Variationally Regularized Graph-based Representation Learning for Electronic Health Records

BDH Reproducibility Challenge

ALEKSANDR ROGACHEV

CAMILO VELÁSQUEZ A

Introduction

- **General Task:** Learn effective graph representation of medical concepts by training Graph-based Neural Network on predictive tasks

- **Paper's contribution:**
  - Utilized ***Graph Attention Networks*** to learning graph representation of medical concepts in EHR
  - Solve the problem of uniformly distributed attention weights by introducing ***variational regularization*** between encoder and decoder
  - Analyze and interpret the effect of variational regularization on the graph representation

- The intent of this work is to attempt to reproduce and confirm the above points

Datasets

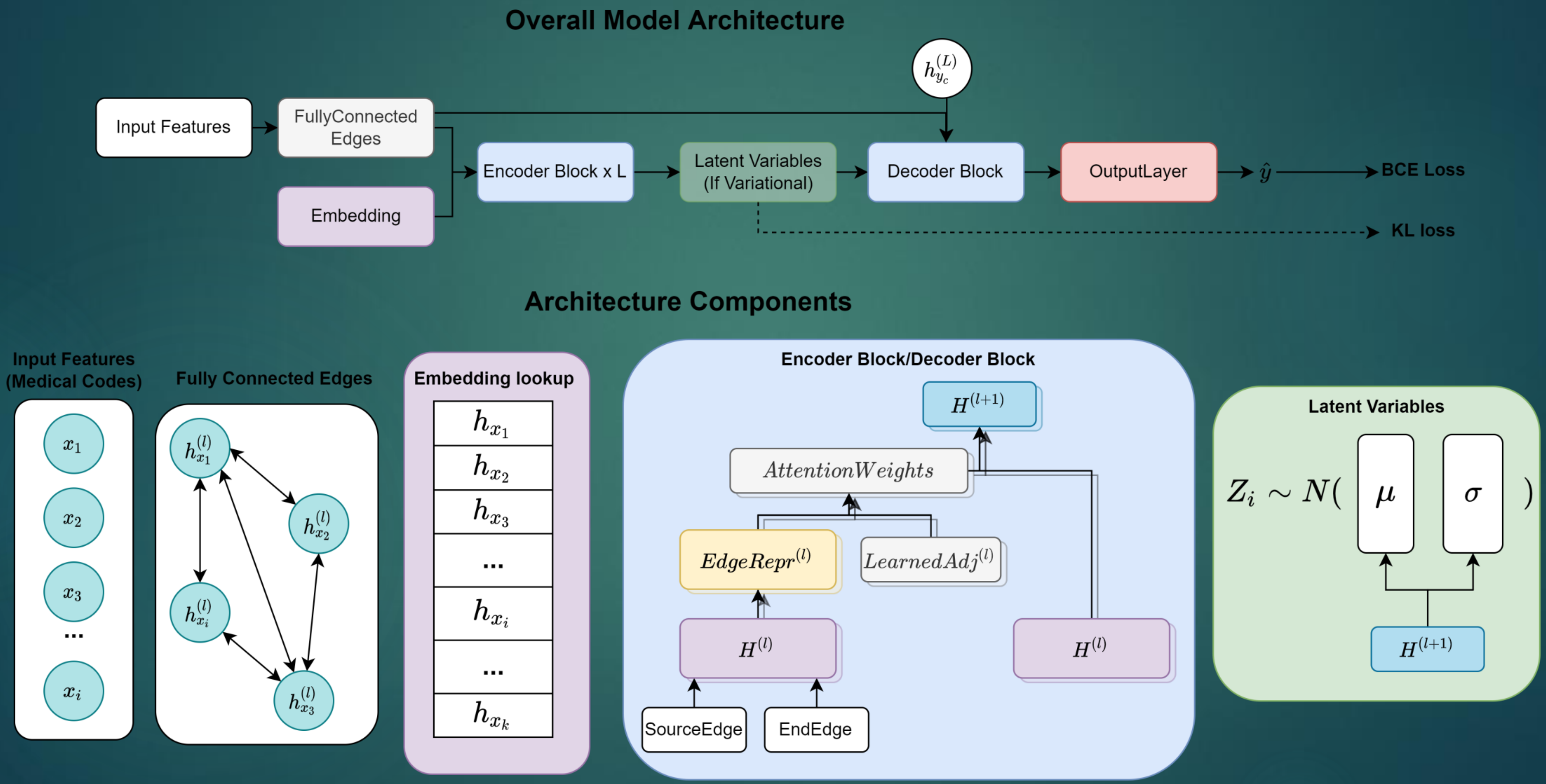
- MIMIC-III. Mortality prediction
- eICU. Readmission prediction
- Alzheimer's Disease EHR from NYU Langone Health (AD-EHR). Not used in this work since it is not publicly available

Total Samples (N° Positive)	MIMIC	eICU
Train	44449 (8440)	38441 (11242)
Valid	5042 (543)	4103 (676)
Test	5043 (552)	4103 (754)

Data was split between train/val/test sets at 80/10/10 ratio

Note: Oversampling was used, duplicating positive samples for the training set

Encoder-Decoder/VGNN Model



Encoder-Decoder/VGNN Model

Problem with Encoder-Decoder: after training, self-attention weights tend to be uniformly distributed, which reduces model expressiveness.

Like Variational Auto-Encoders, VGNN treats encoder output as a standard distribution and forces it to be close to normal. It has a regularization effect.

Loss function:

$$L = L_{bce}(y, \hat{y}) + \beta \cdot KL[q(Z|X)||p(Z)]$$

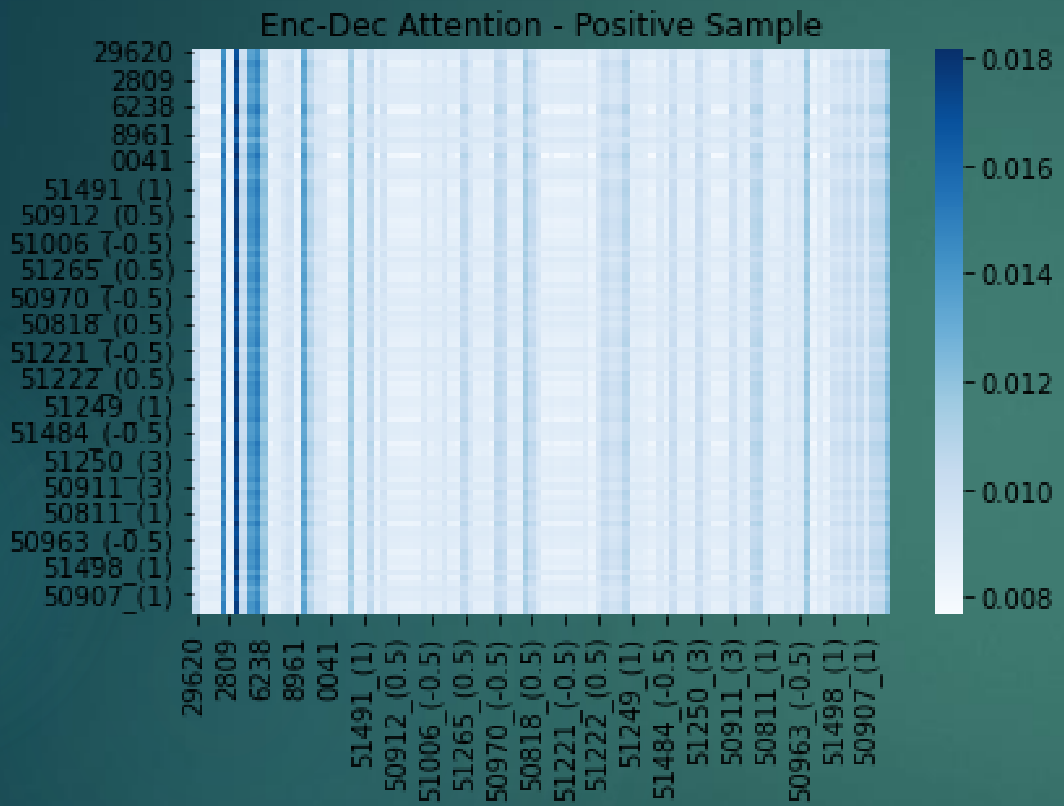
Where KL is Kullback–Leibler divergence between the learned encoded distribution and normal gaussian and β is its weight. The original paper does not use this weight, but we found that it was necessary because KL term dominates BCE. The weight was tuned to get optimal performance for both eICU and MIMIC.

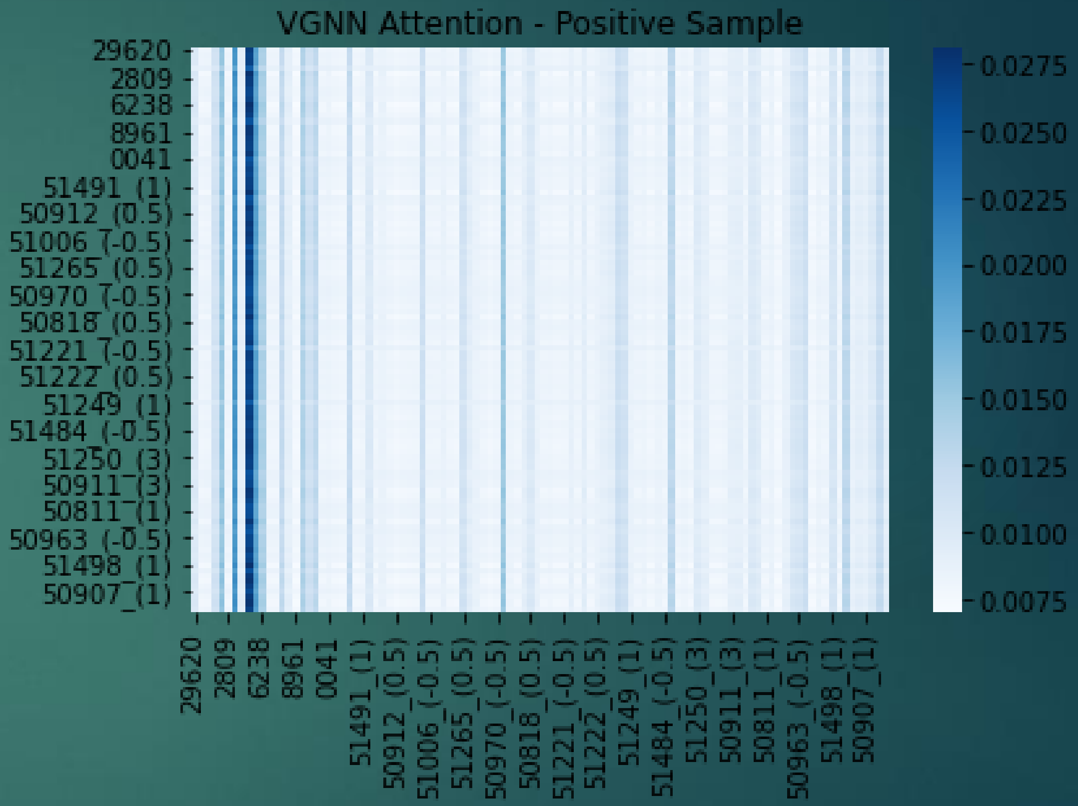
Performance Comparison

Dataset	Method	AUPRC val (Ours)	AUPRC test (Ours)	AUPRC test (Original)
MIMIC	Encoder-Decoder, embedding 256	0.6834	0.6975	0.6962
	VGNN, embedding 256: KL weight 0.1 **KL weight 0.02** KL weight 0.005	0.6842 **0.6960** 0.6843	0.6978 **0.7029** 0.6861	0.7102
eICU	Encoder-Decoder, embedding 128	0.3730	0.4240	0.3881
	VGNN, embedding 128: KL weight 0.02 KL weight 0.005 KL weight 0.0001	0.3112 0.3284 0.3377	0.3629 0.3870 0.3911	0.3986

VGNN Increased AUPRC score on MIMIC, but performed worse on eICU (unlike paper)

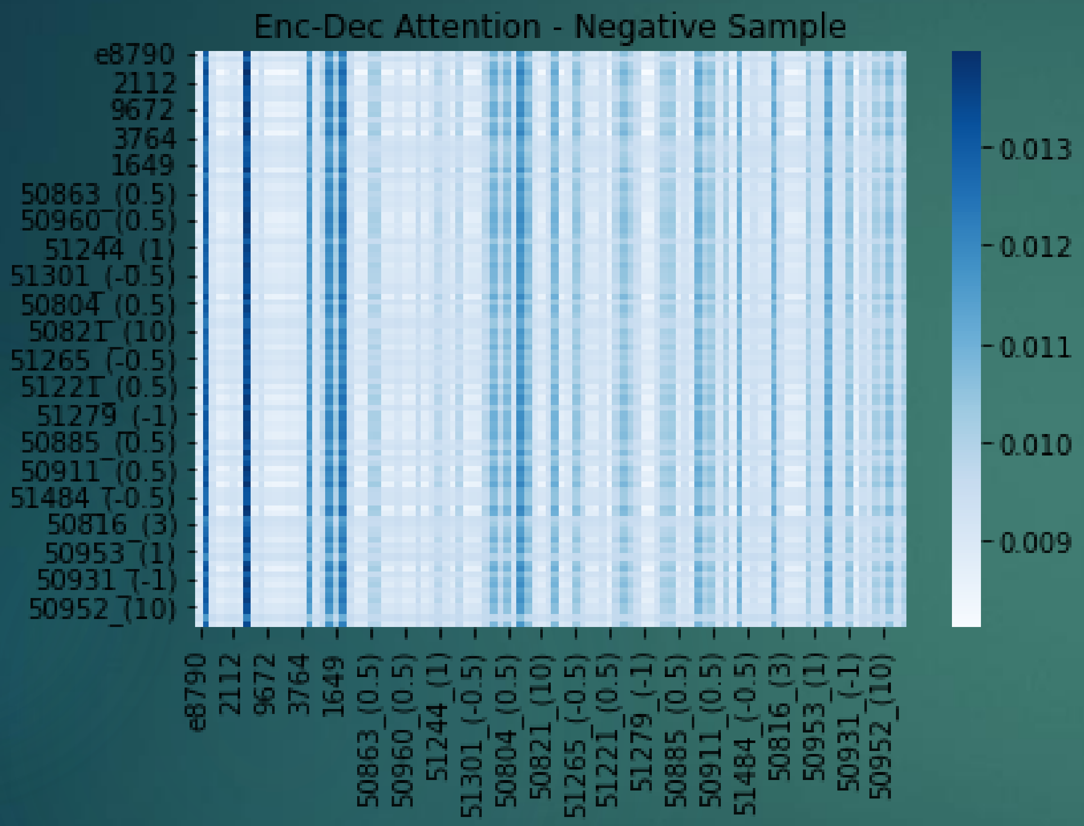
Analysis: Attention Visualization

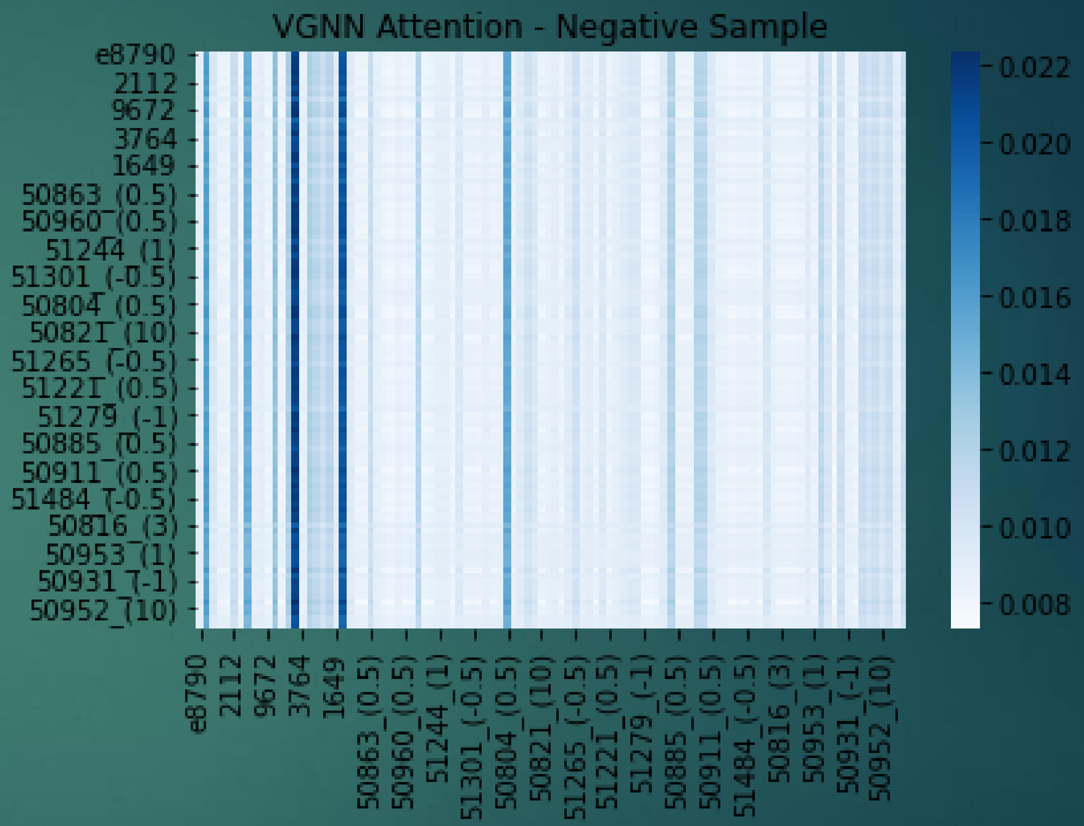




Encoder-Decoder's attentions are more uniform, and it tends to select too many features.
VGNN focuses on specific features, and it leads to more effective representation.

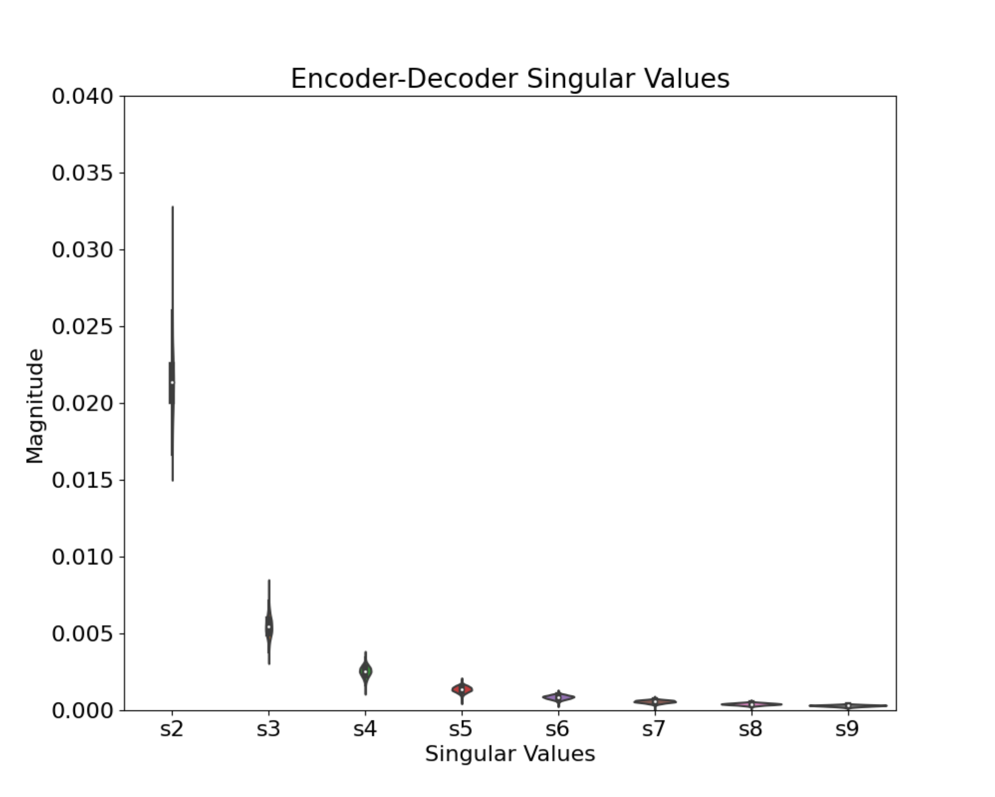
Analysis: Attention Visualization

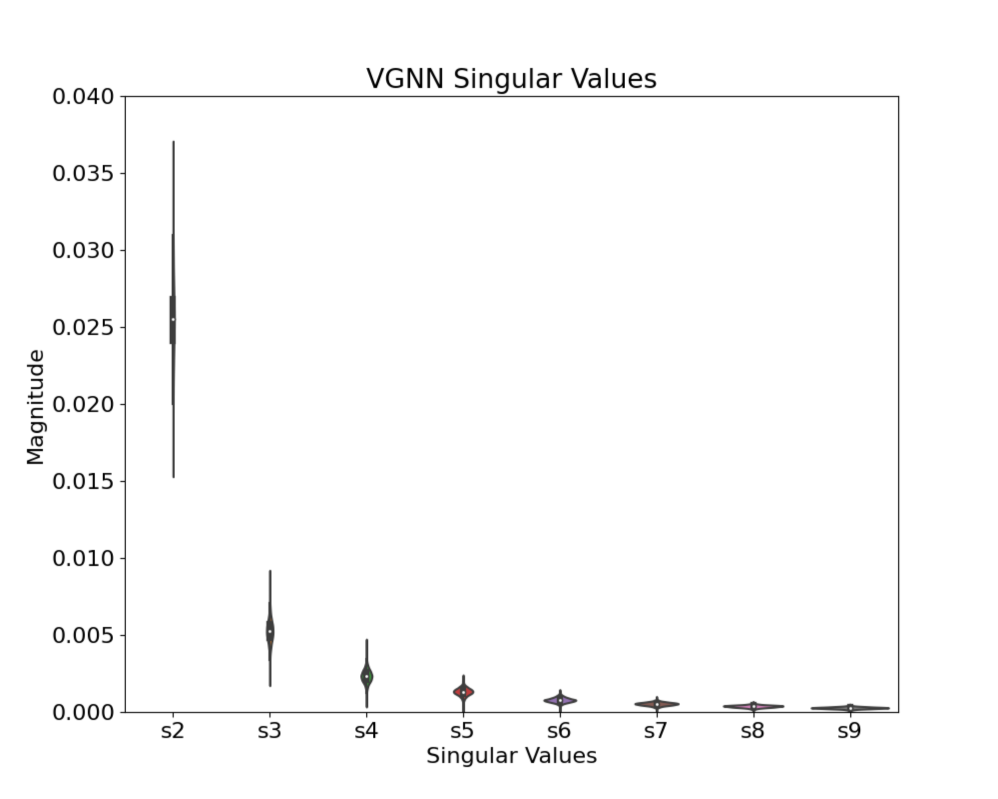




Encoder-Decoder's attentions are more uniform, and it tends to select too many features.
VGNN focuses on specific features, and it leads to more effective representation.

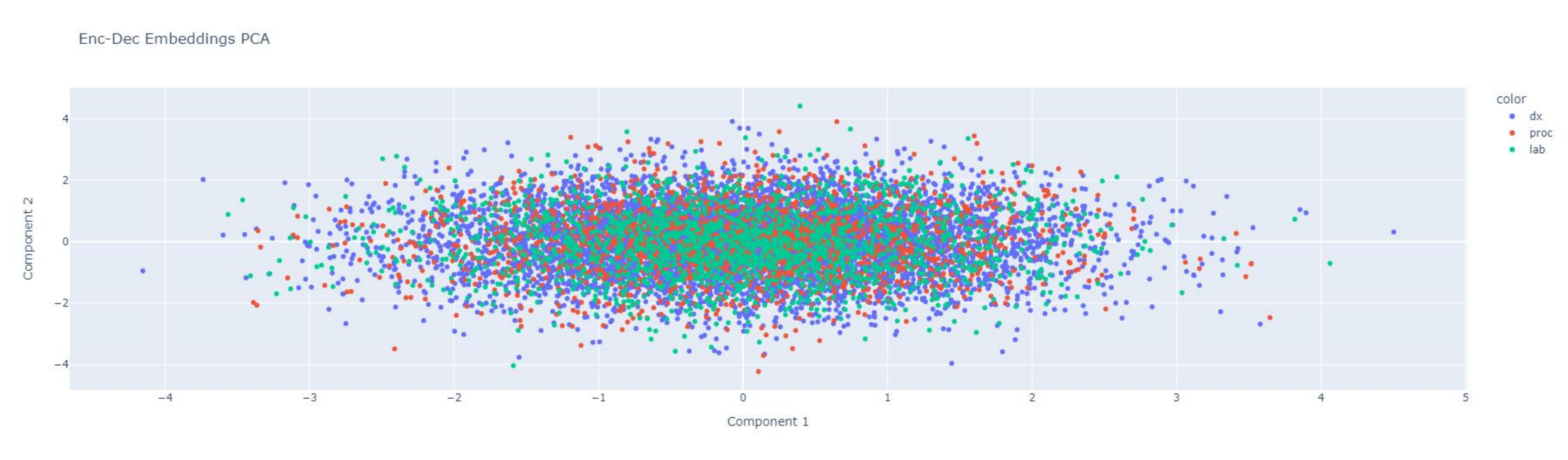
Analysis: SVD of Adjacency Matrix

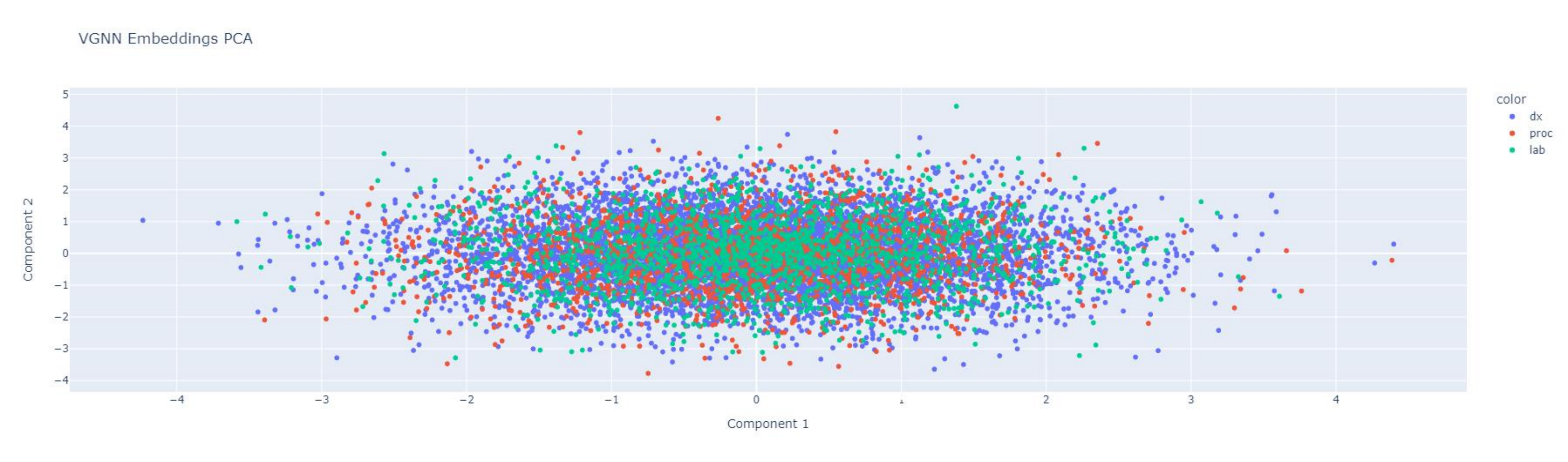




Singular values of attention matrix show effectiveness of message passing among nodes. VGNN has slightly higher singular values with greater range. However, unlike paper, the actual values are on the same scale.

Analysis: Embeddings





Both Enc-Dec and VGNN have noisy 2D PCA representation of embeddings.

No explicit pattern (unlike paper)

Discussion

Successfully Reproduced:

- The proposed variational regularization improved performance on MIMIC dataset as compared to Encoder-Decoder
- Confirmed that VGNN can make attention weights **less uniform**, and it leads to richer graph representation

Unable to Reproduce

- VGNN performed worse than Encoder-Decoder on eICU dataset
- Singular values of the learned adjacency matrix are similar in Encoder-Decoder and VGNN
- No clear pattern in learned embeddings
- Found issues in the provided code. Main issue - the code applies variational regularization only to two out of thousands of nodes

References

- Zhu, Weicheng, and Narges Razavian. "Variationally regularized graph-based representation learning for electronic health records." *Proceedings of the Conference on Health, Inference, and Learning*. 2021.
- Choi, Edward, et al. "Learning the graphical structure of electronic health records with graph convolutional transformer." *Proceedings of the AAAI conference on artificial intelligence*. Vol. 34. No. 01. 2020.

- Github: https://github.com/NYUMedML/GNN_for_EHR
- Our Github: https://github.gatech.edu/arogachev3/vgnn-reproduce