Supplementary Materials

ForecastICU: A Prognostic Decision Support System for Timely Prediction of Intensive Care Unit Admission

1. Proof of Theorem 1

In order to prove the Theorem, we first hold the following assumptions on the physiological stream stopping time and the patients' hospitalization time (time of admission to hospital), and hospitalization period (time between admission to hospital and transfer to ICU or discharge). We assume that the maximum hospitalization period for any patient is $\bar{T}_H$, the hospitalization time t_H is random, and the stopping time τ_s is random where the distributions of hospitalization and stopping times are given by $f_{t_H}(t_H)$, $f_{\tau_s}(\tau_s \,|\mathcal{H}_0)$ and $f_{\tau_s}(\tau_s \,|\mathcal{H}_1)$, where $\text{supp}\left(f_{t_H}(t_H)\right) = [0, \bar{T}_H]$, $\text{supp}\left(f_{\tau_s}(\tau_s \,|t_H)\right) = [t_H, \bar{T}_H]$.

Let B_t^* and B_t be the belief processes of a truthful and a non-truthful belief systems respectively. A truthful belief system has access to the joint distributions of the physiological data stream $(\mathbb{P}_0, \mathbb{P}_1)$ and knows the stopping time τ_s, whereas the non-truthful belief system maintains estimates of the joint distribution of the physiological data stream $(\mathbb{Q}_0, \mathbb{Q}_1)$, where $d\left(\mathbb{P}_m, \mathbb{Q}_m\right) > 0$ for a probability metric d. In the following, we show that both B_t^* and B_t are martingales with respect to the filtration $\mathcal{F}_t$. Note that

$$B_t^*\left(\mathcal{H}_1 \,|\mathcal{F}_t\right) = \frac{\mathbb{P}\left(\{X_\tau\}_{\tau=t_H}^{t} \,|\mathcal{H}_1\right)\mathbb{P}\left(\mathcal{H}_1\right)}{\sum_{i\in\{0,1\}}\mathbb{P}\left(\{X_\tau\}_{\tau=t_H}^{t} \,|\mathcal{H}_i\right)\mathbb{P}\left(\mathcal{H}_i\right)} = \frac{B_{t-1}^*\left(\mathcal{H}_1 \,|\mathcal{F}_{t-1}\right)\mathbb{P}\left(X_t \,|\mathcal{H}_1\right)}{\sum_{i\in\{0,1\}} B_{t-1}^*\left(\mathcal{H}_i \,|\mathcal{F}_{t-1}\right)\mathbb{P}\left(X_t \,|\mathcal{H}_i\right)}. \tag{1}$$

Thus, we have that

$$\begin{aligned}\mathbb{E}\left[B_{t+1}^* \,|\mathcal{F}_{t+1}\right] &= \mathbb{E}\left[\frac{B_t^*\left(\mathcal{H}_1 \,|\mathcal{F}_t\right)\mathbb{P}\left(X_t \,|\mathcal{H}_1\right)}{\sum_{i\in\{0,1\}} B_t^*\left(\mathcal{H}_i \,|\mathcal{F}_t\right)\mathbb{P}\left(X_t \,|\mathcal{H}_i\right)}\right] \\ &= \sum_{X_t\in\mathcal{X}_t} \frac{B_t^*\left(\mathcal{H}_1 \,|\mathcal{F}_t\right)\mathbb{P}\left(X_t \,|\mathcal{H}_1\right)\mathbb{P}(X_t)}{\sum_{i\in\{0,1\}} B_t^*\left(\mathcal{H}_i \,|\mathcal{F}_t\right)\mathbb{P}\left(X_t \,|\mathcal{H}_i\right)} \\ &= \sum_{X_t\in\mathcal{X}_t} B_t^*\left(\mathcal{H}_1 \,|\mathcal{F}_t\right)\mathbb{P}\left(X_t \,|\mathcal{H}_1\right) \\ &= B_t^*\left(\mathcal{H}_1 \,|\mathcal{F}_t\right)\sum_{X_t\in\mathcal{X}_t}\mathbb{P}\left(X_t \,|\mathcal{H}_1\right) \\ &= B_t^*\left(\mathcal{H}_1 \,|\mathcal{F}_t\right).\end{aligned} \tag{2}$$

Since $\mathbb{E}\left[B_{t+1}^* \,|\mathcal{F}_{t+1}\right] = B_t^*\left(\mathcal{H}_1 \,|\mathcal{F}_t\right)$, then the truthful belief process is martingale. Now we focus on the non-truthful belief process B_t, which we can write as

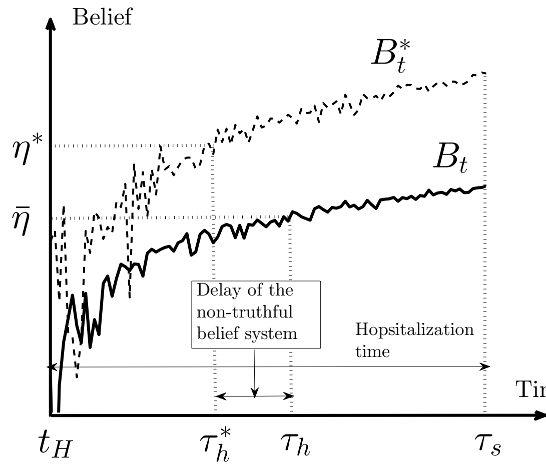

Figure 1. Depiction for the belief process as computed by a truthful and a non-truthful belief systems.

$$B_t\left(\mathcal{H}_1 \,|\mathcal{F}_t\right) = \frac{B_{t-1}\left(\mathcal{H}_1 \,|\mathcal{F}_{t-1}\right)\mathbb{Q}\left(X_t \,|\mathcal{H}_1\right)}{\sum_{i\in\{0,1\}} B_{t-1}\left(\mathcal{H}_i \,|\mathcal{F}_{t-1}\right)\mathbb{Q}\left(X_t \,|\mathcal{H}_i\right)}. \tag{3}$$

Thus, we have that

$$\begin{aligned}\mathbb{E}\left[B_{t+1} \,|\mathcal{F}_{t+1}\right] &= \mathbb{E}\left[\frac{B_t\left(\mathcal{H}_1 \,|\mathcal{F}_t\right)\mathbb{Q}\left(X_t \,|\mathcal{H}_1\right)}{\sum_{i\in\{0,1\}} B_t\left(\mathcal{H}_i \,|\mathcal{F}_t\right)\mathbb{Q}\left(X_t \,|\mathcal{H}_i\right)}\right] \\ &= \sum_{X_t\in\mathcal{X}_t} \frac{B_t\left(\mathcal{H}_1 \,|\mathcal{F}_t\right)\mathbb{Q}\left(X_t \,|\mathcal{H}_1\right)\mathbb{P}(X_t)}{\sum_{i\in\{0,1\}} B_t\left(\mathcal{H}_i \,|\mathcal{F}_t\right)\mathbb{Q}\left(X_t \,|\mathcal{H}_i\right)} \\ &= B_t\left(\mathcal{H}_1 \,|\mathcal{F}_t\right).\end{aligned} \tag{4}$$

Now define the threshold type strategies η^* (a threshold on B_t^*) and $\bar{\eta}$ (a threshold on B_t) as follows:

$$\eta^* = \arg\sup_{\eta\in[0,1]} \mathbb{E}_{\mathbb{P}}\left[g\left(\{X_\tau\}_{\tau=0}^{\tau_h(\eta)}\right)\mathbf{1}_{\{\tau_h(\eta)<\tau_s\}}\right], \tag{5}$$

and

$$\bar{\eta} = \arg\sup_{\eta\in[0,1]} \mathbb{E}_{\mathbb{Q}}\left[g\left(\{X_\tau\}_{\tau=0}^{\tau_h(\eta)}\right)\mathbf{1}_{\{\tau_h(\eta)<\tau_s\}}\right]. \tag{6}$$

As shown in Fig. 1, the non-truthfulness of the belief system may lead for instance to a delay in the ICU alarm.

Now we focus on a certain realization of the stopping time τ_s. Since $\sup_t B_t = \sup_t B_t^* = 1$, and $\inf_t B_t = \inf_t B_t^* = 0$, then $\mathbb{E}\left[B_t\right] < \infty$ and $\mathbb{E}\left[B_t^*\right] < \infty$, i.e. B_t and B_t^* are bounded martingales. Thus, by Doob's martingale convergence theorem, we know that $B_t \rightarrow B_\infty$ and $B_t^* \rightarrow B_\infty^*$ almost surely, where $\mathbb{E}\left[B_\infty^*\right] < \infty$, and $\mathbb{E}\left[B_\infty\right] < \infty$. It is easy to show that the sequence $B_t^* - B_t$ is also a martingale with respect to the filtration $\mathcal{F}_t$, i.e. $\mathbb{E}\left[B_t^* + 1 - B_t + 1 \,|\, \mathcal{F}_t\right] = B_t^* - B_t$. Now recall that we want to show that $\mathbb{P}\left(|V^* - V(\bar{\eta})| < \epsilon\right) > 1 - \delta$. To prove this, it suffices to show that there exists $\epsilon^{'} \in [0, 1]$, such that $\mathbb{P}\left(|\eta^* - \bar{\eta}| < \epsilon^{'}\right) > 1 - \delta$. This is equivalent to show that the martingale sequence $B_t^* - B_t$ converges to a value less than $\epsilon^{''} \in [0, 1]$ with a probability $1 - \delta$. This is satisfied if for $N^*(\epsilon, \delta)$, there exists an algorithm $\mathcal{A}^D$ that if used to estimate $\mathbb{Q}$, it will prompt a distribution that is within a Kolmogorov-Smirnov distance of $\Delta(\epsilon)$ from the true distribution $\mathbb{P}$. By Dvoretzky-Kiefer-Wolfowitz inequality, we know that if the algorithm $\mathcal{A}^D$ just computes $\mathbb{Q}$ as the empirical distribution, then we have that

$$\Pr\left(\sup_{t \in [t_H, t_H + \tau_s]} \left|\mathbb{Q}_m^t - \mathbb{P}_m^t\right| > \Delta(\epsilon)\right) \leq 2\exp\left(-2N\Delta^2(\epsilon)\right).$$

Thus, we can find $N^*(\epsilon, \delta)$ by equating $1 - \delta$ with the RHS in the equation above, and for any $N > N^*(\epsilon, \delta)$, we have that $\mathbb{P}\left(|V^* - V(\bar{\eta})| < \epsilon\right) > 1 - \delta$.

2. Pseudo-code of ForecastICU

Offline Stage:
Input: $\mathbf{X}_0^{ref}, \mathbf{X}_1^{ref}, T_s$
1) Data Reconstruction
for i = 1 **to** N **do**
$\tilde{\mathbf{X}}_{(i)}^{ref} = h_{spline}(\{\mathbf{X}_{(i)}^{ref}(m, n)\}_{n=0}^{K-1}, T_s)$
end for
2) Relevant Feature Selection
$\tilde{\mathbf{Y}}^{ref} = CFS(\tilde{\mathbf{X}}^{ref})$,
3) Parametric density estimation
$[\hat{\mu}_m^t(j)]_{j=1}^R = \frac{1}{N_m}\sum_{i=1}^{N_m} \tilde{\mathbf{Y}}_{(i),m}^{ref}(j, t)$
$[\hat{\mathbf{\Sigma}}_m^t]_{k,l} = \frac{1}{N_m - 1}\sum_{i=1}^{N_m} \bar{\mathbf{Y}}_{(i),m}^{ref}(k, t)\bar{\mathbf{Y}}_{(i),m}^{ref}(l, t)$

Real-time Stage:
Input: $\{\mathbf{X}_\tau\}_{\tau=0}^{T^H}, \gamma, \eta, W$
for t = 1 **to** T^H **do**
1) Current State Estimation
for m = 0 **to** 1 **do**
$T_m^*(t) = \arg\max_\tau \mathbb{Q}_m(\{X_k\}_{k=\tau-t+1}^{\tau} | \mathcal{H}_m)$
end for
2) Belief Update Algorithm
$B_t(\mathcal{H}_1|\mathcal{F}_t) = \mathbb{Q}_1(\mathcal{H}_1|\{X_\tau\}_{\tau=t_o}^{t})$
$= \frac{N_1\mathbb{Q}_1(\{X_\tau\}_{\tau=t_o^1}^{t}|\mathcal{H}_1, T_1^*(t))}{N_0\mathbb{Q}_0(\{X_\tau\}_{\tau=t_o^0}^{t}|\mathcal{H}_0, T_0^*(t)) + N_1\mathbb{Q}_1(\{X_\tau\}_{\tau=t_o^1}^{t}|\mathcal{H}_1, T_1^*(t))}$
$\tilde{B}_t(\mathcal{H}_1|\mathbb{F}_t) = \frac{1}{W}\sum_{\tau=t-W}^{t} B_\tau(\mathcal{H}_1|\mathcal{F}_\tau)$
3) Sequential Decision Making
$\text{Decision}(t) = \begin{cases} \mathcal{H}_1 & \text{if } \hat{B}_t(\mathcal{H}_1|\mathcal{F}_t) \geq \eta \\ \mathcal{H}_0 & \text{otherwise} \end{cases}$
end for

Figure 2. Pseudo-code of ForecastICU

3. Features of the Dataset

3.1. Entire feature information

Table 1. Entire feature information

NO	FEATURE NAME
	Time Dependent Continuous Features
1	SYSTOLIC BLOOD PRESSURE
2	DIASTOLIC BLOOD PRESSURE
3	HEART RATE
4	RESPIRATORY RATE
5	TEMPERATURE
6	O2 SATURATION
7	WHITE BLOOD CELL
8	HEMOGLOBIN
9	PLATELET COUNT
10	SODIUM
11	POTASSIUM
12	CHLORINE
13	CO2 SATURATION
14	BLOOD UREA NITROGEN
15	CREATINE
16	GLUCOSE
	Time Dependent Discrete Features
17	O2 DEVICE (BINARY)
18	BREATH ASSIST DEVICE (49 CATEGORIES)

3.2. Relevant Features for ICU Admission Prediction

Table 2. Relevant features for ICU admission prediction

Rank	Acronym	Relevant Features
1	RR	Respiratory Rate
2	HR	Heart Rate
3	BUN	Blood Urea Nitrogen
4	GLU	Glucose
5	Breath	Oxygen Supply Device (Binary)
6	DBP	Diastolic Blood Pressure
7	SPO2	O2 Saturation

Based on the correlation feature selection (CFS) algorithm with minimum redundancy and maximum relevance (mRMR) criterion, we discover 7 relevant temporal features among the entire 18 temporal features which are highly correlated with ICU admission but poorly correlated with each other. Table 2 explicitly lists 7 relevant features and these can be justified by the medical references (Andrew Egol, 1999) (Bruijns, 2013) (Alexander Olaussen, 2014). Note that all of the relevant features are time dependent features.

4. Model Justifications

In this paper, we also assume that the joint distribution of the physiological data streams can be modeled as a Multivariate Gaussian process. This assumption is validated by a Kolmogorov-Smirnov goodness-of-fit test. Fig. 3 illustrates the histogram of the systolic blood pressure and heart rate extracted by the reconstructed dataset of ICU and DIS patients, respectively. As it can be seen, these can be indeed modeled as Gaussian distributions - the fitting error is less than 10%. Fig. 4 shows that the joint distributions between the physiological features can indeed be modeled using a Multivariate Gaussian distribution.

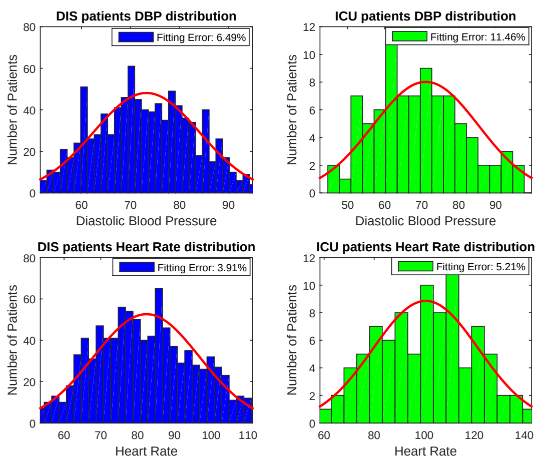

Figure 3. Histograms of diastolic blood pressures and heart rates at 10 hours before ICU/DIS events.

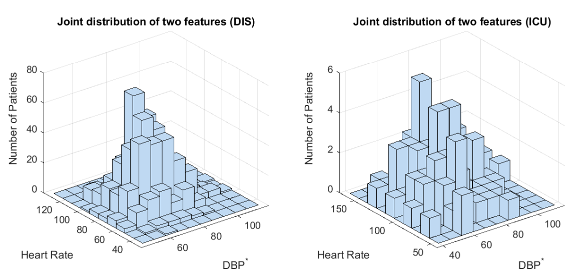

Figure 4. Joint distribution of diastolic blood pressure and heart rates

5. Extension of ForecastICU: Patient Risks Tracking Systems (PRTS)

ForecastICU can be extended to patients risks tracking systems (PRTS) which keeps tracking the ICU belief (risks of ICU admission) until the actual ICU admission or discharge event. This system is useful in real clinical setting because PRTS helps doctors to focus on the real-time high risk patients based on the ICU belief provided by the algorithm. In this subsection, we illustrate the performance of ForecastICU in PRTS setting.

ForcastICU has a consistently higher PPV in comparison to other benchmarks which is represented in Table 3 and Fig. 5. For instance, given 70% TPR, ForecastICU achieves 80.1% PPV which is 5.2% better than the second best algorithm (Lasso Regularization). Moreover, with 70% PPV, Forecast ICU achieves 78.0% TPR which is 4.7% better than the second best algorithm. AUC values are also 1.5% higher than the second best algorithm and the p-value of the hypothesis test comparing ForecastICU and the second best algorithm is ≤ 0.01.

Table 3. Performance comparison of ICU prediction in PRTS setting

Algorithms	TPR(%)	PPV(%)
ForecastICU	70.3± 1.75%	80.1± 1.23%
Logistic Regression	70.5± 1.13%	73.5± 2.09%
Lasso Regularization	70.1± 1.49%	74.9± 1.98%
Random Forest	70.7± 1.34%	56.1± 1.24%
SVMs	70.0± 1.28%	44.9± 1.74%

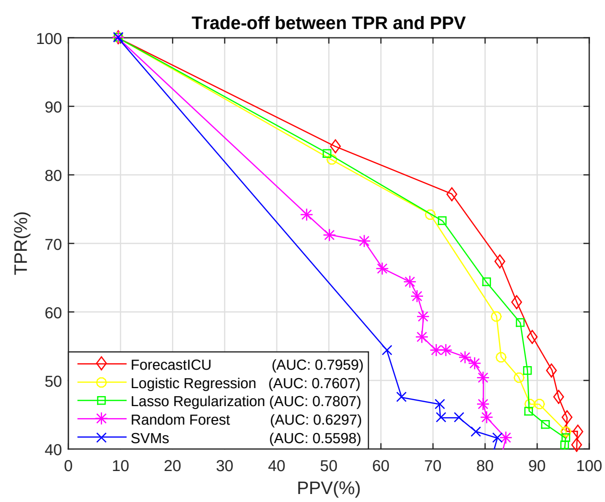

Figure 5. Trade-off between TPR and PPV in PRTS setting

6. Additional Experiment Results

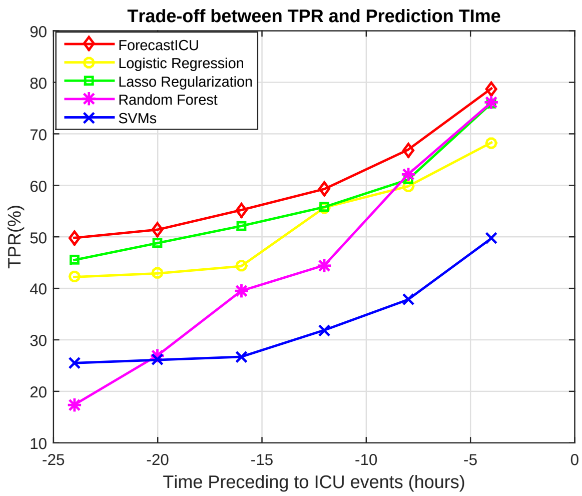

Figure 6. Trade-off between TPR and the prediction time (fix PPV 30%)

References

Alexander Olaussen, Todd Blackburn, Biswadev Mitra Mark Fitzgerald. Review article: shock index for prediction of critical bleeding post-trauma: a systematic review. *Emergency medicine Australasia*, 26(3):223–228, 2014.

Andrew Egol, Robert Fromm, Kalpalatha K Guntupalli Melissa Fitzpatrick David Kaufman Stanley Nasraway David Ryon Jack Zimmerman. Guidelines for intensive care unit admission, discharge, and triage. *Critical Care Medicine*, 27(3):633–638, 1999.

Bruijns, Stevan R, Guly Henry R Bouamra Omar Lecky Fiona Lee Wallis A. The value of traditional vital signs, shock index, and age-based markers in predicting trauma mortality. *Journal of Trauma and Acute Care Surgery*, 74(6):1432–1437, 2013.