



Machine Learning-based Prediction of Hypotension During Anesthesia Induction

Sana Hashemi  ¹, Ahmad Ali Abin  ¹, Hossein Sheikhi ², Azar Ejmalian ³, Shahabedin Nabavi  ¹, Maryam Vosoughian  ⁴, Soudeh Tabashi  ⁴, Ardesir Tajbakhsh  ^{5,*}

¹ Faculty of Computer Science and Engineering, Shahid Beheshti University, Tehran, Iran

² School of Medicine, Shahid Beheshti University of Medical Sciences, Tehran, Iran

³ Department of Anesthesiology, School of Medicine, Iran University of Medical Sciences, Tehran, Iran

⁴ Department of Anesthesiology & Critical Care, Anesthesiology Research Center, Taleghani Hospital, Shahid Beheshti University of Medical Sciences, Tehran, Iran

⁵ Anesthesiology Research Center, Shahid Beheshti University of Medical Sciences, Tehran, Iran

*Corresponding Author: Anesthesia Research Center, Shahid Beheshti University of Medical Sciences, Tehran, Iran. Email: ardesir.tajbakhsh@gmail.com

Received: 3 August, 2024; Revised: 16 August, 2025; Accepted: 18 August, 2025

Abstract

Background: Hypotension is a common complication during the induction of anesthesia, leading to adverse outcomes such as acute kidney injury (AKI), myocardial injury, and, in high-risk patients, death.

Objectives: This study aimed to predict post-induction hypotension (PIH) by considering clinical interventions using machine learning (ML) methods.

Methods: Prior to the induction phase of anesthesia, patient data were collected, and cardiac monitoring was set to measure non-invasive blood pressure (NIBP) at 1-minute intervals. Afterwards, induction was performed by the anesthesiologist. Hypotension was assessed 30 minutes after induction, defined as either a 20% drop in mean arterial pressure (MAP), an absolute MAP below 65 mmHg, or a systolic blood pressure (SBP) below 90 mmHg. The ML techniques were employed to develop a real-time hypotension predictor. These models utilize data gathered from five minutes to predict occurrences of hypotension in the next 10 minutes. Feature selection methods such as dimension reduction and sequential feature selection algorithms were utilized to provide more informative inputs to the ML models. Static features such as clinical features and dynamic features like vital signs were collected from patients undergoing general anesthesia across multiple hospital centers. Among the 215 patients, 110 developed PIH.

Results: Without employing feature selection methods, the best performance belongs to the random forest (RF) model, with an accuracy of 88.3%, precision of 87.6%, recall of 85%, and an area under the curve of the receiver operating characteristic (AUC-ROC) at 0.945. Moreover, when utilizing feature selection methods, the RF model retained its status as the best model, with accuracy, precision, recall, and AUC-ROC values of 88.1%, 88.1%, 85.5%, and 0.947, respectively.

Conclusions: We discovered that ML models hold the potential to predict PIH within the subsequent 10 minutes by utilizing data collected five minutes prior. Furthermore, considering clinical interventions, such as the patient's position and type of anesthetic drug injection, have a positive impact on the performance of ML models.

Keywords: Anesthesia, Forecasting, Hypotension, Machine Learning

1. Background

Post-induction hypotension (PIH), which refers to hypotension occurring after the induction of anesthesia, is a significant concern within anesthesia practice due to its association with adverse patient

outcomes. It can result in complications such as prolonged hospital stays and even mortality (1). Various factors contribute to the development of PIH, including pre-induction systolic blood pressure (SBP), patient age, and emergency surgeries (2). Additionally, premedication drugs and the type of anesthetic

induction can also contribute to PIH. Despite efforts to prevent hypotension, it remains challenging due to the complex interactions of various patient and anesthesia-related factors (3). The early prediction of PIH could potentially enhance patient outcomes by enabling timely intervention and the prevention of adverse events (4). However, predicting PIH is intricate due to the multifactorial causes (5). Recent studies have explored the application of machine learning (ML) techniques to predict hypotension during anesthesia induction (5, 6). These investigations have yielded promising results, indicating that ML models can accurately predict hypotension based on patient data collected during anesthesia induction. Nonetheless, most of these studies have primarily focused on predicting hypotension solely based on patient data and have not accounted for the potential influence of clinical interventions on hypotension prediction.

In this particular study, hypotension during anesthesia induction was predicted while considering both clinical interventions and patient data. Clinical experts' experiences at the bedside suggest that the type of anesthetic drug injection can impact the occurrence of hypotension. Administering the anesthetic all at once (bolus) increases the likelihood of hypotension, whereas administering it gradually (titrated) reduces the chances. Consequently, the type of anesthetic drug injection is one of the clinical interventions investigated in this study, which hasn't been explored by previous research. During anesthesia, patients lose the ability to breathe spontaneously, and it becomes the anesthesiologist's responsibility to facilitate breathing. This is achieved by inserting an endotracheal tube into the patient's airway. The insertion of this tube often necessitates the patient to lie flat on their back. Consequently, in surgeries like those involving the back, spine, and hips, where the patient needs to be in the prone position for the procedure, they must initially lie on their back for anesthesia induction. After anesthesia is administered, a position change is performed to reposition the patient's face. This change in positioning on the bed frequently leads to a decrease in blood pressure – a phenomenon that is examined as a subsequent clinical intervention under a predefined framework. In this research, we utilized ML techniques to build a real-time model for predicting PIH.

Indeed, a crucial aspect of preparing inputs for ML models is determining an appropriate labeling approach. Given the absence of a unique hypotension definition and the varying criteria used in different studies, such as mean arterial pressure (MAP) below 65 mmHg (5, 7-9) or 55 mmHg (10) or SBP below 90 mmHg

(11), the selection of a sensitive and accurate hypotension definition becomes essential for constructing an effective model. Additionally, the incorporation of feature selection techniques serves as another strategy to furnish the ML models with more informative inputs.

In related works, Kendale et al. aimed to predict hypotension 10 minutes after induction by conducting multiple ML analyses using features extracted from electronic health records (EHRs) data. They employed the Recursive Feature Selection (RFS) function to select relevant features for ML models, and gradient boosting machine (GBM) prediction showed the best performance on both the training and test sets (12). Moghadam et al. developed an algorithm for predicting hypotension in intensive care units (ICU) based on the mean and standard deviation of 11 variables. These variables included peripheral capillary oxygen saturation (SpO_2), arterial blood pressure (ABP), diastolic blood pressure (DBP), SBP, respiration rate (RR), heart rate (HR), pulse pressure (PP), MAP, mean arterial pressure-to-heart rate ratio (MAP2HR), cardiac output (CO), and average respiratory rate intervals from electrocardiogram (ECG) time series (7).

In another study, Sudfeld et al. examined the association between PIH, occurring during the first 20 minutes following anesthesia induction, and early intraoperative hypotension (eIOH), occurring during the first 30 minutes after surgery. They utilized data from EHRs, which is considered one of the most reliable sources for medical research, particularly because they include time series of physiological signals (2). Chen et al. presented PHASE, a method for converting time series signals into input features for predictive ML models. By using physiological signals, researchers can more accurately predict the adverse effects of surgery. PHASE was evaluated using EHRs data from two operating room (OR) datasets and an ICU dataset. It demonstrated that the PHASE model is interpretable and valuable for clinical applications (13).

To predict critical conditions such as hemodynamic instability, critical care clinicians analyze multiple physiological parameters simultaneously. Cherifa et al. developed the multi-task learning physiological deep learner (MTL-PDL), which predicts HR and MAP simultaneously (9).

2. Methods

2.1. Problem Definition

Hypotensive events during the induction of anesthesia are characterized by a drop in blood pressure below a clinically significant threshold. This threshold can be defined by considering a decrease in SBP, DBP, or MAP values. To predict hypotension during induction using ML models at time t , we aim to utilize patient information within the time range of $t-5$ to t to make predictions about the occurrence of PIH at the time of $t+10$. Our learning model in this task performs binary classification, categorizing instances into two classes: Hypotension and non-hypotension. This is achieved by calculating the probability of a hypotensive event in the next 10 minutes based on the patient's information collected five minutes ago. A significant aspect of this study is to account for the impact of clinical interventions, such as position and type of anesthetic drug injection, on the occurrence of hypotension during the specified time period. In essence, the main purpose of this study is to employ ML models for predicting PIH, taking into consideration the influence of clinical interventions.

2.2. Data Collection

In this study, we collected a dataset of 215 patients from multiple hospital centers in Iran, between 10 May 2024 and 19 June 2024. The dataset includes the demographic information and vital records of patients who underwent general anesthesia. Demographic information, including age, gender, weight, NPO time, fluid before induction, American Society of Anesthesiologists (ASA) physical status classification, type of surgical procedure, type of surgery, Losartan usage, background diseases, muscle relaxants, premedication drugs, IV anesthetic drugs, and details of anesthesia interventions such as type of anesthetic drug injection and position, were retrieved from the data collection forms. Vital signs were collected from SAADAT B9 vital signs monitors. We extracted the vital records of patients during the first 30 minutes of anesthesia induction, including MAP, SBP, DBP, and HR, which were recorded at 1-minute intervals. The dataset was collected without any intervention in the patients' treatment process and was anonymized to ensure the privacy of patients. However, it's important to note that the authors did not have access to information that could potentially identify individual participants, both during and after the data collection process.

2.3. Features

2.3.1. Feature Types

In this study, we carefully selected several features by incorporating insights from related studies (5, 6) and obtaining confirmation from anesthesiologists to ensure the inclusion of relevant information. These features were categorized into static and dynamic groups. The static features consist of clinical features and clinical interventions. Clinical features such as age, weight, and gender are included in this study as part of the static features. Additionally, clinical interventions implemented during the surgery included position and type of anesthetic drug injection. By investigating these interventions, we aimed to capture the effectiveness of clinical actions in mitigating or resolving hypotensive episodes.

The dynamic features include vital records or intraoperative hemodynamic measurements. Parameters such as SBP, DBP, MAP, and HR were collected. These measurements provide real-time information about the patient's cardiovascular status and are crucial indicators of hypotensive episodes. By incorporating hemodynamic measurements as features, our models aimed to capture the dynamic changes in blood pressure and HR that precede and accompany hypotension. Both types of features can influence the risk of hypotension during anesthesia. By considering these features, our ML models were able to capture complex patterns and interactions, allowing us to predict the occurrence of PIH events. Table 1 shows all the static and dynamic features used in this study.

2.3.2. Feature Vectors

In this section, our objective is to prepare all feature values for utilization as inputs in ML models, effectively forming feature vectors. These feature vectors encompass both static and dynamic features. Each feature vector consists of 20 elements, and elements 1 to 16 correspond to the static features, which remain constant throughout the 30-minute induction period. These static features can have both numerical and non-numerical values. Non-numerical values will be transformed into numerical values using encoding methods. On the other hand, elements 17 to 20 of the feature vectors are designated for representing statistical values (such as mean or variance) of dynamic features that evolve over time. Thus, all the elements within the feature vectors are transformed into numerical values. As a result, this process yields feature vectors containing 20 numerical elements, effectively consolidating all pertinent information. These prepared feature vectors are then utilized for subsequent ML analysis.

Table 1. Features Description

Data Sources	Data Type	Category	Features
Data collection forms	Static features	Clinical features	Age
			Gender
			Weight
			NPO time
			Fluid before anesthesia induction
			ASA (E1, E2, or E3)
			Type of surgical procedure (head and neck surgery, thoracic surgery, abdominal and pelvic surgery, or extremities surgery)
			Type of surgery (elective or emergency)
			Fluid in early 30 minutes of induction
			Losartan usage
			Background diseases (diabetes, hypertension, cardiac ischemia, or others)
			Muscle relaxants drugs (succinyl choline, atracurium, cisatracurium, or rocuronium)
			Premedication drugs (benzodiazepins, opioids, or lidocaine)
		Clinical interventions	IV anesthetics drugs (propofol, sodium thiopental (nesdonal), etomidate, or ketamine)
			Position (supine, prone, lateral, or semi-sitting)
			Type of anesthetic drug injection (bolus or titrated)
Vital recorders	Dynamic features	Intraoperative hemodynamic measurements	HR
			SBP
			DBP
			MAP

Abbreviations: ASA, American Society of Anesthesiologists; HR, heart rate; SBP, systolic blood pressure; DBP, diastolic blood pressure; MAP, mean arterial pressure.

2.3.3. Feature Selection

In this study, we extracted 20 features from both static and dynamic features to predict hypotension during anesthesia induction using ML models. Given that the dataset contains around 215 samples, it is important to maintain a well-balanced ratio of features to samples for effective model training. An excessive number of features can lead to challenges related to dimensionality. High dimensionality can make model training more complex and might necessitate a larger dataset (14). To address these concerns, we employed two feature selection methods: Firstly, using dimensionality reduction techniques, such as Linear Discriminant Analysis (LDA), Principal Component Analysis (PCA), and Singular Value Decomposition (SVD) to reduce the feature count from 20 to 10. Next, we employed sequential feature selection methods in both forward and backward forms to choose the top 10 informative features.

2.4. Labeling Approach

Labeling data with appropriate blood pressure thresholds is crucial for training and evaluating ML models. In this section, we present a medical definition

of predicting hypotension during anesthesia induction and describe the data labeling approach. Hypotensive events during the induction of general anesthesia are defined by a reduction in blood pressure below a clinically significant threshold, which can be either absolute or relative. This binary classification task encompasses two classes: The hypotension class (labeled as 1) and the non-hypotension class (labeled as 0). The determination of these classes relies on three dynamic variables: The MAP, DBP, and SBP, each representing an individual's level of hypotension. Distinct threshold values have been established for these variables to identify instances of hypotension, with MAP less than 65 mm Hg, along with SBP less than 90 mm Hg, serving as hypotensive indicators. To assign labels to each sample, at time t , we examine the values of these three variables at a time point $n+10$ minutes ahead. If any of these values fall below their respective threshold (either absolute or relative), the sample is labeled as 1, denoting hypotension. Conversely, if none of these three values drop below their thresholds, the label 0 is assigned to the sample, indicating the absence of hypotension.

2.5. Data Monitoring and Predicting Interval

In the current study, the prediction of PIH is carried out by considering two-time intervals: The data

Table 2. Results of Using All Dynamic and Static Data ^a

Models	Accuracy	Precision	Recall	AUC-ROC
LR	63	63	46.4	0.524
SVM	67	52.1	40.1	0.486
K-NN	47.4	43.6	41.3	0.458
DT	85.2	82.9	82.7	0.848
RF	88.3	87.6	85	0.945
GBM	85.3	86.6	78.4	0.929
XGBoost	86.6	85.5	83.1	0.941
LightGBM	87.1	85.7	84.2	0.942

Abbreviations: AUC-ROC, area under the curve of the receiver operating characteristic; LR, logistic regression; SVM, support vector machine; K-NN, K-nearest neighbor; DT, decision tree; RF, random forest; GBM, gradient boosting machine; XGBoost, eXtreme gradient boosting; LightGBM, light gradient boosting.

^aValues are expressed as percentage.

Table 3. Results of Performing Sequential Feature Selection Methods ^a

Models	Accuracy	Precision	Recall	AUC-ROC
RF	85	83.5	84.1	0.933
GBM	84.9	87.79	78.1	0.924
XGBoost	86.3	85.1	85.1	0.939
LightGBM	86.7	86.5	84.6	0.943

Abbreviations: AUC-ROC, area under the curve of the receiver operating characteristic; RF, random forest; GBM, gradient boosting machine; XGBoost, eXtreme gradient boosting; LightGBM, light gradient boosting.

^aValues are expressed as percentage.

monitoring interval and the prediction interval for the occurrence of PIH. Relevant studies indicate that the length of these intervals varies based on the patient's anesthesia location. For instance, in the OR, both the data monitoring and prediction intervals are reported as 3, 5, 10, and 15 minutes (12, 15-17). Conversely, for patients in the ICU, the prediction interval tends to be longer, even around 30 minutes (7). In this research, guided by multiple studies and expert anesthesia validation, we selected a data monitoring interval of five minutes and a prediction interval of 10 minutes. In essence, the model will predict PIH in the next 10 minutes, based on information obtained five minutes prior.

3. Results

3.1. Experiment 1: Using All Static and Dynamic Features

In the first experiment, all static and dynamic features were provided to the ML models, and the outcomes are detailed in [Table 2](#).

As observed, the random forest (RF) model has demonstrated the most favorable performance with accuracy, precision, recall, and area under the curve of

the receiver operating characteristic (AUC-ROC) of 88.3%, 87.6%, 85%, and 0.945, respectively. Additionally, the GBM, XGBoost, and LightGBM models also stand out as high-performing models.

3.2. Experiment 2: Using Sequential Feature Selection Methods

In the subsequent experiments, the number of input features for the learning model was intentionally reduced from the initial 20 features through a process of feature selection. This method was aimed at enhancing the model's performance by identifying the most informative features. Proceeding, we will present the results of this technique in preparing inputs for the ML models. It's essential to note that in this experiment, we will report the outcomes of the top four performing models along with their corresponding evaluation metrics.

During the feature selection process using the sequential feature selection method, which involves both forward and backward algorithms, we extracted the top 10 features from the original set of 20 features. Finally, we considered the top-ranked features as input for the best-performing models. These features include

Table 4. Results of Performing Dimensionality Reduction Methods ^a

Models	LDA				PCA				SVD			
	Accuracy	Precision	Recall	AUC-ROC	Accuracy	Precision	Recall	AUC-ROC	Accuracy	Precision	Recall	AUC-ROC
RF	67	63.8	64.7	0.759	87.8	85.6	84	0.944	88.1	88.1	85.5	0.947
GBM	72.3	69	72.4	0.813	83.3	84.5	78.1	0.920	83.2	84.8	77.6	0.915
XGBoost	71.2	68.7	68.5	0.800	87.3	87.6	87.3	0.943	87.3	86.9	85.4	0.946
LightGBM	72.4	69.7	70.8	0.806	87.7	87.7	85.2	0.946	87.2	86.6	85	0.947

Abbreviations: LDA, linear discriminant analysis; PCA, principle component analysis; SVD, singular value decomposition; AUC-ROC, area under the curve of the receiver operating characteristic; RF, random forest; GBM, gradient boosting machine; XGBoost, eXtreme gradient boosting; LightGBM, light gradient boosting.

^a Values are expressed as percentage.

age, ASA, gender, position, type of surgery, Losartan usage, amount of fluid in the early 30 minutes, weight, HR, muscle relaxants, and IV anesthetic. The results of the top-performing models with these features as inputs can be seen in [Table 3](#).

As you can see, LightGBM is the model that has demonstrated the best performance by receiving the selected features through sequential feature selection algorithms. This model achieved accuracy, precision, recall, and AUC-ROC values of 86.7%, 86.5%, 84.6%, and 0.943, respectively.

3.3. Experiment 3: Using Dimensionality Reduction Methods

Dimensionality reduction is another feature selection method used in this study. By applying three kinds of dimensionality reduction methods, including LDA, PCA, and SVD, we obtained the results presented in [Table 4](#).

In this experiment, the RF model achieved the best result when its input was the features selected using SVD. The RF model demonstrated an accuracy of 88.1%, precision of 88.1%, recall of 85.5%, and an AUC-ROC of 0.947.

3.4. Experiment 4: Changing the Length of Data Monitoring Intervals

In the present study, the data monitoring interval is set to five minutes prior. In this experiment, we varied the length of this interval to 3, 10, and 15 minutes to assess the impact of this interval on the performance of ML models as predictors of hypotension. [Table 5](#) presents the results of this experiment.

As evident from the results, the GBM model attained the highest performance with a data monitoring interval of 3 minutes. The corresponding evaluation metrics values for accuracy, precision, recall, and AUC-ROC were 90.2%, 88.8%, 87.5%, and 0.961, respectively.

3.5. Experiment 5: Changing the Length of Post-induction Hypotension Prediction Intervals

As previously mentioned, our goal was to build ML models for predicting PIH within the subsequent 10 minutes. Therefore, the prediction interval in this study remained at 10 minutes. However, in the current experiment, we altered this interval to 3, 5, and 15 minutes to assess its impact. The outcomes of this experiment are detailed in [Table 6](#).

With a prediction interval of 3 minutes, the GBM model demonstrated the most favorable outcome. This model demonstrated an accuracy of 89.5%, precision of 88.8%, recall of 86.3%, and an AUC-ROC of 0.963.

4. Discussion

As previously stated, the aim of this research is to develop a predictive model using ML algorithms to ascertain whether a patient will experience hypotension in the subsequent 10 minutes based on a five-minute data input taken earlier. The data monitoring interval, signifying the five-minute window used for model monitoring, holds paramount importance in this study. In Experiment 4, we explored the impact of altering the data monitoring interval while keeping other parameters constant on the model's performance. We considered three distinct data monitoring intervals: Three, 10, and 15 minutes, in contrast to the original 5-minute interval.

The obtained results indicate that reducing the data monitoring interval while maintaining a consistent prediction interval significantly enhances the model's performance. In other words, the improvement in model performance resulting from the shortened data monitoring interval suggests that, under these conditions, the model places a greater emphasis on the available patient-specific information, particularly dynamic features that evolve over time. This, in turn, facilitates more effective learning by the model, leading

Table 5. Results of Changing the Length of Data Monitoring Intervals ^a

Models	RF				GBM				XGBoost				LightGBM			
	3 min	5 min	10 min	15 min	3 min	5 min	10 min	15 min	3 min	5 min	10 min	15 min	3 min	5 min	10 min	15 min
Accuracy	89.6	88.3	89.2	87.3	90.2	85.3	89.2	86.4	89.2	86.6	88.9	84.4	89.6	87.1	89.2	83.8
Precision	88.3	87.6	88.5	87.3	88.8	86.6	89.5	88.5	87.5	85.5	88	82	88.2	85.7	88.4	84.8
Recall	86.6	85	85.7	83.1	87.5	78.4	84.9	82.2	87.1	83.1	86.1	82.6	87.2	84.2	86.3	82.6
AUC-ROC	0.963	0.945	0.958	0.931	0.961	0.929	0.957	0.933	0.96	0.941	0.958	0.91	0.962	0.942	0.958	0.92

Abbreviations: RF, random forest; GBM, gradient boosting machine; XGBoost, eXtreme gradient boosting; LightGBM, light gradient boosting; AUC-ROC, area under the curve of the receiver operating characteristic.

^a Values are expressed as percentage.

to improved overall efficiency and enhanced prediction of hypotension.

After exploring the impact of altering the data monitoring interval, the next experiment (experiment 5) involves investigating the effect of the prediction interval on the model's predictive capabilities. In this study, the prediction interval is initially set at 10 minutes. However, the mentioned experiment aims to evaluate the model's performance using prediction intervals of 3, 5, and 15 minutes instead of the original 10-minute interval. The central hypothesis in this section posits that, while maintaining a fixed data monitoring interval (which is five minutes in this study), shorter prediction intervals will enhance the accuracy of the model's predictions. This hypothesis is based on the notion that reducing the time gap between these two intervals is likely to result in more accurate predictions by minimizing information gaps. Longer intervals, on the other hand, might lead to the exclusion of critical data, ultimately weakening the model's predictive capability. The results validate the aforementioned hypothesis, demonstrating that shorter prediction intervals indeed contribute to improved model performance in predicting hypotension.

The process of labeling the collected samples, on which the model's learning and evaluation are based, is one of the most important parts of this research. This task involves considering the concept of hypotension and the threshold associated with it. As mentioned, hypotension can be defined in two ways: Absolute or relative. In this study, the determination of the hypotension threshold was done in an absolute manner, using three dynamic variables: The MAP, DBP, and SBP. While the relative concept of hypotension can also be utilized to establish this threshold, we conducted an experiment by employing relative thresholds instead of absolute ones for data labeling. In this case, by considering the values of the three mentioned dynamic variables at time zero, if more than 20% of this value

decreases in any of the subsequent minutes and in any of these three variables, we labeled the respective sample as "1", indicating the occurrence of a hypotensive event.

After training and evaluating the models with relative labeled data, it was observed that compared to the absolute labeling results, the RF model exhibited the best performance. Through relative labeling, the RF model achieved an accuracy of 82.1%. However, employing the absolute labeling approach resulted in an accuracy value of 82.7%. Therefore, no significant improvement was observed by adopting the relative labeling approach. Furthermore, it's worth noting that the study by Putowski et al. (18) also highlighted the superiority of defining absolute hypotension over the relative definition. Given the confidence in the correctness of the obtained results and the demonstrated superiority of the absolute labeling method in this study, we can confidently assert that absolute labeling outperforms the relative labeling approach.

In the present study, as previously mentioned, the features can be categorized into two groups: Dynamic and static. Clinical interventions, such as position and type of anesthetic drug injection, are considered static features. To evaluate the impact of information obtained from these two clinical interventions in the prediction of ML models, we proceeded with two main analyses.

First, the ML models were trained solely using dynamic features. The best-performing model, RF, achieved an accuracy of 80.3%. Subsequently, we added the two mentioned clinical interventions to the dynamic feature set and compared the outcomes. The best-performing model, RF, reported an accuracy of 81.2%. The results clearly demonstrate that the inclusion of data pertaining to position and type of anesthetic drug injection as clinical interventions has a positive effect on the model's performance.

Table 6. Results of Changing the Length of Post-induction Hypotension Prediction Intervals^a

Models	RF			GBM			XGBoost			LightGBM		
	3 min	5 min	10 min	15 min	3 min	5 min	10 min	15 min	3 min	5 min	10 min	15 min
Accuracy	89.1	88.3	88.3	87.4	89.5	88.2	85.3	87.4	89.1	88.9	86.6	86.8
Precision	87.8	87.2	87.6	86.8	88.8	87.6	86.6	88.1	87.9	88	85.5	86
Recall	86.3	85.1	85	84.1	86.3	84.6	78.4	82.4	86.2	85.9	83.1	83.6
AUC-ROC	0.959	0.952	0.945	0.944	0.963	0.952	0.929	0.941	0.96	0.954	0.941	0.945
											0.962	0.955
											0.942	0.947

Abbreviations: RF, random forest; GBM, gradient boosting machine; XGBoost, eXtreme gradient boosting; LightGBM, light gradient boosting; AUC-ROC, area under the curve of the receiver operating characteristic.

^a Values are expressed as percentage.

Second, we trained ML models using all static and dynamic features. The best-performing model, RF, reached an accuracy of 88.3%. Next, we removed position and type of anesthetic drug injection from the feature set and re-evaluated the models. The best-performing model, RF, reported an accuracy of 82%. Once again, the results indicate that the omission of these clinical interventions leads to a decline in performance.

To conclude, our study employed various feature selection techniques, including sequential feature selection and dimensionality reduction. The most notable enhancement in performance was observed when applying the SVD method, which is a dimensionality reduction technique. However, it is imperative to note that in methods like SVD, the selected features remain a black box and lack clarity. Given the paramount importance of model interpretability and comprehending the influential input variables to achieve desirable outcomes, we recommend exploring alternative feature selection methods that provide these attributes.

Footnotes

Authors' Contribution: Study concept and design: A. T.; Acquisition of data: H. S. and A. E.; Analysis and interpretation of data: A. A. A. and S. N.; Drafting of the manuscript: S. H.; Critical revision of the manuscript for important intellectual content: M. V.; Statistical analysis: S. H., A. A. A., and S. N.; Administrative, technical, and material support: S. T.; Study supervision: A. T.

Conflict of Interests Statement: The authors declare no conflict of interest.

Data Availability: The dataset presented in the study is available on request from the corresponding author during submission or after publication. The data are not publicly available.

Ethical Approval: This study was approved by the Research Ethics Committees of the School of Medicine at Shahid Beheshti University of Medical Sciences, in June 2024 (ethics ID: IR.SBMU.MSP.REC.1403.163).

Funding/Support: This research received no funding/support.

Informed Consent: Informed consent was obtained from all individual participants included in the prospective study.

References

- Wong GTC, Irwin MG. Post-induction hypotension: a fluid relationship? *Anaesthesia*. 2020;76(1):15-8. <https://doi.org/10.1111/anae.15065>.
- Sudfeld S, Brechnitz S, Wagner JY, Reese PC, Pinnschmidt HO, Reuter DA, et al. Post-induction hypotension and early intraoperative hypotension associated with general anaesthesia. *Br J Anaesth*. 2017;119(1):57-64. [PubMed ID: 28974066]. <https://doi.org/10.1093/bja/axi217>.
- Bijker JB, van Klei WA, Kappen TH, van Wolfswinkel L, Moons KG, Kalkman CJ. Incidence of intraoperative hypotension as a function of the chosen definition: literature definitions applied to a retrospective cohort using automated data collection. *Anesthesiology*. 2007;107(2):213-20. [PubMed ID: 17667564]. <https://doi.org/10.1097/01.anes.0000270724.40897.8e>.
- Maheshwari K, Shimada T, Fang J, Ince I, Mascha EJ, Turan A, et al. Hypotension Prediction Index software for management of hypotension during moderate- to high-risk noncardiac surgery: protocol for a randomized trial. *Trials*. 2019;20(1):255. [PubMed ID: 31053082]. [PubMed Central ID: PMC6499972]. <https://doi.org/10.1186/s13063-019-3329-0>.
- Kang AR, Lee J, Jung W, Lee M, Park SY, Woo J, et al. Development of a prediction model for hypotension after induction of anesthesia using machine learning. *PLoS One*. 2020;15(4). e0231172. [PubMed ID: 32298292]. [PubMed Central ID: PMC7162491]. <https://doi.org/10.1371/journal.pone.0231172>.
- Lee J, Woo J, Kang AR, Jeong YS, Jung W, Lee M, et al. Comparative Analysis on Machine Learning and Deep Learning to Predict Post-Induction Hypotension. *Sensors (Basel)*. 2020;20(16). [PubMed ID: 32824073]. [PubMed Central ID: PMC7472016]. <https://doi.org/10.3390/s20164575>.
- Moghadam MC, Abad EMK, Bagherzadeh N, Ramsingh D, Li GP, Kain ZN. A machine-learning approach to predicting hypotensive events

in ICU settings. *Comput Biol Med*. 2020;118:103626. [PubMed ID: 32174328]. <https://doi.org/10.1016/j.combiomed.2020.103626>.

8. Cherifa M, Blet A, Chambaz A, Gayat E, Resche-Rigon M, Pirracchio R. Prediction of an Acute Hypotensive Episode During an ICU Hospitalization With a Super Learner Machine-Learning Algorithm. *Anesth Analg*. 2020;130(5):1157-66. [PubMed ID: 32287123]. <https://doi.org/10.1213/ANE.0000000000004539>.
9. Cherifa M, Interian Y, Blet A, Resche-Rigon M, Pirracchio R. The Physiological Deep Learner: First application of multitask deep learning to predict hypotension in critically ill patients. *Artif Intell Med*. 2021;118:102118. [PubMed ID: 34412841]. <https://doi.org/10.1016/j.artmed.2021.102118>.
10. Solomon SC, Saxena RC, Neradilek MB, Hau V, Fong CT, Lang JD, et al. Forecasting a Crisis: Machine-Learning Models Predict Occurrence of Intraoperative Bradycardia Associated With Hypotension. *Anesth Analg*. 2020;130(5):1201-10. [PubMed ID: 32287127]. <https://doi.org/10.1213/ANE.0000000000004636>.
11. Gratz I, Baruch M, Takla M, Seaman J, Allen I, McEniry B, et al. The application of a neural network to predict hypotension and vasopressor requirements non-invasively in obstetric patients having spinal anesthesia for elective cesarean section (C/S). *BMC Anesthesiol*. 2020;20(1):98. [PubMed ID: 32357833]. [PubMed Central ID: PMC7195764]. <https://doi.org/10.1186/s12871-020-01015-9>.
12. Kendale S, Kulkarni P, Rosenberg AD, Wang J. Supervised Machine-learning Predictive Analytics for Prediction of Postinduction Hypotension. *Anesthesiology*. 2018;129(4):675-88. [PubMed ID: 30074930]. <https://doi.org/10.1097/ALN.0000000000002374>.
13. Chen H, Lundberg SM, Erion G, Kim JH, Lee SI. Forecasting adverse surgical events using self-supervised transfer learning for physiological signals. *NPJ Digit Med*. 2021;4(1):167. [PubMed ID: 34880410]. [PubMed Central ID: PMC8654960]. <https://doi.org/10.1038/s41746-021-00536-y>.
14. Jain A, Zongker D. Feature selection: evaluation, application, and small sample performance. *IEEE Transactions on Pattern Analysis and Machine Intelligence*. 1997;19(2):153-8. <https://doi.org/10.1109/34.574797>.
15. Hatib F, Jian Z, Buddi S, Lee C, Settels J, Sibert K, et al. Machine-learning Algorithm to Predict Hypotension Based on High-fidelity Arterial Pressure Waveform Analysis. *Anesthesiology*. 2018;129(4):663-74. [PubMed ID: 29894315]. <https://doi.org/10.1097/ALN.0000000000002300>.
16. Lin CS, Chiu JS, Hsieh MH, Mok MS, Li YC, Chiu HW. Predicting hypotensive episodes during spinal anesthesia with the application of artificial neural networks. *Comput Methods Programs Biomed*. 2008;92(2):193-7. [PubMed ID: 18760495]. <https://doi.org/10.1016/j.cmpb.2008.06.013>.
17. Jeong Y, Kang AR, Jung W, Lee SJ, Lee S, Lee M, et al. Prediction of Blood Pressure after Induction of Anesthesia Using Deep Learning: A Feasibility Study. *Applied Sciences*. 2019;9(23). <https://doi.org/10.3390/app9235135>.
18. Putowski Z, Czajka S, Krzych Lj. Association between Intraoperative Blood Pressure Drop and Clinically Significant Hypoperfusion in Abdominal Surgery: A Cohort Study. *J Clin Med*. 2021;10(21). [PubMed ID: 34768530]. [PubMed Central ID: PMC8584611]. <https://doi.org/10.3390/jcm10215010>.