Question Type: Machine Learning Design

Duration: 40 Minutes
Difficulty: Medium
Domains: Marketing

Problem

LinkedIn wants to improve the conversion rate of its premium subscription using email campaigns. The current email campaign is generic across users.

#1 - How would you improve the email conversion rate?

#2 - How would you run an A/B testing on your model?

Solution

> **#1 -** How would you improve the email conversion rate?

[Candidate] When you say that the email is generic, does that mean that there's no personalized content embedded in the email?

[Interviewer] Yes, you can assume that.

[Candidate] Also, when users click the call-to-action button in the email, are they directed to the premium page for sign-ups?

[Interviewer] That's correct.

[Candidate] Lastly, can I get a clarification on when the generic email is dispatched to users?

[Interviewer] Currently, the campaign is an one-time email sent one-week after a user signs up for LinkedIn.

[Candidate] Thank you for the clarification. I see two opportunities of improving the conversion: (1) personalization and (2) timing. The email content can be personalized based on user profile and activity. Secondly, the email campaign should be sent to users when they are most likely to convert.

[Interviewer] What kind of data would you need to build personalization?

[Candidate] I would need four types of historical data:

Data Sources	Variables
User Profile	1. Occupation(s) 2. YoE(s) 3. Industry(s) 4. Education 5. Skills 6. Location
User Activity	1. Actively searching indicator 2. Months before conversion 3. Sign up date 4. Log-in timestamp 5. Log-off timestamp

User Connections	1. Connections 2. Messages 3. Views
Conversion	1. Conversion date 2. Product purchased

[Interviewer] Can you elaborate on how the data sources are applied?

[Candidate] The idea is to understand what kind of users converted to which of the premium product at what time point. Understanding this pattern can help build a prediction model to identify when a user is likely to convert.

[Interviewer] Interesting approach. How would you design your model?

[Candidate] I would create panel data that shows week to week snapshot of user's profile, actions, connections and conversions. For instance, the panel data resembles the following:

User	Week_since_sign_up	sales_occupation_ind	….	Viewed_job_post	Convert
1	1	1	….	0	0
1	2	1	….	0	0
1	3	0	….	0	0
1	4	0	….	0	0
….	….	….	….	...	...
2	1	0	….	1	0
2	2	0	….	1	1
2	3	0	….	0	0

For each user, there are a set of features, probably 40 to 60, that will predict whether a user converted to the premium subscription plan. As a baseline model, I can apply a model with high-interpretability, such as logistic regression or random forest, to predict that a user will convert on a given week.

Furthermore, I can break it down to the types of products they converted to. Such that the label set actually becomes:

1. Converted_professional
2. Converted_business
3. Converted_sales
4. Converted_hiring

Essentially, the problem becomes a multi-class modeling problem.

[Interviewer] How would you leverage the conversion model in the email campaign?

[Candidate] On a weekly basis, the model will predict the user's conversion status. If the prediction is that a user will convert, then dispatch an email campaign with a recommendation on a product the user will most likely convert to.

Interviewer Comments

The candidate demonstrated strong knowledge in machine learning design. The classification model that signals when an email should be dispatched makes sense. One quick feedback is that he could have taken a simpler approach that personlizes either the email timing or email content. When both are accommodated in a single model, it's difficult to isolate whether the improvement came from the improved content, timing, or the interaction of both. On the plus side, in terms of development and deployment, a simpler approach as an MVP is usually ideal. Nonetheless, the candidate displayed a strong performance.

> **#2 -** How would you run an A/B testing on your model?

[Candidate] Since conversions of groups with or without personalized emails are compared, I propose using T-test for comparing two-sampled proportions.

[Interviewer] How would select users for running you experimentation?

[Candidate] I would randomly sample users who recently signed up. Running the experiment on users who have not seen the product emails removes confounding effects such as the novelty effect. For instance, if current users receive generic emails they've already seen before, they may consider the emails as spam and less likely to convert.

[Interviewer] Okay, how long would you run your experiment?

[Candidate] I would run the experiment based on sample size and traffic required to achieve a statistical power of 0.80 in the experiment at 1% to 3% minimum detectable effect (MDE). If I had to pick a reasonable timeline, it would ideally be somewhere around one to two weeks.

[Interviewer] What would be your observation period?

[Candidate] Given that the model predicts conversion weekly, I believe four weeks should be reasonable. But, I think it's worth looking at the probability curve of user conversion since sign-up, then decide a week that makes the most sense given business constraints and statistical estimations.

[Interviewer] Can you elaborate more on your control vs treatment groups?

[Candidate] Certainly, the control group is the one that receives generic email while the treatment group is the one that receives personalized email based on intent + timing.

[Interviewer] Can you see a potential issue there? All the users in the control receive generic email while some users in the treatment group receive the personalized email. Then, given that the control saw more emails, wouldn't the conversion be higher for the generic version?

[Candidate] That's certainly be a possibility. Hence, I propose, on the fourth week for users in treatment group who have yet to receive emails, just send a generic email.

Interviewer Comments

The candidate understands how to run A/B testing to validate a model. First, he chose the right statistical testing to evaluate whether the personalized content on conversion has statistical significance. Second, he demonstrated that he understands fundamentals in A/B testing (i.e. power and MDE). Lastly, when asked about a potential issue with treatment group, he provided a solution that made sense.

Interviewer Assessment

In the statistics section, a candidate is assessed based on correctness and soundness of machine learning methodology, product sense and communication. For each criterion, the candidate is rated in the following scale: (5) superior, (4) good, (3) adequate, (2) marginal, (1) not competent.

Assessments	Ratings	Comments
Machine Learning Methodology	5	The candidate demonstrated strong knowledge in machine learning design. In the first problem, he proposed improving the email campaign that contains tailored message and dispatched when a user is likely to convert. In the second problem, he provided thorough explanations on how he would run A/B testing. When asked about potential gotchas, he was able to provide responses readily.
Product Sense	5	The candidate demonstrated that he understood product offerings on LinkedIn and user flow. He provided a design that fits the business and product models of LinkedIn.
Communication	5	The candidate ensured that he understood the problem with clarifying questions. He explained his thought process with clarity and thoroughness.