Annals of R.S.C.B., ISSN: 1583-6258, Vol. 25, Issue 4, 2021, Pages. 18515-18527
Received 05 March 2021; Accepted 01 April 2021

Optimization Based Feature Selection Method For Weather Dataset

T. MALATHI[1], DR. M. MANIMEKALAI [2]

[1]research Scholar, Department Of Computer Science, Shrimati Indira Gandhi College (Affiliated Bharathidasan University), Tiruchirappalli
[2] Head , Department Of Computer Science, Shrimati Indira Gandhi College (Affiliated Bharathidasan University), Tiruchirappalli

Abstract

Weather Forecasting Is An Emerging Domain That Predicts The Weather Condition At A Location At A Time. Weather Forecasting Is Considered As The Most Sensitive Research Field Which Facing A Lot Of Real-Time Issues Such As Inaccurate Prediction, Lack Of Handling In Huge Data Volume And Inadequate In Technology Advancement. Forecasting Weather Conditions Is Important For, E.G., Operation Of Hydro Power Plants And For Flood Management. Mechanistic Models Are Known To Be Computationally Demanding. Hence, It Is Of Interest To Develop Models That Can Predict Weather Conditions Faster Than Traditional Meteorological Models. The Field Of Machine Learning Has Received Much Interest From The Scientific Community. Due To Its Applicability In A Variety Of Fields, It Is Of Interest To Study Whether An Artificial Neural Network Can Be A Good Candidate For Prediction Of Weather Conditions In Combination With Large Data Sets. The Availability Of Meteorological Data From Multiple Online Sources Is An Advantage. In This Work, An Ant Colony – Information Gain Based Feature Selection Method Is Proposed Using Optimization Technique And Filter-Based Feature Selection Method.

Abstrak: (Abstract In Malysian Language) Gunakan Gaya Yang Diperuntukkan Dalam Dokumen Ini Untuk Menulis Kertas Anda. Anda Akan Mendapati Nama Gaya Antara Kurungan Pada Akhir Sifat Atau Perenggan. (Normal).

Key Words: Weather Prediction, Feature Selection, Ant Colony Optimization, Information Gain, Classification

1. Introduction

Weather Forecasting Has Been One Of The Most Scientifically And Technologically Challenging Problems Around The World In The Last Century [1]. This Is Due Mainly To Two Factors: First, It's Used For Many Human Activities And Secondly, Due To The Opportunism Created By The Various Technological Advances That Are Directly Related To This Concrete Research Field, Like The Evolution Of Computation And The Improvement In Measurement Systems. To Make An Accurate Prediction Is One Of The Major Challenges Facing Meteorologist All Over The World. Since Ancient Times, Weather Prediction Has Been One Of The Most Interesting And Fascinating Domain. Scientists Have Tried To Forecast Meteorological Characteristics Using A Number Of Methods, Some Of These Methods Being More Accurate Than Other [2].

To Predict The Weather By Numerical Means, Meteorologists Have Developed Atmospheric Models That Approximate The Atmosphere By Using Mathematical Equations To Describe How Atmospheric Temperature, Pressure, And Moisture Will Change Over Time. The Equations Are Programmed Into A Computer And Data On The Present Atmospheric Conditions Are Fed Into The Computer. The Computer Solves The Equations To Determine How The Different Atmospheric Variables Will Change Over The Next Few Minutes. The Computer Repeats This Procedure Again And Again Using The Output From One Cycle As The Input For The Next Cycle. For Some Desired Time In The Future (12, 24, 36, 48, 72 Or 120 Hours), The Computer Prints Its Calculated Information. It Then Analyzes The Data, Drawing The Lines For The Projected Position Of The Various Pressure Systems. The Final Computer-Drawn Forecast Chart Is Called A Prognostic Chart, Or Prog [3]. A Forecaster Uses The Progs As A Guide To Predicting The Weather. There Are Many Atmospheric Models That

Annals of R.S.C.B., ISSN: 1583-6258, Vol. 25, Issue 4, 2021, Pages. 18515-18527
Received 05 March 2021; Accepted 01 April 2021

Represent The Atmosphere, With Each One Interpreting The Atmosphere In A Slightly Different Way.

Climate Is The Long-Term Effect Of The Sun's Radiation On The Rotating Earth's Varied Surface And Atmosphere [4]. The Day-By-Day Variations In A Given Area Constitute The Weather, Whereas Climate Is The Longterm Synthesis Of Such Variations. Weather Is Measured By Thermometers, Rain Gauges, Barometers, And Other Instruments, But The Study Of Climate Relies On Statistics. Nowadays, Such Statistics Are Handled Efficiently By Computers. A Simple, Long-Term Summary Of Weather Changes, However, Is Still Not A True Picture Of Climate. To Obtain This Requires The Analysis Of Daily, Monthly, And Yearly Pattern [5].

2. Related Works

Singh, Nitin, Saurabh Chaturvedi, And Shamim Akhter [6] 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.

Wang, Bin, Et Al [7] The Authors Design A Data-Driven Method Augmented By An Effective Information Fusion Mechanism To Learn From Historical Data That Incorporates Prior Knowledge From Numerical Weather Prediction (Nwp). The Authors Casted The Weather Forecasting Problem As An End-To-End Deep Learning Problem And Solve It By Proposing A Novel Negative Log-Likelihood Error (Nle) Loss Function. The Metrics Like Mean Absolute Error (Mae) And Mean Squared Error (Mse) Are Used In This Paper.

Hua, Yuxiu, Et Al [8] The Authors Gave A Brief Introduction To The Structure And Forward Propagation Mechanism Of Lstm. The Authors Aimed At Reducing The Considerable Computing Cost Of Lstm, The Authors Put Forward A Rclstm Model By Introducing Stochastic Connectivity To Conventional Lstm Neurons. Therefore, Rclstm Exhibits A Certain Level Of Sparsity An Leads To A Decrease In Computational Complexity. The Metrics Like Root Mean Squared Error (Rmse) And Prediction Accuracy Are Considered In This Paper.

Hewage, Pradeep, Et Al [9] Proposed A Novel Lightweight Data-Driven Weather Forecasting Model By Exploring Temporal Modelling Approaches Of Long Short-Term Memory (Lstm) And Temporal Convolutional Networks (Tcn). The Proposed Deep Learning Networks With Lstm And Tcn Layers Are Assessed In Two Different Regressions, Namely Multi-Input Multi-Output And Multi-Input Single-Output. Mse Is Considered As The Evaluation Metrics In This Paper.

Karevan, Zahra, And Johan Ak Suykens [10] The Authors Proposed Transductive Lstm (T-Lstm) Which Exploits The Local Information In Time-Series Prediction. In This Study, A Quadratic Cost Function Is Considered For The Regression Problem. Localizing The Objective Function Is Done By Considering A Weighted Quadratic Cost Function At Which Point The Samples About The Test Point Have Larger Weights. Mae And Mse Are Used As The Performance Metrics In This Paper.

Wan, Renzhuo, Et Al [11] The Authors Utilized Various Deep Learning Models Based On Recurrent Neural Network (Rnn) And Convolutional Neural Network (Cnn) Methods. To Improve The Prediction Accuracy And Minimize The Multivariate Time Series Data Dependence For Aperiodic Data, The Dataset Are Analyzed By A Novel Multivariate Temporal Convolution Network (M-Tcn) Model. In This Model, Multi-Variable Time Series Prediction Is Constructed As A Sequence-To-Sequence Scenario For Non-Periodic Datasets. The Multichannel Residual Blocks In Parallel With Asymmetric Structure Based On Deep Convolution Neural Network Is Proposed. The Metrics Like Rmse, Rrse, And Correlation Are Used In This Paper.

Annals of R.S.C.B., ISSN: 1583-6258, Vol. 25, Issue 4, 2021, Pages. 18515-18527
Received 05 March 2021; Accepted 01 April 2021

Mehrkanoon, Siamak [12] Introduced Novel Data-Driven Predictive Models Based On Deep Convolutional Neural Networks (Cnn) Architecture For Temperature And Wind Speed Prediction In Weather Data. In Particular, The Proposed Deep Learning Framework Employs Different Upgrading Versions Of The Convolutional Neural Networks I.E. 1d-, 2d- And 3d-Cnn. The Introduced Models Exploit The Spatio-Temporal Multivariate Weather Data For Learning Shared Representations Using Historical Data And Forecasting Weather Elements For A Number Of User Defined Weather Stations Simultaneously In An End-To-End Fashion.

Dabrowski, Joel Janek, Yifan Zhang, And Ashfaqur Rahman [13] The Authors Argued That Time Invariance Can Reduce The Capacity To Perform Multi-Step-Ahead Forecasting, Where Modelling The Dynamics At A Range Of Scales And Resolutions Is Required. The Authors Proposed Forecast Net Which Uses A Deep Feed-Forward Architecture To Provide A Time Variant Model. An Additional Novelty Of Forecast Net Is Interleaved Outputs, Which The Authors Showed Assist In Mitigating Vanishing Gradients.

3. Information Gain Feature Selection Method

In This Module, An Information Gain [14][15] Methodology Is Used To Select The Features For Improving The Business Strategy By Analyzing The Weather Dataset.

Entropy Has Frequently Utilized In The Information Theory Measure, Which Characterizes The Purity Of An Arbitrary Collection Of Examples. It Is The Foundation Of The Information Gain (Ig) And Gain Ratio (Gr). The Entropy Measure Is Considered A Measure Of The System's Unpredictability. The Entropy Of Y Is

$$H(Y) = \sum_{y \in Y} p(y) \operatorname{Log}_2(p(y)) \qquad (1)$$

The Random *Feature* Y, The Marginal Probability Density Function Is P(Y). If The Detected Values Of *Y* In The Training Data Set *S* Has Divided Consorting To The Values Of A Next Feature *X*, And The Entropy Of *Y* Concerning The Split Induced By *X* Is Minimum Than The Entropy Of *Y* Before Dividing, Then There Is An Association Among Features X And Y. *Y* Later Observing *X* And Y Entropy Is Given By :

$$H(Y|X) = \sum_{x \in X} P(x) \sum_{y \in Y} P(y|x) \operatorname{Log}_2(p(y|x)) \qquad (2)$$

The Conditional Probability Of *Y* Is P(Y|X) For Given *X*

Given The Entropy Is A Condition Of Impurity In A Training Set S, It Can Describe A Measure Excogitating Additional Information About Y Rendered By X That Represents The Amount By Which The Entropy Of Y Decreases. This Measure Is Known As Ig. It Is Given By

$$\text{Ig} = H(Y)-H(Y \mid X) = H(X)-H(X \mid Y) \qquad (3)$$

In This Work, A New Feature Selection Method Is Proposed To Choose The Most Ideal Feature Subset For Removing The Irrelevant And Redundant Features From The Weather Dataset. In This Method, Aco Is Opted For Choosing The Most Ideal Feature Subsets With Information Gain Whereas Information Gain Is Used As A Heuristic Function For Choosing Most Appropriate Structures. Information Gain In This Contribution Is Used To Find The Fitness Function Of Aco Or Evaluation Of The Subset.

4. Ant Colony Optimization Based Feature Selection Method

Ant Colony Optimization (Aco) [16][17] A Population-Based Probabilistic Meta-Heuristics Aco Is Based On Ants Foraging Behavior. Foraging Behavior Of Ant Is An Interesting Phenomenon By Which Ant Colonies Find The Shortest Path Between Food Source And Nest Through Indirect Communication Called Stigmergy. Ants, Like Many Other Social Insects, Communicate With Each Other By Dropping A Chemical Substance On Their Path. This Chemical Substance Is Called Pheromone. It Provides A Positive Feedback Mechanism To Attract Other Ants. Those Paths Which Have A Higher Value Of Pheromone Have A High

Annals of R.S.C.B., ISSN: 1583-6258, Vol. 25, Issue 4, 2021, Pages. 18515-18527
Received 05 March 2021; Accepted 01 April 2021

Probability Of Being Selected. Whereas The Paths That Are Not Selected Their Pheromone Is Decreased By An Evaporation Process.

In Aco Each Ant Constructs A Complete Solution Using Two Things (1) Node Transition Probability Function Which Is Based On The Quantity Of Pheromone Spread By Ants And Heuristic Information About The Importance And Quality Of Each Individual Solutions And (2) Already Traversed Solutions Memory. As Generations Get Completed, Solutions Constructed By Each Ant Are Evaluated Using Some Evaluation Criteria. After That Pheromone Evaporation And Update, Mechanism Is Also Used Which Evaporates Intensity Of Pheromone From The Paths With Low Fitness Value And Hence Discarded Gradually. The Aco Algorithm Requires Specifying The Following Aspects For Implementation:

1) Representation Of The Problem Domain In Such A Way That It Lends Itself To Incrementally Building A Solution For The Problem, Usually In The Form Of A Graph.

2) Node Transition Probability Rule Based On The Amount Of Pheromone Value And Of The Heuristic Function We Have Employed Information Gain As A Heuristic Function. Following Is The Equation For Calculating The Probability Of Each Node:

$$P_j^i = \frac{[\tau(i,j)]^\alpha \, [\eta(i,j)]^\beta}{\sum_{k \, \epsilon S}[\tau(i,j)]^\alpha \, [\eta(i,j)]^\beta}$$

Where P_j^i Is A Probability Of The Ith Ant To Move From Node I To Node J At Time T. $P_j^i(t)$ =0 Means That Ants Are Not Allowed To Move To Any Node In The Neighbour. $[\tau(i,j)]^\alpha$ Is The Amount Of Pheromone On The Edge Connecting I And J, Where Is A Constant Which Is Used To Control Relative Importance Of Pheromone Information. After Each Iteration, This Pheromone Information Is Updated By All The Ants And In Some Versions Of Aco Only Best Ant Can Update Pheromone. $[\eta(i,j)]^\beta$ Is The Heuristic Function That Denotes The Heuristic Value Of Edge Connecting I And J. Usually, The Heuristic Value Does Not Change During Execution Of The Algorithm. In This Paper We Have Used Information Gain To Denote Heuristic Value. β Is A Constant Which Is Used To Control Relative Importance Of Heuristic Value.

3) A Heuristic Evaluation Function Called Fitness Function Dependent On The Problem, Which Provides A Goodness Measurement For The Different Solution Components. We Have Used Fitness Function Is Based On Information Gain To Normalize The Biasness Of Information Gain And Mutual Information Towards Multi-Valued Attributes. Following Formula Being Used To Compute The Value Of The Selected Subset.

$$Fitness\ (S) = \frac{(F - S) * (\sum_{i=1}^{n}(IG))}{F}$$

Where S Is Reduced Subset Selected By Aco, Ig Is The Information Gain Of Feature I In The Subset S And F Is The Total Number Of Features Present In The Dataset . It Will Select Feature Subset With High Information Gain Value And With Less Number Of Features.

4) Pheromone Evaporation And Updating Rule Which Takes Into Account The Evaporation And Reinforcement Of The Paths. Once Subsets Are Evaluated Using Fitness Function, Pheromone Trails Are Updated. Firstly Using An Evaporation Rate P The Pheromone Trails On The Edges Are Evaporated Or Decreased To Minimize The Effect Of A Sub-Optimal Feature To Which The Ants Have Previously Converged. Secondly Amount Of Pheromone On The Edges Is Updated With Amounts Proportional To The Fitness Of The Solution. Some Approaches For Pheromone Updating Allowed All The Ants To Update Their Paths According To The Fitness Of Their Solution And In Some Approaches Only Best Ant Is Allowed To Update Pheromone Value On Its Path.

For The Pheromone Evaporation And Updating Following Equations Are Used.

$$\tau = (1 - P) * \ \tau \ where \ p \ is \ 0.15$$

$$\tau = \tau + (\tau * \ Q)$$

Annals of R.S.C.B., ISSN: 1583-6258, Vol. 25, Issue 4, 2021, Pages. 18515-18527
Received 05 March 2021; Accepted 01 April 2021

$$\text{And } Q = \left[1 - \left(\frac{1}{1+Fitness}\right)\right]$$

The Above Equations, Used For Pheromone Evaporation And Updating.

5) Where —"Fitness" Is The Value Of The Selected Subset Through An Independent Statistical Measure.

6) Stopping/Convergence Criterion That Decides When The Algorithm Terminates Usually Depends On Maximum Number Of Iterations.

5. Proposed Ant Colony – Information Gain Based Feature Selection Method

In This Approach Aco Is Used For Selecting Most Optimal Feature Subsets Along With Information Gain Where Information Gain Is Used As Heuristic Function For Selecting Most Relevant Features. Fitness Function Or Subset Evaluation Is Also Based On Information Gain. It's A Pure Filter Approach, Along With It We Have Also Used Classifier Ensemble to Improve Predictive Performance of Filter Approaches Comparable to the Wrapper Approaches.

Figure 1 Depicts Proposed Ac-Ig Feature Selection Method. It Is Filter Approach Which Selects Features On The Basis Independent Information Gain Measure. So, Some Features That Might Be Less Important In Terms Of Independent Relevance To Class But For A Classifier, Such Features Could Be Important. Therefore, Ac-Ig Uses Classifier Ensemble On Different Convergence Threshold Values And Classification Accuracy Of Subsets Is Used To Provide Final Feature Subset.

In Proposed Approach First Dataset Is Loaded. Once Dataset Is Loaded, Information Gain Of Each Feature/Attribute In Data Set Is Computed. Then All The Parameters Of Ant Colony Optimization Algorithm Are Initialized. Such As Number Of Ants, A And B Values Of Node Transition Probability Function, Path Convergence Threshold Value, Pheromone Evaporation Rate P. And Maximum Number Of Generations. A Search Space Is Constructed That Consists Of Nodes Proportional To The Number Of Features In The Dataset. Fixed Numbers Of Ants Are Generated In Each Iteration Where Each Ant Generates A Candidate Solution. After Each Generation, Generated Solutions Are Evaluated Using A Subset Evaluator. Subset Evaluator Is Based On Information Gain Between Selected Features And The Class. After Subset Evaluation Best Solution Is Gained On The Basis Of Maximum Fitness Value And Is Preserved. Then Termination Criteria Of The Algorithm Are Checked Which Is Based On Two Conditions I.E. On A Maximum Number Of Generations And Convergence Threshold. If Termination Criteria Are Not Met Each Ant Updates Its Pheromone Value To The Quality Of Solution Generated By Each Ant. Otherwise, If Any Termination/Stopping Criterion Is Met Algorithm Outputs Ten Best Subsets. These Subsets Are Provided To Decision Tree Classifier Techniques Like J48, Id3 And Classification And Regression Tree (Cart).

Annals of R.S.C.B., ISSN: 1583-6258, Vol. 25, Issue 4, 2021, Pages. 18515-18527
Received 05 March 2021; Accepted 01 April 2021

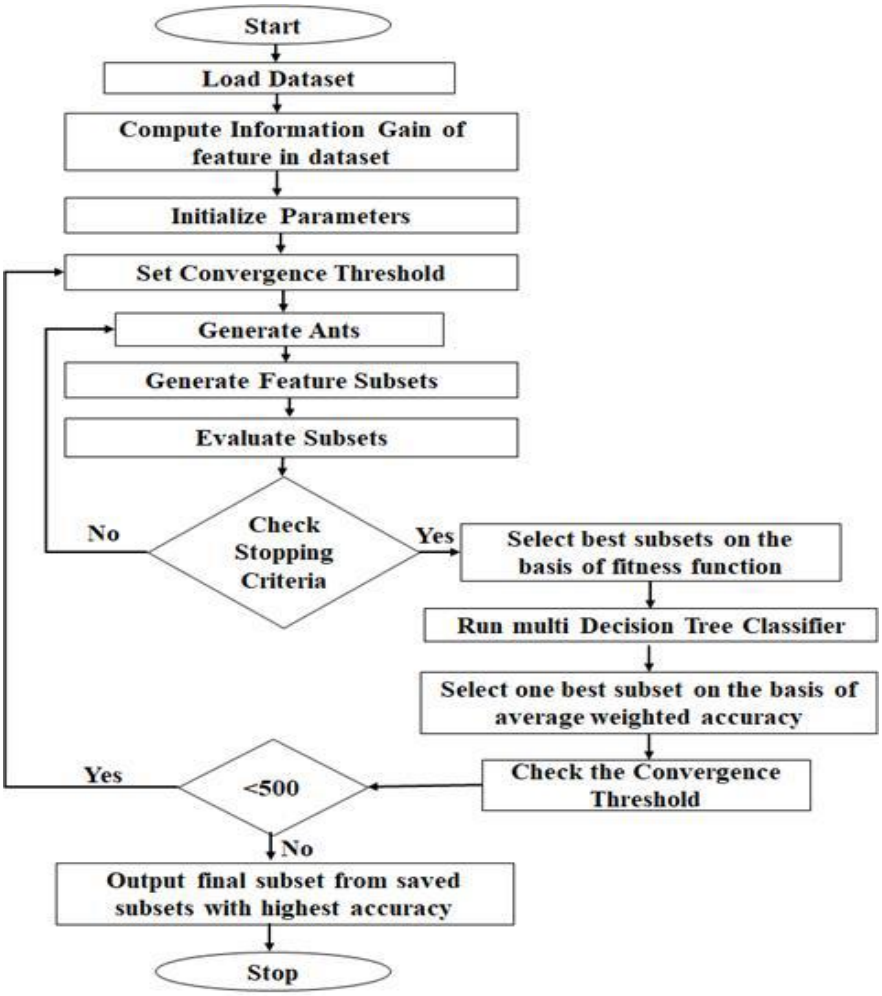

Fig. 1. Proposed Ant Colony – Information Gain Based Feature Selection Method

Then One Subset Is Selected Based On The Highest Average Weighted Accuracy Of Classifier Ensemble And Saved. Then Again Convergence Threshold Is Checked. If It Is Less Than 500, The Whole Process Is Repeated. Otherwise, The Algorithm Stops And One Best Subset With The Highest Accuracy Is Selected From All Saved Subsets And Is Considered As Final Subset Then New Ants Are Produced And This Complete Process Goes Iteratively Till Highest Number Of Epochs Is Reached Or The Algorithm Convergence To A Solution.

Annals of R.S.C.B., ISSN: 1583-6258, Vol. 25, Issue 4, 2021, Pages. 18515-18527
Received 05 March 2021; Accepted 01 April 2021

Input: Weather Dataset
Output: Pre-Processed Dataset
Step 1: Start
Step 2: Load the dataset
Step 3: Compute information Gain (heuristic value) of all the attributes/features in a dataset.
Step 4: Initialize ACO parameters
Step 5: Set convergence threshold (50:50:500)
Step 6: Do 1: maximum generations
Step 6.1: Each ant generates solutions
Step 6.2: Evaluate each solution
Step 6.3: Keep track of best solutions
Step 6.4: Check stopping criteria (yes: go to 10)
Step 7: Update pheromone for all ants.
Step 8: Generate new ants
Step 9: Go to 6
Step 10: Select 10 best subsets of converged /maximum generations
Step 11: Run multi classifier ensemble
Step 12: Select subset with the highest average accuracy
Step 13: Save it
Step 14: Check if convergence threshold >500(yes: go to 18)
Step 15: Go to 5
Step 16: Select one best subset from all the saved subsets that have got the highest average classification accuracy of classifier ensemble
Step 17: Output best subset as a final subset
Step 18: Stop.

6. Result And Discussion

In This Research Paper, The Evaluation Of The Feature Selection Techniques Like Ig, Aco, And Proposed Ant Colony – Information Gain (Ac-Ig) Based Feature Selection Method Is Done With The Classification Techniques Like Id3, Cart And J48 For Three Different Weather Datasets. These Weather Datasets Are Considered And Download In Kaggle Repository [18]. The Evaluation Metrics Like Accuracy (In %), True Positive Rate (In %), Precision (In %), False Positive Rate (In %), Miss Rate (In %), And False Discovery Rate (In %) Are Considered In This Work To Evaluate The Performance Of The Feature Selection Techniques With Different Classifiers.

The Efficiency Of The Proposed Ac-Ig Based Feature Selection Method Is Evaluated With The Classification Techniques Like Cart, Id3, J48 Classical Classification Techniques With The Above-Mentioned Performance Metrics. Table 1 Gives Number Of Features Obtained By The Existing Ig, Aco And Proposed Ac-Ig Feature Selection Method. From The Table 1, It Is Clear That The Proposed Ac-Ig Feature Selection Method Gives A Smaller Number Of Features When It Is Compared With Other Existing Feature Selection Techniques.

Table 1: Number Of Features Obtained By Ig, Aco And Proposed Ac-Ig Feature Selection Method

Feature Selection Techniques	Number Of Features Obtained
Information Gain (Ig)	35
Ant Colony Optimization	33
Proposed Ac-Igfs Method	24

2.1. Performance Analysis Of The Feature Selection Methods

Annals of R.S.C.B., ISSN: 1583-6258, Vol. 25, Issue 4, 2021, Pages. 18515-18527
Received 05 March 2021; Accepted 01 April 2021

The Classification Techniques Like Id3, Cart And J48 Are Used To Analyse The Performance Of The Feature Selection Techniques Using Evaluation Metrics. Table 2 Gives Classification Accuracy (In %) Of The Ig, Aco, And Proposed Ig-Ac Feature Selection Using Id3, Cart And J48 Classifier. Figure 2 Depicts The Graphical Representation Of The Classification Accuracy (In %) Of The Original Dataset, Ig, Aco And Proposed Ig-Ac Based Feature Selection Method Using Id3, Cart And J48 Classifiers. From The Table 2 And Figure 2, The Proposed Ig-Ac Based Feature Selection Method With Cart Classifiers Gives More Accuracy Than Other Techniques.

Table 2: Classification Accuracy (In %) Of The Original Dataset, Ig, Aco And Proposed Ig-Ac Based Feature Selection Method Using Id3, Cart And J48 Classifiers

Feature Selection Techniques	Classification Accuracy (In %)		
	Id3	Cart	J48
Original Dataset	56.25	58.63	55.95
Information Gain (Ig)	78.45	82.54	72.96
Ant Colony Optimization (Aco)	80.12	83.63	73.85
Proposed Ig-Ac Based Feature Selection Method	86.42	87.64	83.36

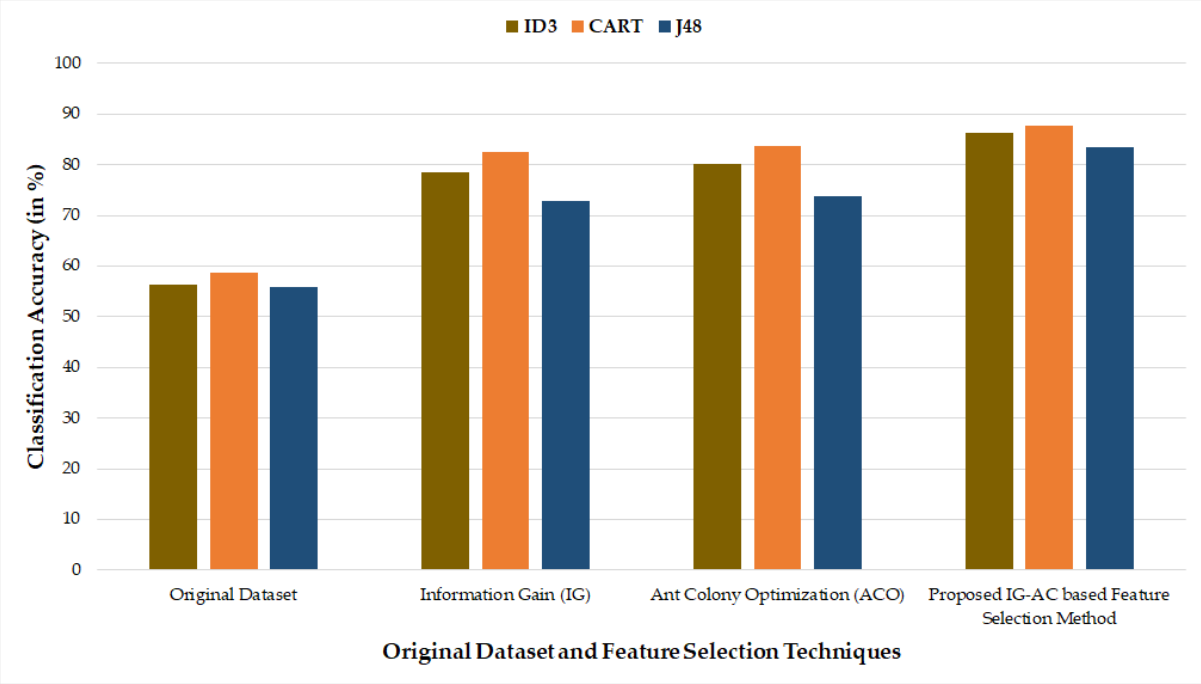

Figure 2: Graphical Representation Of The Classification Accuracy (In %) Of The Original Dataset, Ig, Aco And Proposed Ig-Ac Based Feature Selection Method Using Id3, Cart And J48 Classifiers

Table 3 Gives True Positive Rate (In %) Of The Ig, Aco, And Proposed Ig-Ac Feature Selection Using Id3, Cart And J48 Classifier. Figure 3 Depicts The Graphical Representation Of The True Positive Rate (In %) Of The Original Dataset, Ig, Aco And Proposed Ig-Ac Based Feature Selection Method Using Id3, Cart And J48 Classifiers. From The Table 3 And Figure 3, The Proposed Ig-Ac Based Feature Selection Method With Cart Classifiers Gives More True Positive Rate Than Other Techniques.

Table 3: True Positive Rate (In %) Of The Original Dataset, Ig, Aco And Proposed Ig-Ac Based Feature Selection Method Using Id3, Cart And J48 Classifiers

Feature Selection Techniques	True Positive Rate (In %)		
	Id3	Cart	J48
Original Dataset	54.47	56.53	51.73

Annals of R.S.C.B., ISSN: 1583-6258, Vol. 25, Issue 4, 2021, Pages. 18515-18527
Received 05 March 2021; Accepted 01 April 2021

Information Gain (Ig)	75.4	79.8	76.4
Ant Colony Optimization (Aco)	76.9	81.7	77.4
Proposed Ig-Ac Based Feature Selection Method	**84.64**	**85.42**	**81.58**

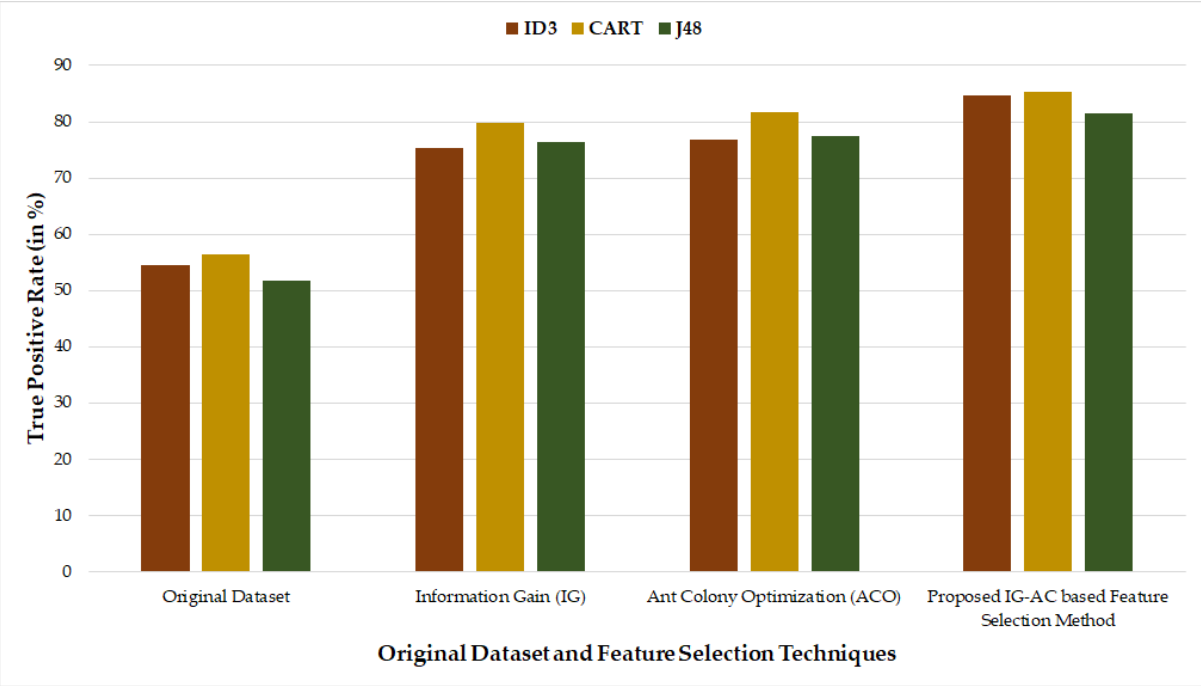

Figure 3: Graphical Representation Of The True Positive Rate (In %) Of The Original Dataset, Ig, Aco And Proposed Ig-Ac Based Feature Selection Method Using Id3, Cart And J48 Classifiers

Table 4 Gives Precision (In %) Of The Ig, Aco, And Proposed Ig-Ac Feature Selection Using Id3, Cart And J48 Classifier. Figure 4 Depicts The Graphical Representation Of The Precision (In %) Of The Original Dataset, Ig, Aco And Proposed Ig-Ac Based Feature Selection Method Using Id3, Cart And J48 Classifiers. From The Table 4 And Figure 4, The Proposed Ig-Ac Based Feature Selection Method With Cart Classifiers Gives More Precision Than Other Techniques.

Table 4: Precision (In %) Of The Original Dataset, Ig, Aco And Proposed Ig-Ac Based Feature Selection Method Using Id3, Cart And J48 Classifiers

Feature Selection Techniques	**Precision (In %)**		
	Id3	**Cart**	**J48**
Original Dataset	41.27	47.69	45.82
Information Gain (Ig)	68.57	73.72	69.74
Ant Colony Optimization (Aco)	79.34	82.41	75.76
Proposed Ig-Ac Based Feature Selection Method	**87.31**	**88.86**	**84.58**

Annals of R.S.C.B., ISSN: 1583-6258, Vol. 25, Issue 4, 2021, Pages. 18515-18527
Received 05 March 2021; Accepted 01 April 2021

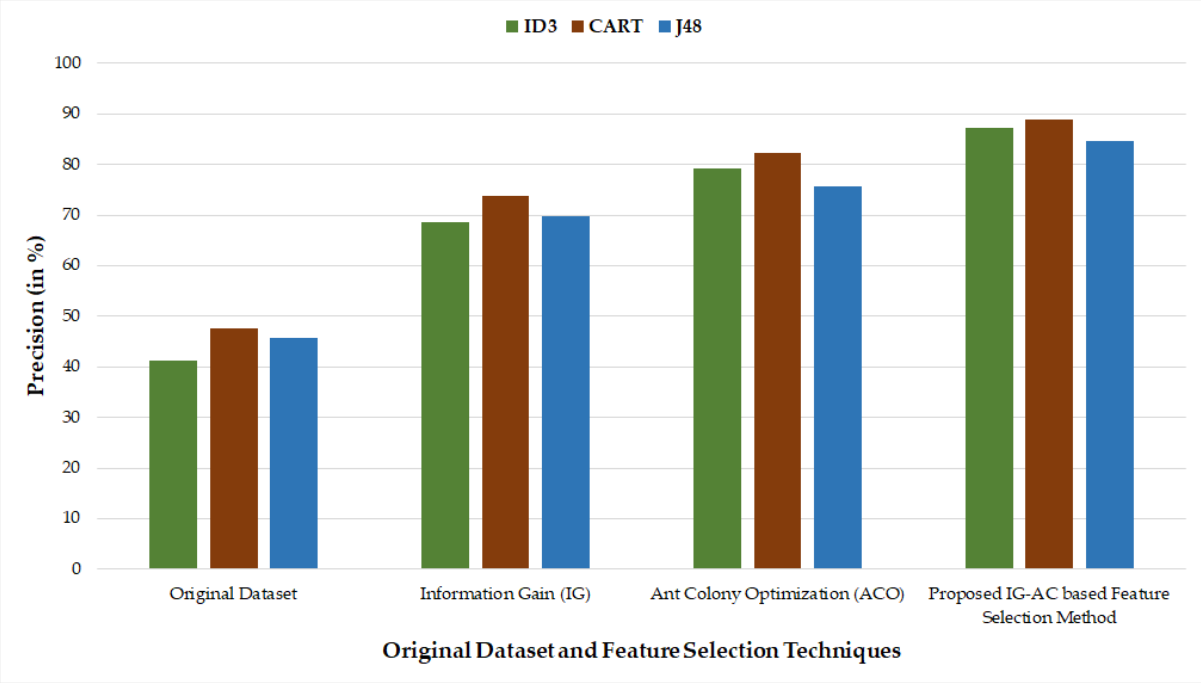

Figure 4: Graphical Representation Of The Precision (In %) Of The Original Dataset, Ig, Aco And Proposed Ig-Ac Based Feature Selection Method Using Id3, Cart And J48 Classifiers

Table 5 Gives False Positive Rate (In %) Of The Ig, Aco, And Proposed Ig-Ac Feature Selection Using Id3, Cart And J48 Classifier. Figure 5 Depicts The Graphical Representation Of The False Positive Rate (In %) Of The Original Dataset, Ig, Aco And Proposed Ig-Ac Based Feature Selection Method Using Id3, Cart And J48 Classifiers. From The Table 5 And Figure 5, The Proposed Ig-Ac Based Feature Selection Method With Cart Classifiers Gives Less False Positive Rate Than Other Techniques.

Table 6 Gives Miss Rate (In %) Of The Ig, Aco, And Proposed Ig-Ac Feature Selection Using Id3, Cart And J48 Classifier. Figure 6 Depicts The Graphical Representation Of The Miss Rate (In %) Of The Original Dataset, Ig, Aco And Proposed Ig-Ac Based Feature Selection Method Using Id3, Cart And J48 Classifiers. From The Table 6 And Figure 6, The Proposed Ig-Ac Based Feature Selection Method With Cart Classifiers Gives Less Miss Rate Than Other Techniques.

Table 7 Gives False Discovery Rate (In %) Of The Ig, Aco, And Proposed Ig-Ac Feature Selection Using Id3, Cart And J48 Classifier. Figure 7 Depicts The Graphical Representation Of The False Discovery Rate (In %) Of The Original Dataset, Ig, Aco And Proposed Ig-Ac Based Feature Selection Method Using Id3, Cart And J48 Classifiers. From The Table 7 And Figure 7, The Proposed Ig-Ac Based Feature Selection Method With Cart Classifiers Gives Less False Discovery Rate (In %) Than Other Techniques.

Table 5: False Positive Rate (In %) Of The Original Dataset, Ig, Aco And Proposed Ig-Ac Based Feature Selection Method Using Id3, Cart And J48 Classifiers

Feature Selection Techniques	False Positive Rate (In %)		
	Id3	Cart	J48
Original Dataset	58.24	55.69	59.32
Information Gain (Ig)	32.1	24.1	38.1
Ant Colony Optimization (Aco)	29.1	21.2	47.6
Proposed Ig-Ac Based Feature Selection Method	21.9	19.32	27.47

Annals of R.S.C.B., ISSN: 1583-6258, Vol. 25, Issue 4, 2021, Pages. 18515-18527
Received 05 March 2021; Accepted 01 April 2021

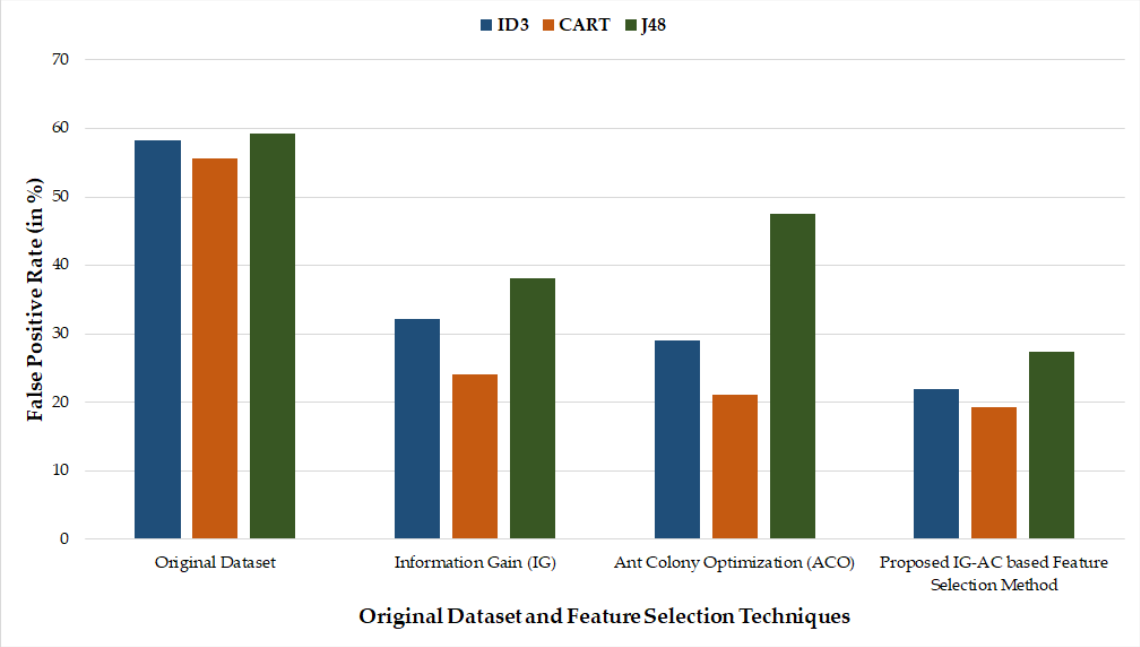

Figure 5: Graphical Representation Of The False Positive Rate (In %) Of The Original Dataset, Ig, Aco And Proposed Ig-Ac Based Feature Selection Method Using Id3, Cart And J48 Classifiers

Table 6: Miss Rate (In %) Of The Original Dataset, Ig, Aco And Proposed Ig-Ac Based Feature Selection Method Using Id3, Cart And J48 Classifiers

Feature Selection Techniques	Miss Rate (In %)		
	Id3	**Cart**	**J48**
Original Dataset	45.53	43.47	48.27
Information Gain (Ig)	24.6	20.2	23.6
Ant Colony Optimization (Aco)	23.1	18.3	22.6
Proposed Ig-Ac Based Feature Selection Method	**15.36**	**14.58**	**18.42**

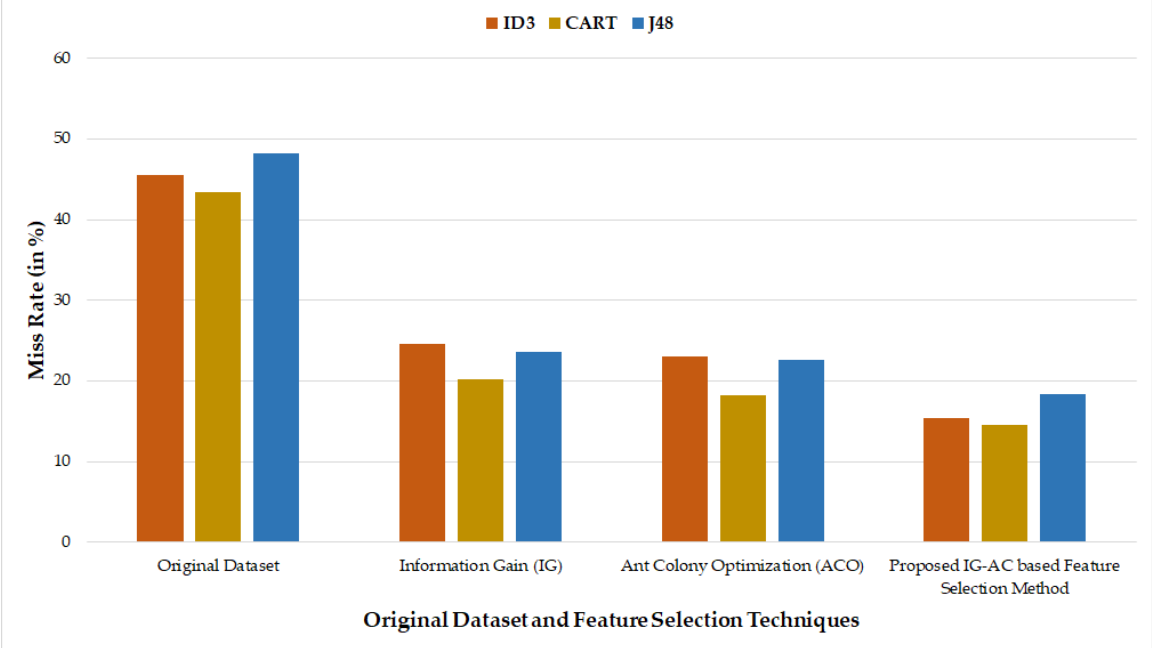

Figure 6: Graphical Representation Of The Miss Rate (In %) Of The Original Dataset, Ig, Aco And Proposed Ig-Ac Based Feature Selection Method Using Id3, Cart And J48 Classifiers

Annals of R.S.C.B., ISSN: 1583-6258, Vol. 25, Issue 4, 2021, Pages. 18515-18527
Received 05 March 2021; Accepted 01 April 2021

Table 7: False Discovery Rate (In %) Of The Original Dataset, Ig, Aco And Proposed Ig-Ac Based Feature Selection Method Using Id3, Cart And J48 Classifiers

Feature Selection Techniques	Miss Rate (In %)		
	Id3	**Cart**	**J48**
Original Dataset	58.73	52.31	54.18
Information Gain (Ig)	31.43	26.28	30.26
Ant Colony Optimization (Aco)	20.66	17.59	24.24
Proposed Ig-Ac Based Feature Selection Method	**12.69**	**11.14**	**15.42**

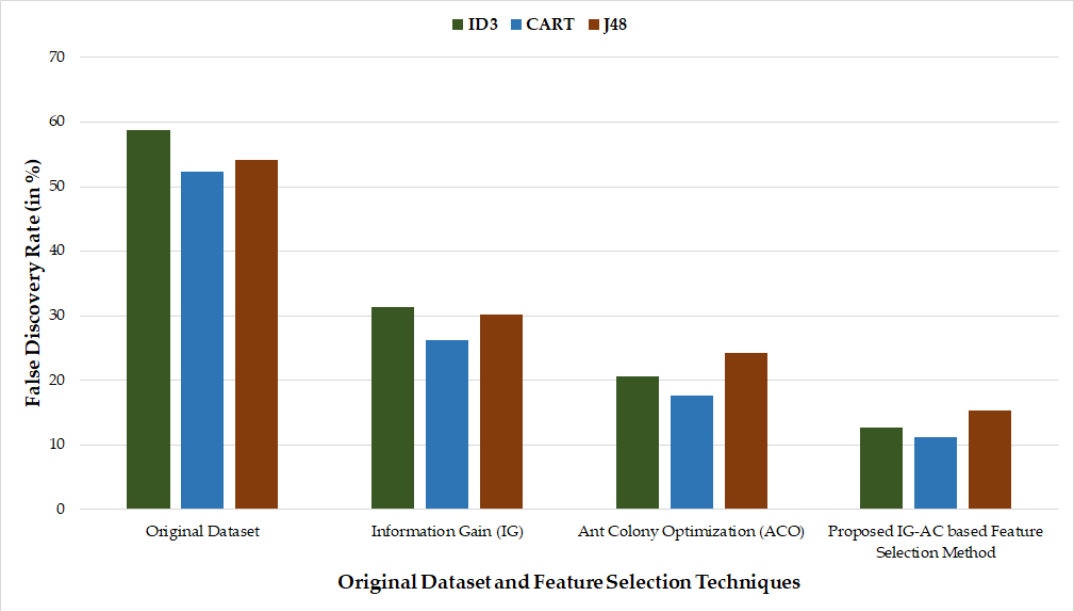

Figure 7: Graphical Representation Of The False Discovery Rate (In %) Of The Original Dataset, Ig, Aco And Proposed Ig-Ac Based Feature Selection Method Using Id3, Cart And J48 Classifiers

7. Conclusion

Ant Colony – Information Gain (Ac-Ig) Based Feature Selection Method Has Been Proposed In This Paper. Ac-Ig Is Proposed By Combining The Ig And Aco With A Classifiers And Has Been Used To Optimize The Feature Subset Selection Process. Results Showed That Proposed Approach Has Performed Well In Terms Of Dimensionality Reduction And Classification Accuracy As Compared To Other Approaches. The Information Gain Is Used As A Heuristic Measure In Ac-Ig Which Has Normalized The Biases Of Another Heuristic Measure Towards Multi-Valued Attributes And Selected Features That Are Highly Relevant To The Class. Secondly, The Classifier Ensemble Has Been Used In A Novel Way With Aco. It Checks The Classification Accuracy Of Subsets Achieved On Different Convergence Threshold. Classifiers Helps To Opt Important Features That Are Not Selected By Independent Measure. It Has Not Used Classifier For Optimizing Results Rather We Have Used It Only For Selecting A Subset With The Highest Accuracy, So The Proposed Approach Is Not Computationally Costly.

References

[1] Abdel-Kader, Hatem, Mustafa Abd-El Salam, And Mona Mohamed. "Hybrid Machine Learning Model For Rainfall Forecasting." *Journal Of Intelligent Systems And Internet Of Things* 1.1 (2021): 5-12.

[2] Diao, Li, Et Al. "Short-Term Weather Forecast Based On Wavelet Denoising And Catboost." *2019 Chinese Control Conference (Ccc)*. Ieee, 2019.
[3] Kale, Shivani S., And Preeti S. Patil. "Data Mining Technology With Fuzzy Logic, Neural Networks And Machine Learning For Agriculture." *Data Management, Analytics And Innovation*. Springer, Singapore, 2019. 79-87.
[4] Gad, Ibrahim, And Doreswamy Hosahalli. "A Comparative Study Of Prediction And Classification Models On Ncdc Weather Data." *International Journal Of Computers And Applications* (2020): 1-12.
[5] Hill, Aaron J., Gregory R. Herman, And Russ S. Schumacher. "Forecasting Severe Weather With Random Forests." *Monthly Weather Review* 148.5 (2020): 2135-2161.
[6] Singh, Nitin, Saurabh Chaturvedi, And Shamim Akhter. "Weather Forecasting Using Machine Learning Algorithm." *2019 International Conference On Signal Processing And Communication (Icsc). Ieee*, 2019.
[7] 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.
[8] Hua, Yuxiu, Et Al. "Deep Learning With Long Short-Term Memory For Time Series Prediction." *Ieee Communications Magazine* 57.6 (2019): 114-119 .
[9] Hewage, Pradeep, Et Al. "Deep Learning-Based Effective Fine-Grained Weather Forecasting Model." *Pattern Analysis And Applications* (2020): 1-24.
[10] Karevan, Zahra, And Johan Ak Suykens. "Transductive Lstm For Time-Series Prediction: An Application To Weather Forecasting." *Neural Networks* 125 (2020): 1-9.
[11] Wan, Renzhuo, Et Al. "Multivariate Temporal Convolutional Network: A Deep Neural Networks Approach For Multivariate Time Series Forecasting." *Electronics* 8.8 (2019): 876.
[12] Mehrkanoon, Siamak. "Deep Shared Representation Learning For Weather Elements Forecasting." *Knowledge-Based Systems* 179 (2019): 120-128.
[13] Dabrowski, Joel Janek, Yifan Zhang, And Ashfaqur Rahman. "Forecastnet: A Time-Variant Deep Feed-Forward Neural Network Architecture For Multi-Step-Ahead Time-Series Forecasting." *Arxiv Preprint* Arxiv:2002.04155 (2020).
[14] Ahuja, Ravinder, Et Al. "Classification And Clustering Algorithms Of Machine Learning With Their Applications." *Nature-Inspired Computation In Data Mining And Machine Learning*. Springer, Cham, 2020. 225-248.
[15] Singer, Gonen, Roee Anuar, And Irad Ben-Gal. "A Weighted Information-Gain Measure For Ordinal Classification Trees." *Expert Systems With Applications* 152 (2020): 113375.
[16] Pan, Mingzhang, Et Al. "Photovoltaic Power Forecasting Based On A Support Vector Machine With Improved Ant Colony Optimization." *Journal Of Cleaner Production* 277 (2020): 123948.
[17] Mehdizadeh, Saeid, Et Al. "Implementing Novel Hybrid Models To Improve Indirect Measurement Of The Daily Soil Temperature: Elman Neural Network Coupled With Gravitational Search Algorithm And Ant Colony Optimization." *Measurement* 165 (2020): 108127.
[18] Https://Www.Kaggle.Com/Datasets?Search=Weather+Dataset