



Data Mining Techniques: To Predict and Resolve Breast Cancer Survivability

Vikas Chaurasia¹, Saurabh Pal²

¹Research Scholar, Sai Nath University, Ranchi, Jharkhand, India

²Head Dept. of MCA, VBS Purvanchal University, Jaunpur (U.P.), India

Abstract— *Breast cancer is one of the deadliest disease, is the most common of all cancers and is the leading cause of cancer deaths in women worldwide, accounting for >1.6% of deaths and case fatality rates are highest in low-resource countries. The breast cancer risks are broadly classified into modifiable and non-modifiable factors. The non modifiable risk factors are age, gender, number of first degree relatives suffering from breast cancer, menstrual history, age at menarche and age at menopause. While the modifiable risk factors are BMI, age at first child birth, number of children, duration of breast feeding, alcohol, diet and number of abortions.*

This paper presents a diagnosis system for detecting breast cancer based on RepTree, RBF Network and Simple Logistic. In test stage, 10-fold cross validation method was applied to the University Medical Centre, Institute of Oncology, Ljubljana, Yugoslavia database to evaluate the proposed system performances. The correct classification rate of proposed system is 74.5%. This research demonstrated that the Simple Logistic can be used for reducing the dimension of feature space and proposed Rep Tree and RBF Network model can be used to obtain fast automatic diagnostic systems for other diseases.

Keywords— *Breast cancer survivability; Rep Tree; RBF Network; Simple Logistic; Diagnostic systems*

Full Text: <http://www.ijcsmc.com/docs/papers/January2014/V3I1201402.pdf>