

IMPLEMENTING MACHINE LEARNING TECHNIQUES FOR CANCER PROGNOSIS

Name: Gauresh Chavan, Kaushal Chaudhary, Mohit Ruke

Date: 02/09/2018

INTRODUCTION

Breast Cancer is the most common form of cancer in women, affecting almost 12% of all women across the world. In recent years, the incidence rate keeps increasing and data show that the survival rate is 88% after first five years from diagnosis and 80% after 10 years from diagnosis. Early prediction of cancer is one of the most crucial steps in the follow-up process [3].

Over past few decades, scientists applied different methods, such as screening in early stage, so that types of cancer can be identified before their symptoms. With advent of new technologies, a large amount of cancer data is available to the research community. However, an accurate prediction of a disease outcome is still one of the challenging tasks for the researchers and physicians around the world [2]. In this paper, we will develop, using machine learning techniques, methods that will allow accurate prognosis of cancer.

BACKGROUND

Today, despite the many advances in early detection of diseases, cancer patients have a poor prognosis and the survival rates in them are low [1]. In cancer prediction/prognosis one is concerned with the following [4]:

- 1) the prediction of cancer susceptibility (i.e. risk assessment)**
- 2) the prediction of cancer recurrence and**
- 3) the prediction of cancer survivability.**

In the first case, one is trying to predict the likelihood of developing a type of cancer prior to the occurrence of the disease. In the second case one is trying to predict the likelihood of redeveloping cancer. In the third case one is trying to predict an outcome (life expectancy, survivability, progression, tumor-drug sensitivity) after the diagnosis of the disease. In the latter two situations the success of the prognostic prediction is obviously dependent on the success or quality of the diagnosis [4].

DATA SOURCES

UCI Machine Learning Repository : Wisconsin Diagnostic Dataset for Breast Cancer (WDBC)
[https://archive.ics.uci.edu/ml/datasets/Breast+Cancer+Wisconsin+\(Diagnostic\)](https://archive.ics.uci.edu/ml/datasets/Breast+Cancer+Wisconsin+(Diagnostic))

ALGORITHMS

We will be employing different algorithms to perform prediction at different stages of the prognosis.

Neural Network:

<https://www.kdnuggets.com/2016/10/artificial-intelligence-deep-learning-neural-networks->

[explained.html](#)

<http://www.wildml.com/2015/09/implementing-a-neural-network-from-scratch/>

Decision Trees:

<https://towardsdatascience.com/decision-trees-in-machine-learning-641b9c4e8052>

SVM:

<http://scikit-learn.org/stable/modules/svm.html>

Automatic Feature Selection – CFS & Wrapper Algorithm:

<https://machinelearningmastery.com/feature-selection-in-python-with-scikit-learn/>

REFERENCES:

[1] Azadeh Bashiri, Marjan Ghazisaeedi,* Reza Safdari, Leila Shahmoradi, and Hamide Ehtesham .Improving the Prediction of Survival in Cancer Patients by Using Machine Learning Techniques: Experience of Gene Expression Data: A Narrative Review

[2] Konstantina Kourou, Themis P. Exarchos, Konstantinos P. Exarchos, Michalis V. Karamouzis, Dimitros I. Fotiadis. Machine learning applications in cancer prognosis and prediction.

[3] Using Three Machine Learning Techniques for Predicting Breast Cancer Recurrence Ahmad LG*, Eshlaghy AT, Poorebrahimi A, Ebrahimi M and Razavi AR

[4] Applications of Machine Learning in Cancer Prediction and Prognosis Joseph A. Cruz and David S. Wishart

[5] On Breast Cancer Detection: An Application of Machine Learning Algorithms on the Wisconsin Diagnostic Dataset Abien Fred M. Agarap