Revisiting Perceptron: Efficient and Label-Optimal Learning of Halfspaces

Songbai Yan  yansongbai@eng.ucsd.edu  UC San Diego
Chicheng Zhang  chicheng.zhang@microsoft.com  UC San Diego -> Microsoft Research New York City

ABSTRACT

- We propose an efficient Perceptron-based algorithm for actively learning homogeneous halfspaces. Specifically:
  - Under the bounded noise condition, our algorithm achieves computational efficiency and label-optimality, improving over the state-of-the-art algorithms [1-3].
  - Under the adversarial noise condition, our algorithm achieves near-optimal label complexity and computational efficiency over the state-of-the-art method [2].
  - In addition, our algorithm can be converted to an efficient passive learning algorithm with near-optimal sample requirement.

SETTING

- Active Learning
  - Goal: Find an \( h \in H \) such that \( P_0(h(X) \neq Y) \) is small while making only a few label queries.

- Noise Models:
  - \( \eta \)-bounded noise (\( 0 \leq \eta < \frac{1}{2} \)): there is a halfspace \( \mathcal{H} \) such that for all \( x \), \( P_0(Y \neq \text{sign}(\langle u \cdot x \rangle | X = x) \) \( \leq \eta \).
  - \( \nu \)-adversarial noise (\( 0 \leq \nu < 1 \)): there is a halfspace \( \mathcal{H} \) such that \( P_0(Y \neq \text{sign}(\langle u \cdot x \rangle | X = x) \) \( \leq \nu \).

- Label Complexity: the number of labels required to output a halfspace \( \mathcal{H} \) such that \( P_0(\langle \text{sign}(\langle u \cdot x \rangle | \mathcal{H} \rangle \neq \text{sign}(\langle u \cdot x \rangle | X = x) \)) \( \leq \epsilon \).

RELATED WORK

- Noise-free (\( \eta = 0 \) or \( \nu = 0 \))
  - Efficient and label-optimal solutions have been proposed (e.g. [3,5]).

- Bounded noise
  - [3]: a margin-based algorithm which is label-optimal but computationally inefficient.
  - [1]: combining the idea of [3] and polynomial regression. Efficient but requires \( \tilde{O}(d^{(1-2\gamma)^*} \ln \frac{1}{\epsilon}) \) labels.

- Adversarial noise
  - [4]: learning halfspaces with agnostic noise is computationally hard with unbounded \( \nu \), even if the unlabeled distribution is uniform.

- [2]: computationally efficient and label-optimal algorithms that tolerate a noise level of \( \nu = \Theta(\epsilon) \).

REFERENCES


ALGORITHM

- Input: target error \( \epsilon \); Output: learned halfspace \( w \).
  1. Initialize \( w \) uniformly at random from the unit sphere.
  2. Set sample schedule \( m_0, b_k, k \geq 1 \).
  3. In phases \( k = 1, 2, \ldots, \lceil \log \frac{1}{\epsilon} \rceil \): Repeat \( m_k \) times:
     - Sample \( x \) from \( \mathcal{D}_k \) and query its label \( y \);
     - Perform modified Perceptron update [5]: \( w \leftarrow w - 2(yw \cdot x \leq 0)(w \cdot x)x \).
  4. Return \( w \).

- Sample Schedule:
  - (\( \eta \)-Bounded Noise): \( m_k = \tilde{O}(\frac{d}{\eta} \ln \frac{1}{\epsilon}) \), \( b_k = \tilde{O}(\frac{d}{\eta} \ln \frac{1}{\epsilon}) \).
  - (Adversarial Noise): \( m_k = \tilde{O}(d) \), \( b_k = \tilde{O}(d) \).

- The Modified Perceptron Update: \( w_{\text{new}} \leftarrow w_{\text{old}} - 2(yw_{\text{old}} \cdot x \leq 0)(w_{\text{old}} \cdot x)x \)
  - Perceptron update with a careful tuning of step size
  - In the noiseless setting, the angle between \( w \) and \( w^* \) never increases; in the noisy setting, the angle never increases in expectation.

PERFORMANCE GUARANTEES

- \( \eta \)-Bounded Noise
  - Our algorithm has a lower running time than the state-of-the-art algorithms
  - \( O(\frac{d}{\eta} \ln \frac{1}{\epsilon}) \) labels.

- \( \nu \)-Adversarial Noise
  - Lower Bound: \( \Omega(\frac{d}{\nu} \ln \frac{1}{\epsilon}) \)

  - Our algorithm achieves optimal label complexity and computational efficiency simultaneously.

OPEN PROBLEMS

- Design efficient and label-optimal halfspace learning algorithms that:
  - adapt to unknown bounded noise parameter \( \eta \)
  - Work under broader unlabeled distributions, e.g. log-concave distributions
  - Work under weaker noise assumptions, e.g. Tsybakov noise condition