Incentives and Machine Learning

John Hegeman

December 12, 2008

In 2007, US paid search ad spend was $8.9 billion - most of which used a pay per click (PPC) billing model. In PPC, advertisers only pay the search engine when their ad receives a click. Hence, the ability to accurately estimate an ad's click-through rate (CTR) is extremely valuable. The more likely that an ad is clicked, the more likely that the search engine is paid.

An important but often ignored feature of this environment, is the adversarial nature of the estimation problem. The cost per click (CPC) that an advertiser pays is often inversely proportional to the estimated CTR of the ad. Hence, an advertiser always prefers that the search engine assign a higher CTR estimate rather than a lower estimate. This is true regardless of the ad's true CTR. The conflicting objectives of the advertiser and the search engine wouldn't necessarily be a problem except that the advertiser often has control over the inputs to the learning algorithm and has some ability to distort them.

One might be tempted to think that this shouldn't matter at all. If the machine learning algorithm is continually trained on new data as it comes in, then it should automatically adjust to this behavior. The problem with this argument is that the resulting "naive" equilibrium results in a distribution of x that is less informative about y than the undistorted distribution is. The following example illustrates this point.

Example 1 *Ads are of two types. With probability p_l an ad is of type $\theta = l$ and has CTR y_l. With probability p_h an ad is of type $\theta = h$ and has CTR equal to y_h (we assume $y_h > y_l$). An ad's type is not observable. Instead, the learning algorithm must estimate y based on the landing page type, x. Ads of type l normally have landing page $x = h$ while ads of type h have landing page $x = h$. However, advertisers are able to send traffic through a landing page of the alternate type for a cost $c < y_h - y_l$. Advertisers maximize the utility function $U(\hat{y}) = E[\hat{y}] - c \cdot 1_{\{x \neq \theta\}}$ and the search engine has an objective function $E[(y - \hat{y})^2]$. Ignoring incentives, the optimal estimation strategy is to choose $\hat{y}$ to approximate $E[y|x]$. However, in our setting, the equilibrium resulting from this estimation strategy has all advertisers using landing page $x = h$ and consequently the search engine is no better off than if it hadn't used the landing page data at all.*

In contrast, the optimal learning algorithm for this example would choose $\hat{y}$ to approximate $\alpha E[y|x] + (1 - \alpha)E[y]$ with α chosen so that the benefit to

advertisers of type l from distorting their landing page is slightly less than the cost of distorting the landing page: $\alpha(E[y|h] - E[y|l]) = c - \epsilon$.

In what follows, we generalize this example to allow for a distribution of distortion costs and show that the optimal learning algorithm in this environment still seeks to approximate the weighted average $\alpha E[y|x] + (1-\alpha)E[y]$ rather than the standard $E[y|x]$. Unfortunately, the correct choice of α will depend on the distribution of the manipulation cost which is both unknown and impossible to estimate statically. However, this result can still provide some guidance in practical applications since the principal, if given enough time, could make small adjustments to α and then observe whether the resulting equilibrium is better or worse than the previous equilibrium.

1 The Model

A principal (the search engine in the above example) must estimate an unobserved characteristic y as a function of a characteristic x reported by an agent. The distribution of y is a function of the agent's type, $\theta \in \{0, 1\}$ which is unobserved by the principal. We assume $y \in [y_l, y_h]$ and $y_l < E[y|\theta = 0] < E[y|\theta = 1] < y_h$. After the principal observes x, she makes an estimate of y which we denote $\hat{y}(x)$. We will refer to $\hat{y}(x)$ as the estimation function. The principal's payoff is a function of the expected accuracy of the estimate:

$$U_p = -E[(y - \hat{y}(x))^2].$$

The agent's payoff is equal to the principal's estimate, $\hat{y}$, minus the distortion cost if the agent reported misreported x:

$$U_a(x; c, \theta) = \hat{y}(x) - c \cdot 1_{\{x \neq \theta\}}.$$

The cost of distortion, c, is distributed according to the distribution function $F(x)$. We assume $F(x)$ is continuous with density function $f(x)$ such that $F(y_l) = 0$ and $f(c) > 0$ for $c > 0$.

We will evaluate estimation functions in terms of their equilibrium performance. That is, we assume agents choose $x^*(c, \theta, \hat{y}) = \arg\max_x U_a(x; c, \theta, \hat{y})$. The principal's equilibrium payoff, U_p^* is her expected utility when agents choose $x = x^*$: $U_p^*(\hat{y}) = -E[(y - \hat{y}(x^*(c, \theta, \hat{y})))^2]$. We will use $\hat{y}_M$ to denote an estimation function that maximizes U_p^*.

2 Results

Our initial concern in Section 3.1 will be to analyze the optimal estimation function $\hat{y}(x)$ without regard for how $\hat{y}(x)$ might be approximated by a principal without knowledge of the distribution of y given θ. Section 3.2 will then turn

to methods for approximating the optimal estimation function using a converging series of estimation functions. We refer to such a series as an estimation strategy.

2.1 Estimation Functions

While we cannot solve for the optimal estimation function explicitly, the following proposition provides some insight into the form of $\hat{y}_M$ which will prove useful later.

Proposition 2 *There exists $\alpha \in [0,1]$ such that $\hat{y}_M(x) = \alpha E[y|x^*(c,\theta,\hat{y}_M) = x] + (1-\alpha)E[y]$.*

Proof. *First, observe that adding a constant to $\hat{y}$ has no effect on an agent's optimal report: $x^*(c,\theta,\hat{y}) = x^*(c,\theta,\hat{y}+a)$. This allows us to prove the following lemma:*

Lemma 3 *$E[\hat{y}_M(x^*(c,\theta,\hat{y}_M))] = E[y]$.*

Proof. *Let $\hat{y}'(x) = \hat{y}_M(x) + E[y] - E[\hat{y}_M(x^*(c,\theta,\hat{y}_M))]$. It follows that*

$$\begin{aligned}
U_p^*(\hat{y}') &= -E[(y - \hat{y}'(x^*(c,\theta,\hat{y})))^2] \\
&= -E[(y - \hat{y}_M(x^*(c,\theta,\hat{y})) - E[y] + E[\hat{y}_M(x^*(c,\theta,\hat{y}_M))])^2] \\
&= -E[((y - \hat{y}_M(x^*(c,\theta,\hat{y})))^2] + (E[y] - E[\hat{y}_M(x^*(c,\theta,\hat{y}_M))])^2 \\
&= U_p^*(\hat{y}_M) + (E[y] - E[\hat{y}_M(x^*(c,\theta,\hat{y}_M))])^2 \\
&\geq U_p^*(\hat{y}') + (E[y] - E[\hat{y}_M(x^*(c,\theta,\hat{y}_M))])^2
\end{aligned}$$

Thus, we must have that $(E[y] - E[\hat{y}_M(x^*(c,\theta,\hat{y}_M))])^2 = 0$ which implies $E[\hat{y}_M(x^*(c,\theta,\hat{y}_M))] = E[y]$. ■

Since only at agents of at most one type will pay the distortion cost, we will always have $E[\hat{y}_M(x^*(c,\theta,\hat{y}_M))] \neq E[y]$. This ensures that we can select a value of $\alpha \in (-\infty,\infty)$ such that $\hat{y}_M(0) = \alpha E[y|x^*(c,\theta,\hat{y}_M) = 0] + (1-\alpha)E[y]$. For a fixed value of $\hat{y}_M(0)$, there is a unique value of $\hat{y}_M(1)$ that ensures Lemma 1 is satisfied. Moreover, observe that $E[\alpha E[y|x^*(c,\theta,\hat{y}_M) = x] + (1-\alpha)E[y]] = E[y]$. Thus, if $\hat{y}_M(0) = \alpha E[y|x^*(c,\theta,\hat{y}_M) = 0] + (1-\alpha)E[y]$ then Lemma 1 requires that $\hat{y}_M(1) = \alpha E[y|x^*(c,\theta,\hat{y}_M) = 1] + (1-\alpha)E[y]$. It follows that there must exist α such that $\hat{y}_M(x) = \alpha E[y|x^*(c,\theta,\hat{y}_M) = x] + (1-\alpha)E[y]$. To complete the proof of the proposition, it remains to show that $\alpha \in (0,1)$.

We first rule out the case of $\alpha < 0$:

$$\begin{aligned}
&U_p^*(\hat{y}_M) = E[(y - \alpha E[y|x^*(c,\theta,\hat{y}_M)] - (1-\alpha)E[y])^2] \\
&= E[(y - E(y) - \alpha(E[y|x^*(c,\theta,\hat{y}_M)] - E[y])^2] \\
&= E[(y - E(y))^2] + \alpha(1-\alpha)E[(E[y|x^*(c,\theta,\hat{y}_M)] - E[y])^2] \\
&\geq E[(y - E(y))^2]
\end{aligned}$$

Thus, if α were less than 0 then the optimal estimation strategy would perform worse than simply using $\hat{y}(x) = E[y]$ which is a contradiction.

We now rule out the possibility of $\alpha > 1$. First, note that if $\alpha > 1$ then $\hat{y}_M(0) < E[y|x^*(c,\theta,\hat{y}_M) = 0]$. Consider the estimation strategy $\hat{y}'$ formed

by setting $\hat{y}'(0) = E[y|x^*(c,\theta,\hat{y}_M) = 0]$ *and* $\hat{y}'(1) = \hat{y}_M(1)$. *Note that the incentive for type* $\theta = 0$ *to report* $x = 1$ *has been strictly decreased and thus fewer agents will misreport. We can now group the agents into three categories: (1) those of type* $\theta = 0$ *who report* $x = 0$ *in response to either estimation strategy, (2) those of type* $\theta = 0$ *who report* $x = 1$ *in response to* $\hat{y}_M$ *but* $x = 0$ *in response to* $\hat{y}'$*, and (3) those who report* $x = 1$ *in response to either estimation strategy. Observe that for agents in groups (1) and (2)* $\hat{y}'$ *is closer than* $\hat{y}_M$ *to the expected value of* y *given which of the above three groups the agent is in. Since estimation accuracy is unaffected for agents in the third category this implies that* $U_p^*(\hat{y}') > U_p^*(\hat{y}_M)$ *which is a contradiction. It follows that* $\alpha \in [0,1]$ *for the optimal estimation strategy.* ■

2.2 Estimation Strategies

Proposition 2 tells us that the optimal estimation function always results in an equilibrium in which the predicted value of y is a simple weighted average of the unconditional expectation of y and the expectation of y conditional on the agent's reported x. Let $\tilde{E}[y|x = x_0]$ be the average value of y across all observations prior to i with $x = x_0$ (or y_H if x_0 has not been observed) and let $\tilde{E}[y]$ be the average value of y across all observations prior to i. This suggests that the optimal estimation function can be approximated by the estimation strategy $\{g^{(i)}(x)\}_{i=1}^{\infty}$ where $g^{(i)}(x) = \alpha\tilde{E}[y^{(i)}|x^{(i)}] + (1-\alpha)\tilde{E}[y^{(i)}]$. Due to space constraints we will not formally discuss convergence issues for estimation strategies and will simply assume that an estimation strategy of this form will converge to some estimation function $\hat{g}$. Note that by the definition of $g^{(i)}$, $\hat{g}$ must satisfy $\hat{g}(x) = \alpha E[y|x^*(c,\theta,\hat{g}) = x] + (1-\alpha)E[y]$. The following proposition ensures that if α is chosen correctly then this estimation strategy will converge to the optimal equilibrium.

Proposition 4 *Let* $\hat{y}$ *and* $\hat{y}'$ *satisfy* $\hat{y}(x) = \alpha E[y|x^*(c,\theta,\hat{y}) = x] + (1-\alpha)E[y]$ *and* $\hat{y}'(x) = \alpha E[y|x^*(c,\theta,\hat{y}') = x] + (1-\alpha)E[y]$. *It follows that* $\hat{y}(x) = \hat{y}'(x)$.

Proof. *Subtracting the expression for* $\hat{y}(0)$ *from the expression for* $\hat{y}(1)$ *yields*

$$\hat{y}(1) - \hat{y}(0) = \alpha(E[y|x^*(c,\theta,\hat{y}) = 1] - E[y|x^*(c,\theta,\hat{y}) = 0]) \qquad (1)$$

and similarly

$$\hat{y}'(1) - \hat{y}'(0) = \alpha(E[y|x^*(c,\theta,\hat{y}') = 1] - E[y|x^*(c,\theta,\hat{y}') = 0]). \qquad (2)$$

Observe that $\hat{y}(1) - \hat{y}(0) > \hat{y}'(1) - \hat{y}'(0)$ *implies that more agents of type* $\theta = 0$ *report* $x = 1$ *in response to* $\hat{y}$ *than in response to* $\hat{y}'$ *which implies that* $E[y|x^*(c,\theta,\hat{y}) = 1] < E[y|x^*(c,\theta,\hat{y}') = 0]$. *However,* $E[y|x^*(c,\theta,\hat{y}) = 0] = E[y|x^*(c,\theta,\hat{y}') = 0]$ *and thus subtracting equation (2) from (1) yields*

$$(\hat{y}(1) - \hat{y}(0)) - (\hat{y}'(1) - \hat{y}'(0)) = E[y|x^*(c,\theta,\hat{y}) = 1] - E[y|x^*(c,\theta,\hat{y}') = 0] \quad ■$$

Let α^* be the value of α such that $\hat{y}_M(x) = \alpha E[y|x^*(c,\theta,\hat{y}_M) = x] + (1-\alpha)E[y]$ (Proposition 2 guarantees that α^* exists).

Corollary 5 *The estimation strategy* $g_{a^*}^{(i)}(x) = \alpha^*\widetilde{E}[y^{(i)}|x^{(i)}] + (1-\alpha^*)\widetilde{E}[y^{(i)}]$ *converges to the optimal estimation function.*

Proof. *g_{a^*} converges to some $\hat{g}_{a^*}$ such that $\hat{g}_{a^*}(x) = \alpha^* E[y|x^*(c,\theta,\hat{g}_{a^*}) = x] + (1-\alpha^*)E[y]$. By proposition 4, this sufficient to ensure that the equilibrium outcome of the estimation strategy g_{a^*} is identical to that of $\hat{y}_M$.* ■

The following proposition improves on the range of possible values of α by ruling out $\alpha = 0$ and $\alpha = 1$:

Proposition 6 *The naive estimation strategy $g_1 = \tilde{E}[y|x]$ always outperforms the trivial estimation strategy $g_0 = \tilde{E}[y]$. However, it is never optimal. In particular, there exists $\alpha^* < 1$ such that $g_\alpha = \alpha\tilde{E}[y|x] + (1-\alpha)\tilde{E}[y]$ outperforms g_1 for every $\alpha \in (\alpha^*, 1)$.*

Proof. *For brevity, we provide only a sketch of the proof. The first claim (that g_1 outperforms g_0) follows immediately from the observation that not all agents choose to manipulate in the equilibrium resulting from g_1.*

For the second claim, consider the equilibrium resulting from the naive estimation strategy ($\alpha = 1$). Switching to the estimation strategy $\alpha = 1 - \epsilon$ has a cost of order ϵ^2 as $\hat{y}$ moves away from $E[y|x]$. However, it also causes a mass of order ϵ of type l advertisers from $x = x_h$ to $x = x_l$ which has a benefit of order ϵ. Thus, for sufficiently small ϵ the impact will be positive. ■

2.3 Agent Surplus

The previous analysis has been focused on the principal's equilibrium payoff. However, in a richer model, the principal might also care to some extent about the payoff to the agent. In our example setting of a search engine and advertisers, this could arise in a model with entry costs since a higher payoff to advertisers will result in more advertisers entering the market which will in turn could yield a higher payoff to the principal. We briefly touch on the agent's surplus with Proposition 7:

Proposition 7 *For estimation functions of the form $\hat{y}(x) = \alpha\widetilde{E}[y|x] + (1-\alpha)\widetilde{E}[y]$, equilibrium advertiser surplus is strictly decreasing in α for $\alpha > 0$.*

Proof. $E[U_a(x;c,\theta)] = E[\hat{y}(x) - c \cdot 1_{\{x \neq \theta\}}] = E[\hat{y}(x)] - c\Pr[x \neq \theta]$

First, note that $E[\alpha\widetilde{E}[y|x] + (1-\alpha)\widetilde{E}[y]] = E[y]$ is independent of α. Since the equilibrium value of $\hat{y}(1) - \hat{y}(0)$ is increasing in α, $\Pr[x \neq \theta]$ is also increasing in α and thus $E[U_a(x;c,\theta)]$ is decreasing in α. ■

3 Conclusion

In many practical machine learning applications, agents with objectives contrary to the objectives of the principal may have the ability to manipulate some or all of the input data. Ignoring incentive issues in such environments will result in a sub-optimal estimation accuracy. Our contribution is to demonstrate that the optimal learning algorithm in such an environment under-utilizes the manipulable input by using an objective function that is a weighted average of the conditional and unconditional expectations of the output variable.