A Novel Multi-omics Based Data Integration Framework for Biological Age Prediction and Correlation With Chemotherapy Dosage for Personalized Therapeutics

Digania, Vanshika (School: Bharatiya Vidya Bhavan's R.K. Sarda Vidya Mandir)

Aging involves a gradual decline in physiological functions, heightening morbidity and mortality risks through diminished organ function, muscle mass, bone density, and a compromised immune system. Chronological age (CA), the conventional criterion, fails to capture diverse aging experiences among individuals of the same CA. To seek a better index, biological age (BA) is considered. Existing BA prediction methods are limited—single omic-specific, computationally intensive, prone to overfitting, and struggle with missing data. To address this, I propose a multi-omics-based machine learning model integrating genomics, epigenomics, metabolomics, phenomics, and proteomics biomarkers to predict the biological age of female breast cancer patients. Using meta-analysis, pathway analysis, and enrichment analysis, 21 biomarkers and 7 age-related clinical parameters were identified, showcasing their significance in aging. Trained on an augmented dataset, various machine learning algorithms, including Linear Regression, Random Forest, Gradient Booster Tree, XGBoost, AdaBoost, SVM, Quantile Regression, and a stacking ensemble were tested. The ensemble model demonstrated enhanced accuracy, surpassing the current gold standards. To address missing data challenges, Multiple Imputation by Chained Equations (MICE) was implemented, ensuring real-world clinical feasibility with large missing data sets. IBA57, F5, and ELOVL2 emerged as the top three crucial biomarkers in the aging process. This machine learning model for BA prediction holds promise in mitigating aging-related disorder risks and correlating with chemotherapy dosage for personalized treatment based on individual BA profiles, presenting a transformative impact on the intersection of aging and healthcare.