

## عنوان مقاله:

Using machine learning techniques for predicting breast cancer

## محل انتشار:

چهاردهمین کنگره بین المللی سرطان پستان (سال: 1397)

تعداد صفحات اصل مقاله: 1

## نویسندگان:

Mohammad Noorchenarboo - *Master student of Department of Epidemiology and Biostatistics, School of Public Health, Tehran University of Medical Sciences, Tehran, Iran*

Mahboubeh Parsaeian - *Assistant Professor of Department of Epidemiology and Biostatistics, School of Public Health, Tehran University of Medical Sciences, Tehran, Iran*

## خلاصه مقاله:

Introduction: Breast cancer is the most common cancer among women. The prediction of breast cancer type has been a challenging research problem. The main objective of this article is to develop accurate prediction models for diagnosing breast cancer using recent machine learning techniques. Methods: Data included 569 tumor samples obtained from the Breast Cancer Wisconsin (Diagnostic) Database. Tumor features were extracted from a digitized image of a fine needle aspirate. Based on these features, we developed machine learning classification algorithms to distinguish between the type of cancer; benign or malignant. We used cross-validation techniques to assess the predictive ability of classification algorithms. Friedman test was used to compare these algorithms (accuracy, sensitivity, and specificity). Results: The mean accuracy of Support Vector Machine was 0.9543 (sensitivity 0.9594 and specificity 0.9537), the decision tree model was 0.9279 (sensitivity 0.8947, specificity 0.949), the neural network model was 0.9613 (sensitivity 0.9463, specificity 0.972), and the Random Forest model was 0.9631 (sensitivity 0.9415, specificity 0.9768). Friedman test showed significant differences between methods for accuracy ( $p$ -value=0.018) and sensitivity ( $p$ -value=0.001), but it wasn't a significant difference between methods on specificity ( $p$ -value=0.186). Conclusion: This study provided insight into the predictive ability of different machine learning methods. All models had high predictive validity which indicates the promising performance of machine learning methods in the breast cancer studies. Also, we used nonparametric statistical analysis for multiple comparisons of machine learning classification algorithms.

## کلمات کلیدی:

لینک ثابت مقاله در پایگاه سیویلیکا:

<https://civilica.com/doc/912377>

