World Academy of Science, Engineering and Technology
International Journal of Biomedical and Biological Engineering
Vol:17, No:08, 2023

Unveiling Comorbidities in Irritable Bowel Syndrome: A UK BioBank Study utilizing Supervised Machine Learning

Authors : Uswah Ahmad Khan, Muhammad Moazam Fraz, Humayoon Shafique Satti, Qasim Aziz

Abstract : Approximately 10-14% of the global population experiences a functional disorder known as irritable bowel syndrome (IBS). The disorder is defined by persistent abdominal pain and an irregular bowel pattern. IBS significantly impairs work productivity and disrupts patients' daily lives and activities. Although IBS is widespread, there is still an incomplete understanding of its underlying pathophysiology. This study aims to help characterize the phenotype of IBS patients by differentiating the comorbidities found in IBS patients from those in non-IBS patients using machine learning algorithms. In this study, we extracted samples coding for IBS from the UK BioBank cohort and randomly selected patients without a code for IBS to create a total sample size of 18,000. We selected the codes for comorbidities of these cases from 2 years before and after their IBS diagnosis and compared them to the comorbidities in the non-IBS cohort. Machine learning models, including Decision Trees, Gradient Boosting, Support Vector Machine (SVM), AdaBoost, Logistic Regression, and XGBoost, were employed to assess their accuracy in predicting IBS. The most accurate model was then chosen to identify the features associated with IBS. In our case, we used XGBoost feature importance as a feature selection method. We applied different models to the top 10% of features, which numbered 50. Gradient Boosting, Logistic Regression and XGBoost algorithms yielded a diagnosis of IBS with an optimal accuracy of 71.08%, 71.427%, and 71.53%, respectively. Among the comorbidities most closely associated with IBS included gut diseases (Haemorrhoids, diverticular diseases), atopic conditions(asthma), and psychiatric comorbidities (depressive episodes or disorder, anxiety). This finding emphasizes the need for a comprehensive approach when evaluating the phenotype of IBS, suggesting the possibility of identifying new subsets of IBS rather than relying solely on the conventional classification based on stool type. Additionally, our study demonstrates the potential of machine learning algorithms in predicting the development of IBS based on comorbidities, which may enhance diagnosis and facilitate better management of modifiable risk factors for IBS. Further research is necessary to confirm our findings and establish cause and effect. Alternative feature selection methods and even larger and more diverse datasets may lead to more accurate classification models. Despite these limitations, our findings highlight the effectiveness of Logistic Regression and XGBoost in predicting IBS diagnosis.

Keywords : comorbidities, disease association, irritable bowel syndrome (IBS), predictive analytics
Conference Title : ICBDAH 2023 : International Conference on Big Data Analytics in Healthcare
Conference Location : New York, United States
Conference Dates : August 10-11, 2023