Efficient machine learning algorithms for investigating and comparing the risk factors of breast and prostate cancers among males and females

SAMANEH RIKHTEHGARAN, WAZIR MUHAMMAD, Florida Atlantic University — Breast and prostate cancers are among the most common invasive cancers in the world. Although, these two cancers are different they share similarities in some aspects. Breast cancer is the leading of cancer in terms of the incidence among women worldwide. In breast cancer, the most important risk factor is probably family history. Risk factors other than family history are age, race, smoking, weight and alcohol consumption. On the other hand, prostate cancer is the third leading cause of cancer death in North America. Prostate cancer risk increases with the number of relatives diagnosed with this cancer and decreases with the relative’s age at diagnosis. Other factors than family history can be listed as age, ethnicity and genetic factors. In this project, we use several machine learning algorithms such as Logistic Regression, Decision Trees, and etc. to investigate risk factors of breast and prostate cancers among males and females and also to compare these risk factors in order to find the possible similarities or differences. For this reason, we use National Health Interview Survey (NHIS) datasets. Comprehensive comparison of risk factors leading to these two important cancers can have a huge impact on early detection and progressive improvement in survival.

Samaneh Rikhtehgaran
Florida Atlantic University

Date submitted: 12 Mar 2021

Electronic form version 1.4