International Journal of Recent Technology and Engineering (IJRTE)
ISSN: 2277-3878 (Online), Volume-8 Issue-4, November 2019

Optimized Hybrid Soft Computing Model for Weather Predictions in Delhi

Suvendra Kumar Jayasingh, Jibendu Kumar Mantri

Abstract*: Soft computing techniques have become very popular now-a-days as these techniques have replaced the traditional and statistical prediction mechanisms in weather forecasting, stock market prediction, crop prediction, solar energy prediction, and predictions in physics and chemistry etc. Each model has its advantages and disadvantages. Hybrid soft computing model is the mechanism of designing the models by exploiting the advantages of two or more models and suppressing their disadvantages. If the advantages of two or more number of models will be taken together in the new proposed model, then the accuracy in the prediction will be enhanced with decrease in error rate. This paper intends to design a hybrid model by taking the advantages of J48 Decision Tree and Fuzzy Logic andit is used to predict the weather parameters in Delhi with better accuracy.

Keywords*: Decision Tree, Fuzzy Rules, Fuzzy Rule based Decision Tree.

I. INTRODUCTION

A hybrid system is an intelligent system which is framed by combining at least two intelligent technologies like Fuzzy Logic, Neural networks, Genetic algorithm, reinforcement learning, etc. The combination of different techniques in one computational model makes these systems possess an extended range of capabilities. These systems are capable of reasoning and learning in an uncertain and imprecise environment. These systems can provide human-like expertise like domain knowledge, adaptation in noisy environment etc. The hybrid climate modeling systems will combine the features of Fuzzy logic, Decision Tree, Artificial Neural Network(ANN), Support Vector machine(SVM) etc. to get a new model which can outperform the constituent soft computing models in terms of accuracy.

II. LITERATURE REVIEW

Forecasting the weather is a challenging task basically the rain fall prediction which is most essential in saving the mankind and their properties. It helps in ongoing constructions, flight operations, transportation etc. Aftab [1] et al have made a study on different articles published during 2011 and 2017 and made a comparison on the different data mining techniques to be used as the base line for the early prediction of rainfall. Data mining techniques can be a better tool than the traditional methods for predicting the rainfall at an early stage. Amanullah [2] et al have noted that the soft computing techniques comprising of fuzzy logic, artificial neural network, evolutionary algorithms and genetic algorithms could be used to predict the weather just like a human like expertise. Amato[3] has compared Rough Set Theory with Multiple Regression Analysis as automated valuation methodologies to make international real estate review. Appelquist[4] has demonstrated a Genetic framework for meso scale assessment of climate change hazards in coastal environments. Baboo[5] et al have developed an efficient weather forecasting system using artificial neural network. They have compared the result with the realtime dataset and it is found that the prediction is approximately good. Babu[6] et al have made a comparison of ANFIS and ARIMA model for weather forecasting. Bautu[7] et al have made an empirical study on meteorological time series. Biradar[8] et al have made the prediction of weather using data mining. Bushara[9] et al have made a review in the use of computational intelligence in weather forecasting. Bushara[10] et al have made forecasting of weather in Sudan using Machine learning schemes. Caskey[11] has demonstrated the use of numerical methods in weather prediction. Chawsheen[12] et al have made seasonal time series modeling and forecasting of monthly mean temperature for decision making in the Kurdistan region of Iraq. Gocken[13] et al have made application of soft computing models for daily average temperature analysis. Gupta[14] et al have used normal equation method and linear regression techniques for prediction of weather. Hamidi[15] have shown a comparative study SVM and ANN for precipitation prediction in Iran. Holbig[16] et al have corrected statistical result of weather forecast by applying the model output calibration. Honarbakhsh[17] et al have demonstrated application of soft computing methods in predicting evapotranspiration. Isa[18] have made weather forecasting using photovoltaic system and neural network. Jain[19] et al have made a study of time series models ARIMA and ETS . Jayasingh[20] et al have shown the application of text mining for faster weather forecasting and suggested that the text mining would help in faster prediction of weather in Delhi in comparison to the numeral prediction by use of same soft computing models like MLP, SVM and J48 decision tree. Jayasingh[21] have demonstrated the comparison between J48 Decision Tree, SVM and MLP in weather prediction and it was found that the SVM out performs the other two in predicting the weather on the basis of different statistical parameters like RAE, RMSE, MAE and RRSE. Joseph[22] have shown that the neural networks can perform better using the functional graph for weather forecasting.

Manuscript published on November 30, 2019.
* Correspondence Author
Dr.Jibendu Kumar Mantri, Reader, PG Department of Computer Application, North Orissa University, Baripada, India, Email: jkmantri@gmail.com

Khajure[23] et al have made prediction of future weather using artificial neural network. Khajure[24] et al have shown the future weather forecasting using soft computing techniques. Jayasingh[25] et al have made a review of soft computing approaches for climate modeling and weather predictions.

The study above concludes that the prediction of weather using soft computing techniques out performs the tradition numerical and statistical method. Every soft computing model has some advantages and disadvantages. If we can go for the hybrid models by taking the advantages of two soft computing models and discarding their individual disadvantages, it may emerge a new model that can perform better than the individual models. This paper intends to design the following hybrid models to get better accuracy in weather prediction. The proposed new hybrid soft computing models will take the same input as it was given to the constituent soft computing models and the output will be analyzed in terms of accuracy rate, MAE, RMSE, RAE and RRSE. The proposed hybrid model is Fuzzy Rule based Decision Tree. In this model the benefits of Decision Tree and Fuzzy Set are taken into consideration by discarding their individual disadvantages.

III. DATA

The weather data of Delhi, collected from 1996 to 2017, are taken into consideration for deployment of the proposed hybrid soft computing models. The data has been collected from AccuWeather website for the following weather parameters. The data contains the following 19 attributes. They are Maximum Temperature, Average Temperature, Low Temperature, Maximum Dew Point, Average Dew Point, Low Dew Point, Maximum Humidity, Average Humidity, Low Humidity, Maximum Sea Level Pressure, Average Sea Level Pressure, Low Sea Level Pressure, Maximum Visibility, Average Visibility, Low Visibility, Maximum Wind Speed, Average Wind Speed, Low Wind Speed and Events.

The above mentioned parameters, present in the database, are numeric values. The one month sample data of the weather is shown in Table - 1.

Table -1
One month sample data of the weather

2016	Temp. (°C)			Dew Point (°C)			Humidity (%)			Sea Level Press. (hPa)			Visibility (km)			Wind (km/h)			Events
Jan	high	avg	low	high	avg	low	high	avg	low	high	avg	low	high	avg	low	high	avg	high	
1	23	16	8	11	9	7	94	71	34	1024	1021	1019	2	1	0	14	2	-	Fog
2	22	15	8	11	8	6	88	67	35	1024	1021	1019	2	1	0	16	3	-	Fog
3	22	15	8	13	8	6	100	72	35	1021	1019	1016	2	1	0	14	5	-	Fog
4	24	16	8	12	9	6	100	66	32	1022	1019	1017	2	1	0	10	2	-	Fog
5	27	19	11	13	11	7	94	65	28	1020	1017	1014	2	1	0	10	5	-	Fog , Tornado

Out of the 19 parameters, we have taken into consideration only 7 parameters such as Average Temperature, Average Dew Point, Average Humidity, Average Sea Level Pressure, Average Visibility, Average Wind Speed and Event as presented in Table - 2.

Table -2
One month sample data of the weather for selected parameters

Date /Jan 2016	temp_avg	dew_avg	humidity_avg	press_avg	Visibilty_avg	wind_avg	Events
1	16	9	71	1021	1	2	Fog
2	15	8	67	1021	1	3	Fog
3	15	8	72	1019	1	5	Fog
4	16	9	66	1019	1	2	Fog
5	19	11	65	1017	1	5	Fog , Tornado

The weather data is converted into fuzzy database by use of fuzzy membership function. The fuzzy membership function is defined by taking into consideration the values of each parameter present in the database. The fuzzy rules will convert the database into fuzzy database by use of the following algorithm.

Algorithm

Input: *Weather database*
Fuzzy Rule Definition
Output: *Fuzzy Weather database*
Start

Retrieve one record from database
Get the retrieved result in array 1
Retrieve the metadata from fuzzy rule set
Get the retrieved result in array 2
Continue until the end of result array 1
Extract numeric data in variable
Extract each variable
Execute each variable to find the fuzzy variable value
Store in a variable
Create a new record to store the generated fuzzy dataset
Go to Next line

End of loop
Close database
End

After running the above algorithm, the numeric weather database is converted into fuzzy database using the mapping fuzzy metadata definition. The result of the output after the execution of the algorithm is shown in Table - 3.

Table -3One month fuzzy sample data of the weather

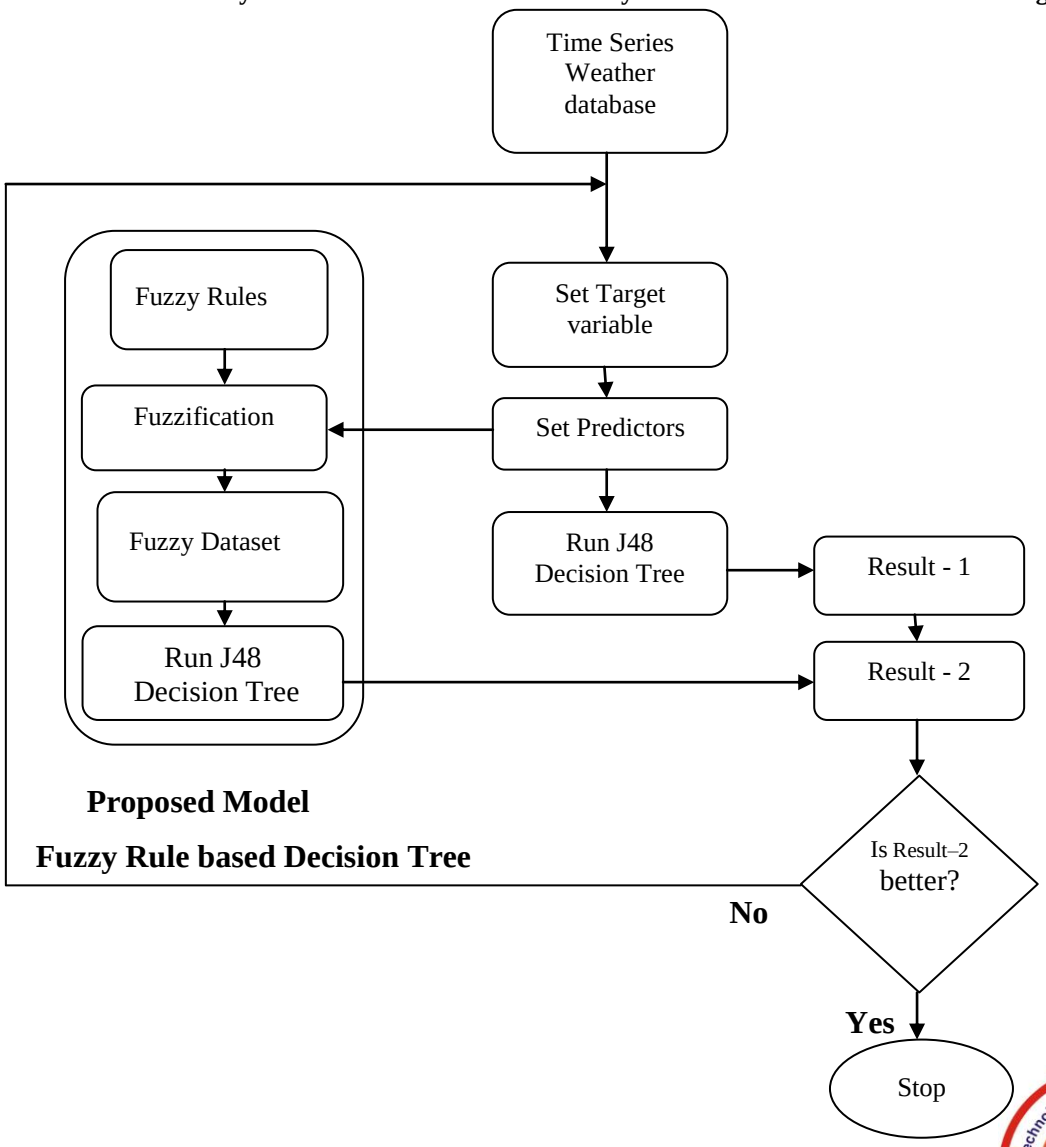
Date	Temp	Dew	Humidity	Pressure	Visibility	Wind_speed	Events
1	very low	very low	high	very high	medium	very low	Fog
2	very low	very low	high	very high	medium	very low	Fog , Rain
3	very low	very low	very high	very high	medium	very low	Fog , Rain
4	very low	very low	high	very high	medium	very low	Fog
5	very low	very low	medium	very high	medium	Medium	Fog

II. FUZZY RULE BASED DECISION TREE (FR-DT)

a. Proposed Model

The Fuzzy rule based decision tree is designed by taking the Fuzzy Set and J48 decision Tree for mining the knowledge from the database. The Fuzzy set is based upon the membership function to express the data present in the database. The fuzzy rules are defined as per the requirement in the fuzzy membership function. The J48 Decision Tree is a classification technique based on the supervised database. The flow chart of the proposed model is shown in Figure – 1.

Figure - 1
Fuzzy Rule based J48 Decision Tree Hybrid Model for Weather Forecasting



b. Methodology

The proposed hybrid model combines the features of Fuzzy set and J48 Decision Tree to predict the weather.
It generates the predicted values as per the algorithm mentioned below.
Step -1: Generate the fuzzy database from the existing database by using appropriate mapping rule
Step -2: Input the fuzzy database to the J48 Decision Tree
Step -3: Generate the Decision Tree for prediction of weather parameter value
Step -4: The decision tree is pruned by applying tree pruning mechanism to optimize the Decision Tree

c. Output Analysis

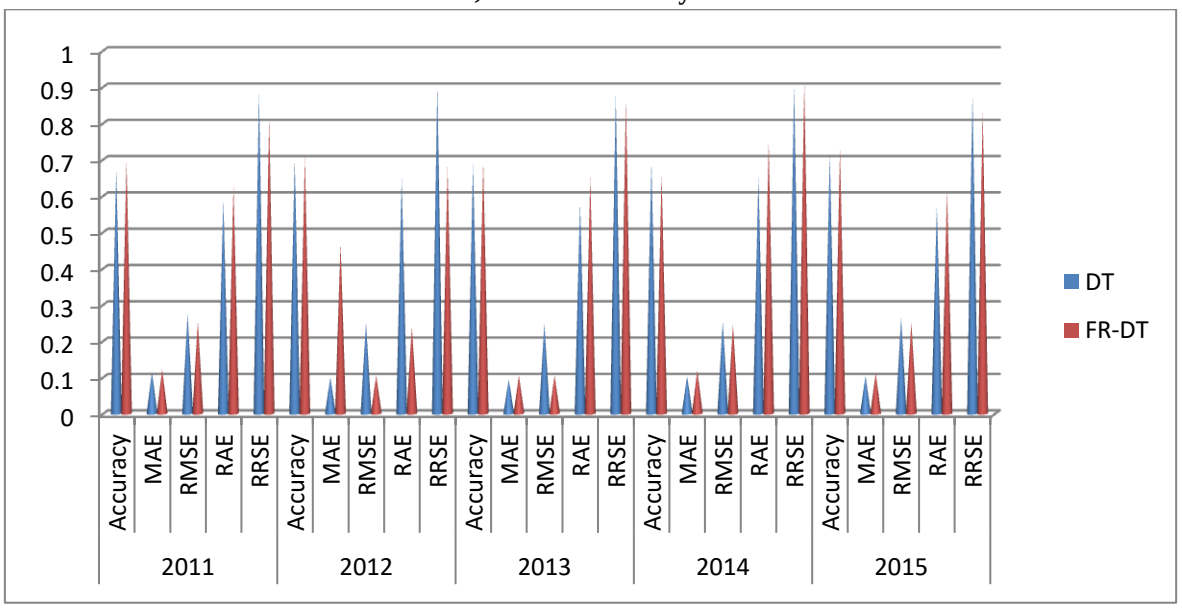
The weather database is fed to the models J48 decision Tree and the proposed Fuzzy based J48 Decision Tree. The output is compared on the basis of accuracy, MAE, RMSE, RAE and RRSE which is summarized in Table - 4.

Table – 4
Comparison between J48 decision Tree and the proposed Fuzzy based J48 Decision Tree year wise

	2011					2012				
	Accuracy	**MAE**	**RMSE**	**RAE**	**RRSE**	**Accuracy**	**MAE**	**RMSE**	**RAE**	**RRSE**
J48 DT	0.6739	0.1148	0.2752	0.59	0.8845	0.6994	0.0998	0.2498	0.6497	0.9056
FR-DT	0.6986	0.1219	0.2534	0.6265	0.8146	0.7103	0.4705	0.1042	0.2392	0.6783
	2013					**2014**				
	Accuracy	**MAE**	**RMSE**	**RAE**	**RRSE**	**Accuracy**	**MAE**	**RMSE**	**RAE**	**RRSE**
J48 DT	0.6877	0.0937	0.2483	0.5805	0.8778	0.6931	0.1033	0.2539	0.6529	0.9067
FR-DT	0.6931	0.1051	0.1051	0.6514	0.8659	0.6575	0.1187	0.2555	0.7503	0.9125
	2015									
	Accuracy	**MAE**	**RMSE**	**RAE**	**RRSE**					
J48 DT	0.717802	0.104	0.2644	0.571	0.8793					
FR-DT	0.7326	0.1119	0.251	0.6144	0.8347					

On comparison between J48 decision tree and fuzzy rule based J48 Decision Tree, it is found that the accuracy of the proposed new model is better than that of the J48 Decision tree. Similarly, if we look into the MAE, RMSE, RAE and RRSE, the error values are less in case of the proposed model. So, it is proved that the proposed model gives promising result with more accuracy and less error rates. The pictorial representation of the different parameters is expressed in Figure – 2.

Figure – 2 Graphical comparison of J48 Decision Tree and the proposed hybrid model on the basis of Accuracy, MAE, RMSE, RAE and RRSE year wise



The pruned tree of **Fuzzy based J48 Decision Tree is shown below.**

```
J48 pruned tree
------------------

pressure = very high: Fog (90.0/16.0)
pressure = high
|   humidity = high: Fog (3.0/2.0)
|   humidity = very high: Fog (2.0)
|   humidity = medium
|   |   visibility = medium: Fog (10.0/1.0)
|   |   visibility = low: Fog (0.0)
|   |   visibility = high: Rain , Thunderstorm (2.0)
|   |   visibility = very high: Fog (0.0)
|   humidity = low
|   |   dew = very low: Fog (10.0)
|   |   dew = low: No Rain (10.0/4.0)
|   |   dew = medium: No Rain (2.0/1.0)
|   |   dew = high: Fog (0.0)
|   humidity = very low: No Rain (51.0/14.0)
pressure = medium: No Rain (60.0/10.0)
pressure = low: No Rain (92.0/25.0)
pressure = very low
|   temp = very low: No Rain (0.0)
|   temp = low: No Rain (0.0)
|   temp = Medium: Rain , Thunderstorm (16.0/10.0)
|   temp = High
|   |   visibility = medium: Fog (3.0/1.0)
|   |   visibility = low: No Rain (0.0)
|   |   visibility = high: No Rain (8.0/1.0)
|   |   visibility = very high: No Rain (0.0)
|   temp = Very High: No Rain (6.0)
```

Increasing the confidence in training data increases the tree size, which increases the accuracy of the system but decreases the interpretability of the system. Reversing the process by decreasing the confidence, decreases the tree size, reduces the accuracy and increases the interpretability of the system. The interpretability of fuzzy rules is an important consideration for fuzzy model generation. Here, we have used number of rules as indicator to ascertain the accuracy versus the interpretability of fuzzy systems. We have shown that Anais Congresso Brasileiro de Sistemas Fuzzy increase in fuzzy rules is based on decreasing the pruning (CBSF). (By increasing confidence limit) which results increases accuracy and decreases interpretability. Similarly, Learning", decrease in fuzzy rules is based on increasing the pruning (by decreasing the confidence limit) which results in discovery", decreasing the accuracy and increasing the Knowledge Discovery, interpretability.

The comparison between J48 Decision Tree and the proposed hybrid model on the basis of RMSE is presented in Table – 5. Also the graphical comparison between J48 Decision Tree and the proposed hybrid model on the basis of RMSE is expressed in Figure – 3.

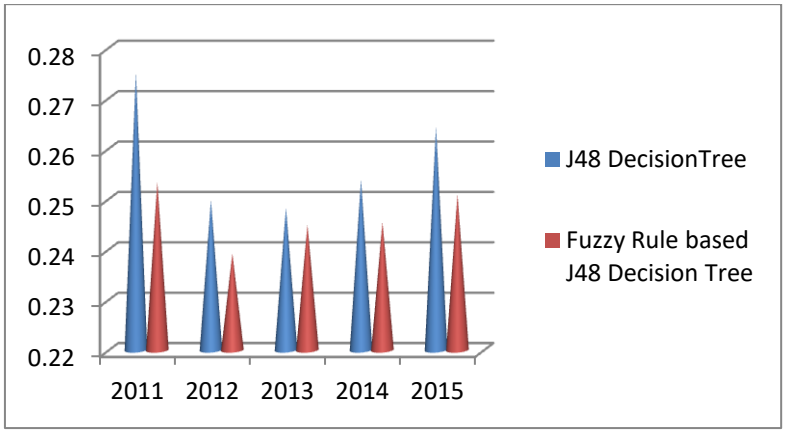
This comparison between different statistical error values obtained after the analysis depicts that the model having less Root Mean Squared Error (RMSE) is the better model for prediction of weather in Delhi.

Table – 5Comparison between J48 Decision Tree and the proposed hybrid model on the basis of RMSE

	J48 Decision Tree	Fuzzy Rule based J48 Decision Tree
2011	0.2752	0.2534
2012	0.2498	0.2392
2013	0.2483	0.2449
2014	0.2539	0.2455
2015	0.2644	0.251

Figure – 3

Graphical comparison of J48 Decision Tree and the proposed hybrid model on the basis of RMSE



III. CONCLUSION

Weather forecasting is a very complex phenomenon for which the researchers make use of the different soft computing models to solve the problem. Hybrid of soft computing climate models is a new attempt to predict the weather with better accuracy. In this paper we have used the fuzzy rule based Decision Tree hybridized model of soft computing techniques which could give better result than the existing models. So, the new hybrid model may be used for any other prediction problems like stock market prediction, medical diagnosis etc.

IV. FUTURE SCOPE

The hybrid model, proposed in this paper, the Fuzzy Rule based J48 Decision Tree, can perform better prediction than Fuzzy Set as well as the J48 Decision Tree. Hence, the weather prediction using the model outperforms than the individual constituent models of the hybrid model. The same hybrid model can be used for other forecasting problem like stock market, business prediction and predictions in Physics and Chemistry etc. The precision in the hybrid model may also be enhanced by modifying the Decision Tree by applying Tree pruning as per the requirements in prediction.

REFERENCES

1. Aftab S., Ahmad M., Hameed N., Bashir M. S., Ali I., Nawaz Z.(2018), Rainfall prediction using data mining techniques: a systematic literature review, International Journal of Advanced Computer Science and Applications, Vol 9(5), pp 143-150.
2. Amanullah M., Khanaa V. K.(2014), Application of soft computing techniques in weather forecasting: ANN approach, Vol 2(1), pp 212-219.
3. Amato M. D. (2007), Comparing Rough Set Theory with Multiple Regression Analysis as Automated Valuation Methodologies, International Real Estate Review, 10(2), pp 42-65.
4. Appelquist L. R.(2013), Genetic framework for meso scale assessment of climate change hazards in coastal environments, Journal Coast Conserv, Vol 17, pp 59-74.
5. Baboo S. S., Shereef I. K.(2010), An efficient weather forecasting system using artificial neural network, International Journal of Environmental Science and Development, Vol 1(4), pp 321-326.
6. Babu N. R., Babu C. B. A., Reddy D. R., Gowtham M. (2015) Comparison of ANFIS and ARIMA Model for Weather Forecasting, Indian Journal of Science and Technology, 8(S2), 70-73.
7. Bautu E., Barbulescu A.(2013), Forecasting meteorological time series using soft computing methods: an empirical study, Applied mathematics & Information Sciences, International journal, Vol 7(4), pp 1297-1306.
8. Biradar P., Ansari S., Paradhar Y., Lohiya S.(2017), Weather Prediction using Data Mining, International Journal of Engineering Development and Research, Vol 5(2), pp 213-214.
9. Bushara N. O., Abraham A.(2013), Computational Intelligence in Weather Forecasting: A review, Journal of Network and Innovative Computing, 1, 320-331.
10. Bushara N. O., Abraham A.(2014), Weather forecasting in Sudan using Machine Learning schemes, Journal of Network and Innovative Computing, Vol 2(2014), pp 309-317.
11. Caskey J.E.(1957), Numerical Methods in weather prediction, Monthly Weather Review, 8(5), pp 329-332.
12. Chawsheen T. A., Broom M.(2017) Seasonal time series modelling and forecasting of monthly mean temperature for decision making in the Kurdistan Region of Iraq, Journal of Statistical Theory and Practice, 11(4), 604 – 633.
13. Gocken M., Boru A., Dosdogru T., Berber N. (2015) Application of Soft Computing Models to Daily Average Temperature Analysis, International Journal of Engineering Technologies, 1(2), 56 – 64.
14. Gupta S., Indumathy K., Singhal G.(2016), Weather Prediction using normal equation method and linear regression techniques, International Journal of Computer Science and Information Technologies, Vol 7(3), pp 1490-1493.
15. Hamidi O., Poorolajal J., Sadeghifar M., Abbasi H., Maryanaji Z., Faridi H. R., Tapak L. (2015) A comparative study of Support Vector machines and Artificial Neural Networks for predicting precipitation in Iran, Theor Appl Climatol (Springer), 119, 723 – 731.
16. Holbig C. A., Machado V. L., Pavan W.(2016), Statistical correction of the result of weather forecast by applying the model output calibration,
17. Honarbakhsh A., Dashpagerdi M. M., Vagharfard H.(2013), Application of soft computing methods in predicting evapotranspiration, Open Journal of Geology, Vol 3, pp 397-403.
18. Isa I. S., Omar S., Saad Z., Noor N. M., Osman M. K.(2010), Weather forecasting using photovoltaic system and neural network, 2010 Second International Conference on Computational Intelligence, Communication Systems and networks, pp 96-100.
19. Jain G., Mallick B., (2017), A study of Time Series Models ARIMA and ETS, International Journal of Modern Education and Computer Science, 4, 57-63.
20. Jayasingh S.K., Mantri J.K., Gahan P. (2016), Application of Text Mining for Faster Weather Forecasting, International Journal of Computer Engineering and Information Technology, Vol 8(11) 213-219
21. Jayasingh S.K., Mantri J.K., Gahan P. (2016), Comparison between J48 Decision Tree, SVM and MLP in weather prediction, SSRG International Journal of Computer Science and Engineering(IJCSE), 3(11), 42 – 47.
22. Joseph R. V.(2008), Better performance of neural networks using functional graph for weather forecasting, 12th WSEAS International conference on computers, Harekliton, Greece, pp 826-831.
23. Khajure S. R., Mohod S. W.(2015), Prediction of future weather forecasting using artificial neural network, International Journal of Engineering and Applications, pp 58-60.
24. Khajure S., Mohod S. W.(2015), Future Weather forecasting using soft computing techniques, International conference on Information Security & Privacy, Science Direct, Procedia Computer Science 78(2016), pp 402-407.
25. Jayasingh S.K., Mantri J.K. (2019), Soft computing approaches for climate modeling and weather predictions: A review, International Journal of New Innovations in Engineering and Technology, Vol 11(2) pp 109-113.

AUTHORS PROFILE

Mr. Suvendra Kumar Jayasingh has completed his Bachelor of Engineering in Computer Science and Engineering (CSE) with Honours from University College of Engineering (UCE), Burla, Odisha (Now VSSUT, Burla) in the year 2003. He got selected in Odisha Public Services Commission (OPSC) in the year 2005 to serve in Institute of Management and Information Technology (IMIT), Cuttack, a constituent College of Biju Patnaik University of Technology (BPUT), Odisha, India. He has completed M. Tech. in Computer Science and Engineering(CSE) from JRN RVU, Udaipur. Now, he is pursuing his Ph. D. in Computer Science and IT in North Orissa University, Baripada. His main research work focuses on Data Mining, Database Design Computational Intelligence and Soft Computing. He has 16 years of teaching experience.

Dr J. K. Mantri pursed Master of Technology in Computer Sc. from Utkal University in year 2002 and Ph. D degree in Computer Application from Sambalpur University in year 2010. He is currently working as Reader in Department of Computer Application, North Orissa University, India. He is a life member of ISTE and ISCA. He has published 5 books and more than 58 research papers in reputed international journals including Thomson Reuters (SCI and Web of Science) and conferences including IEEE and Springer. His main research work focuses on Cryptography, Software Engineering and Computational Intelligence. He has 28 years of teaching experience.