Machine Learning Canvas

PREDICTIONS

IDEA

Context

Who will use the predictive system / who will be affected by it? Provide some background.

We're a company selling monthly subscriptions to a SaaS product and we want to detect customers at risk of not renewing, based on their profiles and on how they use the product, to take action and make them stay.

SPECS

Problem

Question to predict answers to (in plain English)

Is this customer going to leave us within 1 montl

Input (i.e. question "parameter")

Customer

Possible outputs (i.e. "answers")

"Yes" or "No"

Type of problem (e.g. classification, regression, recommendation...)

Binary classification

Baseline

What is an alternative way of making predictions (e.g. manual rules based on feature values)?

See if time since last time customer connected is above a given threshold (e.g. 15 days)

OBJECTIVES

Value Proposition

What are we trying to do? E.g. spend less time on X, increase Y...

Decrease churn rate and thus increase the total number of customers (and revenue)

Performance evaluation

Domain-specific / bottom-line metrics for monitoring performance in production

- Churn rate
- Return On Investment = [insert formula that takes into account both benefits and costs of the associated customer retention program, by integrating the number of True Positives (TP) and False Positives (FP)]

Prediction accuracy metrics (e.g. MSE if regression; % accuracy, #FP for classification)

- % accuracy
- F-measure
- #FP (treating false positives as churners costs money for nothing — time spent and special offers given) and #FN (not detecting churners is a loss of potential revenue)

Offline performance evaluation method (e.g. cross-validation or simple training/test split)

Take all data until X months ago and evaluate on last X months

DATA

Data Sources

Where do/can we get data from? (internal database, 3rd party API, etc.)

- Internal customer database (transactions: who bought what)
- Customer support tool (e.g. Zendesk)
- Website analytics (e.g. Mixpanel)

Dataset

How do we collect data (inputs and outputs)? How many data points?

Snapshots of customers taken 1 month ago, associated to the fact that they have/haven't churned now. This information is derived from the entries in the transactions table.
There are as many data points as there were customers 1 month ago.

Features

Used to represent inputs and extracted from data sources above. Group by types and mention key features if too many to list all.

- Customer "profile": age, income, education...
- Usage: number of times customer used web app, features used, ...
- Customer support interactions: number of requests, topics, satisfaction ratings, ...
- Other contextual: browser, smartphone, ...

DEPLOYMENT

Using predictions

When do we make predictions and how many?

Every time a new model is created (happens once per month), we separate all current customers in two halves. The first half is the holdout set, and we only make churn predictions for customers in the second half.

What is the time constraint for making those predictions?

6 hours max (so that everything can be done overnight)

How do we use predictions and confidence values?

- We consider customers who were detected at risk of churning and feed them to another model that predicts whether we can do something about it and, if so, which action.
- We prioritize customers by taking into account revenue for each of them and confidence of churn prediction.
- We pass on the prioritized list of customers to the support team, along with the predicted actions, and the explanations given by the descriptive models.

Learning predictive models

When do we create/update models? With which data / how much?

Every month we create a new model from the customers who were in the previous month's holdout set only (so that the data we learn from wasn't affected by actions taken after making predictions — which were on the other set of customers).
The 1st model we create can be from all customers there were a month before.

What is the time constraint for creating a model?

6 hours max

Criteria for deploying model (e.g. minimum performance value — absolute, relative to baseline or to previous model)

- Better than baseline on all metrics by at least ...
- Estimated return on investment is positive and at least ...

Note: We might perform worse than the previous model because churn might get more difficult to spot as conditions change (market, product features, customer behaviour)

Reset Form

Machine Learning Canvas v0.1