

CS6208 : Advanced Topics in Artificial Intelligence

Graph Machine Learning

Lecture 1: Introduction to Graph Deep Learning

Semester 2 2022/23

Xavier Bresson

<https://twitter.com/xbresson>

Department of Computer Science
National University of Singapore (NUS)



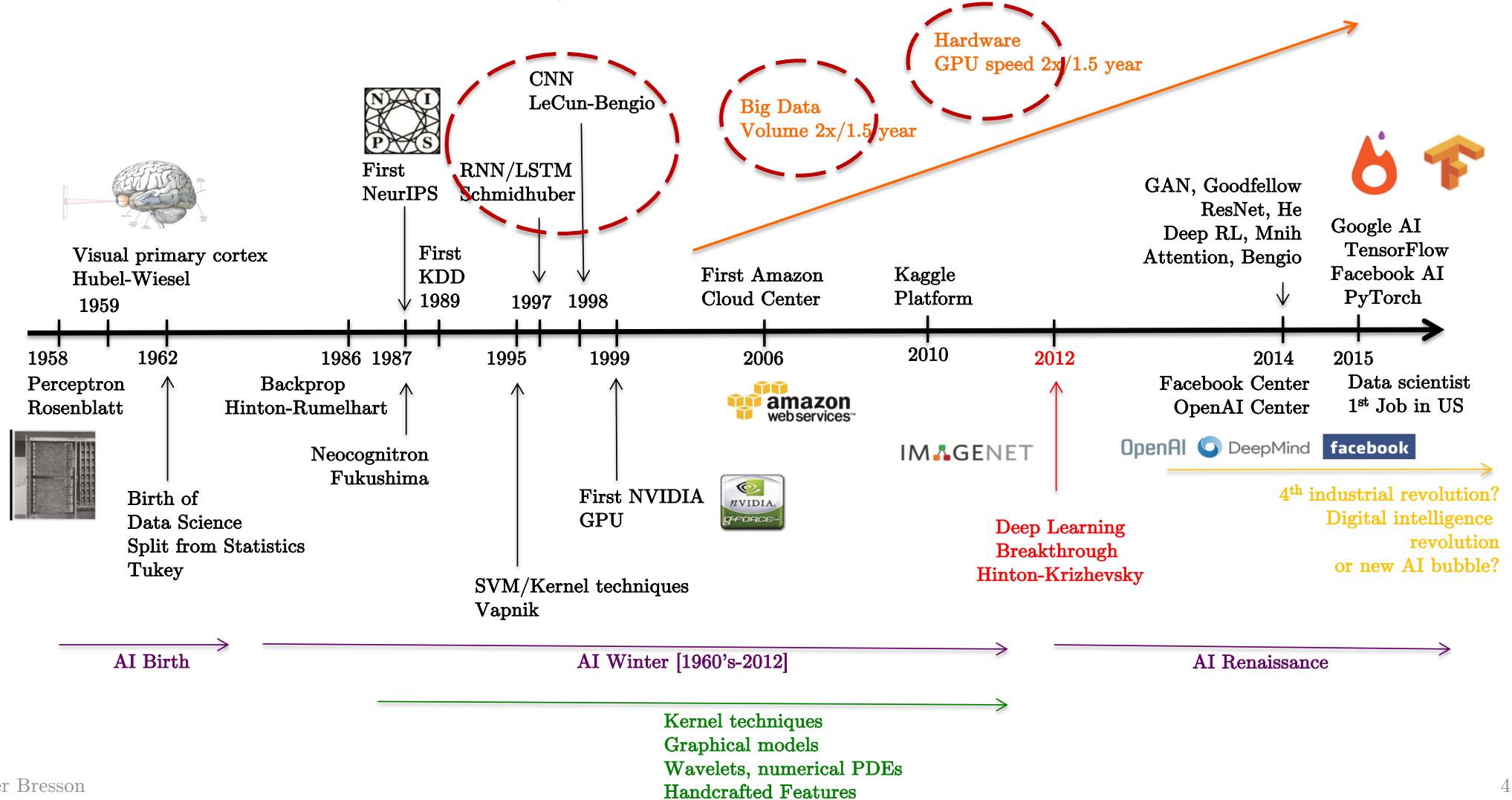
Outline

- The Deep Learning (DL) revolution
- Limitation of DL
- Graph-Structured data
- Graph Neural Networks (GNNs)
- GNN case studies
- GNN for industry
- GNN libraries
- Conclusion

Outline

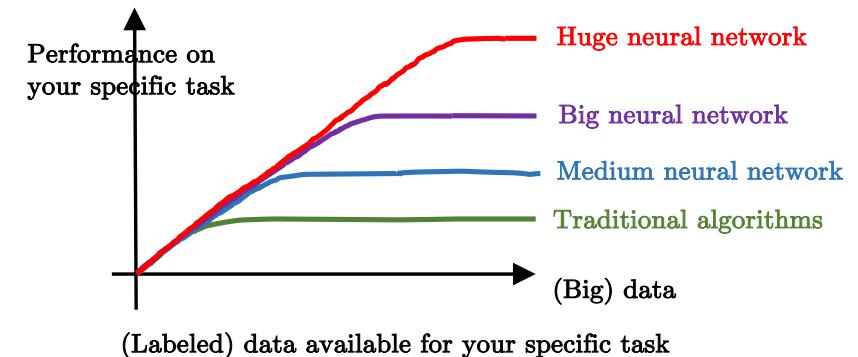
- The Deep Learning (DL) revolution
- Limitation of DL
- Graph-Structured data
- Graph Neural Networks (GNNs)
- GNN case studies
- GNN for industry
- GNN libraries
- Conclusion

AI : Birth, Winter and Renaissance



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 mostly supervised and hence 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.



BROWSE | ADULT APPAREL KIDS APPAREL HOME GOODS CASES & STICKERS

The revolution will not be televised (3D) T-Shirt
Supervised Revolution T-Shirt Design by Apparatus
The shirt for those working on unsupervised learning.
Color: White

Gender: Male

Style: Classic T-Shirt

Size: S M L XL
2XL 3XL 4XL 5XL

\$13 ~~\$20~~

ADD TO CART

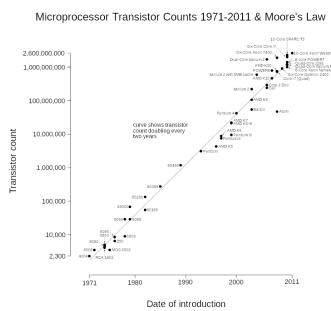
SAVE 35% for the next:
0 Days 0 Hours 18 Mins 38 Secs

FRONT FRONT ART SIZE CHART

A white t-shirt is displayed with the text "The revolution will not be supervised. (x,y)" printed on the front. Below the shirt, there are three small buttons labeled "FRONT", "FRONT ART", and "SIZE CHART".

Computational Power

- 1975 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

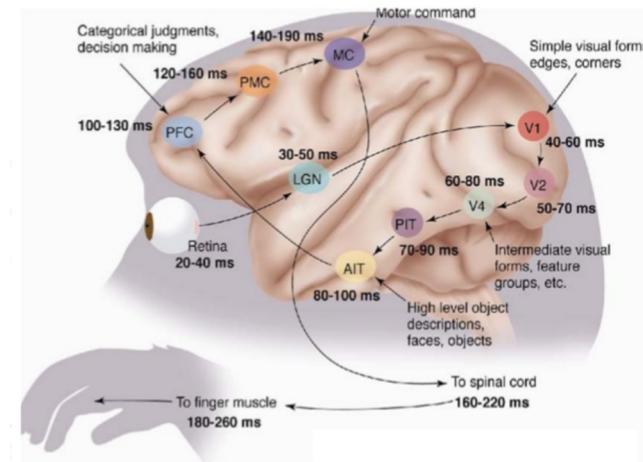


2030
=

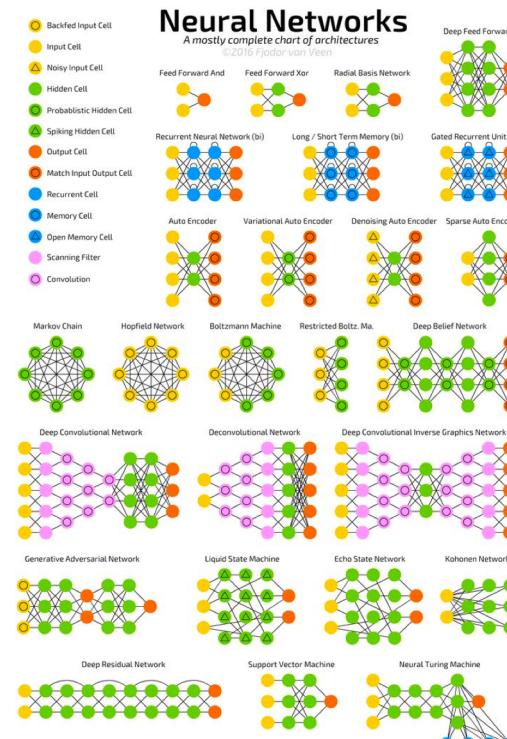


Neural Network Architectures

- Most research/academic works focus on deep learning architectures.
 - Industry labs leverage academic ideas to large-scale/impressive experiments.
- Still, no architecture can solve simultaneously many tasks like the human brain (but progress is made).



Biological Neural Network



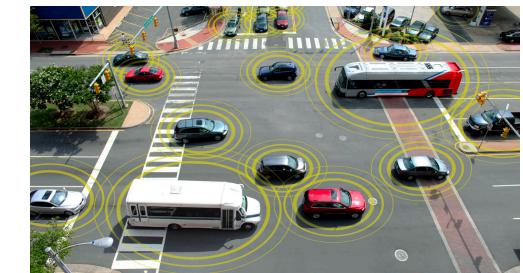
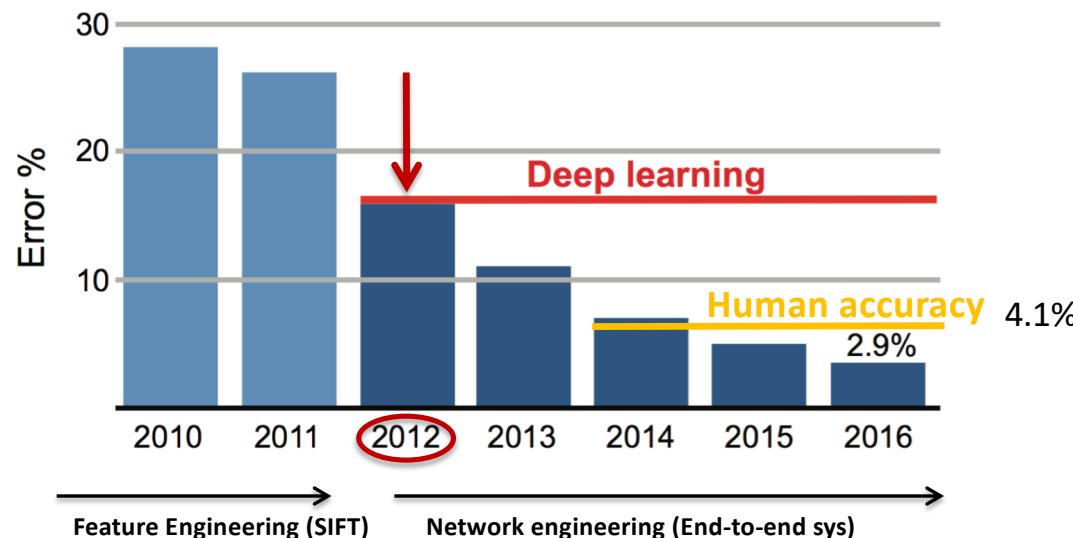
Artificial Neural Networks

The Deep Learning Revolution

- Breakthrough in Computer Vision (CNNs)

LeCun, Bottou, Bengio, Haffner 1998
Krizhevsky, Sutskever, Hinton, 2012

IMAGENET



- Breakthrough in Speech and Natural Language Processing (RNNs/Transformers)
- Hochreiter, Schmidhuber, 1997 and Vaswani et-al, 2017

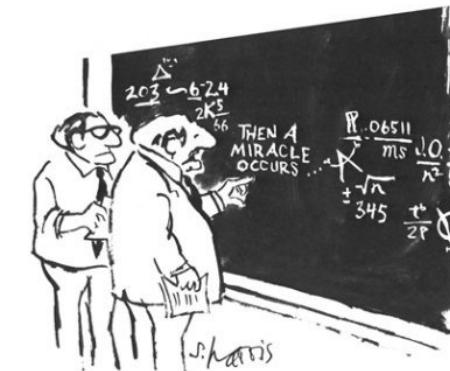
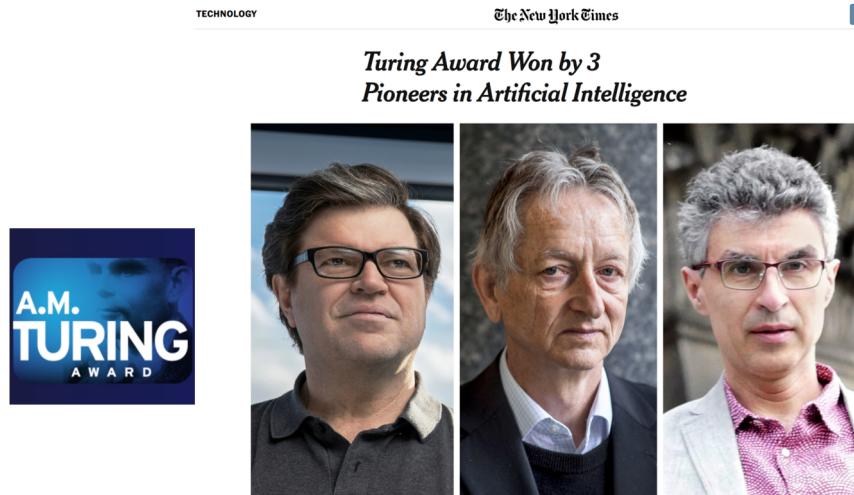
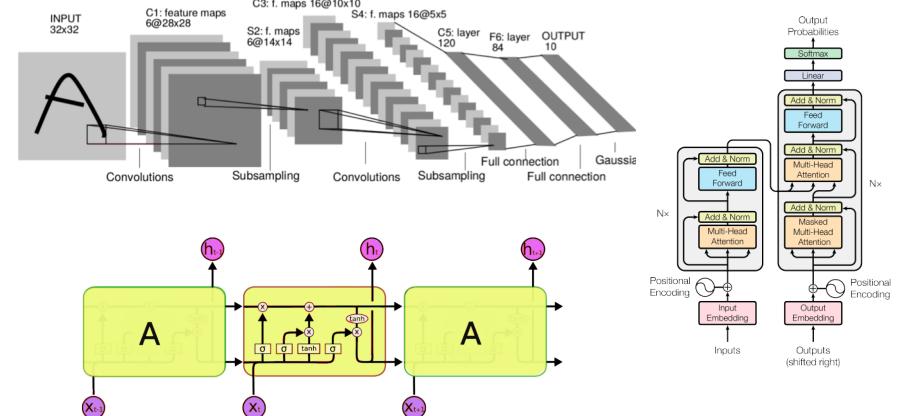
Xavier Bresson

Oliver Schröbe [DE] Ich verstehe nicht, weshalb sich alle plötzlich darüber aufregen, dass wir Steuergelder nach Griechenland "verschwendeln". Als wären Steuergelder bisher sinnvoll angelegt worden...
May 5 at 10:06am via TweetDeck · Comment · Like
· Translate

Oliver Schröbe [DE] I do not understand why all of a sudden get e about the fact that we are wasting taxpayers' money to Greece. W tax money would have been useful to create ...
May 5 at 10:06am via TweetDeck · Comment · Like
· Untranslate

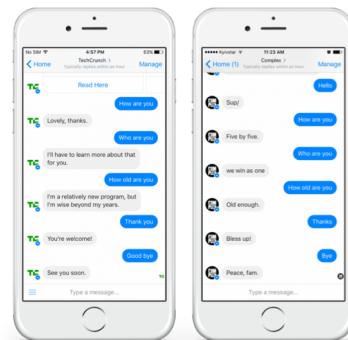
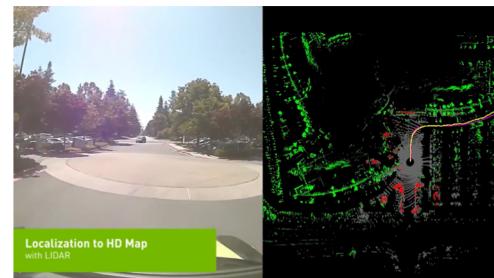
The Power of Deep Learning

- CNNs/RNNs/Transformers are powerful architectures to solve high-dimensional learning problems.
- Curse of dimensionality :
 $\text{dim(image)} = \text{dim}(1000 \times 1000) = 10^6$
For N=10 samples/dim $\Rightarrow 10^{1,000,000}$ points



Applications

- Computer Vision : Autonomous driving, Face recognition
- Natural Language Processing : Machine Translation, Q&A Chatbot, Text generation
- Speech Recognition : Virtual assistants (Alexa/Siri/Google/Cortana)



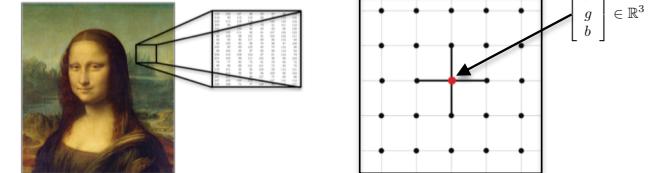
A complex image block. On the left is a colorful heatmap visualization of text data. To its right is a snippet of text from a blog post: 'Better Language Models and Their Implications'. The text discusses training a large-scale unsupervised language model that achieves state-of-the-art performance on benchmarks and performs tasks like reading comprehension, machine translation, and question answering without task-specific training. At the bottom right is a small watermark: 'FEBRUARY 14, 2019 24 MINUTE READ'.

Outline

- The Deep Learning (DL) revolution
- **Limitation of DL**
- Graph-Structured data
- Graph Neural Networks (GNNs)
- GNN case studies
- GNN for industry
- GNN libraries
- Conclusion

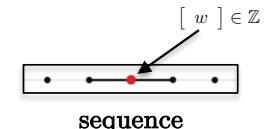
Limitations of Deep Learning

- Standard DL such as:
 - CNNs, RNNs, Transformers
 - Modern DL toolboxes (TensorFlow, PyTorch)
 - Successful industries for CV/NLP/SR (Google, Meta, etc.)



Computer Vision (CV)

grid



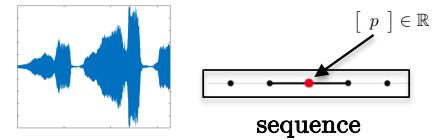
sequence

- Do we have data NOT based on grid/sequence?
Yes, quite a lot! (see next slide)



Natural Language Processing (NLP)

- How to generalize NNs beyond CV/NLP/SR?
 - How to design universal and broadly applicable architectures?
 - A solution is Graph Neural Networks.
 - With universal architecture (unified all NNs) – New NN generation



Speech Recognition (SR)

Outline

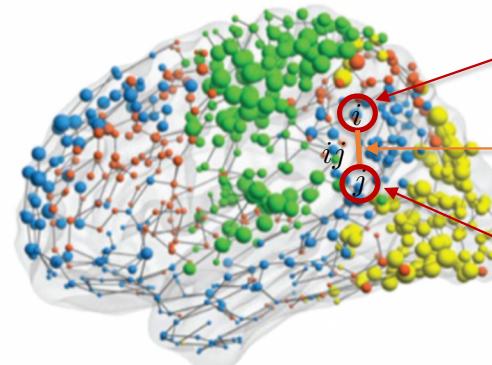
- The Deep Learning (DL) revolution
- Limitation of DL
- **Graph-Structured data**
- Graph Neural Networks (GNNs)
- GNN case studies
- GNN for industry
- GNN libraries
- Conclusion

Graph-Structured Data



Social networks
(Advertisement/
recommendation)

Brain connectivity
(sMRI)

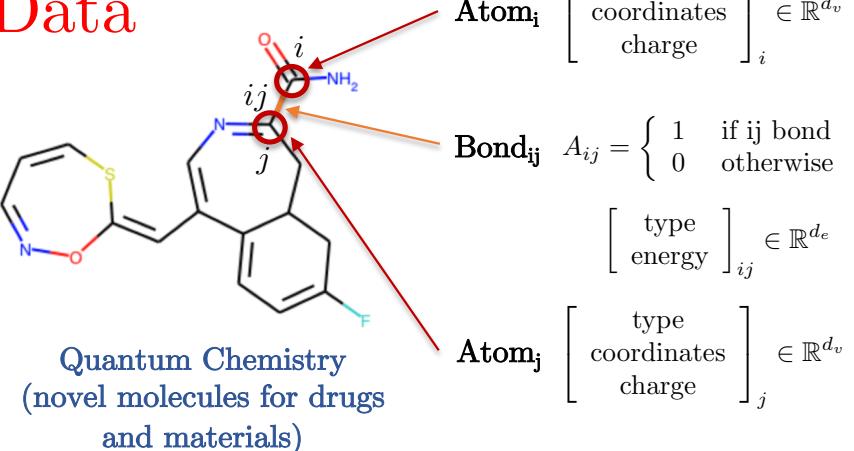


Brain analysis
(Neuroscience/neuro-diseases)

User_i $\begin{bmatrix} \text{messages} \\ \text{images} \\ \text{videos} \end{bmatrix}_i \in \mathbb{R}^d$

User connection_{ij} $A_{ij} = \begin{cases} 1 & \text{if } ij \text{ friends} \\ 0 & \text{otherwise} \end{cases}$

User_j $\begin{bmatrix} \text{messages} \\ \text{images} \\ \text{videos} \end{bmatrix}_j \in \mathbb{R}^d$



Quantum Chemistry
(novel molecules for drugs
and materials)

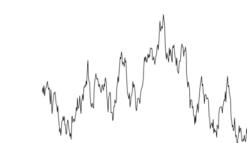
Atom_i $\begin{bmatrix} \text{type} \\ \text{coordinates} \\ \text{charge} \end{bmatrix}_i \in \mathbb{R}^{d_v}$

Bond_{ij} $A_{ij} = \begin{cases} 1 & \text{if } ij \text{ bond} \\ 0 & \text{otherwise} \end{cases}$
Atom_j $\begin{bmatrix} \text{type} \\ \text{coordinates} \\ \text{charge} \end{bmatrix}_j \in \mathbb{R}^{d_v}$

ROI_i $\begin{bmatrix} a_1 \\ \vdots \\ a_T \end{bmatrix}_i \in \mathbb{R}^T$

Cerebral
connection_{ij} $A_{ij} \in \mathbb{R}_+$

ROI_j $\begin{bmatrix} a_1 \\ \vdots \\ a_T \end{bmatrix}_j \in \mathbb{R}^T$

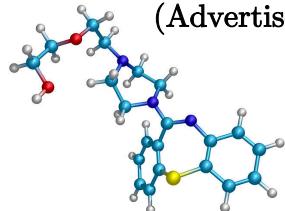


Functional
activations (fMRI)

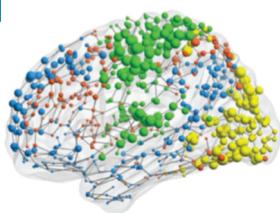
Graph-Structured Data



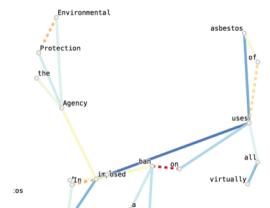
Social networks
(Advertisement)



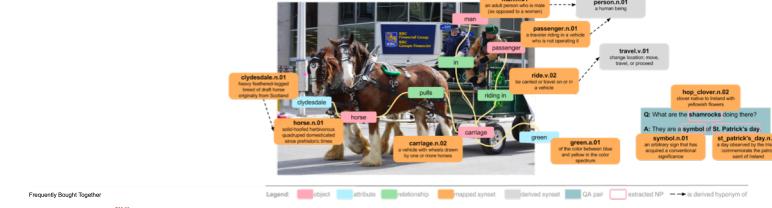
Drug/Material
molecules
(Chemistry)



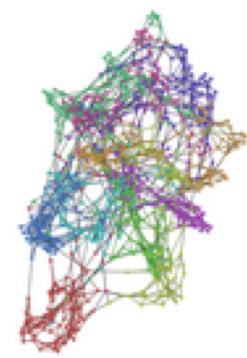
Brain
connectivity
(Neuroscience)



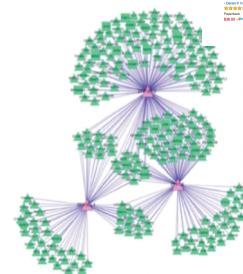
Word relationships
(NLP)



Scene understanding



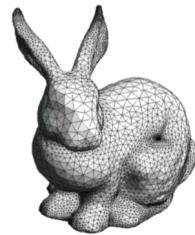
Graph



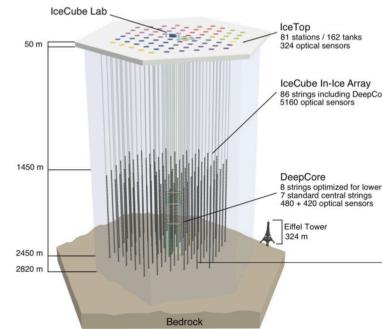
Recommender
systems (Amazon,
Netflix)



Transportation
networks



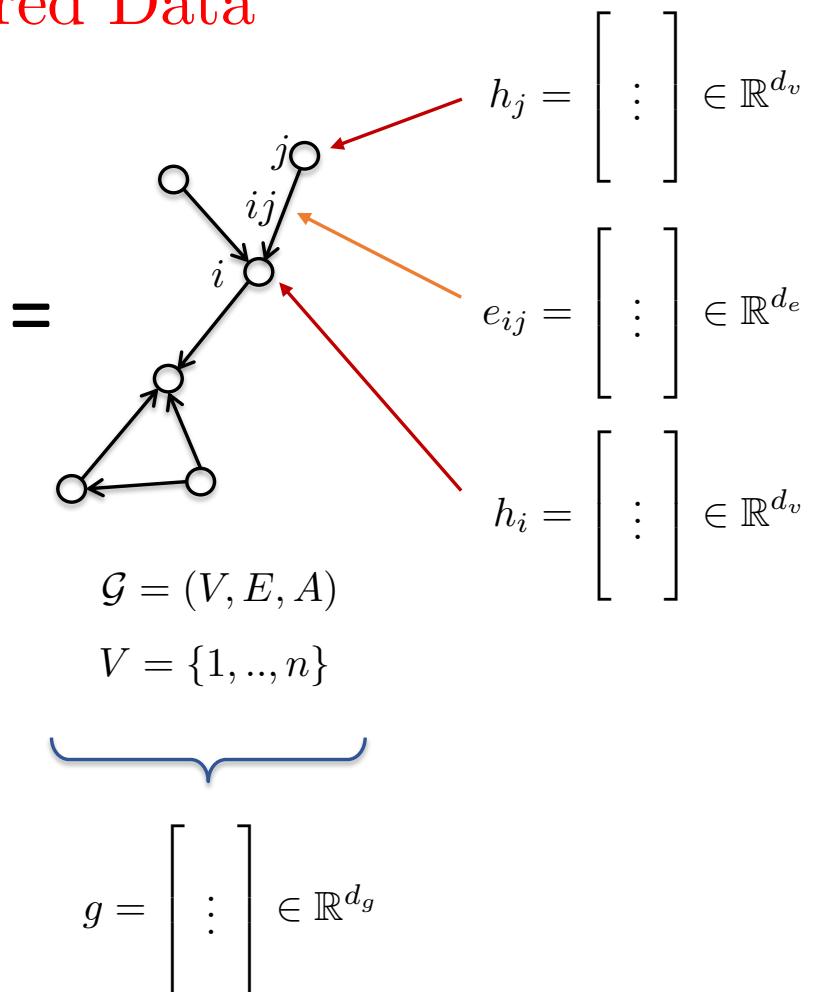
3D Meshes
(Computer Graphics)



Neutrino
detection (High-
energy Physics)

Graph-Structured Data

- A graph G is defined by :
 - Vertices V
 - Edges E
 - Adjacency matrix A
- Graph features :
 - Node features : h_i, h_j (atom type)
 - Edge features : e_{ij} (bond type)
 - Graph features : g (molecule energy)



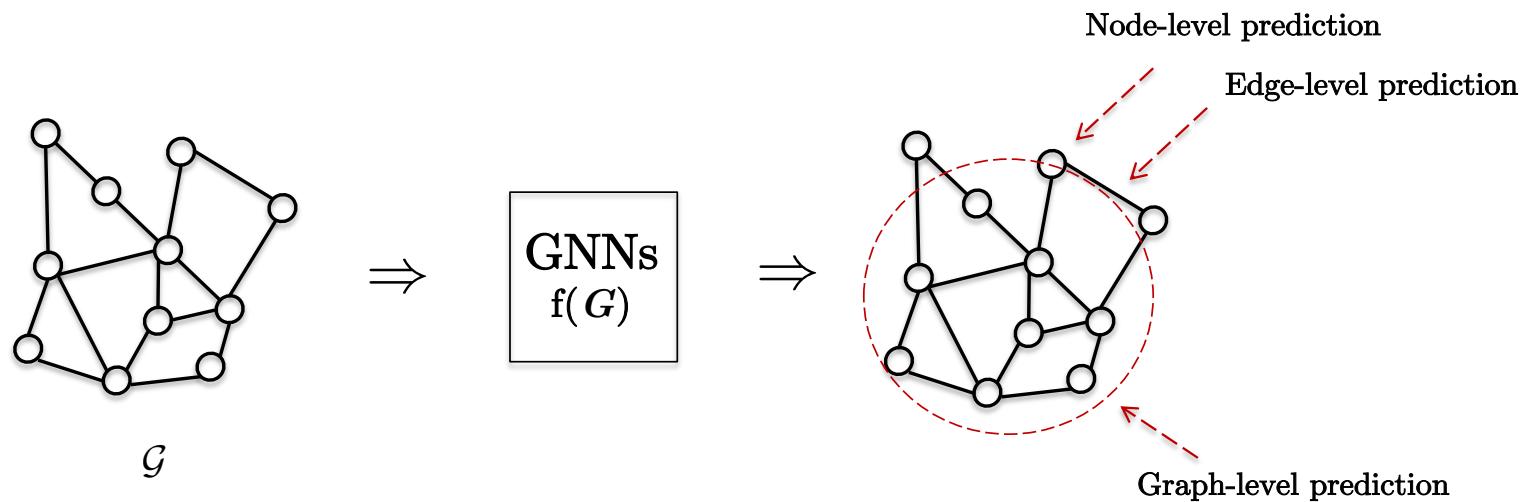
Outline

- The Deep Learning (DL) revolution
- Limitation of DL
- Graph-Structured data
- **Graph Neural Networks (GNNs)**
- GNN case studies
- GNN for industry
- GNN libraries
- Conclusion

Graph Neural Networks

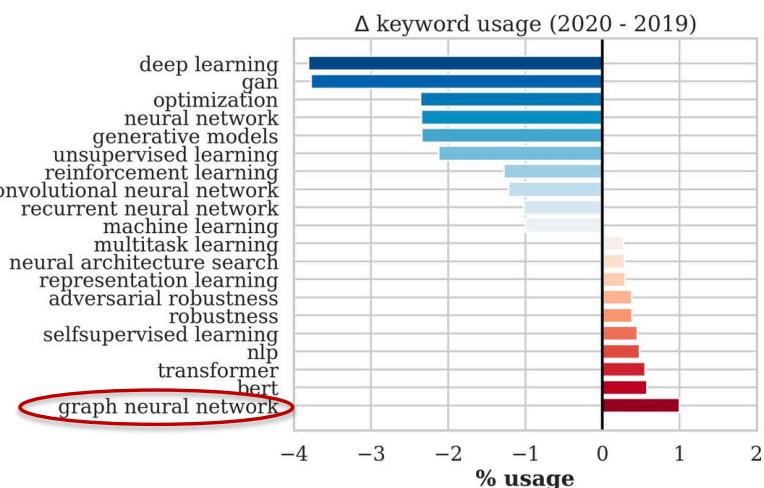
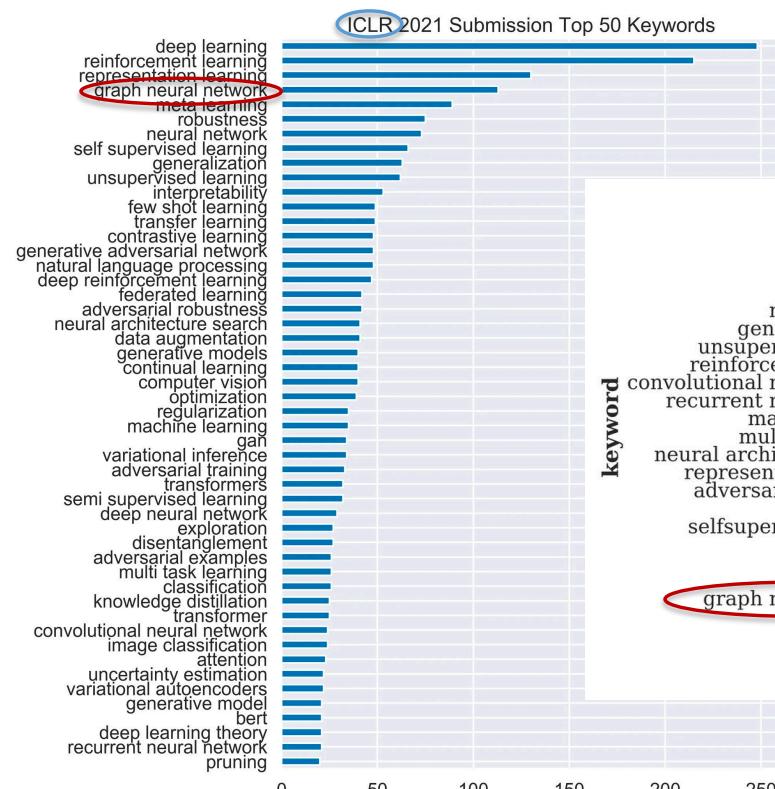
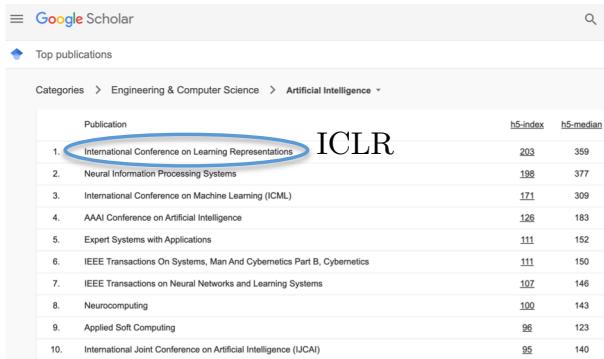
Why generalizing Deep Learning to graphs is hard ?

- Irregular grid/topology: the size of graphs changes and the number of neighbors varies
- No node ordering: no coordinates/directions on graphs like up, down, right or left
- Lots of symmetries:
 - Permutation equivariance/invariance
 - Isomorphic nodes, i.e. nodes with the same neighborhood structures
- Dynamic topology: node/edge can appear and disappear



GNNs in Research

- One of the hottest machine learning topics since 2021 !
- Standard toolkit for analyzing and learning from data on graphs
- Flexible to adapt to complex data structure and combine distinct modalities



My research: Designing powerful GNNs



- Focus on expressive GNNs for a wide range of applications
- Co-inventor of Graph Convolutional Networks in 2016 (10th most cited paper at NeurIPS)
- Popularized GNNs with Yann LeCun (NYU/Meta/Turing) w/ tutorials s.a. NeurIPS'17
- Developed with Yoshua Bengio'21 (MILA/Turing) and Yann LeCun'22 highly expressive GNNs
- Invited speaker at NeurIPS'22, AAAI'21, KDD'21, ICLR'20, ICML'20 (top AI conferences)
- Conference/workshop organizer at LoG'22, UCLA'23'21'19'18, ICLR'22, MLSys'21
- Selected as the first Amazon Scholar in Asia in 2021
- Awarded the US\$2M NRF Fellowship in 2017 (largest individual grant in Singapore)
- 16,000+ citations, 60+ articles, 70,000+ YouTube views, 8,000+ Twitter followers



NeurIPS'17
1,000-2,000
participants



CVPR'17
500-1,000
participants

Outline

- The Deep Learning (DL) revolution
- Limitation of DL
- Graph-Structured data
- Graph Neural Networks (GNNs)
- **GNN case studies**
- GNN for industry
- GNN libraries
- Conclusion

GNN case studies

- Chip design (Google)
- Resource management
- Scene reasoning (Meta)
- Recommendation (UberEats/ Pinterest/ Alibaba)
- Fake news detection
- Finance
- Natural Language Processing
- Knowledge graphs (Amazon)
- Transportation (Google Map ETA)
- Autonomous driving (NVIDIA)
- Protein & drug design (Google DeepMind/ Microsoft/ AstraZeneca)
- Energy physics & simulations (Google DeepMind)
- Code bug detection
- Genomics

GNNs for chip design



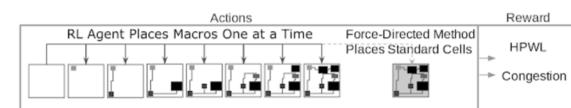
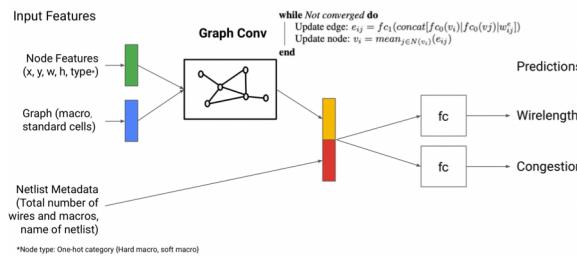
The latest news from Google AI

<https://ai.googleblog.com/2020/04/chip-design-with-deep-reinforcement.html>

Chip Design with Deep Reinforcement Learning

Thursday, April 23, 2020

Posted by Anna Goldie, Senior Software Engineer and Azalia Mirhoseini, Senior Research Scientist, Google Research, Brain Team



During each training iteration, the macros are placed by the policy one at a time and the standard cell clusters are placed by a force-directed method. The reward is calculated from the weighted combination of approximate wirelength and congestion.



Google Invents AI That Learns a Key Part of Chip Design

AI helps designs AI chip that might help an AI design future AI chips

By Samuel K. Moore



Google is using AI to design chips that will accelerate AI



Google is using AI to design AI processors much faster than humans can

By Paul Lilly 10 days ago
Chips making chips...
Comments



Google Proposes AI as Solution for Speedier AI Chip Design

Google uses artificial intelligence to optimize AI chip production

By Maria Konnikova - April 2, 2020



Google Hoping The Next AI Chips Will Be Designed By AI

By Maria Konnikova - April 2, 2020



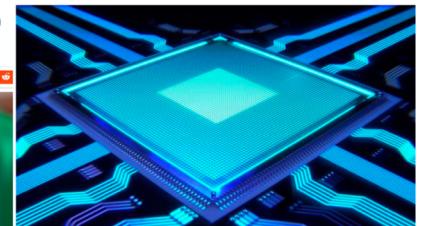
Google Researchers Create AI-ception with an AI Chip That Speeds Up AI

Using a reinforcement-learning algorithm, the AI has learned to optimize the placement of components on a computer chip.



Google trains chips to design themselves

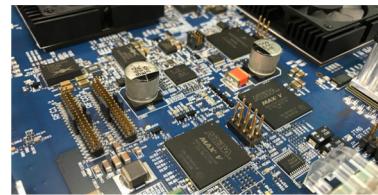
by Peter Gradi , Tech Xplore



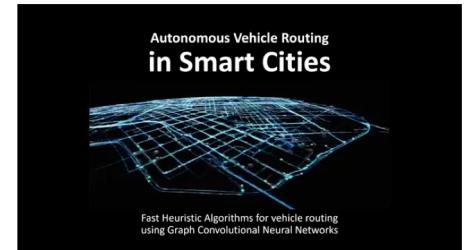
<https://www.theverge.com/2021/6/10/22527476/google-machine-learning-chip-design-tpu-floorplanning>

GNNs for resource allocation

- 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.



Hardware pieces placement



Uber taxis allocation

Can Transformers Solve This 90-Year-Old Classic Computer Science Problem Better Than Human Algorithms?

Deep learning can't beat human solutions — yet

 Andre Ye · Mar 19 · 6 min read ★

...

The Travelling Salesman Problem was formulated in 1930, and is a classical computer science problem for optimization. It's a simple problem:

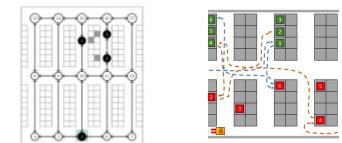
<https://towardsdatascience.com/can-transformers-solve-this-90-year-old-classic-computer-science-problem-better-than-human-8ef78bb2e9d8>



Amazon warehouse management

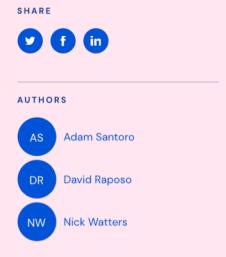


Solving the TSP for Warehouses



GNNs for scene reasoning

A neural approach to relational reasoning



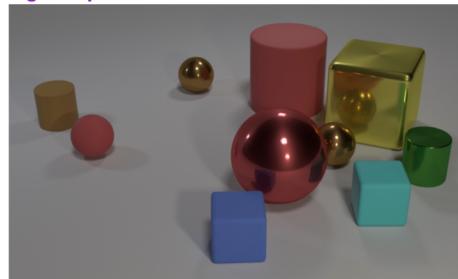
<https://deepmind.com/blog/article/neural-approach-relational-reasoning>



Consider the reader who pieces together the evidence in an Agatha Christie novel to predict the culprit of the crime, a child who runs ahead of her ball to prevent it rolling into a stream or even a shopper who compares the relative merits of buying kiwis or mangos at the market.

We carve our world into relations between things. And we understand how the world works through our capacity to draw logical conclusions about how these different things – such as physical objects, sentences, or even abstract ideas – are related to one another. This ability is called relational reasoning and is central to human intelligence.

Questions in CLEVR test various aspects of visual reasoning including **attribute identification**, **counting**, **comparison**, **spatial relationships**, and **logical operations**.



- Q: Are there an **equal number** of **large things** and **metal spheres**?
Q: **What size** is the **cylinder** that is **left of** the **brown metal** thing that is **left of** the **big sphere**?
Q: There is a **sphere** with the **same size as** the **metal cube**; is it **made of the same material as** the **small red sphere**?
Q: **How many** objects are **either small cylinders or red things**?

Graph R-CNN for Scene Graph Generation

Jianwei Yang^{1*}, Jiasen Lu^{1*}, Stefan Lee¹, Dhruv Batra^{1,2}, and Devi Parikh^{1,2}

¹Georgia Institute of Technology ²Facebook AI Research
{jw2yang, jiasenlu, steflee, dbatra, parikh}@gatech.edu

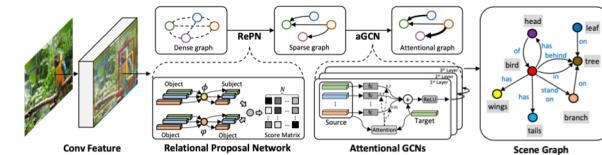
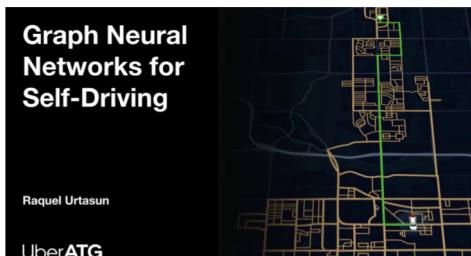
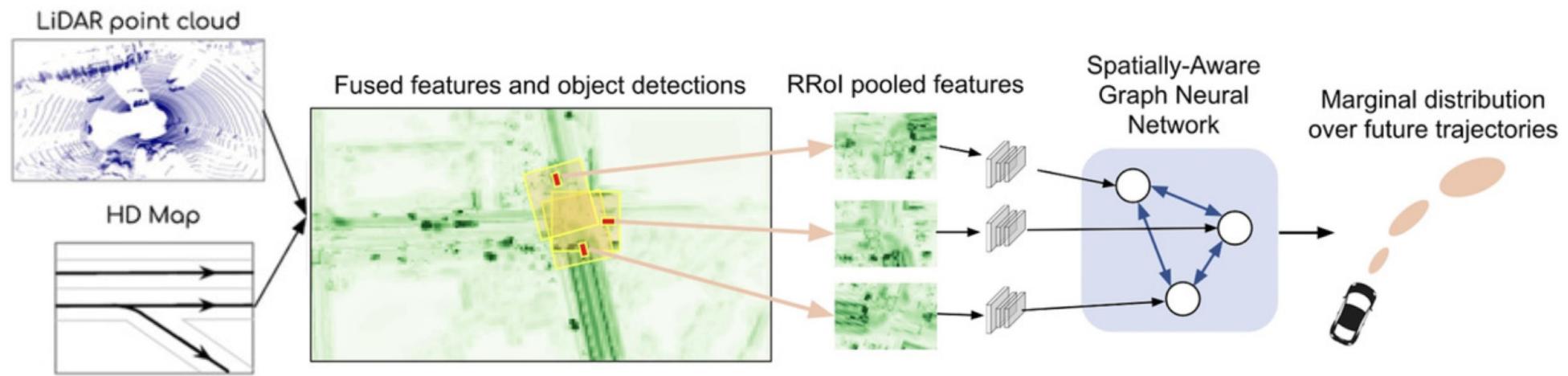
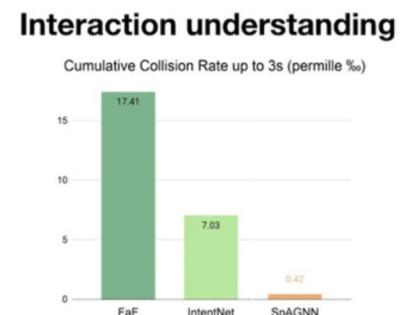
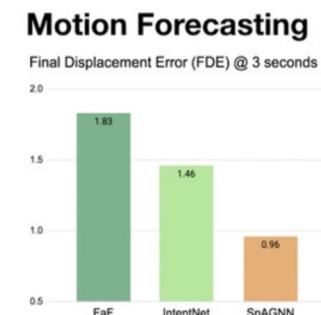
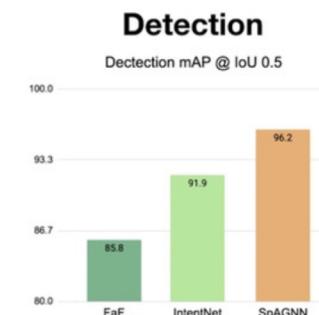


Fig. 2. The pipeline of our proposed Graph R-CNN framework. Given an image, our model first uses RPN to propose object regions, and then prunes the connections between object regions through our relation proposal network (RePN). Attentional GCN is then applied to integrate contextual information from neighboring nodes in the graph. Finally, the scene graph is obtained on the right side.

GNNs for autonomous driving

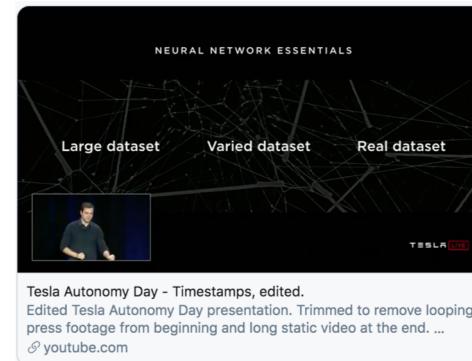
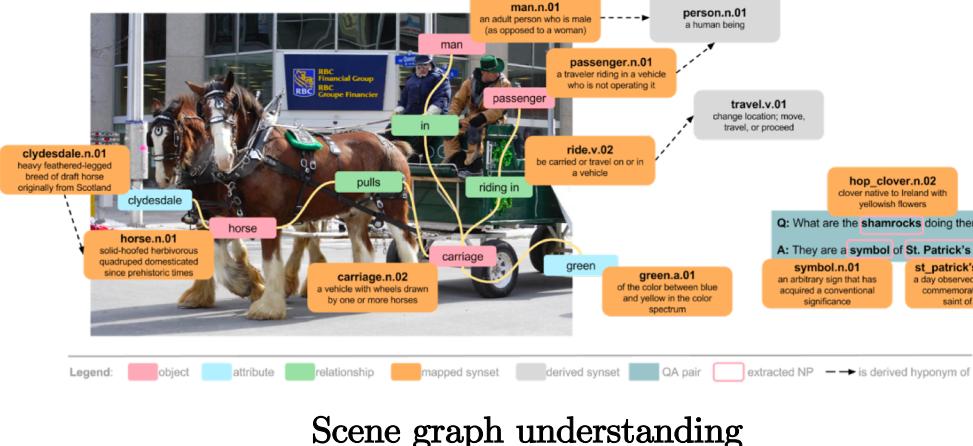


<https://slideslive.com/38930570/graph-neural-networks-for-selfdriving>



GNNs for autonomous driving

- Computer Vision with visual common sense
 - Reasoning on scene graph



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



GNNs for recommendation

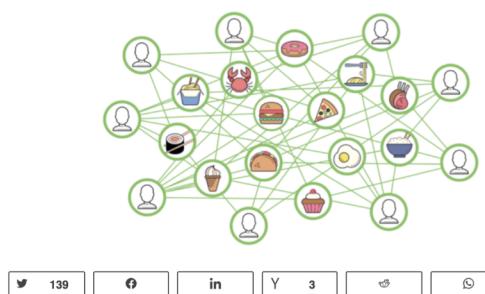


AI General Engineering

Food Discovery with Uber Eats: Using Graph Learning to Power Recommendations

Ankit Jain, Isaac Liu, Ankur Sarda, and Piero Molino

December 4, 2019



The Uber Eats app serves as a portal to more than 320,000 restaurant-partners in over 500 cities globally across 36 countries. In order to make the user experience more seamless and easy-to-navigate, we show users the dishes, restaurants, and cuisines they might like up front. To this end, we previously developed ML models to [better understand queries](#) and for [multi-objective optimization](#) in Uber Eats search and recommender system in Uber Eats searches and surfaced food options.

<https://eng.uber.com/uber-eats-graph-learning>



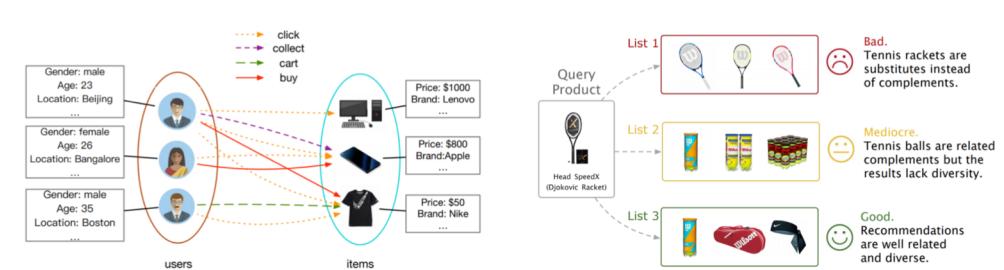
arXiv.org > cs > arXiv:1902.08730

Computer Science > Distributed, Parallel, and Cluster Computing

[Submitted on 23 Feb 2019]

AliGraph: A Comprehensive Graph Neural Network Platform

Rong Zhu, Kun Zhao, Hongxia Yang, Wei Lin, Chang Zhou, Baole Ai, Yong Li, Jingren Zhou



Pinterest Engineering Blog MACHINE LEARNING INFRASTRUCTURE OPEN SOURCE DATA SCIENCE MOBILE

PinSage: A new graph convolutional neural network for web-scale recommender systems

Pinterest Engineering Follow
Aug 16, 2018 - 8 min read

Twitter LinkedIn Facebook More



Figure 3: Examples of pins recommended by different algorithms. The image to the left is the query pin. Recommended items to the right are computed using Visual embeddings, Annotation embeddings, Pixie (purely graph-based method), and PinSage.

<https://medium.com/pinterest-engineering/pinsage-a-new-graph-convolutional-neural-network-for-web-scale-recommender-systems-88795a107f48>

GNNs for fake news detection

- Social networks :
 - Fake News (2016 US Presidential Election) / (human) adversarial attacks

Fake News Detection on Social Media using Geometric Deep Learning

Federico Monti^{1,2} Fabrizio Frasca^{1,2} Davide Eynard^{1,2} Damon Mannion^{1,2}

Michael M. Bronstein^{1,2,3}

¹Fabula AI
United Kingdom

²USI Lugano
Switzerland

³Imperial College
United Kingdom

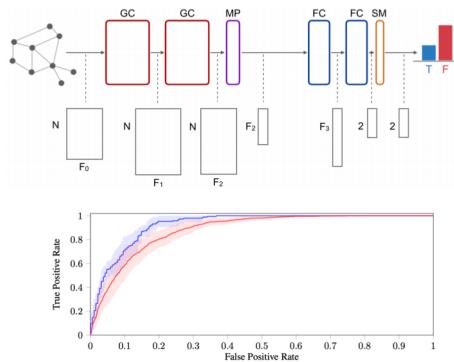
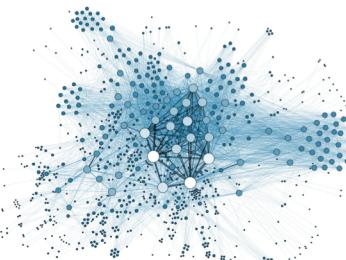


Figure 6: Performance of URL-wise (blue) and cascade-wise (red) fake news detection using 24hr-long diffusion time. Shown are ROC curves averaged on five folds (the shaded areas represent the standard deviations). ROC AUC is $92.70 \pm 1.80\%$ for URL-wise classification and $88.30 \pm 2.74\%$ for cascade-wise classification, respectively. Only cascades with at least 6 tweets were considered for cascade-wise classification.



Twitter network



Facebook network



A screenshot of a TechCrunch article titled "Twitter acquires Fabula AI, a machine learning startup that helps spot fake news". The article is by Paul Sawers and was published on June 3, 2019, at 4:42 AM. The image shows a colorful, abstract geometric pattern with the word "FABULA" integrated into it.

TC
Join Extra Crunch
Login

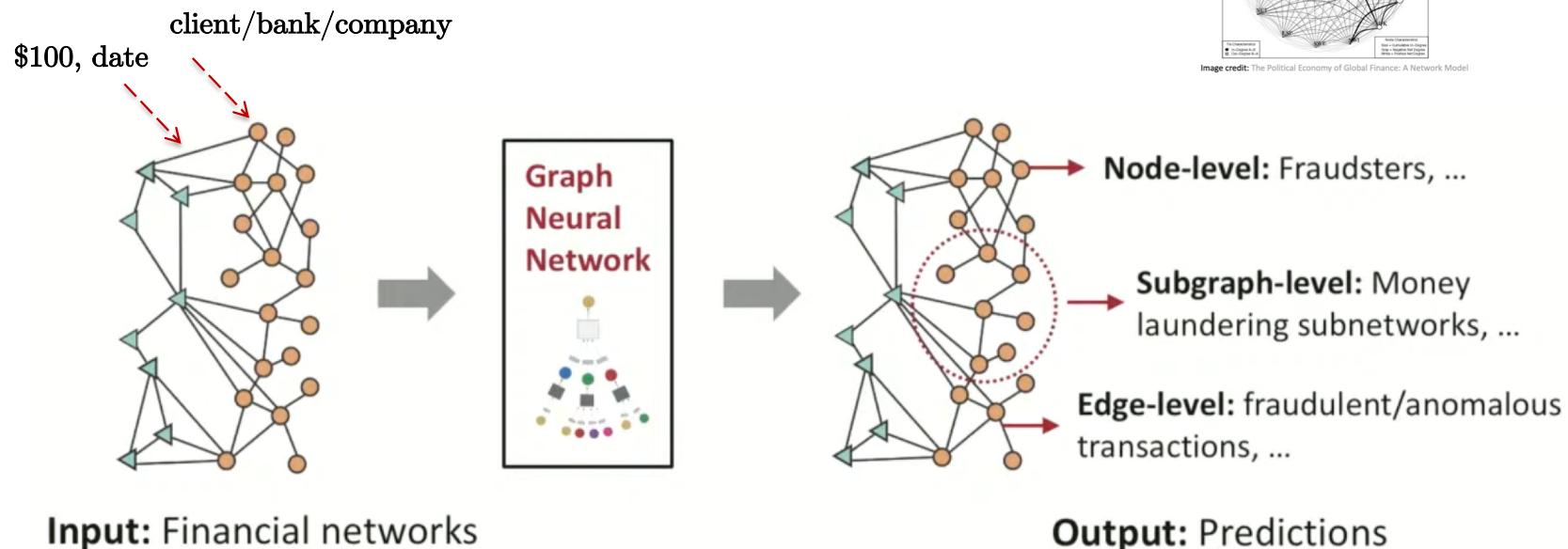
Search Q. TC Early Stage 2021
Startups
Videos
Audio
Newsletters
Extra Crunch
EC-1s
The TC List
Advertise
Events
More

A screenshot of a TechCrunch article titled "Fabula AI is using social spread to spot 'fake news'". The article is by Natasha Lomas and was published on February 7, 2019, at 7:07 AM GMT+0. The image shows a close-up of a mechanical device, possibly a printer or scanner, with the words "Fake news" written on its surface.

UK startup Fabula AI reckons it's devised a way for artificial intelligence to help user generated content platforms get on top of the disinformation crisis that keeps rocking the world of social media with antisocial scandals.

GNNs for finance

- Financial networks model financial entities and their relationships.
- Applications :



International banking
• Nodes: Countries
• Edges: Capital flows

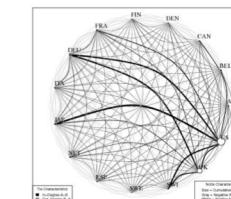


Image credit: The Political Economy of Global Finance: A Network Model

Bitcoin transactions
• Nodes: BTC wallets
• Edges: Transactions



Image credit: <http://dashblockchain.github.io/>

GNNs for knowledge graphs



INFORMATION AND KNOWLEDGE MANAGEMENT

Combining knowledge graphs, quickly and accurately

Novel cross-graph-attention and self-attention mechanisms enable state-of-the-art performance.

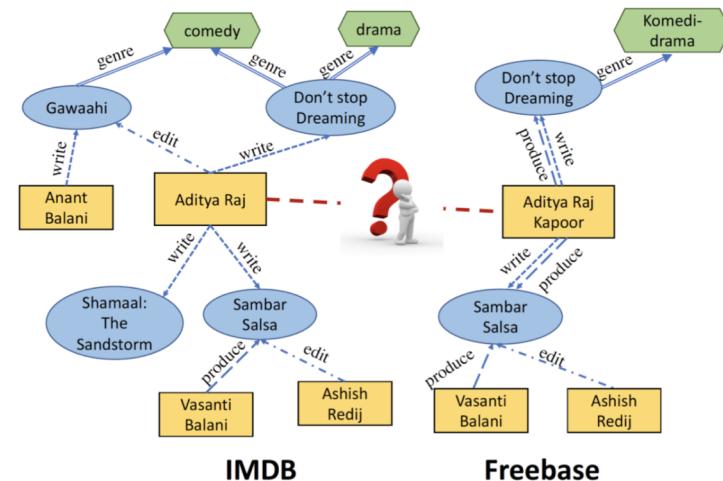
By Hao Wei
March 19, 2020



Knowledge graphs are a way of representing information that can capture complex relationships more easily than conventional databases. At Amazon, we use knowledge graphs to represent the hierarchical relationships between product types on amazon.com; the relationships between creators and content on Amazon Music and Prime Video; and general information for Alexa's question-answering service — among other things.

<https://www.amazon.science/blog/combining-knowledge-graphs-quickly-and-accurately>

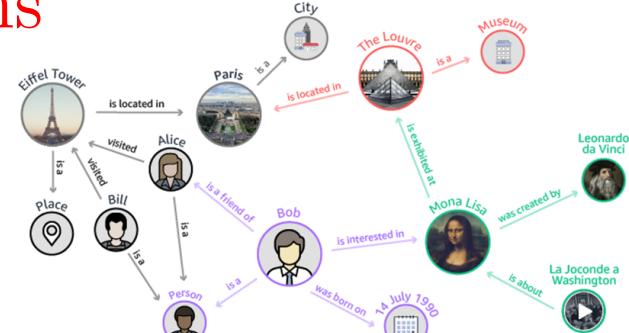
- Knowledge graphs represent large-scale data information in the form of triples.
- Triples are two entities and their type of relationship.



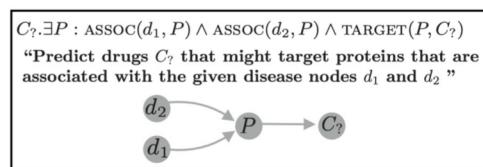
This example illustrates the challenge of entity alignment. IMDB lists the writer of the movie *Don't Stop Dreaming* as Aditya Raj, but the (now defunct) Freebase database lists him as Aditya Raj Kapoor. Are they the same person?

GNNs for knowledge graphs

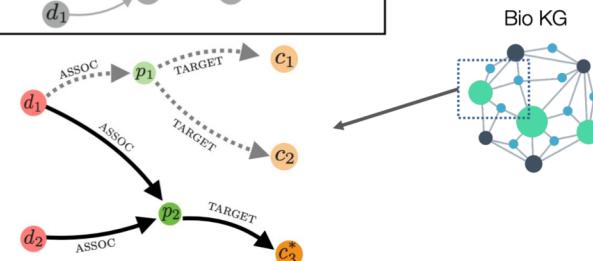
- Reasoning on KGs can provide critical new information.
 - Predict drugs that target proteins associated to Covid-19.
 - Where did Canadian Turing Award winners graduate?
- Limitations of non-ML reasoning on KG :
 - KGs are incomplete, noisy.
 - Answer query on large graphs can become intractable with standard graph search algorithms.
- GNNs can represent complex queries in a continuous embedding space.
 - Answer is given by the closest entity (w.r.t. the Euclidean distance).



Credit: Amazon Neptune ML

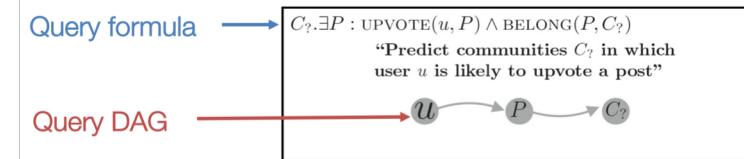


First-order logic Query Formula

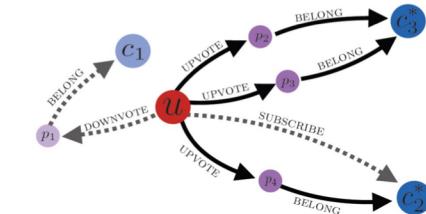


Xavier Bresson

Hamilton et al, NeurIPS 2018



Example subgraphs
that satisfy the query



GNNs for NLP

Attention Is All You Need

Ashish Vaswani*
Google Brain
avaswani@google.com

Noam Shazeer*
Google Brain
noam@google.com

Niki Parmar*
Google Research
nikip@google.com

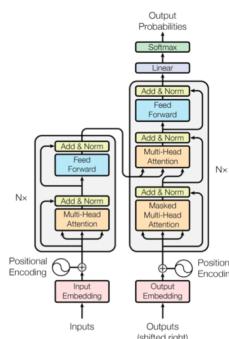
Jakob Uszkoreit*
Google Research
usz@google.com

Llion Jones*
Google Research
llion@google.com

Aidan N. Gomez* †
University of Toronto
aidan@cs.toronto.edu

Łukasz Kaiser*
Google Brain
lukasz.kaiser@google.com

Illia Polosukhin* ‡
illia.polosukhin@gmail.com



Transformers are
Graph Neural
Networks
12.SEP.2020

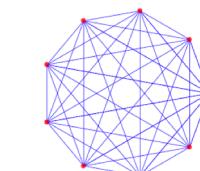
My engineering friends often ask me: deep learning on graphs sounds great, but are there any real applications?

While Graph Neural Networks are used in recommendation systems at Pinterest, Alibaba and Twitter, a more subtle success story is the [Transformer architecture](#), which has taken the NLP world by storm. Through this post, I want to establish a link between [Graph Neural Networks \(GNNs\)](#) and Transformers. I'll talk about the intuitions behind model architectures in the NLP and GNN communities, make connections using equations and figures, and discuss how we can work together to drive future progress. Let's start by talking about the purpose

<https://thegradient.pub/transfomers-are-graph-neural-networks>

Xavier Bresson

Transformers are (fully-connected) Graph Neural Networks.



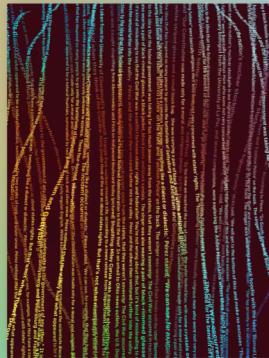
OpenAI

API PROJECTS BLOG ABOUT

Better Language Models and Their Implications

We've trained a large-scale unsupervised language model which generates coherent paragraphs of text, achieves state-of-the-art performance on many language modeling benchmarks, and performs rudimentary reading comprehension, machine translation, question answering, and summarization—all without task-specific training.

February 14, 2019
24 minute read



<https://openai.com/blog/better-language-models>

MIT Technology Review

Artificial intelligence / Machine learning

OpenAI's new language generator GPT-3 is shockingly good—and completely mindless

The AI is the largest language model ever created and can generate amazing human-like text on demand but won't bring us closer to true intelligence.

by Will Douglas Heaven

July 20, 2020

Forbes

Oct 5, 2020, 12:21am EDT | 49,667 views

What Is GPT-3 And Why Is It Revolutionizing Artificial Intelligence?

Bernard Marr Contributor @ Enterprise Tech

There's been a great deal of hype and excitement in the [artificial intelligence \(AI\)](#) world around a newly developed technology known as GPT-3. Put simply, it's an AI that is better at creating content that has a language structure – human or machine language – than anything that has come before it.



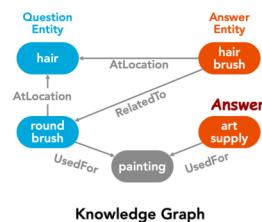
Limitation of Transformers

- Impressive results at large-scale
 - 2020 GPT-3, 175B parameters, 285,000 CPUs, 10,000 GPUs, 400Gb/sec network connectivity, 500 billion tokens, US\$12 Million to train
 - Transformers capture dynamic word representation depending on the context.
 - Example : The vase broke. The news broke. Sandy broke the world record. Sandy broke the law. We broke even. The burglar broke into the house. Etc.
 - DL has not yet reached human performance (no common sense).
- What is missing to get to human's level?
 - More data? Yes but not sufficient.
 - Reasoning, but with what inductive bias?
 - Knowledge graphs



xkcd: Machine Learing

Textual Context
If it is **not** used for hair, a round brush is an example of what?
A. hair brush B. bathroom C. art supplies* D. shower



Credit: Antoine Bosselut, EPFL

Xavier Bresson

DETECT LANGUAGE	ENGLISH	FRENCH	ENGLISH	CHINESE (SIM)
J'ai mangé mon avocat pour le déjeuner à la cafétéria du palais de justice.	<input checked="" type="checkbox"/>		I ate my lawyer for lunch in the courthouse cafeteria.	<input type="checkbox"/>

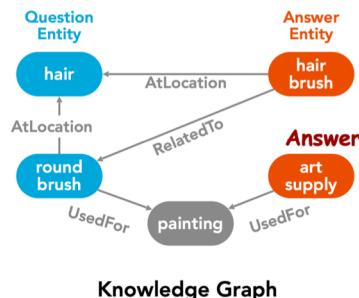
35

GNNs for NLP

- Transformers vs GNNs with knowledge graphs

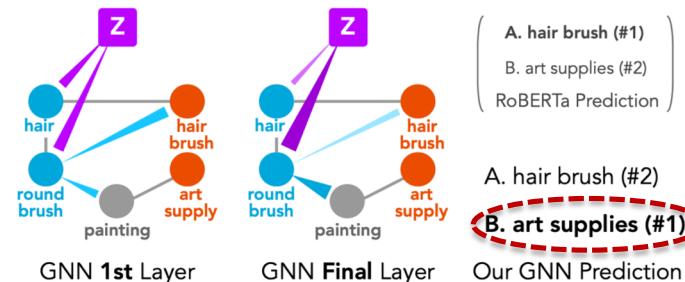
Textual Context

If it is **not** used for **hair**, a **round brush** is an example of what?
 A. hair brush B. bathroom C. art supplies* D. shower



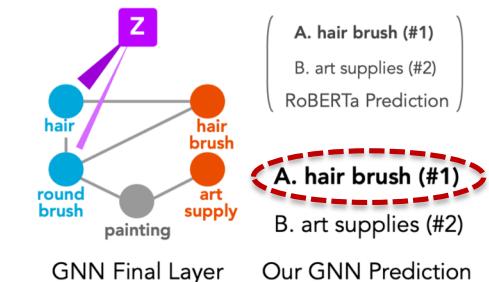
Original Question

If it is **not** used for **hair**, a **round brush** is an example of what?
 A. hair brush B. art supplies*



Negation Removed

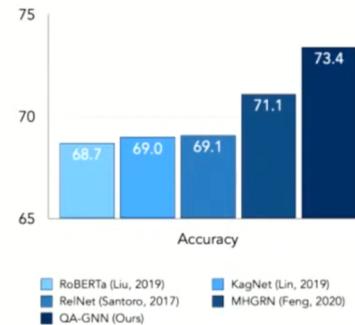
If it is used for **hair**, a **round brush** is an example of what? A. hair brush B. art supplies



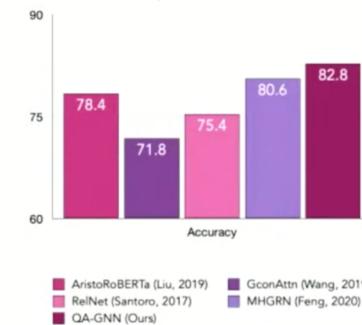
Yasunaga et al, NAACL, 2021

Xavier Bresson

CommonsenseQA



OpenbookQA



GNNs for transportation

BLOG POST
RESEARCH

03 SEP 2020

Traffic prediction with advanced Graph Neural Networks

SHARE

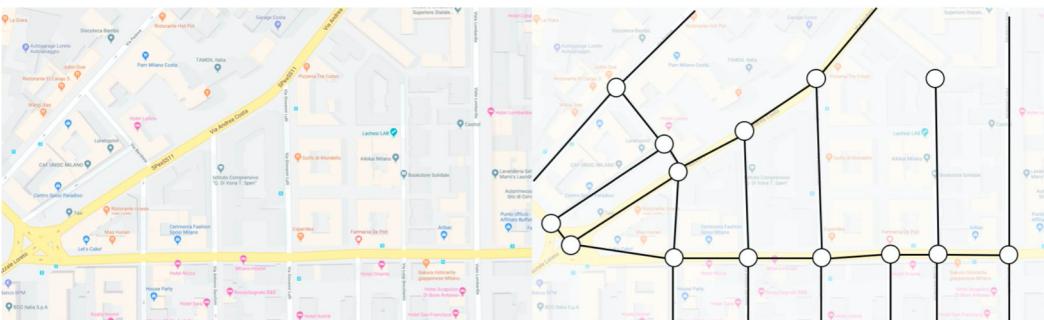
By partnering with Google, DeepMind is able to bring the benefits of AI to billions of people all over the world. From reuniting a speech-impaired user with his [original voice](#), to helping users discover [personalised apps](#), we can apply breakthrough research to immediate real-world problems at a Google scale. Today we're delighted to share the results of our latest partnership, delivering a truly global impact for the more than one billion people that use Google Maps.



Authors

OL Oliver Lange*

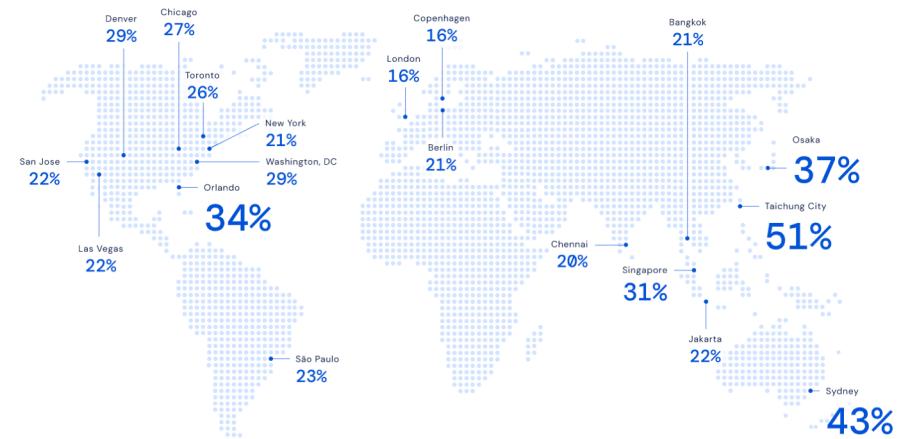
LP Luis Perez



<https://deepmind.com/blog/article/traffic-prediction-with-advanced-graph-neural-networks>

Xavier Bresson

Google Maps ETA Improvements Around the World



GNNs for protein folding

AlphaFold: a solution to a 50-year-old grand challenge in biology

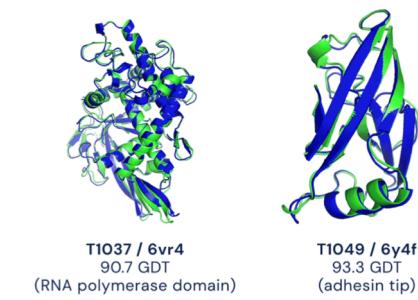
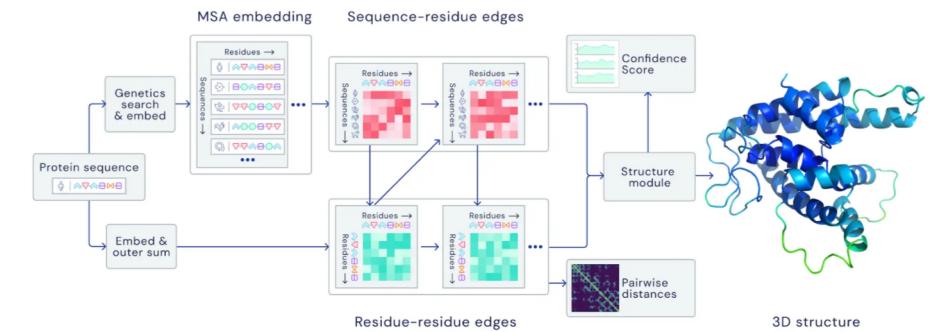
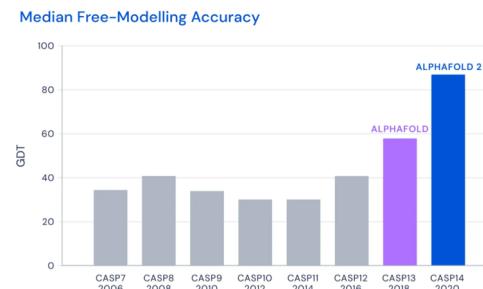
SHARE
[Twitter](#) [Facebook](#) [LinkedIn](#)

AUTHORS
 The AlphaFold team



Proteins are essential to life, supporting practically all its functions. They are large complex molecules, made up of chains of amino acids, and what a protein does largely depends on its unique 3D structure. Figuring out what shapes proteins fold into is known as the “protein folding problem”, and has stood as a grand challenge in biology for the past 50 years. In a major scientific advance, the latest version of our AI system AlphaFold has been recognised as a solution to this grand challenge by the organisers of the biennial Critical Assessment of protein Structure Prediction (CASP). This breakthrough demonstrates the impact AI can have on scientific discovery and its potential to dramatically accelerate progress in some of the most fundamental fields that explain and shape our world.

<https://deepmind.com/blog/article/alphafold-a-solution-to-a-50-year-old-grand-challenge-in-biology>



GNNs for protein function & interaction

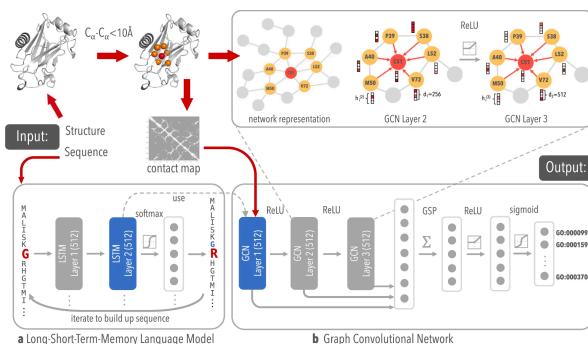
[nature](#) > [nature communications](#) > [articles](#) > [article](#)

Article | Open Access | Published: 26 May 2021

Structure-based protein function prediction using graph convolutional networks

Vladimir Gligorjević , P. Douglas Renfrew, Tomasz Kosciolek, Julia Koehler Leman, Daniel Berenberg, Tommi Vatanen, Chris Chandler, Bryn C. Taylor, Ian M. Fisk, Hera Vlamakis, Ramnik J. Xavier, Rob Knight, Kyunghyun Cho & Richard Bonneau 

[Nature Communications](#) 12, Article number: 3168 (2021) | [cite this article](#)



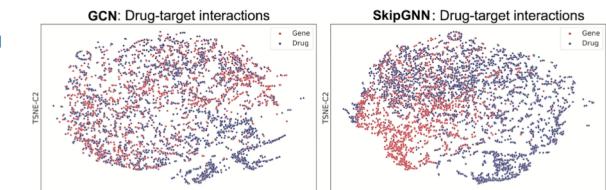
[nature](#) > [scientific reports](#) > [articles](#) > [article](#)

Article | Open Access | Published: 03 December 2020

SkipGNN: predicting molecular interactions with skip-graph networks

Kexin Huang, Cao Xiao, Lucas M. Glass, Marinka Zitnik & Jimeng Sun 

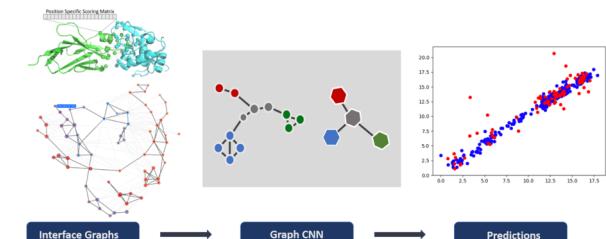
[Scientific Reports](#) 10, Article number: 21092 (2020) | [cite this article](#)



DeepRank-GNN: A Graph Neural Network Framework to Learn Patterns in Protein-Protein Interfaces

M. Réau^{1,2}, N. Renaud^{2,3}, L. C. Xue³, A. M. J. J. Bonvin^{1,*}

¹Computational Structural Biology Group, Department of Chemistry, Bijvoet Centre, Faculty of Science, Utrecht University, Utrecht, 3584CH, The Netherlands., ²Netherlands eScience Center, Science Park 140, 1098 XG, Amsterdam, The Netherlands, ³Department of Biochemistry and Molecular Biology, University of Illinois Urbana-Champaign, Urbana, IL 61801, USA



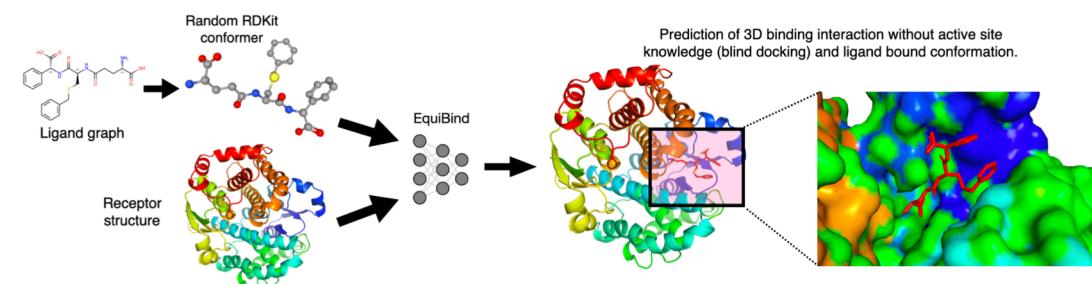
 arXiv > q-bio > arXiv:2202.05146
Quantitative Biology > Biomolecules

(Submitted on 7 Feb 2022 (v1), last revised 4 Jun 2022 (this version, v4))

EquiBind: Geometric Deep Learning for Drug Binding Structure Prediction

Hannes Stark, Octavian-Eugen Ganea, Lagajit Pattaikul, Regina Barzilay, Tommi Jaakkola

Predicting how a drug-like molecule binds to a specific protein target is a core problem in drug discovery. An extremely fast computational binding method would enable key applications such as fast virtual screening or drug engineering. Existing methods are computationally expensive as they rely on heavy candidate sampling coupled with scoring, ranking, and fine-tuning steps. We challenge this paradigm with EquiBind, an SE(3)-equivariant geometric deep learning model performing direct-shot prediction of both i) the receptor binding location (blind docking) and ii) the ligand's bound pose and orientation. EquiBind achieves significant speed-ups and better quality compared to traditional and recent baselines. Further, we show extra improvements when coupling it with existing fine-tuning techniques at the cost of increased running time. Finally, we propose a novel and fast fine-tuning model that adjusts torsion angles of a ligand's rotatable bonds based on closed-form global minima of the von Mises angular distance to a given input atomic point cloud, avoiding previous expensive differential evolution strategies for energy minimization.



GNNs for drug design

Cell

Volume 180, Issue 4, 20 February 2020, Pages 688-702.e13



Article

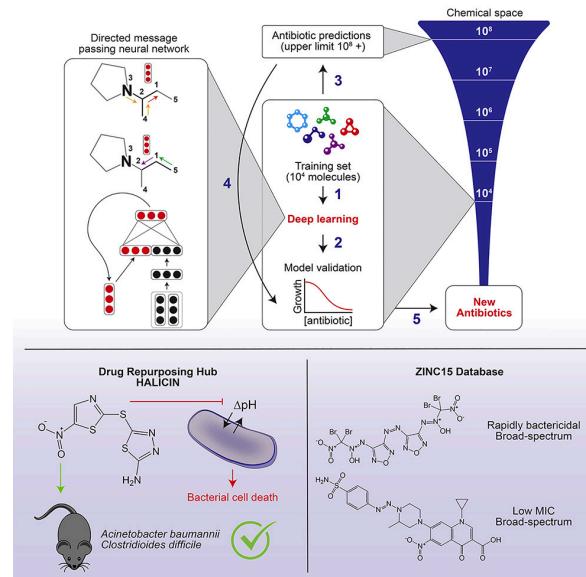
A Deep Learning Approach to Antibiotic Discovery

Jonathan M. Stokes^{1, 2, 3}, Kevin Yang^{3, 4, 10}, Kyle Swanson^{3, 4, 10}, Wengong Jin^{3, 4}, Andres Cubillos-Ruiz^{1, 2, 5}, Nina M. Donghia^{1, 5}, Craig R. MacNair⁶, Shawn French⁶, Lindsey A. Carfrae⁶, Zohar Bloom-Ackermann^{2, 7}, Victoria M. Tran², Anush Chiappino-Pepe^{5, 7}, Ahmed H. Badran², Ian W. Andrews^{1, 2, 5}, Emma J. Chory^{1, 2}, George M. Church^{5, 7, 8}, Eric D. Brown⁶, Tommi S. Jaakkola^{3, 4} ... James J. Collins^{1, 2, 5, 8, 9, 11}

Highlights

- A deep learning model is trained to predict antibiotics based on structure
- Halicin is predicted as an antibacterial molecule from the Drug Repurposing Hub
- Halicin shows broad-spectrum antibiotic activities in mice
- More antibiotics with distinct structures are predicted from the ZINC15 database

<https://www.sciencedirect.com/science/article/pii/S0092867420301021>



nature

Explore content ▾ Journal information ▾ Publish with us ▾ Subscribe Sign up

nature > news > article

NEWS · 20 FEBRUARY 2020

Powerful antibiotics discovered using AI

Machine learning spots molecules that work even against 'untreatable' strains of bacteria.

FINANCIAL TIMES

Artificial intelligence

AI discovers antibiotics to treat drug-resistant diseases

Machine learning uncovers potent new drug able to kill 35 powerful bacteria

Quanta magazine

Physics Mathematics Biology Computer Science All Articles

ARTIFICIAL INTELLIGENCE

Machine Learning Takes On Antibiotic Resistance

To combat resistant bacteria and refill the trickling antibiotic pipeline, scientists are getting help from deep learning networks.



MIT
Technology
Review

77 Mass Ave

AI vs. bacteria

GNNs for energy physics

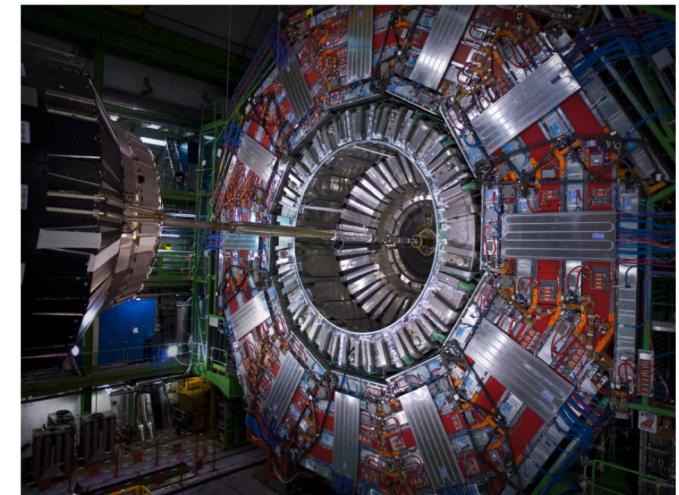
The screenshot shows the Fermilab Newsroom website. The header includes the Fermilab logo, navigation links for Home, About, Science, Jobs, Contact, Phone Book, and a search bar. Below the header, there are links for Newsroom, DUNE at LBNF, Come visit us, and Resources for. A large blue "News" button is prominently displayed. On the left, a sidebar menu lists Newsroom, News and features, Press releases, and Fermilab in the news. The main content area features a news article titled "The next big thing: the use of graph neural networks to discover particles" by Zack Savitsky on September 24, 2020. The article includes social sharing buttons for Facebook, Twitter, and Email.

"What was a week ago just an object of research is now a widely usable tool that could transform our ability to analyze data from particle physics experiments." – Lindsey Gray

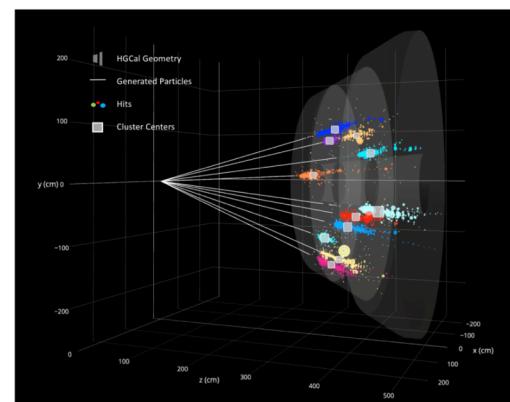
"You can still apply all of the same things we're learning about graph neural networks in the HGCAL to other detectors in other experiments," Gray said. "The rate at which we're adopting machine learning in high-energy physics is not even close to saturated yet. People will keep finding more and more ways to apply it."

<https://news.fnal.gov/2020/09/the-next-big-thing-the-use-of-graph-neural-networks-to-discover-particles>

Xavier Bresson



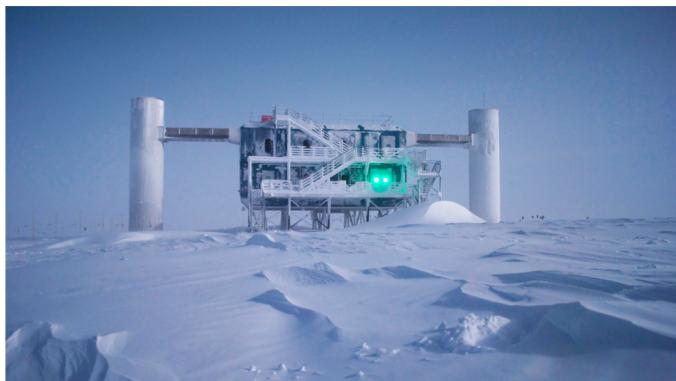
The CMS detector at the Large Hadron Collider takes billions of images of high-energy collisions every second to search for evidence of new particles. Graph neural networks expeditiously decide which of these data to keep for further analysis. Photo: CERN



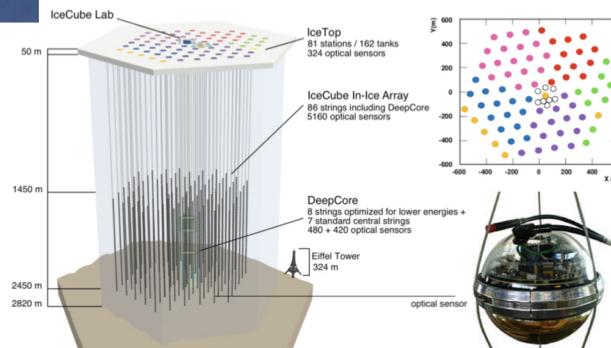
The upgraded high-granularity calorimeter — a component of the CMS detector at the Large Hadron Collider — produces complicated images of particles generated from collisions. Researchers are working to implement graph neural networks to optimize the analysis of this data to better identify and characterize particle interactions of interest. Image courtesy of Ziheng Chen, Northwestern University

GNNs for energy physics

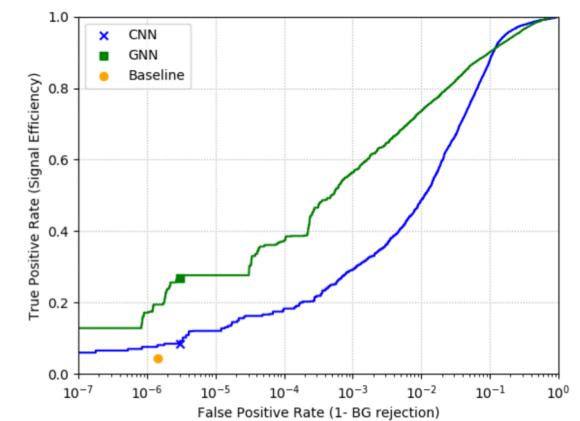
- High-energy physics with neutrino detection (hard to detect because they have a very small chance of interacting with regular matter).



IceCube Neutrino Observatory



Xavier Bresson



GNNs for physics simulation

DeepMind > Research > Learning to Simulate Complex Physics with Graph Networks

OPENSOURCE

14 SEP 2020

SHARE

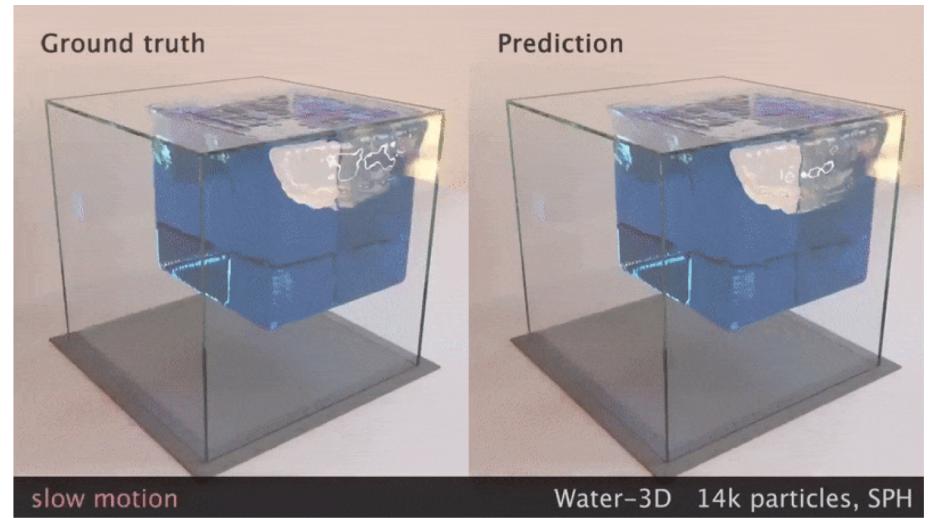
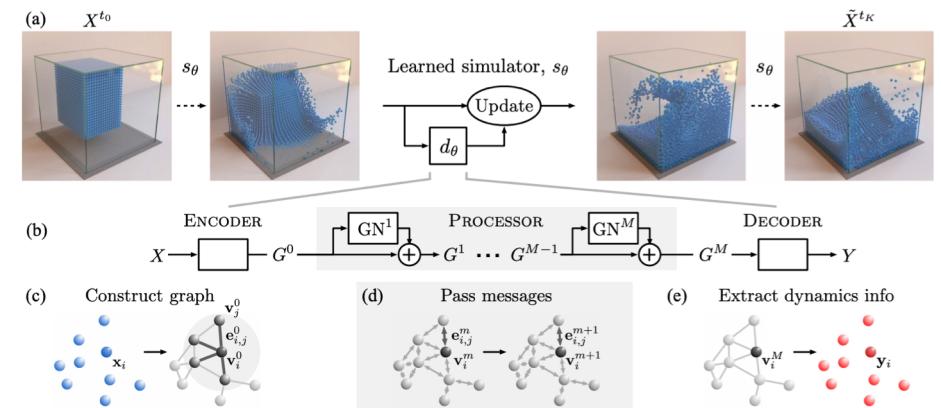
VIEW SOURCE

VIEW PUBLICATION

Learning to Simulate Complex Physics with Graph Networks

Here we present a machine learning framework and model implementation that can learn to simulate a wide variety of challenging physical domains, involving fluids, rigid solids, and deformable materials interacting with one another. Our framework---which we term "Graph Network-based Simulators" (GNS)---represents the state of a physical system with particles, expressed as nodes in a graph, and computes dynamics via learned message-passing. Our results show that our model can generalize from single-timestep predictions with thousands of particles during training, to different initial conditions, thousands of timesteps, and at least an order of magnitude more particles at test time. Our model was robust to hyperparameter choices across various evaluation metrics: the main determinants of long-term performance were the number of message-passing steps, and mitigating the accumulation of error by corrupting the training data with noise. Our GNS framework advances the state-of-the-art in learned physical simulation, and holds promise for solving a wide range of complex forward and inverse problems.

<https://deepmind.com/research/open-source/Learning-to-Simulate-Complex-Physics-with-Graph-Networks>



GNNs for code bug detection

- Experiments conducted on a large dataset with 4.9M methods in 92 different project versions show that GNNs have a relative improvement up to 160% on F-score when comparing with the state-of-the-art bug detection approaches.

```
1  private void print(InputStream body, String jobName) throws PrintException {
2      if (printerOperations.getPrintService().isDocFlavorSupported(printerOperations.
3          getFlavor())) {
4          PrintDocument printDoc = new PrintDocument(body, printerOperations.getFlavor());
5          printerOperations.print(printDoc, config.getCopies(), config.isSendToPrinter(),
6              config.getMimeType(), jobName);
7          printerOperations.print(printDoc, config.isSendToPrinter(), config.getMimeType(),
8              jobName);
9      }
10 }
11 }
```

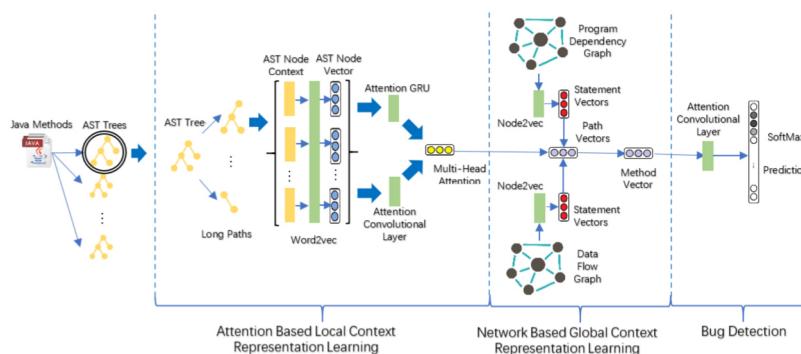
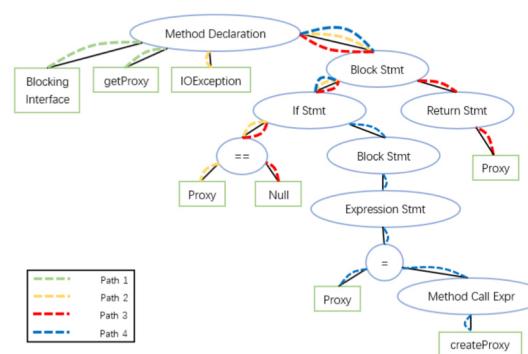


Fig. 2. Overview of our Approach.



(b) The AST of the code in Figure 3a.

<https://dl.acm.org/doi/pdf/10.1145/3360588>

GNNs for genomics

- We proposed to use GNN to replace human heuristics \Rightarrow AI-based genome assembler
- Advantage : Solve genome assembly independently of any type of sequencing machine and no hand-crafting of genome assemblers.
- Given a state-of-the-art genome assembler (Raven), we demonstrated that learned heuristics with GNN outperforms human engineered rules.



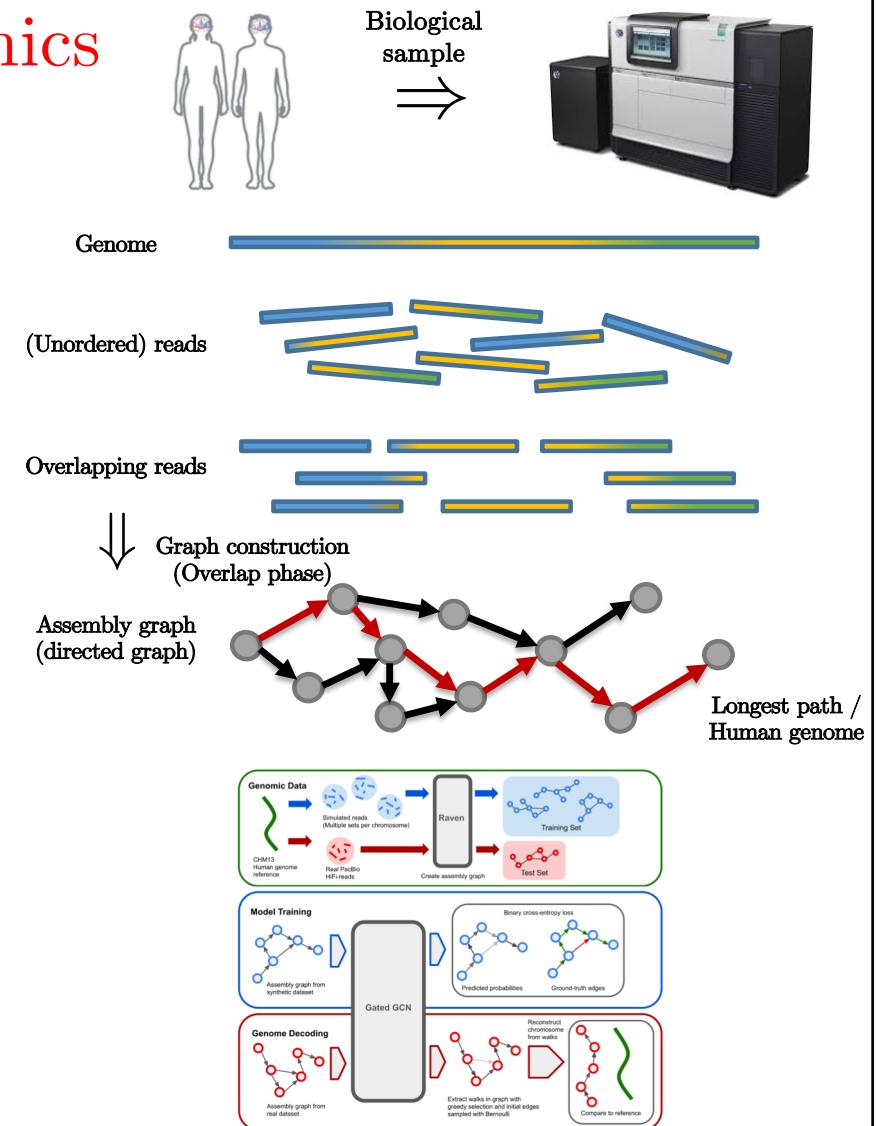
2001



2022

<https://arxiv.org/pdf/2206.00668.pdf>

Xavier Bresson



Outline

- The Deep Learning (DL) revolution
- Limitation of DL
- Graph-Structured data
- Graph Neural Networks (GNNs)
- GNN case studies
- **GNN for industry**
- GNN libraries
- Conclusion

GNNs for Industry

Market potential has a large landscape: GNN is a general technology that can be applied to several tasks

- Finance: fraud/anomaly detection, forecast prediction for e.g. sales, investments, credit risk
- Recommendation: better targeted ads, improved user experience/loyalty
- Knowledge graphs: enhanced CRM, query relationships/interactions with company and customers
- Healthcare: drug design, new diagnostic tools for doctors s.a. brain analysis
- Robotics: better 3D point representation, planning and reasoning
- NLP: improved Q&A chatbot with contextual graphs
- Resource management: supply chain and warehouse/inventory optimization
- Transportation: more accurate and dynamic delivery time
- Etc.

GNNs for Industry

- Recent technology ⇒ No dominant actor yet
- Kumo: <https://kumo.ai>
 - Start-up raised \$18.5 Million from Sequoia in July 2022
 - A co-funder is Jure Leskovec (GNN top expert)
- Amazon Neptune:
<https://aws.amazon.com/neptune/machine-learning>



A screenshot of the Kumo website. The header features the Kumo logo and navigation links for Capabilities, Solutions, Resources, and About Us. A prominent section highlights "Unprecedented time to value" with the tagline "Deliver predictions into production in days not months". Below this, there's a section titled "Comprehensive integrations" with the subtext "to wherever you store your data".

A screenshot of the Amazon Neptune ML website. The header includes the AWS logo and navigation links for Products, Solutions, Pricing, Documentation, Learn, Partner Network, AWS Marketplace, Customer Enablement, Events, Explore More, Contact Us, Support, English, My Account, and Sign In. The main content area is titled "Amazon Neptune ML" with the subtext "Easy, fast, and accurate predictions for graphs". It mentions "Amazon SageMaker Now Supports Shadow Testing | Validate performance of new ML models against production models". The page also discusses the use of Graph Neural Networks (GNNs) and the Deep Graph Library (DGL). A "Benefits" section at the bottom states "Make predictions on graph data without ML expertise" and "Improve the accuracy of most predictions by over 50%".

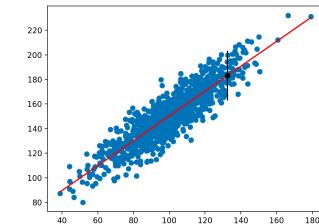
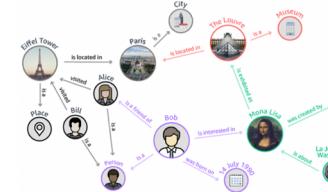


GNN pipeline

- Collect data (user/customer features, product features, etc.)
 - Generate graphs from data relationships and features
 - Train/validate GNN for given predictive tasks
 - Cloud storage, security, computing and deployment
 - Customers query an API to get predictions
- Easy to use for non-expert DL/GNN customers



	A	B	C	D	E	F	G	H	I
	Userid	Name	Age	Height	City	Gender	Degree	Occupation	Interests
1	10001	JORDAN	21	157	Mumbai	Male	PostGraduate	Intermediate	Automotive
2	10001	JMAL	71	138	Mumbai	Female	PostGraduate	Intermediate	Agriculture
3	10002	DPROYN	18	81	Mumbai	Male	MobileAndroid	Graduate	Fashion
4	10002	PNOQX	41	90	Mumbai	Male	SmartTV	PostGraduate	Technical
5	10003	WVJYD	16	156	Mumbai	Male	MobileAndroid	Graduate	Legal
6	10003	VWWDHV	19	97	Mumbai	Male	MobileAndroid	Graduate	Hardware
7	10005	DKKEZ	35	111	Pune	Female	SmartTV	PostGraduate	Construction
8	10006	CAQH	57	137	Pune	Female	MobileAndroid	PostGraduate	Legal
9	10006	QKJTY	72	90	Noida	Female	SmartTV	Graduate	Textile
10	10008	CUPNN	44	139	Noida	Female	SmartTV	Graduate	Textile
11	10009	OKSHU	39	172	Pune	Female	PlayStation	PostGraduate	Information
12	10010	QDRW	12	123	Delhi	Male	PostGraduate	Intermediate	Information
13	10012	MTQGW	32	145	Pune	Female	PlayStation	PostGraduate	Construction
14	10013	NUNHT	31	106	Mumbai	Female	MobileAndroid	Graduate	FinancialServices
15	10014	TBWC	13	52	Pune	Female	Desktop	Intermediate	Fashion

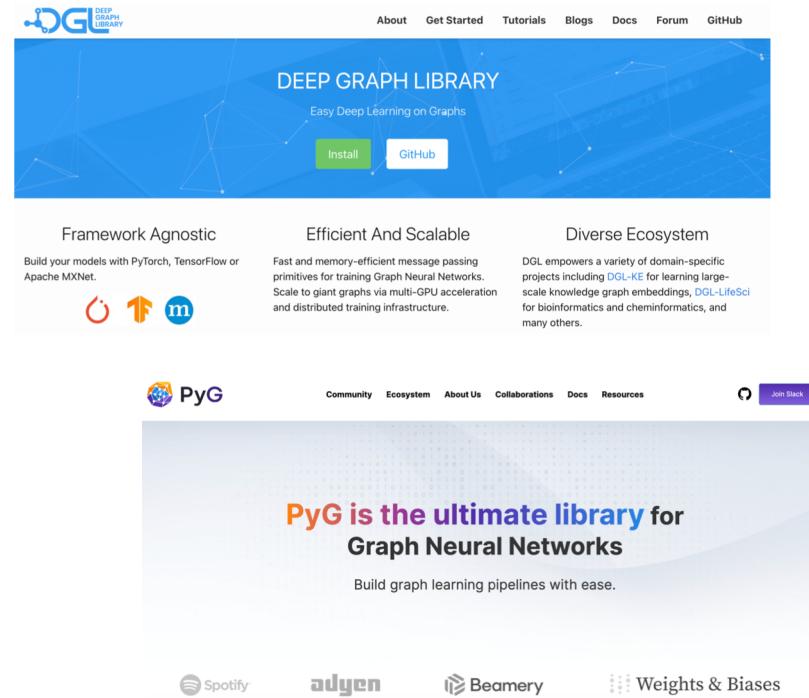


Outline

- The Deep Learning (DL) revolution
- Limitation of DL
- Graph-Structured data
- Graph Neural Networks (GNNs)
- GNN case studies
- GNN for industry
- **GNN libraries**
- Conclusion

GNN libraries

- Amazon DGL (Deep Graph Library)
 - First released on Dec 2018 (11k+ stars)
 - PyTorch / TensorFlow / MxNet
 - <https://www.dgl.ai>



- Kumo PyG (PyTorch Geometric)
 - First released on Mar 2019 (16k+ stars)
 - PyTorch
 - <https://www.pyg.org>

- TensorFlow GNN
 - First released on Nov 2021 (1k+ stars)
 - <https://github.com/tensorflow/gnn>

Outline

- The Deep Learning (DL) revolution
- Limitation of DL
- Graph-Structured data
- Graph Neural Networks (GNNs)
- GNN case studies
- GNN for industry
- GNN libraries
- Conclusion

Conclusion

- Graph Neural Networks are a breakthrough in ML.
 - GNNs have become the standard toolkit for analyzing graph-structured data.
 - They generalize CNNs/RNNs/Transformers from grids/sequences to complex relational data structures.
 - Graphs are everywhere because everything is connected.
 - GNNs are highly flexible and have been/will be applied to a large variety of applications.
 - Large-scale training (linear complexity) with distributed computing.
 - Supervised and self-supervised training.
 - GNNs will boost business analytics.



Questions?