DEEP GENERATIVE MODEL WITH BETA BERNOULLI PROESS FOR MODELING AND LEARNING CONFOUNDING FACTORS

R·I·T B. THOMAS GOLISANO *College of* COMPUTING AND INFORMATION SCIENCES

Prashnna K. Gyawali[1*], Cameron Knight[1], Sandesh Ghimire[1], John L. Sapp[2], Milan Horacek[2] , Linwei Wang[1]

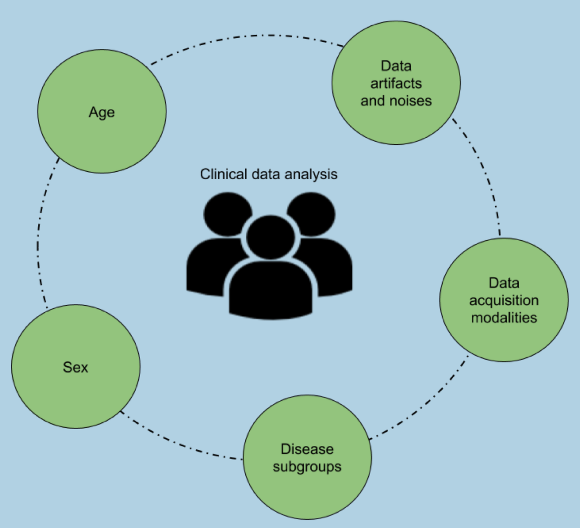
[1]Rochester Institute of Technology, Rochester, USA; [2]Dalhousie University, Halifax, Canada

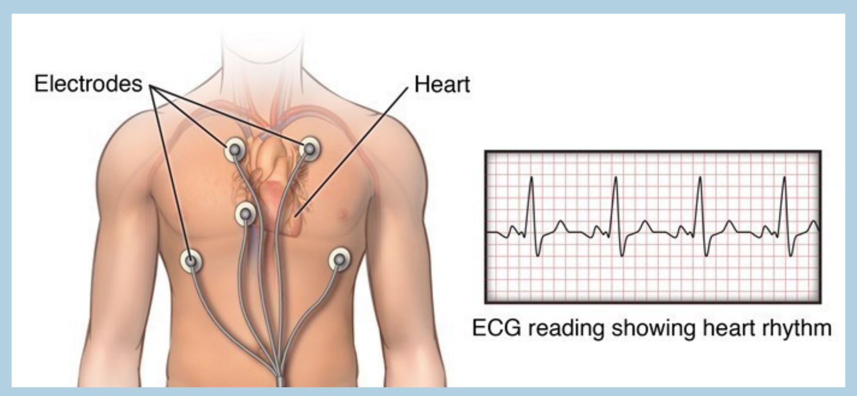
*www.pkgyawali.com pkgyawali@mail.rit.edu @pkgyawali

Computational Biomedicine Lab

Introduction

- Two important challenges exist to address the confounding factors (an inherent property in most data analyses):
  - Unknown and sometimes infinite number of confounding factors.
  - Some of the confounding factors are not easily observable.





- *Disentangled representation* learning aims to separate all the factors of variations within the data.
  - Challenges: Modeling and inference of an infinite number of confounding factors

- Conditional IBP-VAE (cIBP-VAE)
  A deep conditional generative model to disentangle a task-relevant representation from an unknown number of confounding factors that may grow infinite.

Beta-Bernoulli Process

- A stochastic process defining a probability distribution over sparse binary matrices Z indicating feature activation for K features.
- Taking the infinite limit as $K \rightarrow \infty$, we get, Indian Buffet Process (IBP).

$$v_k \sim Beta(\alpha, \beta); z_{n,k} \sim Bernoulli(\pi_k); \pi_k = \prod_{i=1}^{k} v_k$$

where, $z_{n,k}$ is the element of $Z \in \{0,1\}^{\{NXK^+\}}$, α and β are the shape parameters for Beta process.

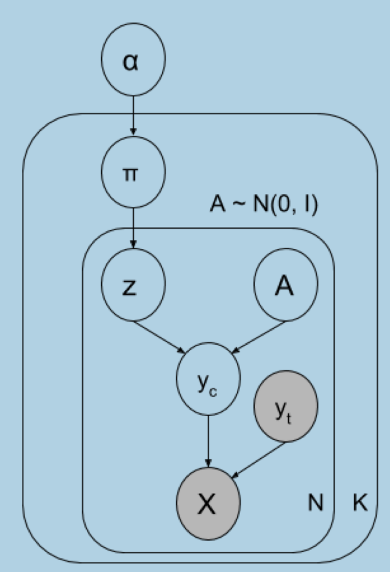
Conditional Generative Model

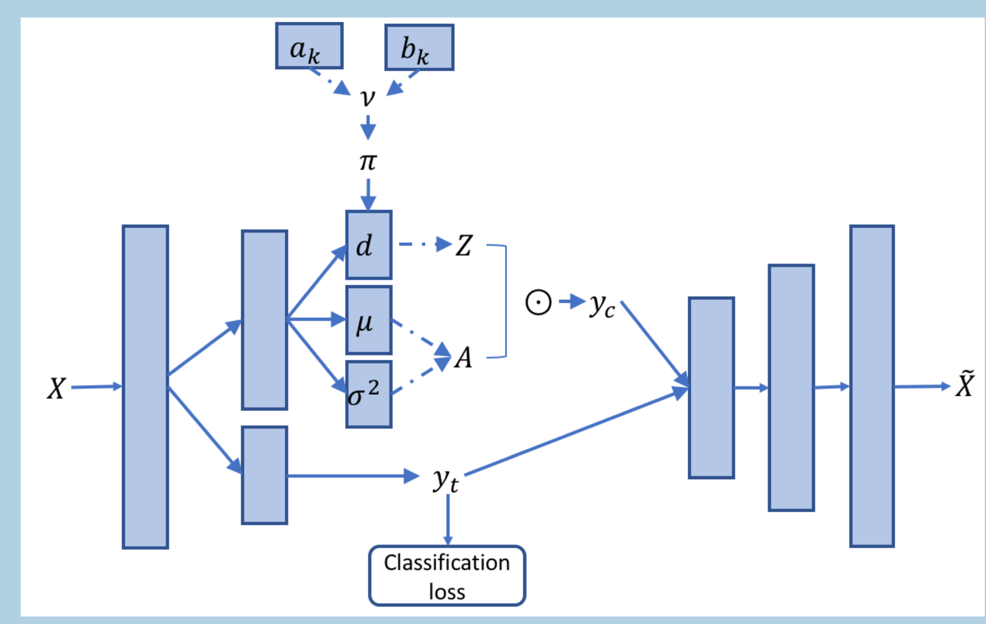
- Conditional probabilistic model admits two sources of variations:
  - the task-related representations, y_t
  - the confounding representation, y_c

- To model an unbounded number of unobserved confounders, we model y_c with an IBP prior :

$$Z, \nu \sim IBP(\alpha); A_n \sim N(O, I); y_c = Z \odot A;$$
$$X \sim p_\theta(X|Z \odot A, y_t)$$

The likelihood function $p_\theta(X|Z \odot A, y_t)$ is defined by a neural networks parameterized by θ.





(Left) Conditional generative model. **(Right)** Overall network architecture.

Inference: Structured stochastic variational inference (SSVI) to infer the latent variables Z, A and ν.

- The objective is to maximize the following lower bound $\mathcal{L}$:

$$KL(q(v_k)||p(v_k)) + \sum_{n=1}^{N} \begin{array}{c}(\mathbb{E}_q[\log p(x_n|Z_n, A_n, y_{tn})] - \\ KL(q(Z_n|\nu, x_n)||p(Z_n|\nu)) - \\ KL(q(A_n|x_n)||p(A_n)))\end{array}$$

- The task-representation y_t is encoded from deterministic encoder $q(y_t|\, X)$ supervised using a classification loss. Hence, final objective is:

$$\mathcal{L}^\gamma = \mathcal{L} + \zeta \cdot \mathbb{E}_{p(X,y_t)}[-\log q_{\phi_2}(y_t|X)]$$

Experiments

Colored MNIST: b/w MNIST dataset augmented with RGB color

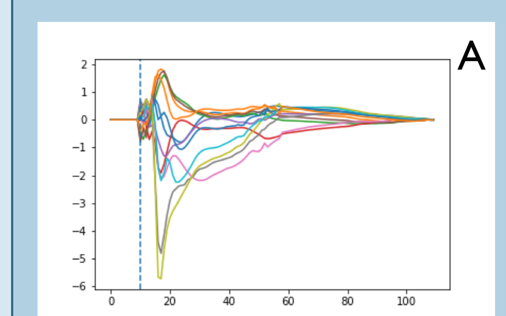
(Left) represents swapping between y_t and y_c. **(Right)** *triggering* units within binary features Z responsible for recognizing presence or absence of color.

Clinical ECG dataset: A large pace-mapping ECG dataset collected from 39 scar-related ventricular tachycardia (VT) patients *(IRB approved)*

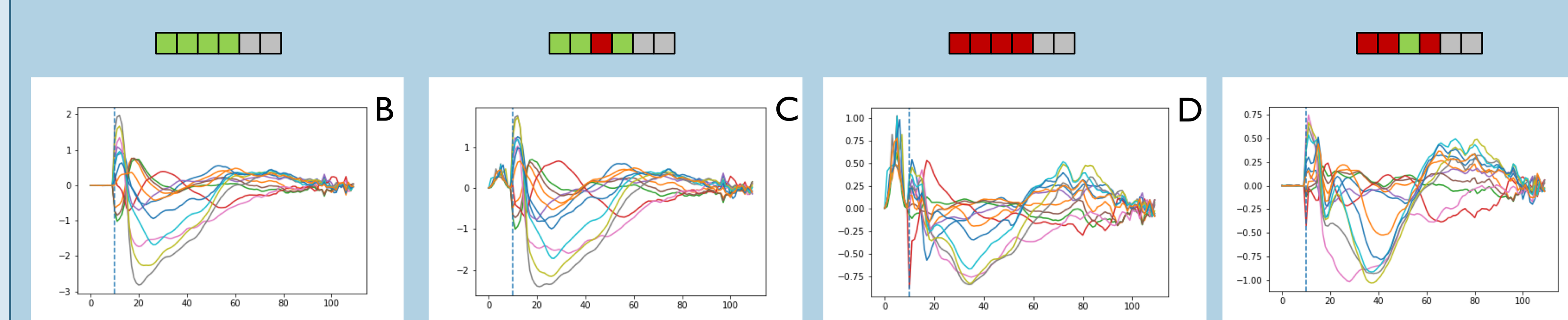
Model	Segment classification
QRS Int	47.61 %
CNN	53.89 %
c-VAE	55.97 %
cIBP-VAE	57.53 %

(Left) Segment classification accuracy for the clinical task of VT localization. **(Right)** Reconstruction errors.

Model	all signal	only artifacts		
		all	no stimulus	stimulus
c-VAE	2293.23	3.20	3.91	2.49
cIBP-VAE	2273.65	0.45	0.19	0.72



(A) Original signal, **(B)** reconstructed signal, **(C) (D) (E)** generated signals by manipulating *triggering* units within binary feature Z.



Conclusion

- A deep conditional generative model for disentangling and learning the unobserved and unbounded number of confounding factors.

- **Acknowledgment:** This work is supported by the National Institutes of Health [grant no: NIH HL140500]