CS747: Foundations of Intelligent and Learning Agents **July - Nov 2017**

Lecture 11: August 23

Lecturer: Manjesh K. Hanawal *Scribe: Ansuma Basumatary*

11.1 Online Learning (Adversarial Setting)

- What is online learning
  - "Online Machine Learning is a method of machine learning in which data becomes available in a sequential order and is used to update our best predictor for future data at each step, as opposed to batch learning techniques which generate the best predictor by learning on the entire training data set at once." - Wikipedia
- Algorithm A

  Input: Hypothesis Class $\mathcal{H}$
  for t:=1,2,3,...
  Receive sample, x_t.
  Select hypothesis $h \in \mathcal{H}$.
  Predict label, $\hat{y}_t = h(x_t)$
  Suffer loss, $|\hat{y}_t - y_t|$
  Return a hypothesis, $h \in \mathcal{H}$

- For any given sequence, $\mathcal{S} = \{(x_i, y_i) : i = 1, 2, ..., T\}$, where T is an integer.
  Let, $M_{\mathcal{S}}(A)$ be the number of mistakes algorithm A makes on $\mathcal{S}$

  Definition $\rightarrow$ Let $M_{\mathcal{H}}(A) = \sup_{\mathcal{S}} M_{\mathcal{S}}(A)$ denote the maximum number of mistakes.
  A bound of the form, $M_{\mathcal{H}}(A) \leq B < \infty$

  Definition $\rightarrow$ (Online Learnability).
  We say that a hypothesis class in "learnable" if $\exists$ an algorithm A & $B < \infty$ such that $M_{\mathcal{H}}(A) < B$

  Assumption $\rightarrow$ All labels are generated by some hypothesis $h^* \in \mathcal{H},\ y_t = h^*(x_t)$

- Consistent Algorithm

  Input: Hypothesis Class $\mathcal{H}$
  Initialize: $V_1 = \mathcal{H}$
  for t:=1,2,3,...
  Receive sample, x_t.
  Select hypothesis $h \in V_t$.
  Predict label, $\hat{y}_t = h(x_t)$
  Receive true label $y_t = h^*(x_t)$
  Update, $V_{t+1} = \{h \in V_t : h(x_t) = y_t\}$

  $\rightarrow$ $|\mathcal{H}| < T$
  $1 \leq |V_t| \leq |\mathcal{H}| - 1$
  $M_{\mathcal{H}}(consis) \leq |\mathcal{H}| - 1$

- <u>Halving Algorithm</u>
  Input: Hypothesis Class $\mathcal{H}$
  Initialize: $V_1 = \mathcal{H}$
  for t:=1,2,3,...
    Receive sample, x_t.
    Predict label, $\hat{y}_t = argmax_{\gamma \in \{0,1\}} |\{h \in V_t : h(x_t) = \gamma\}|$
    Receive true label $y_t = h^*(x_t)$
    Update, $V_{t+1} = \{h \in V_t : h(x_t) = y_t\}$

  $\rightarrow$ $1 \leq |V_{t+1}| \leq |V_t|/2$
  $1 < |V_{t+1}| \leq |\mathcal{H}| \cdot 2^{-M}$
  $M \leq \log_2 |\mathcal{H}|$

11.2 Online Learning under Unrealizable Case

- We dont have $h^* \in \mathcal{H}$

$$R_A(h,T) = \sup_{(x_1,y_1),\ldots,(x_T,y_T)} \left[\sum_{t=1}^{T} |\hat{y}_t - y_t| - \sum_{t=1}^{T} |h(x_t) - y_t| \right]$$

$$\Rightarrow R_A(T) = \sup_{h \in \mathcal{H}} R_A(h,T)$$

$\rightarrow$ Learner's goal is to achieve the lowest regret

$$\lim_{T \to \infty} \frac{R_A(T)}{T} \to 0$$

If the above condition is true, the regret will be sublinear.
And hence, the algorithm will be learning.

- <u>Conditions</u>

  1. We allow the learner to randomize his predictions generated by
  2. The adversary has to decide of h_t without knowing the actual outcome of learner's random predictions.

- To find, $\min_A \mathbb{E}[R_A(T)]$

$$P_t = Pr\{\hat{y}_t = 1\}, \ P_t \in [0,1]$$

$$\mathbb{E}|\hat{y}_t - y_t| = |P_t - y_t|$$

$$\mathbb{E}[R_A(h,T)] = \sup \left[\sum_{t=1}^{T} |P_t - y_t| - \sum (h(x_t) - y_t) \right]$$

- Is there an Algorithm that gives a sub-linear regret?

  Theorem $\rightarrow$ For every hypothesis $\mathcal{H}$, $\exists$ an algorithm for Online classification whose predictions come from [0,1] and has regret bound such that,

$$\forall h \in \mathcal{H}, \ R_A(h,T) \leq \sqrt{2\log(|\mathcal{H}|) \cdot T}$$

- Weighted-Majority Algorithm

  Input: $\mathcal{H},\ T$
  Parameter: $\eta = \sqrt{2\log(|\mathcal{H}|)/T}$
  Initialize: $\tilde{w}^{(1)} = (1, 1, ..., 1)$
  for t:=1,2,3,...
  Set $w_i^{(t)} = \frac{\tilde{w}_i^{(t)}}{\sum \tilde{w}_i^{(t)}}\ \forall i = 1, 2, ..., d$ Where $|\mathcal{H}| = d$
  Choose hypothesis h_t according to distribution $w_i^{(t)}$
  Receive cost vector $l_t, i \in [0, 1]^d$
  Compute expected cost $< w^{(t)}, l_t >$
  Update rule: $\forall i,\ \tilde{w}_i^{(t+1)} = \tilde{w}_i^{(t)} \cdot e^{-\eta l_{t,i}}$

- Theorem $\rightarrow\ R_{WM}(T) \leq O(\sqrt{T})$
  $\rightarrow$ WM is order optimal.

- The above the algorithms are based on the setting called **Full Information** - loss of all information is known.