

- [15] Ziwei Ji, Tiezheng Yu, Yan Xu, Nayeon Lee, Etsuko Ishii, and Pascale Fung. 2023. Towards Mitigating Hallucination in Large Language Models via Self-Reflection. arXiv:2310.06271 [cs.CL] <https://arxiv.org/abs/2310.06271>
- [16] Zhengyao Jiang, Dixing Xu, and Jinjun Liang. 2017. A Deep Reinforcement Learning Framework for the Financial Portfolio Management Problem. arXiv:1706.10059 [q-fin.CP] <https://arxiv.org/abs/1706.10059>
- [17] Guolin Ke, Qi Meng, Thomas Finley, Taifeng Wang, Wei Chen, Weidong Ma, Qiwei Ye, and Tie-Yan Liu. 2017. LightGBM: A Highly Efficient Gradient Boosting Decision Tree. In *Advances in Neural Information Processing Systems*, I. Guyon, U. Von Luxburg, S. Bengio, H. Wallach, R. Fergus, S. Vishwanathan, and R. Garnett (Eds.), Vol. 30. Curran Associates, Inc. https://proceedings.neurips.cc/paper_files/paper/2017/file/6449f44a102fde848669bdd9eb6b76fa-Paper.pdf
- [18] Kemal Kirtac and Guido Germano. 2024. Sentiment trading with large language models. *Finance Research Letters* 62 (2024), 105227. <https://doi.org/10.1016/j.frl.2024.105227>
- [19] Kelvin J.L. Koa, Yunshan Ma, Ritchie Ng, and Tat-Seng Chua. 2024. Learning to Generate Explainable Stock Predictions using Self-Reflective Large Language Models. In *Proceedings of the ACM on Web Conference 2024 (WWW '24)*. ACM. <https://doi.org/10.1145/3589334.3645611>
- [20] Thanos Konstantinidis, Giorgos Iacovides, Mingxue Xu, Tony G. Constantinides, and Danilo Mandic. 2024. FinLlama: Financial Sentiment Classification for Algorithmic Trading Applications. arXiv:2403.12285 [cs.CL] <https://arxiv.org/abs/2403.12285>
- [21] Kausik Lakkaraju, Sara E Jones, Sai Krishna Revanth Vuruma, Vishal Pallagani, Bharath C Muppasani, and Biplav Srivastava. 2023. LLMs for Financial Advice: A Fairness and Efficacy Study in Personal Decision Making. In *Proceedings of the Fourth ACM International Conference on AI in Finance* (Brooklyn, NY, USA) (ICAIF '23). Association for Computing Machinery, New York, NY, USA, 100–107. <https://doi.org/10.1145/3604237.3626867>
- [22] Harrison Lee, Samrat Phatale, Hassan Mansoor, Thomas Mesnard, Johan Ferret, Kellie Lu, Colton Bishop, Ethan Hall, Victor Carbune, Abhinav Rastogi, and Sushant Prakash. 2023. RLAI: Scaling Reinforcement Learning from Human Feedback with AI Feedback. arXiv:2309.00267 [cs.CL] <https://arxiv.org/abs/2309.00267>
- [23] Jean Lee, Nicholas Stevens, Soyeon Caren Han, and Minseok Song. 2024. A Survey of Large Language Models in Finance (FinLLMs). arXiv:2402.02315 [cs.CL] <https://arxiv.org/abs/2402.02315>
- [24] Yinheng Li. 2023. A Practical Survey on Zero-shot Prompt Design for In-context Learning. In *Proceedings of the Conference Recent Advances in Natural Language Processing - Large Language Models for Natural Language Processings (RANLP)*. INCOMA Ltd., Shoumen, BULGARIA. https://doi.org/10.26615/978-954-452-092-2_069
- [25] Yinheng Li, Junhao Wang, and Yijie Cao. 2023. A General Framework on Enhancing Portfolio Management with Reinforcement Learning. arXiv:1911.11880 [q-fin.PM] <https://arxiv.org/abs/1911.11880>
- [26] Yinheng Li, Shaofei Wang, Han Ding, and Hang Chen. 2024. Large Language Models in Finance: A Survey. arXiv:2311.10723 [q-fin.GN] <https://arxiv.org/abs/2311.10723>
- [27] Yang Li, Yangyang Yu, Haohang Li, Zhi Chen, and Khaldoun Khashanah. 2023. TradingGPT: Multi-Agent System with Layered Memory and Distinct Characters for Enhanced Financial Trading Performance. arXiv:2309.03736 [q-fin.PM] <https://arxiv.org/abs/2309.03736>
- [28] Haotian Liu, Chunyuan Li, Qingyang Wu, and Yong Jae Lee. 2023. Visual Instruction Tuning. arXiv:2304.08485 [cs.CV] <https://arxiv.org/abs/2304.08485>
- [29] Alejandro Lopez-Lira and Yuehua Tang. 2023. Can ChatGPT Forecast Stock Price Movements? Return Predictability and Large Language Models. arXiv:2304.07619 [q-fin.ST] <https://arxiv.org/abs/2304.07619>
- [30] Gilles Louppe. 2015. Understanding Random Forests: From Theory to Practice. arXiv:1407.7502 [stat.ML] <https://arxiv.org/abs/1407.7502>
- [31] Dat Mai. 2024. StockGPT: A GenAI Model for Stock Prediction and Trading. arXiv:2404.05101 [q-fin.CP] <https://arxiv.org/abs/2404.05101>
- [32] Nidhi Mehandru, Brian Y. Miao, Eric R. Almaraz, et al. 2024. Evaluating large language models as agents in the clinic. *npj Digital Medicine* 7, 84 (2024). <https://doi.org/10.1038/s41746-024-01083-y>
- [33] Ben Mezrich. 2021. *The Antisocial Network: The GameStop Short Squeeze and the Ragtag Group of Amateur Traders That Brought Wall Street to Its Knees*. Grand Central Publishing, New York. <https://www.hachettebookgroup.com/titles/ben-mezrich/the-antisocial-network/9781549112591/?lens=grand-central-publishing>
- [34] Volodymyr Mnih, Koray Kavukcuoglu, David Silver, Alex Graves, Ioannis Antonoglou, Daan Wierstra, and Martin Riedmiller. 2013. Playing Atari with Deep Reinforcement Learning. arXiv:1312.5602 [cs.LG] <https://arxiv.org/abs/1312.5602>
- [35] OpenAI. 2023. GPT-4V(ision) System Card. <https://api.semanticscholar.org/CorpusID:263218031>
- [36] OpenAI. 2024. GPT-4 Technical Report. arXiv:2303.08774 [cs.CL] <https://arxiv.org/abs/2303.08774>
- [37] Long Ouyang, Jeff Wu, Xu Jiang, Diogo Almeida, Carroll L. Wainwright, Pamela Mishkin, Chong Zhang, Sandhini Agarwal, Katarina Slama, Alex Ray, John Schulman, Jacob Hilton, Fraser Kelton, Luke Miller, Maddie Simens, Amanda Askell, Peter Welinder, Paul Christiano, Jan Leike, and Ryan Lowe. 2022. Training language models to follow instructions with human feedback. arXiv:2203.02155 [cs.CL] <https://arxiv.org/abs/2203.02155>
- [38] Joon Sung Park, Joseph C. O'Brien, Carrie J. Cai, Meredith Ringel Morris, Percy Liang, and Michael S. Bernstein. 2023. Generative Agents: Interactive Simulacra of Human Behavior. arXiv:2304.03442 [cs.HC] <https://arxiv.org/abs/2304.03442>
- [39] Joon Sung Park, Joseph C. O'Brien, Carrie J. Cai, Meredith Ringel Morris, Percy Liang, and Michael S. Bernstein. 2023. Generative Agents: Interactive Simulacra of Human Behavior. arXiv:2304.03442 [cs.HC] <https://arxiv.org/abs/2304.03442>
- [40] Anubhav Sarkar, Swagata Chakraborty, Sohoni Ghosh, and Sudip Kumar Naskar. 2023. Evaluating Impact of Social Media Posts by Executives on Stock Prices. In *Proceedings of the 14th Annual Meeting of the Forum for Information Retrieval Evaluation* (Kolkata, India) (FIRE '22). Association for Computing Machinery, New York, NY, USA, 74–82. <https://doi.org/10.1145/3574318.3574339>
- [41] Jérémy Scheurer, Mikita Balesni, and Marius Hobbhahn. 2024. Large Language Models can Strategically Deceive their Users when Put Under Pressure. In *ICLR 2024 Workshop on Large Language Model (LLM) Agents*.
- [42] Sander Schulhoff, Michael Ilie, Nishant Bolepur, Konstantine Kahadze, Amanda Liu, Chenglei Si, Yinheng Li, Aayush Gupta, Hyojung Han, Sevien Schulhoff, Pranav Sandeep Dulepet, Saurav Vidyadhara, Dayeon Ki, Sweta Agrawal, Chau Pham, Gerson Kroiz, Feilean Li, Hudson Tao, Ashay Srivastava, Hevander Da Costa, Saloni Gupta, Megan L. Rogers, Inna Goncearenco, Giuseppe Sarli, Igor Galyuker, Denis Peskoff, Marine Carpuat, Jules White, Shyamal Anadkat, Alexander Hoyle, and Philip Resnik. 2024. The Prompt Report: A Systematic Survey of Prompting Techniques. arXiv:2406.06608 [cs.CL] <https://arxiv.org/abs/2406.06608>
- [43] John Schulman, Filip Wolski, Prafulla Dhariwal, Alec Radford, and Oleg Klimov. 2017. Proximal Policy Optimization Algorithms. arXiv:1707.06347 [cs.LG] <https://arxiv.org/abs/1707.06347>
- [44] William F. Sharpe. 1966. Mutual Fund Performance. *The Journal of Business* 39, 1 (1966), 119–138. <https://finance.martinswell.com/fund-performance/Sharpe1966.pdf>
- [45] S&P Dow Jones Indices. 2024. S&P 500 Index. <https://www.spglobal.com/spdji/en/indices/equity/sp-500/> Accessed: 2024-07-10.
- [46] Lei Wang, Chen Ma, Xueyang Feng, Zeyu Zhang, Hao Yang, Jingsen Zhang, Zhiyuan Chen, Jiakai Tang, Xu Chen, Yankai Lin, Wayne Xin Zhao, Zhewei Wei, and Jirong Wen. 2024. A survey on large language model based autonomous agents. *Frontiers of Computer Science* 18, 6 (March 2024). <https://doi.org/10.1007/s11704-024-40231-1>
- [47] Meiyun Wang, Kiyoshi Izumi, and Hiroki Sakaji. 2024. LLMFactor: Extracting Profitable Factors through Prompts for Explainable Stock Movement Prediction. arXiv:2406.10811 [cs.CL] <https://arxiv.org/abs/2406.10811>
- [48] Saizhuo Wang, Hang Yuan, Lionel M. Ni, and Jian Guo. 2024. QuantAgent: Seeking Holy Grail in Trading by Self-Improving Large Language Model. arXiv:2402.03755 [cs.AI] <https://arxiv.org/abs/2402.03755>
- [49] Saizhuo Wang, Hang Yuan, Leon Zhou, Lionel M Ni, Heung-Yeung Shum, and Jian Guo. 2023. Alpha-gpt: Human-ai interactive alpha mining for quantitative investment. *arXiv preprint arXiv:2308.00016* (2023).
- [50] Ruoxu Wu. 2024. Portfolio Performance Based on LLM News Scores and Related Economical Analysis. (January 30 2024). <https://ssrn.com/abstract=4709617>
- [51] Frank Xing. 2024. Designing Heterogeneous LLM Agents for Financial Sentiment Analysis. arXiv:2401.05799 [cs.CL] <https://arxiv.org/abs/2401.05799>
- [52] Aiyuan Yang, Bin Xiao, Bingning Wang, Borong Zhang, Ce Bian, Chao Yin, Chenxu Lv, Da Pan, Dian Wang, Dong Yan, Fan Yang, Fei Deng, Feng Wang, Feng Liu, Guangwei Ai, Guosheng Dong, Haizhou Zhao, Hang Xu, Haoze Sun, Hongda Zhang, Hui Liu, Jiaming Ji, Jian Xie, JunTao Dai, Kun Fang, Lei Su, Liang Song, Lifeng Liu, Liyun Ru, Luyao Ma, Mang Wang, Mickel Liu, MingAn Lin, Nuolan Nie, Peidong Guo, Ruiyang Sun, Tao Zhang, Tianpeng Li, Tianyu Li, Wei Cheng, Weipeng Chen, Xiangrong Zeng, Xiaochuan Wang, Xiaoxi Chen, Xin Men, Xin Yu, Xuehai Pan, Yanjun Shen, Yiding Wang, Yiyu Li, Youxin Jiang, Yuchen Gao, Yupeng Zhang, Zenan Zhou, and Zhiying Wu. 2023. Baichuan 2: Open Large-scale Language Models. arXiv:2309.10305 [cs.CL] <https://arxiv.org/abs/2309.10305>
- [53] Yangyang Yu, Haohang Li, Zhi Chen, Yuechen Jiang, Yang Li, Denghui Zhang, Rong Liu, Jordan W. Suchow, and Khaldoun Khashanah. 2023. FinMem: A Performance-Enhanced LLM Trading Agent with Layered Memory and Character Design. arXiv:2311.13743 [q-fin.CP] <https://arxiv.org/abs/2311.13743>
- [54] Chong Zhang, Xinyi Liu, Mingyu Jin, Zhongmou Zhang, Lingyao Li, Zhengting Wang, Wenyue Hua, Dong Shu, Suiyuan Zhu, Xiaobo Jin, Sujian Li, Mengnan Du, and Yongfeng Zhang. 2024. When AI Meets Finance (StockAgent): Large Language Model-based Stock Trading in Simulated Real-world Environments. <https://doi.org/10.13140/RG.2.2.29976.20489>
- [55] Feng Zhang, Ruite Guo, and Honggao Cao. 2020. Information Coefficient as a Performance Measure of Stock Selection Models. arXiv:2010.08601 [q-fin.CP] <https://arxiv.org/abs/2010.08601>

- [56] Haohan Zhang, Fengrui Hua, Chengjin Xu, Hao Kong, Ruiting Zuo, and Jian Guo. 2024. Unveiling the Potential of Sentiment: Can Large Language Models Predict Chinese Stock Price Movements? arXiv:2306.14222 [cs.CL] <https://arxiv.org/abs/2306.14222>
- [57] Wentao Zhang, Lingxuan Zhao, Haochong Xia, Shuo Sun, Jiaze Sun, Molei Qin, Xinyi Li, Yuqing Zhao, Yilei Zhao, Xinyu Cai, Longtao Zheng, Xinrun Wang, and Bo An. 2024. A Multimodal Foundation Agent for Financial Trading: Tool-Augmented, Diversified, and Generalist. arXiv:2402.18485 [q-fin.TR] <https://arxiv.org/abs/2402.18485>
- [58] Xue Zhang, Hauke Fuehres, and Peter A. Gloor. 2011. Predicting Stock Market Indicators Through Twitter “I hope it is not as bad as I fear”. *Procedia - Social and Behavioral Sciences* 26 (2011), 55–62. <https://doi.org/10.1016/j.sbspro.2011.10.562>
- [59] Zheyuan Zhang, Daniel Zhang-Li, Jifan Yu, Linlu Gong, Jinchang Zhou, Zhiyuan Liu, Lei Hou, and Juanzi Li. 2024. Simulating Classroom Education with LLM-Empowered Agents. arXiv:2406.19226 [cs.CL] <https://arxiv.org/abs/2406.19226>
- [60] Qichang Zheng. 2024. *How can we use ChatGPT better: A research of API-Enhanced ChatGPT in stock prediction*. M.A. Thesis. University of Chicago. Advisor(s) Joseph Hardwick. <https://doi.org/10.6082/uchicago.11888> Record appears in Social Sciences Division > Computational Social Sciences (MACSS), Social Sciences Division > MA Thesis Archive.

Received 26 July 2024