Research Statement

Sudeep Salgia

My research focuses on **provably efficient** and **resource-aware** data-driven decision-making in sequential learning problems with applications in healthcare, food industry, training foundation models, and robotics. I am particularly interested in studying the interplay between the efficiency of decision-making and practical constraints and systemic challenges such as communication overhead, with a focus on developing strategies that offer collateral qualitative benefits like interpretability that enhance practical deployability.

The past two decades have witnessed an unprecedented increase in the availability and collection of data. This has stoked the development of data-driven solutions to problems across a variety of scientific disciplines. The ever-increasing data, while crucial to the feasibility of data-driven solutions, gives rise to new challenges in designing practically deployable solutions for real-world problems.

Systemic Constraints. The recent explosion in the availability of data has necessitated its decentralization for ease of storage. This decentralization has fueled the need to design collaborative decision-making strategies. The design of such strategies is often dictated by **communication constraints imposed by systemic bandwidth limitations**. Secondly, a significant portion of data generated today comes from personal devices (cell phones, smartwatches, etc.) and thus contains sensitive information that needs to be protected. Hence, it is imperative to develop solutions that **preserve the privacy of the underlying data**. A third constraint for practically effective solutions is their computational cost, specifically when they are to be designed for deployment on **local devices with limited computational resources**.

Enhancing Practical Deployability. Classical learning algorithms often assume access to problem-specific auxiliary knowledge such as function properties and noise models. Such knowledge not only is integral to their design but also lends an intrinsic explainability to their decision strategies. However, acquiring such knowledge in today's big data regime is practically infeasible. Thus, it is desirable to develop strategies that **adapt to the underlying problem instance**. Furthermore, we also need new techniques to **efficiently extract intelligible information** from raw data to help steer **interpretable decision-making**.

Rooted in theory, and inspired by real-world applications and challenges, my research [1–17] is geared towards facilitating effective data-driven decision-making arising in the fields of **Reinforcement Learning, Distributed Learning,** and **Stochastic Optimization** which find applications in disciplines like **healthcare and robotics** and even in **recent advancements in AI** like alignment of foundation models and distributed training of modern day models. My research agenda consists of two main thrusts:

- Establish lower bounds on feasible performance in order to identify and characterize the key factors governing the trade-off between learning efficiency and practical and systemic constraints;
- Design provably optimal algorithms that methodically encapsulate the identified key trade-off factors and offer qualitative benefits such as interpretability that enhance practical deployability.

Thus, through a focus on lower bounds, efficient algorithms, and resource awareness, my research offers a **unique perspective into decision making and machine learning** that is based on a **blend of information-theoretic, statistical, and systemic design** aspects. My research employs tools from a wide array of fields including **high-dimensional statistics, probability theory, information theory, optimization, and machine learning**. Below, I highlight my research contributions across three threads.

Communication Efficient Distributed and Federated Learning

Distributed and Federated Learning (FL) refer to collaborative approaches where data from several decentralized sources, referred to as the participating agents, are used for decision-making to achieve a

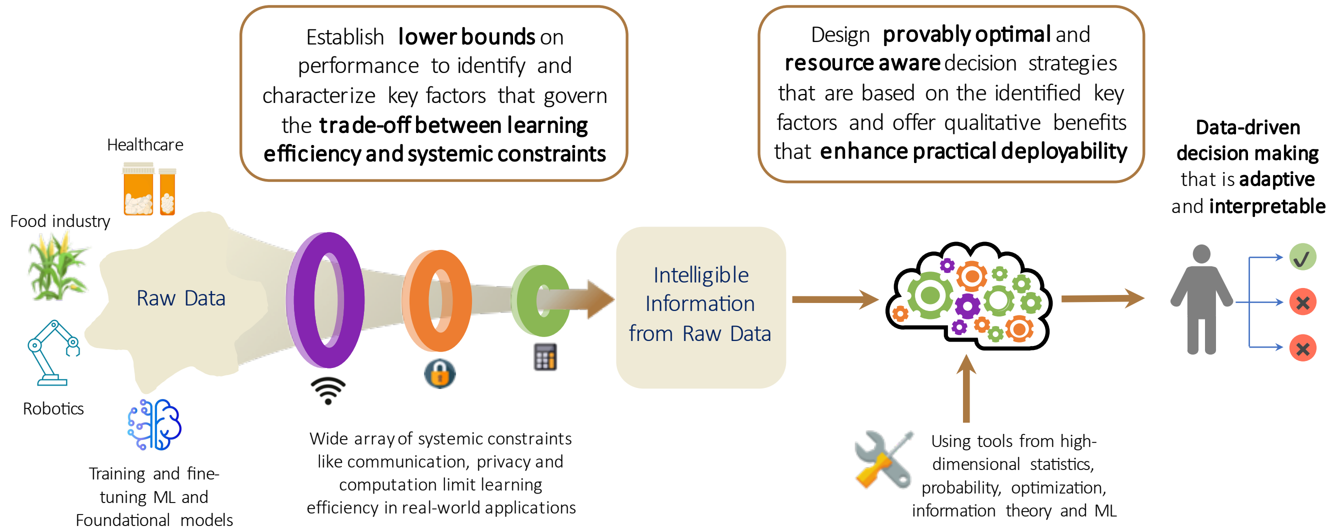

Figure 1: A pictorial representation of a typical pipeline in data-driven decision making. My research agenda, highlighted in the top two boxes, aims at studying both the main steps in the pipeline — efficiently extracting information from raw data and translating the acquired information into interpretable strategies.

common objective. The ability to communicate and share information in distributed learning is an integral aspect of collaboratively working towards a common objective. In recent years, FL has evolved as the de facto approach for **training large scale machine learning models**. There is a natural tension between learning efficiency and communication efficiency in distributed learning. This is because achieving higher learning efficiency, i.e. lower error rate, entails the participating agents to share more information which results in higher communication costs. There is an extensive line of work aimed at designing communication efficient algorithms [18–23], however, principled studies that rigorously characterize the accuracy-communication trade-off are lacking.

Key Takeaways. In a series of studies [1, 3, 4, 9, 11], we rigorously characterized the trade-off between learning and communication efficiency for a wide range of distributed learning problems of increasing complexity — linear bandits, stochastic optimization, kernelized bandits, Q-learning. Each study identifies the **key aspects that determine the fundamental trade-off between accuracy and communication** thereby providing insights into how the trade-off changes with different problems. We also **designed novel algorithms that operate at the optimal frontier** of the trade-off characterized in the previous step. Our latest work on Q-learning was **accepted as a NeurIPS oral presentation**.

Technical Contributions. For linear bandits and stochastic convex optimization, we derived the **first information-theoretic lower bounds** on communication complexity of $\Omega(d)$ and $\Omega(d^2)$ bits respectively required to achieve optimal accuracy (d denotes the dimension of the decision variable). We proposed a novel **interpretable** and **adaptive** algorithm for distributed linear bandits [9] based on **progressive learning and sharing** of the unknown reward vector — one bit at a time. For stochastic convex optimization, our algorithms [3, 4] build upon **classic plane cutting methods** and a **novel adaptive mean estimation routine** to achieve optimal performance. In [1], we showed that the communication complexity of Federated Q-learning (reinforcement learning) is inherently **determined by the interplay between the bias and variance of Q-learning updates**. Based on these insights, we designed the **first** Federated Q-learning algorithm that achieves provably optimal learning and communication complexity.

Computationally Efficient High-Dimensional Inference

Reproducing Kernel Hilbert Spaces (RKHS) offers a highly expressive modeling framework for **high-dimensional black-box functions with low-rank structure** that: (i) finds applications in **experiment design, hyperparameter tuning and scientific simulations**; and (ii) serves as a bridge to **better understand neural networks** using the NTK theory [24]. As a result, optimization of RKHS functions, a.k.a., kernelized bandits, has received significant traction in recent years. Popular query strategies for kernel bandits involve a computationally expensive auxiliary optimization of a non-convex acquisition function at each time step which leads to high computational requirements.

Key Takeaways. We proposed two novel algorithms for kernelized bandits that are not only the **first algorithms to achieve order-optimal learning efficiency in noisy and noise-free settings** respectively, but also **significantly improve upon the computational costs** over state-of-the-art algorithms. Such improved computational efficiency is particularly beneficial for applications in **swarm robotics and distributed hyperparameter tuning of large neural networks** due to limited computing power at individual nodes. The first algorithm [13] is based on a **computationally efficient tree-based domain shrinking methodology** to quickly identify subsets of the domain with high function values. The second algorithm [2] completely **bypasses the auxiliary optimization strategy** and replaces it with an open-loop, computationally inexpensive query strategy of sampling a point uniformly at random from the domain. Our novel analytical results in this work **resolved an open COLT problem** [25].

Technical Contributions. The acquisition function is typically optimized using a grid search over the domain with a grid size that **scales polynomially in the number of query points**. In our tree-based domain shrinking algorithm, we eliminate low-performing nodes (subsets of the domain) by performing the auxiliary optimization in each node up to a precision that is carefully chosen based on its depth. This allowed us to **use a constant-sized grid** at each time step and **significantly reduce the computational cost**. In the second algorithm, we proved the **equivalence between our open-loop strategy and the best closed-loop strategy** and the optimality of our proposed algorithm in both noisy and noise-free settings. We established these results through a **novel concentration result for infinite-dimensional covariance operators** in an RKHS.

Adapting to Underlying Problem Instances

A desirable perquisite for practically deployable algorithms is **adaptivity**, i.e., the ability to adjust to underlying problem instances. It is particularly beneficial for time-sensitive applications and applications with non-stationary environments such as field experimentation and networks.

Key Takeaways. I am a key contributor to a **robust** stochastic optimization algorithm [16] that **has no tuning parameters**, **adapts** to the unknown function characteristics and achieves optimal performance. I developed the first group testing algorithm [14] that achieves optimal learning efficiency and works **for general observation models**, **adapts to the noise level** in the tests and is **agnostic to the noise distribution**. In group testing, which was **widely used during the COVID-19 pandemic**, adaptivity to noise fluctuations **offers an inherent robustness for field applications**. I also proposed a sequential uniformity testing algorithm [5] that achieves **instance-optimal sample complexity** by a novel termination condition that adapts to the distance to the uniform distribution. Such tests find extensive applications in **anomaly detection in networks and power systems**, where severe anomalies demand quicker attention.

Technical Contributions. The proposed stochastic optimization and group testing algorithms are based on a novel strategy of inducing a **random-walk over a carefully constructed graph** that separates the intermediate and final estimation stages in order to achieve order-optimal performance. For the uniform testing algorithm, we present a **novel analysis of the coincidence test based on the saddle-point method** that enables us to extend the test from offline to sequential setting.

Future Research Directions

I am enthusiastic to further my work on pushing the frontiers for resource-aware, efficient data-driven decision-making by exploring the impact of a wider class of systemic constants and practical qualitative benefits. Below, I outline several directions that I am excited to explore.

Privacy. As referred to earlier, the percolation of sensitive information into data these days has stoked the development of algorithms that systematically guarantee user privacy. In our latest work [3], we study the **three-way trade-off among privacy, communication, and accuracy** and design a novel algorithm that operates at the **jointly optimal frontier** of this trade-off. In another work [8], we investigate the design of a **differentially private algorithm for kernel bandits** that builds upon our proposed random sampling-based query strategy [2]. We also investigate the role of privacy in designing effective **collaborative solutions for food safety** [12]. Building upon these initial studies, I am keen to develop differentially private algorithms for a variety of other sequential learning problems, like RL, with a specific focus on the **connection between policy design, data collection, and privacy constraint**.

Personalization and Fairness. A common challenge in Federated Learning (FL) applications is the heterogeneity in data distributions of the agents which dampens the collaborative advantage and requires developing solutions that incorporate agent-level personalization. In our work [11], we demonstrate that in order to obtain optimal personalized performance in the long run, the agents must take **altruistic decisions at the cost of locally optimal ones**. This clearly demonstrates the need for careful algorithm design in the presence of heterogeneous agents. On the other hand, designing personalized policies for individual agents in Federated RL applications in critical sectors like **autonomous driving and clinical trials** can often be prohibitively expensive. Moreover, the policies designed for such critical application often need to pass certain regulatory requirements, posing an additional hurdle for designing fully personalized policies. Thus, it is feasible to learn only a small set of representative policies instead of a personalized policy for each agent. Two natural questions to investigate in such a setting are how to characterize the set of policies that **fairly represent the interests of the diverse set of agents**, and how to learn this set of policies. Leveraging my initial insights about personalization and Federated RL, I aim to systematically study the **interplay among personalization, fairness, and learning efficiency**.

Constrained Reinforcement Learning with Human Feedback (RLHF). RLHF has emerged as a promising new paradigm for fine-tuning large generative models. While our understanding of RLHF has progressed rapidly in a short period of time, there are several exciting directions yet to be explored, particularly in terms of understanding the theoretical underpinnings of the empirical success of RLHF. There are two directions that I am keen to explore in this field. Firstly, I want to theoretically investigate **how policy design and efficiency are affected by different fine-tuning objectives** like safety, factual correctness, and reasoning capabilities. This will allow us to develop improved models that offer superior performance across several objectives simultaneously without having to fine-tune different models for different tasks. Secondly, in light of the increasing popularity of multi-modal generative models, I intend to explore **policy designs that adapt across different data domains** like language and images.

References

[1] **Sudeep Salgia** and Yuejie Chi. The sample-communication complexity trade-off in federated q-learning. In *Proceedings of the 38th Annual Conference on Neural Information Processing Systems*, volume 37, 2024.

[2] **Sudeep Salgia**, Sattar Vakili, and Qing Zhao. Random exploration in bayesian optimization: Order-optimal regret and computational efficiency. In *Proceedings of the 41st International Conference on Machine Learning*. PMLR, 2024.

[3] **Sudeep Salgia**, Nikola Pavlovic, Yuejie Chi, and Qing Zhao. Characterizing the accuracy-communication-privacy trade-off in distributed stochastic convex optimization. *Submitted to the 28th International Conference on Aritifical Intelligence and Statistics, AISTATS*, 2024.

[4] **Sudeep Salgia**, Tamir Gabay, Qing Zhao, and Kobi Cohen. A communication-efficient adaptive algorithm for federated learning under cumulative regret. *IEEE Transactions on Signal Processing*, 2024.

[5] **Sudeep Salgia**, Xinyi Wang, Qing Zhao, and Lang Tong. Adaptive binning coincidence test for uniformity testing. *IEEE Transactions on Signal Processing*, 72:2421–2435, 2024.

[6] Nikola Pavlovic, **Sudeep Salgia**, and Qing Zhao. Order-optimal regret in distributed kernel bandits using uniform sampling with shared randomness, 2024.

[7] Tonghe Zhang, **Sudeep Salgia**, and Yuejie Chi. Distributed policy optimization under partial observability: Tractability, linear speedup, and communication efficiency. *Submitted to the 28th International Conference on Aritifical Intelligence and Statistics, AISTATS*, 2024.

[8] Nikola Pavlovic, **Sudeep Salgia**, and Qing Zhao. Differentially private kernelized contextual bandits. *Submitted to the 28th International Conference on Aritifical Intelligence and Statistics, AISTATS*, 2024.

[9] **Sudeep Salgia** and Qing Zhao. Distributed linear bandits under communication constraints. In *Proceedings of the 40th International Conference on Machine Learning, ICML*, pages 29845–29875. PMLR, 2023.

[10] **Sudeep Salgia**, Sattar Vakili, and Qing Zhao. Provably and practically efficient neural contextual bandits. In *Proceedings of the 40th International Conference on Machine Learning*, pages 29800–29844. PMLR, 2023.

[11] **Sudeep Salgia**, Sattar Vakili, and Qing Zhao. Collaborative learning in kernel-based bandits for distributed users. *IEEE Transactions on Signal Processing*, 2023.

[12] Chenhao Qian, Yuhan Liu, Cecil Barnett-Neefs, **Sudeep Salgia**, Omer Serbetci, Aaron Adalja, Jayadev Acharya, Qing Zhao, Renata Ivanek, and Martin Wiedmann. A perspective on data sharing in digital food safety systems. *Critical Reviews in Food Science and Nutrition*, 63(33):12513–12529, 2023.

[13] **Sudeep Salgia**, Sattar Vakili, and Qing Zhao. A domain-shrinking based Bayesian optimization algorithm with order-optimal regret performance. In *Proceedings of the 35th Annual Conference on Neural Information Processing Systems*, volume 34, 2021.

[14] **Sudeep Salgia** and Qing Zhao. An order-optimal adaptive test plan for noisy group testing under unknown noise models. In *ICASSP 2021 - 2021 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pages 4035–4039, 2021.

[15] **Sudeep Salgia**, Qing Zhao, and Sattar Vakili. Stochastic coordinate minimization with progressive precision for stochastic convex optimization. In *37th International Conference on Machine Learning, ICML 2020*, pages 8396–8406, Mar 2020.

[16] Sattar Vakili, **Sudeep Salgia**, and Qing Zhao. Stochastic gradient descent on a tree: an adaptive and robust approach to stochastic convex optimization. In *2019 57th Annual Allerton Conference on Communication, Control, and Computing (Allerton)*, pages 432–438. IEEE, 2019.

[17] **Sudeep Salgia** and Animesh Kumar. Bandlimited spatiotemporal field sampling with location and time unaware mobile sensors. In *2018 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pages 4574–4578, 2018.

[18] Sebastian Stich. Local SGD converges fast and communicates little. In *7th International Conference on Learning Representations, ICLR*, 2019.

[19] Farzin Haddadpour, Mohammad Mahdi Kamani, Mehrdad Mahdavi, and Viveck R Cadambe. Local SGD with periodic averaging: Tighter analysis and adaptive synchronization. In *Advances in Neural Information Processing Systems*, volume 32, 2019.

[20] Farzin Haddadpour, Mohammad Mahdi Kamani, Aryan Mokhtari, and Mehrdad Mahdavi. Federated learning with compression: Unified analysis and sharp guarantees. In *International Conference on Artificial Intelligence and Statistics*, pages 2350–2358. PMLR, 2021.

[21] Jakub Konečný, H. Brendan McMahan, Felix X Yu, Peter Richtárik, Ananda Theertha Suresh, and Dave Bacon. Federated Learning: Strategies for Improving Communication Efficiency, 2016.

[22] Zhenheng Tang, Shaohuai Shi, Xiaowen Chu, Wei Wang, and Bo Li. Communication-efficient distributed deep learning: A comprehensive survey. *arXiv preprint arXiv:2003.06307*, 2020.

[23] Seyedamin Pouriyeh, Osama Shahid, Reza M. Parizi, Quan Z. Sheng, Gautam Srivastava, Liang Zhao, and Mohammad Nasajpour. Secure Smart Communication Efficiency in Federated Learning: Achievements and Challenges. *Applied Sciences (Switzerland)*, 12(18):10996, 2022.

[24] Arthur Jacot, Franck Gabriel, and Clément Hongler. Neural tangent kernel: Convergence and generalization in neural networks. In *Proceedings of the 32nd Annual Conference on Neural Information Processing Systems*, pages 8571–8580, 2018.

[25] Sattar Vakili. Open problem: Regret bounds for noise-free kernel-based bandits. In *Proceedings of Thirty Fifth Conference on Learning Theory*, volume 178 of *Proceedings of Machine Learning Research*, pages 5624–5629. PMLR, 02–05 Jul 2022.