

A Comparative Study of Machine Learning Algorithms and Explainable AI for Fitness Club Attendance Classification

Divyashri S

Department of Computer Science & Engineering
Amrita School of Computing
Amrita Vishwa Vidyapeetham
Bangalore, India
bl.sc.p2dsc23009@bl.students.amrita.edu

Deepa Gupta

Department of Computer Science & Engineering
Amrita School of Computing
Amrita Vishwa Vidyapeetham
Bangalore, India
g_deepa@blr.amrita.edu

Abstract— The growing interest in health and fitness has resulted in an increase in the number of people joining fitness clubs. However, maintaining member engagement remains a challenge for fitness club operators. The goal of this study is to create a binary classification model that predicts whether a member is likely to attend a fitness club or not. In this paper, we propose a framework to explain how good the fitness club attendance classification performance will be under various Machine learning models by understanding its metrics. In this study by employing several classification models to determine Fitness Club attendance Classification using machine learning algorithms such as Random Forest, Gradient Boosting, Light GBM, etc. The Random Forest classifier performs with a higher accuracy of 0.87, and Explainable AI has been used to reveal the most significant feature.

Keywords— Fitness club attendance, Machine learning, Random Forest, Gradient Boosting, Explainable AI

I. INTRODUCTION

In recent years, the global surge in health and wellness consciousness has resulted in an unprecedented increase in Fitness Club memberships. While the influx of members paints a positive picture for the Fitness Industry, the challenge lies in retaining and maximizing the engagement of these members' overtime. Predicting whether a fitness club member will attend on a given day is critical for optimizing Club management, resource allocation and member experience. Engaging in fitness activities is integral to maintaining good health and preventing diseases. Regular visit to a fitness club contributes significantly to reducing unhealthy habits, promoting overall well-being. Exercise is known to be beneficial for cardiovascular health, reducing the risk of heart diseases & stroke. It can also aid in managing & preventing conditions like obesity, type 2 diabetes and certain types of cancers.

Predicting Fitness Centre Dropout using machine learning classifiers by Pedro Sobereiro et.al, 2021 [1] is the study evaluated machine learning algorithms on a dataset of 5209 members from Portuguese fitness Centre to predict dropout, Gradient Boosting Classifier outperformed predicting dropout (Sensitivity=0.986) and the Random Forest Classifier was the best at predicting non-dropout (Specificity=0.780) so the most relevant variables predicting dropout were "non - attendance days", "Total length of Stay", and "Total amount billed".

The Churn prediction for Gym members using Artificial Neural Networks Assisted with the psychological concept by of Habit formation in the fitness Industry by May Aldosary et al. in 2021 [2] advocates for gym retention strategies in the competitive fitness industry and proposes a high-performing multi-layer perceptron artificial neural networks with 92.1% accuracy, emphasizing the incorporation of habit formation concepts. A Portable Smart Fitness Suite for Real-Time Exercise Monitoring and posture correction by Abdul Hannan et al., 2022 [3], signifies the wearable technology in fitness and introduces a proposed fully portable smart fitness suite with the virtual guidance, achieving 89% accuracy using a KNN model for exercise like T-bar and bicep curls.

A prediction model of retention in a Spanish fitness Centre using machine learning classifier by San Emeterio by et al., 2016 [4] is the study developed a dropout prediction model for Spanish Sport Centre using actual customer behavior data, achieving over 70% effectiveness with a logistic regression model. Key variable included age, attendance frequency, membership length and economic aspects, emphasizing the value of historical behavior records for accurate dropout predictions. Building an Explainable Diagnostic Classification Model for Brain Tumor using Discharge Summaries by Nair, P.C at.al, 2022 [5] classifies suprasellar lesions using machine learning on clinical features from 422 discharge summaries from NIMHANS. Classifiers like XGBoost, LightGBM, and CatBoost are used with interpretability enhanced by the ELI5 tool for clinician understanding. Automatic Symptom Extraction from Unstructured Web Data for Designing Healthcare Systems by Nair, P.C et.al [6] extracts symptoms for diseases from medical websites to design a clinical Decision Support System, validated using SNOMED-CT using NLP techniques for processing data from popular sites. These studies on classifying medical conditions and extracting disease symptoms from websites demonstrate the power of ML and NLP in data-rich environments. Similarly, our work leverages these techniques to classify fitness club attendance patterns, enhancing predictive accuracy and actionable insights for fitness management.

In this work, the contribution follows a comprehensive approach where multiple classifiers were implemented and compared model's accuracy, precision, recall, and f1- score. The challenges posed by data imbalance were effectively addressed and to validate the model's performance, a meticulous 5-fold cross validation was carried out. Explainable AI (XAI) was utilized to identify the most influential features that contributes to the prediction.

There are four sections to the paper. The study's related works are covered in Section II. Section III presents the

experiment's materials and methodology. Section IV gives the result and analysis. The work is concluded in Section V.

II. RELATED WORKS

A model for predicting dropouts from physical activity interventions in leisure centers by Ivan Clavel San Emeterio et al., 2020 [7] is the study examined the behavior of 14,522 adult members in Spanish leisure centers for a year, identifying younger individuals with shorter memberships and lower facility usage as more likely to dropout. Exercise Motion Classification from Large Scale wearable sensor data using convolutional neural networks by Terry Taewoong et al., [8] employs a forearm-worn wearable sensor and a convolutional neural network (CNN) to classify large-scale exercise motion data formatted as images, achieving 92.1% accuracy in classifying 50 gym exercises. The Study explores the effect of image formatting and different CNN architectures on performance.

Fitness Movement types and Completeness Detection using a Transfer-Learning-Based Deep Neural Network by Kuan- Yuchen et.al.,2022[9] addresses the importance of fitness for health, promoting home fitness using minimal equipment It introduces a deep transfer learning-based approach, achieving high accuracy 98.56% and completeness classification. Saeed Ali Alsareii et.al [10] using the ML classifier performance highlighting importance of class Imbalance. Nuno M. Rodrigues et.al [11] addresses the lack of understanding of the global structure of hyperparameters spaces in neural networks.Ivan Cleavel San Emeterio et al [12] logistic regression models based on historical behavior demonstrated around 70% effectiveness in predicting abandonment. Jas Samrl et.al [13] conducts experiments using off-the-shelf machine learning platforms to predict customer behavior and enhance gym retention.

Rui Zhao et.al [14] in the study examined in applying Deep Learning to Machine Health monitoring, leveraging low-cost sensors and internet- connected data. Deep Belief Network, Deep Boltzmann Machines, and Recurrent Neural Networks are used for better object recognition, image segmentation and Machine Translation. Jussi Tohka et.al [15] emphasizing the challenges arising from the unrealistic assumption's issues in development, validation and acceptance exists for better evaluation of ML algorithms.

A.Sabool et.al [16] reviews the latest trends in gait analysis, emphasizing the synergy between wearable sensors & Machine Learning Methods for efficient & cost -effective data collection and accurate gait feature extraction.

M.S. Singh et.al [17] introduces novel approach to overcome the limitations of CNNs in sensor data analysis, particularly in pervasive & wearable computing. D. Ravi et.al

[18] introduces a novel deep learning methodology for real time activity classification using wearable devices. Raj et.al [19] employs autoML to predict functional outcomes in patients post-mechanical thrombectomy, comparing its performance and explainability with traditional ML models. Babu et.al [20] introduces novel approach using machine learning and deep learning classifiers with resampling and data augmentation to detect spoofing speaker systems.

The literature review demonstrates how machine learning is widely used to study the fitness club attendance classification.however, there is a significant gap in assessing diverse ensemble models, including Gradient Boosting, LightGBM, and Random Forest, in fitness club attendance classification.

Moreover, incorporation of Explainable AI along with ensemble models is underexplored. There exists a notable lackof explainable AI and a wide range of classifiers—from basic to ensemble models—as well as hyperparameter tuning for modelling fitness class attendance classification.

Investigating the incorporation of explainable AI and machine learning Fitness Club attendance classification Dataset may lead to evaluations that are more precise and faster, improving our capacity to handle issues while classification is done.

III. METHODOLOGY

The architecture for the suggested models, which include explainable AI for creating the classification model and data collecting, preprocessing model construction, and model assessment, is depicted in Fig 1.

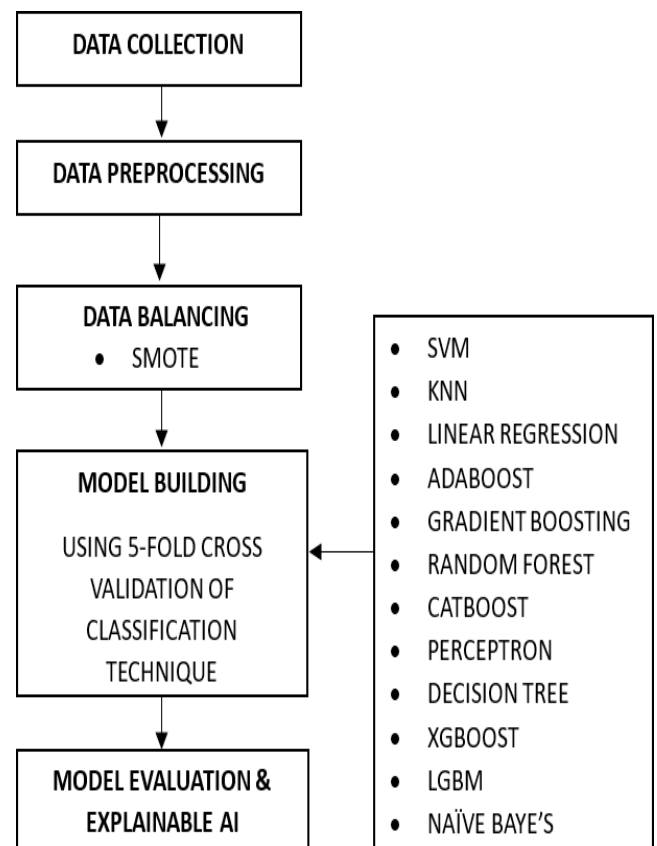


Fig 1. The architecture for Fitness Club Attendance Classification

A. Data Collection

The public dataset available in Kaggle source [21] has been used to determine the Fitness Club Attendance classification with 7 features and one target variable. There are a total of 1501 instances. 80% used for Training and 20%used for Training.

The Features in the dataset and their Data type are mentioned below in Table 1.

S.NO	Feature	Description
1	months_as_member	number of months as this fitness club member
2	weight	Weight in kg
3	days_before	number of days before the class the members registered
4	day_of_week	day of the week of the class
5	time	Time of day of the class. Either AM or PM
6	category	Category of fitness class such as Strength, HIIT, Cycling yoga
7	attended	Whether the member attended the class (1) or not class (0)

Table.1 Features in the fitness club attendance classification Dataset

B. Data Preprocessing

The target variable is the fitness club attendance which has two classes 0 and 1. The dataset used has no missing values. As the Dataset consists of categorical feature It could be useful while applying in classifiers therefore used Label Encoding to convert categorical data into numerical data to show better performance and accuracy. Random Forest provides higher accuracy that classifier is considered the heat map of the features in Fig 2. displays no highly correlated features. Thus, all the features are significant and no features can be omitted.

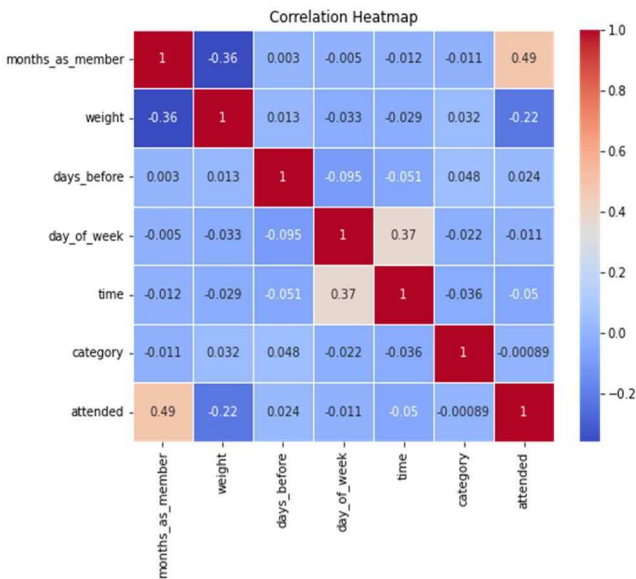


Fig 2. Heatmap of the features for Fitness Club Attendance Classification

C. Data Balancing

The distribution of the classes - attended and not attended

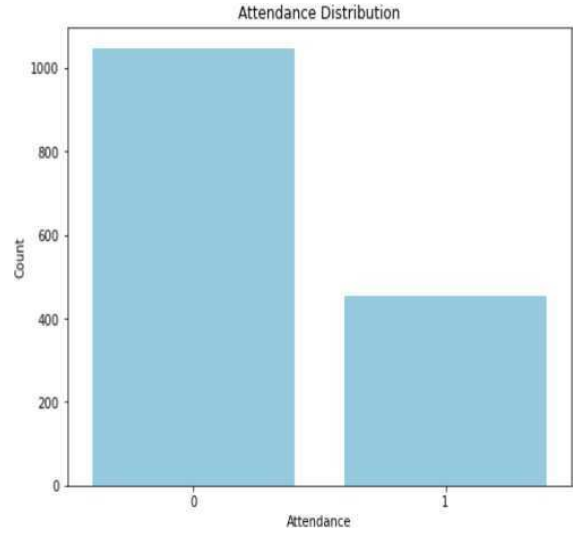


Fig 3. The unbalanced distribution of fitness club attendance target class

From Fig 3, the imbalanced data distribution clearly depicts the class imbalance. Since there is less learning in the minority class, the model is likely to perform badly since there are more not-attended than attended. Consequently, the dataset is balanced using the oversampling approach known as SMOTE (Synthetic Minority Over-sampling approach). SMOTE is used to create synthetic examples for the minority class by interpolating between the available samples.

D. Model Building

The aim of the present investigation is to categorize fitness club attendance as attended or not attended. Because there are just two possible results for this categorization task— "attended" and "not-attended"—thus making it a binary classification problem. Classification algorithms can be used to extract a model that can predict the class labels. To initiate the classification process, the data needs to be split into training and testing sets. The splitting is executed using the k- cross validation. The dataset is partitioned into k folds. Various sets of k-1 folds are utilized for training and testing in each iteration, with 1-fold being used for testing. To improve the performance of the classification models in this study, the dataset is divided into five folds, which averages the model's performance over various data subsets.

E. Implementation of Classifiers

Fitness club attendance classification plays a major rule to classify the attendance distribution. In the present study, a total of Twelve classifier algorithms have been implemented in the Fitness Club dataset for ML Classification - Random Forest, Gradient Boosting Decision Tree, KNN, XGBoost, AdaBoost, LightGBM, Perceptron model, CAT Boost, Support Vector Machine, Logistic Regression, and Naive Bayes. The dataset from the k-fold cross-validation is subjected to these Twelve classifiers, and the best fold performance is employed for prediction.

The study employs a number of machine learning models, which include the Random Forest Classifier, which, in order to lessen overfitting, combines predictions from Decision Tree, each decision tree in the Random Forest is trained using a random selection of attributes and data, which increases the model's resilience.

LightGBM's leaf-wise tree structure makes it possible to model the data in a more intricate and accurate way, increasing accuracy. Furthermore, LightGBM is a well-liked option for big data applications due to its effectiveness in managing the sizable datasets. In order to generate predictions, the KNN method measures the separation between the input data point and its closest neighbors. This method allows it to handle both numerical and categorical data with ease, making it appropriate for datasets with intricate patterns. It's also important to remember that the KNN algorithm is flexible and can be applied to a wide range of problem domains because it doesn't make any assumptions about the underlying data distribution.

XGBoost is an optimized gradient-boosting method that uses parallelization and regularization techniques to achieve scalability and accuracy. AdaBoost is an ensemble technique that builds a robust classifier by progressively enhancing weak classifiers. Gaussian Naïve Bayes was chosen due to its ease of use and efficiency when applying the Bayes theorem and conditional independence. Additionally, logistic regression – which is commended for its interpretability, simplicity, and resilience to outliers – is used to handle the binary classification problem. Support Vector Machine (SVM) is used to optimize the hyperplanes in high- dimensional spaces, and the Perceptron model's integration offers information.

F. Model Evaluation and Explainable AI

To assess each model's performance, the study takes into account a variety of assessment criteria, including the F1 score, Accuracy, Precision, Recall, and Area Under the Receiver Operating Characteristic Curve (ROC). Because the classification problem is binary in nature, the assessment process makes particular use of micro-averaged recall, accuracy, and F1 score.

The Explainable AI (XAI) model provides transparent and understandable explanations for its predictions, highlighting the feature importance and decision-making process. The XAI used in the data set are LIME (Local Interpretable Model-agnostic Explanations) and SHAP (SHapely Additive exPlanations) for the purpose of explaining the predictions and decisions made by machine learning models.

III. RESULT AND ANALYSIS

As Class imbalance is there in this dataset before applying SMOTE. The accuracy and performance were lower, Table 2 shows the classifier before applying of SMOTE and Table 3 shows the classifier after applying SMOTE.

Out of the Twelve models that were assessed, the Random Forest model had the greatest performance, with high accuracy of 0.87, and a remarkable AUC value of 0.88. Following closely after, Gradient Boosting has exceptional performance, demonstrating an accuracy of 0.86 and an AUC value of 0.85. KNN produces competitive results with an accuracy of 0.85 and an AUC value of 0.84 when used to the balanced dataset.

Table. 2 Model Comparison of fitness club attendance classification Before SMOTE

Classifier	Precision	Recall	F1 score	Accuracy
Random Forest	0.72	0.66	0.67	0.73
LightGBM	0.70	0.66	0.67	0.73
XGBoost	0.67	0.64	0.64	0.70
KNN	0.71	0.64	0.65	0.72
Decision Tree	0.65	0.64	0.65	0.69
Gradient Boosting	0.69	0.64	0.64	0.72
Adaboost	0.72	0.67	0.68	0.74
Naive Bayes	0.75	0.64	0.64	0.73
Logistic Regression	0.75	0.65	0.65	0.74
SVM	0.75	0.62	0.62	0.72
Perceptron	0.52	0.52	0.52	0.55
CatBoost	0.72	0.66	0.66	0.73

With an AUC value of 0.85, the CatBoost model exhibits good discriminatory power in line with its accuracy score of 0.85. The Decision Tree classifier shows balanced performance with an AUC value of 0.74 in addition to its accuracy score of 0.84. Table 2 shows the detailed evaluation metrics of each model with precision, recall, F1 score, and accuracy.

The LightGBM model has an accuracy of 0.84 and AUC value of 0.88. The XGBoost classifier, on the other hand, has an accuracy of 0.84 with an AUC value of 0.87. The accuracy of the Adaboost model is 0.76, with a moderate AUC value of 0.84. Having an accuracy of 0.75 and a higher AUC value of 0.83 the Support Vector Model (SVM) trails behind. Then we have the logistic regression of accuracy 0.74 with AUC value of 0.83.

Table.3 Model Comparison of fitness club attendance classification After SMOTE

Classifier	Precision	Recall	F1 score	Accuracy
Random Forest	0.88	0.87	0.87	0.87
LightGBM	0.84	0.84	0.84	0.84
XGBoost	0.85	0.84	0.84	0.84
KNN	0.85	0.84	0.84	0.85
Decision Tree	0.84	0.83	0.83	0.84
Gradient Boosting	0.87	0.86	0.86	0.86
AdaBoost	0.76	0.76	0.76	0.76
Naive Bayes	0.75	0.72	0.70	0.71
Logistic Regression	0.74	0.74	0.74	0.74
SVM	0.76	0.75	0.75	0.75
Perceptron	0.74	0.72	0.71	0.72
CatBoost	0.85	0.85	0.85	0.85

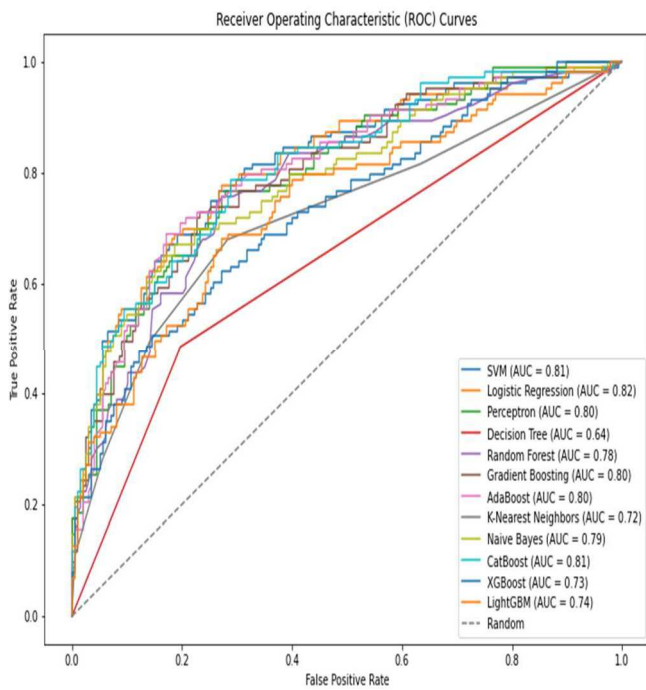


Fig.4 ROC curve for the classifiers for fitness club attendanceBefore SMOTE.

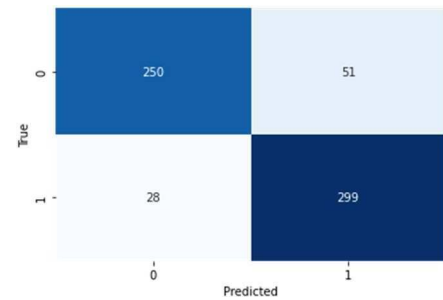


Fig.6 Confusion Matrix of best Random Forest Classifier afterSMOTE for fitness club Attendance Classification

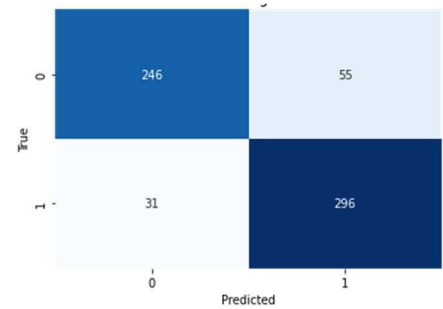


Fig.7 Confusion Matrix of best Random Forest Classifier afterSMOTE for fitness club Attendance Classification

Fig.6 shows the model correctly predicted 299 instances of class 1 with an accuracy of 87% after applying SMOTE & AUC value of 0.88 and Fig.7 shows that the model correctly predicted 299 instances of class 1 & correctly predicted 246 instances of class 0 with an accuracy of 86% & AUC value of 0.85. After seeing this analysis Random Forest classifier performs exceptionally well.

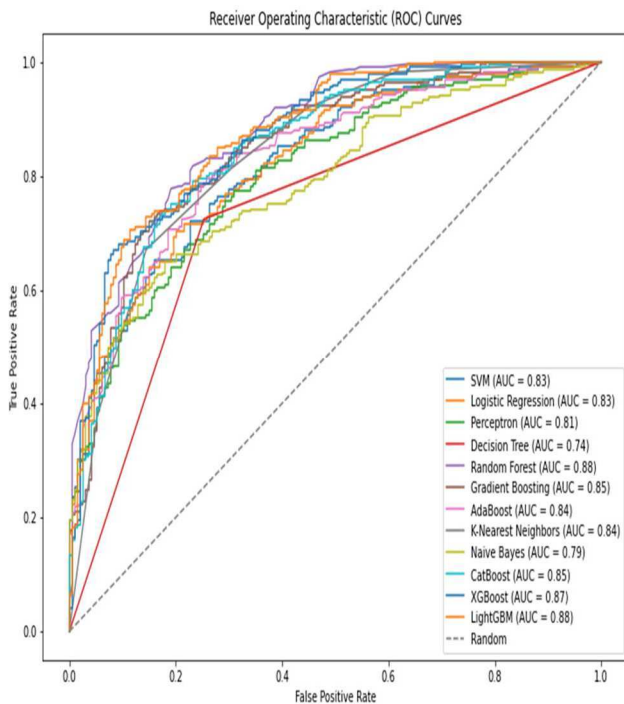


Fig. 5 ROC curve for the classifiers for fitness club attendance After SMOTE.

Fig.4 shows the ROC curve for the fitness class attendance classification without SMOTE, with an AUC value of 0.82, linear regression outperforms other classifiers, CatBoost comes in second with an AUC value of 0.81, while KNN is the least effective with an AUC value of 0.72. Fig.5 shows the ROC Curve for the fitness club attendance classification after applying the SMOTE, with an AUC value of 0.88, Random Forest and LGBM outperforms other classifiers, XGBoost comes in second with an AUC value of 0.87, while Decision Tree and Naïve Bayes's is least effective with an AUC value of 0.72 and 0.74.

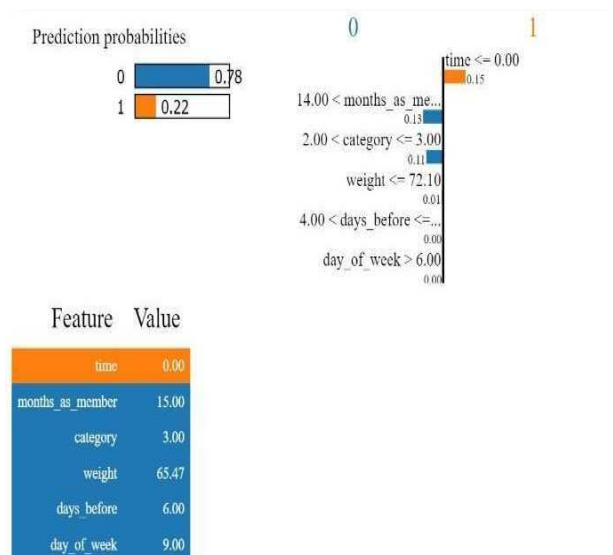


Fig. 8 performance of LIME on Random Forest for fitness club attendance classification

Fig. 8 shows of using the Random Forest classifier to visualize the LIME explanations for the data instances and display the final feature contribution in a tabular manner.

The result is made up of three main parts, which are listed from left to right: (1) the prediction made by the model; (2) the contribution of the features; and (3) the actual value for each feature. This results in a class 0 prediction of 0.78 and a class 1 prediction of 0.22. The remaining features contribute to class 0.

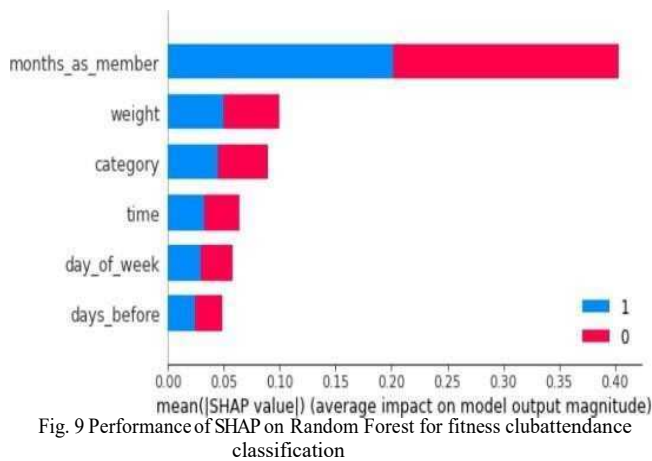


Fig.9 shows the performance of SHAP on Random Forest for fitness club attendance classification with, class 1 Shap value class 0 of 0.40 months as members is the most important feature and the least important feature is days before feature. The Fitness club attendance classification is more significantly influenced by these two characteristics. The red and blue hues, which stand for the Shapley values, are prominent when it comes to months as members and weight. This indicates that variations in these characteristics have a substantial impact on the model's output and are thus important factors in the categorization process. The interpretability is improved by the nuanced knowledge of feature significance that SHAP offers, which aids in identifying and prioritizing the factors that are essential for precise fitness club attendance classification.

LIME and SHAP differs based on LIME was fitted based on the interpretable model around individual prediction, SHAP is based on Global understanding and feature importance

IV. CONCLUSION

The main focus of this study is fitness club attendance classification, and it is clear from a thorough investigation that the Random Forest Classifier performs exceptionally well in terms of classification accuracy. Out of all the classifiers that are used, the Random Forest is the one that performs the best all the time. Notably, the Naïve Bayes model performs less well than ideal according to the accuracy testing findings. When examining the assessment metrics in further detail, the Receiver Operating Characteristic (ROC) curve offers insightful information. The ROC curve's graphic depiction highlights the Random Forest superiority in classifying fitness club attendance classification. The Random Forest classifier, which has an Area Under the Curve (AUC) value of 0.88 and performs better than all other research classifiers, supports this. Together, these results highlight how well the Random Forest model predicts Fitness club attendance classification demonstrates how well it performs when compared to other classifiers.

REFERENCES

- [1] Sobreiro, Pedro, Pedro Guedes-Carvalho, Abel Santos, Paulo Pinheiro, and Celina Gonçalves. "Predicting Fitness Centre Dropout." *International journal of environmental research and public health* 18, no. 19, 2021.
- [2] Aldosary, May, and Abdalla Alrashdan. "Churn prediction for gym members using artificial neural networks assisted with the psychological concept of habit formation in the fitness industry." *Oliveira* 2021.

- [3] Hannan, Abdul, Muhammad Zohaib Shafiq, Faisal Hussain, and Ivan Miguel Pires. "A portable smart fitness suite for real-time exercise monitoring and posture correction." *Sensors* 21, no. 19 (2021): 6692.
- [4] Clavel San Emeterio, Iván, Eliseo Iglesias-Soler, Leonor Gallardo, Sergio Rodríguez-Cañamero, and Jorge García-Unanue. "A prediction model of retention in a Spanish fitness centre." *Managing sport and leisure* 21, no. 5, 2016.
- [5] Nair, Priyanka C., Deepa Gupta, Bhagavatula Indira Devi, and Vani Kanjirangat. "Building an explainable diagnostic classification model for brain tumor using discharge summaries." *Procedia Computer Science* 218 (2023): 2058-2070.
- [6] Nair, Priyanka C., Deepa Gupta, and B. Indira Devi. "Automatic Symptom Extraction from Unstructured Web Data for Designing Healthcare Systems." In *Emerging Research in Computing, Information, Communication and Applications: ERCICA 2020*, Volume 2, pp. 599-608. Springer Singapore, 2022.
- [7] Clavel San Emeterio, Ivan, Eliseo Iglesias-Soler, Jorge Garcia-Unanue, Leonor Gallardo, and Xian Mayo. "A model for predicting dropouts from physical activity interventions in leisure centres." *Sport Sciences for Health* 16, 2020.
- [8] Um, Terry Taewoong, Vahid Babakeshizadeh, and Dana Kulić. "Exercise motion classification from large-scale wearable sensor data using convolutional neural networks." In *2017 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, pp. 2385-2390. IEEE.
- [9] Chen, Kuan-Yu, Jungpil Shin, Md Al Mehedi Hasan, Jiun-Jian Liaw, Okuyama Yuichi, and Yoichi Tomioka. "Fitness Movement Types and Completeness Detection Using a Transfer-Learning-Based Deep Neural Network." *Sensors* 22, no. 15 2022.
- [10] Saeed Ali Alsareii, Muhammad Awais, Abdulrahman Manaa Alamri, Mansour Yousef AlAsmari, Muhammad Irfan, Nauman Aslam, and Mohsin Raza. "Physical activity monitoring and classification using machine learning techniques." *Life* 12, no. 8 2022.
- [11] Rodrigues, Nuno M., Katherine M. Malan, Gabriela Ochoa, Leonardo Vanneschi, and Sara Silva. "Fitness landscape analysis of convolutional neural network architectures for image classification." *Information Sciences* 609, 2022.
- [12] San Emeterio, Iván Clavel, Jorge García Unanue, "Drop out prediction in sport centres: definition of models and reproducibility." *Retos: nuevas tendencias en educación física, deporte y recreación* 37 (2020): 54-61.
- [13] Semrl, Jas, and Alexandru Matei. "Churn prediction model for effective gym customer retention." In *2017 International Conference on Behavioral, Economic, Socio-cultural Computing (BESOC)*, pp. 1-3. IEEE, 2017.
- [14] Zhao, Rui, Ruqiang Yan, Zhenghua Chen, Kezhi Mao, Peng Wang, and Robert X. Gao. "Deep learning and its applications to machine health monitoring." *Mechanical Systems and Signal Processing* 115 2019.
- [15] Tohka, Jussi, and Mark Van Gils. "Evaluation of machine learning algorithms for health and wellness applications: A tutorial." *Computers in Biology and Medicine* 132, 2021.
- [16] Flah, Majdi, Itzel Nunez, Wassim Ben Chaabene, and Moncef L. Nehdi. "Machine learning algorithms in civil structural health monitoring: A systematic review." *Archives of computational methods in engineering* 28 2021.
- [17] M. S. Singh, V. Pondenkandath, B. Zhou, P. Lukowicz and M. Liwicki, "Transforming sensor data to the image domain for deep learning — An application to footstep detection," *2017 International Joint Conference on Neural Networks (IJCNN)*, Anchorage, AK, USA, 2017.
- [18] D. Ravi, C. Wong, B. Lo and G. -Z. Yang, "A Deep Learning Approach to on-Node Sensor Data Analytics for Mobile or Wearable Devices," in *IEEE Journal of Biomedical and Health Informatics*, vol. 21, no. 1, pp. 56-64, Jan. 2017.
- [19] Raj, Rishi, Santhosh Kumar Kannath, Jimson Mathew, and P. N. Sylaja. "AutoML accurately predicts endovascular mechanical thrombectomy in acute large vessel ischemic stroke." *Frontiers in Neurology* 14 (2023): 1259958.
- [20] Babu, Naseem, Prattipati Kumar, Jimson Mathew, and Udit Satija. "Exploration of Bonafide and Spoofed Audio Classification Using Machine Learning Models." In *2022 IEEE 19th India Council International Conference (INDICON)*, pp. 1-6. IEEE, 2022.
- [21] <https://www.kaggle.com/datasets/ddosad/datacamps-data-science-associate-certification>