International Journal of Computational Intelligence and Informatics, Vol. 1 : No. 1, April - June 2011

# Supervised Feature Subset Selection using Extended Fuzzy Absolute Information Measure for Different Classifiers

**K. Sarojini**
Department of MCA
S.N.R. Sons College, Coimbatore, Tamilnadu, India – 641 006.
saromaran@gmail.com

## Abstract

Feature subset selection plays an important role in data mining and machine learning applications. The main aim of feature subset selection is reducing dimensionality by removing irrelevant and redundant features and improving classification accuracy. This paper presents a supervised feature selection method called as Extended Fuzzy Absolute Information Measure (EFAIM) for different classifiers. In this process, first discretization algorithm is applied to discretize numeric and nominal features of a database to construct fuzzy sets of a feature. Then the method EFAIM is applied to select feature subset focusing on boundary samples. To verify the effectiveness of this method, several experiments are conducted for different classifiers, such as, LMT, Naïve Bayes, SMO, C4.5, JRIP, PART and Simple Cart with different UCI datasets. The Experimental results indicates that the proposed algorithm have achieved better classification accuracy for all datasets, that is, almost above 75% of accuracy. For WINE dataset, it gets 96% of classification accuracy for Naïve Bayes classifier. For Ionosphere dataset, it gives almost 89% of classification accuracy for maximum of classifiers with minimum selected feature subset. Thus improved classification accuracy is obtained with selected subset of minimum number of features at minimum processing time.

**Keywords**: Boundary samples, Classification, Data mining, Discretization, Extended Fuzzy Absolute Information Measure, Feature selection, Fuzzy sets, Membership function.

## 1. Introduction

Data preprocessing is an important and critical step in the data mining process. It has huge impact on the successes of data mining projects. The purpose of data preprocessing is to cleanse the noise data, extract and merge the data from different sources, and then transform and convert the data into a proper format.

Feature subset selection is one of the data preprocessing steps in data mining. It refers to choosing subset [4] of attributes from the set of original attributes. Aim of feature subset selection is to reduce the number of features in the dataset and to increase the classification accuracy rate. It is obvious that a data set might have irrelevant and relevant features. If relevant features are selected properly, then it is possible to increase the classification accuracy rates [6, 18]. This paper presents a supervised feature subset selection based on EFAIM for different classifiers. First, discretization algorithm, Fuzzy C Means (FCM) is used to discretize numeric features to construct the membership function of each fuzzy set of a feature. Then, it selects the feature subset, based on the method EFAIM focusing on boundary samples. This method can select feature subset with minimum number of features, which are relevant to get higher average classification accuracy for different datasets. The selected features are validated by using classifiers, which is available in a free software WEKA (http://www.cs.waikato.ac.nz/ml/weak/). Average classification accuracy rates of EFAIM method is compared for different classifiers such as LMT, Naïve Bayes, SMO, C4.5, JRIP, PART and CART. The experiments are conducted for different data sets taken from UCI Repository of Machine Learning Databases (Data source: UCI Repository of Machine Learning Database and Domain Theories, (ftp://ftp.ics.res.uci.edu/pub/machine-learning-databases/).

Hence, the feature subset selection method can select feature subset with minimum number of features, which are relevant to get higher average classification accuracy rate, that is, above 75% for all datasets. For WINE dataset using Naïve Bayes classifiers, it improves the classification accuracy rate to 96% for minimum number of selected subset. In Ionosphere dataset almost for maximum classifiers, it gives 89% of classification accuracy with minimum selected feature subset. For Ecoli dataset, classification is improved by 80%. Thus improved accuracy is obtained for all datasets with selected subset of minimum number of features at minimum processing time.

The rest of this paper is organized as follows. Section 2 briefly reviews related work of feature subset selection. Section 3, presents EFAIM based method for feature subset selection focusing on the boundary samples. An experimental analysis is performed in section 4. Section 5 concludes this paper.

## 2. Literature review

### 2.1 Fuzzy Entropy Measure

Fuzzy sets and logic are powerful mathematical tools for modelling and controlling uncertain systems in industry, humanity, and nature; they are facilitators for approximate reasoning in decision making in the absence of complete and precise information. Their role is significant when applied to complex phenomena not easily described by traditional mathematical tools. Entropy is a measure, which measures the impurity of a collection. Some of the existing entropy measures are found in [4, 6, 9, 11, 13, 15, 16]. Let X be a discrete random variable with a finite set containing n elements, where X = {$x_1$, $x_2$ ... $x_n$}. If an element $x_i$ occurs with a probability $P(x_i)$, then the amount of information $I(x_i)$ associated with $x_i$ is defined as follows:

$$I(x_i) = -\log_2 p(x_i) \tag{1}$$

The entropy H(X) of X is defined as follows:

$$H(X) = -\sum_{I=1}^{n} P(x_i) \log_2 P(x_i) \tag{2}$$

where n denotes the number of elements and $P(x_i)$ denotes the occurring probability of the element $x_i$. Zadeh defined a fuzzy entropy on a fuzzy set Ã for a finite set X = {$x_1$, $x_2$… $x_n$} with respect to the probability distribution P = {$p_1$, $p_2$,… $p_n$}, shown as follows:

$$H(X) = -\sum_{i=1}^{n} \mu_{\tilde{A}}(x_i) P_i \log P_i \tag{3}$$

where $\mu_{\tilde{A}}$ denotes the membership function of Ã. $\mu_{\tilde{A}}(x_i)$ denotes the grade of membership of $x_i$ belonging to the fuzzy set Ã , $p_i$ denotes the probability of $x_i$ , and $1 \leq i \leq n$. In [6], Lee et al. presented a fuzzy entropy measure of an interval, based on Shannon's entropy measure and Luca's axioms.

Feature selection method selects relevant features to get higher average classification accuracy. MIFS method [1], the FQI method [2], Dong-and-Kothari's method [3] and the OFFSS method [4] are various feature selection method can select features to get higher average classification accuracy rates.

### 2.2 Information Gain Measure

Information Gain Measure [8, 10] is a quantitative measure of an attribute. Entropy measures the impurity of a collection, whereas Information Gain measures the purity of a collection with respect to class attribute. The amount by which the entropy of X decreases reflects additional information about X provided by Y and is called as information gain.

### 2.3 Fuzzy Absolute Information Measure [FAIM]

The FAIM [20] can be used to measure the degree of similarity between two fuzzy sets A and B, and so they are useful for further development of the theory of similarity measures.

Suppose X is a discrete universe of discourse, and A,B ∈ ξ(X), then the absolute difference value, denoted by R(A, B), is called the fuzzy absolute information measure of B to A, where ξ(X) is a set consisting of all fuzzy subsets of discrete universe of discourse X, and is defined as,

$$R(A,B) = H(A \cap B) = H(A) - H(A/B) \tag{4}$$

is called the fuzzy absolute mutual information of fuzzy set A to fuzzy set B. Here H(A), H(B) denotes the fuzzy entropy measure of attribute set and H(A/B) is joint entropy of the attribute set. R(A, B) shows an influence degree of the fuzzy set A to fuzzy set B in fuzzy information processing aspects. It is a generalized absolute fuzzy entropy measure, and is called the Fuzzy Absolute Information Measure (FAIM). The FAIM can be used for clustering analysis, picture processing, similarity measures etc.

### 2.4 Discretization Process

In data mining, discretization process is known to be one of the most important data preprocessing tasks. Most of the existing machines learning algorithms are capable of extracting knowledge from databases that store discrete attributes. If the attributes are continuous, the algorithms can be integrated with discretization algorithms, which transform them into discrete feature. Discretization methods are used to reduce the number of values for a given continuous attributes by dividing the range of the attribute into intervals. Discretization makes learning more accurate and faster [5, 7, 12, 19].

Normal discretization process specifically consists of the following four steps:

(i) Sort all the continuous values of the feature to be discretized.

(ii) Choose a cut point to split the continuous values into intervals.

(iii) Split or merge the intervals of continuous values

(iv) Choose the stopping criteria of discretization process [19].

2.4 Boundary samples

Feature subset selection problem is regarded as a dimension reduction problem [9, 17]. In dimension reduction problems, boundary samples take important part in affecting the results. In [14], EFAIM based feature selection focusing on boundary samples gives minimum subset of features. The two dimensional feature spaces are reduced to one dimensional feature space. Reduced dimension will increase the entropy of data because some information will be omitted while reducing the dimension. So, this would decrease the classification accuracy. In dimension reduction problem, each feature might have incorrectly classified samples. Thus, an optimal feature subset is a set of correlated features. It means that other features could correctly classify the samples incorrectly classified by a feature. Boundary samples are incorrectly classified samples of features. So, feature subset selection should focus on boundary samples. Thus entropy of feature subset is calculated only by using the boundary samples instead of full set of samples. This improves the classification accuracy.

## 3. Proposed Work

3.1 Methodology

This section presents a method for feature subset selection based on EFAIM is a dimension reduction problem, which focuses on the "boundary samples" instead of a full set of samples to select the feature subset. However, using the boundary samples directly to calculate a feature subset is not possible. So, an indirect method is used to simplify the feature subset selection process described as follows.

Considering both fuzzy entropy measure of a feature using class degree and based on entropy measure of class, the proposed EFAIM is setup. It measures the purity of samples. EFAIM(C, f), is called as Extended Fuzzy Absolute Information Measure of a feature (f) with respect to class C. In this process, EFAIM is applied only for boundary samples. So, it is called as Extended Fuzzy Absolute Information Measure for Boundary Samples EFAIMBS. It is defined as:

$$EFAIM(C, f) = H(C) - FE(C / f) \quad (5)$$

In this process, Fuzzy Entropy measure of a feature FE(C/f) is defined as:

$$FE(C / f) = \sum_{\bar{A} \in V} \frac{S_{\bar{A}}}{S} FE\left(\bar{A}\right) \quad (6)$$

Where FE(Ã) denotes the fuzzy entropy of fuzzy set Ã, V denotes the set of fuzzy set of feature f; $S_{\tilde{A}}$ denotes the summation of the membership grades of the samples belonging to the fuzzy set Ã. S denotes the sum of the membership grades of samples belonging to each fuzzy set of a feature f. In this process, fuzzy entropy measure FE(Ã) was defined based on the following definitions 3.1, 3.2, and 3.3.

**Definition 3.1** A set X of samples is divided into a set C of classes. The class degree [9] $CD_c(\tilde{A})$ of the samples of class c, where c ϵ C, belonging to the fuzzy set Ã is defined by:

$$CD_C(\tilde{A}) = \frac{\sum_{x \epsilon X_c} \mu_{\tilde{A}}(x)}{\sum_{x \epsilon X} \mu_{\tilde{A}}(x)} \quad (7)$$

where $X_c$ denotes the samples of class c, c ∈ C, $\mu_{\tilde{A}}$ denotes the membership function of the fuzzy set, $\mu_{\tilde{A}}(x)$ denotes the membership grade of x belongs to the fuzzy set Ã, $\mu_{\tilde{A}}(x) \in [0, 1]$.

**Definition 3.2** The fuzzy entropy $FE_c(\tilde{A})$ [9] of the samples of class c, where c ϵ C, belonging to the fuzzy set Ã is defined as follows:

$$FE_C(\tilde{A}) = -CD_C(\tilde{A})(\log_2 CD_C(\tilde{A})) \quad (8)$$

**Definition 3.3** The fuzzy entropy FE(Ã) [9] of a fuzzy set Ã is defined by:

$$FE(\tilde{A}) = \sum_{c \in C} FE_C(\tilde{A}) \quad (9)$$

Similarly, The Entropy of class C with respect to different values is defined as:

$$H(C) = -\sum_{i=1}^{n} P_i \log_2 P_i \quad (10)$$

Thus EFAIM was constructed for the proposed algorithm using fuzzy entropy measure of a feature and class.

3.2 The EFAIM Algorithm

The proposed work and its overall structure are represented in this algorithm. The proposed work starts with discretization process. Discretization process is one of the most important data preprocessing tasks. For the continuous attributes, the algorithms can be integrated with discretization algorithms to transform the continuous

attributes into discrete feature. For this process, it uses Fuzzy C means algorithm to discrete numerical attributes and constructs the triangular membership function to fuzzify all numeric features. For each feature (f), EFAIM is calculated. The steps are given here under.

**Step1:** Initially, set the number K of clusters as 2.

**Step2:** Use Fuzzy C means clustering algorithm to generate K clusters centers based on the values of a feature, where $K \geq 2$.

**Step3:** Construct the membership function of the fuzzy sets using triangular membership functions based on these K cluster centers, respectively.

**Step4:** Calculate fuzzy entropy of feature f using class degree.

$$FE(C / f) = \sum_{A \in V} \frac{S_{\overline{A}}}{S} FE\left(\overline{A}\right)$$

**Step5:** Calculate Entropy of class C.

$$H(C) = -\sum_{i=1}^{n} P_i \log_2 P_i$$

**Step6:** Calculate EFAIM for feature f using fuzzy entropy measure.

$$EFAIM(C, f) = H(C) - FE(C / f)$$

**Step7:** If the decreasing rate of the EFAIM of feature f is larger than the threshold value $T_c$ given by the user, where $T_c \in [0, 1]$, then let K = K+1 and go to Step 2. Otherwise, let K = K−1 and Stop. The threshold value $T_c$ is used in the algorithm for constructing the membership functions of the fuzzy sets of a numeric feature. In this process, thus EFAIM is calculated for each feature.

3.3 Application of EFAIM based on boundary samples for feature subset selection

This section presents the application of EFAIM for feature subset selection focusing on boundary samples. Feature subset selection used in this process considers only the boundary samples instead of full set of samples. While constructing EFAIM, a threshold value $T_r$ is used, where $T_r \in [0, 1]$, which omit the fuzzy set of a feature whose maximum class degree is larger than or equal to the threshold value $T_r$ given by the user for feature subset selection. According to (1), there are n "class degrees" of a set of samples belonging to a fuzzy set with respect to n classes. The maximum class degree of a fuzzy set is defined as the maximum among the n "class degrees". If the maximum class degree of a fuzzy set is larger than or equal to the given threshold value $T_r$, then the fuzzy set will be omitted to reduce the number of fuzzy sets of the feature. These fuzzy sets of a feature having maximum class degrees are considered as correctly classified samples of a feature. So, these samples are omitted, and the remaining samples are considered as boundary samples. Therefore, these samples having lesser class degree than $T_r$ are only considered as boundary samples and included in this process. To consider only boundary samples, the Extension Matrix (EM) is constructed for each feature based on EFAIM. Features having highest EFAIM values are considered as best features and these features are combined using Combined Extension Matrix (CEM) [9]. Then fuzzy entropy measure BSFFE(f1, f2) [9] of a feature subset {f1, f2} focusing on boundary samples is calculated. Considering both fuzzy entropy measure, BSFFE(f1, f2) and entropy measure of class, the EFAIMBS is setup. It measures the purity of boundary samples.

In this process, a set R of samples is divided into a set C of classes, where $R = \{r_1, r_2, \ldots, r_c\}$, F denotes a set of candidate features and FS denotes the selected feature subset. The algorithm for feature subset selection focusing on Boundary samples are presented as follows:

**Step1:** Construct the extension matrix $EM_f$. For each feature f membership grades of values of fuzzy sets are defined in $EM_f$. For these fuzzy sets, calculate class degree with respect to class and fuzzy entropy. Then calculate fuzzy entropy for each feature. Using these measures, calculate Extended Fuzzy Absolute Information Measure EFAIM(f) for each feature.

**Step2:** Put the feature with the maximum EFAIM(f) into the selected feature subset FS and remove it from the set F of candidate features.

Let $E_{FS}$ = EFAIM(f), where fs = arg max $_{f \in F}$ EFAIM(f)

Let FS = {fs};

Let F = F - {f};

**Step3:** Repeatedly put the feature, which can increase the EFAIMBS of the feature subset into FS until no such a feature exists.

Repeat
{
For each $f \in F$, do

{

Calculate $EM_{FS \cup \{f\}}$ according to the maximum class degree threshold value $T_r$ given by the user, where $T_r \in [0, 1]$.

$EM_{FS\ U\ \{f\}} = CEM(FS, f, T_r)$;

Calculate Fuzzy Entropy BSFFE(FS, f) of the feature subset FS U {f} focusing on boundary samples. Calculate Extended Fuzzy Absolute Information Measure EFAIMBS(FS, f) of the feature subset FS U {f} focusing on boundary samples.

};

Let $f = \text{argmax}_{f \in F}$ EFAIMBS(FS, f) which returns one of such a feature f that maximizes the function EFAIMBS(FS, f);

Let $D = EFAIMBS(FS, f) - E_{FS}$;

Let EFS = EFAIMBS(FS, f);

Let FS = FS U {f};

Let F = F - {f};

};

Until ( $E_{FS} = 0$ (or) $D \leq 0$ (or) F = { });

Let FS is a selected feature subset.

Thus Feature Subset Selection is generated using EFAIMBS focusing on boundary samples.

**4. Experimental Analysis and Discussion**

The proposed method has been implemented using MatLab version 7.0 and the experimental analysis is presented. First, the proposed method is selecting minimum number of features for feature subset in minimum time. Second, the selected feature subset is validated using different classifiers to improve the classification performance of the attributes selected by the EFAIM method.

The proposed algorithm gives improved results for the datasets taken from UCI Repository of machine learning databases. The data sets used here have two categories of features that are nominal and numerical attributes, both of them have their corresponding membership functions of fuzzy sets. Each value of a nominal feature can be regarded as a fuzzy set, where its membership function is defined as follows:

$\mu_v(x) = 1$, if x = v,

else

$\mu_v(x) = 0$, otherwise

where $v \in U$, where U denotes a set of values of a nominal feature, and $\mu_v$ denotes the membership function of the fuzzy set v.

For example, the set of values of the feature “Score” is {pass, fail}. When the value of the feature “Score” is “pass”, the membership grades are:

$\grave{I}_{pass}$ (pass) = 1 and $\grave{I}_{fail}$ (pass) = 0.

A numeric feature can be discretized into finite fuzzy sets. The number of fuzzy sets will affect the result of classification. Therefore, the discretization of a numeric feature is an important process. Using unsupervised learning techniques to discretize a numeric feature is a good method, so they can be discretized to transform the numerical data into categorical one. Fuzzy C Means [FCM] is used to discretize the data. Then the algorithm EFAIM is implemented on the discretized dataset to select minimum number of features for feature subset. To validate the selected features, the selected features are compared with seven different classifiers such as LMT, Naïve Bayes, SMO, C4.5, JRIP, PART and CART.

The experiments are carried out by using data sets taken from UCI Repository of Machine Learning Databases (Data source: UCI Repository of Machine Learning Database and Domain Theories, (ftp://ftp.ics.res.uci.edu/pub/machine-learning-databases/). The selected data are used to train classifiers, which is available in WEKA to evaluate the performance of the selected feature subsets by different methods. The classification powers of the selected data are used 10 fold cross-validation, in which, it divides each data set into 10 subsets of approximately equal size and execute 10 times. Each time it selects one of the 10 subsets as the testing data set and train the classifier with the remaining 9 subsets to get the classification accuracy rate with respect to each selected feature subset.

**Table 1.** Feature Selection with different classifiers

| S. No. | Data sets | Raw Data | Selected features | Time(Sec) | Classifiers | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | | | LMT | NB | SMO | C4.5 | JRIP | PART | CART |
| 1 | WPBC | 33 | 2 | 4.13 | 78.79 | 78.79 | 76.26 | 75.76 | 77.78 | 74.75 | 76.26 |
| 2 | SONAR | 60 | 2 | 7.55 | 76.92 | 69.72 | 75.49 | 76.44 | 74.52 | 76.44 | 76.44 |
| 3 | WINE | 13 | 4 | 3.156 | 92.14 | 96.07 | 93.82 | 92.70 | 91.57 | 94.38 | 90.45 |
| 4 | ECOLI | 7 | 3 | 3.88 | 83.04 | 83.33 | 79.76 | 79.46 | 80.95 | 80.95 | 81.85 |
| 5 | IONO | 34 | 2 | 12.23 | 88.89 | 84.05 | 82.08 | 89.18 | 88.89 | 88.6 | 88.6 |

After executing 10 times, it gives the average classification accuracy rates of feature subset selection algorithm. The average classification accuracy with different classifiers is tabulated in Table-1.

The proposed feature subset selection method for different classifiers can select feature subset with minimum number of features, which are relevant to get higher average classification accuracy for all datasets. Almost for all datasets it gets above 75% of accuracy. For WINE dataset using Naïve Bayes classifiers, it improves the classification accuracy rate to 96% and for Ionosphere dataset almost for maximum classifiers; it gives 89% of classification accuracy with minimum selected feature subset. For Ecoli dataset, classification is improved by 80%. Thus improved accuracy is obtained with selected subset of minimum number of features at minimum processing time.

The threshold value $T_c$ is used in the algorithm for constructing the membership functions of the fuzzy sets of a numeric feature and the threshold value $T_r$ is used in the algorithm for feature subset selection. The threshold value $T_c$ and $T_r$ used in this method for different datasets are shown in Table-2. The subset of selected features of different dataset for EFAIM method is displayed in Figure-1. For the selected Subset of features, the classification accuracy rate of different datasets for different classifiers is displayed in Figure-2.

## 5. Conclusion

Data Mining is not the only answer to all problems and sometimes it has been over emphasized. It is expensive to carry out the entire process and therefore has to be thought out clearly. Feature selection approaches reduce the complexity of the overall process by allowing the data mining system to focus on what is really important. Thus, the data mining knowledge produced is found more meaningful. Also the new users / end users will get better results quickly.

A novel method called EFAIM has been proposed for feature subset selection focusing on boundary samples for different classifiers. The proposed method deals with numeric and nominal features. First, the numerical attributes have been discretized using Fuzzy C Means. Then the method EFAIM was applied for feature subset selection focusing boundary samples. The performance evaluation of the feature subset selection method based on EFAIM was compared with different classifiers.

The experimental results showed that the proposed EFAIM method is very efficient for different size of datasets, in selecting very minimum features for feature subset, giving higher classification accuracy with less time. To verify the effectiveness of this method, several experiments are carried out for classifiers LMT, Naïve Bayes, SMO, C4.5, JRIP, PART and Simple Cart with different UCI datasets. The Experimental result indicates that the proposed algorithm achieves better accuracy for all datasets, that is, almost above 75% of accuracy. For WINE dataset, it gets 96% of classification accuracy for Naïve Bayes classifier. For Ionosphere dataset, it gives almost 89% of classification accuracy for maximum of classifiers with minimum selected feature subset. Thus improved accuracy is obtained with selected subset of minimum number of features at minimum processing time.

**Table 2.** The threshold value $T_c$ and $T_r$ for different datasets

| Data set | Threshold value($T_c$) | Threshold value($T_r$) |
|---|---|---|
| WPBC | 0.50 | 0.80 |
| SONAR | 0.20 | 0.75 |
| WINE | 0.20 | 0.50 |
| ECOLI | 0.20 | 0.30 |
| IONO | 0.20 | 0.70 |

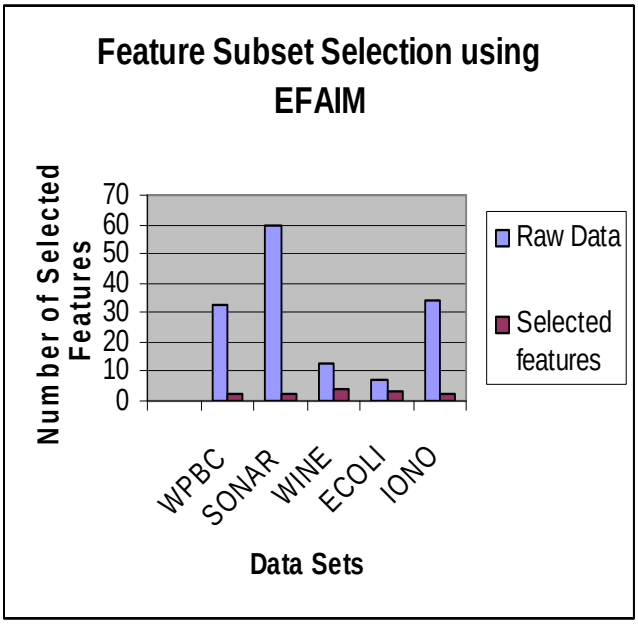


Fig. 1 Feature Subset Selection using EFAIM Method

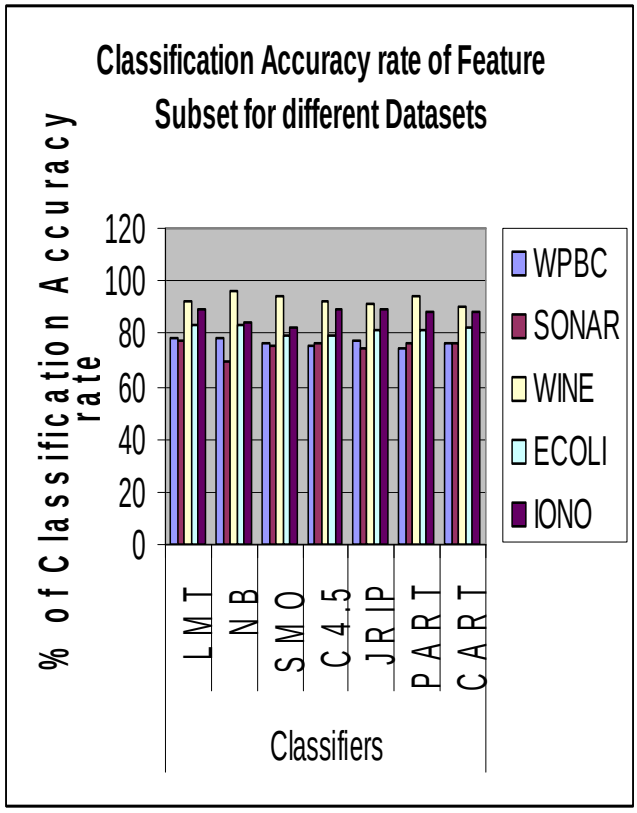


Fig. 2 Classification Accuracy Rate of EFAIM method for Different Classifiers

**References**


[1] Battiti R, "Using Mutual Information for selecting features in supervised neural net learning", IEEE Trans Neural Netw5(4), 1994, pp. 537-550.

[2] De RK, Basak J, Pal SK, "Neuro-fuzzy feature evaluation with theoretical analysis", Neural Netw 12(10), 1999, pp. 1429-1455.

[3] Dong M, Kothari R "Feature subset selection using a new definition of classifiability", Pattern Recognit Lett 24(9), 2003, pp. 1215-1225.

[4] ECC Tsang , Yeung DS, Wang XZ "OFFSS: optimal fuzzy-valued feature subset selection", IEEE Trans Fuzzy Syst 11(2), 2003, pp. 202-213.

[5] H Liu, etal: Discretization: "An Enabling Technique. Data Mining and Knowledge Discovery", Data Mining and Knowledge Discovery, Kluwer Academic Publishers. Manufactured in The Netherlands , 393-423, 2002.

[6] H.M.Lee, C.M.Chen, J.M.Chen, Jou YL, "An efficient fuzzy classifier with feature selection based on fuzzy entropy", IEEE Trans Syst Man Cybern Part B Cybern 31(3), 2001,426–432.

[7] J.A.Hartigan, M.A.Wong, "A k-means clustering algorithm", Journal of the Royal Statistical Society. Series C, 1979, pp. 100–108.

[8] J.D.Shie, S.M.Chen, "A new approach for handling classification problems based on fuzzy information gain measures", In: Proceedings of the 2006 IEEE international conference on fuzzy systems, Vancouver, BC, Canada, 2006, pp 5427–5434.

[9] Jen-Da shie, Shyi-Ming Chen, "Feature subset selection based on fuzzy entropy measures for handling classification problems", Appl Intell , 2008, 28, pp.69-82.

[10] Lei Yu, Huan Liu, "Redundancy Based Feature Selection for Micro array Data", Proceedings of the 2004 ACM SIGKDD, Mylaysia, 2004, pp. 737–742.

[11] Luca AD, Termini S, "A definition of non-probabilistic entropy in the setting of fuzzy set theory", Inf Control 20(4):, 1972, pp.301-312.

[12] Marzuki, F. Ahmad,"Data mining discretization methods and Performance", Proceedings of International conference on Electrical engineering and Informatics, institute Technology Bandung, Indonesia, June 2007.

[13] Nicolaie popescu-bodorin, "Fast K-Means Image Quantization Algorithm and Its Application Iris Segmentation", Bulletin Scientific-2007.

[14] Sarojini. K, Thangavel. K, "Supervised feature subset selection using Extended fuzzy absolute information measure for handling different discretized datasets ", In proceedings of the International Conference and Exhibition on Biometrics Technology, Science Direct , Elsevier, Procedia computer science, Vol. 2, PP. 256-264, 2010.

[15] S.M.Chen, "A new approach to handling fuzzy decision making problems", IEEE Trans Syst Man Cybern 18(6), 1998, pp. 1012–1016.

[16] S.M.Chen, C.H.Chang, "A new method to construct membership functions and generate weighted fuzzy rules from training instances", Cybern Syst 36(4), 2005, pp. 397– 414.

[17] S.M.Chen, J.D.Shie,"A new method for feature subset selection for handling classification

problems", In: Proceedings of IEEE international conference on fuzzy systems, Reno, NV, pp 183-188, 2005.

[18] S.M.Chen, Kao CH, Yu CH," Generating fuzzy rules from training data containing noise for handling classification problems", Cybern Syst 33(7), 2002, pp.723–748.

[19] S.M.Chen, Y.C.Chen, "Automatically constructing membership functions and generating fuzzy rules using genetic algorithms", Cybern Syst 33(8):841–862, 2002.

[20] SHI-Fei Ding, Shi-xiong Xia, Feng-Xiang Jin and Zhong-Zhi Shis, "Novel fuzzy information proximity measures", Journal of Information Science, Volume 33 , Issue 6 , 2007, pp. 678-685.

**K. Sarojini** received her MCA from Bharathidasan University, Trichy, in 1996. She received her M.Phil degree in Computer Science in the year 2003 from Manonmaniam Sundaranar University, Thirunelveli. Currently she is pursuing her PhD at Mother Teresa University for Women. She is also working as an Associate Professor in the Department of MCA, SNR Sons College, Coimbatore. She has 14 years of teaching experience. She has presented papers in various national and international conferences and journals. She has published a book on Fundamentals of Computers. Her area of specialization is Dimensionality Reduction in Data mining.