Case-Based Decision Support System with Contextual Bandits Learning for Similarity Retrieval Model Selection

Booma Devi Sekar[(✉)] and Hui Wang

School of Computing, Ulster University, Jordanstown Campus,
Newtownabbey, Northern Ireland, UK
{b.sekar,h.wang}@ulster.ac.uk

Abstract. Case-based reasoning has become one of the well-sought approaches that supports the development of personalized medicine. It trains on previous experience in form of resolved cases to provide solution to a new problem. In developing a case-based decision support system using case-based reasoning methodology, it is critical to have a good similarity retrieval model to retrieve the most similar cases to the query case. Various factors, including feature selection and weighting, similarity functions, case representation and knowledge model need to be considered in developing a similarity retrieval model. It is difficult to build a single most reliable similarity retrieval model, as this may differ according to the context of the user, demographic and query case. To address such challenge, the present work presents a case-based decision support system with multi-similarity retrieval models and propose contextual bandits learning algorithm to dynamically choose the most appropriate similarity retrieval model based on the context of the user, query patient and demographic data. The proposed framework is designed for DESIREE project, whose goal is to develop a web-based software ecosystem for the multidisciplinary management of primary breast cancer.

Keywords: Case-based reasoning · Clinical decision support system
Similarity retrieval · Contextual bandits learning

1 Introduction

Recent advances in healthcare industry show that there is a growing demand for personalized medicine, which aims to customize treatment to an individual patient based on his/her likelihood of response to the therapy. The move towards personalized medicine is supported by various technological advancements, especially in the area of data science, machine learning and artificial intelligence [1]. One such pathway is the development of personalized diagnostic model based on patient similarity. Case-based Reasoning (CBR), an artificial intelligent approach is very close to human reasoning, and has become a well-adapted methodology in medicine for developing personalized diagnostic model based on patient similarity measure [2]. CBR methodology adapts instance based learning, which aims to learn and derive insights from patients similar to the query patient and then analyze the derived insights in the diagnostic model to

provide personalized diagnostic/treatment recommendations to the query patient. In this paper, we present a Case-based Decision Support System (CB-DSS) for DESIREE[1], which is a European Union funded project focusing on developing a web-based software ecosystem for the personalized, collaborative, and multidisciplinary management of primary breast cancer (PBC), from diagnosis to therapy and follow-ups.

The main difference between a case-based and a rule-based system is that, the knowledge base of a case-based system is populated with cases that incorporates experts experience rather than rules defined using clinical guidelines. Secondly, in a rule-based system it is difficult to pre-define rules that explicitly match any problem, and therefore often fails to solve some of the complex problems. In a case-based approach however, a partial matching is built within the system, which allows it to provide an approximate solution to a problem. Advantage of this is that a CBR system could provide a solution to any given problem, but the challenges allies in building a CBR system that could provide a more reliable solution. This in fact, mainly relies on the first step of the CBR cycle (retrieve, reuse, revise and retain), the similarity retrieval. Building a good similarity retrieval algorithm involves various factors, from assigning proper weights to the description variables, adapting the appropriate similarity function (e.g. cosine, Euclidean, distance correlation) and incorporating general domain knowledge using ontology. Depending on different combination of these factors, the similarity retrieval algorithm could retrieve completely different cases from the case-base and thus provide varying solution to the same problem. Thus, the main challenge of building a CBR system is that there are various uncertainties involved to develop a more reliable model.

In this paper, to address the above challenge, we first present a case-based decision support system (CB-DSS) framework with multi-similarity retrieval models and propose contextual bandits learning algorithm [3] to dynamically choose the appropriate model relevant to the context of the user, query patient and demographic setting. In the following section, we first present the framework and workflow of the proposed CB-DSS. Then, we present the contextual bandit learning methodology and it's adaptation in the framework to dynamically choose between different similarity retrieval model based on the context data. Finally, we draw the conclusion.

2 Case-Based Decision Support System with Contextual Bandit Learning

2.1 Framework and Workflow

One of the main objectives of DESIREE project is to provide decision support for diversity of therapeutic options in BUs, including surgical, radiotherapy, adjuvant systemic therapies etc. With the aim to provide personalized state-of-the-art clinical decision support system to BUs, the project aims at providing guideline-DSS (GL-DSS) [4], experience-DSS (EX-DSS) [5] and CB-DSS. In this paper, we present the

---

[1] http://www.desiree-project.eu.

proposed CB-DSS using CBR methodology and contextual bandits learning algorithm. Figure 1 shows the framework and workflow of the proposed CB-DSS.

In order to incorporate decisional criteria beyond the limitations of current guidelines from breast cancer management, the CB-DSS incorporates the experience of clinicians on previous cases, by collecting description of patients, and the decision made by the clinicians, as the case representation in the data model. Also, to incorporate the knowledge from explicit domain (breast cancer), the data acquired from the clinical partners, clinical practice guidelines and clinical documentation are represented as breast cancer knowledge model (BCKM) in a Web Ontology Language (OWL), which can then be applied in similarity retrieval model using semantic similarity functions. Finally, with the feature selection made, feature weighting matrix is defined, which is also applied in the similarity retrieval model.

Next, as shown in Fig. 1, it provides a tool for querying former patient cases using similarity retrieval model. As briefed above, various factors are involved in the design of the similarity retrieval model, from feature selection and weighting, case representation, and similarity function matrix. Thus, combination of these different factors could lead one to build a completely different similarity retrieval model. For example, for a surgeon certain clinical attributes are more important than for radiologist or a general physician, therefore when defining the feature weighting matrix, one has to take into consideration the context of the user. Likewise, an oncologist will consider various non-clinical attributes of the patient such as the race, family history, and insurance status in recommending a treatment plan. The context of the patient plays an important role in making a decision. Thus, it is critical to consider the contextual information in building a similarity retrieval model.

To address the above challenges, the proposed CB-DSS framework shown in Fig. 1 is built with N number of similarity retrieval models and contextual bandit learning is proposed to exploit the context data of the user, patient and demographic data to dynamically choose the most appropriate similarity retrieval model.

Now, during runtime, the user first enters his/her details, demographic data and query patient case to the CB-DSS. The context extraction module, aggregates the context data, such as clinician's practice data, demographic data, patient family history, race etc., which are then used by the contextual bandits learning algorithm. Meanwhile, the query patient data also enters the different similarity retrieval models present in the CB-DSS. Based on the defined similarity functions and weight matrix, the similarity retrieval model compares the query case with the patient cases present in the data model to retrieve similar patient cases. The contextual bandits learning algorithm will make the decision on which similarity retrieval model will be executed by the DSS. Next section, will discuss the details on how the contextual bandits learning algorithm selects the best performing similarity retrieval model based on the contextual information.

2.2   Contextual Bandit Learning

Determining the best similarity retrieval model can be viewed as a multi-armed bandit (MAB) problem [6], where the clinician has to choose amongst a set of available arms (retrieved patient cases) and he/she can only receive the reward (see the outcome) of the
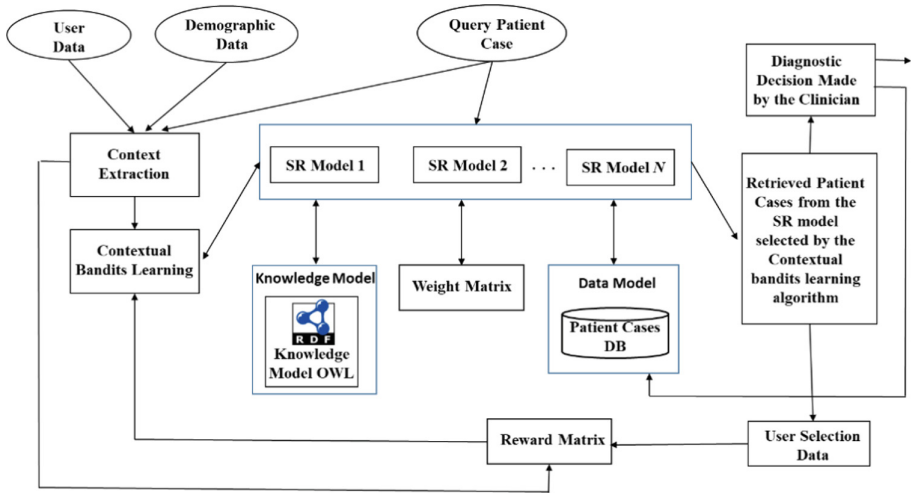
Fig. 1. Framework and workflow of CB-DSS with contextual bandits learning

action (diagnostic decision) that was taken. The clinician will not be able to know the possible outcome, if his/her decision was based on the choice of the patient case retrieved from a different case retrieval model.

To solve the above problem, the proposed algorithm should learn to choose the actions that can maximize the rewards (choose the decision from best performing model). A contextual bandit learning algorithm addresses such problem by providing the context (a hint about the reward) before the action is actually taken. The context in our model that can determine the reward can be derived from various factors, such as the demographic data (breast unit in the hospital), user (e.g. radiologist, oncologist or general physician), and query patient (physiological data, such as race).

Now with the above information, a contextual bandit framework can be defined as follows. Let X be the context set and A be the arms set (action). In each round of the algorithm t, $t = 1, \ldots, T$, T is the time zone, the following events is executed in succession:

1. a context $x_t \in X$ is observed by the learner,
2. based on the observed context, a reward vector $r_t \in [0, 1]^K$ is chosen, but not received by the learner,
3. learner chooses an arm (action) $a_t \in \{1, \ldots, K\}$,
4. learner receives the reward $r_t(a_t)$.

Now, the goal of a bandit algorithm is to maximize the total reward $\sum_{t=1}^{T} r_t(a_t)$. So, in order to maximize the reward, the algorithm should execute a good policy π (e.g. decision rule) to allow the learner to choose an action based on the context. The algorithm will have to work in a rich policy space $\Pi = \{\pi : X \to A\}$ that could be extremely large. Thus, it has to efficiently learn about all policies and choose the best policy. Therefore, when the arm is selected, the learner will observe reward for policies

that would have chosen the same arm. Now, the aim is to obtain a high total reward relative to the best policy $\in \Pi$, computed as minimum contextual regret C_r as shown in Eq. (1). Where the first term in Eq. (1) is the average reward for the best policy and the second term is the learner's average reward.

$$C_r = \max_{\pi \in \Pi} \frac{1}{T} \sum\nolimits_{t=1}^{T} r_t(\pi(x_t)) - \frac{1}{T} \sum\nolimits_{t=1}^{T} r_t(a_t) \tag{1}$$

The goal of the above Eq. (1) is to bring the C_r quickly to zero. Various contextual bandits learning algorithm, including ε-greedy, ε-first, ε-decreasing, contextual ε-greedy, bagging, upper confidence bound, lower confidence bound, Thompson sampling, and bandit forest are present in literature [7–9]. Among which, ε-greedy is the most fastest and simplest approach that can be adapted, which exploits the best strategy with probability of $(1 - \varepsilon)$ and uniformly exploits over all the other actions with probability of (ε). The regret computed with ε-greedy algorithm is shown in Eq. (2).

$$r_t = O\left(\left(\frac{K \ln |\Pi|}{T} \right)^{1/3} \right) \tag{2}$$

As the regret is to the power of 1/3, it may not be the most optimal bandits learning algorithm. However, it is computationally efficient, when working with a larger data set. Thus, as the next step ε-greedy algorithm will be applied as the contextual bandits learning algorithm in the proposed framework of CB-DSS to optimize the selection of the similarity retrieval model.

2.3   A Running Example

In this section, with an example, we will demonstrate on how contextual bandits learning can help in identifying the optimal similarity retrieval model in the CB-DSS for breast cancer management.

The main goal of contextual bandits learning algorithm is to maximize the total reward achieved by the learner, i.e. obtain the minimum contextual regret. As there exists, a policy π (decisional rule) that can give high rewards, the contextual bandits learning algorithm has to efficiently learn from all policies and choose the best policy. In our example let's assume 'n' number of policies are defined using decisional rules (IF-THEN statements) for different contexts. For example, "IF Surgeon THEN SR Model 1", "IF Radiologist THEN SR Model 2", where each SR model is assigned with a different weight matrix and similarity function model.

Now, during run time, the query case is sent to the SR model to retrieve similar cases from the patient case-base. Simultaneously, the context of the user, demographic information and patient data are sent to the contextual learning algorithm to enable it to select an optimal SR model. The selected SR model will now retrieve 10 similar cases to the query case. Based on the patient case selected by the user to make the clinical decision, the learner receives the corresponding reward. As shown in Table 1, we assign 1.0 for the most similar patient case and 0.1 to the 10th similar case.

Table 1. Reward value assigned for the learner's action

Context	Action (Retrieved Patient Cases)							
	1	2	3	4	…	8	9	10
GP, Outpatient, Asian	1.0	**0.9**	<u>0.8</u>	0.7	…	0.3	0.2	0.1
Radiologist, Radiology, European	1.0	0.9	0.8	0.7	…	**<u>0.3</u>**	0.2	0.1
Oncologist, Oncology, African	1.0	0.9	0.8	<u>0.7</u>	…	0.3	**0.2**	0.1
…	…							

The reward received by the learner is marked in **Bold.**
The reward received by the best policy is <u>Underlined</u>.

In the Bandits setting, as the learner could only observe the reward for the action taken, from Table 1, the learner's total reward can be computed as '0.9 + 0.3 + 0.2 + …'. Meanwhile, as the learner and the best policy have chosen the same arm for the second user, only the policy's reward of 0.3 is known. Here the best policy is determined to be the one, which would have chosen the same case with possibly a higher reward.

The contextual bandits learning algorithm is applied to exploit and explore, i.e. exploit the information available and explore from the action taken to learn and choose the best policy that gives the minimum regret and therefore the optimal result. In ε-greedy contextual bandits learning algorithm, it exploits the best strategy with probability of $(1 - \varepsilon)$ and uniformly exploits over all the other actions with probability of (ε), until optimal solution is achieved.

3 Conclusion

In this work, we have developed a CB-DSS for DESIREE project, aimed at providing web-based software for breast cancer diagnosis and management. The proposed CB-DSS provides a tool for querying former cases in order to retrieve similar patient cases from the case-base. As we note that the design of similarity retrieval model involves various factors from feature selection and weighing, similarity function, case representation and knowledge model, developing an optimal similarity retrieval model is challenging. To address such challenge, we presented a CB-DSS framework with multi-similarity retrieval models. We propose contextual bandits learning algorithm to dynamically choose between different similarity retrieval models by learning from the contextual information extracted from the user, patient and demographic data. The paper presents the overall framework of the proposed CB-DSS and systematically describes its workflow with a running example.

References

1. Parra-Calderón, C.L.: Patient similarity in prediction models based on health data: A scoping review, JMIR Med Inform. **5**(1) (2017)
2. Alexandrini, F., Krechel, D., Maximini, K., Wanggenheim, A.: Integrating CBR into the health caser organization. In: 16th IEEE Symposium on Computer-Based Medical Systems (2003)
3. Mary, J., Gaudel, R., Preux, P.: Bandits and recommender systems. In: Pardalos, P., Pavone, M., Farinella, G.M., Cutello, V. (eds.) MOD 2015. LNCS, vol. 9432, pp. 325–336. Springer, Cham (2015). https://doi.org/10.1007/978-3-319-27926-8_29
4. Séroussi, B, et al.: Reconciliation of multiple guidelines for decision support: a case study on the multidisciplinary management of breast cancer within the DESIREE project. In: 2017 Proceedings of the AMIA Annual Symposium, Washington DC, 4–8 November 2017 (2017)
5. Larburu, N., et al.: Augmenting guideline-based CDSS with experts' knowledge. In: 10th International Conference on Health Informatics, Porto, Portugal, 21–23 February 2017 (2017)
6. Bouneffouf, D., Feraud, R.: Multi-armed bandit problem with known trend. Neurocomputing **205**, 16–21 (2016)
7. Langford, J., Zhang, T.: The Epoch-Greedy algorithm for multi-armed bandits with side information. In: Advances in Neural Information Processing System, pp. 817–824 (2008)
8. Bouneffouf, D., Bouzeghoub, A., Gançarski, A.L.: A contextual-bandit algorithm for mobile context-aware recommender system. In: Huang, T., Zeng, Z., Li, C., Leung, C.S. (eds.) ICONIP 2012. LNCS, vol. 7665, pp. 324–331. Springer, Heidelberg (2012). https://doi.org/10.1007/978-3-642-34487-9_40
9. Agrawal, S., Goyal, N.: Thompson sampling for contextual bandits with linear payoffs. In: ICML (3), pp. 127–135 (2013)