Frontiers in Biomedical Technologies Vol. 13, No. 2 (Spring 2026) 395-406

ORIGINAL ARTICLE

Stacking Ensemble Learning Approach for Non-Alcoholic Fatty Liver Disease Identification: Leveraging Explainable Machine Learning for Enhanced Prediction Models

Dolley Srivastava [1*], Himanshu Pandey [2], Ambuj Kumar Agarwal [3], Richa Sharma [4]

[1] *Amity Institute of Information Technology, Amity University, Lucknow, Uttar Pradesh, India*

[2] *Department of Computer Science and Engineering, Faculty of Engineering and Technology, University of Lucknow, Lucknow, India*

[3] *Department of Computer Science and Engineering, School of Engineering and Technology, Sharda University, Greater Noida, India*

[4] *Department of Computer Science & Engineering, Shri Ramswaroop Memorial College of Engineering & Management, Lucknow, India*

*Corresponding Author: Dolley Srivastava
Email: dolley89@gmail.com

Received: 02 April 2024 / Accepted: 08 April 2025

Abstract

Purpose: The prevalence of Non-Alcoholic Fatty Liver Disease (NAFLD) has significantly increased over the past two decades, becoming a leading cause of liver disease in industrialized nations, particularly among individuals who do not consume alcohol. This study aims to develop an efficient detection method for NAFLD using a stacked ensemble learning approach, which integrates multiple machine learning models to enhance predictive accuracy.

Materials and Methods: The dataset utilized in this research was sourced from an open platform and includes critical attributes such as age, gender, Body Mass Index (BMI), height, time to death or last follow-up, and survival status. We implemented a variety of machine learning algorithms, including XGBoost, CatBoost, Decision Trees, and AdaBoost, within a stacking framework to optimize performance. The proposed methodology involved several steps: data preparation, feature engineering, model training, and evaluation. Additionally, we employed explainable AI techniques to identify the most influential features contributing to NAFLD prediction, thereby enhancing the model's interpretability.

Results: The stacked ensemble model achieved an impressive classification accuracy of 95.9%, outperforming individual models and demonstrating the robustness of the ensemble approach. Confusion Metrics, ROC curves, and Calibration Curves are used to evaluate the proposed approach with state-of-the-art approaches. The suggested stacking methodology demonstrates superior performance in all contexts. When Explainable Machine Learning is applied to the proposed approach, it reveals that NAFLD is more common in middle-aged and elderly individuals, but is also present in younger age groups to some extent. Also, the prevalence of NAFLD is higher in males.

Conclusion: The results underscore the potential of a stacked ensemble approach for clinical applications in NAFLD screening and diagnosis, highlighting its importance in healthcare decision-making. By combining various machine learning techniques, we have developed a reliable and resilient model that improves detection accuracy and offers transparency in its predictions. Combining XGBoost, CatBoost, Decision Tree, and AdaBoost in a stacked ensemble model yielded the best results. The findings also indicate that age and gender are significant predictors of NAFLD, with the model providing valuable insights into the underlying patterns associated with the disease. This research contributes to the growing body of knowledge on machine learning applications in gastroenterology and emphasizes the need for explainable models in clinical settings.

Keywords: Machine Learning; Ensemble Learning; Stacking Models; Explainable Artificial Intelligence; Interpretable Machine Learning; Feature Importance.

DOI: https://doi.org/10.18502/fbt.v13i2.21937

1. Introduction

Fatty liver results from lipid buildup within hepatocytes. Conventional wisdom holds that fatty liver is a harmless, reversible disease that the liver develops as a general reaction to metabolic stress. In the past, heavy alcohol use was thought to be the main cause of fatty liver [1]. With an estimated global incidence, Nonalcoholic Fatty Liver Disease (NAFLD) is a leading cause of liver disease and will emerge as a leading cause of end-stage liver disease in the next decades [2].

1.1. Risk Factors for NAFLD

Many variables cause the development of NAFLD. Insulin resistance, type 2 diabetes, dyslipidemia, metabolic syndrome, socioeconomic status, sedentary lifestyle, age, and gender are among these variables [3].

Age: Despite its presence in younger age groups, NAFLD is more prevalent among middle-aged and elderly people. Multiple studies have shown that the severity of NAFLD increases with age, as does the prevalence of its hepatic and extra-hepatic complications, such as non-alcoholic steatohepatitis, cirrhosis, and hepatocellular carcinoma, as well as cardiovascular disease and extrahepatic neoplasms [4]. One possible explanation is that the risk factors for NAFLD tend to rise with age. Research has shown that lipid buildup in non-adipose organs, such as the liver, rises with age, which might contribute to the higher incidence of NAFLD.

Gender: Many studies on NAFLD have shown it to be more prevalent in males. It has been suggested that female hormones protect against NAFLD. This is supported by the evidence observed in different studies where the prevalence of NAFLD was found to be twice as common in postmenopausal women when compared to premenopausal women. Studies have also shown that women receiving hormone replacement therapy are less likely to develop NAFLD when compared to those who do not [5].

Obesity: Obesity, either as a high Body Mass Index (BMI) or visceral obesity, increases the likelihood of developing NAFLD. Insulin resistance and inflammation of the liver are outcomes of obesity-related changes in adipocyte cytokine and adipokine production. Multiple studies have linked NAFLD to an elevated Body Mass Index (BMI), which is a measure of overweight and obesity [6].

Diet: The impact of diet on the development of NAFLD is significant. People with NAFLD tend to eat a lot of nutrients, according to the research. Additional factors that contribute to the development of NAFLD include total caloric intake, macronutrients, and micronutrients. Patients with NAFLD ate a lot of carbs and fat, according to the research. They ate very few fruits, vegetables, and whole grains. They also failed to meet the necessary daily allowances for antioxidant vitamins, calcium, and vitamin D. A greater body mass index and waist circumference were also indicators of obesity and overweight in individuals with NAFLD [7].

Sedentary Lifestyle: The prevalence of NAFLD is inversely related to the amount of physical activity recorded, also known as cardiorespiratory fitness. People with NAFLD are less likely to participate in total, aerobic, and resistance exercises during their free time.

Socio-economic factors: Cardiorespiratory fitness, a measure of how active a person claims to be, has been shown to correlate inversely with the incidence of NAFLD. Compared to the general population, people with NAFLD are less active in their free time, and this is true across all forms of exercise: aerobic, resistance, and total.

1.2. Stacking Ensemble Machine Learning Architecture

Invasive, expensive, and subjective liver biopsies and imaging techniques are common NAFLD diagnosis approaches. These limitations emphasize the need for more efficient and trustworthy diagnostic methods for early identification and intervention.

Machine learning can examine big and complicated information to reveal hidden patterns and correlations, revolutionizing medical diagnoses. Machine learning models can improve diagnostic accuracy and patient health insights by learning from data [8, 9].

Ensemble learning is a popular machine learning method since it combines different models to enhance prediction. A kind of ensemble learning called stacking blends the capabilities of multiple base

models to make more accurate predictions [11]. In NAFLD, where age, gender, BMI, and metabolic problems might confound diagnosis, this technique is very helpful.

To improve model interpretability, the study applies explainable AI approaches. Decision trees and ensemble approaches offer more insight into feature significance and decision-making than deep learning models, which are generally considered "black boxes". This research has used traditional machine learning methods despite deep learning networks' impressive performance gains across benchmarks.

Using a large dataset of demographic and clinical data, this research will create a stacked ensemble model to predict NAFLD. Additionally, we aim to apply explainable AI techniques to enhance the interpretability of our model, providing insights into the key features influencing NAFLD predictions. By addressing the limitations of traditional diagnostic methods and harnessing the power of machine learning, this research seeks to contribute to the advancement of NAFLD screening and diagnosis in clinical practice.

The remaining structure of the paper is as follows: Section 2 explains why previous research on ensemble learning is important. Details of the data set identification and preparation procedures are provided in Section 3. Moreover, this part explains the prospective method. Section 4 details how the solution was put into action and compares the new model to the ones that were already in use. Section 5 wraps up the study and has some thoughts on where the scope may go from here.

1.3. Related Work

Using decision trees, random forests, extreme gradient boosting, and support vector machines, the authors of [12] created classification models to identify individuals with and without NAFLD. The classifier using the Support Vector Machine had the most effective performance, achieving the greatest level of accuracy (0.801).

For their study on the optimal NAFLD diagnosis model, Ma H *et al.* [13] used 10508 patients and eleven different machine learning methods. They found that the SVM model had the best specificity (0.946) and precision (0.725), whereas the logistic regression (LR) model recognized 83.41% of the time. The AODE model has the highest sensitivity, measuring at 0.680. In this research, the prediction models were evaluated using the F-measure. The BN model had the greatest F-measure at 0.655, while the Fatty Liver Index (FLI) model had the lowest at 0.318. Based on an improvement of 9.17% in the F-measure score, the authors concluded that the BN model had the greatest performance.

In their study, Yip *et al.* [14] compared ridge regression, AdaBoost, and logistic regression with 922 cases. After all that work, the logistic regression model was able to get six important parameters, including insulin resistance, triglycerides, and alanine aminotransferase, to an accuracy of 87 to 88%. Sorino *et al.* [15] evaluated eight distinct machine-learning techniques. Based on their analysis, the authors concluded that the SVM algorithm was the best fit and produced the best results. To detect NAFLD, Cheng *et al.* [16] created many models using KNN, RF, and support vector machines (SVM). They found that RF achieved an accuracy of 80% in women and SVM an accuracy of 86.9% in males. Factors associated with insulin resistance and cholesterol were among the relevant characteristics chosen by both models.

1.4. Weaknesses of Past Works

- Limited Model Performance: Many previous studies may have relied on single machine learning models, which can limit predictive accuracy. Some research used Support Vector Machines (SVM) or logistic regression with good accuracy, but they seldom used ensemble approaches that integrate many models for better performance.
- Lack of Interpretability: Healthcare machine learning applications often lack interpretability. Critics call many models, especially deep learning ones, "black boxes," making decision-making difficult for therapists.
- Unbalanced Datasets: Previous efforts have ignored class imbalance, resulting in skewed forecasts for the majority class. In NAFLD investigations, healthy people may outnumber diseased people.

1.5. Contributions

- This work presents stacked ensemble learning, which combines XGBoost, CatBoost, Decision Trees, and AdaBoost to improve NAFLD prediction accuracy. This strategy uses each model's strengths to increase performance measures against individual models.
- With its high classification accuracy, the suggested model outperforms several current models in the literature. In clinical contexts, the ensemble technique works.
- This study emphasizes explainability in machine learning, unlike many others. The work finds the most significant NAFLD prediction characteristics using explainable AI approaches, improving model interpretability. Clinical adoption and healthcare professional trust depend on this.
- The study uses a well-defined dataset with varied variables, including age, gender, BMI, and mortality status, to analyze NAFLD risk factors more accurately. Using SMOTE to solve class imbalance increases the model's dependability

2. Materials and Methods

2.1. Dataset

The NAFLD dataset used in this research is taken from the open-source platform [17]. It contains attributes like age, gender, body mass index, height, time to death or last follow-up, and status (0= alive at last follow-up, 1 = dead). The dataset includes a population cohort of all adult NAFLD subjects from 1997 to 2014

2.2. Method

Figure 1 shows the architecture of the stacking ensemble model that has been presented. In stacking ensemble learning, the meta-classifier and base-classifiers work together. To train models and provide predictions, the meta-classifier makes use of the training set. Metaclassifiers, which rely on metadata, and base-classifiers are the two main approaches to data categorization. This tactic was used by the Netflix group called "The Ensemble," which was tied for first place in the accuracy category of the submission. Diverse base learning techniques are used by heterogeneous ensemble methods to ensure that the ensemble is diverse. Its members strive to achieve the greatest level of human performance by making use of a wide range of fundamental learning techniques,

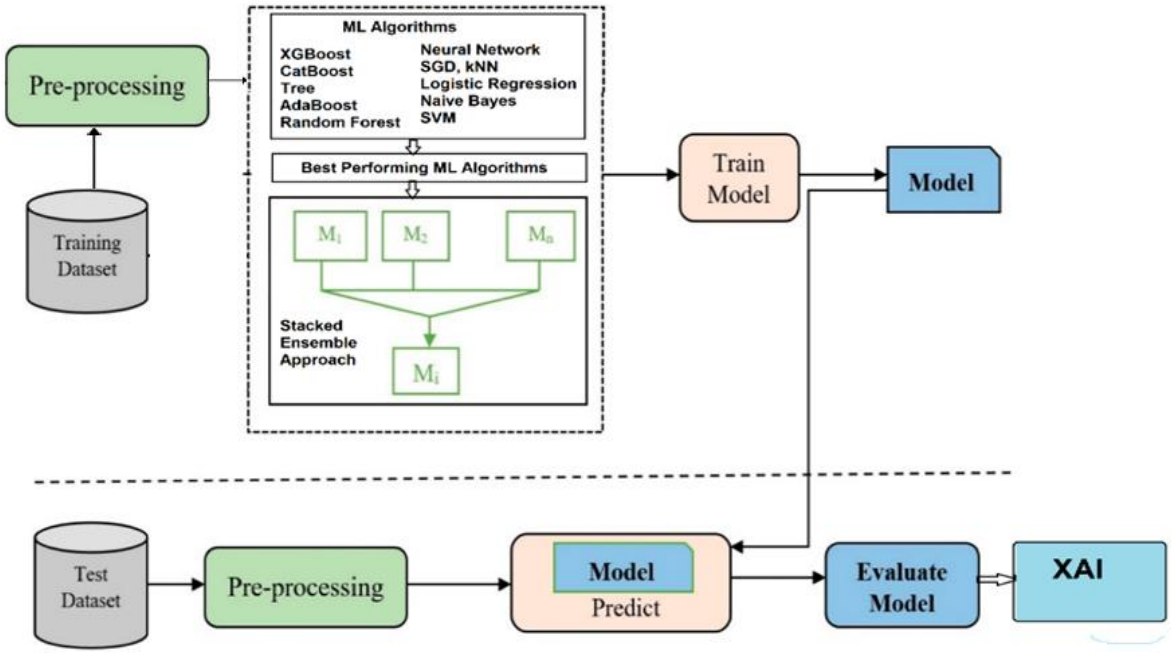

Figure 1. Proposed Stack Ensemble Approach

including decision trees, artificial neural networks, and support vector machines.

The proposed methodology is described as follows:

Proposed Methodology

Step 1: Data Preparation

- Load the NAFLD dataset
- Split the dataset into training and testing sets

Step 2: Feature Engineering

- Apply feature scaling, normalization, or other preprocessing steps
- Balancing of the dataset

Step 3: Machine Learning Models

- Use 10-fold stratified cross-validation for train and test dataset splitting
- Initialize XGBoost, CatBoost, Decision Tree, AdaBoost, Random Forest, Neural Network, SGD, kNN, Logistic Regression, Naive Bayes, SVM

Step 4: Model Training and Evaluation

for each model in [XGBoost, CatBoost, Decision Tree, ...]:

- Train the model on the training set
- Evaluate performance on the testing set

Step 5: Identify Best Performing Models

- Identify models with the highest accuracy, sensitivity, or other relevant metrics

Step 6: Stacking Ensemble

stacked_model = StackModels ([best_models])

Step 7: Train Stacked Model

- Train stacked_model on the training set

Step 8: Evaluate Stacked Model

- Evaluate stacked_model on the testing set

Step 9: Explainable Machine Learning

- ExplainModel (stacked_model) # Utilize Explainable ML methods for interpretation

Step 10: Results and Analysis

- Analyze and compare performance metrics of individual models and stacked models
- Discuss the interpretability of the stacked model using Explainable ML

Incorporating explainable machine learning (XML) techniques is crucial for enhancing the interpretability of our predictive models, especially in healthcare applications where understanding the rationale behind predictions is essential. XML allows clinicians to trust and validate the model's outputs, facilitating better decision-making.

We employed feature importance analysis to identify the most influential variables affecting the model's predictions. This analysis helps in understanding which attributes, such as age, BMI, and gender, significantly contribute to the detection of NAFLD. Techniques like permutation importance were utilized to quantify the impact of each feature on the model's performance. We also utilized SHAP (SHapley Additive exPlanations) values to provide a unified measure of feature importance. SHAP values offer insights into how each feature contributes to the final prediction for individual instances, allowing for a more granular understanding of the model's behavior.

3. Results & Discussion

3.1. Distribution of Different Attributes

The distribution of different attributes in the dataset is shown in Figures 2, 3. It helps understand the data so that it can be used and interpreted correctly.

The highest number of patients falls within the 40 to 70 years bin. The age ranges from under 20 to almost 100. The highest number of patients falls within the 60 to 100-kg bin. The weight ranges from under 40 to 180. The highest number of patients falls within the 160 to 180 cm bin. The weight ranges from under 140 to 200. The highest number of patients falls within the 20 to 40 bin. The BMI ranges from under 20 to 60.

3.2. Balancing of the Dataset

As Figure 2 clearly shows, the dataset is imbalanced. There are 16185 cases of alive at last follow-up and 1364 dead cases in the dataset. SMOTE [18] is applied to balance the dataset as follows:

Input:

C_{min}: Minority class.

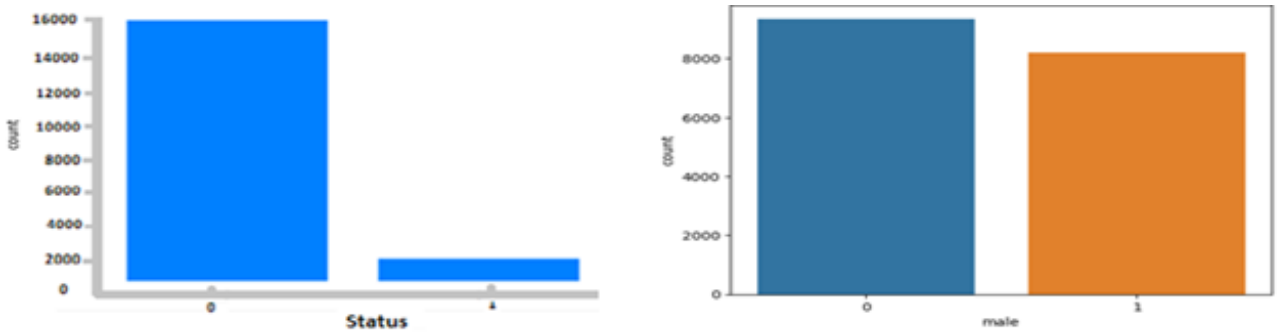

Figure 2. Distribution of Attributes (a) Status (b) Gender

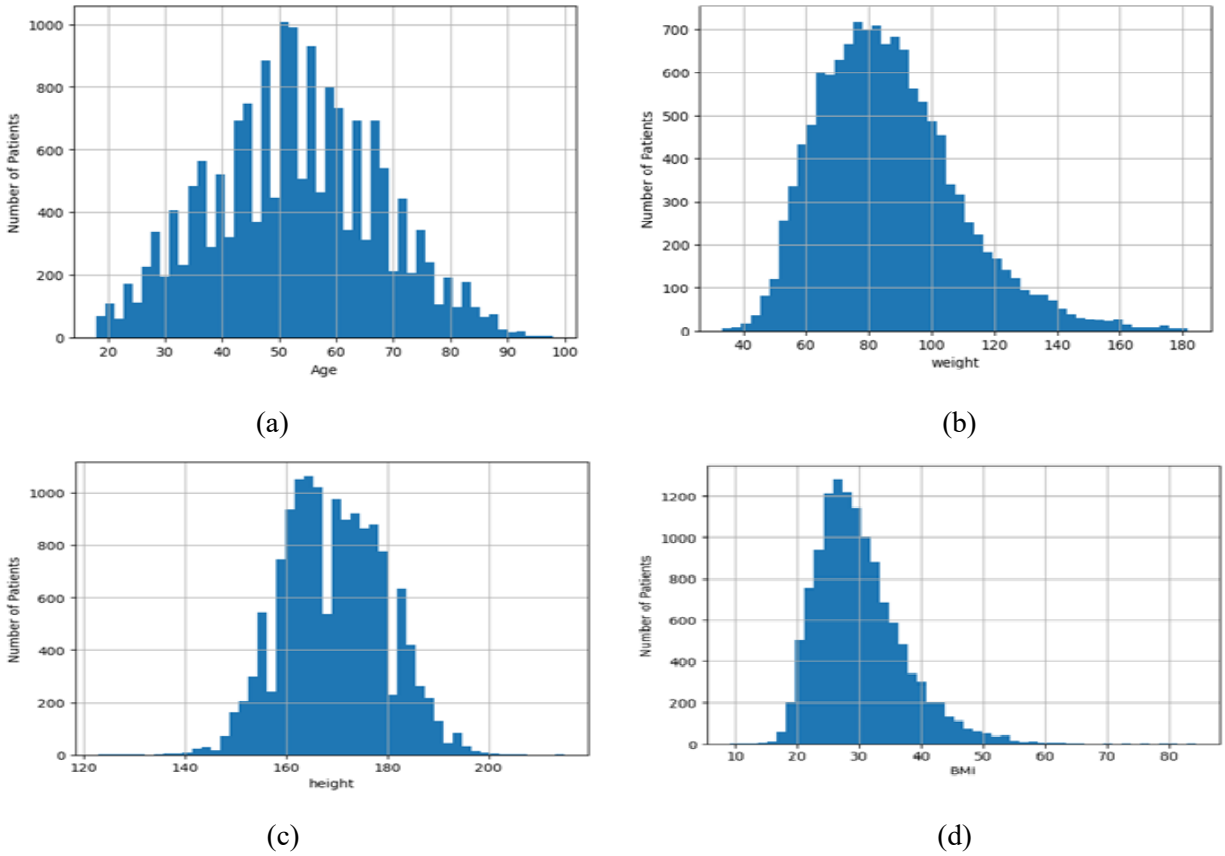

Figure 3. Distribution of other Attributes (a) Age (b) Weight (c) Height (d) BMI

C_{maj}: Majority class.

N_{min}: Number of minority class instances.

N_{maj}: Number of majority class instances.

k: Number of nearest neighbors to consider.

Output:

C_{min_new}: Synthetic minority class instances.

Procedure:

a. For each minority class instance x_i in C_{min}:

Compute the k nearest neighbors of x_i within C_{min}, denoted as $NN(x_i)$.

Randomly select one of the nearest neighbors, denoted as x_{nn}.

For each feature j, compute the difference $\Delta_{ij}=x_{nn,j}-x_{i,j}$.

Generate synthetic instances x_{new} using the formula (Equation 1):

$$x_{new,j} = x_{i,j} + rand(0,1) \times \Delta_{ij} \quad (1)$$

Append x_{new} to C_{min_new}.

b. The size of C_{min_new} is determined by a user-defined oversampling ratio.

3.3. Results

The result of the proposed stacked ensemble approach and other basic machine learning and boosting approaches are shown in Table 1 in terms of Area Under Curve (AUC), Classification Accuracy (CA), F1 score, Precision, Recall, Matthews Correlation Coefficient (MCC), Specificity (Spec), and Logistic Loss (LogLoss).

The "Stack" model, an ensemble of models, outperforms the individual models on most measures, suggesting that the NAFLD detection process is more accurate and resilient as a consequence of the use of many algorithms. Both XGBoost and CatBoost stand on their own as very effective models, surpassing competitors across a range of performance indicators (Figure 4).

Figure 5 shows the confusion metrics for all the machine learning models used in this research.

It is depicted in Figure 5 that the proposed stacking ensemble approach has the highest number of correct classification instances and the fewest instances of misclassification. The fact is also validated in Figures 6, 7, which show the ROC curve, which is a plot that displays the sensitivity and specificity of applied models [19]. The more that the ROC curve hugs the top left corner of the plot, the better the model does at classifying the data.

Figures 8, 9 show the calibration, and it is a known fact that a perfectly calibrated model would have a calibration curve closely aligned with the 45-degree diagonal line on the plot [20]. The proposed stacking approach also outperforms here.

Figure 10 depicts the important features that contribute most towards correct classification. It is evident that age is the most important feature.

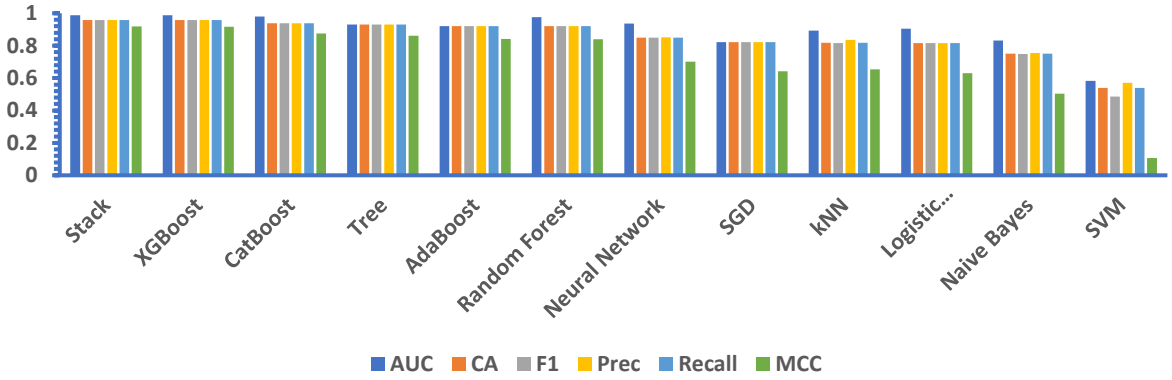

Figure 4. Performance Comparison

Table 1. Performance Comparison

Model	Area Under Curve	Classification Accuracy	F1 Score	Precision	Recall	Matthews Correlation Coefficient	Specificity	Log Loss
Stack	**0.988**	**95.90%**	**0.959**	**0.959**	**0.959**	**0.918**	**0.924**	**0.124**
XGBoost	0.988	95.80%	0.958	0.958	0.958	0.917	0.917	0.127
CatBoost	0.98	93.80%	0.938	0.938	0.938	0.876	0.876	0.191
Decision Tree	0.931	93.10%	0.931	0.931	0.931	0.862	0.862	0.231
AdaBoost	0.921	92.10%	0.921	0.921	0.921	0.842	0.842	0.241
Random Forest	0.976	92.20%	0.922	0.922	0.922	0.844	0.844	0.239
Neural Network	0.937	85.00%	0.85	0.85	0.85	0.7	0.7	0.31
SGD	0.821	82.10%	0.821	0.821	0.821	0.62	0.62	0.374
kNN	0.893	81.80%	0.818	0.818	0.818	0.62	0.62	0.366
Logistic Regression	0.904	81.50%	0.815	0.815	0.815	0.63	0.63	0.385
Naive Bayes	0.831	75.00%	0.749	0.749	0.749	0.5	0.5	0.42
SVM	0.583	54.00%	0.486	0.54	0.54	0.08	0.08	0.6

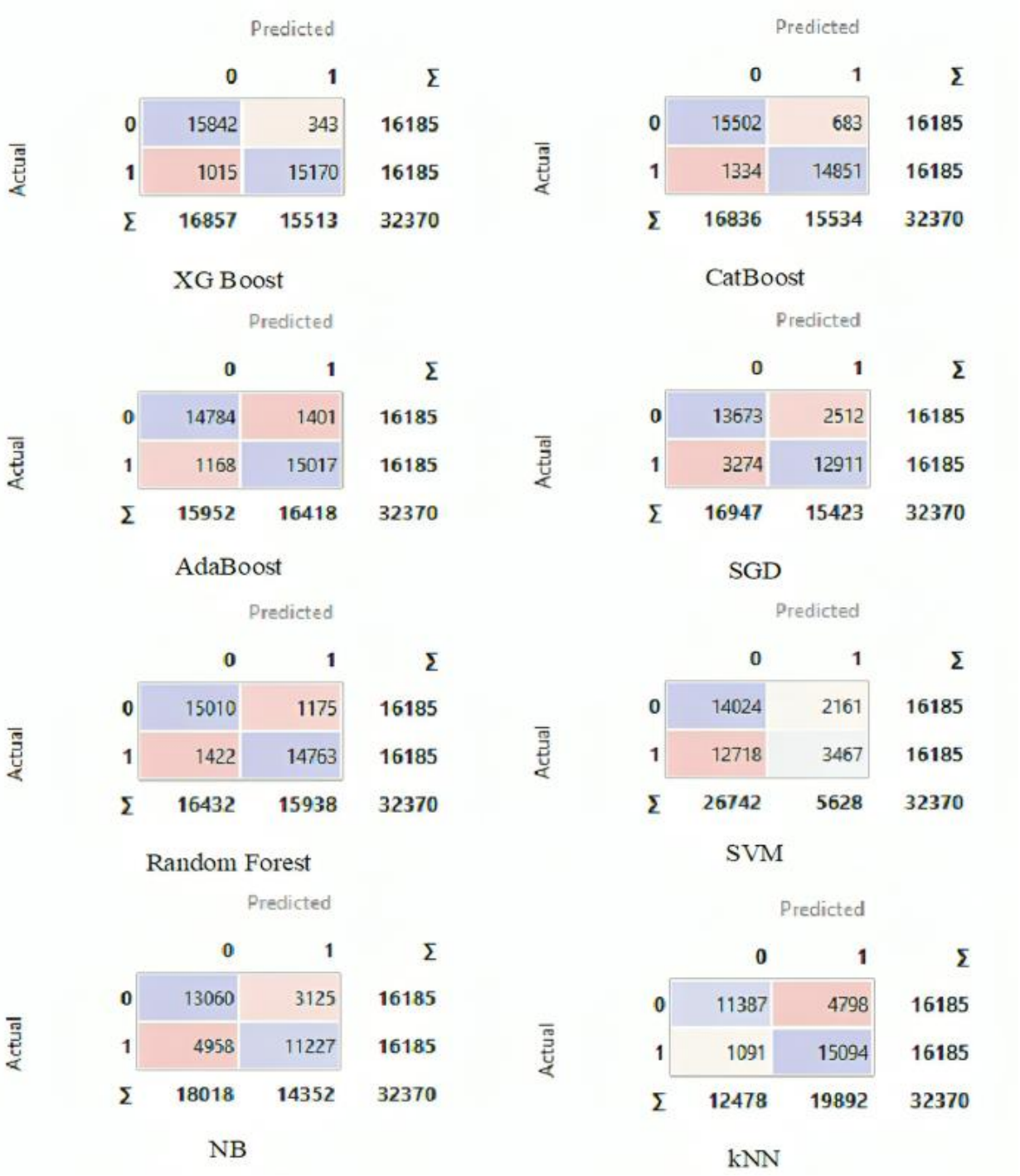

Figure 5. Confusion Matrix for Proposed and state- of- art ML Approaches

Figure 11a states that despite its presence in younger age groups, NAFLD is more prevalent among middle-aged and elderly people. Figure 11b shows that NAFLD is more prevalent in males. It has been suggested that female hormones protect against NAFLD.

4. Conclusion

This study used a wide variety of machine-learning methods to improve NAFLD detection. Finding the best models and making use of ensemble learning to get the best results was the goal of this study. Combining XGBoost, CatBoost, Decision Tree, and AdaBoost in a stacked ensemble model yielded the

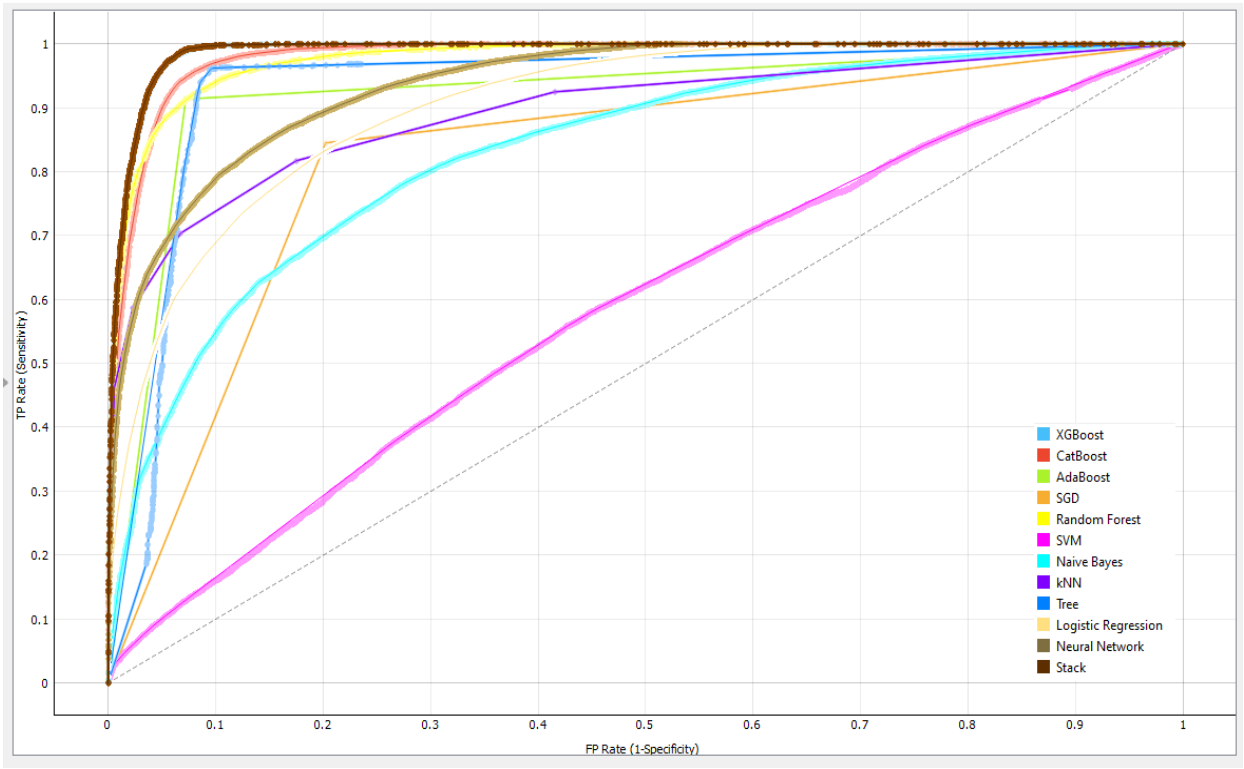

Figure 6. ROC for Target Class 0

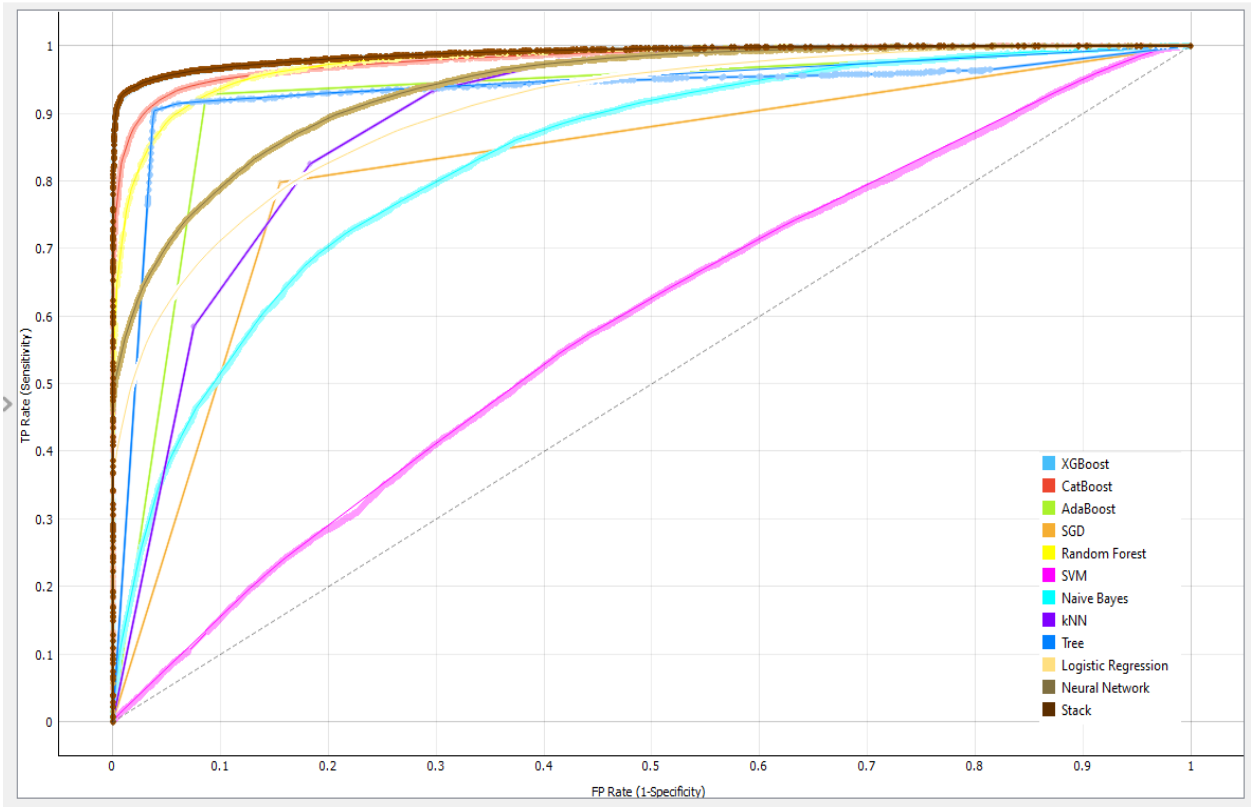

Figure 7. ROC for Target Class 1

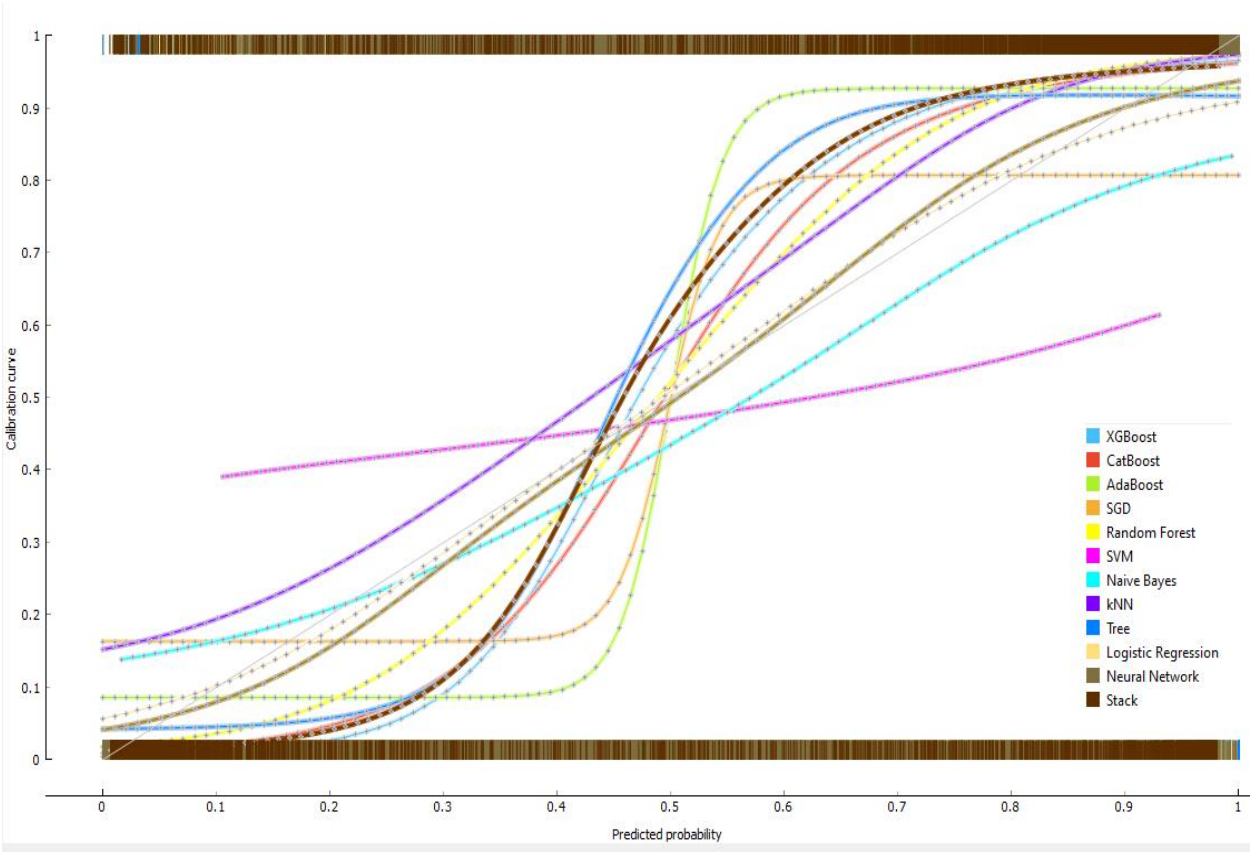

Figure 8. Calibration Curve for target class 0

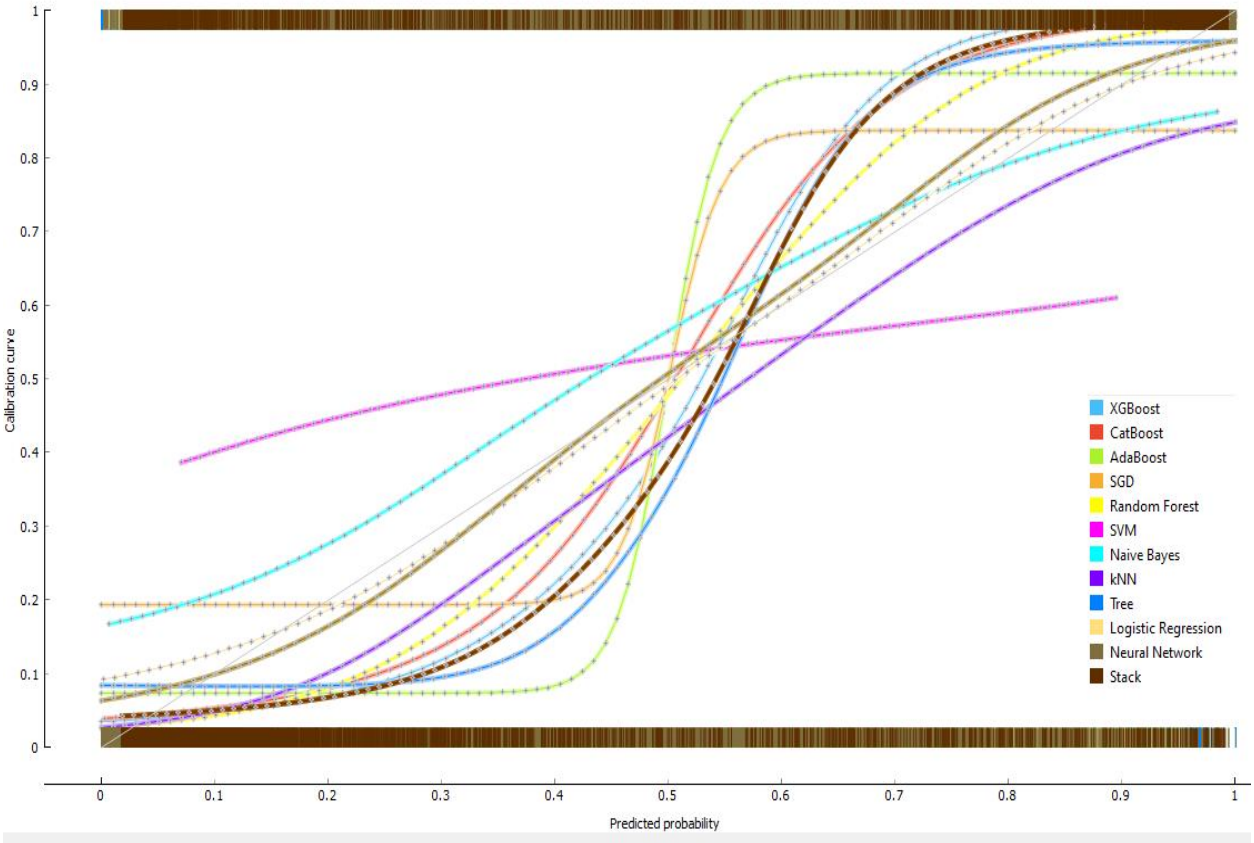

Figure 9. Calibration Curve for target class 1

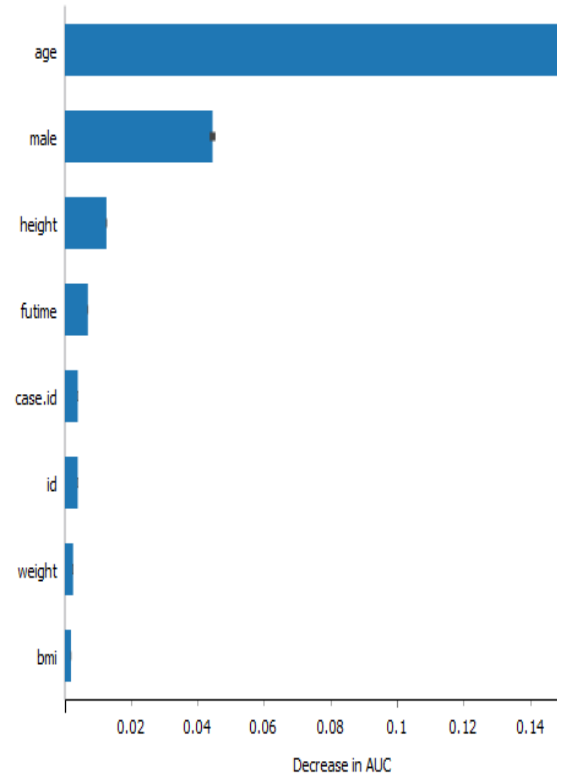

Figure 10. Feature Importance

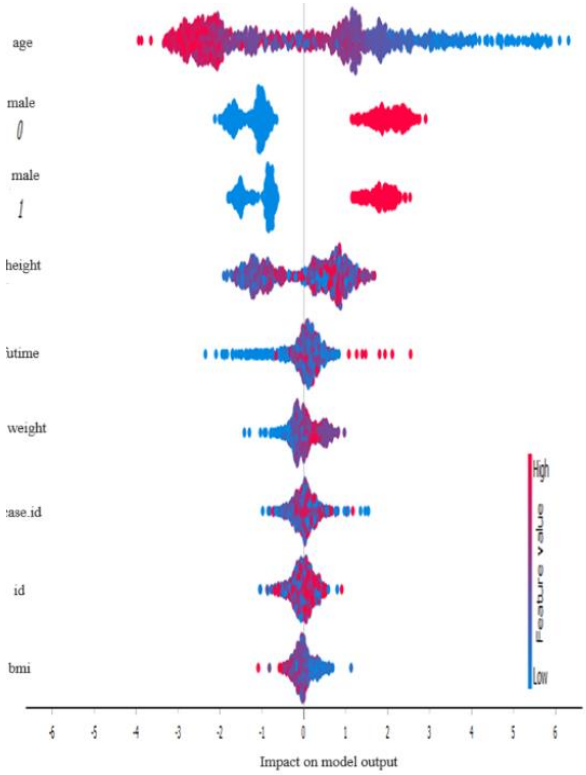

Figure 11b. Model Explanation for Target Class 0

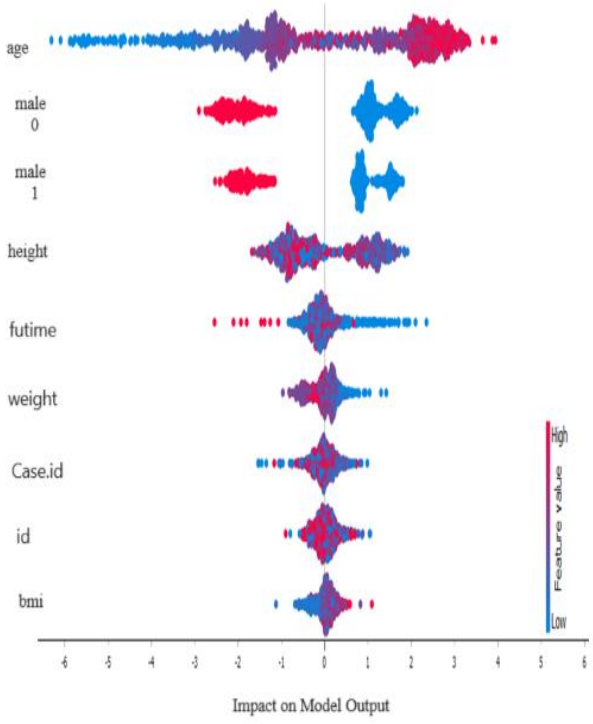

Figure 11a. Model Explanation for Target Class 1

best results, with an astounding accuracy of 95.9%. In comparison to state-of-the-art studies, the proposed approach demonstrates a notable advancement in accuracy and reliability, positioning it as a valuable tool for healthcare decision-making. Important metrics, including confusion matrices, ROC curves, and calibration curves, were used to validate the findings. These tests showed that the suggested ensemble method is reliable and resilient. The high level of accuracy indicates that the proposed approach might be used in clinical settings to identify NAFLD. When explainable machine learning is applied to the proposed approach, it reveals that age and gender are the most important characteristics in predicting NAFD. In the future, this research can be extended to broad datasets that include a wider range of individuals and various clinical environments.

References

1- Díaz LA, Arab JP, Louvet A, Bataller R, Arrese M. "The intersection between alcohol-related liver disease and nonalcoholic fatty liver disease." *Nat Rev Gastroenterol Hepatol*.; 20(12): 764-83, (2023).

2- Paik JM, Henry L, Younossi Y, Ong J, Alqahtani S, Younossi ZM. "The burden of nonalcoholic fatty liver disease (NAFLD) is rapidly growing in every region of the world from 1990 to 2019." *Hepatol Commun*. 7(10), (2023).

3- Riazi K, Azhari H, Charette JH, Underwood FE, King JA, Afshar EE, *et al.* "The prevalence and incidence of NAFLD worldwide: A systematic review and meta-analysis." *Lancet Gastroenterol Hepatol*; 7(9): 851-61, (2022).

4- Zhu X, Huang Q, Ma S, Chen L, Wu Q, Wu L, *et al.* "Presence of sarcopenia identifies a special group of lean NAFLD in middle-aged and older people." *Hepatol Int*; 17(2): 313-25, (2023).

5- Wong VW, Ekstedt M, Wong GL, Hagström H. "Changing epidemiology, global trends and implications for outcomes of NAFLD." *J Hepatol*.; 79(3): 842-52, (2023)

6- Radu F, Potcovaru CG, Salmen T, Filip PV, Pop C, Fierbințeanu-Braticievici C. "The Link between NAFLD and Metabolic Syndrome." *Diagnostics (Basel)*; 13(4), (2023)

7- Bianco A, Franco I, Curci R, Bonfiglio C, Campanella A, Mirizzi A, *et al.* "Diet and Exercise Exert a Differential Effect on Glucose Metabolism Markers According to the Degree of NAFLD Severity." *Nutrients*; 15(10), (2023).

8- Ujjwal AS, A. K. Jain, and R. Tiwari G. "Exploiting Machine Learning for Lumpy Skin Disease Occurrence Detection," *2022 10th International Conference on Reliability, Infocom Technologies and Optimization (Trends and Future Directions), (2022)*.

9- RG Tiwari SY, A Misra, A Sharma. "Classification of Swarm Collective Motion Using Machine Learning." *Human-Centric Smart Computing: Proceedings of ICHCSC 2022*, (2022).

10- N. K. Trivedi RGT, A. K. Agarwal, and V. Gautam. "A Detailed Investigation and Analysis of Using Machine Learning Techniques for Thyroid Diagnosis." *International Conference on Emerging Smart Computing and Informatics (ESCI)*; (2023).

11- Khullar V, Tiwari R, Agarwal A, Dutta S. "Physiological Signals Based Anxiety Detection Using Ensemble Machine Learning." ,p. 597 – 608, (2022).

12- Qin S, Hou X, Wen Y, Wang C, Tan X, Tian H, *et al.* "Machine learning classifiers for screening nonalcoholic fatty liver disease in general adults." *Sci Rep.*; 13(1): 3638, (2023).

13- Ma H, Xu CF, Shen Z, Yu CH, Li YM. "Application of Machine Learning Techniques for Clinical Predictive Modeling: A Cross-Sectional Study on Nonalcoholic Fatty Liver Disease in China." *Biomed Res* Int.;2018:4304376, (2018).

14- Yip TC, Ma AJ, Wong VW, Tse YK, Chan HL, Yuen PC, *et al.* "Laboratory parameter-based machine learning model for excluding non-alcoholic fatty liver disease (NAFLD) in the general population." *Aliment Pharmacol Ther.*; 46(4): 447-56, (2017).

15- Sorino P, Caruso MG, Misciagna G, Bonfiglio C, Campanella A, Mirizzi A, *et al.* "Selecting the best machine learning algorithm to support the diagnosis of Non-Alcoholic Fatty Liver Disease: A meta-learner study." *PLoS One*; 15(10): e0240867, (2020).

16- Cheng H, Chou CY, Hsiung Y. "Application of Machine Learning Methods to Predict Non-Alcohol Fatty Liver Disease in Taiwanese High-Tech Industry Workers." *Proceedings of the International Conference on Data Science (ICDATA)*, (2017).

17- Kaggle. "Non-alcoholic fatty liver disease (NAFLD): Kaggle", *Bioengineering*, (2024).

18- Xu Z, Shen D, Nie T, Kou Y, Yin N, Han X. "A cluster-based oversampling algorithm combining SMOTE and k-means for imbalanced medical data." (2021).

19- Çorbacıoğlu Ş K, Aksel G. "Receiver operating characteristic curve analysis in diagnostic accuracy studies: A guide to interpreting the area under the curve value." *Turk J Emerg Med*; 23(4): 195-8, (2023).

20- Loco JV, Elskens M, Croux C, H. Beernaert. "Linearity of calibration curves: Use and misuse of the correlation coefficient," *Accreditation and Quality Assurance. Springer Link*, (2002).

.