World Academy of Science, Engineering and Technology
International Journal of Biotechnology and Bioengineering
Vol:16, No:09, 2022

Identification of Biological Pathways Causative for Breast Cancer Using Unsupervised Machine Learning

Authors : Karthik Mittal

Abstract : This study performs an unsupervised machine learning analysis to find clusters of related SNPs which highlight biological pathways that are important for the biological mechanisms of breast cancer. Studying genetic variations in isolation is illogical because these genetic variations are known to modulate protein production and function; the downstream effects of these modifications on biological outcomes are highly interconnected. After extracting the SNPs and their effect on different types of breast cancer using the MRBase library, two unsupervised machine learning clustering algorithms were implemented on the genetic variants: a k-means clustering algorithm and a hierarchical clustering algorithm; furthermore, principal component analysis was executed to visually represent the data. These algorithms specifically used the SNP's beta value on the three different types of breast cancer tested in this project (estrogen-receptor positive breast cancer, estrogen-receptor negative breast cancer, and breast cancer in general) to perform this clustering. Two significant genetic pathways validated the clustering produced by this project: the MAPK signaling pathway and the connection between the BRCA2 gene and the ESR1 gene. This study provides the first proof of concept showing the importance of unsupervised machine learning in interpreting GWAS summary statistics.

Keywords : breast cancer, computational biology, unsupervised machine learning, k-means, PCA
Conference Title : ICBMI 2022 : International Conference on Bioinformatics and Medical Informatics
Conference Location : San Francisco, United States
Conference Dates : September 27-28, 2022

Open Science Index, Biotechnology and Bioengineering Vol:16, No:09, 2022 publications.waset.org/abstracts/148748.pdf