

Determining the stage of breast cancer by data mining algorithms

Maryam Sadat Mahmoodi: Department of Computer, Payame Noor University, I.R of Iran

Seyed Abbas Mahmoodi: Department of Computer Engineering, College of Engineering and Technology, Yazd Science and Research Branch, Islamic Azad University, Yazd, Iran

Fatemeh Haghighi: Associate Professor, Department of Oral Pathology, Faculty of Medicine, Birjand University of Medical Sciences, Birjand, Iran

Seyed Mostafa Mahmoodi: Assistant Professor, Department of Oral Pathology, Faculty of Dentistry, Birjand University of Medical Sciences, Birjand, Iran

Corresponding Author: Seyed Abbas Mahmoodi, Sa_mahmoodi_85@yahoo.com

Abstract

Background: Breast cancer is the most common female cancer in Iran. Type of treatment, prognosis and survival rate are depends on the clinical stage.

The main objective of this paper is to select an appropriate algorithm that can determine the stage of disease in breast cancer patients.

Methods: In this descriptive - analytical study, data of tumor size, lymph node involvement, and metastasis of breast cancer in 732 patients in Vali Asr Hospital, Birjand was used. These data were used to select the appropriate algorithm of the decision tree algorithms, artificial neural networks, Bayesian networks, and K-nearest neighbor, to determine the stage of the disease a new person.

Results: Studies show that the K-nearest neighbor algorithm can detect accurately stage disease patients, with 96% probability of measuring tumor size, lymph node involvement and the extent of the amount metastasis

Conclusion: K-nearest neighbor algorithm has the most accurate for the detection of breast cancer.

Keywords: Breast Cancer, Data Mining Algorithms, Decision Tree, Artificial Neural Network.