A Desicion-Theoretic Generalization of On-Line Learning and an Application to Boosting

Yoav Freund Robert E. Schapire

AT&T Bell Laboratories
600 Mountain Avenue
Murray Hill, NJ 07974-0636
{yoav, schapire}@research.att.com

Abstract. We consider the problem of dynamically apportioning resources among a set of options in a worst-case on-line framework. The model we study can be interpreted as a broad, abstract extension of the well-studied on-line prediction model to a general decision-theoretic setting. We show that the multiplicative weight-update rule of Littlestone and Warmuth [10] can be adapted to this model yielding bounds that are slightly weaker in some cases, but applicable to a considerably more general class of learning problems. We show how the resulting learning algorithm can be applied to a variety of problems, including gambling, multiple-outcome prediction, repeated games and prediction of points in $\mathbb{R}^n$. We also show how the weight-update rule can be used to derive a new boosting algorithm which does not require prior knowledge about the performance of the weak learning algorithm.

1 Introduction

A gambler, frustrated by persistent horse-racing losses and envious of his friends' winnings, decides to allow a group of his fellow gamblers to make bets on his behalf. He decides he will wager a fixed sum of money in every race, but that he will apportion his money among his friends based on how well they are doing. Certainly, if he knew psychically ahead of time which of his friends would win the most, he would naturally have that friend handle all his wagers. Lacking such clairvoyance, however, he attempts to allocate each race's wager in such a way that his total winnings for the season will be reasonably close to what he would have won had he bet everything with the luckiest of his friends.

In this paper, we describe a simple algorithm for solving such dynamic allocation problems, and we show that our solution can be applied to a great assortment of learning problems. Perhaps the most surprising of these applications is the derivation of a new algorithm for "boosting," i.e., for converting a "weak" PAC learning algorithm that performs just slightly better than random guessing into one with arbitrarily high accuracy.

We formalize our *on-line allocation model* as follows. The allocation agent A has N options or *strategies* to choose from; we number these using the integers $1, \ldots, N$. At each time step $t = 1, 2, \ldots, T$, the allocator A decides on a distribution $\mathbf{p}^t$ over the strategies; that is $p_i^t \geq 0$ is the amount allocated to strategy i, and $\sum_{i=1}^N p_i^t = 1$. Each strategy i then suffers some *loss* ℓ_i^t which is determined by the (possibly adversarial) "environment." The loss suffered by A is then $\sum_{i=1}^N p_i^t \ell_i^t = \mathbf{p}^t \cdot \boldsymbol{\ell}^t$, i.e., the average loss of the strategies with respect to A's chosen allocation rule. We call this loss function the *mixture loss*.

In this paper, we always assume that the loss suffered by any strategy is bounded so that, without loss of generality, $\ell_i^t \in [0,1]$. Besides this condition, we make no assumptions about the form of the loss vectors $\boldsymbol{\ell}^t$, or about the manner in which they are generated; indeed, the adversary's choice for $\boldsymbol{\ell}^t$ may even depend on the allocator's chosen mixture $\mathbf{p}^t$.

The goal of the algorithm A is to minimize its cumulative loss relative to the loss suffered by the best strategy. That is, A attempts to minimize its *net loss*

$$L_A - \min_i L_i$$

where

$$L_A = \sum_{t=1}^{T} \mathbf{p}^t \cdot \boldsymbol{\ell}^t$$

is the total cumulative loss suffered by algorithm A on the first T trials, and

$$L_i = \sum_{t=1}^{T} \ell_i^t$$

is strategy i's cumulative loss.

In Section 2, we show that Littlestone and Warmuth's [10] "weighted majority" algorithm can be generalized to handle this problem, and we prove a number of bounds on the net loss. For instance, one of our results shows that the net loss of our algorithm can be bounded by $O\left(\sqrt{T \ln N}\right)$ or, put another way, that the average per trial net loss is decreasing at the rate $O\left(\sqrt{(\ln N)/T}\right)$. Thus, as T increases, this difference decreases to zero.

Our results for the on-line allocation model can be applied to a wide variety of learning problems, as we describe in Section 3. In particular, we generalize the results of Littlestone and Warmuth [10] and Cesa-Bianchi et al. [1] for the problem of predicting a binary sequence using the advice of a team of "experts." Whereas these authors proved worst-case bounds for making on-line randomized decisions over a binary decision and outcome space with a $\{0,1\}$-valued discrete loss, we prove (slightly weaker) bounds that are applicable to any bounded loss function over any decision and outcome spaces. Our bounds express explicitly the rate at which the loss of the learning algorithm approaches that of the best expert.

Related generalizations of the expert prediction model were studied by Vovk [12], Kivinen and Warmuth [9], and Haussler, Kivinen and Warmuth [8]. Like us, these authors focused primarily on multiplicative weight-update algorithms. Chung [2] also presented a generalization, giving the problem a game-theoretic treatment.

Finally, in Section 4, we show how a similar algorithm can be used for boosting, i.e., for converting any weak PAC learning algorithm into a strong PAC learning algorithm. Unlike the previous boosting algorithms of Freund [5, 6] and Schapire [11], the new algorithm needs no prior knowledge of the accuracy of the hypotheses of the weak learning algorithm. Rather, it adapts to the accuracies of the generated hypotheses and