Turkish Online Journal of Qualitative Inquiry (TOJQI)

Volume 12, Issue 3, June 2021: 4320 - 4335

Research Article

STREAMLINING NODE CENTRALITY THROUGH USING MACHINE LEARNING

D.B.Shanmugam[1]*,K.Anuradha[2], B.Sridevi[3], Dr.M.Kavitha[4], S.Pradeep[5]

Abstract

Information is typically portrayed by countless highlights. A considerable lot of these highlights might be inconsequential and repetitive for wanted information mining application. The presence of a large number of these inconsequential and repetitive highlights in a dataset adversely influences the presentation of the AI calculation and furthermore expands the computational multifaceted nature. In this way, diminishing the component of a dataset is a central assignment in information mining and AI applications. The fundamental goal of this investigation is to join the hub centrality rule and differential advancement calculation to expand the precision of highlight choice. The proposed strategy just as the exhibition dataset planning of the proposed technique was contrasted and the latest and notable element choice strategies. Various models, for example, grouping precision, number of chosen highlights, just as usage time were utilized to think about various strategies. The examination aftereffects of the various techniques were introduced in different structures and tables and the outcomes were totally broke down. From the factual perspective and utilizing distinctive measurable tests like Friedman various techniques were contrasted and one another. The outcomes indicated that the chose developmental differential calculation for bunching, rather than discovering all the components of the group communities present in the informational collection, discovered just a predetermined number of DCT coefficients of these focuses and afterward by utilizing similar restricted coefficients, bunch focuses reproduced.

***Keyword**:* Machine Learning, Node Centrality Criterion, High Dimensional Data

[1]Assistant Professor, Department of Computer Applications, SRM IST, Ramapuram Campus, Chennai.

[2]PhD Research Scholar,Dr.MGR Educational & Research Institute, Maduravoyal, Chennai

[3]Assistant Professor, Department of Electronics and communication Engineering, MAM College of Engineering and Technology, Siruganur, Trichy.

[4]Professor, Department of Electronics and communication Engineering, K.Ramakrishnan College of Technology, Samayapuram, Trichy.

[5]Assistant Professor, Department of Electronics and communication Engineering, M.Kumarasamy College of Engineering, Karur, India

Corresponding author Email: shanmugd@srmist.edu.in

1. Introduction

Information is typically portrayed by an enormous number of highlights. A significant number of these highlights might be random and repetitive for wanted information mining application. The presence of a large number of these irrelevant and excess highlights in a dataset contrarily influences the exhibition of the AI calculation and furthermore expands the computational unpredictability. In this way, lessening the component of a dataset is a principal task in information mining and AI applications (Leo et al., 2017).

The high element of information in the AI calculation additionally makes significant issues. High-dimensional pictures in AI incorporate highlights for disconnected and excess discovering that lessen the presentation of the learning calculation and consider preparing these information that requires a lot of time and computational assets dependent on choosing the proper component to improve AI execution regarding diminishing time for the learning model and expanding the precision in the learning interaction. Since choosing the fitting component for AI with high dimensional information is important to show the legitimate presentation of AI by choosing the suitable element (Hamidi et al., 2016). Highlight choice has a significant job in AI and example acknowledgment. A large number of the chose highlights are chosen with a learning calculation to create total characterization properties. Be that as it may, a considerable lot of the highlights chose for the learning task are disconnected and lessen the effectiveness of the learning calculation and face the calculation wellness with trouble. The exactness of learning and the learning speed might be essentially more regrettable with repetitive highlights. Consequently, the choice of related and fundamental highlights at the pre-handling stage is significant and significant (Marsan et al., 2016).

High components of information are utilized in various territories of AI, man-made reasoning and information mining. For instance, its applications can be in the ordering records, which brings about the recovery of suitable data, however high dimensionality of the component spaces for choosing highlight is led by include dimensionality decrease approach. Not all highlights are significant for text extraction in archive classifying, and a few highlights should be chosen by the ideal rules. Diminishing the element of highlight and season of text order, upgrading the exactness of text arrangement, and how to look for the sub-set space of potential highlights in classification of archives is of specific significant. Highlight determination expands the learning calculation and improves the precision of the information results. Choosing a component rearranges the model and diminishes computational expense on the grounds that a model with low sources of info has less assessment for new highlights (Oldham et al., 2018).

Additionally order messages in AI because of the high volume of words mess up extricating information and fitting highlights. This issue can be settled by word displaying and determination of suitable component and improved word characterization proficiency in the content. This demonstrating empowers the highlights to be chosen fittingly and straightforward arrangements can create great outcomes. Since by text arrangement can choose the high dimensional highlights, which it lessens measurements in content classification (Masoudian et al., 2015).

In the dataset, time intricacy is one of the basic issues in information mining. The time intricacy of high-dimensional information is made when the information is put away in a data set to remove highlights. Since the component is utilized to quantify the recognizable cycles. To choose the proper highlights of the information, can group its AI calculations. By choosing an element in this information, the calculation gear just as dimensional impact of information and its productivity increment. Subsequently, the time unpredictability of this sort of information can be compelling with AI calculations to zero in on the highlights chose from the dataset (Esfandiarpour, 2015). By choosing a component, a subset of highlights is chosen dependent on the advancement rules. Since the exhibition of conventional measurable techniques has diminished, this lessening in information measurement is because of the expansion in the quantity of perceptions and the quantity of highlights identified with the perception. Since the component determination is utilized to improve execution forecast, better comprehension of datasets related information to AI, etc. Likewise by eliminating inconsequential and repetitive highlights can diminish the excess organization execution and increment the exactness and lessen the multifaceted nature of the information. Time arrangement examination in high dimensional information is significant in light of the fact that it utilizes the grouping strategy to choose the component in various time arrangement. This presentation depends on information similitude models at comparative or close to time stretches, which streamlines information approaches and assumes a significant part in information mining and information disclosure in dataset. Because of the extension of search space at high measurements, the capacity of developmental calculations to manage complex and multi-target issues, streamlining and absence of underlying data in dataset grouping prompting absence of need to displaying and their comprehend in information bunching for complex datasets. The effectiveness and execution of the differential developmental enhancement calculation can be utilized as a quick and productive quest strategy for grouping information to choose the fitting element. Likewise, because of the high volume of information in these organizations, the intricacy in the hubs additionally happens and removes the primary information in organization; yet eliminating every hub doesn't influence the whole organization. In these organizations, the significance of hubs relies upon the centrality of the hub. Truth be told, dataset is the fundamental standards of a marker for unsurprising conduct that the centrality rules are rely upon the chart structure. Appropriately, mix of differential calculation and hub centrality rule can be utilized to build the precision of highlight choice in high-dimensional organizations to improve their presentation.

In this exploration, an endeavor is made to introduce an element determination technique dependent on element grouping and transformative differential calculation. In the proposed technique after element grouping by utilizing diagram bunching and local area recognition calculations, each group is distinguished by developmental differential calculation dependent on centrality standard of proper highlights. Highlight determination dependent on the connections between highlights is an old thought for diminishing elements of issue that has for some time been thought of. Highlights bunching can direct essential investigation on information highlights and eliminate numerous redundancies in the essential highlights. Thus, the chose highlights will have the most noteworthy relationship with the objective class and the least repetition. The benefit of this technique is that it first outcomes in diminished computational unpredictability. Since it is costs under huge number of highlights while characterizing on a little subset of highlights. Furthermore, because of the decrease in measurement, the execution time is diminished.

2. Machine learning

In AI, imaging strategies contain exact data that interaction challenges a lot of information, and grows new ideas and advances in far off detecting just as AI techniques in equal. By completely understanding the fundamental calculations and strategies in AI from profound association with framework improvement takes care of the issue, which by taking care of this issue can ideally remove the data and order AI calculation [15]. In high-dimensional pictures, numerous highlights of inconsequential repetitive and loud learning task lessen and the preparing of this information requires computational time and assets. Thusly, highlight choice assumes a significant part in improving the presentation of AI calculations regarding diminishing time, making a learning model, and expanding exactness in the learning cycle (Fraiwan and Lweesy, 2017).

Profound learning is a sort of profound neural organization of AI and a bunch of calculations. This learning can demonstrate significant level ideas by learning at various levels and layers [16]. This learning helps a sort of connection builds up among various classes and datasets, which is a typical worldview highlight, so it tends to be applied to the planned framework conditions subsequent to characterizing worldview and the model to be prepared model on the grounds that the actual framework finds worldview (Wu and Prasad, 2018).

As one of the expansive and broadly utilized parts of computerized reasoning, AI is change and finding strategies and calculations that empower PCs and frameworks to learn. The target of AI is that the PC (in its most broad idea) can progressively and with expanding information have higher effectiveness in the ideal assignment.

3. Clustering

Information bunching is a significant information mining method in information examination. It is like a solitary dataset in an information base and is important in wide scope of examination fields and their applications. In worldview recognition, information grouping is frequently utilized as solo characterization of unlabeled information. Information grouping is likewise quite possibly the most mainstream apparatuses for separating a subset of information that is a fascinating perspective on huge data sets[17]. Information bunching procedures are reasonable for information pressure and are known as quantization vectors. Grouping calculations are iterative refinement. By and large it is added to discover proper bunches from a dataset. Be that as it may, the computational expense of the calculations increments with the size and measurements of the dataset, and is a significant issue. Subsequently, to decrease the expense for grouping high-dimensional information, there are numerous strategies for bunching equal information (Bharara et al., 2018).

In order of gestures or bunching all the hubs in the organization are first partitioned into groups, and in each group, one hub is chosen as the head, and the remainder of the hubs are called typical hubs [18]. The strategy for choosing head in every technique considers various models. In group based strategies, its fundamental goal is to disseminate energy utilization across all hubs. Indeed, the techniques that are working dependent on hubs grouping they may have distinctive energy utilization. These techniques are among the best directing calculations in these organizations and are enlivened by new strategies to upgrade the life expectancy of the organization (Bendechache and Kechadi, 2018).

Grouping can be considered as the main issue in unaided learning. Grouping endeavors to split the information into various bunches so the similitude between the information inside each group is expanded and the comparability between the information inside the various bunches is limited[19].

4. Research Methodology

4.1 Clustering Nodes Based on the Nodes Centrality Profiles

The estimation of centrality estimation can rely upon bunches of hubs and subsets with independent centrality profiles and organization geography exhibitions. Various leveled grouping utilized for an organization with centrality estimation esteems across all hubs initially by utilizing standardization sigmoid capacity.

$$f(x) = \left[1 + \exp\left(\frac{-(x-\mu)}{\sigma}\right)\right]^{-1}$$

Which μ is the mean and σ is the standard deviation of the values for measuring data centrality across the node of a network and then it is calculated linearly in unit distance. Hierarchical clustering is performed by using the least ward variance method for the Euclidean distance between pairs of centrality (normalization) criteria. Assuming that the Davies-Bouldin index is used to determine the specific resolution of thedendrogram slice and it is investigate the resulting clusters [20]. Davies-Bouldin index is a different inter-cluster similarity ratio for a clustering solution. Low values of Davies-Bouldin represent an effective clustering solution.

Additional Methods Definitions of Centrality

Each network as (A) adjacent matrix $\mathrm{N} \times \mathrm{N}$ indicates that $A_{ij} = 1$. If the nodes of i and j are connected $A_{ij} = 0$. Adjacent matrix indicates the W weighted network, which encodes weight element of W_{ij} from the edge between the nodes of i and j. These definitions of weight networks were simply replaced for A_{ij} by W_{ij}.

- **Degree/Strength (DC)**

The simplest measurement of centrality is called the degree of centrality, which is defined as the number of edges connected to a node:

$$DC_i = k = \sum_{j \neq i} A_{ij}$$

For weighting networks, similar degrees of weighting degree are used; otherwise it is used as s node strength, which is the sum of all edge weights connected to the node.

$$DC_i = s = \sum_{j \neq i} W_{ij}$$

- **Harmonic Index Centrality (HC)**

While it is commonly used to measure the productivity and impact of scientists' work, the index of h has also recently been used as a centralized matrix for analyzing complex networks. If so, a set of i node neighbors that having value equal to or greater than h, index of h from i node can be defined as follows:

$$HC_i = \max_{1 \leq h \leq : k_i} \min\left(\left|N_{\geq h}(i)\right|, h\right)$$

Where h is the value between 1 and the degree of the *i* node. Therefore, the h index of a node defines the maximum value of h for the h neighbors of the *i* node at least degree of h.

- **Leverage Centrality (LC)**

It is the another measure of centrality, which is considered in the leverage centrality relationship of a neighbor node, in addition to the measurement of the other centrality, is impossible; leverage centrality can allocate negative values to a node. The node has less connection than its neighbors. A node is affected by its neighbors. A node of positive values can have more connections than its neighbors, that is, it affects the influence of its neighbors. Leverage centrality is defined as follows:

$$LC_i = \frac{1}{k_i} \sum_{j \in N(i)} \frac{k_i - k_j}{k_i - k_j{}'}$$

Which N (i) is the set of neighbors of i node. In the weighting networks strength used instead of degree.

- **Eigenvector Centrality (EC)**

Eigenvector centrality has high value among the nodes and determineswith the neighbors that are has high degree. This measurement as Eigenvector, v, with the greatest value of Eigenvector λ_1 is related to the adjacent matrix and can be expressed as follows

$$EC_i = v_i = \frac{1}{\lambda_1} \sum_j A_{ji} v_j$$

- **Katz Centrality (KC)**

In a network connected with a large and flexible module, the exclusive main center of high score for the nodes within the module and a low score (if not zero) is allocates for the nodes outside the module. Therefore, it is not appropriate to measure node detection outside the module. Two parameters Katz centrality α and β are added as Eigenvector centrality definition. α Parameteris the contribution of distant dependencies(i.e., neighboring nodes of neighbors) intervenes to the node centrality stage. The parameter allocates a certain amount of centrality to each node, so confidence of each node as thecentrality value is non-zero. As Katz centrality allocates each node a small amount of centrality, it also ensures that highly connected nodes in other clusters have high centrality score. Katz centrality can be expressed as follows:

$$KC_i = \alpha \sum_j A_{ji} v_j + \beta$$

Or it is defined in the matrix form as follows:

$$KC = \vec{\beta}(I - \alpha A)^{-1}$$

Which $\vec{\beta}$ is a vector of N size with each element of β and I is the identity matrix of A. In all analyzes, α is the set is less than 10% inverse of the largest eigenvalue (usually a value close to the largest eigenvalue usable) from the network and β is the set of 1.

- **PageRank Centrality (PR)**

With the Eigenvector and Katz centrality, low-degree nodes may gain a high score, only because with their low-degree are connected to more nodes. Correct PageRank centrality for this behavior with the scale of neighbors nod contribution of i and j to node centrality of i by i degree is expressed as follows:

$$PR_i = \alpha \sum_j A_{ji} \frac{v_j}{k_j} + \beta$$

Matrix of this definition can be written as follows:

$$PR = \vec{\beta}(I - \alpha D^{-1} A)^{-1}$$

That the D is an oblique matrix and D_{ii} is the degree of i node (in the weighting network of S instead of oblique S_{ii} the power of node is i). Parameters of α and β are the same functions in Katz centrality. Or all of analyzes β is the set of 1 to 0.85 of set.

- **Closeness Centrality (CC)**

Closeness centrality defines a node as the center, if it has at least the meanof minimum path length to any other node in the network. It is assumed that nodes with amean short path length to other nodes can transmit or receive information in a relatively short period of time. Since the mean of a shorter path is introduced as a reversed central node, most central nodes have higher values and define as below:

$$CC_i = \frac{N}{\sum_j l'_{ij}}$$

This l_{ij} is shortest topology distance between the nodes of i and j. Weighting networks of computed l_{ij} is as the shortest weighting path (the path with the shortest set of weighting edges) is determined by using weighting inverse matrix.

- **Information Centrality (IC)**

This criterion is also known as Closeness centrality of the main stream can investigate all possible paths between two nodes as well as overlap in these paths and their weight in each information value that includes the path. Information in the path identifies as the inverse topology of that path.

To estimate the Information centrality, firstly $C=(L+J)^{-1}$ matrix is identifies, which Llaplacianof A and J is the $N \times N$ matrix by computing all elements to an element. Information centrality is defines as below:

$$IC_i = \left(C_{ii} + \frac{\sum_j C_{jj} - 2\sum_j C_{ij}}{N} \right)^{-1}$$

In the weighting networks L is the laplacianof W.

- **RandomWalk Closeness Centrality (RWCC)**

Random Walk Closeness centrality computes the mean time to a random step is taken at the start of each node in the network to reach the i node and inverse of the first average time mean to a particular node. The first average time mean can be computed from the main Z matrix.

$$Z = (I - P + \Pi)^{-1}$$

That the (I) is the identity matrix, the transfer matrix $P = D^{-1}A$ (or $P = S^{-1}W$ in the weighting networks), and Π is a $N \times N$ matrix that each column, the π vector of stable distributed probable state is from the transfer matrix (like $\Pi_{ij} = \pi_j$). π Vector can be obtained by solving the system of linear equations $\pi P = \pi$ and $\sum_i^N \pi_i = 1$. The matrix of H first average time mean can be defined as follows:

$$H_{ij} = \frac{Z_{jj} - Z_{ij}}{\pi_j}, i \neq j$$

Element of H_{ij} is the mean matrix of first average time from the i node to j node. Random Walk Closeness centrality simply computed as follows:

$$RWCC_i = \frac{N}{\sum_j H_{ji}}$$

- **Sub Graph Centrality** (SC)

Like other criteria, Sub graph centrality counts a number of steps, but instead of counting steps with other nodes, considers the method of steps closeness. Thus, the criterion of Sub graph centrality consists of several sub graphs, which are determined by the close steps, to which the node belongs, to the smaller sub graphs. Longer steps (and thus larger sub graphs) computed by the weight of each step with $\frac{1}{n!}$ coefficient that the n is the length of step. Therefore Sub graph centrality can be computed as follows:

$$SC_i = \sum_{n=0}^{\infty} \frac{\left[A^n\right]ii}{n!} = \left[e^A\right]_{ii}$$

In the weighted networks reduction of adjacent matrix $S^{-1/2}wS^{-1/2}$ is used instead of A.

Data Analysis

Clustering

Clustering is one of the unsupervised learning branches and is an automated process in which the samples are divided to the groups that its members are similar to each other, referred to as cluster. Cluster is therefore a set of objects in which the objects are similar to each other and dissimilar to the objects in the other clusters [21]. For similarity, a different criterion can be considered, for example, the distance criterion can be used for clustering, and objects that are closer to each other can be considered as a cluster that this type of clustering is called distance- based clustering.

4.3 Centrality

Obviously, the importance of features is not the same among available data. Some of them are more important because of the different positions in the type of subject. This importance allows for greater access to information or a greater role in its transmission for individuals. That is why we consider certain data to be influential and determinative in the dataset. The importance and popularity of these data in different contexts are determined by different criteria. In this study we use the centrality of eigenvalues:

This method computes the importance of nodes based on adjacent nodes. The computation happens on graphs with strong connectivity. If a node is connected to the nodes that are of high importance, its importance is increased by their influence [22]. This method considers the same importance iteratively for the computation of node. First, all nodes are given an initial score. This concession goes on as long as the chain that achieves stability continues. The scoring in this method is based on this concept that high-connectivity nodes help the nodes that follow them in terms of score.

This part of the simulation is used as an innovation by combining the differential evolutionary algorithm and the node centrality criterion to reduce thedimensions of problem and feature selection. The laplacian centrality criterion algorithm is used to cluster features. And some kind of social networking algorithms coupled with differential evolution algorithm for feature clustering are first investigated in this research. The advantages of using these algorithms in combination are as follows. In these algorithms the number of clusters is automatically determined and the number of clusters is determined by the user. On the other hand, in most clustering methods the features are predetermined, in which the dispersion of the features in each cluster is not taken into account in the clustering process [23]. Therefore this method will not be able to detect the optimal clusters. In the population detection algorithm used in this study, both the dispersion of features within each cluster as well as the degree of features connection across different clusters is considered. Therefore, this algorithm will be able to find optimal clusters.

4.4 Datasets Used

The UCI reference is used to evaluate the method presented in the previous sections from the datasets in the real world. This library contains datasets for evaluating machine learning algorithm.

In order to evaluate the proposed model in this simulation, in this part of the study, first the neural network is used to increase the accuracy of the data and the support vector machine is used as well.

4.5 Artificial Neural Network

This classification algorithm is presented based on non-parametric approaches and is widely used in scoring problems. ANN can be used for nonlinear problems. In the present study, the ANN has three layers, the inner layer consisting of neurons related to the input variables and the output layer having one neuron [24].

5. Support Vector Machines

Support vector machines are a supervised classification method based on the theory of statistical learning theory. The basic idea behind this classifier is to find an optimal page cloud as a decision level to maximize the margin between the two classes [25]. If the data is not linearly separated, the data is transferred to a higher dimensional space with a nonlinear kernel data and the optimal page cloud is determined in that space.

5.1Compare and Evaluation of Previous Articles

Abolghasemi and Momtazi (2018) in a research used a machine learning algorithm to select text features and improve text mining, which resulted in accessing to increased feature selection in the Persian text categorization in this regard. The method presented in this section demonstrates the ability to identify the feature of the connections classifications in the dataset with high accuracy and error reduction compared to the presented research method in this paper.

Ismaili and Abbasi (2017) in a research by using closest neighboring classification method and genetic algorithm have investigated the reduction of feature dimensions which have resulted in attaining increased accuracy in data classification. In the method considered in the Ismaili and Abbasi paper, the data of one cluster is similar to the data close to it as possible and is different from the data of the other clusters. Similarity between records is measured by the distance function. The distance function receives two input records and returns a value that shows the similarities of the two, one of the disadvantages of this method is that usually the output of a clustering algorithm consists among each cluster. To implement the clustering algorithm, it is necessary to determine in advance the number of clusters that require prior information of the existing data because without them, the prerequisite for determining the clusters is difficult. Each data sample is assigned to a cluster of data that is the least distance to the center of that cluster. At the end, the centers of each cluster are determined as the output of the algorithm, by using the genetic algorithm to determine the most optimal value, and the feature is selected based on the selection of the best value. From disadvantages of this method can mention the time of accessing to best value that is it is long and expensive in some cases, but the algorithm proposed in this study has a higher accuracy than the algorithm presented in Ismaili and Abbasi article in addition to being optimally low in time because the differential evolutionary algorithm is presented as the multi objective optimization algorithms, which havethe ability to find near-optimal solutions for math problems.

In fact, the accuracy of selecting the differential evolutionary algorithm is higher than the genetic algorithm and the error rate in the proposed algorithm is lower than that of the genetic algorithm.

Chadhari and Agarowal (2018) in a research by using the EBQPSO algorithm concluded that its effect is resulted in improved accuracy and invocation of particle and quantum optimization algorithms. In this study, EBQPSO algorithm was used in gene dataset to classify cancer. One of the main problems of the algorithm used in this study is the lack of local search algorithms. All data may not be examined and the accuracy can be reduced to some extent, resulting in an increased error rate in the simulation, but the differential evolutionary algorithm used combined with the node centrality criterion algorithm eliminates this defection and resulting in high accuracy and reduced error rate in the simulation.

Hiwa et al (2018) in a study by using support vector classification machine and genetic algorithm optimization, proposed a new method for automatic extraction of specific brain network and graph theoretical properties. The method used in this study has limitations such as the fact that, for example, it is not yet clear how parameters can be determined for a mapping function. Support vector-based machines are requiring complex and time consuming computations and due to the computational complexities consume a lot of memory. Discrete and non-numeric data are also incompatible with this method and must be converted. But the proposed algorithm in this study solves these problems and is able to select the high data feature in the less time.

6.Conclusion

After extracting the feature, the clustering is conducted on the obtained features. Given the number of clusters with the clustering algorithm, the features in each cluster are ranked. From each cluster numbers of highest scored features are selected. The features selected from all clusters are sorted by the score at which the cluster is assigned andthe obtained features are evaluated. The use of feature values in clusters or classes causes local feature behavior to be applied to different categories.The clustering process consists of two stages. In the first step, the optimal number of clusters is determined based on the validation indices. The clustering algorithm depends on several factors such as the number of clusters and the distance between the clusters. After determining the optimal number of clusters, the mean-k algorithm is used for cluster categorization. The main objective of the optimization head clusters phase is to select each cluster based on a series of defined criteria of a cluster center. The following three criteria are used to select cluster centers:

- 1. The energy sum of selected cluster centers
- 2. The distance sum of the nodes of that cluster from the center of the cluster
- 3. The distance sum of the selected node centers from each other
- 4. The node centrality sum of the selected cluster centers

As can be seen in this respect, the higher the first and the fourth criteria and the two other criteria are lower is indicate better cluster centers.

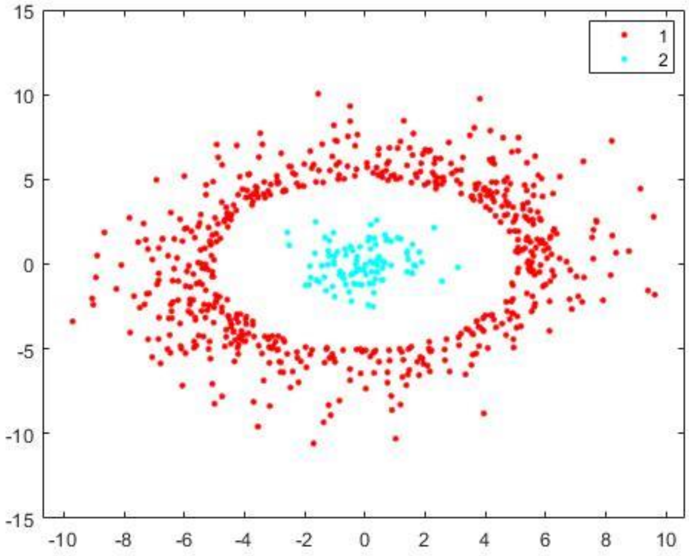
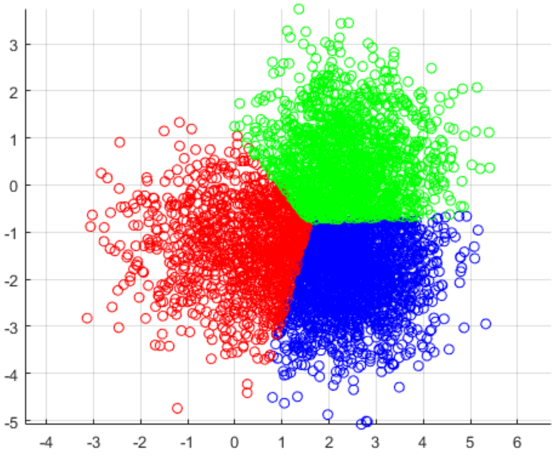


Figure1. Clustering

Feature clustering can perform a basic analysis on the features and can eliminate many redundancies in the primary features. As a result, the features selected will have the most relevance to the target class and the least redundancy.

Solve of clustering problem in general and automated clustering issue specifically can be out of common clustering algorithm power. One of the solutions that considered for this subject is that converting clustering issue to an optimization issue and its resolve can be carried out by intelligence and revolutionary optimization algorithms.



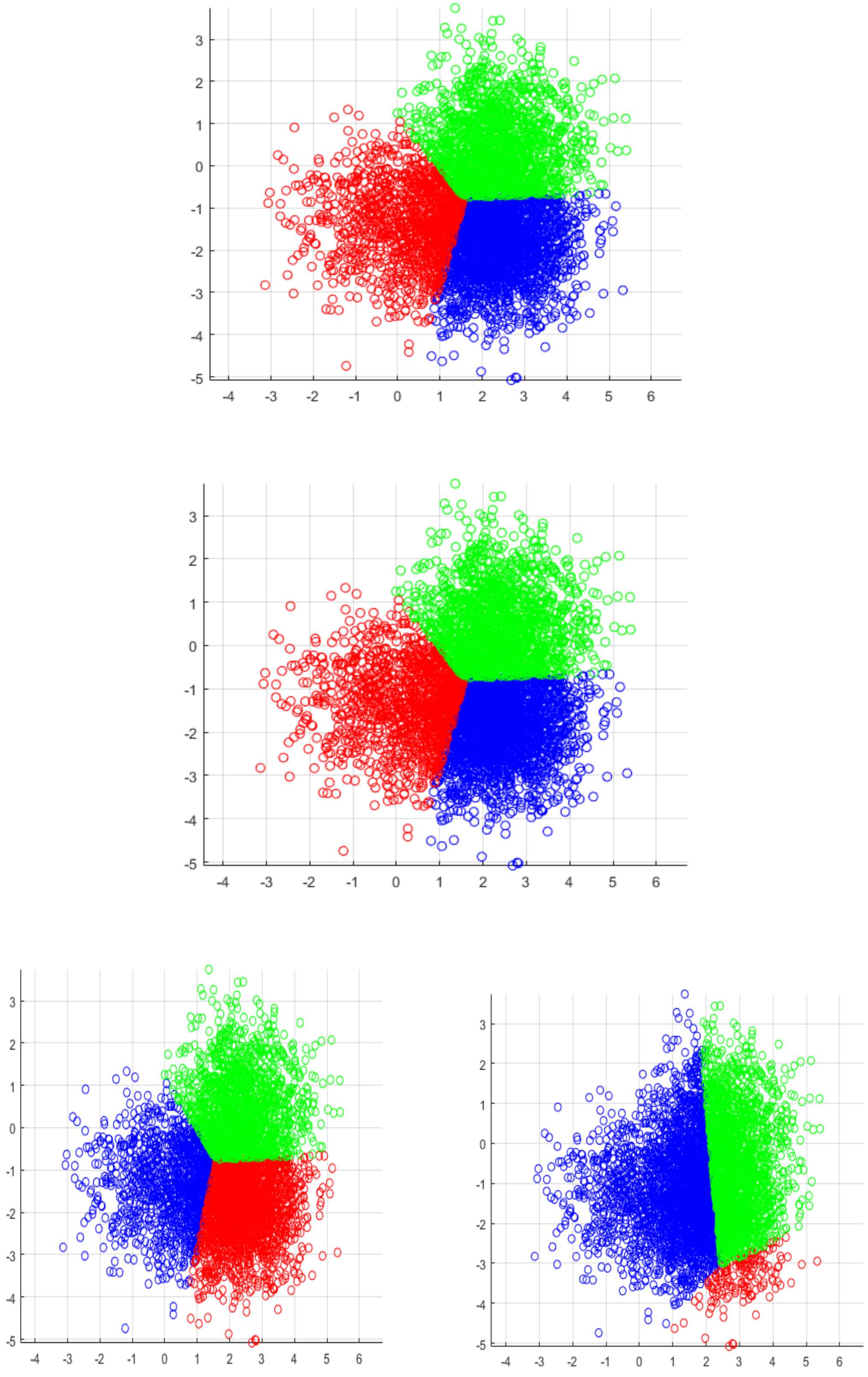


Figure2. Diagrams of differential evolutionary clustering

As a result, the proposed method is a multi-objective algorithm that needs to be optimized.

In this algorithm, each solution has the same number of dimension to the cluster centers and must be select the best cluster centers. Each particle of the hybrid algorithm has a k dimension that k denotes the number of clusters. In this algorithm the(i)dimension of each particle represents the center of the selected cluster for the nodes in that cluster. The algorithm employs a differential operator to generate new solutions that exchange information between population members. One of the benefits of this algorithm is having a memory that stores the information for the right answers in the current population. Another advantage of this algorithm is its selection operator. In this algorithm, all members of a population have an equal chance of being selected as a parent. That is, the infant's generation is compared to the parent's generation in terms of the extent of competence measured by the objective function. Then the best members move on to the next stage as the next generation.

The most important features of the DE algorithm are its high speed, simplicity and robustness. This method only starts by setting three parameters. NP parameter, population extent, parameter F mutation weight and parameter C are multiplied to the difference of two vectors and added to the third vector. The F parameter is usually set to 0 to 2, r and the C parameter is set to 0 to 1.

The differential evolution algorithm is presented to overcome the main problem of genetic algorithms, namely the lack of local search in these algorithms. The main difference between genetic algorithms and DE algorithm is in the order of mutation and coupling operators as well as in the mode of selection operator. The selected evolutionary differential algorithm for clustering, instead of finding all the elements of the cluster centers in the dataset, finds only a finite number of DCT coefficients of these centers and then reconstructs them by using the same finite coefficients of the cluster centers. After clustering features by using graph clustering and community detection algorithms, from each cluster, appropriate features is identified by evolutionary differential algorithm based on centrality criteria.

References

1. Hamidi, M., Ebadi, H., Kiani, A. 2016, A Comprehensive Review on Machine Learning and Feature Selection Methods with Emphasis on Classification in Remote Sensing Applications, 2nd National Conference on Geospatial Information Technology Engineering, KhajehNasireddinTusi University of Mapping Engineering
2. Abolghasemi, M., Momtazi, S. 2018, Text Mining Improvement by Selecting Feature Words, Fourth International Web Research Conference, Tehran.
3. Masoudian, S., Derharmi, V., Zarifzadeh, S. 2015, Investigation of Feature Selection Methods and Text Subject Classification Methods by Using Persian News Data, 46th Iranian Mathematical Conference, Yazd University.
4. Esfandiarpour, S. 2015, Feature Selection by Using Colonial Competition Algorithm, Master Thesis of Computer Science, ShahidBahonar University of Kerman.
5. Ismaili, Z., Abbasi, E. 2017, Application of Hybrid Algorithm Feature Selection Method in Predicting Short Term Performance of Corporate Initial Public Offerings in Securities Exchange of Shiraz University Central Conference.

6. Liu, Y., J.-W. Bi, and Z.-P. Fan, Multi-class sentiment classification: The experimental comparisons of feature selection and machine learning algorithms. Expert Systems with Applications, 2017. 80: p. 323-339.
7. Wu, H., & Prasad, S. (2018). Semi-Supervised Deep Learning Using Pseudo Labels for Hyperspectral Image Classification. IEEE Transactions on Image Processing, 27(3), 1259-1270.
8. Bharara, S., Sabitha, S., & Bansal, A. (2018). Application of learning analytics using clustering data Mining for Students' disposition analysis. Education and Information Technologies, 23(2), 957-984.
9. Marsan, G. A., Bellomo, N., &Gibelli, L. (2016). Stochastic evolutionary differential games toward a systems theory of behavioral social dynamics. Mathematical Models and Methods in Applied Sciences, 26(06), 1051-1093.
10. Oldham, S., Fulcher, B., Parkes, L., Arnatkeviciute, A., Suo, C., & Fornito, A. (2018). Consistency and differences between centrality metrics across distinct classes of networks. arXiv preprint arXiv:1805.02375.
11. Fraiwan, L., & Lweesy, K. (2017, March). Neonatal sleep state identification using deep learning autoencoders. In Signal Processing & its Applications (CSPA), 2017 IEEE 13th International Colloquium on (pp. 228-231). IEEE.
12. Chaudhari, P., & Agarwal, H. (2018). Improving Feature Selection Using Elite Breeding QPSO on Gene Data set for Cancer Classification. In Intelligent Engineering Informatics (pp. 209-219). Springer, Singapore.
13. Bendechache, M., &Kechadi, M. (2018). Distributed clustering algorithm for spatial data mining. arXiv preprint arXiv:1802.00304.
14. Hiwa, S., Obuchi, S., & Hiroyasu, T. (2018). Automated Extraction of Human Functional Brain Network Properties Associated with Working Memory Load through a Machine Learning-Based Feature Selection Algorithm. Computational intelligence and neuroscience.
15. A Manjunathan, A Lakshmi, S Ananthi, A Ramachandran, C Bhuvaneshwari, "Image Processing Based Classification of Energy Sources in Eatables Using Artificial Intelligence", Annals of the Romanian Society for Cell Biology,vol.25, issue.3, pp.7401-7407, 2021.
16. P Matheswaran, C Navaneethan, S Meenatchi, S Ananthi, K Janaki, A Manjunathan," Image Privacy in Social Network Using Invisible Watermarking Techniques", Annals of the Romanian Society for Cell Biology, vol.25, issue.5, pp.319-327, 2021
17. Bhuvaneshwari C, Manjunathan A, "Advanced gesture recognition system using long-term recurrent convolution network", Proc. ICONEEEA, 2019 pp. 1-8.
18. V.Kavitha, V.Palanisamy, "New Burst Assembly and Scheduling T technique for Optical Burst Switching Networks",Journal of Computer Science, Vol. 9, Issue 8, pp.1030-1040, 2013.
19. Sujatha K., Nandagopal C, Realization of gateway relocation using admission control algorithm in mobile WiMAX networks,4th International Conference on Advanced Computing, ICoAC 2012, 2012, 6416831.
20. Nandagopal C, Ramesh S.M,An Effcient data gathering technique using optimal minimum coverage spanning tree algorithm in WSN,Journal of Circuits, Systems and Computers, 2020, 29(14), 2050225
21. C Bhuvaneshwari, A Manjunathan, "Reimbursement of sensor nodes and path optimization", Materials Today: Proceedings, 2020.

22. C Bhuvaneshwari, G Saranyadevi, R Vani, A Manjunathan, "Development of High Yield Farming using IoT based UAV", IOP Conference Series: Materials Science and Engineering 1055 (1), 012007
23. R Dineshkumar, P Chinniah, S Jothimani, N Manikandan, A Manjunathan, M Dhanalakshmi, "Genomics FANET Recruiting Protocol in Crop Yield Areas UAV", Annals of the Romanian Society for Cell Biology,vol.25, issue.5, pp.1515-1522, 2021.
24. M Ramkumar, C Ganesh Babu, K Vinoth Kumar, D Hepsiba, A Manjunathan, R Sarath Kumar, "ECG Cardiac arrhythmias Classification using DWT, ICA and MLP Neural Networks", Journal of Physics: Conference Series, vol.1831, issue.1, pp. 012015.
25. K Balachander, G Suresh Kumaar, M Mathankumar, A Manjunathan, S Chinnapparaj, "Optimization in design of hybrid electric power network using HOMER", Materials Today: Proceedings, 2020.
26. Rahim, R., Murugan, S., Manikandan, R., & Kumar, A. (2021). Efficient Contourlet Transformation Technique for Despeckling of Polarimetric Synthetic Aperture Radar Image. Journal of Computational and Theoretical Nanoscience, 18(4), 1312-1320.