The GM OnStar Capstone Project developed a personalized product offer strategy for OnStar customers to increase free trial activation rates, the first step in gaining a paying customer.

### Marketing
- Campaign touchpoint frequency, subscription status, method of contact

### Demographics
- Marital status, age, number of children, income bracket

### Application Usage
- Usage of OnStar apps, such as wireless hotspot, remote door lock

### Re-marketing campaigns

- **Which Package?**
  - Premium
  - Essentials
  - Safety & Security
  - Remote Access
  - Connected Vehicle

### Prediction

- **PCA to reduce dimensionality**
- **65% of variance explained**
- **Clustering on principal components**
- **Non-spherical shape detection**

### Business Impact
- Recommendations from the model will be piloted on a cohort of OnStar re-marketing customers in the next six weeks.
- GM Team is currently creating marketing materials informed by features relevant to our model.
- If rolled out to the entire re-marketing cohort, our work has the potential for a $1M revenue increase compared to current practice.

### Segmentation Model – Grouping Together Similar Customers

We implemented two segmentation methods, an unsupervised and a supervised approach. The unsupervised model utilized principal component analysis (PCA) and Gaussian mixture clustering, and the supervised approach implemented decision trees.

### Predictive Model – Predicting True Package Preference

We iterated on several models to determine the predictive approach for this task. Our final iteration was a LightGBM model under an unsegmented scenario. Hyperparameters were tuned using an automated optimization tool (HyperOpt) in Python with 5-fold cross validation. Our final model improved accuracy by 14.8% against baseline and had an AUC of 0.74.

**Metric**

- **Training**
  - Accuracy: 0.5715
  - Bal. Accuracy: 0.3773
  - F1: 0.5371
  - Recall: 0.5715
  - Precision: 0.5786

- **Testing**
  - Accuracy: 0.5331
  - Bal. Accuracy: 0.3446
  - F1: 0.4988
  - Recall: 0.5331
  - Precision: 0.5036

**Key features include:** Remote Start, Get Vehicle Data, Number of Marketing Touchpoints

**Feature usage insights used by GM Marketing for pilot copy**

**SHAPley Additive exPlanations (SHAP)**

- **totaltouchpoints**
- **getvehicledata30day**
- **remotestart30day**
- **hotspot30day**
- **remoteaccess30day**
- **WINFIRSTmax**
- **tmin**
- **CONNDORMax**
- **AMAZONMax

**Out-of-Sample ROC Curve**

- **Cumulative Explained Variance**
- **Clustering on principal components**
- **Non-spherical shape detection**

**Cluster Assignment Using Gaussian Mixture**

- 3-dimensional plot of first 3 principal components with color noting cluster assignment

**Decision Tree**

- Predicting package preference
- 9 trees trained constraining number of leaves from 2 – 10