Deep Learning on Graphs

Xavier Bresson

School of Computer Science and Engineering
Data Science and AI Research Centre
Nanyang Technological University (NTU), Singapore

QuantumBlack, McKinsey

Meetup "Machine Learning on Graphs"

Organizers : Jonathan Lofgren, Paul Beaumont, Andrew Tan

Nov 12th 2019

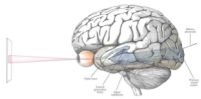
Outline

- Motivation
- Graph Deep Learning
- Applications
- DGL
- Tutorials
- Trainings

Outline

- Motivation
- Graph Deep Learning
- Applications
- DGL
- Tutorials
- Trainings

A Brief History of NNs/DL/AI



Visual primary cortex
Hubel-Wiesel
1959

First
NIPS

First
KDD
1989

RNN/LSTM
Schmidhuber

CNN
LeCun-Bengio

1997

1998

Hardware
GPU speed 2x/1.5 year

Big Data
Volume 2x/1.5 year

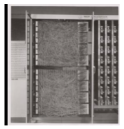
First Amazon
Cloud Center

Kaggle
Platform

GAN, Goodfellow
ResNet, He
Deep RL, Mnih
Attention, Bengio

Google AI
TensorFlow
Facebook AI
Torch

1958
Perceptron
Rosenblatt



1962
Birth of
Data Science
Split from Statistics
Tukey

1982
Backprop
Werbos-Hinton

1987
Neocognitron
Fukushima

1995
SVM/Kernel techniques
Vapnik

1999
First NVIDIA
GPU

2006

2010

2012
Deep Learning
Breakthough
Hinton

2014
Facebook Center
OpenAI Center

2015
Data scientist
1st Job in US

4th industrial revolution?
Digital intelligence
revolution
or new AI bubble?

AI Birth

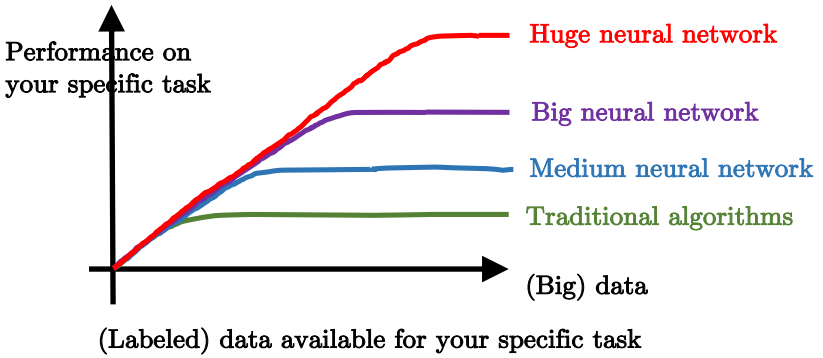
AI Winter [1960's-2012]

AI Renaissance

Kernel techniques
Graphical models
Wavelets, numerical PDEs
Handcrafted Features

Big Data

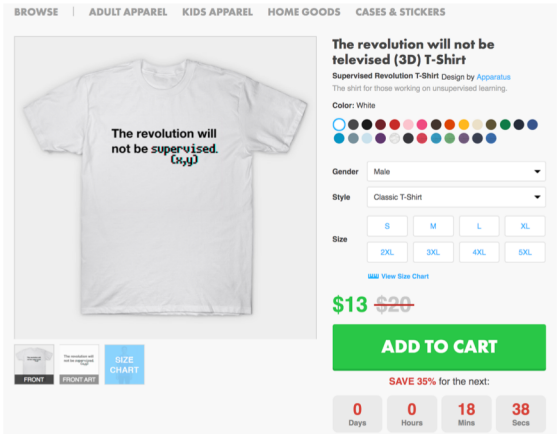
- Google's Law of Data : Volume of data double every 1.5 years.
  - Eric Schmidt : "Every 2 days we create as much information as we did up to 2003."
- Data generation/collection is easy but data labeling is challenging.
  - It is time and money consuming.
  - It requires humans (machine ability bounded by human intelligence).
- Why do we need labeled data? Because Deep Learning is supervised and it requires a lot of data.



Big Data

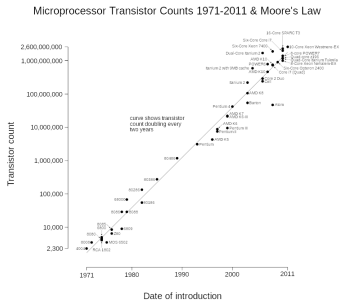
- How much data to get super-human performances?
  - ImageNet for image recognition has 1.4M images (machine error is 3% and human error is 5%).
  - AlphaGo Zero played 4.9 million games.
  - Self-driving cars : Tesla's self-driving cars collected 2 billion miles in 2019.

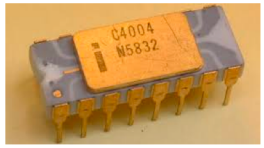
- The next AI revolution will not be supervised.



Computational Power

- Moore's Law of Computers :
  - 100x faster every decade
  - CPU, GPU/TPU, Cloud computing
  - Brain : 10^{14} synapses/10^3 Tflops = (regular) computers in 2030 (100 times faster than today)





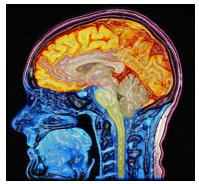
1st Intel CPU 1971

1st Nvidia GPU in 1999

1st Amazon Cloud in 2006

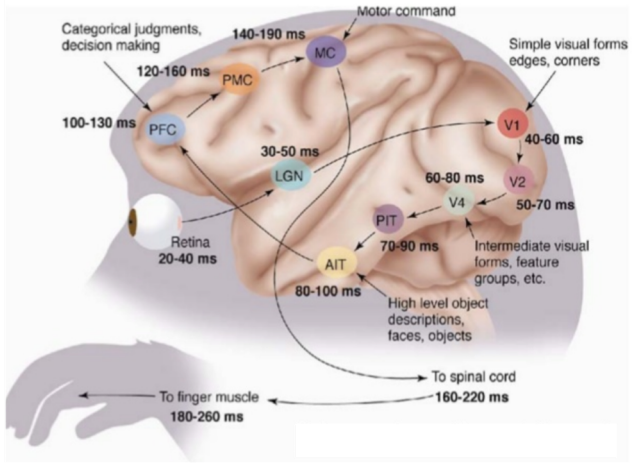
2030

=

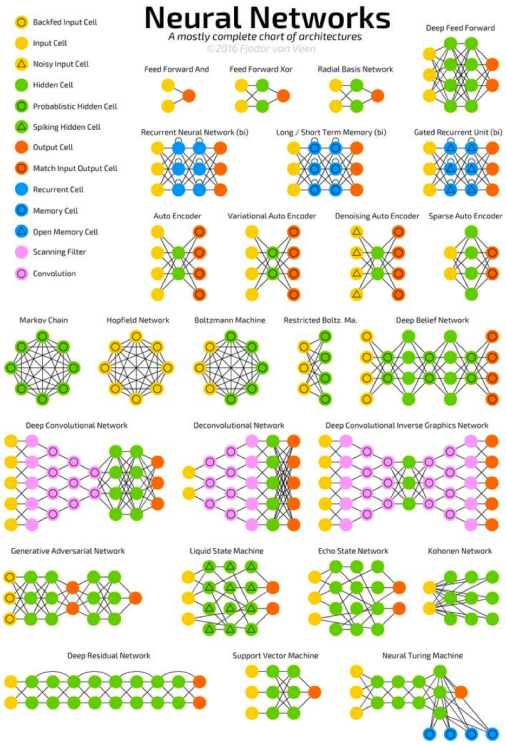


Neural Network Architectures

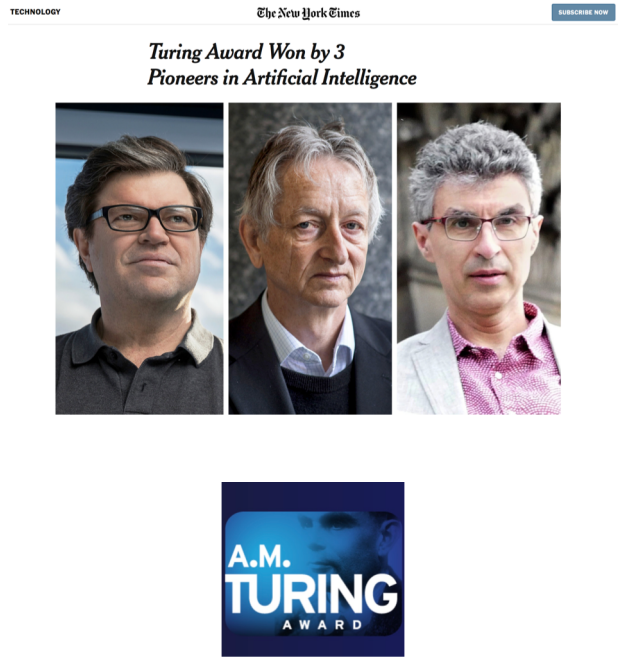
- Most research works focus on deep learning architectures.
- Spoiler : SkyNet does not exit - no architecture can solve simultaneously several tasks like the human brain.



Biological Neural Network

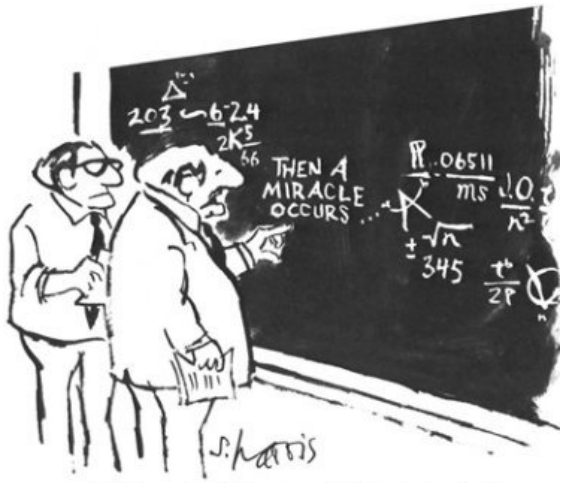


Artificial Neural Networks



Deep Learning Shortcomings

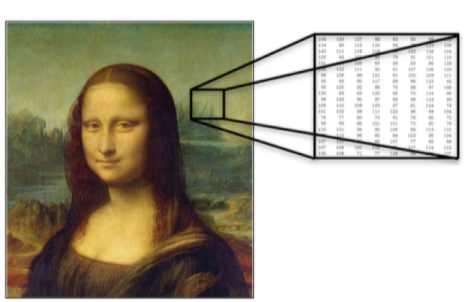
- Immature field
  - Engineering tricks, trials and errors.
  - Will take a decade to make it an engineering field (even more for developing a science).
- No math framework
  - Why performances are revolutionary?
  - New mathematical tools are needed from optimization, statistics, high-dimensional inference, non-linear algebra, non-convex analysis, etc.
- No real breakthrough beyond Computer Vision, Speech and NLP
  - How to generalize to complex data?

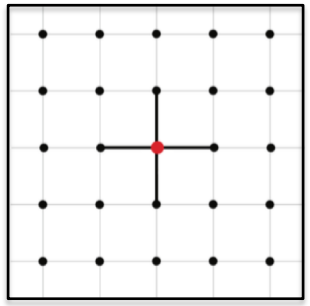


"I think you should be more explicit here in step two."

Regular Data

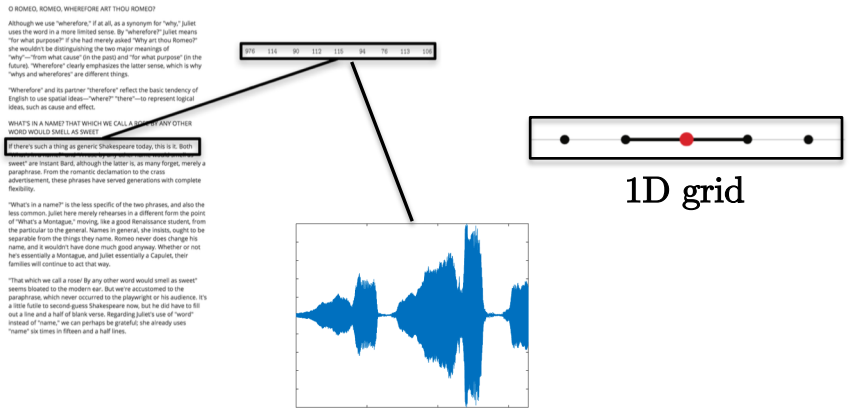
- Computer Vison:
  - Image, volume, video: 2D, 3D, 2D+1 Euclidean domains





2D grids

- NLP, speech recognition:
  - Sentence, sound: 1D Euclidean domain



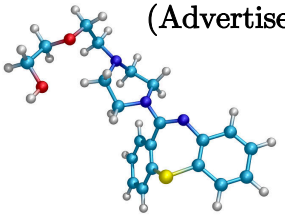
1D grid

- These domains have nice regular spatial structures :

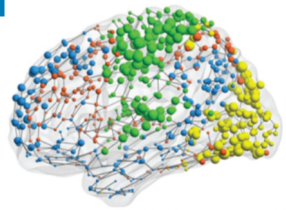
  ⇒ All operations are math well defined and fast O(n).

Graph Structured Data

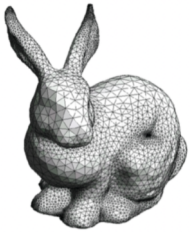
Social networks (Advertisement)



Drug/Material molecules (Chemistry)

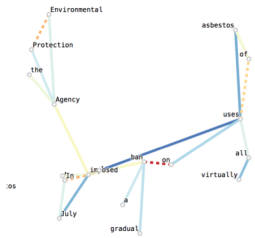


Brain connectivity (Neuroscience)

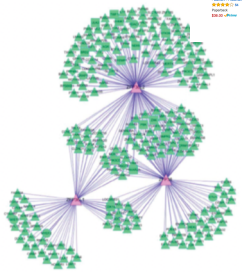


3D Meshes (Computer Graphics)

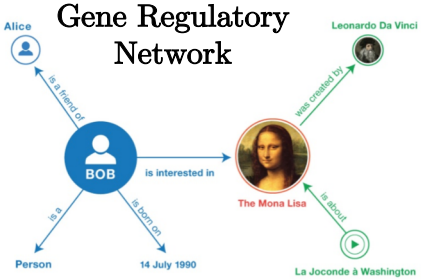
Transportation networks



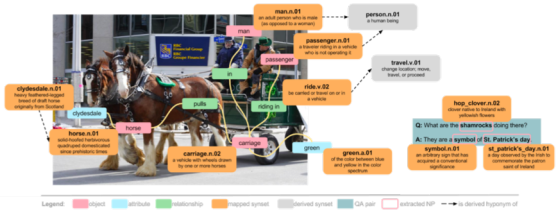
Words relationships (NLP)



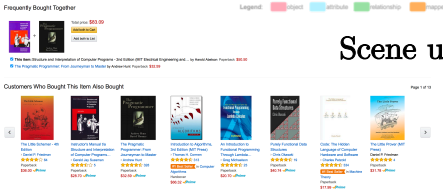
Gene Regulatory Network



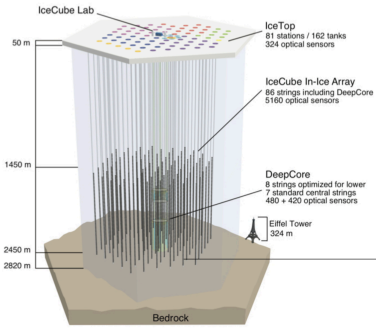
Knowledge graph (Causality)



Scene understanding

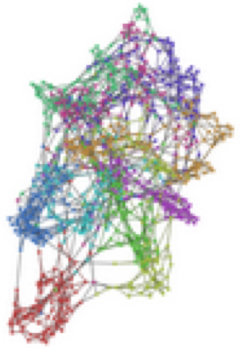


Recommender systems (Amazon, Netflix)



Neutrino detection (High-energy Physics)

=



Graphs/ Networks

Outline

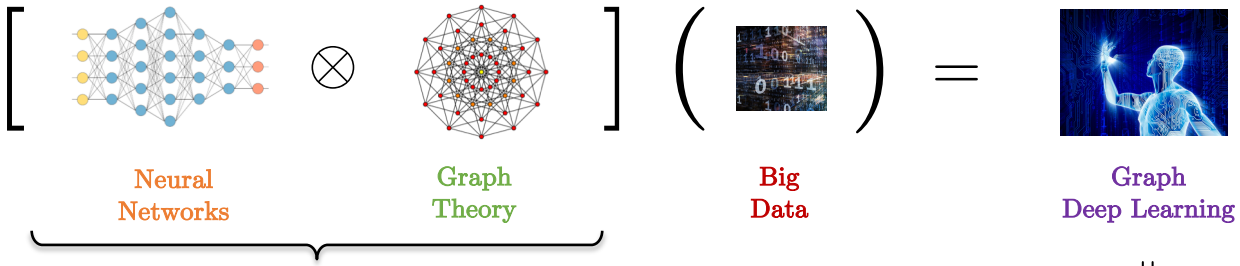
- Motivation
- Graph Deep Learning
- Applications
- DGL
- Tutorials
- Trainings

Deep Learning on Graphs

- New deep learning techniques :
  - Generalizing deep learning to graph-structured data.
- Challenges :
  - How to extend deep learning to graphs?
  - How to define invariances on graphs? (convolution and pooling on graphs)
  - How to make them fast? (linear complexity)

NATIONAL RESEARCH FOUNDATION
PRIME MINISTER'S OFFICE
SINGAPORE

NRF Fellowship 2017
2.5M$ funding



New computational framework for big data science

⇓

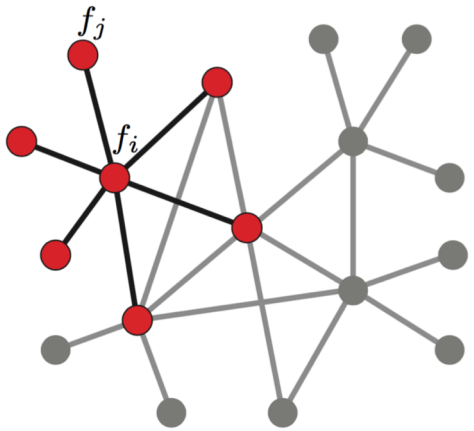
Tackle new real-world problems

Spectral GNNs

- Laplacian operator :

$$(\Delta f)_i = \frac{1}{b_i} \sum_{j:(i,j)\in\mathcal{E}} a_{ij}(f_i - f_j)$$

Difference between local f_i value and its local average f_j (2^{nd} derivative on graphs)



$\mathcal{G} = (\mathcal{V}, \mathcal{E})$

- Core operator in Spectral Graph Theory.
  - Normalized Laplacian :

$$\mathbf{\Delta} = \mathbf{I} - \mathbf{D}^{-1/2}\mathbf{A}\mathbf{D}^{-1/2}$$

where $\mathbf{A} = (a_{ij})$ and $\mathbf{D} = \text{diag}(\sum_{j\neq i} a_{ij})$

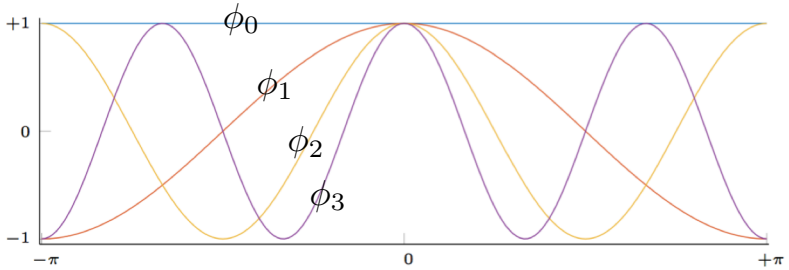
  - Eigen-decomposition :

$$\mathbf{\Delta} = \mathbf{\Phi}^T \mathbf{\Lambda} \mathbf{\Phi}$$

where $\mathbf{\Delta}\boldsymbol{\phi}_k = \lambda_k \boldsymbol{\phi}_k, \quad k = 1, 2, \ldots$ (Graph) Fourier modes !

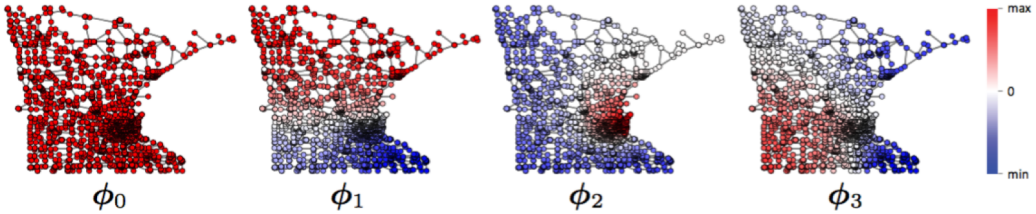
Spectral GNNs

- Regular domain :



First eigenvectors of 1D Euclidean Laplacian = standard Fourier basis

- Graph domain :



First Laplacian eigenvectors of a graph

Fourier modes related to graph geometry

(s.a. communities, hubs, etc)

Spectral Graph Clustering

Spectral GNNs

- Graph convolution can be computed in the Fourier/spectral domain :

$$(f \star g)_i = \underbrace{\sum_{k \geq 1} \underbrace{\langle f, \phi_k \rangle_{L^2(\mathcal{V})} \langle g, \phi_k \rangle_{L^2(\mathcal{V})}}_{\text{product in the Fourier domain}} \phi_{k,i}}_{\text{inverse Fourier transform}}$$

- In matrix-vector notation :

Inverse Fourier Transform

Fourier Transform

$$\mathbf{f} \star \mathbf{g} = \mathbf{\Phi} \left(\mathbf{\Phi}^\top \mathbf{g} \circ \mathbf{\Phi}^\top \mathbf{f} \right) = \hat{g}(\mathbf{\Delta}) \mathbf{f}$$

Spectral filter

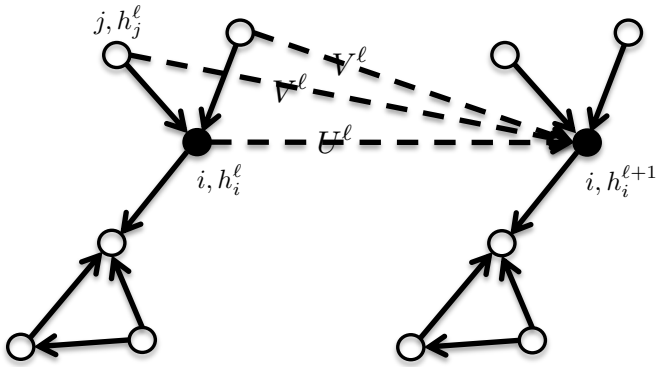
- YouTube video (20k views) :

  https://www.youtube.com/watch?v=v3jZRkvIOIM

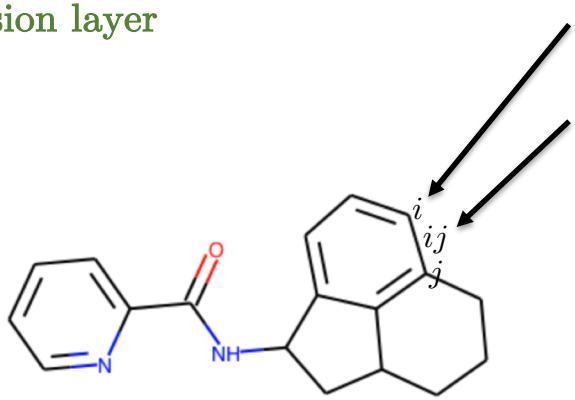
Spatial GNNs

- Vanilla GraphConvNets :

$$h_i^{\ell+1} = \mathrm{ReLU}\Big(U^\ell h_i^\ell + V^\ell \sum_{j \to i} h_j^\ell\Big), \quad h_i^{\ell=0} = x_i$$



- Application : Molecule property regression
  - Vertex, edge and graph embedding layers
  - Regression layer

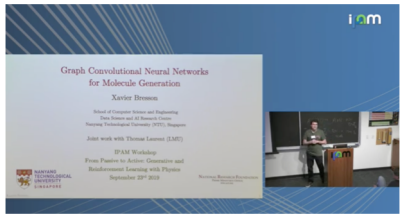


$$h_i^{\ell+1} = \mathrm{ReLU}\Big(U^\ell h_i^\ell + \sum_{j \to i} \sigma(e_{ij}^\ell) \odot V^\ell h_j^\ell\Big), \quad h_i^{\ell=0} = \text{atom\_type}$$

$$e_{ij}^{\ell+1} = \mathrm{ReLU}\Big(A^\ell e_{ij}^\ell + B^\ell h_i^{\ell+1} + C^\ell h_j^{\ell+1}\Big), \quad e_{ij}^{\ell=0} = \text{bond\_type}$$

$$z = \sum_i W_1 h_i^L + \sum_{ij} W_2 e_{ij}^L \quad \text{Molecule embedding}$$

$$p = \mathrm{MLP}(z) \quad \text{Molecule property}$$



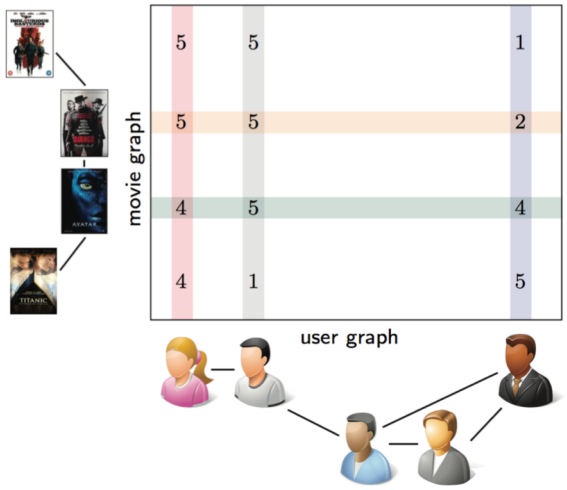
https://bit.ly/2mXrcVr

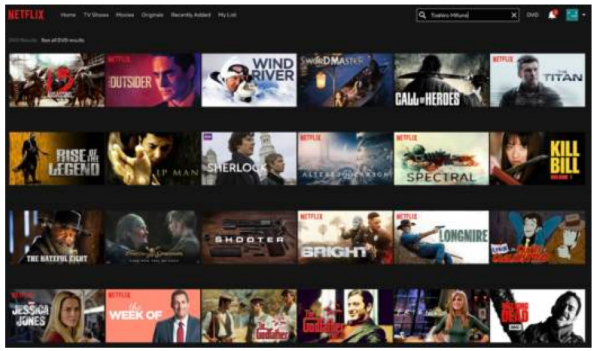
Outline

- Motivation
- Graph Deep Learning
- Applications
- DGL
- Tutorials
- Trainings

Applications

- Recommender systems





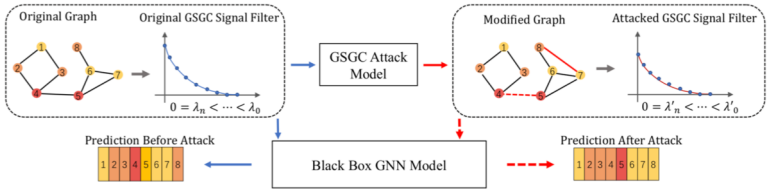
U.S. Patent on Graph Deep Learning
sold to Twitter on May 2019.

Applications

- Social networks :
  - Advertisement
  - Fake News (2016 US Presidential Election) /Adversarial attacks

Twitter network

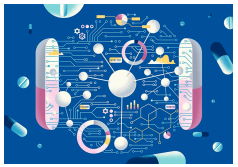
Facebook network

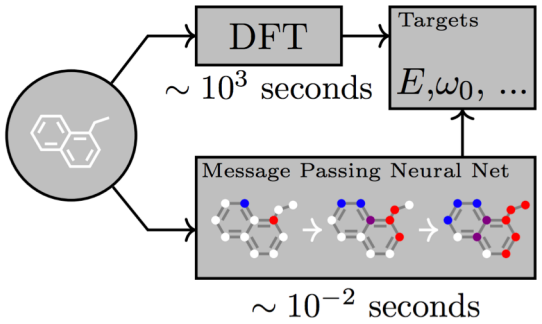


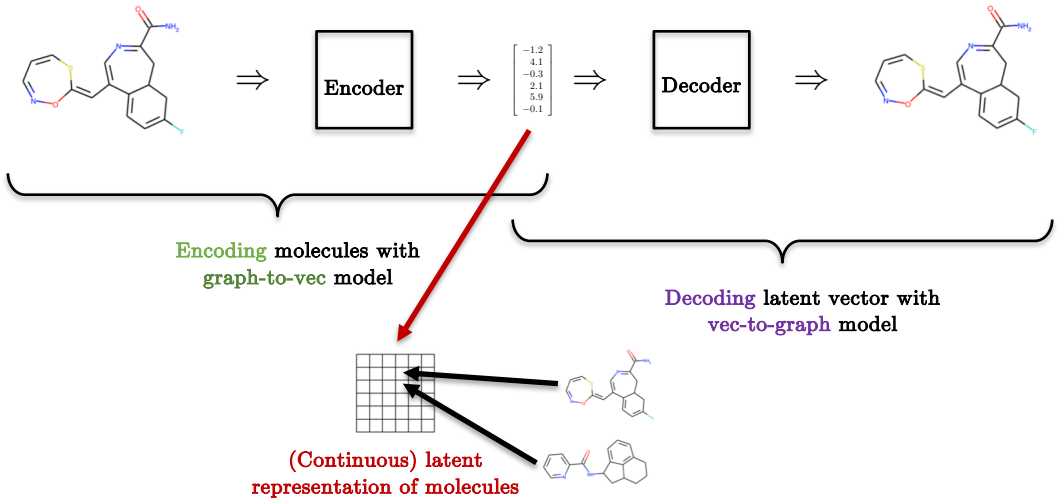
Applications

- Chemistry : Drugs and Materials Design
  - Self-driving lab to create drugs and materials quicker and less expensively (takes a decade and costs $350M-$2.7B to bring a new drug to market)

    https://fortune.com/2019/09/17/artificial-intelligence-insilico-deepmind

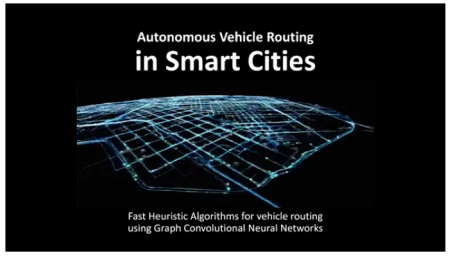




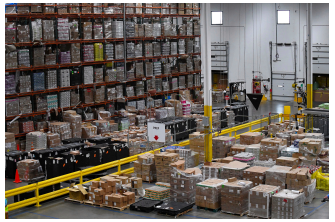


Applications

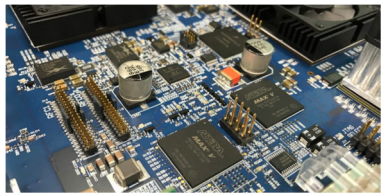
- Operations Research/Combinatorial Optimization s.a. assignment, routing, planning, supply chain, scheduling are used every day in revenue management, transportation, manufacturing, supply chain, public policy, hardware design, etc.
- Most OR problems are NP-hard.
- Traditional OR solvers are hand-crafted algorithms with years of research work and significant specialized knowledge.
- DL can learn universal high-quality algorithms to solve OR problems with GNNs and DRL.



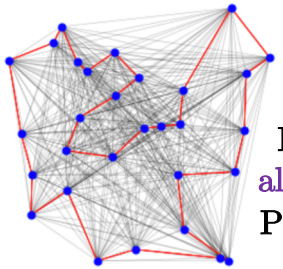
Uber taxis allocation



Amazon warehouse management



Hardware pieces placement



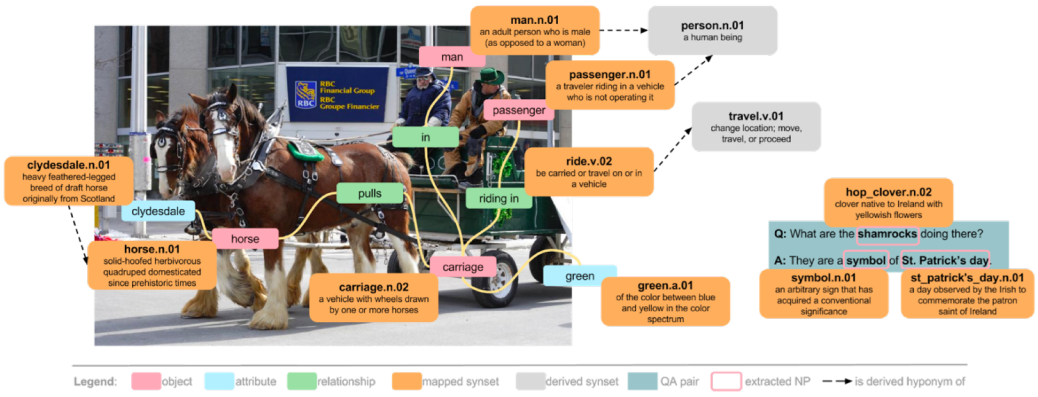
Recent efforts to design learning algorithms for Travelling Salesman Problem (TSP) (proof-of-concept).

Workshop "Deep Learning and Combinatorial Optimization" at IPAM-UCLA, Feb 2021.

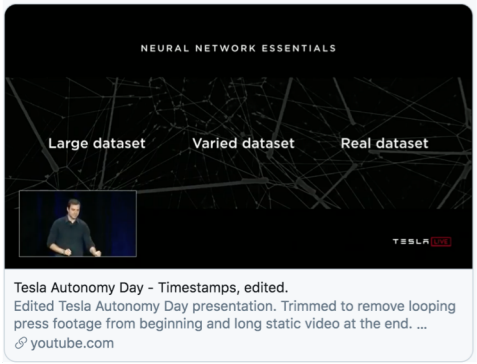
Y. LeCun (NYU, Facebook)
O. Vinyals (DeepMind)
P. Battaglia (DeepMind)
A. Lodi (Montréal)
M. Welling (Amsterdam, Qualcomm)
X. Bresson (NTU, Singapore)

Applications

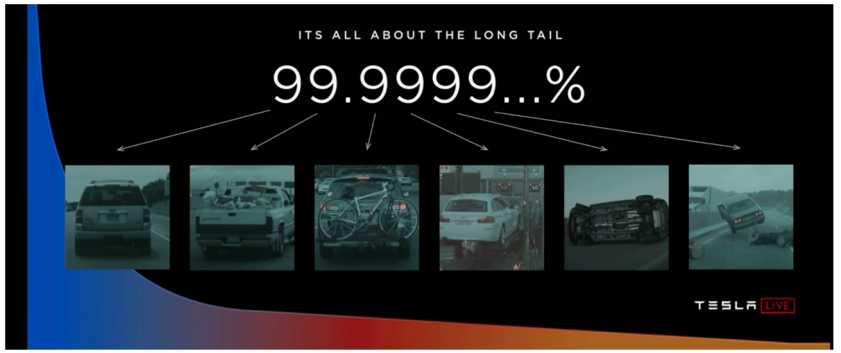
- Computer Vision with common sense : Autonomous driving



Scene graph understanding

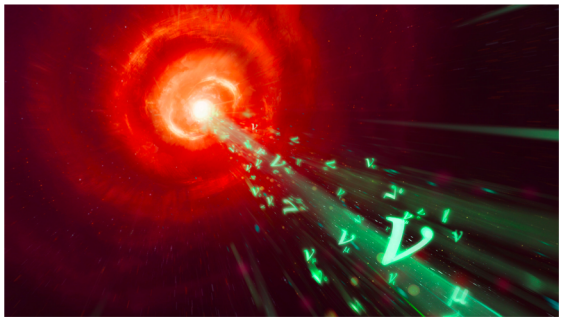


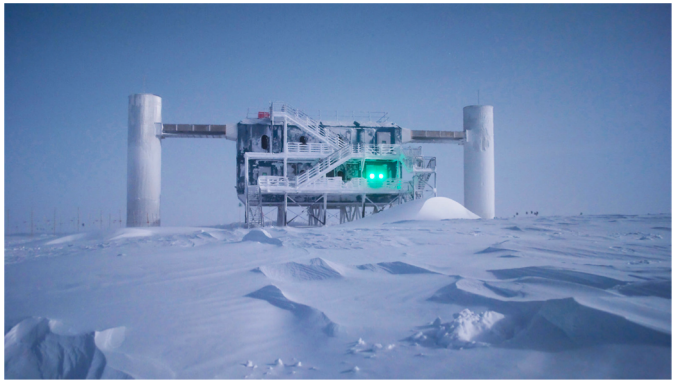
A. Karpathy (Tesla AI Autopilot Director)
https://youtu.be/2PpNmSdFP7Q?t=3369



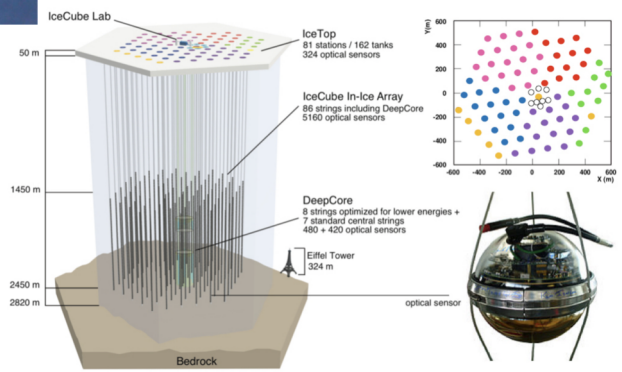
Applications

- High-energy physics with neutrino detection (hard to detect because they have a very small chance of interacting with regular matter).

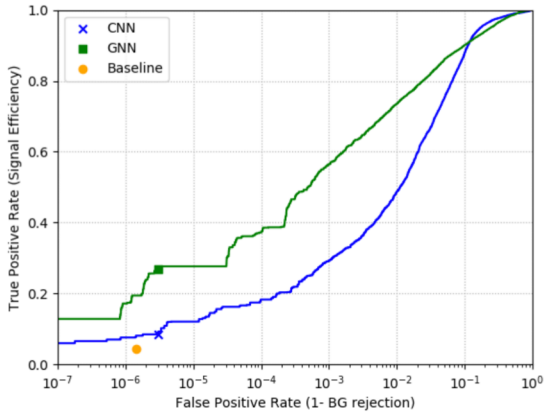




IceCube Neutrino Observatory

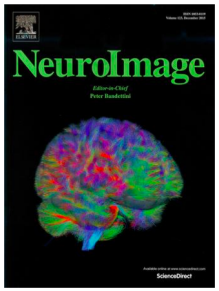


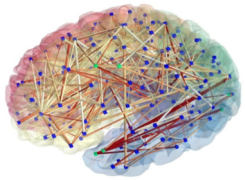
5,160 sensors



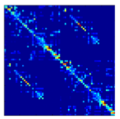
Applications

- Neuroscience : Neuro-degenerative diseases
- Understanding the fundamental relationship brain structure-function using GNNs.
  - Brain analysis (resting states related to vision, body motor, language, etc).
  - Classification/prediction of neuro-diseases.

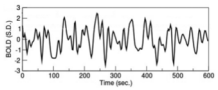




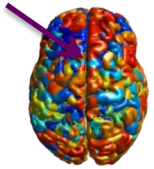
Structural MRI
Brain connectivity



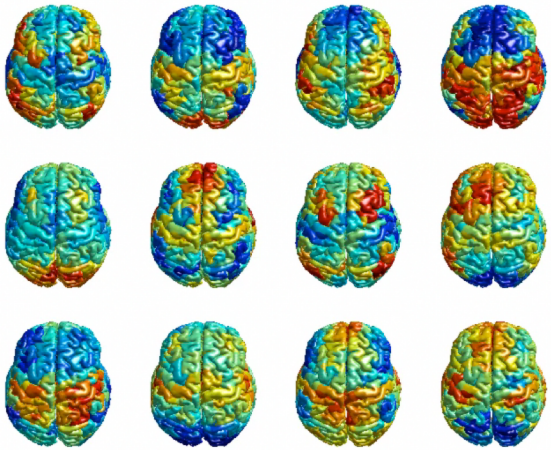
Graph
Connectivity



Time series
at this location



Functional MRI
Dynamic brain activity



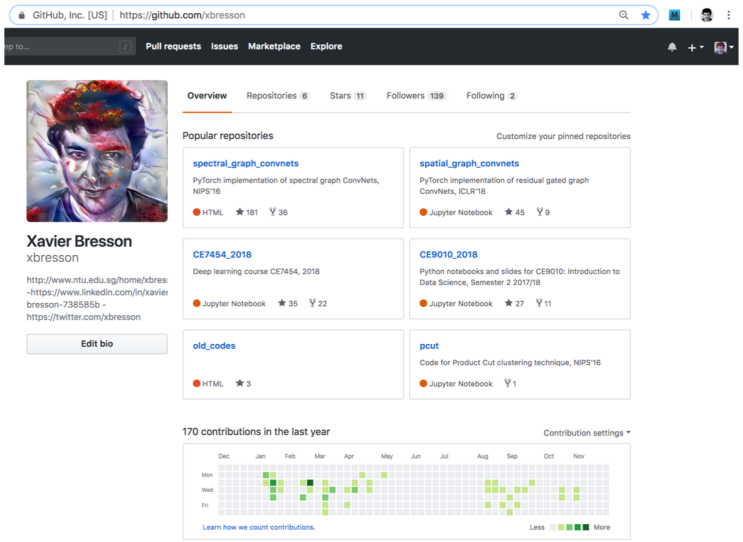
Resting States

Outline

- Motivation
- Graph Deep Learning
- Applications
- **DGL**
- Tutorials
- Trainings

Codes

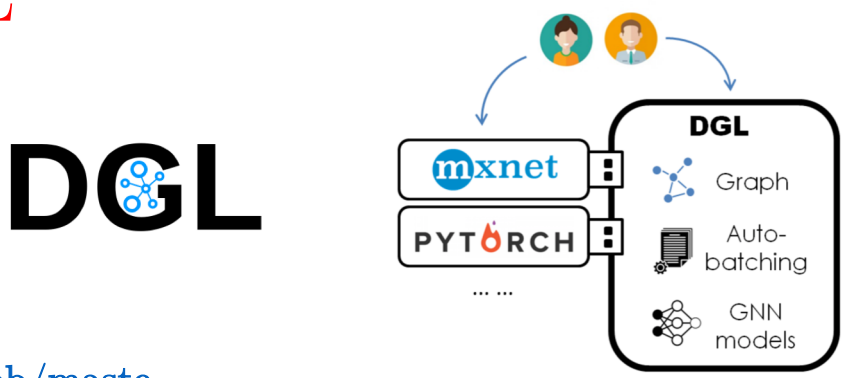
- My GitHub
  - Spectral GNNs
    - https://github.com/xbresson/spectral_graph_convnets
  - Spatial GNNs
    - https://github.com/xbresson/spatial_graph_convnets



DGL

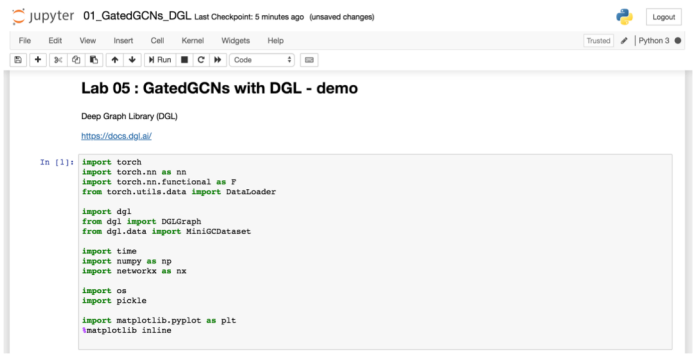
- GNNs with DGL (Deep Graph Library), NYU

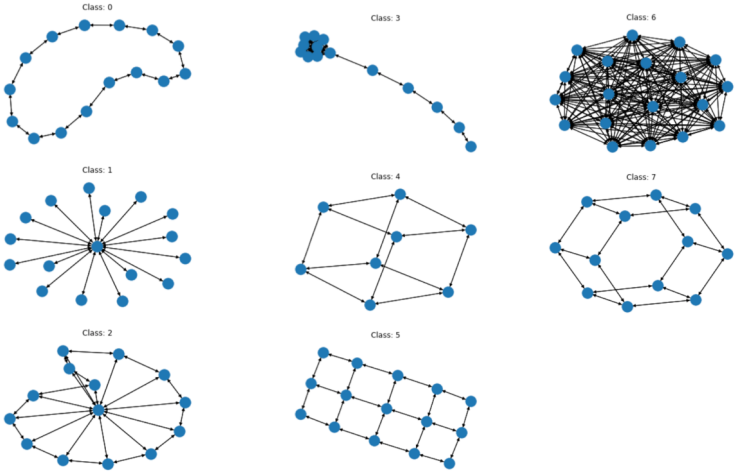
  https://www.dgl.ai



- A simple example to start : Graph classification

  https://github.com/xbresson/CE7454_2019/blob/master/codes/labs_lecture14/lab05_DGL/01_GatedGCNs_DGL.ipynb





Outline

- Motivation
- Graph Deep Learning
- Applications
- DGL
- **Tutorials**
- Trainings

Tutorials on Graph Deep Learning

NeurIPS'17

1,000-2,000 participants

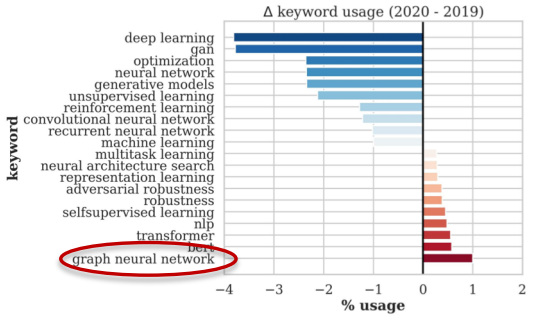
CVPR'17

500-1,000 participants

SIAM Annual Meeting'18

- Organizers :
  - Y. LeCun (NYU, Chief AI Scientist at Facebook, 2018 Turing Award)
  - M. Bronstein (Imperial, Head of Deep Learning at Twitter London)
  - J. Bruna (NYU)
  - A. Szlam (Facebook AI Research)
  - X. Bresson (NTU, Singapore)

- Top AI/DL conferences organize workshops/tutorials on "Graph Neural Networks" :
  - NeurIPS'19, ICML'19, ICLR'19
  - CVPR'19, ICCV'19
  - Etc



ICLR'20

2018 IPAM-UCLA Workshop on GDL

Organizing committee:

- Y. LeCun (NYU, Chief AI Scientist at Facebook, 2018 Turing Award)
- M. Bronstein (Imperial, Head of Deep Learning at Twitter London)
- J. Bruna (NYU)
- A. Szlam (Facebook AI Research)
- S. Osher (UCLA, U.S. National Academy of Sciences)
- X. Bresson (NTU, Singapore)

Workshops

Programs > Workshops > New Deep Learning Techniques

New Deep Learning Techniques

FEBRUARY 5 - 9, 2018

OVERVIEW | SPEAKER LIST | SCHEDULE | APPLY/REGISTER

Overview

In recent years, artificial neural networks a.k.a. deep learning have significantly improved the fields of computer vision, speech recognition, and natural language processing. The success relies on the availability of large-scale datasets, the developments of affordable high computational power, and basic deep learning operations that are sound and fast as they assume that data lie on Euclidean grids. However, not all data live on regular lattices. 3D shapes in computer graphics represent Riemannian manifolds. In neuroscience, brain activity (fMRI) is encoded on the structural connectivity network (sMRI). In genomics, the human body functionality is expressed through DNA, RNA, and proteins that form the gene regulatory network (GRN). In social sciences, people interact through networks. Eventually, data in communication networks are structured by graphs like the Internet or road traffic networks.

Deep learning that has originally been developed for computer vision cannot be directly applied to these highly irregular domains, and new classes of deep learning techniques must be designed. This is highly challenging as most standard data analysis tools cannot be used on heterogonous data domains. The workshop will bring together experts in mathematics (statistics, harmonic analysis, optimization, graph theory, sparsity, topology), machine learning (deep learning, supervised & unsupervised learning, metric learning) and specific applicative domains (neuroscience, genetics, social science, computer vision) to establish the current state of these emerging techniques and discuss the next directions.

This workshop will include a poster session; a request for posters will be sent to registered participants in advance of the workshop.

Video talks : https://lnkd.in/fY2-9UU

Speaker List

Alán Aspuru-Guzik (Harvard University)
Samuel Bowman (New York University)
Xavier Bresson (Nanyang Technological University, Singapore)
Michael Bronstein (USI Lugano, Switzerland, / Tel Aviv University, Israel / Intel Perceptual
Joan Bruna (New York University)
Pratik Chaudhari (University of California, Los Angeles (UCLA))
Kyle Cranmer (New York University)
Michael Elad (Technion - Israel Institute of Technology, Computer Science Department)
Sanja Fidler (University of Toronto)
Emily Fox (University of Washington)
Tom Goldstein (University of Maryland)
Leonidas Guibas (Stanford University, Computer Science)
Yann LeCun (New York University, Canadian Institute for Advanced Research)
Jure Leskovec (Stanford University)
Stéphane Mallat (École Normale Supérieure)
Federico Monti (Universita della Svizzera Italiana)
Stanley Osher (University of California, Los Angeles (UCLA), Mathematics)
Ellie Pavlick (University of Pennsylvania)
Daniel Rueckert (Imperial College)
Ruslan Salakhutdinov (Carnegie-Mellon University)
Zuowei Shen (National University of Singapore, mathematics)
Stefano Soatto (University of California, Los Angeles (UCLA))
Sainbayar Sukhbaatar (New York University)
Arthur Szlam (Facebook)
Raquel Urtasun (University of Toronto)
Wei Zhu (Duke University, Mathematics)

2019 IPAM-UCLA Workshop on GDL

Organizing committee:

- Y. LeCun (NYU, Chief AI Scientist at Facebook, 2018 Turing Award)
- Rene Vidal (Johns Hopkins, Director Data Science Institute)
- Rebecca Willett (Chicago U.)
- S. Osher (UCLA, U.S. National Academy of Sciences)
- X. Bresson (NTU, Singapore)

Video talks : https://bit.ly/2w8EtLV

Speaker List

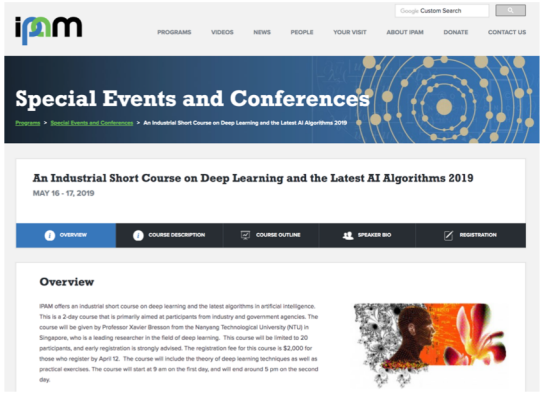
Mikhail Belkin (Ohio State University)
Xavier Bresson (Nanyang Technological University, Singapore)
Soumith Chintala (Facebook AI Research)
Taco Cohen (Qualcomm AI Research)
Kostas Daniilidis (University of Pennsylvania)
Tom Goldstein (University of Maryland)
Bahram Jalali (University of California, Los Angeles (UCLA))
Thomas Kipf (Universiteit van Amsterdam)
Roy Lederman (Yale University, Applied Mathematics)
Jure Leskovec (Stanford University)
Federico Monti (Universita della Svizzera Italiana)
Mathias Niepert (NEC Laboratories Europe)
Stanley Osher (University of California, Los Angeles (UCLA), Mathematics)
Hamed Pirsiavash (University of Maryland Baltimore County)
Marc Pollefeys (ETH Zurich)
Srikumar Ramalingam (University of Utah)
Thiago Serra (Mitsubishi Electric Research Laboratories (Merl))
Jeremias Sulam (Johns Hopkins University)
Arthur Szlam (Facebook)
Jian Tang (HEC Montréal)
Luc Van Gool (ETH Zurich)
Rene Vidal (Johns Hopkins University)
Ersin Yumer (Uber ATG)
Hongyang Zhang (Carnegie Mellon University)

Outline

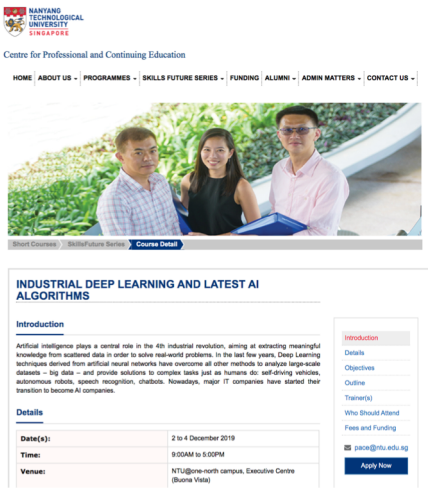
- Motivation
- Graph Deep Learning
- Applications
- DGL
- Tutorials
- **Trainings**

Trainings

- I have offered since 2014 "Industrial Short Course on Deep Learning and the Latest AI Algorithms".
  - 1-day/2-day/3-day in-house trainings for companies,
    - Introductory or advanced course
    - Former clients : Fortune500 like Deloitte, UnitedHealth, Nestle, UBS, SwissCom, Total, AMD, Apple, etc.
  - 3-day trainings at NTU, as part of SkillsFuture.
- Balance between theory and coding :
  - Lectures on main ideas and technical details.
  - Coding exercises (integrated to lectures) to apply the new concepts.



A recent professional course at UCLA, May 2019
https://bit.ly/2NWzcQe



Professional course at NTU,
Dec 2-4 2019
https://bit.ly/33FXbd6
Registration closes on Nov 18th.

Questions?

Xavier Bresson

xbresson@ntu.edu.sg

https://twitter.com/xbresson

https://github.com/xbresson

http://www.ntu.edu.sg/home/xbresson

https://www.facebook.com/xavier.bresson.1

https://www.linkedin.com/in/xavier-bresson-738585b