Collaborative Information Acquisition (Talk)

Danxia Kong Danxia.Kong@phd.mccombs.utexas.edu
Maytal Saar-Tsechansky maytal@mail.utexas.edu
The University of Texas at Austin, McCombs School of Business

Abstract

Most information acquisition policies are accuracy-centric–they aim to improve the predictive accuracy of a given model. However, often in practice, a predictive model is used along with other models to inform arbitrarily complex decisions. In light of such practices, this paper discusses the need for a new kind of collaborative information acquisition (CIA) policies, where multiple predictive models which inform a common set of decisions collaboratively prioritize information acquisitions to promote the decisions they inform. We present a framework and a particular CIA policy and we demonstrate it yields superior decision performance as compared to alternatives.

1. Introduction

Work on active information acquisition addresses the development of policies with the objective to improve predictive models through cost-effective information acquisitions. This research has produced effective *accuracy-centric* policies that improve the accuracy of a variety of models through the cost-effective acquisition of different types of information, including class labels, feature values, and different sets thereof (e.g., Cohn et al., 1994; Roy and McCallum, 2001; Lizotte et al., 2003; Saar-Tsechansky and Provost, 2004; Melville et al., 2004). However, very often in practice, an estimate from a predictive model is used to inform routine decisions. In business, for example, predictive models often inform the selection of one among alternative courses of action. Because the objective of accuracy-centric information acquisitions policies is to improve predictive accuracy, such policies do not consider the impact of prospective acquisitions on the decisions that the learned models inform. It is therefore important to question whether costly improvements in model accuracy are always desirable for improving decision quality. If greedy improvement in model accuracy is not always beneficial, it is important to understand whether and how information about the decisions can be exploited to identify acquisitions that improve these decisions, rather than the models, cost-effectively. Saar-Tsechansky and Provost (2007) have identified the need for *decision-centric* information acquisition policies. For a simple, general decision task, they show that some costly improvements in predictive accuracy pursued by accuracy-centric policies do not always yield improvement in decision performance. They also propose a new policy, GOAL, which exploits knowledge on the decisions informed by a class probability estimation model to identify advantageous acquisitions. Importantly, the acquisitions pursued by GOAL improve the decisions better, per acquisition cost, than what can be achieved through a greedy improvement of the model's accuracy.

GOAL aims to gauge how improvements in a class probability estimation model impacts the decisions the model informs. However, in many settings, estimates from multiple predictive models collectively inform a common decision. Tax audit decisions, for example, are derived from the predicted probability of tax fraud, but also from the predicted revenue that might be recovered. If one aims to acquire audits to improve future audit decisions, it is important to acknowledge that prospective acquisitions may affect both models. Interestingly, for each model, a different audit acquisition might be more desirable. Hence, in this paper, we study a new kind of *collaborative* information acquisitions (CIA), in which multiple modeling tasks collaborate to prioritize information acquisitions that benefit the decisions the most. We study a particular setting exemplified by the tax fraud problem above, where a given acquisition (e.g., an audit) can augment the training data of multiple predictive models, all of which inform each decision . We develop a framework for addressing this challenge and present a policy de-

Appearing in *Proceedings of the 27^{th} International Conference on Machine Learning*, Haifa, Israel, 2010. Copyright 2010 by the author(s)/owner(s).

signed for the setting we explore. Our empirical results for different decision settings suggest the our CIA policy yields better decisions per number of acquisitions as compared to GOAL as well as compared to decision-centric policies which aim to improve the underlying predictive models' accuracies.

2. Framework for Collaborative Information Acquisition

Assume a domain of decisions D, in which each decision $i \in D$ is informed by multiple predictive models. The general problem of collaborative information acquisition is to identify acquisitions that will augment the training data from which the models are induced so as to improve a given decision-making objective the most, per acquisition cost. Possible decision-making objectives include maximizing the *number* of correct decisions in D, or maximize the total economic utility (such as profit) from decisions in D. For simplicity, we will refer to the decision making objective as *utility*. Note that a key distinction from existing frameworks for information acquisition is that CIA is not concerned with improving any particular model's predictive accuracy. Rather, all the modeling tasks collaborate to select acquisitions that improve any of the underlying models informing the decisions in D, so as to promote a common goal.

Our approach aims to examine how different prospective acquisitions may affect all the models informing the decisions, and, consequently, how these changes may influence the decisions themselves.

To estimate changes in the models' estimations and in the subsequent decisions following possible acquisitions, our approach acknowledges the dynamic nature of sequential information acquisition. In particular, rather than consider the implications of augmenting the current training data with a given acquisition, we aim to evaluate the impact of augmenting different versions of the training data with prospective acquisitions. In addition, we adopt a statistical perspective on the inferences of the models induced from the different instantiations of the training data and the decisions they inform. Specifically, we treat the contribution that a prospective acquisition q will have on a particular decision i as a random variable whose distribution is determined by possible instantiations of the current training data. Future versions of the training data T are determined by the current acquisition under consideration, but also by future acquisitions. Therefore, our objective is to estimate how a prospective acquisition will affect the *expected* utility derived from decisions in D, given the inductive techniques used to inform these decisions, and where expectation is calculated over the possible instantiations of the training data. Formally, given a probability density function $f(T)$ over the possible instantiations of the current training set, the total expected benefit over the domain of decisions D from acquiring q is given by $B(q,D) = \sum_{i \in D} \int_T \widehat{U}\left(i|T_q\right) f(T) d(T)$, where $\widehat{U}\left(i|T_q\right)$ is the estimated utility derived from decision i given the training data T augmented with acquisition q. Lastly, we include the estimated utility from a decision i, $\widehat{U}\left(i|T_q\right)$, since the actual utility is not known at the time an acquisition is made. In principle, our approach for estimating the benefit to decision-making from each prospective acquisition is independent of the modeling techniques used or type of information that can be acquired. An implementation of this framework requires various estimations, including estimation of the distribution of training sets, and estimation of the decision utility for any training set augmented with acquisition q.

3. A Policy for Collaborative Information Acquisition

In this section we consider a specific class of decision problems exemplified by the sales tax audit problem discussed earlier, and we develop a collaborative information acquisition policy for improving the corresponding decisions the most per number of acquisitions. We later discuss future work addressing other decision-making settings.

We consider a set of decisions D, such that for each decision $i \in D$ one has to select one of two actions. As noted above, this problem corresponds to many common business decisions, including customer retention, direct marketing, and tax audit selection. For concreteness, assume that for each customer i, a decision-maker has to decide whether or not to make a costly offer so as to maximize the expected revenue. One possible approach for estimating the expected benefit from an offer is to estimate the purchase amount this amount would be zero whenever a customer does not make a purchase and the actual purchase amount otherwise. As has been noted in prior work (Zadrozny and Elkan, 2001), this approach often does not yield good estimations. This is because only a fraction of training instances include a non-zero purchase amount; in addition, purchase amounts are negatively correlated with the likelihood of purchase. Hence, instances with high purchase amounts are often substantially underestimated. An alternative approach which is employed widely in practice and which often yields superior decisions is to view the customer's behavior as a two-phase process. Specifically, one first estimates the probability any given customer i will respond to the offer once made, and another model estimates the amount the customer will spend, should the customer responds to the offer.

The theoretic-optimal choice of whether or not to initiate an offer is determined by the action with the highest expected value. Viewing the customer's behavior as a two-phase process, the expected value from a solicitation is estimated as $(p * s) - C$, where p is the probability of purchase, s is the purchase amount, and C is the solicitation cost.

In the setting we consider here, each customer corresponds to a solicitation decision, and, similarly, an information acquisition corresponds to obtaining a customer's response to a costly solicitation. We assume that a set of predictors is available for each customer, and that information on whether or not a customer will respond to the offer, along with the customer's spending amount can be acquired through a solicitation. A customer's response corresponds to the dependent variable of the model estimating the probability of response (henceforth refers to the response model), and the spending mount pertains to the dependent variable of the spending model. Acquisition of tax audits with the objective to maximize future revenue collections, constitute a similar information acquisition setting, in which whether or not a case is fraudulent the fraud amount can be obtained through a costly audit.

As we note earlier, different instantiations of the training data might lead to different estimations and consequently different decisions. Thus, we aim to select acquisitions that will increase the sum of the *likelihood* of correctly selecting the highest expected value action for decisions in D. Formally, the benefit from acquisition q is given by: $B(q, D) = \sum_{i \in D} \int\limits_{T} \left(P(\hat{d}_i = d_i | T_q) \right) f(T) d(T)$, where $\hat{d}_i$ and d_i are the estimated and actual optimal choice for decision i, respectively.

In our implementation of the policy we employ several heuristics to estimate the benefit from an acquisitions. Specifically, we approximate possible instantiations of the training data by generating multiple bootstrap samples of the current training data. In addition, we use a proxy to gauge the impact of each customer solicitation on decision in D. In particular, we acquire information on cases (customers) for which decisions inferred from the possible instantiations of the training data are more likely to be incorrect. Thus, in our implementation, we implicitly assume that acquiring information on incorrect decisions is likely to improve the likelihood of correct decisions in D, over different possible versions of the training data.

For each version of the data T_j, we induce the set of models $M_j^1, M_j^2, ..., M_j^m$, informing the decisions in D. For each set of models induced from T_j and for each instance i, we estimate the expected value from each alternative action to yield an estimate $\hat{d}_i^j$ of the theoretic-optimal decision with the highest estimated expected value. Note that for each instance i and data version T_j, the selection of the theoretic-optimal action is determined by estimations from all the models informing the decision, $M_j^1, M_j^2, ..., M_j^m$, i.e., the response model and the spending model in our setting. Because the actual theoretic-optimal choice is not known for each i, we capture the likelihood of inferring the incorrect choice from the diversity among the decisions produced from different data version T_j. In particular, in the extreme case that the same decision is inferred from all different versions of the training data, we assume the decision is likely to be correct and no additional information is necessary to improve it. By contrast, the more diverse the decisions produced from different versions of the data, we infer a higher likelihood of not selecting the optimal action for i. We capture the diversity among decisions by the absolute difference between the number of data versions yielding each decision.

Note that our policy does not aim to identify which of the models $M_j^1, M_j^2, ..., M_j^m$ can benefit from additional acquisitions; rather, our policy aims to gauge the likelihood that collectively, the models are likely to yield an incorrect decision and hence benefit from additional information. A pseudo code of our collaborative information acquisition (CIA) policy is show in Algorithm 1.

4. Empirical Evaluation

To study the effectiveness of our policy it is necessary to use data on decisions informed by multiple predictive models. One public data set with this property is the 1998 KDD-cup data (Hettich and Bay,1999). This data includes information on donors who received a solicitation in a recent mailing campaign, along with each donor's response and amount donated, whenever a donation was made. For this campaign, we aim to acquire solicitations to improve the identification of profitable candidates for solicitations, so as to maximize the campaign profit. Specifically, for each prospective

Algorithm 1 CIA: Collaborative Information Acquisition Policy

```
Input: Set of prospective acquisitions $UL$ each corresponding to a decision, an initial training set $L$, induction techniques $I_1, I_2, ..., I_k$, batch size $Q$, a stopping criteria
While (stopping criteria is not met) Do
    1. Generate $m$ bootstrap samples $T_1, T_2, ..., T_m$, from $L$
       # Generate models to inform decisions:
    2. For $j = 1$ to $m$ Do
       For $s = 1$ to $k$ Do
       1) Apply $I_s$ on $T_j$, resulting in model $M_j^s$;
       2) Apply $M_j^s$ to each instance $i$ in $UL$
       End For
       End For
    3. For each data sample $T_j$ and instance $i$ in $UL$ estimate the theoretic-optimal decision $\hat{d}_i^j$
    4. Acquire the set $B$ of the top instances for which the decisions $\hat{d}_i^j$, $j = 1, 2, ..., m$, are most split
    5. Remove $B$ from $UL$, acquire $B$ and add acquired information to $L$
End While
Induce models $I_1, I_2, ..., I_k$ from $L$ to in inform decisions
Output: model $M_1$, model $M_2$, ..., model $M_k$ generated from $L$
```

donor we compare the expected value from a solicitation with the expected value from no solicitation. To inform these decisions, we induce two predictive models: a response model which estimates the probability of response for each donor, and a revenue model estimating the amount one will contribute in the case of donation. We use j48, WEKA's (Witten and Frank, 1999) implementation of the C4.5 classification tree induction algorithm (Quinlan, 1993), to induce the response probability estimation model, and WEKA's implementation of simple linear regression to estimate the revenue in the case of a donation. We also followed prior work for the choice of predictors and used only instances with actual donations to model the donation amount (Zadrozny and Elkan, 2001).

Differently from work on accuracy-centric information acquisition policies, we are not concerned with the predictive accuracy of the models we induce, but with the quality of the decisions they inform. Thus, we compare the campaign profit obtained by acquisitions of the CIA policy with the profits obtained by four alternative policies. First, because decisions are informed by the estimations of a regression (real value estimation) model and a classification model, we compare CIA with an information acquisition policy that aims to improve the regression model's estimation (Krogh and Vedelsby, 1995), as well as with the Uncertainly Sampling policy (Lewis and Gale, 1994), which aims to improve the estimations of the donors' response. We also compare CIA with GOAL (Saar-Tsechansky and Foster, 2007). As we discuss above, GOAL is a decision-centric acquisition policy which aims to improve the class probability estimations whenever such improvements promote better decisions. Importantly, GOAL pursues improvements in the classification model's estimation only when these improvements are likely to yield better decisions. However, GOAL does not consider how other estimations which inform the decisions might be affected by information acquisition. Specifically for our setting, GOAL does not consider how its acquisitions might impact the regression model's performance and thereby the decisions it informs. Finally, we also evaluate the performance of a uniform acquisition policy which assumes the benefit from acquiring information on any prospective donors is uniform. The uniform acquisition policy acquires solicitations by drawing uniformly at random from the pool of prospective donors. As we will see, because multiple models are induced from the training data acquired by each policy, and because the acquired data may be biased to benefit one model (possibly at the expense of the other model), a uniform policy often performs quite well in this setting.

The solicitation task defined by the 1998 KDD cup competition included a very small, fixed solicitation cost for each donor. In this setting, the trivial solution of soliciting to all customers yields good performance, and improvements using estimations of the theoretic-optimal choices yield only marginal profit benefits. To increase the risk (loss) incurred by futile solicitations and to allow us to easily capture the differences in performance among different policies, we created different versions of the problem with increasing solicitation costs. The set of solicitation costs we consider are: \$1, \$3, \$5, \$7, \$9, and \$11. Because the objective of this empirical evaluation is to compare among the relative decision performance enabled by different acquisition policies, and given the limited public data available on decisions informed by multiple models, we created different versions of the data by varying the proportion of donor in each. We have recently acquired an additional, proprietary data set on fraud audits, and hope to include new results for this data in a later version of

Table 1. Data Sets Information

data sets information		
	20% set	40% set
number of donors	1002	2048
number of non-donors	4019	3083
percent of donors	0.1996	0.3992
maximum donation	200	200
minimum donation	1	1
average donation	31.67	31.175
median donation	13	13

this paper. Table 1 presents statistics of the two data sets we produced.

The results we report are averaged over 100 random experiments. In each experiment, we randomly partition the data into an initial training set, a pool of donors whose solicitations can be acquired at a cost, and a test set. To reduce experimental variance, the same partitions are used to evaluate all the policies. At each iteration, each policy selects a sample of B donors whose responses and donation amounts are acquired and added to the corresponding policy's training set. For each policy we induce the response and revenue models from the corresponding policy's training data after each batch acquisition; the induced models are then used to estimate theoretic-optimal actions for prospective donors in the test set.

For each data set and decision setting, we present the average profit per capita obtained by each policy after each acquisition batch. In addition, we compute the average difference between the profit per capita obtained by CIA and each of the alternative policies across all acquisition phases. For each pair-wise comparison, we note whether the reported average difference is statistically significant according to a paired t-test ($p \leq 0.05$) .

4.1. A comparison between cia, goal and a uniform acquisition policy

Perhaps the most interesting comparison is between CIA and the alternative decision-centric policy, GOAL (Saar-Tsechansky and Provost, 2007). Recall that GOAL considers the decision for each donor and pursues acquisitions that can improve the response probability estimation if the improvement is likely to impact the decision. However, GOAL does not consider other estimations that may be influenced by each acquisition, nor their impact on selecting the theoretic-optimal choice.

Figure 1 through 6 show the profit obtained by CIA, GOAL and the uniform policy for different problem settings. Table 2 presents the average increase in profit per capita obtained by CIA as compared to GOAL and the uniform policy on 20 percent donor set. Table 3 presents the average increase in profit per capita obtained with CIA against Goal on 40 percent donor dataset across different cost settings.

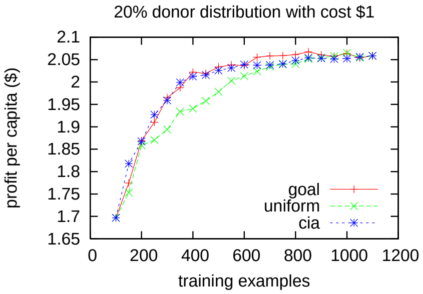

Figure 1. CIA vs. Goal vs. Uniform on 20% donor dataset with solicitation cost of $1

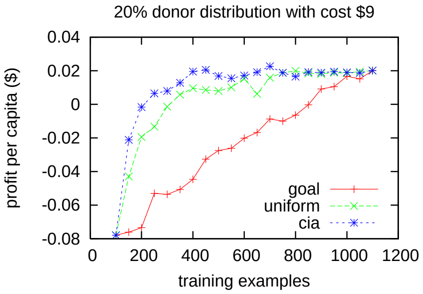

Figure 2. CIA vs. Goal vs. Uniform on 20% donor dataset with solicitation cost of $9

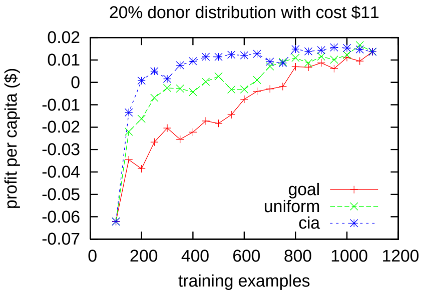

Figure 3. CIA vs. Goal vs. Uniform on 20% donor dataset with solicitation cost of $11

Overall, CIA consistently acquires informative acquisitions for the different decision tasks. As shown, we find that CIA is either comparable or substantially better than GOAL for all the decision settings we explore. We find that CIA's advantage is particularly substantial in

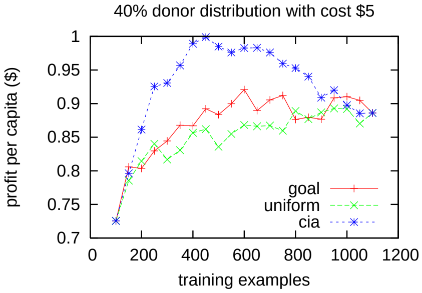

Figure 4. CIA vs. Goal vs. Uniform on 40% donor dataset with solicitation cost of $5

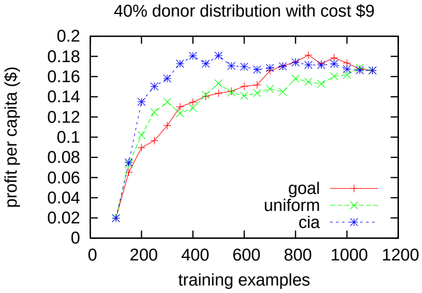

Figure 5. CIA vs. Goal vs. Uniform on 40% donor dataset with solicitation cost of $9

the more challenging decision settings in which the cost of sub-optimal decisions (i.e., the cost of solicitation) is high. In addition, because of the high solicitation costs, even actual donors are not profitable if the donation amount does not exceed the solicitation cost. In these settings, an accurate estimation of the donation amount as well as of the likelihood of response are both critical for identifying the optimal decision and avoiding costly sub-optimal ones. GOAL is more competitive when the solicitation risk (cost) is low and the accuracy of the estimated donation amount is not likely to have a significant impact on the collected revenue. When the cost of sub-optimal decisions is low, GOAL's selective improvement of donors responses can be sufficient to exhibit comparable performance to that of CIA.

Table 2. CIA versus Goal versus Uniform (unf) on 20 Percent Donor Dataset

cost ($)	CIA vs. Goal	CIA vs. unf
1	-0.0035	**0.0264**
3	-0.0009	**0.0289**
5	**0.0317**	**0.0082**
7	**0.0344**	**0.0057**
9	**0.0376**	**0.0069**
11	**0.0191**	**0.0078**

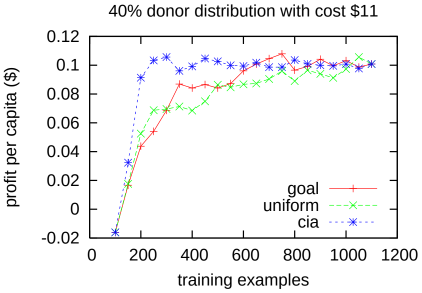

Figure 6. CIA vs. Goal vs. Uniform on 40% donor dataset with solicitation cost of $11

Table 3. CIA versus Goal versus Uniform (unf) on 40 Percent Donor Dataset

cost ($)	CIA vs. Goal	CIA vs. unf
1	**0.048**	**0.0249**
3	**0.0832**	**0.0607**
5	**0.0604**	**0.0821**
7	**0.0269**	**0.0264**
9	**0.0185**	**0.0228**
11	**0.0112**	**0.0162**

4.2. A comparison of CIA with accuracy-centric policies for revenue and response modeling

In this section we present results of a comparison between our collaborative decisions-centric policy and two accuracy-centric policies. The accuracy-centric policies aim to acquire information on customers if they are likely to improve either the revenue estimation or the estimation of a donor's response to a solicitation. Specifically, we compare CIA with an accuracy-centric policy by Krogh and Vedelsby (1995) which is designed to improve real-value estimation (regression) models. This policy empirically estimates the estimation variance of the model's predictions, and prefers acquisitions for which the variance is the highest. We also compare CIA with Uncertainly Sampling (Lewis and Gale, 1994) which aims to improve the donor response model.

Figure 7 and Figure 8 present the average campaign profit obtained by CIA, a regression accuracy-centric policy, and the uniform acquisition policy. Table 4 and Table 5 show the average increases in profit per capita obtained by CIA as compared to the regression accuracy-centric and the uniform policies. Figure 9, Figure 10, Table 4 and Table 5 show the same set of results for a comparison between CIA and the Uncer-

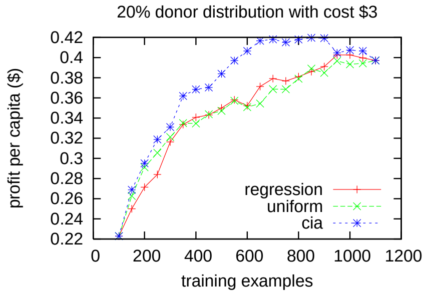

Figure 7. CIA vs. Regression vs. Uniform on 20% donor dataset with solicitation cost of $3

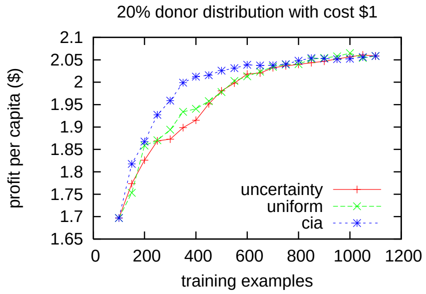

Figure 9. CIA vs. Uncertainty vs. Uniform on 20% donor dataset with solicitation cost of $1

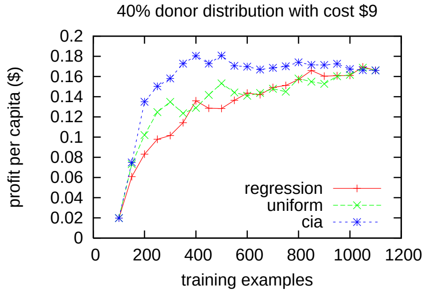

Figure 8. CIA vs. Regression vs. Uniform on 40% donor dataset with solicitation cost of $9

Table 4. CIA versus Regression (reg) versus Uncertainty (unc) on 20 Percent Donor Dataset

cost ($)	CIA vs. reg	CIA vs. unc
1	**0.0231**	**0.0333**
3	**0.0283**	**0.0315**
5	**0.0044**	**0.01**
7	**0.0008**	**0.0037**
9	**0.0058**	**0.0083**
11	**0.0061**	**0.0096**

tainty Sampling.

As shown, we find that for all the settings we explore CIA yields better decisions and consequently higher profits than those obtained by the accuracy centric policies. The benefits shown by CIA as compared to the accuracy-centric policies demonstrate that an accuracy-centric policy which improves the estimation of one of the models clearly ignores opportunities to improve decisions through improvements in the other models informing the decision. We also find that the accuracy-centric policies often obtain comparable decision performance to that of the uniform policy. These results suggest that greedy improvements in only one of the models may in fact undermine the predictive accuracy of other models informing the decision. Accuracy-centric information acquisitions introduce a bias into the training data. Thus, we conjecture that while such bias is beneficial for one of models, it may undermine other modeling tasks, yielding suboptimal choices.

Finally, note that as the potential for higher profits increases with a higher proportion of actual donors, the benefits from correctly identifying profitable donors via CIA are more substantial as compared to the alternatives. Overall, we find that as the risk is higher and when there are more opportunities for deriving profits, CIA's acquisitions yield more substantial benefits over the alternatives.

5. Conclusions, Limitations, and Future Work

We discuss a new kind of collaborative information acquisition policies with the objective to cost-effectively promote decisions informed by multiple predictive models. We develop a specific CIA policy and demonstrate the potential benefits for different decision settings. Our results also suggest several directions for future work. We find that CIA sometimes obtains its pick performance with only a subset of the acquisitions. This result suggests it would be beneficial to follow the framework we propose closely to develop a CIA policy which directly estimates the impact of each acquisition on the decision objective. We compare our CIA policy with the only alternative decision-centric policy; however, there are other accuracy-centric policies that can be considered. Finally, because of the lack of relevant, publicly available data, we produce different decision problems from a single original domain. It would be desirable to examine the performance of our policy on additional decision settings.

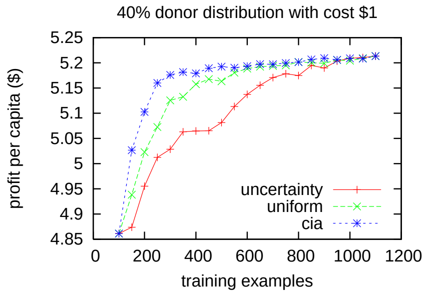

Figure 10. CIA vs. Uncertainty vs. Uniform on 40% donor dataset with solicitation cost of $1

Table 5. CIA vs. Regression (reg) vs. Uncertainty (unc) on 40 Percent Donor Dataset

cost ($)	CIA vs. reg	CIA vs. unc
1	**0.0325**	**0.0707**
3	**0.0681**	**0.1214**
5	**0.0745**	**0.1155**
7	**0.0297**	**0.0453**
9	**0.0288**	**0.0396**
11	**0.0112**	**0.0273**

References

Blake, C. L. and Merz, C. J. Uci repository of machine learning databases. Technical report, University of California, Department of Information and Computer Science, Irvine, CA, 1998.

Cohn, D., Atlas, L., and Ladner, R. Improving generalization with active learning. *Machine Learning*, 15(2):201–221, 1994.

Krogh, A. and Vedelsby, J. Neural network ensembles, cross validation, and active learning. In *Advances in Neural Information Processing Systems*, pp. 231–238. MIT Press, 1995.

Lewis, D. and Gale, W. A. A sequential algorithm for training text classifiers. In *Proceedings of the 6th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 1994.

Lizotte, D., Madani, O., and Greiner, R. Budgeted learning of naive-bayes classifiers. In *Proceedings of the 19th Conference on Knowledge Uncertainty in Artificial Intelligence (UAI-2003)*, 2003.

Melville, P., Saar-Tsechansky, M., Provost, F., and Mooney, R. Active feature-value acquisition for classifer induction. In *Proceedings of the 3rd International Conference on Data Mining*, 2004.

Quinlan, J. R. *C4.5: Programs for Machine Learning*. Morgan Kaufmann, 1993.

Roy, N. and McCallum, A. Toward optimal active learning through sampling estimation of error reduction. In *Proceedings of the 18th International Conference on Machine Learning (ICML-2001).*, pp. 441–448, 2001.

Saar-Tsechansky, M. and Provost, F. Active sampling for class probability estimation and ranking. *Machine Learning*, 54(2):153–178, 2004.

Saar-Tsechansky, M. and Provost, F. Decision-centric active learning of binary-outcome models. *Information Systems Research*, 18(1):1–19, 2007.

Witten, I. H. and Frank, E. *Data Mining: Practical Machine Learning Tools and Techniques with Java Implementations*. Morgan Kaufmann, 1999.

Zadrozny, B. and Elkan, C. Learning and making decisions when costs and probabilities are both unknown. In *Proceedings of the 7th International Conference on Knowledge Discovery and Data Mining*, pp. 204–212, CA, 2001.