Evaluating the Impact of Carboxyhemoglobin Level in the Human Body in Lagos State Using Machine Learning Algorithms

Omotayo Oluwatusin[1], Nsikan Ime Obot[2], Oloruntoba Ekun[3], Olabode Thomas Olakoyejo[1], Kehinde Orolu[4], Okwisilieze Uwadoka[1], Hussein Babalola[1], Adekunle Omolade Adelaja[1], Comfort Oluwaseyi Folorunso[4*], Idowu O. Fadeyibi[5], Ayo A. Oyediran[1] and Oluyemi Akinloye[3, 6]

[1]Department of Mechanical Engineering, Faculty of Engineering, University of Lagos, Akoka, Lagos, Nigeria.
[2]Department of Physics, Faculty of Science, University of Lagos, Akoka, Lagos, Nigeria.
[3]Department of Medical Laboratory Science, Faculty of Basic Medical Sciences, College of Medicine, University of Lagos, Idi-araba, Lagos, Nigeria.
[4]Department of Systems Engineering, Faculty of Engineering, University of Lagos, Akoka, Lagos, Nigeria.
[5]Department of Surgery, College of Medicine, Lagos State University, Ikeja, Lagos, Nigeria.
[6]Center for Genomics of Non-communicable Diseases and Personalized Healthcare, University of Lagos, Akoka, Lagos, Nigeria
Corresponding Author's email: cfolorunso@unilag.edu.ng

Abstract
The detection of carboxyhaemoglobin (COHb), a remarkably stable yet harmful complex present in body cells, presents a significant challenge. Elevated COHb levels can cause symptoms like headaches, nausea, dizziness, and, in severe cases, coma or death. This study utilised thirteen predictive variables, including sex, body mass index, glucose, and blood pressure. The COHb levels in Lagos State, Nigeria, were classified using various machine learning algorithms and variables. Evaluation metrics such as accuracy, precision, and confusion matrices were employed for assessment. Highly varied but negatively correlated factors significantly influenced ML predictions of COHb. Glucose was identified as the most influential predictor, due to food oxidation, it combines with oxygen and dissociates carbon monoxide from the blood. While seven out of twelve models that did not overfit during the training phase were retained, the best-performing model was an artificial neural network (ANN) with seven hidden layers of six neurons each. Apart from being the only model that correctly classified the rare individual of the fourth group by avoiding misplacement into the first group of many persons in the confusion matrix, the ANN scheme achieved the highest scores of 70% and 64% in accuracy and precision, respectively, during generalisation, alongside other optimal performances.

Keywords: Artificial Neural Network, Carboxyhemoglobin, Machine Learning, Lagos State

1.0 INTRODUCTION

Red blood cells contain a crucial protein known as haemoglobin, which has the unique ability to bind with carbon monoxide (CO), forming the carboxyhaemoglobin (COHb) complex—a highly stable molecule. Haemoglobin also interacts with oxygen (O_2), leading to the formation of oxyhaemoglobin, a less stable compound responsible for oxygen transport within the body. O_2 is vital for oxidative phosphorylation, the final step in cellular respiration, which produces adenosine triphosphate (ATP), the primary energy source for cells. This process relies on the diffusion of oxygen across the alveolar membrane into blood tissues, facilitating essential metabolic functions in all body cells, including those in muscles, organs, and the brain (Crecelius *et al.,* 2015; Prockop and Chichkova, 2007; Piantadosi, 2004). The affinity of haemoglobin for oxygen is influenced by various factors, including the partial pressure of oxygen in the blood, the presence of gases like carbon monoxide and carbon dioxide, and the pH of the blood. Elevated levels of CO in the body often result from incomplete combustion of carbonaceous materials, such as wood, natural gas, and gasoline, leading to an excess of CO in the atmosphere. Notably, haemoglobin exhibits an affinity for CO approximately 250 times greater than its affinity for oxygen. Consequently, elevated CO levels can pose serious health risks, as the stability of COHb retains bound CO in areas of the body that require oxygen. Prolonged exposure to elevated CO concentrations and low oxygen levels can significantly increase COHb levels, resulting in adverse effects such as headaches, nausea, dizziness, cardiac dysfunction, and, in severe cases, coma or even death (Mattiuzzi and Lippi, 2020; Huzar *et al.,* 2013; Jones and Kennedy, 1982; Rodkey *et al.,* 1974).

The monitoring of COHb levels primarily relies on established methodologies rooted in spectrometry, spectrophotometry principles, and chromatography. Prominent techniques include pulse oximetry, laboratory co-oximeters, UV-Vis spectrophotometry, and gas chromatography (Bemtgen *et al.,* 2021; Samuel *et al.,* 2021; McNair *et al.,* 2019**;** Bickler and Rhodes, 2018). Pulse oximetry offers an indirect method for determining COHb levels by measuring oxygen saturation in haemoglobin within the pulsating capillary bed, illuminated under light. However, it faces limitations in differentiating accurately between haemoglobin bound to oxygen and haemoglobin bound to other gases, despite the advantage of non-invasiveness. Spectrometry-based lab co-oximeters can assess various haemoglobin forms, including COHb concentrations, with notable accuracy. Nevertheless, these devices are associated with significant costs, require expert operation, demand substantial blood samples, and yield results over an extended period. UV-Vis spectrophotometry utilises energetic radiation and tailored analytical protocols to determine compound concentrations in blood samples. In contrast, pocket-sized spectrometers employ low-energy infrared radiation, available in smartphones, to scan blood samples for COHb levels. These compact instruments offer affordability, ease of use, and accessibility, albeit with slightly lower precision compared to lab co-oximeters. Gas chromatography, a distinct analytical tool, separates or releases gases from blood before quantification using light-sensitive equipment, often involving heat detectors. However, this method, while accurate, is relatively sophisticated, time-consuming, and costly.

Given the limitations of clinical techniques for monitoring COHb levels, including high cost, time consumption, and invasiveness due to blood sample collection, alternative approaches like electronic sensors or wearables and regression analysis have gained prominence (Lee *et al.,* 2021; Adjiski et al, 2019; Wu, 2019; Oluwatusin *et al.,* 2019). While it is known that patients with premedicated heart and brain issues are more susceptible to CO poisoning, researchers have also shown that high COHb concentration in the blood can be liked to certain factors such as gender, age, BMI, smoking behaviour, ambient serenity, and maternal and foetal CO levels (Abbey *et al.,* 2022; Hampson and Hauff, 2008; Prockop and Chichkova, 2007, Piantadosi, 2004). Various mathematical models, including the Alternating-Direct Implicit Scheme, Finite volume, Gaussian distribution, Fick's law of diffusion, Monte Carlo simulation, time-series analysis, and non-Newtonian mechanics, have been applied to the prediction of COHb as well as carbon monoxide diffusion (Oluwatusin *et al.,* 2019; Sobamowo, 2016; Guarnaccia *et al.,* 2014; Isa *et al.,* 2013; Oghenejoboh and Adiotomre, 2012; Maynard and Robert, 1999). Machine learning (ML), a subset of artificial intelligence, empowers electronic circuits and computer systems to learn and enhance their performance through experience. ML algorithms identify patterns in data for tasks such as classification, clustering, regression, and identification. They require training before task execution, with different learning approaches available, including supervised, unsupervised, semi-supervised, and reinforcement learning (Obot *et al.,* 2023; Murphy, 2012). In recent times, machine learning algorithms, such as support vector machines, random forest, and artificial neural networks, have been used in medical diagnosis (Wang *et al.,* 2023; Hanis *et al.,* 2022, Shi *et al.,* 2022).

This study is particularly relevant to Lagos State, Nigeria, a rapidly developing and densely populated city with over 21 million residents, although its findings can be applied to other locations. The rapid industrialisation and urbanisation in Lagos have led to severe air pollution, primarily driven by transportation emissions, increasing the risk of respiratory illnesses such as asthma and chronic bronchitis (Adedokun and Owode, 2019; Oluwole *et al.,* 2016). Given the high population density and pollution levels, this study focuses on the application of machine learning algorithms to address the health risks associated with elevated COHb concentrations. To the best of our knowledge, in addition to the knowledge dearth of studies on COHb from developing countries like Nigeria, there has been no previous attempt to employ machine learning for modelling COHb levels in this specific context, despite the existence of various experimental and mathematical models for gas pollution and its impact on humans (Abbey *et al.,* 2022; Adedokun and Owode, 2019; Oluwatusin *et al.,* 2019; Oluwole *et al.,* 2016; Oghenejoboh and Adiotomre, 2012).

2.0 MATERIALS AND METHODS

2.1 Data collection

Stringent guidelines govern the handling of blood samples containing carboxyhaemoglobin to ensure the accurate detection of low concentrations of pollutants in the specimens. Health Research Ethics Committee of the Lagos State University Teaching Hospital with the approval number HREC: 19/12/2008a and Lagos State Ministry of Health Research Ethics Committee.

A total of 516 consented volunteers (Table 1) from all 20 local government areas of Lagos State (Fig 1) were recruited for this study. Various instruments, such as a weighing machine, metre rule, manometer, glucometer, and other necessary materials, were used for data collection during visits to the different sites. This data collection period began on 6th June 2016 and concluded on 19th August 2016. Each participant provided 8ml of blood for COHb lab analysis. To preserve the blood specimens, fluoride oxalate was used in securely stoppered tubes. These tubes were filled and stored in the dark in a deep freezer refrigerator, maintaining a temperature of 4°C. It's important to note that in the presence of oxygen at atmospheric partial pressure, these samples may undergo exchange with the carbon monoxide present in the samples, a process accelerated by exposure to light and heat.

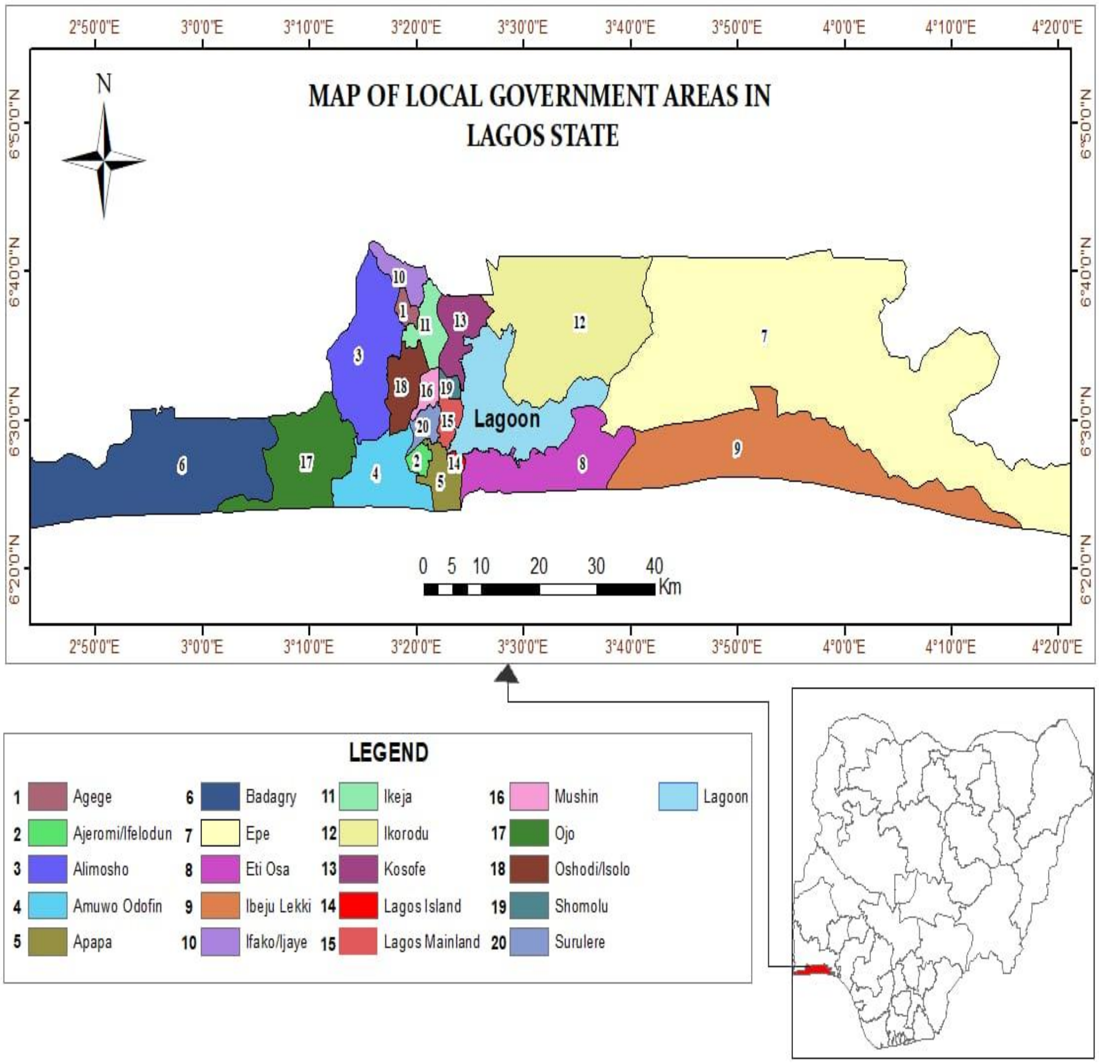

Figure 1: Location areas of the study population.

Table 1: Number of blood samples collected from consented volunteered from each LGA

S/no	LGA	No of participants
1	Agege	20
2	Ajeromi/Ifelogun	16
3	Alimosho	30
4	Amuwo Odofin	16
5	Apapa	24
6	Badagry	22
7	Epe	18
8	Eti Osa	28
9	Ibeju Lekki	21
10	Ifako/Ijaye	20
11	Ikeja	29
12	Ikorodu	25
13	Kosofe	21
14	Lagos Island	32
15	Lagos Mainland	21
16	Mushin	18
17	Ojo	20
18	Oshodi/Isolo	20
19	Shomolu	21
20	Surulere	20

2.2 Machine Learning Algorithms and Error Assessment

In this study, the ML models were implemented through a four-step pipeline that included imputation, data normalisation, feature selection, and classification. Python, utilising essential libraries such as Scikit-learn, TensorFlow, Pandas, NumPy, and specific packages like extreme gradient boost and cat-boost, served as the programming language. The data collected during the field experiment encompassed information such as sex, age, height, weight, waist circumference, body mass index (BMI), glucose level, systolic blood pressure (SYS), and diastolic blood pressure (DIA). Additional variables were derived through mathematical manipulations of the initial data to enhance the stability and performance of machine learning algorithms. These included SYS/DIA, DIA/SYS, weight/age, and height/age ratios. In total, thirteen factors were used as predictors in classifying the measured COHb concentration, with missing values imputed using the median of the observations. The classification groups were composed of four different ranges of COHb percentages in the blood samples. The first group included cases with less than 1% COHb, the second group consisted of individuals with COHb levels between 1% and 1.8%, the third group covered values greater than 1.8% and up to 3.5%, while the final group included those with COHb percentages exceeding 3.5%.

A comprehensive evaluation of various classification algorithms was conducted, including random forest (RF), artificial neural network (ANN), naive bayes, logistic regression, support vector machines (SVM) (using kernels like linear, polynomial (poly), radial basis function (RBF), and sigmoid), k-nearest neighbours (KNN), extreme gradient boost (XGB), light gradient boost machine (LGBM), cat-boost, and bagging classifier. The selection of these algorithms was based on their established state-of-the-art performance in data analysis and their interpretability when employing feature-attribution techniques. Hyper-parameters were transformed using the Yeo-Johnson transformation to make variables more normally distributed, reducing skew in the raw data. Additionally, a data scaling approach normalised the limits between -1 and 1, in addition to the power transformation procedure in Scikit-learn. Further details on the training techniques for the ML algorithms can be found in Table 2.

Table 2: Training modalities of the ML algorithms in Python, where except for the mentioned features, the rest of the systems were left at default.

S/no	ML Model	Training Techniques
1	logistic regression	{penalty = l2, maximum iteration = 100}
2	mnlogit_mod	{maximum iteration = 1000}
3	SVM (linear, poly, RBF, sigmoid)	{probability = True, random state = 0}
4	naïve bayes	{Guassian_NB}
5	KNN	{neighbours = 3}
6	ANN	{various hidden layers (varied between 3 and 13), number of neurons (varied between 1 and 10), batch sizes = 265, maximum iteration = 1000, validation fraction = 0.02, random state= 0}
7	decision trees	{random state = 0}
8	RF	{n-estimators (varied between 1 and 10), random state = 0}
9	bragging classifier	{base estimator = rfc2, iteration = 15, max samples = 0.75, max features = 0.75, random states = 0}
10	XBG	{random state = 0}
11	LGBM	{random state = 0}
12	cat-boost	{verbose = 0, random state = 0}

A meticulous two-step process was adopted for model selection, training, and evaluation to mitigate the risk of overfitting. Initially, a stratified approach was employed to partition the dataset into a training set (72%) and a hold-out test set (28%). The entire training set was then utilised for model training and calibration, while the hold-out test set was used to assess the calibrated models. To evaluate the generalisability of models to new settings with limited data, an internal-external validation procedure was implemented using a bootstrap-based approach. This methodology facilitated the estimation of model performance when exposed to new data in different settings, offering valuable insights into their robustness and applicability beyond the original dataset.

Several metrics like accuracy, F1, precision, and confusion matrix were used to evaluate the performance of the machine learning models. A confusion matrix is a table that is often used to describe the performance of a classification model on a set of test data for which the true values are known. Accuracy measures how often the classifier correctly predicts. It is the ratio of the number of correct predictions and the total number of predictions. It is expressed as:

$$Accuracy = (TP + TN) / (TP + TN + FP + FN)$$

where TP is the true-positive, TN is the true-negative, FP is the false-positive, and FN is the false-negative. Precision explains how many of the correctly predicted cases actually turned out to be positive. It is defined as the number of true positives divided by the number of predicted positives. Its expression is:

$$Precision = TP / (TP + FP)$$

Recall or sensitivity explains how many of the actual positive cases we were able to predict correctly with our model. It is defined as the number of true positives divided by the total number of actual positives.

Recall = TP / (TP + FN)
The F1 score is the harmonic mean of precision and recall. It is maximum when precision is equal to recall.
F1 = 2 × (Precision × Recall) / (Precision + Recall)
Finally, the ROC-AUC, which means Receiver Operator Characteristic (ROC) - Area Under the Curve (AUC). While ROC is a probability curve that plots the TPR (true positive rate) against the FPR (false positive rate) at various threshold values and separates the signal from the noise, AUC is the measure of the ability of a classifier to distinguish between classes.

3.0 RESULTS AND DISCUSSION

3.1 Predictors Correlations

Following the coded procedure, the 516 individuals from the field experiment were automatically split into two sets: the training set, which included 254 individuals in the first group, 64 in the second group, 55 in the third group, and 3 in the fourth group; and the testing set, comprising 94, 24, 21, and 1 individual in the respective four groups. Although the extent by which a factor predicts depends on the given ML algorithm, COHb specifically exhibited minimal dependence on gender, waist circumference, and height but demonstrated a stronger correlation with glucose levels, DIA/SYS, and BMI. However, other factors that could significantly impact COHb prediction included weight, age, systolic blood pressure, and diastolic blood pressure (Fig 2). Furthermore, the derived variables exerted a more pronounced influence than individual predictors. For instance, while height and age had a limited impact on COHb, their combined effect contributed to a more accurate prediction.
The correlation coefficients between COHb and the predictors are consistently weak, with values generally falling below 10%, whether positive or negative (Fig 3). Furthermore, COHb displays a positive relationship with only four out of the thirteen variables, which include height, waist, SYS/DIA, and height/age. In order of importance, the five factors out of the lot that indirectly impact on COHb include glucose level, BMI, weight, sex, and DIA. The positive correlation values range from 0.0019 to 0.04, whereas the negative correlations range from -0.0045 to -0.089. Notably, the comparison between Figs 2 and 3 suggests that machine learning algorithms appear to rely more on negatively correlated predictors than the positive ones in most cases. For instance, a typical ML model exhibits a higher dependence on glucose, DIA/SYS, BMI, DIA, weight, and SYS, which have negative relationship with COHb compared to height, waist, SYS/DIA, and height/age.

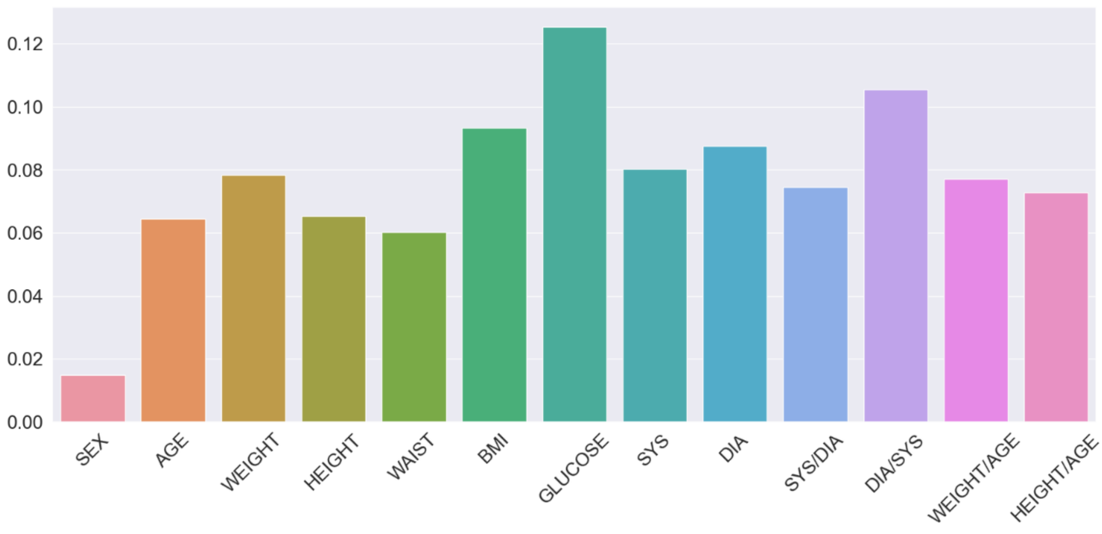

Figure 2: A typical impact of the predictors on the outcome of COHb from random forest

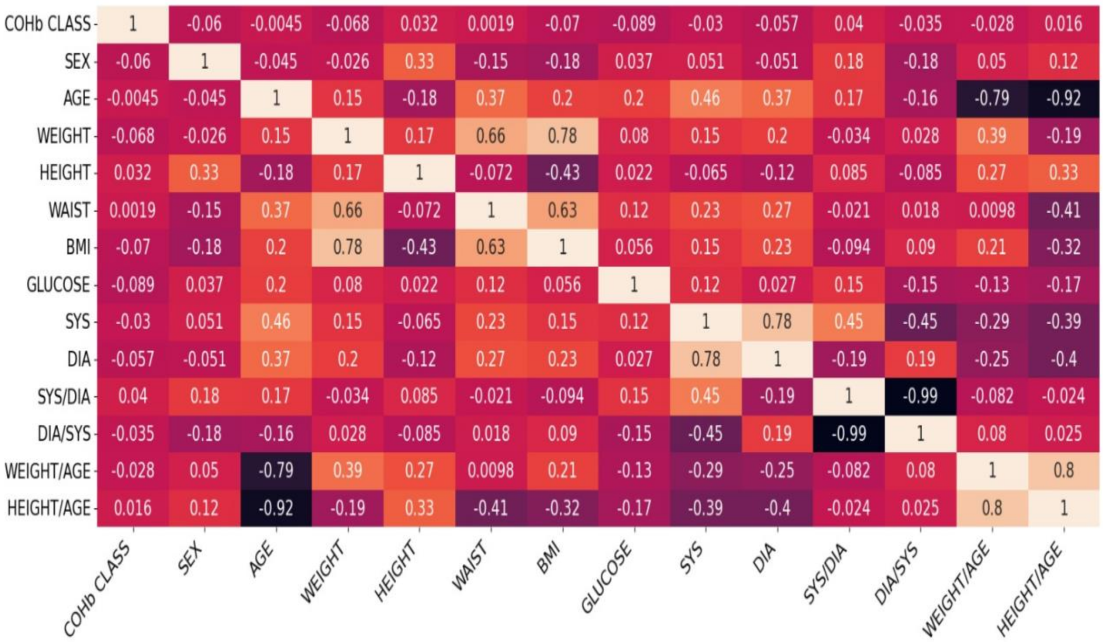

Figure 3: Correlation coefficient values between all the variables used in this study.

3.2 Retained ML Models

Since this is a classification problem designed to function as a diagnostic tool rather than an imaging task, ML algorithms achieving 100% in error functions like accuracy and precision during the training phase were considered prone to overfitting and subsequently excluded. The affected algorithms include decision trees, XBG, LGBM, and cat-boost. However, the optimal configurations for RF and the ANN were identified and retained. The random forest performed best with 10 n-estimators, while the artificial neural network, which, aside from the input and output layers, had seven hidden layers, each containing six neurons, demonstrated superior performance. Additionally, as the hold-out test section indicates generalisation capacity, models with low performance, scoring below a cutoff of 65% accuracy during this phase, were also excluded, although a score of 70% and above is deemed desirable.

At times, the error terms exhibited disparities among themselves by displaying different trends, especially when disregarding the overfitting models that were dropped, in which all error terms were consistently 100%. So, because the testing phase reflects responses to unfamiliar circumstances and generalisation capabilities, greater emphasis is placed on it than on the training phase. In the testing phase (Table 3), accuracy scores range between 70.00% and 65.00%, F1 scores fall within 63.00% and 53.13%, ROC-AUC values range between 67.48% and 49.91%, while precision varies between 63.58% and 44.93%. Throughout this phase, the ANN consistently achieves the top results in all error terms. However, while the log regression ML algorithm maintains the lowest grades in F1 and precision, it pairs with SVM-poly in having the lowest score in accuracy, while SVM-linear has the lowest value in ROC-AUC. Nevertheless, during the training phase (Table 4), accuracy ranges between 97.87% and 67.55%, F1 values range from 97.84% to 54.17%, ROC-AUC varies between 99.92% and 47.35%, and precision falls between 97.91% and 45.63%. While all the highest values pertain to random forest, all the lowest values do not necessarily belong to a single ML model.

Table 3: Error terms for the retained ML models at the testing phase

S/no	Model	Accuracy (%)	F1 (%)	ROC-AUC (%)	Precision (%)
1	Log regression	65.00	53.13	57.59	44.93
2	SVM-linear	67.14	53.94	49.91	45.08
3	SVM-poly	65.00	54.63	59.86	48.78
4	SVM-RBF	67.14	53.94	63.43	45.08
5	ANN	70.00	63.00	67.48	63.58
6	random forest	66.43	61.39	67.07	60.25
7	bragging classifier	65.71	57.77	61.00	60.92

Table 4: Error terms for the retained ML models at the training phase

S/no	Model	Accuracy (%)	F1 (%)	ROC-AUC (%)	Precision (%)
1	log regression	68.62	57.25	64.49	74.92
2	SVM-linear	67.55	54.47	47.35	45.63
3	SVM-poly	73.94	67.17	78.62	80.36
4	SVM-RBF	67.55	54.47	81.03	45.63
5	ANN	74.20	70.01	81.45	72.03
6	random forest	97.87	97.84	99.92	97.91
7	bragging classifier	88.03	87.10	99.38	88.57

Due to the skewness of the data in this study, where a greater number of individuals can be found in a given group compared to the others, a high score in error terms can be misleading. The confusion matrix (Fig 4) can help reveal the extent to which the groups with few individuals were predicted. Fig 4(a) presents the ideal confusion matrix for the problem at hand, and the comparison of the perfect situation with the rest reveals the inadequacy of the ML models. The worst scenario arose from models that failed to classify individuals into any other groups except placing everyone in the 1st class. The affected ML algorithms were the respective SVM with linear and RBF kernels (Figs 4(b) and 4(c)). The apt generalisation of the ANN model with the hold-out test set regarding the error terms is also evident in the confusion matrix. Given that members of the first class are 94 in number, the ANN (Fig 4(g)) and log regression (Fig 4(e)) algorithms predicted the closest to 94, identifying 91 persons. In the second group where 24 are expected, the closest to that is the random forest, followed by both ANN and bragging classifier (Fig 4(h)) at 5 and 2 persons, respectively. In the third group where 21 individuals belong, ANN (Fig 4(g)), random forest (Fig 4(d)), bragging classifier (Fig 4(h)), and SVM-poly (Fig 4(f)) placed 5, 3, 2, and 1 in it, respectively. However, in the last group with a lone individual, no ML model predicted it correctly.

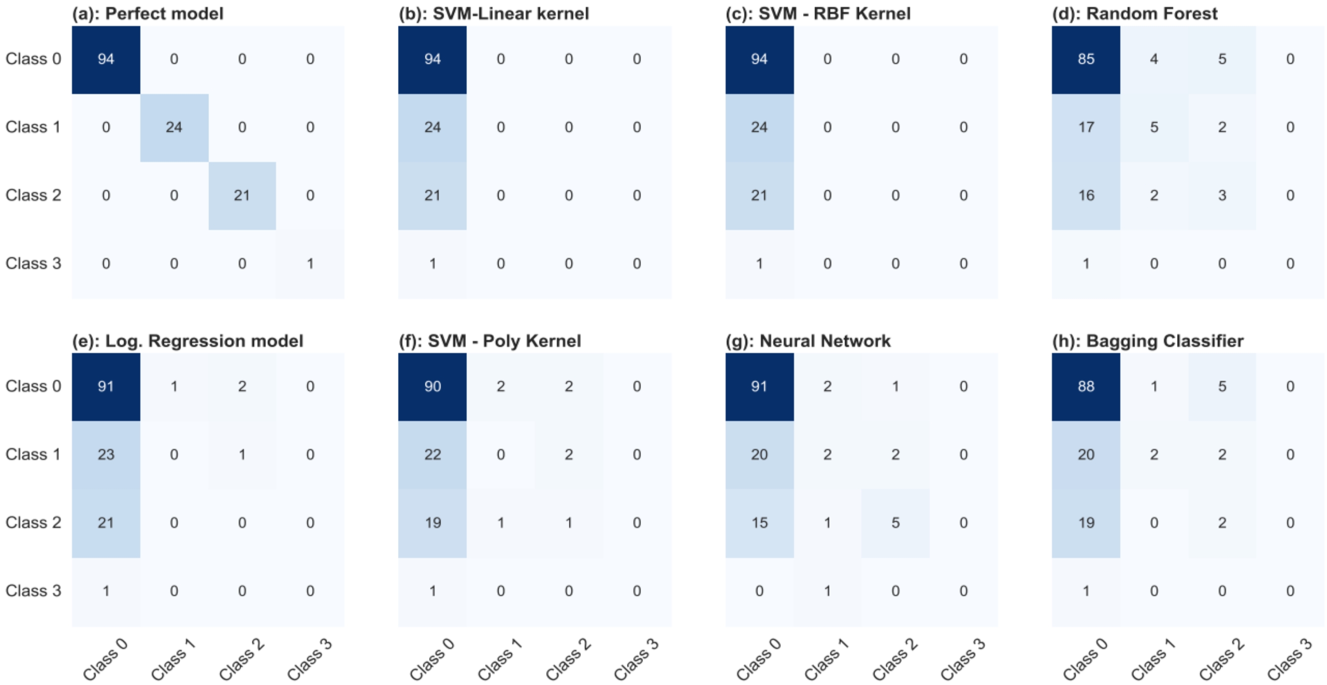

Figure 4(a-h): Multiple confusion matrixes for COHb classifications for individuals in Lagos State, Nigeria.

3.3 Discussion

Clearly, factors demonstrating a negative relationship with COHb respond appropriately to machine learning modelling. Glucose, which exhibits the highest negative correlation with COHb, represents the natural sugar content or a type of carbohydrate found in foods. Cells derive energy from glucose for survival. During food oxidation under various reaction steps, oxygen reacts with glucose to form carbon dioxide, water, and energy for the eventual production of ATP, as illustrated in Equation 1:

$$C_6H_{12}O_6 + 6O_2 = 6CO_2 + 6H_2O + Energy \quad (1)$$

As indicated in Equation 1, a relatively low glucose level suggests sufficient oxygen content in the body, supporting metabolism that releases carbon dioxide during respiration. In addition to tissue oxygenation, O_2 eliminates CO from haemoglobin based on their partial pressures, lifespan stages, or exposure time, concentrations of other gases such as nitric oxide, and other factors including body temperature and enzyme or protein actions (Crecelius *et al.,* 2015; Ryter *et al.,* 2006; Jones and Kennedy, 1982; Rodkey *et al.,* 1974). Given the role of glucose in cell metabolism and its implications for oxygen availability in blood cells, it is unsurprising that glucose significantly influences the prediction of COHb. However, while SYS/DIA and DIA/SYS share a 15% relationship with glucose (Fig 3), the latter, with a notably negative correlation, has a more pronounced effect on the prediction of COHb compared to the former, as depicted in Fig 2. Importantly, while SYS/DIA holds a correlation coefficient value of 4% with COHb, DIA/SYS, which predicts the molecule more accurately, has a correlation coefficient value of -3.5%. Moreover, the relatively significant impact of BMI or weight on COHb prediction may be attributed to a larger body's increased capacity to retain substantial amounts of food, fat, or glucose. Consequently, the inverse correlation between O_2 and CO allows machine learning algorithms to easily recognise glucose, BMI, and weight as robust predictors of COHb. Notably, these three factors—glucose level, BMI, and weight—exhibit the highest negative correlation coefficient values with COHb (Fig 3). Inevitably, maintaining a healthy diet or administering synthesised glucose to patients can positively influence the symptoms of CO poisoning.

Except for gender, factors that correlate relatively well with COHb, whether positively or particularly otherwise, are likely to have an impactful tendency to predict the molecule (Figs 2 and 3). Gender, as an exception to this trend, is likely due to its lack of variation, notwithstanding its comparatively high negative correlation with COHb. Although gender has little impact on diagnosing COHb concentration in this study, it could be essentially impactful under the same height and age with mathematical modelling (Oluwatosin *et al.,* 2019).

This study has several limitations, primarily stemming from its narrow scope. Although the study targeted individuals at risk of CO poisoning, such as roadside traders and open-market traders exposed to highly polluted air in Lagos, the participants were volunteers from the community who were neither hospitalised due to CO poisoning nor systematically exposed to CO gas. As a result, the results may not fully capture the experiences of all individuals who are unwell or exhibiting symptoms of elevated COHb concentration, despite some individuals being found to have elevated levels of CO poison in their bodies. Furthermore, the absence of laboratory animal testing limits the researchers' ability to comprehensively explore various possibilities, such as evaluating the associated risks of CO poisoning with different exposure times.

Given that a high concentration of CO significantly impacts vital organs like the heart and brain, individuals with pre-existing medical conditions require special attention. Furthermore, apart from factors like smoking habits, workplace incidents, and fire accidents, elevated CO exposure can occur during activities such as cooking with kerosene, wood, coal, charcoal, and cooking gas, as well as travelling in vehicles running on fossil fuels. It is noteworthy that the effects of CO poisoning tend to diminish after a few days with proper oxygen intake.

It is possible that if models like decision trees, XBG, LGBM, and cat-boost were expansively trained (Wang *et al.,* 2023), they would have been suitable for this exercise. This is because RF with 2 n-estimators (though not shown here) fell into such a category, achieving 100% accuracy, F1, ROC-AUC, and precision scores during the training phase. The RF model used in the study performed optimally during the training section yet could not sustain such performance during the testing phase, indicating some degree of overfitting. Future studies could evaluate the impacts of tuning system parameters and consider hybridising either ML with mathematical models or optimisers like genetic algorithm and particle swamp optimisation to ascertain if there is a better approach (Obot, 2024).

Given the escalating adoption of machine learning algorithms in the medical sciences, a study of this nature holds invaluable potential for applications in the clinical treatment of CO-related issues, such as burn treatment, sleeping disorders, mental health problems, and respiratory diseases. Notably, there is a growing trend in the development of wearables—medical devices that integrate artificial intelligence techniques with sensors for non-invasive diagnosis. The significance of wearables is particularly pronounced in remote regions with limited access to advanced clinical examination facilities.

Furthermore, the inherent unpredictability in the recurrence of medical conditions, coupled with constraints such as limited hospital facilities, bed space, and manpower, as well as frequent changes in weather, lifestyles, and dietary patterns, underscores the necessity for physicians to remotely monitor their patients. Wearables or oximeters equipped with alert systems can play a crucial role in providing timely alarms whenever necessary (Lee *et al.,* 2021; Adjiski *et al.,* 2019; Wu, 2019).

Those with a relatively high percentage of COHb in their blood samples are at risk of CO poisoning and its effects, but they fall into the 3rd and 4th groups with relatively few numbers. As such, ML models are somewhat incapacitated in handling the issue at hand because enough data is required for them to learn. Unfortunately, only one person falls into the fourth class (with a percentage of COHb in the blood sample higher than 3.5) in the hold-out test set, and the only ML model that could reveal that the individual does not belong to the first group is the ANN (Fig 4(g)). To err on the side of safety, precautionary measures or treatment may be necessary when a subject does not fall into the first group when using the recommended ANN scheme.

4.0 CONCLUSION

To address the constraints associated with time wastage, high costs, painful body piercing, and inaccuracies linked to clinical and mathematical methods for determining carboxyhaemoglobin, this study employed a comprehensive use of machine learning algorithms to classify the complex levels in the densely populated Lagos State of Nigeria for a possible development of wearable. Thirteen inputs, namely sex, age, height, waist circumference, BMI, glucose, diastolic BP (DIA), systolic BP (SYS), DIA/SYS, SYS/DIA, weight/age, and height/age, were utilised to classify four levels of COHb. The machine learning algorithms included artificial neural networks (with various hidden layers), support vector machines (employing different kernels such as linear, radial basis function, sigmoid, and poly), k-nearest neighbour, decision trees, random forest (with varied n-estimators), naive bayes, and well-known classifiers like bagging, XGB, LGBM, and cat-boost. Using a 72:28 ratio to partition the data into training and testing groups, statistical measures for evaluating the ML models included accuracy, confusion matrix, precision, and F1. Models with 100% accuracy during the training phase and those scoring below 65% during the testing phase were excluded. The conclusions in this study are as follow:

i. Machine learning (ML) demonstrates the ability to identify factors strongly linked to oxygen, particularly in the context of oxidation, such as glucose and BMI, when modelling COHb. Additionally, this capability could be extended to capture the inherent relationship between the quantity of carbon monoxide and oxygen in human blood. Therefore, administration of glucose may be helpful in the management of CO poisoning.
ii. Negatively correlated variables wield a more pronounced impact than positively related factors in ML-based COHb modelling.
iii. In instances where a predictor lacks adequate variation, its contribution to ML modelling is minimal, regardless of the level of negative correlation with COHb.
iv. Frequently, the influence of a variable derived through the mathematical manipulation of two variables surpasses that of each individual variable.
v. An artificial neural network structure with 7 hidden layers, each comprising 6 neurons, demonstrated optimal performance during testing and is consequently recommended for COHb analysis in the Python environment for the city of Lagos, Nigeria. This model stood out with a 70% accuracy score during testing and yielded the most favourable confusion matrix output.

ACKNOWLEDGEMENTS

The authors are grateful to health educators in all twenty local government areas for their support during the time of this study.

Funding The study was funded by the CRC Grant of the University of Lagos, Nigeria (CRC. 2015/24).

REFERENCES

Abbey, M., Amadi, S.C., Ocheche, U.S., Wekere, F.C.C., Altraide, B.O., Oloyede, O.A., Inimgba, N.M., and Akani, C.I. (2022). Maternal exposure to carbon monoxide in the first trimester (7-13^{+6} weeks) of pregnancy in the core Niger Delta. International Journal of Reproduction, Contraception, Obstetrics and Gynecology 11, 1839-1847.

Adedokun, O.A., and Owoade, O.K. (2019). Source apportionment of vehicular exhaust emissions in Lagos, Nigeria. Sustainable Cities and Society, 47, 101484.

Bemtgen, X., Rilinger, J., Jäckel, M., Zotzmann, V., Supady, A., Benk, C., Bode, C., Wengenmayer, T., Lother, A., and Staudacher, D.L. (2021). Admission blood glucose level and outcome in patients requiring venoarterial extracorporeal membrane oxygenation. Clinical Research in Cardiology 110, 1484–1492.

Bickler, M.P., and Rhodes, L.J. (2018). Accuracy of detection of carboxyhemoglobin and methemoglobin in human and bovine blood with an inexpensive, pocket-size infrared scanner. PloS one 13, e0193891.

Crecelius, A.R., Kirby, B.S., and Dinenno, F.A. (2015). Intravascular ATP and regulation of blood flow and oxygen delivery in humans. Exercise and Sport Sciences Review 43, 5–13.

Guarnaccia, C., Cerón Bretón, J.G., Quartieri, J., Tepedino, C. and Cerón Bretón, R.M. (2014). An application of time series analysis for forecasting and control of carbon monoxide concentrations. International Journal of Mathematical Models and Methods in Applied Sciences 8, 505–515.

Hampson, N.B., and Hauff, N.M. (2008). Carboxyhemoglobin levels in carbon monoxide poisoning: do they correlate with the clinical picture? The American Journal of Emergency Medicine 26, 665–669.

Hanis, T.M., Islam, M.A., and Musa, K.I. (2022). Diagnostic accuracy of machine learning models on mammography in breast cancer classification: a meta-analysis. *Diagnostics 12*, 1643.

Huzar, T.F., George, T., and Cross, J.M. (2013). Carbon monoxide and cyanide toxicity: etiology, pathophysiology, and treatment in inhalation injury. Expert Review of Respiratory Medicine 7, 159–170.

Isa, R.O., Odigure, J.O., Okafor, J.O., Abdulkareem, A.S., Abdulfatai, J., and Afolabi, A.S. (2013). Mathematical modeling of road traffic carbon monoxide pollutant: a case study of Minna Metropolis, Nigeria. International Review of Chemical Engineering 5, 57– 65.

Jeon, S.B., Sohn, C.H., Seo, D.W., Oh, B.J., Lim, K.S., Kang, D.W., and Kim, W.Y. (2018). Acute brain lesions on magnetic resonance imaging and delayed neurological sequelae in carbon monoxide poisoning. JAMA Neurology 75, 436–443.

Jones, D.P., and Kennedy, F.G. (1982). Intracellular oxygen supply during hypoxia. American Journal of Physiology 243, C247.

Lee, I., Park, N., Lee, H., Hwang, C., Kim, J.H. and Park, S., (2021). Systematic review on human skin-compatible wearable photoplethysmography sensors. Applied Sciences *11*, 2313.

Mattiuzzi, C., and Lippi, G. (2020). Worldwide epidemiology of carbon monoxide poisoning. Human and Experimental Toxicology 39, 387–392.

Maynard, R.L., and Robert, W. (1999). Carbon monoxide. In Air pollution and Health (pp. 749-796). Academic Press.

McNair, H.M., Miller, J.M., and Snow, N.H. (2019). Basic gas chromatography. John Wiley & Sons.

Murphy, K.P. (2012). Machine learning: a probabilistic perspective. The MIT Press.

Obot, N.I. (2024). Regression models and hybrid intelligent systems for estimating clear-sky downward longwave radiation in Equatorial Africa. Journal of Earth and Space Physics 49, 143–159.

Obot, N.I.**,** Olubgon, B., Humphrey, I., and Akeem, R.A. (2023). Equatorial all-sky downward longwave radiation modelling. Communication in Physical Sciences 9, 111–124.

Oghenejoboh, K. M., and Adiotomre, K. O. (2012). Finite volume modelling of carbon monoxide from vehicular emissions at Oshodi market in Lagos Metropolis of Nigeria. Journal of Computations and Modelling *2*, 55–77.

Oluwatusin, O., Orolu, K., Ekun, O., Akinloye, O., and Oyediran, O.O. (2019). Mathematical model to evaluate the effect of carbon monoxide exposure as a function of gender, age, and height. Journal of Applied Sciences and Environmental Management *23,* 799–803.

Oluwole, O., Arinola, G., Ologun, G., and Fatusi, A. (2016). The impact of air pollution on respiratory and cardiovascular health in Lagos, Nigeria. Journal of Environmental and Public Health 8086069.

Piantadosi, C.A. (2004). Carbon monoxide poisoning. Undersea and Hyperbaric Medicine 31, 167–177.

Prockop, L.D., and Chichkova, R.I. (2007). Carbon monoxide intoxication: an updated review. Journal of the Neurological Sciences 262, 122–130.

Rodkey, F.L., O'Neal, J.D., Collison, H.A., and Uddin, D.E. (1974). Relative affinity of hemoglobin S and hemoglobin A for carbon monoxide and oxygen. Clinical Chemistry 20, 83–84.

Ryter, S.W., Alam, J., and Choi, A.M., (2006). Heme oxygenase 1/carbon monoxide: from basic science to therapeutic applications. Physiological Reviews *86*, 583–650.

Samuel, J.M., Kahl, J.H., Zaney, M.E., Hime, G.W., and Boland, D.M. (2021). Comparison of spectrophotometric methods for the determination of carboxyhemoglobin in post-mortem blood. Journal of Analytical Toxicology 45, 885–891.

Shi, B., Zhou, T., Lv, S., Wang, M., Chen, S., Heidari, A.A., Huang, X., Chen, H., Wang, L. and Wu, P. (2022). An evolutionary machine learning for pulmonary hypertension animal model from arterial blood gas analysis. Computers in Biology and Medicine 146, 105529.

Sobamowo, M. G. (2016). Prediction of the effects of combustion–generated pollutant on human health: mathematical models and numerical solutions. *Iranian (Iranica) Journal of Energy and Environment 7*, 233–242.

Wu, A.H., (2019). “On vivo” and wearable clinical laboratory testing devices for emergency and critical care laboratory testing. Journal of Applied Laboratory Medicine *4*(2), 254–263.