0.21026952935196724	99.79
Testing	5.824254227913425	94.28

Experimental Results

The Back propagation neural network was used in this work which established accuracy 99.79%, in the training stage we got accuracy of 94.28%. The corresponding Mean Square Error for each data type is as shown in the Table.

Data type	Mean Square Error	Accuracy
Training	0.21026952935196724	99.79
Testing	5.824254227913425	94.28

Conclusion

For rainfall prediction, Artificial Neural Network was applied and the rainfall was predicted in India. According to the results backpropagation neural network were acceptably accurate and can be used for predicting the rainfall. So by using this method for prediction we can find the amount of rainfall in the region by using the attributes like humidity, dew point and pressure.

References

[1] India twelfth wealthiest boss in 2005: World Bank"". The Economic Times. http://economictimes.indiatimes.com/Mr_Rupee_pulls_India_into_1_trillion_GDP_gang/articleshow/1957520.cms. Retrieved 2006-07-08.

[2] "CIA — The World Factbook — Rank Order — GDP (purchasing power parity)". Cia.gov. 2009-03-05. https://www.cia.gov/library/publications/the-world-factbook/rankorder/2001rank.html. Retrieved 2009-03-13.

[3] India Vision 2020" (PDF). http://planningcommission.gov.in/reports/genrep/pl_vsn2020.pdf. Retrieved 2009-12-12.

[4] "CIA - The World Fact book - India". CIA. 2007-09-20. https://www.cia.gov/library/publications/the-world-factbook/geos/in.html#Econ. Retrieved 2007-10-02.

[5] S.Lee,S.Cho and P.M.Wong 1998,"Rainfall prediction using Artificial neural networks". journal of geographic information and Decision Analysis,Vol .2,No2,pp.233-242

[6] Stuart Russell and Peter Norvig. *Artificial Intelligence A Modern Approach*. p. 578. "The most popular method for learning in multilayer networks is called Back-propagation. It was first invented in 1969 by Bryson and Ho, but was more or less ignored until the mid-1980s."

[7] Arthur Earl Bryson, Yu-Chi Ho (1969). *Applied optimal control: optimization, estimation, and control.* Blaisdell Publishing Company or Xerox College Publishing. pp. 481.

[8] Jiawei Han and Micheline Kamber, "Data Mining Concepts and Techniques " ELSEVIR Second Edition pp.71