# Research Statement

Yihan Du

I am broadly interested in the area of machine learning, including online learning (in particular, multi-armed bandit), reinforcement learning (RL), representation learning and meta learning. My research mainly focuses on designing both computationally and statistically efficient bandit and RL algorithms, establishing rigorous theoretical guarantees, and expanding bandit and RL theories.

During my Ph.D. study, I published 7 $1^{st}$-author papers (including one with alphabetical order) at the top conferences in machine leaning, including ICML [1, 2], NeurIPS [4, 6], AAAI [5, 8] and AAMAS [7], and currently have 2 $1^{st}$-author papers under submission [3, 9]. Motivated by real-word decision applications with complex decision space and rigid risk requirements, my research [1, 2, 3, 4, 5, 6, 7, 8, 9] focuses on three important branches of online decision making (online learning and RL), i.e., risk-aware online decision making, online decision making with combinatorial action space and preference-based online decision making.

Despite my accomplishments in online learning and RL, there are still many important fields of machine learning that I am eager to further explore. Below I present three future directions that I plan to further investigate.

- **Safety in Machine Learning.** With increasing demands of the practicability of machine learning algorithms in real-world applications such as autonomous driving and robotics, safety has attracted a large amount of attention in the machine learning community. For example, how can we prevent robots from destroying surrounding environments and hurting workers during work due to control or decision failures, and how can we ensure the safety of passengers when deploying RL-based algorithms in real-world autonomous driving? I am interested in building theoretical foundations for the safety of machine learning and employ theoretical findings to inspire the design of practical methods.

- **Representation Learning and Meta Learning.** Representation learning and meta learning have obtained great empirical success in many fields of machine learning, including computer vision and natural language processing. However, theoretical understandings and supports behind such empirical success are still underdeveloped, e.g., how large benefits learning expressive representations (meta) will bring to RL. I am interested in establishing theoretical evidence and quantification for the power of representation learning and meta learning in practical applications.

- **Applications of Machine Learning.** Besides contributing to theoretical foundations of machine learning, I am also eager to combine theoretical findings and practical applications, to design practical approaches that both possess optimal statistical guarantees and achieve state-of-the-art empirical performance. For example, I am interested in making use of theoretical findings in online learning and RL to facilitate the development of real-world learning applications, such as clinical trials, autonomous deriving, robotics and finance.

## References

[1] Wei Chen, Yihan Du, Longbo Huang, and Haoyu Zhao. Combinatorial pure exploration for dueling bandit. In *International Conference on Machine Learning (ICML)*, pages 1531–1541, 2020.

[2] Yihan Du and Wei Chen. Branching reinforcement learning. *International Conference on Machine Learning (ICML)*, 2022.

[3] Yihan Du, Wei Chen, Yuko Kuroki, and Longbo Huang. Collaborative pure exploration in kernel bandit. *arXiv preprint arXiv:2110.15771*, 2021.

[4] Yihan Du, Yuko Kuroki, and Wei Chen. Combinatorial pure exploration with bottleneck reward function. *Advances in Neural Information Processing Systems (NeurIPS)*, 34:23956–23967, 2021.

[5] Yihan Du, Yuko Kuroki, and Wei Chen. Combinatorial pure exploration with full-bandit or partial linear feedback. In *Proceedings of the AAAI Conference on Artificial Intelligence (AAAI)*, volume 35, pages 7262–7270, 2021.

[6] Yihan Du, Siwei Wang, Zhixuan Fang, and Longbo Huang. Continuous mean-covariance bandits. *Advances in Neural Information Processing Systems (NeurIPS)*, 34:875–886, 2021.

[7] Yihan Du, Siwei Wang, and Longbo Huang. Dueling bandits: From two-dueling to multi-dueling. In *Proceedings of the International Conference on Autonomous Agents and Multiagent Systems (AAMAS)*, pages 348–356, 2020.

[8] Yihan Du, Siwei Wang, and Longbo Huang. A one-size-fits-all solution to conservative bandit problems. In *Proceedings of the AAAI Conference on Artificial Intelligence (AAAI)*, volume 35, pages 7254–7261, 2021.

[9] Yihan Du, Siwei Wang, and Longbo Huang. Risk-sensitive reinforcement learning: Iterated CVaR and the worst path. *arXiv preprint arXiv:2206.02678*, 2022.