DYNACARE-OP: Dynamic Cardiac Arrest Risk Estimation Incorporating Ordinal Features

Joyce C. Ho[1], Yubin Park[1], Carlos M. Carvalho[2], Joydeep Ghosh[1]

University of Texas - Austin
[1] Department of Electrical and Computer Engineering
[2] McCombs School of Business

ICML WHEALTH '13

Cardiac Arrest Overview

Cardiac arrest occurs when the heart stops beating and blood circulation ceases

Motivation

- Survival to discharge rate $\sim 20\%$
- $\sim 62\%$ of cardiac arrests can be prevented
- Rapid response teams (RRT) can increase survival rate up to $\sim 40\%$
- Major RRT weakness is inability to identify patients with sufficient intervention time

Model Overview

DYNACARE-OP is a dynamic cardiac arrest risk estimation model that:

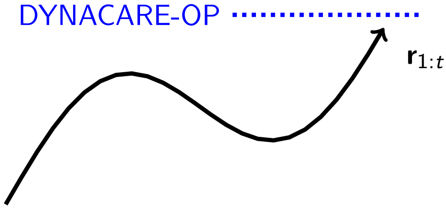
- utilizes continuous and ordinal clinical measurements
- continually tracks a patient's cardiac arrest trajectory
- allows interpretability and predictability of a cardiac arrest event
- provides a dynamic hazard function unobtainable via traditional survival analysis
- generalizes to detecting other medical events

DYNACARE-OP: Components

1. Single latent factor r_t governs clinical observations
2. Two states, healthy and risky state, influence the mean of r_t
3. Correlations in fluctuations of r_t captured by a stochastic volatility term
4. Semi-supervised framework to connect cardiac arrest event to r_t
5. Data augmentation trick to incorporate ordinal features

DYNACARE-OP: Cardiac Arrest Trajectory

Model cardiac arrest trajectory (CAT) as a *single latent factor* r_t



CAT (r_t) governs clinical observations

Define cardiac arrest as a temporal signature in CAT

DYNACARE-OP: Simple Dynamic Linear Model

General Linear Model

$$r_t = r_{t-1} + \epsilon_t \qquad \epsilon_t \sim N(0, \lambda_t)$$
$$\mathbf{y}_t = \boldsymbol{\beta}_t r_t + \boldsymbol{\eta}_t \qquad \boldsymbol{\eta}_t \sim N(0, \boldsymbol{\Sigma})$$

Assume $\mathbf{y}_t$ continuous and knowledge of $\lambda_t, \boldsymbol{\beta}_t, \boldsymbol{\Sigma}$

Kalman filter is closed form solution to estimate latent variables

Limitations:

- Unable to fully capture the data
- How to encode temporal CAT signature?

DYNACARE-OP: Markov Switching Model

Markov switching model (MSM) is a widely adopted financial economics modeling framework to incorporate heterogenous dynamics

Our model assumes two states, healthy and risky, govern the mean of r_t

Cardiac arrest occurs when patient is in the risky state

$$
\begin{aligned}
u_t &\sim \text{MarkovChain}(u|u_{t-1}) & u_t &\in \{s_h, s_c\} \\
r_t &= r_{t-1} + \epsilon_t & \epsilon_t &\sim N(0, \lambda_t) \\
\mathbf{y}_t &= \boldsymbol{\beta}_t r_t + \boldsymbol{\eta}_t & \boldsymbol{\eta}_t &\sim N(0, \boldsymbol{\Sigma})
\end{aligned}
$$

How to model correlated fluctuations around the mean?

DYNACARE-OP: Stochastic Volatility Model

Stochastic volatility models can capture "volatility clustering", where large residuals are clustered together, using an underlying stochastic process

Stochastic volatility term λ_t captures time-to-time correlations as well as patient-to-patient correlations

$$\begin{aligned}
\lambda_t &= \lambda_{t-1} + \delta_t & \delta_t &\sim N(0, k^2) \\
u_t &\sim \text{MarkovChain}(u|u_{t-1}) & u_t &\in \{s_h, s_c\} \\
r_t &= r_{t-1} + \epsilon_t & \epsilon_t &\sim N(0, \exp(\lambda_t)) \\
\mathbf{y}_t &= \boldsymbol{\beta}_t r_t + \boldsymbol{\eta}_t & \boldsymbol{\eta}_t &\sim N(0, \boldsymbol{\Sigma})
\end{aligned}$$

How to estimate r_t with non-linearity term?

DYNACARE-OP: Particle Filters and Smoothers

Particle filtering and smoothing can be used to perform sophisticated model estimation based on simulation

Particle filting characterizes current state given observations up to the current time

Particle smoothing estimates the distribution of current state given all observations up to some *future* time

Excellent tutorial by Arnaud Doucet:
`http://www.cs.ubc.ca/~arnaud/doucet_johansen_tutorialPF.pdf`

How to connect latent factor to cardiac arrest event?

DYNACARE-OP: Semi-Supervised Learning Framework

Use partial knowledge to connect cardiac arrest risk to latent factor

For each patient:

- Known - cardiac arrest event at specific time point
- Unknown - period in which patient was healthy
- Unknown - unrecorded or unobserved cardiac arrest events

Semi-supervised particle smoother to approximate $E[r_{1:T} | \mathbf{y}_{1:T}, T = CA]$

DYNACARE-OP: Semi-supervised particle smoother

Forward filter and backward constrainining framework

- Forward filter to approximate $(r_{1:T} | \mathbf{y}_{1:T})$
- Impose the constraint $T = CA$: $p(u_t = s_c | \mathbf{y}_t) \approx 1$
- Backward smooth to get $(r_{1:T} | \mathbf{y}_{1:T}, T = CA)$

Use fixed-lag approach to prevent particle degeneracy

How to utilize ordinal features?

DYNACARE-OP: Data Augmentation Trick

Ordered profit trick introduced by Albert & Chib 1993

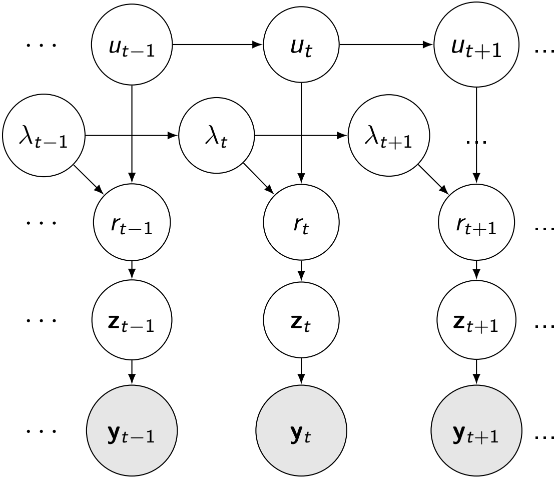
Introduce new latent continuous variable $\mathbf{z}_t$

For ordinal feature y_ρ with J ordered categories:

$$
\begin{aligned}
&y_\rho = j, \quad \gamma_{j-1} < z_\rho \leq \gamma_j \\
&z_\rho \sim N(\beta_\rho r, \sigma_\rho) \\
&\gamma_0 = -\infty \\
&\gamma_J = \infty
\end{aligned}
$$

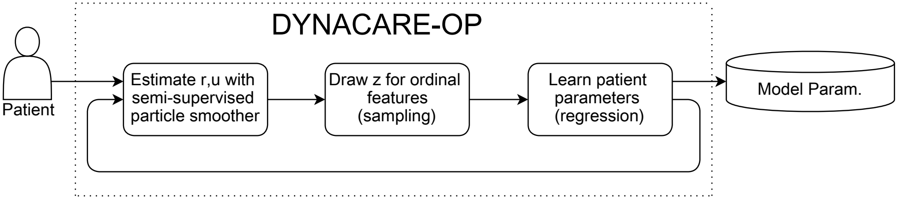
Global bin boundaries γ^o defined using empirical distribution of patient measurements

DYNACARE-OP: Graphical Model



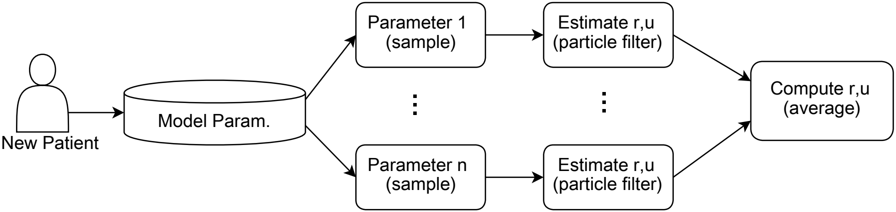
DYNACARE-OP: CAT Estimation Algorithm

Combination of expectation maximization and particle smoothing to estimate parameters and latent factors



1. Given a patient's parameters, $\boldsymbol{\beta}$, estimate latent variables r, u with semi-supervised particle smoother
2. Estimate z for ordinal features by drawing from truncated normal centered around $\boldsymbol{\beta}r$ with appropriate bounds
3. Learn $\boldsymbol{\beta}, \boldsymbol{\Sigma}$ given r, z
4. Repeat until convergence occurs

DYNACARE-OP: Estimating a New Patient



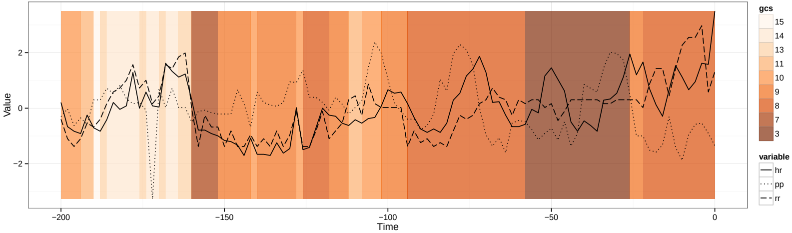
1. Draw new patient's $\boldsymbol{\beta}, \boldsymbol{\Sigma}$, from learned parameter stratum
2. Estimate r, u for each parameter pair using unsupervised particle filter
3. Take average of r, u for new patient's cardiac arrest risk

Results: Experimental Setup

298 cardiac arrest patients from MIMIC-II database

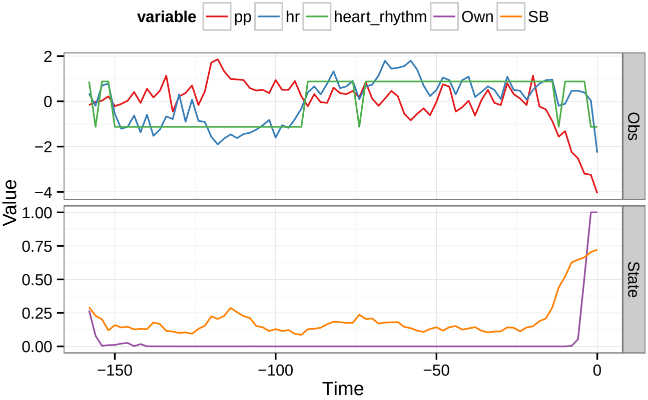
8 clinical measurements: pulse pressure, pulse pressure index, respiratory rate, heart rate, temperature, Glasgow Coma Score, respiratory pattern, heart rhythm

Data prior to cardiac arrest is discretized into 2-hour bins and normalized



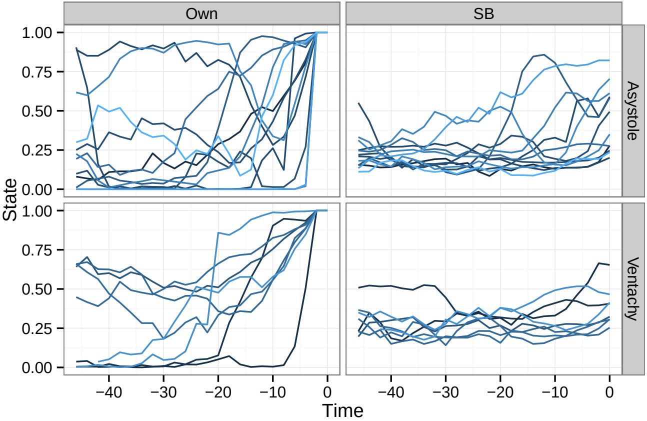
Results: Single Patient Estimation

Continuously tracking a patient's CAT based on the clinical measurements



Results: Estimating CAT

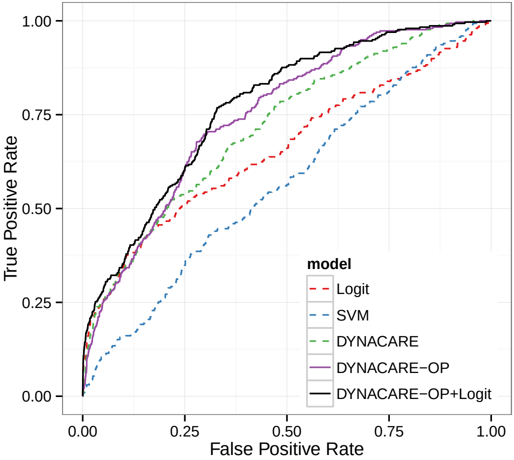
Comparing CAT using individually learned (Own) or stratified bootstrapping (SB) β parameters for two types of CA events



Results: Predictive Performance

Evaluation of classification task on 20 time periods before CA

- 1 for last time period, 0 for others



Conclusions

Our contributions:

- Dynamic cardiac arrest risk estimation model combining financial econometric and semi-supervised learning frameworks
- Incorporation of ordinal features for improved cardiac risk tracking
- Predictive power of temporal signature over static features
- DYNACARE-OP framework can be used for other diagnosis

Survival Analysis

Traditional hazard function:

$$h(t)dt = p(t < t_{\text{event}} < t + dt | t_{\text{event}} \geq t)$$

Inability to align time across patients is problematic for classical hazard models

DYNACARE-OP survival analysis:

- New patient without observations $p(\text{CA at time } t) = \pi_c$
- Sequence of measurements $h(t) = p(\hat{u}_t = s_c | \mathbf{y}_{1:t})$

DYNACARE-OP: Kalman Filter

Notation: $D_t = \{D_{t-1}, \mathbf{y}_t\}$ is the information set up to time t

Forward recursions to estimate current state

Posterior at $t-1$: $$(r_{t-1}|D_{t-1}) \sim N(m_{t-1}, C_{t-1})$$

Prior at t :
$$\begin{aligned}(r_t|D_{t-1}) &\sim N(m_{t-1}, C_{t-1} + \lambda_t) \\ &\sim N(a_t, W_t) \\ (\mathbf{y}_t|D_{t-1}) &\sim N(\boldsymbol{\beta}_t m_t, \boldsymbol{\beta}_t W_t \boldsymbol{\beta}_t^\mathsf{T} + \boldsymbol{\Sigma}) \\ &\sim N(\mathbf{f}_t, \mathbf{Q}_t)\end{aligned}$$

Posterior at t :
$$\begin{aligned}(r_t|D_t) &\sim N(a_t + \mathbf{G}_t\mathbf{e}_t, W_t - \mathbf{G}_t\boldsymbol{\beta}_t W_t) \\ &\text{where } \mathbf{G}_t = W_t\boldsymbol{\beta}_t^\mathsf{T}\mathbf{Q}_t^{-1}, \mathbf{e}_t = (\mathbf{y}_t - \mathbf{f}_t) \\ (r_t|D_t) &\sim N(m_t, C_t)\end{aligned}$$

DYNACARE-OP: Kalman Filter

Backward smoothing to update state given future observations

$$\text{Let:} \quad (r_{t+1}|D_T) \sim N(h_{t+1}, V_{t+1})$$

$$\begin{aligned}\text{Then:} \quad (r_t|D_T) &\sim N(m_t + B_t(h_{t+1} - a_{t+1}), P_t) \\ \text{where } B_t &= C_t W_{t+1}^{-1} \\ P_t &= C_t - B_t(W_{t+1} - P_{t+1}) W_{t+1}^{-1} C_t\end{aligned}$$