

Question Type: Product

Duration: 20 Minutes

Difficulty: Medium

Domains: Product

Problem

How would you measure **driver** satisfaction on Lyft?

Solution

How would you measure **driver** satisfaction on Lyft?

[Candidate] To address this question, I want to first clarify assumptions. Can I assume that a survey can be used to measure satisfaction?

[Interviewer] Sure, if you are using a survey, what would be your methodology?

[Candidate] I suppose a questionnaire can be dispatched to the drivers. The survey could consist of questions that ask on a likert scale - 1 being the worst and 5 being the best - *how satisfied are you being a Driver on Lyft?*

[Interviewer] What do you think is the purpose of measuring satisfaction the first place?

[Candidate] I suppose satisfaction serves as a proxy for the behavior that Lyft would not want, which is driver churn. The acquisition cost of a driver is high and the driver must be retained for long enough to break even. Understanding driver satisfaction could help the company intervene, perhaps implement a new promotion or dispatch a driver customer service to check on the driver.

[Interviewer] Very good. How would you monitor satisfaction?

[Candidate] I can think of two possible approaches. The first approach involves sending out a survey based on how long a driver has been on Lyft. Basically, send out the survey at fixed time points - the first week, first month, third month and e.t.c. since sign-up. So, you can average the scores across the lifecycle of the driver. The second approach is to random sample drivers to send surveys on a daily basis. Ultimately, you would get a daily average score of driver satisfaction.

[Interviewer] Okay, that sounds good with me. Now, how would you monitor satisfaction without the survey?

[Candidate] Hmm... that basically means that the satisfaction score of a driver is measured implicitly. One approach could be using the existing survey data as a label set and using the driver log data to predict on the labels. This could be an implicit way to extract driver satisfaction.

[Interviewer] Sounds good. What's the next step?

[Candidate] I would gather the data, perform preprocessing, build a model and evaluate the model. If the model performs good enough in train, valid and test splits then productionize the model to score driver satisfaction in real-time.

[Interviewer] Gotcha, walk me through your choice of data. What kind of data source and variables do you imagine would be helpful for this modeling exercise?

[Candidate] I'm going to presume that I have access to the following data sources - driver logs (i.e. time stamps when the driver provided rides, distance covered, commission, geography), ratings data (i.e. rider reviews of the driver) and profile data (i.e. date when enrolled, vehicle type, date of birth)

[Interviewer] Great. Now, how would you build your model?

[Candidate] So, assuming that I've performed preprocessing, feature engineering and feature selection, I'd start with either a logistic regression or random forest model. It's a relatively simple model to begin with, and models provide interpretability of the variables which is a plus. In terms of the design, I could design a model that generates daily satisfaction scores of a driver based on a moving window features. For instance, given a driver's profile features, past 60-day ratings and past 60-day log data, what's the satisfaction score of a driver today?

[Interviewer] Quick question on that, given that the survey values are multiclass, how would you accommodate for this?

[Candidate] So, I would keep the design, but either use a multi-class version of the logistic regression and random forest or use a set of five binary-class models - each predicting one of the 5 scores in the survey response.

[Interviewer] Makes sense. How would you evaluate your model?

[Candidate] So, I'm a bit hazy on the multi-class metric, but I believe there's a way to evaluate multi-class model using the average precision score. And, of course, for closer inspection on which score is predicted poorly, I'd look at the f1 score or precision/recall on that score.

[Interviewer] Great. Now, one last question. Suppose Lyft doesn't have survey data. How would you have predicted satisfaction then?

[Candidate] Hmm. Great question. As I'm thinking about this, the purpose of measuring satisfaction is to ensure that drivers continue to operate for Lyft. Dissatisfied drivers will leave the platform which is an opportunity loss for Lyft while satisfied drivers will stay on the platform and generate a revenue stream for the business. I believe one approach is to build a churn model that predicts whether a driver is likely to churn. In a sense, churn serves as a proxy for the driver satisfaction.

Interviewer Solution

The best explanation bridges the technical solution to the business problem. Consider the importance of driver satisfaction based on cost. The acquisition cost of new drivers is high considering advertisement and on-boarding expenses.

To recuperate the initial capital spent to recruit new drivers, the rideshare platform must ensure that drivers are partners in the long-term. Considering the importance of longevity, gauging driver satisfaction is vital.

If you start your explanation with this business-sense, then you hooked the interviewer. Now, the recruiter could ask for the technical solution to measure.

A common response could be: "To measure satisfaction, survey on a sample of drivers is a decent approach. The survey contains multiple questions that asks drivers to rate their experience that gauge satisfaction."

That's not a bad approach. But suppose that the interviewer asks another follow-up: "What would be an alternative approach to survey? Suppose you can't identify this explicitly."

In many analytics and A/B testing questions, you need to define a proxy that measures a quality. **Before proceeding further, consider a proxy that could serve as a measurement of "satisfaction."**

One approach is using driver activity data as a proxy for satisfaction. You could hypothesized that drivers satisfied on the platform stay in the long-term while those less satisfied are likely to churn. One stream of data that platforms collect is rides activity.

The underlying assumption is that drivers with positive satisfaction offer more rides while drivers with less satisfaction offer less rides.

You could apply the univariate statistics approach to bin the rides provided by each driver per each month. Any downward trend is a suggestion that the satisfaction is lowered.

This approach is better, but could still use an improvement. "The approach you suggested not a bad proxy. One of the programs, the platform plans to enroll is 'intervention.' The idea is to provide tailored calls or incentives for drivers signaling low satisfaction through the data. Obviously, the survey will only provide satisfaction as a whole, but not at the individual-level. How would you approach this?"

The best approach will combine survey data, proxies, and model to score satisfaction on each driver.

Suppose a survey received a response count of 300K (5% response rate). Use behavioral and driver profile data to as features to train a model such as random forest and predict on the satisfaction score measured in the survey. Assuming that the model performance is decent, you could predict the satisfaction for each driver on a montly level. At any month, the predicted satisfaction score is “low,” trigger an alert to the dispatch team to service the driver.

Interviewer Assessment

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

Assessments	Rating	Comments
Statistical Methodology	5	<p>The interview was designed to test the candidate's statistical skills in surveying and modeling when survey data doesn't exist.</p> <p>The candidate did a phenomenal job in providing sound and accurate responses. The candidate understands the basics of a survey design and possesses a strong foundation in machine learning. She proposed using the existing survey labels to implicitly predict satisfaction of drivers. This is a good design. When asked a curved ball question on how she would handle cases when the satisfaction label is absent, she figured out that a proxy, churn, can be used to gauge whether a driver is satisfied or not.</p>
Product Sense	5	<p>The expectation from this interview question was whether the candidate understands enough about the Lyft driver experience to address a statistical question.</p> <p>The candidate demonstrated a strong product-sense about the Lyft platform on the driver side. She understands why measuring satisfaction is vital for driver retention (i.e. "Understanding driver satisfaction could help the company intervene, perhaps implement a new promotion or dispatch a driver customer service to check on the driver"). She also anticipated relevant variables that are collected from drivers.</p>
Communication	5	<p>This portion assesses whether the candidate can engage in a discussion with comprehension, coherence and clarity.</p> <p>From the gecko, the candidate asked questions to ensure that she understands the constraints and assumptions about the problem. This is a great approach to start to ensure to avoid miscommunication further down the discussion. She also provided answers that were comprehensive and clear.</p>