ISYE 6501 Course Project

Honda Warranty and Service Automation

Stone Hayden

November 19, 2020

Introduction:

This course project is based on an article from the SAS website (link below on reference page) detailing the successes that Honda America has had automating the warranty claim process and service visit preparation. This project will cover what theoretical models may have been used in automating this process and will not cover the specific software solutions Honda America implemented to achieve better efficiencies.

Case Details:

Honda America receives tens of thousands of service visits each year across it's 1,200 dealerships, and authorized service partners. Many of these service visits will include warranty claims that will make the cost of service/repair the responsibility of Honda America. A majority of these warranty claims are accurate and Honda America is rightly responsible for the cost; but there are certain visits where a warranty claim should be denied, or could be fraudulent. It is in Honda America's best interest to identify these illegitimate claims to prevent the financial and time loss associated with servicing vehicles for free. Each one of Honda America's service visits are detailed through the company's own database and provide a vast amount of data for each individual service visit. Previously, Honda America had to manually inspect each warranty claim with the service data and decide on whether the warranty should be approved or not. With its

new model, Honda America has automated this process and cut down on warranty review time by 67%.

In addition to warranty claims, Honda America has created a model to aid in efficiency for all service visits. Honda America has taken all the data from previous service records to forecast what parts should be kept in stock during various times of the year. Additionally, this model aids in predicting the amount of service technicians working during each shift over various times of the year. Keeping service visits as short as possible helps with keeping staff costs regulated and preventing any additional costs such as storing a customer's car for days at a time. Additionally, poor service can have long term harm on Honda America by damaging customer's faith in the brand and not purchasing their next vehicle from Honda; and persuading friends and family to avoid the brand as well.

Model Objective:

The company's objective is to maximize profit in the immediate and long terms. For these problems, the objective of Honda America's model (and our theoretical model) is to minimize the cost of performing unnecessary warranty service, and to keep the appropriate service parts stocked during the right times of year.

Data Collection:

For data collection, we will split up the data needed into the two different issues:

1) Warranty Selection Data:

   The necessary data to create our Warranty Selection model will come from previous warranty data. Luckily, Honda America already has all this data so there is no cost of collection. Within the previous warranty data, we would want factors such as vehicle

model, vehicle age, model year, past recalls, type of service performed, mileage, title status, condition of vehicle. With all of this data, we would also need the response data on whether a warranty claim was approved or not.

2) Service Parts Data:

The necessary data to create our Service Parts model will come from previous service data. However, unlike warranty claims, Honda America may not have all past service records for each type of vehicle. Honda America would most likely be able to access all service data from it's 1st party owned dealerships, but there are also independent mechanics that would service Honda vehicles without sending service records to Honda America. Could Honda America create a model based solely on it's own data? I wouldn't think so because typically, independent mechanics will be cheaper than a dealership and those with older cars may tend to have their cars serviced with mechanics. Thus, Honda America's 1st party data may skew towards newer cars and not older ones. However, it's also not feasible to get service data from every independent mechanic in the country. So, Honda America may need to request records from random mechanics in different parts of the country to accurately supplement their 1st party data. The service data would need all of the same factors listed earlier in the warranty data; however, the desired response would be wait time.

Theoretical Case Model:

One theoretical way to classify the warranty claims could be to start by clustering past cases. However, before clustering, we may want to split our data to account for cyclical changes (different seasons). Considering that some issues may be prevalent during different times of year. We could use a CUSUM model on the temperature in the area to see our

"unofficial seasons" (similar to previous HW) and which claims occur during which season. Once we have claims coming from similar climate conditions, we start by using k-means clustering on past claims and we will most likely see clusters based on model type (AKA Accord, Pilot, Ridgeline, etc.). Once we have our clusters, we could use an Support Vector Machine (SVM) model within each cluster to classify whether a claim should be approved or rejected. So a new claim would come in, be classified into a dataset of specific climate conditions, be categorized as a certain vehicle based on k-means clustering, and then approved/denied based on an SVM of that vehicle type. While not considered here due to not knowing exact data, it would also be important to weight the impacts of classification errors. If a warranty claim is incorrectly approved, it equates to an unnecessary cost to Honda America. If a warranty claim is wrongly denied, it could have a significant impact to the image of Honda America in that customer's mind and prevent them from buying another Honda.

For the service parts model, we would want to prepare our data in a similar fashion to the warranty model. We would want to know what kind of climate conditions the vehicle broke down in since different parts may break down in different conditions. And then we would want to perform the same k-means clustering to divvy up our data by vehicle model. Now the service parts model differs here from the warranty model as an SVM model wouldn't help here. With our clustered data, we would want to use variable selection, perhaps Elastic Net, to find relevant factors in our service data. Once we have simplified our factors, we would want to create an optimization simulation to find the optimal parts and their quantities to have stocked in order to reduce wait times. We may also need to weight some factors in our

simulation. Perhaps someone with a more expensive repair should be prioritized. Or maybe services should be performed first come, first serve (FIFO).

Model Maintenance:

The Warranty Selection model should be adjusted somewhat regularly. Some warranty claims may occur that are warranted but are denied due to old data. For example, maybe Honda Accords don't typically need their radiators replaced for 8 years. So if a vehicle model is approximately 6 or 7 years old, there won't be a lot of prior radiator replacements and a new data point requiring a radiator replacement may get denied. Therefore it will be important to update our model to account for the aging or vehicles that were once considered newer.

Similar to above, for the Service Parts model, we will need to consider the vehicle aging process. As certain vehicles age, different parts may break. While we may not have a lot of data for 2012 Honda vehicles becoming 8 years old, we do have this data for older Honda vehicles and we can probably assume that parts that failed in 8 years on a 2005 Honda Accord will also fail within the same timeframe on a 2012 Honda Accord.