Research Article

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

Ruziicka Similarity Feature Selection Based Generalized Linear Regression Analysis Technique

Because of size and complexity of weather patterns, it very complex to forecast weather in advance. A lot of research works have been devloped for weatehr forecasting using different machine learning techniques. However, prediction accuracy and time complexity of was not at required level. Therefore, Ruziicka similarity Feature Selection Based Generalized Linear Regression Analysis (RSFS-GLRA) technique is designed to improve the performance of weather forecasting with higher accuracy and minimal time while taking a big data as input. RSFS-GLRA technique is proposed by Ruziicka Similarity-Based Feature Selection (RS-FS) algorithm and Generalized Linear Regression Analysis based Weather Forecasting (GLRA-WF) algorithm. The architecture diagram of RSFS-GLRA technique is presented in Figure 1.

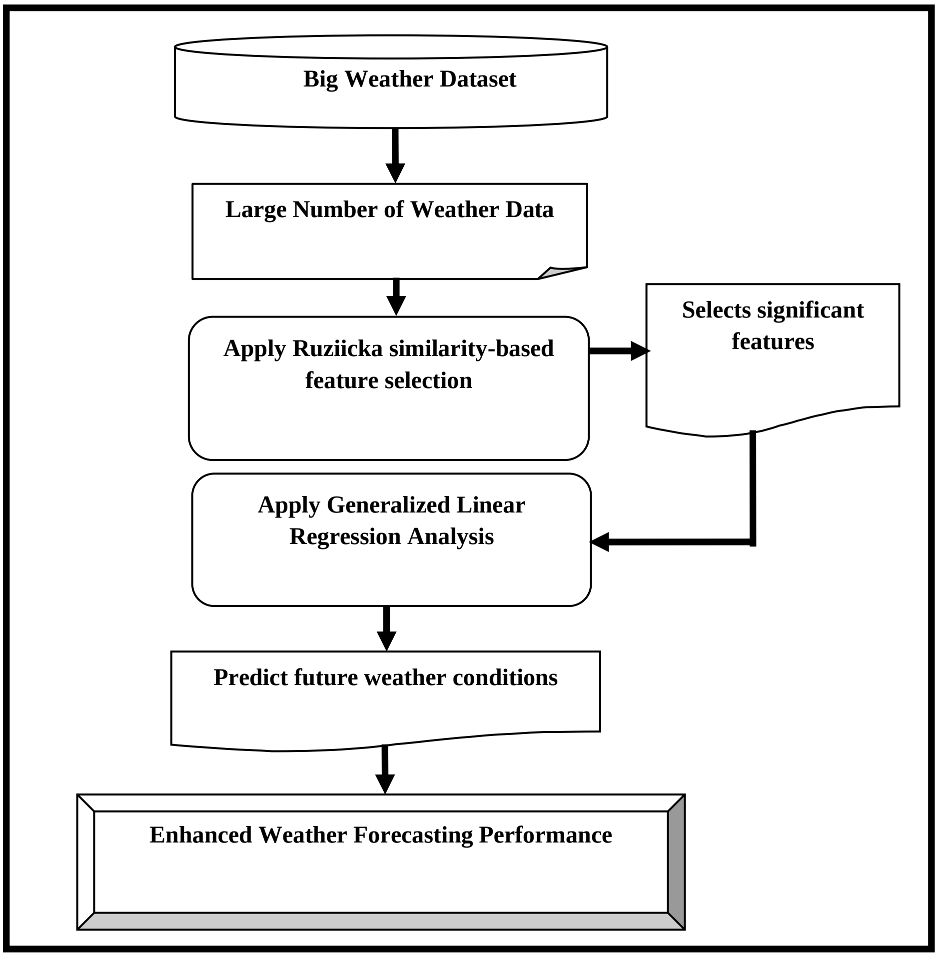

Figure 1 Architecture Diagram of RSFS-GLRA Technique For Weather Forecasting

Figure 1 demonstrates overall processes of RSFS-GLRA technique to get enhanced weather forecasting performance. As presented in the below figure, the RSFS-GLRA technique includes of two main processes namely feature selection and prediction for accurate weather forecasting. Initially, RSFS-GLRA technique obtains big weather dataset as input which comprises numerous numbers of weather data and features. During feature selection process, RSFS-GLRA technique identifies most relevant features in input dataset by using Ruziicka similarity-based feature selection to perform weather forecasting process with a minimal amount of time consumption. During the prediction process, RSFS-GLRA technique finds relationship between the input weather data by using Generalized Linear Regression Analysis. RSFS-GLRA technique achieves better prediction accuracy for weather forecasting as compared to existing works. The detailed processes of RSFS-GLRA technique are described in below subsections.

1. Ruziicka Similarity-Based Feature Selection

In RSFS-GLRA technique, Ruziicka Similarity-Based Feature Selection (RS-FS) algorithm is proposed with aiming at increasing the performance of feature selection for weather forecasting. Ruziicka similarity-based feature selection (RS-FS) initially selects the relevant features from the big weather dataset. The big input dataset contains the large number of features for processing the input data. This causes the more complexity during the weather prediction process. As a result, the proposed RS-FS selects the relevant features to reduce the complexity of weather forecasting process. Feature selection is a preprocessing technique which helps to remove the irrelevant features from the big weather dataset. The RS-FS uses the Ruziicka similarity Coefficient for discovering the similarity between the features. The process involved in RS-FS algorithm is demonstrated in Figure 2.

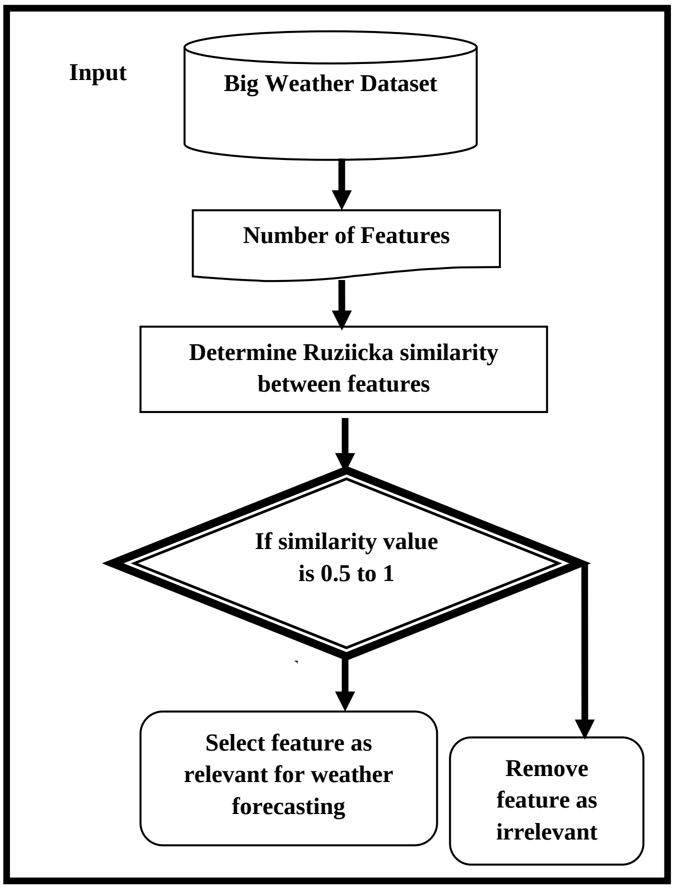

Figure 2 RS-FS algorithm Process for Relevant Feature Selection

Figure 2 depicts the block diagram of RS-FS algorithm. As depicted in the above figure, Let us consider number of features $\mathrm{fr}_1, \mathrm{fr}_2, \mathrm{fr}_3, \ldots . \mathrm{fr}_n$ in big weather database 'WD' as input. In RS-FS algorithm, Ruziicka similarity between features is determined by,

$$\mathrm{R_S}(\mathrm{fr}_1, \mathrm{fr}_2) = \frac{\sum_i \min(\mathrm{fr}_1, \mathrm{fr}_2)}{\max(\mathrm{fr}_{1_i}, \mathrm{fr}_2)} \qquad (1)$$

From (1), 'min' and 'max' represent the pointwise operators. When the similarity value lies between '0' to '0.5', then the feature is not relevant to weather prediction. When the similarity value lies between '0.5' to '1', then feature is relevant to weather prediction. The high similarity features are selected for accurate weather forecasting. The algorithmic process of RS-FS is described in below,

// Ruziicka Similarity-Based Feature Selection Algorithm

Input: Big weather dataset WD, Number of features $fr_1, fr_2, fr_3, \ldots . fr_n$

Output: Select relevant features

Begin

1. Determine Ruziicka Similarity between the two features fr_1, fr_2
2. **If** $R_S\ (fr_1, fr_2) =' 0.5\ to + 1$' **then**
3. High similarity between the features
4. Select the features as relevant for performing weather forecasting process
5. **Else**
6. Lower similarity between the features
7. Remove the features as irrelevant
8. **End if**

End

Algorithm 1 Ruziicka Similarity-Based Feature Selection

Algorithm 1 presents algorithmic processes of features selection using Ruziicka similarity measurement. Initially, numbers of features are taken from big weather database. Ruziicka similarity between features is computed to find high and low similarity features. RS-FS effectively selects similar features and removes irrelevant features. With selected features, weather prediction is performed with lesser prediction time.

2. Generalized Linear Regression Analysis based Weather Forecasting

In RSFS-GLRA technique, Generalized Linear Regression Analysis based Weather Forecasting (GLRA-WF) algorithm is introduced in order to improve the weather forecasting performance with minimal time complexity. The GLRA-WF algorithm designed in RSFS-GLRA technique is a powerful statistical method that examines relationship between two or more variables of interest for predicting the future event of weather conditions based on collected historical data. This supports for GLRA-WF algorithm for performing an accurate weather forecasting process. The GLRA-WF algorithm is a machine learning technique. GLRA-WF algorithm accomplishes a regression analysis in which it determines the relationship between the input weather data and thereby finds future weather conditions accurately.

In RSFS-GLRA technique, the GLRA-WF algorithm is a statistical model for performing weather prediction process. GLRA-WF algorithm initially evaluates the relationship between dependent variable and one or more independent variables. Here, the independent variables are input big weather data with chosen key features and the dependent variable is the output (i.e. future weather conditions). The proposed GLRA-WF algorithm to regression allows us to determine the parameters of a linear model. GLRA-WF algorithm is to discover linear model that reduces sum of the squared errors between observations in dataset and those predicted by model.

GLRA-WF algorithm exactly identifies future weather conditions with lesser time consumption. The process involved in GLRA-WF algorithm is presented in Figure 3.

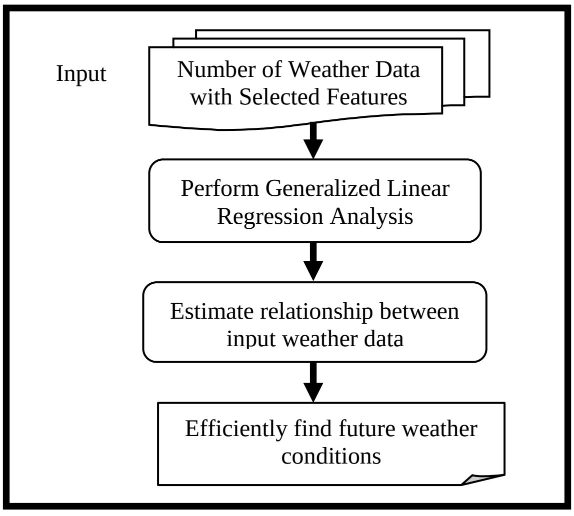

Figure 3 Processes of GLRA-WF algorithm for Weather Forecasting

Figure 3 shows the flow process of GLRA-WF algorithm for effective weather forecasting. As depicted in above block diagram, GLRA-WF algorithm initially takes a number of weather data with chosen features as input. Let us consider a large number of weather data are collected from a cloud are represented as '$\alpha_1, \alpha_2, \alpha_3, \ldots, \alpha_n$'. After that, GLRA-WF algorithm carried outs the Generalized Linear Regression Analysis in order to find out a dependent variable value 'β' (i.e. future weather conditions) based on a given independent variable 'α' (i.e. input weather data with selected features). During the regression analysis process, GLRA-WF algorithm models the linear relationship between input weather data. Thus, Generalized Linear Regression Analysis is mathematically performed as follows,

$$\beta = a_1 + a_2.\alpha \qquad (2)$$

From the above mathematical expression (2), 'β' denotes result of GLRA-WF algorithm i.e. future weather conditions and 'α' refers to an input training data (i.e. input weather data with selected relevant features).

While performing regression process, GLRA-WF algorithm fits the best line to predict the 'β' value for an input weather data 'α'. The GLRA-WF algorithm applied in RSFS-GLRA technique obtains the best regression fit line by means of estimating the best 'a_1' and 'a_2' values. Here, 'a_1' symbolize intercept and 'a_2' signifies coefficient of 'α'. By using the best 'a_1' and 'a_2' values, GLRA-WF algorithm determines the best fit line. Thus, GLRA-WF algorithm makes a decision of future weather conditions. In this manner, GLRA-WF algorithm precisely carried out weather forecasting process in order to predict the future weather conditions.

Through determining the best-fit regression line, the GLRA-WF algorithm finds 'β' value such that the error dissimilarity between predicted output and actual output is very lower. As a result, it is very essential to update the 'a_1' and 'a_2' values in GLRA-WF algorithm to get best value that decreases the error between discovered output and actual output. From that, error function is mathematically determined using below,

$$\varepsilon rr = \frac{1}{n}\sum_{i=1}^{n}(\rho_O - \tau_O) \qquad (3)$$

From the above mathematical formula (3), error function 'εrr' is measured as the Root Mean Squared Error between predicted output 'ρ_O' and actual output 'τ_O'. After computing the error, GLRA-WF algorithm updates 'a_1' and 'a_2' values in order to diminish error function and thereby finding the best fit line by using gradient descent which is mathematically expressed as follows,

$$\beta = arg\ min\ \varepsilon rr \qquad (4)$$

$$\beta = arg\ min \frac{1}{n}\sum_{i=1}^{n}(\rho_O - \tau_O) \qquad (5)$$

From (4) and (5), proposed GLRA-WF algorithm accurately detects future weather conditions with a minimal error rate when compared to conventional works. The algorithmic processes of GLRA-WF is explained in below,

// **Generalized Linear Regression Analysis Algorithm**
Input: Number of weather data '$\alpha_1, \alpha_2, \alpha_3, \ldots, \alpha_n$'; Selected features
Output: Achieve higher accuracy for weather prediction

- **Step 1: Begin**
- **Step 2: For** weather data with chosen key features 'α_i'
- **Step 3:** Carry out Generalized Linear Regression Analysis using (2)
- **Step 4:** Predict future weather conditions
- **Step 5:** Determine error function 'εrr' using (3)
- **Step 6:** Reduce error function using (4) and (5)
- **Step 7: If** '($\beta == 1$)', **then**
- **Step 8:** Future weather conditions is correctly predicted
- **Step 9: Else**
- **Step 10:** Future weather conditions is not predicted correctly
- **Step 11: End If**
- **Step 12: End For**
- **Step 13: End**

Algorithm 2 Generalized Linear Regression Analysis

Algorithm 2 presents the step by step processes of GLRA-WF algorithm. By using the above algorithmic process, GLRA-WF algorithm at first applies Generalized Linear Regression Analysis for each weather data with chosen key features. Next, GLRA-WF algorithm predicts future weather conditions by estimating relation between the data. After that, GLRA-WF algorithm measures error between discovered output and actual output. Finally, GLRA-WF algorithm reduce error function and thereby finding the best fit line in order to accurately find the future weather conditions. Hence, RSFS-GLRA technique enhances weather prediction accuracy as compared to conventional works.

Experimental Settings

In order to estimate the performance of proposed RSFS-GLRA technique is implemented in Java language with help of Hurricanes and Typhoons 1851-2014 dataset [21]. This dataset comprises of many data acquired from tropical cyclones in Atlantic Ocean and Eastern Pacific Ocean with 22 attributes. Also, this dataset contains 49,105 instances from the Atlantic Ocean and 26,138 instances from the eastern Pacific Ocean. For performing the experimental process, RSFS-GLRA technique takes various numbers of big weather data in the range of 1000-10000 from an input dataset. By using this weather data, RSFS-GLRA technique performs feature selection and regression analysis in order to find future cyclones. The performance of RSFS-GLRA technique is evaluated in terms of prediction accuracy, error rate and prediction time. The performance result of RSFS-GLRA technique is compared against with two conventional methods namely LTP_DSEF model [1] and SVR approach [2].

Results and Discussion

The experimental result of proposed RSFS-GLRA technique is evaluated using below metrics with table and graph.

1. Prediction Accuracy (PA)

PA estimates ratio of number of weather data that are exactly predicted to the total number of weather data. The prediction accuracy is determined as,

$$PA = \frac{n_{EP}}{n} * 100 \qquad (6)$$

From (6), 'n_{EP}' refers the number of weather data accurately predicted in which 'n' point outs a total number of weather data. The prediction accuracy is evaluated in percentage (%). The experimental result analysis of prediction accuracy is presented in below Table 1.

As shown in table 1, proposed RSFS-GLRA technique achieves improved prediction accuracy with increasing number of weather data as input to exactly discover future weather conditions i.e. cyclone occurrence when compared to conventional LTP_DSEF model [1] and SVR approach [2]. This is due to application of RS-FS algorithm and GLRA-WF algorithm in proposed RSFS-GLRA technique on the contrary to existing works. The proposed RSFS-GLRA technique extracts features that are more important to accurate forecasting the cyclone with increasing number of weather data as input.

After performing feature selection process, RSFS-GLRA technique employs GLRA-WF algorithm that compute the relationship between the input weather data for efficiently performing weather forecasting process with higher accuracy when compared to existing works. Thus, RSFS-GLRA technique improves a ratio of number of weather data that are accurately predicted when compared to other existing LTP_DSEF model [1] and SVR approach [2]. The proposed RSFS-GLRA technique increases prediction accuracy by 13 % and 19 % as compared to LTP_DSEF model [1] and SVR approach [2] respectively.

Table 1 Comparative Result of Prediction Accuracy

No. of weather data (n)	Prediction Accuracy (%)		
	RSFS-GLRA technique	LTP_DSEF model	SVR approach
1000	93	86	74
2000	94	88	80
3000	93	80	81
4000	92	84	80
5000	93	85	80
6000	92	84	79
7000	91	82	78
8000	89	81	75
9000	95	79	76
10000	96	77	75

2. Prediction Time (PT)

PT determines the amount of time required to find the future weather conditions. The prediction time is calculated as,

$$PT = n * time\ (PSWD) \quad (7)$$

From (7), '$time\ (PSWD)$' represents time taken for predicting a single weather data and 'n' refers a total number of weather data considered for experimental work. The prediction time is estimated in terms of milliseconds (ms). The performance result analysis of prediction time with respect to diverse number of weather data is depicted in below figure 4.

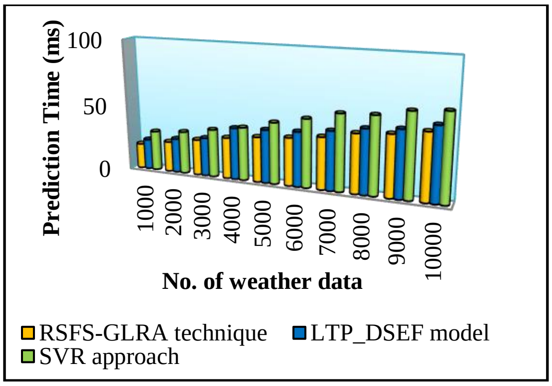

Figure 4 Comparative result of time complexity versus number of weather data

As shown in figure 4, proposed RSFS-GLRA technique gives lower amount of time complexity with increasing number of weather data as input to accurately predict future cyclones when compared to conventional LTP_DSEF model [1] and SVR approach [2]. This is because of application of Ruziicka Similarity-Based Feature Selection (RS-FS) algorithm and Generalized Linear Regression Analysis based Weather Forecasting (GLRA-WF) algorithm in proposed RSFS-GLRA technique. The designed RS-FS algorithm predicts key features in input big dataset that are more related to perform weather forecasting with minimal time consumption when compared to state-of-the-art works. RSFS-GLRA technique correctly predicts future cyclones by finding relationship between input weather data with lower amount of time consumption. Thus, RSFS-GLRA technique decreases the amount of time required to find cyclone when compared to other conventional LTP_DSEF model [1] and SVR approach [2]. As a result, proposed RSFS-GLRA technique minimizes the prediction time by12% and 27% when compared to LTP_DSEF model [1] and SVR approach [2] respectively.

3. Error Rate

Error Rate 'ER' determines the ratio of a number of weather data incorrectly predicted to the total number of weather data as input. The error rate is measured as,

$$ER = \frac{n_{IP}}{n} * 100 \quad (8)$$

From (8), 'n_{IP}' denote number of weather data wrongly predicted in which 'n' denotes number of weather data. The error rate is measured in percentage (%). The tabulation analysis of error rate depends on different number of weather data using three methods is presented in below Table 2.

Table 2 Comparative Result of Error Rate

No. of weather data (n)	Error Rate (%)		
	RSFS-GLRA technique	LTP_DSEF model	SVR approach
1000	8	14	26
2000	6	13	20
3000	7	20	19
4000	8	16	20
5000	7	15	20
6000	8	16	21
7000	9	18	22
8000	11	19	25
9000	6	21	24

10000	**5**	**22**	**25**

As shown in above table, proposed RSFS-GLRA technique gives minimal error rate with increasing number of weather data as input to perfectly forecasts future cyclones when compared to LTP_DSEF model [1] and SVR approach [2]. This is owing to application of Ruziicka Similarity-Based Feature Selection (RS-FS) algorithm and Generalized Linear Regression Analysis based Weather Forecasting (GLRA-WF) algorithm in proposed RSFS-GLRA technique. The objective of GLRA-WF algorithm is to find out the linear model that decreases the sum of the squared errors between the observations in a dataset and those predicted by the model. From that, proposed GLRA-WF algorithm exactly identifies future weather conditions with higher accuracy. RSFS-GLRA technique reduces the ratio of a number of weather data incorrectly predicted. Accordingly, proposed RSFS-GLRA technique reduces the error rate by 66 % and 70 % when compared to LTP_DSEF model [1] and SVR approach [2] respectively.

Literature Survey

A review of various methods for weather prediction was examined in [7]. Sliding Window Algorithm was employed in [8] to increase day's weather conditions forecasting performance. However, weather prediction performance was not sufficient. A non-parametric spearman correlation based technique was utilized in [9] to choose important features for accurately performing weather prediction. But, time complexity involved was high. Incremental K-Means Clustering was performed in [10] to determine weather related events with lesser time consumption. However, prediction accuracy of weather data was poor. A visual comparison of weather features was presented in [11] to resolve data with meteorological forecasting problems. However, features selection time during the weather forecasting process was lower. An examination of various data mining techniques introduced for carried outing weather forecast analysis was presented in [12].A backpropagation neural network (BPNN) was applied in [13] for discovering the weather changes. However, ratio of number of weather data that are correctly predicted was low. A multilayer perception (MLP) and radial basis function (RBF) was employed in [14] to minimize time taken for performing weather forecasting process. But, computational complexity of weather prediction was very higher.

A dynamic fine-tuning stacked auto-encoder neural network was introduced in [15] to reduce processing time for weather prediction. However, misclassification error involved during weather forecasting process was not reduced. Clustering and classification methods were developed in [16] for predicting severe climate conditions with a lower time. But, the weather prediction performance was not adequate. A novel machine learning technique was developed in [17] for enhancing accuracy of weather prediction. However, error rate of weather forecasting was very higher.

A method was introduced in [18] for accomplishing data modeling and optimizing weather research and forecasting performance. But, execution time of optimization process was more. Big Data Prediction Framework was implemented in [19] to increase weather prediction accuracy with MapReduce algorithm. However, the amount of time taken for predicting weather was high. Gridded probabilistic weather forecasts was accomplished in [20] with analog ensemble method. But, feature selection process was not carried out which impacts weather prediction accuracy.

Conclusion

The RSFS-GLRA technique is proposed with the aim of increasing the performance of weather forecasting for cyclone prediction with a minimal error rate. The aim of RSFS-GLRA technique is obtained with the application of Ruziicka Similarity-Based Feature Selection (RS-FS) algorithm and Generalized Linear Regression Analysis based Weather Forecasting (GLRA-WF) algorithm on the contrary to traditional works. The proposed RSFS-GLRA technique increases the ratio of number of weather data that are correctly predicted by using GLRA-WF algorithm when compared to conventional works. Additionally, proposed RSFS-GLRA technique minimizes the time utilized to predict future cyclone data with the support of RS-FS and GLRA-WF algorithmic process when compared to other state-of-the-art works. Also, proposed RSFS-GLRA technique reduces ratio of number of weather data wrongly predicted to accurately discover future cyclone by using GLRA-WF algorithm when compared to other existing works. Therefore proposed RSFS-GLRA technique renders better weather forecasting performance in terms of accuracy, time and error rate when compared to conventional works. The experimental result shows that the proposed RSFS-GLRA technique presents better cyclone forecasting performance with enhancement of prediction accuracy and minimization of time when compared to state-of-the-art works.

References

[1] Zuo Jia, Fumin Ren, Dalin Zhang, Chenchen Ding, Mingjen Yang, Tian Feng, Boyu Chen, Hui Yang. An application of the LTP_DSEF model to heavy precipitation forecasts of land falling tropical cyclones over China in 2018. Science China Earth Sciences,Springer 2019,63(1)**,** 27-36, https://doi.org/10.1007/s11430-019-9390.

[2] Michael B.Richman, Lance M.Leslie, Hamish A.Ramsay, Philip J.Klotzbach. Reducing Tropical Cyclone Prediction Errors Using Machine Learning Approaches, Procedia Computer Science, Elsevier, 2017 114, 314-323.

[3] Alexander Kumpf, Bianca Tost, Marlene Baumgart, Michael Riemer, R¨udiger Westermann, and Marc Rautenhaus. Visualizing Confidence in Cluster-based Ensemble Weather Forecast Analyses, IEEE Transactions on Visualization and Computer Graphics, 2018, 24(1),109-119.

[4] Florian Ferstl, Mathias Kanzler, Marc Rautenhaus, and Rudiger Westermann. Time-hierarchical Clustering and Visualization of Weather Forecast Ensembles. IEEE Transactions on Visualization and Computer Graphics, 2017, 23(1), 831 – 840.

[5] Abir Jaafar Hussain, Panos Liatsis, Mohammed Khalaf, Hissam Tawfik, Haya Al-Asker. A dynamic neural network architecture with immunology inspired optimization for weather data forecasting. Big Data Research, Elsevier, 2018,1-38.

[6] Ashkan Zarnani, Petr Musileka and Jana Heckenbergerova. Clustering numerical weather forecasts to obtain statistical prediction intervals. Meteorological Applications, 2014, 21, 605–618.

[7] Sara khan, Mohd Muqeem, Nashra Javed. A Critical Review of Data Mining Techniques in Weather Forecasting. International Journal of Advanced Research in Computer and Communication Engineering, 2016, 5(4), 1091-1094.

[8] Piyush Kapoor and Sarabjeet Singh Bedi. Weather Forecasting Using Sliding Window Algorithm.Hindawi Publishing Corporation, ISRN Signal Processing, 2013, 1-5.

[9] Eduardo Soares, Pyramo Costa Jr, Bruno Costa and Daniel Leite(2018).Ensemble of evolving data clouds and fuzzy models for weather time series prediction. Applied Soft Computing, 64, 445-453.

[10] Sanjay Chakraborty, N. K. Nagwani, Lopamudra Dey. Weather Forecasting using Incremental K-means Clustering, CiiT International Journal of Biometrics and Bioinformatics, 2012,4(5), 1-6.

[11] P. Samuel Quinan and Miriah Meyer. Visually Comparing Weather Features in Forecasts. IEEE Transactions on Visualization and Computer Graphics, 2016, 22(1), 389-398.

[12] Nikita Gupta, Rashmi Narayanan, Anagha Chaudhari. Implementation and Analysis of Data Mining Techniques for Weather Prediction. International Journal of Innovations & Advancement in Computer Science, 2017, 6(11), 101-104.

[13] Prasanta Rao Jillella S.S, P Bhanu Sai Kiran, P. Nithin Chowdary, B. Rohit Kumar Reddy, Vishnu Murthy.Weather Forecasting Using Artificial Neural Networks and Data Mining Techniques. International Journal of Innovative Technology And Research, 2015, 3(6), 2534 – 2539.

[14] Tanzila Sabar, Amjad Rehman, Jarallah S. AlGhamdi. Weather forecasting based on the hybrid neural model. Applied Water Science, Springer, 2017, 7(7),3869–3874.

[15] Szu-Yin Lin, Chi-Chun Chiang, Jung-Bin Li, Zih-Siang Hung, Kuo-Ming Chao. Dynamic fine-tuning stacked auto-encoder neural network for weather forecast. Future Generation Computer Systems, Elsevier, 2018, 89, 446-454.

[16] Glauston R.Teixeira de Lima and Stephan Stephany. A new classification approach for detecting severe weather patterns. Computers & Geosciences, Elsevier, 2013, 57, 158-165.

[17] Shivangi Raj, Deepak Arora, Pooja Khanna, Bramah Hazela. Implication of Machine Learning towards Weather Prediction. International Journal of Research in Advent Technology, 2019,7(5), 199-203.

[18] Yonghua Xie, Xiaoyong Kou, Ping Li. A simulation method of three-dimensional cloud over WRF big data. EURASIP Journal on Wireless Communications and Networking, Springer, 2019,252.

[19] Khalid Adam, Mazlina A. Majid, Mohammed Adam Ibrahim Fakherldin and Jasni Mohamad Zain. A Big Data Prediction Framework for Weather Forecast Using MapReduce Algorithm. Journal of Computational and Theoretical Nanoscience, 2017, 23(11), 11138-11143.

[20] Simone Sperati, Stefano Alessandrini, Luca Delle Monache. Gridded probabilistic weather forecasts with an analog ensemble. Quarterly Journal of the Royal Meteorological Society,2017,143(708), 2874-2885.

[21] Hurricanes and Typhoons, 1851-2014 dataset: https://www.kaggle.com/noaa/hurricane-database