

Chapter 11

Data Mining and Machine Learning Approaches in Breast Cancer Biomedical Research

Gunavathi Chellamuthu
VIT University, India

Kannimuthu S.
Karpagam College of Engineering, India

Premalatha K.
Bannari Amman Institute of Technology, India

ABSTRACT

Breast cancer is the most common invasive cancer in females worldwide. Breast cancer diagnosis and breast cancer prognosis are the two important challenges for the researchers in the medical field and also for the practitioners. If the cells in the breast start to grow without any control, it leads to cancer. Normally, the growth of the lump can be seen using x-ray. The benign and malignant breast lumps are distinguished during breast cancer diagnosis. The prognosis process predicts the period at which the breast cancer is likely to reappear in patients who have had their cancers removed. Data mining techniques and machine learning algorithms are mostly used in the whole process of breast cancer diagnosis and treatment. They utilize the large volume of breast cancer data for extracting knowledge. The application of data mining and machine learning methods in biomedical research is presently vital and crucial in efforts to transform intelligently all available data into valuable knowledge.

INTRODUCTION

Cancer is featured by an irregular, unmanageable growth that may demolish and attack the neighbouring healthy body tissues. Cancer categorization by medical practitioners and radiologists was depend on morphological and clinical characteristics and had restricted diagnostic ability in the earlier period. The most prevalent and dangerous disease among the females in worldwide is breast cancer. It was

DOI: 10.4018/978-1-5225-4999-4.ch011

found that 22.9% of evident cancers in women and 16% of all female cancers are related to breast cancer disease. 18.2% of men and women are passed away worldwide because of breast cancer (Mandelblatt et al. 2004). Researcher faces two major challenges such as Breast Cancer identification and prognosis in medical applications. Data mining techniques and machine learning algorithms are mostly used in the whole process of breast cancer diagnosis and treatment and they changed the entire process simple in nature. The benign and malignant breast lumps are distinguished during Breast Cancer diagnosis. The prognosis process predicts the period at which the breast cancer is likely to reappear in patients who have had their cancers removed. Applying data analytics and machine learning methodologies in cancer research is one of the key approaches to utilize large volume of Breast cancer data for extracting knowledge. The noteworthy methods in biotechnology and health care have shown to an outstanding production of data, such as high throughput genetic and clinical data, generated from large volume of E-health records. The application of data mining and machine learning methods in biomedical research is recently essential and crucial in efforts to convert intelligently all available data into valuable facts (Diz et al., 2015). The primary intension of the contemporary study is to take a systematic investigation of the applications of machine learning, data mining algorithms and tools in the field of Breast cancer research with reference to i) Breast cancers complications, ii) Prediction, Diagnosis and Prognosis and iii) Identification of biomarker genes.

BACKGROUND

In human body, if the cells in the breast are started to grow without any control, it leads to cancer. Normally, the growth of the lump can be seen using x-ray. The affected cells can propagate into nearby tissues or blowout all over the body. More than 100 types of cancer are identified in the medical field and it became one of the major roots of death in the world. There are many factors that influence the formation and spreading of cancers which are listed below:

1. Genetics
2. Gender
3. Age
4. Life style
5. Environment
6. Marital status

Breast cancer is the deadliest disease which is a malevolent cell development in the breast. The cancer cells extend to other parts of the body, if the patient left untreated. The occurrence of breast cancer upsurges after 40 years of age. The incidence of this disease is even high with women over age 50. Breast cancer causes deaths in women which has the second place among other diseases. According to the survey, mortality rates dropped significantly during 1992-1998 with the major decreases in younger women. Breast cancer is most prevalent in women which affect over 13% of all women in the world. There is more number of younger women vulnerable to breast cancer in most of the nations (Richie & Swanson, 2003).

Breast cancer is of two types. They are:

28 more pages are available in the full version of this document, which may be purchased using the "Add to Cart" button on the publisher's webpage:
www.igi-global.com/chapter/data-mining-and-machine-learning-approaches-in-breast-cancer-biomedical-research/210970

Related Content

A Multi-Methodological Approach to Rare Association Rule Mining

Yun Sing Koh and Russel Pears (2010). *Rare Association Rule Mining and Knowledge Discovery: Technologies for Infrequent and Critical Event Detection* (pp. 76-97).

www.irma-international.org/chapter/multi-methodological-approach-rare-association/36901

Creating Risk-Scores in Very Imbalanced Datasets: Predicting Extremely Violent Crime among Criminal Offenders Following Release from Prison

Markus Breitenbach, William Dieterich, Tim Brennan and Adrian Fan (2010). *Rare Association Rule Mining and Knowledge Discovery: Technologies for Infrequent and Critical Event Detection* (pp. 231-254).

www.irma-international.org/chapter/creating-risk-scores-very-imbalanced/36910

Incorporating Correlations among Gene Ontology Terms into Predicting Protein Functions

Pingzhao Hu, Hui Jiang and Andrew Emili (2011). *Ontology Learning and Knowledge Discovery Using the Web: Challenges and Recent Advances* (pp. 154-173).

www.irma-international.org/chapter/incorporating-correlations-among-gene-ontology/53885

Hierarchical Clustering Using Evolutionary Algorithms

Monica Chis (2008). *Mathematical Methods for Knowledge Discovery and Data Mining* (pp. 146-156).

www.irma-international.org/chapter/hierarchical-clustering-using-evolutionary-algorithms/26138

Grief and the Psychological Transition Process Among Information Technology Team Members

Letitia Larry (2020). *Optimizing Data and New Methods for Efficient Knowledge Discovery and Information Resources Management: Emerging Research and Opportunities* (pp. 1-31).

www.irma-international.org/chapter/grief-and-the-psychological-transition-process-among-information-technology-team-members/255749