Neurocomputing 149 (2015) 1177–1187

Contents lists available at ScienceDirect

Neurocomputing

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

Variable weight neural networks and their applications on material surface and epilepsy seizure phase classifications

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

[a] Department of Informatics, King's College London, Strand, London WC2R 2LC, 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 5 May 2014
Received in revised form
21 July 2014
Accepted 6 September 2014
Communicated by Ligang Wu
Available online 16 September 2014

Keywords:
Material recognition
Epilepsy signals
Bayesian decision
KNN
Neural networks
Variable-weight neural networks

ABSTRACT

This paper presents a novel neural network having variable weights, which is able to improve its learning and generalisation capabilities, to deal with classification problems. The variable weight neural network (VWNN) allows its weights to be changed in operation according to the characteristic of the network inputs so that it can adapt to different characteristics of input data resulting in better performance compared with ordinary neural networks with fixed weights. The effectiveness of the VWNN is tested with the consideration of two real-life applications. The first application is on the classification of materials using the data collected by a robot finger with tactile sensors sliding along the surface of a given material. The second application considers the classification of seizure phases of epilepsy (seizure-free, pre-seizure and seizure phases) using real clinical data. Comparisons are performed with some traditional classification methods including neural network, k-nearest neighbours and naive Bayes classification techniques. It is shown that the VWNN classifier outperforms the traditional methods in terms of classification accuracy and robustness property when the input data is contaminated with noise.

1. Introduction

Classification is a process that takes samples from objects and assigns each one of them to a pre-defined group or class label. This is a promising and important field of research which provides a solution to a wide range of applications e.g., classification of different investments and lending opportunities as acceptable or unacceptable risk [1], classification of electrocardiogram (ECG) arrhythmias [2], classification of ECG beat [3], facial recognition [4,5], hand-writing recognition [6–10], heart sound classification [11], human body posture classification [12], speaker verification [13], speech recognition [14,15] and text classification [16,17].

In general, a classification process usually consists of three main stages. In the first stage, data from the objects have to be collected for the design of classifiers. In the second stage, feature extraction is performed to extract characteristics from the collected data to be classified such that redundant information is removed and representative information is extracted resulting in reduction of input dimensions and improved classification accuracy. In the third stage, a classifier is designed using the extracted feature data.

Existing methods from the literature review conducted include traditional and machine learning methods which are listed and briefly described below. An example of traditional methods is the cover linear discriminant analysis [18], logic based method (e.g., decision trees [19]), statistical approach (e.g., Bayesian classification [20]) and instance-based methods (e.g., nearest neighbour algorithm [21,22]). Machine learning methods include the support vector machines and neural network (NN) [4,9,13,17].

Bayesian decision theory [20] is one of the most important methods in statistical classification which offers a primary model for further classification procedures. Naive Bayesian classifier is based on the assumption that equal prior probabilities exists for all classes [23] this reduces the complexity of analysis and helps in resolving conflicts that occur when two or more classes are not well separable, resulting in improving the classification accuracy. Although Bayesian decision is simple and powerful, the posterior probabilities cannot be

* Corresponding author.
E-mail addresses: hak-keung.lam@kcl.ac.uk (H.K. Lam), udeme.ekong@kcl.ac.uk (U. Ekong), 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. Ho Ling).

http://dx.doi.org/10.1016/j.neucom.2014.09.011
0925-2312/

determined directly [24]. A recent development of the Bayesian classifier has been the proposal of a hybrid Bayesian classifier [23]. The Bayesian classifier has also been successfully implemented to many real-world applications e.g, weeds identification [20]. The k-nearest neighbour technique (kNN) [21] is easy to apply and good in dealing with text based problems such as visual category recognition [22]. However, kNN has its intrinsic limitations, the main disadvantages of the kNN technique are the large memory requirements and the lack of a logical way to choose the best 'k', this would introduce difficulties to a classification application as different data sets require an optimised value of 'k' to improve the performance of this method [25]. Furthermore, the precision accuracy of kNN will be declined when there are too many classes to deal with or when an uneven density of training samples is presented.

Machine learning methods include single layer perceptron [26], artificial neural networks [7,24,27], support vector machines, neural-fuzzy networks [2,3,8,12,15] and self-organisation map [6,11,16,28]. The first neural networks were designed based on mathematics and algorithms in the 1940 s. The McCulloch–Pitts neuron was proposed in 1943, this laid the foundation of modern neural networks [27]. However, there was no effective neural network training algorithm, so the development of neural networks was stagnated for some years. After that, a trainable network with adaptive elements, which are the building blocks, was designed [27]. A single layer perceptron [26] neural network model was first introduced by Rosenblatt in 1962, which cast a huge impact on the artificial intelligence field, and then different types of perceptron-based techniques have emerged in large numbers. A single layer perceptron has a simple structure which can be seen as a component that just weighted the inputs and then computes the sum to the output of the system. After that, the outputs are used to compare with the corresponding targets to verify the accuracy. With the information of difference between outputs and targets, the weights can be adjusted to achieve higher level of accuracy. However, there is a major limitation that restrains the applications, only linearly separable problems can be solved by the single layer perceptron. Although it has the major limitation, the single layer perceptron has also been implemented well to finger print matching [26] and image detection [29] applications. In addition to applications, the system stability of systems [30–32] invoking the NN are also investigated. Traditionally, a feedforward neural network [33] has three layers (input, hidden, and output layers) of nodes connected in a layer-to-layer manner. Neural Networks have various applications due to its favourable approximation performance and convenient modelling process. However, ordinary neural networks suffer from the 'over-fitting' and 'local optimisation' problems [34].

The Support Vector Machine (SVM) [35,36] is a kernel method that is used to map non-linear and inseparable data from an input space into a higher dimensional feature space where the data would then be linearly separable [37], this is done with the aid of the separating hyperplane [38]. The benefit of the kernel method is that the use of kernel functions enables the user to save time and computational power as the mapping is no longer compulsory [39]. The SVM algorithm aims to maximize the margin (the region separating the support vectors on either side of the hyperplane). This would result in an optimal classification accuracy of the hyperplane. Although the SVM sometimes suffers from high complexity and long computational times, it is shown to be very resistant to the problem of over-fitting the data. The SVM has a good generalisation ability and also performs well in a high dimensional feature space.

Neural-fuzzy network (NFN) involves the merging of fuzzy logic with the neural network. Neural networks are typically useful for non-linear mapping of inputs to outputs whilst fuzzy systems are designed based on the fuzzy set theory which processes data and has the ability to perform human-like reasoning when classifying data [36]. The method has been highly successful in applications due to the low-level learning and computational efficiency of neural networks coupled with the high-level human-like reasoning of fuzzy systems [40,41]. They use the well known backpropagation algorithm for the learning of the membership functions and fuzzy rules from the training data [36]. One limitation of using this algorithm is its inability to concurrently minimise the training and test error of the system. This therefore limits the attainable classification performance at the testing phase. The method involves the combination of multiple neural network classifiers with the aim that the fusion of their output would produce a higher classification accuracy compared to just utilising a single classifier [42].

Self-organising map (SOM) is an unsupervised competitive learning technique that was proposed by Kohonen [43] in 1982. In the SOM technique, neural maps transform data from a high-dimensional input space onto a lower dimensional output space [44] in a way that would preserve the architecture. Neighbouring neurons in the output space correspond to neighbouring data points in the input space [45]. The SOM technique is competent for dimensionality reduction and topology preservation [46]. This technique however utilises a fixed network architecture which mostly has to be defined by the user before the commencement of training and thus creates a dilemma concerning the size of the pre-defined output layer. If the fixed size is too small, the model is unable to express the input data effectively. If the output size is too large, the model would take a considerably long time to converge and also produce many redundant neuron units in the output layer [47]. The growing self-organising map (GSOM) was proposed as a solution to this problem. It provides a dynamic structure to the network model instead of the previously fixed structure used in the SOM [45,48].

Compared with traditional classification techniques, the machine learning approach in general uses a black-box approach. It demonstrates an appealing advantage as the designer does not need to know much about the problem due to the fact that the characteristic and information about the problem are obtained through learning algorithm. This is a prominent factor in the motivation to employ neural networks to deal with some classification problems. However, the traditional feed-forward neural networks demonstrate a drawback as the weights are fixed after training which limits the learning and generalisation capabilities. In handling a large amount of data in a large spacial domain, a sufficiently large size of neural network is required. In this paper, we consider a variable weight neural network (VWNN) to improve the learning and generalisation capability of the traditional neural networks. A VWNN consists of two traditional neural networks, namely tuning and tuned neural networks. The tuned neural network is the one which actually classifies the input data. The tuning neural network is to provide the weights to the tuned neural network according to the characteristic of the input data. Theoretically, the VWNN can be viewed as an infinite number of traditional neural networks with fixed weights. In the operation, the VWNN will use the best traditional neural network to handle the input data for classification.

Previous works relating to this study of the VWNN include [49] where the weight coefficients are obtained by means of an adaptive genetic algorithm and [50] where the number of hidden neurons in the network is varied based on the complexity of the input pattern to the system. A VWNN in [51] is applied to a handwriting pattern recognition problem, however, where the weights are predicted in only one layer in [51] whereas the weight is predicted over multiple layers in this paper. Also this research is applied to a material surface and epilepsy seizure phase classification, these are new applications of the VWNN technique when compared to the previous paper.

In this paper, we consider two real-life applications which are surface material recognition and epilepsy seizure phases recognition. In the application of surface material recognition, we develop classifiers to recognise the surface material of an unknown object