A Bandit Framework for Strategic Regression

Yang Liu and Yiling Chen
School of Engineering and Applied Science & CRCS, Harvard

Introduction

Training data for a series of regression problems are collected from strategic sources, such as via crowdsourcing or survey.

Such data suffers from quality issues:
- Intrinsic noise: different worker expertise
- Strategic noise: due to lack of monitoring, incentives etc.

How to control the quality for learning?
- Lack of prior knowledge of workers
- Lack of ground truth for quality verification
- Lack of monitoring: e.g., effort is not observable
- Lack of incentives

Objectives

Elicit high quality data for training regression model with performance guarantee.
- How to incentivize effort from workers?
- Any other incentives besides one-step payment?
- A robust mechanism or algorithm.
- Easy to implement.

Acknowledgment

We acknowledge the support of NSF grant CCF-1301976.

References


Strategic regression and our goal

The learner has a regression problem in mind.
- Assign data to each worker to label
- Targeting a good regression with training data collected from workers

Workers are effort sensitive, and can be incentivized via monetary payment
- Higher effort leads to smaller variance in data
- Utility function: Payment – Effort over tasks.

One-step payment function can be designed to pay each contributed data point to elicit effort [2,3]

Our approach: long-term incentive... instead of immediate payment

Future job opportunity
- In a stable market, workers care about future job opportunities.
- Or in other words, they care about their reputation.

Future selection ▶ Bandit (Multi-Armed Bandit)
- Each worker will be taken as an “arm”.
- Maintain online “score”, which the future selection will be based on
- But we do not directly observe workers’ performance => scoring rule

SR-UCB: a scoring rule aided UCB

Step 1. For each worker, train a reference estimator using data from other workers.

\[ \hat{f}_{\text{UCB}}(x) = \frac{1}{n} \sum_{i=1}^{n} y_i \]

Step 2. Then compute the following index for worker

\[ I_t(x) = \frac{1}{n} \sum_{i=1}^{n} (y_i - \hat{f}_{\text{UCB}}(x) - \frac{1}{n} \sum_{i=1}^{n} y_i)^2 \]

Step 3. Select the following workers to assign data

\[ I_t(x) \geq (\gamma + \beta) \sum_{i=1}^{n} (y_i - \hat{f}_{\text{UCB}}(x) - \frac{1}{n} \sum_{i=1}^{n} y_i)^2 \]

For workers with homogeneous expertise:
- Pay each worker
- Form a competition: Exerting efforts guarantees linear number of selections.

Lemma 1. If every worker exerts effort level \(e_i(t) = \gamma \), then there exists a constant \(b_i > 0\) such that for any \(i, j\) that \(i \neq j\), we have probability at least \(1 - O(1)\), \(n_i(t) \leq (1 + b_i) n_j(t)\).

Slack of: selection bounded sub-linearly

Theorem: Workers exerting \(e_i\) is an appropriate BNE.

With heterogeneous workers:
- Targeting the best two: sufficient and necessary
- Form competition & data is coming from the most competitive workers

Extension, computation & privacy

Ridge regression
- \( \hat{\beta} = \arg\min_{\beta} \sum_{i=1}^{n} (y_i - x_i^T \beta)^2 + \lambda \|\beta\|_2^2 \)
- Biased reference model
- To account for bias: a larger confidence term

Non linear regression
- Inspired by consistency of M-estimator.
- Works for parametric families such that \( \|\hat{f}_t(x) - f(x)\| \leq b \|\hat{f}_{t-1}(x) - \theta\| \)

Computational issues
- Online model update
- Online score update

Privacy preserving
- Preserving privacy in a sequential index system
- Partial sum idea: separate indexes into partial sums; only add noise to each partial sum.

What we achieve
- Show a bandit framework can help provide long-term incentive for such regression problems.
- A long-term, quantitative reputation system.
- Robust to different regression models & can be maintained efficiently.
- Preserve privacy in workers’ data.