Online Learning Summer School Copenhagen 2015 Lecture 3

Shai Shalev-Shwartz

School of CS and Engineering,
The Hebrew University of Jerusalem

Online Learning

Recall: The Online Classification Game

For $t = 1, 2, \ldots$

- Environment presents a question x_t
- Learner predicts an answer $\hat{y}_t \in \{\pm 1\}$
- Environment reveals true label $y_t \in \{\pm 1\}$
- Learner pays 1 if $\hat{y}_t \neq y_t$ and 0 otherwise

Recall: The Online Classification Game

For $t = 1, 2, \ldots$

- Environment presents a question x_t
- Learner predicts an answer $\hat{y}_t \in \{\pm 1\}$
- Environment reveals true label $y_t \in \{\pm 1\}$
- Learner pays 1 if $\hat{y}_t \neq y_t$ and 0 otherwise

Goal of the learner: Make few mistakes

When can we guarantee to make few mistakes (or to have low regret w.r.t. $\mathcal{H}$)?

Online Learnability

When can we guarantee to make few mistakes (or to have low regret w.r.t. $\mathcal{H}$)?

For finite $\mathcal{H}$:

- Realizable case: Halving makes at most $\log(\mathcal{H})$ mistakes
- Non-realizable case: Weighted Majority (normalized-EG with the "randomization" approach) have expected regret of $\sqrt{\log(d)\,T}$

Realizable Case (no noise)

Realizable Assumption: Environment answers $y_t = h(\mathbf{x}_t)$, where $h \in \mathcal{H}$ and the hypothesis class, $\mathcal{H}$, is known to the learner

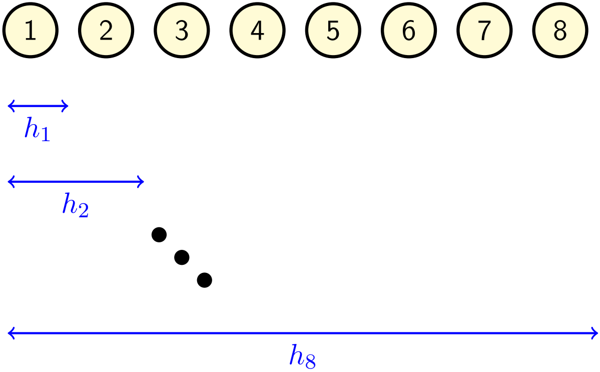
Theorem (Littlestone'88)

A combinatorial dimension, $\mathrm{Ldim}(\mathcal{H})$*, characterizes online learnability:*

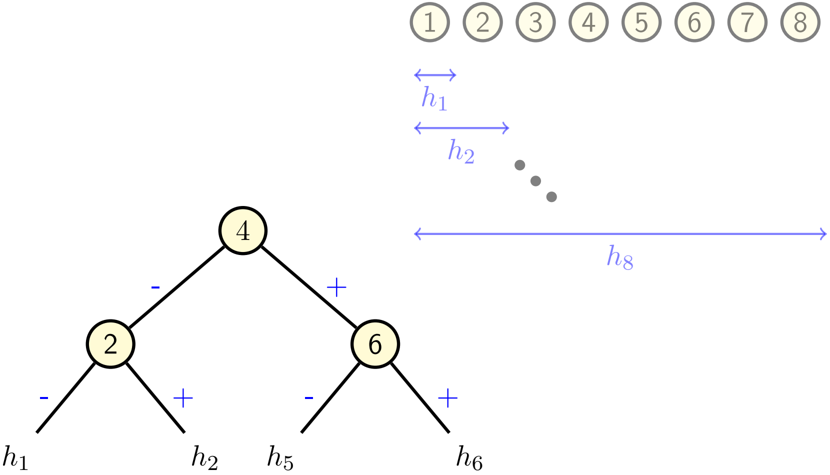
- *Any algorithm might make at least* $\mathrm{Ldim}(\mathcal{H})$ *mistakes*
- *Exists algorithm that makes at most* $\mathrm{Ldim}(\mathcal{H})$ *mistakes*

$\mathrm{Ldim}(\mathcal{H}) \leq \log(|\mathcal{H}|)$ and the gap can be substantial

Littlestone's dimension – Motivation



Littlestone's dimension – Motivation



Littlestone's dimension

Definition

$\mathrm{Ldim}(\mathcal{H})$ is the maximal depth of a full binary tree such that each path is "explained" by some $h \in \mathcal{H}$

Lemma

Any learner can be forced to make at least $\mathrm{Ldim}(\mathcal{H})$ *mistakes*

Proof.

Adversarial environment will "walk" on the tree, while on each round setting $y_t = -\hat{y}_t$. □

Standard Optimal Algorithm (SOA)

initialize: $V_1 = \mathcal{H}$
for $t = 1, 2, \ldots$
  receive $\mathbf{x}_t$
  for $r \in \{0, 1\}$ let $V_t^{(r)} = \{h \in V_t : h(\mathbf{x}_t) = r\}$
  predict $\hat{y}_t = \arg\max_r \mathrm{Ldim}(V_t^{(r)})$
  receive true answer y_t
  update $V_{t+1} = V_t^{(y_t)}$

Standard Optimal Algorithm (SOA)

initialize: $V_1 = \mathcal{H}$
for $t = 1, 2, \ldots$
  receive $\mathbf{x}_t$
  for $r \in \{0, 1\}$ let $V_t^{(r)} = \{h \in V_t : h(\mathbf{x}_t) = r\}$
  predict $\hat{y}_t = \arg\max_r \mathrm{Ldim}(V_t^{(r)})$
  receive true answer y_t
  update $V_{t+1} = V_t^{(y_t)}$

Theorem

SOA makes at most $\mathrm{Ldim}(\mathcal{H})$ *mistakes.*

Proof.

Whenever SOA errs we have $\mathrm{Ldim}(V_{t+1}) \leq \mathrm{Ldim}(V_t) - 1$. □

Randomized Prediction and Expected Regret

- Environment should decide on y_t before seeing $\hat{y}_t$
- Learner can randomize his predictions
- We analyze expected regret

$$\sum_{t=1}^{T} \mathbb{E}[|\hat{y}_t - y_t|] \; - \; \min_{h \in \mathcal{H}} \sum_{t=1}^{T} |h(\mathbf{x}_t) - y_t|$$

Weighed Majority (Warmuth&Littlestone 89)

WM for learning with d experts

initialize: assign weight $w_i = 1$ for each expert
for $t = 1, 2, \ldots, T$
each expert predicts $f_i \in \{0, 1\}$
environment determines y_t without revealing it to the learner
predict $\hat{y}_t = 1$ w.p. $\propto \sum_{i:f_i=1} w_i$
receive label y_t
foreach wrong expert: $w_i \leftarrow \eta\, w_i$

Weighed Majority (Warmuth&Littlestone 89)

WM for learning with d experts

initialize: assign weight $w_i = 1$ for each expert

for $t = 1, 2, \ldots, T$

each expert predicts $f_i \in \{0, 1\}$

environment determines y_t without revealing it to the learner

predict $\hat{y}_t = 1$ w.p. $\propto \sum_{i: f_i = 1} w_i$

receive label y_t

foreach wrong expert: $w_i \leftarrow \eta\, w_i$

Theorem

WM achieves expected regret of at most: $\sqrt{\ln(d)\, T}$

WM and Online Learnability

- WM regret bound $\Rightarrow$ a finite $\mathcal{H}$ is learnable with regret $\sqrt{\ln(|\mathcal{H}|)\, T}$
- Is this the best we can do ? And, what if $\mathcal{H}$ is infinite ?
- Solution: Combing WM with SOA

WM and Online Learnability

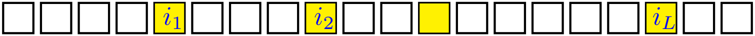
- WM regret bound $\Rightarrow$ a finite $\mathcal{H}$ is learnable with regret $\sqrt{\ln(|\mathcal{H}|)\, T}$
- Is this the best we can do ? And, what if $\mathcal{H}$ is infinite ?
- Solution: Combing WM with SOA

Theorem

- *Exists learner with expected regret* $\sqrt{\mathrm{Ldim}(\mathcal{H})\, T\, \log(T)}$
- *No learner can have expected regret smaller than* $\sqrt{\mathrm{Ldim}(\mathcal{H})\, T}$

Therefore: $\mathcal{H}$ *is agnostic online learnable* $\iff \mathrm{Ldim}(\mathcal{H}) < \infty$

Proof idea



Expert($i_1, \ldots, i_L$)

initialize: $V_1 = \mathcal{H}$

for $t = 1, 2, \ldots$

  receive $\mathbf{x}_t$

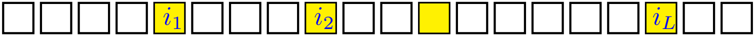
  for $r \in \{0, 1\}$ let $V_t^{(r)} = \{h \in V_t : h(\mathbf{x}_t) = r\}$

  define $\hat{y}_t = \arg\max_r \mathrm{Ldim}(V_t^{(r)})$

  if $t \in \{i_1, \ldots, i_L\}$ flip prediction: $\hat{y}_t \leftarrow \neg \hat{y}_t$

  update $V_{t+1} = V_t^{(\hat{y}_t)}$

Proof idea



Expert($i_1, \ldots, i_L$)

initialize: $V_1 = \mathcal{H}$
for $t = 1, 2, \ldots$
  receive $\mathbf{x}_t$
  for $r \in \{0,1\}$ let $V_t^{(r)} = \{h \in V_t : h(\mathbf{x}_t) = r\}$
  define $\hat{y}_t = \arg\max_r \mathrm{Ldim}(V_t^{(r)})$
  if $t \in \{i_1, \ldots, i_L\}$ flip prediction: $\hat{y}_t \leftarrow \neg \hat{y}_t$
  update $V_{t+1} = V_t^{(\hat{y}_t)}$

Lemma: $\forall h \in \mathcal{H},\ \exists i_1, \ldots, i_L,\ L < \mathrm{Ldim}(\mathcal{H})$, s.t. Expert($i_1, \ldots, i_L$) agrees with h on the entire sequence.