Body Mass Index Prediction with Machine Learning

Degree programme : MAS Data Science

Overweight and obesity are pressing global health challenges with serious health and economic implications. This work explores the potential of machine learning to predict BMI using Swiss and Latin American datasets. We achieved balanced accuracies of 0.67, with xAI improving interpretability and feature selection. The results demonstrate the practicality of machine learning and the potential of AI in Swiss healthcare.

Introduction

Overweight and obesity are critical global health challenges that pose significant health risks and economic costs. Effective prediction of body mass index (BMI) is essential to identify at-risk individuals and enable targeted preventive and therapeutic interventions.

Objectives

This thesis explores the potential of modern machine learning (ML) methods to accurately predict BMI using data from two different sources: a representative Swiss dataset (menuCH) and a Latin American dataset (LatAm).

Methodology

Several ML approaches were applied, including:
– Traditional algorithms: Standard ML methods for benchmarking,
– Neural networks: Advanced models for performance evaluation,
– Explainable AI (xAI): Techniques to improve interpretability,
– Regression models: For continuous BMI prediction.

Results

The HistGradientBoostingClassifier delivered the best results, with balanced accuracies of 0.67 (menuCH) and 0.68 (LatAm). For the obese class, it achieved recall scores of 0.71 (menuCH) and 0.88 (LatAm), and for the underweight class, 0.93 (menuCH) and 0.79 (LatAm). While neural networks produced comparable results, their black-box nature limited interpretability. In contrast, xAI methods improved feature selection and transparency, including actionable counterfactual explanations.

The features identified were highly consistent with established risk factors from leading healthcare organizations. These findings underscore the potential of ML for BMI prediction and broader healthcare applications, particularly in the early detection of conditions such as cardiovascular disease, type 2 diabetes, and mental health disorders.

Recommendations

To fully harness the potential of AI and ML in Swiss healthcare, it is essential to prioritize data quality, enhance interoperability through improved data integration across systems, and drive digital transformation by accelerating the adoption of AI technologies.

Conclusion

This research demonstrates the practicality of ML methods, particularly xAI-powered approaches, in predicting BMI and highlights their broader relevance in advancing healthcare initiatives.

Dr. Hans Rudolf Keller
hrk@ventivo.ch
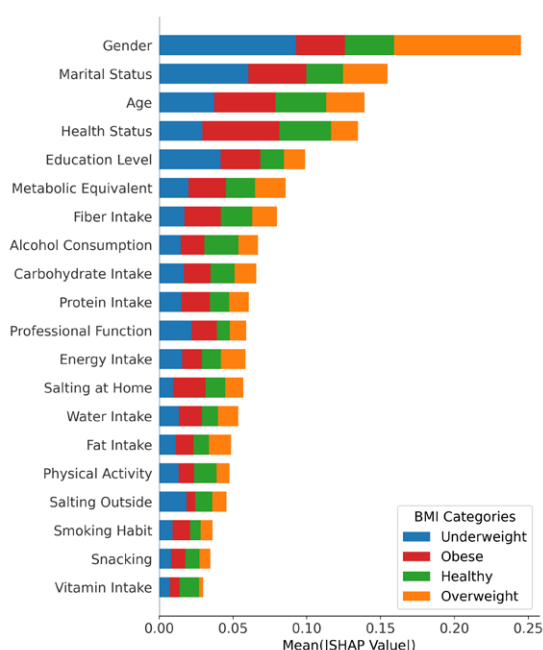
SHAP values (feature relevance) for menuCH dataset