Using Value of Information to Learn and Classify under Hard Budgets

Russell Greiner*
Dep't of Computing Science; University of Alberta Edmonton, AB T6J 2E8
greiner@cs.ualberta.ca

While a doctor may have the option of using a wide variety of medical tests (including MRI scans, blood work, etc.) to diagnose a patient, each of these tests has some (typically known) costs.[1] There are also penalties for misdiagnoses; typically one value for false positives and another (here, larger) value for false negatives. Of course, the physician should prefer a diagnostic process that minimizes the total (expected) cost+penalty of diagnosis. If we know the underlying probabilistic distribution, relating test outcomes to one another and to the diagnosis, we can view this as a Markov decision process (MDP), and the diagnostic process as a "policy": at each stage ("state") we know the outcomes of a subset of the tests and must decide whether to return a diagnosis (and accept the penality if wrong) or obtain the value of some other test (at cost). In the simplest model, when the costs of the tests are independent, this corresponds to a decision tree, called an "active classifier" in [GGR02].

There are many computational challenges in finding the best such tree; e.g., [GGR02] shows that it is NP-hard in general. That paper also shows, however, that this task can be solved efficiency (via Dynamic Programming, DP) if we consider only "bounded-horizon" policies — i.e., policies that can perform only a fixed number of tests. These are "bounded active classifiers". This special case is quite relevant, as it corresponds directly to many standard examples, including capitation costs of health care — e.g., where the physician can spend only $100 per patient visit.

Of course, this assumes the research team (who is designing the optimal policy) knows the complete probabilistic distribution. [GGR02] shows that one can efficiently get a sufficient approximation to (PAC-)learn these bounded active classifiers, by observing a polynomial number of complete instances. However, this learning framework is also problematic: if the *physican* applying the treatment has to pay to observe each feature, one would expect the researcher to also have to purchase each feature used to compute the classifier. This motivates the "Budgeted Learning" model: The learner BL begins with a fixed budget (e.g., $1000) and access to a set of instances, but initially knows only the class label ("diagnosis") of each, but none of their test values. At each step, BL will also know the values of only the already-purchased specified features of specified instances, and can then purchase the value of some specified feature of some instance (e.g., can run blood-test#7 on instance#31, at some *cost*(blood-test#7, instance#31) $\in \Re^+$. This process continues until BL has spent its entire budget, at which point BL returns a classifier, that can then return labels for novel instances.

[LMG04] explores a simple version of this task, "active model selection" (AMS), in the context of the "coins problem": Given a set of n coins (each with a known prior disbriution) and a total budget of k flips, sequentially decide when to flip which coin to determine the coin that has the highest head probability. (While this task is clearly related to standard bandit problems [BF85], it has significant differences — enough that the standard methods do not apply. Moreover, that paper also shows how our task is different from standard experimental design, on-line learning, standard active learning, and so forth.) It shows that this task (when worded in a Bayesian framework) corresponds to an MDP, and so can be solved using DP, albeit using expential time and space. It then provides this task is NP-hard in general, and considers a variety of obvious heuristics, including "round robin" (flip each coin $\approx n/k$ times), "random", and greedy look ahead, as well as novel algorithms, based on Gittins sets, interval estimation, "play-the-winner" (PTW) and single-feature-look-ahead (SFLA). After showing that none of these algorithms is an approximation algorithm (i.e., the "expected expected regret" of each can be arbitrarily worse than optimal), it provides empirical evidence that PTW and SFLA perform well in general, and that PTW is significanly more efficient.

[LMG03] then returns to the task of actually learning a classifier — here a Naïve Bayes classifier. We ran variants of those AMS algorithms, on UCIrvine datasets, and found that the same heuristic algorithms continue to work well.

However, the framework in that paper is problematic, as it assumed that the researcher, learning the distribution, had to pay for information, but the physician, who would use that learned classifier, would not. We come full-circle in [KG05], which considers the situation where there are hard

*I gratefully acknowledge NSERC, iCORE and the Alberta Ingenuity Centre for Machine Learning for their generous support; and Aloak Kapoor, Dan Lizotte and Omid Madani for their insights on various related problems.

[1] In general, this can involve many components: financial, time delays, suffering, etc. We will assume these can all be accumulated into a single scalar value for each test.

budgetary constraints on *both the learner and the classifier*: e.g., the learner can spend at most $b_L = \$1000$ to learn the best $b_C = \$30$/patient classifier.

This task can be solved using a "double dynamic program": We can view this BL task (what to purchase, when) as a fixed-depth MDP: At any point, the "state" corresponds to the remaining budget b'_L and the set of $\langle$ instance, feature$\rangle$ probes performed and their results, typically encoded as a posterior distribution over features and labels. The set of actions correspond to the (remaining) $\langle$ instance, feature$\rangle$ probes. The rewards are all 0 for all internal states; when the budget reaches $b'_L = 0$, we use a DP to compute the policy that is optimal for the resultant distribution over test outcomes and diagnoses; the reward then is (the negative of) the expected regret of this policy. That is, we use a DP to traverse the possible purchases for $\approx b_L$ steps (assuming unit-cost features), then use a DP to compute the reward associated with each final state.

The resulting policy will be a (fixed-depth) decision tree with minimum expected regret, of the form: E.g., first probe $\langle$ i#21, f#2$\rangle$. If the response is $+$, then probe $\langle$ i#4, f#7$\rangle$, otherwise probe $\langle$ i#21, f#5$\rangle$, and so forth, until reaching the prescribed depth.

Of course, the obvious implementation here is too slow to be practical for any but the smallest of problems. (This is not suprising, given that this task inherits the NP-hardness of the simpler AMS task.) We therefore explore a variety of heuristic approaches for determining which $\langle$ instance, feature$\rangle$ in each situation, and observed empirically that PTW and randomized SFLA were consistently among the best.

Value of Information issues

There are issues related to VOI in both of these tasks:

Active Classifier: Here we are trying to find the "decision-theoretic best" classification for an instance, given complete information about the distribution over feature values and classification labels. We consider the challenges of identifying (i) the best such "active classifier" in general, and (ii) the best of the set of active classifiers constrained to purchase at most a fixed number of features. In both cases, it is critical to determine how much information about the class label is associated with each feature, which VOI computations can provide.

Budgeted Learning: Here we are trying to determine what sequence of $\langle$ instance, feature$\rangle$ information we need to extract *at learning time*, to produce the best classifier, focusing on the situation where we have only a limited budget to spend acquiring this information. At each stage, it is critical to determine how much additional information each $\langle$ instance, feature$\rangle$ provides about the feature-to-label *classifier*, which again corresponds to VOI.

References

D. A. Berry and B. Fristedt. *Bandit Problems: Sequential Allocation of Experiments*. Chapman and Hall, London, 1985.

Russell Greiner, Adam Grove, and Dan Roth. Learning cost-sensitive active classifiers. *Artificial Intelligence*, 139:137–174, Sept 2002.

A. Kapoor and R. Greiner. Learning and classifying under hard budgets. In *European Conference on Machine Learning (ECML)*, Porto, Portugal, October 2005.

Daniel J. Lizotte, Omid Madani, and Russell Greiner. Budgeted learning of naive-bayes classifiers. In *Proceedings of Uncertainty in Artificial Intelligence*, 2003.

Daniel J. Lizotte, Omid Madani, and Russell Greiner. Active model selection. In *Proceedings of Uncertainty in Artificial Intelligence*, 2004.