

Development of a Deep Learning-Based Prediction Model for Metabolic Syndrome using genomic data

Minhyuk Lee¹, Ji-Yeon Shin, and Mira Park^{3*}

¹*Department of Statistics, Korea University*

²*Department of Preventive Medicine, School of Medicine, Kyungpook National University*

²*Department of Preventive Medicine, School of Medicine, Eulji University*

*Corresponding author: mira@eulji.ac.kr

Metabolic syndrome (MetS) is a complex disorder characterized by a cluster of metabolic abnormalities, including abdominal obesity, hypertension, elevated triglycerides, reduced high-density lipoprotein cholesterol, and impaired glucose tolerance. It poses a significant public health concern, as individuals with MetS are at an increased risk of developing cardiovascular diseases and type 2 diabetes. Early and accurate identification of individuals at risk of MetS is essential.

In this study, we propose a methodology for predicting MetS using a multi-task deep learning approach. We compare the performance of this methodology to that of several single-task approaches, including support vector machine, CatBoost, LightGBM, and XGBoost. To build our prediction model, we utilize data from the Korean Association Resource (KARE) project, which consists of 327,872 SNPs on 8,840 individuals. In addition to genomic data, we incorporate lifestyle, dietary, and socio-economic factors that influence chronic diseases. By evaluating various measures such as accuracy, precision, and AUC, we demonstrate that our multi-task learning model outperforms conventional single-task machine learning models in MetS prediction.