Use of artificial intelligence techniques for diagnosis of malignant pleural mesothelioma

Malign pleural mezotelyoma tanısı için yapay zeka teknikleri kullanımı

Orhan Er¹, A. Çetin Tanrikulu², Abdurrahman Abakay²

ABSTRACT

Objective: Malignant pleural mesothelioma is a highly aggressive tumor of the serous membranes, which in humans results from exposure to asbestos and asbestosiform fibers. The incidence of malignant mesothelioma is extremely high in some Turkish villages where there is a low-level environmental exposure to erionite, a fibrous zeolite. Therefore epidemiological studies are difficult to perform in Turkey.

Methods: In this paper, a study on malignant pleural mesothelioma disease diagnosis was realized by using artificial immune system. Also, the artificial immune system result was compared with the result of the multi-layer neural network focusing on malignant pleural mesothelioma disease diagnosis and using same database. The malignant pleural mesothelioma disease dataset were prepared from a faculty of medicine’s database using patient’s hospital reports.

Results: 97.74% accuracy performance is obtained by artificial immune system. The accuracy results of artificial immune system algorithm are much better than the accuracy results of multi-layer neural network algorithm.

Conclusion: This system is capable of conducting the classification process with a good performance to help the expert while deciding the healthy and patient subjects. So, this structure can be helpful as learning based decision support system for contributing to the doctors in their diagnosis decisions.

Key words: malignant pleural mesothelioma disease diagnosis, artificial immune system, machine learning based decision support system.

ÖZET

Amaç: İnsanların beyin zarında bulunan, asbestos ve asbestosiform liferine maruz kalmakla oluşan kötü huylu pleval Mezotelyoma, oldukça saldıran bir tümördür. Düşük seviyeli çervesel erionite fibrous zeolite’e maruz bırakılmış Türkiye’deki bazı kasabalarda Mezotelyoma görülme oranı oldukça yüksektir.


Bulgular: Yapay bağışıklık sistemi tarafından hastalık teşhisi için %97,74 doğruluk oranında bir performans elde edilmiştir. Yapay bağışıklık sistemi algoritmalarının doğruluk sonuçları çok katmanlı yapay sinir ağı algoritmalarından çok daha iyi olduğunu göstermiştir.

Sonuç: Bu sistem uzmana, sağlıklı ve hasta kişiyi sınıflandırma sürecinde doğru teşhisı bulma yönünde iyi bir performans sağlar. Böylece bu yapı ile doğru teşhis sonuçcnda ulaşmadakı doktorlara bir karar destek sistemi olarak yardımcı olur.

Anahtar kelimeler: Kötű huylu plevral mezotelyoma hastalığının teşhisi, yapay bağışıklık sistemi, makine öğrenme tabanlı karar destek sistemi.

¹ Bozok University, Department of Electrical And Electronics Engineering, Yozgat, Turkey
² Dicle University, Faculty of Medicine, Department Of Chest Diseases, Diyarbakir, Turkey

Yazışma Adresi /Correspondence: Orhan Er,
Bozok university, Electrical-Electronics Engineering, Yozgat, Turkey   Email: orhan.er@bozok.edu.tr
Geliş Tarihi / Received: 12.01.2015, Kabul Tarihi / Accepted: 06.02.2015
Copyright © Dicle Tip Dergisi 2015, Her hakkı saklıdır / All rights reserved
**INTRODUCTION**

Malignant Pleural Mesothelioma (MPM) is a disease originating from pleura, pericardium, peritoneum or tunica vaginalis and it is since the early 1960s recognized to be strongly related to asbestos exposure [1]. MPM is generally caused by environmental and occupational exposure to asbestos. Also, erionite, a natural fibrous zeolite, which can be found in volcanic tuffs, has been found to induce MPM. MPM due to environmental exposure to asbestos and erionite is a relatively common cancer in Turkey [2-5].

MPM is a fatal cancer of increasing incidence associated with asbestos exposure [6]. MPM is a malignancy that is resistant to the common tumor directed therapies, but again individual patients might respond to chemotherapy, radiotherapy or immunotherapy, and selected patients might benefit from radical surgery and multimodality treatment [7].

Malignant mesothelioma is very aggressive tumors of the pleural which are responsible for approximately 15,000–20,000 deaths annually worldwide [8]. Estimated 1000 patients have MPM in Turkey per year. The annual incidence of pleural mesothelioma was 22.4/1,000,000 in Anatolia [9].

Several studies were carried out about MPM epidemiology, clinics in our region [10-13]. But there isn’t any study on MPM disease diagnosis using artificial immune systems (AIS) and artificial neural networks (ANN) with prognostic data.

In most published series of patients, the median survival for MPM was reported to be about one year [14-17]. Although it is claimed that multi-modality regimens slightly prolonged survival for relatively few patients in whom it is possible to perform radical surgery [18-19], most patients have unrespectable disease at presentation and systemic therapy has been the only treatment option for them [20].

Patient groups can be discriminated with mainly good or poor prognostic factors, and individual patients within these groups are likely to have a better or worse survival. The median survival of MPM patients differs from four to nine months depending on the presence of mainly poor or good prognostic factors, the two year survival ranges from 0 to 10% [7,21,22].

MPM has bad prognosis and low survival due to no curative treatment was implemented. Prognostic affect on MPM prognosis of various factors that clinical and laboratory were studied in several reports. In these studies, optimal treatment options and affect of survival were investigated. [2,7,16,20-23].

MPM disease diagnosis is an important classification issue. Classification is often a very important part of process in many different fields like medicine. The use of artificial intelligence methods in medical diagnosis is increasing gradually. There is no doubt that evaluation of data taken from patients and decisions of experts are the most important factors in diagnosis. However, experts systems and different artificial intelligence techniques for classification also help professionals in a great deal [24].

Artificial Immune Systems (AIS) can be defined as computational systems inspired by theoretical immunology, observed immune functions, principles and mechanisms in order to solve complex problems [25]. The biological immune system (BIS) is a subject of great research interest because of its powerful information processing capabilities; in particular, understanding the distributed nature of its memory, self-tolerance and decentralized control mechanisms from an informational perspective, and building computational models believed to better solve many science and engineering problems [26]. AIS can provide an alternative, efficient way for solving disease diagnosis problems like MPM disease diagnosis.

The multilayer neural networks (MLNNs) have been successfully used in replacing conventional pattern recognition methods for the disease diagnosis systems [27-29]. The back-propagation (BP) algorithm [30] is widely recognized as a powerful tool for training of the MLNNs. But, since it applies the steepest descent method to update the weights, it suffers from a slow convergence rate and often yields suboptimal solutions [31-32]. A variety of related algorithms have been introduced to address that problem. A number of researchers have carried out comparative studies of MLNN training algorithms [33-35]. Levenberg-Marquardt (LM) algorithm [33] used in this study provides generally faster convergence and better estimation results than other training algorithms [35].
In this paper, a comparative study of AIS on MPM disease diagnosis was realized. Also, the AIS results were compared with the results of the MLNN focusing on MPM disease diagnosis and using same database. The MPM disease dataset were prepared from a faculty of medicine’s database using patient’s hospital reports. The study aims also to provide machine learning based decision support system for contributing to the doctors in their diagnosis decisions.

METHODS

Data source

In order to perform the research reported in this manuscript, the patient’s hospital reports taken from Dicle University, Faculty of Medicine’s Hospital from southeast of Turkey was used. One of the special characteristics of this diagnosis study is to use the real dataset using patient reports gathered from this hospital in 2010. The study included 324 patients suffering from variety of MPM disease. The study was retrospectively, only investigated patients file.

In this dataset, all samples have thirty four features because it is more effective than other feature subsets by doctor’s guidance. These features are: age, gender, city, asbestos exposure, type of MPM, duration of asbestos exposure, diagnosis method, keep side, cytology, duration of symptoms, dyspnoea, ache on chest, weakness, habit of cigarette, performance status, White Blood cell count (WBC), haemoglobin (HGB), platelet count (PLT), sedimentation, blood lactic dehydrogenase (LDH), Alkaline phosphatase (ALP), total protein, albumin, glucose, pleural lactic dehydrogenase, pleural protein, pleural albumin, pleural glucose, dead or not, pleural effusion, pleural thickness on tomography, pleural level of acidity (pH), C-reactive protein (CRP), class of diagnosis. Diagnostic tests of each patient were recorded by an attending physician.

Diagnosis of the MPM disease using artificial immune system

The artificial immune system has been formed on the basis of the working principles of the natural immune system found in the human body. Tissues and organs related with the natural immune system in the body are the thymus gland, the bone marrow, the lymph nodes, the spleen and the tonsils. A central organ coordinating the functions of the associated tissue, the organ, the molecule and the cells does not exist in the immune system. The immune system, via its special cells, recognizes the foreign (external) cells filtering through the body and neutralizes them. The basic immunity cell is the lymphocyte [36]. The lymphocytes are grouped into two categories: “T” and “B” cells. The “B” cells can recognize the antigens without restraint in liquid solutions whereas the “T” cells need the recognition of antigens by means of other assisting cells [37].

Two different selection methods are utilized for purposes of reaching a solution in different types of problems as regards to artificial immune systems functioning on the basis of the natural immune system. The negative selection mechanism is used for problems such as pattern recognition, anomaly detection, computer and network security and time series analysis. The clonal selection mechanism, on the other hand, is particularly used for problems such as multi-purpose and combinatory optimization, disease diagnosis, computer and network security and error detection [38].

These problem-solving methods that are used in artificial immune systems thoroughly imitate the mechanisms found in the natural immune system that the human body possesses.

In this study, an artificial immune system model was used for the MPM disease diagnosis. The algorithmic steps of AIS model used for this purpose are,

Step 1. Create the antibody population and determine the suppression threshold.

Step 2. Generate clones (new antibody / antigen) for each antibody.

Step 3. Calculate the affinity among antibody cells and kill the antibodies whose affinities are less than the suppression threshold and determine the number of antibodies after suppression.

Step 4. If not ensure that memory population is constant, return to step-2

Step 5. Classify the given values

The antibody values are normal, MPM’s classes at the algorithm. For generating clones of the
antibodies, the antibody cells are mutated. In this study, each antibody has 34 antibody cells. In other words, 34 features were used as 34 antibody cells. An example of the generating clones of antibodies which used in AIS model shown in Table 1.

<table>
<thead>
<tr>
<th>Antibody cells</th>
<th>1</th>
<th>2</th>
<th>28</th>
<th>29</th>
<th>30</th>
<th>31</th>
<th>32</th>
<th>33</th>
<th>34</th>
</tr>
</thead>
<tbody>
<tr>
<td>47.0</td>
<td>1</td>
<td>79.0</td>
<td>0.0</td>
<td>0.0</td>
<td>34</td>
<td>2</td>
<td>1.0</td>
<td>0.0</td>
<td>1.0</td>
</tr>
<tr>
<td>55.0</td>
<td>1</td>
<td>6.0</td>
<td>1.0</td>
<td>1.0</td>
<td>1.0</td>
<td>42</td>
<td>3</td>
<td>42</td>
<td></td>
</tr>
<tr>
<td>55.0</td>
<td>1</td>
<td>6.0</td>
<td>1.0</td>
<td>0.0</td>
<td>1.0</td>
<td>42</td>
<td>3</td>
<td>42</td>
<td></td>
</tr>
</tbody>
</table>

As the Ag–Ab affinity is related to their distance, it can be estimated via any distance measure between two strings or vectors, such as the Euclidean, the Manhattan, or the Hamming distance. Hence, if the coordinates of an antibody are given by \( Ab = \{ Ab_1, Ab_2, \ldots, Ab_N \} \) and those of an antigen are given by \( Ag = \{ Ag_1, Ag_2, \ldots, Ag_N \} \), then the distance \( D \) between them can be defined as

\[
D = \sqrt{\sum_{i=1}^{N} (Ab_i - Ag_i)^2} \quad (1)
\]

\[
D = \sum_{i=1}^{N} |Ab_i - Ag_i| \quad (2)
\]

\[
D = \sum_{i=1}^{N} \delta_i, \quad \text{where } \delta_i = \begin{cases} 1 & \text{if } Ab_i = Ag_i \\ 0 & \text{Otherwise} \end{cases} \quad (3)
\]

where Eq. (1) is the Euclidean distance, Eq. (2) the Manhattan distance and Eq. (3) the Hamming distance.

For measuring affinity of generated antibody cells, Hamming model was used because of that the used features indicate the situations generally. Detailed calculations of the step-3 which was used in the artificial immune system based model algorithm are, Where, is numbers of training patterns, is existing antibody, is new clone, is previous affinity, is final affinity, is estimated class, is desired class, is suppression threshold, is training pattern index.

The step-3 of the artificial immune system coded in C# program codes of AIS model used for this purpose are shown Figure 1.

The step-5 of the artificial immune system based model was used for the classification of the test patterns. The calculation details of this step are similar to the calculation details of the step-3. These calculation details are shown Figure 2.

The MLNN structure (with one input layer, two hidden layers, and one output layer) was used for the classification of the
compared results with AIS structure on the MPM disease diagnosis. The hidden layer neurons (35 neurons for each hidden layer) and the output layer neurons use nonlinear sigmoid activation functions. In this system, thirty four inputs were features, and eight outputs are index of eight classes (MPM type and phase). Detailed computational issues about the application of the MLNN structures can be found in references [39].

Subsequently, the artificial immune system result was compared with the result of the multi-layer neural network focusing on malignant pleural mesothelioma disease diagnosis and using same methods and database.

MEASURES FOR PERFORMANCE EVALUATION

Classification accuracy

Classification accuracy [40] has been used for the study on MPM disease diagnosis.

Equations which used in the classification accuracies are shown in (4) and (5):

\[
\text{classification accuracy}(N) = \frac{\sum_{n} \text{assess}(n)}{|N|}, \quad n \in N
\]  

(4)

\[
\text{assess}(n) = \begin{cases} 
1 & \text{if } \text{classify}(n) = nc \\
0 & \text{otherwise}
\end{cases}
\]  

(5)

where is the set of data items to be classified (the test set), is the class of the item , and returns the classification of by AIS and NN.

k-Fold cross-validation

In order to minimize the bias associated with the random sampling of the training and holdout data samples in comparing the predictive accuracy of two or more methods, researchers tend to use k-fold cross-validation. In k-fold cross-validation, whole data are 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 10 results gives the test accuracy of the algorithm [42-46].

RESULTS AND CONCLUSIONS

This work presents an application for artificial immune systems on MPM disease diagnostic and using artificial neural network (multilayer neural network structure) for comparing results. The classification accuracies obtained by AIS and NN structures for MPM disease was presented in Table 2.

As seen in the table, average 97.74 % classification accuracy was obtained by using AIS algorithm for MPM disease dataset. From the same table it can easily seen that the accuracy results of AIS algorithm are much better than the accuracy results of MLNN algorithm. According to the same table, it can be seen also that the best results for the classification accuracy were obtained from the AIS structure used in this study.

There is not any study on MPM disease diagnosis using artificial immune systems (AIS) and artificial neural networks (ANN) with prognostic data.
So we could not compare our results with the other studies. That’s why, we compared two classification methods with each other for MPM disease diagnosis.

**Table 2.** Average of classification accuracies of test dataset for malignant pleural mesothelioma by 10-fold cross validation

<table>
<thead>
<tr>
<th>Test Folds</th>
<th>AIS accuracy %</th>
<th>Average</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>89.29</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>96.55</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>100</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>100</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>100</td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>100</td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>97.14</td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>100</td>
<td></td>
</tr>
<tr>
<td>9</td>
<td>94.44</td>
<td></td>
</tr>
<tr>
<td>10</td>
<td>100</td>
<td></td>
</tr>
</tbody>
</table>

AIS: Artificial immune systems, NN: Neural networks

Consequently, the following results can be summarized;

- This classification accuracy is highly reliable for such a problem because only a few samples were misclassified by the system.
- AIS algorithm is better than NN algorithm for the MPM disease diagnoses problem.
- The results obtained using artificial immune system structure is also quite good for MPM diagnostic problem. This system is capable of conducting the classification process with a good performance to help the expert while deciding the healthy and patient subjects. So, this structure can be helpful as learning based decision support system for contributing to the doctors in their diagnosis decisions.

**REFERENCES**


