# Model Selection in Exploration

John Langford (Yahoo! Research)

New Frontiers in Model Order Selection Workshop, 2011

# Supervised ERM Learning

$H$ = set of hypotheses $\{h\}$ where $h(x) \in \{0,1\}$.
$D(x,y)$ = distribution over events
where $x$ = features and $y \in \{0,1\}$ = label.

# Supervised ERM Learning

$H$ = set of hypotheses $\{h\}$ where $h(x) \in \{0, 1\}$.
$D(x, y)$ = distribution over events
where $x$ = features and $y \in \{0, 1\}$ = label.

For each $h$, there is some true error rate:
$e(D, h) = \Pr_{(x,y)\sim D}(h(x) \neq y)$
Given samples $S$, we have an empirical error rate:
$\hat{e}(S, h) = E_{(x,y)\in S} I(h(x) \neq y)$

# Supervised ERM Learning

$H$ = set of hypotheses $\{h\}$ where $h(x) \in \{0, 1\}$.
$D(x, y)$ = distribution over events
where $x$ = features and $y \in \{0, 1\}$ = label.

For each $h$, there is some true error rate:
$e(D, h) = \Pr_{(x,y)\sim D}(h(x) \neq y)$
Given samples $S$, we have an empirical error rate:
$\hat{e}(S, h) = E_{(x,y)\in S} I(h(x) \neq y)$

Let $\hat{h}_H(S) = \arg\min_{h\in H} \hat{e}(s, h)$

**Theorem**

For all $H, D$, let $\hat{h} = \hat{h}_H(S)$. If samples in $S$ are IID from $D$, with probability at least $1 - \delta$:
$|\hat{e}(S, \hat{h}) - e(D, \hat{h})| \leq \sqrt{\frac{\ln |H| + \ln \frac{2}{\delta}}{2|S|}}$

# Supervised Model Selection

Suppose we have a set of hypothesis sets $M = \{H\}$.
Let $\hat{H}_M(S) =$
$\arg\min_{H \in M} \hat{e}(S, \hat{h}_H(S)) + \sqrt{\frac{\ln|H| + \ln|M| + \ln 2/\delta}{2|S|}}$

# Supervised Model Selection

Suppose we have a set of hypothesis sets $M = \{H\}$.
Let $\hat{H}_M(S) =$
$\arg\min_{H \in M} \hat{e}(S, \hat{h}_H(S)) + \sqrt{\frac{\ln|H| + \ln|M| + \ln 2/\delta}{2|S|}}$

**Theorem**

For all $M, D$, let $\hat{h} = \hat{h}_{\hat{H}_M(S)}(S)$, with probability at least $1-\delta$: $|\hat{e}(S, \hat{h}) - e(D, \hat{h})| \leq \sqrt{\frac{\ln|H| + \ln|M| + \ln 2/\delta}{2|S|}}$

# Supervised Model Selection

Suppose we have a set of hypothesis sets $M = \{H\}$.
Let $\hat{H}_M(S) =$
$\arg\min_{H \in M} \hat{e}(S, \hat{h}_H(S)) + \sqrt{\frac{\ln|H| + \ln|M| + \ln 2/\delta}{2|S|}}$

**Theorem**

For all $M, D$, let $\hat{h} = \hat{h}_{\hat{H}_M(S)}(S)$, with probability at least $1 - \delta$: $|\hat{e}(S, \hat{h}) - e(D, \hat{h})| \leq \sqrt{\frac{\ln|H| + \ln|M| + \ln 2/\delta}{2|S|}}$

cost of model selection = cost of not knowing right $H$ in advance.

$$O\left(\sqrt{\frac{\ln|M|}{|S|}}\right)$$

.

# What happens when there is exploration?

1. Realizable Active Learning
2. Agnostic Active Learning
3. Agnostic Contextual Bandits
4. Realizable Contextual Bandits

# Realizable Active Learning (RAL)

We have a pool of examples $x$, and we must request labels $y$ until we find an $\epsilon$ optimal $h \in H$. One $h \in H$ is guaranteed to have zero error.

# Realizable Active Learning (RAL)

We have a pool of examples $x$, and we must request labels $y$ until we find an $\epsilon$ optimal $h \in H$. One $h \in H$ is guaranteed to have zero error.

Label complexity is dependent on $H$ and $D$ [D05]. $S(H, D) = \frac{\ln |H|}{\rho} \ln 1/\epsilon$
Here $\rho =$ splitting rate. Let
$Q_\epsilon = \{(h, h') : h, h' \in H \text{ and } d_D(h, h') > \epsilon\}$ where $d_D(h, h') = \Pr_{x \sim D}(h(x) \neq h'(x))$. Then $x$ $\rho$-splits $Q$ if the label $y$ always eliminates a $\rho$ fraction of $Q$. In other words:

$$|Q'| \leq (1 - \rho)|Q|$$

Splitting rate = largest guaranteed split amongst $x$s.

# Model Selection for RAL

Only unlabeled data is required to estimate $d_D(h, h')$.

# Model Selection for RAL

Only unlabeled data is required to estimate $d_D(h, h')$.

MS_Active($M$)

1. Sample many unlabeled points.
2. For each $H \in M$, compute $S(H, D) = \frac{\ln |H|}{\rho} \ln 1/\epsilon$.
3. Choose $\arg\min_{H \in M} S(H, D)$ and active learn on it.

# Model Selection for RAL

Only unlabeled data is required to estimate $d_D(h, h')$.

MS_Active($M$)

1. Sample many unlabeled points.
2. For each $H \in M$, compute $S(H, D) = \frac{\ln |H|}{\rho} \ln 1/\epsilon$.
3. Choose $\arg\min_{H \in M} S(H, D)$ and active learn on it.

The cost of model selection for realizable active learning is $0$. (Unlabeled complexity generally grows though.)

# Model Selection for RAL

Only unlabeled data is required to estimate $d_D(h, h')$.

MS_Active($M$)

1. Sample many unlabeled points.
2. For each $H \in M$, compute $S(H, D) = \frac{\ln |H|}{\rho} \ln 1/\epsilon$.
3. Choose $\arg\min_{H \in M} S(H, D)$ and active learn on it.

The cost of model selection for realizable active learning is $0$. (Unlabeled complexity generally grows though.)

It's also $0$ for realizable model selection in supervised learning.

# Agnostic Active Learning

[BBL06, H07, DHM08, BDL09, BHLZ10]
We have a stream of examples $x$, and we must request labels $y$ online until we find an $\epsilon$ optimal $h \in H$.

# Agnostic Active Learning

[BBL06, H07, DHM08, BDL09, BHLZ10]
We have a stream of examples $x$, and we must request labels $y$ online until we find an $\epsilon$ optimal $h \in H$.

Online Active Learning($H$)

1. $H_0 = H$
2. for t = 1...
   1. Observe $x$
   2. If there exists $h, h' \in H_{t-1}$ s.t. $h(x) \neq h'(x)$ query.
   3. $H_t = \{h \in H_{t-1} :$ $h$ not eliminated by sample complexity bound$\}$

# Label Complexity

Label complexity is dependent on $H$ and $D$.
$S(H, D) = \tilde{O}(\theta(\nu m + \ln m(\ln |H| \ln m + \ln 1/\delta)))$
where $m =$ supervised label complexity
$\nu \equiv \min_{h \in H} e(D, h) =$ minimum error rate $\theta =$ disagreement coefficient.

# Label Complexity

Label complexity is dependent on $H$ and $D$.
$S(H, D) = \tilde{O}(\theta(\nu m + \ln m(\ln |H| \ln m + \ln 1/\delta)))$
where $m =$ supervised label complexity
$\nu \equiv \min_{h \in H} e(D, h) =$ minimum error rate $\theta =$ disagreement coefficient.

The disagreement coefficient is
$\max_r \frac{\Pr_{x \sim D}(\exists h \in B(h^*, r) : h(x) \neq h^*(x))}{r}$.
If you only care about hypotheses within $r$ of the optimal $h^*$, what fraction of the $x$ are disagreed upon?

# Model Selection for AAL

Labeled data is required to evaluate $\nu$!

# Model Selection for AAL

Labeled data is required to evaluate $\nu$!

The best solution I know:
RoundRobin($M$)

1. For each $H \in M$ start $\text{AAL}(H)$
2. For $x_t$ query label if $\text{AAL}(H_{t \bmod |M|}, x)$ queries.
   1. If query $y$ received, update $\text{AAL}(H_{t \bmod |M|}, x, y)$.

You also need to choose a single hypothesis amongst the survivors. The best approach I know is to form $H' = \{h \in \text{AAL}(H) : H \in M\}$ and run $\text{AAL}(H')$.

# Model Selection for AAL

Labeled data is required to evaluate $\nu$!

The best solution I know:
RoundRobin($M$)

1. For each $H \in M$ start $\text{AAL}(H)$
2. For $x_t$ query label if $\text{AAL}(H_{t \bmod |M|}, x)$ queries.
   1. If query $y$ received, update $\text{AAL}(H_{t \bmod |M|}, x, y)$.

You also need to choose a single hypothesis amongst the survivors. The best approach I know is to form $H' = \{h \in \text{AAL}(H) : H \in M\}$ and run $\text{AAL}(H')$.

Label complexity of model selection is at least a factor of $|M|$ worse.

# Agnostic Contextual Bandits

[ACFS95, BLLRS11, DHKKLRZ11]

We have a stream of examples $x$. At each step we choose an action $a$, and get a loss $l_a$ for that action. We want to compete with the hypothesis having the smallest average loss.

EXP4P($H$)

1. Let $w_h = 1$
2. for t = 1...
   1. Choose $P(h)$ according to normalized weights.
   2. Observe $x$
   3. Draw $h \sim P$ and act as $h(x)$.
   4. Observe $r_a$
   5. Update $w_{h'} \leftarrow w_{h'} e^{p \min \frac{r_a}{p_a} I(h(x)=h'(x)) + \text{a bit}}$.

(approximately)

# EXP4P analysis

**Theorem**

For all $H$, for all sequences, with probability $1 - \delta$ the average per round regret of EXP4P is:

$$O\left(\sqrt{\frac{\ln |H| + \ln(1/\delta)}{T}}\right)$$

(In general, number of actions divides $T$, but we only have two actions.)

# Model Selection for CBs

Suppose we have $M = \{H\}$. Again, there is no way to tell in advance which is best. Best combination method = ?

# Model Selection for CBs

Suppose we have $M = \{H\}$. Again, there is no way to tell in advance which is best. Best combination method = ?

RoundRobin($M$)

1. Initialize EXP4P($H$)
2. for t = 1...
   1. Choose $P(h)$ according to EXP4P($H_{t \bmod |M|}$)
   2. Draw $h \sim P$ and act as $h(x)$.
   3. Update all instances.

# Model Selection for CBs

Suppose we have $M = \{H\}$. Again, there is no way to tell in advance which is best. Best combination method = ?

RoundRobin($M$)

1. Initialize EXP4P($H$)
2. for t = 1...
   1. Choose $P(h)$ according to EXP4P($H_{t \bmod |M|}$)
   2. Draw $h \sim P$ and act as $h(x)$.
   3. Update all instances.

Conjecture: For all $M$, for all sequences, with probability $1 - \delta$ the average regret is:

$$O\left(|M|\sqrt{\frac{\ln|H| + \ln\delta}{T}}\right).$$

# Realizable Contextual Bandits [DHKLZ12]

We have a stream of examples $x$. At each step we choose an action $a$, and get a loss $l_a$ for that action. We compete with a set of regressors $f(x, a)$ where one satisfies $E_{D|x} l_a = f(x, a)$.

# Realizable Contextual Bandits [DHKLZ12]

We have a stream of examples $x$. At each step we choose an action $a$, and get a loss $l_a$ for that action. We compete with a set of regressors $f(x, a)$ where one satisfies $E_{D|x} l_a = f(x, a)$.

Regressor Elimination($F$)

1. $F_0 = F$
2. for t = 1...
   1. Choose $P(f)$ so as to achieve near-uniformity over $F_{t-1}$'s actions.
   2. Observe $x$
   3. Draw $f \sim P$ and act as $\arg\max_a f(x, a)$.
   4. Observe $r_a$
   5. $F_t = \{f \in F_{t-1} :$ $f \text{ not eliminated by sample complexity bound}\}$

# Realizable CB analysis

**Theorem**

For all $F$, for all $D$, with probability $1-\delta$ the average regret of Regressor Elimination is:

$$O\left(\sqrt{\frac{\ln|F| + \ln\delta}{T}}\right)$$

# Realizable CB analysis

**Theorem**

For all $F$, for all $D$, with probability $1-\delta$ the average regret of Regressor Elimination is:

$$O\left(\sqrt{\frac{\ln|F|+\ln\delta}{T}}\right)$$

and sometimes *much* better.

# Model Selection for RCB

For model selection, we can pick the smallest set of $F$, incurring average regret:

$$O\left(\min_F \sqrt{\frac{\ln |F| + \ln \delta}{T}}\right)$$

# Model Selection for RCB

For model selection, we can pick the smallest set of $F$, incurring average regret:

$$O\left(\min_F \sqrt{\frac{\ln|F| + \ln\delta}{T}}\right)$$

There is no cost of model selection here.

# The Cost of Model Selection

| | Supervised | Active | Cont. Band. |
|---|---|---|---|
| Realizable | 0 | 0 | 0 |
| Agnostic | $+\log\|M\|$ | $*\|M\|$? | $*\|M\|$? |

# The Cost of Model Selection

| | Supervised | Active | Cont. Band. |
|---|---|---|---|
| Realizable | 0 | 0 | 0 |
| Agnostic | $+\log\|M\|$ | $*\|M\|$? | $*\|M\|$? |

What I learned:

1. Realizability makes model selection trivial.
2. Agnostic model selection with exploration is hard!