Clustering to Enhance Case-Based Reasoning

Abdelhak Mansoul and Baghdad Atmani

Abstract In this article, we propose an approach to improve CBR processing mainly in its retrieval task. A major difficulty arise when founding several similar cases and consequently several solutions, hence a choice must be done involving an appropriate strategy focusing the best solution. This main difficulty has a direct impact on the adaptation task. To overcome this limitation many works related to the retrieval task were conducted as hybridizing CBR with data mining methods. Through this study, we provide a combining approach using CBR and clustering to reduce the search space in the retrieval step. The objective is to consider only the most interesting cases and the most interesting solution to support decision and provide an intelligent strategy that enables decision makers to have the best decision aid. We also present some preliminary results and suggestions to extend our approach.

Keywords Decision · Support · Case-based reasoning · CBR · Clustering

1 Introduction

The traditional CBR approach has been widely used for enhancing decision aid systems. However it presents some drawbacks mainly in the retrieval task [1–3]. A major drawback is that if process found several similar cases and consequently several solutions which involve looking for a strategy to choose the best solution. Several works were conducted to overcome this drawbacks [2, 4–8] by using

A. Mansoul (✉)
Department of Computer Science, University of Skikda, Skikda, Algeria
e-mail: mansoul21@gmail.com

A. Mansoul
Lio Laboratory, University of Oran 1 Ahmed Ben Bella, Oran, Algeria

B. Atmani
University of Oran 1 Ahmed Ben Bella, Oran, Algeria
e-mail: atmani.baghdad@gmail.com

S. Chikhi et al. (eds.), *Modelling and Implementation of Complex Systems*, Lecture Notes in Networks and Systems 1,
DOI 10.1007/978-3-319-33410-3_10

different strategies with the aim to impact positively the CBR process. In this work, we experiment collaboration between CBR and clustering. Case-based reasoning approach emerged, it was appropriated and widely used to solve problems and support decision in health care, however it presents also some drawbacks in the two mean tasks: the retrieval and the adaptation (reuse and revision) tasks [1–3]. Thus, and as the two tasks are interrelated several works have been conducted on the retrieval task using different strategies which deals with suitable solutions with the aim to impact positively the adaptation task. These solutions are range from simple sequential calculation to non-sequential indexing, classification algorithms such as ID3 and Nearest Neighbor matching. For the retrieval task, a major drawback is that if process found several similar cases and consequently several solutions, hence choice arises for the process and involves looking for a strategy to choose the best solution. Schmidt et al. suggested clustering cases into prototypes and remove redundant ones to avoid an infinite growth of case base, the retrieval searches only among these prototypes [2]. This solution can simplify the adaptation task. Missing similar cases can also occur and may lead to less robust decision due to large number of features. Marling et al. suggested a solution at retrieval task using a three matching algorithms and combined three different measures and fuzzy similarity and they also proposed another solution using a reutilisability measure to select and retrieve a case in addition to check of constraints and a scoring. This method gives the easiest case to adaptation task. This solution was used to propose a menu planer system based on CBR and RBR [6]. Thus, it becomes imperative to review traditional approaches of knowledge processing and to propose new solutions based on new paradigms.

In this work, we experiment a new method by using collaboration between CBR and clustering to propose an available strategy at retrieval task which permits choosing the best solution from a set of solutions found by clustering a case base. We present some preliminary results and suggestions to extend and improve our approach. The rest of this paper is structured as follows: in Sect. 2 we give some notions on CBR and data mining. In Sect. 3 we give a survey on most important related works showing particularly the use of case-based reasoning and data mining that have contributed in decision aid. We continue by presenting our approach in Sect. 4. In Sect. 5 we give a presentation of experimentation and interpretation of results and finally in Sect. 6 we give the conclusion which summarizes the paper and point out a possible trend.

2 Background

Before describing our approach we will give some notions related to decision support to help for understanding the continuation of the paper:

Decision support: "Decision support is the activity that is supported by models clearly explained but not necessarily completely formalized, which helps to obtain some answers to questions asked by an intervener in a decision process,…" [9].

This decision aid is often supported by methods such as statistics, operations research, multi-criteria methods, etc. The decision support involves the development of an action plan or a decision model.

CBR: Reasoning by reusing past cases is a powerful and frequently applied approach to solve problems. The case-based reasoning uses this principle. It is conventionally based on four tasks: retrieve, reuse, revise and retain as shown in Fig. 1. It is widely used in medicine, because of the reasoning method used, which is close to the physician reasoning against a pathological situation. Indeed, finding a medical solution is based on reminding previous cases already experienced for being guided toward a similar situation [10]. Figure 1 shows four steps that a case-based reasoner must perform according to Aamodt and Plaza [11]:

(a) retrieve cases which are similar to the problem description from a case base, this task involves analogy-making or case matching, it is based on reminding,
(b) reuse a solution suggested by the retrieved cases in order to make it applicable to the current case,
(c) revise or adapt the solution to better fit the new case,
(d) retain the new solution once it has been confirmed or validated.

Data mining: Data mining uses a variety of methods and a large volumes of data in order to discover useful knowledge for decision making. Thus, it constitutes a decision aid support in various sectors. We mention here some data mining methods commonly used in the medical field:

- Decision trees. These are structures that represent sets of decisions. These decisions generate rules for the classification of a data set.
- Neural networks. They are at the origin of mathematical modeling of the human brain. They use existing data with a known result to form a model that can later be used to make predictions upon data with unknown result.
- Bayesian networks. They are a directed acyclic graph where each node represents a random variable and the arcs represents probabilistic dependencies between a node and its parents.

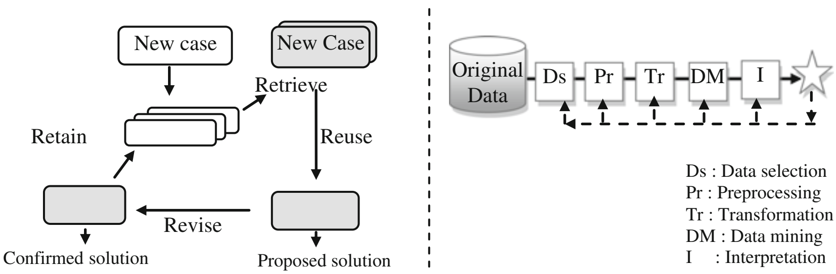

Fig. 1 The CBR cycle and standard data mining process

- The K-nearest neighbors. It is a case reasoning method dedicated to the classification. It allow making decisions by searching for similar cases already resolved.
- Logistic regression. It is a method derived from statistics. It is an extension of ordinary regression.
- Clustering. Its divides data into meaningful groups (clusters) that share common characteristics. It is used in information retrieval.

2.1 Data Mining and CBR

Data Mining can be used for a variety of purposes in Case-Based Reasoning. Some uses are listed here:

- Find features for a case (in a case base): It might be interest to classify the cases for use.
- Find features for a case (in a database): A data base can be searched to complete the information given in a case.
- Find domain knowledge: Domain knowledge might be mined from the data in the form of functions, rules or causal graphs which can be used later by a CBR process.
- Construct "artificial cases": We should be able to build cases from a database.

3 Related Works

As mentioned in introduction, there is a large body of work attempting to use CBR in medical fields particularly in the diagnosis of diseases. Many works have addressed directly the use of CBR while other works were conducted with other approaches and many recent works have emerged and have contributed to the advancement of research. We notice the following concerns.

3.1 Use of Case-Based Reasoning

Many studies concerning CBR in decision support were conducted. Marling et al. presented an approach for treatment of patients suffering from diabetes [12]. Jha et al. presented a work for diabetes detection and care [13]. Bresson and Lieber created CASIMIR, a system for treatment of breast cancer [14]. Shanbezadeh et al. proposed an oriented decision support system for treatment of asthma [15]. Song et al. proposed a system for dose planning in radiotherapy for prostate cancer [16].

Begum et al. presented a decision support system to help physicians to diagnose the state of stress [17]. De Paz et al. presented a decision support system for the diagnosis of different types of cancer [18]. Schwartz et al. used also CBR to enhance care on insulin therapy [19]. This list is not exhaustive but it shows the diversity of utilization of CBR and underlines the interest for this approach to improve the care of patients by providing physicians with information technology tools for medical decision support.

3.2 Use of Data Mining

Owing to the large volume of data manipulated by health structures, it became imperative to take into account this mass of medical data to improve medical practice and at the same time improve the care practiced by clinicians. These methods and particularly the decision trees and the neural networks have been used in many studies in medicine [17, 20]. We give a non-exhaustive list of some research works which uses data mining methods in a medical field. Sivakumar presented a method based on the neural networks to classify subjects with diabetic retinopathy (common complications of diabetes) [21]. This algorithm generates a set of rules for the diagnosis and prediction of diabetes. Kiezun et al. have used logistic regression to help clinicians in the diagnosis of myocardial infarction (heart attack) in patients presenting chest pain [22]. Malyshevska studied the problem of cancer diagnosis using the neural networks [20]. The aim of this study is to classify different types of tissue that are used to determine the risk of cancer. Sung and Seong recently done a study based on the building of a hybrid method, combining methods of data mining (association rules, classification trees) to help clinicians to classify more faster and more accurate chest pain diseases [23]. Xu Li used an indexing/matching module based on retrieving only cases that match the important indices of the new case, calculate an aggregate match score for the comparable cases and retrieve only those comparable cases with higher aggregate match scores [7]. Kumar et al. used two distances (Weighted Euclidean, Mahalanobis) to perform retrieval task and eliminate bad cases with an eliminating score [8]. Macura and Macura used a retrieval-only system to avoid the adaptation task, because the users wish to see and interpret all specific details of the cases themselves without going until adaptation task [24]. Bichindaritz and Marling have used combination of RBR and model based components but this strategy can't be seen as solution for CBR drawbacks (for retrieval or adaptation), but as an opportunity to enhance CBR subtasks instead of using an older technique. Recently, a major trend seems to be the extending of CBR applications beyond the traditional diagnosis or treatment toward the applicability of CBR to new reasoning tasks [1].

3.3 Combining CBR with Other Approaches

Combining CBR with other approaches (Multi-Modal Reasoning) represents another way used to avoid the adaptation problem, mainly by combining the retrieval task with other reasoning strategies, to provide decision support. The interest in multi-modal approaches involving CBR has reached the medical areas [2]. This is an issue of current concern in CBR research in different fields as medicine and others [10]. The first multi-modal reasoning system in the health care was CASE, it integrated CBR with model-based reasoning (MBR) for diagnosis of heart failures [6]. Araujo et al. have combined rule-based reasoning (RBR) and CBR to recommend neuroleptic drugs for Alzheimer's patients [25]. Althoff et al. have integrated induction and CBR for diagnostic [4]. Janetzko and Strube also tried to combine CBR with knowledge processing [26]. Li and Sun hybridized multi-criteria and CBR to enhance a data mining process for improving detection of disease [27]. Armaghan and Renaud used also combination of CBR and multi-criteria to study diabetes [5]. Angehrn and Dutta used also this combination to study diabetes [28]. Royes used multi-criteria and CBR for strategic planning support [29]. Araujo de Castro et al. used a hybrid model based on multi-criteria and CBR to diagnosis of Alzheimer's disease [25]. Other researchers have proposed hybrid solution by combining CBR with other techniques as reasoning by rules and many works emerged among these studies [6, 30].

4 The Proposed Approach

The medical situation we advocate is described by the decision maker (physician) is in front of a diagnosis of a situation and will have to explore possible options (diagnosis) to choose the best therapy. This situation is characterized by: a problem definition more or less complete, an exhaustive survey of possible diagnosis and the existence of specific signs for each patient such as for example "elderly patient", "allergy to penicillin", etc. these specific signs can indicates that a desired therapy will be more or less compliant as an elderly patient may be less compliant with a salt diet for example. Moreover, it is well recognized today that diagnosis decisions related to each patient must take into account particular signs (drug risk, medication side effects, dosages, etc.). Moreover, the physician reasons when searching for a therapy such a system which uses old situations (cases) in order to propose a similar or a best therapy.

Thus, the physician defines his medical situation with a set of symptoms and a environmental context described by the specific signs. The medical situation becomes a medical case composed of u specific signs, and v symptoms which are descriptive of the case and w descriptors giving the "diagnosis/therapy" considered for the case in question.

Thus, the medical situation will be defined as follows:

Medical situation={Specific_Signs, Symptoms, Diagnosis}.

Medical_case
  Specific_Signs : ss_1=$value_1$, …., ss_u=$value_u$
  Symptoms : $symptom_1$=$value_1$, …., $symptom_v$=$value_v$
  Diagnosis : $Diagnosis_1$, …., $Diagnosis_w$
End_medical_case.

The contribution of data mining. However, specific signs can guide search space reduction while using a clustering technique. Clustering does not aim at labeling the cases in a group with a specific tag (as it happens in classification), where the tag represents a piece of generalized domain knowledge, extracted from the subsumed cases. In clustering specific signs remains enables collecting the most similar cluster (s) allow the identification of the cases collected under similar circumstances, and the limitation of retrieval just to them. In result we can have exploitation of prototypes which are a generalization from single to clustered typical cases. Their main purposes are to structure the case base and to guide and speedup the retrieval process.

4.1 Processing

In this work, we experiment a new method by using collaboration between CBR and data mining to propose an available strategy at retrieval task which permit choosing the best solution from a set of solutions found by mining cases by a so called constraint based clustering.

The constraint based clustering process: We chose a rational approach for retrieval task: instead of a massive retrieval of cases that is the classic recipe of reasoning, we analyze the cases rationally and we focus research on particular perimeters with specific cases that are the subject of suspicion or potential cases. Our aim is to find closest cases on all previously treated cases in order to avoid complication at the adaptation phase and make it arduous. Indeed, it doesn't matter to collect all the closest cases, but it should rather focus on a small perimeter of cases. So, we must proceed otherwise than by classical method. Thus, our approach focuses on reducing the perimeter of research, and then retrieve. We will call it: *MineR* for Mining and Retrieve.

From there, the clustering operation is guided by specific signs to select a subset of cases contained in the case base. Thus, reducing the search space solutions (diagnoses) for the retrieval step becomes an essential operation for the CBR process. This choice can clearly make retrieval computationally better regarding to only interesting solution for the case being processed and hopefully more meaningful, since only cases taken under comparable circumstances are retrieved. The set of closest diagnoses (*Closest_Diags*) is received from the CBR_Process to join it to proposed diagnoses (*Proposed_Diags)* that the user has already defined in the medical situation as described below (see also step C). This process will be handled by the following steps:

Step A: Cases selection

A this step, the specific signs help to filter the cases so as to keep only cases which meet only those constraints, then we proceed successively by steps B and C.

Step B: Preprocessing

At this step, we will only prepare data for clustering as cleaning data or appropriate treatment.

Step C: Clustering

At this step, we will launch the suitable clustering method. This entire step will be initiated by the following pseudo_algorithm which will generate the best cluster (Best_cluster) for processing later the best diagnosis (Best_Diag) by the CBR.

The CBR process: This process has a main task: the matching. It consists in finding the n closest cases to the proposed case by using a similarity measure. We used the K-nn method for the simplicity of its implementation. The process will select closest or similar cases from the best cluster (Best_cluster) proposed by clustering process, and will extract the preliminary closest diagnoses (Closest_Diags) that have been considered for the *n* similar cases, and then those preliminary closest diagnoses are considered to determine the best diagnosis (Best_Diag).

All this two process will be handled by the following pseudo-algorithm.

```
Pseudo_algorithm : CBR_Miner_Process
 1: Input : Closest_Diags • Ø
 2: New_Case(Ss,Sy,Proposed_diags)
 3: Begin MineR_Process()
     a) Best_cluster • Clustering()
     b) Closest_cases • Extraire(Best_cluster)
     c) Accept_or_refuse(Best_cluster)
     d) If accept = "yes" then go to step 5
          Else Review any step: 2 or 3
       Endif
 4: End MineR_Process()
 5: Initialize k
 6: Closest_Diags=Proposed_diags
 7: Closest_cases•Retrieve(New_Case,Best_Cluster,K-
    nn)
 8: If Closest_Cases•Ø then
      For each  Current_Case in Closest_Cases
      For  i=1 to n
       {Closest_Diags• Closest_Diags U Current_Case(
       Diagosis_i)}
      Endfor
      Endfor
          Else
  EndIf
 9: Reuse()
10: Revise()
11: If confirmed_solution="yes" then
     Specific_Signs =Ss
     Symptoms=Sy
     Diagnosis=Best_Diag
     Produce_New_Case(Specific_Signs,Symptoms,Diagnosis)
     Retain_New_Case()
    Endif
      New_Case(Ø,Ø,Ø)
12: End
```

5 Implementation and Experimentation

The proposed approach has been applied to a medical datasets the Vertebral Column Data Set of orthopaedic patients, we project to use the presumptive diagnosis of diseases of orthopaedic patients data set which is a data set containing values for six biomechanical features used to classify orthopaedic patients into 3 classes (normal, disk hernia or spondylolisthesis) [31]. Each patient is represented in the data set by six biomechanical attributes derived from the shape and orientation of the pelvis and lumbar spine (in this order): pelvic incidence, pelvic tilt, lumbar lordosis angle, sacral slope, pelvic radius and grade of spondylolisthesis. The following convention is used for the class labels: DH (Disk Hernia), Spondylolisthesis (SL), Normal (NO) and Abnormal (AB).

5.1 Data Description

This data contains information about diseases of orthopaedic patients (normal, disk hernia or spondylolisthesis) of a patient based on his characteristics. Figure 2 gives an overview of data set sample.

Each patient is described by seven descriptors, with the last attribute that contains the results of the diagnosis. It contains [31]:

1. pelvic incidence (numerical)
2. pelvic tilt (numerical)
3. lumbar lordosis angle (numerical)
4. sacral slope (numerical)
5. pelvic radius (numerical)
6. grade of spondylolisthesis (numerical)
7. diagnosis: DH (Disk Hernia), Spondylolisthesis (SL), Normal (NO) and Abnormal (AB).

Data Set sample
39.05695098,10.06099147,25.01537822,28.99595951,114.4054254,4.564258645,Hernia
68.83202098,22.21848205,50.09219357,46.61353893,105.9851355,-3.530317314,Hernia
83.93300857,41.28630543,61.99999999,42.64670314,115.012334,26.58810016,Spondylolisthesis
78.49173027,22.1817978,59.99999999,56.30993248,118.5303266,27.38321314,Spondylolisthesis

Fig. 2 Overview of presumptive diagnosis of diseases of orthopaedic patients sample [31]

Table 1 Case base descriptors

Descriptor	Label
X_1	Pelvic incidence
X_2	Pelvic tilt
X_3	Lumbar lordosis angle
X_4	Sacral slope
X_5	Pelvic radius
X_6	Grade of spondylolisthesis
Y	Diagnosis

For the purposes of our experimentation we have transformed Presumptive diagnosis of diseases of orthopaedic patients data set descriptors into a case base descriptors and each case will be described by the set of variables X_1, X_2,…, $X_{6,}$ called descriptive variables and we have associated a target attribute denoted Y corresponding to diagnosis. The following Table 1 shows case base descriptors.

5.2 Construction of Case Base Ω_N and Partial Case Bases Ω_L, Ω_T

To evaluate the efficacy of our approach, we have transformed the presumptive diagnosis of diseases of orthopaedic patients date set into a case base named $\Omega_{N.}$ It contains a number of cases $\omega_i \cdot \Omega_N = \{\omega_1, \omega_2, \ldots, \omega_n\}$, each case is described by the set of variables X_1, X_2,…, $X_{6,}$ called descriptive variables. For each case ω_i we associate a target attribute denoted Y, which takes its values in the set Y = {DH, SL, NO, AB} corresponding to diagnosis where DH = “Disk Hernia”, SL = “Spondylolisthesis”, NO = “Normal” and AB = “Abnormal”. Table 2 shows some cases noted ω_1, ω_2,…, ω_n of Vertebral Column Case Base.

After construction of case base Ω_N, we subdivide it on a learning case base Ω_L (80 % of Ω_N) and a test case base Ω_T (20 % of Ω_N) by separating the population Ω_N as follows in Table 3.

Table 2 Ω_N

ω	$X_1(\omega)$	$X_2(\omega)$	$X_3(\omega)$	$X_4(\omega)$	$X_5(\omega)$	$X_6(\omega)$	Y (ω)
ω_1	63.0278175	22.55258597	39.60911701	40.47523153	98.67291675	−0.254399986	Hernia
…							
ω_i	44.529051	9.433234213	51.99999999	35.09581679	134.7117723	29.10657504	Spondylolisthesis
…							
ω_n							

Table 3 Partial case bases

Case base Ω_N	Learning case base Ω_L 80 %	Testing case base Ω_T 20 %
310	248	62

5.3 Implementation/Experimentation

Experiments are conducted on an interactive system developed in JAVA with an interconnection module to JCOLIBRI system [32]. This system is essentially based on an engine described by Fig. 3. We use the JCOLIBRI platform to build the case base and all the relative operations for CBR process. In first step clustering process is initiated, this operation is done under WEKA platform, helped with the specific signs to reach the different clusters, then the results of this platform is the generation of the best cluster (Best_cluster) that will be returned to CBR process for collaboration to improve the decision support. The final objective of the whole process is to collaborate for deciding about the best diagnosis to each new case (medical situation) given as input.

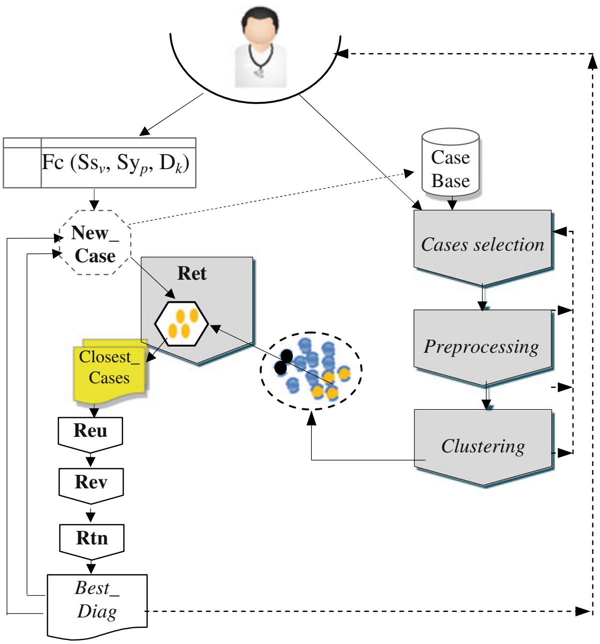

Fig. 3 Overview of the different steps involved by the approach. *Rtv* Retrieve similar cases, *Reu* Reuse a suggested solution, *Rev* Revise or adapt the solution, *Rtn* Retain the new solution, *Fc* (Ss_v, Sy_p, D_k) Formulation_of_case (*v* Specific_Signs, *p* Symptoms, *k* Diagnosis)

Table 4 Comparison results of three data sets using the error rate (ER)

Number of Cases tested from Ω_T	Type of diagnosis for tested cases	Misclassified cases	ER (%)
20	Normal	*4*	*20*
20	Disk Hernia	*7*	*35*
20	Spondylolisthesis	*8*	*40*

We have considered 20 cases randomly taken from the testing case base Ω_T without any hypothesis of diagnosis. A comparison of each case under Ω_T is done with the learning case base Ω_L as follows:

$$\begin{array}{c}\forall \omega_i \in \Omega_T \\ \text{and} \\ \forall \omega_j \in \Omega_L\end{array} \quad \text{if} \begin{cases} y(X(\omega_{i)} = y(X(\omega_j)) & \text{then good matching} \\ \quad \text{else} \quad \text{mismatch} \end{cases} \tag{1}$$

Thereafter, with the conditional structure (1), we calculate the rate of correct matching. This rate represents the number of correctly identified cases in learning case base Ω_L and identically diagnosed in testing case base Ω_T. The test results are presented in Table 4.

5.4 Evaluation of Results

To evaluate the efficiency of our approach, a comparison of each case tested from Ω_T is made in Ω_L. We calculate the error rate of each type of diagnosis. From results presented in Table 4, we note that the error rate is relatively low (lower than average) which indicates that our system tends to give answers close to reality.

According to these results, we note that the rate of correct matching (similar diagnosis) is relatively high compared to the average which indicates that the system provides results closer to the reality as declared in testing case base Ω_T answers particularly for good matching cases. We note also that the rate of good matching is more over than the average which indicates that our system tends to recognize and make a good matching of diagnosis.

6 Conclusion and Future Trends

This study provides the theoretical basis of an approach that tends to solve a problem of CBR reasoning. Later, we intend to evolve our approach in another orientation by using the concept of clustering in a form a rule and where each rule define a cluster then make a comparative study.

References

1. Bichindaritz, I., Marling, C.: Case-based reasoning in the health sciences: foundations and research directions. Computational Intelligence in Healthcare 4, pp. 127–157. Springer, Berlin (2010)
2. Schmidt, R., Montani, S., Bellazzi, R., Portinale, L., Gierl, L.: Cased-based reasoning for medical knowledge-based systems. Int. J. Med. Inform. **64**(2), 355–367 (2001)
3. Montani, S.: Exploring new roles for case-based reasoning in heterogeneous AI systems for medical decision support. Appl. Intell. **28**(3), 275–285 (2008)
4. Althoff, K.D., Bergmann, R., Maurer, F., Wess, S., Manago, M., Auriol, E., Conruyt, N., Traphoner, R., Brauer, M., Dittrich, S.: Integrating inductive and case-based technologies for classification and diagnostic reasoning. In: Proceedings ECML-93 Workshop on Integrated Learning Architectures (1993)
5. Armaghan, N., Renaud, J.: An application of multi-criteria decision aids models for case-based reasoning. Inf. Sci. **210**, 55–66 (2012)
6. Marling, C., Rissland, E., Aamodt, A.: Integrations with case-based reasoning. Knowl. Eng. Rev. **20**(3), 241–245 (2005)
7. Xu, L.D.: An integrated rule-and case-based approach to AIDS initial assessment. Int. J. Biomed. Comput. **40**(3), 197–207 (1996)
8. Kumar, K.A., Singh, Y., Sanyal, S.: Hybrid approach using case-based reasoning and rule-based reasoning for domain independent clinical decision support in ICU. Expert Syst. Appl. **36**(1), 65–71 (2009)
9. Roy, B.: Méthodologie multicritère d'aide à la decision. Paris Economica (1985)
10. Begum, S., Ahmed, M., Funk, P., Xiong, N., Folke, M.: Case-based reasoning systems in the health sciences: a survey of recent trends and developments. IEEE Trans. Syst. Man Cybern. Part C Appl. Rev. **41**(4), 421–434 (2011)
11. Aamodt, A., Plaza, E.: Case-based reasoning: foundational issues, methodological variations, and system approaches. AI Commun. **7**(1), 39–59 (1994)
12. Marling, C., Jay, S., Schwartz, F.: Towards case-based reasoning for diabetes management: a preliminary clinical study and decision support system prototype. Comput. Intell. **25**(3), 165–179 (2009)
13. Jha, M.K., Pakhira, D., Chakraborty, B.: Diabetes detection and care applying CBR techniques. Int. J. Soft Comput. Eng. **2**(6), 132–137 (2013)
14. Bresson, B., Lieber, J.: Raisonnement à partir de cas pour l'aide au traitement du cancer du sein. Actes des journées ingénierie des connaissances, pp. 189–196 (2000)
15. Shanbezadeh, M., Soltani, T., Ahmadi, M.: Developing a clinical decision support model to evaluate the quality of asthma control level. Middle-East J. Sci. Res. **14**(3), 387–393 (2013)
16. Song, X., Petrovic, S., Sundar, S.: A case-based reasoning approach to dose planning in radiotherapy. In: 7th International Conference on Case-based Reasoning ICCBR, pp. 348–357 (2007)
17. Begum, S., Ahmed, M.U., Funk, P., Xiong, N., Von Schéele, B.: A case-based decision support system for individual stress diagnosis using fuzzy similarity matching. Comput. Intell. **25**, 180–195 (2009)
18. De Paz, F.J., Rodriguez, S., Bajo, J., Corchado, M.J.: Case-based reasoning as a decision support system for cancer diagnosis: a case study. Int. J. Hybrid Intell. Syst. **6**(2), 97–110 (2009)
19. Schwartz, F.L., Shubrook, J.H., Marling, R.: Use of case-based reasoning to enhance intensive management of patients on insulin pump therapy. J. Diab. Sci. Technol. **2**(4), 603–611 (2008)
20. Malyshevska, K.: The usage of neural networks for the medical diagnosis. International Book Series. Inf. Sci. Comput 77–80 (2009)
21. Sivakumar, R.: Neural network based diabetic retinopathy classification using phase spectral periodicity components. ICGST-BIME J. **7**(1), 23–28 (2010)

22. Kiezun, A., Lee, I.T.A., Shomron, N.: Evaluation of optimization techniques for variable selection in logistic regression applied to diagnosis of myocardial infarction. Bioinformation **3**, 311–313 (2009)
23. Ha, S.H., Joo, S.H.: A hybrid data mining method for the medical classification of chest pain. Int. J. Comput. Inf. Eng. **4**(1) 33–38 (2010)
24. Macura, R.T., Macura, K.J.: Macrad: radiology image resource with a case-based retrieval system. Case-Based Reasoning Research and Development, pp. 43–54. Springer, Berlin (1995)
25. Araujo de Castro, A.K., Pinheiro, P.R., Dantas Pinheiro, M.C.: Towards the neuropsychological diagnosis of Alzheimer's disease: a hybrid model in decision making. WSKS, CCIS **49**, 522–531 (2009)
26. Janetzko, D., Strube, G.: Case-based reasoning and model-based knowledge-acquisition. Contemp. Knowl. Eng. Cogn. Lect. Notes Comput. Sci. **622**, 97–114 (1992)
27. Li, H., Sun, J.: Hybridizing principles of the Electre method with case-based reasoning for data mining: Electre-CBR-I and Electre-CBR-II. Eur. J. Oper. Res. **197**(1), 214–224 (2009)
28. Angehrn, A.A., Dutta, S.: Integrating case-based reasoning in multi-criteria decision support systems. INSEAD (1992)
29. Royes, G.F.: A hybrid fuzzy-multicriteria-CBR methodology for strategic planning support. In: Processing NAFIPS'04, Annual Meeting of the Fuzzy Information, vol. 1, pp. 208–213 (2004)
30. Verma, L., Srinivasan, S., Sapra, V.: Integration of rule based and case-based reasoning system to support decision making. In: International Conference on Issues and Challenges in Intelligent Computing Technics (ICICT), pp. 106–108. IEEE (2014)
31. http://archive.ics.uci.edu/ml/datasets/Vertebral+Column#
32. Bello-Tomás, J.J., González-Calero, P.A., Díaz-Agudo, B.: Jcolibri: an object-oriented framework for building CBR systems. Advances in Case-Based Reasoning, pp. 32–46. Springer, Berlin (2004)
33. Bouhana, A., Abed, M., Chabchoub, H.: An integrated case-based reasoning and AHP method for personalized itinerary search. In: 4th International Conference on Logistics, pp. 460–467. IEEE (2011)
34. John, D.A., John, R.R.: A framework for medical diagnosis using hybrid reasoning. In: Proceedings of the International Multi Conference of Engineers and Computer Scientists, vol. 1 (2010)
35. Bichindaritz, I., Montani, S.: Introduction to the special issue on case-based reasoning in the health sciences. Comput. Intell. **25**(3), 161–194 (2009)
36. Pandey, B., Mishra, R.B.: Data mining and CBR integrated methods in medicine: a review. Int. J. Med. Eng. Inform. **2**(2) (2010)
37. Zhuang, Z.Y., Churilov, L., Burstein, F., Sikaris, K.: Combining data mining and case-based reasoning for intelligent decision support for pathology ordering by general practitioners. Eur. J. Oper. Res. **195**(3), 662–675 (2009)
38. Yuan, G., Hu, J., Yinghong, P.: Research on CBR system based on data mining. Appl. Soft Comput. **11**(8) 5006–5014 (2011)
39. Hall, M., Frank, E., Holmes, G., Pfahringer, B., Reutemann, P., Witten, I.H.: The WEKA data mining software: an update. ACM SIGKDD Explor. Newslett. **11**(1), 10–18 (2009)
40. Holmes, G., Donkin, A., Witten, I.H.: Weka: a machine learning workbench. In: Proceedings of the 1994 Second Australian and New Zealand Conference on Intelligent Information Systems, pp. 357–361. IEEE (1994)