A Prediction Market Approach to Learning with Sequential Advice

Sindhu Kutty
Department of Computer Science and Engineering
University of Michigan, Ann Arbor
skutty@umich.edu

Rahul Sami
School of Information
University of Michigan, Ann Arbor
rsami@umich.edu

Abstract

We consider a new class of online classification problems motivated by Internet recommendation and forecasting applications in which the learner receives advice sequentially over time from experts who may be adversarial or genuine. We show that, for this set of problems, the use of a *market trading* metaphor is useful in constructing a learning algorithm. We illustrate this by considering the concrete problem of learning prediction sequences under partial monitoring. We use a non-traditional definition of regret under certain analytical assumptions. In a setting with m items, we prove that a measure of regret with respect to the collective information held by n experts is bounded by $O(n\sqrt{m}\log m)$.

1 Introduction

Many online ranking, recommendation, and personalization systems rely on input from multiple forecasters or experts. Combining multiple forecasters' inputs appropriately is the central goal of a rich machine learning literature, and computational learning algorithms underpin many of these systems, but social media domains present a unique combination of challenges to effective aggregation. In this paper, we introduce a new learning model that captures some of the characteristic features of this forecast setting, and we present a technique to construct efficient learning algorithms for this class of problems.

In particular, expert forecasts and recommendations in Internet settings present the following challenges: First, the identities and motives of the individual forecasters are not always known. Some forecasters may provide best-effort forecasts, but in some cases the forecasters or their inputs may have vested interests in manipulating the system. For such attackers, it is often easy to create a sizeable number of *shill* or *sybil* accounts in order to manipulate the system. We argue that, for these domains, the best formal model of the set of experts is neither purely adversarial, nor purely stochastic, but a hybrid of the two. For forecasters with unknown motives, assuming that the forecasts are governed by a stationary stochastic generative process would be unrealistic. On the other hand, for genuine best-effort forecasters, a stochastic model of forecasts and forecast error is appropriate, and may lead to stronger performance guarantees.

Second, for any given item, forecast inputs from different sources often arrive haphazardly over time, and not all sources produce forecasts for each item. Critically, later forecasters may have access to information from earlier forecasts about the same event or item. Such a setting is vulnerable to *cloning* attacks, where a potentially harmful expert imitates the advice of a genuine, informative one. Further, a prediction may need to be made before all of the experts have reported their advice. In prior work, a partial availability of expert forecasts is typically handled by modeling so-called "sleeping experts", who may be inactive in certain rounds. However, these algorithms cannot distinguish between genuine forecasters and clones, even though the clones are forced to make forecasts later than the genuine forecasters they copy.

In this paper, we describe a technique to develop machine learning algorithms for forecast aggregation systems based on the metaphor of a *prediction market*. Prediction markets are markets that allow traders to bet on securities whose value depends on a future event; for example, the Iowa Electronic Market predicts the outcome of a presidential election. One form of prediction markets, which is rapidly gaining in popularity, is the *market scoring rule* [1]. In a market scoring rule, traders earn rewards proportional to the reduction in "loss" (measured using a proper scoring rule) caused by their trades; in other words, the difference in the loss of forecasting based on market price after their trade as compared to the market price after the previous trade. Our approach involves designing a learning algorithm by tracking a budget for each trader, and simulating a prediction market: for each input, the algorithm carries out a "trade" on the forecasters' behalf, and then later updates the budgets by treating received feedback as the prediction market outcomes.

Algorithms based on prediction markets are attractive for the particular features of the domains we are interested in, because of the following reasons: First, traders' budgets allow us to control the total net impact of a single identity. By coupling traders' payoffs to the effects of their actions, and limiting their effect so that their budget is never negative, we can provide worst-case bounds against adversarial forecasters. Second, in a setting with honest agents but stochastic outcomes, a budget-proportional betting scheme (the Kelly criterion [2]) leads to exponential growth in traders' budgets (in expectation), and thus the small initial budgets are not crippling to honest agents in the long run. Third, betting protocols have been used before in machine learning algorithms, for the reasons above (see, *e.g.*, [3]). Prediction markets are a natural extension of betting protocols to the sequential forecasting setting. Traders' profits are based on the extent to which they *change* forecasts, thus ensuring that merely cloning previous information is not profitable.

We illustrate this technique in the context of a specific recommendation problem: Consider a system that has to predict how attractive each of a set of items will be to a target user (or group of users). The system has access to forecasts from a set of experts, some of whom may actually be controlled by an attacker. Not all experts provide forecasts on every item, and they need not provide forecasts in a fixed order for every item. All experts are assumed to factor in previous information into their forecasts. Importantly, the system has limited access to feedback on recommendations: items that are not recommended highly might never be inspected by the target user. In section 2, we formally model this problem, and develop a prediction market-based learning algorithm for it.

Related Work: Chen and Vaughan [4] explore the connection between prediction markets and no-regret learning. In one direction, this can be used to develop efficient prediction markets, and in the other, to develop a particular class of learning algorithms. In contrast, our work exploits the sequential nature of a prediction market, and we argue by example that this feature of prediction markets can be used to develop a new class of learning algorithms. Prediction under limited feedback has been studied under different modeling assumptions in prior work [5, 6]. Our feedback model differs from the previous models of partial information. In particular, the multi-armed bandit setting differs from ours in that the loss of the actions are intrinsically independent of each other. The most natural definition of an action in our case is the prediction. Receiving a feedback on any prediction for an item label is tantamount to receiving feedback on *all* predictions for that item. With the general forecaster algorithm [6], the problem can only be posed with strong assumptions to yield weaker bounds than those we achieve here. Kleinberg et. al. [7] present the sleeping experts problem. This is similar to our model in capturing the fact that not all experts provide feedback on all items. Unlike our model, however, they assume that the expert advice is available simultaneously. A hybrid analytical approach to the multi-armed bandit problem has been studied by Lazaric and Munos [8]. They consider a model with stochastic input space and adversarial labels; in contrast, our hybrid model features adversarial experts and stochastic experts. Resnick and Sami [9] consider a full-feedback variant of the specific recommendation problem studied here. Although our algorithm is similar to theirs, we have elaborated on the deeper connection between prediction markets and learning algorithms. Additionally, as we study a limited feedback setting, our analysis and bounds are somewhat different. Yu et. al. [10] specifically consider a system to thwart sybil attacks in recommender systems. They do not consider a sequential ordering on the experts; further, unlike our model, their model assumes a strong similarity between the actual labels and the forecasts of *some* expert.

2 Model and Results

For concreteness of analysis, we focus on a specific problem motivated by a recommender setting: Suppose that every day, a single news article is created, that will potentially be assigned a binary label $l_i \in \{0, 1\}$ by the target, if inspected. Throughout the day, input forecasts on classification come in one at a time from a set of experts, some subset of which are controlled by an attacker, while the rest are honest. Each expert has access to all previous forecasts. We assume that genuine experts can perfectly aggregate previous information and hence that the last genuine expert's forecast is an unbiased estimate of the true label based on information from all honest experts. The challenge is to weed out the misleading forecasts by malicious experts while retaining the information provided by genuine forecasters. We make no assumption about the number or fraction of adversarial experts in the pool. The forecast output by the learning algorithm is a real value between 0 and 1. We assume that articles for which the target got a forecast below a certain threshold q_T will not be inspected, and thus we will never find out the true label of these items. This feedback model seems most natural for this class of problems, but it differs from the traditional partial-monitoring settings, necessitating a new algorithm and analysis.

We will now describe our solution to this learning problem. We use the metaphor of a quadratic market scoring rule to construct the algorithm. λ and ϵ are algorithm parameters that determine different components of loss. We track reputations R_j for each expert j corresponding to budgets of traders in a market. Initial reputations are fixed at $e^{-\lambda}$ where the algorithm parameter λ can be used to control the compromise between the damage caused by attackers and loss of information from genuine experts. ϵ helps control the feedback mechanism in our model. Expert forecasts arrive sequentially, and each expert provides an aggregation of all previous expert forecasts. When an expert with reputation $R_j < 1/\epsilon$ makes a forecast, we move the (market) probability by an amount proportional to R_j, ensuring that the budgets will always remain positive. Thus, this expert is influence-limited by his reputation. For $R_j \geq 1/\epsilon$, the expert has maximum reputation, and hence his aggregation is used undiluted. The term ϵR_j is known as his influence limit; those experts with $\epsilon R_j \geq 1$ are not influence limited *i.e.,* their influence limit is 1. In order to control for selection bias in the feedback that is received, we randomly raise the forecast for a small fraction ϵ of items to the threshold q_T. We assume that all items that have predictions above this threshold receive feedback. If and when feedback is received, the reputations of the experts are updated according to a modified market scoring rule. For each expert who made a prediction, his reputation is increased in expectation by his influence limit times the incremental decrease in prediction loss due to him. Note that we score experts on their aggregations rather than their individual predictions. The algorithm is similar in structure to the algorithm by Resnick and Sami [9], but modified for the particular partial monitoring feedback model we consider.

Using the fact that budget changes are tied to the actual increase or decrease in loss due to an expert, we immediately get the following result:

Theorem 1 *(Limited Expected Damage) If an attacker controls η experts, the total expected increase in prediction loss is at most $\eta e^{-\lambda}$ on any one item where the expectation is over the random coin flips of the algorithm. This result is with respect to an adversarial model of experts whose forecasts may be arbitrarily chosen.*

We will now define the model in which we analyze how much information is lost from truly informative experts. In this model, the world is assumed to be divided into states with each state corresponding to a $\{0, 1\}$ prediction. The states of the world occur with some fixed and known *prior* probability as used by our algorithm. Each expert sees the world as partitioned, based on his private information. Thus, he has a prediction probability based on the partition he identifies. For an honest expert, this corresponds exactly to the true probability in that partition. This hypothesis is at the core of our hybrid analysis. Note that the states of the world could be rich enough to include item features, historical data, and any other information the expert may use to come up with his prediction. The *cumulative* information up to an honest expert j corresponds to a prediction denoted by q_j which corresponds exactly to a sequential partition refinement based on all honest experts' information up to expert j. This assumption precludes a situation in which no stochastic state space describes the pattern of predictions by honest experts. One such instance can occur when honest experts condition their information on predictions made by attacker identities, which may be chosen according to a dynamic strategy rather than a stationary distribution.This admittedly strong assump-

tion comes from our myopic definitions of damage and information loss: since these are defined incrementally, loss due to imperfect aggregation of information (or the effects of attackers on honest raters' aggregate predictions) are not accounted for.

In the algorithm, the movement of predictions due to an expert's advice (corresponding to a trade) is proportional to his current reputation (or budget), when he is influence-(or budget-) limited. This allows us to prove a result on the expected growth of reputation among honest experts, thereby bounding the information lost from these experts:

Theorem 2 *(Limited Expected Information Loss) Suppose expert j has made predictions for m items. Let h_j denote the expected reduction in prediction loss due to j's prediction. Then, for all sufficiently large m, expert j's expected reputation (with MSE loss) is bounded below by*

$$E(R_j) \geq mh_j - (2\lambda/\epsilon) + (2/\epsilon)\log\epsilon - (2/\epsilon)\log[mh_j\epsilon - 2\lambda + 2\log\epsilon]$$

Here the expectation is with respect to a stationary distribution on the forecasts of honest experts.

From the definition of h_j we see that ideally, we would like to be able to extract $\sum_{j=1}^{n} mh_j$ information in its entirety from the total number, n, of experts. However, even if every honest expert performed a loss-less aggregation of the previous honest experts' information, we would still lose some of the information due to influence limiting. Theorem 2 bounds this loss. We can combine the two bounds above in a metric of regret: we define the regret of an algorithm as the difference between the actual loss and the ideal loss achievable if the algorithm knew which subset of experts was honest. Intuitively, this definition of regret captures the fact that, in hindsight, we would be able to pinpoint the truly informative experts and *not* influence-limit them, while not suffering any damage due to the other experts. We provide an upper bound on regret that is sublinear in m, the number of items. Under this definition, the Partial Feedback Influence Limiter achieves an asymptotic regret bound of $O(n\sqrt{m}\log m)$ for an appropriate choice of algorithm parameters ϵ and λ where regret is measured with respect to the expected reduction in prediction loss due to n experts across m items.

3 Conclusions and Future Work

In this work, we presented a model and an algorithm that uses sequentially arriving expert advice to make predictions on the class label of an item under partial monitoring conditions. Our model for partial feedback assumes availability of feedback as a function of the predicted probability. The main technique for algorithm construction involves a reputation system that implements a market scoring rule. We treat every expert as responsible for aggregating the predictions received so far, and thus score them based on this reported aggregation. It is with respect to these reports that we measure the regret of the algorithm. We showed that for particular choices of algorithm parameters, we are able to achieve $O(n\sqrt{m}\log m)$ regret.

We also proved separate information loss and damage bounds. Different choices of the algorithm parameter allow us to shrink one bound at the expense of the other. This can be exploited with specific domain knowledge to thwart attacks from malicious entities. Although this sequential analysis is useful, we do not claim it to be universally applicable: one negative consequence of using the sequence information is that, as the first contributor of a piece of information is disproportionately rewarded, it could create incentives to race if information is public and freely accessible.

One strong assumption in our model is on the aggregation process: our analysis treats each honest expert's prediction as if it ideally aggregated all prior information. It would be useful to relax this assumption and take into account the actual process by which the experts aggregate prior information (from attackers as well as honest raters) into their forecast. We are currently exploring a model in which expert aggregates are derived by Maximum Likelihood Estimates (MLE) of sufficient statistics of a probability distribution. In this case, preliminary results show that MLE is in fact equivalent to a prediction market of traders with infinite budgets whose beliefs are reflected by the means of the reported sufficient statistics. It would be interesting to ask whether in this case, in addition to a hybrid bound, we can also recover a bound for the traditional definition of regret.

Acknowledgement

This material is based upon work supported by the National Science Foundation under grants IIS-0812042 and CCF-0728768. We are grateful to Satinder Singh and Paul Resnick for providing valuable feedback and suggestions.

References

[1] Robin Hanson. Combinatorial information market design. In *Information Systems Frontiers*, pages 107–119, 2003.

[2] J. L. Kelly. A new interpretation of information rate. *Bell System Technical Journal*, 35:917–926, 1956.

[3] Glenn Shafer and Vladimir Vovk. *Probability and Finance: It's Only a Game!* John Wiley and Sons, 2001.

[4] Yiling Chen and Jennifer Wortman Vaughan. A New Understanding of Prediction Markets Via No-Regret Learning. In *Proceedings of the ACM Conference on Electronic Commerce (EC'10)*, 2010.

[5] D. P. Helmbold, N. Littlestone, and P. M. Long. Apple tasting and nearly one-sided learning. In *SFCS '92: Proceedings of the 33rd Annual Symposium on Foundations of Computer Science*, pages 493–502. IEEE Computer Society, 1992.

[6] Nicolo Cesa-Bianchi and Gabor Lugosi. *Prediction, Learning, and Games*. Cambridge University Press, 2006.

[7] Robert D. Kleinberg, Alexandru Niculescu-Mizil, and Yogeshwer Sharma. Regret bounds for sleeping experts and bandits. In *Conference on Learning Theory (COLT)*, pages 425–436, 2008.

[8] Alessandro Lazaric and Rémi Munos. Hybrid stochastic-adversarial on-line learning. In *Conference on Learning Theory (COLT)*, 2009.

[9] Paul Resnick and Rahul Sami. The influence limiter: Provably manipulation-resistant recommender systems. In *Proceedings of the ACM Recommender Systems Conference (RecSys07)*, 2007.

[10] Haifeng Yu, Chenwei Shi, Michael Kaminsky, Phillip B. Gibbons, and Feng Xiao. Dsybil: Optimal sybil-resistance for recommendation systems. In *IEEE Symposium on Security and Privacy*, pages 283–298, Los Alamitos, CA, USA, 2009. IEEE Computer Society.