Temporal Trends and Survival Determinants in Capnography: A Machine Learning Study

Banu Priya Raya Krishnamoorthy

Graduate Program in Computer Science and Engineering

The Ohio State University, 2024

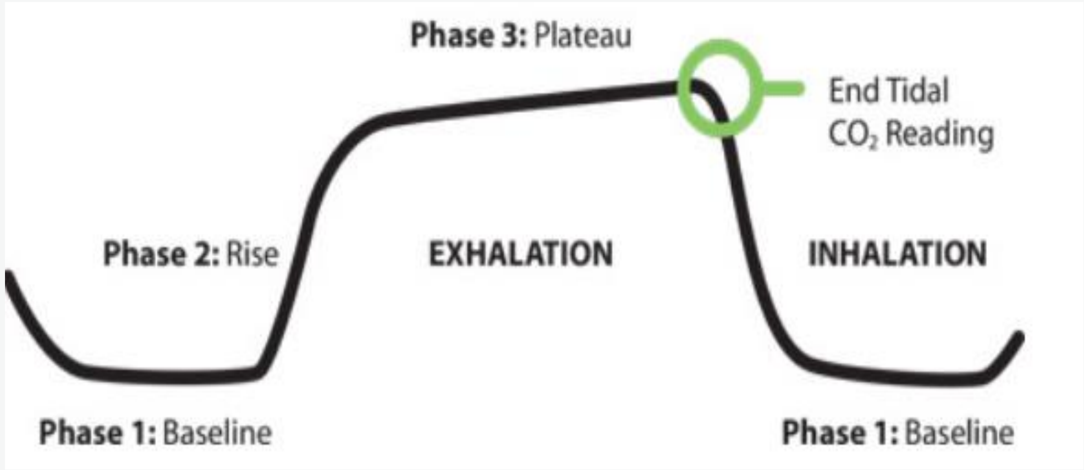
Dr. Rajiv Ramnath, Advisor

Dr. Michelle Nassal, Principal Investigator

Dr. Emre Ertin, Project Mentor

End Tidal Capnography & ROSC

- End-tidal capnography waveform is a simple graphic measurement of how much CO_2 a person is exhaling.
- The measurement at the end of the tide of respiration, the peak measurement at the very end of phase 3, is the **$EtCO_2$** reading.
- This measurement is widely used in emergency and critical care to assess ventilation, perfusion, and the effectiveness of resuscitation.



ROSC stands for "Return of Spontaneous Circulation'', it refers to the moment when a heart's rhythm resume and blood begins to flow through the body after cardiac arrest.

https://www.jems.com/patient-care/airway-respiratory/how-to-read-and-interpret-end-tidal-capnography-waveforms/

Problem Statement

- Every year, half a million Americans suffer from out-of-hospital cardiac arrest (OHCA), with survival rates as low as 7-10%.
- Current resuscitation algorithms apply fixed interventions at set time intervals, without considering individual patient characteristics or evolving cardiac conditions.
- This highlights the urgent need for more adaptive and personalized resuscitation strategies.

Proposed Solution

Investigation of $EtCO_2$ Variations as Predictors of ROSC

- Analyzing temporal variations in $EtCO_2$ during resuscitation.
- Assessing patient responses to epinephrine as predictors of Return of Spontaneous Circulation (ROSC).
- Using statistical methods and machine learning to identify critical survival indicators.

Creation of Medical Datasets Optimized for AI Processing

- Building datasets for $EtCO_2$-guided resuscitation.
- Enhancing real-time clinical decision-making through AI-driven predictive models.

Dataset

Source: Pragmatic Airway Resuscitation Trial (PART).

Population: Adult patients (18+ years) with non-traumatic out-of-hospital cardiac arrest (OHCA) from 27 EMS agencies across 5 communities.

Data Collection Period: December 1, 2015 – November 4, 2017.

Sample Size: 3004 patients.

Demographic Breakdown:

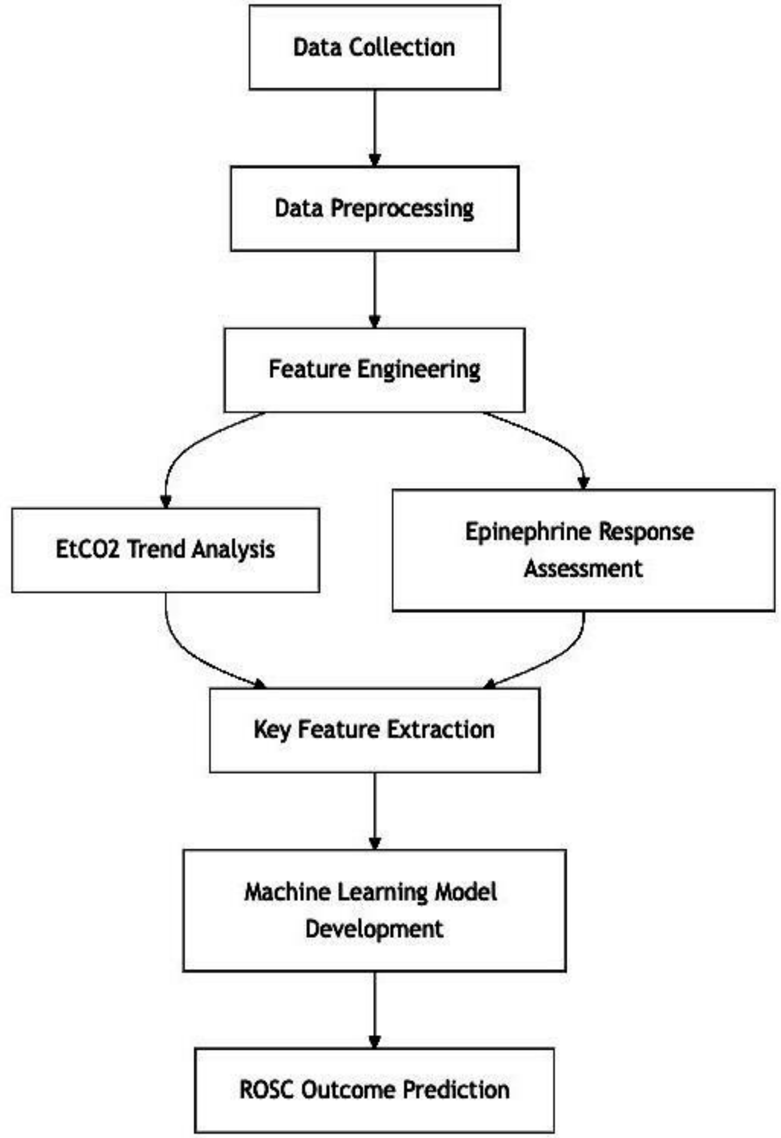
- Gender: 62.4% males, 37.6% females.
- Race: 25.6% Black/African American, 53.2% White, 21.2% other.

Attributes of the Dataset

Column Name	Description
caseid	Unique ID for the person.
surv72	Indicates whether the person survived or not. 0 = did not survive, 1 = survived.
mnrosc	Return of Spontaneous Circulation (ROSC). 0 = no ROSC, 1 = ROSC achieved. ROSC occurs when a patient's heart resumes a sustained rhythm after cardiac arrest.
EtCO2	End-tidal carbon dioxide ($EtCO_2$) measurement in exhaled air, used to guide care during cardiopulmonary resuscitation (CPR).
age_yrs	The patient's age in years.
sexp	Gender of the patient. Categories include Male, Female, and Unknown.
witbys	Indicates whether the cardiac arrest was witnessed. Categorical values: "Witnessed" and "Not Witnessed".
frhyem_shock	Indicates if the first rhythm was shockable or non-shockable.
bystander_cpr	Whether CPR was initiated by a bystander. 0 = no, 1 = yes.
bystander_aed	Whether AED was initiated by a bystander. 0 = no, 1 = yes.
last_succ_air_rev	The method used to establish the airway: "LT" for Laryngeal Tube, "ETI" for Endotracheal Tube.
Dose	The number of epinephrine doses administered. 1 = one dose, 2 = two or more doses.
epi_time	The time when the first dose of epinephrine was administered.
Race	Race of the patient. Categories include White, Black, and Other.

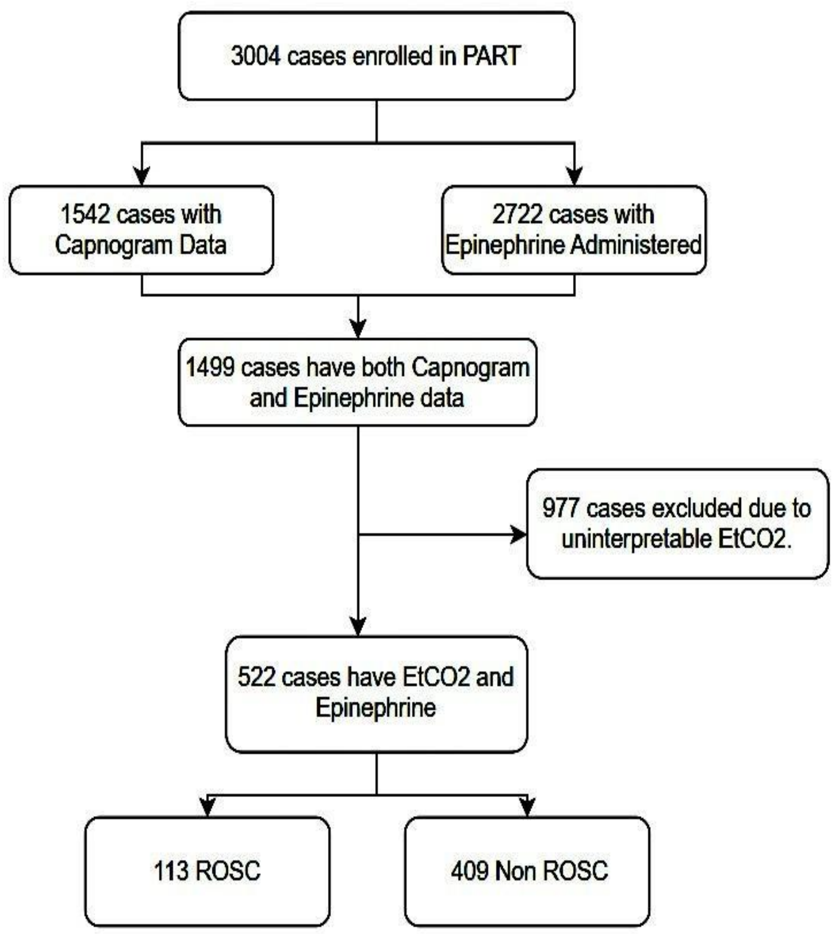
Project Pipeline

This flowchart illustrates the step-by-step process for predicting Return of Spontaneous Circulation (ROSC) using $EtCO_2$ trends, from data collection and preprocessing to feature extraction and machine learning model development for real-time clinical decision-making.



Data Preprocessing

- This flowchart illustrates the data cleaning process, where all non-interpretable $EtCO_2$ data was removed to ensure the accuracy and consistency of the dataset, minimizing discrepancies for the analysis.
- Categorical variables (sex, witnessed status, rhythm type) were encoded into numerical formats for model compatibility.
- Class imbalance between ROSC and non-ROSC cases was addressed using SMOTE(Synthetic Minority Oversampling Technique) to ensure balanced and accurate predictions.



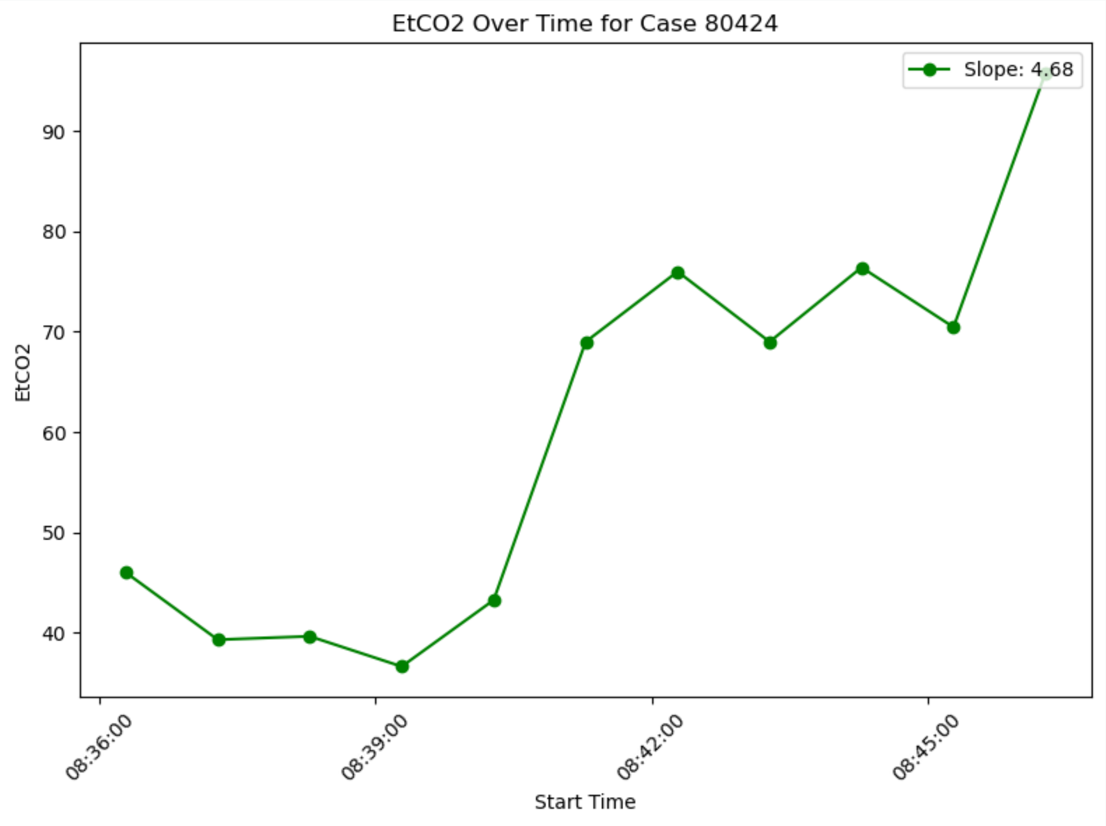
Feature Engineering – $EtCO_2$ Slope

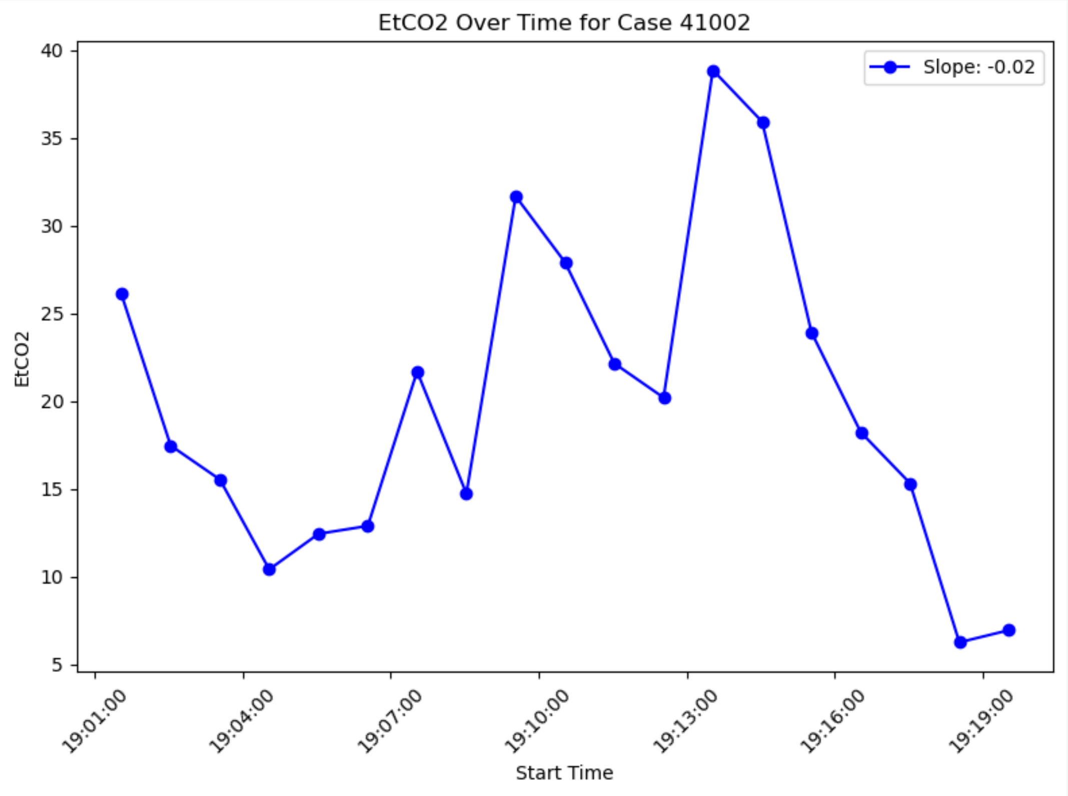
EtCO2 Slope Calculation: Analyzing the change in EtCO2 over time during resuscitation to predict ROSC.

Slope Formula: ΔEtCO2 = EtCO2(t2) - EtCO2(t1), where the slope indicates the rate of change in EtCO2, calculated over time intervals, typically 1 minute.

- The relationship between EtCO2 and time is modeled using **linear regression** (**polynomial fitting of degree 1**), which fits a straight line to the EtCO2 data points for each patient. The resulting slope indicates the average rate of change in EtCO2 per unit of time, providing insights into the patient's likelihood of achieving ROSC.

Result of Slope Calculation

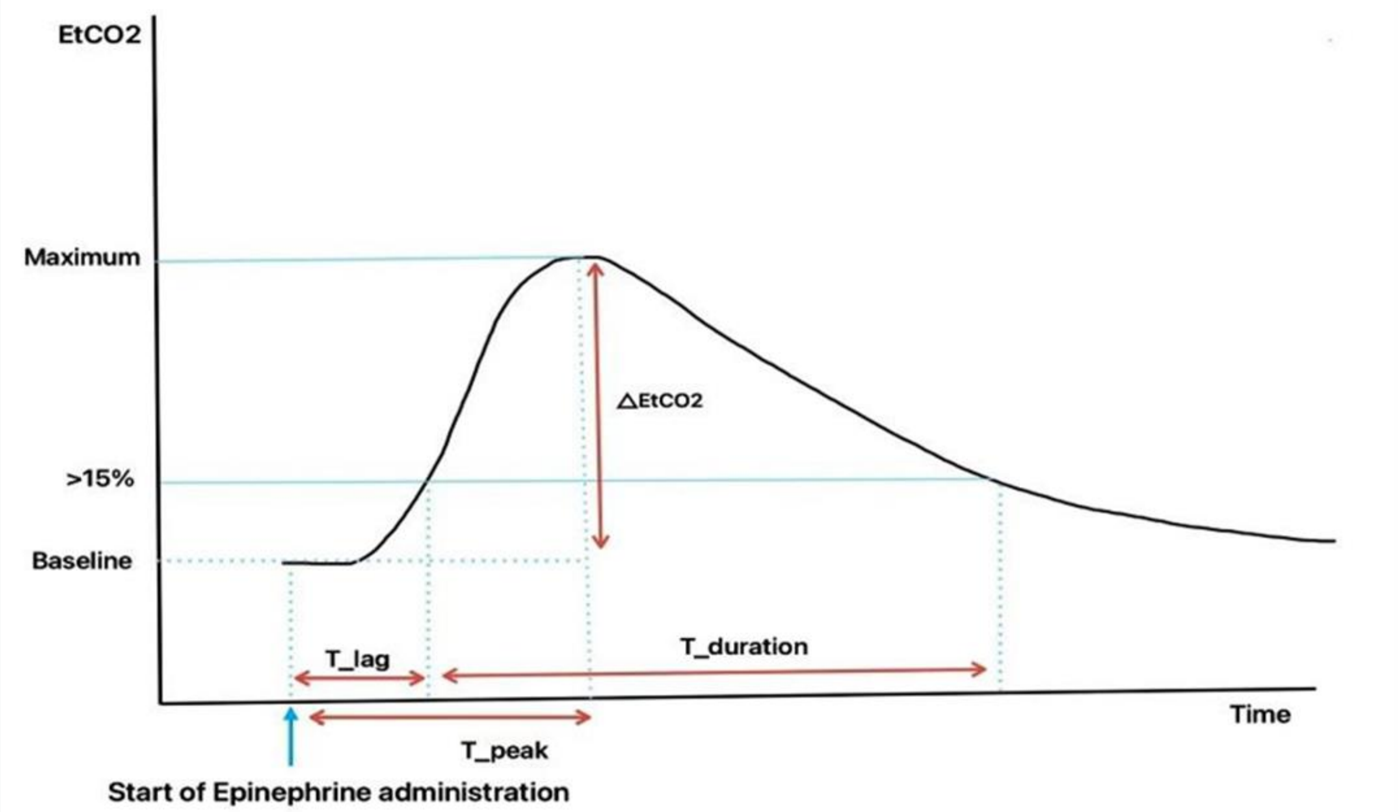




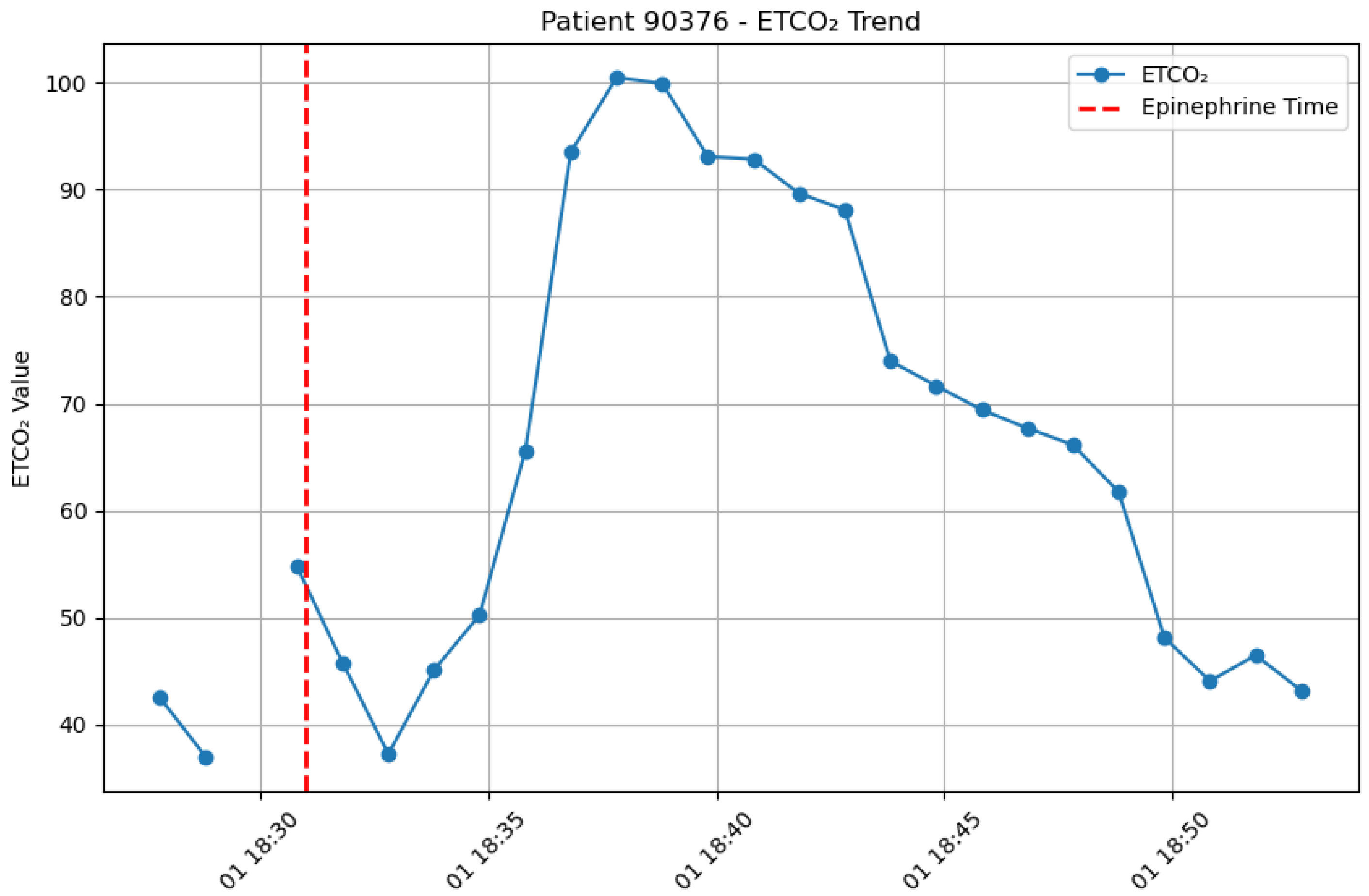
The slope of EtCO2 indicates the direction of a patient's resuscitation progress, with a positive slope suggesting improvement and a negative slope indicating worsening conditions.

Epinephrine’s Temporal Effects on EtCO2

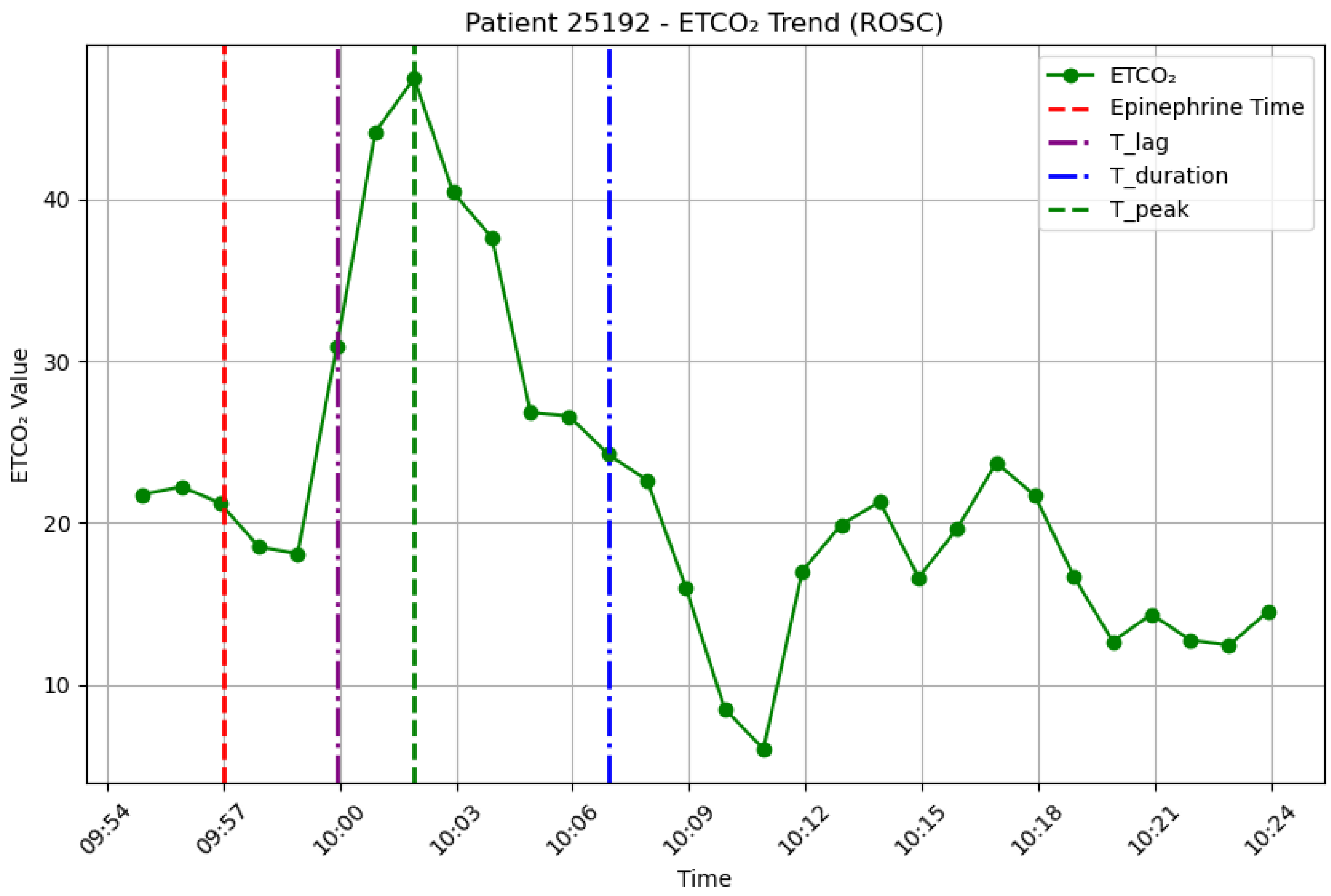
- Following Roh et al.'s methodology, this study investigates how epinephrine affects $EtCO_2$ dynamics during cardiac resuscitation, analogous to their research on sodium bicarbonate's impact.



Examples of $EtCO_2$ waveform Post Epinephrine



Example plot for a ROSC patient, showing the calculation of ΔEtCO2 peak, Tlag, Tpeak, and Tduration.



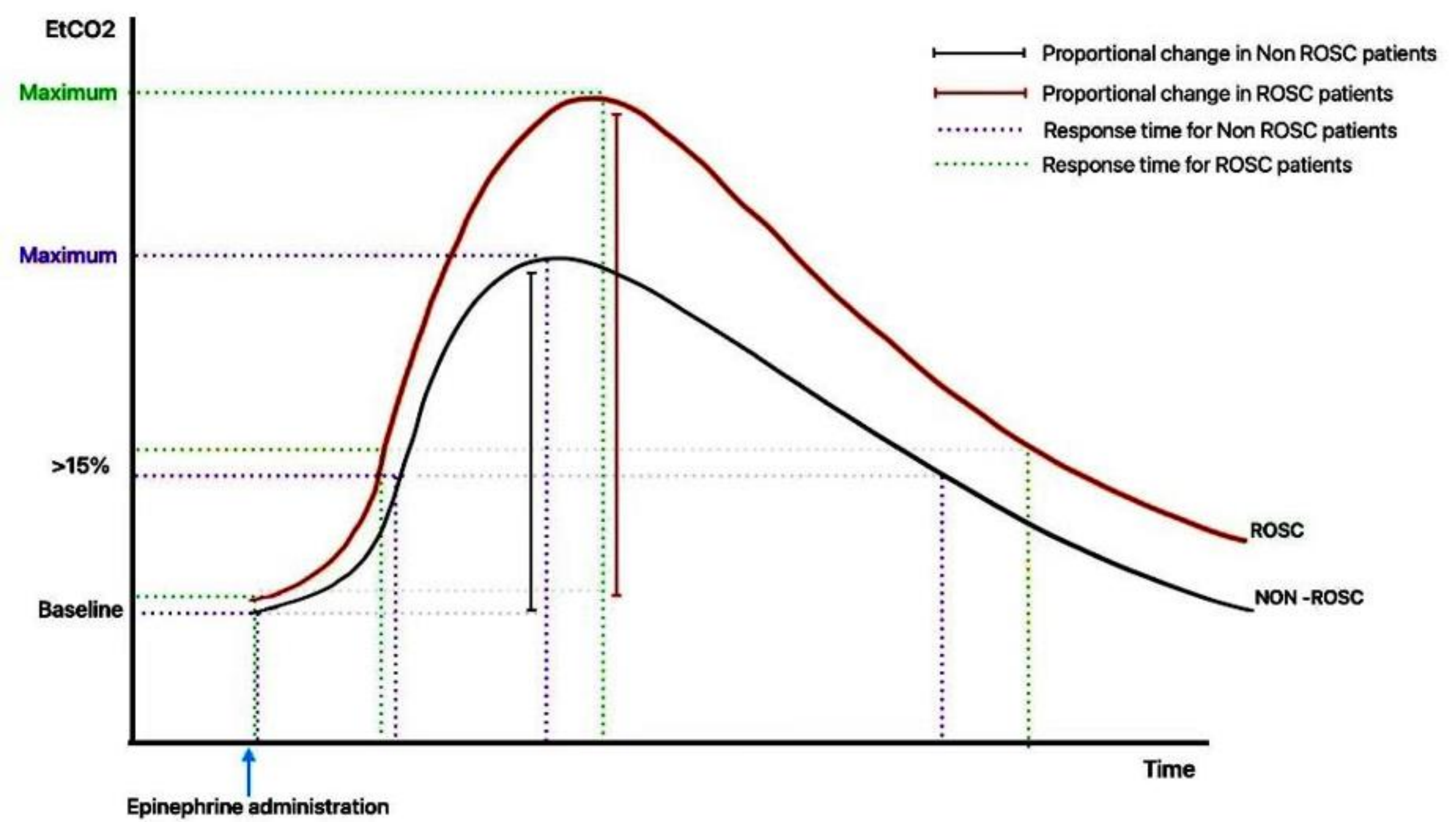
Observations

Variables	ROSC (n = 139)	Non-ROSC (n = 378)	p-value
T_{lag} (min)	1.87 (0.88–3.38)	1.93 (0.77–3.50)	0.177
T_{peak} (min)	3.43 (1.44–6.67)	2.70 (1.10–4.99)	0.129
$T_{duration}$ (min)	4.00 (0.50–10.00)	1.00 (0.50–5.75)	< 0.0001
Proportional Change (%)	43.33 (15.01–105.88)	24.67 (4.86–64.10)	0.00027

Table 2: Changes in end-tidal CO_2 after epinephrine administration.
(Data are expressed as medians and interquartile ranges.)

The Mann-Whitney test showed significant differences in **Tduration** and **Proportional Change**, while no significant differences were found for Tlag and Tpeak, with a p-value threshold of 0.05.

Schematic Representation of Post Epi Results for ROSC and Non ROSC group.



Feature Extraction

Along with the existing dataset attributes, additional features like slope, Tlag, Tpeak, Tduration, and proportional change were introduced.

Identifying important features was done using two methods:

Odds Ratio - Logistic Regression

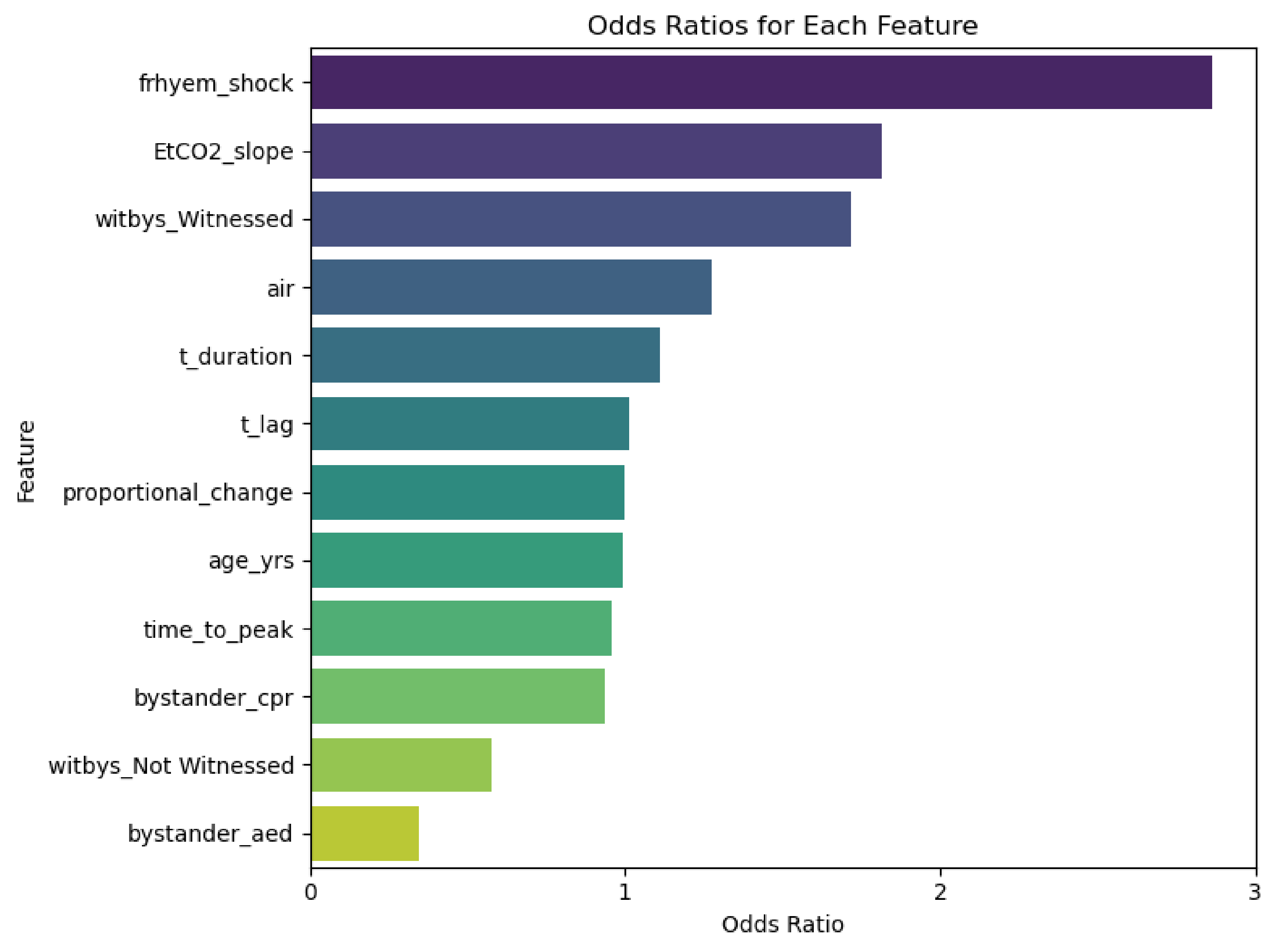
Odds ratio is obtained by exponentiating the regression coefficient from logistic regression, with the formula OR=eβ, where β is the estimated coefficient for the predictor variable.

Feature Importance - Random Forest

Obtained by evaluating how much each feature contributes to the reduction in model error, based on the improvement in prediction accuracy when splitting the data at each feature.

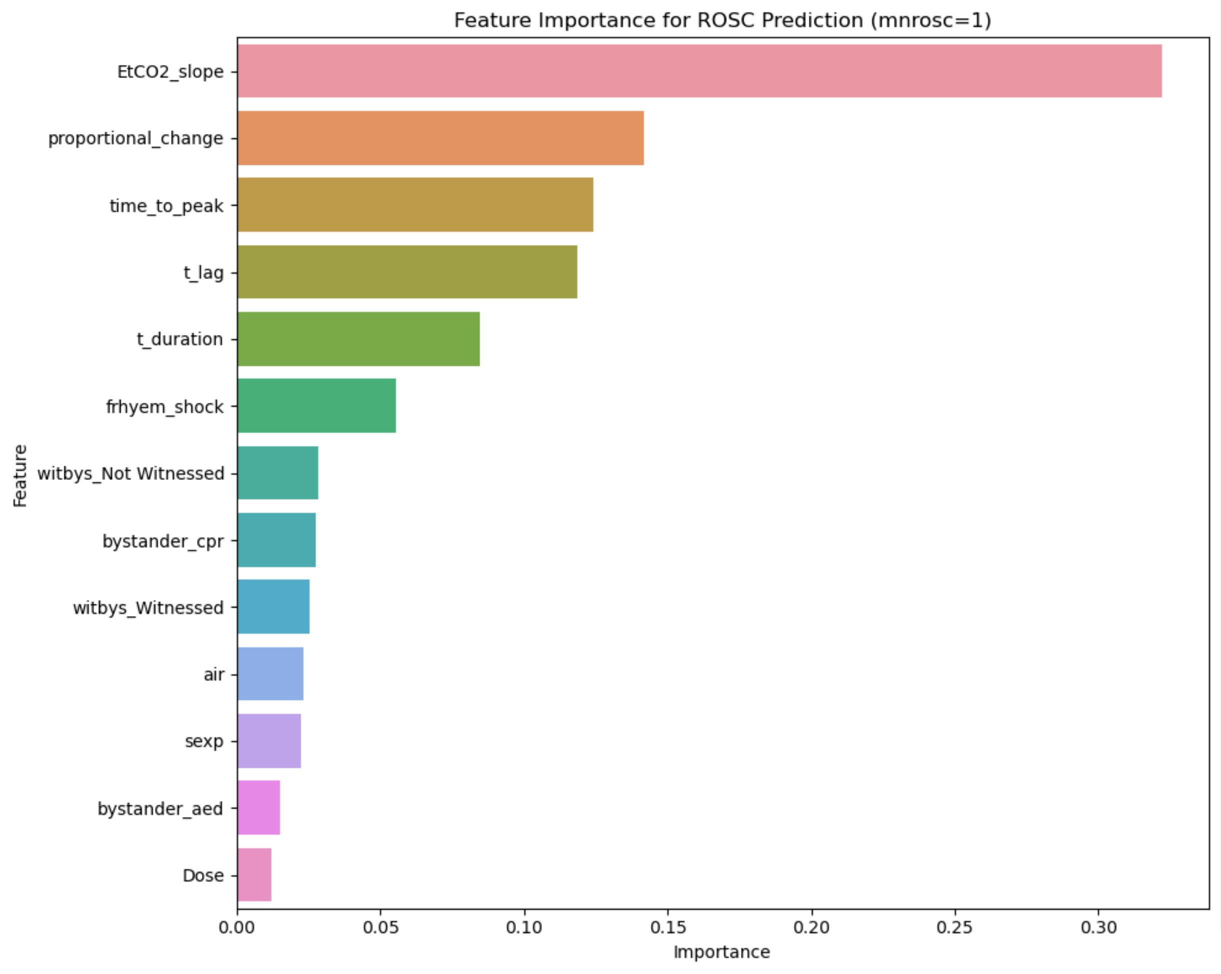
Odds Ratio

The chart illustrates that "frhyem_shock[2.8]“ and "EtCO2_slope[1.8]" are the most significant features, with higher odds ratios for predicting ROSC.



Feature Importance

- The first five features are all derived from EtCO2 values, emphasizing the significance of EtCO2 dynamics in predicting resuscitation outcomes, while the sixth feature, "fryhem_shock," pertains to the initial rhythm type.
- These results align with the findings from the odds ratio analysis.



Predictive Modelling for ROSC/Non ROSC

Model Selection:

- **Logistic Regression**: A linear model trained with balanced class weights for comparison.
- **Random Forest Classifier**: Trained using the preprocessed data, capturing complex, non-linear relationships.

Model Training:

- The models were trained using the processed dataset, ensuring the inclusion of features like **EtCO2 Slope**, **Tpeak**, **Proportional Change**, and other temporal EtCO2 metrics.

Model Evaluation:

- Performance was measured using **accuracy**, **precision**, **recall**, **F1-score**, and **AUC-ROC**.

The **ROC curve** was plotted for comparison between Random Forest and Logistic Regression.

Classification Report

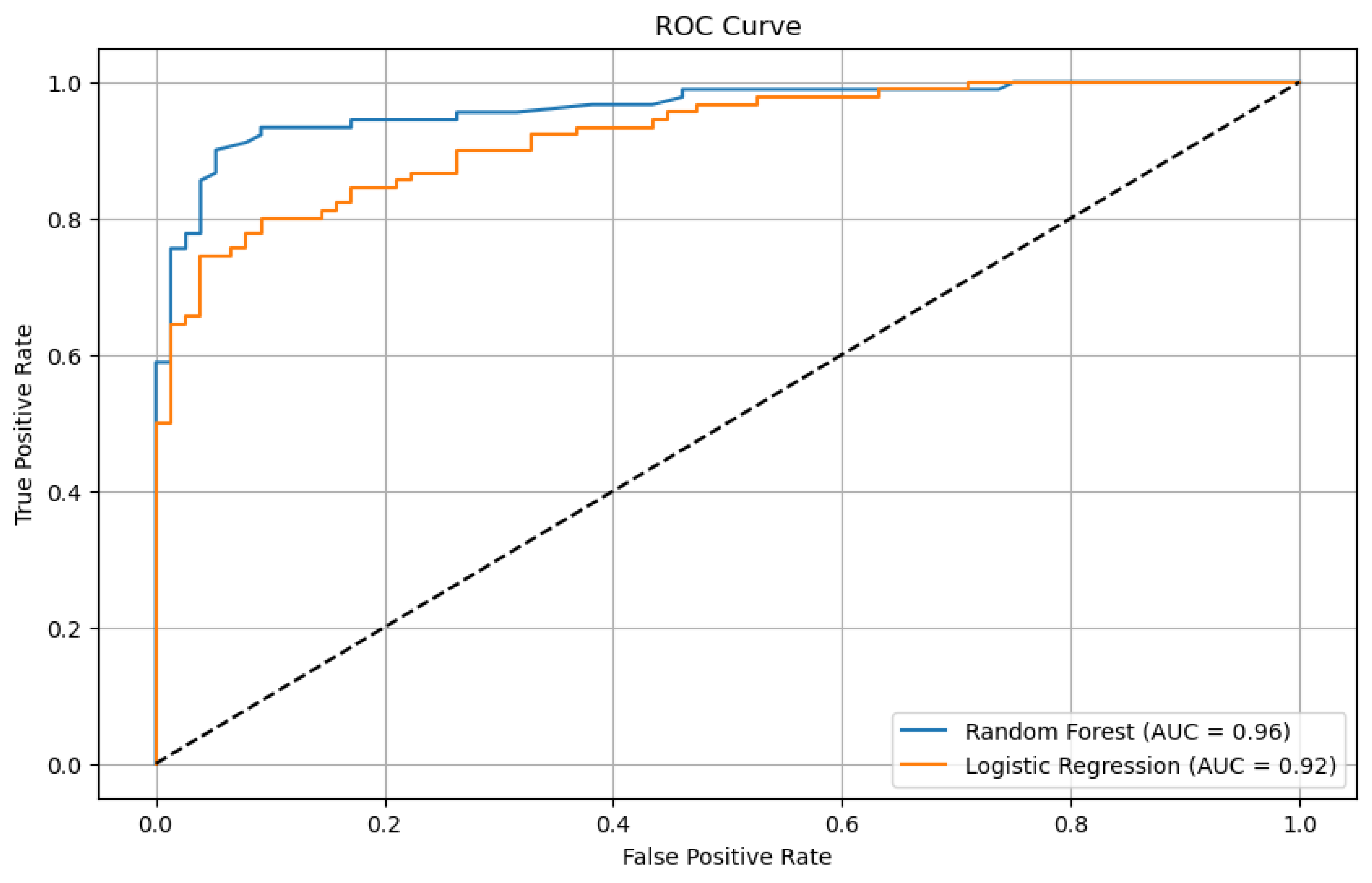
Logistic Regression [Accuracy: 85]

Class	Precision	Recall	F1-Score	Support
Non-ROSC (0)	0.79	0.91	0.85	76
ROSC (1)	0.91	0.80	0.85	90

Random Forest Classifier [Accuracy: 90]

Class	Precision	Recall	F1-Score	Support
Non-ROSC (0)	0.84	0.96	0.90	76
ROSC (1)	0.96	0.84	0.90	90

Results


ROC Curve
True Positive Rate
False Positive Rate
0.0
0.2
0.4
0.6
0.8
1.0
Random Forest (AUC = 0.96)
Logistic Regression (AUC = 0.92)

Flask Application for Cardiac Arrest Data Processing

Key Features:

Standardized Data Extraction: Harmonizes data from multiple organizations (e.g., Columbus Fire Department, Peel Data Center) into a consistent format.

Customized Filtering: Enables targeted analysis by selecting specific clinical interventions (e.g., epinephrine, defibrillation).

Data Integration: Consolidates disparate datasets into a unified structure with unique identifiers for each patient.

Code and Access:

The complete code and live application are deployed at banupr15.pythonanywhere.com/ ensuring efficient data preparation for the study on $EtCO_2$-guided resuscitation.

Limitations & Conclusion

- The dataset, collected between 2015 and 2017, may not fully reflect current medical practices or advancements in technology.
- The study only examined the first dose of epinephrine, missing potential variations in response to multiple doses.
- EtCO2 dynamics, especially the slope and proportional change, effectively predict ROSC and can guide clinical decisions during resuscitation.
- The Flask-based tool facilitates the efficient processing of medical data, enhancing real-time decision-making and future research applications.

Thank You