程序代写代做 CS编程辅导

CS861: Theoretical Foundations of Machine Learning Lecture 22 - 10/25/2023

University of Wi[illegible] Fall 2023

Lect[illegible]earning, The experts problem

Lecturer: Kirthe[illegible] *Scribed by: Xinyan Wang, Zhifeng Chen*

Disclaimer: *These n[illegible]ected to the usual scrutiny reserved for formal publications. They may be distributed outside this class only with the permission of the instructor.*

In this lecture, we will introduce Online Learning and the experts problems. We will first complete the proof of the martingale concentration result from the last lecture.

WeChat: cstutorcs

Proof Pick a round $t \in \{d+1, ..., T\}$ and any $a \in \mathcal{A}$. By the L-Lipschitz property of f, We know

$$\left|f(\theta_*^T a) - f(\hat{\theta}_t^T a)\right| \leq L\left|(\theta_* - \hat{\theta}_t)^T a\right| \tag{1}$$

Assignment Project Exam Help

Now we bound $\theta_* - \hat{\theta}_t$. Using the assumption that $f' \geq c$, we have

$$\nabla g_{t-1}(\theta) = \sum_{s=1}^{t-1} A_s A_s^T f'(A_s^T \theta) \geq c \sum_{s=1}^{t-1} A_s A_s^T \tag{2}$$

Email: tutorcs@163.com

$$\geq cI \tag{3}$$

As f' is continuous, by the fundamental theorem of calculus,

QQ: 749389476

$$g_{t-1}(\theta_*) - g_{t-1}(\hat{\theta}_{t-1}) = G_{t-1} * (\theta_* - \hat{\theta}_{t-1})$$

Where $G_{t-1} = \int_0^1 \nabla g_{t-1}\left(s\theta_* + (1-s)\hat{\theta}_{t-1}\right) ds$.

By (3), G_{t-1} is invertible $\Rightarrow (\theta_t - \hat{\theta}_{t-1}) = G_{t-1}^{-1}\left(g_{t-1}(\theta_t) - g_{t-1}(\hat{\theta}_{t-1})\right)$. We therefore have,

https://tutorcs.com

$$\begin{aligned}
\left|(\theta_\star - \hat{\theta}_t)^T a\right| &= \left|\left(g_{t-1}(\theta_\star) - g_{t-1}(\hat{\theta}_{t-1})\right)^T G_{t-1}^{-1} a\right| \\
&\leq \left\|g_{t-1}(\theta_\star) - g_{t-1}(\hat{\theta}_{t-1})\right\|_{G_{t-1}^{-1}} \|a\|_{G_{t-1}^{-1}}^{-1}, \text{ as } \left\|A^T B\right\| \leq \|A\| \|B\| \\
&= \frac{1}{c}\left\|g_{t-1}(\theta_\star) - g_{t-1}(\hat{\theta}_{t-1})\right\|_{V_{t-1}^{-1}} \|a\|_{V_{t-1}^{-1}}, \text{ as } G_{t-1} \geq cV_{t-1} \Rightarrow G_{t-1}^{-1} \leq \frac{1}{c} V_{t-1} \\
&\leq \frac{2}{c}\left\|\sum_{s=1}^{t-1} A_S \epsilon_s\right\|_{V_{t-1}^{-1}} \|a\|_{V_{t-1}^{-1}} \qquad (4)
\end{aligned}$$

We will apply the martingale concentration result with $t_0 = d, V_{t-1} = \sum_{s=1}^{t-1} A_s A_s^T, c^2 = \max_{a \in A} a^T a = d$ and finally $\delta = 1/T^2$. Then, with probability $\geq 1 - T^2$, by (1) and (4)

$$\forall a \in \mathcal{A}, \left|f(\theta_\star^T a) - f(\hat{\theta}_{t-1}^T a)\right| \leq \underbrace{\frac{2L\sigma\gamma}{c}\sqrt{2d\log(t)\log(dT^2)}\,\|a\|_{V_{t-1}^{-1}}}_{\rho(t)}.$$

Applying a union bound over all $t \in d+1, ..., T$ we have that $\forall a \in \mathcal{A}$ and $\forall t \in d+1, ..., T$,

$$\left|f(\theta_\star^T a) - f(\hat{\theta}_{t-1}^T a)\right| \leq \rho(t)\,\|a\|_{V_{t-1}}^{-1}.$$

程序代写代做 CS编程辅导

□

1 The Expert

To motivate the ensuin with two examples.

Example 1 (Online spa hypothesis class $\mathcal{H}$ of binary classifiers, where $\mathcal{H} \in \{h : \mathcal{X} \to \{0,1\}\}$. Consider the fo unds:

1. A learner receives an input email $n_t \in \mathcal{X}$ on round t.
2. The learner chooses some $h_t \in \mathcal{H}$ and predicts $h_t(n_t)$ (spam or not-spam).
3. Learner sees the label y_t and incurs loss $1\{h_t(n_t) \neq y_t\}$.

WeChat: cstutorcs

Note that the learner knows the loss for all $h \in \mathcal{H}$.

Example 2 (Weather forecasting). Given a set of models $\mathcal{H}$. Consider the following game over T rounds:

Assignment Project Exam Help

1. Learner (weather forecaster) chooses some model $h \in \mathcal{H}$ and predicts the number $\hat{y}_t$.
2. Learner observes the true weather y_t and incurs loss $l(y_t, \hat{y}_t)$.

We can now introduce **Expert Problem**, which proceeds over T rounds in the following fashion:

Email: tutorcs@163.com

1. We are given a set of K experts, denoted $[K]$.
2. On each round, the learner chooses an expert(action) $A_t \in [K]$. Simultanously, the environment picks a loss vector $\ell_t \in [0,1]^K$, where $\ell_t(i)$ is the loss for expert i.
3. Learner incurs loss $\ell_t(A_t)$.
4. Learner observes ℓ_t (losses for all experts).

QQ: 749389476

https://tutorcs.com

This type of feedback, where we observe feedback for all actions is called full information feedback. In contrast, when we observe losses or rewards only for the action we took, it is called bandit feedback.

Unlike in the stochastic bandit setting, we will **not** assume that the loss vectors are drawn from some distribution. Then how do we define regret? Recall that in the stochastic setting, we let $a_\star = \arg\min_{i \in [K]} \mathbb{E}_{X \sim \nu_i}[X]$ be the action with the highest expected reward and defined the regret as follows:

$$R_T^{Stochastic}(\pi, \nu) = \mathbb{E}\left[\sum_{t=1}^{T} X_t\right] - T\,\mathbb{E}_{X \sim V_{G^\star}}[X]$$

We did this for bandit feedback, but can define the regret similarly for full information feedback.

Here, in the non-stochastic setting, where loss vectors are arbitrary, we will compete against the best fixed action in hindsight. For a policy π, and a sequence of losses, $\ell =_1, ..., \ell_t$, define

$$R_T'(\pi, \ell) = \sum_{t=1}^{T} \ell_t(A_t) = \min_{a \in [K]} \sum_{t=1}^{T} \ell_t(a)$$

For a stochastic policy, we will consider

$$R_T(\pi, \ell) = \mathbb{E}\left[R_T'(\pi, \underline{l})\right] = \mathbb{E}\left[\sum_{t=1}^{T} \ell_t(A_t)\right] - \min_{a \in [K]} \sum_{t=1}^{T} \ell_t(a),$$

程序代写代做 CS编程辅导

where $\mathbb{E}$ is with respect to the randomness of the policy.

For a given policy π, we wish to bound $R_T(\pi, \underline{l})$ for all loss sequences. That is $\sup_{\underline{l}} R_T(\pi, \underline{l})$. We wish to do well even if the losse[…] **adversary** which had full knowledge of our policy π. Here, we are concerned with [… w]ho can choose ℓ_t to only be a function of the current action, and not previous action[…]

2 The Hedge […]

The most intuitive app[…]lem is to choose the action $A_t = \arg\min_{a \in [K]} \sum_{s=1}^{t-1} l_s(a)$ on round t. This is called […](FTL). For instance, for binary classification example, this would simply be empirical risk minimization, as we will choose

WeChat: cstutorcs

$$h_t = \arg\min_{h \in \mathcal{H}} \sum_{s=1}^{t-1} \mathbb{1}(h(x_t) \neq y_t)$$

Unfortunately, this does not work. To see why, suppose $K = 2$, and define the loss vectors as follows:

Assignment Project Exam Help

$$\ell_t = \begin{cases} (0.5, 0) & \text{if } t = 1 \\ (1, 0) & \text{if } t \text{ is odd and } t > 1 \\ (0, 1) & \text{if } t \text{ is even} \end{cases}$$

Then, FTL will choo[se]

Email: tutorcs@163.com

$$A_t = \begin{cases} 1 & \text{on odd rounds} \\ 2 & \text{on even rounds} \end{cases}$$

Then the total loss of FTL will be at least $T - 1$ while the best action in hindsight will have loss at most $T/2$. Hence, the regret of FTL is at least $T/2 - 1 \in \Omega(T)$.

QQ: 749389476

In the Hedge algorithm, we will instead use a soft version of the minimum, where we will sample from a distribution which samples arms with small losses more frequently. We hae summarized the Hedge algorithm below.

https://tutorcs.com

Algorithm 1 The Hedge Algorithm(a.k.a multiplicative weights, a.k.a exponential weights)

Given time horizon T, learning rate η
Let $L_0 \leftarrow \underline{0}_K$ (all zero vector in $\mathbb{R}^K$)
for $t = 1, ..., T$ **do**
  Set $P_t(a) \leftarrow \frac{e^{-2L_{t-1}(a)}}{\sum_{j=1}^{K} e^{-2L_{t-1}(j)}}, \forall a \in [K]$
  Sample $A_t \sim P_t$ (note that $P_t \in \Delta^K$)
  Incur loss $\ell_t(A_t)$, observe ℓ_t
  Update $\ell_t(a) \leftarrow L_{t-1}(a) + \ell_t(a), \forall a \in [K]$
end for