Neurocomputing 199 (2016) 66–76

Contents lists available at ScienceDirect

Neurocomputing

journal homepage: www.elsevier.com/locate/neucom

Classification of epilepsy seizure phase using interval type-2 fuzzy support vector machines

Udeme Ekong [a], H.K. Lam [a], Bo Xiao [a], Gaoxiang Ouyang [b], Hongbin Liu [a], Kit Yan Chan [c], Sai Ho Ling [d]

[a] Department of Informatics, King's College London, Strand, London WC2R 2LS, United Kingdom
[b] State Key Laboratory of Cognitive Neuroscience and Learning, School of Brain and Cognitive Sciences, Beijing Normal University, No. 19, XinJieKoWai St., HaiDian District, Beijing 100875, PR China
[c] Department of Electrical and Computer Engineering, Curtin University, Perth, Australia
[d] Centre for Health Technologies, Faculty of Engineering and Information Technology, University of Technology, Sydney, NSW, Australia

ARTICLE INFO

Article history:
Received 14 October 2015
Received in revised form
11 January 2016
Accepted 14 March 2016
Available online 25 March 2016

Keywords:
Classification
Epilepsy
Fuzzy support vector machine
Interval type-2 fuzzy sets

ABSTRACT

An interval type-2 fuzzy support vector machine (IT2FSVM) is proposed to solve a classification problem which aims to classify three epileptic seizure phases (seizure-free, pre-seizure and seizure) from the electroencephalogram (EEG) captured from patients with neurological disorder symptoms. The effectiveness of the IT2FSVM classifier is evaluated based on a set of EEG samples which are collected from 10 patients at Peking university hospital. The EEG samples for the three seizure phases were captured by the 112 2-s 19 channel EEG epochs, where each patient was extracted for each sample. Feature extraction was used to reduce the feature vector of the EEG samples to 45 elements and the EEG samples with the reduced features are used for training the IT2FSVM classifier. The classification results obtained by the IT2FSVM are compared with three traditional classifiers namely Support Vector Machine, k-Nearest Neighbor and naive Bayes. The experimental results show that the IT2FSVM classifier is able to achieve superior learning capabilities with respect to the uncontaminated samples when compared with the three classifiers. In order to validate the level of robustness of the IT2FSVM, the original EEG samples are contaminated with Gaussian white noise at levels of 0.05, 0.1, 0.2 and 0.5. The simulation results show that the IT2FSVM classifier outperforms the traditional classifiers under the original dataset and also shows a high level of robustness when compared to the traditional classifiers with white Gaussian noise applied to it.

1. Introduction

A classification problem can be best illustrated when an object or group of objects have to be assigned into a pre-defined group or class where the assignment is made based on a number of observed features/attributes pertaining to that particular object. Classification is a very important field of research due to the advantageous nature that a classifier with high generalization ability would benefit the economical, industrial and medical fields [1]. As a result of this, extensive research has been carried out over the years and this has resulted in a large number of applications such as risk classification of loan clients [2], hand-writing recognition [3], image classification [4] and speech recognition [5].

Literature review shows that classification methods can be categorized by four types namely logic based approach (e.g. decision trees) [2], statistical approach (e.g. Bayesian classification) [6], instance-based approach (e.g. nearest neighbor algorithm [7]) and machine learning (e.g. single layer perceptrons, neural networks [8,9] and support vector machine (SVM) [10]).

The decision tree method is carried out by categorizing the inputs based on the feature values in the input [7]. A drawback of this method is that once the splitting rule makes a wrong decision, it is impossible to produce the correct path and this would therefore generate an accumulation of errors. Bayesian classifier is based on the assumption that equal prior probabilities exist for all classes [6]. The main limitation of the Bayesian classifier is that the posterior probabilities cannot be determined directly [8]. An example of the instance-based method is the k-Nearest Neighbor

E-mail addresses: udeme.ekong@kcl.ac.uk (U. Ekong), hak-keung.lam@kcl.ac.uk (H.K. Lam), bo.xiao@kcl.ac.uk (B. Xiao), ouyang@bnu.edu.cn (G. Ouyang), hongbin.liu@kcl.ac.uk (H. Liu), kit.chan@curtin.edu.au (K.Y. Chan), steve.Ling@uts.edu.au (S.H. Ling).

http://dx.doi.org/10.1016/j.neucom.2016.03.033
0925-2312/

(kNN) [7] technique which is based on the principle that objects in a dataset generally exist in the neighborhood of other objects with similar properties. The technique finds the k nearest objects to the particular input and determines its class by identifying the most frequent class label.

The single layer perceptron can be simply described as a component that computes the sum of weighted inputs and then feeds to the system outputs. A major limitation of the single layer perceptron is that it can only learn linearly separable problems and is therefore incompatible when considering non-linear problems [9]. This problem is solved by the introduction of the Neural Network (NN). The Neural Network can be divided into 3 distinct segments: the input units which have the primary responsibility of receiving information; the hidden units which contain neurons to carry out the input–output mapping and the output units which store the processed results [7]. When the optimal connection weights and transfer functions are determined, the NN can be used as a universal approximator [11] which is able to approximate any continuous functions (e.g., hyperplanes) to any arbitrary precision in a compact domain.

The Support Vector Machine (SVM) was first proposed by Vapnik in 1995 [7] as a machine learning model which can be applied to various supervised and unsupervised learning applications [12–14]. The SVM approach can be redeveloped as Support Vector Classification (SVC) which are used for task such as pattern recognition and Support Vector Regression (SVR) which is mainly applicable to time-series applications [12]. The SVM uses the hyperplane to separate two data classes. The SVM attempts to maximize the margin between the hyperplane and the input samples which is being separated by it thereby reducing the generalization error. Data that is difficult to separate on the input space is mapped into a higher dimensional feature space for the ease of separation. The higher dimensional feature space computations are done with the use of a kernel function [7]. This feature illustrates a very important trait of the SVM which is its ability to perform well in a high dimensional feature space [15,16].

The SVM performs structural risk minimization (SRM) in order to find a trade-off between model complexity and generalization capability [7]. Therefore the SVM can achieve good generalization ability for classification problems as it can simultaneously minimize the empirical risk [10]. The SRM principle is grounded on the fact that the generalization error of the model is bounded by the sum of the empirical error and a confidence interval which is based on the Vapnik–Chervonenkis (VC) dimension [7], a higher classification performance is achieved by minimizing this bound. The SVM also provides a global optimization solution to the problem at hand and therefore provides a more credible output when compared to the neural network which provides a local optimization solution [10]. One of the drawbacks of the SVM method is its sensitivity to outliers, this stems from the fact that the same penalty weight is assigned to each sample and an outlier would significantly distort the representation of the input and therefore affect the classification performance. Another drawback is that when the SVM is applied to a classification problem with imbalanced dataset (i.e. negative samples significantly outweighs the positive samples) the optimum separating hyperplane is skewed towards the positive with the consequence that the SVM could be very ineffective in identifying targets that should be mapped to the positive class [12,15].

A relatively recent classification method is based on fuzzy logic [17] which is the theory of fuzzy sets used to handle fuzziness or imprecision in datasets. The approach attempts to assign each variable with membership functions with respect to its relative distance to the class [17,4]. There are two main types namely type-1 and type-2 fuzzy sets [18–20]. In type-1 fuzzy sets, the membership values are precise numbers in the range between 0 and 1 whilst the membership grade of a type-2 fuzzy set is a type-1 fuzzy set due to the imprecision in assigning a membership grade. As a result, type-2 fuzzy sets are effective in modeling higher level uncertainty in the human decision making process when compared to the type-1 fuzzy set, where the membership grade is distinct. In fuzzy logic, classification rules are specified by the user instead of being inherently decided upon by the machine learning method like in the SVM or NN. Therefore fuzzy logic is not a black-box method and the decision rules are clearly visible. Incorporating the mechanisms of fuzzy logic, NN and SVM, two hybrid machine learning methods namely Neural Fuzzy Network (NFN) and Fuzzy Support Vector Machine (FSVM) [13] were developed. The NFN works effectively when the amount of sample data provided is sufficient but suffers from a significantly reduced generalization performance when the amount of sample data is not enough. The FSVM however works effectively even when the amount of sample data is limited and is proven to provide higher generalization performance [13]. There are many complex systems used in industry that are prone to abrupt changes such as the random failure of components or sudden environmental disturbances. Markov jump systems (MJS) can be used to represent these systems [21–23]. In Markovian jump systems, each event governed by a Markov process corresponds to the jump in finite operation modes of practical systems. This method is used to estimate the probability of an object moving from one state to another. This is done by using observations in the historical data to estimate the probability of transition [24]. In the literature we also see fuzzy logic being applied Markovian jump systems [3–5]. Interval type-2 fuzzy logic systems can also be applied to deal with complex non-linear MJS [25].

When considering a real world application of the SVM, it is important to account for the difficulty in obtaining a precise measurement of the input data. A main deficiency of the SVM technique was its sensitivity to outliers and sample noise. This SVM deficiency is caused by the same penalty cost setting to each sample. The FSVM attempts to resolve this deficiency by assigning membership to each sample with respect to the relative importance of this sample. Hence, it reduces the impact of outliers in the input dataset [26].

The application being considered in this paper is the classification of the phases involved in the onset of an epileptic seizure, where the epilepsy signals obtained from the Electroencephalograph (EEG) using real clinical data is subjected to the novel classification technique [27,28]. This is a very challenging classification problem as the EEG has multiple features and is also contaminated with noise and distortion [29,30]. The classification technique is designed to differentiate between the 3 seizure phases namely seizure-free, pre-seizure and seizure. The early detection of seizure phases is a potentially life-saving application/research field and this is a major motivation for this research. The accurate classification/differentiation between the 3 seizure phases would give doctors and other healthcare professionals ample time to be able to prepare for the oncoming seizure. Therefore the main objective of the research carried out in this paper is to propose an adequate classifier to deal with this problem. As a result of this, an interval type-2 fuzzy support vector machine (IT2FSVM) is being proposed to deal with this problem. The IT2FSVM will be utilized to differentiate between the 3 seizure phases. The IT2FSVM is proposed due to its superior ability at dealing with uncertainties and unbalanced data [26]. This therefore provides a higher level of classification accuracy than the traditional SVM and forms the basis for the implementation of this classifier. The classification performance of the IT2FSVM technique will be compared to some traditional classifiers including the kNN technique [7], SVM [12] and naive Bayes classifier [6].