Predicting Hypotension by Learning from Multivariate Mixed Responses

Jodie Ritter, Xiaoyu Chen, Lihui Bai*, *Member, IAENG,* Jiapeng Huang

***Abstract*— Blood pressure is the main determinant of blood flow to organs. Hypotension is defined as a systolic blood pressure less than 90 mmHg or a diastolic blood pressure less than 50 mmHg. The severity and duration of hypotension is associated with low blood flow to organs, which often result in organ damage and a high mortality rate. Predicting hypotension prior to and during surgery can reduce the incidence and duration of hypotension, thus improve patient outcomes. This paper uses preoperative bloodwork and vital signs as well as perioperative vital signs in 5-minute increments as inputs to forecast hypotension. Hypotension can be represented by multivariate mixed responses for systolic and diastolic blood pressures and hypotension classification, which follow both continuous and binary distributions, respectively. The main objective of this paper is to apply a new machine learning method known as an "Interpretable Neural Network" (INN) to this clinical predictive application by simultaneously modeling mixed hypotension responses considering experts' domain knowledge. A novel data pipeline is proposed to conduct variable selection, clustering, and missing data imputation to address the issues of missing value and heterogeneous samples that are common in medical records. The customized INN method was developed and tested with a dataset containing 588 hysterectomy surgeries. Computational results suggest that the Gaussian mixture model produced better clustering results, compared to a simple clustering based on patients' lab work-up records. The novel INN method was successfully applied to the hypotension prediction, providing a prediction with reasonable accuracy and high interpretability for the prediction. The binary response has a testing accuracy of 92~95%, while the continuous responses have a root mean square error in the range of 10~25. Finally, the mixed response model outperformed the pure classification model in predicting hypotension by exploiting the hidden relationship between hypotension and the actual measures of diastolic and systolic blood pressures.**

***Index Terms*—hypotension, perioperative medicine, machine learning, neural network**

I. Introduction

Perioperative hypotension is associated with adverse outcomes in patients undergoing surgery [1]. Hypotension during noncardiac surgery can cause postoperative complications such as renal insufficiency, myocardial injury, and increased mortality. Predicting hypotension prior to the episode and taking preventative measures early can be crucial to improving patient outcomes.

Currently, management of perioperative hypotension is reactive [2]. Many factors contribute to perioperative hypotension such as patient comorbidities, preoperative medications, and medications used for induction of anesthesia [3]. Additionally, hypotension during surgery is preceded by subtle hemodynamic changes. These changes are hard to detect because the cardiovascular system is interdependent, has complicated networks, and is influenced by compensatory mechanisms [4]. In Lin et al. [1], when anesthesiologists used the prevalent methods in practice to predict perioperative hypotensive episodes, they scored an average accuracy of 51.6%. Realizing the complexity of prediction and its importance in improving patient outcomes, researchers increasingly resort to artificial intelligence (AI) and machine learning (ML) due to their ability to incorporate large amounts of data and develop robust predictive analytics (e.g., [2] and [5]).

This paper aims to use pre-operative bloodwork and vital signs as well as intra-operative vital signs as inputs to forecast hypotension. In particular, the perioperative medical data is collected at 5-minute interval. This research contributes to the literature of applying machine learning to hypotension prediction in two ways. First, we propose a novel data pipeline to deal with medical records data in time series but with significant number of missing values. The pipeline consists of a Least Absolution Shrinkage and Selection Operator (LASSO) based variable selection, a Gaussian Mixture Model based clustering, and a Principal Component Analysis (PCA) based missing data imputation. Second, we propose a mixed response model that considers binary response (hypotension or non-hypotension) and continuous response (diastolic and systolic blood pressures) jointly. This allows for better prediction accuracy by utilizing hidden relationship between these types of responses. Third, we explore a new machine learning method known as the "interpretable neural network" (INN) in predicting perioperative hypotension, thus integrating anesthesiologists' expert opinion into the prediction [6]. This increases the interpretability of the machine learning model and therefore its potential adoption by clinicians.

The research utilized a dataset consisting of medical records for 1,463 hysterectomy patients at University of Louisville Health (Louisville, Kentucky, USA) from June 2018 to June 2021. The customized INN method is developed and tested with a dataset containing 588

Manuscript received March 26, 2023; revised April 15, 2023.

J. Ritter obtained her Master of Science in Industrial Engineering from the University of Louisville in 2022 (e-mail: jodie.ritter@louisville.edu).

X. Chen is an assistant professor of Industrial Engineering, University of Louisville, Louisville, KY 40292 USA (e-mail: xiaoyu.chen@louisville.edu).

L. Bai is an associate professor of Industrial Engineering, University of Louisville, Louisville, KY 40292 USA (e-mail: lihui.bai@louisville.edu).

J. Huang is a professor of Anesthesiology, University of Louisville, Louisville, KY 40292 USA (e-mail: Jiapeng.huang@louisville.edu).

hysterectomy surgeries.

The remainder of this paper is organized as follows. Section II reviews literature on predictive models for perioperative hypotension and novel artificial neutral network methodology. Section III presents exploratory data analysis and ANOVA results for the 1,463 medical records. Section IV discusses the methodology including data pipeline, the architecture of the INN and INN with mixed response. Section V describes parameter settings and computation results for the INN with mixed response model. Decision rules are presented to demonstrate the interpretability of the prediction model. Finally, Section VI concludes the paper with future research directions.

II. LITERATURE REVIEW

A. Hypotension Prediction Using Machine Learning

Predicting hypotension using pre- and intra-operative data is a young field and the literature on this is rather scant. One closely related work by Kang et al. [7] conducted a binary classification of hypotension using pre-operative and intra-operative data as we do. Intra-operative data was recorded from induction to incision and Naïve Bayes, logistic regression, random forest, and ANN models were used to predict hypotension. The random forest model performed the best with an area under the receiver operating characteristic (AUROC) curve of 0.84. Similarly, Kendale et al. [3] used pre-operative data and intra-operative vital signs from anesthesia induction to 10 minutes post induction for the classification of hypotension using similar machine learning models as in [7]. The random forest model in [3] again performed the best with an AUROC of 0.74. Unlike the previous two studies, Hatib et al. [8] used arterial blood pressure waveform data in a logistic regression model for the classification of hypotension. This model was successful in predicting intraoperative hypotension 15 minutes before it occurred with a sensitivity of 88% and specificity of 87%.

B. Novel Approaches in Machine Learning

Due to the project needs, we focus our review on ML research that addresses missing data imputation, clustering, interpretable ML methods and models with mixed responses.

A common challenge in healthcare analytics is to handle missing data that occurs due to either clinicians' simply not collecting them, or a monitoring device malfunctioning, or random glitches in the electronic health record system. Methods for missing data imputation include zero imputation, means imputation, *k*-nearest neighbor imputation, and Expectation-Maximization (EM) imputation. Hegde et al. [9] simulated missing data in healthcare records and compared Principal Component Analysis (PCA) to Multiple Imputation for Chained Equations (MICE) for imputation of healthcare data. They used PCA to implement feature reduction and the EM algorithm to fill missing data. They concluded that PCA outperformed MICE in overall missing value imputation accuracy and root mean square error. Regarding imputation of missing data not at random (MNAR) commonly seen in healthcare data, Le and Tan [10] concluded that the use of more information of the same medical context improves the imputation of missing values. In the literature, PCA has also been used with Gaussian Mixture Models (GMM) to improve performance. Guo and Chen [11] used PCA with GMM in a HVAC fault diagnosis model using a Bayesian network that yielded improved accuracy compared to using GMM alone. Additionally, PCA with GMM was used in population stratification for ancestry estimation and demonstrated superior performance, compared to GMM, *k*-means, and *k*-means with PCA [12].

Artificial neural network (ANN) is arguably one of the most implemented ML methods. One drawback of ANN is the "black box" effect. To address this, Chen et al. [13] developed an "Interpretable Neural Network" (INN) to enable transparency in the neural network. The INN takes rules established using human domain knowledge and optimizes their decision thresholds for better prediction performance.

Lastly, modeling mixed response is prevalent in statistics, and is being introduced to ML. Kang et al. [14] jointly modeled binary and continuous responses, and concluded that the joint predictions were more accurate than if modeled separately. Hwang and Pennell [15] confirmed similar advantages of jointly modeling, especially when the two types of responses have underlying relations.

III. EXPLORATORY DATA ANALYSIS

Data was collected for 1,463 hysterectomy patients at the University of Louisville Health between September 2018 and June 2021. Of these patients, 642 patients or 44% of the total patient population experienced intra-operative hypotension. This data was manually recorded using patient records where any systolic blood pressure (SBP) reading less than 90 or diastolic blood pressure (DBP) reading less than 50 were defined as hypotension during surgery.

The Analysis of Variance (ANOVA) and logistic regression reveal that age, pre-operative diastolic blood pressure, pre-operative heart rate, hypertension and congestive heart failure were significant factors for hypotension classification.

We analyzed 15 binary variables, including hypertension, use of anti-hypertensive medication, dyslipidemia, angiotensin-converting enzyme inhibitors (ACEI), beta blockers (BB), angiotensin receptor blockers (ARBs), dysrhythmia, coronary artery disease (CAD), emergency surgery, congestive heart failure (CHF), abnormal electrocardiogram (ECG), valvular heart disease, hypotension, peripheral vascular disease, and syncope. We note on several observations: 43% of the patient population had hypertension, 30% were taking medications to control hypertension, and 17% had dyslipidemia. Patients with heart pathologies took three different medications: 13% ACEI, 13% BB and 7% ARBs.

We also studied 31 continuous variables, including demographic and pre- and intra-operative measures. Boxplots (not included here), broken down by hypotensive and non-hypotensive cases, for age, preoperative systolic blood pressure (SBP) and diastolic blood pressure (DBP) and bilirubin reveal that for the hypotensive cases, the median age is slightly higher while the preoperative SBP and DBP are slightly lower. Bilirubin has a larger spread in the hypotensive

ISBN: 978-988-14049-4-7
ISSN: 2078-0958 (Print); ISSN: 2078-0966 (Online)

class. Additionally, scatterplots and correlation matrices (not included here) with preoperative continuous variables confirm that there was no separation of hypotensive and non-hypotensive patients for any of the variables. This suggests that intra-operative data must be combined with the pre-operative data for high quality prediction of hypotension.

IV. METHODOLOGY

A. Overview of the Machine Learning Pipeline

In order to predict a mixed response (binary classification of hypotension, continuous response of SBP and DBP) through a neural network with interpretable decisions using large, multivariate datasets with random and blocks of missing values, we propose a data pipeline as illustrated in Fig. 1. The pipeline contains five sequential steps: data preprocessing, missing value imputation, GMM clustering, one-step ahead forecasting and finally INN prediction.

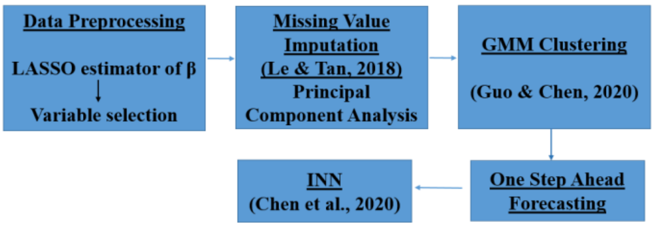

Figure 1. Overview of Pipeline to Predict Mixed Response

B. Data Preprocessing

To model the relationship between the input data and the response variable, a linear regression model was used to reveal their correlation. To estimate this model, each blood pressure reading was considered as one sample. While our feature size did not exceed the sample size, the proposed pipeline was developed to accommodate large datasets where the feature size is greater than the sample size. Therefore, a Least Absolution Shrinkage and Selection Operator (LASSO) variable selection method (see e.g., [16]) was used to identify a smaller set of predictors.

The data for the linear regression model is represented by $\boldsymbol{X}$, which is a matrix of m samples and q features. The response variable $\boldsymbol{Y}$ represents the SBP and DBP readings for time i. The relationship between $\boldsymbol{X}$ and $\boldsymbol{Y}$ can be modeled as $\boldsymbol{Y}=\boldsymbol{X\beta}+\varepsilon$ where $\boldsymbol{\beta}$ is the model coefficients and ε is the model error which is independently an identically distributed and follows a normal distribution. The LASSO estimation of the model coefficients is formulated as follows (e.g., [16]):

$$\beta=\operatorname{argmin}_{\beta}\{\tfrac{1}{2}\textstyle\sum_i (y_i-\beta_0-\sum_j x_{ij}\beta_j)^2+\lambda\sum_j|\beta_j|\}, \quad (1)$$

where λ is a tuning parameter. In our experiment, λ was selected by LASSO cross validation on the training data. The linear regression model was applied to the intra-operative features in the prediction of binary and mixed responses.

C. Missing Value Imputation

Because the INN cannot accept null values, we must impute the missing values first. Since missing values permeated the data randomly and in blocks, filling these values with interpolation or other methods like means filling could not be used. In this paper, we used the PCA due to its demonstrated performance for latent variables (e.g., [17]). Specifically, the Expectation–Maximization (EM) algorithm was used in this paper to iteratively calculate the Maximum Likelihood Estimates (MLE) of missing values.

The data, x_{ij} is reduced to p_{ij} by reducing the Euclidean distance between the original data points and the estimated data using a set number of components. The number of components was determined so that one would be able to explain 95% of the variance. Standardization was performed prior to using the PCA to fill the missing values.

D. Gaussian Mixture Model Clustering

Using large, multivariate datasets with the INN can yield rather long computational time. In the literature, Guo & Chen [11] suggested that imputing missing values with PCA followed by the Gaussian Mixture Model (GMM) for clustering can improve accuracy and computation time. Therefore, we utilized the GMM to identify the optimal number of clusters for greater computational efficiency.

The GMM clustering can identify multiple distributions within a dataset and assign a probability of each sample belonging to each cluster. The GMM algorithm uses the EM algorithm iteratively to determine the best mean and variance for a specified number of clusters. The minimum Bayesian Information Criterion (BIC) determines the optimal number of clusters, as in Wang and Liu [18].

The GMM was applied to the dataset that is the output from the PCA. The BIC and BIC gradient versus the number of clusters were graphed. The optimal number of clusters was chosen where the minimum BIC occurs and before the BIC gradient upward trend levels off.

E. One-Step Ahead Forecasting

Our intra-operative medical data naturally forms time series; however, data can be missing at randomly distributed timestamps. To make better predictions using such "incomplete" time series data, we propose a one-step ahead forecasting model. Multivariate irregular time series data cause random missing values in data where imputation is not appropriate. A common technique used in prediction models is one-step ahead forecasting that estimates the next time period from previous feature values (e.g., [19]).

Let x_{nj} be the nth summary statistics for feature j at the time of i-1 minus a window size or lag of previous readings. We used them to predict y_{ik} for response k at time i. In this way, a set of features is used to predict the next time period.

The one-step ahead forecasting approach was applied to the intra-operative data, which had irregular time series for all features. The window size was set to 6 representing approximately 30 minutes. The summary statistics used were min, max, mean, mode, median, range, standard deviation, variance and entropy. Additionally, two variables were used to measure the change in response from the previous reading.

ISBN: 978-988-14049-4-7
ISSN: 2078-0958 (Print); ISSN: 2078-0966 (Online)

F. *Interpretable Neural Nets (INN) with Mixed Responses*

As stated in the literature review, ANNs are widely used in ML but do not offer insight into how decisions are made. The INN in [13] addresses this lack of transparency. It takes rules established based on human domain knowledge and optimizes their thresholds for better prediction performance.

The input data for the INN is denoted by x_{nj} where *n* is the summarized statistics for time *i*-1 through the window size and *j* is the feature. Each feature *j* is associated with rule *l* in the architecture. Each rule *l* has a threshold associated with feature *j* represented by α_{jl}. From the input layer, hidden layer 1 makes a rule-based conclusion, t_{jl}, which is 1 if $x_j > \alpha_{jl}$. From there, hidden layer 2 has a combination logic that states if $x_1 > \alpha_{11}$ AND $x_2 > \alpha_{21}$, then z_l is 1. In other words, if both input variables associated with a rule are greater than the corresponding thresholds, then the rule conclusion in hidden layer 2 is 1. The response variable in the output layer is then determined to be 1 if any z_l is 1. For example, if z_1 =1 OR z_2 =1, then the response variable would be 1. Hidden layer 1initializes α_{jl} while hidden layer 2 optimizes α_{jl}.

Hypotension modeling requires an extension of the INN model from single continuous or categorical responses to multiple responses that follows different distributions (i.e., mixed responses). Therefore, we propose to jointly model: 1) two continuous responses: systolic and diastolic blood pressures, and 2) one binary response: indicator of hypotension status. In this way, the hidden association between continuous responses and binary response can be learned by the INN, which is expected to enhance the hypotension classification accuracy by sharing information.

V. COMPUTATIONAL RESULTS

A. *Input Data and Data Processing*

The intra-operative data that were added to the existing preoperative data included oxygen saturation, pulse rate, heart rate, respiratory rate and temperature at 5-minute intervals throughout surgery. Due to limitations of the case study data, only 588 of the 1,463 patients were used for the remainder of this research. Of these patients, 499 out of 7,233 readings, or 7%, met the criteria for hypotension.

Within the intra-operative data, there were inaccurate readings caused by various reasons (e.g., the arterial line or non-invasive line not connected to patients, the patient having severe hypotension or cardiac arrest or artifacts). These inaccurate readings were removed from the data by the consulting anesthesiologist. Examples include: SBP less than 30 mmHg, difference between systolic and diastolic arterial blood pressure in the same measurement less than 15mmHg, respiratory rate less than 5 or more than 40, pulse rate less than 30 or more than 150, and temperature less than 34 degrees Celsius. All values of oxygen saturation were kept.

A cross validation technique (see e.g., [20]) was used in Python to optimize the tuning parameter, λ, of the LASSO model. Subsequently, the best λ associated with the lowest cross validation error was used to estimate the LASSO model coefficients. The top twenty features selected by the 5-fold cross validation in the order of coefficient magnitude are shown in Table I. These features were used in all testing.

TABLE I
TOP 20 FEATURES SELECTED BY COEFFICIENT MAGNITUDE

Feature	Coefficients
DBP Delta	7.28
Temperature_ mean	-1.64
Pulse Rate _max	1.60
Respiratory Rate_entropy	-0.64
Temperature _var	-0.59
Pulse Rate _var	-0.58
Pulse Rate _entropy	0.58
Temperature_entropy	0.56
SBP Delta	0.52
Respiratory Rate_stdev	0.48
Temperature_range	0.47
Temperature_median	-0.32
Pulse Rate _min	0.28
Respiratory Rate_mode	0.26
SPO2_max	-0.19
Temperature_min	-0.13
SPO2_min	-0.11
Respiratory Rate_max	0.11
SPO2_mode	0.09
Temperature_stdev	0.07

B. *Gaussian Mixture Model Clustering*

Proper data clustering is important to achieving high quality prediction. One simple approach for this study is to separate patients based on the amount of lab work-up in their pre- and intra-operative medical records. The requirement for more lab work-up could indicate a sicker patient raising stability concerns by physicians. This benchmark approach is set as lab work-up low, medium and high, in our presentation.

Alternatively, we explored the GMM clustering. To determine the number of clusters to use, the Bayesian Information Criterion (BIC) and BIC gradient were graphed versus the number of clusters ranging from 2 to 12 shown in Figure 2a and 2b, respectively.

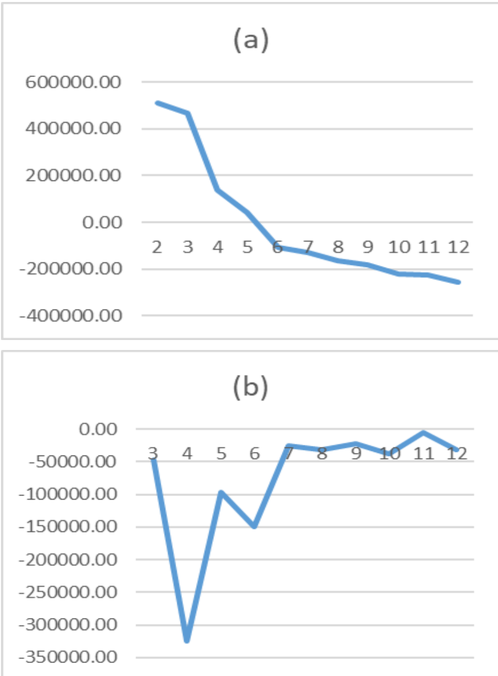

Figure 2. Number of Clusters vs. (a) BIC and (b) BIC Gradient

In Fig. 2a, a BIC closest to one signifies the best model. In evaluating the gradient in Fig. 2b, the optimal number of clusters is right before the upward trend levels off. In both graphs, the optimal number of clusters was five.

Fig. 3, a scatterplot between "pre-operative bilirubin" and "heart rate," shows clear separation of five clusters.

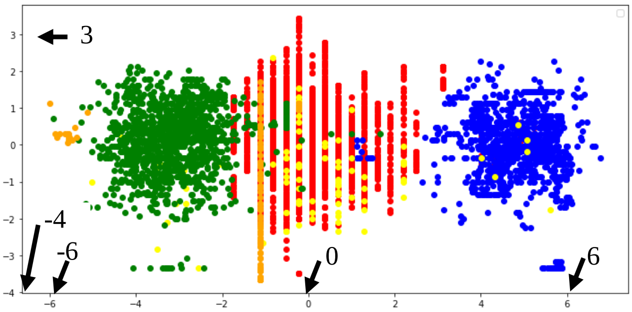

Figure 3. Preoperative Bilirubin vs. Preoperative Heart Rate. Cluster 1: Red; 2: Blue; 3: Yellow; 4: Green, and 5: Orange.

Results in Table II suggest using these five clusters (Clusters 1-5 in the table) yields better prediction than using the amount of records of lab work-up (Lab workup-low, medium and high in the table). We used regular ANN in this comparison study. Table II shows consistent lower average test accuracy and false positive and negative percentages when using our identified clusters.

TABLE II
CLASSIFICATION ACCURACY BY RISK GROUP AND CLUSTER

Clustering	N	Avg. Test Accuracy	False Positive	False Negative
Lab workup-Low	1,253	85.1%	11%	4%
Lab workup -Medium	1,096	95.2%	3%	1%
Lab workup-High	4,884	90.2%	7%	3%
1	4,620	**89.5%**	10%	1%
2	1,090	**99.2%**	0.4%	0.5%
3	96	**99.0%**	0%	1%
4	1,264	**97.3%**	1%	1%
5	163	**97.0%**	2%	1%

C. *Computational Results of INN with Mixed Responses*

A Synthetic Minority Oversampling Technique (SMOTE) method was used to balance the hypotensive and non-hypotensive cases since only 7% of the patient population had hypotensive readings. Besides, stratified *k*-fold cross validation splitted the data to preserve the percentage of samples for each class. Within each fold, SMOTE was applied to the training data.

Prior to fitting the model, the AdamWarmup optimizer was used for a warmup set to 0.01 and a decay set to 0.0001. Total steps and warmup steps for the optimizer was done using a warmup proportion of 0.1 and 1,000 epochs. The batch size was set to 10% of the number of samples in each cluster. The INN models were compiled using this optimizer with a loss set to binary cross entropy and metrics set to accuracy. The model was then fit to the training data with 3,000 epochs. The testing accuracy was averaged over the five folds.

In the INN, the output layer was customized to predict three response variables or a mixed response as described in Section IV. F. The INN activation function was changed in hidden layer 2 to Rectified Linear Unit (ReLU) from the sigmoid function. In addition to the activation and output changes, the loss function was changed to include a binary cross entropy loss for the binary response variable and the mean squared error loss for the continuous response variables. This new loss function, as in Equation (2), assigns weights to account for lack of binary hypotensive readings:

$$Loss = L_{CE} + wL_{MSE}, \qquad (2)$$

where L_{CE} is the binary cross entropy loss that ensures the classification accuracy; L_{MSE} is the mean squared error loss to enforce the regression accuracy for continuous variables; and w is the weight to balance the classification and regression performance. The weight was set to 100 in this study.

The INN architecture was modified to incorporate the number of input variables and rule assignments. Each rule was assigned no more than four input variables. All five clusters used the same rules involving 50 input variables and 15 rules. Specific rules and involving variables are shown in Table III. Note that some variables may be adopted in more than one rules. For example, the Preop Temperature variables were used in both Rules 2, 6 and 7.

TABLE III
LABWORK GROUP RULES FOR THE INN

Rule	**Variables**
1	Age
2	Preop Temperature, SBP and DBP
3	Preop Heart Rate, Oxygen Saturation and Respiration Rate
4	Preop Labs – Comprehensive blood counts
5	Oxygen Saturation – min, max, mode
6	Temperature – min, mean, median, range
7	Temperature - standard deviation, variance, entropy
8	Pulse Rate – min, max, variance, entropy
9	Respiratory Rate - max, mode, standard deviation, entropy
10	SBP and DBP Delta

These rules are integrated into the neural network in the mixed response prediction as illustrated in Fig. 4, where the decision thresholds were estimated by learning from data during the training phase of INN.

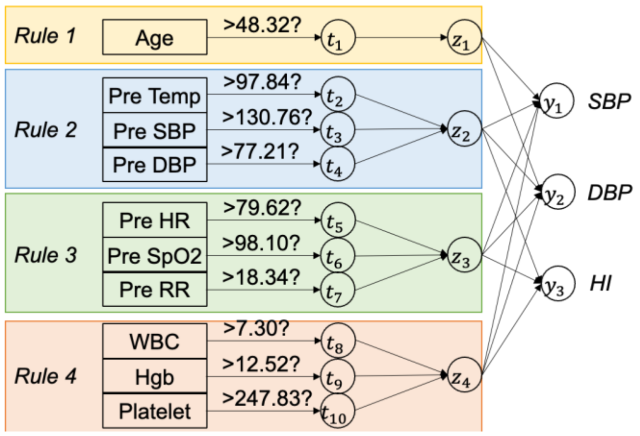

Figure 4. Rule assignments integrated in the INN

Table IV presents the testing results for INN for predicting the continuous responses, i.e., SBP and DBP. From Table IV, the prediction quality for SBP and DBP are similar with the average root mean square error (RMSE) in the range of 10~25. Although these are not ideal with RMSE, the benefits

ISBN: 978-988-14049-4-7
ISSN: 2078-0958 (Print); ISSN: 2078-0966 (Online)

of having interpretable rules as presented in Fig. 4 may outweigh the slightly higher errors. After all, our predictions are envisioned to be a guide to clinician's medical decision making. Therefore, interpretability is rather important.

TABLE IV
INN results on Continuous Prediction

Method	Cluster	SBP Avg. RMSE	DBP Avg. RMSE
INN	1	16.40	11.80
	2	13.60	9.60
	3	25.80	19.40
	4	15.20	12.20
	5	20.00	9.40

Table V displays two INN prediction results on binary classification of hypotension. One is from the mixed response model, and the other from a pure binary prediction model. From the table, the INN with mixed response yielded the testing accuracy within the range of 92~95%. Additionally, the mixed model outperformed the pure binary model for Clusters 1, 2 and 5. We attribute this superior performance to the ability to exploit the hidden association between continuous response (systolic and diastolic blood pressure) and hypotension classification.

TABLE V
INN results on Binary Prediction vs. Mixed Response Prediction

Method	Cluster	Mixed Response Avg. Test Accuracy	Binary Response Avg. Test Accuracy
INN	1	**93.8%**	78.1%
	2	**94.9%**	92.9%
	3	92.7%	**93.8%**
	4	**91.6%**	78.8%
	5	93.9%	**95.7%**

VI. Conclusion

This paper develops a new hypotension prediction model using the "interpretable neural network," where human domain knowledge is incorporated (as rules) into the neural networks and greatly improves the interpretability of the prediction. Specifically, we propose a novel data pipeline to address missing values and heterogeneous distributions of input variables that are common in perioperative medical data. The pipeline includes the use of LASSO for feature selection, GMM for clustering, PCA for feature reduction and missing value imputation, one-step ahead forecasting and the INN with a mixed response. The paper also demonstrates the use of a mixed response model to predict perioperative hypotension, where binary and continuous variables co-exist.

Computational results based on 1,463 hysterectomy cases are encouraging. First, the GMM produced reliable clustering results, compared to lab work-up based clustering. Second, the novel INN method was successfully applied to the hypotension prediction, providing high quality forecast (e.g., binary response with a testing accuracy of 92~95%) and high interpretability for the prediction. Third, the mixed response model outperformed the pure binary model due to exploitation of the hidden association between hypotension and the actual measures of DBP and SBP.

One future research is to develop the INN using a fully connected architecture. Each input variable would comprise one rule and the optimized thresholds would reflect the logic that anesthesiologists use in practice. Furthermore, acquiring intra-operative data in one-minute intervals will deliver a better prediction window in the first 15 minutes of surgery, a crucial period that has much greater clinical impact.

References

[1] Lin, C. S., Chang, C. C., Chiu, J. S., Lee, Y. W., Lin, J. A., Mok, M. S., & Li, Y. C. (2011). Application of an artificial neural network to predict postinduction hypotension during general anesthesia. Medical Decision Making, 31(2), 308-314.
[2] Wijnberge, M., Geerts, B. F., Hol, L., Lemmers, N., Mulder, M. P., Berge, P. & Veelo, D. P. (2020). Effect of a machine learning–derived early warning system for intraoperative hypotension vs standard care on depth and duration of intraoperative hypotension during elective noncardiac surgery: the HYPE randomized clinical trial. Jama, 323(11), 1052-1060.
[3] Kendale, S., Kulkarni, P., Rosenberg, A. D., & Wang, J. (2018). Supervised machine-learning predictive analytics for prediction of postinduction hypotension. Anesthesiology, 129(4), 675-688.
[4] Saugel, B., Kouz, K., Hoppe, P., Maheshwari, K., & Scheeren, T. W. (2019). Predicting hypotension in perioperative and intensive care medicine. Best Practice & Research Clinical Anaesthesiology, 33(2), 189-197.
[5] Briganti, G., & Le Moine, O. (2020). Artificial intelligence in medicine: today and tomorrow. Frontiers in medicine, 7, 27.
[6] Chen, X., & Jin, R. (2020). Adapipe: A recommender system for adaptive computation pipelines in cyber-manufacturing computation services. IEEE Transactions on Industrial Informatics, 17(9), 6221-6229.
[7] Kang, A. R., Lee, J., Jung, W., Lee, M., Park, S. Y., Woo, J., & Kim, S. H. (2020). Development of a prediction model for hypotension after induction of anesthesia using machine learning. PloS one, 15(4), e0231172.
[8] Hatib, F., Jian, Z., Buddi, S., Lee, C., Settels, J., Sibert, K., ... & Cannesson, M. (2018). Machine-learning algorithm to predict hypotension based on high-fidelity arterial pressure waveform analysis. Anesthesiology, 129(4), 663-674.
[9] Hegde, H., Shimpi, N., Panny, A., Glurich, I., Christie, P., & Acharya, A. (2019). MICE vs PPCA: Missing data imputation in healthcare. Informatics in Medicine Unlocked, 17, 100275.
[10] Le, T. D., Beuran, R., & Tan, Y. (2018, November). Comparison of the most influential missing data imputation algorithms for healthcare. In 2018 10th International Conference on Knowledge and Systems Engineering (KSE) (pp. 247-251). IEEE.
[11] Guo, Y., & Chen, H. (2020). Fault diagnosis of VRF air-conditioning system based on improved Gaussian mixture model with PCA approach. International journal of refrigeration, 118, 1-11.
[12] Budiarto, A., Mahesworo, B., Hidayat, A. A., Nurlaila, I., & Pardamean, B. (2021). Gaussian Mixture Model Implementation for Population Stratification Estimation from Genomics Data. Procedia Computer Science, 179, 202-210.
[13] Chen, X., Zeng, Y., Kang, S., & Jin, R. (2022). INN: An Interpretable Neural Network for AI Incubation in Manufacturing. ACM Transactions on Intelligent Systems and Technology (TIST).
[14] Kang, L., Kang, X., Deng, X., & Jin, R. (2018). A Bayesian hierarchical model for quantitative and qualitative responses. Journal of Quality Technology, 50(3), 290-308.
[15] Hwang, B. S., & Pennell, M. L. (2014). Semiparametric Bayesian joint modeling of a binary and continuous outcome with applications in toxicological risk assessment. Statistics in medicine, 33(7), 1162-1175.
[16] Tibshirani, R. (1996). Regression shrinkage and selection via the lasso. Journal of the Royal Statistical Society: Series B (Methodological), 58(1), 267-288.
[17] Dray, S., & Josse, J. (2015). Principal component analysis with missing values: a comparative survey of methods. Plant Ecology, 216(5), 657-667.
[18] Wang, Y., & Liu, Q. (2006). Comparison of Akaike information criterion (AIC) and Bayesian information criterion (BIC) in selection of stock–recruitment relationships. Fisheries Research, 77(2), 220-225.
[19] Kantardzic, M. (2011). Data mining : concepts, models, methods, and algorithms (Second). IEEE Press.
[20] Kirkland, L. A., Kanfer, F., & Millard, S. (2015, January). LASSO tuning parameter selection. In Annual Proceedings of the South African Statistical Association Conference (Vol. 2015, No. con-1, pp. 49-56). South African Statistical Association (SASA).