The Economics of Machine Learning

Haifeng Xu
Department of Computer Science
University of Chicago
haifengxu@uchicago.edu

Abstract

This survey overviews a new research agenda on the *economics of machine learning*, pursued at the **S**trategic **I**ntelli**G**ence for **M**achine **A**gent (SIGMA) Lab at UChicago. This overall research agenda has two themes: machine learning for economics and, conversely, economics for machine learning (ML). The first theme focuses on designing and analyzing ML algorithms for economic problems, ranging from foundational economic models to influential real-world applications such as recommender systems and national security. The second theme employs economic principles to study machine learning itself, such as the valuation and pricing of data, information and ML models, and designing incentive mechanisms to improve large-scale ML research peer reveiws. While our research focuses primarily on developing methodologies, in each theme we also highlight some real-world impacts of these works, including ongoing large-scale live experiments and potential deployments for various applications.

1 Machine Learning for Economics

From a technial point of view, research in machine learning can be roughly divided into two categories: learning to detect patterns and learning to act in the unknowns.[1] Our lab's research falls primarily into the second category — i.e., learning to optimize *decisions*, particularly in non-cooperative multi-agent setups with complex information environments such as asymmtric or limited access to information. Both optimization and learning in such game-theoretic problems exhibit significant challgenges due to *uncommon* knowledge among agents and potenial deceptive behaviors from opponents, and lead to facinating research questions. Next, we selectively overview some of our efforts in addressing these challenges in both foundational economic models and real world applications.

[1]Though in application, the boundaries between the two have become more and more vague nowadays since most successful technologies have to combine both (e.g. ,ChatGPT or self-driving cars).

1.1 Multi-Agent Learning in Foundational Economic Models

Dominated actions are a basic concept in game theory. It is also a natural (and perhaps the simplest possible) multi-agent generalization of sub-optimal actions as in standard single-agent decision making. Thus similar to standard bandit learning, a basic learning question in multi-agent systems is whether agents can learn to efficiently eliminate all iteratively dominated actions in an unknown game if they can only observe noisy bandit feedback about the payoff of their played actions. Surprisingly, despite a seemingly simple task, in our recent work [Wu *et al.*, 2022], we show a quite negative result; that is, standard no regret algorithms — including the entire family of Dual Averaging algorithms — provably take exponentially many rounds to eliminate all iteratively dominated actions. Moreover, algorithms with the stronger no swap regret also suffer similar exponential inefficiency. To overcome these barriers, we develop a new algorithm that adjusts Exp3 with Diminishing Historical rewards (termed Exp3-DH); Exp3-DH gradually "forgets" history at carefully tailored rates. We prove that when all agents run Exp3-DH (a.k.a., self-play in multi-agent learning), all iteratively dominated actions can be eliminated within polynomially many rounds. Our experimental results further demonstrate the efficiency of Exp3-DH, and that state-of-the-art bandit algorithms, even those developed specifically for learning in games, fail to eliminate all iteratively dominated actions efficiently

Another basic game-theoretic framework that is gaining significant recent interest in economics, CS and operation research is the *Bayesian persuasion* problem [Kamenica and Gentzkow, 2011], which captures the strategic information communication between a sender and a receiver. In a recent work [Zu *et al.*, 2021], we study a natural online learning variant of the basic Bayesian persuasion setup in a repeated setting, where at each time t, the sender observes a payoff-relevant state drawn independently and identically from an unknown prior distribution, and shares state information with the receiver, who then myopically chooses an action. As in the standard setting, the sender seeks to persuade the receiver into choosing actions that are aligned with the sender's preference by selectively sharing information about the state. However, in contrast to the standard models, the sender does not know the prior, and has to persuade while gradually learning the prior on the fly. We study the sender's learning problem of making persuasive action recommendations to achieve low regret against the optimal persuasion mechanism with the knowledge of the prior distribution. Our main positive result is an algorithm that, with high probability, is persuasive across all rounds and achieves $\sqrt{T \log T}$ regret, where T is the horizon length. The core philosophy behind the design of our algorithm is to leverage robustness against the sender's ignorance of the prior. Intuitively, at each time our algorithm maintains a set of candidate priors, and chooses a persuasion scheme that is simultaneously persuasive for all of them. To demonstrate the effectiveness of our algorithm, we further prove that no algorithm can achieve regret better than $\Omega(\sqrt{T})$, even if the persuasiveness requirements were significantly relaxed. Therefore, our algorithm achieves optimal regret for the sender's learning problem up to terms logarithmic in T.

1.2 Instantiation in Recommender Systems: Incentives and Dynamics

Multi-agent learning behaviors are ubiquitous in economic systems, particularly with sophisticated revenue-driven players. One important example is recommender systems. Content creators compete for exposure on recommendation platforms, and such strategic behavior leads to a dynamic shift over the content distribution. However, how the creators' competition impacts user welfare and how the relevance-driven recommendation influences the dynamics in the long run are still largely unknown. In our recent work [Yao *et al.*, 2023a], we provide theoretical insights into these research questions. We model the creators' competition under the assumptions that: 1) the platform employs an innocuous top-K recommendation policy; 2) user decisions follow the Random Utility model; 3) content creators compete for user engagement and, without knowing their utility function in hindsight, apply arbitrary no-regret learning algorithms to update their strategies. We study the user welfare guarantee through the lens of *Price of Anarchy* [Koutsoupias and Papadimitriou, 1999] and show that the fraction of user welfare loss due to creator competition is always upper bounded by a small constant depending on K and randomness in user decisions; we also prove the tightness of this bound. Our result discloses an intrinsic merit of the myopic approach to the recommendation, i.e., relevance-driven matching performs reasonably well in the long run, as long as users' decisions involve randomness and the platform provides reasonably many alternatives to its users.

The reward mechanisms employed by RS platforms create competition among creators which affect their production choices and, consequently, content distribution and system welfare. Following the PoA analysis in the above work, we then turn to study how to change the PoA for the better — that is, how to "pro-actively" design the platform's reward mechanism in order to steer the creators' competition towards a desirable welfare outcome in the long run. Our work [Yao *et al.*, 2023b] makes two major contributions in this regard: first, we uncover a fundamental limit about a class of widely adopted mechanisms, coined Merit-based Monotone Mechanisms, by showing that they inevitably lead to a constant fraction loss of the welfare. To circumvent this limitation, we introduce Backward Rewarding Mechanisms (BRMs) and show that the competition games resulting from BRM possess a potential game structure, which naturally induces the strategic creators' behavior dynamics to optimize any given welfare metric. In addition, the class of BRM can be parameterized so that it allows the platform to directly optimize welfare within the feasible mechanism space even when the welfare metric is not explicitly defined.

1.3 Instantiation in Adversarial Domains: Deception-Aware Learning of Equilibria

Another representative game-theoretic environment of strategic learning is to play against intelligent adversaries, with applications to border controls, national security and military settings [Tambe, 2011]. One of the key challenges in this case is that the learning from a strategic opponent – in fact an adversary in our domains — may intentionally manipulate his behaviors in order to mislead our learning. Our recent work [Nguyen and Xu, 2019] focuses on understanding how such attacker deception affects the game equilibrium. We examine a basic deception strategy termed imitative deception, in which the attacker simply pretends to have a different payoff assuming his true payoff is unknown to the defender. We provide a clean characterization about the game equilibrium as well as optimal algorithms to compute the equilibrium. In a follow-up paper, we further study how the defender can pro-actively deceive the adversary, by attempting to alter the adversary's perception of the defender's patrolling strategies so as to influence the attacker's decision making [Nguyen and Xu, 2019; Nguyen and Xu, 2022]. We are interested in understanding the complexity and effectiveness of optimal defender deception under different attacker behavior models. Specifically, we consider three different attacker strategies of response (to the defender's deception) with increasing sophistication, and design efficient polynomial-time algorithms to compute the equilibrium for each. Moreover, we prove formal separation for the effectiveness of patrol deception when facing an attacker of increasing sophistication, until it becomes even harmful to the defender when facing the most intelligent attacker we consider. Our results shed light on when and how deception should be used in adversarial domains.

Besides optimizing strategic decisions in game-theoretic problems, in many real-world situations, we may face the exact opposite of this problem – instead of prescribing equilibrium of a given game, we may directly observe the agents' equilibrium behaviors but want to infer the underlying parameters of an unknown game. This research question, also known as inverse game theory, has been studied in multiple recent works in the context of Stackelberg games. Unfortunately, existing works exhibit quite negative results, showing statistical hardness and computational hardness, assuming follower's perfectly rational behaviors. Our recent work [Wu *et al.*,] relaxes the perfect rationality agent assumption to the classic quantal response model, a more realistic behavior model of bounded rationality. Interestingly, we show that the smooth property brought by such bounded rationality model actually leads to provably more efficient learning of the follower utility parameters in general Stackelberg games.

Systematic empirical experiments on synthesized games confirm our theoretical results and further suggest its robustness beyond the strict quantal response model.

Remarks on Real-World Deployment

Besides developing methodologies, we have also been attempting to apply some of our algorithms to real-world problems in order to show how it could work in reality. For instance, in collaboration with researchers and engineers from a large social media recommendation platform, we are currently live testing the new incentive mechansims of Section 1.2 that we designed to reward content creators in order improve the system's social welfare. The initial results based on 3 weeks of live experiments show very promising performance, though the test will likely last for a few more months in order to generate more convincing statistics.

For the multi-agent learning in adversarial environments, as mentioned in Section 1.3, we are currently developing a systematic testbed for evaluating multi-agent learning algorithms in highly non-cooperative game-theoretic environmentsy. This testbed (a project we called *SigGym*) will serve similar purpose as OpenAI Gym, but with simulated complex and large-scale strategic games that are of critical importance of national security. We plan to open source our simulation environment by the end of the year and invite researchers to tackle these important challenges together.

2 Economics for Machine Learning

The impact of machine learning on our society has now grown to be so large that it starts to require systematic economic and societal studies about ML. The societal aspect of ML has already attracted extensive attention recently, including the born and popularity of new research conferences such as AISE, FaaCCT. However, the economic aspect about ML has received relatively less attention. In this survey, we will overview some of our initial studies in this space and highlight many facinating big open problems, such as how to democratize ML to make this technology accessible to small entities like small businesses or even single person and how to induce high-quality innovation in machine learning at the era of numerous publications. We argue why resolving these crucial problems will require researchers from various disciplines, and wish to bring together interdisciplinary researchers and encouraging more works into this field.

2.1 Incentive-Aware Machine Learning

Despite its significant success in recognition-style tasks, machine learning often suffers from additional challenges of being gamed when applied to non-cooperative multi-agent domains for strategic decision making, e.g., for deciding loan approval or which ad or content to recommend. To address these questions, ML algorithms need to be "incentive-aware" and study of such algorithms has attracted significant recent interest in various learning tasks (supervised vs unsupervised, online vs offline) under various situations of manipulations (adversarial vs strategic, test vs training time).

In the study of strategic manipulation of testing data for classification, most previous works have focused on two extreme situations where any testing data point either is completely adversarial or always equally prefers the positive label. Our recent work [Sundaram *et al.*, 2021] generalizes both of these through a unified framework for strategic classification and introduce the notion of strategic VC-dimension (SVC) to capture the PAC-learnability in our general strategic setup. SVC provably generalizes the recent concept of adversarial VC-dimension (AVC) introduced by [Cullina *et al.*, 2018]. We instantiate our framework for the fundamental strategic linear classification problem. We fully characterize: (1) the statistical learnability of linear classifiers by pinning down its SVC; (2) it's computational tractability by pinning down the complexity of the empirical risk minimization problem. Interestingly, the SVC of linear classifiers is always upper bounded by its standard VC-dimension. This characterization also strictly generalizes the AVC bound for linear classifiers in [Cullina *et al.*, 2018].

Another of our recent work concerns the strategic manipulation in a fundamental problem of online learning, i.e., the stochastic multi-armed bandits problem [Bubeck *et al.*, 2012]. Motivated by economic applications such as recommender systems, we study a situation where each arm is a *self-interested* strategic player who can modify its own reward whenever pulled, subject to a cross-period budget constraint, in order to maximize its own expected number of times of being pulled. We analyze the robustness of three popular bandit algorithms: UCB, ϵ-Greedy, and Thompson Sampling. We prove that all three algorithms achieve a regret upper bound $O(\max\{B, K \ln T\})$ where B is the total budget across arms, K is the total number of arms and T is the running time of the algorithms. This regret guarantee holds for *arbitrary adaptive* manipulation strategy of arms. Our second set of main results shows that this regret bound is *tight*— in fact, for UCB, it is tight even when we restrict the arms' manipulation strategies to form a *Nash equilibrium*. We do so by characterizing the Nash equilibrium of the game induced by arms' strategic manipulations and show a regret lower bound of $\Omega(\max\{B, K \ln T\})$ at the equilibrium.

2.2 Towards a Market for Data and ML

An important functionality of machine learning is to transform data to information, i.e., distilled data. Our recent works start to investigage the problem of pricing the information generated by machine learning algorithms [Chen *et al.*, 2020; Liu *et al.*, 2021], or directly pricing data [Chen *et al.*, 2022].

In [Chen *et al.*, 2020], we consider a monopoly information holder selling information to a budget-constrained decision maker, who may benefit from the seller's information. The decision maker has a utility function that depends on his action and an uncertain state of the world. The seller and the buyer each observe a private signal regarding the state of the world, which may be correlated with each other. The seller's goal is to sell her private information to the buyer and extract maximum possible revenue, subject to the buyer's budget constraints. We show that the optimal information selling mechanisms are simple in the sense that they can be naturally interpreted, have succinct representations, and can be efficiently computed. The optimal mechanism has the format of acting as a consultant who recommends the best action to the buyer but uses different and carefully designed payment rules for different settings. Our optimal mechanisms can be easily computed by solving a single polynomial-size linear program. This result significantly simplifies exponential-size LPs solved by the Ellipsoid method in the previous work, which computes the optimal mechanisms in the same setting but without budget limit. In a followup work, we characterize closed-form format of the optimal mechasnim in the special case of binary buyer actions [Liu *et al.*, 2021].

In [Chen *et al.*, 2022], we consider a new problem of selling data to a machine learner who looks to purchase data to train his machine learning model. A key challenge in this setup is that neither the seller nor the machine learner knows the true quality of data. When designing a revenue-maximizing mechanism, a data seller faces the tradeoff between the cost and precision of data quality estimation. To address this challenge, we study a natural class of mechanisms that price data via costly signaling. Motivated by the assumption of i.i.d. data points as in classic machine learning models, we first consider selling homogeneous data and derive an optimal selling mechanism. We then turn to the sale of heterogeneous data, motivated by the sale of multiple data sets, and show that 1) on the negative side, it is NP-hard to approximate the optimal mechanism within a constant ratio $\frac{e}{e+1} + o(1)$; while 2) on the positive side, there is a $1/k$ -approximate algorithm, where k is the number of the machine learner's private types.

2.3 Mechanism Design for Better ML Peer Review

In recent years, major machine learning conferences such as NeurIPS and ICML have faced a concerning decline in the quality of peer review—a development posing a significant challenge to the global machine learning community. For instance, the NeurIPS 2021 experiment highlighted that nearly half to two-thirds of accepted papers would likely face rejection if subjected to review by an alternate set of referees. This inconsistency in review outcomes was further aggravated at NeurIPS 2021, a trend partly attributable to the exponential increase in submission volumes. To mitigate this issue, there has been a progressive trend to propose various strategies aimed at enhancing the peer review process in machine learning. An emergent approach, termed the "Isotonic Mechanism", employs mechanism design to solicit private information from authors, thereby enabling more accurate estimation of review scores [Su, 2021].

Our recent work [Wu *et al.*, 2023] extends the original Isotonic Mechanism in an elegant paper by [Su, 2021] from single-owner to multiple-owner settings, in order to make it applicable to peer review where a paper often has multiple authors. Our approach starts by partitioning all submissions of a machine learning conference into disjoint blocks such that each block of submissions shares a common set of co-authors. We then employ the Isotonic Mechanism to elicit a ranking of the submissions from each author and to produce adjusted review scores that align with both the reported ranking and the original review scores. The generalized mechanism uses a weighted average of the adjusted scores on each block. We show that, under certain conditions, truth-telling is a Nash equilibrium for all authors for any valid partition of the overlapping ownership sets. While the calibration performance of the mechanism depends on the partition structure, it is computationally intractable in general to find the optimal partition. We develop a quadratic-time greedy-based algorithm that provably finds a good partition with appealing approximation guarantees. Extensive experiments on both synthetic data and real-world conference review data demonstrate the effectiveness of the proposed mechanism.

Remarks on Real-World Deployment

We have also been actively seeking to apply our methods to real-world problems. For instance, we are currently in conversations with leading ML conference organizations in applying the new generalized Isotonic mechansim. Though this deployment is less mature than those mentioned at the end of Section 1, we are hopeful and believe that our mechansims are both simply enough for real-world deployment and strong enough for guaranteeing the performance.

On the market design for ML algorithms as mentioned in Section 2.2, we are currently implementing such a data-centric market for machine-learning-as-a-service in collaboration with researchers from systems and database. Specifically, in our recent work [Galhotra *et al.*, 2023], we observe a gap in today's ML industry: many ML users can benefit from new data in possession of others whom they do not know about, whereas these data owners sit on piles of data without knowing whom can benefit from their data. This gap creates the opportunity for building a marketplace that can automatically connect supply with demand. To fill this gap, we developed new techniques to tackles two core challenges in designing such a market: (a) to efficiently match demand with supply, we develop an algorithm to automatically discover useful data for any ML task from a pool of thousands of datasets, achieving high-quality (data, ML model) matching; (b) to encourage participation from ML users, particularly those small task owners without much ML expertise, we design a carefully tailored pricing mechanism for selling data-augmented ML models. The following Figure 1 shows the structure of the marketplace we designed. Compared to existing markets like Vertex AI or Sagemaker, our pricing mechanism significantly reduces ML users' participation risk. We are currently working on deploying a prototype of this platform.

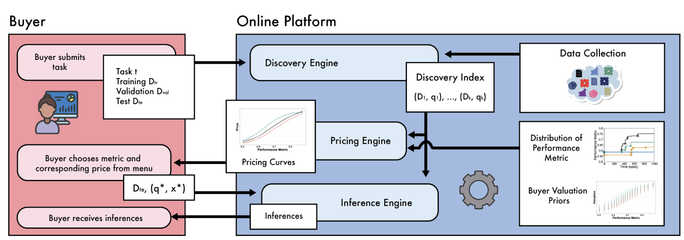

Figure 1: Overview of the Designed Data Market Architecture

Acknowledgments

Research at the SIGMA lab has been generously supported by an NSF Award CCF-2132506, an Army Research Office Award W911NF-23-1-0030, a Google Faculty Research Award, and a 3CAV seed grant.

References

[Bubeck *et al.*, 2012] Sébastien Bubeck, Nicolo Cesa-Bianchi, et al. Regret analysis of stochastic and nonstochastic multi-armed bandit problems. *Foundations and Trends® in Machine Learning*, 5(1):1–122, 2012.

[Chen *et al.*, 2020] Yiling Chen, Haifeng Xu, and Shuran Zheng. Selling information through consulting. In *Proceedings of the Fourteenth Annual ACM-SIAM Symposium on Discrete Algorithms*, pages 2412–2431. SIAM, 2020.

[Chen *et al.*, 2022] Junjie Chen, Minming Li, and Haifeng Xu. Selling data to a machine learner: Pricing via costly signaling. In *International Conference on Machine Learning*, pages 3336–3359. PMLR, 2022.

[Cullina *et al.*, 2018] Daniel Cullina, Arjun Nitin Bhagoji, and Prateek Mittal. Pac-learning in the presence of adversaries. *Advances in Neural Information Processing Systems*, 31, 2018.

[Galhotra *et al.*, 2023] Sainyam Galhotra, Minbiao Han, Jonathan Light, Steven Xia, Raul Castro Fernandez, and Haifeng Xu. A data-centric market for machine learning: From discovery to pricing. In *in submission*, 2023.

[Kamenica and Gentzkow, 2011] Emir Kamenica and Matthew Gentzkow. Bayesian persuasion. *American Economic Review*, 101(6):2590–2615, 2011.

[Koutsoupias and Papadimitriou, 1999] Elias Koutsoupias and Christos Papadimitriou. Worst-case equilibria. In *Stacs*, volume 99, pages 404–413. Springer, 1999.

[Liu *et al.*, 2021] Shuze Liu, Weiran Shen, and Haifeng Xu. Optimal pricing of information. In *Proceedings of the 22nd ACM Conference on Economics and Computation*, pages 693–693, 2021.

[Nguyen and Xu, 2019] Thanh Nguyen and Haifeng Xu. Imitative attacker deception in stackelberg security games. In *IJCAI*, pages 528–534, 2019.

[Nguyen and Xu, 2022] Thanh Nguyen and Haifeng Xu. When can the defender effectively deceive attackers in security games? In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 36, pages 9405–9412, 2022.

[Su, 2021] Weijie Su. You are the best reviewer of your own papers: An owner-assisted scoring mechanism. *Advances in Neural Information Processing Systems*, 34:27929–27939, 2021.

[Sundaram *et al.*, 2021] Ravi Sundaram, Anil Vullikanti, Haifeng Xu, and Fan Yao. Pac-learning for strategic classification. In *International Conference on Machine Learning*, pages 9978–9988. PMLR, 2021.

[Tambe, 2011] Milind Tambe. *Security and game theory: algorithms, deployed systems, lessons learned*. Cambridge university press, 2011.

[Wu *et al.*,] Jibang Wu, Weiran Shen, Fei Fang, and Haifeng Xu. Inverse game theory for stackelberg games: the blessing of bounded rationality. In *Advances in Neural Information Processing Systems*.

[Wu *et al.*, 2022] Jibang Wu, Haifeng Xu, and Fan Yao. Multi-agent learning for iterative dominance elimination: Formal barriers and new algorithms. In *Conference on Learning Theory*, pages 543–543. PMLR, 2022.

[Wu *et al.*, 2023] Jibang Wu, Haifeng Xu, Yifan Guo, and Weijie Su. An isotonic mechanism for overlapping ownership. In *in submission*, 2023.

[Yao *et al.*, 2023a] Fan Yao, Chuanhao Li, Denis Nekipelov, Hongning Wang, and Haifeng Xu. How bad is top-k recommendation under competing content creators? In *International Conference on Machine Learning*. PMLR, 2023.

[Yao *et al.*, 2023b] Fan Yao, Chuanhao Li, Karthik Abinav Sankararaman, Yiming Liao, Yan Zhu, Qifan Wang, Hongning Wang, and Haifeng Xu. Rethinking incentives in recommender systems: Are monotone rewards always beneficial? In *in submission*, 2023.

[Zu *et al.*, 2021] You Zu, Krishnamurthy Iyer, and Haifeng Xu. Learning to persuade on the fly: Robustness against ignorance. In *Proceedings of the 22nd ACM Conference on Economics and Computation*, pages 927–928, 2021.