Artificial Intelligence in Medicine 63 (2015) 119–133

Contents lists available at ScienceDirect

Artificial Intelligence in Medicine

journal homepage: www.elsevier.com/locate/aiim

Application of a two-stage fuzzy neural network to a prostate cancer prognosis system

Ren-Jieh Kuo[a,*], Man-Hsin Huang[b], Wei-Che Cheng[a], Chih-Chieh Lin[c,d], Yung-Hung Wu[e]

[a] Department of Industrial Management, National Taiwan University of Science and Technology, No. 43, Section 4, Kee-Lung Road, Taipei 106, Taiwan
[b] Hon Hai Precision Industry Co., Ltd., No. 2, Zihyou St., Tucheng Dist., New Taipei City 236, Taiwan
[c] Department of Urology, Taipei Veterans General Hospital, Institute of Clinical Medicine, 201, Section 2, Shih-Pai Road, Taipei 112, Taiwan
[d] Department of Urology, School of Medicine, National Yang-Ming University, Taipei, Taiwan
[e] Superintendent Office, Taipei Veterans General Hospital, 201, Section 2, Shih-Pai Road, Taipei 112, Taiwan

ARTICLE INFO

Article history:
Received 2 February 2014
Received in revised form
13 December 2014
Accepted 20 December 2014

Keywords:
Optimization version of an artificial immune network
Particle swarm optimization algorithm
Fuzzy neural network
Prostate cancer
Prognosis

ABSTRACT

Objective: This study intends to develop a two-stage fuzzy neural network (FNN) for prognoses of prostate cancer.
Methods: Due to the difficulty of making prognoses of prostate cancer, this study proposes a two-stage FNN for prediction. The initial membership function parameters of FNN are determined by cluster analysis. Then, an integration of the optimization version of an artificial immune network (Opt-aiNET) and a particle swarm optimization (PSO) algorithm is developed to investigate the relationship between the inputs and outputs.
Results: The evaluation results for three benchmark functions show that the proposed two-stage FNN has better performance than the other algorithms. In addition, model evaluation results indicate that the proposed algorithm really can predict prognoses of prostate cancer more accurately.
Conclusions: The proposed two-stage FNN is able to learn the relationship between the clinical features and the prognosis of prostate cancer. Once the clinical data are known, the prognosis of prostate cancer patient can be predicted. Furthermore, unlike artificial neural networks, it is much easier to interpret the training results of the proposed network since they are in the form of fuzzy IF–THEN rules. These rules are very important for medical doctors. This can dramatically assist medical doctors to make decisions.

1. Introduction

Prostate cancer is one of the most common causes of death for men in most industrialized countries [1]. Many studies have considered the introduction of diagnoses for prostate cancer, but few studies have aimed at prognoses for the prostate. Other studies have evaluated the use of artificial neural networks (ANNs) to increase prostate cancer detection rates and reduce unnecessary biopsies [2] and used a neuro-fuzzy system based on both serum data (total prostate-specific antigen; tPSA, percent free PSA; %fPSA), and clinical data (age) to enhance the performance of tPSA in discriminating prostate cancer [3]. However, none of these studies have discussed prognoses of prostate cancer. In Taiwan, the rates of morbidity and mortality of prostate cancer have been rising in recent years. Patients often wonder how much time they have left and how serious their disease is. Doctors can only point to a five-year survival rate ratio depending on patients' clinical data. Therefore, this study intends to build a prognosis system for prostate cancer to estimate how many years patients have remaining, not just classify patients into two categories (those who can survive more than five years and those who cannot).

Nowadays, many business decisions are made based on the results of data analysis. Another important application area of this type of analysis is in medical diagnoses. Models obtained from data analysis that are applied in practice require transparency and interpretability in terms of the attributes they process. Only then can the user can understand what the results mean easily. Thus, employing appropriate algorithms to build a prognosis system is a very important issue. This kind of algorithm can enable more accurate prediction of prostate cancer prognoses and translate the results of the proposed scheme into comprehensible information for users, or doctors.

In view of the excellent performance of artificial immune systems (AIS) in many fields [4], this study proposes a novel two-stage FNN and applies it to prostate cancer prognoses. The proposed approach includes the following steps: (1) cluster the features and

* Corresponding author. Tel.: +886 2 27376328; fax: +886 2 27376324.
E-mail address: rjkuo@mail.ntust.edu.tw (R.-J. Kuo).

http://dx.doi.org/10.1016/j.artmed.2014.12.008
0933-3657/

get the initial means and variances of FNN and (2) integrate Opt-aiNET and PSO-based FNN (IOAP-FNN) to forecast the prognoses of prostate cancer patients. The solution is first found through Opt-aiNET due to its good searching capability and then fine-tuned by using PSO. Since the proposed algorithm combines both the advantages of Opt-aiNET and PSO, it has good searching capability, avoids getting stuck to the local optimal solutions and is able to converge rapidly. Unlike an artificial neural network, it is much easier to interpret the training results when using the fuzzy neural network since they are in the form of fuzzy IF–THEN rules.

In order to verify the proposed FNN, three benchmark functions are first employed. Then, the proposed network is further applied to prostate cancer data in order to build a prognosis system for medical doctors who can use these results to improve the diagnoses of prostate cancer patients.

The remainder of this study is organized as follows. Section 2 provides a review of the related literatures. Section 3 presents the proposed prognosis system which employs two-stage FNN, while the simulation results are presented in Section 4. Section 5 presents the model evaluation results for the prostate cancer prognosis system. Discussions and concluding remarks are finally made in Sections 6 and 7.

2. Background

This section presents some background related to the current study, including discussions of prostate cancer, fuzzy theory, FNNs, artificial immune networks, PSO and applications of soft computing techniques to FNN.

2.1. Prostate cancer

The prostate gland is one organ of male reproduction. It can be found below the bladder and in front of the rectum. The definition of prostate cancer is a disease in which cancer develops in the prostate gland. Cancer occurs when the cells of the prostate mutate and begin to multiply out of control. These cells may metastasize from the prostate to other organs, especially the bones and lymph nodes. Prostate cancer is a significant cause of morbidity and mortality in Western countries [5]. In the United States, prostate cancer is the most common cancer and the second most common cause of cancer related death [6]. In 2011, an estimated 240,890 new cases and 33,720 deaths from prostate cancer were recorded. Similar data have been reported in Europe and Canada. In Taiwan, the morbidity of prostate cancer is the fifth for men with cancer, and the mortality ranks seventh. According to the statistical analysis published by the Department of Health, Executive Yuan, there were 3603 cases and 1052 deaths from prostate cancer recorded in 2008. The mean age of prostate cancer patients is 75 years old. This indicates that older men have a higher possibility of getting prostate cancer.

2.1.1. Process of diagnosis

Prostate cancer is one of the most common types of cancer found in men. Risk factors for prostate cancer include age, the family's cancer history and ethnic background [7]. The increasing rate of prostate cancer for men in Taiwan has led doctors and patients seek more information regarding the probability of prostate cancer diagnosis, and the prognosis of prostate cancer is important to patients as well. Definitive diagnosis of prostate cancer can be made through a biopsy. An initial diagnosis is made after patients' transrectal ultrasonography, rectal examination results and the amount of prostate-specific antigen (PSA) are assessed by a specialist doctor. The PSA level in the blood has become one of the most common methods as a result of studies conducted in recent years for early diagnosis of prostate cancer [7].

PSA levels which are below 4 ng/ml in the blood are considered normal, while levels between 4 and 10 ng/ml are considered limit values and levels above 10 ng/ml are high. It has been stated that the higher the PSA level is, the higher the prostate cancer risk is [8]. However, PSA values may not yield conclusive results about existence of prostate cancer because PSA levels can be increased by inflammation of the prostate and benign prostate hyperplasia (BPH). Therefore, patients are also given rectal examinations. If anomalies are observed at the end of a rectal examination, even if PSA results may seem normal, it is recommended that a prostate biopsy be performed and definitive diagnosis be made [7].

The probability of prostate cancer can be estimated by logistic regression analysis and ANNs, which can be trained to predict diagnostic outcomes. However, none of these tools have resolved the problem of low specificity for prostate cancer diagnosis [3].

2.1.2. Classification of prostate cancer

Since prostate cancer is one of the most common causes of cancer death among men in most industrialized countries, people are becoming more concerned about this disease. There already have been some studies regarding the classification of prostate cancer by using different methods and features. ANNs which have been applied in many different areas due to their promising results have also been employed to solve medical diagnoses problems recently. In 2007, Keles et al. [1] applied the neuro-fuzzy classification (NEFCLASS) tool to classify whether the patients have prostate cancer or benign prostatic hyperplasia (BPH) since these two illnesses have similar symptoms. Çınar et al. [9] presented a classifier-based expert system for early diagnosis of the prostate in the constraint phase to facilitate informed decision making without biopsy by using some selected features without biopsy examination. These features include weight, height, body mass index (BMI), PSA, free PSA, age, prostate volume, density, smoking, systolic, diastolic, pulse, and Gleason score. Saritas et al. [7] applied an artificial neural network to predict whether patients have cancer or not. The features used are free PSA, total PSA and age data. Though the system does not diagnose cancer conclusively, yet it helps doctors decide whether a biopsy is necessary or a waiting policy should just be carried to see whether the patient has prostate cancer or according to the information provided [7].

2.1.3. Predictive application for prostate cancer

Benecchi [3] developed a neuro-fuzzy system to predict the presence of prostate cancer. The proposed neuro-fuzzy system uses tPSA, %fPSA and clinical data (age) as the input features. This can enhance the performance of tPSA in discriminating prostate cancer. The predictive accuracy of the neuro-fuzzy system is superior to that of tPSA and %fPSA. Ecke et al. [2] evaluated the use of the ANN program "Prostate Class" in a daily routine to increase the prostate cancer detection rate and reduce unnecessary biopsies. A total of 204 patients were included in the study. Individual ANN predictions were generated with the use of the ANN application for the Beckman Access PSA and free PSA assays, relying on the features of age, PSA, %fPSA, prostate volume, and DRE. The results show that ANN is a very helpful parameter in the daily routine to increase the prostate cancer detection rate and reduce unnecessary biopsies.

2.2. Fuzzy neural networks

An ANN is a system derived from neurophysiology models. In general, an ANN consists of a collection of simple, nonlinear computing elements, whose inputs and outputs are tied together, to form a network [10]. However, a disadvantage of ANNs, which is an impediment to their more widespread acceptance, is the absence of capability to explain to the user, in a form comprehensible to humans, how the network arrives at a particular decision. Neither