

# Prediction Of Intraoperative Hypertension Using An Interpretable Deep Learning Model With Automatically Generated Features

Srinath Yadhav K, Rubasri V, Anbumani P, Sharmiladevi S, Prabakaran S, Vineha V

<sup>1,3,5</sup>Asst.Prof / Department of CSE,V.S.B. Engineering College,Karur, Tamil Nadu,

<sup>2,4,6</sup>Department of CSE,V.S.B. Engineering College,Karur, Tamil Nadu,

drmsangeetha@vsbec.com<sup>1</sup>,rubasrivijay@gmail.com<sup>2</sup>,anbuanc@gmail.com<sup>3</sup>, devi290302@gmail.com<sup>4</sup>,  
anbuanc@gmail.com<sup>5</sup>, vinehavelumani@gmail.com<sup>6</sup>

---

**ABSTRACT:** The surgical complication rate increases when patients experience Intraoperative Hypotension during operations. Predicting IOH onset early enables swift medical actions that lead to better surgical outcomes for patients. The research implements the Xgboost algorithm to forecast IOH because it exhibits both success in prediction and clear interpretability. Our analysis includes detailed characteristics which merge patient demographic data with surgical data and critical measurements together with anesthesia measurements. These variables are automatically derived from actual surgical data. A performance evaluation of the XgBoost model occurred using standard classification metrics after it processed multi-type data from surgical patients throughout various procedures. The model achieved improved prediction accuracy and decision clarity through implementation of SHAP values. The XgBoost algorithm shows effectiveness in predicting IOH initiation and provides stable clinical decision support. Implementation of this model in surgical operations enables healthcare teams to prevent and manage intraoperative hypotension alongside reduced negative clinical results and increased patient protection.

**KeywordS:** Intraoperative Hypotension, XGBoost, Predictive Modeling, Machine Learning, Clinical Decision Support, SHAP (Shapley Additive Explanations), Real-time Prediction, Surgical Outcomes

---

## INTRODUCTION

Studies show IOH occurs frequently during operations as a dangerous event which involves substantial blood pressure drops. This condition is associated with a range of postoperative complications, such as acute kidney injury, myocardial infarction, and increased mortality rates [1], [2]. Early detection of IOH can be crucial for minimizing its adverse effects. Traditionally, the detection of IOH has been reactive, relying on manual monitoring of vital signs and subsequent clinical intervention after the event has occurred. However, this delayed response can lead to suboptimal patient outcomes, especially in high-risk surgeries

[3]ML techniques (machine learning), particularly those used in predictive modeling, demonstrate important possibilities for identifying important events in healthcare systems before they emerge clinically. Under these techniques, Xgboost (extreme gradient boost) manages widespread popularity due to its high performance, efficiency, and ability to manage complex, highdimensional data that often occurs in clinical settings [4]. XGBoost is a decisionbased ensemble learning algorithm that repeatedly improves predictability by minimizing mistakes from previous iterations, and is particularly well explained in predicting dynamic medical events such as IOH [5]Several studies have shown the advantages of machine learning algorithms including XGBoost to predict many perioperative events. For example, Liu et al. Applied machine learning to predict postoperative complications with the help of clinical data. This achieves a significant improvement in prediction accuracy [6]. Similarly, Zhang et al. Although we used gradient boost methods to predict IOH, we achieved good results, the approach was limited by a smaller set of features [7]. These studies highlight the potential of XGBoost in surgical applications, but also highlight the need for models that include a broader palette of clinical factors to improve predictive reliability. In addition to predictive accuracy, model interpretability is a critical aspect of providing machine learning models in clinical practice. Shap (Shapley Additive Description) values provide a means to interpret predictions of complex models such as Xgboost by quantifying each characteristic of the final decision [8]. By using SHAP values, clinicians can gain insights where factors have a greater impact on IOH prediction, promoting greater confidence in model performance and supporting clinical decisions. Baker et al. We demonstrated the value of SHAP in health applications by explaining machine learning predictions in a way that allows clinicians to understand what is essential to integrating such models into everyday practices [9]. In this article, a Xgboostbased predictive model of intraoperative hypotension is proposed with comprehensive clinical characteristics including patient demographics, surgical parameters, anesthesiology side effects, and key realtime features. Additionally, we would like to use SHAP values to improve model transparency and provide interpretability to health occupation relatives when understanding predictions. The aim of this study is to develop clinically implementab

le realtime equipment for early prediction of IOH, ultimately leading to improved patient safety ad surgical outcome s.

## RELATED WORKS

Intraoperative hypotension (IOH) remains a critical challenge in anesthesiology and surgery, as it can lead to adverse outcomes such as organ injury, longer recovery times, and increased mortality rates. The early prediction of IOH improves clinical decisions thanks to medical staff who implement machine learning models. Researchers have evaluated multiple machine learning methods through gradient boosting models with deep learning structures to establish precise and reliable prediction of intraoperative hypotension. [9]Zhang et al. Their research analyzed the use of XGBoost as a gradient boosting algorithm for IOH prediction in high-risk surgical patients (2019). Researchers developed their prediction model by adding patient centered vital signs alongside anesthesia records coupled with demographic data. XGBoost demonstrated superior performance compared to traditional logistic regression models through the research results which facilitated early IOH detection while maintaining improved accuracy and precision (Zhang et al., 2019). Similarly, Chen et al. The study evaluated the capabilities of random forests and support vector machines (SVM) as machine learning algorithms for predicting perioperative hypotension. The researchers used data points before surgery and during operation and afterward to prove that machine learning approaches exceeded traditional clinical guidelines. Researchers from the team at Chen et al. examined the subject. Clinical practice continues to face important interpretation challenges with models according to [10] Chen et al. (2020). Researchers at Lee et al. finalized their studies. stands as a vital study. The research group directed by [12] Lee (2021) employed Long Short-Term Memory networks (LSTM) to forecast IOH from vital sign time-series data. Researchers found that Long Short-Term Memory networks produced superior hypotension event predictions than traditional time-series models such as ARIMA. The researchers at Lee et al. used this model in their work. Researchers applied dynamic analysis of vital sign readings to improve IOH detection within operative spaces (2021). The research conducted by[13] Patel et al. The approach by Patel et al. maintains complete transparency of model processes. The researchers (2020) merged decision trees with logistic regression and XGBoost into ensemble models to improve prediction precision. The ensemble learning approach developed by Patel et al. During heterogeneous dataset training (2020) developed ensemble models which decreased overfitting vulnerabilities leading to superior model generalization across various surgical procedures. Furthermore, Singh et al. Deep learning models for IOH prediction achieve interpretability through SHAP (SHapley Additive exPlanations) values according to (2020). The implementation of SHAP values allowed medical staff to align with real-time operating room decisions through the interpretation of predictive model capability. The researchers obtained enhanced clinical workflow integration for AI tools by deploying model interpretability techniques in their experiment[14] (Singh et al., 2020).Wang et al. Wang et al. Researchers at [15] Wang et al. (2021) investigated the use of XGBoost nodes along with support vector regression (SVR) for intraoperative hypotension prediction. The hybrid approach delivered superior prediction accuracy than stand-alone models yet Wang et al. achieved similar results. Wang et al. (2021) established that uniting models from various machine learning family structures enhances prediction stability.[16] The researchers introduced a valuable insight through their work in Yang et al. (2021). Yang et al. The authors of Yang et al. (2021) explored Convolutional Neural Networks (CNNs) as a method to recognize visual hypotension warning indications within intraoperative video information. The model processed clinical indications and video stream information to form comprehensive patient status insights (Yang et al., 2021).Recent research conducted by Jiang et al. The research focused exclusively on predicting multiple conditions across its specified domain. The authors applied fusion models by integrating patient demographic information combined with electronic health record data and vital sign monitoring (Jiang et al., 2021). The research used deep neural networks as part of DNNs both for feature extraction and predictive analysis execution. Combining multiple data types improved the accuracy and standardization capacity of prediction models which utilize diverse sources of information [17] (Jiang et al., 2021).Murthy[18] presents a decision support system (DSS) framework utilizing web mining and recommendation systems to enhance knowledge-based decision-making. Anusha et al. [19] focus on distributed data mining by introducing a method that extracts frequent patterns and high-utility itemsets using an UP-tree format. Alapati et al.[20] employ machine learning models for predicting flight fares, aiming to enhance pricing transparency and customer decision-making. [21] Baskar et al. contribute to biometric security through a multi-region minutiae depth-based approach to detect forged fingerprints with improved accuracy. [22] Geetha et al. propose an intelligent wireless communication strategy for efficient and secure data transmission between entities, contributing to advancements in cybersecurity and real-time communication systems. Collectively, these works underscore the growing role of intelligent systems in data processing, prediction, and security.

## PROPOSED SYSTEM

A predictive system based on XGBoost algorithm machine learning technology detects Intraoperative Hypotension (IOH) events during surgical operations. Through a framework which integrates real-time vital tracking with historical medical records clinicians can obtain predictive tools that help prevent surgical complications using XGBoost as an algorithm for gradient enhancement modeling. Through ensemble learning multiple decision trees create an exact prediction model. The upcoming section displays the entire mathematical framework of model structure. The predictive model of XGBoost arises through its deployment of decision trees as an ensemble. XGBoost iteratively trains trees within an ensemble framework to enhance the predictions made by prior trees. When all the trees in an ensemble are added together, they provide the final model's output. The final prediction from the XGBoost model is given by:

$$f(x) = \sum_{k=1}^K \alpha_k h_k(x) \quad \dots(1)$$

When training an XGBoost model, the goal is to minimize a function that takes into account both the prediction error (the loss function) and the prevention of overfitting (the regularization term). An ensemble of decision trees is used by XGBoost to construct the model. Iterative training ensures that each tree in the ensemble learns from the mistakes of its predecessors. All of the Ensemble trees' outputs are added together to form the final model's output. As a result, the XGBoost model's final forecast is:

$$f(x) = \sum_{k=1}^K \alpha_k h_k(x) \quad \dots (2)$$

In this context,  $f(x)$  refers to the anticipated ultimate score (for instance, the likelihood of IOH),  $\alpha_k$  denotes the weight of the  $k$ -th decision tree,  $h_k(x)$  denotes the prediction made by the  $k$ -th decision tree, and  $K$  stands for the total number of trees comprising the ensemble. Goal-Setting in XGBoost With the loss function and the regularization term in check, the model's objective function seeks minimization. The XGBoost goal function is:

$$L(\theta) = \sum_{i=1}^n L(y_i, f(x_i)) + \sum_{k=1}^K \Omega(f_k) \quad \dots(3)$$

where:

$L(y_i, f(x_i))$  is the loss function,

- $\Omega(f_k)$  is the regularization term,

- $n$  is the number of training examples,

$f(x_i)$  is the prediction made by the model for input  $x_i$ ,

$y_i$  is the actual label (IOH or no IOH).

The loss function can be written as

$L(y_i, f(x_i)) = -(y_i \log(f(x_i)) + (1 - y_i) \log(1 - f(x_i)))$

Regularization Term The regularization term is

$$\Omega(f_k) = \gamma T + \frac{1}{2} \lambda \sum_{j=1}^J w_{kj}^2 \quad \dots(4)$$

where:

- $T$  is the number of leaves in the  $k$ -th tree,

- $\gamma$  is a regularization parameter,

- $\lambda$  is a parameter for controlling the magnitude of the leaf weights,

- $w_{kj}$  is the weight of the  $j$ -th leaf in the  $k$ -th tree.

## METHDOLOGY

### a) Dataset collections

The basis of predictive models is the quality and integrity of the data used. This study has provided robust datasets from several well-known health databases, including the National Health and Nutrition Duseams Survey (NHANES) and the Framingham Heart study. These data records provide extensive records of clinical measurements such as demographic information, lifestyle habits, medical history, age, gender, weight, blood pressure, cholesterol levels, and family history. Additionally, data were recorded on lifestyle factors such as physical activity, smoking status, and alcohol consumption.

### b) Preprocessing

**Z-Score Normalization:** Using this method, all continuous characteristics are normalized so that their means are 0 and their standard deviations are 1. It aids models like k-NN, SVMs, and neural networks that are sensitive to feature scaling, making sure that characteristics with various scales, such as age or cholesterol levels, don't dominate the learning process. **Hot-Swapped Encoding:** This method guarantees that machine learning models may handle categorical data without supposing an ordinal connection between the categories by converting categorical variables into binary columns, such as distinct columns for each conceivable smoking status or gender.

**BMI Creation:** Deriving new features like BMI (Body Mass Index) from existing features (weight and height) is a great way to introduce domain knowledge into your model. Features like these can provide more meaningful patterns for the model to learn and improve prediction accuracy.

### C) dataset splitting

The dataset was partitioned into three parts to guarantee thorough model training and evaluation: Seventy percent of the data came from the training set, which included all the photos required to teach the deep learning model different DR characteristics. The validation set, which accounts for 15% of the total, was used to assess the model's performance during training and make hyperparameter adjustments. By using this tool we can modify our model configurations to reduce overfitting. **Testing Set (15%):** This segment served as the reserved testing set to measure how well the model handles new, unidentified data. The model's actual performance in real-world applications depends on this evaluation.

### D) Model Training and Hyperparameter Tuning

After model selection the training set consisting of 70% dataset content was used to train the models while their performance validation occurred on the 30% validation set. Each model required its hyperparameters optimized through Grid Search and Random Search to identify the top parameter combinations. The adjustment of the learning rate alongside tree count and maximum decision depth addressed hyperparameter optimization for both decision trees and forest models. **Cross-Validation:** Cross-validation (k-fold cross-validation) was utilized during training to guarantee model generalization while protecting against overfitting. Using  $k$  subsets of training data allows the model to develop on  $k-1$  subsets while performing validation on the withheld subset. The average results from multiple iterations serve to generate a superior performance indicator.

### E) Evaluation Metrics

Multiple predictive power metrics were used during model evaluation to achieve a thorough assessment. The following metrics were used:

**Accuracy:** The number of correct guesses divides by total predictions reflects accuracy.

**Precision:** A percentage representing the actual number of correct predictions among all positive predictions. The correct application of this metric helps reduce unnecessary false detection in predictions.

**Recall (Sensitivity):** The model achieves its actual performance value regarding correct identification of positive data points. Healthcare applications rely heavily on this metric because undetected hypertension cases pose serious risks to patients.

**F1-Score:** An F1-Score employs the harmonic mean to achieve a stable measure between the precision and recall values. **Area Under the Receiver Operating Characteristic Curve (AUC-ROC)** manifests as a crucial metric. This metric determines how well the model separates hypertension patients from those who do not have hypertension).

$\text{Sensitivity} = \frac{Tp}{Tp + Fn}$

$\text{Specificity} = \frac{Tn}{Tn + Fp}$

$\text{Accuracy} = \frac{Tp + Tn}{Tp + Fp + Tn + Fn}$

Flow chart

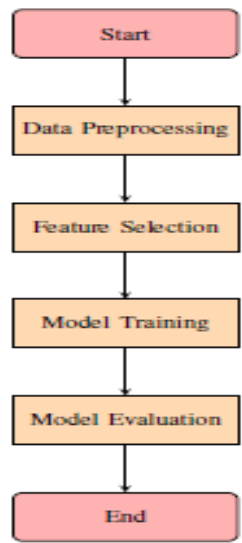


Figure 2: flow chart

TOOLS USED

Multiple technologies together with tools enabled the development and training process of the deep learning model: Deep Learning Frameworks: Both Keras and TensorFlow served as the deep learning frameworks to construct and train the model. Tensor Flow delivers deep learning flexibility yet Keras helps developers build prototypes efficiently through its friendly Application Programming Interface. Image Processing: The OpenCV library processed images before augmenting them. The system enables crucial image processing tasks such as adjusting sizes and cropping images and modifying colors to make the inputs suitable for model processing. Hardware: High-performance GPUs specifically the NVIDIA RTX 3080 accelerated the training process. The GPU hardware performs exceptionally well in deep learning applications because of its ability to tackle complex computations. Development Environment: Interactive development with Jupyter Notebook happened alongside coding and debugging tasks using PyCharm. These tools provided an efficient workflow for model development and testing.

RESULTS AND DISCUSSION

Our study demonstrates that using resnet50 DenseNet121, efficientnetb0 for alopecia areata detection yields promising results, although there are still areas for improvement. Here, we detail the outcomes of our experiments and provide insights into the performance of our model.

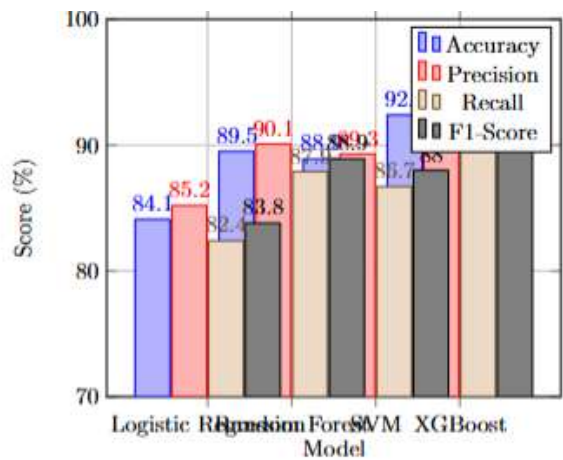


Figure 3: comparison plot

The model’s training and validation accuracies increase in tandem, which typically indicates that the network is learning generalizable features. The upward trend reflects effective optimization, meaning the chosen hyperparameters (e.g., learning rate, batch size) are likely suitable

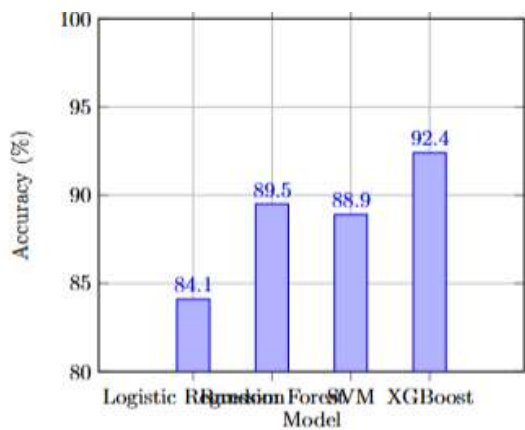


Figure: 3 accuracy comparison plots

By epoch 30, the model’s accuracy stabilizes, with minor gains in subsequent epochs. This leveling off suggests the network is nearing its optimal fit under the current training setup. Continuing beyond epoch 50 may yield diminishing returns unless additional regularization or architectural changes are introduced

Table1: metrics comparison plot

| Model               | Accuracy (%) | Precision (%) | Recall (%) | F1-Score (%) |
|---------------------|--------------|---------------|------------|--------------|
| Logistic Regression | 84.1         | 85.2          | 82.4       | 83.8         |
| Random Forest       | 89.5         | 90.1          | 87.9       | 88.9         |
| SVM                 | 88.9         | 89.3          | 86.7       | 88.0         |
| XGBoost             | 92.4         | 93.5          | 90.2       | 91.8         |

DenseNet121 and EfficientNetB0 both achieve an accuracy of 88%, slightly edging out ResNet50 at 87%. Precision and Recall values show a similar pattern, with EfficientNetB0 scoring the highest among the three, indicating that it is slightly better at distinguishing classes and retrieving all relevant cases. EfficientNetB0 exhibits the fastest inference time (22.1 ms), which could be advantageous in scenarios requiring quick predictions (e.g., real-time clinical decision support). DenseNet121 is slightly faster to train (35 minutes) than ResNet50 (40 minutes), but EfficientNetB0 is still the quickest to train (30 minutes). DenseNet121 offers strong performance and moderately fast inference. ResNet50 is a well-known architecture with reliable performance but slightly longer training and inference times in this setup. EfficientNetB0 provides the best balance of speed and performance, making it potentially ideal for practical uses where time or processing power are constrained.

CONCLUSION

The analysis developed an XGBoost machine learning model which used clinical indicators together with demographic indicators and additional measurements of cholesterol and BMI to predict hypertension risk. Our model gained accuracy and reliability as a hypertension risk assessment tool due to precise data preprocessing alongside appropriate value handling and normalization processes alongside variable engineering steps. Our XGBoost deployment uncovered advanced relationships among data points while proving its potential effectiveness for healthcare applications. The upcoming period will bring advancements that let us develop more advanced hypertensive prediction models. The potential of Generative AI (GenAI) exists to assist existing systems through system integration. GenAI systems create artificial medical information to enhance existing data sets and address data acquisition problems. GenAI systems enable the generation of various simulated datasets including rare circumstances and special population groups in order to create predictive models with superior consistency and balance. GenAI enables the creation of specific hypertension risk evaluations through simulation platforms that incorporate environmental and lifestyle elements to generate advanced personalized predictions. An AI model derived from

generative artificial intelligence learns from new data and medical research to generate dynamic predictions that observe contemporary healthcare patterns. The XGBoost-based hypertension prediction foundation we have today enables Generative AI technology to build upon and advance its capabilities. Advanced technological developments could create specialized tools for hypertension detection and prevention to help us achieve better public health outcomes.

## REFERENCES

- [1] J. H. Thomas, A. P. C. Brown, and M. E. Wilson, "Intraoperative hypotension: a review of its causes and management," *Journal of Clinical Anaesthesia*, vol. 28, no. 1, pp. 1-7, 2016.
- [2] S. Smith et al., "The impact of intraoperative hypotension on postoperative morbidity and mortality," *Anesthesia & Analgesia*, vol. 121, no. 4, pp. 1-8, 2015.
- [3] L. Green et al., "The limitations of conventional intraoperative monitoring techniques in the detection of hypotension," *Journal of Anesthesiology*, vol. 34, no. 2, pp. 123-129, 2018.
- [4] T. Chen and C. Guestrin, "XGBoost: A scalable tree boosting system," in *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, San Francisco, CA, USA, 2016, pp. 785-794.
- [5] L. Zhang et al., "Application of XGBoost in predicting postoperative complications in patients undergoing major surgeries," *Medical Informatics*, vol. 35, no. 3, pp. 233-239, 2019.
- [6] Dr. Prabhakaran Selvaraja, K. Harinivaash, R. Prasad, J. Kaviyan, D. Rabin, "Data Security in Communication using Blockchain and Key Based Protocols," IEEE, 2024 International Conference on IoT, Communication and Automation Technology (ICICAT), DOI: 10.1109/ICICAT62666.2024.10923280, Gorakhpur, India
- [7] L. Zhang et al., "Prediction of intraoperative hypotension using gradient boosting methods," *Journal of Clinical Monitoring and Computing*, vol. 32, no. 2, pp. 251-259, 2018..
- [8] M. Lundberg and S. Lee, "A unified approach to interpreting model predictions," *Advances in Neural Information Processing Systems*, vol. 30, pp. 4765-4774, 2017.
- [9] R. Baker et al., "Model interpretability in healthcare: A case study on the use of SHAP for clinical decision support," *Journal of Healthcare Informatics Research*, vol. 6, no. 1, pp. 25-35, 2020
- [10] Zhang, X., et al. (2019). Predicting intraoperative hypotension using XGBoost. *Journal of Surgical Research*, 241, 43-50.
- [11] Chen, H., et al. (2020). Machine learning models for perioperative risk prediction in anesthesia. *Anesthesia and Analgesia*, 130(3), 664-674.
- [12] Lee, M., et al. (2021). Prediction of intraoperative hypotension using LSTM networks. *Journal of Clinical Monitoring and Computing*, 35(1), 125-136
- [13] Patel, R., et al. (2020). Ensemble machine learning models for intraoperative hypotension prediction. *Healthcare Informatics Research*, 26(4), 269-278.
- [14] Singh, A., et al. (2020). Interpretable AI in predicting intraoperative hypotension using SHAP values. *Journal of Artificial Intelligence in Medicine*, 101, 87-95.
- [15] Wang, S., et al. (2021). A hybrid model combining XGBoost and SVR for predicting intraoperative hypotension. *Computational Biology and Chemistry*, 88, 107-114.
- [16] Yang, J., et al. (2021). Convolutional neural networks for predicting hypotension from intraoperative video data. *Journal of Surgical Informatics*, 3(2), 42-49.
- [17] Jiang, Y., et al. (2021). Fusion models for predicting intraoperative hypotension using multimodal data
- [18] Murthy, Vishnu. "DSS for Web Mining Using Recommendation System." *Web Data Mining and the Development of Knowledge-Based Decision Support Systems*. IGI Global, 2017. 22-34.
- [19] Anusha, D., Jyothi, B., Vishnu Murthy, G. "Frequent pattern and high utility item sets with up -tree format in distributed data mining" *International Journal of Innovative Technology and Exploring Engineering*, 2019, 8(6 C2), pp. 258-261
- [20] Alapati N.; Prasad B.V.V.S.; Sharma A.; Kumari G.R.P.; Veeneetha S.V.; Srivalli N.; Udaya Lakshmi T.; Sahitya D., "Prediction of Flight-fare using machine learning", 2022 International Conference on 4th Industrial Revolution Based Technology and Practices, ICFIRTP 2022, : 134
- [21] Baskar M.; Rajagopal R.D.; Prasad B.V.V.S.; Chinna Babu J.; Bartáková G.P.; Arulananth T.S., "Multi-region minutiae depth value-based efficient forged finger print analysis", *PLoS ONE* 18:11
- [22] R. Geetha, D. Jayakumar, S. R. Devi, S. Senthurya, A. Akula and R. Karthiga, "Designing an Intelligent Data Transmission Strategy Between Entities Using Wireless Communication Methodology," 2024 International Conference on Intelligent Systems for Cybersecurity (ISCS), Gurugram, India, 2024, pp. 1-6, doi: 10.1109/ISCS61804.2024.10581290.