Case-Based Reasoning System for Chronic Obstructive Pulmonary Disease

Zun Myat Myat Soe, Thin Thin Htike, Mar Mar Min
Computer University (Pathein)
zunmyatmyatsoe@gmail.com

Abstract

Case-Based Reasoning (CBR) is the process of solving new problems based on the solution of similar past problems. The system stores the knowledge of medical experts and the medical records of the previous cases. In medicine, it is very difficult to get correct diagnosis because there are many possible diseases in each case. And it is also important to get diagnosis as soon as possible to save lives of people. It is not easy for a person to consult a medical specialist whenever she feels a symptom of a disease. In this situation, a computer program giving health information provides users to decide whether the symptom is serious or not. The fundamental idea in CBR is problem solving based on the retrieval of similar cases, in its simplest form this is k-nearest neighbor classification. This paper has presented Chronic Obstructive Pulmonary Diagnosis (COPD) System which helps to diagnose a disease stages and treatments by using the basic idea behind CBR and classify the severity of COPD.

1.Introduction

Case-based Reasoning is a recognized and well established method for building medical systems. Case-based Reasoning is memory based, thus reflecting human use of remembered problems and solution as a starting point of new problems solving. Artificial Intelligence, also known as AI can be considered as tools that emulate human thought to help in solving problems. They are robotics, vision, natural language understanding, sound recognition and knowledge-based systems. In these five branches, experts are more interested in knowledge systems because these systems can be applied to any kinds of knowledge in order to solve problems within their domain. There are several approaches using knowledge in solving problems of various domains. These are rule-based reasoning, case-based reasoning and model-based reasoning. Many classification and predication methods are proposed by researchers in machine learning, expect systems, statistics and neurobiology. Some other approaches of classification, such as K-nearest neighbor classifiers, Case-based reasoning, genetic algorithms, rough set and fuzzy logic techniques are applications in a wide range of domains including financial analysis, risk assessment, technical maintenance, process control quality control, medical diagnosis software support systems, forecasting planning, design, classification of objects, photo-interpretation, electronic commerce, customer support, knowledge management and software engineering. CBR in which the diagnosis of new patients is based on physicians' past experience. In this situation, a case could represent a person's symptoms together with the associated treatments. Extraction of meaningful information from large biological datasets is a central theme of many bioinformatics research problems. We have demonstrated in the past a hybrid algorithm consisting of a nearest-neighbors classifier in conjunction with a case-base.

2. Related work

The system describes the CBR status and future prospects of applying diagnosis approaches to data mining. In case-based reasoning, the diagnosis of new patients is based on physicians' past experience and a case could represent a person's symptoms together with the associated treatment [1] and [2]. The authors in [3] and [4], the case retrieved is not identical to the current case, an adaptation phase occurs. During adaptation, differences between the current and retrieved cases are first identified and then the solution associated with the case retrieved is modified, taking differences into account. In [5] and [6], information can be added to a system for two purposes: first, the more information that is stored in a case-base, second, adding information to the case-base generally improves the solution that the system is able to create. In [7]and [8], the k-nearest neighbors principle involves search for the k-NN to the current input case using a distance measure. In [8] and [9], the authors discuss knowledge based system and expert system. In [10] and [11], case retrieval for

proposed system and database design. In [12], discuss briefly about COPD.

3. Case-Based Reasoning Approach

Case-Based Reasoning (CBR) classifiers are instance-based. When given a new case to classify, a case-based reasoned will first check if an identical training case exists. If one is found, then the accompanying solution to that case is returned. If no identical case is found, then the case-based reasoner will search for training cases having components that are similar to those of the four new cases. Conceptually, these training cases may be considered as neighbors of the new case. The case-based reasoner tries to combine the solutions of the neighboring training cases in order to propose a solutions for the new case. If incompatibilities arise with the individual solutions, then backtracking to search for other solutions may be necessary. The case-based reasoning may employ background knowledge and problem-solving strategies in order to propose a feasible combined solution. A previously experienced situation, feature problems is referred to as a past case, previous case, stored case, or retained case. Correspondingly, a new case or unsolved case is the description of a new problem to be solved. Case-based reasoning is in effect a cyclists and integrated process of solving a problem, learning from this experience, solving a new problem.

The nearest-neighbor algorithm is one of the best known classification algorithms and an enormous body of research exit on the subject.

3.1 Case Retrieval of Proposed System

This system is implemented to classify the grading severity of Chronic Obstructive Pulmonary Disease (COPD). The system can be used two types of user depending on their roles. The authorized user can get diagnostic result, see the patients' information, and view the performance analysis with disease stages and treatments.

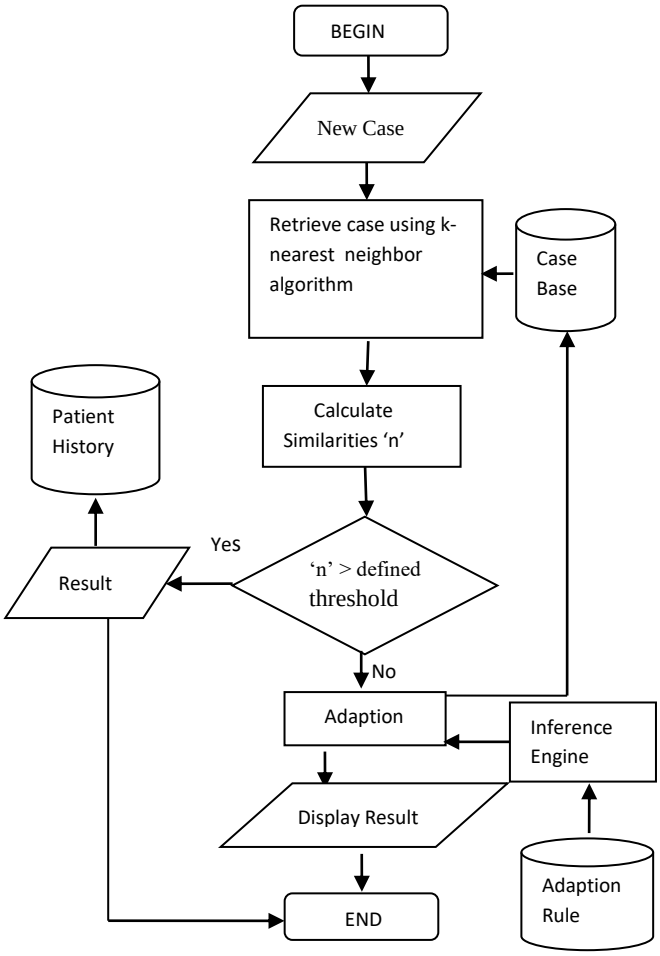

Figure 1. The flow chart of Proposed System

The user can use the system to provide his/her input case by entering the case features on the diagnosis interface. To get the diagnosis of the new case, K-Nearest Neighbor algorithm is used to retrieve in the case-base. In this process, the system will retrieve the case either identical match or the nearest match which is greater than the predefined threshold. If an identical case is found, the system takes this result and presents to user directly. If not the system will retrieve the case which match the new case with the cases which is stored is the greater than threshold and gives result to user. Otherwise, the system uses adaptation rules to get result and gives this result to user.

4. Overview of Case-Based Reasoning

Case-based Reasoning has become successful technique for developing medical systems. Case-based Reasoning is an approach to knowledge-based problem solving and decision support. Case-based

Reasoning based on the intuition that new problems are often similar to previously encountered problems and, therefore, that past solutions may be used in the current situation. Cases are stored and derived from legacy databases. Most CBR system include the following steps: case representation and storage, case matching and retrieval, adaptation of the retrieved solution, forming of new solution and finally, updating of case-base with new solution.

5. System Design

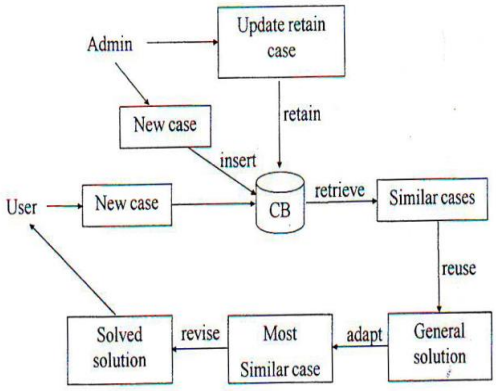

Figure 2. Proposed System Design

6. The CBR cycle

At the highest level of generality CBR cycle may be described by the four process:

1. RETRIEVE the most similar cases.
2. REUSE the information and knowledge in that case to solve the problem.
3. REVISE the proposed solution.
4. RETAIN the parts of experience likely to be useful for future problem solving.

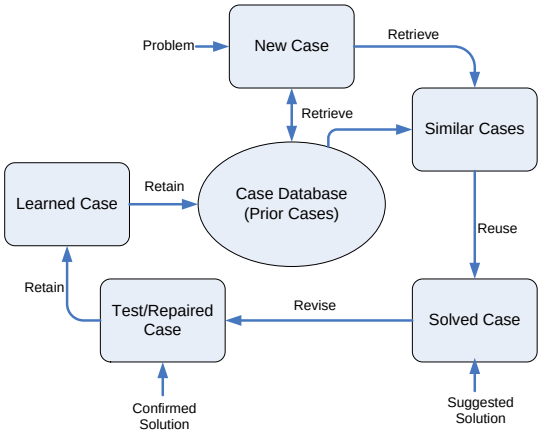

Figure 3. CBR cycle

6.1 Retrieval using Nearest Neighbors Technique

The features of the input case are assigned as indices characterized the case. These indices are used to retrieve a similar past case(s) from the case-memory. The system uses the nearest-neighbor algorithm that finds the closest matches of the cases already stored in the database to the new case to determine how similar two cases are by comparing their features.

The main distance measure equation is

$$Similarity(T,S) = \frac{\sum_{i=1}^{n} f(T_i,S_i) x W_i}{\sum_{i=1}^{n} W_i} \quad (1)$$

Where,

T is the target case.
S is the source case.
n is the number of attributes in each case.
i is an individual attribute from 1 to n.
f is a similarity function for attribute i in cases T and S.
wi is the importance weighing of attribute i.

Thus, the weight is introduced in the case retrieval and the similarity between cases is considered to be weighted summation of the similarity between attributes.

- **Distance for numeric features**

  Distance (ai,bi) = |ai-bi|

- **Distance for symbolic features**

Distance (ai,bi) = { 0 if ai = bi }
{ 1 if ai != bi }

Where, case a has a set of attribute values{a1,….. ,an}.

case b has a set of attribute values{b1,….,bn}.

7. Implementation of the System

This system not only classifies the grading of Chronic Obstructive Pulmonary Disease (COPD) but also gives guidelines for each grade. The system must display the result by calculating with k-NN. The system focuses case representation and then retrieve records sets in this work. These are attributes in the datasets, moreover, patient's ID also include. The datasets are 250 records and 20 attributes. The system only uses 8 attributes, for example, Age, Cough day, Tachypnoeic, Rarely Cyanosed, Heart Oedema,

Peripheral Vasodilation, Severe Hypercapnia, Papilloedema when diagnose the cases into individual groups. Because they are very important features of these disease. There are five stages according to the symptoms is serious or not. These diseases are AVI (Acute Viral Infection), ARI (Acute Respiratory-tube Infection), Asthma, Emphysema, Chronic Bronchitics. At this system, age level is standard under 25, 25 to 40, 40 to 65 and then cough day level is under 35days, 35days to 5months, 5months to 2years. The system retrieves similar cases from the case library by using k nearest neighbors techniques as shown in figure 4 and 5. In a CBR problem information, users can select number of symptoms for generating this system. The system stores the patient's information in the case library. This system has two major parts. admin and user. In administrator part, the admin can insert new cases their corresponding successful solution into the case-base and can update retained cases and can look in the patient's history. In the user part, he/she can use the system to apply the diagnosis of their input case by entering the symptoms and get the treatments concerning with disease stages.

Figure 4. Similar cases

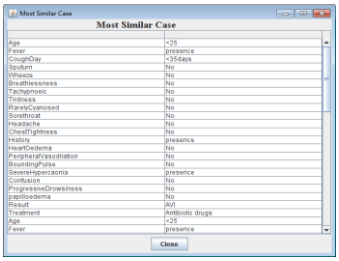

Figure 5. Most similar case

8. Conclusion

Case-based reasoning (CBR) has long been successfully applied in different medical fields. CBR is appropriate for methodology for all medical domains and tasks for the following reasons: cognitive adequateness, explicit experience, duality of objective and subjective knowledge and system integration. CBR presents an essential technology of building intelligent CBR systems for medical diagnose that can aid significantly in improving the decision of the physicians. This thesis has developed a system that uses the advantages of CBR method and supports the decision making in medical diagnosis. The system will be useful as an aid to medical staffs for diagnosis in COPD.

References

[1] J.Kolodner " Case-based Reasoning" , Morgan Kaufmann, San Francisco, pp.27-28, 1993.

[2] D. L. Medin, R, L, Goldstone, and D. Gentner, Respect for similarity, Psychological Review, vol.100, no. 2, pp.254-278, 1993.

[3] B. Smyth and M. T. Keane, Adaptation-guided retrieval: questioning the similarity assumption in reasoning, Artificial Intelligence, vol.102, pp.249-293,1998.

[4] D. Aha at the U. S. Naval Research Laboratory maintains the http://www.aic.nrl.navy.mil/~aha.

[5] K. Racine and Q. Yang, Maintaining unstructured case bases, in Proceedings of the Second International Conference on Case-based Resoning (ICCBR-97), Providence, RI, Springer-Verlag, Berlin, pp. 553-564, 1997.

[6] Q. Xiaoli "A CBR system for Bearing Design".

[7] M. Jaulent, C. Bozec, E. Zapletal, P. Degoulet "A CBR method for Computer Assisted Diagnosis in Histography "AI in Medicine, AIME 97, pp. 239-242.

[8] W. Jan, M. Farhi "Case-based Reasoning: A Review"

[9] I. Watson, Case-based reasoning is a methodology, not a technology, Knowledge Based Systems, vol. 12, pp. 303-308, 1999.

[10] Min Min Ei "Development of Breast Cancer Diagnosis system by using case-based reasoning approach", University of Computer studies, Yangon.

[11] Aye Aye San "Dengue Haemorrhagic Fever Diagnosis System by using case-based reasoning", Computer Studies, Yangon.

[12] Global initiative for Chronic Obstructive Pulmonary Disease (GOLD guidelines). National institute of health, April 2001.