LIONS EDIC Intro

Volkan Cevher, Associate Professor EPFL

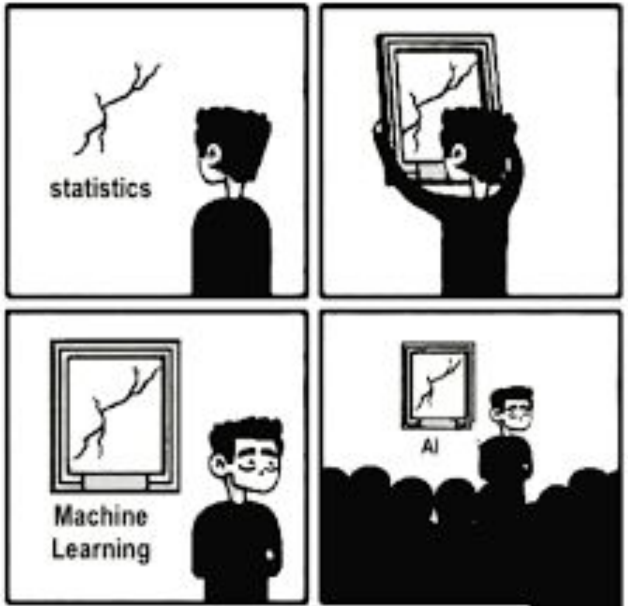
Preface

My research:

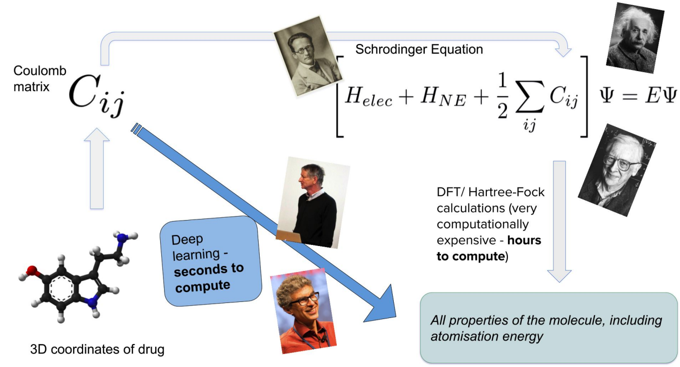
- Machine Learning (ML)
- Optimization
- Signal Processing
- Information Theory
- Statistics
- Automatic Control

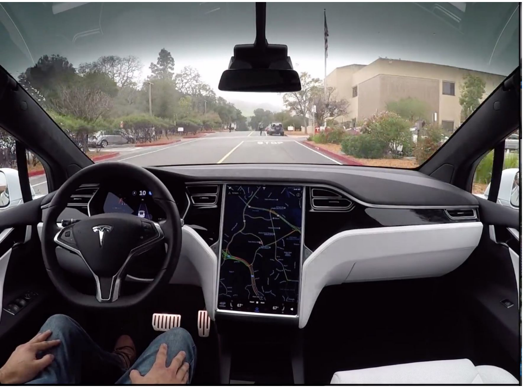
My current courses:

- Mathematics of Data
- Reinforcement Learning
- Advanced Topics in ML
- Online Learning in Games

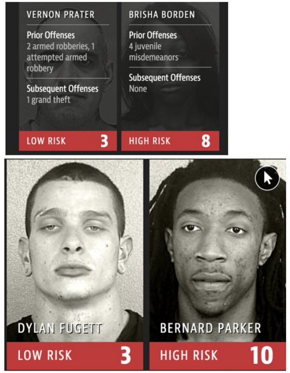


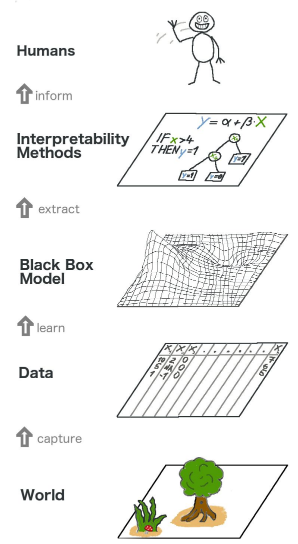
Machine Learning (ML) works





Machine Learning (ML) works... but many new challenges A-RISE!



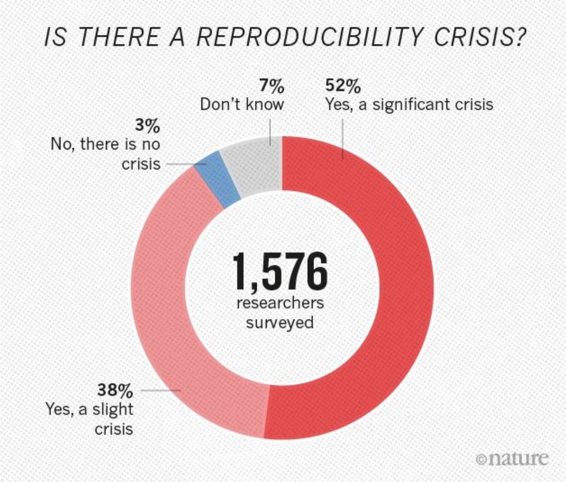


GPT-3 medical chatbot tells suicidal test patient to kill themselves

Rob Beschizza 2 days ago

Researchers experimenting with GPT-3, the AI text-generation model, found that it is not ready to replace human respondents in the chatbox.

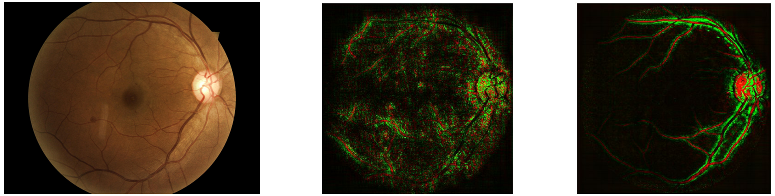
1. **R**obustness
2. **I**nterpretability
3. Bia**s** & Fairness
4. R**e**producibility

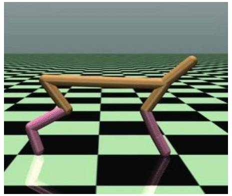


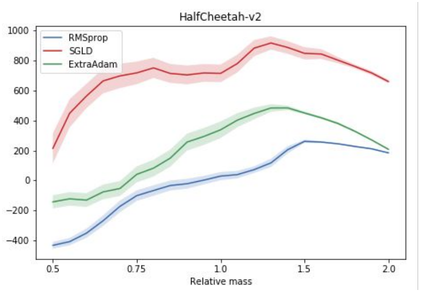
Robustness and interpretability are entangled

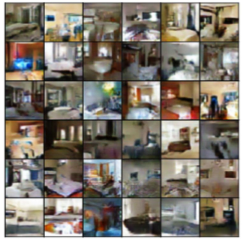
$$\min_{x:x\in\mathcal{X}} \max_{y:y\in\mathcal{Y}} \Phi(x,y)$$

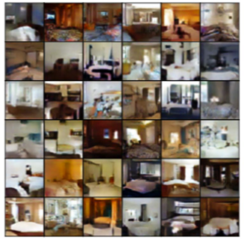
- Robust ML problems
  a. Adversarial training
  b. Robust reinforcement learning
  c. GAN training
- Interpretability
  a. Integrated gradients
  b. Sparse features
  c. Disentangled representations

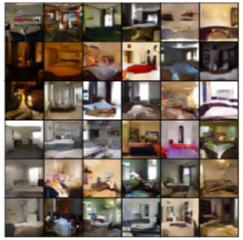








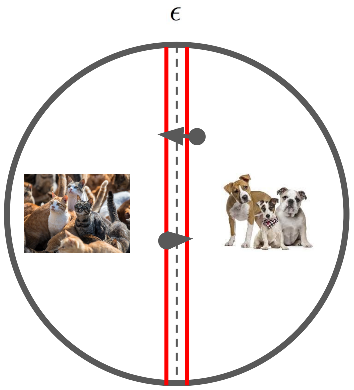


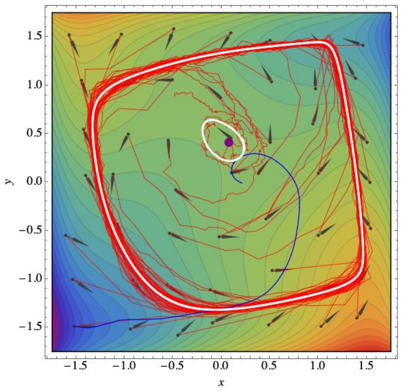


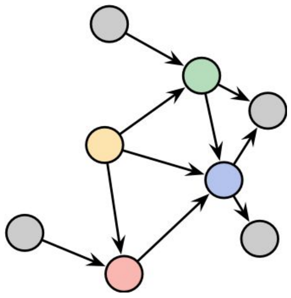
1. Finding mixed Nash equilibria of generative adversarial networks. Hsieh, Liu and Cevher. ICML'19.
2. Robust reinforcement learning via adversarial training with langevin dynamics. Kamalaruban, Huang, Hsieh, Rolland, Shi and Cevher. NeurIPS'20.
3. A natural actor-critic framework for zero-sum Markov games. Alacaoglu, Viano, He, and Cevher. ICML'22
4. Score matching enables causal discovery of nonlinear additive noise models. Rolland, Cevher, et al., ICML'22.

Challenges to robustness & interpretability

$$\min_{x:x\in\mathcal{X}} \max_{y:y\in\mathcal{Y}} \Phi(x,y)$$







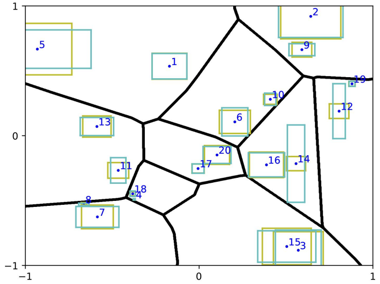
- Critical challenges to robustness
    a. High-dim. geometry
    b. Spurious attractors
    c. Reproducibility & Verification

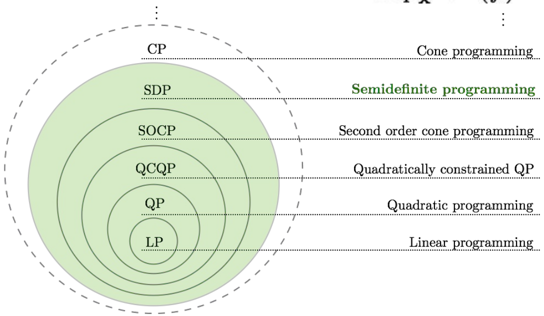
1. Are adversarial examples inevitable? Shafahi, Huang, Studer, Feizi, and Goldstein. ICLR'19.
2. The limits of min-max optimization algorithms: Convergence to spurious non-critical sets. Hsieh, Mertikopoulos, Cevher. ICML'20.
3. Score matching enables causal discovery of nonlinear additive noise models. Rolland, Cevher, et al., ICML'22.
4. Escaping limit cycles: Global convergence for constrained nonconvex-nonconcave minimax problems. Pethick, Latafat, Patrinos, Fercoq, Cevher. ICLR'22.

Trust but verify

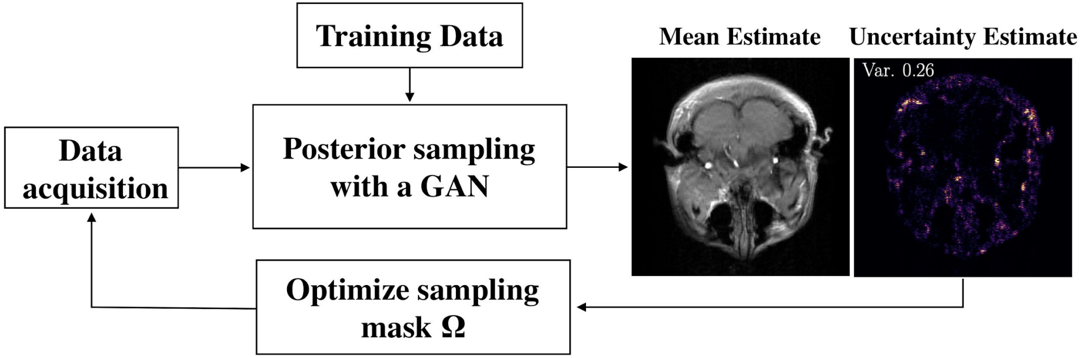
- Trustworthiness requires effort
  a. Geometric boundary
  b. Algebraic formulations
  c. Uncertainty estimation

$$|f(\mathbf{x}+\epsilon) - f(\mathbf{x})| \leq \|\epsilon\| \underbrace{\|\nabla f(\mathbf{x})\|}_{\sup_{\mathbf{x}} \to L(f)}$$

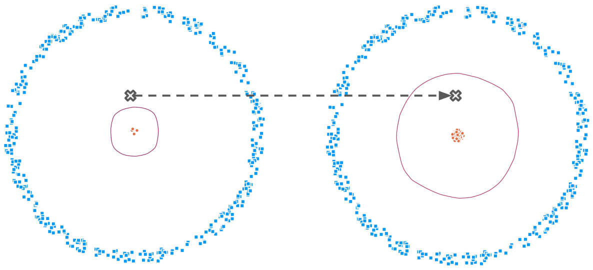


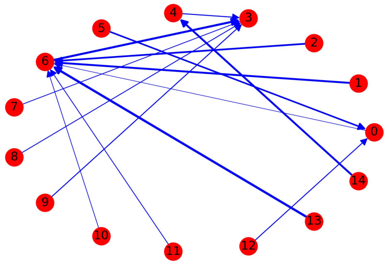


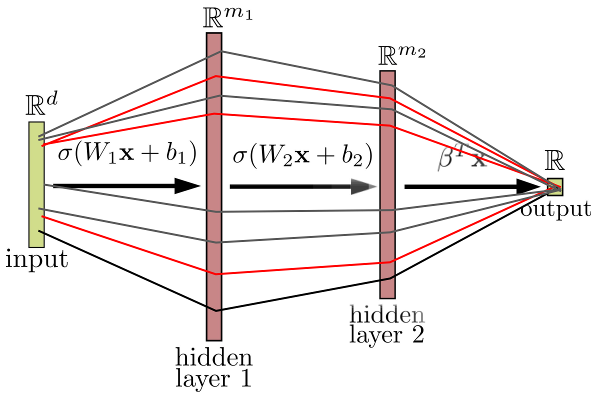
1. On certifying non-uniform bounds against adversarial attacks. Liu, Tomioka and Cevher. ICML'19.
2. Lipschitz Constant Estimation of Neural Network via Sparse Polynomial Optimization. Latorre, Rolland and Cevher. ICLR'20.
3. Scalable semidefinite programming. Yurtsever, Tropp, Fercoq, Udell and Cevher. SIAM Journal on Mathematics of Data Science, 2021.
4. Uncertainty-Driven Adaptive Sampling via GANs, Sanchez, Krawczuk, Sun and Cevher. NeurIPS Workshop, 2020.

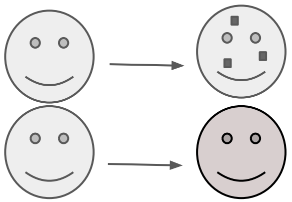


Challenges to Fairness & Bias







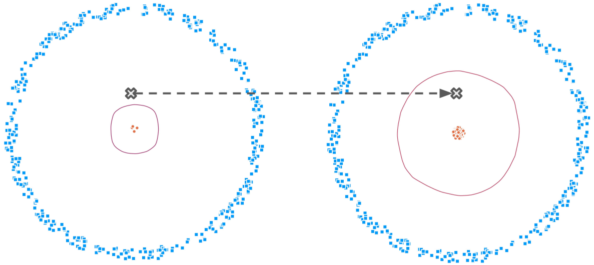


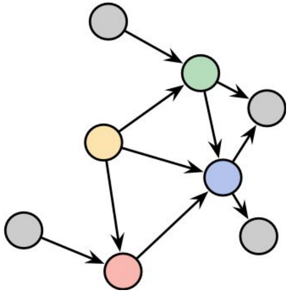
- Models
  a. Causal & disentangled reps.
  b. Generative models
  c. Manifold models…
- Regularizers (and optimizers!)
  a. Robustness: Path norm.
  b. Structured (graph) sparsity
  c. Learning-based regularizers

1. Efficient Proximal Mapping of the 1-path-norm of Shallow Networks. Latorre, Rolland, Hallak and Cevher. ICML'20.
2. An inexact augmented Lagrangian framework for nonconvex optimization with nonlinear constraints. Sahin, Eftekhari, Alacaoglu, Latorre and Cevher. NeurIPS'19.
3. The Rademacher Complexity of Quadratic Classifiers with a nuclear-norm constraint. Latorre, Dadi, Rolland, and Cevher. NeurIPS'20.
4. Subquadratic overparameterization for shallow neural networks. Song, Ramezani-Kebrya, Pethick, Eftekhari, Cevher, NeurIPS'21.

New projects: Mathematical foundations for RISE of AI

$$\min_{x:x\in\mathcal{X}} \max_{y:y\in\mathcal{Y}} \Phi(x,y)$$





Game Theory

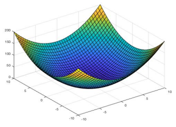
1. Distributionally robust
2. Multi-player setting
3. (Mixed) Nash equilib.
4. Universal & adaptive

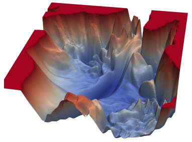
Geometry

1. Trust but verify
2. NAS & Algorithm
3. Explicit bias
4. Limit cycles

Adaptation

1. Enhance interpretability
2. Model-based adaptation
3. Reinforcement learning
4. Representation theory

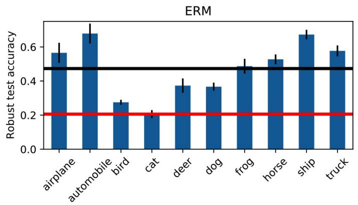


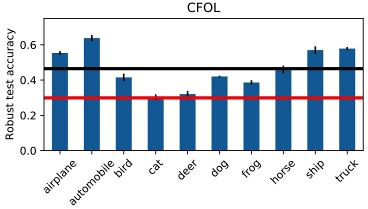


Scalable, rigorous, and reproducible optimization

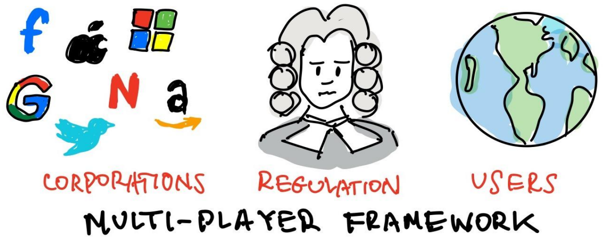
Game Theory

1. Distributionally robust
2. Multi-player games
3. (Mixed) Nash equilib.
4. Universal & adaptive

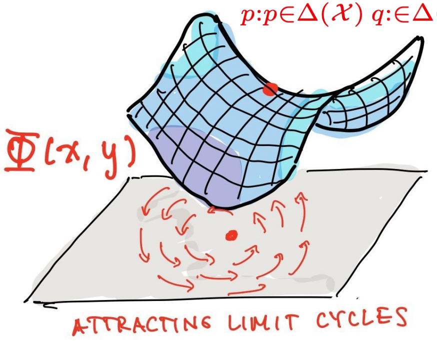




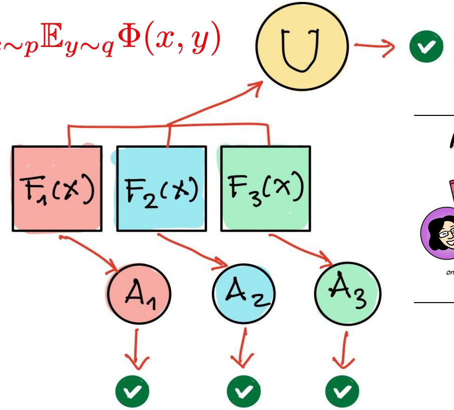
$$\min_{x:x\in\mathcal{X}} \max_{y:y\in\mathcal{Y}} \Phi(x,y)$$



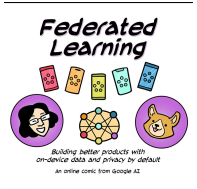
$$\min_{p:p\in\Delta(\mathcal{X})} \max_{q:\in\Delta(\mathcal{Y})} \mathbb{E}_{x\sim p}\mathbb{E}_{y\sim q}\Phi(x,y)$$



CHALLENGE: NON-CONCAVE NON-CONVEX MIN-MAX OPTIMIZATION



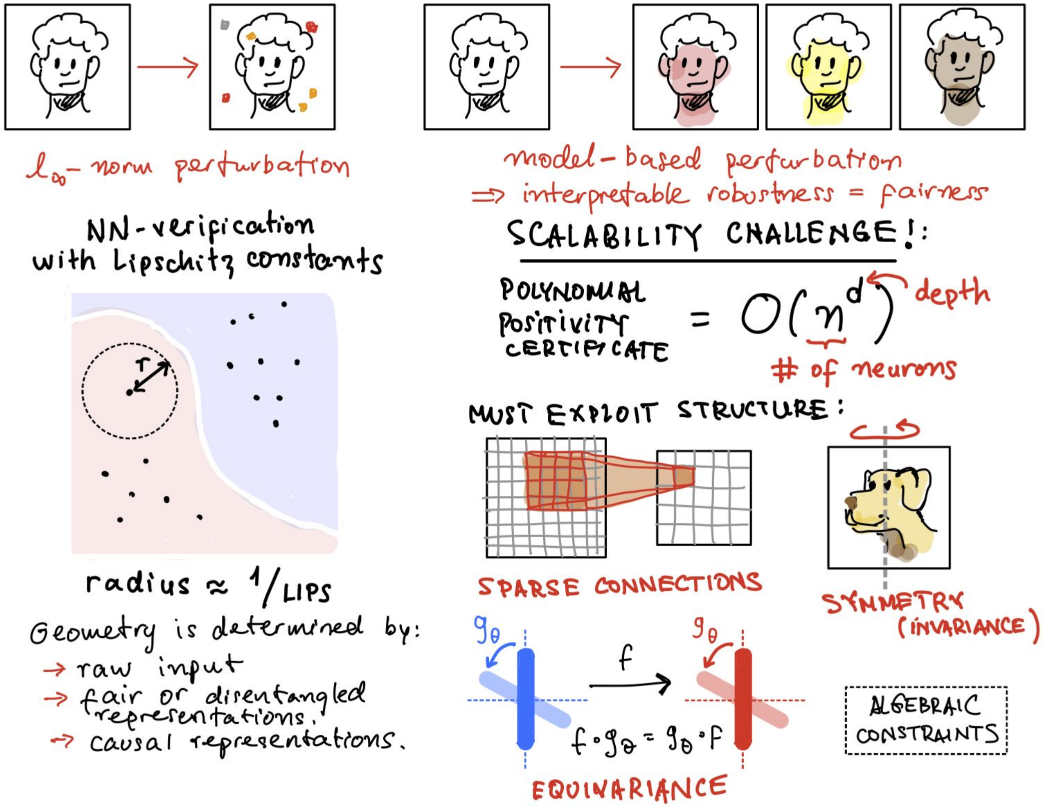
UNIVERSAL OPTIMIZATION ALGORITHMS



Geometric and algebraic foundations for NNs

Geometry

1. Trust but verify
2. NAS & Algorithm
3. Explicit bias
4. Limit cycles

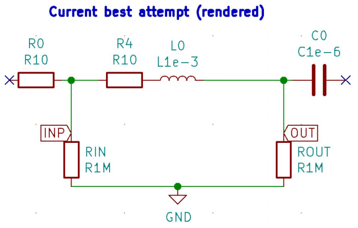


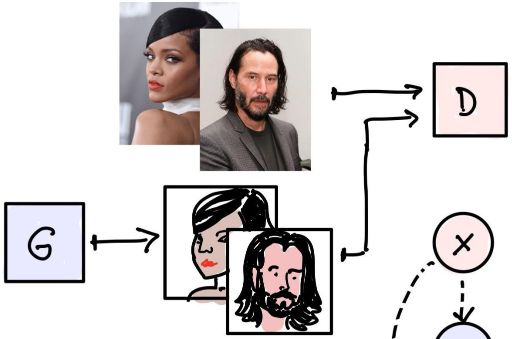
Interpretability: New technical developments

$$\Delta\chi\Delta\rho \geq \frac{\hbar}{2}$$

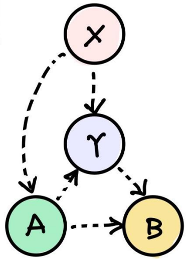
Adaptation

1. Enhance interpretability
2. Model-based adaptation
3. Robust RL
4. Representation theory

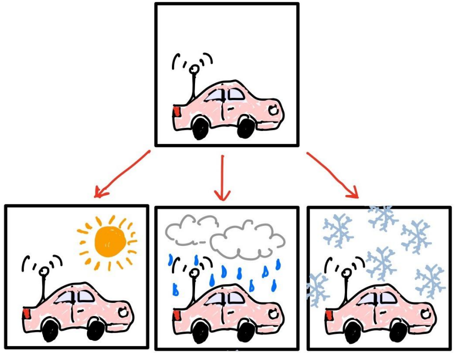




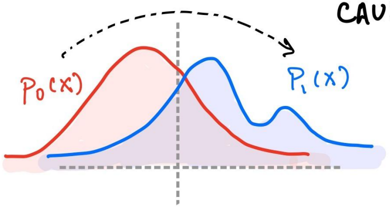
GENERATIVE NETWORKS



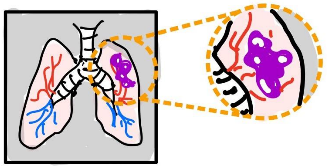
CAUSAL GRAPH



ROBUST REINFORCEMENT LEARNING



MODEL-BASED ADAPTATION



UNCERTAINTY QUANTIFICATION

New applications: Robust domain adaptation (DA), Causality, GNNs

Source domain

labels

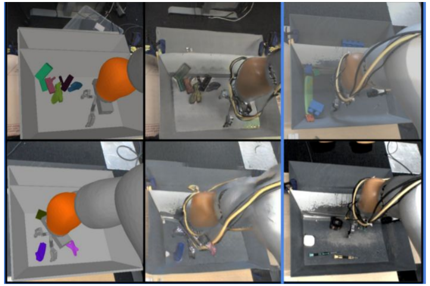
Target domain

No labels
Few labels

1. Using simulation and domain adaptation to improve efficiency of deep robotic grasping . K. Bousmalis, et. al., IEEE ICRA 2018.
2. CYCADA: cycle-consistent adversarial domain adaptation. J. Hoffman et. al., ICML 2018.

Robotics applications:

Simulated data (*SD*) | Domain Adapted synthetic data | Real images (*TD*)

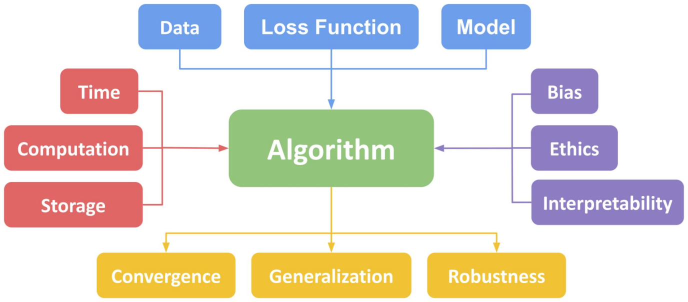


Computer vision applications:

Simulated images | Domain adapted synthetic data

- Mathematical formulations for Robust DA
  - formulation challenges
  - algorithmic challenges
- Causality
  - Joint with Amazon Research
  - Nonlinear models
- Knowledge graphs & graph generation
  - Scalable graph inference
  - Scientific ML & circuit applications

Ideal profile



- Pareto frontier of theory & crack coding
  - super collaborative group (within EPFL and outside)
  - strong connections to industry (Amazon, Google,...)
- Academic track & research labs
  - Faculty@Rice, NUS, Umea (2), Zhejiang, UNC, Linkoping, AIMS, UoCB, VNU, NTU, Technion, UCL, Alborg
  - Postdoc@ETHZ (3), McGill, Turing, UWisc, University of Geneva
  - Others@Kandou bus, SwissRE, TUM, Meta

Ideal profile

- Basic, future-proof research
  - NeurIPS, ICML, ICLR, SIAM Opt, Math Prog,…
    - https://lions.epfl.ch
    - http://csrankings.org/#/index?mlmining&world
    - https://ml.epfl.ch

#	Publication	h5-index	h5-median
1.	Nature	414	607
2.	The New England Journal of Medicine	410	704
3.	Science	391	564
4.	IEEE/CVF Conference on Computer Vision and Pattern Recognition	356	583
5.	The Lancet	345	600
6.	Advanced Materials	294	406
7.	Cell	288	459
8.	Nature Communications	287	389
9.	Chemical Reviews	270	434
10.	International Conference on Learning Representations	253	470
11.	JAMA	253	446
12.	Neural Information Processing Systems	245	422
13.	Proceedings of the National Academy of Sciences	245	337
14.	Journal of the American Chemical Society	245	330
15.	Angewandte Chemie	235	314
16.	Chemical Society Reviews	234	339
17.	Nucleic Acids Research	233	512
18.	Renewable and Sustainable Energy Reviews	225	294
19.	Journal of Clinical Oncology	213	297
20.	Physical Review Letters	209	297
21.	Advanced Energy Materials	206	267
22.	Nature Medicine	205	356
23.	International Conference on Machine Learning	204	370

CSRankings: Computer Science Rankings

CSRankings is a metrics-based ranking of top computer science institutions around the world. **Click on a tria** faculty member's home page. **Click on a chart icon** (the after a name or institution) to see the distribution **Scholar icon** ( ) to see publications, and **click on the DBLP logo** ( ) to go to a DBLP entry. *Applying to gra* **CSrankings on GitHub.**

Rank institutions in the world by publications from 2015 to 2022

All Areas [off | on]

AI [off | on]

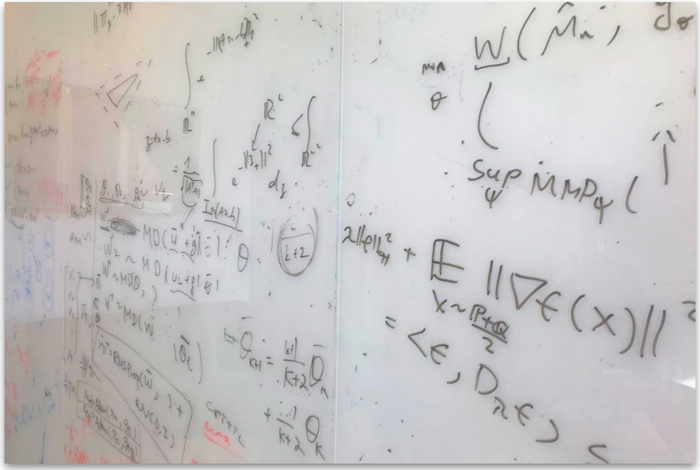
- ▶ Artificial intelligence ☐
- ▶ Computer vision ☐
- ▶ Machine learning & data mining ☑
- ▶ Natural language processing ☐
- ▶ The Web & information retrieval ☐

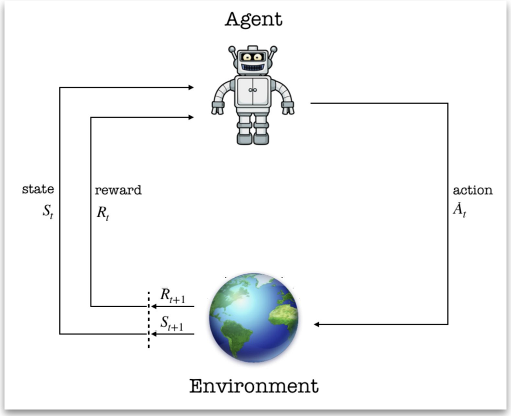
#	Institution	Count	Faculty
1	▶ Carnegie Mellon University	138.3	52
2	▶ Univ. of California - Berkeley	127.9	45
3	▶ Stanford University	115.8	43
4	▶ Massachusetts Institute of Technology	112.4	54
5	▶ Tsinghua University	84.3	50
6	▶ Univ. of Illinois at Urbana-Champaign	82.7	39
7	▶ Univ. of California - Los Angeles	80.5	19
8	▶ Cornell University	75.0	29
9	▶ Univ. of California - San Diego	71.9	38
10	▼ EPFL	70.8	27

Ideal profile

- Basic, future-proof research
  - NeurIPS, ICML, ICLR, SIAM Opt, Math Prog,...
    - https://lions.epfl.ch
    - http://csrankings.org/#/index?mlmining&world
    - https://ml.epfl.ch

- Interest in integrated teaching & research
  - Mathematics of data
  - Reinforcement learning
  - Advanced topics in ML
  - Online learning in games





Thank you!

Visit us at ELD 243 and have coffee with my group members!

Also, open house next Friday in our lab!