Early Prediction of Cognitive Impairment in Adults Aged 20 Years and Older using Machine Learning and Biomarkers of Heavy Metal Exposure

Authors : Ali Nabavi, Farimah Safari, Mohammad Kashkooli, Sara Sadat Nabavizadeh, Hossein Molavi Vardanjani

Abstract : Cognitive impairment presents a significant and increasing health concern as populations age. Environmental risk factors such as heavy metal exposure are suspected contributors, but their specific roles remain incompletely understood. Machine learning offers a promising approach to integrate multi-factorial data and improve the prediction of cognitive outcomes. This study aimed to develop and validate machine learning models to predict early risk of cognitive impairment by incorporating demographic, clinical, and biomarker data, including measures of heavy metal exposure. A retrospective analysis was conducted using 2011-2014 National Health and Nutrition Examination Survey (NHANES) data. The dataset included participants aged 20 years and older who underwent cognitive testing. Variables encompassed demographic information, medical history, lifestyle factors, and biomarkers such as blood and urine levels of lead, cadmium, manganese, and other metals. Machine learning algorithms were trained on 90% of the data and evaluated on the remaining 10%, with performance assessed through metrics such as accuracy, area under curve (AUC), and sensitivity. Analysis included 2,933 participants. The stacking ensemble model demonstrated the highest predictive performance, achieving an AUC of 0.778 and a sensitivity of 0.879 on the test dataset. Key predictors included age, gender, hypertension, education level, urinary cadmium, and blood manganese levels. The findings indicate that machine learning can effectively predict the risk of cognitive impairment using a comprehensive set of clinical and environmental exposure data. Incorporating biomarkers of heavy metal exposure improved prediction accuracy and highlighted the role of environmental factors in cognitive decline. Further prospective studies are recommended to validate the models and assess their utility over time.