Automatic Threshold Search for Heat Map Based Feature Selection: A Cancer Dataset Analysis

Authors : Carlos Huertas, Reyes Juarez-Ramirez

Abstract : Public health is one of the most critical issues today; therefore, there is great interest to improve technologies in the area of diseases detection. With machine learning and feature selection, it has been possible to aid the diagnosis of several diseases such as cancer. In this work, we present an extension to the Heat Map Based Feature Selection algorithm, this modification allows automatic threshold parameter selection that helps to improve the generalization performance of high dimensional data such as mass spectrometry. We have performed a comparison analysis using multiple cancer datasets and compare against the well known Recursive Feature Elimination algorithm and our original proposal, the results show improved classification performance that is very competitive against current techniques.