IJCRT.ORG ISSN : 2320-2882

INTERNATIONAL JOURNAL OF CREATIVE RESEARCH THOUGHTS (IJCRT)

An International Open Access, Peer-reviewed, Refereed Journal

FMFQR Algorithm: An Improved Hybrid Optimization Technique for Brain-Computer Interface

[1]M.JEYANTHI, [2]Dr.C.VELAYUTHAM, [3]Dr. S.JOHN PETER

[1]Research Scholar, [2]Research Supervisor & Research Joint Supervisor

[1]Department of Computer Applications & [2&3]Department of Computer Science

[1&3]St.Xavier College (Autonomous), Palayamkottai & [2]Aditanar College of Arts and Science, Tiruchendur

Abstract:

The Friedman-based fuzzy rough quick reduct algorithm(FMFQR) is a feature selection method designed to enhance the efficiency and accuracy of machine learning models. By integrating Friedman's statistical test with fuzzy rough set theory, the algorithm identifies and eliminates irrelevant or redundant features while preserving essential data characteristics. This approach combines the strengths of fuzzy rough sets in handling uncertainty with Friedman's test to prioritize features based on their significance. The quick reduct process ensures computational efficiency, making it suitable for high-dimensional datasets. In this paper, First, Statistical feature extraction method is applied for raw EEG data. Then the extracted features are normalized by Z-Score normalization approach, Further, the normalized features are selected by Friedman-based fuzzy rough quick reduct algorithm. Finally, the classification task is done using three classification algorithms such as Fuzzy-Rough Nearest Neighbor (FRNN), Fuzzy Logic-based Neural Networks (FNN), Fuzzy Decision Tree (FDT). The experimental works are discussed and analyzed. Results indicate that the proposed approach significantly improves a FRNN based classification method in terms of classification accuracy. The proposed method improves the accuracies when compared to other state of the methods.

Keywords: BCI, Feature Extraction, FMFRQR, Classification

I. Introduction:

Brain-Computer Interfaces (BCIs) are systems that establish a direct communication pathway between the human brain and external devices, bypassing traditional neuromuscular output channels. This technology is designed to interpret neural signals generated by the brain and convert them into commands that control devices such as computers, prosthetics, or wheelchairs. BCIs have evolved significantly since their inception in the 1970s and are now being developed for a wide range of applications, from assisting individuals with disabilities to enhancing cognitive abilities in healthy individuals[1].

The foundation of BCI technology lies in the acquisition and processing of brain signals, most commonly through electroencephalography (EEG) or invasive neural implants. These signals are translated into meaningful outputs using machine learning algorithms and signal processing techniques. In recent years, BCIs have garnered substantial attention in the fields of medicine, particularly in neurorehabilitation and communication for individuals with severe motor impairments such as amyotrophic lateral sclerosis (ALS) or spinal cord injuries[2]. This advancing technology holds great promise for bridging the gap between neural function and external control, paving the way for revolutionary developments in human-computer interaction.

The organization of the paper as follows: Segment II curtly express the Literature Review, Segment III express the materials and methods which include the BCI data set, Feature Extraction, Normalization, FMRQR Algorithm, Classification, Segment IV gives the Results and Discussions, Segment V concludes the paper.

II. Literature Review

Ashima Gawar et.al.,[3] discussed the comparison analysis of the Quickreduct and the Relative QuickReduct algorithm. The Relative QuickReduct algorithm finds reducts based on backward elimination of attributes and the QuickReduct algorithm finds reducts based on forward elimination. We also found out that Quick Relative Reduct was better than the QuickReduct algorithm.

K.Anitha et.al.,[4] proposed a Rough Set based Quick-Reduct Algorithm have been used to reduce the gene expression data. It gives the minimal reduct set for the given data set. We can use it for Car data set, Mammogram Image Analysis, Iris- Thyroid information system and so on.

A. Chinnaswamy et.al.,[5] proposed a novel method that employs correlation based filter for dimensionality reduction followed by fuzzy rough quick reduct for feature selection on a particle swarm optimization search space. The first phase removed the redundant genes using correlation coefficient filter on a particle swarm optimization search space. The second phase produced a fuzzy rough quick reduct that would be used for classification. The genes obtained after feature selection are subjected to classification using traditional classifiers. It has been determined that the proposed method contributes to reduction in the total number of genes and improvement in the classifier accuracy compared to gene selection and classification using correlation coefficient and traditional fuzzy rough quick reduct algorithm. This approach also reduces the number of misclassifications that might occur in other approaches.

J. R. Anaraki et.al.,[6] proposed a new method based on Fuzzy Lower Approximation-Based Feature Selection is proposed which selects smaller subset of features, makes better classification accuracy and run faster than the base method, especially on big datasets.

P. S. V. S. Sai Prasad et.al.,[7] proposed a Quick Reduct Algorithm (QRA) for reduct computation is most popular since its discovery. The QRA has been modified in this paper by sequential redundancy reduction approach. The performance of this new improved Quick Reduct (IQRA) is discussed in this work.

III. MATERIALS AND METHODS

A. METHODOLOGY

The Brain Machine Interface's Motor Imagery based EEG Signals encloses feature extraction, normalisation, feature selection, and classification. The workflow of the proposed methodology is shown in Fig.1. This methods is applied to different motor imagery EEG signals in BCI Datasets. More datasets with two class mental tasks with different subjects are available.

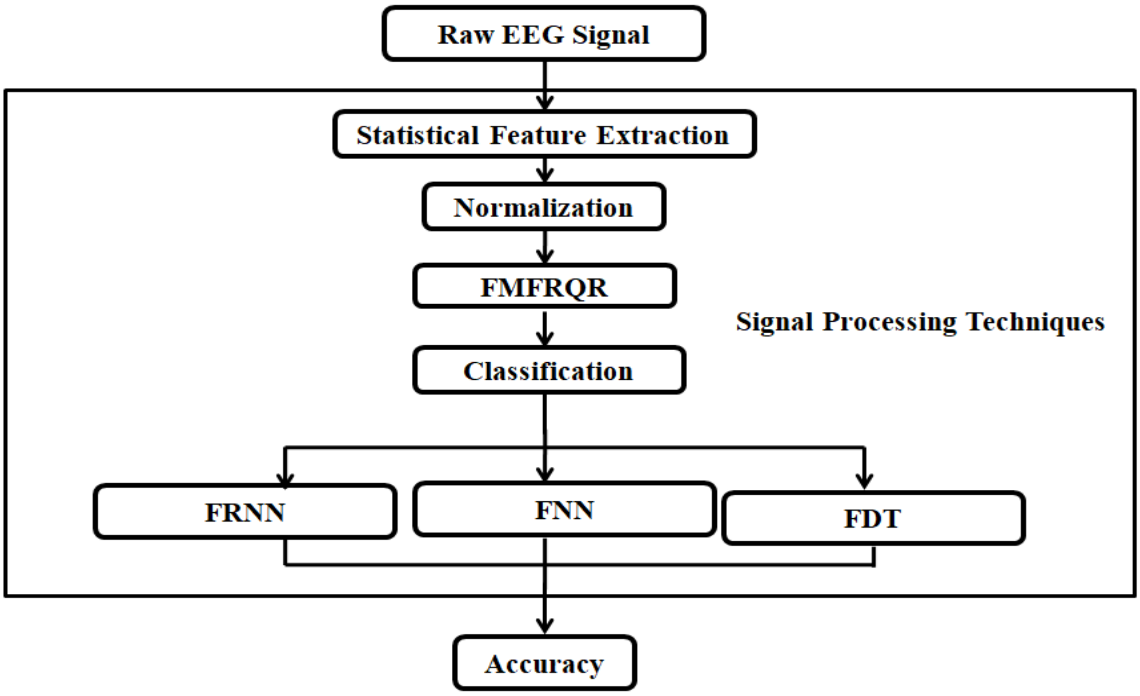

Fig. 1.Proposed Methodology

B. DATASET DESCRIPTION

Datasets are furnished with the aid of Professor. Cichocki's Lab (Lab. for Advanced Brain Signal Processing), [8].

Datasets recording:

At starting trail, the subject became sitting in front of a blank computer display. 2 seconds after the trail commenced out, a cue seemed on the display as an arrow, left arrow in favor of imagining left hand motion and right arrow in favor of imagining proper right hand motion. The cue period is shown in Table. 1 within the duration column. g.tec (g.USBamp) became utilized to record the electroencephalography (EEG) at sample rate equals to 256 Hz. The recordered EEG signals had been bandpass filtered within the range from 2 Hz to 30 Hz. 50 Hz notch filter out become carried out too. Six electrodes had been utilized: C4, Cp4, C3, Cp3, Cz and Cpz. Fig. 2 and Fig. 3 illustrate the time scheme of EEG signal recording and six electrodes arrangement. Table 1. describes the details of the used datasets in details.

Fig.2.The time design of EEG signal recording

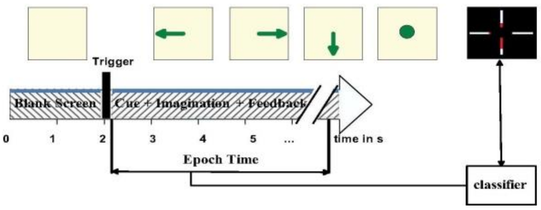

Fig.3. The electrodes placement

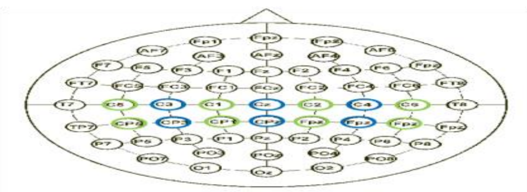

Table 1. Detail Information of Dataset

Dataset	Subject	Class	Channel	Duration (sec)	Trial number	10x10 CV (Acc.±std.)	Sample rate	Device
SubA_6chan_2LR_s1	A	LH/RH	6	3s	130	0.88±0.01	256Hz	g.tec
SubA_6chan_2LR_s2					134	0.84±0.01		
SubB_6chan_2LR	B	LH/RH	6	4s	162	0.88±0.01	250Hz	Neuroscan
SubC_6chan_2LR_s1	C	LH/RH	6	3s	170	0.86±0.01	256Hz	g.tec
SubC_6chan_2LR_s2				3s	158	0.85±0.01		
SubC_6chan_2LR_s3				5s	48	0.92±0.01		
SubC_6chan_2LR_s4				3s	120	0.89±0.01		
SubC_6chan_2LR_s5				3s	90	0.93±0.01		
SubD_5chan_2LR	D	LH/RH	5	4s	80	0.75±0.02	256Hz	g.tec
SubE_5chan_2LR	E	LH/RH	5	4s	48	0.86±0.02	256Hz	g.tec
SubF_6chan_2LR	F	LH/RH	6	4s	80	0.71±0.03	256Hz	g.tec
SubG_6chan_2LR	G	LH/RH	6	4s	120	0.81±0.01	256Hz	g.tec
SubH_6chan_2LR	H	LH/RH	6	3s	150	0.71±0.02	256Hz	g.tec
SubA_6chan_2LF	A	LH/RF	6	3s	150	0.84±0.004	256Hz	g.tec
SubC_6chan_2LF_s1	C	LH/F	6	3s	330	0.96±0.002	256Hz	g.tec
SubC_6chan_2LF_s2					180	0.91±0.004		
SubC_6chan_2LF_s3					102	0.86±0.02		

C. FEATURE EXTRACTION

Statistical feature extraction involves calculating key statistical properties from data to describe its essential characteristics. These methods simplify raw data into meaningful attributes that can be used in analysis and machine learning. Below are commonly used statistical feature extraction methods such as mean [9], variance [10], standard deviation, skewness [11], kurtosis [12], maximum, minimum ,range[13],entropy [14], correlation coefficients[15] etc.

D. Normalization

Normalization is the process of adjusting data to a standard scale, typically by rescaling features to a common range or transforming them to have a mean of 0 and a standard deviation of 1. This technique ensures that each feature contributes equally to the analysis or model, preventing certain features from dominating due to differences in scale [16].

Z-Score Normalization (Standardization)

Z-score normalization rescales data to have a mean of 0 and a standard deviation of 1.

$$z = \frac{x - \mu}{\sigma}$$

Where μ is the mean, and σ is the standard deviation of the dataset.

E. Friedman Based Fuzzy Rough Quick Reduct Algorithm

This algorithm combines Friedman's statistical rank-based measure with fuzzy rough set theory for feature selection, enabling effective classification in BCI (Brain-Computer Interface) systems [17][18].

Algorithm: Friedman Based Fuzzy Rough Quick Reduct (FMFRQR)

Inputs

- U: Set of training instances.
- A: Full set of features.
- D: Decision attribute (class labels).
- $\gamma R(D)$: Dependency degree using fuzzy rough sets for subset R.
- F(a;D): Friedman measure for a feature a based on D.

Output

- R: A reduced subset of significant features.

Algorithm Steps

Step 1: Initialization

1. Start with an empty subset R:
   R=∅
2. Initialize the dependency degree for the current subset R:
   γR(D)=0
3. Compute initial Friedman scores F(a;D)F(a; D)F(a;D) for all a∈A:
   F(a;D)=$\sum_{i=1}^{k}[r_i(a) - \bar{r}(a)]^2$
4. where $r_i(a)$ is the rank of feature a for class i, k is the number of classes, and $\bar{r}(a)$ is the mean rank of a across all classes.

Step 2: Feature Selection Loop

4. While γR(D) <γA(D):
   a. Compute the gain of adding a feature a∈A\R using the fuzzy rough dependency:
   $\Delta\gamma_{RU\{a\}}$= $\gamma_{RU\{a\}}$(D)- γ_R(D)

   b. Use Friedman's measure to select the most significant feature:
   a*=arg max (F(a;D)· $\Delta\gamma_{RU\{a\}}$)
   a∈ $A\backslash R$
   Add a* to the reduct set R:
   R=R∪{a*}

Step 3: Stopping Criteria

6. Stop when adding further features does not significantly increase γR(D):
   γR(D)≈γA(D)

Output

- R: Final selected feature subset.

This formulation ensures robust feature selection by integrating Friedman's ranking-based significance measure with fuzzy rough dependency theory, tailoring it for effective classification in BCI applications.

F. Classification

Classification is a type of supervised learning in machine learning, where the goal is to predict the categorical label or class of a given input based on prior training data. The process involves learning a mapping from input data (features) to output classes. The system is trained using a labeled dataset, and the learned model can then classify new, unseen data into one of the predefined categories.

Fuzzy-Rough Nearest Neighbor (FRNN)

Fuzzy-Rough Nearest Neighbor (FRNN) is an advanced classification technique that integrates fuzzy set theory and rough set concepts to improve decision-making in uncertain and imprecise environments. Unlike the traditional k-Nearest Neighbors (k-NN) algorithm, which assigns a crisp class label based on the majority of neighbors, FRNN employs fuzzy-rough approximations to assess the degree of membership of an instance to different classes. This method enhances classification performance by effectively handling overlapping and ambiguous data points. The approach relies on computing the fuzzy-rough lower and upper approximations of classes, using the similarity between data instances to determine class membership in a more flexible and accurate manner[19].

Fuzzy Logic-based Neural Networks (FNN)

Fuzzy Logic-based Neural Networks (FNN) are hybrid computational models that integrate the reasoning capabilities of fuzzy logic with the learning abilities of neural networks. These models leverage fuzzy logic to handle uncertainty and imprecision in data while using neural networks to adaptively learn patterns and relationships. Unlike conventional neural networks that rely on precise numerical inputs, FNNs incorporate fuzzy rules and membership functions to process vague or ambiguous information effectively. This combination enhances interpretability, robustness, and decision-making, making FNNs suitable for applications such as pattern recognition, control systems, and decision support[20].

Fuzzy Decision Tree (FDT)

A Fuzzy Decision Tree (FDT) is an extension of traditional decision trees that incorporates fuzzy logic to handle uncertainty and imprecise data in classification and decision-making tasks. Unlike conventional decision trees, which use sharp decision boundaries, FDTs employ fuzzy membership functions to allow partial belonging of instances to multiple categories. This approach improves flexibility and interpretability, especially when dealing with real-world data that may be noisy or ambiguous. The construction of an FDT involves selecting attributes, defining fuzzy partitions, and recursively splitting the dataset while maintaining fuzzy membership values, leading to more robust and adaptive decision-making[21].

IV. Result and Discusion

Seventeen BCI datasets taken from UCI repository of machine learning are utilized for performing experimental results. These datasets characteristics are presented in Table 1. Also experimental results are shown in Table 2. In Table 2. report the classification results obtained by using classification algorithm on seventeen BCI datasets.. It shows the accuracy and error rate for three classification algorithm. Our proposed method provides the best classification accuracy using statistical feature extraction techniques. As we expected, Fuzzy-Rough Nearest Neighbor (FRNN) produces the highest accuracy for all BCI datasets compared among other algorithms such as FNN and FDT. The results are satisfied compared with other classification algorithms.In Figure 4: Comparative Analysis of various Classification Algorithms such as FRNN, FNN, FDT. Our results shown that FRNN reached the high level of accuracy compared with FNN and FDT Thus the FRKNN classification algorithm is suitable for BCI dataset.

Table 2. Classification Measures

S.No	Data Set	AC	ER	AC	ER	AC	ER
1	SubA_6chan_2LR_s1	97.3	2.7	96.6	3.4	96.2	3.8
2	SubA_6chan_2LR_s2	96.4	3.6	95.3	4.7	95.2	4.9
3	SubB_6chan_2LR	94.2	5.8	93.3	6.7	93.2	6.9
4	SubC_6chan_2LR_s1	93.1	6.9	92.4	7.6	92.1	7.9
5	SubC_6chan_2LR_s2	91.5	8.5	90.5	9.5	90.1	9.9
6	SubC_6chan_2LR_s3	92.3	7.7	92.4	7.6	92.1	7.9
7	SubC_6chan_2LR_s4	99.5	0.5	97.1	2.9	97.1	2.9
8	SubC_6chan_2LR_s5	97.4	2.6	96.5	3.5	96.1	3.9
9	SubD_5chan_2LR	99.8	0.2	98.6	1.4	98.3	1.7
10	SubE_5chan_2LR	95.3	4.7	94.7	5.3	94.1	5.9
11	SubF_6chan_2LR	94.6	5.4	93.8	6.2	93.2	6.8
12	SubG_6chan_2LR	96.2	3.8	94.6	5.4	94.5	5.8
13	SubH_6chan_2LR	99.5	0.5	98.4	1.6	98.2	1.8
14	SubA_6chan_2LF	95.9	4.1	93.2	6.8	93.4	6.6
15	SubC_6chan_2LF_s1	97.8	2.2	95.5	4.5	95.3	4.7
16	SubC_6chan_2LF_s2	98.2	1.8	96.3	3.7	96.3	3.8
17	SubC_6chan_2LF_s3	98.5	1.5	97.5	2.5	97.3	2.8
Mean		96.3	3.7	95.1	4.9	94.9	5.1

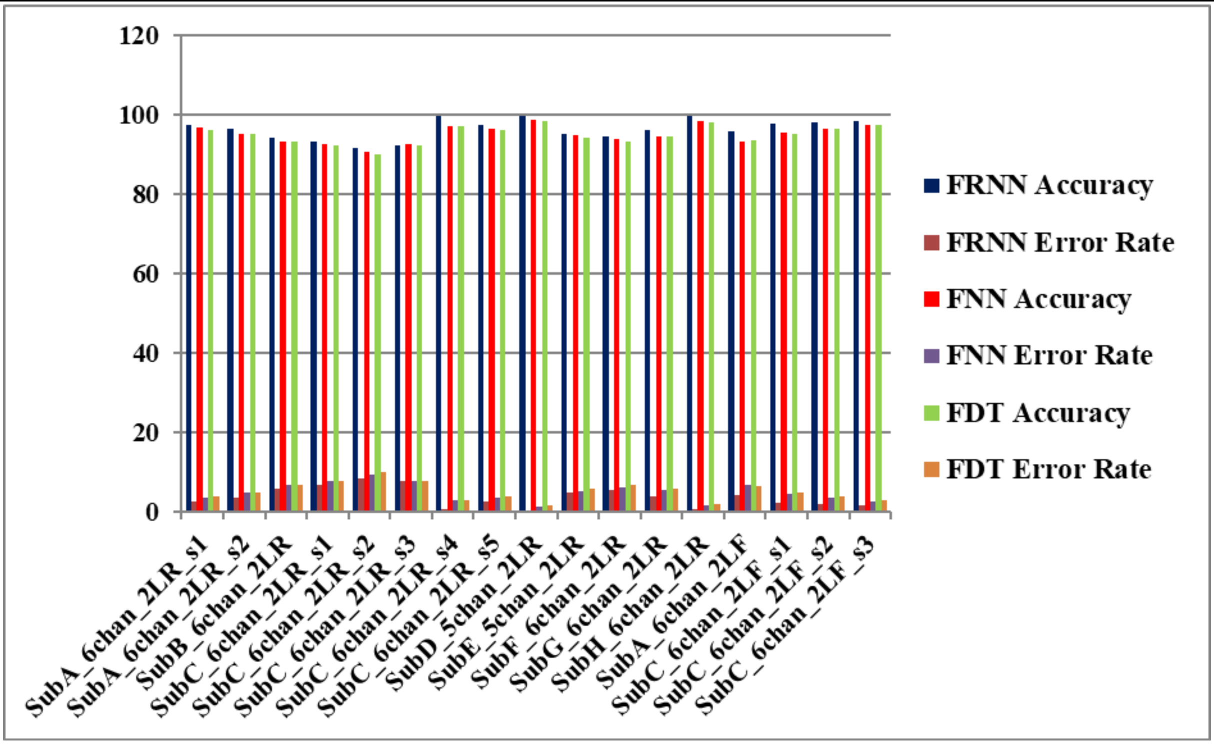

Figure 4: Comparative Analysis of various Classification Algorithms

V. Conclusion:

In this paper Friedman-based fuzzy rough quick reduct algorithm (FMFRQR) Algorithm have been used to reduce the BCI data. This approach runs quicker than the current FMFRQR algorithm, improves accuracy, and reduces the number of picked features. Proposed method has been applied on seventeen BCI Datasets taken from UCI and the results confirm the performance and applicability of the method in terms of selected features , classification accuracy and error rate. Then those results are applied to classification algorithm with ten-fold cross- validation. Finally, the classified results are obtained by classification methods. The future work is aimed at a better understanding of attribute selection for the machine learning techniques through the combination of some other attribute Selection methods. And also we reduce the computational time with use of deep learning methods.

References

[1] Wolpaw, J. R., et al., "Brain-computer interfaces for communication and control", Clinical Neurophysiology, 113(6), pp.767-791,2002.

[2] Nicolelis, M. A. L.," Brain-machine interfaces to restore motor function and probe neural circuits", Nature Reviews Neuroscience, 10(6), pp.417-422,2019.

[3] Ashima Gawar," Performance Analysis of Quickreduct, Quick Relative Reduct Algorithm and a New Proposed Algorithm", Global Journal of Computer Science and Technology: C Software & Data Engineering,pp.1-7,2014.

[4] K.Anitha , Dr.P.Venkatesan ," Feature Selection By Rough – Quick Reduct Algorithm", International Journal of Innovative Research in Science, Engineering and Technology,pp. 3989-3993,2013.

[5] A. Chinnaswamy and Ramakrishnan, S., "Modified Fuzzy Rough Quick Reduct Algorithm for Feature Selection in Cancer Microarray Data", Asian Journal of Information Technology, vol. 15, no. 2, pp. 199–210, 2016.

[6] J. R. Anaraki, M. Eftekhari," Improving fuzzy-rough quick reduct for feature selection", Iranian Conference on Electrical Engineering,May 2011.

[7] P. S. V. S. Sai Prasad and C. Raghavendra Rao," IQuickReduct: An Improvement to Quick Reduct Algorithm, Lecture Notes in Computer Science,pp. 152–159,2009.

[8] http://www.bsp.brain.riken.jp/~qibin/homepage/Datasets.html(Accessed:02/03/2018 20:30)

[9] **Alessio Ferone** and **Antonio Maratea**," Adaptive Quick Reduct for Feature Drift Detection", Algorithms,pp.1-13,2021.

[10] Jain, A., Murty, M. N., & Flynn, P. J.,"Data clustering: A review. ACM Computing Surveys", 31(3), 264-323,1995.

[11] Pechenizkiy, M., Tsymbal, A., & Puuronen, S.," PCA-based feature transformation for classification", Proceedings of the 19th IEEE Symposium on Computer-Based Medical Systems,2006.

[12] Tharwat, A., et al. Metrics to measure the performance of machine learning algorithms. Journal of Applied Computing and Informatics,2017.

[13] DeCarlo, L. T.," On the meaning and use of kurtosis", Psychological Methods, 2(3), 292.,1997.

[14] Shannon, C. E., "A Mathematical Theory of Communication". Bell System Technical Journal,1948.

[15] Guyon, I., and Elisseeff, A..," An introduction to variable and feature selection", Journal of Machine Learning Research, pp.1157-1182,2003.

[16] Wright, M., " Scaling and normalizing data in machine learning ", Statistical Methods for Data Analysis, p.91-110,2008.

[17] Bishop, C. M, "Pattern Recognition and Machine Learning", Springer, 2006.

[18] Hastie, T., Tibshirani, R., & Friedman, J., "The Elements of Statistical Learning: Data Mining, Inference, and Prediction". Springer.

[19] Jensen, R., and Cornelis, C., "Fuzzy-rough nearest neighbour classification. In: Rough Sets and Knowledge Technology", Lecture Notes in Computer Science,2011.

[20] Jang, J. S. R., "ANFIS: Adaptive-network-based fuzzy inference system". IEEE Transactions on Systems, Man, and Cybernetics, pp.665–685, 1993.

[21] Pal, S. K., & Chakraborty, D.," Fuzzy decision tree, linguistic rules and fuzzy knowledge-based network: Generation and evaluation". Fuzzy Sets and Systems, pp.97-114,2001.