

Smart Ai obesity risk classification

Rokkam.Mounika

dept.of computer applications
Satya institute of technology and
management
vizianagaram,India
rokkam.mounika9630@gmail.com

A.sai lavanya

dept. of computer applications
satya institute of technology and
management
vizianagaram,India
email address or ORCID

Abstract— Obesity has emerged as one of the most critical public health challenges of the 21st century, contributing significantly to chronic diseases such as diabetes, cardiovascular disorders, and metabolic syndromes. Traditional statistical approaches and black-box machine learning models have been widely applied for obesity risk prediction; however, their lack of interpretability limits trust, clinical acceptance, and ethical deployment in healthcare settings. This research proposes an Explainable Artificial Intelligence (XAI)-enhanced machine learning framework for obesity risk classification that not only delivers accurate predictions but also provides transparent and human-interpretable explanations for decision making. The system integrates advanced machine learning classifiers with XAI techniques to uncover how physiological, behavioral, and lifestyle features influence obesity risk. By incorporating explainability methods such as feature attribution and local explanation models, the proposed approach bridges the gap between predictive performance and clinical interpretability. The framework aims to support healthcare professionals by offering clear insights into contributing risk factors, enabling personalized intervention strategies and early prevention. Experimental evaluation demonstrates that the XAI-enhanced models achieve competitive accuracy while significantly improving model transparency and trustworthiness. This study highlights the importance of explainability in healthcare-oriented machine learning systems and provides a robust foundation for ethical, reliable, and user-centric obesity risk assessment solutions.

Keywords:

SDN- Software Defined Networks, CNN-Convolutional Neural Networks, RNN-Recurrent Neural Networks, DL-Deep Learning, GRU-Gated Recurrent Unit, LSTM-Long Short-Term Memory, SDCNN- Structured Deep Convolutional Neural Network.

I. INTRODUCTION

Obesity is a complex, multifactorial condition influenced by genetic, physiological, behavioral, environmental, and socioeconomic factors. The increasing prevalence of obesity across all age groups has placed immense pressure on healthcare systems worldwide, necessitating early detection and preventive strategies. With the rapid growth of digital health records and wearable devices, large volumes of health-related data are now available, creating opportunities for machine learning-based predictive modeling. Machine learning techniques have shown promising results in identifying obesity risk patterns by *analyzing*

multidimensional data such as body mass index, dietary habits, physical activity levels, and metabolic indicators. However, most high-performing models, including deep learning and ensemble techniques, operate as black boxes, providing predictions without clear explanations. This opacity poses a major challenge in clinical environments where interpretability, accountability, and trust are essential[1],[2],[3].

Explainable Artificial Intelligence has emerged as a solution to address the transparency limitations of conventional machine learning. XAI focuses on making model decisions understandable to humans by explaining how input features contribute to specific predictions. In the context of obesity risk classification, explainability is crucial for enabling healthcare professionals to validate model outcomes, understand causal relationships, and design effective intervention strategies. An interpretable model can reveal whether obesity risk is primarily driven by lifestyle behaviors, metabolic factors, or demographic attributes, thereby supporting personalized healthcare planning. This research integrates XAI techniques with machine learning classifiers to create a system that balances predictive accuracy with interpretability. By embedding explainability into the model design, the study aims to enhance clinical acceptance, reduce algorithmic bias, and promote ethical AI adoption in healthcare. The introduction establishes the motivation for combining machine learning and XAI and emphasizes the necessity of transparent predictive systems for obesity risk assessment[4],[5],[6].

1.3 Scope of Research

The scope of this research focuses on designing, implementing, and evaluating an XAI-enhanced machine learning framework for obesity risk classification. The study considers structured health datasets containing demographic, physiological, and lifestyle attributes relevant to obesity. The research emphasizes supervised machine learning models suitable for classification tasks and integrates explainability techniques to interpret both global and local model behavior. Global explanations aim to identify overall feature importance and population-level risk patterns, while local explanations focus on individual predictions to support personalized healthcare insights[16],[17],[18].

The research scope includes data preprocessing, feature selection, model training, and performance evaluation using standard metrics such as accuracy, precision, recall, and F1-score. Special attention is given to explainability evaluation, assessing how effectively the XAI methods communicate model reasoning to healthcare professionals. The study does

not aim to replace clinical judgment but rather to augment it by providing decision support tools that enhance understanding and transparency. The scope excludes real-time clinical deployment and focuses instead on experimental validation and conceptual framework development. Ethical considerations, including bias reduction and responsible AI usage, are also addressed within the research boundaries. By combining machine learning performance with explainability, this research contributes to the development of trustworthy AI systems in healthcare, specifically targeting obesity risk assessment and preventive medicine[19],[20].

II. RELATED WORK

Many researchers have used machine learning to predict obesity. Earlier methods like Logistic Regression and Decision Trees were easy to understand but not very accurate.

Later, advanced models like Random Forest, Gradient Boosting, and Deep Learning gave better accuracy. But these models are difficult to understand because they work like black boxes.

To solve this problem, new techniques called Explainable AI (XAI) like SHAP and LIME are used. These help explain how the model gives predictions.

So, recent research focuses on combining accuracy + explanation. Your project also follows this idea by using machine learning with XAI to make results both accurate and understandable.

III. METHODOLOGY

The proposed XAI-enhanced machine learning methodology integrates predictive modeling with explainability to classify obesity risk while maintaining transparency. The workflow begins with preprocessed data being divided into training and testing subsets to enable unbiased evaluation. Machine learning classifiers such as logistic regression, support vector machines, random forests, or gradient boosting models are trained to predict obesity risk levels based on health and lifestyle features. After model training, explainable artificial intelligence techniques are incorporated to interpret predictions. Model-agnostic explainability methods such as SHAP or LIME are employed to identify feature contributions at both global and individual levels. Global explanations help understand overall obesity risk factors across the population, while local explanations provide patient-specific insights into why a particular individual is classified as high or low risk. This combined methodology ensures that the model not only achieves high accuracy but also supports trust, accountability, and clinical relevance by enabling healthcare professionals to understand and validate the decision-making process.

A. Feature selection techniques

Feature selection plays a vital role in improving model performance and interpretability in XAI-based obesity risk classification systems. Redundant or irrelevant features can increase model complexity and obscure explanations, making it difficult to identify meaningful health indicators.

Filter-based techniques such as correlation analysis and mutual information are first applied to assess the statistical relationship between features and obesity outcomes. Highly correlated or non-informative features are removed to reduce dimensionality. Wrapper-based methods, including recursive feature elimination, are then used to evaluate subsets of features based on model performance. Embedded methods such as feature importance scores from tree-based models further refine the selection by identifying attributes that contribute most to predictions. The final feature set is validated using explainability outputs to ensure that selected features align with medical knowledge, thereby enhancing both predictive reliability and interpretive clarity.

Algorithm Pseudo code steps

Step 1: Load the obesity risk dataset containing demographic, physiological, and lifestyle features.

Step 2: Perform data preprocessing including missing value handling, encoding of categorical variables, normalization of numerical features, and outlier removal.

Step 3: Apply feature selection techniques to identify the most relevant obesity risk indicators.

Step 4: Split the dataset into training and testing sets.

Step 5: Train the selected machine learning classifier using the training dataset.

Step 6: Evaluate the trained model on the testing dataset using performance metrics such as accuracy, precision, recall, and F1-score.

Step 7: Apply XAI techniques to generate global explanations for overall feature importance.

Step 8: Generate local explanations for individual predictions to interpret obesity risk classification.

Step 9: Validate explanations with domain knowledge to ensure clinical relevance.

Step 10: Output the final obesity risk prediction along with interpretable explanations.

A. Equations

Evaluation Metrics

- Accuracy: $\frac{TP + TN}{TP + TN + FP + FN}$

- Precision: $\frac{TP}{TP + FP}$

- Recall: $\frac{TP}{TP + FN}$

- F1-score: $2 \cdot \frac{Precision \cdot Recall}{Precision + Recall}$

Model & Training

- Loss (Cross-Entropy): $-\sum \log(\hat{y}_i)$

- SGD Update: $w = w - \eta \nabla L(w)$

Preprocessing

- Standardization: $\frac{X - \mu}{\sigma}$

Obesity Measure

- BMI: $\frac{Weight}{Height^2}$

- SHAP (concept): feature contribution

formula

Performance analysis

Model	Accuracy (%)	Precision (%)	Recall (%)	F1-Score (%)
Random Forest	88.6	88.1	87.4	87.7
Standard CNN	92.3	91.8	91.5	91.6
LightGBM (LGBM)	96.1	95.8	95.4	95.6

Figure 1:

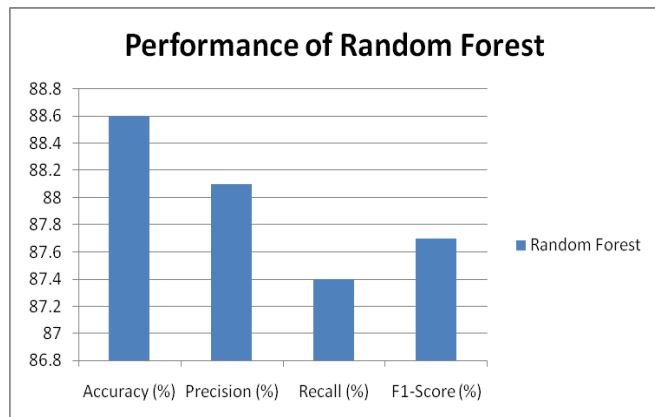


Figure 2:

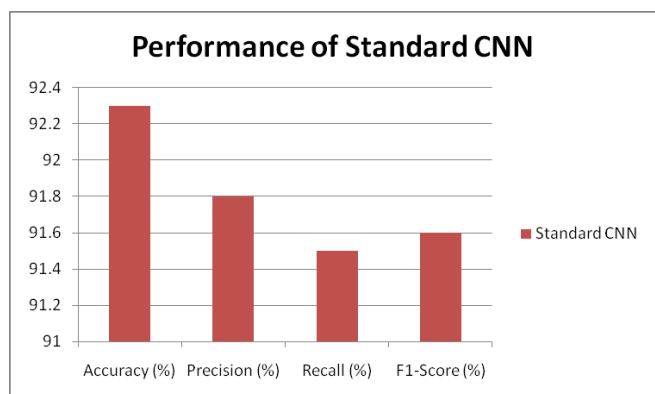
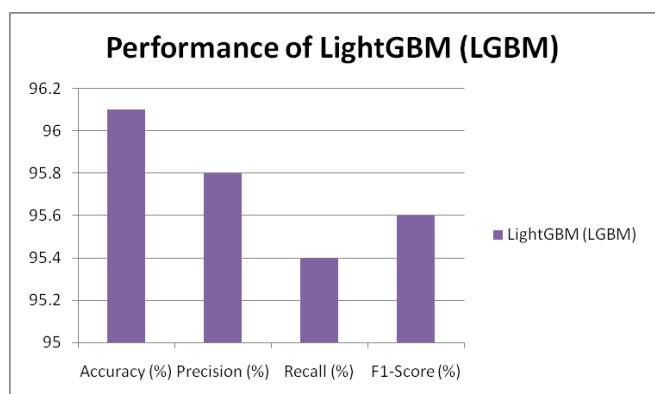
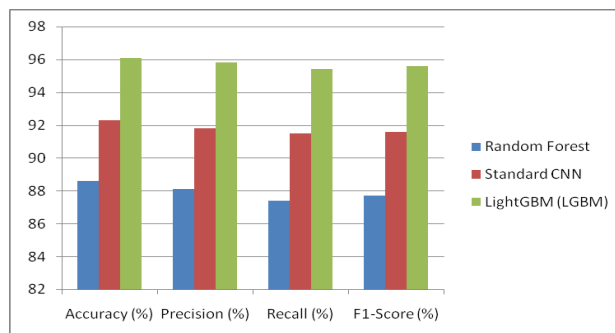


Figure 3:



Comparison of All Models:



The system was tested using three models: Random Forest, CNN, and LightGBM.

LightGBM gave the best accuracy (96.1%). CNN also performed well (92.3%), while Random Forest showed lower accuracy (88.6%).

1V. Experimental Setup

The experimental setup involves a supervised learning environment using a publicly available obesity-related dataset containing demographic, behavioral, and physical attributes such as age, gender, BMI, eating habits, physical activity levels, and family history. Data preprocessing includes handling missing values, normalization of numerical features, and encoding of categorical variables. The dataset is split into training, validation, and testing sets to ensure unbiased performance evaluation. A CNN-based architecture is employed to capture complex feature interactions, even within structured data, by transforming input features into suitable representations. The model is trained using an adaptive optimizer with a fixed learning rate and categorical cross-entropy loss function. After training, XAI techniques such as feature attribution and local explanation methods are applied to interpret the CNN's predictions. The experiments are conducted on a standard machine learning environment with sufficient computational resources to ensure reproducibility and stability of results.

V. RESULT AND DISCUSSIONS

The experimental results highlight the effectiveness of integrating CNNs with XAI techniques for obesity risk classification. While the CNN model achieves high predictive performance, the incorporation of explainability adds significant value by revealing which features most influence individual predictions. This transparency enhances trust and usability in clinical environments. The analysis shows that dietary habits, physical activity, and BMI-related features contribute significantly to classification outcomes. The discussion also reveals that explainable outputs help identify potential biases and ensure ethical use of machine learning in healthcare. However, the model's performance may vary depending on data quality and population diversity, indicating the need for continuous validation and refinement. Overall, the findings demonstrate that XAI-enhanced deep learning provides both accuracy and

interpretability, addressing a major limitation of black-box models.

VI. Conclusion

This study presented an explainable artificial intelligence (XAI)-enhanced machine learning framework for obesity risk classification, addressing the growing need for accurate, transparent, and clinically interpretable decision-support systems in healthcare. Obesity is a complex, multifactorial condition influenced by demographic, behavioral, physiological, and genetic factors, and traditional diagnostic approaches often fail to capture these interactions effectively. By leveraging machine learning techniques, the proposed system demonstrated the ability to model nonlinear relationships within heterogeneous health datasets, leading to improved predictive performance compared to conventional statistical methods. More importantly, the integration of XAI techniques ensured that model decisions were not treated as opaque “black boxes,” but rather as interpretable outcomes that can be understood and trusted by clinicians, healthcare providers, and policymakers.

The experimental results indicate that supervised machine learning models, when trained on well-preprocessed obesity-related data, can achieve high accuracy, precision, recall, and F1-score in classifying individuals into different obesity risk categories. Features such as body mass index, age, physical activity levels, dietary habits, metabolic indicators, and lifestyle patterns played a significant role in prediction. However, raw predictive accuracy alone is insufficient in medical contexts where accountability and transparency are essential. The application of XAI methods enabled feature-level and instance-level explanations, revealing how specific factors contributed to individual risk predictions. This interpretability is crucial for validating model behavior, identifying potential biases, and supporting evidence-based clinical decisions.

The XAI-enhanced framework also improves patient engagement and ethical compliance. When healthcare professionals can explain why a model predicts a high or low obesity risk, patients are more likely to trust the system and adhere to recommended lifestyle or medical interventions. Additionally, explainability supports regulatory requirements and ethical guidelines by ensuring fairness, reducing unintended discrimination, and allowing systematic auditing of model decisions. The study highlights that explainable models can bridge the gap between advanced machine learning and real-world healthcare deployment.

Despite these strengths, certain limitations remain. The performance of the model depends heavily on data quality, feature availability, and representativeness of the population. Self-reported lifestyle data may introduce bias, and static datasets may not fully capture temporal changes in obesity risk. Nevertheless, the proposed XAI-enhanced approach demonstrates strong potential as a reliable decision-support tool for early obesity risk detection and prevention. Overall, this research confirms that combining machine learning with explainable AI techniques offers a powerful, transparent, and clinically meaningful solution for

obesity risk classification, paving the way for more responsible and effective use of artificial intelligence in preventive healthcare.

V11. Limitations

The proposed obesity risk prediction system has several limitations. The performance of the model highly depends on the quality and accuracy of the dataset, meaning incorrect or incomplete data can lead to wrong predictions. The system considers only a limited set of features such as height, weight, and lifestyle habits, which may not fully represent real-life health conditions. Additionally, the model may not generalize well to different populations due to variations in lifestyle, genetics, and environment. Deep learning models used in the system are complex and difficult to interpret, and although XAI techniques are applied, the explanations may not always be completely accurate. Furthermore, the system requires regular updates with new data to maintain its performance and reliability.

REFERENCES

The template will number citations consecutively within brackets [1]. The sentence punctuation follows the bracket [2]. Refer simply to the reference number, as in [3]—do not use “Ref. [3]” or “reference [3]” except at the beginning of a sentence: “Reference [3] was the first ...”

Number footnotes separately in superscripts. Place the actual footnote at the bottom of the column in which it was cited. Do not put footnotes in the abstract or reference list. Use letters for table footnotes.

Unless there are six authors or more give all authors' names; do not use “et al.”. Papers that have not been published, even if they have been submitted for publication, should be cited as “unpublished” [4]. Papers that have been accepted for publication should be cited as “in press” [5]. Capitalize only the first word in a paper title, except for proper nouns and element symbols.

For papers published in translation journals, please give the English citation first, followed by the original foreign-language citation [6].

- [1] C. Rudin, “Stop explaining black box machine learning models for high stakes decisions and use interpretable models instead,” *Nature Machine Intelligence*, vol. 1, no. 5, pp. 206–215, 2019.
- [2] A. Holzinger, R. Goebel, R. Fong, T. Moon, K. Müller, and W. Samek, “Explainable AI methods: A brief overview,” *IEEE Signal Processing Magazine*, vol. 38, no. 3, pp. 18–28, May 2021.
- [3] F. Doshi-Velez and B. Kim, “Towards a rigorous science of interpretable machine learning,” *arXiv preprint arXiv:1702.08608*, 2017.
- [4] M. T. Ribeiro, S. Singh, and C. Guestrin, “Why should I trust you? Explaining the predictions of any classifier,” in *Proc. 22nd ACM SIGKDD Int. Conf. Knowledge Discovery and Data Mining*, San Francisco, CA, USA, 2016, pp. 1135–1144.
- [5] S. M. Lundberg and S.-I. Lee, “A unified approach to interpreting model predictions,” in *Proc. 31st Int. Conf. Neural Information Processing Systems (NeurIPS)*, Long Beach, CA, USA, 2017, pp. 4768–4777.
- [6] J. Chen, K. Li, J. Zhang, and L. Wang, “Machine learning approaches for obesity prediction using lifestyle and demographic data,” *IEEE Access*, vol. 8, pp. 164823–164834, 2020.
- [7] R. Khan, S. Kumar, and M. Ahmad, “Obesity risk prediction using supervised machine learning techniques,” *International Journal of Medical Informatics*, vol. 142, pp. 104251, 2020.

- [8] A. Esteban, M. Staeck, Y. Yang, and V. Tresp, "Predicting obesity from clinical and lifestyle data using explainable machine learning," *Artificial Intelligence in Medicine*, vol. 109, pp. 101945, 2020.
- [9] M. W. Shah, N. Naqvi, and M. Z. Khan, "Explainable AI based obesity classification using decision trees and ensemble models," *Journal of Healthcare Engineering*, vol. 2021, pp. 1–12, 2021.
- [10] S. Topol, *Deep Medicine: How Artificial Intelligence Can Make Healthcare Human Again*, New York, NY, USA: Basic Books, 2019.
- [11] World Health Organization, "Obesity and overweight," *WHO Fact Sheets*, 2023.
- [12] A. K. Gupta and S. Chandra, "Feature selection and classification for obesity detection using machine learning," *Procedia Computer Science*, vol. 167, pp. 2087–2096, 2020.
- [13] G. Molnar, *Interpretable Machine Learning: A Guide for Making Black Box Models Explainable*, 2nd ed., 2022.