

MACHINE LEARNING TO HANDLE MISSING VALUES, PREDICTION AND DETECTION OF RISK FACTOR FOR ATRIAL FIBRILLATION, DIABETES MELLITUS TYPE 2, AND PREDIABETES IN OBSTRUCTIVE SLEEP APNEA (OSA) PATIENTS

Kurnia Susvitasari¹, Yuni Rosita Dewi², Titin Siswantining¹, Diana Nur Vitasari¹, Felicia Erinna Puspitaningtyas¹, Telly Kamelia^{3,4}

¹ Department of Mathematics, Universitas Indonesia, Jakarta, Indonesia

² Department of Data Science, Universitas Negeri Surabaya, Surabaya, Indonesia

³ Division of Pulmonology, Department of Internal Medicine, Universitas Indonesia/

⁴ dr. Cipto Mangunkusumo Hospital, Indonesia

e-mail: titin@sci.ui.ac.id

DOI: 10.14710/medstat.18.2.1-12

Article Info:

Received: 13 January 2026

Accepted: 3 Agustus 2026

Available Online: 8 Agustus 2026

Keywords:

Missing Values, Multilayer Perceptron, Obstructive Sleep Apnea (OSA), Preprocessing, Random Forest, Support Vector Machine, XGBoost

Abstract: Machine learning is a discipline of artificial intelligence. It enables computational systems to exhibit behaviors derived from empirical data, without explicit programming for the task. Algorithms are systematically applied to various tasks in the biomedical and healthcare sectors, including classification, regression, clustering, dimensionality reduction, model selection, and data preprocessing. This clinical study employs machine learning techniques to resolve missing data anomalies and to create predictive models for the risk of Atrial Fibrillation (AF), Type 2 Diabetes Mellitus (DM), and Prediabetes in patients with Obstructive Sleep Apnea, in a case study at dr Cipto Mangunkusumo Hospital. The assessment of various imputation and deletion methodologies (listwise deletion, zero-imputation, central tendency substitutions, and fuzzy c-means clustering) was conducted to address the structural incompleteness of the dataset. The disease risk was categorized using several algorithms: Support Vector Machines, Random Forests, Extreme Gradient Boosting, and Multilayer Perceptrons. A missing value proportion of 25% or less was observed during the preliminary diagnostic. The empirical findings show that the classification accuracies for detecting AF were 0.908, for Type 2 DM were 0.9245, and for Prediabetes were 0.7924. The results indicate that the optimal machine learning architecture and data imputation strategy vary depending on the specific clinical targets.

1. INTRODUCTION

Obstructive Sleep Apnea (OSA) is defined by the occurrence of at least five episodes of apneic events per hour of sleep (Basner, 2007). The main symptoms of OSA are nocturnal respiratory disorders, mainly snoring and sleep disturbance, and daytime symptoms such as chronic fatigue and cognitive impairments caused by disturbed sleep patterns (Motamedi et al., 2009). Genetic factors influencing craniofacial anatomy, for instance, are associated with

heightened risk and severity of disease, including an increased susceptibility to OSA (Punjabi, 2008). Dysregulation of the autonomic nervous system often leads to cardiac arrhythmias, resulting in cardiovascular disorders such as Atrial Fibrillation (AF).

Recent studies by Alomari et al. (2025) and Ledziński et al. (2025) demonstrated the efficacy of machine learning in predicting AF and its associated risks across complex clinical populations. Specifically, advanced predictive models have become critical for identifying the development of AF in patients with concurrent comorbidities such as Type 2 DM and OSA (Xu et al., 2025).

One of the challenges in clinical datasets is the occurrence of missing values. The structure of the missingness in data, whether it is completely at random or not, should be mitigated using appropriate imputation strategy. The failure to adequately handle this issue can severely affect the accuracy of machine learning predictions (Rios et al., 2022). Commonly, methods such as listwise deletion and statistical imputation using measures of central tendency are implemented. However, these strategies are often inadequate for complex medical data. For instance, listwise deletion drastically reduces statistical power by discarding potentially valuable patient observations, while central tendency substitutions (such as mean or median) ignore the multivariate relationships between clinical features, thus reducing data variance. Aristiawati et al. (2019) showed a more sophisticated approach, Fuzzy C-Means (FCM) clustering. Unlike the conventional methods, FCM operates as a soft-clustering algorithm that allows individual data points to exhibit partial, probabilistic associations across multiple clusters, providing a more robust estimation.

Building upon our previous framework (Amalia et al., 2024), we address missing clinical data from dr. Cipto Mangunkusumo Hospital by evaluating conventional imputation strategies alongside Fuzzy C-Means. Previous research utilized the FCM technique to impute missing values in Chronic Obstructive Pulmonary Disease (COPD) data, subsequently classifying the imputed data using Random Forest (RF). In contrast, this study employs a variety of imputation techniques. Furthermore, because the clinical prediction of cardiovascular and metabolic comorbidities is a nonlinear problem, there are no single machine learning algorithm universally dominates all datasets. Therefore, it is imperative to compare multiple classifiers simultaneously. The classification methods utilized include RF, Support Vector Machine (SVM), Multilayer Perceptron (MLP), and XGBoost. During the pre-processing stage, various combinations of training and testing data were tested to identify the optimal model for predicting AF, pre-diabetes, and Type-2 DM, conditioned on OSA. The accuracy as well as the F1-score are then measured for each model and imputation strategy across difference data partition scenarios.

2. LITERATURE REVIEW

In machine learning, classification is a fundamental supervised learning technique used to predict categorical outcomes. Numerous researchers have applied these methods across various medical classification tasks. For instance, Siswantining et al. (2019) utilized the K-Nearest Neighbors (K-NN) algorithm to classify thalassemia, while Rustam et al. (2020) applied Support Vector Machines (SVM) for tumor classification. Additionally, Huljanah et al. (2019) employed a Random Forest (RF) classifier to predict prostate cancer.

Prior empirical studies conducted by Vitasari et al. (2019) assessed the factors influencing AF in a cohort of high-risk OSA patients at RSCM, utilizing a decision tree classifier applied to primary data collected via the Berlin Questionnaire. Their findings indicated that age is the primary factor affecting AF in this patient group. Furthermore, their

CHAID model identified blood pressure, waist circumference, and neck circumference as significant contributing factors, whereas their CART model highlighted blood pressure, waist circumference, body mass index, coronary heart disease, and smoking habits. This classification approach achieved an accuracy of approximately 71.7%. Similarly, Puspitaningtyas et al. (2019) examined the simultaneous interdependencies among OSA, diabetes mellitus type 2, and Prediabetes using a least-squares modeling framework. By establishing this comprehensive relationship model, the risk factors for all three conditions can be identified concurrently.

3. MATERIAL AND METHOD

This section outlines the analytical framework to predict the risk of Atrial Fibrillation (AF), Prediabetes, and Type 2 Diabetes Mellitus in Obstructive Sleep Apnea (OSA) patients.

3.1. Data

The study population comprised OSA patients at RSCM who satisfied the predefined inclusion criteria and exhibited no exclusion criteria. We established some criteria to the subjects of interest to ensure the cohort represents patients with confirmed clinical diagnosis of OSA. Conversely, we eliminate patients with highly incomplete medical records and diagnosis. This research protocol was reviewed and approved by the Health Research Ethics Committee of the Faculty of Medicine, Universitas Indonesia, dr. Cipto Mangunkusumo Hospital. A total of 348 patients were selected for the sample by a purposive sampling method. This is a non-probabilistic sampling technique, hence inherently arises potential sampling bias. Because the data is sourced from a single referral hospital (RSCM), the cohort may overrepresent patients with serious complexities of OSA compared to the general population. Consequently, this study cannot be generalized into a broader epidemiological situation.

We collected some clinical features from the OSA patients including age, gender, weight, height, Body Mass Index (BMI), waist circumference, neck circumference, systolic and diastolic blood pressure, smoking and alcohol consumption habits, Congestive Heart Failure (CHF), stroke, and Coronary Heart Disease. The primary goals are to determine the presence of Atrial Fibrillation (AF), Type 2 DM, and prediabetes within this patient cohort. The descriptive statistics for the dataset are summarized in Table 1.

Table 1. Data Describe about OSA Patient

Name of Feature	Count	Mean	Std	Min	25%	50%	75%	Max	Data Type	Note
Name	329								Object	Character
Gender	348	0.5229	0.5002	0.0	0.0	1.0	1.0	1.0	Float64	Categoric
Age	348	50.029	15.759	17.0	42.0	51.5	63.0	83.0	Float64	Numeric
Body Mass Index (BMI)	348	26.03	5.547	11.22	22.845	25.71	28.323	64.13	Float64	Numeric
Waist Circumference (WC)	348	92.764	15.227	61.0	82.0	92.76	100.0	194.0	Float64	Numeric
Neck Circumference (NC)	348	38.458	3.74	27.0	37.0	38.45	40.0	55.0	Float64	Numeric

Name of Feature	Count	Mean	Std	Min	25%	50%	75%	Max	Data Type	Note
Systole	262	134.13	19.198	86.0	120.0	130.0	150.0	190.0	Float64	Numeric
Diastole	262	80.752	11.912	53.0	70.0	80.0	90.0	140.0	Float64	Numeric
Smoke	348	0.5201	0.5003	0.0	0.0	1.0	1.0	1.0	Float64	Categoric
Coffee	348	0.1983	0.3993	0.0	0.0	0.0	0.0	1.0	Float64	Categoric
Pre-Diabetes	348	0.1552	0.3626	0.0	0.0	0.0	0.0	1.0	Float64	Categoric 1 st target
Diabetes Mellitus (DM)	348	0.4339	0.4963	0.0	0.0	0.0	1.0	1.0	Float64	Categoric 2 nd target
Atrial Fibrillation (AF)	347	0.1268	0.3332	0.0	0.0	0.0	0.0	1.0	Float	Categoric 3 rd target

3.2. Data Preprocessing

Prior to model building, we conducted a diagnostic to identify missing values. The target variable, AF, presents a very low missing value proportion of less than 1%. The patient identifier (ie., “Name”) exhibits a missing data rate of approximately 5.5%. Conversely, features such as blood pressure (systole and diastole) show the highest rates of missingness, each approaching 25%. This distribution shows the necessity for imputation strategies. During the preprocessing, the “Name” feature was removed. Subsequently, multiple imputation and deletion strategies were evaluated. To assess model robustness, the dataset was partitioned into varying training and testing subsets: 60-40, 70-30, 75-25, 80-20, and 85-15 ratios.

3.3. Missing Value Imputation

Structural incompleteness is often caused by data corruption or incomplete extraction techniques. A dataset with a high missing value rate can seriously affect statistical analysis, often leading to biased parameter estimates and inflated standard errors (Finch, 2010). To solve this issue, we evaluated listwise deletion and statistical substitutions based on measures of central tendency (mean, median, and mode) (Somasundaram, 2016), alongside a more technique: Fuzzy C-Means (FCM) clustering. As a soft-clustering algorithm, FCM tackles missing data by partitioning the feature space and estimating missing observations via the weighted centroids of the formed clusters. Prior research by Li et al. (2010) demonstrated that this probabilistic association yields significantly more robust clustering performance when processing incomplete datasets compared to traditional methods. A detailed mathematical formulation of the FCM algorithm is provided in the next section.

3.4. Fuzzy C-Means

As introduced in the previous section, the algorithm partitions a given dataset $X = \{x_1, x_2, \dots, x_n\} \subset \mathbb{R}^S$ into c fuzzy clusters by minimizing the following objective function:

$$f(U, V) = \sum_{i=1}^c \sum_{k=1}^n u_{ik}^m \|x_k - v_i\|_2^2 \quad (1)$$

subject to

$$\sum_{i=1}^c u_{ik} = 1, \quad \text{for } k = 1, 2, \dots, n \quad (2)$$

where $x_k \in \mathbb{R}^S$ represents an individual data observation, $v_i \in \mathbb{R}^S$ denotes the i -th centroid, and $V = [v_1, v_2, \dots, v_c] \in \mathbb{R}^{S \times c}$ is the cluster matrix. The parameter $m \in (1, \infty)$ dictates the degree of fuzzification, and $\|\cdot\|_2$ denotes the Euclidean norm.

FCM algorithm uses the Lagrange multiplier method to solve this optimization problem. The Lagrange function is defined as follows

$$J_\lambda(U, V) = \sum_{i=1}^c \sum_{k=1}^n u_{ik}^m \|x_k - v_i\|_2^2 - \sum_{k=1}^n \lambda_k \left(\sum_{i=1}^c u_{ik} - 1 \right) \quad (3)$$

where $\lambda = [\lambda_1, \lambda_2, \dots, \lambda_n]^T$ is the vector of Lagrange multipliers. Minimizing equation (1) subject to the constraint in equation (2) yields the necessary update conditions. The cluster centroids and membership degrees are iteratively updated according to the following equations:

$$v_i = \frac{\sum_{k=1}^n u_{ik}^m x_k}{\sum_{k=1}^n u_{ik}^m}, \quad \text{for } i = 1, 2, \dots, c \quad (3)$$

and

$$u_{ik} = \left[\sum_{j=1}^c \left(\frac{\|x_k - v_i\|_2}{\|x_k - v_j\|_2} \right)^{\frac{2}{m-1}} \right]^{-1}, \quad \text{for } i = 1, 2, \dots, c \text{ and } k = 1, 2, \dots, n \quad (4)$$

This procedure is iterated until the objective function attains a local minimum.

3.5. Some Machine Learning Techniques

After data preprocessing and imputation process, we proceed to the predictive modeling phase using a variety of machine learning classification. We assess Support Vector Machines (SVM), Random Forests (RF), Extreme Gradient Boosting (XGBoost), and Multilayer Perceptrons (MLP) for the accurate prediction of clinical objectives.

3.5.1. Support Vector Machine (SVM)

Support Vector Machines (SVM), originally formulated by Vapnik, represent a class of supervised learning models used for classification (SVC) and regression (SVR). The objective of an SVM is to identify the optimal hyperplane that maximizes the distance between the nearest data points of distinct classes, known as support vectors, within a high-dimensional feature space.

Given a dataset $D = \{(x_i, y_i)\}_{i=1}^n$ where $x_i \in \mathbb{R}^p$ and $y_i \in \{-1, 1\}$, SVM aims to find the hyperplane $w \cdot x + b = 0$ that solves:

$$\min_{w, b} \frac{1}{2} \|w\|^2 + C \sum_{i=1}^n \xi_i$$

subject to $y_i(w \cdot x_i + b) \geq 1 - \xi_i$, where $\xi_i \geq 0$ are slack variables and C is the regularization parameter. To handle non-linear decision boundaries, SVM employs kernel functions $K(x_i, x_j)$ to map input vectors into a higher-dimensional feature space, such as linear kernel,

polynomial kernel, radial basis function kernel, and sigmoid kernel. The SVM algorithm is given as follows. A detailed algorithm of SVM can be seen in (Al-Mejibli, 2020).

The SVM was implemented using the scikit-learn library in Python. Because this algorithm is highly sensitive to feature magnitudes, we rescale all continuous variables prior to training. As well, we construct the model using the Radial Basis Function (RBF) kernel with default parameters ($C = 1$ and $\gamma = 'scale'$) to strictly isolate the performance impact of the imputation techniques without interference from algorithmic tuning.

3.5.2. Random Forest (RF)

Random Forest (RF) is an ensemble learning method that constructs multiple decision trees via bootstrap aggregation, determining the final classification through majority voting (Breiman, 1997). During the training phase, the empirical input dataset is split into many bootstrap samples, and an independent decision tree is trained on each randomized subset. For the classification task, the final prediction is obtained by a majority voting over the entire generated ensemble of trees. In this study, the RF model was implemented in Python using the scikit-learn library. We set the architecture of this algorithm in this study with the parameters as follows: 100 trees and a maximum tree depth of 2.

3.5.3. Extreme Gradient Boosting (XGBoost)

Extreme Gradient Boosting (XGBoost) is an optimized ensemble technique that aggregates weak decision trees to iteratively minimize a predetermined loss function via gradient descent (Chen, 2016). In each iteration, the technique builds a new model to the residual errors of the previous ensemble, iteratively improving the overall predicted accuracy via gradient descent in function space. In this study, we configured the model with 100 sequential trees, a maximum tree depth of 3, and a step-size shrinkage of 0.1.

3.5.4. Multilayer Perceptron (MLP)

Multilayer Perceptrons (MLP) represent a fundamental class of feedforward artificial neural networks. We structured the architectural topology of our MLP with three separate strata of interconnected nodes: an input layer, hidden layers, and an output layer, see Figure 1. We applied non-linear activation functions to all computational neurons beyond the input layer and trained the network using backpropagation.

To implement the model, we utilized the Python scikit-learn library and configured the network architecture to predict each clinical target. We designed the network with two hidden layers containing five and two neurons, respectively (`hidden_layer_sizes=(5, 2)`). We optimized the model weights using the L-BFGS algorithm, which we selected for its computational efficiency on smaller clinical datasets. To prevent the model from overfitting the training data and to ensure robust generalization, we applied an L2 regularization penalty (`alpha=1e-5`).

3.6. Evaluation

We evaluated the predictive performance of the machine learning models using two primary metrics: Accuracy and F1-Score. We selected Accuracy to measure the overall proportion of correct clinical predictions across the dataset. Additionally, because clinical datasets often exhibit class imbalances, we utilized the F1-Score to calculate the harmonic mean of precision and recall, ensuring a balanced and robust assessment of the models' predictive capabilities.

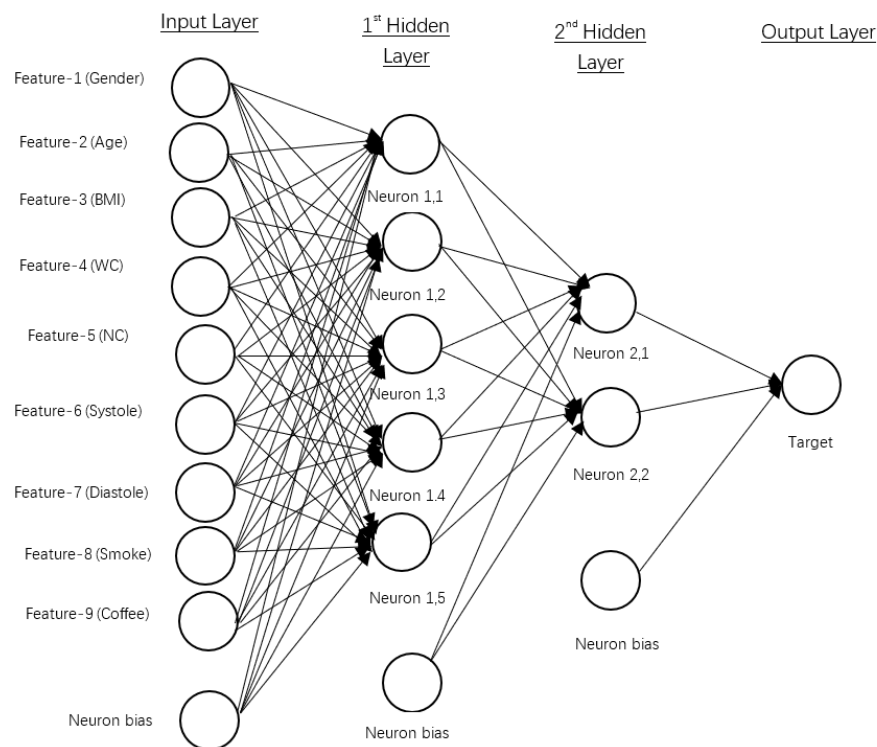


Figure 1. Architecture on MLP Model

4. RESULTS AND DISCUSSION

In our methodology, we initiate a learning process to optimize model parameters before implementing the model to the testing dataset. All model implementations, Support Vector Machine (SVM), Random Forest (RF), Extreme Gradient Boosting (XGBoost), and Multilayer Perceptron (MLP), utilize standard libraries from scikit-learn.org.

4.1. Result using Listwise Deletion Imputation

In this section, we assess the performance metrics of the machine learning architectures when missing values are handled via listwise deletion. We provide the overall classification accuracies and F1-score across all training-testing partitions evaluated in Appendix; see Table A1a and A1b. We found that the models are highly sensitive to both the training-testing partition ratios and the specific clinical targets. For the prediction of AF, XGBoost achieved the highest accuracy of 0.8868 using an 85-15 split, which is corroborated by a peak F1-score of 0.8786 (see Table A1b). Similarly, XGBoost outperformed the other models for DM prediction, reaching an accuracy of 0.7586 on a 75-25 split and a peak F1-score of 0.7549 on a 70-30 split. Conversely, for Prediabetes, the SVC, RF, and MLP all achieved an identical accuracy of 0.9245 (supported by an F1-score of 0.9159) on an 80-20 partition.

Based on the best model used for predicting each target variable, the features that influence the target variables are quite similar in hierarchical order. For instance, patient's age has the highest importance score for predicting each target variable, then followed by waist circumference (WC), diastole blood pressure, Body Mass Index (BMI), systole blood pressure, and neck circumference (NC) for AF; NC, BMI, WC, and blood pressure for Pre-Diabetes; and BMI, WC, NC, and blood pressure for DM. Across all target variables, smoking habit, coffee consumption, and patient's gender have the lowest importance scores.

4.2. Result using Zero-Imputation

Here, we evaluate the accuracy of the selected machine learning architectures when missing values are imputed via a zero-imputation strategy (replacing missing cells with a zero entry). The detailed results are provided in Appendix; see Table A2a and Table A2b. For AF prediction, the SVC, RF, and MLP models all achieved an accuracy of 0.9080 on a 75-25 split. This performance is supported by an F1-score of 0.9151 on an 80-20 split for these same models (see Table A2b). Conversely, XGBoost outperformed the other architectures in predicting DM, achieving the highest accuracy of 0.7714 on an 80-20 split, although RF yielded the highest supporting F1-score for DM (0.7316) on a 70-30 partition. For Prediabetes, the RF and MLP models reached the highest accuracy of 0.9245 on an 85-15 split, corroborated by an F1-score of 0.8883 achieved by the SVC, RF, and MLP models on that exact same partition.

The feature importance score based on the best models corresponding to each target variable reveals distinct orders. For AF, it is dominated by diastolic blood pressure and Body Mass Index (BMI), followed by waist circumference (WC), systolic blood pressure, and neck circumference (NC). Like AF, for the Pre-Diabetes, it is also mostly influenced by diastolic blood pressure and BMI, but sequentially followed by systolic blood pressure, NC, and WC. For DM, BMI and patient age have the highest score, subsequently followed by WC, NC, and diastolic blood pressure. Across all variables, the lifestyle (smoking habit and coffee consumption) as well as patient's gender are rank the lowest.

4.3. Result using Mean

In this section, we evaluate the predictive accuracy of the selected machine learning architectures under the mean substitution strategy. See Table A3a and Table A3b for results. We found that for AF prediction, the SVC, RF, and MLP models all achieved an accuracy of 0.9000 on an 80-20 split, although XGBoost yielded the highest supporting F1-score (0.8799) on a 70-30 partition (see Table A3b). For DM, the RF architecture outperformed the others, achieving the highest accuracy of 0.7924 on an 85-15 split, while XGBoost produced the optimal F1-score (0.8934) on a 70-30 partition. For Prediabetes, XGBoost achieved an accuracy of 0.8851 on a 75-25 split, whereas RF recorded the highest corresponding F1-score (0.7885) on an 85-15 partition.

Based on the best models predicting the target variables, it shows that patient age, BMI, and WC highly influence AF and DM, followed sequentially by blood pressure and NC. The best feature in hierarchical order for Pre-Diabetes variables are BMI, WC, and diastolic blood pressure. Across all target variables, patient gender and lifestyle variables (smoking status and coffee consumption) exert the least influence.

4.4. Result using Median

In this section, we assess the performance of the selected models when missing values are mitigated using median substitution. The accuracy and F1-Score metrics for each target variable, model, and splitting ratios are detailed in Table A4a and Table A4b (Appendix). For AF prediction, all evaluated models achieved an identical accuracy of 0.8857 on an 80-20 split (with SVC and RF also reaching this maximum on the 70-30 partition). For both Prediabetes and DM, the XGBoost architecture outperformed the other models, reaching an accuracy of 0.9056 on an 85-15 split and 0.7048 on a 70-30 split, respectively. Supporting these accuracy metrics, Table A4b reveals that XGBoost generated the optimal F1-scores across all three clinical targets under median substitution: 0.8716 for AF (60-40 split), 0.8914 for Prediabetes (85-15 split), and 0.7488 for DM (70-30 split).

Based on best model for each target variable, it indicates that systolic and diastolic blood pressure have the highest scores for predicting AF, followed by WC, age, BMI, and NC. In contrast, for Pre-Diabetes is heavily influenced by BMI, WC, and NC, whereas the DM is primarily influenced by patient age, WC, and BMI. Consistent with prior imputation strategies, lifestyle variables (smoking habit and coffee consumption) alongside patient gender rank lowest in the hierarchy.

4.5. Result using Mode

We evaluate the classification performance of the selected models when missing data are addressed via mode substitution. See Table A5a Table A5b (Appendix) for detailed results. For AF prediction, the SVC, RF, and MLP models all achieved an accuracy of 0.8736 on a 75-25 split. These same three architectures also yielded the highest accuracy for Prediabetes, reaching 0.9057 on an 85-15 partition. For DM, the SVC and MLP models outperformed the others, achieving a maximum accuracy of 0.9080 on a 75-25 split. Despite these specific accuracy peaks, Table A5b reveals that XGBoost generated the optimal F1-scores across all three clinical targets under mode substitution: 0.8811 for AF (60-40 split), 0.8929 for Prediabetes (80-20 split), and 0.7575 for DM (80-20 split).

For AF variable, the feature importance scores in descending order, are systolic and diastolic blood pressure, sequentially followed by patient age, BMI, WC, and NC. In contrast, the important features for Pre-Diabetes are NC and smoking status, followed by age, BMI, patient gender, and coffee consumption. For the DM variable, it is predominantly influenced by BMI, age, WC, and NC, then subsequently followed by smoking habit, blood pressure metrics, and gender.

4.6. Result using Fuzzy C-Means

We now implement Fuzzy C-Means (FCM) clustering for missing data imputation. Table A6a shows the accuracy metrics and Table A6b shows F1-score for the each target variable, model, and splitting proportion. Table A6a demonstrates that for AF prediction, the SVC and MLP architectures achieved an accuracy of 0.9080 on a 75-25 split. However, XGBoost emerged as the superior model for both diabetic outcomes, achieving an accuracy of 0.9245 for Prediabetes on an 85-15 split and 0.7714 for DM on an 80-20 split. Furthermore, Table A6b reveals that XGBoost generated the optimal F1-scores across all three clinical targets under FCM imputation: 0.8958 for AF (80-20 split), 0.8979 for Prediabetes (60-40 split), and 0.7488 for DM (85-15 split).

For the AF variable, the model is mainly influenced by diastolic blood pressure and BMI, sequentially followed by smoking habit, systolic blood pressure, patient age, WC, and NC. The hierarchy importance for Pre-Diabetes are systolic and diastolic blood pressure, BMI, patient age, NC, and WC. Finally, for the DM variable, the model emphasizes BMI, diastolic blood pressure, age, and systolic blood pressure as the primary predictors, subsequently followed by WC and NC. Across all three target variables, patient gender and coffee consumption rank lowest.

Compared with previous studies that primarily demonstrated the effectiveness of individual machine-learning or imputation approaches in clinical prediction (Aristiawati et al., 2019; Rios et al., 2022; Amalia et al., 2024; Alomari et al., 2025), the present findings show that no single imputation–classifier combination consistently performs best across different clinical outcomes. In particular, the finding that FCM did not consistently outperform simpler imputation strategies highlights an important empirical gap between its theoretical advantages and its performance in moderately incomplete clinical data. These

results suggest that the selection of imputation and classification methods should be target-specific rather than based on the assumed superiority of a particular technique.

5. CONCLUSION

The empirical evaluations conducted in this study reveal that the optimal combination of missing data imputation and machine learning architecture is highly contingent upon the specific clinical target and the ratio of training-to-testing data. For the prediction of Atrial Fibrillation (AF), the Support Vector Classification (SVC) and Multilayer Perceptron (MLP) demonstrated strong predictive capabilities across most imputation strategies. However, ensemble tree-based models, particularly Extreme Gradient Boosting (XGBoost), consistently provided the most robust overall performance, optimizing both accuracy and F1-scores for the classification of Prediabetes and Diabetes Mellitus (DM).

The main results of this study are the comparative performance of the imputation approaches. Interestingly, the Fuzzy C-Means (FCM) clustering did not outperform the traditional approaches on this dataset. Specifically, AF prediction was best with zero-imputation (0.9080), Prediabetes classification was best with listwise deletion and zero-imputation (0.9245), and DM classification was best with mean substitution (0.7924). This suggests that, although the theoretical advantages of FCM for variance preservation, most of its advantages are achieved in high-dimensional, continuous feature spaces. In discrete clinical tabular data with moderate missingness, the non-linear classifiers like XGBoost and RF can handle this issue better.

From a clinical point of view, this study offers an early prediction of health risks in Obstructive Sleep Apnea (OSA) patients. Within our specific patient group, we found that age, Body Mass Index (BMI), and blood pressure are the most important predictors. The machine learning models proposed here can help doctors to prioritize early cardiovascular and metabolic screenings. However, because our study used non-probabilistic sampling, these results cannot be automatically applied to all OSA patients. The models must be tested on larger, randomized patient groups across multiple hospitals before they can be safely used in everyday clinical practice.

There are several limitations in this study. First, the models' performance varied depending on the training and testing data ratios. This suggests the dataset is too small to ensure completely stable learning. Second, the results are highly dependent on the patient characteristics and clinical practices of one hospital (dr. Cipto Mangunkusumo Hospital) where all patient data were collected. Thus, these results are best viewed as a baseline for comparing missing data methods rather than as definitive predictions that can be automatically applied to the larger OSA population.

Future research should test these models on new patients across multiple hospitals to prove they work reliably in different settings. Additionally, before doctors can actually use these tools in real clinics, future studies must include explainable AI methods (like SHAP). This will allow doctors to clearly see exactly why the model made a specific prediction for an individual patient.

REFERENCES

Al-Mejibli, I. S., Alwan, J. K., & Abd, D. H. (2020). The Effect of Gamma Value in Support Vector Machine Performance with Different Kernels. *International Journal of*

- Electrical and Computer Engineering (IJECE)*, 10(5), 5497–5506.
<https://doi.org/10.11591/ijece.v10i5.pp5497-5506>
- Alomari, L., Jarrar, Y., Al-Fakhouri, Z., Otabor, E., Lam, J., & Alomari, J. (2025). A Machine Learning–Based Risk Prediction Model for Atrial Fibrillation in Critically Ill Patients. *Heart Rhythm O2*, 6(5), 652–660.
<https://doi.org/10.1016/j.hroo.2025.02.008>
- Amalia, S. N., Siswantining, T., & Sarwinda, D. (2024). Applications of Missing Values Imputation Using Ensemble Fuzzy C-Means Model with Majority Voting and Averaging for Chronic Obstructive Pulmonary Disease (COPD) Data. *AIP Conference Proceedings*, 3163(1), 030004. <https://doi.org/10.1063/5.0215830>
- Aristiawati, K., Siswantining, T., Sarwinda, D., & Soemartojo, S. M. (2019). Missing Values Imputation Based on Fuzzy C-Means Algorithm for Classification of Chronic Obstructive Pulmonary Disease (COPD). *Proceedings of the 8th SEAMS-UGM International Conference on Mathematics and its Applications*, 2192(1), 1551–17616.
<https://doi.org/10.1063/1.5139149>
- Basner, R. C. (2007). Continuous Positive Airway Pressure for Obstructive Sleep Apnea. *New England Journal of Medicine*, 356(17), 1751–1758.
<https://doi.org/10.1056/NEJMc066953>
- Bezdek, J. C. (1981). Pattern Recognition with Fuzzy Objective Function Algorithms. Plenum Press. <https://doi.org/10.1007/978-1-4757-0450-1>
- Biau, G. (2012). Analysis of a Random Forests Model. *Journal of Machine Learning Research*, 13(38), 1063–1095.
- Breiman, L. (1997). *Arcing the Edge* (Technical Report No. 486). Department of Statistics, University of California, Berkeley.
- Chen, T., & Guestrin, C. (2016). XGBoost: A Scalable Tree Boosting System. *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 785–794. <https://doi.org/10.1145/2939672.2939785>
- Dewi, Y. R. (2019). *Analisis Akurasi Metode Seleksi Fitur Gram-Schmidt Orthogonalization pada Support Vector Machine untuk Big Data-Studi Kasus Masalah Prediksi Klaim*. [Unpublished Master's thesis]. Universitas Indonesia.
- Finch, W. H. (2010). Imputation Methods for Missing Categorical Questionnaire Data: A Comparison of Approaches. *Journal of Data Science*, 8(3), 361–378.
[https://doi.org/10.6339/JDS.2010.08\(3\).612](https://doi.org/10.6339/JDS.2010.08(3).612)
- Huljanah, M., Rustam, Z., Utama, S., & Siswantining, T. (2019). Feature Selection Using Random Forest Classifier for Predicting Prostate Cancer. *9th Annual Basic Science International Conference 2019*, 4546(1), 1–8. <https://doi.org/10.1088/1757-899X/546/5/052031>
- Ledziński, Ł., Grzešek, E., Ledzińska, M., & Grzešek, G. (2025). Predicting Mortality in Atrial Fibrillation Patients Treated with Direct Oral Anticoagulants: A Machine Learning Study Based on the MIMIC-IV Database. *Journal of Clinical Medicine*, 14(11), Article 3697. <https://doi.org/10.3390/jcm14113697>

- Li, D., Gu, H. & Zhang, L. (2010). A Fuzzy C-Means Algorithm Based on Nearest-Neighbor Intervals for Incomplete Data. *Expert Systems with Applications*, 37(10), 6942–6947. <https://doi.org/10.1016/j.eswa.2010.03.028>
- Motamedi, K. K., MacClary, A. C., & Amedee, R. G. (2009). Obstructive Sleep Apnea: A Growing Problem. *Ochsner Journal*, 9(3), 149-53.
- Punjabi, N. M. (2008). The Epidemiology of Adult Obstructive Sleep Apnea. *Proceedings of the American Thoracic Society*, 5(2), 136–143. <https://doi.org/10.1513/pats.200709-155MG>
- Puspitaningtyas, F. E., Siswantining, T., & Kamelia, T. (2019). Analysis of Obstructive Sleep Apnea, Diabetes Mellitus Type 2, and Prediabetes at dr. Cipto Mangunkusumo Hospital using Partial Least Squares. *Journal of Physics: Conference Series*, 1321(2), 1–7. <https://doi.org/10.1088/1742-6596/1321/2/022062>
- Rios, R., Miller, R. J. H., Manral, N., Sharir, T., Einstein, A. J., Fish, M. B., Ruddy, T. D., Kaufmann, P. A., Sinusas, A. J., Miller, E. J., Bateman, T. M., Dorbala, S., Di Carli, M., Van Krieking, S. D., Kavanagh, P. B., Parekh, T., Liang, J. X., Dey, D., Berman, D. S., & Slomka, P. J. (2022). Handling Missing Values in Machine Learning to Predict Patient-Specific Risk of Adverse Cardiac Events: Insights from REFINE SPECT Registry. *Computers in Biology and Medicine*, 145, Article 105449. <https://doi.org/10.1016/j.combiomed.2022.105449>
- Rustam, Z., Hartini, S., Siswantining, T., Utami, D. A., & Putri, N. K. (2020). Comparison Between Fuzzy Kernel C-Means, Fuzzy Kernel Possibilistic C-Means and Support Vector Machines in Soft Tissue Tumor Classification. *AI2SD 2019: Advanced Intelligent System for Sustainable Development*, 92–105. https://doi.org/10.1007/978-3-030-36664-3_11
- Siswantining, T., Kamalia, A., Rustam, Z., Subroto, F., & Semendawai, A. S. (2019). Classification of Thalassemia Data using K-nearest Neighbor and Naïve Bayes. *International Journal of Advances Science and Technology*, 8(28), 15–19.
- Somasundaram, A. & Reddy, U. S. (2016). Data Imbalance: Effect and Solution for Classification of Large and Highly Imbalanced Data. *Proceedings of the 1st International Conference on Research in Engineering, Computers and Technology*. 28–36
- Vitasari, D. N., Siswantining, T., & Kamelia, T. (2019). Identification of Factor Affecting Atrial Fibrillation in A Patient with Risk of Obstructive Sleep Apnea at Rumah Sakit dr. Cipto Mangunkusumo using Decision Tree Method. *Journal of Physics: Conference Series*, 1321, 1–6. <https://doi.org/10.1088/1742-6596/1321/2/022108>
- Wuest, T., Weimer, D., Irgens, C., & Thoben, K.-D. (2016). Machine Learning in Manufacturing: Advantages, Challenges, and Applications. *Production & Manufacturing Research*, 4, 23–45. <https://doi.org/10.1080/21693277.2016.1192517>
- Xu, Y., Duan, S., Yin, W., Yang, Y., Zhang, X., & Wang, J. (2025). Machine Learning-Enhanced Prediction Model for Atrial Fibrillation Development in Patients with Concurrent Type 2 Diabetes and Obstructive Sleep Apnea Syndrome: A Comorbidity Perspective. *Nutrition & Metabolism*, 22(1), Article 158. <https://doi.org/10.1186/s12986-025-01046-4>

Appendix A: Accuracy and F-1 Metric for each imputation method and Machine Learning classification model.

Table A1a. Accuracy using Listwise Deletion (DropNaN) Imputation

Method	Accuracy with Proportion of training-testing data (%)				
	60 – 40	70 – 30	75 – 25	80 – 20	85 – 15
Accuracy Metrics for Target Variable OSA-AF					
SVC	0.8667	0.8607	0.8736	0.8113	0.8679
RF	0.8667	0.8607	0.8736	0.8113	0.8679
XGBoost	0.8381	0.8101	0.8506	0.8113	0.8868
MLP	0.8667	0.8607	0.8736	0.8113	0.8679
Accuracy Metrics for Target Variable OSA-Pre-Diabetes					
SVC	0.9238	0.8607	0.8161	0.9245	0.8679
RF	0.9238	0.8607	0.8161	0.9245	0.8491
XGBoost	0.7238	0.8607	0.8276	0.8868	0.8491
MLP	0.9143	0.8607	0.8161	0.9245	0.8679
Accuracy Metrics for Target Variable OSA-Diabetes Melitus					
SVC	0.6190	0.6582	0.6897	0.6604	0.5849
RF	0.6381	0.6456	0.6552	0.7169	0.5283
XGBoost	0.7238	0.5949	0.7586	0.6792	0.6038
MLP	0.6095	0.6329	0.5977	0.7169	0.5660

Table A1b. F1-Score Using Listwise Deletion (DropNaN) Imputation

Method	F-1 Score with Proportion of training-testing data (%)				
	60 – 40	70 – 30	75 – 25	80 – 20	85 – 15
F1-Score Metrics for Target Variable OSA-AF					
SVC	0.7243	0.8144	0.8146	0.8066	0.8065
RF	0.7243	0.8144	0.8146	0.8066	0.8065
XGBoost	0.7604	0.7955	0.8447	0.7982	0.8786
MLP	0.7243	0.8144	0.8146	0.8066	0.8065
F1-Score Metrics for Target Variable OSA-Pre-Diabetes					
SVC	0.8457	0.9060	0.7334	0.9159	0.8346
RF	0.8457	0.9060	0.7334	0.9159	0.7971
XGBoost	0.8660	0.8734	0.7803	0.8967	0.8716
MLP	0.8457	0.9060	0.7334	0.9159	0.8065
F1-Score Metrics for Target Variable OSA-Diabetes Melitus					
SVC	0.5579	0.6048	0.7003	0.4546	0.5622
RF	0.6618	0.6419	0.6453	0.4546	0.5134
XGBoost	0.7286	0.7549	0.7309	0.7508	0.5849
MLP	0.5579	0.6048	0.5996	0.4546	0.5604

Table A2a. Accuracy using Zero-Imputation (FillNaN) Strategy

Method	Accuracy with Proportion of training-testing data (%)				
	60 – 40	70 – 30	75 – 25	80 – 20	85 – 15
Accuracy Metrics for Target Variable OSA-AF					
SVC	0.8429	0.8952	0.9080	0.8857	0.8491
RF	0.8429	0.8952	0.9080	0.8857	0.8491
XGBoost	0.8071	0.8571	0.8621	0.8857	0.8491
MLP	0.8429	0.8952	0.9080	0.8857	0.8491

Accuracy Metrics for Target Variable Pre-Diabetes					
SVC	0.8357	0.8571	0.7816	0.8571	0.9067
RF	0.8214	0.8476	0.7931	0.8571	0.9245
XGBoost	0.8429	0.8476	0.8276	0.8286	0.8868
MLP	0.8071	0.8476	0.7931	0.7571	0.9245
Accuracy Metrics for Target Variable Diabetes Melitus					
SVC	0.6786	0.7048	0.5287	0.7000	0.5283
RF	0.6926	0.7048	0.5517	0.6286	0.6415
XGBoost	0.7357	0.6667	0.6896	0.7714	0.7547
MLP	0.5857	0.4762	0.4942	0.6714	0.5094

Table A2b. F1-Score Using Zero-Imputation (FillNaN) Strategy

Method	F1-Score with Proportion of training-testing data (%)				
	60 – 40	70 – 30	75 – 25	80 – 20	85 – 15
F1-Score Metrics for Target Variable OSA-AF					
SVC	0.8423	0.7912	0.8643	0.9151	0.7797
RF	0.8423	0.7912	0.8643	0.9151	0.7797
XGBoost	0.8526	0.8264	0.8609	0.9052	0.8414
MLP	0.8423	0.7912	0.8643	0.9151	0.7797
F1-Score Metrics for Target Variable Pre-Diabetes					
SVC	0.7509	0.7912	0.7016	0.6914	0.8883
RF	0.7509	0.7912	0.7016	0.6914	0.8883
XGBoost	0.8198	0.8299	0.8041	0.7432	0.8592
MLP	0.7509	0.7912	0.7016	0.6914	0.8883
F1-Score Metrics for Target Variable Diabetes Melitus					
SVC	0.5587	0.6834	0.5253	0.6409	0.5819
RF	0.5587	0.7316	0.5253	0.6420	0.5974
XGBoost	0.6049	0.7153	0.6635	0.6432	0.6686
MLP	0.3987	0.4943	0.3422	0.3820	0.3439

Table A3a. Accuracy using Mean Missing Values for OSA-AF

Method	Accuracy with Proportion of training-testing data (%)				
	60 – 40	70 – 30	75 – 25	80 – 20	85 – 15
Accuracy Metrics for Target Variable OSA-AF					
SVC	0.8643	0.8952	0.8736	0.9000	0.8491
RF	0.8643	0.8952	0.8736	0.9000	0.8679
XGBoost	0.8714	0.8952	0.8851	0.8286	0.8868
MLP	0.8643	0.5048	0.8736	0.9000	0.8679
Accuracy Metrics for Target Variable Pre-Diabetes					
SVC	0.8500	0.8571	0.8621	0.8571	0.8113
RF	0.8500	0.8476	0.8621	0.8571	0.8113
XGBoost	0.8500	0.8381	0.8851	0.8286	0.8302
MLP	0.8500	0.8571	0.8621	0.8571	0.8113
Accuracy Metrics for Target Variable Diabetes Melitus					
SVC	0.6071	0.6762	0.6092	0.7714	0.7358
RF	0.6286	0.6571	0.7126	0.7571	0.7924
XGBoost	0.6000	0.6762	0.7356	0.7428	0.7358
MLP	0.5428	0.5048	0.5402	0.6143	0.6226

Table A3b. F1-Score using mean missing values for OSA-AF

Method	F1-Score with Proportion of training-testing data (%)				
	60 – 40	70 – 30	75 – 25	80 – 20	85 – 15
F1-Score Metrics for Target Variable OSA-AF					
SVC	0.7811	0.8595	0.8146	0.8320	0.8065
RF	0.7811	0.8595	0.8146	0.8320	0.8065
XGBoost	0.8374	0.8799	0.8657	0.8000	0.8665
MLP	0.7811	0.8595	0.8146	0.8320	0.8065
F1-Score Metrics for Target Variable Pre-Diabetes					
SVC	0.7912	0.8047	0.7982	0.8320	0.7268
RF	0.7912	0.8047	0.7982	0.8320	0.7268
XGBoost	0.8299	0.8934	0.8712	0.8571	0.7770
MLP	0.7912	0.8047	0.7982	0.8320	0.7268
F1-Score Metrics for Target Variable Diabetes Melitus					
SVC	0.3655	0.3931	0.3789	0.3493	0.4778
RF	0.7309	0.6113	0.7088	0.6745	0.7885
XGBoost	0.6563	0.7036	0.6556	0.5283	0.6775
MLP	0.3655	0.3931	0.3789	0.3493	0.4778

Table A4a. Accuracy using Median Substitution Strategy

Method	Accuracy with Proportion of training-testing data (%)				
	60 – 40	70 – 30	75 – 25	80 – 20	85 – 15
Accuracy Metrics for Target Variable OSA-AF					
SVC	0.8357	0.8857	0.8161	0.8857	0.8302
RF	0.8286	0.8857	0.8276	0.8857	0.8491
XGBoost	0.85	0.8571	0.8161	0.8857	0.7736
MLP	0.8357	0.8476	0.8276	0.8857	0.8491
Accuracy Metrics for Target Variable Pre-Diabetes					
SVC	0.8429	0.8476	0.8161	0.8429	0.8868
RF	0.8429	0.8476	0.8276	0.8429	0.8868
XGBoost	0.8429	0.819	0.8161	0.8143	0.9056
MLP	0.7071	0.8476	0.8276	0.8429	0.8868
Accuracy Metrics for Target Variable Diabetes Melitus					
SVC	0.7	0.6286	0.6782	0.7	0.6038
RF	0.6786	0.6476	0.6667	0.6	0.6981
XGBoost	0.6929	0.7048	0.6552	0.6857	0.6604
MLP	0.5714	0.5619	0.6667	0.6857	0.5471

Table A4b. F1-Score using Median Substitution Strategy

Method	F1-Score with Proportion of training-testing data (%)				
	60 – 40	70 – 30	75 – 25	80 – 20	85 – 15
F1-Score Metrics for Target Variable OSA-AF					
SVC	0.8116	0.8183	0.7495	0.7912	0.7797
RF	0.8116	0.8183	0.7495	0.7912	0.7797
XGBoost	0.8716	0.8556	0.7838	0.8639	0.7237
MLP	0.8116	0.8183	0.7495	0.7912	0.7797
F1-Score Metrics for Target Variable Pre-Diabetes					
SVC	0.7210	0.7509	0.7334	0.8320	0.8336
RF	0.7210	0.7461	0.7334	0.8320	0.8336
XGBoost	0.7210	0.7797	0.7971	0.8189	0.8914
MLP	0.7210	0.7509	0.7334	0.8320	0.8336

F1-Score Metrics for Target Variable Diabetes Melitus					
SVC	0.4515	0.3931	0.5333	0.3987	0.3870
RF	0.6847	0.6443	0.6750	0.7538	0.6941
XGBoost	0.7015	0.7488	0.6872	0.6708	0.6384
MLP	0.4432	0.3931	0.5333	0.5671	0.3870

Table 5Aa. Accuracy using Modus Imputation Strategy

Method	Accuracy with Proportion of training-testing data (%)				
	60 – 40	70 – 30	75 – 25	80 – 20	85 – 15
Accuracy Metrics for Target Variable OSA-AF					
SVC	0.8643	0.8476	0.8736	0.8429	0.7736
RF	0.8643	0.8476	0.8736	0.8429	0.7924
XGBoost	0.85	0.5905	0.8506	0.8571	0.7924
MLP	0.8643	0.8476	0.8736	0.8429	0.7924
Accuracy Metrics for Target Variable Pre-Diabetes					
SVC	0.8786	0.8762	0.8391	0.8714	0.9057
RF	0.8786	0.8762	0.8391	0.8571	0.9057
XGBoost	0.85	0.8857	0.8161	0.8571	0.8868
MLP	0.8786	0.8762	0.8391	0.8714	0.9057
Accuracy Metrics for Target Variable Diabetes Melitus					
SVC	0.8571	0.8476	0.9080	0.8857	0.8302
RF	0.8571	0.8476	0.8276	0.8857	0.8302
XGBoost	0.8357	0.8571	0.8965	0.8714	0.7547
MLP	0.8571	0.8476	0.9080	0.8857	0.8302

Table A5b. F1-Score using Modus Imputation Strategy

Method	F1-Score with Proportion of training-testing data (%)				
	60 – 40	70 – 30	75 – 25	80 – 20	85 – 15
F1-Score Metrics for Target Variable OSA-AF					
SVC	0.8628	0.8320	0.8146	0.8116	0.7007
RF	0.8628	0.8320	0.8146	0.8116	0.7007
XGBoost	0.8811	0.8478	0.8082	0.8601	0.6714
MLP	0.8628	0.8320	0.8146	0.8116	0.7007
F1-Score Metrics for Target Variable Pre-Diabetes					
SVC	0.8014	0.8184	0.7657	0.8733	0.8608
RF	0.7978	0.8184	0.7657	0.8733	0.8608
XGBoost	0.8510	0.8691	0.7946	0.8929	0.8613
MLP	0.8014	0.8184	0.7657	0.8733	0.8608
F1-Score Metrics for Target Variable Diabetes Melitus					
SVC	0.3655	0.3602	0.4754	0.2718	0.4137
RF	0.6502	0.6910	0.6807	0.7093	0.6749
XGBoost	0.6711	0.6460	0.6400	0.7575	0.7159
MLP	0.3655	0.3602	0.4754	0.2718	0.4317

Table A6a. Accuracy using Fuzzy Missing Values Imputation Strategy

Method	Accuracy with Proportion of training-testing data (%)				
	60 – 40	70 – 30	75 – 25	80 – 20	85 – 15
Accuracy Metrics for Target Variable OSA-AF					
SVC	0.8571	0.8476	0.908	0.8857	0.8302
RF	0.8571	0.8476	0.8276	0.8857	0.8302
XGBoost	0.8357	0.8571	0.8965	0.8714	0.7547

MLP	0.8571	0.8476	0.908	0.8857	0.8302
Accuracy Metrics for Target Variable Pre-Diabetes					
SVC	0.8714	0.8381	0.8276	0.9000	0.8868
RF	0.8714	0.8381	0.8276	0.9000	0.8868
XGBoost	0.8929	0.9048	0.8851	0.9140	0.9245
MLP	0.8714	0.8381	0.8276	0.9000	0.8868
Accuracy Metrics for Target Variable Diabetes Melitus					
SVC	0.5500	0.5809	0.5747	0.5000	0.5283
RF	0.6929	0.6476	0.6092	0.6714	0.6604
XGBoost	0.7071	0.7238	0.7126	0.7714	0.6981
MLP	0.5500	0.5714	0.5632	0.5000	0.5283

Table A6b. F1-Score using Fuzzy Missing Values Imputation Strategy

Method	F1-Score with Proportion of training-testing data (%)				
	60 – 40	70 – 30	75 – 25	80 – 20	85 – 15
F1-Score Metrics for Target Variable OSA-AF					
SVC	0.8218	0.8457	0.8643	0.8733	0.7532
RF	0.8218	0.8457	0.8643	0.8733	0.7532
XGBoost	0.8138	0.8667	0.8769	0.8958	0.7342
MLP	0.8218	0.8457	0.8643	0.8733	0.7532
F1-Score Metrics for Target Variable Pre-Diabetes					
SVC	0.8395	0.8041	0.8406	0.6771	0.8608
RF	0.8014	0.7643	0.7495	0.6771	0.8336
XGBoost	0.8979	0.8706	0.8879	0.7329	0.8766
MLP	0.8014	0.7643	0.7495	0.4852	0.8336
F1-Score Metrics for Target Variable Diabetes Melitus					
SVC	0.6695	0.6411	0.6315	0.7467	0.7349
RF	0.68714	0.6357	0.6038	0.6857	0.6662
XGBoost	0.7435	0.7111	0.6292	0.7394	0.7488
MLP	0.3903	0.3710	0.4058	0.4852	0.3652