

PAC Learning and Online Learning

Dimitris Fotakis

SCHOOL OF ELECTRICAL AND COMPUTER ENGINEERING
NATIONAL TECHNICAL UNIVERSITY OF ATHENS, GREECE

PAC Learning

Domain \mathcal{X} , binary labels $\mathcal{Y} = \{-1, +1\}$,

hypothesis class $\mathcal{H} = \{h : (\mathcal{X} \rightarrow \mathcal{Y})\}$

(Fixed unknown) distribution \mathcal{D} over domain \mathcal{X}

Labeled training data $(x_1, y_1), \dots, (x_m, y_m) \in \mathcal{X} \times \mathcal{Y}$

The training set **distributed** according to \mathcal{D} : $S = (x_1, \dots, x_m) \sim \mathcal{D}^m$

Realizability assumption: $\exists f \in \mathcal{H}$ that **correctly** determines the **labels** of all $x \in \mathcal{X}$, i.e., $\forall x_i \in \mathcal{X}, y_i = f(x_i)$.

PAC Learning

Domain \mathcal{X} , binary labels $\mathcal{Y} = \{-1, +1\}$,
hypothesis class $\mathcal{H} = \{h : (\mathcal{X} \rightarrow \mathcal{Y})\}$

(Fixed unknown) distribution \mathcal{D} over domain \mathcal{X}

Labeled training data $(x_1, y_1), \dots, (x_m, y_m) \in \mathcal{X} \times \mathcal{Y}$

The training set **distributed** according to \mathcal{D} : $S = (x_1, \dots, x_m) \sim \mathcal{D}^m$

Realizability assumption: $\exists f \in \mathcal{H}$ that **correctly** determines the **labels** of all $x \in \mathcal{X}$, i.e., $\forall x_i \in \mathcal{X}, y_i = f(x_i)$.

Loss of hypothesis $h \in \mathcal{H}$: $L_{\mathcal{D},f}(h) = \mathbb{P}_{x \sim \mathcal{D}}[h(x) \neq f(x)]$

Class \mathcal{H} is **PAC learnable** if for all ε, δ , there is # samples = $m_{\mathcal{H}}(\varepsilon, \delta)$ and algorithm A so that for any $m \geq m_{\mathcal{H}}(\varepsilon, \delta)$, \mathcal{D} and f ,

$$\mathbb{P}_{S \sim \mathcal{D}^m} [L_{\mathcal{D},f}(A(S)) \leq \varepsilon] \geq 1 - \delta$$

Empirical Risk Minimization (ERM): output any hypothesis h **not suffering any loss** on S (recall realizability!)

PAC Learning

Domain \mathcal{X} , binary labels $\mathcal{Y} = \{-1, +1\}$,
hypothesis class $\mathcal{H} = \{h : (\mathcal{X} \rightarrow \mathcal{Y})\}$

(Fixed unknown) distribution \mathcal{D} over domain \mathcal{X}

Labeled training data $(x_1, y_1), \dots, (x_m, y_m) \in \mathcal{X} \times \mathcal{Y}$

The training set **distributed** according to \mathcal{D} : $S = (x_1, \dots, x_m) \sim \mathcal{D}^m$

Realizability assumption: $\exists f \in \mathcal{H}$ that **correctly** determines the **labels** of all $x \in \mathcal{X}$, i.e., $\forall x_i \in \mathcal{X}, y_i = f(x_i)$.

Loss of hypothesis $h \in \mathcal{H}$: $L_{\mathcal{D},f}(h) = \mathbb{P}_{x \sim \mathcal{D}}[h(x) \neq f(x)]$

Class \mathcal{H} is **PAC learnable** if for all ε, δ , there is # samples = $m_{\mathcal{H}}(\varepsilon, \delta)$ and algorithm A so that for any $m \geq m_{\mathcal{H}}(\varepsilon, \delta)$, \mathcal{D} and f ,

$$\mathbb{P}_{S \sim \mathcal{D}^m} [L_{\mathcal{D},f}(A(S)) \leq \varepsilon] \geq 1 - \delta$$

Empirical Risk Minimization (ERM): output any hypothesis h **not suffering any loss** on S (recall realizability!)

VC dimension :

- \mathcal{H} **shatters** $C \subseteq \mathcal{X}$ if each of the $2^{|C|}$ possible labelings of C can be produced by some $h \in \mathcal{H}$.
- VC dimension of $\mathcal{H} = \sup\{|C| : \mathcal{H} \text{ shatters } C\}$

Agnostic PAC Learning

Domain \mathcal{X} , labels \mathcal{Y} , hypothesis class $\mathcal{H} = \{h : (\mathcal{X} \rightarrow \mathcal{Y})\}$

(Fixed unknown) distribution \mathcal{D} over $\mathcal{X} \times \mathcal{Y}$

Training set $S = \{(x_1, y_1), \dots, (x_m, y_m)\} \sim \mathcal{D}^m$

Loss of hypothesis $h \in \mathcal{H}$: $L_{\mathcal{D}}(h) = \mathbb{P}_{(x,y) \sim \mathcal{D}}[h(x) \neq y]$

Class \mathcal{H} is **agnostically PAC learnable** if for all ε, δ , there is #samples $= m_{\mathcal{H}}(\varepsilon, \delta)$ and algorithm A so that for any $m \geq m_{\mathcal{H}}(\varepsilon, \delta)$ and \mathcal{D} ,

$$\mathbb{P}_{S \sim \mathcal{D}^m} \left[L_{\mathcal{D}}(A(S)) \leq \varepsilon + \min_{f \in \mathcal{H}} L_{\mathcal{D}}(f) \right] \geq 1 - \delta$$

Empirical Risk Minimization (ERM): $\arg \min_{h \in \mathcal{H}} L_S(h)$

Uniform convergence: ERM on $\frac{\varepsilon}{2}$ -representative training sets

For finite hypothesis class \mathcal{H} , $\lceil \frac{2 \ln(2|\mathcal{H}|/\delta)}{\varepsilon^2} \rceil$ samples suffice for $\frac{\varepsilon}{2}$ -representative training set.

Online Learning: Setting

Two actions: H and L (binary classification).

On each day $t = 1, \dots, T$:

- 1 Learner **picks action** $i_t \in \{H, L\}$
- 2 Adversary **picks loss** vector $\ell_t = (\ell_t^H, \ell_t^L) \in [0, 1]^2$
- 3 Learner learns ℓ_t and **incurs loss** $\ell_t^{i_t}$

Online Learning: Setting

Two actions: H and L (binary classification).

On each day $t = 1, \dots, T$:

- 1 Learner **picks action** $i_t \in \{H, L\}$
- 2 Adversary **picks loss** vector $\ell_t = (\ell_t^H, \ell_t^L) \in [0, 1]^2$
- 3 Learner learns ℓ_t and **incurs loss** $\ell_t^{i_t}$

Goal is to minimize **regret** (loss wrt. **best fixed** action in hindsight):

$$\text{Regret}(T) = \sup_{\ell_1, \dots, \ell_T} \left(\sum_{t=1}^T \ell_t^{i_t} - \min_{i \in \{H, L\}} \sum_{t=1}^T \ell_t^i \right)$$

(Online learning) algorithm is **no-regret** if $\text{Regret}(T)/T \rightarrow 0$ at $T \rightarrow \infty$

Online Learning: Follow the Leader

Follow the Leader (FTL):

$$i_t = \arg \min_{i \in \{H, L\}} \sum_{\tau=1}^{t-1} \ell_{\tau}^i$$

Online Learning: Follow the Leader

Follow the Leader (FTL):

$$i_t = \arg \min_{i \in \{H,L\}} \sum_{\tau=1}^{t-1} \ell_{\tau}^i$$

Two obvious caveats with FTL:

- 1 Deterministic action choice, given the past (randomness always helps against the unknown).
- 2 Action choices can be very unstable (different choice each day).

Online Learning: Follow the Leader

Follow the Leader (FTL):

$$i_t = \arg \min_{i \in \{H, L\}} \sum_{\tau=1}^{t-1} \ell_{\tau}^i$$

Two obvious caveats with FTL:

- 1 Deterministic action choice, given the past (randomness always helps against the unknown).
- 2 Action choices can be very unstable (different choice each day).

Lower bound: Any deterministic algorithm has **linear**, i.e., $\Omega(T)$, regret.

Online Learning: Follow the Leader

Follow the Leader (FTL):

$$i_t = \arg \min_{i \in \{H,L\}} \sum_{\tau=1}^{t-1} \ell_{\tau}^i$$

Two obvious caveats with FTL:

- 1 Deterministic action choice, given the past (randomness always helps against the unknown).
- 2 Action choices can be very unstable (different choice each day).

Lower bound: Any deterministic algorithm has **linear**, i.e., $\Omega(T)$, regret.

Proof: loss for action i_t (chosen by the algorithm) = 1, and loss for other action = 0.

Any deterministic algorithm incurs loss = T , while best action incurs loss $\leq T/2$.

Online Learning: Randomization

Two actions: H and L (binary classification).

On each day $t = 1, \dots, T$:

- 1 Learner picks action H with probability p_t (and L with probability $1 - p_t$).
- 2 Adversary picks loss vector $\ell_t = (\ell_t^H, \ell_t^L) \in [0, 1]^2$
- 3 Learner learns ℓ_t and incurs **expected loss**

$$f(p_t; \ell_t) = p_t \ell_t^H + (1 - p_t) \ell_t^L$$

Online Learning: Randomization

Two actions: H and L (binary classification).

On each day $t = 1, \dots, T$:

- 1 Learner picks action H with probability p_t (and L with probability $1 - p_t$).
- 2 Adversary picks loss vector $\ell_t = (\ell_t^H, \ell_t^L) \in [0, 1]^2$
- 3 Learner learns ℓ_t and incurs **expected loss**

$$f(p_t; \ell_t) = p_t \ell_t^H + (1 - p_t) \ell_t^L$$

Goal is to minimize **expected regret**:

$$\text{Exp-Regret}(T) = \sup_{\ell_1, \dots, \ell_T} \left(\sum_{t=1}^T f(p_t; \ell_t) - \min_{p \in [0, 1]} \sum_{t=1}^T f(p; \ell_t) \right)$$

Randomization potentially allows for improved **stability**.

Online Learning: (Randomized) Follow the Leader

Follow the Leader (FTL):

$$p_t = \arg \min_{p \in [0,1]} \sum_{\tau=1}^{t-1} f(p; \ell_\tau) = \arg \min_{p \in [0,1]} F_{t-1}(p)$$

Online Learning: (Randomized) Follow the Leader

Follow the Leader (FTL):

$$p_t = \arg \min_{p \in [0,1]} \sum_{\tau=1}^{t-1} f(p; \ell_\tau) = \arg \min_{p \in [0,1]} F_{t-1}(p)$$

Is randomized FTL **really different** from deterministic FTL?

Online Learning: (Randomized) Follow the Leader

Follow the Leader (FTL):

$$p_t = \arg \min_{p \in [0,1]} \sum_{\tau=1}^{t-1} f(p; \ell_\tau) = \arg \min_{p \in [0,1]} F_{t-1}(p)$$

Is randomized FTL **really different** from deterministic FTL?

For any loss sequence ℓ_1, \dots, ℓ_T , FTL has:

$$\text{Exp-Regret}_{FTL}(T) = \underbrace{\sum_{t=1}^T f(p_t; \ell_t) - \min_{p \in [0,1]} \sum_{t=1}^T f(p; \ell_t)}_{\text{expected regret}} \leq \underbrace{\sum_{t=1}^T |p_t - p_{t+1}|}_{\text{stability}}$$

Online Learning: (Randomized) Follow the Leader

Follow the Leader (FTL):

$$p_t = \arg \min_{p \in [0,1]} \sum_{\tau=1}^{t-1} f(p; \ell_\tau) = \arg \min_{p \in [0,1]} F_{t-1}(p)$$

Is randomized FTL **really different** from deterministic FTL?

For any loss sequence ℓ_1, \dots, ℓ_T , FTL has:

$$\text{Exp-Regret}_{FTL}(T) = \underbrace{\sum_{t=1}^T f(p_t; \ell_t) - \min_{p \in [0,1]} \sum_{t=1}^T f(p; \ell_t)}_{\text{expected regret}} \leq \underbrace{\sum_{t=1}^T |p_t - p_{t+1}|}_{\text{stability}}$$

For the analysis, we define **Be the Leader** (BTL):

$$p_t^* = \arg \min_{p \in [0,1]} \sum_{\tau=1}^t f(p; \ell_\tau) = \arg \min_{p \in [0,1]} F_t(p)$$

Regret of FTL Against BTL

Lemma: For any loss sequence ℓ_1, \dots, ℓ_T ,

$$\text{Regret}_{FTL}(T) \leq \text{Regret}_{BTL}(T) + \underbrace{\sum_{t=1}^T |p_t - p_{t+1}|}_{\text{stability}}$$

Regret of FTL Against BTL

Lemma: For any loss sequence ℓ_1, \dots, ℓ_T ,

$$\text{Regret}_{\text{FTL}}(T) \leq \text{Regret}_{\text{BTL}}(T) + \underbrace{\sum_{t=1}^T |p_t - p_{t+1}|}_{\text{stability}}$$

$$\begin{aligned} \sum_{t=1}^T f(p_t; \ell_t) &= \sum_{t=1}^T f(p_t^*; \ell_t) + \sum_{t=1}^T (f(p_t; \ell_t) - f(p_t^*; \ell_t)) \\ &= \sum_{t=1}^T f(p_t^*; \ell_t) + \sum_{t=1}^T (p_t - p_t^*)(\ell_t^H - \ell_t^L) && \text{by dfn of } f(p_t; \ell_t) \\ &\leq \sum_{t=1}^T f(p_t^*; \ell_t) + \sum_{t=1}^T |p_t - p_t^*| && \text{losses } \ell_t \in [0, 1]^2 \\ &= \sum_{t=1}^T f(p_t^*; \ell_t) + \sum_{t=1}^T |p_t - p_{t+1}| && \text{by dfn, } p_t^* = p_{t+1} \end{aligned}$$

Regret of Be the Leader

Lemma: For any loss sequence ℓ_1, \dots, ℓ_T , $\text{Regret}_{\text{BTL}}(T) \leq 0$

Regret of Be the Leader

Lemma: For any loss sequence ℓ_1, \dots, ℓ_T , $\text{Regret}_{\text{BTL}}(T) \leq 0$

By **induction** on t , we show that for any $t \geq 1$:

$$\underbrace{\sum_{\tau=1}^t f(p_\tau^*; \ell_t)}_{\text{loss of BTL up to } t} \leq \underbrace{\min_{p \in [0,1]} F_t(p)}_{\text{loss of best fixed action up to } t} = \underbrace{F_t(p_t^*)}_{\text{by definition of } p_t^*}$$

Regret of Be the Leader

Lemma: For any loss sequence ℓ_1, \dots, ℓ_T , $\text{Regret}_{\text{BTL}}(T) \leq 0$

By **induction** on t , we show that for any $t \geq 1$:

$$\underbrace{\sum_{\tau=1}^t f(p_{\tau}^*; \ell_{\tau})}_{\text{loss of BTL up to } t} \leq \underbrace{\min_{p \in [0,1]} F_t(p)}_{\text{loss of best fixed action up to } t} = \underbrace{F_t(p_t^*)}_{\text{by definition of } p_t^*}$$

$$\begin{aligned} \sum_{\tau=1}^{t+1} f(p_{\tau}^*; \ell_{\tau}) &= f(p_{t+1}^*; \ell_{t+1}) + \sum_{\tau=1}^t f(p_{\tau}^*; \ell_{\tau}) \\ &\leq f(p_{t+1}^*; \ell_{t+1}) + \min_{p \in [0,1]} F_t(p) && \text{induction hypth.} \\ &\leq f(p_{t+1}^*; \ell_{t+1}) + F_t(p_{t+1}^*) && F_t(p_t^*) \leq F_t(p_{t+1}^*) \\ &= F_{t+1}(p_{t+1}^*) && \text{by dfn of } F_{t+1}(p) \end{aligned}$$