

Abstract citation ID: ckae144.342

Predictive machine learning algorithms for metabolic syndrome among Tunisian adults

Mariem Zribi

*M Zribi^{1,2,3}, F Zaier², N Zoghlami^{1,6}, J Elati^{3,4,6}, P Traissac^{5,6},
H Aounallah-Skhiri^{1,2,3}*

¹National Institute of Health, Tunis, Tunisia

²Faculty of Medicine of Tunis, University Tunis El Manar, Tunis, Tunisia

³SURVEN Research, Tunis, Tunisia

⁴INNTA, National Institute of Nutrition and Food Technology, Tunis, Tunisia

⁵IRD, UMR 266 MoISA - University Montpellier, CIRAD, CIHEAM-IAMM, INRAE, Montpellier, France

⁶JEAL_TANIT, Tunis, Tunisia

Contact: zribi.mariem.fmt@gmail.com

Background: Metabolic syndrome (MS), a cluster of interconnected risk factors for cardiovascular disease and type 2 diabetes, poses a significant health burden globally. In Tunisia, amidst an ongoing epidemiological transition, the prevalence of MS presents a growing concern. Our study aimed to identify the potential risk factors of MS and propose better machine learning (ML) based models for predicting MS among Tunisian adults.

Methods: Data was sourced from the Tunisian Health Examination Survey THES-2016, encompassing individuals aged 20 years and older. MS was defined based on the criteria set by the International Diabetes Federation. Logistic regression (LR) was employed to determine risk factors for MS. Adjusted odds ratios (aOR) and 95% confidence intervals (CI) were calculated. Five ML algorithms (Naïve Bayes, Support Vector Machine (SVM), Artificial Neural Network (ANN), AdaBoost, and Random Forest (RF)) were utilized to predict MS. Performance evaluation was conducted using accuracy, precision, recall, and area under the curve (AUC) metrics.

Results: Among the 8908 participants, the prevalence of MS was 32.8% (95% CI: 31.4% - 34.6%). LR identified age group 60-69 years (aOR: 18.5, 95% CI: 16.9-19.9, $p < 0.001$), female sex (aOR: 1.7, 95% CI: 1.2-1.9), sedentariness (aOR: 1.6, 95% CI: 1.5-1.8) and low education level (aOR: 1.2, 95% CI: 1.1-1.4) as significant factors associated with MS. Among ML models tested, AdaBoost had the highest accuracy (89.8%), followed by SVM (89.6%). Naïve Bayes had the highest Recall (98.1%) and the most performant AUC (91.4%), while RF had the highest Precision (70.8%). The ANN had an accuracy of 89%, a precision of 70.4%, a recall of 78.7% and an AUC of 85.1%.

Conclusions: Our study highlighted the effectiveness of ML algorithms like AdaBoost and SVM in predicting MS among Tunisian adults, offering potential as early detection tools. However, further validation with larger datasets is necessary to solidify their utility in healthcare settings.

Key messages:

- Metabolic syndrome, a complex of cardiovascular risk factors, presents a substantial global public health challenge and is particularly pronounced in Tunisia.
- Machine learning algorithms, notably AdaBoost and Support Vector Machine, demonstrate promise in forecasting and addressing metabolic syndrome in Tunisian adults.