JURNAL EMACS
[Engineering, MAthematics and Computer Science] Vol.8 No.1 January 2026: 19-25
e-ISSN: 2686-2573
DOI: 10.21512/emacsjournal.v8i1.15268

Machine Learning Approach: A Comparative Analysis of Classifiers in Predicting Obesity Type

Jeffrey Junior Tedjasulaksana[1*], Ferry Jaya Dinata[2], Rafael Krisnadi[3], Matthew S.W. Reksosamudro[4], Wilbert Wen[5], Muhammad Fadlan Hidayat[6]

[1-6] Computer Science Department, School of Computer Science
Bina Nusantara University,
Jakarta, Indonesia 11480
jeffrey.t@binus.ac.id; ferry.dinata@binus.ac.id; rafael.krisnadi@binus.ac.id; matthew.reksosamudro@binus.ac.id; wilbert.wen@binus.ac.id; muhammad.hidayat@binus.ac.id

*Correspondence: jeffrey.t@binus.ac.id

Abstract – *Obesity is a growing global public health concern that increases the risk of chronic diseases and significantly affects quality of life. Traditional diagnostic methods such as Body Mass Index (BMI) have limitations in accurately representing body fat distribution and individual health conditions. This study aims to comparatively evaluate the performance of various machine learning and neural network models in predicting obesity levels using a multiclass classification approach. The dataset consists of 2,111 observations with 12 predictor variables and seven obesity categories, obtained from a publicly available source. Data preprocessing included duplicate removal, outlier handling using the interquartile range method, feature scaling, and categorical encoding, followed by a 60:20:20 train–validation–test split. Several classifiers were implemented, including Logistic Regression, Support Vector Classifier, Random Forest, Extra Trees, Gradient Boosting-based models (XGBoost and LightGBM), Multilayer Perceptron, K-Nearest Neighbors, and TabNet. Model performance was evaluated using macro-average F1-score and confusion matrix analysis. The results indicate that LightGBM achieved the highest predictive performance with an F1-score of 0.96, demonstrating strong generalization across obesity categories. XGBoost and Random Forest also showed strong performance, while Support Vector Classifier exhibited consistent results across training, validation, and cross-validation. These findings suggest that ensemble-based models are highly effective for obesity classification, while model selection should consider accuracy, interpretability, and computational constraints.*

Keywords: *health; machine learning; neural network; hyperparameter tuning; AI application*

I. INTRODUCTION

One of the key factors in keeping good health is maintaining an ideal body weight. Excess body weight, one of which is obesity, can significantly impact our health by increasing the risk of various diseases and health problems. According to a study, Obesity is an increasing, global public health issue. Patients with obesity are at major risk for developing a range of comorbid conditions, including cardiovascular disease (CVD), gastrointestinal disorders, type 2 diabetes (T2D), joint and muscular disorders, respiratory problems, and psychological issues, which may significantly affect their daily lives as well as increasing mortality risks (Bae et al., 2025; Fruh, 2017; Habehh & Gohel, 2021; Mahindru et al., 2023).

A research paper found that in 2015, elevated Body Mass Index (BMI) contributed to approximately 4 million deaths worldwide—and over two-thirds of those deaths were due to cardiovascular disease, with 41% of BMI-related deaths and 34% of BMI-related disability-adjusted life years attributed to cardiovascular diseases (Poirier et al., 2009). This statistical statement represents how obesity can independently cause comorbid conditions including cardiovascular disease, type 2 diabetes, hypertension, joint disorders, respiratory problems, liver disease, gastrointestinal disorders, and mental health challenges (Airlangga, 2025; Fitch & Bays, 2022; Hemmingsson et al., 2022). A 2025 study further highlighted the link between obesity and chronic inflammatory diseases (CIDs), revealing that obese individuals have a 28.48% risk of developing any CID compared to non-obese individuals with a 17.55% risk factor (Khun & Johnson, 2013). These findings indicate that obesity increases the risk for

Copyright © 2026

an individual to develop CIDs, thereby posing a broader threat to overall health and well-being.

Body Mass Index, one of the most popular traditional diagnostic methods is used to measure body fat and applicable to both men and women. It is quick and moderately easy to use but, it is increasingly clear that BMI is a rather poor indicator of percent of body fat (Nuttall, 2015). Body Mass Index (BMI) has clear disadvantages that it does not consider variations such as age, sex, and ethnicity (Shi et al., 2022). Moreover, it also does not differentiate between muscle mass and fat which may lead to inaccuracy of predicting body fat (Wu et al., 2024). These shortcomings underscore the need for a more accurate and comprehensive diagnostic method.

With the rapid advancement of technology and machine learning techniques, scientists have started to apply and integrate AI into the health sector. The application of AI has the capacity to assist with case triage and diagnoses, enhance image scanning and segmentation, support decision making, predict the risk of disease, and in neuroimaging (Habehh & Gohel, 2021; Lin et al., 2023). Creating a model to predict post-stroke pneumonia using deep neural network and utilizing deep learning methods for cancer detection are both examples of the transformative potential in integrating AI into the health sector (Ge et al., 2019; Kumar & Alqahtani, 2022).

Recent advancements have led to the development of technologies that integrate machine learning (ML) techniques with the healthcare sector. One significant application is the use of machine learning models in predicting and classifying chronic diseases such as obesity (Mousavi et al., 2025). Traditional diagnostic methods often struggle to detect complex, non-linear patterns in patient data, limiting their effectiveness in early detection and personalized treatment. This limitation has driven researchers to explore ML classifiers as more accurate, scalable, and data-driven alternatives for addressing obesity and related health issues (Aggarwal & Pandey, 2023).

Several studies have highlighted the effectiveness of traditional machine learning classifiers in the domain of obesity prediction. Logistic Regression (LR) has been widely adopted due to its simplicity and interpretability. However, it also has its own limitations in capturing complex relationships in health data due to its assumption of linearity between features and output limits (Dirik, 2023). In contrast, Decision Trees (DT) and Random Forests (RF) have proven effective at modeling nonlinear interactions, with RF particularly excelling due to its ensemble nature, which mitigates overfitting and improves generalization (Safaei et al., 2021). Support Vector Classifier with a radial basis function (RBF) kernel could also be considered an alternative due to its capability in handling high-dimensional and non-linear data (Poliban, 2025).

Due to recent advancements, ensemble learning techniques such as Gradient Boosting Machines (GBM), XGBoost, and LightGBM have become more popular for their capability to build strong models from multiple weak learners (Jeon et al., 2022). These models have been tested to outperform single classifiers in various healthcare prediction tasks, including obesity classification, by effectively modeling complex feature relationships and handling missing or noisy data (Aggarwal & Pandey, 2023).

The use of deep learning in medical diagnostics has also seen significant growth and has rapidly advanced. A comprehensive review of deep learning applications in cancer detection and diagnosis, outlining various models, methodologies, and their potential for high precision in pattern recognition, was one of the examples provided by a recent study (Kumar & Alqahtani, 2022). Similarly, a researcher explored the effectiveness of Convolutional Neural Networks (CNNs) models when combined with XGBoost or Feedforward Neural Networks in COVID-19 Detection, highlighting improved accuracy when leveraging multiple algorithms in ensemble or hybrid models (Dumakude & Ezugwu, 2023). These findings represent facts that hybrid models can leverage the strengths of different algorithms for enhanced performance and can be applied to obesity prediction.

In addition to ensemble and deep learning models, several classical and modern classifiers offer distinct advantages and limitations in the context of obesity prediction. Multilayer Perceptron (MLP), a type of feedforward artificial neural network, is well-suited for capturing complex, non-linear relationships in data, which is beneficial when modeling multifactorial conditions like obesity (Chauhan & Srivastava, 2025). Similarly, TabNet, a more recent deep learning architecture designed for tabular data, addresses the interpretability issue through sequential attention and built-in feature selection mechanisms, making it particularly attractive for healthcare applications where transparency is critical (Arik & Pfister, 2019). On the other hand, K-Nearest Neighbors (KNN) offers simplicity and ease of implementation, and it makes no assumptions about data distribution—an advantage when working with heterogeneous health data (Zeedhan et al., 2025). Collectively, these models provide a range of trade-offs between accuracy, interpretability, and computational efficiency, highlighting the importance of context-aware model selection in obesity research.

In conclusion, existing literature infers that no single classifier consistently outperforms others across all obesity prediction scenarios. Instead, the choice of model should be guided by the nature of the dataset, interpretability requirements, and computational constraints. The emerging trend favors ensemble and hybrid models, which integrate the advantages of different learning techniques and offer robust, high-performing solutions for obesity classification tasks.

This research paper aims to create and compare a selection of predictive models that can accurately classify individuals based on their likelihood of being obese, using a range of features. By leveraging ML techniques, this research seeks to enhance the accuracy and efficiency of obesity detection, thereby supporting early intervention strategies and contributing to better health outcomes (Safaei et al., 2021).

II. METHODS

2.1 Flowchart

Figure. 1 here shows our model's pipeline, where it starts at data collection, which is collected from a Kaggle dataset.

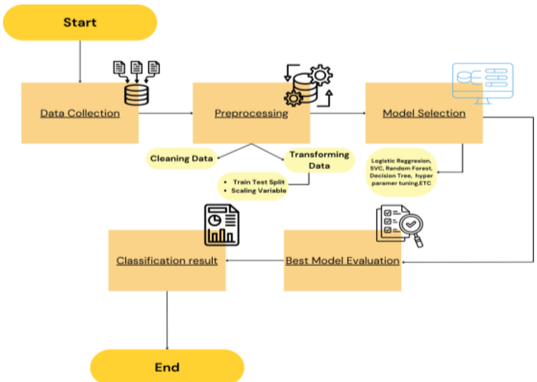

Figure 1. Flowchart

This dataset then went through a preprocessing process, which included cleaning, which handles missing values and removing duplicates. Then, transforming the data, which scales numerical features and encodes categorical variables, and afterwards splitting them into a train-validation-test split. The preprocessed data then goes through various model selection, including ML models and neural net models, using a validation set to optimize the parameters (hyperparameter tuning). The best model out of said fitting is then tested on the test set to find out its evaluation based on the best F1-score on the validation set. We also generate a confusion matrix to evaluate the model. The classification result includes the final prediction produced for the new data and the performance metrics that it reported.

2.2 Dataset

The dataset consists of 2.111 observations with 12 predictors. With a multiclass target variable of obesity levels of seven classes. This dataset then went through the removal of duplicate values and outlier removal with the IQR (InterQuartile Range) method, specifically for both the age feature and the NCP feature. The clean dataset is then ready for splitting, which is split into 60:20:20 proportions for the train, validation, and test sets, respectively. Finally, the dataset goes through scaling of both numerical and categorical encoding, where for categorical encoding, two methods were used: the first label encoding for tree-type ML (machine learning) models, while the second one-hot encoding for the rest. Whereas the standard scaler is used for numerical scaling.

2.3 Preprocessing

Before constructing any of the predictive models, we did a thorough Exploratory Data Analysis (EDA) on the dataset of 2.111 observations with 13 features and found that. There are no class imbalances, whereas the classes have all equal weight. Second, there were no missing values, but there were a few duplicate values. Third, the distribution of a few features contained outliers, namely age and the NCP (Number of main meals daily) feature. Lastly, we found that other than weight and height, there is no significant correlation between the features.

III. RESULTS AND DISCUSSION

3.1 Correlation Matrices

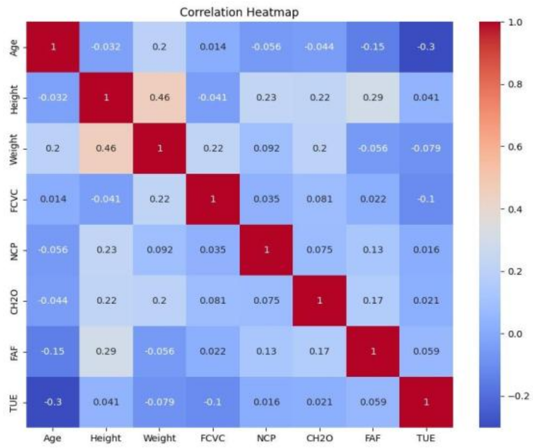

Figure 2. Correlation Matrices of Numerical Variables

From Figure. 2 it can be seen that it is a heatmap that shows the correlation between numeric variables. Immediately we can see that most variables have either weak or low correlation with each other. There are some variables that differ such as Height and Weight. The correlation between the two is a moderate positive correlation of 0.46, which makes sense since taller people will weigh more.

On the other spectrum is the correlation between Age and TUE (Time of Technology Use), which shows a negative correlation of -0.30, which suggests that younger individuals spend more time with technology than older ones.

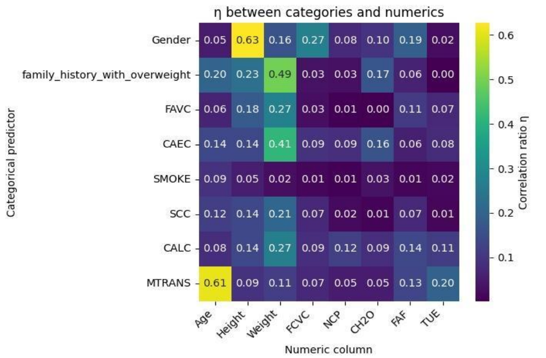

Figure 3. Correlation Matrices of Categorical and Numerical Variables

Figure. 3 is another heat map that shows the strength of the relationship between categorical and numerical variables using the eta coefficient. From the matrix, it can be seen that the correlation between Gender and Height is one of the strongest, with 0.63, since height is usually different between males and females. Another strong one is between MTRANS (Mode of Transport) and Age, with 0.61, showing that different age groups prefer different transport methods. There are also some that show moderate correlation, such as Family History with Overweight, with weight being 0.49, suggesting that genetics or shared family habits might affect body weight.

3.2 Model's Evaluation

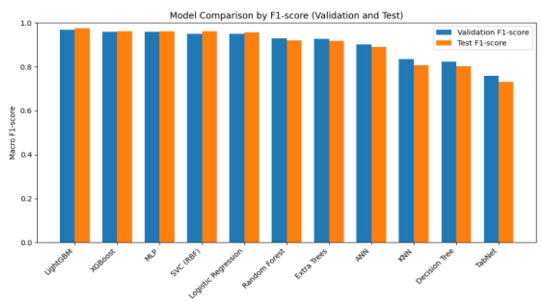

Figure 4. F1-Score Graph of Machine Learning Models

Figure. 4 is a graph that shows the F1-score used by these models, with TabNet having the worst F1-score (around 0.9 F1-score) from all of the models used, with LightGBM having the best F1-score and XGBoost ranking second behind. Extra Trees and Random Forest also performed well with an F1-score of 0.95 on both tests, while traditional methods like logistic regression and SVC scored around 0.94.

It could be seen that out of all the neural networks, LightGBM managed to perform the best with a macro-average F1-score of 0.96 on both the test and validation sets, which indicates strong generalization.

3.3 Confusion matrix of the most accurate model

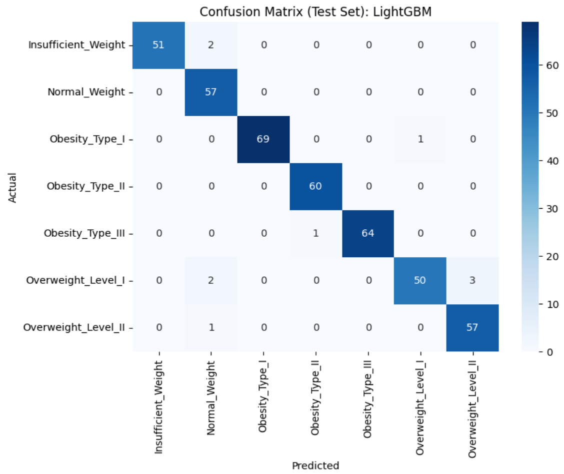

Figure 5. Confusion Matrix of LightGBM

Figure. 5 is a confusion matrix for LightGBM. It can be seen that LightGBM got most of the data correct in each target variable, which is indicated by the diagonal entries (true positives). With Obesity_Type_II and Normal_Weight, LightGBM got no mistakes. While the model has a few errors (indicated by the off-diagonal errors), like Overweight_Level_I and II, and Obesity_Type_I and III. These misclassifications occur mostly between adjacent categories, which is understandable given subtle feature differences. Overall, LightGBM demonstrated excellent per-class accuracy with very few confusions, validating its selection as the best model.

The results of our experiments highlighted the strengths and weaknesses of each machine learning model applied to obesity classification tasks. Among the models, LightGBM proved to be the most accurate model. It has proven to be the best among the models in effectively handling the multiclass nature of the dataset and classifying instances among obesity levels. Similarly, XGboost followed closely in terms of performance, validating the efficiency of gradient boosting frameworks in structured medical data classification. However, Support Vector Classifier (SVC) showcased the best consistency on training, validation, and cross-validation metrics. SVC demonstrated stable generalization across different datasets and highlights better reliability and reproducibility than other models, which are crucial in medical applications and could be a more preferable model compared to other models that focus on performance.

V. CONCLUSION

This study demonstrates the effectiveness of various models for multi-class classification tasks in medical datasets. While models like LightGBM and XGBoost achieved the best in terms of accuracy, SVC shows it is the most consistent in training, validation and cross-validation loss metrics, suggesting strong generalizablility. Our results indicate that each model has individual strengths and weaknesses, therefore the selection of an optimal model depends on the specific requirement of the deployment conditions such as accuracy, reliability, consistency, interpretability, and resource constraints. Future research could focus on hybrid models that utilize the strength of multiple algorithms to further enhance predictive performance or clinical applicability and utilize larger datasets to further improve robustness, generalizability and address problems such as class imbalance and overfitting.

AUTHOR’S CONTRIBUTION

This study was conceptualized and coordinated by Jeffrey Junior Tedjasulaksana as the project chairman, who also conducted the experiments, under the supervision of Muhammad Fadlan Hidayat. Ferry Jaya Dinata, Rafael Krisnadi, Matthew S.W. Reksosamudro, and Wilbert Wen contributed to data curation, methodology development, experimentation, and analysis. Conceptualization, data curation, formal analysis, methodology, validation, visualization, and writing were carried out collaboratively by all authors. All authors have read and approved the final version of the manuscript.

AVAILABILITY DATA AND MATERIALS

The dataset used in this study is publicly available and can be accessed through the Kaggle repository. It provides the data required to reproduce the experiments and analyses conducted in this research. The dataset is available at: *https://www.kaggle.com/datasets/adeniranstephen/obesity-prediction-dataset*

REFERENCES

Aditya Mahindru, Pradeep Patil, Varun Agrawal. (2023). Role of Physical Activity on Mental Health and Well-Being: A Review, 1-7, DOI: 10.7759/cureus.33475

Airlangga, G. (2025). A comparative analysis of machine learning models for obesity prediction. Jurnal Informatika Ekonomi Bisnis, 7(1). https://doi.org/10.37034/infeb.v7i1.1089

Ali, A., Wei, Y., Tyson, J., Akerman, H., Jackson, A. I. R., Lane, R., Spencer, D., & White, N. M. (2024). Enhancing the response of a wearable sensor for improved respiratory rate (RR) monitoring. IEEE Access. https://doi.org/10.1109/ACCESS.2024.3509676

Alimbayeva, Z., Alimbayev, C., Ozhikenov, K., Bayanbay, N., & Ozhikenova, A. (2024). Wearable ECG Device and Machine Learning for Heart Monitoring. Sensors 2024, Vol. 24, Page 4201, 24(13), 4201. https://doi.org/10.3390/S24134201

Angela K. Fitch, Harold E. Bays. (2022). Obesity definition, diagnosis, bias, standard operating procedures (SOPs), and telehealth: An Obesity Medicine Association (OMA) Clinical Practice Statement (CPS) 2022, Obesity Pillars, Volume 1, 100004, ISSN 2667-3681, https://doi.org/10.1016/j.obpill.2021.100004

Arik, S. O., & Pfister, T. (2019). TabNet: Attentive interpretable tabular learning. arXiv. https://doi.org/10.48550/arXiv.1908.07442

Ba, N., Yue, W., Cao, C., Wu, W., & Cheng, P. (2024). Advances in Wearable Smart Chemical Sensors for Health Monitoring. Applied Sciences 2024, Vol. 14, Page 11199, 14(23), 11199. https://doi.org/10.3390/APP142311199

Bae, J. P., Nelson, D. R., Boye, K. S., & Mather, K. J. (2025). Prevalence of complications and comorbidities associated with obesity: A health insurance claims analysis. BMC Public Health, 25(1). https://doi.org/10.1186/s12889-024-21061-z

Chauhan, P., & Srivastava, S. (2025). Comparative analysis of deep learning and machine learning techniques for obesity classification. Atlantis Press.

Dirik, M. (2023). Application of machine learning techniques for obesity prediction: A comparative study. Journal of Complexity in Health Sciences, 6(2), 16–34. https://doi.org/10.21595/chs.2023.23193

Dumakude, A., & Ezugwu, A. E. (2023). Automated COVID-19 detection with convolutional neural networks. Scientific Reports, 13(1). https://doi.org/10.1038/s41598-023-37743-4

Erik Hemmingsson, Paulina Nowicka, Stanley Ulijaszek,Thorkild I. A. Sørensen. (2022). The social origins of obesity within and across generations, 1-11. https://doi.org/10.1111/obr.13514

Fruh, S. M. (2017). Obesity. Journal of the American Association of Nurse Practitioners, 29(S1), S3–S14. https://doi.org/10.1002/2327-6924.12510

Ge, Y., Wang, Q., Wang, L., Wu, H., Peng, C., Wang, J., Xu, Y., Xiong, G., Zhang, Y., & Yi, Y. (2019). Predicting post-stroke pneumonia using deep neural network approaches. International Journal of Medical Informatics, 132, 103986. https://doi.org/10.1016/j.ijmedinf.2019.103986

Habehh, H., & Gohel, S. (2021). Machine learning in healthcare. Current Genomics, 22(4), 291–300. https://doi.org/10.2174/1389202922666210705124359

Ikharo, B. A., & Aliu, D. (2023). Challenges Associated with Wearable Internet-of-Things (IoTs) Monitoring Systems for E-Health. FUOYE Journal of Engineering and Technology, 8(4). https://doi.org/10.46792/FUOYEJET.V8I4.1099

Jakachira, R., Jakachira, R., Yan, W., Burrow, J. A., Toussaint, K. C., & Toussaint, K. C. (2024). Dual-Wavelength, Polarization-Sensitive Wearable Photoplethysmographic Sensor on Diverse Skin Tones. Frontiers in Optics + Laser Science 2024 (FiO, LS) (2024), Paper JW4A.65, JW4A.65. https://doi.org/10.1364/FIO.2024.JW4A.65

Jeon, J., Lee, S., & Oh, C. (2022). Age-specific risk factors for the prediction of obesity using a machine learning approach. Research Square Platform LLC. https://doi.org/10.21203/rs.3.rs-1515734/v1

Joyce, D., De Brún, A., Symmons, S. M., Fox, R., & McAuliffe, E. (2023). Remote patient monitoring for COVID-19 patients: comparisons and framework for reporting. BMC Health Services Research, 23(1), 1–11. https://doi.org/10.1186/S12913-023-09526-0/TABLES/3

Kajzar, M. (2024). Wearable Devices for Training and Patient Monitoring: A Comprehensive Review. Quality in Sport, 29, 55667. https://doi.org/10.12775/QS.2024.29.55667

Kuhn, M., & Johnson, K. (2013). Applied predictive modeling. Springer New York. https://doi.org/10.1007/978-1-4614-6849-3

Kumar, G., & Alqahtani, H. (2022). Deep learning-based cancer detection-recent developments, trend and challenges. Computer Modeling in Engineering & Sciences, 130(3), 1271–1307. https://doi.org/10.32604/cmes.2022.018418

Li, Z., Li, H., & Gianchandani, Y. B. (2024). A Disposable Sensor for PM2.5 and PM10 Based on Wireless Magnetoelastic Resonators. Proceedings of IEEE Sensors. https://doi.org/10.1109/SENSORS60989.2024.10785131

Lin, W., Shi, S., Huang, H., Wen, J., & Chen, G. (2023). Predicting risk of obesity in overweight adults using interpretable machine learning algorithms. Frontiers in Endocrinology, 14. https://doi.org/10.3389/fendo.2023.1292167

Linh, V. T. N., Han, S., Koh, E., Kim, S., Jung, H. S., & Koo, J. (2025). Advances in wearable electronics for monitoring human organs: Bridging external and internal health assessments. Biomaterials, 314, 122865. https://doi.org/10.1016/J.BIOMATERIALS.2024.122865

Mahmood Safaei, Elankovan A. Sundararajan, Maha Driss, Wadii Boulila, Azrulhizam Shapi'i. (2021). A systematic literature review on obesity: Understanding the causes & consequences of obesity and reviewing various machine learning approaches used to predict obesity, Computers in Biology and Medicine, Volume 136, 104754, ISSN 0010-4825, https://doi.org/10.1016/j.compbiomed.2021.104754

Mousavi, S., Bieber, K., Zirpel, H., Vorobyev, A., Olbrich, H., Papara, C., De Luca, D. A., Thaci, D., Schmidt, E., Riemekasten, G., Lamprecht, P., Laudes, M., Kridin, K., & Ludwig, R. J. (2025). Large-scale analysis highlights obesity as a risk factor for chronic, non-communicable inflammatory diseases. Frontiers in Endocrinology, 16. https://doi.org/10.3389/fendo.2025.1516433

Nuttall, F. Q. (2015). Body mass index. Nutrition Today, 50(3), 117–128. https://doi.org/10.1097/nt.0000000000000092

Poirier, P., Eckel, R. H., Lavie, C. J., Van Gaal, L. F., Ross, R., Després, J. P., & Sharma, A. M. (2006). Obesity and cardiovascular disease: A scientific statement from the American Heart Association. Circulation, 113(6), 898–918 https://doi.org/10.1161/CIRCULATIONAHA.106.171016.

Poliban. (2025). Radial basis function model for obesity classification (PDF). Eltikom Poliban.

Safaei, M., Sundararajan, E. A., Driss, M., Boulila, W., & Shapi'i, A. (2021). A systematic literature review on obesity: Understanding the causes & consequences of obesity and reviewing various machine learning approaches used to predict obesity. Computers in Biology and Medicine, 136, 104754. https://doi.org/10.1016/j.compbiomed.2021.104754

Shivani Aggarwal, Kavita Pandey. (2023). Early identification of PCOS with commonly known diseases: Obesity, diabetes, high blood pressure and heart disease using machine learning techniques, Expert Systems with Applications, Volume 217, 119532, ISSN 0957-4174, https://doi.org/10.1016/j.eswa.2023.119532

Wu, Y., Li, D., & Vermund, S. H. (2024). Advantages and limitations of the body mass index (BMI) to assess adult obesity. International Journal of Environmental Research and Public Health, 21(6), 757. https://doi.org/10.3390/ijerph21060757

Xiaoyi Shi, Yuxin Zheng, Haiwen Cui, Yuxi Zhang, Menghui Jiang. (2022). Exposure to outdoor and indoor air pollution and risk of overweight and obesity across different life periods: A review, Ecotoxicology and Environmental Safety, Volume 242, 113893, ISSN 0147-6513, https://doi.org/10.1016/j.ecoenv.2022.113893

Zeedhan, M., Mohamed Ziham, M. M., Abdul Razick, M. S., & Ul Amin, N. (2025). Predicting obesity classification using k-nearest neighbors: A data science approach in Python. Preprints.

Zhang, Y., Wang, H., Cui, J., He, T., Qiu, G., Xu, Y., & Zhang, J. (2024). An ultraviolet photodetector based on conductive hydrogenated TiO2 film prepared by radio frequency atmospheric pressure plasma. Journal of Physics D: Applied Physics, 57(38), 385201. https://doi.org/10.1088/1361-6463/AD584B