Journal of Computational and Applied Mathematics 309 (2017) 683–694

Contents lists available at ScienceDirect

Journal of Computational and Applied Mathematics

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

A neuro-fuzzy classification technique using dynamic clustering and GSS rule generation

Heisnam Rohen Singh *, Saroj Kr. Biswas, Biswajit Purkayastha

NIT, Silchar-788010, Assam, India

ARTICLE INFO

Article history:
Received 19 November 2015
Received in revised form 9 April 2016

Keywords:
Feature selection
Neuro-fuzzy
Dynamic clustering
Linguistic variables
Golden section search method

ABSTRACT

An efficient feature subset selection for predictive and accurate classification is highly desirable in many application domains like medical diagnosis, target marketing etc. Many neuro-fuzzy models were proposed for feature selection and efficient classification. One of such existing neuro-fuzzy models is Enhance Neuro-Fuzzy (ENF) system for classification using dynamic clustering. The major problem of ENF is, huge number of linguistic variables generated for each feature, which results in poor interpretation of the rules generated for classification. Therefore, this paper proposes a neuro-fuzzy model which is an extension of ENF. The novelty of the proposed model lies in determining less number of linguistic variables for each feature and also in generating significant linguistic variables in the rules for classification with better interpretation and accuracy. Six datasets are used to test the performance of the proposed model. 10-fold cross validation is used to compare the performance of the proposed model with others. It is observed from the experimental results that the performance of the proposed model is superior to others.

1. Introduction

Data-driven rule extractions are widely used in machine learning and data mining algorithms for classification, prediction and clustering. These algorithms operate on a huge amount of data with multiple dimensions to extract knowledge. Moreover, most of these data are insignificant to the specific domain. An important concept that helps in classification, clustering and a better understanding of the domain is feature selection [1]. Feature selection is a process of selecting a subset of features from a set of features without losing the characteristics and identity of the original object. There are two factors that affect feature selection: irrelevant features and redundant features. Irrelevant features are those which provide no useful information in a context and redundant features are those which provide no more information than the currently selected features.

Feature selection has been proved an inevitable part of a classifier through numerous researches. In the real-world scenario, to better represent the domain, many candidate features are introduced, which result the existence of irrelevant/redundant features to the target concept [2]. In many classification problems, due to the huge size of data, it is difficult to learn good classifiers before removing these unwanted features. Reducing the number of irrelevant/redundant features can drastically abate the running time of the learning algorithms and yields a more general classifier. Feature selection provides us with the advantages of facilitating data visualization and data understanding, reducing training and utilization times, reducing the measurement and storage requirements and defying the curse of dimensionality; which aids in the elevation of classification performance [1,3]. Hua et al. [4] have reported that feature selection is a part of the classification rule.

* Corresponding author. Tel.: +91 9402612120.
E-mail address: rohenheisnam87@gmail.com (H.R. Singh).

http://dx.doi.org/10.1016/j.cam.2016.04.023
0377-0427/

Feature selection can be done using various techniques like mutual information [5,6], genetic algorithm [6–8], bayesian network [9], artificial neural network (ANN) [10] etc. All these techniques have certain limitations. In mutual information technique, it is hard to calculate mutual information between the features that have continuous values, as it is often difficult to compute the integral in the continuous space based on a limited number of samples. In the bayesian network, the number of structures super-exponentially increases as number of features increases and in this more focus is made on the dependency of the features rather than the importance of features. In the genetic algorithm, some kind of randomness is involved and is very hard to assign importance to more significant features. Among these techniques, ANN is mostly used for feature selection and classification. It is well-known massively parallel computing model that exhibits excellent behavior in input–output mapping and in resolving complex artificial intelligence problems in classification tasks. However, ANN is a black box in nature that does not give any description of how the classification or the operation is done. Moreover, due to the presence of imprecise information, ambiguity or vagueness in input data, overlapping boundaries among classes and indefiniteness in defining features some uncertainties can arise at any stage of data classification task. The fuzzy logic [11–13] is very flexible in handling different aspects of uncertainties or incompleteness about real life situations. Both ANN and fuzzy logic are very adaptable in estimating the input–output relationships, in which ANN deals with numeric and quantitative data while fuzzy logic handles symbolic and qualitative data. Neuro-fuzzy hybridization leads to a crossbreed intelligent system widely known as Neuro-Fuzzy System (NFS) [14–17] that exploits the best qualities of these two approaches efficiently. NFS combines the advantages of both ANN and fuzzy logic, which covers up each other's disadvantages. In such system, the knowledge gained by the network from the linguistic interpretation of data can be used to generate rules that are used for feature selection as well as classification.

The linguistic rules generated by the neuro-fuzzy system are more helpful for understanding and analysis of the features. To generate these rules, the input features need to be labeled with some symbolic representation called linguistic variables. The knowledge extracted from the data is combined with the linguistic variables for rule-based classification.

The neuro-fuzzy schemes proposed for linguistic feature selection and rule-based classification in [18,19] are complicated as the structure of the network keeps on changing during the training phase. The neuro-fuzzy schemes [20,21] use fixed number of linguistic variables i.e. 3, {SMALL, MEDIUM, LARGE} for each feature. However fixing the equal number of linguistic variables for each feature is not a correct way of interpreting features. Therefore, the neuro-fuzzy model [22] has determined significant linguistic variables for each feature by dynamic clustering instead of fixing the number of linguistic variables. The model has established a criterion for dynamic clustering in such a way that generates a huge number of clusters which results the huge number of linguistic variables. But excessive clusters need unnecessary computational effort and provide poor interpretation of the rules. This problem of a huge number of linguistic variables is considered and lightly resolved in [23] by the fuzzy union and Golden section search (GSS), however, the accuracy of classification tasks drops using the fuzzy union. The proposed model resolves the same problem with a different approach. The proposed model uses a modified equation of the threshold in the dynamic clustering algorithm, that reduces the number of linguistic variables and the model also uses GSS to determine fixed significant number of linguistic variables rather than a flexible (different) number of linguistic variables in classification rules.

2. Related studies

Technique for classification tasks using neuro-fuzzy has been continually evolving to ensure efficient classification. There are many neuro-fuzzy techniques for feature selection and classification. Li et al. [24] have selected the important features and calculated the degree with which input pattern matches the memory vector using maximum fuzzy entropy interpretation. Kulkarni et al. [25] have computed feature wise membership of each pattern to its class which is useful in classification when the classes are overlapping and ill-defined. Basak et al. [26] have described a neuro-fuzzy methodology which involves connectionist minimization of a fuzzy feature evaluation index with unsupervised training. Yang et al. [27] have used one triaxial accelerometer to acquire subjects' acceleration data and train the neuro-fuzzy classifier to distinguish different activities/movements. Chen et al. [28] have proposed Quantum Neuro-Fuzzy Classifier (QNFC) model which combines the compensatory-based fuzzy reasoning method with the traditional Takagi–Sugeno–Kang (TSK) fuzzy model. The compensatory-based fuzzy reasoning method uses adaptive fuzzy operations of neuro-fuzzy systems that can make the fuzzy logic system more adaptive and effective. Azar et al. [29] have presented linguistic hedges neuro-fuzzy classifier with selected features (LHNFCSF) for dimensionality reduction, feature selection and classification. Chakraborty et al. [18,19] have integrated feature analysis and system identification which enables online feature selection and also builds a fuzzy rule-based classifier. Eiamkanitchat et al. [20,21] have developed a good classification model using less complicated rule for that. Wongchomphu et al. [22] have proposed a neuro-fuzzy system for classification using dynamic clustering, which is an extension of [20,21]. Napook et al. [23] have further extended [22] using adaptive dynamic clustering algorithm.

3. Proposed neuro-fuzzy model

The proposed neuro-fuzzy model is mainly divided into five main phases — preprocessing, transition, learning, linguistic selection and rule generation phases. The control flow diagram of the proposed model is given in Fig. 1. The details of each phase are given below.