Abstract

The practice of medicine has evolved rapidly of the last 50 years. Due to these advances, the number of parameters that are presented to the medical professional have expanded beyond the capacity for complete analysis by healthcare professionals. Medical data is complex and includes a number of different representations including written notes, numeral data, and images. An automated analysis must not only deal with the multiple data types but must also be able to consolidate all results related to possible diagnoses and treatment options. In this paper several independent methods of analysis are unified into a comprehensive approach to automated medical decision making. It is illustrated with a specific example in diagnosis of congestive heart failure.

Key words: automated decision making, expert systems, neural networks, medical data analysis

1 Introduction

Due to rapid advances in medical care many new parameters have been added to diagnosis and treatment [1]. To fully investigate all possibilities computer assistance is needed [2]. Methods for computer-assisted diagnosis must be defined for each disease. In addition, an automated decision aid must be able to explore the possibility that patient symptoms may be due to more than one disorder. Disease diagnosis relies on a number of different data formats include textual, numeric, and imaging input. In order to address all types of information a number of different approaches are needed. In this article, structures are outlined that provide the ability to include all types of data in an automated decision making system. Rule-based analysis began with the MYCIN program [3] and was one of the earliest attempts to use computers to assist in medical diagnosis. It used rules for analyzing medical questions. An example is MYCIN which used an inference engine comprising a set of rules which would provide a diagnosis accompanied by a certainty factor that indicated the confidence level of the results. This early attempt led to a number of new automated methods for analyzing medical data. [4], including INTERNIST [5], Quick Medical Reference (QMR) [6], Iliad [7], and DXplain [8].

Figure 1 illustrates the complex sources of information that must be combined in automated medical decision making. The Task Manager collects and compiles and allocates findings to the appropriate automated decision making component.

In the work presented here a medical decision making approach is illustrated that can deal with a broad range of data formats as well an methods for determining the level of certainty of the result [9]. It also includes a framework for easy updates when new knowledge becomes available.

2 Methodology

Several methodologies are needed to cover the broad area of medical diagnosis. Each type is illustrated here. In Section 3 specific cases are included that illustrate the need for computer-assisted analysis in the diagnosis of a complex case.

2.1 Rule-Based Analysis

The rule based system, EMERGE, was first used for screening patients to determine the need for hospitalization [10]. EMERGE was shorthand for emergency as the intent was to use the approach in the emergency room environment. Work was expanded to use EMERGE as part of a new approach to diagnosis in general that included not only emergencies but also functioned as a broader diagnostic model. The new model includes not only rule-based analysis but also employs neural networks, and time series analysis [11].
The rule structures for EMERGE are stated as situation-action pairs in which the first part is a list of items to watch for and the second part is a list of things to do:

*If this condition holds then this action is appropriate.*

A set of standard conditions (SC) is used for each rule. Rules take the following forms: The condition part of the rule has any number of conditions. The conditions can take the following forms and substantiated as indicated:

**AND:** All conditions must be confirmed

**OR:** At least one condition must be confirmed

**COUNT** A specific number be confirmed

Rules may use any or all of the three methods to determine if the condition exists. The rules are called standard conditions (SC):

\[ SC_i(m,m) = \text{AND} \]

\[ SC_j(1,m) = \text{OR} \]

\[ SC_k(n,m) = \text{COUNT } n, 1 < n < m \]

The first variable indicates the number of conditions that must be substantiated for the rule to hold and the second number indicates the total number of conditions in the rule. Examples from the EMERGE chest pain rule base are given in Section 3.
2.2 Neural Networks

Complex data requires the use of nonlinear neural networks. The Hypernet [12] neural network designed by the authors is used in this approach to deal with numeric data. The general structure is given by the equation:

\[ D(x) = \sum_{i=1}^{n} w_i x_i + \sum_{i=1}^{n} \sum_{j=1}^{n} w_{ij} x_i x_j \]  

(1)

In this method all variables are considered individually and all pairs of variables are also considered individually. A sample three-layer classification neural network is shown in Figure 2 below.

![Figure 2: Three-Layer Classification Network](image)

2.3 Medical Imaging

Medical image interpretation is usually reliant on experienced radiologists to interpret the imaging results. Inclusion of imaging results in an automated analysis system poses problems. The options for inclusion in the analysis include the findings of the radiologist. In general, the radiologist's analysis of the image that includes specific findings is added to the patient record. These findings can be included in the automated patient record analysis in the form of yes/no findings translated into 1/0 or a textual evaluation which must be interpreted by natural language processing or by human intervention [13].

2.4 Time Series Analysis of Medical Data

For specific diagnoses it may be necessary to observe the patient over a 24-hour time span. The methods require a means for interpreting the data which may consist of more than 100,000 points. The following approach developed by the authors, the Central Tendency Method (CTM), is used [14]. The CTM is defined as

\[ CTM = \frac{\sum_{i=1}^{n} \delta(d_i)}{t-2} \]  

(2)

where

\[ \delta(d_i) = \begin{cases} 1 & \text{if } [(a_{i+2}-a_{i+1})^2 + (a_{i+1}-a_i)^2]^{0.5} < r \\ \text{Otherwise} & \end{cases} \]

Clinical information relevant to the diagnosis of CHF includes: edema, rales, heart rate, and BUN.

2.5 Patient History

Automated notifications are used to alert the physician or other health care professional if changes have occurred in the patient status. This method relies on access to the personal health record of the patient [15]. Since many patients, especially the elderly, are treated by more than one physician, the complete history is most often not available. One solution is to supply the patient with his or her medical history on a USB device. This information would be especially useful for patients seeing multiple physicians for different health issues. Each physician would then know about other conditions as well as current diagnosis, treatments, and medications.

2.6 Automated Historical Comparisons

The algorithm below describes the automated process for evaluation of changes in the patient's health status [16].

**Tracking Changes in the Health of the Patient**

For all confirmed conditions

- If condition i is present with \( \delta(C) = a \)
- If condition i was previously present with \( \delta(C) = b \) and if \(|a-b| > x| \) then send alert
- If condition i previously not present then send alert

For all previously-confirmed conditions

- If condition i not currently present send notification
- If condition i is present with \( \delta(C) = a \)
- If condition i previously present with \( \delta(C) = b \) and if \(|a-b| > x| \) then send notification

This process can be used to effectively determine changes that could have a health impact.
3. Applications

3.1 EMERGE - A Rule-Based Application

The EMERGE system was initially used to evaluate patients coming to the emergency room. Examples of EMERGE rules are given below. They indicate the number of parameters that are sufficient to trigger the conclusion that the patient should be admitted to the hospital [17]. Each rule indicates the level required. For example, SC1 requires only one affirmative answer to admit the patient.

\[ SC_{1}(1,5): \text{ANY OF} \]
\[ \text{Abnormal mental status} \]
\[ \text{Cold, clammy skin} \]
\[ \text{Gray, cyanotic skin} \]
\[ \text{Weak peripheral pulses} \]
\[ \text{Urinary output < 30 cc/hr} \]

\[ SC_{15}(2,3): 2 \text{ of } 3 \]
\[ \text{Sweating} \]
\[ \text{Nausea} \]
\[ \text{Dizziness} \]

\[ SC_{29}(2,2): \text{ALL OF} \]
\[ \text{Pain excruciating} \]
\[ \text{Pain unremitting} \]

The rule base is separate from the engine so that the application can be changed without modifying the software [18]. This approach has the advantage being able to add more rules when new evidence is available as well as removing rules which are obsolete. In addition, if a new application is needed all the existing rules can be easily replaced with a new set. Table illustrates a rule to determine if the patient has myocardial infarction.

3.2 Applications of the Hypernet Neural Network

Hypernet [19] can accommodate any number of antecedents. The only requirement is to have a sufficient number of cases so that a portion can be used to train the network and the remainder can be used to test the network. Each \( x_i \) represents the degree of presence of constituent \( i \) along with a value for \( w_i \) that indicates the relative importance of that constituent. During training each \( w_i \) is adjusted to develop the complete decision surface that can then be used for future classifications.

3.2.1 Rule out Myocardial Infarction

A typical method for disease diagnosis relies on the rule-out method. In this approach data is applied to symptoms of possible diseases to remove those which are not indicated by the available data. In the following case the Hypernet neural network is used to rule out coronary artery disease [20].

The hybrid neural network model along was used to rule out coronary artery disease [21]. The parameters along with weighting factors are shown below. Note that the CTM model is used as the most important factor in this model [22].

**Table I: Rule out Myocardial Infarction (MI)**

<table>
<thead>
<tr>
<th>Weighting of Presence Factor</th>
</tr>
</thead>
<tbody>
<tr>
<td>BP &lt; 100/60</td>
</tr>
<tr>
<td>Abnormal Mental Status</td>
</tr>
<tr>
<td>Cold, clammy skin</td>
</tr>
<tr>
<td>Gray, cyanotic skin</td>
</tr>
<tr>
<td>Weak peripheral pulses</td>
</tr>
<tr>
<td>Urinary output &lt; 30 cc/hr</td>
</tr>
</tbody>
</table>

Then Possible MI \( T \) 0.6

\[ \text{Sum} = 0.5(0.6) + 0.1(0.1) + 0.1(0.1) + 0.1(0.1) + 0.1(0.3) + 0.1(0.6) = 0.48 \]

Results: \( \text{Sum < 0.6, MI not confirmed} \)

3.2.2 Rule out coronary artery disease

Hybrid Neural Network Model/CTM Model

The following are the medical data items that are relevant to the diagnosis of coronary artery disease. Data for \( x_1 \) is obtained from the 24-hour Holter recording. The remainder are test results.

\[ x_1 \text{ Abnormal Holter ECG} \quad 0.75 \]
\[ x_2 \text{ Dyspnea} \quad 0.25 \]
\[ x_3 \text{ Orthopnea} \quad 0.10 \]
\[ x_4 \text{ Edema} \quad 0.40 \]
\[ x_5 \text{ Functional impairment} \quad 0.10 \]
\[ x_6 \text{ PND} \quad 0.10 \]
\[ x_7 \text{ BUN} \quad 0.05 \]

Values are normalized to the interval \([0, 1]\).

3.2.3 Central Tendency Measure (CTM)

This method was applied to evaluate Holter tapes which are 24-hour recordings of cardiac activity [23]. The method is based on the R-R interval which is the time between heartbeats. Figure 3 illustrates the Central Tendency using data obtained.
Two models were examined: the CTM alone and the CTM with clinical data. The CTM alone showed significant differences in the means between normal subjects and CHF subjects although there was significant overlap between the two groups. However, there was no normal subject with a CTM less than 0.62.

Figure 3: Central Tendency Measure Patient with Coronary Artery Disease

Table II shows an example for analysis using the CTM measure alone on a data set with 22 normal patients and 22 patients diagnosed with congestive heart failure (CHF) as well as the analysis using the CTM and clinical measures [24]. The objective was to rule out congestive heart failure. The following scale was used:

- CTM > 90, negative
- CTM between 60 and 90, uncertain
- CTM < 60 CHF

A sample of CTM = 0.75 was labeled as possible coronary artery disease (CAD) with the recommendation to examine clinical data.

Table II: Central Tendency Measure

<table>
<thead>
<tr>
<th>CTM Alone</th>
</tr>
</thead>
<tbody>
<tr>
<td>Normal (22)</td>
</tr>
<tr>
<td>CHF (22)</td>
</tr>
<tr>
<td>Range: 0.623-0.997</td>
</tr>
<tr>
<td>0.15-0.99</td>
</tr>
<tr>
<td>Mean:</td>
</tr>
<tr>
<td>0.90</td>
</tr>
<tr>
<td>0.81</td>
</tr>
</tbody>
</table>

Combination of CTM with clinical parameters Accuracy

| Normal: 84%                   |
| CHF: 82%                      |

3.3 Trend Analysis

The previous sections focused on the diagnosis of current possible diseases. Many patients, especially the elderly, have pre-existing conditions that must be monitored to determine any change in the condition, positive or negative. A positive finding may indicate the need to reduce or discontinue present medications [25]. On the other hand, a negative finding may require additional monitoring or re-evaluation. The following trend analysis approach can be used to make certain the patient is receiving appropriate care for his or her current condition. The algorithm below can alert the medical professional if changes in the health record have occurred.

Trend Analysis in Medical Applications

The following algorithm illustrates the automated analysis to determine changes in a patient’s condition. The analysis is based on the current condition along with the patient history to determine improvement, decay, or stability for the disease under consideration. Although the basic structure of the algorithm does not change sufficient data must be available to use the method in a new medical application.

3.2.1 Trend Analysis Algorithm [26]

For all currently confirmed conditions:

- If condition i is present at time $t_n$ with $\delta(t_n) = a$
- If condition n was previously present with $\delta(t_{n-1}) = b$
- set $\alpha = a - b$
- If condition n was not previously present set $\alpha = a$
- If $\alpha > 0$
  - if ($x_1 < \alpha < x_2$) then send alert level 1
  - if ($x_2 < \alpha < x_3$) then send alert level 2
  - if $\alpha > x_3$ then send alert level 3

For all previously-confirmed conditions:

- If condition is not currently present send notification
- If condition i is present with $\delta(t_n) = a$
- If condition i previously present with $\delta(t_{n-1}) = b$
- and if $(b-a) > x$
  - then send notification of change in degree

4 Conclusions

Recent technological advances are changing the practice of medicine in many ways. Extensive online data assists the physician in diagnosis and potential treatment options. Electronically available disease models along with potential treatments are available for some diseases but not for all and are less available for rare diseases. The use of online information and treatment for all diseases would be useful especially in remote areas without direct access to specialized medical areas. This paper gives an
overview of a number of methods for analyzing medical data to provide potential diagnosis and treatment options. Physicians are beginning to use certified online information to help in both diagnosis and treatment. The work described here addresses several methods of analysis that are necessary to deal with the broad range of medical data formats including numerical, alphanumeric, and images. All types can be combined to present an integrated analysis of the current state of the patient.

**References**


