Journal of Computer Sciences

Investigations

Backpropagation Vs. Radial Basis Function Neural Model: Rainfall Intensity Classification For Flood Prediction Using Meteorology Data

[1]Soo See Chai, [1]Wei Keat Wong and [2]Kok Luong Goh

[1]*Department of Software Engineering and Computing, Faculty of Computer Science and Information Technology, University of Malaysia Sarawak (UNIMAS), 94300, Kota Samarahan, Sarawak, Malaysia*
[2]*Faculty of Cognitive Science and Human Development, UNIMAS, 94300, Kota Samarahan, Sarawak, Malaysia*

Article history
Received: 18-11-2014
Revised: 11-04-2016
Accepted: 14-04-2016

Corresponding Author:
Soo See Chai
Department of Software Engineering and Computing, Faculty of Computer Science and Information Technology, University of Malaysia Sarawak (UNIMAS), 94300, Kota Samarahan, Sarawak, Malaysia
Email: sschai@fit.unimas.my

Abstract: Rainfall is one of the important weather variables that vary in space and time. High mean daily rainfall (>30 mm) has a high possibility of resulting in flood. Accurate prediction of this variable would save human lives and properties. Soft computing methods have been widely applied in this field. Among the various soft computing methods, Artificial Neural Network (ANN) is the most commonly used methodology. While numerous ANN algorithms were applied, the most commonly applied are the Backpropagation (BPN) and Radial Basis Function (RFN) models. However, there was no research conducted to verify which model among these two produces a superior result. Therefore, this study will fill this gap. In this study, using the meteorology data, the two ANN models were trained to classify the rainfall intensity based on four different classes: Light (<10 mm), moderate (11-30 mm), heavy (31-50 mm) and very heavy (>51 mm). The architecture of the neural networks models based on the different combination of inputs and number of hidden neurons to obtain the optimum classification were verified in this study. The influence of the number of training data on the classification results was also analyzed. Results obtained showed, in term of classification accuracy, BPN model performed better than the RFN model. However, in term of consistency, the RFN model outperformed BPN model.

Keywords: Precipitation, Classification, Backpropagation, Radial Basis Function, Neural Networks

Introduction

Weather forecasting is a complicated procedure yet the most essential and vital process for the mankind nowadays, because it severely affect human activities. Highly accurate weather forecast could help to prevent casualties and damages. Amongst all the weather happenings, floods are the leading cause of natural disaster death world-wide and were responsible for 6.8 million death in the 20th century (Doocy *et al*., 2013). Rainfall intensity is important for flood warning system. Considering the alert system involved in heavy rain management, it would be useful to classify the rainfall intensity based on different threshold. The depth or intensity of the rainfall and its distribution in the temporal and spatial dimensions depend on many variables, such as pressure, temperature, wind speed and direction (Luk *et al*., 2001). Understanding the complex physical processes that create the rainfall is very challenging. Large number of attempts has been made by researchers to accurately predict rainfall. However, the accuracy obtained by these techniques is still below satisfactory level due to the nonlinear nature of rainfall (Nayak *et al*., 2013). Artificial Neural Network (ANN), which has the ability in handling complex and non-linear problems, has drawn the attention of researchers in the field of weather forecasting. Among the different architectures of ANN, the Backpropagation Network (BPN) and Radial Basis Function (RBF) network are the two main models that are sufficiently suitable for precipitation prediction (Shrivastava *et al*., 2012). Although a lot of works have been done using these two architectures, the superiority of one architecture over another is not being discussed. Therefore, this paper aims to compare and analyze the performance of these two architectures for rainfall classification.

Data

The study area selected for this study is Kuching city, the capital city of Sarawak located in the east Malaysia. Since Malaysia is located on top of equator, Kuching city has a tropical type of climate with average of five to six hours of sunshine, high temperature and high humidity. The city, located at the southwest of the Sarawak state with latitude 1.6019N and longitude 110.3244E, covers area of 895.09 km^2 and a population of 681,901 (Wikipedia 2015). A collection of historical meteorology parameters of daily measurement was obtained from the Malaysian Meteorological Department. These daily meteorology data from year 2009 to 2013 consisted of seven elements: Minimum temperature (°C), maximum temperature (°C), mean temperature (°C), mean relative humidity (%), mean wind speed (m/s), mean sea level pressure (hPa) and mean precipitation (mm).

Data Pre-Processing

Noise and missing data would affect the performance of ANN models (Sola and Sevilla, 1997). Thus, before training and testing the ANN, it is important to perform data checking and cleansing to maximize the performance of ANN forecasting. In this research, missing and incomplete data which was confirmed by the Sarawak Meteorology Station was deleted from the database.

Data Normalization

The pre-processed database is next gone through normalization. Normalization aims to produce good result and prevent numerical difficulties occurs when performing calculation (Chen *et al*., 2013). Moreover, according to (Chai *et al*., 2009), normalization speeds up the training process of the ANN and reduces the likelihood of the ANN getting stuck in local minima. Adequate data normalization before applying it into the ANN can reduce the estimation error generated by the ANN in a factor between 5 and 10 (Sola and Sevilla, 1997). In this study, the input data was normalized so that the minimum and maximum values for each input row are between +1 and -1.

Rainfall Intensity Classification

The rainfall intensity for year 2009 to 2013 is classified into four classes: Light precipitation (<10 mm), moderate precipitation (11-30 mm), heavy precipitation (31-50 mm) and very heavy precipitation (> 51 mm). In the research to estimate rainfall using radar for the Klang River Basin in Selangor, Malaysia (Ramli *et al*., 2011), three classes namely low (<10 mm), moderate (>10, <30 mm) and heavy (>30 mm) were used. However, in this study, the heavy precipitation class (>30 mm) was sub-divided into 2 classes: Heavy precipitation (31-50 mm) and very heavy precipitation (>51 mm) as rainfall of more than 50 mm could be termed as "hazard precipitation" (Szalińska *et al*., 2014). Table 1 summarized the different classes used in this study.

Methodology

Input and Output of ANN Models

The input of the ANN models included 6 meteorological data obtained from Department of Irrigation and Drainage (DID) Kuching Division of Sarawak, Malaysia. These six parameters were: Daily minimum temperature (°C), maximum temperature (°C), mean temperature (°C), mean relative humidity (%), mean wind speed (m/s), mean sea level pressure (hPa). Each input node of the ANN models consists of an array of different parameter values at a different time period. The output of the ANN models would be the different class of the rainfall intensity as shown in Table 1.

Data Discretization for Training and Testing Process

The ANN models will be trained by providing "examples" for the models to learn. The "well-learned" ANN will next be tested with some unseen data. In order to accomplish this, the meteorology data obtained were divided into training and testing data according to Table 2. Table 2 showed that, in order to make a fair comparison of the ANN models, the same set of data, i.e., 1 month data ranging from 1 to 31 Dec. 2013, was used for testing. The training of the data was divided into 5 groups of different amount of training data.

BPN Architecture

One single hidden layer feed forward network can approximate any measureable function arbitrarily well regardless of the activation function, the dimension of the input space and the input space environment (Hornik *et al*., 1989). Therefore, in this study, one hidden layer Backpropagation neural Network (BPN) with Levenberg-Marquardt learning algorithm is used.

RFN Architecture

Radial functions are class of functions which could be applied in any sort of model (linear or non-linear) and any sort of network (single- or multi-layer) (Orr, 1996). Moreover, the single-weight layer network (the input component is feed-forward to the basis functions whose outputs are linearly combined with weights into the network outputs) is associated with the traditional RFN model (Broomhead and Lowe, 1988). Consequently, in this study, the single-weight layer RFN model with Gaussian learning algorithm is used.