Representation Learning using Graph Neural Nets: A case-study in HMMs

Rajan Kumar Soni, Karthick Seshadri~, Balaraman Ravindran**

** RBCDSAI, IIT Madras, ~ NIT Andhra Pradesh*

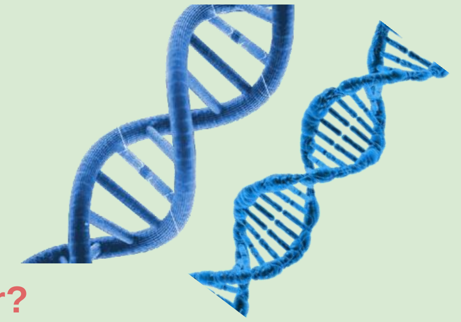
ECML PKDD 2021
VIRTUAL
13-17 September

Use-Cases

To compare sequence models through representation learning.

Are mannerisms similar?

Which surgeon is better?



Are these DNA strands similar?

To enable the above use-cases, we propose two Deep Neural Network based representation learning techniques to learn embeddings for Hidden Markov Sequence Models.

How sequence models are compared typically?

- Metric based on Monte Carlo approximation which is used for measuring entropy divergence.
- Metric based on Kullback–Leibler (KL) divergence.
- Metric based on graph matching.
- Co-emission probability based metric.
- Matrix factorization-based metric.
- Cross-Likelihood metric.

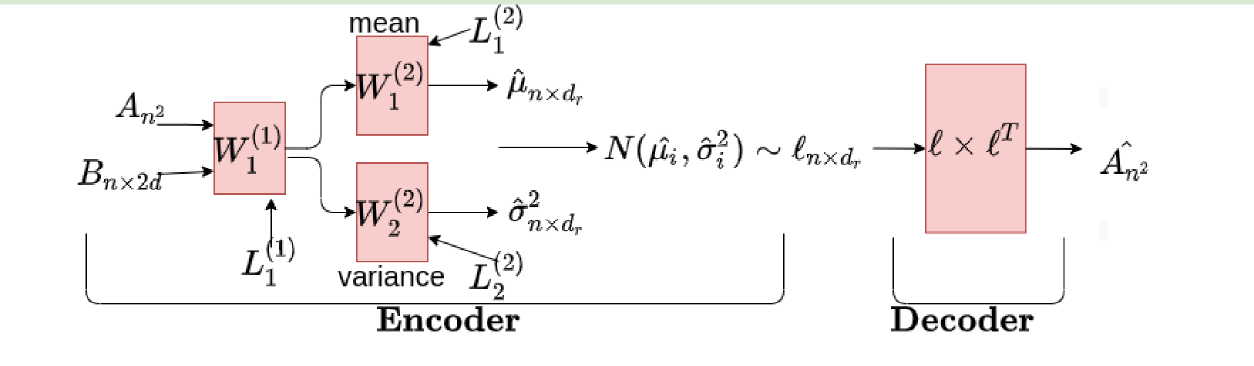
Motivation behind our Work

- Typical metrics perform well only when similar structure implies similar behavior.
- Intractability of graph-matching based metrics.
- Accuracy of the metrics impacted by noise in observations.
- Behavior-based metrics are influenced by the reference sequence.
- Importance of metric-learning in descriptive, predictive and prescriptive analytic tasks.
- Graph Neural Networks have been effective in encoding graphs.
- No exploration has been done for assessing GNNs for representation learning.

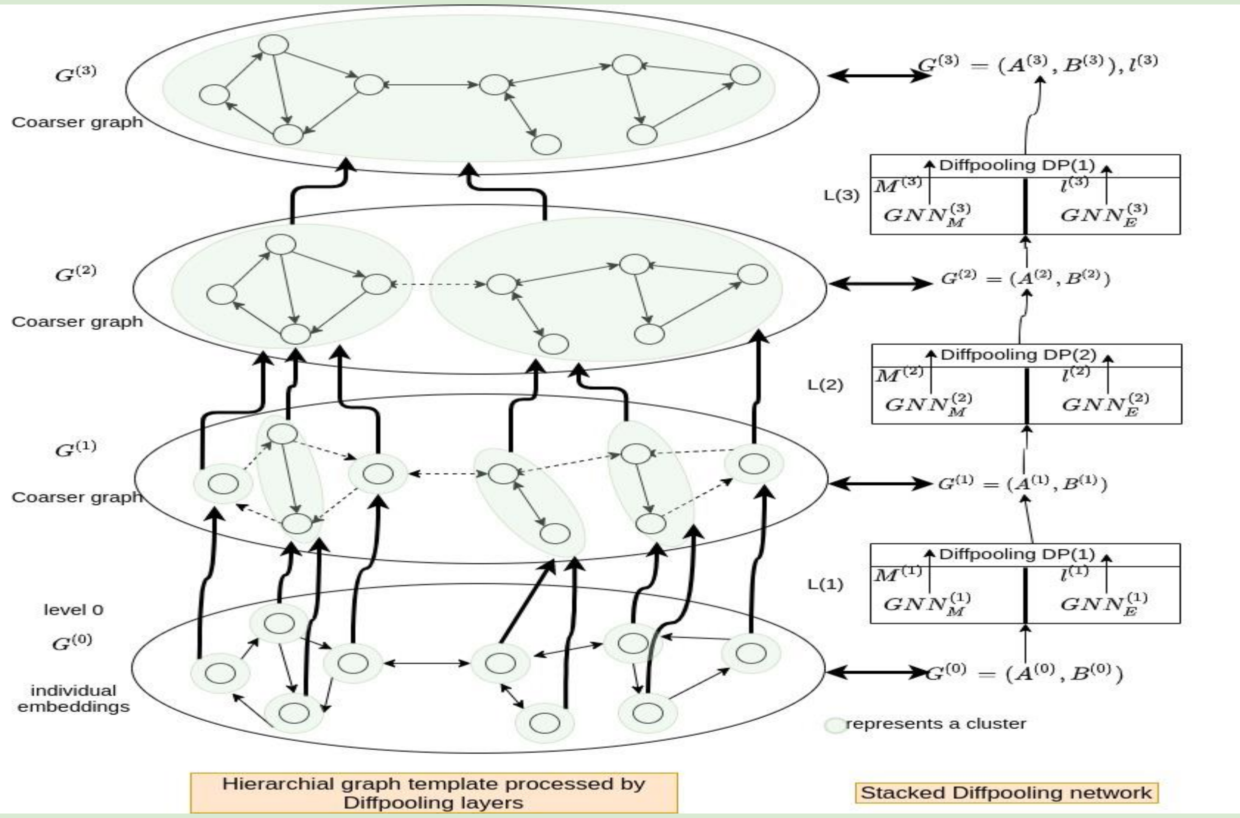
What did we do?

- Proposed a task-agnostic representation learning technique using a Graph Variational Auto Encoder (GVAE).
- Proposed a label-aware representation learning approach using a Diffpooling-based GNN.
- Evaluated the validity and representativeness of the .embeddings learnt using single-linkage & complete-linkage clustering and classification tasks.

GVAE based metric: Preserves the generative ability of a HMM by inferring a low-dimensional latent-space using the architecture below:



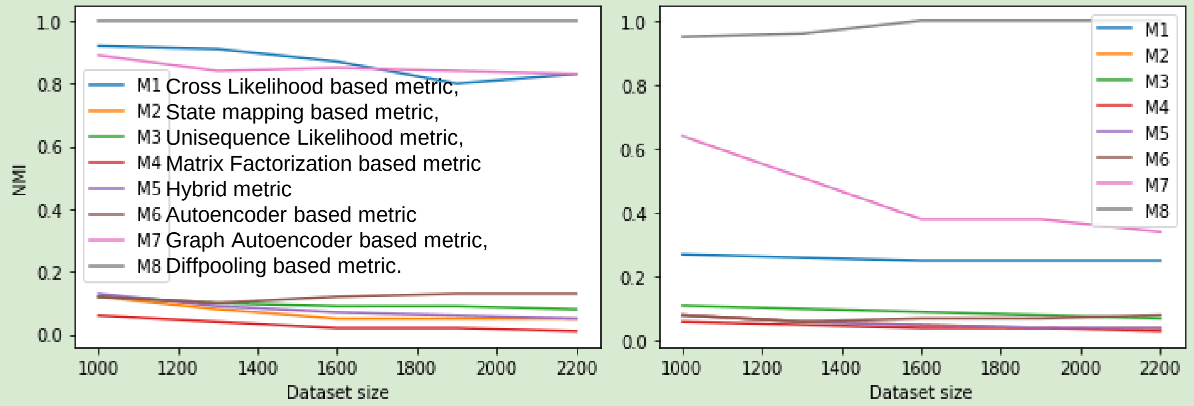
Diffpooling based metric: Learnt by using a stack of layers in which each layer comprises a pair of GNNs. One GNN learns the embedding and the other GNN learns the spectral sub-structures in the HMM.



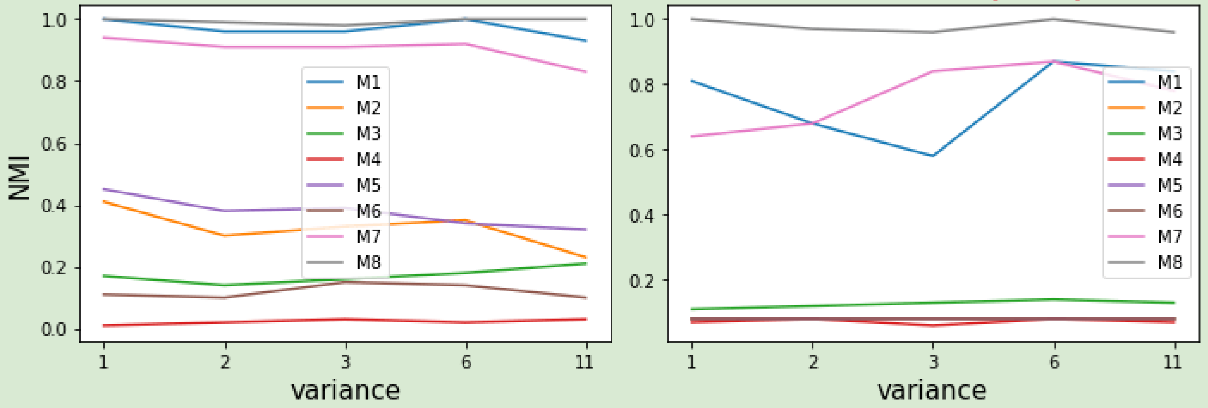
Were we successful?

We evaluated the embeddings learnt by pitching them against six baselines using the FSDD dataset.

Training Set Size vs. Cluster quality

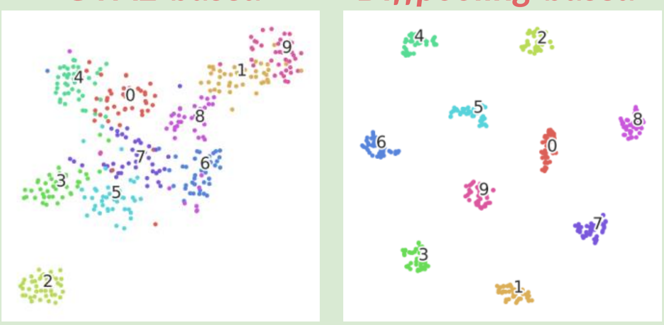


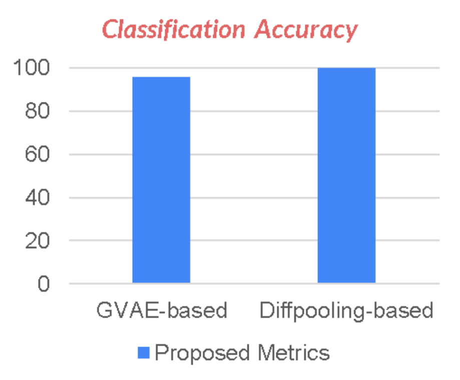
Variance in the number of HMM states vs. Cluster quality



T-sne plot of embeddings.

GVAE-based *Diffpooling-based*





Conclusion and Pointers for Further Research

- GVAE learns regularized and behavior preserving latent embeddings.
- Diffpooling-based embeddings out-perform structure agnostic flat embeddings.
- Diffooling aces the single-linkage clustering task unlike other metrics.
- *Up Next*: Extending our work to other temporal-behavioral models and non-ergodic models.