Prof. Friedrich Eisenbrand
Nicolai Hähnle

Discussions from: December 3, 2009

Mathematics of Machine Learning

Fall 2009

Assignment Sheet 6

Exercises marked with a $\star$ can be handed in for bonus points. Due date is December 17.

Exercise 1

Let us describe a general computational problem in the following way. Let $\mathscr{I}$ be a set of *problem instances*. For every $I \in \mathscr{I}$ there is a non-empty set $\mathscr{S}(I)$ of possible *solutions*. We consider two types of randomized algorithms A for solving such a problem.

Type 1) $A(I)$ always returns a solution from $\mathscr{S}(I)$. $A(I)$ may run arbitrarily long, but its expected running time is bounded by a polynomial $p_1(|I|)$.

Type 2) For every $0 < \delta < 1$, $A(I,\delta)$ always runs in time bounded by a polynomial $p_2(|I|, \log\frac{1}{\delta})$. With probability at least $1-\delta$, it returns a solution from $\mathscr{S}(I)$; otherwise, it indicates failure.

Show that these types of algorithm are equivalent, that is, a Type 1 algorithm exists for the problem if and only if a Type 2 algorithm exists.

Exercise 2

Improve the analysis of the modest accuracy boosting procedure to show: if the weak learning algorithm returns hypotheses h' with $\text{error}(h') \leq p$, then the returned majority hypothesis h has $\text{error}(h) \leq 3p^2 - 2p^3$.

Exercise 3

Show that every concept class can be "learned" with accuracy arbitrarily close to $\frac{1}{2}$. More precisely, show the following: there is an algorithm A that on input $A(\varepsilon, \delta, EX(c,\mathscr{D}))$ returns, with probability at least $1-\delta$, a hypothesis h with $\text{error}(h) \leq \frac{1}{2} + \varepsilon$. Furthermore, the algorithm runs in time polynomial in $\frac{1}{\varepsilon}$ and $\log(\frac{1}{\delta})$.

Exercise 4 ($\star$)

In the lecture on the full recursive accuracy boosting procedure, we assumed that the weak learning algorithm A always returns a hypothesis with error at most p. Suppose now that A takes a parameter δ so that $\Pr[A \text{ returns a hypothesis with error } > p] < \delta$. Determine $\delta > 0$ as large as possible so that $\Pr[\text{Recursive boosting produces a hypothesis with error} \leq \varepsilon] \geq \frac{1}{2}$?

Note: You may ignore the possibility of failures of the sampling oracles (compare exercise 5).

Exercise 5 ($\star$)
In the modest accuracy boosting procedure, we defined a modified distribution $\mathscr{D}_2$ based on a previously computed hypotheses h_1 with error at most p.

We can create the $EX(c,\mathscr{D}_2)$ sampling oracle as follows:

EX$(c,\mathscr{D}_2)$
1 $b \leftarrow_R \{0,1\}$
2 **for** $i = 1 \dots m$
3 **do** $(x, c(x)) \leftarrow$ EX$(c,\mathscr{D})$
4 **if** $(b = 0 \wedge c(x) = h_1(x)) \vee (b = 1 \wedge c(x) \neq h_1(x))$
5 **then return** $(x, c(x))$
6 Abort the accuracy boosting and directly return h_1

Show that this generates the correct distribution $\mathscr{D}_2$ when successful. Furthermore, show that we can choose $m = q(\log(\frac{1}{\delta}), \frac{1}{p})$ (where q is a polynomial) such that if $\text{error}(h_1) > 3p^2 - 2p^3$, the oracle is successful with probability at least $1 - \delta$.

Note: You may assume that $\text{error}_{\mathscr{D}}(h_1) \leq p$ always holds.

Note: If the oracle implementation aborts in line 6, then most likely h_1 is already as good a hypothesis as we can hope to achieve. You can design and analyze an analogous implementation of $EX(c,\mathscr{D}_3)$. Taken together, the respective parameters for δ can be chosen (using union bound) to get a full proof of the correctness of the modest accuracy boosting procedure that takes the confidence parameters into account.

Together with exercise 4 and some more technical details, you could use this to prove the correctness of the recursive accuracy boosting procedure. However, this is not required.