Headache Disease Diagnosis by Using The Clonal Selection Algorithm

G. Tezel¹, U. Köse²

¹Selçuk University, Konya/Turkey, gtezel@selcuk.edu.tr
²Afyon Kocatepe University, Afyonkarahisar/Turkey, utkukose@aku.edu.tr

Abstract—The clonal selection algorithm is one of the well-known artificial immunity approaches and it is based on the clonal selection principle, which is used to express basic functions and features of an adaptive immune response given to an antigenic stimulus. According to the clonal selection principle, only specific cells recognizing the existent antigens are selected to proliferate against any potential future antigen presence. At this point, in order to improve effectiveness of the current immunity system, the selected cells are also put to an affinity maturation process and their affinities to specific antigens are improved for providing better protection. In this study, a headache disease diagnosis system has been designed and developed by using the clonal selection algorithm. The obtained results point a successful artificial immunity system that is capable of diagnosing a headache disease according to the evaluated symptom sets.

Keywords—Artificial immunity, clonal selection algorithm, disease diagnosis, artificial intelligence, classification.

I. INTRODUCTION

Today, artificial intelligence techniques are widely used to develop different types of intelligent systems that can solve real-world problems, with more accurate and advanced approaches. Because of this, the artificial intelligence field is associated with many different areas in the modern world. In this context, some popular artificial intelligence techniques like genetic algorithms, artificial neural networks and fuzzy logic are often used by researchers and scientists to perform different types of works and studies. In addition to these techniques, there are also some works on newer artificial intelligence techniques, which provide better solutions for the encountered real-world problems. At this point, artificial immunity is one of the newest artificial intelligence techniques, which provide more effective and accurate solutions for especially optimization and classification problems.

Artificial Immunity is some kind of artificial intelligence technique, which is based on the biological immune system that protects the organism against any dangerous, antigenic factors. In this context, a developed artificial immune system works according to some predefined algorithms or approaches that have been derived from important functions or features of the biological immune system. In time, it has been seen that the artificial immunity is successful at solving complex problems that can be categorized under optimization or classification. Especially in the last years, there has been an increasing interest in the artificial immunity technique and its related applications [1-6]. At this point, there are some important artificial immunity algorithms like negative selection or clonal selection algorithm and they are widely used by researchers or scientists to solve the related problems and obtain more accurate results.

The clonal selection algorithm is a well-known artificial immunity approach and it is based on the clonal selection principle, which is used to express basic functions and features of an adaptive immune response given to an antigenic stimulus. According to the clonal selection principle, only specific cells recognizing the existent antigens are selected to proliferate against any potential future antigen presence. At this point, in order to improve effectiveness of the current immunity system, the selected cells are also put to an affinity maturation process and their affinities to specific antigens are improved for providing better protection.

This study presents a headache disease diagnosis system, which has been designed and developed by using the clonal selection algorithm. The system is based on a classification mechanism, which can classify the person with headache disease or a healthy one. In order to achieve this, the system needs to be trained according to some special values associated with the foremost symptoms for a possible headache disease. Because of this, a training data set was used to maturate and determine the memory cells, which will be used to diagnose the headache disease according to new test data sets. After the training process, the performance of the system has been evaluated by using some new test data set including symptom sets for people with a headache disease and also people with no potential disease.

The rest of the paper is organized as follows: In the second section, basics of the artificial immunity technique and also a general structure of the clonal selection algorithm are explained briefly. Right after, the developed system is introduced in the third section. In this section, some information about headache diseases and also possible symptoms are also expressed in order to give more idea about function of the performed study. After the third section, the training and the test process of the developed system is explained in the fourth section. Finally, results of the study and some future works related to the context of the system are discussed in the last section.
II. ARTIFICIAL IMMUNITY AND THE CLONAL SELECTION ALGORITHM

In the context of this study, both artificial immunity and the related clonal selection algorithm are important factors to perform the expressed study and obtain the related results to discuss about success and effectiveness of the artificial intelligence field on solving the examined problem. Because of this, it is important to explain biological immune system, artificial immunity and the clonal selection algorithm briefly in order to have more idea about the study.

A. Biological Immune System and the Artificial Immunity

Biological immune system is some kind of natural mechanism that protects the organism against any dangerous, antigenic factors. At this point, a biological immune system is known as responsible for detecting the antigenic factors and performing the related operations to remove them from the living organism body [7]. Generally, a biological immune system can recognize all body cells and categorize them as “self” or “non-self” factors [8]. In this context, “non-self” factors are dangerous and disease-causing cells. On the other hand, “self” factors are known as the harmless cells within the body [9]. Within the biological immune system, there are two major groups of immune cells named as B-cells and T-cells and these cells help the system in recognizing an almost limitless range of antigenic patterns [10]. Additionally, some agents, which are specially produced as responsive for antigenic factors (antigens) and work on removing the related factors, also take place within the biological immune system. As general, these agents are called as antibodies [1].

The artificial immunity is a new artificial intelligence technique and also a computational approach, which was originally inspired from theories and ideas that are included in the context of biological immunology field. The artificial immunity includes some special algorithms and structures (negative selection algorithm, clonal selection algorithm, immune network model...etc.) to be used for solving real-world problems with more effective methods. By using artificial immunity algorithms or approaches, it is easy to design and develop intelligent systems that are capable of solving different domain problems [7]. As mentioned before, the related technique is especially used for optimization and classification problems. In addition to these problem approaches, it is also widely used for fault and anomaly detection, pattern recognition, scheduling, data mining autonomous navigation and searching problems [11].

B. The Clonal Selection Algorithm

The clonal selection algorithm was designed and developed according to the clonal selection theory that can be examined and explained within the biological immunology. With its approaches and functions, the clonal selection algorithm can be evaluated as similar to genetic algorithms. On the other hand, the related algorithm can also be categorized under evolutionary algorithms with its structure and algorithmic elements used within algorithm steps. At this point, it is important to present basics of the clonal selection theory in order to understand functions of the clonal selection algorithm.

As explained before, any cell that can be detected or recognized by the biological immune system is called as an antigen and can be shown with the abbreviation: Ag. In the context of immune mechanism, when a living organism is exposed to an Ag, some subpopulation of special cells responds by producing antibodies, which can be shown with the abbreviation: Ab. Within the mechanism, Ab’s are carried on B cells. B cells are some kind of cells, which try to recognize and bind to Ag’s. As general, each B cell secretes a single type of Ab, whereas the related Ab is specific for an Ag. After recognizing the Ag’s, the most responsive and effective cells can be taken into different cell operations to improve effective of the general immune system. In this aim, cells are taken into mitosis-division process and the related process results in a clone, a cell or set of cells that are the progenies of a single cell. Within the related cell operations, specific B cells can differentiate into improved B memory cells. After the related processes, the obtained memory cells circulate through the body of the living organism and when exposed to a new antigenic stimulus, they can start the process of producing high-responsive, improved Ab’s, which were selected for the specific Ag recognized within the past immune response [12].

The expressed mechanism is shown in the Figure 1.

![Figure 1: The clonal selection mechanism.](image)
past responses. The differences in the related affinities are because of the mechanism, which is also known as the maturation of the immune response. Generally, “the maturation” causes Ab cells to be different from the Ab’s that were included within the past response process. After the affinity maturation process, the Ab’s with higher affinities are selected to the set of memory cells that will be kept for a possible, future antigenic stimulus. It is also important that not only the set is diversified through a hypermutation mechanism, but also mechanisms must exist such that rare B cells with high affinity mutant receptors can be selected to dominate the response [12]. At this point, the cells with low affinity or self-reactive receptors must be efficiently eliminated, become anergic or edited [13-15].

During developing artificial immunity algorithms or approaches, it is also important to design a mathematical model that express the interactions among Ag’s and Ab’s and so provide effective, computational models to solve different types of problems. In this aim, the shape-space model has been introduced to show and express interactions among Ag’s and Ab’s [17]. In the context of the shape-model, both Ag’s and Ab’s are shown by using a codification as binary or real number values. Distance measures calculated for each Ag-Ab pairs are then used to calculate to determine the interaction degree values between the related cells. Mathematically, the generalized shape of a cell ( (m): an Ab or an Ag) can be represented by a set of L attributes, which are associated with coordinate axes such that \( m = \langle m_1, \ldots, m_L \rangle \) can be regarded as a point in an L-dimensional real-valued shape space \( S, m \in S^L \subseteq \mathbb{R}^L \) [12].

According to the explained features and functions, the clonal selection algorithm can be expressed under the following steps [12]:

1. Choose an Ag randomly \((Ag_i \in Ag_{(M)})\) and present it to all Ab’s in the related set \( Ab = Ab_{(r)} \cup Ab_{(m)} \) \((r + m = N)\).
2. Determine the vector \( f_j \) containing the affinity of \( Ag_i \) to all the \( N \) Ab’s in Ab set.
3. Select the highest affinity Ab’s from Ab set to compose a new set \( Ab^i_{(m)} \) of high-affinity Ab’s in relation to \( Ag_i \).
4. The \( n \) selected Ab’s will be cloned independently and proportionally to their antigenic affinities, generating a set \( C^i \) of clones: the higher the antigenic affinity, the higher the number of clones generated for each of the \( n \) selected Ab’s.
5. The set \( C^i \) is submitted to an affinity maturation process inversely proportional to the antigenic affinity, generating a population \( C^{ii} \) of matured clones: the higher the affinity, the smaller the mutation rate.
6. Determine the affinity \( f_j \) of the matured clones \( C^{ii} \) in relation to antigen \( Ag_j \).
7. From this set of mature clones \( C^{ii} \), reselect the one with highest affinity \((Ab^i_j)\) in relation to \( Ag_j \) to be a candidate to enter the set of memory antibodies \( Ab_{(m)} \).

If the antigenic affinity of this Ab in relation to \( Ag_j \) is larger than its respective memory Ab, then \( Ab^i_j \) will replace this memory Ab.

8. Finally, replace the \( d \) lowest affinity Ab’s from \( Ab_{(r)} \), in relation to \( Ag_j \), by new individuals in \( Ab^i_{(d)} \).

The flow diagram of the whole clonal selection algorithm steps is shown in Figure 2.

![Figure 2: The clonal selection algorithm.](image-url)

III. A HEADACHE DISEASE DIAGNOSIS SYSTEM WITH THE CLONAL SELECTION ALGORITHM

In this study, a headache disease diagnosis system was designed and developed with the support of the clonal selection algorithm, which is a well-known artificial immunity algorithm as it was explained before. As general, the function of the system is too simple and it is based on a two-class classification approach enabling user to determine healthy or ill people having a headache disease. In order to achieve this, the system must be fed with a training data that includes some certain symptom sets showing different values for a headache disease or healthy situation. On the other hand, in order to provide a consistent training set, the medical literature about headache diseases must be examined and the system structure or values that will be used must be modeled or determined according to the obtained information.

A. Headache Disease, Their Classification and Symptoms

Generally, headache diseases are classified by the International Headache Society (HIS) into 13 different classes [18]. 4 of these classes are related to primer headache diseases, whereas other remaining 9 classes are known as secondary headache diseases. The diseases included in the secondary headache class are usually because of some medical situations like celebral hemorrhage, brain tumors and eye diseases or some problems on the neck. Because of these
Headache Diagnosis by Using The Clonal Selection Algorithm

Factors, secondary headache diseases are usually treated by different clinics. But primary headache diseases are examined, evaluated and treated in the context of the Neurology clinic [19].

Some headache diseases like migraine, tension based headache and set headache are important diseases and must be specially treated by using different approaches. Especially migraine has some symptoms that characterize the related disease directly. On the other hand, set headache is the most unusual, but the most serious headache disease. Additionally, migraine is the most sighted headache disease among the expressed classes. At this point, it is important evaluate symptoms to determine a headache as a disease or a common situation. Because most of headaches are not categorized as a disease and may be because of environmental factors like stress, air pollution, noise...etc. So, before diagnosing a headache disease, symptoms must be examined and evaluated to make a sagacious decision.

Some important symptoms that are related to headache diseases generally are shown in the Table 1. It is important that existences of the related symptoms are changed according to different types of headache diseases. But eventually, all of them are evaluated as a headache disease. In the context of this study, the important point is to classify healthy and ill people separately.

Table 1: Some important headache disease diagnosis.

<table>
<thead>
<tr>
<th>Question</th>
<th>Yes</th>
<th>No</th>
</tr>
</thead>
<tbody>
<tr>
<td>Is there any vomitus?</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Is there any nausea?</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Is there any photophobia?</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Is there any aura symptom?</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Is the attack time is between 4-72 hours?</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Can the headache be categorized as a throb?</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Does the headache continue on different points?</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Does the intensity of the headache improve after any move?</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

B. Diagnosing Headache Diseases

The developed clonal selection algorithm provides an approach that is capable of diagnosing a possible headache disease. In this aim, the cell sets in the related artificial immune system must be evaluated and trained to provide suitable memory cells, which can determine an ill or healthy person according to given different values included in different symptom sets. After the training process, the system then can be fed by using different types of test values. All of these processes and the evaluation of the system in this sense are explained in detail in the following section.

IV. EVALUATION

The developed diagnosis system has been trained by using a training set including 215 different symptom values associated with different types of headache diseases and healthy situations. 150 of these values are related to headache diseases including different disease types like migraine, tension based headache and set headache...etc. Other remaining 65 values are related to healthy situations. As it can be understood from steps of the clonal selection algorithm, there are some user defined parameters that are used during the algorithm process. At this point, there are six different user defined parameters: clonal factor (CF), Ab pool size (APS), remainder pool ratio (RPR), selection pool size (SPS), number of generations (NG) and total replacements (TR). From these parameters, the CF is known as a scaling element, which determines the number of clones generated for each selected Ab. During the evaluation process, 30 different CF values changing between 0.01 and 0.99 are used to see performance of the developed system. Other remaining parameters are determined as APS = 80, NG = 50, RPR = 0.15, SPS = 30, and TR = 5.

After the training process, the developed system has been tested by using a test set including 250 different symptom values for different people. Like the training set, the test set also included different symptom values for both ill and healthy people. As the evaluation factor, classification accuracy of the system has been calculated. In this aim, number of correctly and incorrectly classified symptom values (symptom sets) are evaluated. The best classification results have been obtained with the CF value 0.92. Table 2 shows the classification results for the related test set.

Table 2: Classification results for the test set.

<table>
<thead>
<tr>
<th>The best classification CF</th>
<th>0.92</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total number of symptom sets</td>
<td>215</td>
</tr>
<tr>
<td>Correctly classified symptom sets</td>
<td>208 (96.74 %)</td>
</tr>
<tr>
<td>Incorrectly classified symptom sets</td>
<td>7 (3.26 %)</td>
</tr>
</tbody>
</table>

As it can be seen from the Table 2, the headache diagnosis system, which was developed with the clonal selection algorithm, shows a remarkable performance on classifying headache disease and healthy situations according to the given values included in different symptom sets. In addition to the related results, the whole algorithm performance (as in number of correctly symptom sets) for 30 different CF values can be represented in a graphic shown in Figure 3.
V. CONCLUSION

In this study, an artificial immunity system, which is based on the clonal selection algorithm, was used to diagnose headache disease according to received values within symptom sets. The system has been trained with a special training data set to include necessary memory cells to determine the class of a symptom set. In this context, the system works as some kind of a two-class classifier. At this point, the system is also some kind of experiment tool to show general performance of the artificial immunity technique in especially classification problems according to the context of this study.

In conclusion, it can be expressed from the obtained results that the developed system can be evaluated as a successful approach in diagnosing the headache diseases among different data sets. More generally, this study also shows that the artificial immunity technique can be used effectively for real-world based classification problems. As it was mentioned before, the system has been performed the best classification approach with the CF value 0.92. For other different CF values, it can be said that the performance of the system is improved with higher values of CF. Additionally; changing other parameters used for the clonal selection algorithm may affect the whole approach in a positive or negative way. The related parameters have been kept constant within this study but the related hypothesis can be tested in future works of this study.

In addition to the testing the algorithm parameters, the system can also be improved to classify the related headache diseases into the mentioned 13 different classes. This future work may require some modifications on the clonal selection algorithm and perhaps, more addition to the current system may be needed to form a more advanced (hybrid) system to classify all types of headache diseases successfully.

REFERENCES