
AFRL-RI-RS-TR-2020-040

# INTERACTIVE LEARNING FROM SPARSE AND DIVERSE FEEDBACK

CARNEGIE MELLON UNIVERSITY

*MARCH 2020*

FINAL TECHNICAL REPORT

***APPROVED FOR PUBLIC RELEASE; DISTRIBUTION UNLIMITED***

STINFO COPY

**AIR FORCE RESEARCH LABORATORY**
**INFORMATION DIRECTORATE**

■ **AIR FORCE MATERIEL COMMAND** ■ **UNITED STATES AIR FORCE** ■ **ROME, NY 13441**

# NOTICE AND SIGNATURE PAGE

Using Government drawings, specifications, or other data included in this document for any purpose other than Government procurement does not in any way obligate the U.S. Government. The fact that the Government formulated or supplied the drawings, specifications, or other data does not license the holder or any other person or corporation; or convey any rights or permission to manufacture, use, or sell any patented invention that may relate to them.

This report is the result of contracted fundamental research deemed exempt from public affairs security and policy review in accordance with SAF/AQR memorandum dated 10 Dec 08 and AFRL/CA policy clarification memorandum dated 16 Jan 09. This report is available to the general public, including foreign nations. Copies may be obtained from the Defense Technical Information Center (DTIC) (http://www.dtic.mil).

AFRL-RI-RS-TR-2020-040 HAS BEEN REVIEWED AND IS APPROVED FOR PUBLICATION IN ACCORDANCE WITH ASSIGNED DISTRIBUTION STATEMENT.

FOR THE CHIEF ENGINEER:

**/ S /**
AMY WONG
Work Unit Manager

**/ S /**
JULIE BRICHACEK
Chief, Information Systems Division
Information Directorate

This report is published in the interest of scientific and technical information exchange, and its publication does not constitute the Government's approval or disapproval of its ideas or findings.

# REPORT DOCUMENTATION PAGE

*Form Approved*
**OMB No. 0704-0188**



| 1. REPORT DATE *(DD-MM-YYYY)* | 2. REPORT TYPE | 3. DATES COVERED *(From - To)* |
|---|---|---|
| MARCH 2020 | FINAL TECHNICAL REPORT | SEP 2017 – SEP 2019 |

**4. TITLE AND SUBTITLE**

INTERACTIVE LEARNING FROM SPARSE AND DIVERSE FEEDBACK

**5a. CONTRACT NUMBER**
FA8750-17-2-0212

**5b. GRANT NUMBER**
N/A

**5c. PROGRAM ELEMENT NUMBER**
62788F

**6. AUTHOR(S)**

Aarti Singh, Artur Dubrawski and Barnabas Poczos

**5d. PROJECT NUMBER**
S2DL

**5e. TASK NUMBER**
UO

**5f. WORK UNIT NUMBER**
WM

**7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)**
Carnegie Mellon University
5000 Forbes Avenue
Pittsburgh PA 15213

**8. PERFORMING ORGANIZATION REPORT NUMBER**

**9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)**

Air Force Research Laboratory/RISC
525 Brooks Road
Rome NY 13441-4505

**10. SPONSOR/MONITOR'S ACRONYM(S)**
AFRL/RI

**11. SPONSOR/MONITOR'S REPORT NUMBER**
AFRL-RI-RS-TR-2020-040

**12. DISTRIBUTION AVAILABILITY STATEMENT**
Approved for Public Release; Distribution Unlimited. This report is the result of contracted fundamental research deemed exempt from public affairs security and policy review in accordance with SAF/AQR memorandum dated 10 Dec 08 and AFRL/CA policy clarification memorandum dated 16 Jan 09

**13. SUPPLEMENTARY NOTES**

**14. ABSTRACT**

Despite the great success of machine learning, most learning algorithms remain primarily non-interactive. Human learning, on the other hand, is highly interactive. As learners we are inquisitive - we ask questions about the data shown to us and employ feedback to decide what questions would yield answers that are most informative for the learning task at hand. In this project, we developed a suite of data-efficient interactive machine learning algorithms that employ judicious choice of what data to collect. This includes development of new algorithms as well as integration of algorithms developed in PIs' prior works.

**15. SUBJECT TERMS**

Active learning, Data-efficient learning, interactive

| 16. SECURITY CLASSIFICATION OF: | | | 17. LIMITATION OF ABSTRACT | 18. NUMBER OF PAGES | 19a. NAME OF RESPONSIBLE PERSON |
|---|---|---|---|---|---|
| a. REPORT | b. ABSTRACT | c. THIS PAGE | | | **AMY WONG** |
| U | U | U | UU | 27 | 19b. TELEPHONE NUMBER (*Include area code*) **N/A** |

## TABLE OF CONTENTS

**Section** **Page**

## LIST OF FIGURES

**Figure** **Page**

## 1.0 SUMMARY

Despite the great success of machine learning, most learning algorithms remain primarily non-interactive. Human learning, on the other hand, is highly interactive. As learners we are inquisitive - we ask questions about the data shown to us and employ a variety of diverse feedback to decide what questions would yield answers that are most informative for the learning task at hand. In this project, we developed a suite of data-efficient interactive machine learning algorithms that employ judicious choice of what data to collect, when to collect it and how much to collect. This includes development of new algorithms as well as integration of algorithms developed in the investigators' prior works.

## 2.0 INTRODUCTION

In this project, we focused on developing machine learning methods that can interact with the data generating process to query the most informative data points, features, representations, labels, and type of labels. We focus on the following four specific tasks which focus on different aspects of interactive machine learning.

Task I. Interactive unsupervised learning. We extended and incorporated our prior work on interactive algorithms for unsupervised problems, where the algorithm sequentially decides which data points, similarities, and features to query that are the most relevant to the given problem [9],[10],[13],[17].

Task II. Interactive supervised learning with diverse oracles. We developed methods for supervised problems such as classification [1],[19],[22], regression [7],[11],[15], optimization [21], etc. which can leverage feedback from multiple types of diverse oracles such as direct labeling oracle, pairwise comparison oracle, ranking oracle, and multi-fidelity oracle, by adapting the queries presented to them according to their (unknown) skill set.

Task III. Interactive transfer learning. We developed interactive methods that selectively query data points and labels from source domain that are most informative to the target domains [3] and can enable best transfer of information between the two domains in a transfer learning setting [2].

Task IV. Interactive deep learning. We developed a solid foundational understanding of methods for optimization of black-box models such as deep neural networks [4],[8],[12],[14],[16] and their generalization properties for deep neural networks [18],[20]. Then we developed methods for active learning and optimization where the function being learnt/optimized is represented using a deep neural network model [23].

## 3.0 METHODS, ASSUMPTIONS AND PROCEDURES

We present below an overview of the interactive learning methods we developed for the four tasks outlines above. Details of the algorithms, assumptions and theoretical analysis can be found in the sections following the overview and the papers we published on these topics that are referenced below.

Task I: Interactive unsupervised learning

As an unsupervised learning task, we studied the **adaptive high-dimensional density estimation** problem. For general data, unlike spatial/temporal data, we do not have known correlations a priori among covariates that may be exploited. In order to model high dimensional data, the main challenge lies in constructing models that are flexible enough while having tractable learning algorithms. A large number of methods have considered auto-regressive models to estimate the conditional factors $p(x_i|x_{i-1}, \ldots, x_1)$, for $i \in \{1, \ldots, d\}$ in the chain rule. If we only model the conditionals, the conditional factors $p(x_i|x_{i-1}, \ldots)$, may become increasingly complicated as i increases to d. On the other hand if we use a complex transformation with restricted conditionals, then the transformation has to ensure that the transformed variables are independent. This requirement of independence on the transformed variables can be very restrictive. Now note that the transformed space is homeomorphic to the original space and a simple relationship between the density of the two spaces exists through the Jacobian. Thus, we can employ conditional modeling on the transformed variables to alleviate the independence requirement, while being able to recover density in the original space in a straightforward fashion. Using these ideas, we proposed a new family of methods, the transformation autoregressive networks for density estimation [9]. They apply complex transformations and autoregressive modeling of the conditionals.

We also studied the problem of **active search on networks**. How can we help an investigator to efficiently connect the dots and uncover the network of individuals involved in a criminal activity based on the evidence of their connections, such as visiting the same address, or transacting with the same bank account? We formulate this problem as Active Search of Connections, which finds target entities that share evidence of different types with a given lead, where their relevance to the case is queried interactively from the investigator. We present RedThread [10], an efficient solution for inferring related and relevant nodes while incorporating the user's feedback to guide the inference. Our experiments focus on case building for combating human trafficking, where the investigator follows leads to expose organized activities, i.e. different escort advertisements that are connected and possibly orchestrated. RedThread is a local algorithm and enables online case building when mining millions of ads posted in one of the largest classified advertising websites. The results of RedThread are interpretable, as they explain how the results are connected to the initial lead. We experimentally show that RedThread learns the importance of the different types and different pieces of evidence, while the former could be transferred between cases.

**Sparse dictionary learning** has become a popular method for adaptively identifying parsimonious representations of a dataset, a fundamental problem in machine learning and signal processing. While most work on sparse dictionary learning assumes a training dataset of independent and identically distributed samples, a variant known as convolutional sparse dictionary learning relaxes

this assumption, allowing more general sequential data sources, such as time series or other dependent data. Although recent work has explored the statistical properties of classical sparse dictionary learning, the statistical properties of convolutional sparse dictionary learning remain unstudied. We aimed to understand the statistical properties of convolutional sparse dictionary learning, in terms of reconstruction (or denoising) error (i.e., the error of reconstructing a data sequence from a learned convolutional sparse dictionary decomposition, a process widely used for denoising and compression. We study this by identifying the minimax convergence rate of convolutional sparse dictionary learning in terms of reconstruction risk, by both upper bounding the risk of an established sparse dictionary learning estimator and proving a matching information-theoretic lower bound [13]. The emphasis in this paper is on proving results under minimal assumptions on the data. As such, we make no assumptions whatsoever on the dictionary or encoding matrices (such as restricted eigenvalue or isometry properties). Compared even to these "assumptionless" results, we consider dropping yet another assumption, namely that noise is independently distributed. In many of the above applications (such as image demosaicing or inpainting, or structured denoising), noise is strongly correlated across the data, and yet dictionary learning approaches nevertheless appear to perform consistent denoising. To the best of our knowledge, this phenomenon has not been explained theoretically. One likely reason is that this phenomenon does not occur in basis pursuit denoising, lasso, and many related compressed sensing problems, where consistency is usually not possible under arbitrarily dependent noise. In the context of sparse dictionary learning, tolerating dependent noise is especially important in the convolutional setting, where coordinates of the data are explicitly modeled as being spatially or temporally related. Under sufficient sparsity, our results imply consistency (in reconstruction error) of convolutional sparse dictionary learning, even under arbitrary noise dependence. Our results indicate that consistency in reconstruction risk is possible precisely in the `ultra-sparse' setting, in which the sparsity (i.e., the number of feature occurrences) is in o(N) in terms of the length N of the training sequence. Notably, our results make very weak assumptions, allowing arbitrary dictionaries and dependent measurement noise. We also verified our theoretical results with numerical experiments.

Many problems in adaptive decision-making under uncertainty fall into the design of experiments framework, where one wishes to design a sequence of experiments and collect data so as to achieve a desired goal. For example, in electrolyte design for batteries, a chemist would like to conduct experiments that measure battery conductivity in order to identify an electrolyte design that maximizes conductivity. On a different day, she would like to experiment with different designs to learn how the viscosity of the electrolyte changes with the design. These two tasks, black-box optimization and active learning, fall under the umbrella of design of experiments and are pervasive in industrial and scientific applications. While several methods exist for specific design of experiments tasks, real world problems are broad and complex, and specialized methods have limited applicability. Continuing with the electrolyte example, the chemist can typically measure both conductivity and viscosity with a single experiment. Since such experiments are expensive, it is wasteful to first perform a set of experiments to optimize conductivity and then a fresh set to learn viscosity. Rather, it is desirable to design a single set of experiments that simultaneously achieves both goals. A second example is metallurgy, where one wishes to identify phase transitions in an alloy by carefully selecting a sequence of X-ray diffraction experiments. Here and elsewhere, the goal of the experimenter is application specific and cannot be simply shoehorned into standard design of experiments formulations such as black-box optimization, active learning,

etc. In addition, domain knowledge about the problem may need to be considered in selecting experiments, as it may significantly reduce the number of experiments needed to achieve the desired goal.

To address these desiderata, we develop a general and flexible framework for goal oriented design of experiments, where a practitioner may specify her desired goal via a reward function that can depend on the data collected during the design of experiments process and unknown system characteristics, and hence cannot be directly computed by a decision-maker. We designed a new myopic strategy [17] for a wide class of **sequential design of experiment** problems, where the goal is to collect data in order to fulfil a certain problem specific goal. Our approach, Myopic Posterior Sampling, is inspired by the classical posterior (Thompson) sampling algorithm for multi-armed bandits and leverages the flexibility of probabilistic programming and approximate Bayesian inference to address a broad set of problems. Our approach has two key advantages. First, our Bayesian formulation allows one to straightforwardly specify domain expertise. Moreover, modern tools for probabilistic programming enable practitioners to apply a Bayesian algorithm such as ours in a fairly straightforward manner. Second, our myopic strategy is simple and computationally attractive in comparison with policies that engage in long-term planning. Nevertheless, borrowing ideas from submodular optimization and reinforcement learning, we derive natural conditions under which our myopic policy is competitive with the globally optimal one. Empirically, this general-purpose strategy is competitive with more specialized methods in a wide array of design of experiments tasks, and more importantly, enables addressing complex Design of experiments goals where no existing method seems applicable. On the theoretical side, we leverage ideas from adaptive submodularity and reinforcement learning to derive conditions under which myopic posterior sampling achieves sublinear regret against natural benchmark policies. Specifically, our contributions are:

(i) We propose a flexible framework for design of experiments that allows a practitioner to describe their system via a probabilistic model, and specify their goal via a reward function. We derive a general purpose algorithm, Myopic Posterior Sampling, for this setting.
(ii) Empirically, we demonstrate myopic posterior sampling performs favorably in a variety of synthetic and real world design of experiments tasks. Despite its generality, myopic posterior sampling is competitive with specialized methods designed for particular settings. More importantly, it enables design of experiments in non-standard application specific settings.
(iii) In our theoretical analysis, we explore conditions under which myopic posterior sampling, which learns about the system over time, is competitive with myopic and globally optimal strategies that have full knowledge of the system.

Task II: Interactive supervised learning with diverse oracles

Given high costs of obtaining labels for big datasets, interactive learning is gaining popularity in both practice and theory of machine learning. On the practical side, there has been an increasing interest in designing algorithms capable of engaging domain experts in two-way queries to facilitate more accurate and more effort-efficient learning systems. On the theoretical side, study

of interactive learning has led to significant advances such as exponential improvement of query complexity over passive learning under certain conditions. While most existing approaches to interactive learning fix the form of an oracle, e.g., the labeling oracle, and explore the best way of querying, some recent work allows for multiple diverse forms of oracles. We focus on this latter setting, also known as active dual supervision. We investigate how to learn a classifier, regressor or optimizer, given limited access to labels on individual instances and pairwise comparisons about which one of two given instances is more likely to belong to a class or have a higher regression or optimization function value.

It is known in prior literature that it is often easier for human labelers to compare or rank inputs rather than predict an output directly. For example, in material design, synthesizing materials for specific conditions requires expensive experimentation, but with an appropriate algorithm we can leverage expertise of material scientists, for whom it may be hard to accurately assess the resulting material properties, but who can quickly compare different input conditions and suggest which ones are more promising. Similarly, in clinical settings, precise assessment of each individual patient's health status can be difficult, expensive and/or risky (e.g. it may require application of invasive sensors or diagnostic surgeries), but comparing relative statuses of two patients at a time may be relatively easy and accurate. In both these scenarios we may have access to a modest amount of individually labeled data, but the bulk of more accessible training information is available via pairwise comparisons. There are many other examples where humans find it easier to perform pairwise comparisons rather than providing direct labels, including content search, image retrieval, ranking, etc.

Despite many successful applications of comparison oracles, many fundamental questions remain. One of them is how to design noise-tolerant, cost-efficient algorithms that can approximate the unknown target hypothesis to arbitrary accuracy while having access to pairwise comparisons. On one hand, while there is theoretical analysis on the pairwise comparisons concerning the task of learning to rank, estimating ordinal measurement models and learning combinatorial functions, much remains unknown how to extend these results to more generic hypothesis classes. On the other hand, although we have seen great progress on using single or multiple oracles with the same form of interaction, classification, regression or optimization using both comparison and labeling queries remains an interesting open problem.

In [1], we developed an interactive classification algorithm that learns from pairwise comparisons/ranking and direct labeling oracles, adaptively choosing which one to query, when and how much. The algorithm interleaves comparison queries to learn a ranking with a binary search using direct queries that classify current sample. Our analysis precisely quantifies how access to such comparison/ranking information significantly reduces label complexity (e.g. the need to perform expensive experiments) as a function of their relative cost with respect to the comparison queries and noise in both types of oracles. Specifically, we make the following contributions:

(i) We analyze our algorithm under Tsybakov and adversarial noise conditions for the labeling oracle, along with the adversarial noise condition for the comparison oracle. Our general framework can augment any active learning algorithm by replacing the batch sampling in these algorithms with our algorithm.

(ii) We propose a specific instantiation of our algorithm, which can learn an arbitrary hypothesis class. The label complexity of the algorithm is as small as learning a threshold function under both Tsybakov and adversarial noise condition, independently of the structure of the hypothesis class. The total query complexity improves over previous best-known results under Tsybakov noise condition which can only access the labeling oracle.

(iii) We derive another instantiation of our algorithm to learn the class of halfspaces. This algorithm has the same label and total query complexity, but is computationally efficient.

(iv) We present lower bounds on total query complexity for any algorithm that can access both labeling and comparison oracles, and a noise tolerance lower bound for our algorithms. These lower bounds demonstrate that our analysis is nearly optimal.

Building on this work where we demonstrated algorithms and guarantees for nonparametric classification using pairwise comparison and direct query oracles, we also developed methods for non-parametric regression using ranking or pairwise comparisons and direct labeling oracles [7]. We develop the Ranking-Regression algorithm for nonparametric regression that can leverage ordinal information, in addition to direct labels. Theoretical analysis and practical experiments show the strength of our algorithm. Unlike classification where active queries help improve sample complexity for both comparisons and direct labels, in regression of smooth functions, active queries are only needed for comparisons to reduce the problem to learning an isotonic one-dimensional function which can be learnt efficiently using passive direct queries. Specifically, we make the following contributions:

(i) To establish the usefulness of ordinal information in nonparametric regression, in we first consider the idealized setting where we obtain a perfect ordering of the unlabeled set. We show that the mean squared error of our algorithm can be shown to scale like $m^{-2/3} + n^{-2/d}$, where m denotes the number of labeled samples, n the number of ranked samples, and d denotes the dimension of the feature space. To achieve a target constant MSE, the number of labeled samples required by the Ranking-Regression algorithm is independent of dimension. This result establishes that sufficient ordinal information of high quality can allow us to effectively circumvent the curse of dimensionality.

(ii) Next, we analyze Ranking-Regression when using a noisy ranking. We show that the mean squared error scales as $m^{-2/3} + n^{-2/d} + v^{1/2}$, where v is the Kendall-Tau distance between the true and noisy ranking.

(iii) As a corollary, we develop results for Ranking-Regression using pairwise comparisons. If the comparison noise is bounded, our algorithm can be combined with algorithms for ranking from pairwise comparisons to obtain an MSE of $m^{-2/3} + n^{-2/d}$ when $d >= 4$.

(iv) We give information-theoretic lower bounds to characterize the fundamental limits of combining ordinal and standard supervision. These lower bounds show that our algorithms are almost optimal. In particular, the Ranking-Regression algorithm under perfect ranking, as well as under bounded noise comparisons, is optimal up to log factors.

(v) In our experiments, we test Ranking-Regression on simulated data, and on various age-estimation tasks. Our experimental results show the advantage of our algorithm over algorithms that only use labeled data when this labeled data is scarce. Our experiments

with the age-estimation data also show the practicality of our proposed Ranking-Regression algorithm. These results are presented in Section 4.0.

We also developed an interactive algorithm [11] that utilizes both labels and comparisons for linear regression, and show that it only requires a very small amount of direct labels to achieve low error. Similar to non-parametric regression, we show that active queries are only needed for comparisons to reduce the problem to a one-dimensional problem of learning the scale in linear regression (akin to learning threshold for classification). Specifically,

(i) Given a budget of m label queries and n comparison queries, we show that MSE of our algorithm decays at the rate of $m^{-1}+e^{-n/d}$. Once again we see that when sufficiently many comparisons are available, the label complexity of our algorithm is independent of the dimension d.

(ii) We also give minimax lower bounds for the problem, showing that our algorithm is optimal up to log factors in the number of direct labels needed.

(iii) We demonstrate the efficacy of this algorithm on a practical task of estimating people's ages from face images.

We further polished our work on regression using pairwise comparison and direct labels, and tested our algorithms in experiments on simulated data, age estimation and price-estimation tasks [15]. Specifically, we curated a new dataset using crowdsourced data obtained through Amazon's Mechanical Turk. We provide workers AirBnB listings and attempt to estimate property asking prices. We obtain both direct and comparison queries for the listings, and also study the time taken by workers to provide these different types of feedback. We find that, our algorithms which combine direct and comparison queries are able to achieve significantly better accuracy than standard supervised regression methods, for a fixed worker time budget.

We have successfully applied our earlier work on comparison oracles to streamline training data acquisition in critical care informatics [19]. The idea is that we can reduce an expert clinician's effort needed to annotate a sufficient sample of data to enable training predictive models of clinical interest, and such reduction is feasible when we can ask simple, fitting types of queries. In our case study, it appeared that comparing two patient cases to assess which one is more severe can be done faster, and more importantly with greater confidence, than labeling the two cases directly into severe or mild. These results are presented in Section 4.0. This work has been published and presented at lead healthcare venues. We are currently generalizing our approach to enable crowdsourcing annotations from multiple labelers.

Zeroth order non-convex optimization, also variously known as continuous multi-armed bandit or black-box optimization, is an important problem that naturally appears in various domains like dynamic pricing, reinforcement learning and material science. With an unknown black-box function f, zeroth order optimization aims to find the optimal point of the function with as few queries to (a noisy version of) f(x) as possible, with no gradient information directly available. Although zeroth order convex optimization is generally efficient, optimizing a non-convex f under smoothness constraints requires the same effort as estimating f almost everywhere, and usually leads to a query complexity exponential in d, where d is the feature space dimensionality. The prohibitive cost for non-convex optimization has motivated research on suitable assumptions, such as linear bandits, convex approximations, and optimization based on level sets. We propose a complementary approach, where in addition to direct queries to f, one can also compare two points

in feature space, and obtain the point with a larger f value. Inspired by dueling bandits in the bandit domain, we call our setting non-convex optimization with dueling-choice bandits; we note that different than dueling bandits, here direct queries and comparisons are both available for optimizing f, and therefore duels are available as an additional choice.

We build our work on GP-UCB, a method for optimizing unknown functions under the Gaussian process (GP) assumption by optimizing the Upper Confidence Bound (UCB). Closest to our setting is a line of recent research on multi-fidelity GP optimization, which assumes that we can query the target functions at multiple fidelities of different costs and precisions. To briefly describe it, our setting is harder since we cannot directly query the function on which comparisons are based. Moreover, the multi-fidelity assumptions such as fidelities being close in sup-norm do not hold for our setting since any constant shift of the comparison function yields the same comparisons. We instead consider an active transfer learning setting where information from a function that can be learned using comparisons is transferred actively to optimize the target function. Optimization with comparisons has been studied under the framework of derivative-free optimization and continuous dueling bandits. However to the best of our knowledge, no previous work has theoretical guarantees on optimizing a non-convex f. Also, these results cannot be applied when the comparisons are biased (i.e., a Condorcet winner on comparisons might not be the best point for direct queries).

We develop and evaluate a new algorithm for non-convex optimization with dueling choices, which we refer to as Comparison-based GP-UCB (COMP-GP-UCB). We can show that if comparisons are based on the same underlying function as a direct label, then the function can be optimized using comparisons only (which are cheaper) unlike the regression and classification settings where some labels are needed to figure out the 'scale' of the parameters being learnt. However, in practice, comparisons are done by experts/humans and are often based on a very crude approximation to the original function. To address this model mismatch and enable active transfer of knowledge from human queries to the true function being learned, we developed a Gaussian process regression based algorithm [21], where instead of directly querying the point with the maximum upper confidence bound, we perform a constrained optimization and use comparisons to filter out suboptimal points. Our theoretical and experimental results show the strengths of our algorithm. Specifically, we make the following contributions:

(i) When we can obtain comparisons based on the target function f, we show that comparisons can be as powerful as direct queries: COMP-GP-UCB can achieve the same rate of convergence as its label-only counterparts, while using only comparisons and no direct queries. This solves an open problem of developing continuous dueling bandit algorithms with no-regret guarantees.

(ii) Next, we assume that comparisons are based on a misspecified function fc, where fc approximates f. COMP-GP-UCB in this case uses comparisons to optimize a function fr which has the same optimizer as fc, and then use direct queries to search in a smaller region for the optimum of the target function. The regret rate of COMP-GP-UCB is then better than the label-only counterparts, and it depends on the difference between fc and f: the better the approximation, the lower the regret we can get from COMP-GP-UCB. In particular, we show that the label complexity of our algorithm depends on an improved information gain corresponding to a comparison based constraint set that

restricts the search space for the optimum. In contrast, in the direct query only setting, the label complexity depends on information gain corresponding to the entire domain. We further demonstrate a version of the algorithm that adapts to the degree of model mismatch.

(iii) Our algorithm also extends multi-fidelity GP optimization to the setting where information is transferred actively from a lower fidelity to a higher fidelity while only assuming that the optimizer of the lower fidelity (source function) is within a constant distance of the optimizer of the higher fidelity (target function), instead of the fidelities being close everywhere.

(iv) In our experiments, we test COMP-GP-UCB on multi-fidelity functions from previous literature and show that it outperforms label-only algorithms and existing multi-fidelity algorithms when comparisons are cheaper than direct labels.

Building on our prior work on classification, we consider a related problem called the Thresholding Bandit Problem which aims to find the set of arms with mean rewards greater than a given threshold. The Thresholding Bandit Problem is an important pure-exploration multi-armed bandit problem and has a wide range of applications, such as anomaly detection, candidate filtering, and crowdsourced classification. For example, a popular crowdsourced classification model assumes that there are K items with the latent true labels $l_i$ in {0,1} for each item. The labeling difficulty of the i-th item is characterized by its soft label $y_i$ in [0 1], which is defined as the probability that a random crowd worker will label the i-th item as positive. It is clear that the item is easy to label when $y_i$ is close to 0 or 1, and difficult when $y_i$ is close to 0.5. In multi-armed bandits, $y_i$ is the mean reward of arm i, and pulling this arm leads to a Bernoulli observation with mean $l_i$. Moreover, it is natural to assume that the soft label $y_i$ is consistent with the true label, i.e., $y_i >= 0.5$ if and only if $l_i = 1$. Therefore, identifying items belonging to class 1 is equivalent to detecting those arms with $y_i > t$ where $t = 0.5$.

We consider a new setting of thresholding bandits which arises in crowdsourced scenarios, where in addition to pulling arms, one can also duel two arms and get the arm with a greater mean since dueling two arms (comparing two choices) can be more cost and time efficient for humans than direct pulls (labeling an input). For example, in crowdsourcing, it will be cheaper and time efficient to ask a worker which image is more relevant to a query as compared to asking for an absolute relevance score of an image. We refer to this problem as thresholding bandit problem with Dueling Choices.

A straightforward way to solve the thresholding bandit problem with dueling choices is to utilize an existing ranking algorithm to rank all the arms, and then use a binary search to find the boundary. However, this method is impractical because it can be very hard to differentiate arms with similar means (e.g., equally good images, similar quality materials). These arms might be far from the threshold and it is actually unnecessary to differentiate them. We instead take an iterative approach and develop an algorithm [22] called Rank-Search for solving thresholding bandit problem with dueling choices by alternating between ranking and binary search. Specifically, our contributions are:

(i) We develop the Rank-Search algorithm for thresholding bandits with dueling choices, which alternates between refining the rank over all items using duels and a binary

search process using pulls to figure out the threshold among ranked items. We interleave the ranking and searching step so that we do not waste time differentiating equally good arms.

(ii) We prove theoretical guarantees for Rank-Search. We analyze the number of duels and pulls required under the fixed confidence setting, i.e., to recognize the set of arms with reward larger than t with high probability. To better illustrate our main idea, we further provide concrete examples, which show that the proposed Rank-Search only requires $O(\log^2 K)$ direct labels, while the classical thresholding bandit problem requires at least $O(K)$ labels.

(i) We show complementary lower bounds which establish that Rank-Search is near-optimal in both duel and pull complexity.

(ii) Experiments show that Rank-Search outperforms previous baseline algorithms that only use pulls or duels.

Task III: Interactive transfer and multitask learning.

We developed an interactive method that selectively query data points and labels from source domain that are most informative to the target domains [3]. For this, we consider a classical design of experiments setting but optimize the design in source domain to enable best transfer of information to a target domain. Classical optimal solutions for design of experiments are combinatorial and we develop a continuous convex relaxation of these criteria followed by intelligent rounding using greedy, sampling and regret minimization approaches that is computationally efficient and near optimal. This work only considers a domain adaptation setting where the distribution of points in each domain are different but they share the same regression function. This can further be extended to the setting of hypothesis transfer learning where the input-output mapping in the two domains is related by a link function, building on our prior work on hypothesis transfer learning setting [2].

Task IV: Interactive deep learning

We started by developing a solid foundational understanding of methods for optimization of black-box models such as deep neural networks and their generalization properties for deep neural networks.

We start with the problem of black-box global optimization of an unknown non-convex smooth function with zeroth-order feedback. In this setup, an algorithm is allowed to adaptively query the underlying function at different locations and receives noisy evaluations of function values at the queried points (i.e. the algorithm has access to zeroth-order information). We precisely identify when improvements in convergence rate are possible for smooth functions; we show that for functions with fast level set growth around the global minimum, carefully designed optimization algorithms can identify a near global minimizer with many fewer queries [16]. In contrast to classical global minimax analysis of nonparametric estimation problems, we adopt a local analysis which characterizes the optimal convergence rate of optimization error when the underlying

function f is within the neighborhood of a "reference" function f0. We provide an intuitive and efficient algorithm that attains the derived upper error bounds. Specifically, our contributions are:

(i) We design an iterative (active) algorithm whose optimization error converges at a rate depending on the reference function f0. When the level sets of f0 satisfy certain regularity and polynomial growth conditions, the local rate can be upper bounded by $n^{-a/(2a+d-ab)}$, where b in [0,d/a] is a parameter depending on f0 that characterizes the volume growth of level sets of f0. The rate matches the global minimax rate $n^{-a/(2a+d)}$ for worst-case f0 where b=0, but has the potential of being much faster when b>0. We emphasize that our algorithm has no knowledge of f0, a or b and achieves this rate adaptively.

(ii) We prove local minimax lower bounds that match the $n^{-a/(2a+d-ab)}$ upper bound, up to logarithmic factors in n. More specifically, we show that even if f0 is known, no (active) algorithm can estimate f in close neighborhoods of f0 at a rate faster than $n^{-a/(2a+d-ab)}$. We further show that, if active queries are not available and $x_1, \ldots, x_n$ are i.i.d. uniformly sampled from X, the $n^{-a/(2a+d)}$ global minimax rate also applies locally regardless of how large b is. Thus, there is an explicit gap between local minimax rates of active and uniform query models.

(iii) In the special case when f is convex, the global optimization problem is usually referred to as zeroth-order convex optimization and this problem has been widely studied. Our results imply that, when f0 is strongly convex and smooth, the local minimax rate is on the order of $n^{-1/2}$, which matches the convergence rates in literature. Additionally, our negative results indicate that the $n^{-1/2}$ rate cannot be achieved if f0 is merely convex, which seems to contradict $n^{-1/2}$ that do not require strong convexity of f. However, it should be noted that mere convexity of f0 does not imply convexity off in a neighborhood of f0. Our results show significant differences in the intrinsic difficulty of zeroth-order optimization of convex and near-convex functions.

As a first step towards optimizing high-dimensional non-convex black-box functions, we consider establishing a baseline of optimizing a high-dimensional convex function using stochastic zeroth-order queries. Under sparsity assumptions on the gradients or function values, we present two algorithms: a successive component/feature selection algorithm and a noisy mirror descent algorithm using Lasso gradient estimates, and show that both algorithms have convergence rates that depend only logarithmically on the ambient dimension of the problem [14]. The first method uses a few noisy samples to select a small subset of "important variables" S in 1, …, d; afterwards, existing low-dimensional zeroth-order optimization techniques are applied to obtain a minimizer of f restricted to S. This requires assuming the function is sparse and only varies on a small subset of coordinates. The second method combines stochastic mirror descent and de-biased Lasso gradient estimates. The stochastic mirror descent based method does not require assuming the function is sparse and is seen to outperform the variable selection based method in simulations.

Moving to first-order methods which are used more prominently for black-box functions such as neural networks, a central challenge to using first-order methods for optimizing nonconvex problems is the presence of saddle points. Although the popular first-order method gradient descent was shown to almost always escape saddle points asymptotically, we show that even with fairly natural random initialization schemes and non-pathological functions, gradient descent can be significantly slowed down by saddle points, taking exponential time to escape [4]. On the other

hand, gradient descent with perturbations is not slowed down by saddle points - it can find an approximate local minimizer in polynomial time. This result implies that gradient descent is inherently slower than perturbed gradient descent, and justifies the importance of adding perturbations for efficient non-convex optimization. The counter-example that supports this conclusion is a smooth function defined in d dimensions, where gradient descent with random initialization will visit the vicinity of d saddle points before reaching a local minimum. While perturbed gradient descent takes a constant amount of time to escape each saddle point, gradient descent will get closer and closer to the saddle points it encounters later, and thus take an increasing amount of time to escape. Eventually, gradient descent requires time that is exponential in the number of saddle points it needs to escape, thus $e^{\Omega(d)}$ steps. While our focus is theoretical, we also present experiments that illustrate our theoretical findings.

Since first-order methods often get stuck at saddle points, greatly deteriorating their performance, an alternative to adding perturbations is to use second-order methods. However, most works on second-order methods rely extensively on expensive Hessian-based computations, making them impractical in large-scale settings. To tackle this challenge, we introduce a generic framework that minimizes Hessian based computations while at the same time provably converging to second-order critical points [12]. Our framework carefully alternates between a first-order and a second-order subroutine, using the latter only close to saddle points, and yields convergence results competitive to the state-of-the-art. Empirical results suggest that our strategy also enjoys a good practical performance.

While the previous analysis is general and applies to any black-box function with certain mild properties, it is not known if neural networks satisfy those properties. So we specifically consider the problem of learning a one-hidden-layer neural network with non-overlapping convolutional layer and rectified linear unit activation. We show that despite the presence of the spurious local minimizer, gradient descent with weight normalization from randomly initialized weights can still be proven to recover the true parameters with constant probability, which can be boosted to probability 1 with multiple restarts [8]. We also show that, with constant probability, the same procedure could also converge to the spurious local minimum, showing that the local minimum plays a non-trivial role in the dynamics of gradient descent. Furthermore, a quantitative analysis shows that the gradient descent dynamics has two phases: it starts off slow, but converges much faster after several iterations.

Next, we analyzed the generalization error of a simple convolution filter and an one-hidden-layer convolutional neural network with linear activation [18] and provide first formal evidence that the success of convolutional neural network is because it is a more compact representation than the fully connected neural network and thus requires fewer samples for learning.

We also formalized a statistical model for generative adversarial networks, starting from studying the problem of estimating a nonparametric probability distribution under a family of losses called Besov integral probability metrics [20]. This family is quite large, including, for example, Lp distances, total variation distance, and generalizations of both Wasserstein (Earthmover's) and Kolmogorov-Smirnov distances. For a wide variety of settings, we determined the mini-max optimal convergence rate of this estimation problem. We also showed that in many cases linear distribution estimates, such as the empirical distribution or kernel density estimator, cannot

converge at the optimal rate. We also showed that integral probability metrics can be used to formalize a statistical model of generative adversarial networks. As a result, the above theorems imply bounds on the statistical error of a generative adversarial network, showing, for example, that, in many cases, generative adversarial networks can strictly outperform the best linear estimator.

Deep neural networks are increasingly being used to model black-box functions. Examples include modeling brain response to stimuli, material properties under given synthesis conditions, and digital art. In these applications, often the model is a surrogate and the goal is rather to optimize the black-box function to achieve the desired brain response, material property, or digital art characteristics. Moreover, resource constraints imply that, rather than training on a passive dataset, one should focus subsequent sampling on the most informative data points. In the Bayesian setting, this can be achieved by utilizing the ability of Bayesian models such as Gaussian processes to model uncertainty in observed data via posterior variance, which can guide subsequent sampling. While a simple approach is to learn a deep model on passively acquired evaluations of the function, and then report its optima, this is wasteful as often the black-box evaluations are expensive (c.f., subject time in a brain scanner is limited, material synthesis experiments are expensive, etc.). Also, often pre-trained models of black-boxes need to be updated subsequently to identify inputs that may lead to novel outputs not explored in training set. For example, in material science a model trained to predict the energy of a pure lattice may need to be updated to understand new low-energy configurations achievable under defects, or deep neural net models of images may need to be updated to achieve desired characteristics of synthetic digital images. Sequential optimization of neural network models re-quires an acquisition function, similar to Bayesian optimization. However, popular acquisition functions (UCB, expectation maximization, Thompson sampling, etc.) are mostly based on an uncertainty measure or confidence. bound which characterizes the variability of the predictions. However, uncertainty estimates for deep neural networks are largely lacking or are very expensive to compute. For example, softmax probabilities are found to be not well-calibrated and bootstrap or cross-validation estimates require re-training the network several times which is often computationally prohibitive. We seek to investigate principled and computationally efficient estimators of uncertainty measures (like variance of prediction), that can then be used to guide subsequent sampling for optimization of black-box functions.

We use the concept of influence functions from classical statistics to estimate the variance of neural network outputs which can be computed efficiently without retraining the network, and design a black-box optimization algorithm similar to confidence bound-based Bayesian algorithms. It is known that if the loss function is twice-differentiable and strictly convex, then the influence function has a closed-form approximation, and the influence-function based estimate provides an asymptotic approximation of the variance of prediction. Even though the loss function of neural networks is non-differentiable and non-convex, it was recently shown that in practice, the approximation continues to hold for this case. Our method [23] can outperform Bayesian optimization approaches for complex synthetic and real-world optimization problems since neural networks provide a better model for such complex scenarios. Our contributions can be summarized as follows:

(i) We use influence functions to obtain a computation-ally efficient approximation of prediction variance in neural networks.

(ii) We propose a computationally efficient method to compute influence functions for neural network pre-dictions. Our approach uses a low-rank approximation of the Hessian, which is represented using an auxiliary network, and trained along with the main network.

(iii) We develop a deep black-box optimization method using these influence function based uncertainty estimates that is valid in the non-Bayesian setting.

(iv) We demonstrate the efficacy of our method with experiments, and show that our method can be com-parable to, and also outperforms Bayesian optimization in settings where neural networks may be able to model the underlying function better than GPs.

(v) We demonstrate that our variance approximation can be used in the setting of active learning with experiments on real-world material synthesis data.

## 4.0 RESULTS AND DISCUSSION

In addition to the theoretical results quantifying the sample complexity of the methods we develop, the data-efficient interactive methods we developed were demonstrated to be practically successful in various simulated settings as well as real world applications including age estimation from face images, price estimation for AirBnB ads, human trafficking as well as critical care informatics, as mentioned above. We refer to the papers cited herein and included in Appendices for details and extensive results, only highlighting some of the results here.

In a critical care setting, we considered the task of estimating a patient's disease severity level from 1 (min) to 4 (max) by using direct queries to clinicians of the form "How sick is this patient?" as well as indirect queries of the form "Which parent is sicker?" given access to their vital signs data. Specifically, 2 expert clinicians were shown 793 real alerts in non-invasive continuous vital signals data as shown below on left (Figure 1), collected in a 24 bed surgical/medical trauma step down unit over a period of 3 months. The patient's ground truth severity level on a scale from 1 to 4 was available in records based on actual outcomes as shown on right (Figure 2).

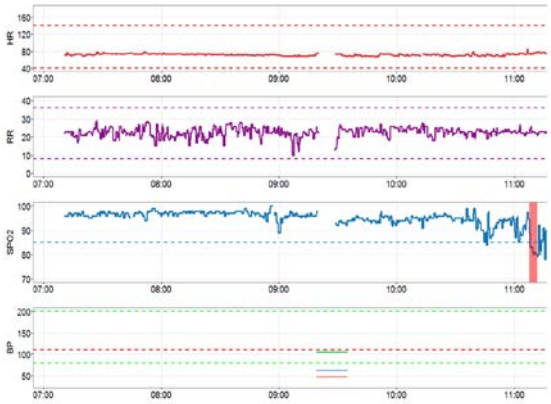


**Figure 1. Continuous vital signals data for a single patient shown to experts.**

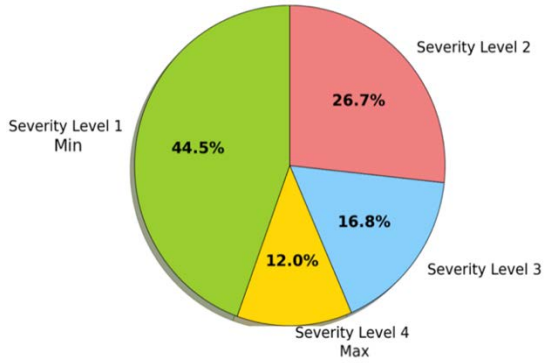


**Figure 2. Distribution of ground truth severity level.**

The hypothesis is that easier questions (such as comparisons) can be carried out with less effort and more confidence, yielding more accurate data for training machine learning models. The results of our analysis are shown below for training machine learning models using direct queries only vs. direct + comparison queries. Lesser effort was needed for comparisons and the x-axis has been adjusted to reflect total effort needed in terms of direct labels (i.e. 384 comparisons required effort equivalent to 180-116 = 64 direct label queries). As can be seen, active learning using comparisons and direct queries is able to save considerably on the total clinician effort needed for the task.

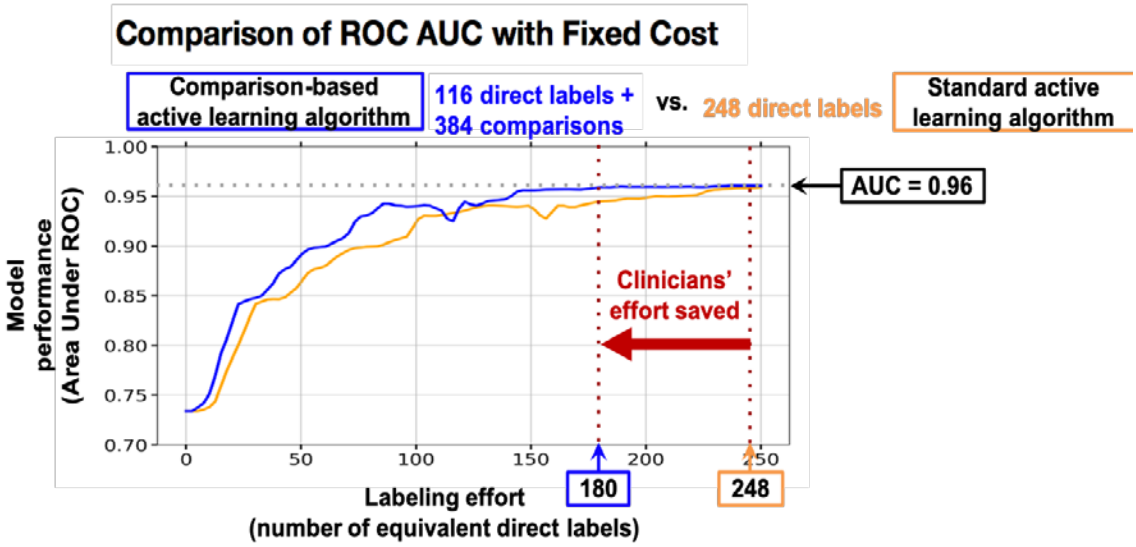


**Figure 3. Comparison of Receiver Operating Characteristic (ROC) Area Under the Curve (AUC) for the standard active learning algorithm and comparison-based learning algorithm with direct and comparison clinician feedback on patient data**

In another application, we considered the task of predicting the biological age of a person given his face image using the Apparent-Real age tagged face image (APPA-REAL) dataset which consists of 7,591 images, and each image is associated with a biological age and an apparent age. The biological age is the person's actual age, whereas the apparent ages are collected by asking crowdsourced workers to estimate their apparent ages. The images are divided into 4,063 train, 1,488 validation and 1,962 test samples. We use the apparent ages to get a ranking of the images, in addition to selective direct biological ages. The figure 4 below shows the error in estimating the biological image using direct queries only (5 and 10-Nearest Neighbor (NN)) and direct queries + ranking information (our method, Ranking-Regression ($R^2$), called $R^2$ 5-NN and $R^2$ 10-NN below). Again, we see significant improvement when using ranking information.

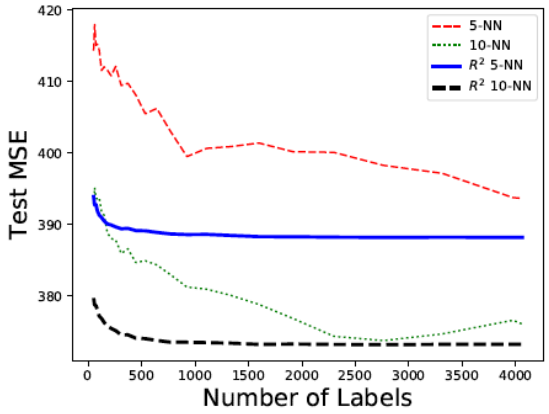


**Figure 4. Error in estimating the biological image using direct queries only (5 and 10-NN) and direct queries + ranking information (our method called R2 5-NN and R2 10-NN).**

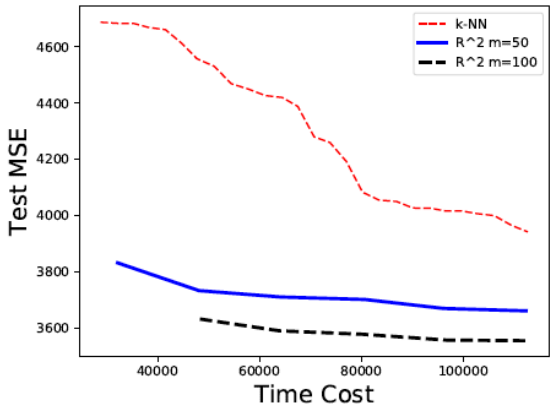


**Figure 5. Error in estimating the AirBnB price using direct queries only (k-NN) and direct queries + ranking information (our method called $R^2$)**

We also applied our methods to predict the price of an AirBnB listing. We use 357 listings as the training set and 93 as the test set. We collected direct labels and comparisons using mechanical turk. We observe that, in this case, both direct queries and comparisons take the same amount of time (probably as the time to read each posting is the dominant time factor), however comparisons were 10% more accurate than direct labels. We again observe that out method $R^2$ had a significant performance both boost over the best nearest neighbor regressor under the same total worker time (see figure 5 above, here m denotes number of direct queries).

Finally, we also report on some simulated and real data results for interactive optimization using deep neural networks. In the figure 6 below, we compare our method that uses influence functions to guide further sampling for optimizing a black-box synthetic function modeled by a neural network to competing Bayesian optimization approaches which use Gaussian process as the model. We observe that even though Neural Networks (NNs) require initial warm-up data to obtain a good initial model, further updates to that model using selective data quickly outperform competing GP based approaches (INF, LCB, MPI and EI), where INF denotes Influence Function, LCB denotes Lower Confidence Bound, MPI denotes Maximum Probability of Improvement and EI denotes Expected Improvement.

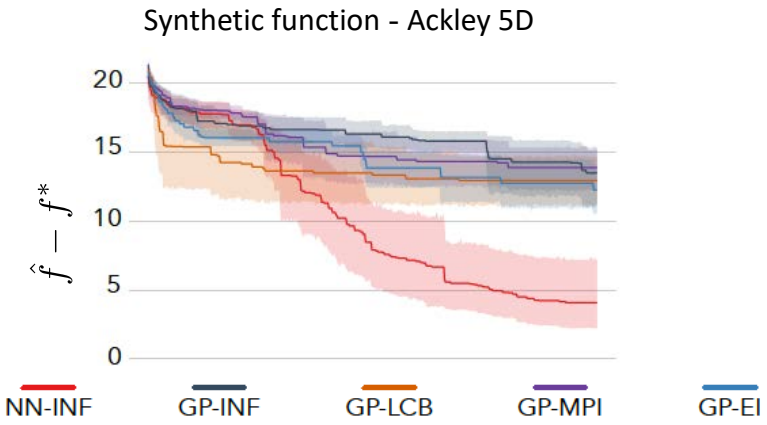


**Figure 6. Error achieved by our method (NN-INF) compared to competing GP based methods for a synthetic function**

Similarly, on a real world Modified National Institute of Standards and Technology database (MNIST) data of hand-written digits (figure 7 below), we see that the influence function (INF) driven neural network model (when pre-trained to classify digits 1 and 8) is able to quickly select input images to maximize the probability of chosen image to be of digit 3 with higher probability than competing GP optimization method (LCB).

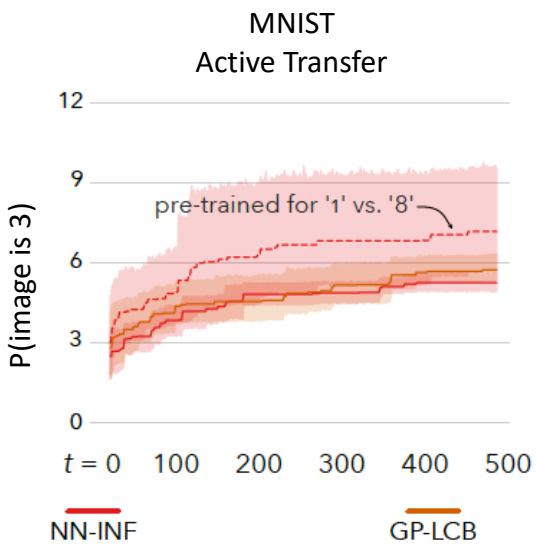


**Figure 7. Likelihood achieved by our method (NN-INF) compared to competing GP based method for MNIST dataset**

## 5.0 CONCLUSIONS

This project developed methods for interactive machine learning where instead of simply learning input-output mappings and associations given data that is collected a priori to any analysis, machine learning algorithms made high-level decisions about which data should be collected that is most informative to the learning task at hand. The methods we developed came with principled theoretical guarantees which illuminated the settings under which data-efficient learning was possible by precisely quantifying the sample complexity improvements over learning from passively collected datasets under reasonable assumptions. A wide variety of learning problems were considered ranging from density estimation, classification, regression, optimization and bandit problems. Furthermore, this was achieved by leveraging multiple diverse forms of interactions of the algorithms with the data-generating process including direct label queries, indirect comparisons and rankings, as well as unsupervised feature and similarity queries. However, such interactive analysis only scratches the surface of learning from diverse forms of interactive feedback – many other modes of interaction that govern human learning include relations, partial knowledge graphs, instructions, demos, etc. To achieve the next level of artificial intelligence it is critical to introduce a new paradigm of intelligent interactive machine learning algorithms that learn continually via such rich and complex feedback and make high-level decisions in collaboration with humans.

## 6.0 REFERENCES


[1] Noise-Tolerant Interactive Learning Using Pairwise Comparisons
Yichong Xu, Hongyang Zhang, Kyle Miller, Aarti Singh, and Artur Dubrawski
Conference on Neural Information Processing Systems (NIPS), 2017.

[2] Hypothesis Transfer Learning via Transformation Functions
S. Du, J. Koushik, A. Singh and B. Poczos.
Conference on Neural Information Processing Systems (NIPS), 2017.

[3] On Computationally Tractable Selection of Experiments in Regression Models.
Yining Wang, Adams Wei Yu and Aarti Singh.
Journal of Machine Learning Research, 18(143):1-41, 2017.

[4] Gradient Descent Can Take Exponential Time to Escape Saddle Points
S. Du, C. Jin, J. Lee, M. Jordan, B. Poczos and A. Singh.
Conference on Neural Information Processing Systems (NIPS), 2017.

[5] Revealing Actionable Simplicity in Data
Gisolfi N, Dubrawski A.
AAAI Symposium Designing the User Experience of Artificial Intelligence, Palo Alto, CA, March 2018.

[6] Internal Bleeding Detection by Multi-View Correlation Clustering of Central Venous Pressure
Lei E, Miller K, Pinsky MR, Dubrawski A.
NIPS Workshop on Machine Learning for Healthcare (ML4H), Long Beach, CA, December 2017.

[7] Nonparametric Regression with Comparisons: Escaping the Curse of Dimensionality with Ordinal Information
Y. Xu, H. Muthakana, S. Balakrishnan, A. Singh and A. Dubrawski.
International Conference on Machine Learning, ICML 2018.

[8] Gradient Descent Learns One-hidden-layer CNN: Don't be Afraid of Spurious Local Minima
S. Du, J. D. Lee, Y. Tian, B. Poczos and A. Singh.
International Conference of Machine Learning, ICML 2018.

[9] Transformation Autoregressive Networks
Oliva, J., Dubey, K., Zaheer, M., Salakhutdinov, R., Xing, E., Schneider, J., and Poczos, B.
International Conference on Machine Learning (ICML), 2018.

[10] RedThread: Active Search of Connections for Case Building and Combating Human Trafficking Online
Rabbany R, Bayani D, Dubrawski A.
Knowledge Discovery in Databases (KDD) 2018.

[11] Interactive Linear Regression with Pairwise Comparisons
Y. Xu, S. Balakrishnan, A. Singh, and A. Dubrawski
Asilomar Conference on Signals, Systems and Computers 2018, invited paper.

[12] A Generic Approach for Escaping Saddle points
Reddi, S., Zaheer, M., Sra, S., Poczos, B., Bach, F., Salakhutdinov, R., and Smola, A.
Artificial Intelligence and Statistics (AISTATS), 2018.

[13] Reconstruction Risk of Convolutional Sparse Dictionary Learning
Singh, S., Poczos, B., and Ma, J.

Artificial Intelligence and Statistics (AISTATS), 2018.
[14] Stochastic Zeroth-order Optimization in High Dimensions
Y. Wang, S. Du, S. Balakrishnan and A. Singh.
Artificial Intelligence and Statistics, (AISTATS), 2018.
[15] Regression with Comparisons: Escaping the Curse of Dimensionality with Ordinal Information
Y. Xu, H. Muthakana, S. Balakrishnan, A. Singh and A. Dubrawski.
Journal of Machine Learning Research (JMLR), *Accepted pending minor revision.*
[16] Optimization of Smooth Functions with Noisy Observations: Local Minimax Rates.
Y. Wang, S. Balakrishnan and A. Singh
Neural Information Processing Sytems (NIPS) 2018.
[17] Myopic Bayesian Design of Experiments via Posterior Sampling and Probabilistic Programming
Kirthevasan Kandasamy, Willie Neiswanger, Reed Zhang, Akshay Krishnamurthy, Jeff Schneider, Barnabas Poczos
[18] How Many Samples are Needed to Learn a Convolutional Neural Network?
S. Du**, Y. Wang**, X. Zhai, S. Balakrishnan, R. Salakhutdinov and A. Singh.
Neural Information Processing Sytems (NIPS) 2018.
[19] Using Comparisons to Reduce Cost of Data Annotation Required to Train Models for Bedside Monitoring.
Sheng J, Chen L, Xu Y, Pinsky MR, Hravnak M, Dubrawski A.
Critical Care Medicine 2019 (in press).
[20] Nonparametric Density Estimation under Besov IPM Losses.
Ananya Uppal, Shashank Singh, Barnabás Póczos.
Conference on Neural Information Processing Systems (NeurIPS), 2019.
[21] Zeroth Order Non-convex optimization with Dueling-Choice Bandits.
Yichong Xu, Aparna Joshi, Aarti Singh and Artur Dubrawski.
arXiv preprint arXiv:1911.00980, 2019.
[22] Thresholding Bandit Problem with both Duels and Pulls.
Yichong Xu**,** Xi Chen, Aarti Singh and Artur Dubrawski.
arXiv preprint arXiv:1910.06368, 2019.
Short version in NeurIPS2019 workshop on Learning with Rich Experience (LIRE).
[23] Deep Black-Box Optimization with Influence Functions.
Jayanth Koushik, Mike Tarr and Aarti Singh.
*In preparation.*

## LIST OF SYMBOLS, ABBREVIATIONS, AND ACRONYMS

| | |
|---|---|
| APPA-REAL | Apparent-Real age tagged face image dataset |
| AUC | Area Under the Curve |
| COMP-GP-UCB | Comparison based Gaussian Process with Upper Confidence Bound |
| EI | Expected Improvement |
| GP | Gaussian Process |
| INF | Influence function based approach |
| LCB | Lower Confidence Bound |
| MNIST | Modified National Institute of Standards and Technology database |
| MPI | Maximum Probability of Improvement |
| $R^2$ | Ranking Regression |
| ROC | Receiver Operating Characteristics |
| UCB | Upper Confidence Bound |