IOSR Journal of Computer Engineering (IOSR-JCE)
e-ISSN: 2278-0661,p-ISSN: 2278-8727,
PP 13-16
www.iosrjournals.org

A Novel Approach to Optimal Feature Selection for Epileptic Seizure Prediction

Anjum Shaikh[1], Dr.Mukta Dhopeshwarkar [1]
[1](*Department of CS&IT Dr.Babasaheb Ambedkar Marathwada University Aurangabad.)*

Abstract: *To improve the performance of predictive model and data quality it is always necessary to use optimal feature to learn the predictive model i.e. always use relevant feature in dataset and discard the irrelevant feature by using the proper feature subset selection method. The feature subset selection method improves the quality of the data, reduce the dimensionality and improve the accuracy of classifier .In the present study different feature selection method is used to select the optimal relevant feature to predict the Epileptic seizure before its onset.*

***Keywords* - :** Feature Selection, Relieff, Statistical Dependency (SD), Mutual Information (MI)

I. Introduction

Epilepsy is the neurological disorder of brain electrical signal causes seizure because of that the body and brains behave strangely, near about 1% population all over the world affected by this disease. Epilepsy is unpredictable in nature posse the risk, so it is necessary to predict the Epilepsy before its onset.EEG is the Diagnostic tool used to record the brain signal to find the abnormalities rerated to the brain. After recording the brain signal of Epileptic patient the number of characteristic feature were extracted that all features are not important so we need to select the more optimal feature subset using different filter method of feature subset selection method and classify using ANN classifier. The filter method determines the importance of correlation between features and the output variable. Filter based approaches are not depend on classifiers and usually faster and more scalable than wrapper based methods. In addition, they have low computational complexity [1]

II. Feature Selection

Feature selection is also known as attribute selection or variable selection is the method totally differ from the process of feature Extraction, this method select the feature which is already extracted from the data. It is method used to improve the quality of data and performance of the classifier by reducing the size of feature vector and considering the optimal feature .Feature selection method broadly classified into three different type 1) Filter Method 2) Wrapper method 3)Embedded Method.

2.1 Filter Method

Filter Method of Feature Selection **i**s used to identify the subset of input columns (features) that have the highest predictive power. It is the process of applying statistical tests to feature values given a specified class (target), to determine how relevant a feature is to the target. which features are more correlated with the class and rank that feature based on the correlation. The method produce two outputs i.e. first is a dataset containing the top rank features as ranked by predictive power. The second output is a transformed dataset containing the numeric scores assigned to the selected variable. Then we are able to select N number of top rank features (subset) as per requirement from M features (N<M).

2.2. Wrapper Method

In the wrapper method , the feature subset selection is perform using the induction algorithm as a black box. [2].Wrapper algorithms, as per the name , "wrap" the feature selection with a specific classifier and select a optimal subset of features based on the classifier's performance accuracy by using cross-validation method . While there are strong arguments in favor of both approaches, wrapper algorithms are generally slower as compare to filter method. Therefore, researcher will mostly consider only filter methods that do not require explicit class labeling [3].

1.3. Embedded Method

Embedded methods differ from other feature selection methods in the way feature selection and learning interact. , in embedded methods the learning part and the feature selection part cannot be separated - the structure of the class of functions under consideration plays a crucial role.

III. Relief Algorithm

The Relieff [4] algorithm is the simplest and effective method of feature selection it is a feature ranking method. It rank and weight the feature in between -1 and 1[5] the output of this algorithm gives two variable that is rank and weight of feature for input data matrix X and target vector Y for classification. For classification, relieff uses K nearest neighbors per class. RANKED are indices of columns in X ordered by feature importance, meaning RANKED(1) is the index of the most important predictor. WEIGHT are attribute weights ranging from -1 to 1 with large positive weights assigned to important feature.

IV. Statistical Dependency (Sd)

The Statistical Dependency(SD) method of feature ranking is more effective way to measure the dependency of feature with related values of class labels. First the each feature value is quantized into one of QS levels, where the feature-specific quantization scale is flexibly determined such that each bin will contain roughly an equal amount of samples over the entire data set. The bins are selected in this way, instead of a conventional uniform quantization scale, in order to contribute some statistical validity to the occurrence of different quantization levels. The statistical dependence between the discretized feature values f and the class labels c is evaluated by using Eq number (1)

$$SD = \sum_{f \in F} \sum_{c \in C} p(f,c) \frac{p(f,c)}{p(f)\,p(c)} \tag{1}$$

If the Eq (1) produce larger value that means the dependency between the feature values and the class labels is high. In the case that Eq (1) produce minimal value that means feature is fully independent of the class labels[6].

V. Mutual Information (MI)

MI is feature ranking algorithm [7] which ranks the features by measuring the dependency between the two features by using some ranking criteria. MI is describe by Shannon definition for entropy given by:

$$H(Y) = -\sum_{y} p(y)\log(p(y)) \tag{2}$$

The above Eq. (2) shows that the uncertainty in output Y, but suppose when we observe a variable X then the uncertainty in output Y decreases by the conditional entropy is given by Eq. (3):

$$H(Y \mid X) = -\sum_{x}\sum_{y} p(x,y)\log(p(y \mid x)) \tag{3}$$

The decrease in uncertainty is given as:

$$I(Y,X) = H(Y) - H(Y \mid X) \tag{4}$$

This Eq.(4) is used to measure dependency between the features which gives the Mutual Information between Y and X.If Eq.(4) produce the difference greater than zero that means X and Y are depend on each other ,and if the difference is zero that means they are not depend on each other. This indicates that one variable can provide information about the other variable thus it proves dependency among the variables.

VI. Experimental Work

In this present study we apply the three different feature ranking algorithm i.e. SD ,MI and ReliefF . for epileptic seizure prediction, these algorithm ranks the features by using statistical criteria and assign high values for relevant features and small for irrelevant ones, these three algorithm apply on 10 different EEG database of epileptic patient which ASPPR(advance seizure prediction via pre_ictal relabeling)algorithm[8] make publicly available in the form of feature vector of 204 features . In present work the same number of top rank feature vector is select from all three algorithm and passes to Artificial Neural Network ANN classifier for classification during that it is observe that average accuracy of classification using SD optimal set is 91.31% ,for MI 89.12% and for Relieff it is 85.07% .

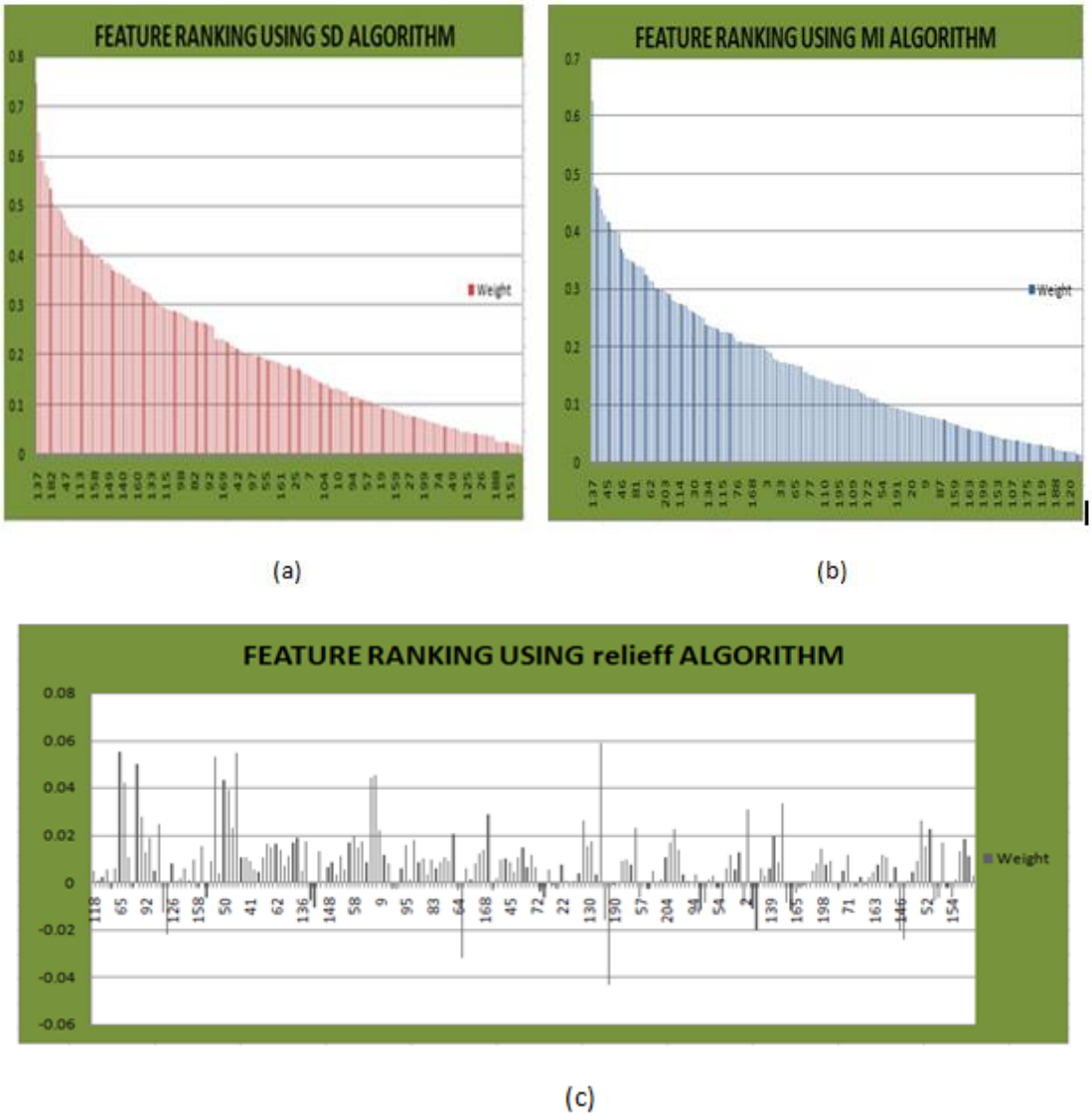

Fig. 1 Feature Ranking **(a)**SD Algorithm (b) MI Algorithm (c) Relieff Algorithm

Table 1. Results of Experiment

Sr.No.	Data set	No. of top rank feature	SD	MI	Relieff
1	Data1	17	98.0%	96.7%	97.6%
2	Data2	7	89.9%	95.1%	94.4%
3	Data3	11	78.5%	75.7%	90.2%
4	Data4	17	97.2%	94.7%	91.8%
5	Data5	08	94.6%	94.6%	84.0%
6	Data6	08	97.2%	94.3%	90.2%
7	Data7	08	93.0%	83.2%	91.6%
8	Data8	09	95.6%	91.9%	60.5%
9	Data9	09	92.9%	92%	84.6%
10	Dta10	10	76.1%	73%	65.8%

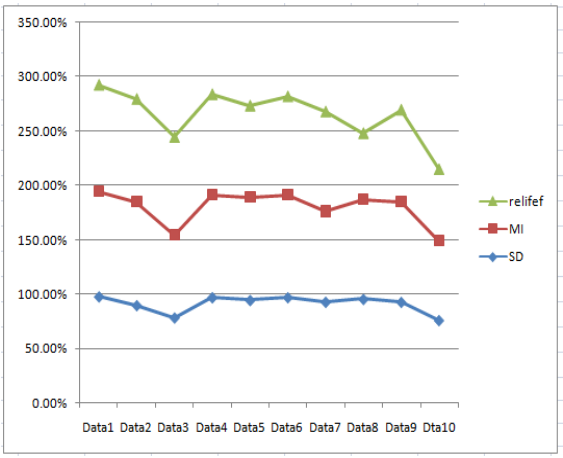

Fig.2 performance of classification for SD,MI and Relieff

VII. Conclusion

In the present work three feature ranking algorithms are used i.e. SD,MI and Relieff Algorithm for optimal feature selection for Epileptic seizure prediction from the experimental result it is observer that that average accuracy of classification using SD optimal set is 91.31% ,for MI 89.12% and for Relieff it is 85.07%.so for present study Statistical Dependency (SD) gives maximum classification accuracy than MI and Relieff .

References

[1] Pinar Yildirim" Filter Based Feature Selection Methods for Prediction of Risks in Hepatitis Disease" International Journal of Machine Learning and Computing, Vol. 5, No. 4, August 2015 DOI: 10.7763/IJMLC.2015.V5.517

[2] R. Kohavi, A study of cross-validation and bootstrap for accuracy estimation and model selection, in: C.S. Mellish, ed., Proceedings IJCAI-95, Montreal, Que. (Morgan Kaufmann, Los Altos, CA, 1995) 1137-l 143.

[3] T. M. Phuong, Z. Lin et R. B. Altman. Choosing SNPs using feature selectio. Proceedings / IEEE Computational Systems Bioinformatics Conference, CSB. IEEE Computational Systems Bioinformatics Conference, pages 301-309, 2005. PMID 16447987

[4] kenji kira **"**a practical approach to feature selection" ml92 proceedings of the ninth international workshop on machine learning pages 249-256

[5] S. Francisca Rosario, Dr. K. Thangadurai "RELIEF: Feature Selection Approach" international journal of innovative research & development Vol 4 Issue 11, 218-224

[6] D. Lakshmi Padmaja et al " Survey of Dimensionality Reduction and Mining Techniques on Scientific Data" International Journal of Computer Science & Engineering Technology (IJCSET) ISSN : 2229-3345 Vol. 5 No. 11 Nov 2014 1062-1066

[7] Chunyong Yin,Lu Feng, Luyu Ma, Zhichao Yin andJin Wang "A Feature Selection Algorithm based on Hoeffding Inequality and Mutual Information" International Journal of Signal Processing, Image Processing and Pattern Recognition Vol.8, No.11(2015), pp.433-444http://dx.doi.org/10.14257/ijsip.2015.8.11.39

[8] Moghim N, Corne DW (2014) Predicting Epileptic Seizures in Advance. PLoS ONE 9(6): e99334. doi:10.1371/journal.pone.00993