



Ovarian cancer diagnosis using a hybrid intelligent system with simple yet convincing rules



Di Wang^a, Chai Quek^{a,*}, Geok See Ng^b

^a School of Computer Engineering, Nanyang Technological University, 639798 Singapore, Singapore

^b School of Information Systems Technology & Design, Singapore University of Technology & Design, 138682 Singapore, Singapore

ARTICLE INFO

Article history:

Received 1 April 2013

Received in revised form

15 December 2013

Accepted 16 December 2013

Available online 25 January 2014

Keywords:

Ovarian cancer diagnosis

Decision support system

Hybrid intelligent system

Interpretability

Feature selection

ABSTRACT

Ovarian cancer is the ninth most common cancer among women and ranks fifth in cancer deaths. Statistics show that the five-year survival rate is greater than 75% if diagnosis occurs before the cancer cells have spread to other organs (stage I), but it drops to 20% when the cancer cells have spread to upper abdomen (stage III). Therefore, it is crucial to detect ovarian cancer as early as possible and to correctly identify the stage of the cancer to prevent any further delay of appropriate treatments. In this paper, we propose a novel self-organizing neural fuzzy inference system that functions as a reliable decision support system for ovarian cancer diagnoses. The system only requires a limited number of control parameters and constraints to derive simple yet convincing inference rules without human intervention and expert guidance. Because feature selection and attribute reduction are performed during training, the inference rules possess a great level of interpretability. Experiments are conducted on both established medical data sets and real-world cases collected from hospital. The experimental results of our proposed model in ovarian cancer diagnoses are encouraging because it achieves the most number of correct diagnoses when benchmarked against other computational intelligence based models. More importantly, its automatically derived rules are consistent with expert knowledge.

© 2014 Elsevier B.V. All rights reserved.

1. Introduction

Ovarian cancer begins with the formation of a tumor in a woman's ovary. It is classified as the epithelial cancer caused by genetic alterations that disrupts the regulation of proliferation, apoptosis, senescence, and DNA repair [1]. Ovarian cancer ranks fifth in cancer deaths among women and accounts for more deaths than any other cancer of the female reproductive system [2]. It is estimated that in United States, 2012, about 22,280 women receive new diagnoses of ovarian cancer and about 15,500 patients die from it [2]. Early detection of ovarian cancer is crucial because statistics [3] show that the five-year survival rate of ovarian cancer is greater than 75% if diagnosis occurs before the cancer cells have spread to other organs (stage I), but it drops to 20% when the cancer cells have spread to upper abdomen (stage III). However, early detection of ovarian cancer is not easy because it is rather hard to tell by the symptoms when a tumor in the ovary turning from benignancy to malignancy. Even annual routine gynecologic and pelvic examinations have only detected 3% of the early stage ovarian cancer cases [4]. Furthermore, the accurate staging of ovarian cancer is also important because the treatment options depend on the type

and the stage that ovarian cancer has advanced to [5]. Therefore, early detection and accurate staging of ovarian cancer are of great interests and importance because they are the keys to improve the survival rate of a patient [6].

There are a number of research works in literature focused on cancer diagnoses using various approaches, such as statistical methods, clustering techniques, decision trees, neural networks, and neural fuzzy inference systems. Valerio et al. [7] used χ^2 -test to identify unique protein peaks in pancreatic cancer diagnoses. Lee [8] applied ANOVA to select gene markers in ovarian cancer diagnoses. Generally speaking, statistical approach identifies important features and usually requires other methods or doctor expertise to perform actual cancer diagnoses. Petricoin et al. [9] proposed a self-organizing clustering technique to separate patients with or without prostate cancer. Poon et al. [10] applied a two-way hierarchical clustering algorithm to differentiate hepatocellular carcinoma from chronic liver disease. Generally speaking, classifications based on the clustering results are conceptually equivalent to similarity-based diagnoses, which do not provide concrete decisive statements. Adam et al. [11] developed a classification decision tree to identify better biomarkers for early detection of prostate cancer. Qu et al. [12] further improved the results on the same prostate cancer data set using boosted decision trees. Decision trees are commonly known as comprehensive inference tools, which employ crisp decision rules and perform feature selections.

* Corresponding author.

E-mail address: ASHCQUEK@ntu.edu.sg (C. Quek).

However, decision trees are sensitive to noisy inputs and the boosted decision trees are difficult to interpret although they increase diagnostic accuracies. Ball et al. [13] proposed a three-layered neural network, which employs back propagation learning to analyze mass spectra for the prediction of astroglial tumor grades. Tsai et al. [14] used *t*-test to cull most of unconcerned genes and then applied the selected oncogenes to a three-layered neural network for ovarian cancer diagnoses. Although neural networks are accurate prediction tools, most of them function as black boxes [15] because users cannot make senses of the reasoning processes. Neural fuzzy inference system (or fuzzy neural network) combines the learning capabilities of neural networks and the transparent properties of fuzzy systems together by performing respective fuzzy or non-fuzzy operations in each layer of the network. Tan et al. [16] proposed neural fuzzy inference systems, which are inspired by the hippocampus and neocortex memory structures to diagnose ovarian cancer. When benchmarked against other models, their experimental results illustrated clear advantages of the proposed neural fuzzy inference systems that achieved higher accuracies and employed smaller number of interpretable fuzzy rules. Tung et al. [17] first used the Monte Carlo evaluative selection method [18] to select important features and subsequently trained the proposed neural fuzzy inference system for ovarian cancer diagnoses. However, their experimental results did not show superior performances even with feature selection applied. In recent years, there were studies focusing on obtaining highly compact fuzzy inference rules [19,20]. The models were able to derive extraordinarily small numbers of rules through the systematic pruning process from both established medical data sets and real-world disease diagnoses. Furthermore, the hybridization of multiple techniques combines the advantages of each individual and alleviates certain limitations. Therefore, hybrid intelligent systems receive increasing attentions in the past few years. For example, evolutionary algorithms are often combined with other techniques to solve optimization problems [21,22] and to find optimal parameters [23]. A recent comparative study is reported in [24]. In this paper, we propose a novel neural fuzzy inference system, which self-organizes its network structure and performs feature selection and attribute reduction during training. Therefore, simple yet convincing fuzzy inference rules are systematically derived to diagnose both established medical data sets and real-world ovarian cancer cases collected from hospital for stage identifications.

The rest of this paper is organized as follows: Sections 2 and 3 describe two techniques, respectively, to introduce the foundations of our proposed model. Section 4 provides the details of our proposed clustering method. Section 5 defines the system architecture of our proposed model, which employs the fuzzy rules automatically derived by the proposed clustering method. Section 6 benchmarks the performances of our proposed model on two established medical data sets. Section 7 presents the experimental results of applying our proposed model to diagnose ovarian cancer cases collected from hospital and analytically benchmarks the results against other computational intelligence based models. Section 8 concludes this paper.

2. Rough set theory for knowledge reduction

Rough sets [25] are always compared to fuzzy sets [26]. To describe belongingness, fuzzy sets use straightforwardly defined membership functions, while rough sets use relative relations named the lower and upper approximations. Both theories aim to achieve the same type of goal [27]. However, it is always better to have them both in one system to take advantages of their complements [28].

Table 1
A simple example to illustrate the concept of a decision table.

U	A		
	C Height	Weight	D Body size
1	Tall	Heavy	Big
2	Short	Light	Small

2.1. Knowledge representation system and decision table

Rough set theory is initiated to model relations in a given data set or knowledge base [25]. It is a formally defined methodology, which can be applied to reduce the dimensionality of a given data set [29], proceeding to train an inference system.

To express mathematically how rough set theory is applied in knowledge reduction, decision logic language is used to model the Knowledge Representation System (KRS). Such a system is represented in a pair $S=(U, A)$, where U is a nonempty and finite set named the universe of discourse and A is a nonempty and finite set of primitive attributes.

Decision tables can be defined in terms of KRS. If we have $S=(U, A)$ and $C, D \subset A$, which denote the condition and decision attributes, respectively, then S with distinguished (C, D) pairs is essentially a decision table, i.e. $T=\{U, C, D\}$ or T is a CD-based decision table. Moreover, each element in U is a CD-based decision rule and it represents a cluster of data. Subsequently, U is the union of all clusters found in the data set. A simple example of decision table is given in Table 1 to illustrate the concept.

2.2. Indiscernible relation and approximations

Indiscernible relation over knowledge K , named $IND(K)$, is defined in Eq. (1). The family of all equivalence classes of the equivalence relation $IND(K)$ is named $U/IND(K)$.

$$IND(K) = \{(x, y) \in U^2 \mid \forall r \in K, r(x) = r(y)\}. \quad (1)$$

Rough set theory approximates knowledge using a pair of relational approximations. The lower and upper approximations of a set, when given an equivalence relation $IND(K)$, are defined in Eq. (2). By presenting knowledge K , lower approximation $\underline{K}Y$ is the set of elements that can be certainly classified by K and Y and upper approximation $\overline{K}Y$ is the set of elements that can be possibly classified by K and Y .

$$\underline{K}Y = \bigcup \{X : X \in U/IND(K), X \subseteq Y, Y \subseteq U\}, \quad (2a)$$

$$\overline{K}Y = \bigcup \{X : X \in U/IND(K), X \cap Y \neq \emptyset, X \subseteq Y, Y \subseteq U\}. \quad (2b)$$

2.3. Attribute reduction and feature selection

Rough set theory performs knowledge reduction with two fundamental concepts named reduct and core. Intuitively, a reduct of knowledge is an essential subset of knowledge suffices to define all basic relations, whereas a core is the most fundamental subset of knowledge consists of the common attributes of all reducts.

Given a decision table $T=\{U, C, D\}$, an attribute a is dispensable if and only if $IND(C)=IND(C-\{a\})$. Otherwise, a is indispensable. The family C is independent if $\forall a \in C$ is indispensable in C . Attribute reduction is performed during the process of finding independent C with minimum cardinality. If for all elements in U , an attribute b is dispensable, then b can be removed from the training data set. This process of eliminating attributes that do not contribute any information for inferences is considered as feature selection, which is extremely useful to handle problems with high dimensionality.

متن کامل مقاله

دریافت فوری ←

ISIArticles

مرجع مقالات تخصصی ایران

- ✓ امکان دانلود نسخه تمام متن مقالات انگلیسی
- ✓ امکان دانلود نسخه ترجمه شده مقالات
- ✓ پذیرش سفارش ترجمه تخصصی
- ✓ امکان جستجو در آرشیو جامعی از صدها موضوع و هزاران مقاله
- ✓ امکان دانلود رایگان ۲ صفحه اول هر مقاله
- ✓ امکان پرداخت اینترنتی با کلیه کارت های عضو شتاب
- ✓ دانلود فوری مقاله پس از پرداخت آنلاین
- ✓ پشتیبانی کامل خرید با بهره مندی از سیستم هوشمند رهگیری سفارشات