

Chapter 65

Quantum–Behaved Particle Swarm Optimization Based Radial Basis Function Network for Classification of Clinical Datasets

N. Leema

College of Engineering Guindy, Anna University, Chennai, India

H. Khanna Nehemiah

College of Engineering Guindy, Anna University, Chennai, India

A. Kannan

College of Engineering Guindy, Anna University, Chennai, India

ABSTRACT

In this article, a classification framework that uses quantum-behaved particle swarm optimization neural network (QPSONN) classifiers for diagnosing a disease is discussed. The neural network used for classification is radial basis function neural network (RBFNN). For training the RBFNN K-means clustering algorithm and quantum-behaved particle swarm optimization (QPSO) algorithm has been used. The K-means clustering algorithm is used to find the optimal number of clusters which determines the number of neurons in the hidden layer. The cluster approximation error is used to find the optimal clusters. The weights between the hidden and the output layer is determined using QPSO algorithm based on the mean squared error (MSE). The performance of the developed classifier model has been tested with five clinical datasets, namely Pima Indian Diabetes, Hepatitis, Bupa Liver Disease, Wisconsin Breast Cancer and Cleveland Heart Disease were obtained from the University of California, Irvine (UCI) machine learning repository.

DOI: 10.4018/978-1-7998-2460-2.ch065

INTRODUCTION

Clinical decision making is a pervasive task that needs to be accurate. Knowledge models extracted from medical data should be novel, interesting and comprehensible to clinicians. Data mining techniques have been widely used in Computer Aided Diagnosis (CAD) systems, because of their ability to discover hidden patterns and relationships to solve the diagnostic problems in medical data (Sadoughi, Ghaderzadeh, Fein, & Standring, 2014). The common data mining tasks are association rule mining, clustering and classification (Han & Kamber, 2000). Association rule mining is a process of finding frequent patterns, associations, correlation among sets of items in transaction databases, relational databases and other information repositories. Clustering is a process of finding similarities between data according to the characteristics found in the data and grouping similar data objects into clusters. Classification is a supervised machine learning technique and is carried out in two steps namely learning and classification. In the learning step the classification model is constructed. In the classification step the constructed model is used to classify the test samples.

To address the problem of classification, the commonly used approaches are Artificial Neural Network (ANN), Decision trees, Support Vector Machine (SVM), Bayesian classifiers and K-Nearest Neighbor classifiers (Neelamegam & Ramaraj, 2013). Researchers have also developed hybrid classifiers to address the classification problem in clinical datasets and microarray gene expression data. A hybrid classifier for classifying leukemia gene expression data is presented in (Susmi, Nehemiah, Kannan, & Christopher, 2015). Mokeddem and Baghdad, (2016) performed an assessment for predicting coronary heart disease using fuzzy logic based clinical decision support system. Dasarathy and Sheela, (1979) introduced an ensemble based classification system. This system partitions the feature space into two or more classifiers. In 1990, Hansen and Salamon proved the generalization performance of a neural network that can be improved using an ensemble based neural network.

ANNs are computational models inspired by the biological nervous system (De Castro, 2006). Radial Basis Function (RBF) network is a feedforward ANN introduced by Broomhead and Lowe (Broomhead & Lowe, 1988). RBF network has several advantages when compared to the other types of ANNs, such as their simple structure, good approximation capabilities and faster learning rate (Qasem, Shamsuddin, Hashim, Darus, & Al-Shammari, 2013). Because of the above capabilities, the RBFNN can be applied in many science and engineering domains, such as function approximation, time series prediction, curve fitting, control and classification problems (Devaraj, Yegnanarayana, & Ramar, 2002; Du & Zhai, 2008; Fu & Wang, 2003; Han & Xi, 2004; Oyang, Hwang, Ou, Chen, & Chen, 2005).

RBFNN has three functional layers, namely input layer, hidden layer and output layer. The neurons in the input layer represent the input features. The hidden layer performs fixed nonlinear transformations of the input features and the connections from the input layer to hidden layer are not weighted. The RBFNN training is a two phase training process. In the first phase of training, number of neurons in the hidden layer is determined. In the second phase, the optimal weights between the hidden and the output layer is determined. In the first phase, finding the number of neurons in the hidden layer plays a major role in classification, as it affects the generalization and complexity of the network. If the number of neurons in the hidden layer is too high, an overfitting of the network may occur and if the number of neurons in the hidden layer is low underfitting of the network may occur. Training radial basis function neural network using particle swarms (Liu, Zheng, Shi, & Chen, 2004) achieves more rational architecture for RBFNN. Hien and Huan (2015) used the training function as multivariate regression function to train the RBFNN. Training function uses grid equally spaced nodes to find the number of neurons in the hid-

22 more pages are available in the full version of this document, which may be purchased using the "Add to Cart" button on the publisher's webpage:

www.igi-global.com/chapter/quantum-behaved-particle-swarm-optimization-based-radial-basis-function-network-for-classification-of-clinical-datasets/252082

Related Content

Spatial Modeling and Geovisualization of Rental Prices for Real Estate portals

Harald Schernthanner, Hartmut Asche, Julia Gonschorek and Lasse Scheele (2020). *Cognitive Analytics: Concepts, Methodologies, Tools, and Applications* (pp. 962-977).

www.irma-international.org/chapter/spatial-modeling-and-geovisualization-of-rental-prices-for-real-estate-portals/252065

Inquiry-Based Learning on the Cloud

Alexander Mikroyannidis, Alexandra Okada, Andre Correa and Peter Scott (2020). *Cognitive Analytics: Concepts, Methodologies, Tools, and Applications* (pp. 529-549).

www.irma-international.org/chapter/inquiry-based-learning-on-the-cloud/252042

Feature Selection Algorithms for Classification and Clustering

Arvind Kumar Tiwari (2020). *Cognitive Analytics: Concepts, Methodologies, Tools, and Applications* (pp. 422-442).

www.irma-international.org/chapter/feature-selection-algorithms-for-classification-and-clustering/252037

Designing Extreme Learning Machine Network Structure Based on Tolerance Rough Set

Han Ke (2020). *Cognitive Analytics: Concepts, Methodologies, Tools, and Applications* (pp. 263-282).

www.irma-international.org/chapter/designing-extreme-learning-machine-network-structure-based-on-tolerance-rough-set/252029

A Novel Bat Algorithm for Line-of-Sight Localization in Internet of Things and Wireless Sensor Network

Mihoubi Miloud, Rahmoun Abdellatif and Pascal Lorenz (2019). *Machine Learning and Cognitive Science Applications in Cyber Security* (pp. 213-239).

www.irma-international.org/chapter/a-novel-bat-algorithm-for-line-of-sight-localization-in-internet-of-things-and-wireless-sensor-network/227583