PAPER LIST

November 7: Training Deep Neural Networks
Improving Sharpness-Aware Minimization by Lookahead
Runsheng Yu, Youzhi Zhang, James Kwok. ICML 2024.
https://github.com/michaelrzhang/lookahead

November 14: Convolutional Neural Networks
InceptionNeXt: When Inception Meets ConvNeXt
Weihao Yu, Pan Zhou, Shuicheng Yan, Xinchao Wang. CVPR 2024.
https://github.com/sail-sg/inceptionnext

November 21: Understanding and Visualizing CNNs
Saliency strikes back: How filtering out high frequencies improves white-box explanations
Sabine Muzellec, Thomas Fel, Victor Boutin, L´eo and´eol, Rufin VanRullen, Thomas Serre. ICML 2024.
https://github.com/sukrutrao/Attribution-Evaluation

November 28: Recurrent Neural Networks
Were RNNs All We Needed?
Leo Feng, Frederick Tung, Mohamed Osama Ahmed, Yoshua Bengio, Hossein Hajimirsadegh. arXiv 2024.
https://github.com/lucidrains/minGRU-pytorch

December 5: Attention and Transformers
FasterViT: Fast Vision Transformers with Hierarchical Attention
Ali Hatamizadeh, Greg Heinrich, Hongxu Yin, Andrew Tao, Jose M. Alvarez, Jan Kautz, Pavlo Molchanov. ICLR 2024.
https://github.com/NVlabs/FasterViT

December 12: Graph Neural Networks
Graph Neural Networks for Learning Equivariant Representations of Neural Networks
Miltiadis Kofinas, Boris Knyazev, Yan Zhang, Yunlu Chen, Gertjan J. Burghouts, Efstratios Gavves, Cees G. M. Snoek, David W. Zhang. ICLR 2024.
https://github.com/mkofinas/neural-graphs

December 19: *No Paper Presentations – Project Progress Presentations*

December 26: Language Model Pretraining
Making Pre-trained Language Models Great on Tabular Prediction
Jiahuan Yan, Bo Zheng, Hongxia Xu, Yiheng Zhu, Danny Chen, Jimeng Sun, Jian Wu, Jintai Chen. ICLR 2024.
https://github.com/jyansir/tp-berta

January 2: Large Language Models
LLM Augmented LLMs: Expanding Capabilities Through Composition
Rachit Bansal, Bidisha Samanta, Siddharth Dalmia, Nitish Gupta, Sriram Ganapathy, Abhishek Bapna, Prateek Jain, Partha Talukdar. ICLR 2024.
https://github.com/google-deepmind/calm

January 9: Efficient LLMs
GaLore: Memory-Efficient LLM Training by Gradient Low-Rank Projection
Jiawei Zhao, Zhenyu Zhang, Beidi Chen, Zhangyang Wang, Anima Anandkumar, Yuandong Tian. ICML 2024.
https://github.com/jiaweizzhao/GaLore