Presentation Paper List - CS583 Section B

1. Unbiased Gradient Estimation In Unrolled Computation Graphs With Persistent Evolution
Paul Vicol, Luke Metz and Jascha Sohl-Dickstein
https://proceedings.mlr.press/v139/vicol21a.html

2. Oops I took a gradient: Scalable sampling for discrete distributions
Will Grathwohl, Kevin Swersky, Milad Hashemi, David Duvenaud and Chris J. Maddison
https://arxiv.org/abs/2102.04509

3. Optimal complexity in decentralised training
Yucheng Lu and Christopher De Sa
https://arxiv.org/abs/2006.08085

4. Understanding self-supervised learning dynamics without contrastive pairs
Yuandong Tian, Xinlei Chen, and Surya Ganguli
https://arxiv.org/abs/2102.06810

5. Solving high-dimensional parabolic PDEs using the tensor train format
Lorenz Richter, Leon Sallandt, and Nikolas Nüsken
https://arxiv.org/abs/2102.11830

6. No-Regret Learning Dynamics for Extensive-Form Correlated Equilibrium.
Andrea Celli, Alberto Marchesi, Gabriele Farina, Nicola Gatti
https://arxiv.org/abs/2004.00603

7. Improved Guarantees and a Multiple-Descent Curve for Column Subset Selection and the Nyström Method.
Michał Dereziński, Rajiv Khanna, Michael W. Mahoney
https://arxiv.org/abs/2002.09073

8. Language Models are Few-Shot Learners.

Tom B. Brown, Benjamin Mann, Nick Ryder, Melanie Subbiah, Jared Kaplan, Prafulla Dhariwal, Arvind Neelakantan, Pranav Shyam, Girish Sastry, Amanda Askell, Sandhini Agarwal, Ariel Herbert-Voss, Gretchen Krueger, Tom Henighan, Rewon Child, Aditya Ramesh, Daniel M. Ziegler, Jeffrey Wu, Clemens Winter, Christopher Hesse, Mark Chen, Eric Sigler, Mateusz Litwin, Scott Gray, Benjamin Chess, Jack Clark, Christopher Berner, Sam McCandlish, Alec Radford, Ilya Sutskever, Dario Amodei
https://arxiv.org/abs/2005.14165

9. Distribution-Independent PAC Learning of Halfspaces with Massart Noise
Ilias Diakonikolas, Themis Gouleakis, Christos Tzamos
https://arxiv.org/abs/1906.10075

10. Uniform convergence may be unable to explain generalization in deep learning
Authors: Vaishnavh Nagarajan, J. Zico Kolter
https://arxiv.org/abs/1902.04742

11. Nonparametric Density Estimation & Convergence Rates for GANs under Besov IPM Losses
Ananya Uppal, Shashank Singh, Barnabás Póczos
https://arxiv.org/abs/1902.03511

12. Fast and Accurate Least-Mean-Squares Solvers
Alaa Maalouf, Ibrahim Jubran, Dan Feldman
https://arxiv.org/abs/1906.04705

13. Putting An End to End-to-End: Gradient-Isolated Learning of Representations
Sindy Löwe, Peter O'Connor, Bastiaan S. Veeling
https://arxiv.org/abs/1905.11786

14. Continuous 3D-Structure-Aware Neural Scene Representations
Vincent Sitzmann, Michael Zollhöfer, Gordon Wetzstein
https://arxiv.org/abs/1906.01618

15. On Learning Sets of Symmetric Elements

Haggai Maron, Or Litany, Gal Chechik, Ethan Fetaya
https://arxiv.org/abs/2002.08599

16. Tuning-free Plug-and-Play Proximal Algorithm for Inverse Imaging Problems
Kaixuan Wei, Angelica I Aviles-Rivero, Jingwei Liang, Ying Fu, Carola-Bibiane Schönlieb, Hua Huang
https://arxiv.org/abs/2002.09611

17. Efficiently sampling functions from Gaussian process posteriors
James Wilson, Slava Borovitskiy, Alexander Terenin, Peter Mostowsky, Marc Deisenroth
https://arxiv.org/abs/2002.09309

18. Generative Pretraining From Pixels
Mark Chen, Alec Radford, Rewon Child, Jeffrey K Wu, Heewoo Jun, David Luan, Ilya Sutskever
https://cdn.openai.com/papers/Generative_Pretraining_from_Pixels_V2.pdf

19. Dual Averaging Method for Regularized Stochastic Learning and Online Optimization
Lin Xiao
https://www.jmlr.org/papers/v11/xiao10a.html

20. Video-aided Unsupervised Grammar Induction
Songyang Zhang, Linfeng Song, Lifeng Jin, Kun Xu, Dong Yu and Jiebo Luo
https://arxiv.org/abs/2104.04369

21. Unifying Cross-Lingual Semantic Role Labeling with Heterogeneous Linguistic Resources
Simone Conia, Andrea Bacciu and Roberto Navigli
https://www.semanticscholar.org/paper/Unifying-Cross-Lingual-Semantic-Role-Labeling-with-Conia-Bacciu/2172c289b97e4d709f1c54683a242ce9a4c2f37c

22. It's Not Just Size That Matters: Small Language Models Are Also Few-Shot Learners
Timo Schick and Hinrich Schütze

https://arxiv.org/abs/2009.07118

23. Beyond Accuracy: Behavioral Testing of NLP Models with CheckList
Marco Tulio Ribeiro, Tongshuang Wu, Carlos Guestrin and Sameer Singh
https://arxiv.org/abs/2005.04118

24. Don't Stop Pretraining: Adapt Language Models to Domains and Tasks
Suchin Gururangan, Ana Marasović, Swabha Swayamdipta, Kyle Lo, Iz Beltagy, Doug Downey and Noah A. Smith

25. Tangled up in BLEU: Reevaluating the Evaluation of Automatic Machine Translation Evaluation Metrics
Nitika Mathur, Timothy Baldwin and Trevor Cohn
https://arxiv.org/abs/2004.10964

26.Attention Is All You Need
Ashish Vaswani, Noam Shazeer, Niki Parmar, Jakob Uszkoreit, Llion Jones, Aidan N. Gomez, Lukasz Kaiser, Illia Polosukhin
https://arxiv.org/abs/1706.03762

27. How Can We Accelerate Progress Towards Human-like Linguistic Generalization?
Tal Linzen
https://arxiv.org/abs/2005.00955

28. A Unified Approach to Interpreting Model Predictions
Scott Lundberg, Su-In Lee
https://arxiv.org/abs/1705.07874

29. Informer: Beyond Efficient Transformer for Long Sequence Time-Series Forecasting
Haoyi Zhou, Shanghang Zhang, Jieqi Peng, Shuai Zhang, Jianxin Li, Hui Xiong, Wancai Zhang
https://arxiv.org/abs/2012.07436

30. Exploration-Exploitation in Multi-Agent Learning: Catastrophe Theory Meets Game Theory

Stefanos Leonardos, Georgios Piliouras
https://arxiv.org/abs/2012.03083

31. Learning from eXtreme Bandit Feedback
Romain Lopez, Inderjit S. Dhillon, Michael Jordan
https://arxiv.org/abs/2009.12947

32. Self-Attention Attribution: Interpreting Information Interactions Inside Transformer
Yaru Hao, Li Dong, Furu Wei, Ke Xu
https://arxiv.org/abs/2004.11207

33. Mitigating Political Bias in Language Models through Reinforced Calibration
Ruibo Liu, Chenyan Jia, Jason W Wei, Guangxuan Xu, Lili Wang, Soroush Vosoughi
https://arxiv.org/abs/2104.14795

34. Fair Division of Mixed Divisible and Indivisible Goods
Xiaohui Bei, Zihao Li, Jinyan Liu, Shengxin Liu, Xinhang Lu
https://arxiv.org/abs/1911.07048

35. Lifelong Learning with a Changing Action Set
Yash Chandak, Georgios Theocharous, Chris Nota, Philip S. Thomas
https://arxiv.org/abs/1906.01770

36. Unsupervised Data Augmentation For Consistency Training.
Qizhe Xie, Zihang Dai, Eduard Hovy, Minh-Thang Luong, Quoc V. Le
https://arxiv.org/abs/1904.12848

37. RandAugment: Practical Automated Data Augmentation With A Reduced Search Space.
Ekin D. Cubuk, Barret Zoph, Jonathon Shlens, Quoc V. Le
https://arxiv.org/abs/1909.13719

38. On Adaptive Attacks To Adversarial Example Defenses
Florian Tramer, Nicholas Carlini, Wieland Brendel, Aleksander Madry
https://arxiv.org/abs/2002.08347

39. Big Self-Supervised Models Are Strong Semi-Supervised Learners
Ting Chen, Simon Kornblith, Kevin Swersky, Mohammad Norouzi, Geoffrey Hinton
https://arxiv.org/abs/2006.10029

40. XLNet: Generalized Autoregressive Pretraining For Language Understanding
Zhilin Yang, Zihang Dai, Yiming Yang, Jaime Carbonell, Ruslan Salakhutdinov, Quoc V. Le
https://arxiv.org/abs/1906.08237

41. Cross-lingual Language Model Pretraining
Guillaume Lample, Alexis Conneau
https://arxiv.org/abs/1901.07291