

Ruziicka Similarity Feature Selection Based Generalized Linear Regression Analysis For Weather Forecasting Using Big Data

Dr.R.Pushpalatha

Assistant Professor, Department of Computer Science, Kongu Arts and Science College (Autonomous), Nanjanapuram, Erode, Tamil Nadu, India. rpljour@gmail.com
Orcid: 0000-0001-8834-7284

Article History: Received: 10 January 2021; Revised: 12 February 2021; Accepted: 27 March 2021; Published online: 28 April 2021

Abstract

Weather forecasting is the process of finding state of atmosphere at a future time and a particular location. Few research works have been developed for weather forecasting with help of various machine learning techniques. But, prediction performance of conventional machine learning technique was not enough to accurately find weather conditions. In order address existing issues, Ruziicka similarity Feature Selection Based Generalized Linear Regression Analysis (RSFS-GLRA) technique is proposed. RSFS-GLRA technique takes weather data from cloud server as input. The designed RSFS-GLRA technique performs feature selection and prediction process. Ruziicka Similarity-Based Feature Selection (RS-FS) process is carried out to select relevant features for performing weather forecasting. After feature selection, Generalized Linear Regression Analysis based Weather Forecasting (GLRA-WF) process predicts future weather conditions. GLRA-WF algorithm is a powerful statistical method that examines relationship between two or more variables for finding future event of weather conditions according to collected historical data. This helps RSFS-GLRA technique to perform accurate weather forecasting process with minimal time and higher accuracy. Experimental evaluation of RSFS-GLRA technique is carried out using parameters such as prediction accuracy, error rate and prediction time with respect to various numbers of weather data.

Keywords: Big Data, Cloud, Feature selection, Generalized Linear Regression Analysis, Ruziicka Similarity, Weather forecasting.

Introduction

Weather forecasting plays an imperative role in daily routine, businesses and their decisions. Weather forecasting is very hard for given the number of attributes involved and complex relations between variables. Dramatic increase in data collection has improved ability of weather forecasters to identify timing and harshness of hurricanes, floods, snowstorms, and other weather events. However, prediction accuracy was not sufficient predict weather conditions with minimal time. A track-similarity-based Dynamical-Statistical Ensemble Forecast (LTP_DSEF) model was introduced in [1] to predict land falling tropical cyclones. However, prediction accuracy using LTP_DSEF model was not adequate. Support vector regression (SVR) approach was applied in [2] with aim of minimizing tropical cyclone prediction errors. But, time complexity was high.

A novel technique was presented in [3] to get better clustering accuracy for accomplishing weather prediction process. However, prediction accuracy was low. Time-hierarchical Clustering was carried out in [4] for analyzing sequential growth of uncertainty in weather forecasts with minimal time complexity. But, false positive rate of weather forecasting were high.

The dynamic self-organized neural network algorithm was employed in [5] for weather data forecasting process with higher accuracy. However, computational complexity during weather prediction process was more. Fuzzy C-means clustering was applied in [6] for bunching different weather data and increasing prediction performance. But, prediction accuracy was not at required level.

In order to overcome above mentioned conventional issues, Ruziicka similarity Feature Selection Based Generalized Linear Regression Analysis (RSFS-GLRA) technique is introduced in this research work. The main contributions of RSFS-GLRA technique are:

- To achieve better weather prediction performance, RSFS-GLRA technique is proposed. RSFS-GLRA technique is introduced by combining Ruziicka Similarity Based Feature Selection (RS-FS) algorithm and Generalized Linear Regression Analysis based Weather Forecasting (GLRA-WF) algorithm.
- To select relevant features for accurate weather prediction, RS-FS Algorithm is introduced in RSFS-GLRA technique. RS-FS Algorithm used Ruziicka Similarity coefficient measurements for feature selection process with lower time complexity.
- To increase weather forecasting accuracy with lesser time complexity, GLRA-WF algorithm is used in RSFS-GLRA technique. GLRA-WF algorithm is a statistical method for predicting future event of weather condition by collected historical data.

The rest of article structure is created as follows. In Section 2, proposed RSFS-GLRA technique is explained with the aid of architecture diagram. In Section 3, experimental settings are presented and the performance result of RSFS-GLRA technique is discussed in Section 4. Section 5 illustrates the literature survey. Section 6 portrays the conclusion of paper.

Materials and Methods