International Journal of Advanced Research in Engineering Technology & Sciences ISSN: 2394-2819
June-2016 Volume 3, Issue-6
Email:editor@ijarets.org www.ijarets.org

REVIEW ON CLASSIFICATION OF EEG SIGNAL IN DATA MINING

Sangeeta Sindhu
M. Tech Scholar
Department CSE
BPS University
KhanpurKalan
Sonipat

Mr. Vinod Saroha
Assistant professor
Department CSE
BPS University
KhanpurKalan
Sonipat

Abstract: **-** Data Mining is a field of search and researches of data. Mining the data means fetching out a piece of data from a huge data block. The basic work in the data mining can be categorized in two subsequent ways. One is called classification and the other is called clustering. Although both refers to some kind of same region but still there are differences in both the terms. In the presented research work, our aim is to find out the maximum number of clusters in a specified region by applying the area searching algorithms .Classification is always based on two things. a) the area which you choose for the classification that is the cluster region b) the kind of dataset which you are going to apply on the selected region.
KEYWORD*:*Data Mining, Classification, Recognition

I. INTRODUCTION

With the enormous amount of data stored in files, databases, and other repositories, it is increasingly important, if not necessary, to develop powerful means for analysis and perhaps interpretation of such data and for the extraction of interesting knowledge that could help in decision-making. Data Mining, also popularly known as Knowledge

Discovery in Databases (KDD), refers to the nontrivial extraction of implicit, previously unknown and potentially useful information from data in databases. Data mining refers to extracting or "mining" knowledge from large amounts of data. Classification (technique to analyses the *frequent itemsets*) is one of the major fields in the area of extracting knowledge from vast data. A *frequent itemset*typically refers to a set of items that frequently appear together in a transactional data set, such as milk and bread (Han &Kamber, 2001). In this chapter, we will briefly review about data mining, its architecture, functionalities, classification, methods of classification etc. We are in an age often referred to as the information age. In this information age, because we believe that information leads to power and success, and thanks to sophisticated technologies such as computers, satellites, etc., we have been collecting tremendous amounts of information. Initially, with the advent of computers and means for mass digital storage, we started collecting and storing all sorts of data, counting on the power of computers to help sort through this amalgam of information. Unfortunately, these massive collections of data stored on disparate structures very rapidly became overwhelming. This initial chaos has led to the creation of structured databases and database management systems (DBMS). The efficient database management systems have been very important assets for management of a large corpus of data and especially for effective and efficient retrieval of particular information from a large collection whenever needed. The proliferation of database management systems has also contributed to recent massive gathering of all sorts of information. Today, we have far more information than we can handle: from business transactions and scientific data, to satellite pictures, text reports and military intelligence. Information retrieval is simply not enough anymore for decision-making. Confronted with huge collections of data, we have now created new needs to help us make better managerial choices. These needs are automatic summarization of data, extraction of the "essence" of information stored, and the discovery of patterns in raw data.

.

II. IMPURITY MEASURES

A split that will separate the data as much as possible in accordance with the class labels So the objective is to obtain nodes that contain cases of a single class only as mentioned before. We define impurity as a function of the relative frequencies of the classes in that node's(t) = fi(p1, p2, ..., pJ) (1) with pj (j = 1, ..., J) as the relative frequencies of the J different classes in that node To compare all the possible splits of the data you have, a quality of a split as the reduction of impurity that the split achieves must be defined. In the example later on the following equation for the impurity reduction of split s in node t will be used: where fi(j) is the proportion of cases sent to branch j by s, and i(j) is the impurity of the node of branch j. Because different algorithms of tree construction use different impurity measures, we will discuss three of them and give a general example later on substitution error This is a measure for the impurity defined by the fraction of the cases in a node that is classified incorrectly if we assign every case to the majority class in that node: i(t) = 1 – max j p(j|t) (3) where p(j|t) is the relative frequency of class j in node t. The substitution error gives a score to a split according to the incorrectly classified cases in a node. It can recognize a better split if it has less error in that node. But the substitution error has one major disadvantage: it does not recognize a split as a better one if one of its resulting nodes is pure. So it does not prefer Split 2 over Split 1 in Figure 2. In such a case we want a split with a pure node to be preferred. Gini index The Gini index does recognize such a split. Its impurity measure is defined as follows: i(t) =X j p(j|t)(1 − p(j|t)) (4) Entropy Finally we have the entropy measure which is used in well-known classification tree algorithms like ID3 and C4.5. The advantage of the entropy measure over the gini-index is that it will reach minimum faster if more instances of the child nodes belong to the same class. Splits to consider we have looked at different impurity criteria for computing the quality of a split. In this section we look at which splits are considered and how we select the best split (for binary splits only). The attributes can have numerical or categorical values. In the case of numerical values, all the values of the attributes occurring in the training set are considered. The possible splits are made between two consecutive numerical values occurring in the training set. If the attribute is categorical with N categories, then 2N−1−1 splits are considered. There are 2N−2 non-empty proper subsets of a set of N elements. The empty set and the complete set do not count. Furthermore a split of the N categories into S and Sc

III. SURVEY

M. Ikonomakis, S. Kotsiantis, V. Tampakas [1] Automated text classification has been considered as a vital method to manage and process a vast amount of documents in digital forms that are widespread and continuously increasing. In general, text classification plays an important role in information extraction and summarization, text retrieval, and question answering.

This paper illustrates the text classification process using machine learning techniques. The references cited cover the major theoretical issues and guide the researcher to interesting research directions.

ICA was first applied to EEG by Makeig et al. in 1996 [2]. ICA separates the artefacts from the EEG signals into independent components based on the characteristics of the data without relying on the reference channels. The data in the recorded trails, each channel data and the frontal data are also preserved during the ICA artefact removal [3]. The ICA algorithm decomposes the multi-channel EEG data into temporal independent and spatial-fixed components. It is computationally efficient. ICA shows high performance when the size of the data to decompose is large [4]. ICA requires more computations to decompose signals [4] [5]. EEGLAB supports various types of ICA algorithms (nearly 20 algorithms) and most used algorithms are Joint Approximate Decomposition of Eigen matrices (JADE), fixed point ICA, Infomax [6].

This method removes the noise by subtracting the common activity from the position of interest. The common activity can be the noise present in the EEG signal [7]. The referencing methods are used to improve the Signal-to-Noise Ratio (SNR). The presence of the artefacts yields low SNR in EEG signals. In CAR method the removal of mean of all electrodes from all the electrodes results in noise free signals. The results in [8] show that CAR outperforms all referencing methods and shows best classification accuracy

results. Finite sample density and incomplete head coverage of EEG electrode arrays cause problems in calculating the averages in referencing methods [9].

An estimate of current density entering or leaving the scalp through the skull is referred to as the Surface Laplacian of the skull. It only considers the outer shape of the volume conductor and does not require any details of volume conduction [10]. Ocular movements can be efficiently eliminated during the signal acquisition. For large artifacts ranging from 50μV (>50μV) visual inspection is needed and by considering shape of the artifacts the gradients of activities are obtained [11]. Hjorth method offers good framework for theoretical explorations. SL is robust against artifacts generated at uncovered regions by the electrode cap and it solves the electrode reference problem [12]. SL is a way of viewing the EEG data with high spatial resolution. SL is sensitive to the choice of spline parameters during spline interpolation [13]. As SL is sensitive to artifacts, care has to be taken during the artifact removal [14].

PCA was invented in 1901 by Karl Pearson and later developed independently by Harold Hotelling in 1930 [15]. The PCA transforms the correlated vectors into linearly uncorrelated vectors. These uncorrelated vectors are called as "Principal Components" [16][17]. This is a classical method of Second Order Statistics. It depends on decomposition of covariance matrix. PCA helps in reduction of feature dimensions. Ranking will be done by using PCA based on the variability of the signal properties. This ranking helps in classification of the data. The application of PCA in a BCI system yields best classification results [18]. The PCA is well but it is not as well as ICA [19].

CSP was first presented by Koles and it can detect abnormal EEG activity. CSP performs transformation of EEG signal into a variance matrix that maximally discriminates between different classes. CSP uses spatial filtering and with spatial information it detects the patterns in EEG. CSP does not require a-priori selection of subject specific frequency bands and knowledge of these bands and requires use of many electrodes. It is sensitive to artefacts and electrode positions. During the training process the identical electrode positions is to be maintained to capture the same signals. The increase in accuracies may obsolete because of the change in electrode positions.

IV. CONCLUSION

Data Mining is a field of search and researches of data. Mining the data means fetching out a piece of data from a huge data block. The basic work in the data mining can be categorized in two subsequent ways. One is called classification and the other is called clustering. Although both refers to some kind of same region but still there are differences in both the terms. The classification of the data is only possible if you have modified and identified the clusters. In the presented research work, our aim is to find out the maximum number of clusters in a specified region by applying the area searching algorithms.

Classification is always based on two things:

a) The area which you choose for the classification that is the cluster region.

b)The kind of dataset which you are going to apply on the selected region .

REFERENCES

1. M. Ikonomakis, S. Kotsiantis, V. Tampakas, "Text Classification Using Machine Learning Techniques" WSEAS TRANSACTIONS on COMPUTERS, Issue 8, Volume 4, August 2005
2. Antara Bhattacharya, Dr. N. G. Bawane, S. M. Nirkhi, Brain Computer Interface Using EEG Signals, and Available at: http://www.schalklab.org/sites/default/files/misc/Abrain-computer interface using electrocorticographic signals in humans.pdf.
3. JosepMouriño, José del R. Millán, FeboCincotti, Silvia Chiappa, RaimonJané, Fabio Babiloni, "Spatial Filtering in the Training Process of A Brain Computer Interface", IEEE Engg, in Medicine and Biology Soc., Oct 2001.
4. Mohammed J. Alhaddad, "Common Average Reference (CAR) Improves P300 Speller", Int. J. of Engg. and Tech., Vol. 2 No. 3, 2012.
5. Thomas C Ferree and Ramesh Srinivasan, "Theory and Calculation of the Scalp Surface Laplacian", Electrical Geodesics, Inc., August 24, 2000.
6. Siyi Deng, William Winter, Samuel Thorpe, Ramesh Srinivasan, "EEG Surface Laplacian using realistic head geometry", Int. J. of Bioelectromagnetism, Vol. 13, No. 4, pp. 173-177, 2011.

7. C. Tandonnet, B. Burle, T. Hasbroucq, F. Vidal, "Spatial Enhancement of EEG traces by Surface Laplacian Estimation: Comparison between Local and Global Methods", Clinical Neurophysiology 116 (2005) 18–24.
8. Claudio G. Carvalhaes, Patrick Suppes, "A Spline Framework for Estimating the EEG Surface Laplacian Using the Euclidean Metric", Neural Computation 23, 2010
9. Principal Component Analysis, available at http://en.wikipedia.org/wiki/ Principal_component_analysis.
10. FengGu, Julie Greensmith, Robert Oates and UweAickelin, PCA 4 DCA: The Application Of Principal Component Analysis To The Dendritic Cell Algorithm, In 9th Annual Workshop on Computational Intell. (UKCI), 7-9 Sept, 2009.
11. Tomas Zeman, BSS – Pre-processing Steps for Separation Improvement, May 2000, Available at http://noel.feld.cvut.cz/vyu/prj-czs-asi/zeman/pca.pdf
12. Bruce A. Draper, KyungimBaek, Marian Stewart Bartlett, J. Ross Beveridge, Recognizing Faces with PCA and ICA, Computer Vision and Image Understanding 91, 2003, available at www.ComputerScienceWeb.com.
13. Dieter Devlaminck, BartWyns, Moritz Grosse-Wentrup, Georges Otte, and Patrick Santens, Multi subject Learning for Common Spatial Patterns in Motor-Imagery BCI, Computational Intell and Neuroscience Volume 2011.
14. Fabien Lotte and Cuntai Guan, Regularizing Common Spatial Patterns to Improve BCI Designs: Unified Theory and New Algorithms, IEEE Trans. On Biomedical Engg., Vol. 58, No. 2, 2011.
15. Guger C, Ramoser H and Pfurtscheller G, Real-time EEG analysis with subject-specific spatial patterns for a brain–computer interface (BCI), IEEE Trans. Rehabil. Eng, 2000.
16. Pfurtscheller G, Neuper C, Guger C, Harkam W, Ramoser H, Schlogl A, Obermaier B and Pregenzer M, Current trends in graz brain–computer interface (BCI) research, IEEE Trans. Rehabil, Eng. 8, 2000.
17. Ramoser H, Muller-Gerking J and Pfurtscheller G, Optimal spatial filtering of single trial EEG during imagined hand movement, IEEE Trans. Rehabil. Engg, 2000.
18. A Garcés Correa , E Laciar , H D Patiño , M E Valentinuzzi, Artifact removal from EEG signals using adaptive filters in cascade, 16th Argentine Bioengineering Cong. and the 5th Conf. of Clinical Engg., IOP Publishing J. of Physics: Conf. Series 90 (2007).
19. ChinmayChandrakar, M.K. Kowar, De-noising ECG Signals Using Adaptive Filter Algorithm, Int. J. of Soft Computing and Engg. (IJSCE), Vol. 2, Issue 1, 2012.