Breast cancer and liver disorders classification using artificial immune recognition system (AIRS) with performance evaluation by fuzzy resource allocation mechanism

Kemal Polat a,*, Seral Şahan a, Halife Kodaz b, Salih Günes a

a Selcuk University, Engineering and Architecture Faculty, Electrical and Electronics Engineering, 42075 Konya, Turkey
b Selcuk University, Computer Engineering, Konya, Turkey

Abstract

Artificial Immune Recognition System (AIRS) classification algorithm, which has an important place among classification algorithms in the field of Artificial Immune Systems, has showed an effective and intriguing performance on the problems it was applied. AIRS was previously applied to some medical classification problems including Breast Cancer, Cleveland Heart Disease, Diabetes and it obtained very satisfactory results. So, AIRS proved to be an efficient artificial intelligence technique in medical field. In this study, the resource allocation mechanism of AIRS was changed with a new one determined by Fuzzy-Logic. This system, named as Fuzzy-AIRS was used as a classifier in the diagnosis of Breast Cancer and Liver Disorders, which are of great importance in medicine. The classifications of Breast Cancer and BUPA Liver Disorders datasets taken from University of California at Irvine (UCI) Machine Learning Repository were done using 10-fold cross-validation method. Reached classification accuracies were evaluated by comparing them with reported classifiers in UCI web site in addition to other systems that are applied to the related problems. Also, the obtained classification performances were compared with AIRS with regard to the classification accuracy, number of resources and classification time. Fuzzy-AIRS, which reached to classification accuracy of 98.51% for breast cancer, classified the Liver Disorders dataset with 83.36% accuracy. For both datasets, Fuzzy-AIRS obtained the highest classification accuracy according to the UCI web site. Beside of this success, Fuzzy-AIRS gained an important advantage over the AIRS by means of classification time. In the experiments, it was seen that the classification time in Fuzzy-AIRS was reduced about 70% of AIRS for both datasets. By reducing classification time as well as obtaining high classification accuracies in the applied datasets, Fuzzy-AIRS classifier proved that it could be used as an effective classifier for medical problems.

Keywords: Fuzzy resource allocation; AIRS; Breast Cancer dataset; Liver Disorders dataset; k-Fold cross-validation

1. Introduction

The use of classifier systems in medical diagnosis is increasing gradually. There is no doubt that evaluation of data taken from patient and decisions of experts are the most important factors in diagnosis. But, expert systems and different artificial intelligence techniques for classification also help experts in a great deal. Classification systems, helping possible errors that can be done because of fatigued or inexperienced expert to be minimized, provide medical data to be examined in shorter time and more detailed.

Breast Cancer is a very common cancer type among women. Today, innovations in cancer treatment have caused higher survival rates in cancer so in Breast Cancer. Especially, early diagnosis can increase these survival rates at considerable amount. Liver Disorders is also an important disease in medicine. Levels of enzymes mixed to blood are analysed in Liver Disorders diagnosis. There can be lots
of possible errors in this diagnosis due to the number of enzymes to be many as well as the effects of different taken alcohol rates to be vary from one patient to the other (Yalcın & Yıldırım, 2003).

While a new artificial intelligence field named as Artificial Immune Systems (AIS) was emerging in late 1990s, performances of proposed methods were not so good especially for classification problems. However, Artificial Immune Recognition System (AIRS) proposed in 2001 has changed this situation by taking attention among other classifiers with its performance (Watkins, 2001). The reason of its success in classification problems can be found in the following properties of it (Goodman, Boggess, & Watkins, 2003):

• AIRS performs good on very different problems such as large dimensioned feature space problems, problems with many classes, . . . etc.
• AIRS is self-adjusting with regard to its architecture in problem space.

In this study, resource allocation of AIRS was changed with its equivalence formed with Fuzzy-Logic to increase its classification performance by means of resource number and classification time more than classification accuracy. The effects of this change were analysed in the applications using medical datasets and obtained classification accuracies were compared with other classifiers used for same datasets. Fuzzy-AIRS has showed a further performance than expected and obtained the highest classification accuracy among the classifiers reported in UCI web site on these medical datasets consisting of Breast Cancer and Liver Disorders taken from UCI database (ftp://ftp.ics.uci.edu/pub/machine-learning-databases, 2005). Fuzzy-AIRS, which proved it self to be used as an effective classifier in medical field by reaching its goal, has also provided a considerable decrease in the number of resources. In all applications conducted, Fuzzy-AIRS required less resource than half of required by AIRS and by this way, classification time has reduced in a great rate.

The rest of the paper is organized as follows. Section 2 gives the background information including natural and artificial immune systems, AIRS, introduction of Breast Cancer and Liver Disorders in brief and previous research in literature. We explained the method in Section 3 with subtitles of Fuzzy Resource Allocation, Data Source, Used Parameters and Measures for Performance Evaluation. In each subsection of this section, the detailed information was given. The results obtained in applications are given in Section 4 both for Breast Cancer and Liver Disorders dataset. Section 5 includes the discussion of these results in specific and general manner. Consequently in Section 6, we conclude the paper with summarization of results by emphasizing the importance of this study and mentioning about some future work.

2. Background

As in other artificial intelligence techniques, AIS has emerged to design problem solving algorithms with high performance. AIRS is one of the classification algorithms that proposed in this area but it has taken a great attention in a short time. It has reached to high classification accuracies both in machine-learning benchmarks and real-world problems including medical data. This is the reason why this classification system was applied to two important medical classification problems after some improvements in this study.

2.1. Natural and artificial immune systems

The natural immune system is a distributed novel-pattern detection system with several functional components positioned in strategic locations throughout the body. Immune system regulates defence mechanism of the body by means of innate and adaptive immune responses. Between these, adaptive immune response is much more important for us because it contains metaphors like recognition, memory acquisition, diversity, self-regulation . . . etc. The main architects of adaptive immune response are lymphocytes, which divide into two classes as T and B Lymphocytes (cells), each having its own function. Especially B cells have a great importance because of their secreted antibodies (Abs) that take very critical roles in adaptive immune response.

The simplified working procedure of our immune system is illustrated in Fig. 1. Specialized Antigen Presenting Cells (APCs) called Macrophages circulates through the body and if they encounter an Antigen, they ingest and fragment them into antigenic peptides (I). The pieces of these peptides are displayed on the cell surface by Major Histocompatibility Complex (MHC) molecules existing in the digesting APC. The presented MHC-peptide combination on the cell surface is recognized by the T cells causing them to be activated (II). Activated T cells secrete some chemicals as alert signals to other units in response to this recognition. B cells, one of the units that take these signals from the T cells become activated with the recognition of Antigen by their Antibodies occurred in the same time (IV). When activated, B cells turn into plasma cells that secrete bound Antibodies on their surfaces (V). Secreted Antibodies bind the existing Antigens and neutralize them signaling other components of immune system to destruct the Antigen–Antibody complex (VI) (De Castro & Timmis, 2002). For detailed information about immune system refer to Abbas and Lichtman (2003).

Artificial Immune Systems emerged in the 1990s as a new computational research area. Artificial Immune Systems link several emerging computational fields inspired by biological behaviour such as Artificial Neural Networks and Artificial Life.

In the studies conducted in the field of AIS, B cell modelling is the most encountered representation type.
Different representation methods have been proposed in that modelling. Among these, shape-space representation is the most commonly used one (Perelson & Oster, 1979).

The shape-space model \( S \) aims at quantitatively describing the interactions among antigens (Ags), the foreign elements that enter the body like microbe...etc., and antibodies (Ag–Ab). The set of features that characterize a molecule is called its generalized shape. The Ag–Ab representation (binary or real-valued) determines a distance measure to be used to calculate the degree of interaction between these molecules. Mathematically, the generalized shape of a molecule \( m \), either an antibody or an antigen, can be represented by a set of coordinates \( m = (m_1, m_2, \ldots, m_L) \), which can be regarded as a point in an \( L \)-dimensional real-valued shape-space \( m \in S^L \). In this work, we used real strings to represent the molecules. Antigens and antibodies were considered of same length \( L \). The length and cell representation depend upon the problem (De Castro & Timmis, 2002).

### 2.2. AIRS classification algorithm

AIRS is a resource limited supervised learning algorithm inspired from immune metaphors. In this algorithm, the used immune mechanisms are resource competition, clonal selection, affinity maturation and memory cell formation. The feature vectors presented for training and test are named as Antigens while the system units are called as B cells. Similar B cells are represented with Artificial Recognition Balls (ARBs) and these ARBs compete with each other for a fixed resource number. This provides ARBs, which have higher affinities to the training Antigen to improve. The memory cells formed after the whole training Antigens were presented are used to classify test Antigens. The algorithm is composed of four main stages, which are initialization, memory cell identification and ARB generation, competition for resources and development of a candidate memory cell and lastly memory cell introduction. The flow chart of the algorithm is shown in Fig. 2.

#### 2.2.1. Initialization

The first step of the algorithm is a data pre-processing stage. Firstly, all of the data are normalized to ensure that the Euclidean distance between two data is in the interval of \([0–1]\). Euclidean distance is used for both affinity and stimulation value calculations (Eq. (1)):

\[
\text{Euclidean distance} = \sqrt{\sum_{i=1}^{n} (x_i - y_i)^2}
\]
where $x$ and $y$ refer feature vectors while $n$ is the number of attributes in data.

### 2.2.2. Memory cell identification and ARB generation

In this step, the algorithm begins to iterate for each training antigen. Training antigen is presented to memory cells and the most stimulated memory cell by that antigen is cloned. The stimulation levels are calculated by Eq. (2). All of the clones with memory cell are added to ARB pool. Here, the number of clones is determined according to the affinity between memory cell and antigen. Calculation of affinity values is done as in Eq. (3), which results higher affinities for lower Euclidean distances:

\[
\text{Stimulation}(x, y) = \begin{cases} 
  \text{affinity}(x, y), & \text{if class of } x = \text{class of } y \\
  1 - \text{affinity}(x, y), & \text{otherwise} 
\end{cases} 
\]  
(2)

\[
\text{affinity}(x, y) = 1 - \text{Euclidean distance}(x, y) 
= 1 - \sqrt{\sum_{i=1}^{n} (x_i - y_i)^2} 
\]  
(3)

#### 2.2.3. Competition for resources and development of a candidate memory cell

After above processes, the training antigen is presented to all ARBs in ARB pool. All of the ARBs are awarded so that the rewards of ARBs in the same class with presented antigen are given according to higher affinity values than the ARBs in different classes. Here the rewards are resource numbers. The required number of resources can exceed the allowed number by the system. In this case, excess resources are removed beginning with lower affinity ARBs and this continues until the required number is equal to the allowed number of resources. The stimulation levels of remaining ARBs are tested and the average value of these levels is determined for each class. If any of these average values is lower than a stimulation threshold determined by user, the ARBs belonging to that class are mutated and resulted clones are added to ARB pool. This step proceeds until average stimulation of all classes is bigger than stimulation threshold. Eq. (4) shows the formula for calculation of average stimulation value for each class:

\[
s_i = \frac{\sum_{j=1}^{\left| \text{ARB}_i \right|} \text{arb}_j \text{stim}_j}{\left| \text{ARB}_i \right|}, \quad \text{arb}_j \in \text{ARB}_i 
\]  
(4)

where $i = 1, \ldots, nc$, $s = \{s_1, s_2, \ldots, s_{nc}\}$, $|\text{ARB}_i|$ is the number of ARBs belonging to $i$th class and $\text{arb}_j \text{stim}$ is the stimulation level of $j$th ARB of $i$th class.

### 2.2.4. Memory cell introduction

After the total stimulation value of ARBs in all classes reaches stimulation threshold, the best ARB in the same class with training antigen is taken as a candidate memory cell. Here, best is synonym of having the highest affinity. If the stimulation value between training antigen and this candidate memory cell is bigger than the stimulation value between training antigen and original memory cell selected for cloning in step 2, the candidate memory cell is added to the memory cell pool.

These steps are repeated for each training antigen. After training, test data are presented only to memory cells. $k$-Nearest neighbour algorithm is used in determination of classes in test phase. For detailed information about AIRS, the reader is referred to Watkins (2001).

### 2.3. Used medical data: Breast Cancer and Liver Disorders

It was reported in Delen, Walker, and Kadam (2005) and http://medlineplus.gov/ (2005) that Breast cancer is most frequent disease in women in United States and therefore it is very important concern in medical field of United States. Cancer begins with uncontrolled division of one cell and results in a visible mass named Tumour. Tumour can be benign or malignant. Malignant Tumour grows rapidly and invades its surrounding tissues causing their damage. Breast cancer, is a malignant tissue beginning to grow in...
the breast. The abnormalities like existence of a breast mass, change in shape and dimension of breast, differences in the colour of breast skin, breast aches, etc. are the symptoms of breast cancer. Cancer diagnosis is performed based on the non-molecular criterions like tissue type, pathological properties and clinical location (Kıyan & Yıldırım, 2003). As for the other cancer types early diagnosis in Breast cancer can be life saving.

Liver is an effective organ in neutralizing toxics and throwing them from the body. If the amount of toxics reaches a level exceeding working capacity of the organ, the cells of related parts in organ are destroyed. Then, some substances and enzymes are appeared and interfere in blood. During diagnosis of the disease, the levels of these enzymes are analysed. Because of the fact that effects of different alcohol dosages vary from one person to the other as well as the fact that there are many enzymes, there can be frequent possible errors in diagnosis (Yağçın & Yıldırım, http://medlineplus.gov/, 2005).


2.4. Previous research

2.4.1. Research on AIRS

As mentioned in Section 2.2, AIRS is a classifier based on principles of resource-limited artificial immune systems, which was proposed by Watkins (2001). In this thesis work AIRS was applied a serial of problems including both machine-learning benchmarks and real-world problems. In some of these AIRS’s performance was very good and this study has brought related works after then.

In the study of Donald E. Goodman et al., they applied AIRS to multi-class problems and compared it with Kohonen’s LVQ (Goodman, Boggess, & Watkins, 2002). In most of the applied problems, they found that AIRS performed better than LVO and Optimized LVQ. Gaurav Marwah and Lois Boggess tried to do some modifications on AIRS for resource allocation and approaches to ARB pool organization (Marwah & Boggess, 2002). They also explored several different algorithms for tie breaking which could increase the accuracy of AIRS and other k-nearest neighbour classifiers. In the Study of Donald E. Goodman et al., AIRS was examined empirically, replacing one of the two likely sources of its classification power with alternative modifications (Goodman et al., 2003). They concluded with the modifications provided fast test versions of AIRS for users to experiment with. Besides of these studies, one more study has conducted recently. Hamaker and Boggess analysed the effects of adding non-Euclidean distance measures to the basic AIRS algorithm (Hamaker & Boggess, 2004). They used iris, Wisconsin Breast Cancer, Cleveland Heart disease and credit screening (Crx) datasets in their experiments.

2.4.2. Research on Breast Cancer and Liver Disorders classification

As for the other clinical diagnosis problems, classification systems have been used for breast cancer diagnosis problem, too. When the studies in the literature related with this classification application are examined, it can be seen that a great variety of methods were used which reached high classification accuracies. Among these, Quinlan reached 94.74% classification accuracy using 10-fold cross-validation (10 × CV) with C4.5 method (Quinlan, 1996). Hamilton et al., obtained 96% accuracy with RIAc method (10 × CV) (Hamilton, Shan, & Cercone, 1996) while Ster and Dobnikar obtained 96.8% with Linear Discriminant Analysis (LDA) method (10 × CV) (Ster & Dobnikar, 1996). The accuracy obtained by Bennett and Blue who used SVM method was 97.2% (5 × CV) (Bennet & Blue, 1997) while by Nauck and Kruse was 95.06% with neuro-fuzzy techniques (10 × CV) (Nauck & Kruse, 1999) and by Pena-Reyes and Sipper was 97.51% using Fuzzy-GA method (train: 75%–test: 25%) (Pena-Reyes & Sipper, 1999). Moreover, Setiono was reached 98.1% by using neuro-rule method (train: 50%–test: 50%) (Setiono, 2000). Goodman et al. applied three different methods to the problem which were resulted with the following accuracies: Optimized-LVQ method’s performance was 96.7%, big-LVQ method reached 96.8% and the last method, AIRS which he proposed depending on the Artificial Immune System, obtained 97.2% classification accuracy (10 × CV) (Goodman et al., 2002). Nevertheless, Abonyi and Szefírt applied Supervised Fuzzy Clustering (SFC) technique and obtained 95.57% accuracy (10 × CV) (Abonyi & Szefírt, 2003).

Like Breast Cancer, there are many studies for classification of Liver Disorders, too. Newton Cheung used some methods for this problem (Cheung, 2001). He obtained 65.50% classification accuracy using C4.5 (10 × CV), 63.39% using Naive Bayes classifier (10 × CV), 61.83% using Bayesian Network with Naive Dependence (BNND) classifier (10 × CV) and 61.42% using Bayesian Network with Naive Dependence and Feature Selection (BNNF) classifier (10 × CV). Tony Van Gestel et al. reached 69.20% classification accuracy with Support Vector Machine (SVM) classifier (10 × CV) (Van Gestel et al., 2002). The two methods that were used by Yuh-Jye and Mangarissan were Smooth Support Vector Machine (SSVM) classifier (10 × CV) (Lee & Mangasarian, 2001a) and Reduced Support Vector Machines (RSM) classifier (10 × CV) (Lee & Mangasarian, 2001b). They obtained 70.33% and 74.86% classification accuracies with these methods respectively. The classification accuracy obtained by Pham et al., using RULES-4 algorithm was 55.90% (train: 40%–test: 60%) (Pham, Dimov, & Salem, 2000). Beside of these studies, Yağçın and Yıldırım used some Neural Network architectures for this problem (Yağçın & Yıldırım, 2003). The classification accuracy obtained with Multilayer Perceptron (MLP) was 73.05%, 42.03% with Probabilistic Neural Networks (PNN), 65.55% with
Generalized Regression Neural Network (GRNN) and 58.55% with Radial Basis Function (RBF) (3 × CV).

3. Method

3.1. Fuzzy resource allocation

The competition of resources in AIRS allows high-affinity ARBs to improve. According to this resource allocation mechanism, half of resources is allocated to the ARBs in the class of Antigen while the remaining half is distributed to the other classes. The distribution of resources is done according to a number that is found by multiplying stimulation rate with clonal rate. In the study of Baurav Marwah and Lois Boggess, a different resource allocation mechanism was tried (Marwah & Boggess, 2002). In their mechanism, the Ag classes occurring more frequently get more resources. Both in classical AIRS and the study of Marwah and Boggess, resource allocation is done linearly with affinities. This linearity requires excess resource usage in the system, which results long classification time and high number of memory cells.

In this study, to get rid of this problem, resource allocation mechanism was done with fuzzy-logic. So there existed a non-linearity because of fuzzy-rules. The difference in resource number between high-affinity ARBs and low-affinity ARBs is bigger in this method than in classical approach.

The input variable of Fuzzy resource allocation mechanism is stimulation level of ARB hence the output variable is the number of resources, which will be allocated to that ARB. As for the other fuzzy-systems, input membership functions as well as output membership functions were formed. The input membership functions are shown in Fig. 3a.

The input variable, ARB.stim, varies between 0 and 1. A membership value is calculated according to this value using input membership functions. In this calculation, two points are get which are the cutting points of membership triangles by the input value, ARB.stim. Also these points are named as membership values of input variable for related membership function. The minimum of these points is taken as the membership value of input variable x, ARB.stim (Eq. (5)):

\[
\mu_{A:B}(x) = \min(\mu_A(x), \mu_B(x)), \quad x \in x
\]

here in Eq. (5), \( \mu_A(x) \) is the membership value of x in A and \( \mu_B(x) \) is the membership value of x in B, where A and B are the fuzzy sets in universe X. The calculated input

![Fig. 3. (a) Input membership function and (b) output membership function.](image-url)
membership value is used to get the output value through output membership functions, which are shown in Fig. 3b. In the $x$-axis of Fig. 3b, allocated resource number that will be calculated using the membership functions for the ARB is shown which changes between 0 and 10. The weight in the $y$-axis, which is the input membership value get as explained above, intersects the membership triangles at several points. The rule base for Fuzzy Resource Allocation is seen in Fig. 4.

Here VS, S, MS...etc. are the labels of input membership triangles and VS', S', MS'...etc. are the labels of output membership values. The rules in Fig. 4 define which points will be taken to average. For example if the input value cuts the triangles VS and S among the input membership functions, then the points to be averaged will be only the ones of VS' and S' triangles in the output membership functions.

Whereas determining membership value and getting output value using fuzzy-rules are of crucial importance, another important point is determination of linguistic values used in the input and output membership functions, which are shown in Table 1.

These linguistic values were determined in such a manner that the allocated resource number for ARBs which have stimulation values between 0 and 0.50 will be less while for ARBs which have stimulation values between 0.50 and 1 will be more.

3.2. Used data source

The used data source in this study is UCI machine learning repository (ftp://ftp.ics.uci.edu/pub/machine-learning-databases, 2005). Both breast cancer and liver disorders datasets were taken from this repository.

The name of the dataset for breast cancer problem is WBCD (Wisconsin Breast Cancer dataset). The dataset consist of 683 samples that were collected by Dr. W.H. Wolberg at the University of Wisconsin—Madison Hospitals taken from needle aspirates from human breast cancer tissue (Setiono, 2000). The WBCD consists of nine features obtained from fine needle aspirates, each of which is ultimately represented as an integer value between 1 and 10. The measured variables are as follows: (1) Clump Thickness ($x_1$); (2) Uniformity of Cell Size ($x_2$); (3) Uniformity of Cell Shape ($x_3$); (4) Marginal Adhesion ($x_4$); (5) Single Epithelial Cell Size ($x_5$); (6) Bare Nucleoli ($x_6$); (7) Bland Chromatin ($x_7$); (8) Normal Nucleoli ($x_8$); and (9) Mitoses ($x_9$). 444 of the dataset with 683 samples belong to benign, and remaining 239 data is of malignant.

The Liver Disorders data is named as BUPA Liver Disorders. BUPA Liver Disorders dataset prepared by BUPA Medical Research Company includes 345 samples consisting of six attributes and two classes. Each sample is taken from an unmarried man. Two hundred of these samples are of one class with remaining 145 are belong to the other. First five attributes of the collected data samples are the results of blood test while the last attribute includes daily alcohol consumption. The attributes are as follows (BUPA Liver Disorders Dataset, 2005):

1. mcv: mean corpuscular volume,
2. alkphos: alkaline phosphotase,
3. sgot: aspartate aminotransferase,
4. sgot: aspartate aminotransferase,
5. gammagt: gamma-glutamyl transpeptidase,
6. drinks: number of half-pint equivalents of alcoholic beverages drunk per day.

3.3. Used parameters

One advantage of AIRS is that it is not necessary to know the appropriate settings for the classifier. The most important element of the classifier is self-determined (Goodman et al., 2003). So, the following used parameters shown in Table 2 in our experiments have little effect on system performance.

Table 1

<table>
<thead>
<tr>
<th>Input</th>
<th>Output</th>
</tr>
</thead>
<tbody>
<tr>
<td>VS—Very Small</td>
<td>VS'—Very Small</td>
</tr>
<tr>
<td>S—Small</td>
<td>S'—Small</td>
</tr>
<tr>
<td>MS—Middle Small</td>
<td>MS'—Middle Small</td>
</tr>
<tr>
<td>LS—Little Small</td>
<td>LS'—Little Small</td>
</tr>
<tr>
<td>LB—Little Big</td>
<td>LB'—Little Big</td>
</tr>
<tr>
<td>MB—Middle Big</td>
<td>MB'—Middle Big</td>
</tr>
<tr>
<td>B—Big</td>
<td>B'—Big</td>
</tr>
<tr>
<td>VB—Very Big</td>
<td></td>
</tr>
</tbody>
</table>

if ARB.stim = VS and S then Output = (VS' + S') / 2
if ARB.stim = S and MS then Output = (S' + MS') / 2
if ARB.stim = MS and LS then Output = (MS' + LS') / 2
if ARB.stim = LS and LB then Output = (LS' + LB') / 2
if ARB.stim = LB and MB then Output = (LB' + MB') / 2
if ARB.stim = MB and B then Output = (MB' + B') / 2
if ARB.stim = B and VB then Output = (B' + VB') / 2

Fig. 4. Rule base for fuzzy resource allocation.
3.4. Measures for performance evaluation

3.4.1. Classification accuracy

In this study, the classification accuracies for the datasets were measured according to Eq. (6) (Watkins, 2001):

\[
\text{accuracy}(T) = \frac{\sum\limits_{t \in T} \text{assess}(t)}{|T|}, \quad t \in T
\]

\[
\text{assess}(t) = \begin{cases} 
1, & \text{if classify}(t) = t.c \\
0, & \text{otherwise} 
\end{cases}
\]

(6)

where \( T \) is the set of data items to be classified (the test set), \( t \in T \), \( t.c \) is the class of the item \( t \), and classify\((t)\) returns the classification of \( t \) by AIRS.

3.4.2. k-Fold cross-validation

For test results to be more valuable, \( k \)-fold cross-validation is used among the researchers. It minimizes the bias associated with the random sampling of the training (Delen et al., 2005). In this method, whole data is randomly divided to \( k \) mutually exclusive and approximately equal size subsets. The classification algorithm trained and tested \( k \) times. In each case, one of the folds is taken as test data and the remaining folds are added to form training data. Thus \( k \) different test results exist for each training–test configuration. The average of these results gives the test accuracy of the algorithm (Delen et al., 2005).

We used this method as 10-fold cross-validation in our applications. But we also conducted our experiments with three fold run for each training–test configuration. The average of these three test results gave us the test result for each fold. So we obtained 30 results in total to average.

4. Results

4.1. Results for Breast Cancer dataset

Fuzzy resource allocation mechanism provided Fuzzy-AIRS to classify WBCD with 98.51% classification accuracy. The accuracy reached by Goodman et al. with the use of AIRS was 97.2% (Goodman et al., 2002).

The relation between resource number and classification accuracy in Fuzzy-AIRS and AIRS for the WBCD is shown in Fig. 5. As can be seen from the Fig. 5, the required resource number for Fuzzy-AIRS is always less than AIRS in most of the points with same classification accuracy.

The advantages obtained by Fuzzy-AIRS are presented in Table 3. The values of used resource number and classification time in Table 3 are recorded for the highest classification accuracy.

As can be seen from Table 3 and Fig. 5, AIRS classified WBCD with the accuracy of 97.2% using 250 resources in 215 s, whereas Fuzzy-AIRS did the same classification process with the 98.51% classification accuracy using only 120 resources and only in 59 s. However the increase of 1.31% in classification accuracy is not so big, it is good to see an

---

Table 2

<table>
<thead>
<tr>
<th>Used parameters</th>
<th>Wisconsin Breast Cancer dataset (WBCD)</th>
<th>BUPA Liver Disorders dataset</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mutation rate</td>
<td>0.15</td>
<td>0.1</td>
</tr>
<tr>
<td>ATS (Affinity Threshold Scalar)</td>
<td>0.1</td>
<td>0.2</td>
</tr>
<tr>
<td>Stimulation Threshold</td>
<td>0.91</td>
<td>0.8</td>
</tr>
<tr>
<td>Clonal rate</td>
<td>10</td>
<td>10</td>
</tr>
<tr>
<td>Hyper clonal rate</td>
<td>2.0</td>
<td>2.0</td>
</tr>
<tr>
<td>Number of resources in AIRS</td>
<td>250</td>
<td>200</td>
</tr>
<tr>
<td>Iteration number</td>
<td>10,000</td>
<td>10,000</td>
</tr>
<tr>
<td>( K ) value for ( k )-nearest neighbour</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Number of resources in Fuzzy-AIRS</td>
<td>120</td>
<td>70</td>
</tr>
</tbody>
</table>

Fig. 5. The relationship between resource number and classification accuracy for WBCD in AIRS and Fuzzy-AIRS.
increase in classification accuracy. Indeed, the aim of changing resource allocation mechanism in AIRS was reducing required resource number so the classification time more than seeing an increase in classification accuracy. Fortunately, Fuzzy-AIRS showed a great performance increase in this manner and reduced classification time about 72% for both datasets. This is because the resource number required by Fuzzy-AIRS decreased more than half of the resource number used in AIRS. This improvement in performance is also very important especially in medical field and in applications that use large datasets.

4.2. Results for Liver Disorders dataset

Fuzzy-AIRS classified BUPA Liver Disorders dataset with the accuracy of 83.38%. There is no study in which BUPA Liver Disorders dataset is classified with AIRS. So, to compare Fuzzy-AIRS classification results with that of AIRS, BUPA Liver Disorders dataset was also classified with AIRS in this study. The classification accuracy of 81% was obtained with AIRS classification system.

The relationship between resource number and classification accuracy in Fuzzy-AIRS and AIRS for the BUPA Liver Disorders dataset is shown in Fig. 6. Like WBCD classification problem, the required resource number for Fuzzy-AIRS is always less than for AIRS in most of the points having same classification accuracy.

The used resource number and classification time for BUPA Liver Disorders dataset in Fuzzy-AIRS and AIRS are showed in Table 4. These values were taken for the highest classification accuracies, too.

It is seen from Table 4 and Fig. 6 that, Fuzzy-AIRS classified BUPA Liver Disorders dataset with 83.38% classification accuracy using 70 resources in 35 s while the used resource number and classification time of AIRS were 200 and 115 s respectively for the classification accuracy of 81%. The increase in classification accuracy was 2.38% this time and it is more satisfactory result than for Breast Cancer dataset thought our aim was not to obtain an increase in classification accuracy primarily. The success reached in classification time for Wisconsin Breast Cancer data was also produced for BUPA Liver Disorders with 70% decrease. Again the importance of this improvement shows itself in medical field and large sized datasets.

5. Discussion

5.1. Comparison of Fuzzy-AIRS with other classifiers

The classification accuracy obtained by Fuzzy-AIRS for Wisconsin Breast Cancer dataset is the highest one among the classifiers reported in UCI web site, in addition to other classifiers applied to corresponding problem. The comparison of Fuzzy-AIRS with these classifiers with respect to
Table 5  
Fuzzy-AIRS’s classification accuracy for WBCD classification problem with classification accuracies obtained by other methods in literature

<table>
<thead>
<tr>
<th>Author (Year)</th>
<th>Method</th>
<th>Classification accuracy (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Quinlan (1996)</td>
<td>C4.5 (10 × CV)</td>
<td>94.74</td>
</tr>
<tr>
<td>Hamilton et al. (1996)</td>
<td>RIAC (10 × CV)</td>
<td>94.99</td>
</tr>
<tr>
<td>Ster and Dobnikar (1996)</td>
<td>LDA (10 × CV)</td>
<td>96.80</td>
</tr>
<tr>
<td>Bennet and Blue (1997)</td>
<td>SVM (5 × CV)</td>
<td>97.20</td>
</tr>
<tr>
<td>Nauck and Kruse (1999)</td>
<td>NECLASS (10 × CV)</td>
<td>95.06</td>
</tr>
<tr>
<td>Pena-Reyes and</td>
<td>Fuzzy-GA1</td>
<td>97.36</td>
</tr>
<tr>
<td>Sipper (1999)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Setiono (2000)</td>
<td>Neuro-Rule 2a</td>
<td>98.10</td>
</tr>
<tr>
<td>Goodman et al. (2002)</td>
<td>Optimized-LVQ (10 × CV)</td>
<td>96.70</td>
</tr>
<tr>
<td>Goodman et al. (2002)</td>
<td>Big-LVQ (10 × CV)</td>
<td>96.80</td>
</tr>
<tr>
<td>Goodman et al. (2002)</td>
<td>AIRS (10 × CV)</td>
<td>97.20</td>
</tr>
<tr>
<td>Abonyi and Szefert (2003)</td>
<td>Supervised Fuzzy</td>
<td>95.57</td>
</tr>
<tr>
<td>Clustering (10 × CV)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Our study (2005)</td>
<td>Fuzzy-AIRS (10 × CV)</td>
<td>98.51</td>
</tr>
</tbody>
</table>

The classification accuracy is shown in Table 5. Although Fuzzy-AIRS provided only an increase of 0.41% in classification accuracy it is satisfactory for us to see Fuzzy-AIRS as a classification system that reached the highest accuracy for this data especially if we remember that AIRS is a new classification system in a new artificial intelligence area. We also think that any increase in classification accuracy for such medical data is important.

The success reached in Wisconsin Breast Cancer data was obtained for BUPA Liver Disorders dataset, too. Again, among the classifiers in UCI web site, Fuzzy-AIRS reached highest classification accuracy. Table 6 includes the comparison of Fuzzy-AIRS and AIRS with other classifiers in UCI web site in addition to other related with BUPA Liver Disorders dataset classification problem.

The considerable difference between the accuracies of Fuzzy-AIRS and the classifier that reached highest accuracy previously can be seen easily from Table 6. We do not include AIRS for this comparison because we want to emphasize the classification power of Fuzzy-AIRS over the other classifiers in the table. This difference of 8.52% in classification accuracy is a valuable improvement for classification of Liver Disorders dataset as well as it is also a great success for a version of AIRS.

5.2. Comparison of Fuzzy-AIRS with AIRS

In classification task, we use Receiver Operating Characteristic (ROC) curves and the area under the curves to compare Fuzzy-AIRS and AIRS classifiers. ROC curves are the reliable technique based on the values of true positives and false positives. Therefore, it provides a trade-off between sensitivity and specificity. The area under the curves is computed and according to value of this area

![Fig. 7. ROC curves for Fuzzy-AIRS and AIRS with the area under the curves for WBCD.](image)

![Fig. 8. ROC curves for Fuzzy-AIRS and AIRS with the area under the curves for BUPA Liver Disorders dataset.](image)
we evaluate these classifiers performances. The bigger area means that we have better classifier than another one, which has smaller area. ROC curves for Fuzzy-AIRS and AIRS with the area under the curves for WBCD is shown in Fig. 7. ROC curves for Fuzzy-AIRS and AIRS with the area under the curves for BUPA Liver Disorders dataset is shown in Fig. 8. In Figs. 7 and 8, Az refers area values for Fuzzy-AIRS and AIRS, respectively.

In addition to, comparing between Fuzzy-AIRS and AIRS based on sensitivity and specificity is shown in Table 7 for WBCD and BUPA Liver Disorders dataset.

### 6. Conclusion

In this study, the resource allocation mechanism of AIRS that is among the most important classification systems of Artificial Immune Systems was changed with a new one that was formed using fuzzy-logic rules.

In the application phase of this study, two important medical datasets, Wisconsin Breast Cancer dataset and BUPA Liver Disorders dataset, were used. In the classifications of these datasets, the analyses were conducted both for the comparison of reached classification accuracy with other classifiers in UCI web site and to see the effects of the new resource allocation mechanism.

According to the application results, Fuzzy-AIRS showed a considerably high performance with regard to the classification accuracy especially for BUPA Liver Disorders dataset. The reached classification accuracies of Fuzzy-AIRS for Breast Cancer dataset and BUPA Liver Disorders dataset were 98.51% and 83.38% respectively, which were the highest ones among the classifiers reported in UCI web site and other classifiers used for related problems. Moreover, Fuzzy-AIRS produced a difference of 8.52% in classification accuracy over the classifiers reported in UCI web site in addition to the other methods applied for BUPA Liver Disorders dataset. Beside of this success, Fuzzy-AIRS reduced the classification time with respect to AIRS by the amount of 71% for both datasets. This was the result of decrease in resource number done by fuzzy resource allocation. If we consider the importance of classification time for medical data and large datasets, this improvement gets more value.

AIRS is going one step ahead among the other classifiers with the aid of improvements done in the algorithm. The proposed change in this study has not only reduced the classification time through decreasing the required resource number but also produced very satisfactory results to use the classifier in other medical datasets. Other application areas are also open for Fuzzy-AIRS to experiment with. One of the further studies can be using the fuzzy mechanisms in other Artificial Immune Systems similar to AIRS.

### Acknowledgement

This study is supported by the Scientific Research Projects of Selcuk University (project no. 05401069).

### References


