Webology (ISSN: 1735-188X)
Volume 18, Number 5, 2021

Optimizations Based Feature Selection Method For Rainfall Forecasting

A. MARY SUBASHINI[1] , Dr. M. RAJA KUMAR[2]

[1]Research Scholar, PG and Research Department of Computer Science, Jamal Mohamed College (Autonomous) (Affiliated to Bharathidasan University), Tiruchirappalli, Tamilnadu, India.

[2]Assistant Professor, PG and Research Department of Computer Science, Jamal Mohamed College (Autonomous) (Affiliated to Bharathidasan University), Tiruchirappalli, Tamilnadu, India.

ABSTRACT

Weather forecasting is a way to predict future weather. It is widely researched area due to the fact that human life on earth is affected by the global climate. The techniques of Data Mining are employed in the act of real information discovery for most domains' problems. This work focuses on developing an optimized system model which predicts future weather. The frequency of natural hazards occurring due to unpredictable weather conditions have been seen to be increasing causing damage to human life. Rainfall is important for food production plan, water resource management and all activity plans in the nature. The occurrence of prolonged dry period or heavy rain at the critical stages of the crop growth and development may lead to significantly reduce crop yield. India is an agricultural country and its economy is largely based upon crop productivity. There are some models that predict weather during real time, or monthly or annual period. Feature Selection plays a role that is vital a large wide range of datasets. In this research work, Information Gain and Differential Evolution hybridized for choosing the most relevant features. Once, the suitable features subset is obtained, the classification algorithm called Artificial Neural Network (ANN), and Random Forest are adopted, that could classify the info in the manner that is effective the selected features.

KEYWORDS: Feature Selection, Classification, Information Gain, Optimization, Decision Tree, Artificial Neural Network

1. INTRODUCTION

In India, where the majority of agribusiness is dependent on precipitation as its standard wellspring of water, the time and measure of precipitation hold high importance and can impact the entire economy of the nation. Climate plays a vital role in our everyday life. From the earliest starting point of the human development, we are occupied with thinking about climatic changes [1][2]. Weather forecasting is one of the most challenging issues seen by the world, in a most recent couple of centuries in the field of science and technology. Prediction is the phenomena of knowing what may happen to a system in the near future. Present weather observations are obtained by ground-based instruments and from the satellite through remote sensing. As India's economy significantly depends on horticulture, precipitation plays an important part.

Rainfall prediction [3] is important in Indian civilization and it plays major role in human life to a great extent. It is demanding responsibility of meteorological department to predict the frequency of rainfall with uncertainty. It is complicated to predict the rainfall accurately with changing climatic conditions. It is challenging to forecast the rainfall for both summer and rainy seasons. Researchers in all over the world have developed various models to predict the rain fall mostly using random numbers and they are similar to the climate data.

1.1 Importance of Rainfall Forecasting

Rainfall is a crucial phenomenon within a climate system, whose chaotic nature has a direct influence on water resource planning, agriculture and biological systems. Within finance, the level of rainfall over a period of time is vital for estimating the value of a financial security. Over recent years, scientists' abilities in understanding and predicting rainfall have increased, due to numerous models developed for increasing the accuracy of rainfall prediction [4]. Subsequently, such efforts in new techniques can lead to the correct predictions of rainfall amounts for weather derivatives. Rainfall derivatives share similar principles with weather derivatives and other regular derivatives. They are defined as contracts between two or more parties, where the value of a contract is dependent upon the underlying financial asset. Hence, in the case of weather derivatives, the underlying asset is a weather type, such as rainfall. One significant difference between common derivatives and weather derivatives is that the underlying asset that governs the price of the contract is not tradable.

Rainfall derivatives is a new method for reducing the financial risk posed by adverse or uncertain whether circumstances [5]. These contracts are aimed towards individuals whose business is directly or indirectly affected by rainfall. For example, farmers whose crops are their main source of income, requiring a certain range of rainfall over a period of time to maximise their income. Moreover, rainfall derivatives are a better alternative than insurance, because it can be difficult to prove that the rainfall has had an impact unless it is destructive, such as severe floods or drought. Therefore, rainfall derivatives offer a simple solution to resolving the problems of financial protection against unfavourable rainfall. Similar contracts exist for other weather variables, such as temperature and wind.

2. RELATED WORKS

Grace, R. Kingsy, and B. Suganya [6] proposed a rainfall prediction model using Multiple Linear Regression (MLR) for Indian dataset. The input data is having multiple meteorological parameters and to predict the rainfall in more precise. MSE, RMSE, and Correlation are the metrics used in this paper.

Pham, Binh Thai, et al [7] studied the main objective is to develop and compare several advanced Artificial Intelligence (AI) models namely Adaptive Network based Fuzzy Inference System with Particle Swarm Optimization (PSOANFIS), Artificial Neural Network (ANN) and Support Vector Machine (SVM) for the prediction of daily rainfall in Hoa Binh Province, Vietnam. Correlation Coefficient (R) and Mean Absolute Error (MAE), Score Skill (SS), Probability of Detection (POD), Critical Success Index (CSI) and False Alarm Ratio (FAR) are used as the performance evaluation metrics.

Ahmed, Kamal, et al [8] employed Multi-Model Ensembles (MMEs) to reduce the uncertainties related to GCM simulations/projections. The objective of this study was to evaluate the performance of MMEs developed using machine learning (ML) algorithms with different combinations of GCMs ranked based on their performance and determine the optimum number of GCMs to be included in an MME.

Le, Vuong Minh, et al [9] proposed a prediction model using Nonlinear Autoregressive Neural in order to forecast daily rainfall at Hoa Binh city, Vietnam. For this aim, eight-year time series of meteorological data were first collected, involving temperature, wind speed, relative humidity, solar radiation as input variables and daily rainfall as output variable Networks with external variables (NARX). The metrics like Correlation Coefficient, MAE, RSME, Error Mean, Median, STD are considered for evaluation.

Sardesh pande, Kaushik D., and Vijaya R. Thool [10] presented a case study on time series prediction as an application of neural networks. The case study was done for the rainfall prediction using the local database in India. The results were obtained by the comparative study of neural network architectures like back propagation (BPNN), generalized regression (GRNN), and radial basis function (RBFNN).

Singh, Nitin, Saurabh Chaturvedi, and Shamim Akhter [11] developed a weather forecasting system that can be used in remote areas is the main motivation of this work. The data analytics and machine learning algorithms, such as random forest classification, are used to predict weather conditions. In this paper, a low-cost and portable solution for weather prediction is devised.

Balan, M. Selva, et al [12] proposed a few statistical analysis techniques and the use of Artificial Neural Network to predict the rainfall. The correlation between the attributes defines the influence on the prediction of any system. Attributes with no correlation can be removed as they do not contribute to the activation of the neuron. The loss functions, activation functions, number of neurons and the number of hidden layers also affect the accuracy of the system.

Parashar, Anubha [13] The main aim of this paper is to monitor and report weather conditions so that one is informed beforehand and necessary actions can be taken to reduce the damage by any calamity by forecasting it. Here the authors are using various sensors in order to collect the data and previous data is used in order to train the system and with current data collection we do the prediction. Explained Variance, Mean Absolute Error, and Median Absolute Error are used as the evaluation metrics in this paper.

Kalteh, Aman Mohammad [14] proposed a rainfall forecasting method based on coupling wavelet analysis and a novel artificial neural network technique called extreme learning machine (ELM). In this way, the unique characteristics of each technique are combined to capture different patterns in the data. At first, wavelet analysis is used to decompose rainfall time series into wavelet coefficients, and then the wavelet coefficients are used as inputs into the ELM model to forecast rainfall. The correlation coefficient (r), root mean square errors (RMSE) and Nash–Sutcliffe efficiency coefficient (NS) statistics are used as the performance metrics.

Lathifah, Siti Nur, et al [15] used Classification and Regression Tree (CART) algorithm to forecast the rainfall in Bandung Regency. Furthermore, an Adaptive Synthetic Sampling (ADASYN) algorithm was added to optimize the model produced due to a class imbalance in the data. The performance metrics like Precision, Recall, Accuracy and F1-Score are used in this paper.

3. PROPOSED INFORMATION GAIN WITH OPTIMIZATION BASED FEATURE SELECTION METHOD

3.1 Differential Evolution

Differential evolution (DE) is merely one of several approaches through evolutionary algorithm where in actuality the features are search and centred on ant colony. An easy and yet effective, DE give you the benefits usually requires like many optimization methods [16] [17]. There are several actions from DE such; 1) ability to handle non-differentiable, nonlinear and value this is certainly multimodal, 2) parallelizability to cope with computation cost that is intensive, 3) simplicity of good use, 4) good convergence properties.

Like GA, DE employ factors which can be same of mutation, selection and crossover. The efficiency of DE depends on the handling of target vector and difference in order to acquire a task vector in exploring procedure. Every real-value this is certainly d-Dimensional, a population of NP members is provided. NP will be the population size and D will be the true range that is wide of to be fine-tuned. Among the members of two population like y_{s2} and y_{s3} added the vector of weight difference to the y_{s1} which is third member for creating a trial vector. This action is termed as mutation. A mutant vector is generating relating to for every target vector $y_{(I,G)}, j = 1,2,3, \dots, M$ a mutant vector using the given equation:

$$w_{j,G+1} = y_{s1,H} + G\left(y_{s2,H} - y_{s2,H}\right) \qquad (2)$$

Where $s_1, s_2, s_3 \in \{1,2,\dots,NP\}$ are integers that are chosen randomly, should be specific from 1 another plus unique through the operating index j. The control rate of Scaling factor F(0,1) that your particular population comprises. In order to improve the variety in connection with perturbed factor vectors, introduction of crossover is takes place. The trial vector:

$$v_{j,H+1} = \left(v_{1,j,H+1}, v_{2,j,H+1}, \dots, v_{E,j,H+!}\right) \qquad (3)$$

Is from where;

$$v_{kj,H+1} = \begin{cases} w_{kj,H+1} \text{ if } rand(0,1) \leq d_s \\ y_{kj,H+1} \text{ otherwise} \end{cases} \quad (4)$$

Where the H is the current population and the trial vector k^{th} j^{th} for the dimension of $v_{(kj,H)}$. The probability of crossover $d_s(0,1)$ is a person described value that operates the portion in connection with parameter values which are often and that can be replicated through the mutant. Selection will be the stage to get the vector among the target vector as well as trial vector making use of the aim of generating an individual in terms of generation this is certainly next. Then your causing vector substitutes the vector with which it absolutely was compared [18] if the recently created vector leads to a lower objective function value (better fitness) as compared to population member that is predetermined. But, many factors from DE are instantly transformative without needed user to see by learning from your own errors strategy. In this work that is ongoing size of generation and population are adaptively identifying predicated on a total of features remained from relief-f. Hence, the buyer doesn't always have to initialize those factor values manually.

3.2 Information Gain

There are numerous means of scoring the features such as for example Information entropy, Correlation, Chi squared ensure that you Gini index [19][20]. Entropy is regarded as several techniques to measure diversity. Impurity of data may be measured by information entropy to quantify the uncertainty of predicting the worth regarding the goal variable. Let y be a discrete random variable with two outcomes that are possible. The entropy that is binary H, expressed in logarithmic base 2, for example. Shannon unit is written by Eq. (1):

$$H(y) = -p(+)\log_2 p(+) - p(-)\log_2\big(p(-)\big) \qquad (1)$$

Where, (+,-) will be the classes, $p(+)$ denotes the probability that a sample $y \in (+)$, and $p(-)$ denotes the probability that $y \in (-)$. Entropy quantifies the uncertainty of each and every feature in the act of decision making.

3.3 Optimization based Feature Selection Method for Road Classification

An important entropy-based filter feature selection [21] method called Information Gain is hybridized with the Differential Evolution optimization algorithm in this proposed optimization-based feature selection method for road classification. The ranked features obtained from the information Gain is given that input into the Optimization algorithm for optimal feature subset. Listed below are important stages when you look at the proposed Optimization based Feature Selection method.

Stage 1: Problem Formulation for Feature Selection

Assume, the M range that is wide of attributes for a presented classifier are described by $G_1, \ldots, G_P$. Let, $B = \{G_j : j = 1; P\}$.

Stage 2: Information Gain

The initial road classification dataset is given as the input to the filter-based feature selection Information Gain algorithm in this step. The information and knowledge entropy of each and every feature when you look at the dataset is calculated utilizing this IG algorithm.

Stage 3: Initialization of population and representation of chromosome

G is the chromosome length. The items of every chromosome are initialized randomly to either 1 or 0 if G is the total amount of features. Here, then it presents that j^{th} feature will not be involved in building the classifier then the j^{th} feature participates in constructing the classifier if it is 1 if the j^{th} position of a chromosome is 0.

Stage 4: Computation of Fitness

The calculation of fitness is performed making use of the M range that is wide of found in a chromosome that is certain for example. you will find total M number of 1's in that chromosome).

Stage 5: Mutation

A mutant vector/donor vector is generated relating to for every single target vector is $y_{j,H}; j = 1,2,3, \dots, NP$

$$w_{j,H+1} = y_{s1,H} + G(y_{s2,H} - y_{s2,H}) \qquad (2)$$

Where s1, s2, s3 Where s1, s2, s3 will be the random variables indices and belong to {1,2,…,NP}. These are generally some integer values, mutually various and G>0. The randomly taken integers s1, s2 and s3 are often selected to be various through the working index j, to help NP must certainly be greater or corresponding to four to allow for this disorder. G is a constant and suitable factor 0.5 [0,1] which handles the amplification in connection with differential variation $(y_{s2,H} - y_{s2,H})$.

Stage 6: Crossover

So that you can raise the diversity regarding the parameter that is perturbed, crossover is introduced. It is well-known given that recombination. The trial vector to this end

$$v_{j,H+1} = (v_{1,j,H+1}, v_{2,j,H+1}, \dots, v_{E,j,H+1}) \qquad (3)$$

Is made where:

$$v_{k,j,H+1} = w_{k,j,H+1} \text{ if } (\text{randb}(k) \leq DS) \text{ or } k = \text{rnbr}(j) \qquad (4)$$

$$= y_{k,j,H+1} \text{ if } (\text{randb}(k) > DS) \text{ or } k \neq \text{rnbr}(j) \qquad (5)$$

For k=1, 2,…,E.

In equation (3), $\text{randb}(k)$ could be the k^{th} evaluation of an uniform random number generator with outcome belongs to [0,1]. CR may be the crossover constant belongs to [0,1] which includes to be decided by the consumer. Here the worth of CR is 0.5. $\text{rnbr}(j)$ is a randomly chosen index x belongs to $v_{j,H+1}$ which ensures that{1,2,…,E} gets one or more parameter from $w_{j,H+1}$.

Stage 7: Selection

The trial vector to decide whether or not it should become a member of generation H+1. $v_{j,H+1}$ is set alongside the target vector $y_{j,H}$ with the criterions that are greedy. If vector $v_{j,H+1}$, otherwise, the value that is old is retained.

Stage 8: Termination Condition

The processes of selection, mutation, computation of fitness, and recombination (or crossover) are executed for a maximum number of generations in this approach. The string that is best seen as much as the very last generation gives the best subset of features. Here the string that is best contains a collection of features.

Algorithm: Optimization based Feature Selection Method

Input: Initial Feature Set

Output: Optimal Feature Subset

Step 1: Apply Information Gain on initial feature set

Step 1.1: for j=0 to m do

Calculate Information Entropy for several features when you look at the set using the equation (1)

End for

Step 1.2: list selected ranked features in ordered (P x M matrices)

Step 2: Applying Differential Evolution algorithm on the selected ranked features.

Step 2.1: Initialization of all parameter (population size, no features (10 to 15), no generation, fitness function)

Step 2.2: Population size=M and number of generations = M

Step 2.3: for j=1 to M(population size) do

Select s1,s2, s3∈ M randomly

For k=1 to E (Dimension) do

Select k_{rand} ∈ E

If $(rand() < DS\ of\ k = k_{rand})$ and calculate $w_{kj,H+1}$ from the equation (2)

End if

If $\left(f\left(v_{kj,h+1}\right) < f(Y_{kj,h})\right)$ then from the equation (3), (4) and (5)

End if

Step 2.4: End for

Step 3: End for

4. RESULT AND DISCUSSION

4.1 Dataset Description

The weather dataset is taken from the Kaggle Repository. Table 1 depicts the features involved in the weather rainfall dataset [22].

Table 1: Description of the Data set

Feature Number	Feature Name
1	meantempm
2	maxtempm
3	mintempm
4	meantempm_1
5	meantempm_2
6	meantempm_3
7	meandewptm_1
8	meandewptm_2
9	meandewptm_3
10	meanpressurem_1
11	meanpressurem_2
12	meanpressurem_3
13	maxhumidity_1
14	maxhumidity_2
15	maxhumidity_3
16	minhumidity_1
17	minhumidity_2
18	minhumidity_3
19	maxtempm_1
20	maxtempm_2
21	maxtempm_3
22	mintempm_1
23	mintempm_2
24	mintempm_3
25	maxdewptm_1
26	maxdewptm_2
27	maxdewptm_3
28	mindewptm_1
29	mindewptm_2
30	mindewptm_3
31	maxpressurem_1
32	maxpressurem_2
33	maxpressurem_3
34	minpressurem_1
35	minpressurem_2
36	minpressurem_3
37	precipm_1
38	precipm_2
39	precipm_3

40	Weather_Class (rain, no rain)

4.2 Number of Features Obtained

Table 2 depicts the number of features obtained by the Proposed Optimization based Feature Selection Algorithm, Information Gain, Differential Evolution algorithm. From the table 2, it is clear that the proposed optimization-based feature selection method gives a smaller number of features than the existing optimization algorithm like IG and DE algorithm. In the table 2, Proposed optimization-based feature selection algorithm gives only 27 features, IG gives 34 features whereas DE algorithm gives 31 features.

Table 2: Number of Features obtained by the Proposed Optimization based Feature Selection Method, Information Gain and Differential Evolution algorithm

S.No	Feature Selection Techniques		
	Proposed Optimization based Feature Selection	**Differential Evolution algorithm**	**Information Gain Algorithm**
1	meantempm	meantempm	meantempm
2	maxtempm	maxtempm	maxtempm
3	mintempm	mintempm	mintempm
4	meantempm_2	meantempm_1	meantempm_1
5	meantempm_3	meantempm_2	meantempm_2
6	meandewptm_3	meantempm_3	meantempm_3
7	meandewptm_1	meandewptm_1	meandewptm_1
8	meanpressurem_2	meandewptm_2	meandewptm_2
9	meanpressurem_3	meandewptm_3	meandewptm_3
10	maxhumidity_1	meanpressurem_1	meanpressurem_1
11	maxhumidity_3	meanpressurem_2	meanpressurem_2
12	minhumidity_1	meanpressurem_3	meanpressurem_3
13	minhumidity_2	maxhumidity_1	maxhumidity_1
14	maxtempm_2	maxhumidity_2	maxhumidity_2
15	maxtempm_3	maxhumidity_3	maxhumidity_3
16	mintempm_2	minhumidity_1	minhumidity_1
17	mintempm_3	minhumidity_2	minhumidity_2
18	maxdewptm_3	minhumidity_3	minhumidity_3
19	mindewptm_3	maxtempm_1	maxtempm_1
20	maxpressurem_2	maxtempm_2	maxtempm_2
21	maxpressurem_3	maxtempm_3	maxtempm_3
22	minpressurem_1	mintempm_1	mintempm_1
22	minpressurem_2	mintempm_2	mintempm_2
23	minpressurem_3	mintempm_3	mintempm_3
24	precipm_1	maxdewptm_1	maxdewptm_1

25	precipm_2	precipm_1	maxdewptm_2
26	precipm_3	precipm_2	maxdewptm_3
27	mindewptm_2	mindewptm_1	mindewptm_1
28		mindewptm_2	minpressurem_3
29		mindewptm_3	mindewptm_3
30		maxpressurem_1	maxpressurem_1
31		minpressurem_3	maxpressurem_2
32			precipm_2
33			minpressurem_1
34			precipm_3

4.3 Performance analysis of the proposed Optimization based Feature Selection Method

The classification techniques like Artificial Neural Network, Random Forest and Naïve Bayes are utilized in this research work to analyze the proposed Optimization based Feature Selection, Information Gain and Differential Evolution algorithms. The metrics like Accuracy, True Positive Rate (TPR or Recall), False Positive, Precision, Miss Rate and Specificity. Table 3 depicts the performance metrics for the evaluation.

Table 3: Performance Metrics

Performance Metrics	Equation
Accuracy (in %)	$\frac{TP + TN}{TP + FN + TN + FP}$
True Positive Rate (in %)	$\frac{TP}{TP + FN}$
False Positive Rate (in %)	$\frac{FP}{FP + TN}$
Precision (in %)	$\frac{TP}{TP + FP}$
Miss Rate (in %)	1-TPR
Specificity (in %)	1-FPR

Table 4 depicts the performance analysis of the Original Dataset, IG processed dataset, DE processed dataset and proposed IGOFS method processed dataset using ANN classification technique.

Table 4: Performance analysis of the original dataset, IG dataset, DE dataset and proposed IGOFS dataset using ANN classification

Performance Metrics (in %)	Feature Selection Techniques with ANN Classification			
	Original Dataset	IG Dataset	DE Dataset	Proposed IGOFS Dataset

Accuracy	45.65	57.05	59.9	89
TPR	54.62	60.77	63.21	93.62
FPR	67.62	51.47	46.38	16.2
Precision	63.13	71.13	73.21	86.78
Miss Rate	26.79	39.23	36.8	6.38
Specificity	32.38	48.53	53.62	83.8

Table 5 depicts the performance analysis of the Original Dataset, IG processed dataset, DE processed dataset and proposed IGOFS method processed dataset using RF classification technique.

Table 5: Performance analysis of the original dataset, IG dataset, DE dataset and proposed IGOFS dataset using RF classification

Performance Metrics (in %)	Feature Selection Techniques with RF Classification			
	Original Dataset	IG Dataset	DE Dataset	Proposed IGOFS Dataset
Accuracy	43.1	53.25	56.64	67.95
TPR	50.46	57.75	60.311	69.83
FPR	70.11	56.16	51.68	35.42
Precision	56.34	66.34	70.69	77.91
Miss Rate	49.54	42.25	39.69	30.17
Specificity	29.89	43.84	48.32	64.55

Table 6 depicts the performance analysis of the Original Dataset, IG processed dataset, DE processed dataset and proposed IGOFS method processed dataset using Naïve Bayes classification technique.

Table 6: Performance analysis of the original dataset, IG dataset, DE dataset and proposed IGOFS dataset using Naïve Bayes classification

Performance Metrics (in %)	Feature Selection Techniques with Naïve Bayes Classification			
	Original Dataset	IG Dataset	DE Dataset	Proposed IGOFS Dataset
Accuracy	41.85	51.2	54.85	65.25
TPR	49.51	56.78	59.18	67.81
FPR	73.82	58.77	54.04	39.28
Precision	57.82	63.30	69.21	75.30
Miss Rate	50.49	43.22	4082	32.19
Specificity	26.18	41.23	45.94	60.72

5. CONCLUSION

A new classification model was developed with the assistance got from the intelligent methods in this paper. Two main phases were utilized in this paper; they are (i) Optimal Feature selection and (ii) Classification. The dimensions regarding the data were high, and therefore the choice regarding the optimal features was a task that is complex that the proposed model used the suitable feature selection technology referred as IGOFS method to pick the suitable features. The investigation was dedicated to the aim of diminishing the correlation amongst the selected features, while they were pertaining to the generation of diverse information which was pertaining to different classes of information. Further, the classification regarding the feature was carried out following the variety of the features that are optimal. The classification regarding the selected features was through with NN, RF and NB which had the capacity to classify the info in a manner that is effective the selected features. Thus, the complete analysis that is experimental the effective performance of proposed IGOFSM way for the weather rainfall dataset classification method.

REFERENCE

[1] Datta, Anisha, Shukrity Si, and Sanket Biswas. "Complete Statistical Analysis to Weather Forecasting." Computational Intelligence in Pattern Recognition. Springer, Singapore, 2020. 751-763.

[2] Wang, Bin, et al. "Deep uncertainty quantification: A machine learning approach for weather forecasting." Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining. 2019.

[3] Refonaa, J., et al. "Machine Learning Techniques for Rainfall Prediction Using Neural Network." Journal of Computational and Theoretical Nanoscience 16.8 (2019): 3319-3323.

[4] Kachwala, Sakina, et al. "Predicting Rainfall from Historical Data Trends." Available at SSRN 3571738 (2020).

[5] Rasp, Stephan, et al. "Weather Bench: A benchmark dataset for data-driven weather forecasting." arXiv preprint arXiv:2002.00469 (2020).

[6] Grace, R. Kingsy, and B. Suganya. "Machine Learning based Rainfall Prediction." 2020 6th International Conference on Advanced Computing and Communication Systems (ICACCS). IEEE, 2020.

[7] Pham, Binh Thai, et al. "Development of advanced artificial intelligence models for daily rainfall prediction." Atmospheric Research 237 (2020): 104845.

[8] Ahmed, Kamal, et al. "Multi-model ensemble predictions of precipitation and temperature using machine learning algorithms." Atmospheric Research 236 (2020): 104806.

[9] Le, Vuong Minh, et al. "Daily Rainfall Prediction Using Nonlinear Autoregressive Neural Network." Micro-Electronics and Telecommunication Engineering. Springer, Singapore, 2020. 213-221.

[10] Sardeshpande, Kaushik D., and Vijaya R. Thool. "Rainfall Prediction: A Comparative Study of Neural Network Architectures." Emerging Technologies in Data Mining and Information Security. Springer, Singapore, 2019. 19-28.

[11] Singh, Nitin, Saurabh Chaturvedi, and Shamim Akhter. "Weather Forecasting Using Machine Learning Algorithm." 2019 International Conference on Signal Processing and Communication (ICSC). IEEE, 2019.

[12] Balan, M. Selva, et al. "Rainfall Prediction using Deep Learning on Highly Non-Linear Data.", International Journal of Research in Engineering, Science and Management, Volume-2, Issue-3, March-2019, pp. 590-592.

[13] Parashar, Anubha. "IoT Based Automated Weather Report Generation and Prediction Using Machine Learning." 2019 2nd International Conference on Intelligent Communication and Computational Techniques (ICCT). IEEE, 2019.

[14] Kalteh, Aman Mohammad. "Modular wavelet–extreme learning machine: a new approach for forecasting daily rainfall." Water Resources Management 33.11 (2019): 3831-3849.

[15] Lathifah, Siti Nur, et al. "Rainfall Forecasting using the Classification and Regression Tree (CART) Algorithm and Adaptive Synthetic Sampling (Study Case: Bandung Regency)." 2019 7th International Conference on Information and Communication Technology (ICo ICT). IEEE, 2019.

[16] Sudeeptha, J., and C. Nalini. "Hybrid Optimization of Cuckoo Search and Differential Evolution Algorithm for Privacy-Preserving Data Mining." International Conference on Artificial Intelligence, Smart Grid and Smart City Applications. Springer, Cham, 2019.

[17] Liang, Jing, et al. "Multimodal multi objective optimization with differential evolution." Swarm and evolutionary computation 44 (2019): 1028-1059.

[18] Brezočnik, Lucija, Iztok Fister, and Grega Vrbančič. "Applying Differential Evolution with Threshold Mechanism for Feature Selection on a Phishing Websites Classification." European Conference on Advances in Databases and Information Systems. Springer, Cham, 2019.

[19] Poornappriya, T. S., and M. Durairaj. "High relevancy low redundancy vague set-based feature selection method for telecom dataset." Journal of Intelligent & Fuzzy Systems Preprint: 1-18.

[20] Durairaj, M., and T. S. Poornappriya. "Choosing a spectacular Feature Selection technique for telecommunication industry using fuzzy TOPSIS MCDM." International Journal of Engineering & Technology 7.4 (2018): 5856-5861.

[21] Durairaj, M., and T. S. Poornappriya. "Why Feature Selection in Data Mining Is Prominent? A Survey." International Conference on Artificial Intelligence, Smart Grid and Smart City Applications. Springer, Cham, 2019.

[22] https://www.kaggle.com/datasets?search=weather+dataset