

Generating fuzzy rules for constructing interpretable classifier of diabetes disease

Nesma Settouti · M. Amine Chikh ·
Meryem Saidi

Received: 7 May 2011 / Accepted: 9 July 2012
© Australasian College of Physical Scientists and Engineers in Medicine 2012

Abstract Diabetes is a type of disease in which the body fails to regulate the amount of glucose necessary for the body. It does not allow the body to produce or properly use insulin. Diabetes has widespread fallout, with a large people affected by it in world. In this paper, we demonstrate that a fuzzy c-means-neuro-fuzzy rule-based classifier of diabetes disease with an acceptable interpretability is obtained. The accuracy of the classifier is measured by the number of correctly recognized diabetes record while its complexity is measured by the number of fuzzy rules extracted. Experimental results show that the proposed fuzzy classifier can achieve a good tradeoff between the accuracy and interpretability. Also the basic structure of the fuzzy rules which were automatically extracted from the UCI Machine learning database shows strong similarities to the rules applied by human experts. Results are compared to other approaches in the literature. The proposed approach gives more compact, interpretable and accurate classifier.

Keywords Interpretable classification · Fuzzy rules · FCM · Neuro-fuzzy ANFIS · UCI machine learning database

Introduction

Diabetes is a chronic disease that occurs when the pancreas does not produce enough insulin, or when the body cannot effectively use the insulin it produces. Hyperglycemia, or raised blood sugar, is a common effect of uncontrolled diabetes and over time leads to serious damage to many of the body's systems, especially the nerves and blood vessels [1]. Diabetes is the most rapidly growing chronic disease of our time (see Fig. 1). It has become an epidemic that affects a large number of people in the world. The two main kinds of diabetes are: Type 1 diabetes, which is characterized by a lack of insulin production and therefore dependent upon insulin administration and Type 2 diabetes, which is characterized by an ineffective use of insulin. Type 1 diabetes is the most common type in young people whereas Type 2 is the most common diabetes and affects primarily but not exclusively adults and it is largely the result of obesity and physical inactivity [2]. Our work will take an interest more specifically in Type 2 of diabetes.

Diagnose of diabetes for medical expert is a difficult task. For this reason a much research effort has been put till today in diagnosis of diabetes disease literature. In Temurtas et al. [4] a multi-layer neural network and a probabilistic neural network were used for diagnosing Pima Indian diabetes. They have reported, respectively 79.62 and 78.05 % in terms of correct classification rate. The recognition rate obtained by Kayaer and Yildirim [5] using general regression neural network was 80.21 %, while using multilayer neural network with LM algorithm was 77.08 % with tenfold CV and 82.37 % accuracy with conventional (one training and one test) validation method. They have also reported 78.05 %.

Purnami et al. [6] have reported better performance in classifying diabetes disease diagnostic with accuracy

N. Settouti (✉) · M. A. Chikh · M. Saidi
Biomedical Engineering Laboratory, Tlemcen University,
Tlemcen, Algeria
e-mail: nesma.settouti@mail.univ-tlemcen.dz

M. A. Chikh
e-mail: mea_chikh@mail.univ-tlemcen.dz

M. Saidi
e-mail: meryem.saidi@mail.univ-tlemcen.dz