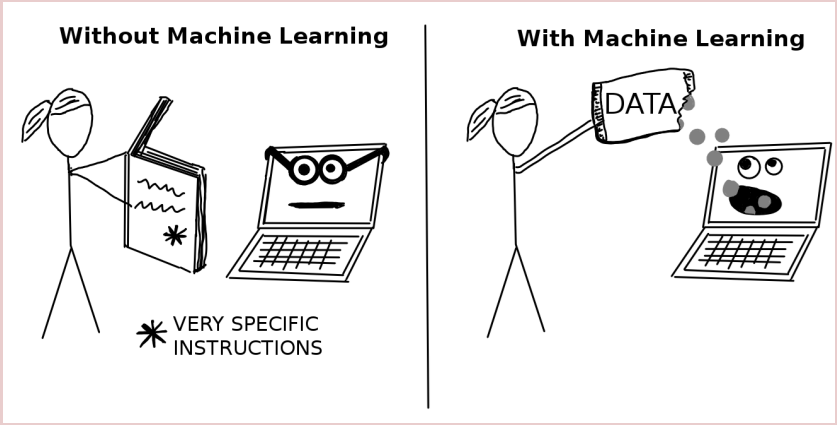
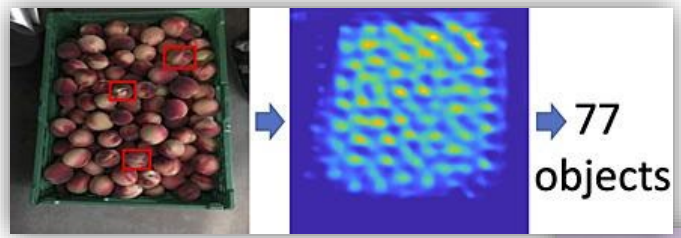
Theoretical Machine Learning @ Stony Brook

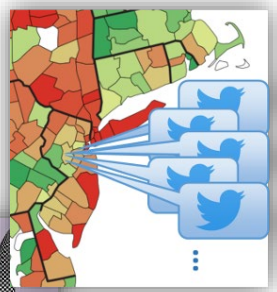
Yifan Sun

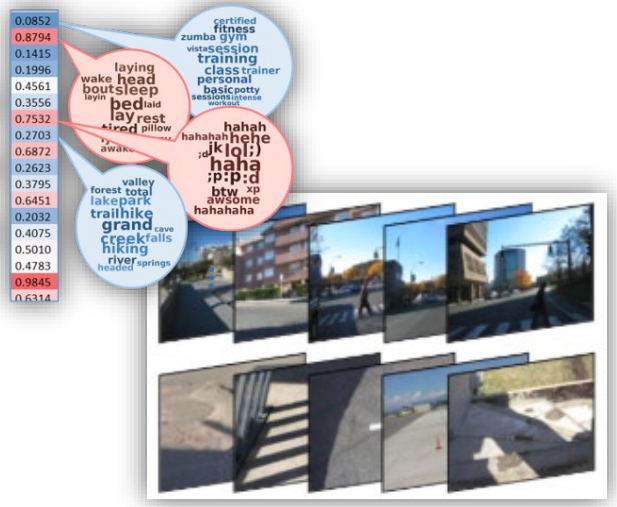
SBU-BNL AI Workshop

June 2024









Stony Brook University

Natural language

H. Andrew Schwartz

Assistant Professor

Large and scalable language analyses for psychological and health discovery; computational social science; natural language processing; lexical semantics;... More

Niranjan Balasubramanian

Assistant Professor

Natural Language Processing (NLP) and information retrieval.

Ritwik Banerjee

Research Assistant Professor

Natural Language Processing (NLP); Computational Linguistics; Information Extraction

Steven Skiena

Distinguished Teaching Professor, SUNY Empire Innovation Professor and Director, AI Institute

Algorithms, Computational Biology, Large-scale Text Analytics and Sentiment Analysis, Social Trends Analysis, Combinatorial

Paul Fodor

Associate Professor of Practice

Artificial Intelligence, Natural Language Processing, Logic Programming, Complex Event Processing, Knowledge Representation for the Semantic Web, Active... More

Praveen Tripathi

Research Assistant Professor

Machine learning, Data mining, Spatio-temporal data analysis, Time series data analysis.

I.V. Ramakrishnan

Professor and Associate Dean for Strategic Initiatives

Artificial Intelligence, Computational Logic, Machine Learning/Computational Logic Combination, Information Retrieval, Computer Accessibility

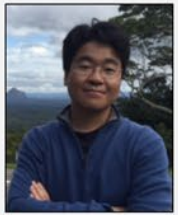
Mining and extraction

Computer vision

Dimitris Samaras

SUNY Empire Innovation Professor

Computer vision; machine learning; computational behavioral sciences; computer graphics; medical imaging; computational photography.



Michael Ryoo

SUNY Empire Innovation Associate Professor

Michael's prime research interest is in the field of deep learning and its applications to computer vision and robotics.

Haibin Ling

SUNY Empire Innovation Professor

Computer Vision, Medical Image Analysis, Augmented Reality and Human-Computer Interaction.

Zhaozheng Yin

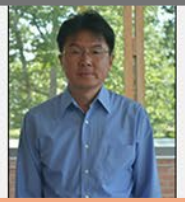
SUNY Empire Innovation Associate Professor

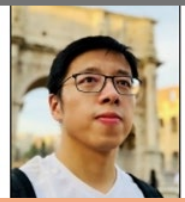
Biomedical Image Analysis, Computer Vision, Machine Learning, Cyber-Physical Systems, Human-Robot Collaboration

Minh Hoai Nguyen

Associate Professor

Computer Vision; Machine Learning; Human Activity Recognition; Semantic Video





Yifan Sun

Assistant Professor

Convex and nonconvex optimization for machine learning and scientific computing

Xianfeng Gu

SUNY Empire Innovation Professor

Computational Conformal Geometry, Computer Graphics, Visualization, Computer Vision, Geometric Modeling, Networking, Medical Imaging, Digital Geometry... More

Stanley Bak

Assistant Professor

Verification of Neural Networks, Cyber-Physical Systems, Formal Analysis of Hybrid Systems

Ting Wang

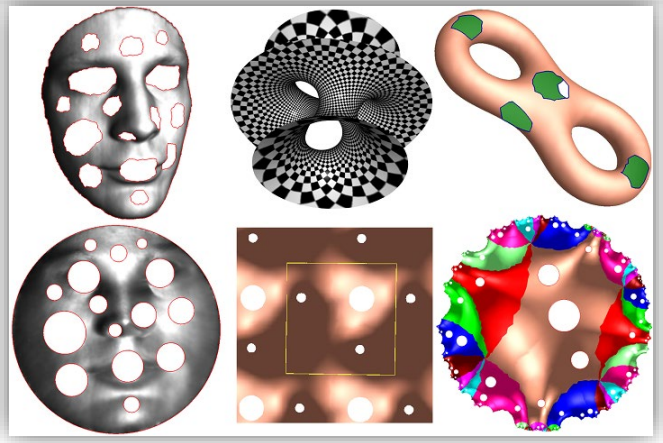
SUNY Empire Innovation Associate Professor

Computer Security and Machine Learning

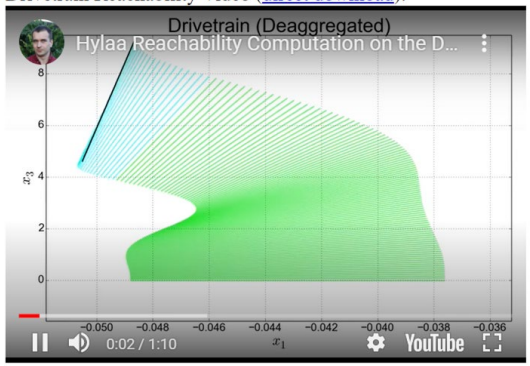
Theoretical foundations and verification

@ SBU

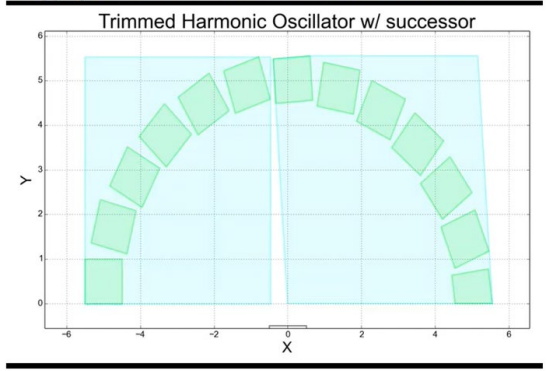
Theoretical foundations and verification



Drivetrain Reachability Video (direct download):



Deaggregation Demonstration Video (direct download):

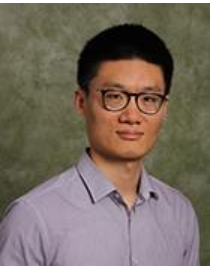


Stony Brook University

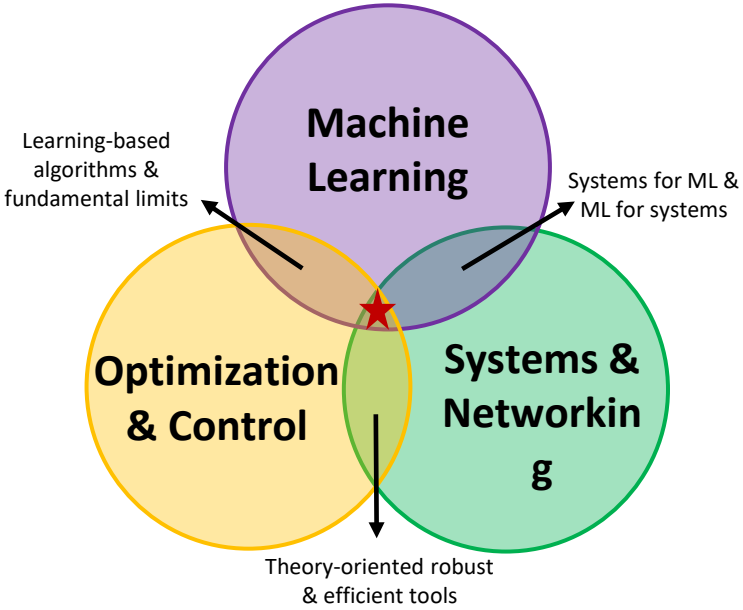
Decision Intelligence & Emerging Networked Systems Laboratory

--- **PI: Jian Li** (jian.li.3@stonybrook.edu)

- **Decision Intelligence:** intersection of sequential decision making (DM) and AI/ML
  - Theory foundation and algorithmic solutions to DM in large-scale AI/ML and data science problems, including algorithm design and analysis, optimization, and implementation, using state-of-the-art mathematical techniques and system technology:
    - Reinforcement learning (RL) and representation learning
      - RL for generative models, e.g., LLMs w/RL, diffusion models
    - Robust/Adversarial and risk-aware online decision making (online learning, e.g., multi-armed bandits)
    - Learning-augmented network optimization
    - Distributed/federated/trustworthy learning and optimization
- **Emerging Networked Systems**
  - Resource allocation in general centered on fundamental issues on optimality, scalability, and trustworthiness (e.g., communication-computation efficiency, resilience)
    - AI for wireless networking: O-RAN, mmWave networks, etc
    - Autonomous systems: robotics, swarm/drones control
    - Edge/Cloud systems: serverless computing, content delivery
    - Cyber-physical systems



★ **Decision Intelligence for Robust and Efficient NextG Data Systems**



- **Multiple funded positions for Ph.D. and/or MS students. Contract Prof. Li for details**
- **Collaboration/internship with industry research labs: AT&T, IBM, MERL, etc**

ML for Cyber-Security

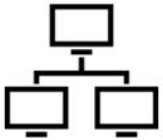
ALPS (Algorithmic, Learning, Privacy and Security) Lab
Principle Investigator: Ting Wang

Thrust 1: Leveraging the advances in ML for cyber-security applications

Software

- Reverse engineering
- Code clone detection
- Fuzzing

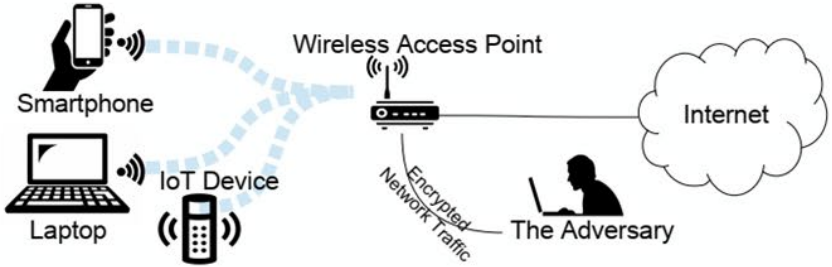
Network



- Attack investigation
- Exploit detection

Web App

- App fingerprinting
- UI analysis



App fingerprinting over encrypted wireless traffic

ML → Security

Trustworthy ML

ALPS (Algorithmic, Learning, Privacy and Security) L
Principle Investigator: Ting Wa

Thrust 2: Investigating security & privacy issues arising in ML techniques

Security

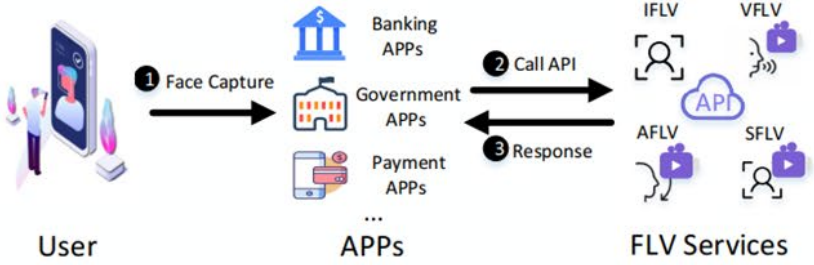
- Adversarial robustness
- Poisoning robustness
- Backdoor robustness

Privacy

- Privacy inference
- Federated learning
- Differential privacy

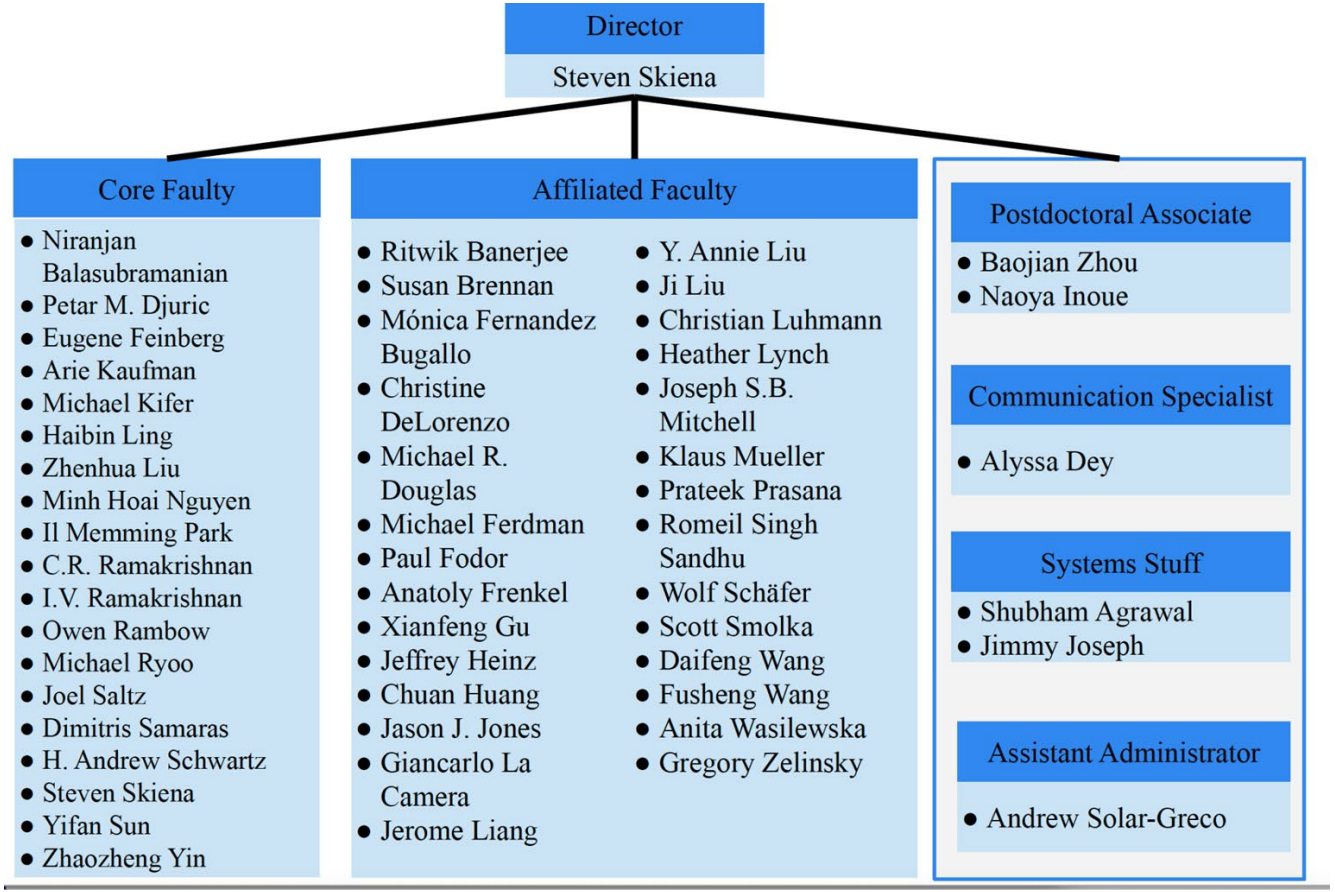
Transparency

- Interpretation reliability
- Human in the loop



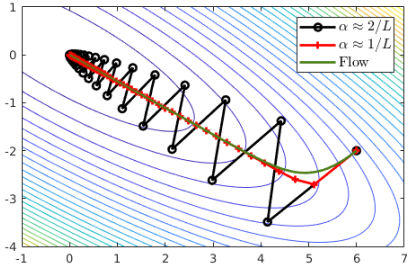
Vulnerability of facial liveness verification

AI Institute

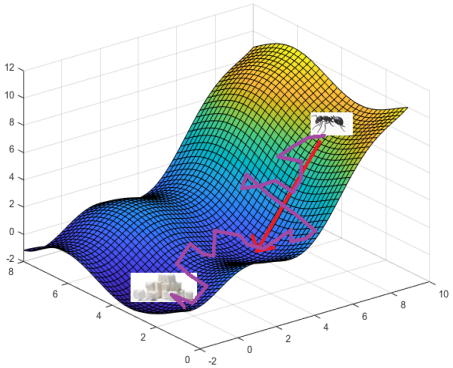


Stony Brook University

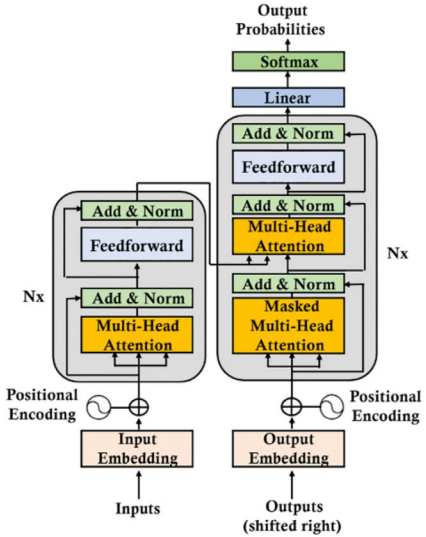
Optimization theory + machine learning



Optimization method behavior



Loss function behavior



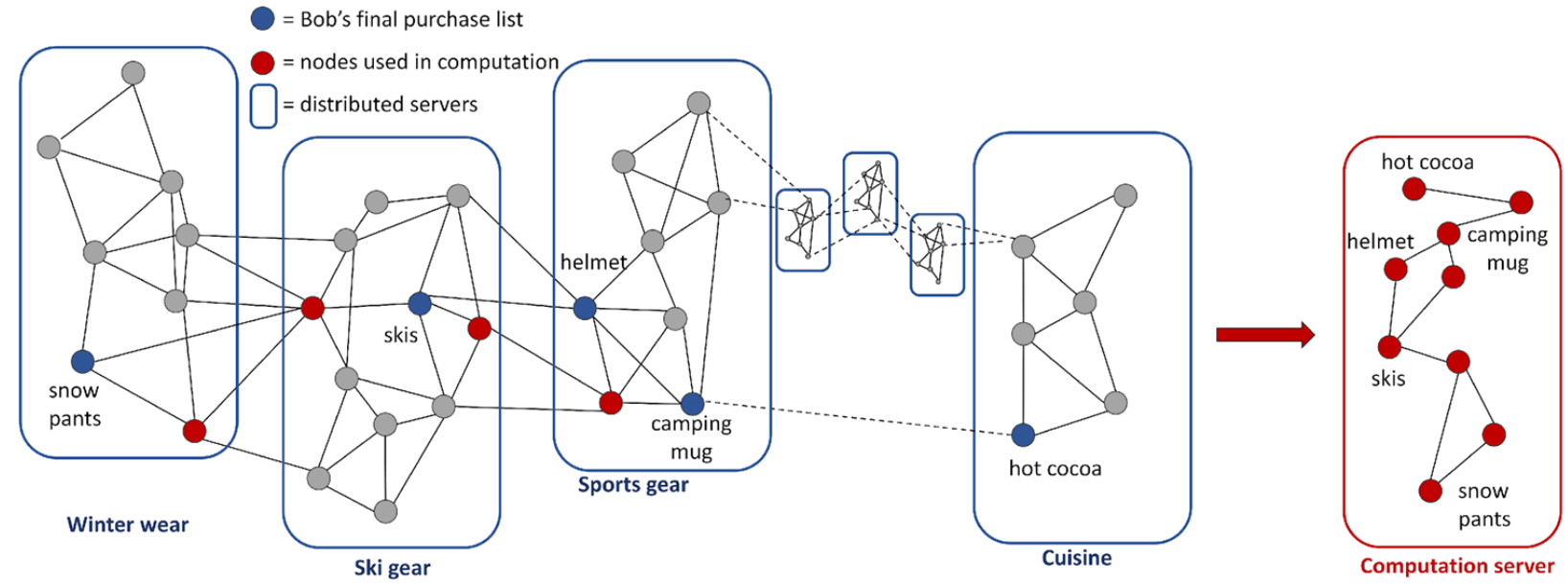
Larger ML models

Security?
Privacy?
Dependability
?

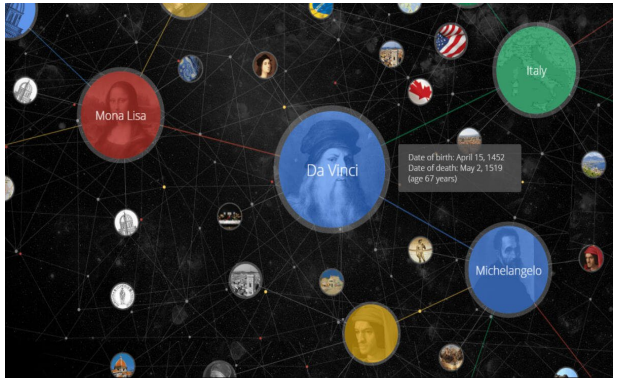
Modern day applications

Stony Brook University

Machine learning over graphs



Machine learning over graphs



- Amazon's product graph: 12 million products
- Google knowledge graph: 500 million entities
- Paypal's fraud detection: 500 million users

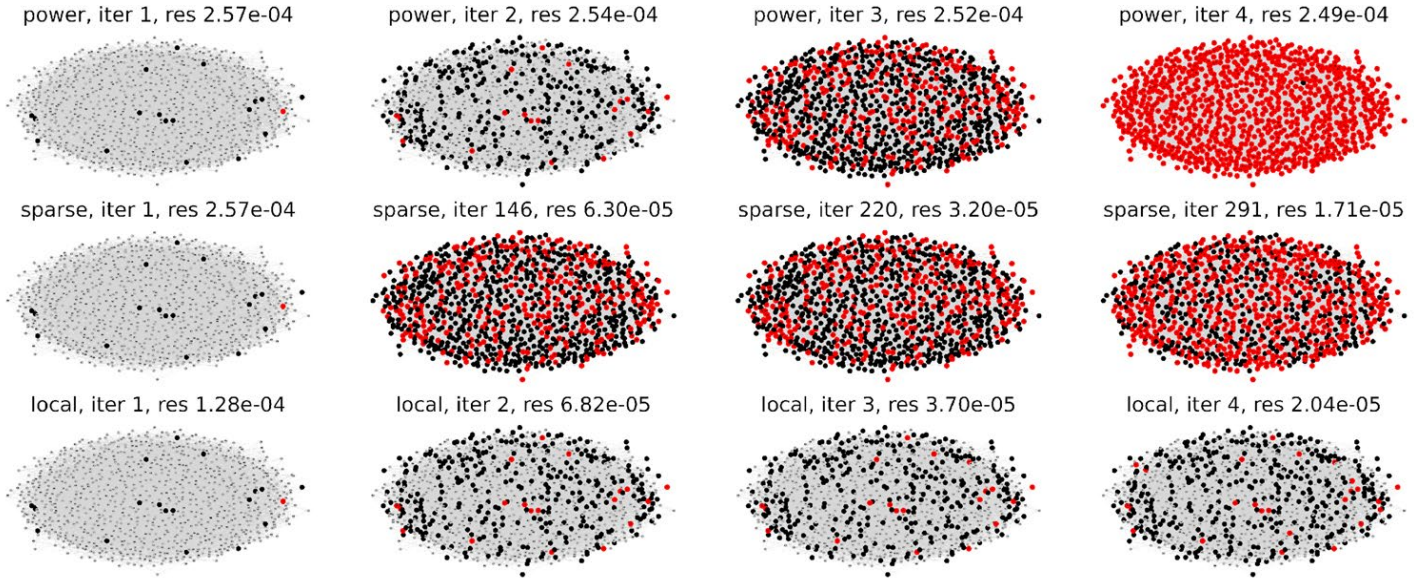
These graphs don't fit on one server!

Machine learning over graphs

- Cybersecurity or web services: Internet consists of billions of websites, connects billions of devices, all heterogeneous nodes
- Complex scenarios, many independent small entities

We want it now, we want it fast

Fast graph solvers with low memory footprint



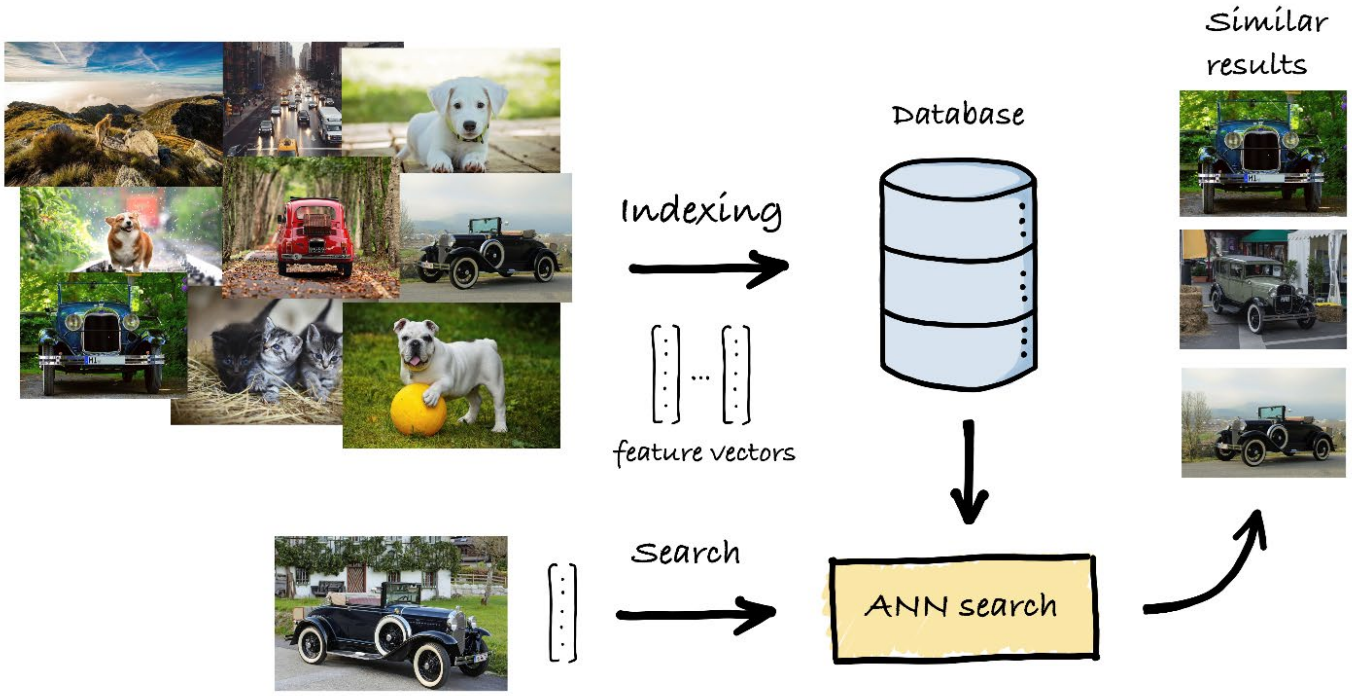
Baojian Zhou and Yifan Sun. Approximate frank-wolfe algorithms over graph-structured support sets. In *International Conference on Machine Learning*, pages 27303–27337. PMLR, 2022.

Baojian Zhou, Yifan Sun, and Reza Babanezhad Harikandeh. Iterative methods via locally evolving set process for large graphs. Under review.

Baojian Zhou, Yifan Sun, and Reza Babanezhad Harikandeh. Fast online node labeling for very large graphs. In *International Conference on Machine Learning*, pages 42658–42697. PMLR, 2023.

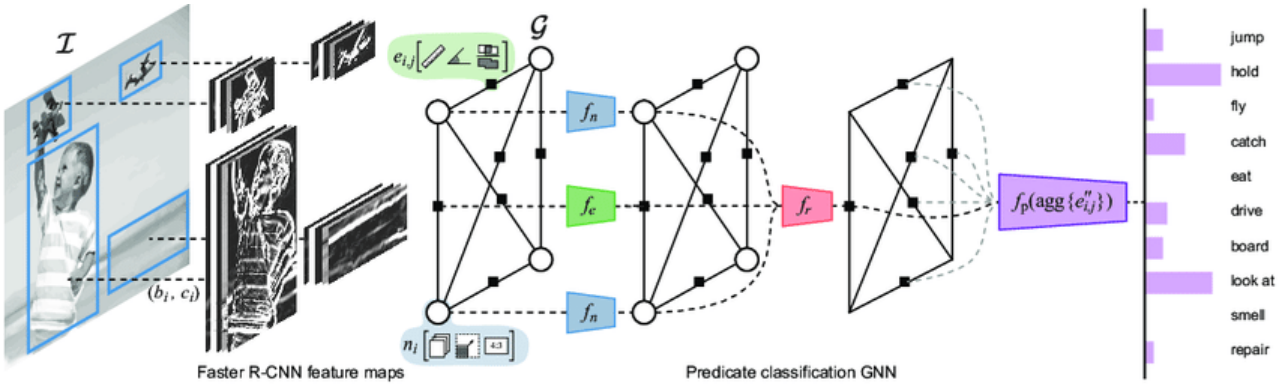
Stony Brook University

Application: Information retrieval



Winning the NeurIPS BillionScale Approximate Nearest Neighbor Search Challenge
Mariano Tepper, Cecilia Aguerrebere, Ted Willke, Sourabh Dongaonkar, Jawad B Khan, Mark

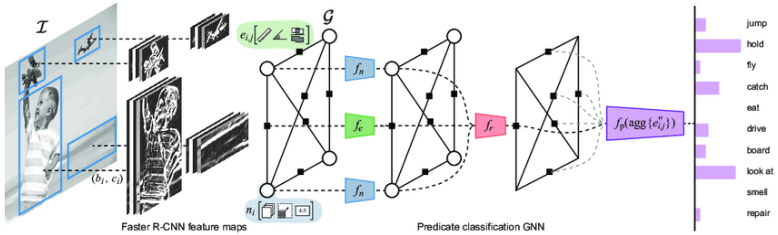
Graphical neural networks



Baldassarre, Federico & Smith, Kevin & Sullivan, Josephine & Azizpour, Hossein. (2020). Explanation-based Weakly-supervised Learning of Visual Relations with Graph Networks.

Stony Brook University

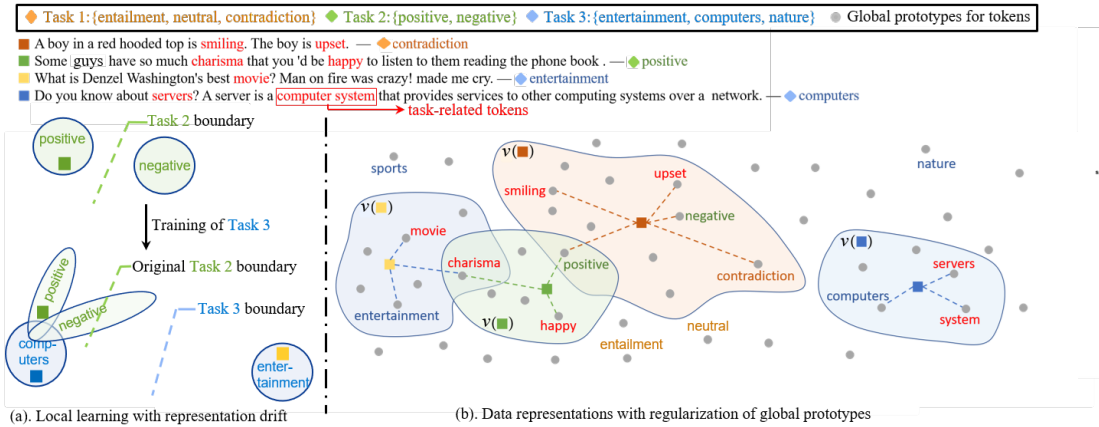
Application: train deeper graphical neural networks



- GNNs cannot be deep
  - 3-4 layers deep
  - In contrast, CNNs are 50-100 layers deep
- Memory complexity
  - “width” = number of nodes
  - Even a sparse graph becomes dense in a few layers
- Oversmoothing
  - Output performance degrades with depth

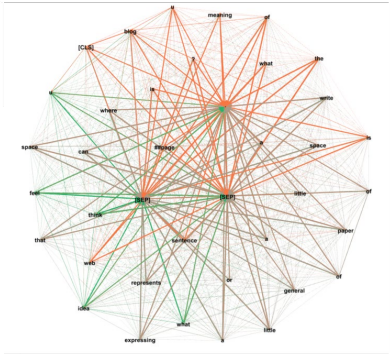
Better message passing protocols for GNNS?

Modeling transformers with graphs

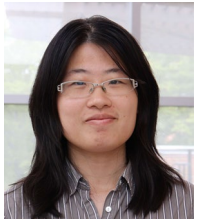


(a). Local learning with representation drift

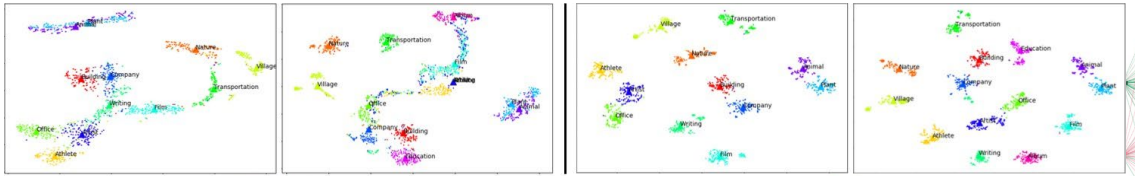
(b). Data representations with regularization of global prototypes



BERT Last Layer

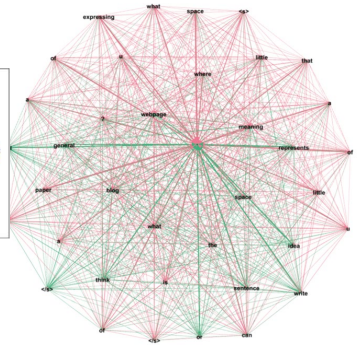


Xueying Bai



(a) Fine-Tuning

(b) NeiAttn



RoBERTa Last Layer

Niranjan Balasubramanian

Neighborhood attention for continual learning

Stony Brook University

Thank you!

Stony Brook University