Optimal Teaching for Online Perceptrons

Xuezhou Zhang, Hrag Gorune Ohannessian, Ayon Sen,
Scott Alfeld, Xiaojin Zhu

Department of Computer Sciences, University of Wisconsin-Madison

December 9, 2016

Optimal Teaching Problem

- **Student**: A machine learner $\mathcal{A}$.
- **Teacher**: A person who knows a target model θ^*, and wants to teach it to the student $\mathcal{A}$ by creating a training set $\mathcal{D}$.
- **Goal**: Find the 'best' training set.
- **General Optimization Formulation**:

$$\min_{\mathcal{D}} \quad \text{loss}(\mathcal{A}(\mathcal{D}), \theta^*) + \text{effort}(\mathcal{D})$$

Alternatively,

$$\begin{aligned} \min_{\mathcal{D}} \quad & \text{effort}(\mathcal{D}), \\ \text{s.t.} \quad & \text{loss}(\mathcal{A}(\mathcal{D}), \theta^*) \leq \epsilon \end{aligned}$$

Aim of this line of work

- Extend optimal teaching problem to **sequential** learners.
- Explore the teaching setting of **uncertainty**, i.e. the lack of information on the teacher's side.

Motivating Example: Online Perceptrons

Algorithm 1 Online Perceptron

```
Learning parameters: Initial weight vector w_0 ∈ R^d, learning rate η > 0.
for t = 1 ... do
    receive x_t
    predict ŷ_t = sign(⟨w_{t−1}, x_t⟩)
    receive y_t
    w_t ← w_{t−1} + 1_{(y_t⟨x_t, w_{t−1}⟩ ≤ 0)} η y_t x_t
```

- **General Setting**: Allow non-zero $\mathbf{w}_0$ and arbitrary learning rate η.
- **Formulation**: In this example, the machine learner $\mathcal{A}$ is the perceptron, and model θ is the linear decision boundary represented by the parameter **w**.

Teaching with Full Knowledge of the Perceptron

Definition

The **Exact Teaching dimension** of perceptron is defined as

$$\begin{aligned} \arg\min_{\mathcal{D}} \quad & |\mathcal{D}|, \\ \text{s.t.} \quad & \mathcal{A}(\mathcal{D}) = \mathbf{w}^* \end{aligned}$$

Theorem

For any target parameter $\mathbf{w}^*$*, a perceptron with any initial weight* $\mathbf{w}_0$ *and learning rate* η *has exact teaching dimension* 1*.*

Approximate Teaching with Unknown W0

Definition

The ϵ-**Approximate Teaching dimension** of perceptron is defined as

$$\arg\min_{\mathcal{D}} \quad |\mathcal{D}|,$$

$$\text{s.t.} \quad \frac{\langle \mathcal{A}(\mathcal{D}), \mathbf{w}^* \rangle}{\|\mathcal{A}(\mathcal{D})\| \|\mathbf{w}^*\|} \geq 1 - \epsilon$$

Theorem

For any target parameter $\mathbf{w}^*$ *and precision* ϵ*, a perceptron with unknown initial weight* $\mathbf{w}_0$ *and known learning rate* η *has* ϵ*-approximate teaching dimension* 3*.*

Discussion and Future Work

- An ‘**interactive**’ or ‘**collaborative**’ learning setting, where the student tries to learn the target model, while the teacher learns to teach.
- One potential solution is through **active learning**. Here the teacher can be formulated as an active learner who learns by probing the student and receive its feedback.
- However, this does not capture the interactive nature in teaching, and only optimizes the teacher’s learning task.
- Better solution?