Webology (ISSN: 1735-188X)
Volume 16, Number 2, 2019

Feature Selection For Crop Yield Prediction Using Optimization Techniques

Chellammal Surianarayanan , Kodimalar Palanivel

Department of Computer Science Bharathidasan University Constituent Arts and Science College Navalurkuttapattu, Srirangam Tk., Affiliated to Bharathidasan University Tiruchirappalli- 620 027, Tamilnadu, India.

Abstract

Agriculture sector forms the basis of Indian economy. Prediction of yield of crops is an important issue in agriculture as the domain is facing a huge challenge of producing effective crops having high yield while maintaining the sustainability of natural resources. In addition, early prediction of yield enable the farmers to take precautionary actions to improve the productivity of crops. Crop yield depends on various factors including weather parameters, soil parameters, irrigation availability, fertilizer's needs, farm capacity, etc. In general, early prediction is being done by analyzing the above data which is archived over several years using machine learning techniques. Identifying relevant and more useful attributes from the set of all available attributes(called feature selection) which really affect the yield becomes important before performing the prediction process using machine learning algorithms. The focus of this work lies with feature selection. In this work, four different feature selection techniques, namely, Binary Cuckoo Search(BCS), Relief, Grey Wolf Optimization(GWO) and Principal Component Analysis(PCA) have been employed over the agricultural data collected from field and preprocessed using Monte Carlo method. The selected features are given for yield prediction using multiple regression. The accuracy of prediction obtained using the above four methods are presented. Results are discussed.

Keywords – crop yield prediction, feature selection, Binary Cuckoo Search, Grey Wolf Optimization, Principal Component Analysis, machine learning techniques, multiple regression

I INTRODUCTION

In agriculture crop yield prediction is an important problem. Every farmer is always trying to know how much yield will get from his expectation. In the past, yield prediction was calculated by analyzing the farmer's previous experience on a particular crop. The Agricultural yield primarily depends on weather conditions, pests and planning of harvest operation. Accurate yield prediction is a major problem that ought to be addressed for making decisions related to agricultural risk management. Early prediction of yield would facilitate the farmers to make precautionary actions to improve productivity. Early prediction is possible through collection and archival of data related to previous experience of the farmers, weather conditions, soil parameters, water parameters and other influencing factors such as rainfall, temperature, humidity, solar radiation, crop population density, fertilizer application, irrigation, tillage, type of soil, depth, farm capacity, and soil organic matter. Manually extracting knowledge from the archived data is tedious and machine learning plays a crucial role in yield prediction

Typically, machine learning algorithms are extensively used for predicting the yield[1]. Two aspects that become more important before predicting the yield. One aspect that should be considered during preprocessing of the data is that different parameters are being collected at different acquisition rate. For example, temperature, pressure, wind speed and wind direction are being routinely collected whereas parameters such as water quality parameters and soil quality parameters are collected at different rate say for example, once in 3 months. So, for an attribute such as temperature, there may be around 365 readings per year (with number of readings taken per days is 1) whereas for an attribute such as soil quality parameters there may be 4 values with number of readings taken per year is 4. In addition, for each crop, yield values are recorded once per year. In its raw form, the archived data has the form of multiple valued function characteristics. Here, Monte Carlo method[2] can be used to select the data at random which enables to arrive as single valued function between various features and yield of the crop.

The second aspect is the selection of optimal features. For any prediction or classification problem, choosing relevant attributes is very crucial in order to achieve adequate accuracy for the problem in hand. The raw data may contain irrelevant and redundant attributes. These attributes need to be eliminated. The theme of the present paper is optimal feature selection for crop yield prediction using different 4 different methods, namely, Relief, Principal Component Analysis (PCA), Grey Wolf Optimization(GWO) and Binary Cuckoo Search(BCS). In this work, a site specific dataset which has been preprocessed by Monte Carlo method and conventional average based method is taken for analysis. The above four feature selection algorithms have been employed over the preprocessed data in order to select the optimal features for prediction. The results of the methods have been compared and better method has been suggested for optimal feature selection

II RELATED WORK

Selection of relevant features can be viewed as optimization problem. In [3], a new hybrid self-adaptive cuckoo search algorithm was proposed that extends the original cuckoo search. A framework for attribute reduction using a correlation-based filtering model for attribute reduction was suggested in [4]. Jain et al., [5] presents a novel approach which selects hybrid features by defining optimum threshold values for selecting important and non-redundant characteristics and reduces the dimensionality of features significantly. It is appropriate for both text and micro-array datasets. A new hybrid binary variant of bat and improved particle swarm optimization algorithm is proposed by Taw hid et al., [6] to solve problems with feature selection. Discriminatory feature selection using Firefly Algorithm (FA) for classification and regression is proposed by Zhang et al., [7].

To search the attribute space with minimal correlation between selected attributes, the cuckoo search (CS) optimization algorithm was used. Cuckoo Search (CS) algorithm [8] is being used for selecting optimized features due to its large generalization capabilities. It imitates the reproduction of the cuckoo and combining the cuckoo nest's behavior with Lévy's preference for flight. Different categories and application of cuckoo search algorithm are discussed in [9]. CS algorithm is found to outperform Differential Evolution algorithm with respect to convergence speed and efficiency of computation [10]. An extended binary version of the Cuckoo Search, namely BCS, is being used for mapping continuous solutions produced by CS to binary ones while performing feature selection. The usefulness of BCS in identifying optimized features which yields to enhanced classification accuracy is discussed in [11].

III PROPOSED WORK

In this work, addition, real agricultural data has been collected from Thanjavur District for the years from 2007 to 2016. The collected data have been preprocessed using Monte Carlo method and conventional average based method. It contains ten files with totally 7245 records. The description of the data set and sample record are given in Table 1 and Table 2 respectively.

Table 1 Data set Description

Variable	Type
Station Code	String
Station Name	String
District	String
Latitude	String
Longitude	String
Year	Number
Month	Number
Day	Number
Hour	Number
A1 - A11 (weather parameters)	Absolute pressure, minimum temperature, maximum temperature, temp dry bulb, temp wet bulb, relative humidity, Instant wind speed , Ave. wind speed, wind direction, evaporation, rainfall
A12 - A28 (Soil and water related parameters)	TDS, No2+No3, Ca, Mg, Na, K, Cl, So4, Co3, HCo3, F, Ph, EC, HAR, SAR, RAC, Na%

As in Table 1, various weather, soil and water parameters are numbered from A1 to A28.

Table 2 Sample Record

THANJAVUR			Thanjavur		Thanjavur		10°46'22"		79°08'09"		
2007	1	1	08:30	1015.80		22.00	29.50	24.20	21.00	74.00	
4.00	NE	0.80	10.00	0.00	391	15	44	16	81	7	124
34	1	114	0.5	7.8	700	175	2.67	0.00	48.95		

The datasets have been preprocessed using Monte Carlo method and conventional average based method. The preprocessed datasets obtained using Monte Carlo method and conventional average method are given in Appendix-A and Appendix-B respectively

Optimal Feature Selection

Selection of features is the process of selecting a subset of the most important features from a given set of features in a collection of data examples by discarding irrelevant, redundant and noisy features, either based on the corresponding example labels or not based on them. Selection of features is distinct from the extraction of features. Extraction of features creates a new feature space by translating the original high-dimensional feature space into a space of low dimensions using a set of selection functions. But both the approaches share the same possible benefits of the computational capacity of the model, decreasing storage space, enhancing learning accuracy etc. In feature selection main issue lies in how to explore the perfect subsets and then analyze the perfectly generated subsets. The current algorithms used as a search technique cannot explore feature selection search space effectively without getting trapped in the local optima. Evolutionary computations have currently been used as search techniques to explore the broad search space in order to arrive at a selected set of features. But most of them are going through early convergence. Cuckoo search converges earlier than many other evolutionary-based techniques.

In the extended version of Cuckoo search, binary cuckoo search of feature selection, the search space is modeled as a Boolean n-dimensional lattice in which binary values of 0 and 1 are used to denote the inclusion or exclusion of a particular feature respectively. Brief overview of cuckoo search and binary cuckoo search algorithms is presented in the remaining part of this section.

Cuckoo Search(CS)

Yang and Deb[8] suggested a new meta-heuristic approach for continuous optimization, called Cuckoo Search (CS), based on the fascinating technique of cuckoo birds for reproduction. It is a process of imitating the reproduction of the cuckoo and combining the cuckoo nest's behavior with Levy's preference for flight. The ultimate aim is to solve the problem of finding the optimum global quickly. In each iteration, the CS replaces the wrong solution with a good solution. The egg of the cuckoo bird may be seen as a new solution to replace the remaining birds. The following three ideal conditions are used to promote the development of a CS mathematical model:

- The way the cuckoos choose the nest is completely random, and each cuckoo produces only one bird's egg in the same nest.
- It should preserve the best bird's nest (solution) until the next generation.
- The probability of the host bird finding the cuckoo bird is P_a, where $P_a \in [0, 1]$, the number of nests is constant at n.

At the same time, if other birds find the cuckoo's eggs, they substitute the solution and randomly delete it by following the Lévy flight preference random walk to produce a new solution, to increase the variety of the solution. The following Equation (4.1) used to update the location of the cuckoo.

$$X_i^{(k+1)} = X_i^{(k)} + \alpha \oplus \text{Lévy}(\lambda) \qquad (1)$$

where $X_i^{(k)}$ represents the Kth solution of the ith generation, and the solution is the position of the bird's nest in the algorithm. α is the step factor, which is a constant greater than zero. $\oplus$ is an operator, which is called a dot product; Lévy (λ). The solution for each generation is updated based on the solution from the previous generation, and the length of the flight path determines the solution's change value. Mathematically search path that satisfies the heavy-tailed distribution of Levi's flight mode is given in Equation (2).

$$\text{Lévy}\ (\lambda) = \frac{\varphi \times u}{|v|^{\frac{1}{\lambda}}}\left(X_i^{(k)} - X_b^{(k)}\right) \qquad (2)$$

where $X_b^{(k)}$ represents the Kth optimal solution, u and v satisfy the normal distribution.
Every iteration produces a solution created by the equation of the position update. The new solution is used to replace the solution that was originally generated at random. The host bird's found the bird's egg after the substitution.

In this method, r, a random number in the range from 0 to 1 is generated randomly for every new solution. The value r is compared with the probability P_a that a host bird finds cuckoo bird's egg. If r is greater than $P_{a\ \text{then}}$ the value of $X_i^{(k)}$ will be changed by preferring random walk.

On the contrary, keep its value unchanged, save the solution with better fitness function value, and record it as $X_i^{(k+1)}$.

$$X_i^{(k+1)} = \begin{cases} X_i^{(k)} + r\left(X_i^j - X_i^e\right) & r > P_a \\ X_i^{(k)} & r \leq P_a \end{cases} \quad (3)$$

In the above equation, X_i^j and X_i^e are the two solutions of the ith generation, which are two arbitrary solutions, representing the positions of two randomly selected bird eggs.

Binary Cuckoo Search (BCS)

This section explains the optimal feature selection based on binary cuckoo search. As mentioned earlier, in conventional cuckoo search, the solutions are modified into continuously valued positions in the search space whereas in the binary cuckoo search for feature selection, the search space is modeled as a Boolean n-dimensional lattice in which the solutions are modified across a hypercube corner. Therefore, because the problem is to pick a given feature or not, a binary vector solution is used, where 1 corresponds to indicate that a feature is selected to compose the new set of data and 0 indicates that a feature is not selected to compose the new set. The binary vectors limit the new solutions to binary values using the equations the following equations

$$S\left(x_i^j(t)\right) = \frac{1}{1+e^{-x_i^j(t)}} \quad (4)$$

$$x_i^j(t+1) = \begin{cases} 1 \text{ if } S\left(x_i^j(t)\right) > \sigma \\ \quad 0 \text{ otherwise} \end{cases} \quad (5)$$

In which $\sigma \sim U(0,1)$ and $x_i^j(t)$ denotes the value of the new eggs at time step t.

The above method is employed to the preprocessed dataset and the attributes selected by the above algorithm along with correlation coefficient obtained for multiple linear regressions are shown in Table 3

Table 3 Selected features using BCS method and correlation co-efficient of Multiple Linear Regression

Data	Selected Attributes	Correlation co-efficient obtained using multiple linear regression
Preprocessed using Monte Carlo based method	Relative humidity, temp wet bulb, Max temperature, temp dry bulb, Min temperature, HAR, Na%, Ca, So_4, Evaporation, Wind direction, RAC, Mg, EC, K, HCo_3, Na, Cl, Instant wind speed	0.681

Preprocessed using average based method	TDS, EC, Cl, Na, temp dry bulb, Relative humidity, HCo_3, Max temperature, Min temperature, temp wet bulb, Mg, Ph, So_4, Rainfall, Ca, Absolute pressure, RAC, Evaporation, Wind direction, Instant wind speed, Na%	0.316

Relief

Relief is an instance-based feature selection method which evaluates a feature by how well its value distinguishes samples that are from different groups but are similar to each other. For each feature X, Relief-F selects a random sample and k of its nearest neighbors from the same class and each of different classes. Then X is scored as the sum of weighted differences in different classes and the same class. If X is differentially expressed, it will show greater differences for samples from different classes, thus it will receive higher score (or vice versa). Relief smoothens the influence of noise in the data by averaging the contribution of K nearest neighbors from the same and opposite class of each sampled instance instead of the single nearest neighbor. Multi-class data sets are handled by finding nearest neighbors from each class that is different from the current sampled instance and weighting their contributions by the prior probability of each class.

The above method is employed to the preprocessed dataset and the attributes selected by the above algorithm along with correlation coefficient obtained for multiple linear regression are shown in Table 4

Table 4 Selected features using Relief method and correlation co-efficient of Multiple Linear Regression

Data	Selected Attributes	Correlation coefficient obtained for Multiple linear regression
Preprocessed using Monte Carlo based method	Ave. wind speed, Min temperature, Na%, Wind direction	0.410
Preprocessed using average based method	Ave. wind speed, Instant wind speed, Wind direction, Relative humidity, Evaporation, Max temperature, temp dry bulb, temp wet bulb, Mg	0.306

Principal Component Analysis (PCA)

The PCA is used to reduce the dimensionality of the feature space of the samples. This concept is achieved by transforming data into a new set of variables by calculating the eigenvectors and Eigen values of the covariance matrix. In the PCA, the principal components are computed for an input matrix X of size m ×n, containing samples of features to find the Eigen values and eigenvectors of the correlation matrix, which is given by

$$\Sigma = Y^{T}Y \qquad (6)$$

$$Y = X - \eta_X \qquad (7)$$

Where η_X is the mean value of the features. Hence, the principal components matrix is given by:

$$R = US \qquad (8)$$

WhereSis a diagonal matrix of the singular values and Uisan×nmatrix, also, the principal component $R_j$is given by:

$$R_j = s_j U_j \qquad (9)$$

This component represents the scaled left-singular vectorusing the standard deviation of the data points in the consistent direction, where the data variance is given by:

$$\sigma = \sum_j \lambda_j \qquad (10)$$

λ_j is the eigen value. The output of the PCA is the principal component, where Rcontains the R_j principal components of the input samples.

The above method is employed to the preprocessed dataset and the attributes selected by the above algorithm along with correlation coefficient obtained for multiple linear regression are shown in Table 5

Table 5 Selected Features Using PCA Method and Correlation Co-Efficient of Multiple Linear Regression

Data	Selected Attributes	Correlation co-efficient obtained for multiple linear Prediction
Preprocessed using Monte Carlo based method	0.248a16+0.248a25+0.248a12+0.248a21+0.248a9=E 0.426a2+0.383a3-0.362a9=SSE+0.334a9=NW+0.288a10 -0.466a9=SW+0.381a10-0.357a1+0.343a9=SSE-0.314a8 -0.508a4+0.473a6+0.413a9=NE-0.345a1-0.292a11	0.1184

	0.633a9=SSW+0.389a8+0.329a10-0.282a9=NW-0.259a9=SW 0.589a9=NE+0.374a28-0.342a6-0.294a9=SW+0.259a3 0.624a28+0.412a2+0.349a8-0.274a9=NE-0.236a5	
Preprocessed using average based method	0.267a1+0.26 a16+0.25 a11+0.246a12+0.245a24 0.385a15+0.338a21+0.301a17-0.282a23+0.239a12 0.338a2+0.324a3+0.292a9+0.248a7+0.235a26 0.415a8-0.369a6+0.364a3+0.33 a20+0.317a4 -0.436a20-0.335a22-0.332a10+0.303a2+0.281a8 0.391a27-0.39a10+0.334a13-0.327a8+0.291a23	0.2674

Grey Wolf Optimization (GWO)

GWO is one of the recently proposed swarm intelligence based algorithms, which is developed by Mirjaliliet al. [12] in 2014. The GWO algorithm is inspired by grey wolves in nature that searching for the optimal way for hunting preys. GWO algorithm applies the same mechanism in nature, where it follows the pack hierarchy for organizing the different roles in the wolf's pack. In GWO, pack's members are divided into four groups based on the type of the wolf's role that help in advancing the hunting process. The four groups are alpha, beta, delta and omega, where the alpha represents the best solution found for hunting so far.

The GWO search process like other SI (Swarm Intelligence)-based algorithms starts with creating random population of grey wolves. After that, the four groups of wolves and their positions are formed and the distances to the target prey are measured. Each wolf represents a candidate solution and is updated through the searching process. Furthermore, GWO applies powerful operations controlled by two parameters to maintain the exploration and exploitation to avoid the local optima stagnation.

In the GWO, α, β and δguides the hunting process and ωwolves follows them. The encircling behavior of GWO can be calculated as follows.

$$\vec{\mathbf{X}}(\mathbf{t}+\mathbf{1}) = \vec{\mathbf{X}}_{\mathbf{p}}(\mathbf{t}) + \vec{\mathbf{A}}.\vec{\mathbf{D}} \qquad \textbf{(11)}$$

$$\vec{\mathbf{D}} = \left|\vec{\mathbf{C}}.\vec{\mathbf{X}}_{\mathbf{p}}(\mathbf{t}) - \vec{\mathbf{X}}(\mathbf{t})\right| \qquad \textbf{(12)}$$

where t indicates the current iteration, $\vec{A}$ and $\vec{C}$ are coefficient vectors, $\vec{X}_p$ is the position vector of the prey, and $\vec{X}$ indicates the position vector of a grey wolf.

The vectors $\vec{A}$ and $\vec{C}$ are calculated as follows:

$$\vec{\mathbf{A}} = \mathbf{2}\vec{\mathbf{a}}.\vec{\mathbf{r_1}} - \vec{\mathbf{a}} \qquad (13)$$

$$\vec{\mathbf{C}} = \mathbf{2}.\vec{\mathbf{r_2}} \qquad (14)$$

where $\vec{r_1}$, $\vec{r_2}$ are vectors randomly in [0, 1]. $\vec{a}$ a set vector linearly decreases from 2 to 0 over iterations.

In the hunting process of grey wolves, alpha is considered as the optimal applicant for the solution, beta and delta expected to be knowledgeable about the prey's possible position. Thus, three best solutions that have been found until a certain iteration are kept and forces others (e.g. omega) to modify their positions in the decision space consistent with the best place. The mechanism of updating positions can be calculated as follows:

$$\vec{\mathbf{X}}(\mathbf{t}+\mathbf{1}) = \frac{\vec{\mathbf{x_1}}+\vec{\mathbf{x_2}}+\vec{\mathbf{x_3}}}{\mathbf{3}} \quad (15)$$

where x_1, x_2, x_3 are defined and calculated as following:

$$\vec{x_1} = \vec{x_\alpha} - A_1.\left(\vec{D_\alpha}\right), \quad (16)$$

$$\vec{x_2} = \vec{x_\beta} - A_2.\left(\vec{D_\beta}\right), \quad (17)$$

$$\vec{\mathbf{x_3}} = \vec{\mathbf{x_\delta}} - \mathbf{A_3}.\left(\vec{\mathbf{D_\delta}}\right) \quad (18)$$

where $\vec{\mathbf{x_1}}$,$\vec{\mathbf{x_2}}$,$\vec{\mathbf{x_3}}$ are the three best wolves (solutions) in the swarm at a given iteration t.

$$\vec{D_\alpha} = \left|\vec{C_1}.\vec{X_\alpha} - \vec{X}\right|, \quad (19)$$

$$\vec{D_\beta} = \left|\vec{C_2}.\vec{X_\beta} - \vec{X}\right|, \quad (20)$$

$$\vec{D_\delta} = \left|\vec{C_3}.\vec{X_\delta} - \vec{X}\right|, \quad (21)$$

In GWO, one of the main components to tune exploration and exploitation is the vector $\vec{a}$. In the main paper of this algorithm, it is suggested to decrease the vector for each of dimension, linearly proportional to the number of iterations from 2 to 0. The equation to update it is as follows:

$$\vec{a} = 2 - t.\frac{2}{\max_i ter} \quad (22)$$

Where t is the iteration number, ter is the optimization total iterations number.

The above method is employed to the preprocessed dataset and the attributes selected by the above algorithm along with correlation coefficient obtained for multiple linear regressions are shown in Table 6

Table 6 Selected Features Using GWO Method and Correlation Co-Efficient of Multiple Linear Regression

Data	Selected Attributes	Correlation co-efficient obtained using multiple linear regression

Preprocessed using Monte Carlo based method	Relative humidity, temp wet bulb, Max temperature, temp dry bulb, Min temperature, HAR, Na%, Ca, So4, Evaporation, Wind direction, RAC, Mg, EC, K, HCo3, Na, Cl, Instant wind speed	0.681
Preprocessed using average based method	TDS, EC, Cl, Na, temp dry bulb, relative humidity, F, Max temperature, Min temperature, temp wet bulb, Mg, Ph, So4, Rainfall, Ca, Absolute pressure, RAC, Evaporation, Wind direction, Instant wind speed, Na%	0.316

IV RESULTS

The correlation co-efficient of multiple linear regression obtained with different feature selection techniques are inter-compared through Table 7 and Fig. 1

Table 7 Correlation Coefficient of Multiple Linear Regressions with Different Feature Selection Methods

Feature Selection method	Correlation co-efficient of regression	
	Monte Carlo Simulation based method	Average based method
Relief	0.410	0.306
PCA	0.118	0.267
GWO	0.681	0.316
Binary Cuckoo Search	0.940	0.124

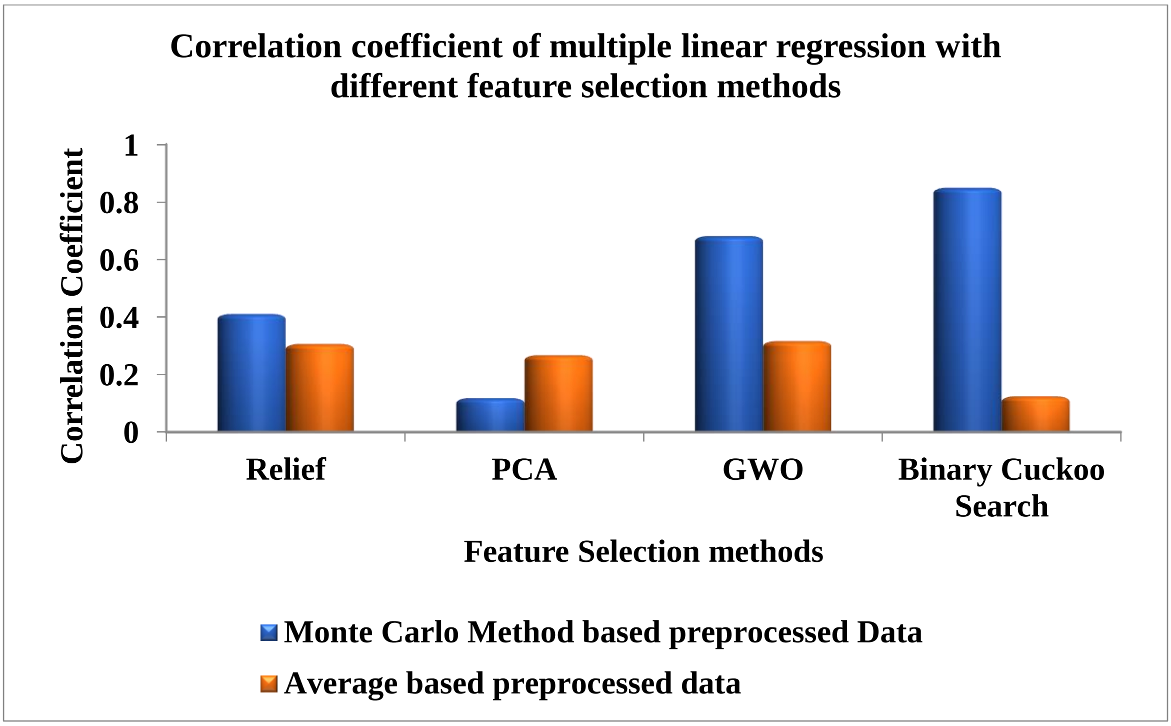

Fig. 1 Correlation Co-Efficient of Multiple Linear Regressions with Different Feature Selection Methods

From Table 7 and Fig. 1, it is found that the accuracy of multiple linear regression is the highest with features selected using BCS method.

V CONCLUSION

In this paper, need for selecting best features for a regression problem is highlighted. Experiments have been conducted to find optimal features from the data which is preprocessed using Monte Carlo Simulation and average based methods using four different feature selection methods namely Relief method, Principle Component Analysis (PCA), Grey Wolf Optimization (GWO) and Binary Cuckoo Search(BCS). The correlation coefficient of multiple linear regression has been computed for the data consisting of selected features with the above methods of feature selection. From experimentation, BCS is found to outperform the other methods.

REFERENCES

[1] Yogesh Gandge, Sandhya, "A study on various data mining techniques for crop yield prediction", Electrical Electronics Communication Computer and Optimization Techniques (ICEECCOT) 2017 International Conference on, pp. 420-423, 2017
[2] Jaroslav Menčík, “Monte Carlo Simulation Method”, in the book “Concise Reliability for Engineers” authored and edited by Jaroslav Mencik, ISBN 978-953-51-2278-4, E-ISBN - 978-953-51-6653-5,2016
[3] U. Mlakar, I. Fister, I. Fister,"Hybrid self-adaptive cuckoo search for global optimization, "Swarm and Evolutionary Computation, Vol. 29,pp. 47-72, 2016

[4] W. Yamany, N. El-Bendary, A. E. Hassanien, E. Emary, "Multi-Objective Cuckoo Search Optimization for Dimensionality Reduction", Procedia Computer Science, Vol. 96,pp. 207-215, 2016
[5] D. Jain, V. Singh, "An Efficient Hybrid Feature Selection model for Dimensionality Reduction", Proscenia Computer Science, Vol. 132,pp. 333-341, 2018
[6] M.A. Tawhid, K. B. Dsouza, "Hybrid Binary Bat Enhanced Particle Swarm Optimization Algorithm for solving feature selection problems", Applied Computing and Informatics, 2018
[7] L. Zhang, K. Mistry, C. P. Lim, S. C. Neoh, "Feature selection using firefly optimization for classification and regression models", Decision Support Systems, Vol. 106, pp. 64-85, 2018
[8] X.S. Yang, S. Deb, "Cuckoo search via Lévy flights", World Congress on Nature & Biologically Inspired Computing (2009), pp. 210-214
[9] Venkata Vijaya Geeta. Pentapalli1 , Ravi Kiran Varma P2 ,"Cuckoo Search Optimization and its Applications: A Review", International Journal of Advanced Research in Computer and Communication Engineering ISO 3297:2007 Certified Vol. 5, Issue 11, November 2016
[10] M.I. Solihin, M.F. Zanil, Performance comparison of Cuckoo search and differential evolution algorithm for constrained optimization, in: International Engineering Research and Innovation Symposium (IRIS), vol. 160(1), 2016, pp. 1–7.
[11] S. Salesi and G. Cosma, "A novel extended binary cuckoo search algorithm for feature selection," 2017 2nd International Conference on Knowledge Engineering and Applications (ICKEA), London, 2017, pp. 6-12, doi: 10.1109/ICKEA.2017.8169893
[12] S. Mirjalili, S.M. Mirjalili, A. Lewis, "Grey wolf optimizer", Adv Eng Software, vol.69 pp.46–61, 2014

Appendix – A

Preprocessed data obtained using Monte Carlo Method

Absolute Pressure	Minimum Temperature	Maximum temperature	Temp dry bulb	Temp wet	Relative humidity	Instant wind speed
Ave. wind speed	Wind direction	Evaporation	Rainfall	TDS	No_2+No_3	Ca
Mg	Na	K	Cl	So_4	Co_3	HCo_3
F	Ph	EC	HAR	SAR	RAC	NA %
Productivity						

1005.4428	27.5764	37.8571	32.2818	25.474	63.5	4.1714	24.0268	N	2.9333	6.9111
	391.0	14.7	20.0	25.51	74.0	7.0	71.0	41.0	1.64	138.3

	0.58	8.1	690.0	175.0	2.67	1.35	46.36	541

1009.4000	28.5	35.0	28.0714	25.1199	72.9777	0.0	4.1877	SW	1.7666	0.0
	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
	0.0	0.0	0.0	0.0	0.0	0.0	6.1666			

1005.85	23.763	38.25	33.028	27.72	65.6666	73.5	3.8513	NW	3.1666	13.042
	715.0	2.0	18.0	16.0	20.0	25.0	269.0	7.0	18.0	64.0
	0.5	8.4	230.0	0.87	0.0	59.1	6.6	645		

1008.2461	26.8714	36.625	2.466	22.400	77.166	3.0	4.92	ENE	3.275	13.143
	618.0	2.0	24.0	55.89	19.0	19.0	152.0	7.0	0.930	0.930
	396.5	0.45	8.0	240.0	0.8269	0.203	39.455	4.96	540	

0.0	26.2400	28.85	26.0285	25.5	79.5	0.0	4.7166	SW	1.5	0.0
	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
	0.0	0.0	0.0	0.0	0.0	0.0	0.0	551		

0.0	26.1599	37.0	29.4666	25.9000	68.5	0.0	4.225	S	1.7	0.0
	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
	0.0	0.0	0.0	0.0	0.0	0.0	0.0	782		

0.0	27.2	35.0	26.6874	27.2000	66.5	0.0	6.2428	SSW	1.6749	0.0
	239.0	1.0	22.0	23.0853	2.0	35.0	5.0	0.7368	124.232	0.0

	7.8	0.0	0.2842	0.0	34.3634	0.0	705	

0.0	26.25	34.0	26.8800	26.3	66.5	2.0	0.0	SSW	2.5	0.0
	253.0	3.0	0.0	13.3653	1.0	67.0	13.0	1.7603	118.16	0.0
	8.2	105.0	2.2084	78	0.0	35.0926	0.0	1100		

0.0	25.56	31.0	27.6	24.0	74.8333	0.0	1.0	SW	1.3666	0.0
	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
	0.0	0.0	0.0	0.0	0.0	0.0	1.5	1350		

0.0	31.6400	38.25	31.45	22.4	71.6666	5.2071	ENE	3.1624	5.4222	85.0
	0.05	18.0	58.32	53.0	16.0	67.0	14.0	4.7259	3.6742	0.07
	8.4	285.0	1.4947	0.0	35.4085	5.3428	12.6			

Appendix – B

Preprocessed data obtained using conventional average based method

Absolute Pressure	Min Temperature	Max temperature	Temp dry bulb	Temp wet	Relative humidity	Instant windspeed	
Ave. wind speed	Wind direction	Evaporation	Rainfall	TDS	No_2+No_3	Ca	
Mg	Na	K	Cl	So_4	Co_3	HCo_3	
F	Ph	EC	HAR	SAR	RAC	NA %	**Productivity**

1008.541	27.2673	4.4742	9.0562	4.7037	0.0	811.60	4.259	NE	1.987	5.464
	377.667	6.673	31.3331	9.03	83	5	102.667	34.333	3.356	153.213
	0.457	8.167	713.333	156.667	2.96	0.45	5238575			

1008.125	27.5983	4.0162	8.9452	4.6176	8.392	1.753	4.273	NE	1.963	2.076
	372	5.25	27	32.1976	7.5	5.55	97.5	35.5	0.409	196.21
	0.11	8.05	200	1.886	0	40.647	5.109	462		

1008.685	24.8833	4.0582	9.3042	4.8086	8.0141	1.747	4.371	NE	2.157	1.458
	324.667	1.667	15.3332	5.3336	8.667	106.333	20	7	135	0.533
	8.267	143.333	2.01	0	32.387	5.324	645			

1008.398	26.7993	3.6822	9.1032	4.9116	9.075	1.486	3.93	NE	1.886	1.68
	409.667	1.333	26.667	6.8555	13	112	16	2.833	232.584	0.25
	8.167	218.333	1.929	0.068	36.57	4.922	540			

0	26.9773	3.8242	8.8372	4.7377	1.329	1.763	4.144	NE	1.999	1.287
	121	0.05	18	12.15	9	0.1	21	19	2.913	61.836
	0.05	8.7	95	0.402	0	17.079	4.515	551		

0	27.9153	4.4672	9.4642	5.1216	9.093	1.652	3.197	NE	2.114	0.626
	154	0.05	14	9.72	30	1	50	9	2.913	61.836
	0.05	8.7	75	1.508	0	46.127	0.534	782		

0	28.4223	4.5722	9.8032	5.4347	0.075	2.531	5.122	NW	2.686	0.754
	206	0.525	23	22.4772	2.5	2.5	31.5	5.5	1.762	155.683
	0.05	8	150	0.799	0.081	22.304	0.539	705		

0	27.36	33.8672	9.1632	5.2987	2.536	2.529	0	NE	2.58	1.149
	245	1.525	20	17.6174	4.5	4	65.5	10.5	1.065	121.387
	0.05	7.85	122.5	1.785	0	43.335	0.486	1100		

0	28.2583	3.94	29.2362	5.2077	1.597	2.472	1.839	NE	2.283	1.579
	255	4.55	17	18.8324	2.5	65.5	7	2.823	119.552	0.05
	8.4	120	1.961	0	46.654	4.713	1350			

0	29.04	34.8352	9.8172	5.2856	8.661	2.304	3.287	NW	1.844	0.786
	325.667	3.717	23.3333	3.21	46.6677	85	9.333	2.342	169.241	0.22
	8.1	195	1.457	0	33.972	4.696	12.6			