International Journal of

DATA SCIENCE AND IOT MANAGEMENT SYSTEM

ISSN: 3068-272X www.ijdim.com Original Research Paper

An Interpretable Voting Ensemble Model for Obesity Classification Using Lifestyle and Physical Health Attributes

KORCHA ASWINI
ASSISTANT PROFESSOR
CSE DEPARTMENT
PVKK INSTITUTE OF TECHNOLOGY

Abstract: Obesity has become a significant global health concern that requires reliable and interpretable classification approaches to support early intervention and reduce long-term health complications. This study presents an explainable ensemble learning framework for multi-class obesity classification using behavioral, lifestyle, and anthropometric attributes obtained from the UCI Machine Learning Repository obesity dataset. The implemented pipeline performs data cleaning, categorical label encoding, feature scaling, Recursive Feature Elimination with Cross Validation (RFECV) using Logistic Regression for feature optimization, and Synthetic Minority Over-sampling Technique for Nominal and Continuous features (SMOTENC) to address class imbalance. A comprehensive comparative evaluation is conducted across multiple machine learning algorithms including Adaboost, Perceptron, GaussianNB, SGD, SVM, KNN, MLP, Decision Tree, ExtraTrees, Bagging Classifier, Random Forest, Gradient Boosting, LogisticRegressionCV, XGBoost, and LightGBM. Based on comparative performance analysis, an Extended Voting ensemble combining Gradient Boosting, XGBoost, LightGBM, and CatBoost is selected as the final prediction framework. To improve transparency and interpretability, Explainable Artificial Intelligence (XAI) techniques including LIME and SHAP are employed to provide local and global explanations of model behavior. The final trained model is deployed through a Flask-based web application to support real-time obesity level prediction. Experimental results demonstrate superior classification capability, achieving 99.3% accuracy with equally strong precision, recall, and F1-score performance and a ROC–AUC score of 1.000, indicating reliable and interpretable obesity classification for practical decision-support applications.

Index Terms: *Obesity Classification, Ensemble Learning, Explainable Artificial Intelligence (XAI), Extended Voting Ensemble, Feature Selection, Flask Deployment, Machine Learning.*

1. INTRODUCTION

Obesity has become one of the most critical public health challenges worldwide due to its increasing prevalence and long-term health consequences. Characterized by excessive accumulation of body fat, obesity contributes significantly to morbidity, mortality, and escalating healthcare expenditure across populations [1]. The growing incidence of obesity is associated with multiple behavioral and environmental factors including sedentary lifestyles, unhealthy dietary habits, reduced physical activity, and changing living patterns [2]. These factors collectively increase the risk of chronic conditions such as cardiovascular disease, diabetes, metabolic disorders, stroke, and several forms of cancer [3]. Consequently, there is a growing demand for reliable analytical approaches that support early obesity identification and assist preventive healthcare decision-making.

Traditional obesity assessment methods primarily depend on anthropometric indicators such as Body Mass Index (BMI) because of their simplicity and ease of application. However, BMI alone is often insufficient to represent the multidimensional characteristics of obesity and may fail to distinguish body fat composition from other physiological conditions [4]. Furthermore, conventional assessment approaches typically overlook demographic, behavioral, and lifestyle-related factors that contribute to obesity development, reducing their effectiveness for comprehensive obesity classification.

Recent advances in data-driven healthcare have demonstrated that machine learning techniques can extract complex relationships from heterogeneous health-related data and improve predictive performance beyond conventional statistical assessment methods [5]. Nevertheless, existing obesity prediction studies frequently emphasize classification accuracy while providing limited interpretability and practical deployment capabilities. In healthcare-oriented applications, predictive performance alone is insufficient unless model decisions can be explained and utilized effectively within accessible decision-support environments. Therefore, there remains a need for an integrated framework that combines predictive strength, interpretability, and practical usability for obesity classification.

This study aims to develop an explainable ensemble learning framework for multi-class obesity classification using behavioral, lifestyle, and anthropometric attributes obtained from the UCI Machine Learning Repository obesity dataset [6], [7]. The proposed framework integrates data preprocessing, feature optimization, class balancing, ensemble-based classification, explainability analysis, and web-based deployment to improve obesity level prediction while maintaining transparency and usability.

The primary contributions of this work are summarized as follows:

- Development of an end-to-end obesity classification pipeline integrating preprocessing, feature selection, and ensemble learning.
- Comparative evaluation of multiple machine learning and ensemble classifiers to identify an effective prediction strategy.
- Design of an Extended Voting ensemble framework for improved obesity classification performance.
- Integration of LIME and SHAP to provide interpretable local and global prediction explanations.
- Deployment of the final prediction model through a Flask-based web application for real-time obesity classification.

2. RELATED WORK

The growing prevalence of obesity has motivated extensive research into predictive and classification frameworks capable of supporting early intervention and improving healthcare decision-making. Recent studies have increasingly adopted machine learning approaches to model obesity-related behaviors and physiological indicators. Thamrin et al. developed obesity prediction models using machine learning techniques on large-scale health survey data and demonstrated that integrating demographic and lifestyle attributes improves predictive capability beyond conventional assessment methods [8]. Similarly, Zheng and Ruggiero explored obesity prediction among high school students and showed that behavioral and physical characteristics can effectively support obesity identification through data-driven classification methods [9]. These studies highlight the growing applicability of machine learning for obesity analysis but primarily focus on prediction performance rather than interpretability and deployment.

Beyond predictive analytics, researchers have explored integrated healthcare environments for obesity management. Machorro-Cano et al. introduced the PISIoT framework, which combines machine learning with Internet of Things (IoT) technologies to support continuous obesity monitoring and personalized intervention strategies [10]. Such systems demonstrate the potential of intelligent health platforms but often depend on complex infrastructures that limit accessibility and reproducibility in lightweight predictive applications.

To improve predictive robustness and address limitations of individual classifiers, ensemble learning techniques have become increasingly prominent in healthcare analytics. Diayasa et al. proposed a stacking-based framework for obesity level prediction by combining multiple machine learning classifiers and demonstrated that ensemble integration can enhance classification performance through complementary learning behavior [11]. Similarly, Solomon et al. developed a hybrid majority voting model for obesity prediction and reported improved predictive reliability by aggregating outputs from multiple classifiers [12]. These findings emphasize that ensemble approaches can outperform standalone algorithms in obesity-related classification tasks.

The effectiveness of ensemble learning has also been demonstrated in broader healthcare applications. Kibria et al. implemented a soft voting ensemble integrated with explainable AI for diabetes prediction and showed that combining ensemble methods with interpretation mechanisms improves both predictive quality and usability [13]. These studies collectively indicate that ensemble strategies provide an effective pathway for improving classification performance while maintaining model stability across complex healthcare datasets.

Although predictive accuracy remains an important objective, modern healthcare systems increasingly require transparency and interpretability to support trustworthy decision-making. Explainable Artificial Intelligence (XAI) techniques have emerged as essential tools for understanding model behavior and improving confidence in machine learning outcomes. Raihan et al. demonstrated the application of SHAP to interpret predictions generated by an XGBoost-based clinical prediction model and highlighted its capability to quantify feature contributions globally across model outputs [14]. Jahan et al. further showed that integrating explainability into healthcare prediction systems improves interpretability and supports practical adoption in medical applications [15].

Among widely adopted explanation approaches, LIME provides local interpretability by approximating prediction behavior around individual instances and enabling understandable explanations for complex models [16]. In contrast, SHAP offers a unified framework for global and local feature attribution using cooperative game theory principles, enabling consistent interpretation of prediction outcomes across multiple observations [17].

Despite the progress achieved in obesity prediction, existing studies typically focus on isolated objectives such as predictive accuracy, ensemble optimization, explainability, or intelligent healthcare environments. Limited attention has been given to developing a unified framework that simultaneously integrates feature optimization, imbalance handling, ensemble-based classification, explainability, and practical deployment for obesity level prediction. To address this limitation, the present study introduces an explainable ensemble learning framework that combines comprehensive preprocessing, Extended Voting-based

International Journal of
DATA SCIENCE AND IOT MANAGEMENT SYSTEM
ISSN: 3068-272X www.ijdim.com Original Research Paper

classification, LIME and SHAP interpretation, and Flask-based deployment to provide accurate, interpretable, and deployable obesity classification.

3. PROPOSED METHODOLOGY

A) Proposed Overview

The proposed framework introduces an explainable ensemble learning pipeline for multi-class obesity classification using behavioral, lifestyle, and anthropometric attributes. The methodology combines structured preprocessing, feature optimization, class balancing, ensemble-based prediction, explainability analysis, and deployment within a unified workflow. Initially, the obesity dataset undergoes cleaning, encoding, and scaling to improve data quality and learning consistency. Feature relevance is optimized using Recursive Feature Elimination with Cross Validation (RFECV), followed by imbalance correction using synthetic oversampling techniques. Multiple machine learning and ensemble models are evaluated comparatively to identify the most effective prediction strategy. Based on overall performance, an Extended Voting ensemble is selected as the final prediction framework. Explainability is incorporated through LIME and SHAP analysis, while deployment is achieved using a Flask-based web application to enable real-time obesity classification.

B) System Architecture

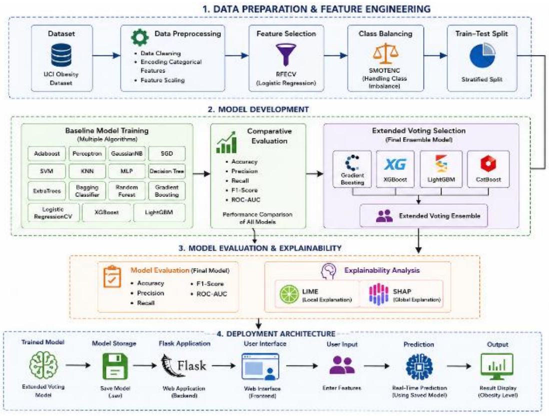

Fig.1 Proposed Architecture

The proposed system architecture, illustrated in Fig. 1, follows a sequential learning and deployment workflow beginning with dataset acquisition and preprocessing operations. The processed data undergoes feature optimization and imbalance correction before being supplied to multiple classification algorithms for comparative evaluation. Ensemble learning techniques are subsequently applied to improve prediction robustness and generalization capability. Explainability mechanisms are performed after model training to interpret prediction behavior and identify influential attributes. The final trained ensemble model is exported and integrated into a Flask-based application that enables real-time user interaction and obesity level prediction through an accessible web interface.

C) Dataset

The experimental analysis utilizes the obesity estimation dataset obtained from the UCI Machine Learning Repository. The dataset contains 2,111 observations collected from individuals aged between 14 and 61 years across Colombia, Peru, and Mexico and includes demographic, dietary, behavioral, and physical attributes. The input space contains variables describing eating habits, physical

International Journal of

DATA SCIENCE AND IOT MANAGEMENT SYSTEM

ISSN: 3068-272X www.ijdim.com Original Research Paper

activity, hydration, transportation behavior, family history, and anthropometric indicators such as age, height, and weight. The target variable, NObeyesdad, represents seven obesity categories: Insufficient Weight, Normal Weight, Overweight Level I, Overweight Level II, Obesity Type I, Obesity Type II, and Obesity Type III. This diverse feature composition supports comprehensive multi-class obesity classification. The structure and source of the dataset are illustrated in Fig. 2.

	Gender	Age	Height	Weight	family_history_with_overweight	FAVC	FCVC	NCP	CAEC	SMOKE	CH2O	SCC	FAF	TUE	CALC	MTRANS	NObeyesdad
0	Female	21.0	1.62	64.0	yes	no	2.0	3.0	Sometimes	no	2.0	no	0.0	1.0	no	Public_Transportation	Normal_Weight
1	Female	21.0	1.52	56.0	yes	no	3.0	3.0	Sometimes	yes	3.0	yes	3.0	0.0	Sometimes	Public_Transportation	Normal_Weight
2	Male	23.0	1.80	77.0	yes	no	2.0	3.0	Sometimes	no	2.0	no	2.0	1.0	Frequently	Public_Transportation	Normal_Weight
3	Male	27.0	1.80	87.0	no	no	3.0	3.0	Sometimes	no	2.0	no	2.0	0.0	Frequently	Walking	Overweight_Level_I
4	Male	22.0	1.78	89.8	no	no	2.0	1.0	Sometimes	no	2.0	no	0.0	0.0	Sometimes	Public_Transportation	Overweight_Level_II

Fig.2 UCI Machine Learning Repository

D) Data Preprocessing

Data preprocessing was performed to improve data quality, ensure learning consistency, and prepare the dataset for predictive modeling. The preprocessing pipeline includes cleaning, categorical transformation, and feature scaling procedures before model training.

Cleaning: The dataset was initially examined for missing values, duplicate records, and structural inconsistencies. No significant missing observations were identified during inspection; however, duplicate instances were detected and removed to prevent repeated learning patterns and improve generalization capability. A total of 24 duplicate samples were eliminated and dataset indexing was reconstructed to maintain consistency across subsequent analytical stages.

Label Encoding: Since several input variables were represented in categorical form, label encoding was applied to transform non-numeric values into numerical representations compatible with machine learning algorithms. Each category was mapped into a unique integer value while preserving dataset structure and maintaining efficient computational processing. This transformation enabled classifiers to analyze behavioral and lifestyle attributes alongside continuous physiological variables.

Feature Scaling: Feature scaling was applied using StandardScaler to normalize numerical feature distributions and reduce variation across measurement ranges. Standardization improves convergence behavior during model learning and prevents attributes with larger numerical magnitudes from dominating prediction outcomes. The scaling process contributed to stable optimization and improved consistency across evaluated classifiers.

E) Feature Selection

Feature optimization was performed using Recursive Feature Elimination with Cross Validation (RFECV) to identify the most informative attributes and eliminate redundant variables. Logistic Regression was selected as the evaluation estimator and Stratified 5-Fold Cross Validation was applied during the recursive selection process to preserve class representation. RFECV iteratively removed low-contribution features and evaluated model performance until the optimal feature subset was identified. As a result, fourteen features were retained for subsequent learning stages while two attributes, namely SMOKE and TUE, were excluded due to limited contribution toward classification performance. The reduced feature space improved computational efficiency, interpretability, and predictive stability.

F) Class Balancing

Class imbalance handling was performed to reduce prediction bias and improve minority class representation across obesity categories.

SMOTENC: The primary balancing strategy utilized Synthetic Minority Over-sampling Technique for Nominal and Continuous features (SMOTENC). This approach generates synthetic observations while preserving relationships between categorical and continuous variables, making it suitable for obesity-related mixed data structures. SMOTENC produced balanced class distributions and improved learning stability across multiple classification algorithms.

SVMSMOTE: As an extended balancing experiment, Support Vector Machine Synthetic Minority Over-sampling Technique (SVMSMOTE) was evaluated to further strengthen class boundary learning. This method generates synthetic observations using support vector principles and improves representation near difficult classification regions. The extension enabled additional robustness analysis and supported improved ensemble performance under balanced learning conditions.

G) Ensemble Model Development

Baseline Models: To establish performance baselines, multiple machine learning algorithms were implemented and evaluated independently. The evaluated models include Adaboost, Perceptron, GaussianNB, SGD, SVM, KNN, MLP, Decision Tree, ExtraTrees, Bagging Classifier, Random Forest, Gradient Boosting, LogisticRegressionCV, XGBoost, and LightGBM. Each classifier was assessed using identical training conditions and standard evaluation metrics.

Comparative Ensembles: Following baseline evaluation, ensemble strategies were explored to improve classification reliability and reduce dependence on individual learner behavior. Multiple ensemble configurations were examined, including boosting-based approaches and stacking architectures. Stacking was investigated as a comparative benchmark to evaluate the benefit of combining complementary learners and to establish performance references for advanced ensemble configurations.

Final Extended Voting Model: Based on comparative analysis, an Extended Voting ensemble was selected as the final prediction framework. The ensemble combines Gradient Boosting, XGBoost, LightGBM, and CatBoost using a soft voting mechanism that aggregates prediction probabilities across classifiers. This strategy leverages complementary strengths among boosting algorithms and improves robustness, predictive consistency, and overall classification accuracy. The trained ensemble demonstrated superior performance and was selected as the deployed model for real-time obesity prediction.

H) Explainability Layer

To improve transparency and interpretability, Explainable Artificial Intelligence techniques were incorporated after model training. Local Interpretable Model-Agnostic Explanations (LIME) was applied to explain individual prediction outcomes by approximating local decision behavior and identifying influential attributes for specific classifications. SHapley Additive exPlanations (SHAP) was employed to provide global interpretation by quantifying feature contributions across the entire prediction space. Together, these techniques enabled interpretation of both local and global model behavior and supported understanding of the influence of demographic, behavioral, and physiological variables on obesity classification. Explainability was conducted as an offline analytical stage and was not integrated into runtime prediction deployment.

International Journal of

DATA SCIENCE AND IOT MANAGEMENT SYSTEM

ISSN: 3068-272X www.ijdim.com Original Research Paper

I) Deployment Architecture

The final trained ensemble model was exported and integrated into a Flask-based web application to support real-time obesity classification. The deployment workflow begins with user interaction through a browser interface, where lifestyle and physical attributes are entered through structured forms. The input data is processed and forwarded to the saved prediction model loaded using Joblib. The inference module generates obesity category predictions and returns the output to the user through a dedicated result interface. User authentication and interface management are supported through SQLite and Flask routing components. The deployment architecture enables accessible and lightweight obesity prediction without requiring cloud infrastructure.

Following methodology implementation and deployment preparation, experimental evaluation was conducted to analyze predictive performance and interpret model behavior.

4. RESULTS AND DISCUSSION

A) Experimental Setup and Comparative Results

The experimental evaluation was conducted to assess the effectiveness of the proposed obesity classification framework across multiple machine learning and ensemble learning algorithms. The implementation environment utilized Python as the primary development language with Jupyter Notebook for experimentation and Flask for deployment. Major libraries included NumPy, Pandas, Scikit-learn, XGBoost, LightGBM, CatBoost, SHAP, LIME, Matplotlib, and Joblib. Model training incorporated preprocessing operations including StandardScaler, RFECV-based feature optimization, and synthetic balancing techniques. Comparative evaluation was performed under identical learning conditions to ensure fairness across all implemented models.

Hyperparameter configurations were selected using iterative experimental tuning and maintained consistently during comparative evaluation. Performance evaluation was conducted using Accuracy, Precision, Recall, F1-Score, and ROC–AUC metrics to provide comprehensive assessment of predictive capability and classification reliability.

Table 1. Comparative Performance Evaluation of Machine Learning and Ensemble Models

ML Model	Accuracy	Precision	Recall	F1-Score	ROC-AUC
Adaboost	0.358	0.633	0.358	0.304	0.779
Perceptron	0.618	0.600	0.618	0.587	0.903
Gaussian NB	0.606	0.643	0.606	0.584	0.915
SGD	0.404	0.581	0.404	0.365	0.872
SVM	0.547	0.535	0.547	0.531	0.915
KNN	0.902	0.906	0.902	0.898	0.977
MLP	0.785	0.788	0.785	0.780	0.964
Decision Tree	0.931	0.931	0.931	0.930	0.960
Extra Trees	0.943	0.944	0.943	0.943	0.995
Bagging Classifier	0.961	0.962	0.961	0.961	0.998
Random Forest	0.951	0.952	0.951	0.951	0.997

International Journal of
DATA SCIENCE AND IOT MANAGEMENT SYSTEM
ISSN: 3068-272X www.ijdim.com Original Research Paper

Gradient Boost	0.980	0.980	0.980	0.980	0.998
Logistic RegressionCV	0.732	0.726	0.732	0.725	0.927
XGBoost	0.980	0.980	0.980	0.980	0.999
Light GBM	0.963	0.964	0.963	0.963	0.996
Stacking Ensemble	0.984	0.984	0.984	0.984	0.999
Extended Voting	0.993	0.993	0.993	0.993	1.000

The results demonstrate that ensemble-based approaches consistently outperform individual classifiers across all evaluation metrics. Traditional linear methods exhibit lower classification capability, while tree-based and boosting models show substantial improvement. Among all evaluated approaches, the Extended Voting framework achieves the highest overall performance and is selected as the final deployed prediction model.

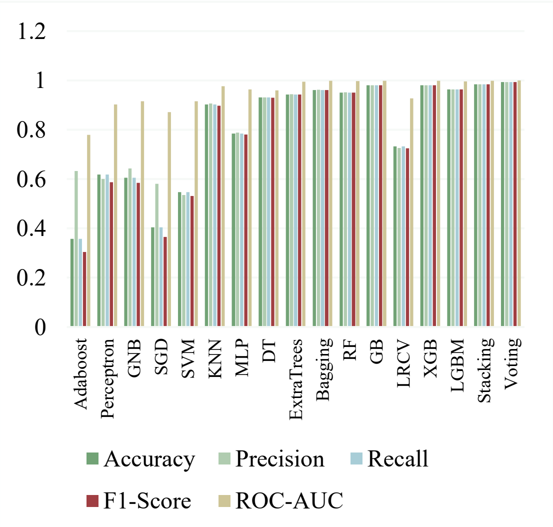

Fig. 3. Comparative Performance Analysis Across Machine Learning and Ensemble Models

As illustrated in Fig. 3, ensemble integration significantly improves classification stability and predictive reliability. The comparison graph indicates clear performance separation between baseline classifiers and advanced ensemble approaches.

B) Ensemble Performance Analysis

The comparative analysis indicates that the Extended Voting ensemble delivers superior predictive performance due to its ability to combine complementary learning characteristics from multiple boosting algorithms. Unlike single classifiers that depend on isolated decision boundaries, the voting mechanism aggregates prediction probabilities generated by Gradient Boosting, XGBoost, LightGBM, and CatBoost to produce a more robust final prediction.

Gradient Boosting contributes iterative error correction, XGBoost improves optimization and regularization, LightGBM enhances computational efficiency through histogram-based learning, and CatBoost strengthens handling of categorical relationships. The integration of these learners reduces model variance while improving generalization across obesity classes.

Although the Stacking Ensemble achieved competitive results with 98.4% accuracy, the Extended Voting framework demonstrated higher consistency across Precision, Recall, F1-Score, and ROC–AUC. Consequently, the Extended Voting model was selected as the final deployed framework and exported for real-time prediction.

C) Explainability Analysis

To improve transparency and support interpretation of prediction outcomes, Explainable Artificial Intelligence techniques were applied after model

International Journal of

DATA SCIENCE AND IOT MANAGEMENT SYSTEM

ISSN: 3068-272X www.ijdim.com Original Research Paper

training. The analysis utilized both local and global explanation approaches to understand how individual features influence obesity classification decisions.

LIME Analysis: The class-wise local interpretation generated using LIME is illustrated in Fig. 4. LIME identifies feature contributions surrounding a specific prediction and provides interpretable insight into model decision behavior.

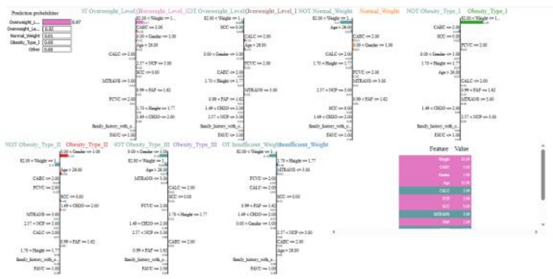

Fig. 4. Class-Wise LIME Explanations for Local Interpretation of Obesity Classification

The LIME analysis indicates that obesity predictions are primarily influenced by anthropometric and behavioral variables. Weight consistently emerged as the dominant predictor, followed by Age, Height, CAEC, CALC, and physical activity indicators across multiple obesity categories. The explanations demonstrate consistent decision behavior across obesity categories and improve understanding of class-specific prediction reasoning.

SHAP Analysis: Global feature contribution analysis obtained through SHAP is presented in Fig. 5. SHAP quantifies the contribution of each feature and enables ranking of variables according to their overall impact on classification outcomes.

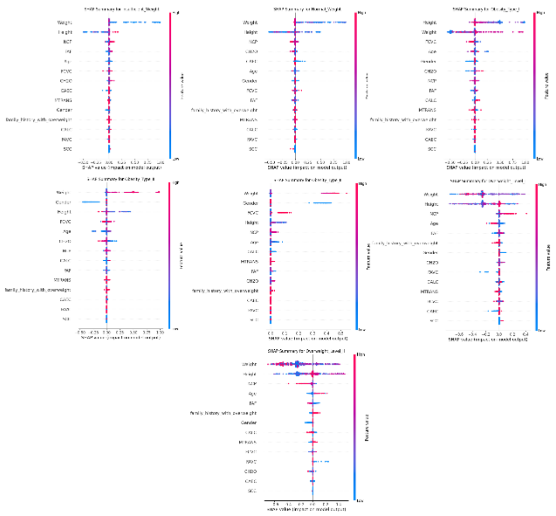

Fig. 5. SHAP-Based Global Interpretation Across Obesity Categories

The SHAP interpretation reveals that Weight, Gender, Age, CAEC, Height, and physical activity attributes contribute strongly to model predictions. This observation supports the ability of the ensemble framework to capture interactions between physiological and behavioral factors during obesity classification.

D) Deployment Demonstration

To demonstrate practical applicability, the final Extended Voting model was integrated into a Flask-based web application that supports real-time obesity classification. The deployment architecture enables users to interact through a browser interface, submit obesity-related attributes, and receive prediction outcomes through a lightweight prediction workflow.

The deployment interface used for user interaction is presented in Fig. 6.

International Journal of
DATA SCIENCE AND IOT MANAGEMENT SYSTEM
ISSN: 3068-272X www.ijdim.com Original Research Paper

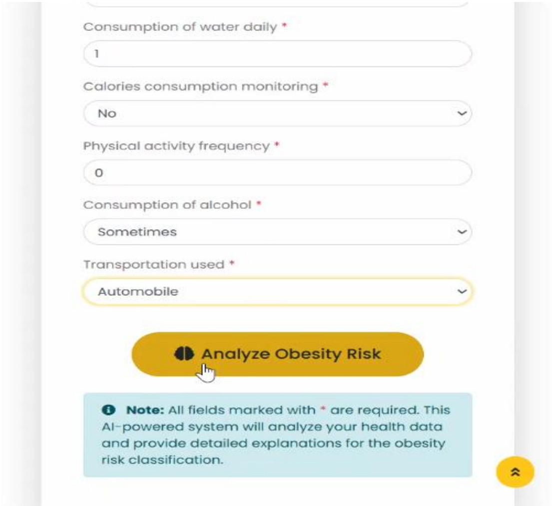

Fig. 6. User Interface for Input Collection and Prediction Request

The deployed application includes authentication functionality and structured user input forms for collecting demographic, behavioral, and physical attributes. Submitted values are processed and forwarded to the saved ensemble model for inference.

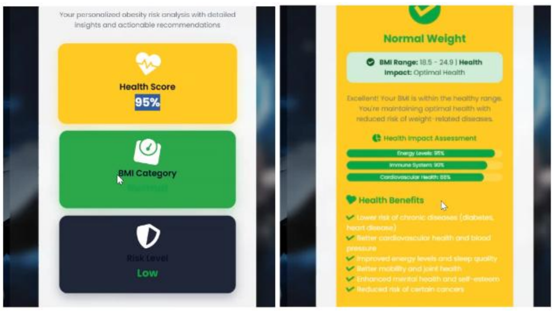

Fig. 7. Prediction Output Generated by the Deployed Extended Voting Model

An example prediction outcome generated by the deployed system is shown in Fig. 7.

The prediction interface returns obesity classification results immediately after inference and demonstrates the practical feasibility of translating ensemble learning and explainability-driven obesity analysis into an accessible user-oriented application environment.

5. CONCLUSION AND FUTURE WORK

This study presented an explainable ensemble learning framework for multi-class obesity classification using behavioral, lifestyle, and anthropometric attributes obtained from the UCI Machine Learning Repository obesity dataset. The implemented pipeline combined structured preprocessing, including duplicate removal, categorical transformation, feature scaling, RFECV-based feature optimization, and imbalance handling using SMOTENC and extended SVMSMOTE analysis to improve learning quality and classification robustness.

A comprehensive comparative evaluation was conducted across traditional machine learning and ensemble classifiers to identify the most effective prediction strategy for obesity level classification. Experimental findings demonstrated that ensemble approaches consistently achieved superior predictive capability compared with standalone classifiers. Among all evaluated methods, the Extended Voting ensemble integrating Gradient Boosting, XGBoost, LightGBM, and CatBoost delivered the strongest overall performance and was selected as the final deployed framework. The model achieved an accuracy of 99.3% with equally strong precision, recall, and F1-score performance, together with a ROC–AUC value approaching 1.0.

Beyond predictive performance, this work incorporated Explainable Artificial Intelligence techniques to improve transparency and interpretability of model behavior. LIME enabled local interpretation of individual prediction outcomes, while SHAP provided global understanding of feature influence across obesity categories. The interpretation results indicated that anthropometric and lifestyle-related attributes contributed substantially to obesity classification decisions. To demonstrate practical applicability, the final trained model was deployed through a Flask-based web application that supports real-time obesity level prediction through an accessible user interface.

Although the proposed framework demonstrated strong predictive performance and practical deployment applicability, several limitations remain. The current implementation relies on a static dataset and offline explainability analysis and does not incorporate continuous health monitoring or external clinical validation.

Future work may focus on integrating real-time health data sources, extending analysis through wearable and mobile health platforms, exploring additional demographic and environmental variables, and improving explainability through advanced interpretation strategies. Further enhancement of deployment scalability and longitudinal obesity monitoring may support broader adoption of intelligent obesity assessment systems in practical healthcare environments.

REFERENCES

[1] Palechor, F. M., & de la Hoz Manotas, A. (2019). Dataset for estimation of obesity levels based on eating habits and physical condition in individuals from Colombia, Peru and Mexico. Data in Brief, 25, 104344.

[2] Chawla, N. V., Bowyer, K. W., Hall, L. O., & Kegelmeyer, W. P. (2002). SMOTE: Synthetic minority over-sampling technique. Journal of Artificial Intelligence Research, 16, 321–357.

[3] Breiman, L. (2001). Random forests. Machine Learning, 45(1), 5–32.

[4] Chen, T., & Guestrin, C. (2016). XGBoost: A scalable tree boosting system. In Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (pp. 785–794).

[5] Ke, G., Meng, Q., Finley, T., Wang, T., Chen, W., Ma, W., Ye, Q., & Liu, T. Y. (2017). LightGBM: A highly efficient gradient boosting decision tree. In Advances in Neural Information Processing Systems (Vol. 30).

[6] Prokhorenkova, L., Gusev, G., Vorobev, A., Dorogush, A. V., & Gulin, A. (2018). CatBoost: Unbiased boosting with categorical features. Advances in Neural Information Processing Systems, 31.

[7] Ribeiro, M. T., Singh, S., & Guestrin, C. (2016). “Why should I trust you?” Explaining the predictions of any classifier. In Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (pp. 1135–1144).

[8] Lundberg, S. M., & Lee, S. I. (2017). A unified approach to interpreting model predictions. In Advances in Neural Information Processing Systems (Vol. 30).

[9] Thamrin, S. A., Arsyad, D. S., Kuswanto, H., Lawi, A., & Nasir, S. (2021). Predicting obesity in adults using machine learning techniques: An analysis of Indonesian basic health research 2018. Frontiers in Nutrition, 8, 663145.

[10] Machorro-Cano, I., Alor-Hernández, G., Paredes-Valverde, M. A., Ramos-Deonati, J. L., Sánchez-Cervantes, J. L., & Rodríguez-Mazahua, L. (2019). PISIoT: A machine learning and IoT-based smart health platform for overweight and obesity control. Applied Sciences, 9(15), 3037.

[11] Zheng, Z., & Ruggiero, K. (2017). Using machine learning to predict obesity in high school students. In Proceedings of the IEEE International Conference on Bioinformatics and Biomedicine (pp. 2132–2138).

[12] Diayasa, I. G. S. M., Idhom, M., Fauzi, A., & Damaliana, A. T. (2022). Stacking ensemble methods to predict obesity levels in adults. In Proceedings of the 8th Information Technology International Seminar (pp. 339–344).

[13] Kibria, H. B., Nahiduzzaman, M., Goni, M. O. F., Ahsan, M., & Haider, J. (2022). An ensemble approach for the prediction of diabetes mellitus using a soft voting classifier with an explainable AI. Sensors, 22(19), 7268.

[14] Hastie, T., Tibshirani, R., & Friedman, J. (2009). The elements of statistical learning: Data mining, inference, and prediction (2nd ed.). Springer.

[15] Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., ... & Duchesnay, E. (2011). Scikit-learn: Machine learning in Python. Journal of Machine Learning Research, 12, 2825–2830.