OCTOBER 2018

Determining Consumer Intent for Automotive Dealership Visits

Placed.

Overview

In the automotive industry, it is often difficult for companies who pay for advertising to get lower funnel metrics based on consumer behavior in the physical world. As a result, automotive advertisers may be blocked from planning and optimizing campaigns for the shopping behavior they want to drive. By using offline attribution, this hurdle is overcome by advertisers who now leverage offline visitation as a lower-funnel performance metric for analyzing and optimizing automotive campaigns.

As a result of working closely with automotive advertisers on offline attribution, Placed identified a need to categorize visits by visit intent. This white paper details our process to develop and validate a machine learning model that provides visibility into dealership visit intent for use in Placed Attribution reports.

Not All Dealership Visits Have Equal Value

In many industries, using offline visits as a proxy for purchases enables businesses and advertising agencies to understand and optimize their media plans and maximize return on ad spend. For some categories of business, such as automotive, there are visits viewed as more valuable than others. Because consumers visit car dealerships for a variety of reasons —shopping, test drives, service & repair, employment, promotions, etc—automotive is one example where understanding the intent behind a visit could help advertisers understand their campaigns' effects on the most valuable types of visits as defined by the campaign objectives.

Approach to Model Development and Validation

Inferring the intended purpose behind a dealership visit can be posed as a machine learning problem whose solution requires two essential ingredients:

1. A source of ground truth data which can be used as both true labels for the training of the visit intent model as well as for validating the performance of the model on new data.
2. Availability of high fidelity location data which can be used to measure dealership visits and whose richness can support the extraction of features which are necessary to discriminate and predict the variety of visit intents.

Ground Truth Data from First Party Surveys

At Placed, we specialize in tying offline visitation behavior back to omni-channel media delivery. In addition to the location data that we observe from over 300MM devices monthly, for ground truth, we have a 1st party audience with millions of mobile application panelists who both share their persistent location data with us as well as answer surveys.

Solving every classification problem begins with gathering ground truth data. To gather the data, we sent a survey every time a panelist visits a dealership asking about why they visited. The answers were not mutually exclusive, so users have the ability to select as many options as are applicable. The wording of the survey is as follows:

- **Question:**
  Which of the following statements describes your reason for visiting the car dealership? If you accompanied someone to the dealership, please choose the main reason for the visit.

- **Answers:**

  a. Test drive

  b. Browsing, researching, or shopping for a vehicle

  c. Purchasing or leasing a vehicle or to complete related paperwork

  d. Repair, parts, or scheduled maintenance

  e. Other

In Figure A, we plot the frequency with which each option was selected. Because the responses were not mutually exclusive, the total adds up to more than 100%; however, the frequency of surveys with multiple options selected was less than 10%. Thus, although users may select as many responses as applicable, choosing only one response was by far the most common pattern, suggesting that consumers engage in a limited set of behaviors when visiting dealership.

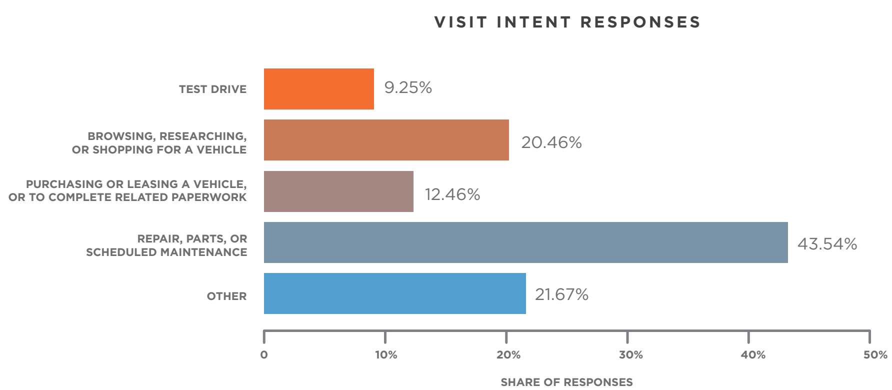

Figure A: Reasons why survey respondents were visiting car dealerships.

Any response in the set *{Browsing, researching, or shopping for a vehicle, Purchasing or leasing a vehicle or to complete related paperwork, Test drive}* automatically qualified a visit as shopping-related. All other responses were placed into a non-shopping category. **Partitioning the responses in this manner provides an estimate that overall, 37.5% of car dealership visits are shopping-related.**

Before we moved forward with using the as-defined shopping categories for the machine learning model, we analyzed which survey visit reasons had overlap. The overlap analysis had two objectives. First, we were able to validate that we are capturing the right kinds of shopping-related behaviors buyers engage in when visiting a dealership. Second, it justifies the earlier partitioning decision for defining a shopping visit.

For example, looking at the *Test Drive* response, we determined the likelihood of also selecting every other option (first row in Figure B). The most likely other reason to be selected was *Browsing, researching, or shopping for a vehicle*, indicating internal consistency in users' answers and that a shopping visit may entail more than one visit intent. Overall, when a user selected one of the shopping-related responses, the most likely other responses to be selected were other reasons defined by us as shopping-related (darker colors in the top left quadrant of the figure below). No natural grouping was observed for the non-shopping category.

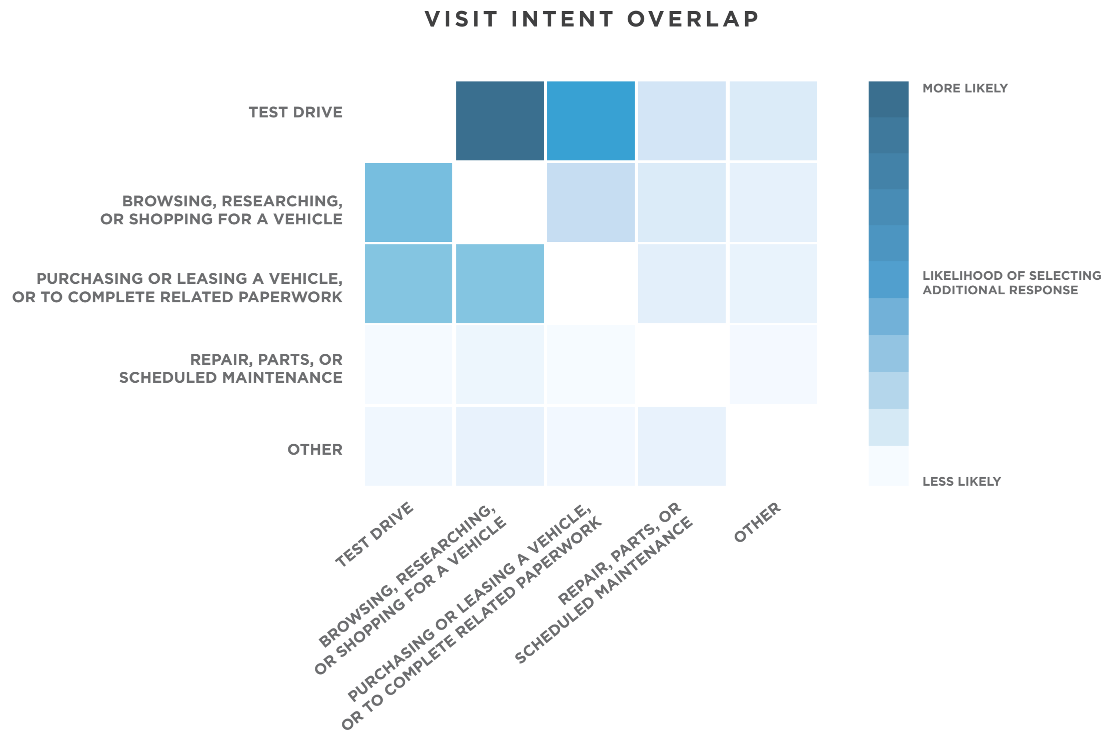

Figure B: Overlap in visit reasons for car dealership visits.

Building the Machine Learning Model

Approach: Model Powered by, and Validated with, High-Fidelity Location Data

Because Placed uses location data from 1st, 2nd, and 3rd party sources (and only panelists of our owned-and-operated apps can be surveyed) we needed to develop an algorithm that would enable us to predict visit intent based on a visit's observed features. In other words, we sought to build a predictive model that would take as an input the features of the dealership visit and output the probability that the visit had a specific visit intent, such as being shopping-related. For this white paper, we will focus on shopping as the high-value visit intent, but the model can be easily extended to predict other types of visit intent, like test drives or parts and service-related visits.

Building an accurate prediction system is a complex process. At the simplest level, machine learning exercises of this kind can be broken down into two main stages:

1. Extract features we believe to be important for discriminating the intent behind a dealership visit.

2. Train a model to make use of these features and make an accurate prediction.

This simplification serves as a useful skeleton around which to frame the discussion. The general idea behind this approach is shown schematically in Figure C below.

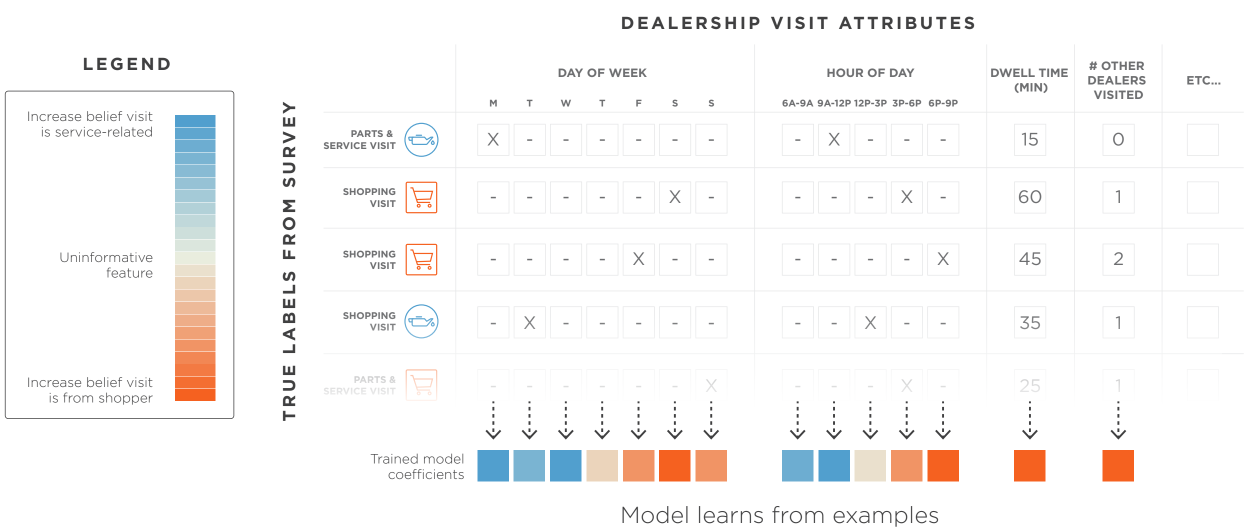

Figure C: Visit-related features are analyzed for their impact on successfully defining a shopping visit.

Identifying and Quantifying the Impact of Features

One challenge in building an accurate predictive model is crafting an informative set of features. No matter how sophisticated the model, if the only feature used to predict visit intent are height and weight of the visitor, accuracy will be limited. At the same time, the relationship between features and model output can help validate the overall approach. That is, it can be instructive to examine whether the features one expects to have significant impact on model output do so. If how the model behaves runs counter to intuition, it can serve as a red flag that the approach needs to be revised. As mentioned earlier, ground truth data is the prerequisite for validation activities, and Placed's first-party survey data played a key role.

At a high level, the features we derived could be grouped into one of three categories:

1. Demographic, geographic, and device information of the visitor
2. Properties of the visit itself, such as the day of week and time of day on which the visit occurred and how long the visit lasted
3. Historical visitation patterns of the visitor, such as the number of distinct and total dealerships visited in the epoch surrounding the current visit and the total time spent at those dealerships

Drawing on intuitive examples from each category, in Figure D, we show how knowing that a feature took on a particular value influences the model's belief that the visit was shopping-related. In all panels below, positive values (above the dashed line) indicate that the presence of that feature increased the probability the visit was shopping-related.

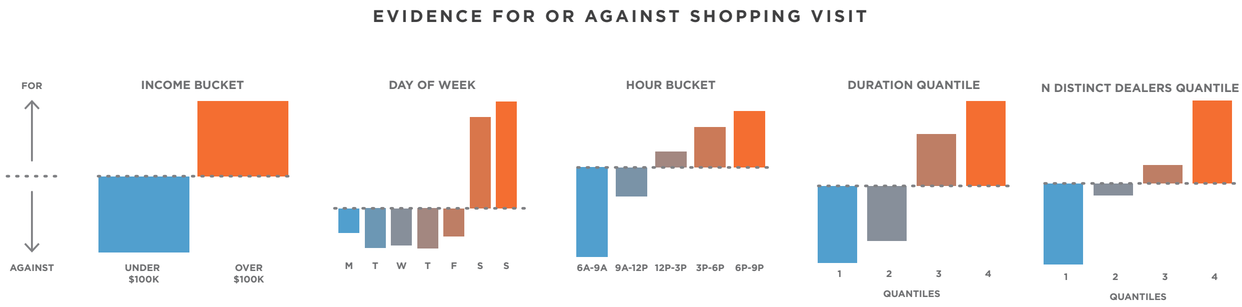

Figure D: A subset of the features evaluated for their effects on the probability a visit was shopping-related.

Figure D illustrates that if a measured visit:

- originates from a high income earner
- happens on the weekend
- is late in the day
- has a longer dwell time
- and is preceded or followed by other dealership visits

—there is a high likelihood that the visit is shopping-related. Returning to our earlier point about correspondence between model behavior and intuition, all of the above observations are consistent with how, a priori, one might expect shopping visits to differ from parts, service and maintenance visits.

Given such intuitive relationships, it could be asked whether we even needed to couple high-fidelity location data with survey responses to derive an accurate model. Our position is that the approach described maximizes the accuracy with which we can discriminate visit intents. First, we cannot underestimate the importance of using the ground truth labels obtained via the survey to teach the model exactly how much to update its predictions when it sees a visit on Saturday at 4PM that took 2 hours. Heuristics would simply not suffice. Second, without high fidelity location data, we would be unable to obtain reliable estimates of crucial quantities such as visit duration and number of other dealerships visited. And last, there are tens of other features included in the model that play a role in improving accuracy that would not only have been very difficult or impossible to intuit without the large ground truth dataset but also depend on accurate location data (e.g., the total area covered by the ambulant visitor while on the lot).

Validation of the Machine Learning Model for Dealership Visit Intent

The real power of a predictive model lies in its ability to accurately predict visit intent for examples that it has not seen before. Indeed, it is almost always the case that a classification model will perform better on its training data than on unseen data. Having a large collection of survey responses offers us an opportunity to assess model performance on unseen data by training the model on a subset of the survey responses and testing it on a held out set. Because the held out set is drawn from the survey responses, we have the true labels for each test example, allowing us to see how the model's accuracy will translate to unseen examples in the future. Figure E shows a schematic of the cross-validation procedure.

Figure E: From the universe of labeled examples, a subset are used to train a model while a held out set are used to validate its performance.

During the validation step, the Placed model performed above chance with 99% confidence, showing the feasibility of inferring visit intent from visit characteristics. One way to assess how well the model performed on the unseen data is to examine the true rate of shopping visits at different values of model output. We expect that with more confident model outputs (i.e., with higher predicted probabilities), the true rate of shopping visits should increase. In Figure F, we see that is indeed the case, where there is a strong and monotonic relationship between predicted probability and the fraction of visits that were in fact shopping-related.

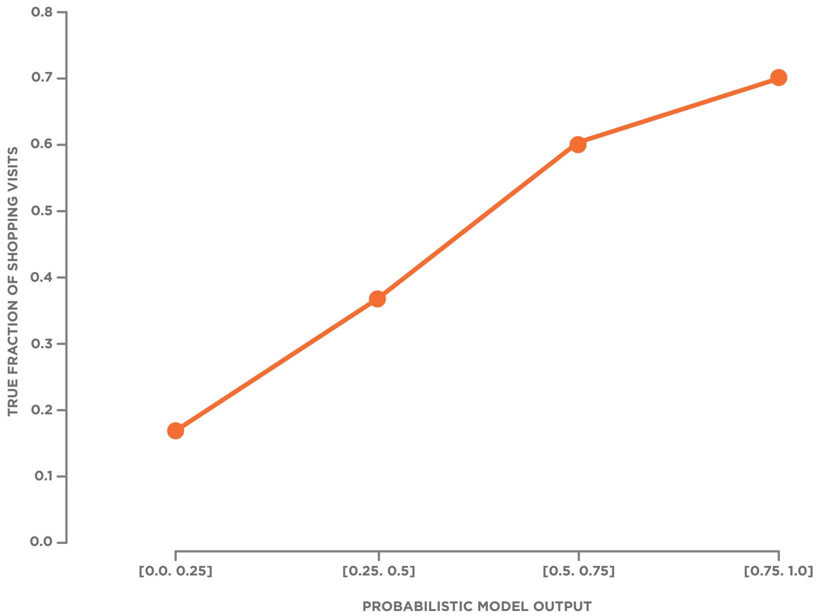

Figure F: Cross-validated rate of shopping visits as function of model output

Leveraging Visit Intent for Offline Attribution

After we built and validated a model to partition dealership visits according to their intent, we started integrating these reclassified visits into Placed Attribution reports. As a result, Placed is able to measure and inform advertisers of the impact of their delivered media on specific types of dealership visits.

The entire approach is flexible enough to be easily extensible to providing attribution reports on test drive visits, where the only change is the target variable the model learns to predict. More generally, the framework opens up new research opportunities.

Research Methodology

The research was based on more than 75,000 completed surveys to Placed's first-party audience. The survey's collection began in June 2017 and continues to be active today. For each completed survey, we joined the survey responses with the visit cluster that led to the survey push and then also with the demographic attributes of the user that generated the visit. This series of joins enabled us to extract all the necessary features to build an accurate predictive model.