Results in other Online Learning Models

Nader H. Bshouty

Department of Computer Science

Technion

Haifa, Israel

bshouty@cs.technion.ac.il

Abstract

Here we give results in other Online Learning Models.

1 The Models

2 Online Models

In the online learning model [L87] the teacher at each trial sends a point $x \in X$ to the learner and the learner has to predict $f(x)$. The learner returns to the teacher the prediction y. If $f(x) \neq y$ then the teacher returns "mistake" to the learner. The goal of the learner is to minimize the number of prediction mistakes. he teacher can be regarded as an adversary with unlimited computational power that must answer honestly but also wants to fail the learner from learning quickly. The teacher does not know the coin tosses of the learner.

Online [L87] In the online model we say that algorithm $\mathcal{A}$ of the learner *Online-learns* the class C if for any $f \in C$ and for any δ, algorithm $\mathcal{A}(\delta)$ with probability at least $1-\delta$ makes bounded number of mistakes. We say that C is *Online-learnable* if the number of mistakes and the running time of the learner for each prediction is $poly(\log(1/\delta), I_f)$.

Online$_f$ This is the online model where the the teacher fixes the target function before running the learning algorithm.

Online$_{f,S}$ [M91] This is the online model where the the teacher fixes the target function and the sequence of points that will be sent to the learner before running the learning algorithm.

Probabilistic Prediction (PP) [?] In the Probabilistic Prediction model the points sent to the learner are chosen from X according to some distribution D. We say an algorithm $\mathcal{A}$ of the learner η-*PP-learns* the class C if for any $f \in C$ and for any δ

the algorithm $\mathcal{A}(\delta)$ with probability at least $1-\delta$ after bounded number of mistakes can predict the answer with probability greater than $1-\eta$. We say that C is *η-PP-learnable* if the number of mistakes and the running time of the learner at each trial is *poly*$(\log(1/\delta), I_f)$.

3 Results

Maass in [M91] investigated the model $\text{Online}_{f,S}$. He shows that the Halving algorithm can be replaced by an algorithm that simply chooses after each mistake a random consistent hypothesis for all the examples seen so far. The expected number of mistakes is $\ln |C| + O(1)$. This result is also true for the weakly oblivious model where the teacher can let its choice of examples depend on earlier reactions of the learner, but is not able to predict future moves of the learner. You can find a simple proof of this fact in Ron Rivest LN. If finding a randomized consistent hypothesis is in **RP** then this algorithm runs in time *poly*(t) after each mistake where t is the trial number.

Open Problem. *Find an algorithm that run in poly*$(\log t)$ *assuming finding a randomized consistent hypothesis in C is in* **RP**. *There is a non-proper learning algorithm. Just use randomized Halving. What about proper?*

Then ignoring the computational complexity he investigates the connection between several online models (including the common one) and finds bounds on the number of mistakes. He show that the number of mistakes of the randomized online is bounded by the deterministic online $/\log|C|$.

References

[M91] W. Maass. On-line learning with an oblivious environment and the power of randomization. In Proceedings of the 4th Annual ACM Workshop on Computational Learning Theory, pages 167-175. Morgan Kaufmann (San Mateo), 1991.

[L87] Nick Littlestone. Learning Quickly When Irrelevant Attributes Abound: A New Linear-threshold Algorithm. Machine Learning 2(4): 285-318 (1987).