Differential Deep Learning on Graphs and its Applications

Chengxi Zang, Fei Wang
Department of Healthcare Policy and Research
Weill Cornell Medicine
425 E 61st St, New York, NY, USA, 10065
chz4001@med.cornell.edu, few2001@med.cornell.edu

Keywords

Differential Deep Learning; Differential Equations; Ordinary Differential Equations; Graph Neural Network; Normalizing Flow; Network Dynamics; Network Science; Graph Mining; Drug Design; Molecular Graph Generation;

Abstract

This tutorial investigates the recent advancements in introducing differential equation theory to the deep learning methods, denoted as differential deep learning, and further broadens the horizon of such methods with an emphasis on graphs. We will show that differential deep learning on graphs are powerful tools for modeling the structures and dynamics of complex systems and generating molecular graphs in drug discovery.

Tutorial information

Modern deep learning methods have achieved significant performance on various tasks by transforming the data in a layer-by-layer manner (LeCun, Bengio, and Hinton 2015). The residual learning (He et al. 2016a), which introduces skip-connection to deep structures, further overcomes vanishing of gradients and empowers the deep learning methods with deeper structures and thus more powerful expression ability. However, in the perspective of researchers in dynamic systems (control, applied physics, systems biology etc.), such a residual-net framework is the approximation of differential equation (DE) systems by Euler method (which is a 1st order numerical method in solving DEs' initial value problems) (Pearlmutter 1995; Weinan 2017; Lu et al. 2017; Chen et al. 2018). Thus, fundamental questions are raised: Can we view (residual-) deep models as dynamic systems and then use differential equation theory as a framework to analyze them? Can we design new and more powerful deep models inspired by dynamic systems? What are the benefits by introducing DE theory to deep learning? What are the potential new applications?

The boldest goal of this tutorial is to bridge the gap between the modern deep learning methods in computer science and DE theory (developed in control, applied math, physics, systems biology, numerical computation, etc.), and further to broaden the horizon of the deep learning methods with an emphasis on deep learning methods on graphs (also termed networks) and their applications. We will show that differential deep learning on graphs are powerful tools to model the structures and dynamics of networks and to generate molecular graphs. We summarize the contents covered in this tutorial as follows.

We will start by introducing the basic idea of Differential Deep Learning (DDL). We motivate the audiences by connecting the dots between deep learning models and dynamic systems modeled by differential equations. Basic ideas of differential equations (DEs) and physics models are illustrated. We will show examples of how to make deep structures into differential equation systems. We will conclude this part by summarizing the importance and values of differential deep learning.

In the perspective of dynamic systems, the residual learning framework, e.g. Residual-Net (He et al. 2016a; 2016b), is the approximation of differential equation (DE) systems by Euler method (1st order numerical method in solving DEs' initial value problems (Dormand 1996)). We show the evolutionary path from CNN to Residual-Net and further to ODE-Net (Chen et al. 2018). Inspired by the CNNs developed for images, Graph neural networks (GNN) (Kipf and Welling 2017) are proposed to deal with combinatorial graph data. We show how to get Residual-GNN, ODE-GNN, and our neural-dynamics-on-complex-network model. The promising experimental results demonstrate our model's capability of jointly capturing the structure, dynamics and semantics of complex systems (Zang and Wang 2019).

Normalizing flows (Rezende and Mohamed 2015) are one of the most promising deep generative models. We show how to use flow-based models for graphs, especially on molecular graph generation. We introduce our MoFlow model which achieved the state-of-the-art performance on molecular graph generation.

"Data Science lacking a model of reality may be statistics but hardly a science. –Judea Pearl". In physics, differential equations are one of the most successful models which describe the time-varying reality well. We motivate the problem by briefly introducing the history of network science. Then we show how to get DEs from statistical distribution functions by constructing dynamic systems. We show the relationships between DEs and the distributions functions.

Many DEs in network science and statistical physics can be found in a principled way (Zang et al. 2019).

Three applications of differential deep learning on graph models are discussed: molecular graph generation, namely to generate novel molecules with optimized properties; learning dynamics on graphs, namely to predict temporal change or final states of complex systems (Zang and Wang 2019), and mechanism discovery in graphs, namely to find dynamical laws of complex systems (Kutz et al. 2017) and distributions (Zang et al. 2019).

We conclude the tutorial by introducing related research topics. The limitations and potential future directions are then discussed. We will leave some time for questions and answers, feedback and further discussions.

Tutorial materials and outline

- Introduction: What's Differential Deep Learning and their graph applications (45 min)
  - Introduction to graphs and differential equations.
  - What is differential deep learning? Integrating differential equations into deep models.
  - Importance and challenges.
  - Graph applications: molecular graph generation, learning dynamics on graphs, and mechanism discovery.
- Molecular graph generation (45 min)
  - Drug discovery.
  - Molecular graph generation and normalizing flow models.
  - Our MoFlow model for molecular graph generation
  - Experiments on molecule generation, reconstruction, visualization and optimization.
- Learning dynamics on graphs. (45 min)
  - Learning dynamics of complex systems.
  - Our differential graph neural networks
  - Interpretations from Residual-Net, GNN, RNN, and temporal GNN.
  - Experiments on continuous-time network dynamics prediction, structured sequence prediction, and node classification
- Mechanism discovery in graphs (45 min)
  - From data to statistics, and further to dynamical models.
  - Density estimation in computer science vs. mechanism discovery in network science.
  - A theorem bridging distributions and differential equations.
  - Data-driven discovery of differential equations and distributions.
- Conclusion: Discussions and Future Directions (30 min)

Duration

3 hours, 30 minutes, plus 30-minute break. This tutorial will be held at The Thirty-Fourth AAAI Conference on Artificial Intelligence (AAAI-20), February 7-12, 2020 at the Hilton New York Midtown, New York, New York, USA.

Audience

All the researchers and practitioners engaged in data mining and machine learning are welcome. Basic knowledge on deep learning, graph mining, and differential equations is preferred but not required. The estimated number of participants is 100.

Presenter

Please kindly refer to the CVs of Dr. Chengxi Zang and Dr. Fei Wang as attached.

References

Chen, T. Q.; Rubanova, Y.; Bettencourt, J.; and Duvenaud, D. K. 2018. Neural ordinary differential equations. In *Advances in Neural Information Processing Systems*, 6571–6583.

Dormand, J. R. 1996. *Numerical methods for differential equations: a computational approach*, volume 3. CRC Press.

He, K.; Zhang, X.; Ren, S.; and Sun, J. 2016a. Deep residual learning for image recognition. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, 770–778.

He, K.; Zhang, X.; Ren, S.; and Sun, J. 2016b. Identity mappings in deep residual networks. In *European conference on computer vision*, 630–645. Springer.

Kipf, T. N., and Welling, M. 2017. Semi-supervised classification with graph convolutional networks. In *5th International Conference on Learning Representations, ICLR 2017, Toulon, France, April 24-26, 2017, Conference Track Proceedings*.

Kutz, J. N.; Rudy, S. H.; Alla, A.; and Brunton, S. L. 2017. Data-driven discovery of governing physical laws and their parametric dependencies in engineering, physics and biology. In *2017 IEEE 7th International Workshop on Computational Advances in Multi-Sensor Adaptive Processing (CAMSAP)*, 1–5. IEEE.

LeCun, Y.; Bengio, Y.; and Hinton, G. 2015. Deep learning. *nature* 521(7553):436.

Lu, Y.; Zhong, A.; Li, Q.; and Dong, B. 2017. Beyond finite layer neural networks: Bridging deep architectures and numerical differential equations. *arXiv preprint arXiv:1710.10121*.

Pearlmutter, B. A. 1995. Gradient calculations for dynamic recurrent neural networks: A survey. *IEEE Transactions on Neural networks* 6(5):1212–1228.

Rezende, D. J., and Mohamed, S. 2015. Variational inference with normalizing flows. *arXiv preprint arXiv:1505.05770*.

Weinan, E. 2017. A proposal on machine learning via dynamical systems. *Communications in Mathematics and Statistics* 5(1):1–11.

Zang, C., and Wang, F. 2019. Neural dynamics on complex networks.

Zang, C.; Cui, P.; Zhu, W.; and Wang, F. 2019. Dynamical origins of distribution functions. In *Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining*, 469–478. ACM.