Papers timeline

JR

March 13, 2025

Contents

1 2025

1.1 March

1.2 February

1.3 January

[1] DeepSeek-AI *et al.*, *DeepSeek-R1: Incentivizing Reasoning Capability in LLMs via Reinforcement Learning*, arXiv:2501.12948 [cs], Jan. 2025. DOI: `10.48550/arXiv.2501.12948`. [Online]. Available: `http://arxiv.org/abs/2501.12948` (visited on 03/14/2025).

2 2024

2.1 December

[1] Qwen *et al.*, *Qwen2.5 Technical Report*, arXiv:2412.15115 [cs], Jan. 2025. DOI: `10.48550/arXiv.2412.15115`. [Online]. Available: `http://arxiv.org/abs/2412.15115` (visited on 03/14/2025).

2.2 September

[1] N. Muennighoff *et al.*, *OLMoE: Open Mixture-of-Experts Language Models*, arXiv:2409.02060 [cs], Mar. 2025. DOI: `10.48550/arXiv.2409.02060`. [Online]. Available: `http://arxiv.org/abs/2409.02060` (visited on 03/14/2025).

2.3 June

[1] G. Penedo *et al.*, *The FineWeb Datasets: Decanting the Web for the Finest Text Data at Scale*, arXiv:2406.17557 [cs], Oct. 2024. DOI: `10.48550/arXiv.2406.17557`. [Online]. Available: `http://arxiv.org/abs/2406.17557` (visited on 03/14/2025).

3 2019

3.1 February

[1] A. Radford, J. Wu, R. Child, D. Luan, D. Amodei, and I. Sutskever, "Language Models are Unsupervised Multitask Learners," 2019. [Online]. Available: `https://www.semanticscholar.org/paper/Language-Models-are-Unsupervised-Multitask-Learners-Radford-Wu/9405cc0d6169988371b2755e573cc28650d14dfe` (visited on 03/14/2025).

4 2018

4.1 June

[1] A. Radford, K. Narasimhan, T. Salimans, and I. Sutskever, "Improving Language Understanding by Generative Pre-Training," en,

5 2017