Chapter 10
Diagnosis of Breast Cancer Using Intelligent Information Systems Techniques

Ahmad Al-Khasawneh
Hashemite University, Jordan

ABSTRACT

Breast cancer is the second leading cause of cancer deaths in women worldwide. Early diagnosis of this illness can increase the chances of long-term survival of cancerous patients. To help in this aid, computerized breast cancer diagnosis systems are being developed. Machine learning algorithms and data mining techniques play a central role in the diagnosis. This paper describes neural network based approaches to breast cancer diagnosis. The aim of this research is to investigate and compare the performance of supervised and unsupervised neural networks in diagnosing breast cancer. A multilayer perceptron has been implemented as a supervised neural network and a self-organizing map as an unsupervised one. Both models were simulated using a variety of parameters and tested using several combinations of those parameters in independent experiments. It was concluded that the multilayer perceptron neural network outperforms Kohonen's self-organizing maps in diagnosing breast cancer even with small data sets.

1. INTRODUCTION

Excluding non-melanoma skin cancers, breast cancer is the most widely spread cancer in women in both developing and developed countries. In developed countries, breast cancer has become a major cause of death among women (American Cancer Society, 2013). It also affects men but with a low percentage (1%) due to the fact that men have lower levels of female hormones which affect the growth of breast cells besides that their breast duct cells are less developed than in women (American Cancer Society, 2013).

On the other hand, breast cancer is the most treatable form of cancer once it is detected early. Thus, breast cancer successful treatment depends largely on the rapid diagnosis process. Breast cancer is a malignant tumor that is developed from cells of the breast. A malignant tumor is a group of cancer cells that may grow into (invade) surrounding tissues or spread (metastasize) to distant areas of the body

[2]. Benign tumors are less dangerous. Because they do not spread, benign tumors do not intimidate life. Indeed, no specific cause of the diseased is known yet. Nevertheless, several risk factors may help in developing the disease in women. These factors include ageing, genetic risk factors, family history, menstrual periods, not having children, and obesity (American Cancer Society, 2013).

Most breast cancers are detected by the patient as a lump in the breast. The role of the physician is to diagnose breast cancer by distinguishing benign lumps from malignant ones. Clinically, the diagnosis of breast cancer is done based on a number of tests done for the patient. Several diagnostic techniques exist; they are based on either Mammography, Fine needle aspiration cytology (FNAC), or surgical biopsy (American Cancer Society, 2013). The presence/absence of malignant tumors is determined by medical practitioners based on the results of these tests. Clinicians recommend doing tumor evaluation regularly to detect any formation early.

Correct diagnosis of the disease implies classifying tumors into malignant and benign (Gayathri, Sumathi, & Santhanam, 2013). Indeed, this is not a trivial task even for expert physicians. Hence, computerized decision support systems are required to give a hand in the diagnosis process and to reduce the time required to diagnose the breast cancer and thus treat it early (Al-Khasawneh & Hijazi, 2014).

Diagnosing diseases correctly is a major problem in medical sciences and bioinformatics analysis. Diagnosing a disease is a classification problem. Classification problems could be solved either by artificial intelligence techniques or statistical techniques. Overall, artificial intelligence ones are more recommended due to their efficiency and lower complexity. Recently, several machine learning algorithms have been widely used in prediction, especially in medical diagnosis.

Machine learning is a branch of artificial intelligence that is concerned with the design and development of algorithms that can learn from data allowing computers to evolve behaviors based on empirical data (Gayathri, Sumathi, & Santhanam, 2013). The result of such algorithms is usually a model that is used to make decisions.

Machine learning algorithms can be supervised, unsupervised, semi-supervised, reinforcement, transduction, and learning to learn. In supervised learning, both inputs and their desired outputs are presented to the system aiming to learn how to map inputs to outputs (Singh, Saini, & Singh, 2012). In unsupervised learning, no output is provided; it is the responsibility of the algorithm to group similar inputs into different classes. This gathering is usually called clustering (Singh, Saini, & Singh, 2012).

Artificial neural network (ANN) is a branch of computational intelligence that employs a variety of optimization tools to "learn" from past experiences and use that prior training to classify new data, identify new patterns or predict an outcome. ANNs are being used efficiently in medical problems and specifically in the diagnosis ones (Amato, López, Peña-Méndez, Vaňhara, & Hampl, 2013). This could be ascribed to their abilities to handle classification problems, represent complex relationships in large datasets, and to approximate nonlinear functions. Imperatively, ANNs aim to support clinicians but not to replace them.

Artificial Neural networks may utilize supervised learning or unsupervised learning. Supervised ANNs usually outperform the unsupervised network. Hence, the majority of neural networks based diagnosis systems use supervised learning. Unsupervised learning is the only solution when there are no previous medical records for old patients of the disease.

In medical research, multilayer perceptron's are the most commonly used artificial neural networks (Burke, et al., 1997). Multilayer perceptron's are supervised neural networks that utilize back propagation training algorithm. Multilayer perceptron's are predictive modeling technique. Self-organizing Maps (SOMs) are unsupervised neural networks that perform clustering. Clustering is a descriptive modeling

Related Content

Intuitionistic Fuzzy 2-Metric Space and Some Topological Properties

Q.M. Danish Lohani (2011). *International Journal of Artificial Life Research (pp. 59-73).*

www.irma-international.org/article/intuitionistic-fuzzy-metric-space-some/56323

A Machine Learning Approach to Tracking and Characterizing Planar or Near Planar Fluid Flow

Mahendra Gooroochurn, David Kerrand Kaddour Bouazza-Marouf (2020). *International Journal of Natural Computing Research (pp. 76-87).*

www.irma-international.org/article/a-machine-learning-approach-to-tracking-and-characterizing-planar-or-near-planar-fluid-flow/258961

Population Symbiotic Evolution in a Model of Industrial Districts

U. Merlone (2007). *Handbook of Research on Nature-Inspired Computing for Economics and Management (pp. 301-316).*

www.irma-international.org/chapter/population-symbiotic-evolution-model-industrial/21136

Solving Facility Location Problems with a Tol for Rapid Development of Multi-Objective Evolutionary Algorithms (MOEAs)

A. L. Medaglia (2007). *Handbook of Research on Nature-Inspired Computing for Economics and Management (pp. 642-660).*

www.irma-international.org/chapter/solving-facility-location-problems-tol/21157

Forward and Backward Chaining with P Systems

Sergiu Ivanov, Artiom Alhazov, Vladimir Rogojinand Miguel A. Gutiérrez-Naranjo (2011). *International Journal of Natural Computing Research (pp. 56-66).*

www.irma-international.org/article/forward-backward-chaining-systems/57971