

# Diagnosis of Diabetes Diseases Using an Artificial Immune Recognition System<sup>2</sup> (AIRS2) with Fuzzy K-nearest Neighbor

Mohamed Amine Chikh · Meryem Saidi ·  
Nesma Settouti

Received: 24 March 2011 / Accepted: 6 June 2011 / Published online: 22 June 2011  
© Springer Science+Business Media, LLC 2011

**Abstract** The use of expert systems and artificial intelligence techniques in disease diagnosis has been increasing gradually. Artificial Immune Recognition System (AIRS) is one of the methods used in medical classification problems. AIRS2 is a more efficient version of the AIRS algorithm. In this paper, we used a modified AIRS2 called MAIRS2 where we replace the K- nearest neighbors algorithm with the fuzzy K-nearest neighbors to improve the diagnostic accuracy of diabetes diseases. The diabetes disease dataset used in our work is retrieved from UCI machine learning repository. The performances of the AIRS2 and MAIRS2 are evaluated regarding classification accuracy, sensitivity and specificity values. The highest classification accuracy obtained when applying the AIRS2 and MAIRS2 using 10-fold cross-validation was, respectively 82.69% and 89.10%.

**Keywords** Pima Indians diabetes data set · Diagnosis · AIRS2 · Fuzzy k- nearest neighbors

## Introduction

Diabetes is a chronic illness that requires continuing medical care and patient self-management education to prevent acute complications and to reduce the risk of long-

term complications. People develop diabetes because the pancreas does not make enough insulin or because the cells do not use insulin properly, or both. As a result, the amount of glucose in the blood increases while the cells are starved of energy. Insulin is a hormone, made in the pancreas by beta cells, that regulates blood sugar. Over the years, high blood glucose, also called hyperglycemia, damages nerves and blood vessels, which can lead to complications such as heart disease, stroke, kidney disease, blindness, nerve problems, gum infections, and amputation. The classification of diabetes includes two main clinical classes: Type 1 diabetes, which results from Beta-cell destruction, usually leading to absolute insulin deficiency, Type 2 diabetes, which results from a progressive insulin secretory defect on the background of insulin resistance [1].

A medical diagnosis is a classification process. Using the computer science to perform this classification is becoming more frequent. Even if the expert's decision is the most important factor in diagnosis, classification systems provide substantial help as they reduce the errors due to fatigue and the time needed for diagnosis. There is a great variety of methods related to classification and diagnosis of diabetes disease in literature. In [2], a principal component analysis and adaptive neuro-fuzzy inference were used for diagnosing Pima Indian diabetes. They have reported 89.47% classification accuracy. Purnami et al. [3], obtained 93.2% classification accuracies using a new smooth support vector machine and its applications in diabetes disease diagnosis. In [4], Polat et al. used attribute weighted artificial immune System with 10-fold cross validation method; they obtained a classification accuracy of 75.87%. Ster and Dobnikar [6] accomplished 77.5%, 76.6%, 76.5%, 76.4%, 75.8% and 75.8% using respectively linear discriminant analysis, ASI, Fisher discriminant analysis, MLP+BP, LVQ and LFC with 10-fold cross validation.

---

M. A. Chikh · M. Saidi (✉) · N. Settouti  
Biomedical Engineering Laboratory, Tlemcen University,  
Tlemcen, Algeria  
e-mail: miryem.saidi@gmail.com

M. A. Chikh  
e-mail: mea\_chikh@mail.univ-tlemcen.dz

N. Settouti  
e-mail: nesma.settouti@gmail.com