ABSTRACT

Diabetes mellitus is a growing health problem in Indonesia, particularly in Central Java Province. This study aims to identify the dominant risk factors of diabetes cases using the Random Forest algorithm and map the results. The data used includes community lifestyle indicators such as the number of sweet drink stores, gyms, sport halls (GOR), transportation availability, karaoke venues, tourist attractions (tourism), fast food outlets, and alcohol stores in each regency/city in Central Java. The increasing prevalence of diabetes has a significant impact on public health. However, a comprehensive analysis of the dominant risk factors influencing diabetes cases in Central Java has yet to be conducted. Therefore, identifying these factors is essential to support effective prevention policies. This study employs Random Forest with optimal hyperparameter combinations to calculate feature importance globally (across the entirety of Central Java) and SHAP techniques to analyze feature importance locally (in each regency/city in Central Java). The results show that the number of tourist attractions and sweet drink stores are the primary global risk factors, with feature importance values of 0.2817 and 0.1502, respectively. Meanwhile, locally, tourism is the dominant factor in most areas of Central Java (18 regencies/cities or 51.4% of the total region). These findings provide valuable insights for designing diabetes prevention strategies in Central Java.

Keywords: *diabetes, risk factors, Random Forest, SHAP, Central Java, prevention.*