World Academy of Science, Engineering and Technology
International Journal of Medical and Health Sciences
Vol:18, No:12, 2024

Application of Deep Learning and Ensemble Methods for Biomarker Discovery in Diabetic Nephropathy through Fibrosis and Propionate Metabolism Pathways

Authors : Oluwafunmibi Omotayo Fasanya, Augustine Kena Adjei

Abstract : Diabetic nephropathy (DN) is a major complication of diabetes, with fibrosis and propionate metabolism playing critical roles in its progression. Identifying biomarkers linked to these pathways may provide novel insights into DN diagnosis and treatment. This study aims to identify biomarkers associated with fibrosis and propionate metabolism in DN. Analyze the biological pathways and regulatory mechanisms of these biomarkers. Develop a machine learning model to predict DN-related biomarkers and validate their functional roles. Publicly available transcriptome datasets related to DN (GSE96804 and GSE104948) were obtained from the GEO database (https://www.ncbi.nlm.nih.gov/gds), and 924 propionate metabolism-related genes (PMRGs) and 656 fibrosis-related genes (FRGs) were identified. The analysis began with the extraction of DN-differentially expressed genes (DN-DEGs) and propionate metabolism-related DEGs (PM-DEGs), followed by the intersection of these with fibrosis-related genes to identify key intersected genes. Instead of relying on traditional models, we employed a combination of deep neural networks (DNNs) and ensemble methods such as Gradient Boosting Machines (GBM) and XGBoost to enhance feature selection and biomarker discovery. Recursive feature elimination (RFE) was coupled with these advanced algorithms to refine the selection of the most critical biomarkers. Functional validation was conducted using convolutional neural networks (CNN) for gene set enrichment and immunoinfiltration analysis, revealing seven significant biomarkers—SLC37A4, ACOX2, GPD1, ACE2, SLC9A3, AGT, and PLG. These biomarkers are involved in critical biological processes such as fatty acid metabolism and glomerular development, providing a mechanistic link to DN progression. Furthermore, a TF–miRNA–mRNA regulatory network was constructed using natural language processing models to identify 8 transcription factors and 60 miRNAs that regulate these biomarkers, while a drug–gene interaction network revealed potential therapeutic targets such as UROKINASE–PLG and ATENOLOL–AGT. This integrative approach, leveraging deep learning and ensemble models, not only enhances the accuracy of biomarker discovery but also offers new perspectives on DN diagnosis and treatment, specifically targeting fibrosis and propionate metabolism pathways.

Keywords : diabetic nephropathy, deep neural networks, gradient boosting machines (GBM), XGBoost
Conference Title : ICDM 2024 : International Conference on Diabetes and Metabolism
Conference Location : Honolulu, United States
Conference Dates : December 30-31, 2024