Notes on Defining the Reward Function

Aleksey Kocherzhenko

January 20, 2021

Broadly speaking, subscribers to email ads fall into the following categories:

- "Unicorn": always open email, always buy (perfect subscribers who don't actually exist);
- "Next-best": (almost) always open email, sometimes buy;
- "Next-next": sometimes open email, sometimes buy;
- "False hope": (almost) always open email, never buy;
- "Occasional browser": sometimes open email, never buy;
- "Dead weight": never open email, never buy.

Typically, a company would want most of its subscribers to fall into the first three categories. In fact, if a customer buys *anything*, no matter how small, the email was probably worth sending, as the per-subscriber cost of email campaigns is very low.

In some cases (e.g., if the purpose is raising brand awareness/creating buzz), the fourth and fifth categories may be of interest too. However, unless "dead weight" subscribers can somehow be converted to other categories, they only waste marketing resourses.

A good reward metric should be flexible enough to account for multiple possible objectives of a marketing campaign. It should allow to emphasize, e.g., opening an email vs. buying a product vs. the average revenue per customer. Ideally, a marketing manager should be able to set the weights for these (and, possibly, other) objectives depending on the nature of a campaign. For instance:

- a campaign promoting premium, high-ticket items may emphasize revenue per customer;
- a blowout sale may emphasize the mere fact of making a purchase;
- a campaign aimed at shaping brand perception may emphasize opening an email, even if no purchase is subsequently made.

Such a composite metric requires a machine learning model trained on subscriber data that returns the relevant parameters: the probability of opening an email, P_o, the probability of making a purchase, P_p, and the predicted purchase amount, A. A separate model may be trained for each of these parameters, and the reward function, R, may be, e.g., the linear combination of their predictions, with weights that are decided on by the marketing department for a specific campaign:

$$R = w_o P_o + w_p P_p + w_A A.$$

Such a metric would help avoid a common engineering pitfall: assuming what a customer (marketing manager) wants. Instead, we would convert an intractable amount of subscriber data to simple, actionable metrics for each subscriber that the marketing manager could use to make the ultimate call on how to act. It could also make the model more modular, so that other metrics, such as the probability of a subscriber unsubscribing, could be easily incorporated, either from the outset or later on.