An Approach to Management of Health Care and Medical Diagnosis Using a Hybrid Disease Diagnosis System

Hodjat Hamidi¹, Maryam Jahanshahifard

1. Department of Information Technology, Faculty of Industrial Engineering, K. N. Toosi University of Technology, Tehran, Iran

ARTICLE INFO

Original

Received: 2 Oct 2016
Accepted: 29 Dec 2016

ABSTRACT

Introduction: In order to simplify the information exchange within the medical diagnosis process, a collaborative software agent’s framework is presented. The purpose of the framework is to allow the automated information exchange between different medicine specialists.

Methods: This study presented architecture of a hybrid disease diagnosis system. The architecture employed a learning algorithm and used soft computing to build a medical knowledge base. These machine intelligences are combined in a complementary approach to overcome the weakness of each other. To evaluate the hybrid learning algorithm and compare it with other methods, 699 samples were used in each experiment, where 60% was for training, 20% was for cross validation, and 20% for testing.

Results: The results were obtained from the experiments on the breast cancer dataset. Different methods of soft computing system were merged to create diagnostic software functionality. As it is shown in the structure, the system has the ability to learn and collect knowledge that can be used in the detection of new images. Currently, the system is at the design stage. The system is to evaluate the performance of hybrid learning algorithm. The preliminary results showed a better performance of this system than other methods. However, the results can be tested with hybrid system on larger data sets to improve hybrid learning algorithm.

Conclusion: The purpose of this paper was to simplify the diagnosis process of a patient by splitting the medical domain concepts (e.g., causes, effects, symptoms, tests) in human body systems (e.g., respiratory, cardiovascular), though maintaining the holistic perspective through the links between common concepts.

Keywords: Medical diagnosis, Diagnosis process, Genetic algorithms, Ontology.
Introduction

A medicine diagnosis is an important process in health care which can be fitted as an important and necessary process on the patient before pharmaceutical phase by a medical practitioner. This attitude requires more efficient health and health services such as medical diagnosis and reliable electronic health system. Various health problems are initiated by combination of different factors such as the environment, genetics, and lifestyle. These factors can influence different systems of human body such as heart, digestive, and respiratory (1). Diagnosis is often based on the results of blood test, urine test, and similar methods. Also, if it is needed to test the internal organs of patients, the doctors refer these patients to radiologists to take basic images through devices such as MRI, CT scan, and x-ray referral (2). Computer-aided detection systems and images are used in detection and diagnosis of radiologists (4-3). Performance of these systems increases cancer detection to 6.19% by radiologist (5). However, apart from the order that the use of computer-aided detection systems achieved, it is difficult to find the hidden cancer cells because these cells are difficult to identify (6). According to reports of radiologists, they don't trust cancer diagnosis just by computer-aided detection systems. On the other hand, in order to simplify the exchange of information and to reduce the complexity of medical diagnosis process, a common framework of operating software is presented. Each expert system of human body has a software agent to facilitate the relationship between data in diagnosis of disease and human body systems information update (Figure 1). A certain list of any auxiliary operating external systems is dependent on the knowledge provided in the knowledge base. After the update of internal documents (the physician documents collected after diagnosis), auxiliary agent assign the information to a list of the operating system-dependent agents. This information can be shared (the individual who smokes takes part in both cardiovascular and respiratory systems). If no outgoing performance is compromised among the agents, operation ends detection process.

![Figure 1](image.png)

**Figure 1.** For each of the professionals, a software agent to each system is assigned for the body to be fitted.

Literature review

The idea of an intelligent hybrid system is that more than one method of techniques are combined together to complement each other (13). Azvine (1997) (12) employed expert system methods for a clinical diagnosis to help the system in detecting the test of ear diseases. They made experiment with the use of a number of soft computing
algorithms (neural networks and fuzzy logic) to do this work. The use of this algorithm supports its medical applications in the future. Economou (2001) \(^{19}\) used computer to help doctors detect lung diseases in their diagnosis which makes reliable factors for final decisions to them. Other systems are suggested by Nazmi \(^{14}\) for liver diseases on the basis of CT diagnosis. The extracted factors and features feed the classification of neural networks.

The system is based on two-dimensional images and for some cases malformed and hidden 3d pictures are required for detection. The system introduced by Raja \(^{15}\) is a computer backup system for early detection of Alzheimer’s disease. It is based on Single Photon Emission Computed Tomography (SPECT). This has 96\% of accuracy in the classification. Table 1 presents some of the works conducted in this area.

<table>
<thead>
<tr>
<th>Author</th>
<th>Title</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Nazmy et al. (^{14})</td>
<td>Adaptive neuro-fuzzy inference System for classification of ECG signals.</td>
<td>This has used ANFIS for detection of ECG signals. It has 97% of accuracy.</td>
</tr>
<tr>
<td>Raja et al. (^{15})</td>
<td>A hybrid Fuzzy-Neural system for computer-aided diagnosis of ultrasound kidney image using prominent features.</td>
<td>In this work the hybrid method was efficient in detection of kidney disease.</td>
</tr>
<tr>
<td>Benamrane et al. (^{16})</td>
<td>A hybrid fuzzy neural networks for the detection of Tumors in medical images.</td>
<td>This work was useful in detection of Tumors by A hybrid fuzzy neural networks.</td>
</tr>
<tr>
<td>Andrės, C. et al (^{17})</td>
<td>A fuzzy-genetic approach to breast cancer diagnosis</td>
<td>This work with a approach got 97% of accuracy in diagnoses.</td>
</tr>
<tr>
<td>Harikumar, R, et al, (^{18})</td>
<td>Genetic algorithm optimization of fuzzy outputs for classification of epilepsy risk levels from EEG signals</td>
<td>The results of this study that used optimization method indicated 90% of accuracy.</td>
</tr>
<tr>
<td>Guler, I, et al. (^{19})</td>
<td>Combining Neural Network and Genetic Algorithm for Prediction of Lung Sounds</td>
<td>The hybrid method in this study could also reach successful results for detection.</td>
</tr>
<tr>
<td>Verma, B., et al. (^{20})</td>
<td>A novel neural-genetic algorithm to find the most significant Combination of features in digital mammograms.</td>
<td>This study used genetic algorithm in mammograms</td>
</tr>
<tr>
<td>Das, A., Bhattacharya, M. (^{23})</td>
<td>GA Based Neuro Fuzzy Techniques for Breast Cancer Identification</td>
<td>This study has also used GA Based Neuro Fuzzy Techniques for Breast Cancer detection</td>
</tr>
<tr>
<td>Yardimci, A. (^{24})</td>
<td>Soft computing in medicine</td>
<td>This work has also used soft computing in medicine.</td>
</tr>
</tbody>
</table>

Table 1. Some research works that used soft computing to detect disease.
Table 2. Some of the researches used the operating software for medical diagnosis

<table>
<thead>
<tr>
<th>Author</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Chena Yang et al. (25)</td>
<td>Web-based expert system to diagnose the nutrition with the use of technology in expert system of artificial intelligence.</td>
</tr>
<tr>
<td>Samuel et al. (26)</td>
<td>The backup system is a Web-based decision-making system with fuzzy logic to detect the disease typhus.</td>
</tr>
<tr>
<td>Pereiraa et al. (27)</td>
<td>Automatic construction of medical system based on the agent. The main objective of this project is development of intelligent Web template to provide e-services to health professionals, patients, and citizens who are involved in the care of the elderly patients at home.</td>
</tr>
</tbody>
</table>

A number of tasks have similar theoretical methods such as ontology, Bayesian network, and the operating software. They are used in the midst of this market. Bayesian network is used to demonstrate a lack of certainty and causality. Ontology is a sustainable model in the system of knowledge and cooperation between the operating software is for specific tasks of modulus.

In this article, these methods are applied with the vision of integrating uncertainty reasoning mechanism using a graphical network model of Bayesian along with ontology views of knowledge domains and software operating as a communication framework. The ontology is related to modeling knowledge. But as a single sample, it is introduced by the whole system. Each disease detection system has its own ontology which is shared amongst the experts operating. Each expert has access only to his expertise (Special node of network is Bayesian itself). Backup software agent shares the information with other professionals. This method is different from all other proposed researches. The idea of joint medical software operating was introduced by Arsen et al., in 2010 (35).

In a novel perspective, Nikovski et al. (36), constructed the bayesian network medical diagnosis system from the incomplete and partially correct numerical probabilistic information by introducing domain-dependent constraints. It aimed to solve the problem of determining combined influences of several diseases on a single test result for individual diseases. While it is a fact that this method encounters the following conditions as more network nodes, complex network structure, as well as high uncertain relationship between parameters and node, the feasibility and calculation performance will descend greatly. Therefore, it is suitable to diagnose for few symptoms. Recent published methods in the intelligent medical diagnosis include an improved genetic algorithm procedure to optimize parameter and select feature of the multi-layer perception network in medical diagnosis of diabetes, heart, and cancer (37). Chattopadhyay et al. (38), proposed a CBR-based expert system that uses the K-nearest neighbor (KNN) algorithm to search k similar cases based on the Euclidean distance measure to enhance diagnostic accuracy for complex diseases. Nonetheless, expert system has little applied value due to the defects in the production structure and serial working manners.

The aim of this paper was to simplify the diagnosis process of a patient by splitting the medical domain concepts in human body systems through the links between common concepts.

Methods

This paper offers a way in which a soft computing method and algorithm are used to create a database of medical information and a conclusion to the assortment of new information.

Over the past few years, the increasing attention to severe challenges in medical diagnosis process such as sharply increased elderly patients and limited medical personnel...
have led to a number of contributions in the areas of the intelligent medical diagnosis methods. Early contributions can be found on the neural networks, they provide a new significant way for intelligent medical diagnosis. Based on this idea, artificial neural networks have been applied in the diagnosis of: (i) pancreatic disease \(^\text{(1)}\), (ii) gynecological diseases \(^\text{(2)}\), (iii) early diabetes \(^\text{(3)}\), (iv) colorectal cancer \(^\text{(4)}\), and (v) multiple sclerosis lesions \(^\text{(5)}\). While this kind of method to set up the achievements of medical diagnostic system is still limited, the main reasons are that the learning algorithm cannot calculate the right results when the required algorithm to set up neural network model solves the larger multi-features disease diagnosis problems.

To this end, a system of soft smart computing is combined with the methods of artificial neural networks, fuzzy logic, and genetic algorithm. Combination ability and characteristics of each of the three methods are applied with regard to the necessity of learning capabilities in the field of diagnosis. The system consisted of two phases - phase and stage of learning diagnosis. This is the learning and training phase before the diagnosis. A collection of video data that represents a disease is suspected for the training phase. This collection of data can be extracted using principal component analysis method attributes. Then, the genetic algorithm has been used through the most important attributes for use in the training algorithm. In this paper, an algorithm is introduced based on the combination of learning methods and artificial neural networks, fuzzy logic, and genetic algorithm. This algorithm produces a database to analyze a new image. If the database fails to detect a new picture of its trained algorithm, it is used again to learn the new examples. This will help knowledge base and its abilities to detect items. This is phase detection of new images as input. The knowledge base is made in the process of learning a new image for diagnostic decision. At this stage, each image can perform feature extraction and classification. Then, the algorithm is for detection as input. Structure of the proposed system is presented in Figure 2.

![Figure 2. Proposed structure for disease detection](image)

After receiving the images, a process will be applied to extract the important features for classification of images. The properties of the tissue are extracted here for a description resulting in a better performance \(^\text{(26)}\). The principal component analysis method was used to extract the features and categories of images. In the process, an image is initially decomposed to reduce resolution and then the desired properties can be extracted with the principal component
analysis method\(^{(27)}\).

The next step after the extraction of features is use of these properties as input for the learning algorithm. The algorithm combines strengths of the three methods to create a powerful algorithm. The strengths are optimization by genetic algorithms, neural networks, learning ability and tolerance of uncertainty, as well as lack of precision of fuzzy logic neural network in education.

The system is expected, to be able to teach her again, and since the whole network has been training again it has to be time consuming. The type of network used in the system must have a memory as the output of network weights after training and save the combined data with new data. Neural networks, as a type of artificial memory or an enclosed tissue layer are returned \(^{(19)}\) to keep or maintain the output hidden layer weights. The weight of this layer is as an input for the hidden layer after a period of time. This method can provide new information once again to be used for training and retraining requirements. In the first stage, the most important features among the large number of features, was selected. Genetic algorithm is used to optimize the features and to reduce the number of algorithm features. In addition, the genetic algorithm is also applied for optimizing initial weights of neural network. Although the capability of learning and training make artificial neural networks, applicable to many applications, but it only deals with crisp values that are not appropriate in programs based on lack of precision or uncertainty. For this reason, to enhance artificial neural networks' capability, when they are under uncertainty conditions, these networks are combined with fuzzy logic to improve the performance of the composition. Fuzzy rules are performed as if-then approach. The role of fuzzy is to deal with uncertainty and accuracy and these features can be combined with neural networks' ability to learn and training. Therefore, by combining these two methods in medical imaging applications (that's a lot of uncertainty regarding the nature of images and information enclosed in it, for example, organizations, etc.), this system will help error detection accuracy. On the other hand, in fuzzy logic, all fuzzy rules do not contribute to better results. Genetic algorithm is used for the optimization of fuzzy rules by selecting the most important rules that make performance of algorithm better.

The components of the human body system (such as the respiratory system, digestive, urinary, cardiovascular, endocrine, immune, muscular, and sight) are placed in separate operating software. Each expert based on his/her expertise requires the software agents to make such interaction. The expert uses the collected evidences from the patient during analysis (such as cholesterol levels). Then, two factors exchange the information based on common concepts among them (for example, respiratory and heart diseases with two concepts of smoking and shortness of breath are common). The issue of network operating was greatly studied \(^{(34-36)}\). This article is a practical framework for a distributed expert system and common development to create a general model for diagnosis of medical assistant. The following general issues are divided into distributed domains as the platform agents. Single network operating framework is also presented by Arsene et al. (2011). This distributed and shared plan has been developed with applying the integrity of multi-factorial system \(^{(39)}\). The method was suggested by Arsene et al. (2011), and modeled in the world by ontology. It indicates a data structure in the rear of the Bayesian network's nodes. Therefore, there is a Bayesian network factor.

The scope of medical professionals was made by ontology experts and can be used by physicians through the graphical user interface. The network is set up to allow Bayesian user to arrange documents based on the variables scope (special system, the cause of the disease, the impact of disease symptoms, and testing for diagnosis of the disease).

According to this method, each specialization is in a particular system of the human body expertise (such as the respiratory and cardiac). All of the
variables of a scope are not only seen in a system of concepts, but also solved in a certain sphere of ontology system. Therefore, the lung specialist physicians will work only with the concepts of respiratory system. The rest of concepts of this field are part of the Bayesian network model.

Medical diagnosis in the case of a certain diseases can be seen as a unique overall process. Using this method, the process of parallel detection system can be optimized by experts without losing optimal communication among them.

Each operating system by a network is encryption and the connection between the shared variables in the background is automatically conducted. The influences of the Exchange System are among them. Just an ontology agent is responsible for the knowledge of the field there (Figure 3). Each factor is a network model of ontology and makes the system on the basis of dependency list associated with the ontology agent. They sent and received the ideas of shared variables in their system (Figure 4). The most important part of the research was in the field of artificial intelligence and intelligent systems as multiple views. The Bayesian network models are distributed mainly on a few factors and system characteristics (the distribution and communication). Factor of a few systems are operating under the scope of knowledge and the entire field. A collection of local track goals and ideas are exchanged with other operating systems.

In this approach, the completely independent agents are employed and issues are distributed based on the existence of agents’ characteristic (Figure 5).

Ontology is a clear and explicit conventional characteristic of common concepts. A concept model of a number of interesting phenomena in the world that identify the concepts related to a phenomenon \(^{(40-45)}\). The ontology is used in artificial intelligence to make it easier to develop and reuse knowledge. The ontology is developed to prepare semantics machines capable of processing information resources to communicate among different agents. The first step to create ontology is to define the main concepts of contractual methods to show concept medical ontology classes offered by Arsene et al. 2011, (Figure 6).

There is a main class: cause and effect. Sub-class environment is assumed. Under the impact classes are: illness, symptoms, symptoms of the disease, and tests to diagnose the condition.

![Figure 3. relationship between ontology and network factors](image1)

![Figure 4. Relationship between network agents](image2)
Results

Experiment preparation: In order to assess the accuracy of hybrid learning algorithm presented in the previous section, the cloud was used by neuro solutions software to test the algorithm. This software is a simulation environment that is a platform for development of neural networks applications. Breast cancer data set that has been used in the testing were obtained from the source \(^ {46-52} \). The data set have 699 samples with 10 features for each sample. As represented in Table 3, the first column in the data set represents the sample number. Columns 2 to 10 indicate the examples that have possible outputs as benign or malign. Column 11 shows the theoretical output in which number 2 represents benign and number 4 malignant. In the test, three different experiments were conducted with different methods including neural network with fuzzy logic, neural networks, and learning algorithm with genetic algorithm. Neural Networks.

Table 3. Sample data

<table>
<thead>
<tr>
<th>Columns</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
</tr>
<tr>
<td>A</td>
</tr>
<tr>
<td>B</td>
</tr>
<tr>
<td>C</td>
</tr>
<tr>
<td>D</td>
</tr>
<tr>
<td>E</td>
</tr>
<tr>
<td>F</td>
</tr>
<tr>
<td>G</td>
</tr>
<tr>
<td>H</td>
</tr>
<tr>
<td>K</td>
</tr>
<tr>
<td>L</td>
</tr>
</tbody>
</table>
Experiment results: Hybrid learning algorithm has been made using a graphical interface neuro solutions. All parts of the system were to form the learning algorithm which is connected to the three test methods.

In order to review and compare the hybrid learning algorithm with other methods, 699 samples were tested in each instance, among which 60% were for training, 20% for cross-training, and 20% for the assessment test. Table 4 represents the results obtained from these tests. As it can be seen, this method functions more efficient than the method resulted from the combination of neural network with fuzzy logic. But the method of neural network is combined with genetic algorithm method to obtain better results. These results can be due to the information of a sample used in the training phase. The neural network program is to produce a better performance to a large data set for training needs and stage with a combination of three methods in a combined system. The sample data for training must be larger than the data used in this experiment. Generally, the performance reported in this experiment indicates some evidences that this is a hybrid system. It may be better than the other methods discussed in the previous sections.

Table 4. Results of experiment on the database of breast cancer

<table>
<thead>
<tr>
<th>Accuracy</th>
<th>False</th>
<th>True</th>
<th>Number of samples</th>
<th>Method</th>
</tr>
</thead>
<tbody>
<tr>
<td>96.4</td>
<td>5</td>
<td>134</td>
<td>139</td>
<td>neural network combined with fuzzy logic</td>
</tr>
<tr>
<td>97.84</td>
<td>3</td>
<td>136</td>
<td>139</td>
<td>neural network combined with genetic algorithm</td>
</tr>
<tr>
<td>97.12</td>
<td>4</td>
<td>135</td>
<td>139</td>
<td>Hybrid method (neural network, fuzzy logic, genetic algorithm)</td>
</tr>
</tbody>
</table>

The following test cases have been proposed for evaluation of the framework:

- Two of human body systems can be considered: heart and respiratory system.
- Respiratory system has the following features: smoking, lung cancer, asthma, bronchitis, shortness of breath, bacteria, pneumonia, and allergies.
- Cardiovascular system has the following features: smoking, alcohol, obesity, cholesterol, arteriosclerosis, drug consumption, diet, serum selenium, blood pressure, heart disease, shortness of breath, angina, gentle exercise, fast heart rate, serum triglyceride, and LDL serum.
- Cardiovascular and respiratory systems have two characteristics in common: shortness of breath and cigarettes.
- Conditional probability tables are in the resource studies (17-18) as the test data set.

Specialists of cardiovascular system can make reasoning for the patient's smoking from the patient's appearance. Thus, they put “yes” for cigarette variable (Figure 11). The idea has been updated through the respiratory system for smoking variables. The relationship between the agent and the agent network heart network has been represented in Figure 7 and 8. Operation of a network of respiratory variables include: smoking, cancer, bronchitis, shortness of breath while in the variable view smoking and shortness of breath are common with previous network operating. The following variables are the heart network: smoking, hypertension, heart disease, shortness of breath, gentle exercise, and MI. Due to the lack of space only a few variables have been shown in Figures 7 and 8.
The relationship between Network Agent and the relationship between the network and operating network have been shown in the shape of a heart in Figures 9 and 10. In Figure 11, new updated ideas in the form of cigar variable, the value of Yes increased from 50.5 to 70. Double agent software is able to convey the idea of common features. Calculation of updated Bayesian network system happens in the device. Instead of having multiple heavy computational efforts, for example, in the case of a single agent (39), the whole process of modal is distinguished between parallel parts of body system. It can be divided into some parts of

information for cooperation. The performance of network load graph stage is tested by Bayesian with quad-core processor 5.2 MHz and memory of 8 GB 64-bit operating system. Deduction of the run-time function is parallel: 95% of overall Cpu time, \( F = \) a total of 2.0 times = 0op. Elapsed time comes down (Table 5). Best CPU is for the processor IV. The difference between values of the elapsed time depends on the degree of complexity of graph. If complexity is higher, better acceleration is by higher number of cores. In order to have the best trigger, we should have more main causes and more parallel strings.
Figure 9. The relationship between the agent and the agent network the cardiac Plexus.

Figure 10. The relationship between a networks operating.

Figure 11. Set document cigarettes in cardiovascular system.
Discussion

The proposed framework is a combination of auxiliary agent architecture. All three technology artificial intelligence, ontology, and Bayesian system are combined with each other and the network creates a common factor. The following important benefits are from the collaboration of AI technology:

- All three layers of the framework are defined as software components. They cooperate in runtime applications and can be integrated in dynamic and flexible methods. The actual implementation of the interface definitions is a separate component. It is for the development and update of individuals.
- Bayesian network is easier for understanding the dependency and the local distribution of human being. There is a relationship between the ideas of a system of infinite loop. The impact of the system is only distributed after the specialist arranged her/his reasons, based on his/her adjusted view about some system variables. Intermediaries to communicate from this point of view are just some shared and observed variables. They can send their ideas, possibilities to integrate other systems to affect the off-set. Thus, the local stability of a network agent can be guaranteed.
- Book system specialists work in parallel without losing the knowledge partners.
- Using the Java technology is considered ready to run on all operating systems without any changes and conversions. The application can be
in the form of handheld devices (smartphones, personal assistant) development.

- All modules depend on the communication framework.
- The structure of ontology is not personal. Application of the targeted models is encrypted.
- From the perspective of a medical field, a specified area must be covered by many specialists. The current research assumes that this is a one by one relationship. The future direction of this work includes implementation and assessment of the entire system components based on the data (benchmarking). This system is expected to progress the health system to provide an accurate diagnosis of the disease and automatically specifies the secret areas that may have existed in the cancer cells. The proposed method is based on the knowledge of following distribution areas of medicine. Therefore, it is a horizontal distribution that shows the level of permission to the individual. The overall cooperation among the experts can guarantee the stability of the level of knowledge. Vertical distribution due to increased accuracy can also reduce intelligence, while multiple knowledge fields together can collaborate to expand the current methods.

Conclusion

This article presents the structure of a hybrid detection system. Different methods of soft computing system (artificial neural networks, fuzzy logic, and genetic algorithm) merged to create diagnostic software functionality. As it is shown in the structure, the system has the ability to learn and collect knowledge that can be used in the detection of new images. Currently, the system is at the design stage. The system is to evaluate the performance of hybrid learning algorithm. The preliminary results showed a better performance of this system than other methods. However, the results can be tested with hybrid system on larger data sets to improve hybrid learning algorithm.

The main section is as follows:

- Software architecture combines use of artificial intelligence technology as the independent model and as a portable application.
- Computational model, i.e., a Java technology, was used for multi-disciplinary to implement a graphical algorithm. Parallel computational model, has the benefits of a multi core computer. The future aim of this project will be focused on the measurement of times using both technologies and compare them (compare description logics using a multi-agent architecture versus a rule based system able to make the same kind of diagnosis) to determine better one. Other research lines will be focused on the design and implementation of distributed inference platforms based on other reasoning techniques and their comparisons.

Acknowledgement

We’d like to thank Dr. Maryam Kooshki for assistance with medical fields, as well as Dr. Taha Mokfi and Roohallah Alizadeh Sani for their guidance and constructive suggestions that greatly improved the manuscript.

Conflict of interest

The authors declare that there is no conflict of interests.

References


