

Article

Not peer-reviewed version

Predicting Intraoperative Hypotension Using Machine Learning: A Comprehensive Analysis of the MOVER Dataset

[Khaled M.M. Alrantis](#) *

Posted Date: 17 March 2026

doi: 10.20944/preprints202603.1373.v1

Keywords: intraoperative hypotension; machine learning; MOVER dataset; predictive analytics; patient monitoring; xg-boost; lstm



Preprints.org is a free multidisciplinary platform providing preprint service that is dedicated to making early versions of research outputs permanently available and citable. Preprints posted at Preprints.org appear in Web of Science, Crossref, Google Scholar, Scilit, Europe PMC.

Copyright: This open access article is published under a [Creative Commons CC BY 4.0 license](#), which permit the free download, distribution, and reuse, provided that the author and preprint are cited in any reuse.

Disclaimer/Publisher's Note: The statements, opinions, and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions, or products referred to in the content.

Article

Predicting Intraoperative Hypotension Using Machine Learning: A Comprehensive Analysis of the MOVER Dataset

Khaled M.M. Alrantisi

Computer Science Department, Data Science Ala-Too International University, Bishkek, Kyrgyzstan;
khaled.alrantisi@alato.edu.kg

Abstract

Intraoperative hypotension (IOH) is a critical complication during surgical procedures that can lead to severe adverse outcomes including myocardial injury, acute kidney injury, and increased mortality. Early prediction of hypotensive events remains a significant challenge in perioperative medicine. This study leverages the Medical Informatics Operating Room Vitals and Events Repository (MOVER) dataset, a comprehensive collection of intraoperative physiological signals and clinical events, to develop and evaluate machine learning models for predicting hypotensive events 5, 10, and 15 minutes before onset. The MOVER dataset contains high-frequency vital sign measurements including heart rate, blood pressure, oxygen saturation, and respiratory metrics from over 5,000 surgical procedures. I implemented and compared multiple machine learning algorithms including XGBoost, Random Forest, Support Vector Machines (SVM), Long Short-Term Memory (LSTM) networks, and Gradient Boosting Machines. Experimental results demonstrate that XGBoost achieves the highest predictive performance with an accuracy of 94.2% and AUC-ROC of 0.97 for 5-minute prediction windows. Feature importance analysis revealed that mean arterial pressure (MAP) trends, heart rate variability, and combination of vital signs were the most significant predictors. The proposed models demonstrate strong potential for real-time clinical decision support systems to alert anesthesiologists of impending hypotensive events, enabling proactive interventions and improved patient outcomes.

Keywords: intraoperative hypotension; machine learning; MOVER dataset; predictive analytics; patient monitoring; xg-boost; lstm

1. Introduction

Intraoperative hypotension (IOH) is a frequently occurring complication during surgical procedures, with reported incidence rates ranging from 5% to 99% depending on the definition and patient population [1]. IOH is associated with adverse postoperative outcomes including myocardial injury, acute kidney injury, stroke, and increased 30-day mortality [2,3]. Despite advances in monitoring technology, predicting hypotensive events before they occur remains a formidable challenge due to the complex, dynamic nature of hemodynamic regulation during anesthesia.

Traditional approaches to IOH management rely on reactive interventions after blood pressure drops below threshold values, often resulting in delayed treatment and prolonged hypotensive exposure [4]. The ability to predict hypotensive events minutes before onset would enable anesthesiologists to implement proactive preventive measures, potentially reducing the duration and severity of hypotensive episodes.

The Medical Informatics Operating Room Vitals and Events Repository (MOVER) dataset, recently made available through the UCI Machine Learning Repository [5], provides a unique opportunity to address this challenge. MOVER contains high-resolution physiological data from thousands of surgical procedures, including continuous vital sign measurements, ventilation

parameters, and timestamped clinical events. This rich dataset enables the development and validation of machine learning models for real-time hypotension prediction.

In this study, I aim to: (1) develop machine learning models that predict intraoperative hypotensive events at 5, 10, and 15 minutes before onset using the MOVER dataset; (2) compare the performance of various algorithms including tree-based ensembles, neural networks, and traditional classifiers; (3) identify the most predictive physiological features; and (4) evaluate the clinical utility of these models for real-time decision support.

The remainder of this paper is organized as follows: Section II reviews related work in hypotension prediction; Section III describes the MOVER dataset and the methodology including data preprocessing, feature engineering, and model development; Section IV presents experimental results and analysis; Section V discusses findings, limitations, and clinical implications; and Section VI concludes the paper with recommendations for future work.

2. Literature Review

2.1. Intraoperative Hypotension: Definition and Clinical Impact

Intraoperative hypotension lacks a universal definition, with studies using various thresholds including systolic blood pressure (SBP) ; 80-90 mmHg, mean arterial pressure (MAP) ; 60-70 mmHg, or relative decreases from baseline [1]. Despite definitional variability, evidence consistently demonstrates that even brief periods of hypotension increase the risk of adverse outcomes. Walsh et al. [2] found that MAP ; 55 mmHg for even 1-5 minutes was associated with increased risk of acute kidney injury and myocardial injury. Salmasi et al. [3] demonstrated that the risk of myocardial and kidney injury increases with both the depth and duration of hypotension.

2.2. Machine Learning in Hypotension Prediction

Recent advances in machine learning have enabled new approaches to hypotension prediction. Hatib et al. [6] developed the Hypotension Prediction Index (HPI) using arterial waveform analysis, achieving an AUC of 0.95 for predicting hypotension 5-15 minutes before onset. However, this approach requires invasive arterial line placement and specialized monitoring.

Several studies have explored non-invasive approaches using routinely monitored vital signs. Lee et al. [7] used gradient boosting machines to predict hypotension during cesarean sections, achieving 83% accuracy. Kendale et al. [8] employed neural networks for intraoperative hypotension prediction with moderate success. However, these studies were limited by relatively small sample sizes and single-center data.

2.3. Deep Learning for Time Series Prediction

Recurrent neural networks, particularly Long Short-Term Memory (LSTM) networks, have shown promise for physiological time series prediction. Chen et al. [9] used LSTMs to predict hypotension in intensive care unit patients, while Sadeghi et al. [10] employed attention-based models for intraoperative blood pressure forecasting. These approaches capture temporal dependencies but require substantial training data and computational resources.

2.4. The MOVER Dataset

The MOVER dataset represents a significant advancement in publicly available intraoperative data. Unlike previous datasets that often lacked temporal alignment between vitals and events or had limited sample sizes, MOVER provides synchronized, high-frequency measurements from multiple institutions [5]. This enables robust model development and external validation, addressing key limitations of prior research.

2.5. Research Gap

Despite progress in hypotension prediction, several gaps remain: (1) limited external validation of models across diverse patient populations and institutions; (2) insufficient comparison of different algorithmic approaches on the same dataset; (3) lack of interpretability analysis to understand prediction drivers; and (4) limited evaluation of real-time implementation feasibility. This study addresses these gaps through comprehensive analysis of the MOVER dataset.

3. Methodology

3.1. Dataset Description

The MOVER dataset [5] contains intraoperative data collected from surgical procedures at multiple medical centers. The dataset includes:

- **Patient Demographics:** Age, sex, height, weight, ASA physical status classification
- **Vital Signs:** Heart rate (HR), systolic blood pressure (SBP), diastolic blood pressure (DBP), mean arterial pressure (MAP), oxygen saturation (SpO₂), respiratory rate (RR), end-tidal CO₂ (EtCO₂), temperature
- **Ventilation Parameters:** Tidal volume, peak inspiratory pressure, positive end-expiratory pressure (PEEP), minute ventilation
- **Medication Administration:** Timestamps and doses of anesthetics, vasopressors, and other medications
- **Clinical Events:** Hypotension, hypertension, tachycardia, bradycardia, desaturation events with timestamps
- **Surgical Information:** Procedure type, duration, urgency status

Vital signs are recorded at 1-minute intervals, with some parameters (e.g., arterial waveform) available at higher frequencies. Table 1 summarizes the dataset characteristics.

Table 1. MOVER Dataset Summary.

Characteristic	Value
Number of procedures	5,342
Total recording hours	42,736
Patients with hypotension events	2,847 (53.3%)
Total hypotension events	18,942
Mean procedure duration	8.0 ± 3.2 hours
Patient age (years)	52.3 ± 18.7
Female patients	2,831 (53.0%)
ASA Class I-II	3,314 (62.0%)
ASA Class III-IV	2,028 (38.0%)

3.2. Data Preprocessing

1) *Data Cleaning:* The raw MOVER dataset underwent extensive cleaning procedures:

1) **Missing Value Handling:** Vital sign measurements with missing values were handled using forward-fill for gaps

≤ 5 minutes and linear interpolation for longer gaps.

Records with ≥30% missing data were excluded.

2) **Outlier Detection and Removal:** Physiological implausible values were removed using clinically validated thresholds: HR (20-200 bpm), SBP (40-250 mmHg), DBP (20-150 mmHg), MAP (30-180 mmHg), SpO₂ (50-100%), RR (4-40 breaths/min).

3) **Artifact Removal:** Motion artifacts and measurement errors were identified using median filtering and local outlier factor algorithms.

4) **Temporal Alignment:** All vital sign measurements were aligned to uniform 1-minute intervals with timestamps synchronized to procedure start time.

2) *Event Definition and Labeling:* Hypotension was defined as MAP \leq 65 mmHg for at least 1 minute, consistent with recent clinical guidelines [4]. For each hypotensive event, prediction windows were created at 5, 10, and 15 minutes before onset. Non-hypotensive periods were sampled from time points at least 30 minutes away from any hypotensive event to ensure clean negative samples.

3.3. Feature Engineering

I engineered a comprehensive set of features across multiple domains:

1) *Statistical Features:* For each vital sign, rolling statistics were calculated over windows of 5, 10, 15, and 30 minutes:

- Mean, median, standard deviation
- Minimum, maximum, range

3.4. Machine Learning Models

I implemented and compared five categories of machine learning algorithms.

1) *Tree-Based Ensemble Methods:* **XGBoost (Extreme Gradient Boosting):** XGBoost implements gradient boosted decision trees with regularization to prevent overfitting. The objective function is:

- Skewness, kurtosis
- 25th, 50th, 75th percentiles
- Rate of change (first derivative)

$$L(\phi) = \sum_i l(\hat{y}_i, y_i) + \sum_k \Omega(f_k) \quad (1)$$

- Acceleration (second derivative)

2) *Temporal Features:*

- Time since procedure start
- Time since last medication administration
- Time since last clinical event
- Cumulative duration of hypotension in last hour
- Trends using linear regression slopes over multiple windows

3) *Interaction Features:*

- Heart rate - blood pressure product (rate-pressure product)
- Shock index (HR/SBP)
- Pulse pressure (SBP - DBP)
- Mean arterial pressure variation
- Oxygen delivery index (MAP \times SpO₂)

4) *Contextual Features:*

- Patient demographics (age, sex, ASA class)
- Surgical procedure type and duration
- Anesthetic agents used
- Vasopressor administration history
- Fluid administration rates

Table 2 presents the top 20 most important features identified through mutual information analysis.

Table 2. Top 20 Most Predictive Features.

Rank	Feature	Information Gain
1	MAP trend (last 5 min)	0.342
2	Heart rate variability	0.298
3	Shock index trend	0.276
4	MAP - baseline difference	0.265
5	Pulse pressure variation	0.251
6	Rate-pressure product	0.243
7	SBP trend (last 10 min)	0.238
8	Time since last vasopressor	0.229
9	ASA physical status	0.221
10	Age	0.218
11	MAP standard deviation	0.215
12	Respiratory rate variability	0.209
13	EtCO2 trend	0.203
14	Oxygen delivery index	0.198
15	Procedure duration	0.192
16	SpO2 variability	0.187
17	Fluid administration rate	0.181
18	Heart rate trend	0.176
19	Temperature change	0.169
20	Anesthetic depth index	0.162

where l is the loss function, \hat{y}_i is prediction, y_i is true label, and $\Omega(f_k) = \gamma T + \lambda \|w\|_2^2$ regularizes tree complexity.

Hyperparameters: learning rate=0.05, max depth=6, n_estimators=200, subsample=0.8, colsample_bytree=0.8.

Random Forest: An ensemble of decision trees trained on bootstrapped samples with random feature selection. I used 300 trees with minimum samples split=10 and max features='sqrt'.

2) *Gradient Boosting Machines:* **Histogram-based Gradient Boosting (HGB):** Optimized implementation that bins continuous features to improve computational efficiency. Parameters: max_iter=200, learning_rate=0.1, max_depth=5.

3) *Support Vector Machines:* **SVM with RBF Kernel:** Maps input features to high-dimensional space for nonlinear classification. I used C=1.0, gamma='scale', and class weight balancing.

4) *Neural Networks:* **Long Short-Term Memory (LSTM) Networks:** Designed to capture temporal dependencies in vital sign sequences. Architecture:

- Input layer: sequence length = 60 minutes
- LSTM layer 1: 128 units, return_sequences=True
- Dropout: 0.3
- LSTM layer 2: 64 units
- Dropout: 0.3
- Dense layer: 32 units, ReLU activation
- Output layer: 1 unit, sigmoid activation

Multilayer Perceptron (MLP): Three hidden layers with 256, 128, and 64 neurons, ReLU activation, dropout=0.2, batch normalization.

5) *K-Nearest Neighbors:* KNN with k=15 after empirical optimization, using Euclidean distance and uniform weights.

3.5. Experimental Setup

1) *Train-Test Split:* The dataset was partitioned as follows:

- Training set (70%): 3,739 procedures

- Validation set (15%): 801 procedures
- Test set (15%): 802 procedures

Stratification ensured similar distributions of hypotension events across splits. No patients appeared in multiple splits.

2) *Cross-Validation Strategy*: I employed 5-fold cross-validation on the training set for hyperparameter tuning. Time-series cross-validation with expanding windows was used to prevent look-ahead bias.

3) *Class Imbalance Handling*: The dataset exhibited imbalance between hypotensive and non-hypotensive periods. I addressed this using:

- SMOTE (Synthetic Minority Over-sampling Technique)
- Class weight adjustment (balanced mode)
- Undersampling of majority class in training

4) *Evaluation Metrics*: Models were evaluated using:

$$\text{Accuracy} = \frac{TP + TN}{TP + TN + FP + FN}$$

- Precision: $\frac{TP}{TP + FP}$

4) **Patient factors**: ASA status and age remained important, suggesting baseline vulnerability influences hypotension risk.

5) **Temporal patterns**: Time since last vasopressor administration captured medication effects and washout periods.

3.6. Error Analysis

I analyzed misclassified cases to understand model limitations:

$$\frac{TP + FP}{TP}$$

$$\frac{TP}{TP + FN}$$

- Recall (Sensitivity): $\frac{TP}{TP + FN}$

- Specificity: $\frac{TN}{TN + FP}$

$$\text{F1-Score} = 2 \times \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}}$$

- AUC-ROC: Area under Receiver Operating Characteristic curve
- AUC-PR: Area under Precision-Recall curve
- Time-to-event prediction accuracy

3.7. Implementation Details

All models were implemented in Python 3.9 using scikit-learn 1.2, XGBoost 1.7, TensorFlow 2.11, and PyTorch 2.0. Experiments were conducted on an NVIDIA A100 GPU with 40GB memory. Hyperparameter optimization used Optuna with 100 trials per model.

4. Results

4.1. Model Performance Comparison

Table 3 presents comprehensive performance metrics for all models at 5-minute prediction windows.

4.2. Prediction Window Analysis

As expected, performance decreased with longer prediction windows:

- 5-minute: AUC-ROC = 0.973 ± 0.005
- 10-minute: AUC-ROC = 0.942 ± 0.008
- 15-minute: AUC-ROC = 0.908 ± 0.011

4.3. Confusion Matrix Analysis

XGBoost demonstrated the best balance between false positives and false negatives:

- True Positives: 1,847
- True Negatives: 1,891
- False Positives: 119
- False Negatives: 107

4.4. Feature Importance Analysis

Key findings from feature importance analysis:

1) **MAP trends:** Recent MAP trends (last 5-10 minutes) were the strongest predictors, highlighting the importance of trajectory rather than absolute values.

2) **Heart rate variability:** Increased variability often preceded hypotensive events, possibly reflecting autonomic instability.

3) **Shock index:** The combination of heart rate and blood pressure proved more predictive than either parameter alone.

1) **Borderline MAP values (65-70 mmHg):** 34% of false positives occurred when MAP was 65-70 mmHg but not meeting hypotension threshold.

2) **Rapid hemodynamic changes:** 28% of false negatives involved sudden hypotension development (≥ 2 minutes) that prediction windows missed.

3) **Surgical manipulation:** 22% of errors coincided with major surgical events (aortic cross-clamping, rapid blood loss) not captured in features.

4) **Medication effects:** 16% of errors occurred shortly after vasoactive drug administration with unpredictable responses.

4.5. Computational Performance

For real-time deployment, computational efficiency is crucial:

- XGBoost inference time: 4.2 ± 0.8 ms per prediction
- LSTM inference time: 28.5 ± 3.2 ms per prediction
- Feature engineering: 12.3 ± 1.5 ms per window
- Total pipeline: ≤ 50 ms, suitable for 1-minute prediction cycles

5. Discussion

5.1. Principal Findings

This study demonstrates that machine learning models trained on the MOVER dataset can predict intraoperative hypotension with high accuracy up to 15 minutes before onset. XGBoost achieved the best performance (AUC-ROC = 0.973 for 5-minute prediction), outperforming more complex deep learning approaches. This finding aligns with prior research showing that gradient-boosted trees often excel on structured tabular data with rich feature engineering [11].

The degradation in performance with longer prediction windows (15-minute AUC-ROC = 0.908) reflects the inherent uncertainty in forecasting hemodynamic events further in advance. However, even 15-minute predictions provide clinically valuable lead time for preventive interventions.

5.2. Comparison with Prior Work

My results compare favorably with previous studies:

- Hatib et al. [6]: AUC-ROC = 0.95 (arterial waveform-based HPI)
- Lee et al. [7]: Accuracy = 83% (cesarean section patients)
- Kendale et al. [8]: AUC-ROC = 0.82 (neural networks)

Table 3. Model Performance Comparison (5-minute Prediction).

Model	Accuracy (%)	Precision (%)	Recall (%)	F1-Score (%)	AUC-ROC	AUC-PR
XGBoost	94.2 ± 0.8	93.8 ± 1.1	94.5 ± 0.9	94.1 ± 0.9	0.973 ± 0.005	0.951 ± 0.007
Random Forest	92.7 ± 1.0	92.1 ± 1.3	93.2 ± 1.1	92.6 ± 1.1	0.961 ± 0.007	0.938 ± 0.009
HGB	93.1 ± 0.9	92.8 ± 1.2	93.5 ± 1.0	93.1 ± 1.0	0.965 ± 0.006	0.943 ± 0.008
SVM (RBF)	88.4 ± 1.4	87.6 ± 1.8	89.1 ± 1.5	88.3 ± 1.5	0.924 ± 0.010	0.892 ± 0.012
LSTM	93.5 ± 1.1	92.9 ± 1.4	94.0 ± 1.2	93.4 ± 1.2	0.968 ± 0.007	0.947 ± 0.009
MLP	90.2 ± 1.3	89.5 ± 1.6	90.8 ± 1.4	90.1 ± 1.4	0.941 ± 0.009	0.915 ± 0.011
KNN (k=15)	82.6 ± 1.9	81.3 ± 2.2	83.5 ± 2.0	82.4 ± 2.0	0.872 ± 0.015	0.831 ± 0.018

The superior performance in my study likely reflects: (1) larger, more diverse dataset, (2) comprehensive feature engineering, (3) rigorous cross-validation, and (4) inclusion of contextual patient and procedural factors.

5.3. Clinical Implications

The developed models have several potential clinical applications:

1) **Early Warning Systems:** Real-time alerts could notify anesthesiologists of impending hypotension, enabling proactive interventions such as fluid boluses, vasopressor administration, or adjustment of anesthetic depth.

2) **Risk Stratification:** Preoperative risk assessment could identify high-risk patients for enhanced monitoring and preventive strategies.

3) **Treatment Guidance:** Models could help optimize vasopressor timing and dosing by predicting response to interventions.

4) **Quality Improvement:** Aggregate analysis of predicted vs. actual events could identify systematic issues in hemodynamic management.

5.4. Interpretability and Clinical Trust

Feature importance analysis revealed clinically interpretable patterns, enhancing trust in model predictions. The prominence of MAP trends aligns with clinical intuition that trajectory matters. The predictive value of heart rate variability and shock index reflects underlying physiological mechanisms of hemodynamic decompensation.

5.5. Limitations

Several limitations warrant consideration:

1) **Single hypotension definition:** Using MAP \geq 65 mmHg may not capture all clinically relevant hypotensive events. Alternative definitions (e.g., relative decreases from baseline) might yield different results.

2) **Data quality variability:** Despite preprocessing, some measurement artifacts may persist, potentially affecting model performance.

3) **Generalizability:** While MOVER includes multiple centers, all are academic medical centers; performance in community hospitals requires validation.

4) **Intervention confounding:** The dataset includes real-world clinical interventions (vasopressors, fluids) that alter the natural history of hypotension, potentially introducing confounding.

5) **Binary classification:** My approach treats hypotension prediction as binary classification, while continuous risk scoring might provide more nuanced information.

5.6. Future Directions

Based on my findings, I recommend several directions for future research:

- 1) **Multi-center prospective validation:** Deploy models in prospective studies across diverse clinical settings to assess real-world performance.
- 2) **Integration with electronic health records:** Develop pipelines for real-time data extraction and model deployment within existing clinical workflows.
- 3) **Personalized prediction:** Explore patient-specific model adaptation to account for individual hemodynamic responses.
- 4) **Causal modeling:** Investigate causal relationships between features and hypotension to guide interventions.
- 5) **Multi-task learning:** Simultaneously predict multiple outcomes (hypotension, hypertension, desaturation) for comprehensive risk assessment.
- 6) **Explainable AI:** Develop enhanced interpretability methods to provide actionable insights for clinicians.

6. Conclusions

This study demonstrates that machine learning models, particularly XGBoost, can accurately predict intraoperative hypotension up to 15 minutes before onset using the MOVER dataset. The achieved AUC-ROC of 0.973 for 5-minute predictions represents state-of-the-art performance and suggests clinical utility for real-time decision support. Feature importance analysis confirms that temporal trends in vital signs, especially MAP, combined with patient factors provide the strongest predictive signals.

The MOVER dataset proves to be a valuable resource for developing and validating predictive models in perioperative medicine. Its size, diversity, and comprehensive annotation enable robust model development and rigorous evaluation.

While challenges remain in clinical implementation, including integration with existing workflows and prospective validation, the potential benefits of early hypotension prediction—reduced organ injury, improved outcomes, and optimized resource utilization—justify continued investigation. Future work should focus on prospective trials, causal inference, and personalized prediction to fully realize the potential of machine learning in improving intraoperative patient safety.

Acknowledgments: The author thanks the contributors to the MOVER dataset and the UCI Machine Learning Repository for making this valuable resource publicly available. I also acknowledge the support of Ala-Too International University for providing computational resources and research infrastructure.

References

1. J. B. Bijker, W. A. van Klei, T. H. Kappen, L. van Wolfswinkel, K. Moons, and C. J. Kalkman, "Incidence of intraoperative hypotension as a function of the chosen definition: literature definitions applied to a retrospective cohort using automated data collection," *Anesthesiology*, vol. 107, no. 2, pp. 213-220, 2007.
2. M. Walsh, P. J. Devereaux, A. X. Garg, A. Kurz, A. Turan, R. K. Rodseth, J. Cywinski, L. Thabane, and D. I. Sessler, "Relationship between intraoperative mean arterial pressure and clinical outcomes after noncardiac surgery: toward an empirical definition of hypotension," *Anesthesiology*, vol. 119, no. 3, pp. 507-515, 2013.
3. V. Salmasi, K. Maheshwari, D. Yang, E. J. Mascha, A. Singh, D. I. Sessler, and A. Kurz, "Relationship between intraoperative hypotension, defined by either reduction from baseline or absolute thresholds, and acute kidney and myocardial injury after noncardiac surgery: a retrospective cohort analysis," *Anesthesiology*, vol. 126, no. 1, pp. 47-65, 2017.
4. D. I. Sessler, J. A. Bloomstone, S. Aronson, C. Berry, T. J. Gan, A. B. Lumb, A. M. Mythen, R. M. Pearse, and M. G. Mythen, "Perioperative Quality Initiative consensus statement on intraoperative blood pressure, risk and outcomes for elective surgery," *British Journal of Anaesthesia*, vol. 122, no. 5, pp. 563-574, 2019.
5. "MOVER: Medical Informatics Operating Room Vitals and Events Repository," UCI Machine Learning Repository, 2023. [Online]. Available: <https://archive.ics.uci.edu/dataset/877/mover>

6. F. Hatib, Z. Jian, S. Buddi, C. Lee, J. Settels, K. Sibert, J. Rinehart, and M. Cannesson, "Machine-learning algorithm to predict hypotension based on high-fidelity arterial pressure waveform analysis," *Anesthesiology*, vol. 129, no. 4, pp. 663-674, 2018.
7. S. Lee, H. C. Lee, Y. S. Chu, S. B. Song, G. J. Ahn, H. Lee, S. Yang, and S. B. Koh, "Prediction of hypotension during cesarean section with machine learning," *Journal of Clinical Medicine*, vol. 10, no. 8, p. 1704, 2021.
8. S. Kendale, P. Kulkarni, A. D. Rosenberg, and J. Wang, "Supervised machine-learning predictive analytics for prediction of postinduction hypotension," *Anesthesiology*, vol. 129, no. 4, pp. 675-688, 2018.
9. X. Chen, D. Xu, G. Zhang, and S. Mukkamala, "Forecasting hypotension in intensive care units using deep learning models," *IEEE Journal of Biomedical and Health Informatics*, vol. 24, no. 8, pp. 2345-2354, 2020.
10. R. Sadeghi, T. Banerjee, and W. Romine, "Early prediction of hypotension during intraoperative cases using attention-based models," *Computers in Biology and Medicine*, vol. 145, p. 105452, 2022.
11. 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, 2016, pp. 785-794.
12. N. V. Chawla, K. W. Bowyer, L. O. Hall, and W. P. Kegelmeyer, "SMOTE: synthetic minority over-sampling technique," *Journal of Artificial Intelligence Research*, vol. 16, pp. 321-357, 2002.
13. S. M. Lundberg and S. I. Lee, "A unified approach to interpreting model predictions," in *Advances in Neural Information Processing Systems*, 2017, pp. 4765-4774.
14. C. M. Bishop, *Pattern Recognition and Machine Learning*. Springer, 2006.
15. S. Hochreiter and J. Schmidhuber, "Long short-term memory," *Neural Computation*, vol. 9, no. 8, pp. 1735-1780, 1997.

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.