Graph Neural Networks with Optimal Transport

Master/Ecole d'Ingénieur internship

Expected starting date: May / Juin 2022 (6 months)

Key-words. Deep learning; Graph neural networks; Optimal Transport; Wasserstein distance; Gromov-Wasserstein distance

Context. Over the past decade, Deep Learning, and more specifically Convolutional Neural Networks (CNNs) and Recurrent Neural Networks (RNNs), had a strong impact in various applications of machine learning, such as image recognition [10] and speech analysis [8]. These successes have mostly been achieved on sequences or im- ages, i.e. on data defined on grid structures that benefit from linear algebra operations in Eu- clidean spaces. However, there are many domains where data (e.g. social networks, molecules, knowledge graph) cannot be trivially encoded into a Euclidean domain but can be naturally represented as graphs. This explains the recent challenge tackled by the machine learning community which consists of transposing the deep learning paradigm into the world of graphs. The objective is to revisit Neural Networks to operate on graph data, in order to benefit from the representation learning ability. In this context, many Graph Neural Networks (GNNs) have been recently proposed in the literature of geometric learning [12, 9]. GNNs are Neural Networks that rely on the computation of hidden representations of nodes using information carried by the whole graph. In contrast to conventional Neural Network, where the architecture of the net- work is related to the known and invariant topology of the data (e.g. a 2-D grid for images), the node features of GNNs are propagated according to the graph topology called Message Passing (Graph) Neural Networks (MPNN). Some recent works studied the spectral expressive power of MPNN and showed majority of them work as loss-pass filter [4] and how to improve this expressive power limits in [3]. It is a well-known issue that, if there are two different point clouds where each has a different number of points in it, the best approach to measure the distance between two groups is Mover distance which we can calculate it by optimal transport [11, 1, 2, 6]. Instead of summarizing the node representation into some fixed-length vector, calculating optimal transport distance from any graph's learned node features to another learnable anchor graph's node represen- tation and uses that optimal distance as the output of the graph readout layer is proposed in [5]. However, since the anchor graph's node are learnable this approach can overfit easily and needs heavy regularization. Another application of OT in graph-based data learning is to use kernels where pair-wise graph distances were calculated by optimal transport in [7, 13]. However, both mentioned method uses the given raw node feature to calculate the opti- mal transport distance and no advantage to learn the new more meaningful representation of nodes for the given certain task.

Scientific objectives and expected achievements. In our proposal, we would like to use an optimal transport distance to create a kernel based graph level prediction model. In the model, the new representation of nodes will be learned by Graph Neural Network. Later, the new representation's distance will be calculated by optimal transport in order to create a better kernel for the prediction model. The

prediction would be done by the Support vector machine or Gaussian Random Process which are fully differentiable.

Research environment/Location. The research will take place within the LMAC laboratory located at Compiègne, France. The internship will be supervised by Mokhtar Z. Alaya (LMAC - UTC) and Muhammet Balcilar (InterDigital, Rennes).

Candidate profile. Candidates are expected to be graduated in computer science and/or machine learning and/or signal image processing and/or applied mathematics/statistics, and show an excellent academic. The internship subject requires skills Python development tools, specillay TensorFlow or Pytorch.

For more details. Feel free to contact by sending an email to Mokhtar Z. Alaya (`elmokhtar.alaya@utc.fr`) and to Muhammet Balcilar (`muhammetbalcilar@gmail.com`).

References

[1] Alaya, M. Z., M. Bérar, G. Gasso, and A. Rakotomamonjy (2019). Screening Sinkhorn algorithm for regularized optimal transport. In H. Wallach, H. Larochelle, A. Beygelzimer, F. Alche-Buc, E. Fox, and R. Garnett (Eds.), *Advances in Neural Information Processing Systems 32*, pp. 12169–12179. Curran Associates, Inc.

[2] Alaya, M. Z., M. Bérar, G. Gasso, and A. Rakotomamonjy (2022). Theoretical guarantees for bridging metric measure embedding and optimal transport. *Neurocomputing 468*, 416–430.

[3] Balcilar, M., P. Héroux, B. Gaüzère, P. Vasseur, S. Adam, and P. Honeine (2021). Breaking the limits of message passing graph neural networks. In *Proceedings of the 38th International Conference on Machine Learning (ICML)*.

[4] Balcilar, M., G. Renton, P. Héroux, B. Gaüzère, S. Adam, and P. Honeine (2021). Analyzing the expressive power of graph neural networks in a spectral perspective. In *International Conference on Learning Representations*.

[5] Benson, C., B. Gary, G. Octavian-Eugen, B. Regina, and J. Tommi (2021). Optimal transport graph neural networks. *arXiv*.

[6] Chapel, L., M. Z. Alaya, and G. Gasso (2020). Partial optimal tranport with applications on positive-unlabeled learning. In H. Larochelle, M. Ranzato, R. Hadsell, M. F. Balcan, and H. Lin (Eds.), *Advances in Neural Information Processing Systems*, Volume 33, pp. 2903–2913. Curran Associates, Inc.

[7] Chen, L., Z. Gan, Y. Cheng, L. Li, L. Carin, and J. Liu (2020, 13–18 Jul). Graph optimal transport for cross-domain alignment. In H. D. III and A. Singh (Eds.), *Proceedings of the 37th International Conference on Machine Learning*, Volume 119 of *Proceedings of Machine Learning Research*, pp. 1542–1553. PMLR.

[8] Graves, A., A. Mohamed, and G. Hinton (2013, May). Speech recognition with deep recurrent neural networks. In *2013 IEEE International Conference on Acoustics, Speech and Signal Processing*, pp. 6645–6649. IEEE.

[9] Justin, G., S. S. Samuel, F. R. Patrick, V. Oriol, and E. D. George (2017, 06–11 Aug). Neural message passing for quantum chemistry. In D. Precup and Y. W. Teh (Eds.), *Proceedings of the 34th International Conference on Machine Learning*, Volume 70 of *Proceedings of Machine Learning Research*, pp. 1263–1272. PMLR.

[10] Krizhevsky, A., I. Sutskever, and G. E. Hinton (2012). Imagenet classification with deep convolutional neural networks. In *Proceedings of the 25th International Conference on Neural Information Processing Systems - Volume 1*, NIPS'12, Red Hook, NY, USA, pp. 1097–1105. Curran Associates Inc.

[11] Peyré, G. and M. Cuturi (2019). Computational optimal transport. *Foundations and Trends® in Machine Learning 11*(5-6), 355–607.

[12] Scarselli, F., M. Gori, A. C. Tsoi, M. Hagenbuchner, and G. Monfardini (2009). The graph neural network model. *IEEE Transactions on Neural Networks 20*(1), 61–80.

[13] Vayer, T., N. Courty, R. Tavenard, L. Chapel, and R. Flamary (2019). Optimal Transport for structured data with application on graphs. In *Proceedings of the 36th International Conference on Machine Learning*, pp. 6275–6284. PMLR.