Course Project

Project topic:

I chose to discuss American Honda Motor Co., Inc. case study after going through lots of case studies. One of the reason I decided to work on this case study is because I have a great passion in cars and I want to be in manufacturing sector if possible "Tesla". I want to work as an Sr. Analyst and I believe that this case study will help me with my understanding from an analytics point of view and a thought this would be a good place to start.

Up to my understanding the case study discusses how Analytics helped the company to reduce time taken to identify suspicious noncompliant claim, and at the same time also increase the accuracy of finding suspicious claim. This issue had led to decrease in labor cost over all the 1200 dealerships nationwide.

Thus, the goal of this case study is to forecast the number of vehicles in operation to predict the number of customers coming to the dealership. This would result in translating to the optimum number of parts needed on hand to meet customer needs.

Solution:

Given:

- Vehicle type and Model
- City, state, and zip code of dealership
- Service Date
- Description of the problem
- Type of Service issue

To:

- Forecast the number of customers coming to a dealership during a time.
- Classify and cluster the type of service repair and thus the component serviced.
- Predict the type of issue a dealership might face in each month.
- Create cluster of issue types by region in different times of a year.

Using:

I would use **SVM(Support Vector Machine) or KNN(K Nearest Neighbor)** to classify service request into different types. Once the service request is classified into let's say Engine or transmission or other class issue, a second classification using SVM will be used to further identify which components failed or were serviced during what period of the year, and at which location. **k-means clustering** could also be used to find clusters of failed/serviced components after the first Classification.

We could achieve the similar result when we apply **CUSUM(Cumulative Sum)** approach on the results of the first classification model. The change detection on the daily service request data is going to show which component is going to need stocked more compared to other components in each season.

In order to predict the number of service requests and the number of customers in a given time of the year, time Series forecasting model like **ARIMA** or **Double Exponential Smoothing** can handle both the trend and seasonality to obtain a very good result.

Given:

- Number of employees available in the location.
- Description of the problem
- City, state, and zip code of dealership
- Type of Service issue
- Time taken for service
- Service Date

To:

Create an optimization model and use simulation to find the optimum number of employees and machines needed without wasting any resources.

Using:

Stochastic Optimization model can be created to find a solution for optimum number of labor and machines that is needed to minimize a service time.

Objective function is used to minimize the service time. Some of the decision variables are number of employees, number of machines, number of service requests. Some constraints would be minimum of "n" number of employees working, number of cars in each bay, maximum number of employees working on each car in each bay, and distance to the nearest node in the community.

Based on an **Exponential distribution** of time between vehicles entering each bay, simulation could determine the optimum number of labor and machines required.

SUMMARY:

Consider we are provided with the historical service data and the geography of the dealership location, using **SVM** the type of service request by engine/drivetrain or other issues can be classified. It will be an excellent starting point to create different clusters based on service requests with respect to region and period of the year. **k-means clustering** could be used to find the clusters of serviced components. Each component that's present will have factors like date and region. **Change detection algorithm using CUSUM** will help to detect which type of component is most likely to be serviced in a given time of the year. This step is important because this information will be used to rank the issues based on severity of the service requests. The output, severity, of the issue is an indicator of repeated occurrence of a service issue. This will help save money for Honda on warranty claims. After creating classification models and clustering issues based on the component of the car, the possible next step would be to predict the type of issues the dealership would encounter in the coming days. Time Series forecasting models like **ARIMA** or **Double Exponential Smoothing** handles both trend and seasonality, that will help to forecast the type of service request the dealership can expect, and to also predict the number of customers that would bring the vehicles to service in a specific period of the year.

The next step would be to establish an optimum number of employees and machines needed to minimize service times and in turn maximize or increase customer satisfaction with a finite amount of floor space available to the dealership. **Stochastic Optimization** is used to minimize service times, constrained by the floor space, labor hours, availability of components in a community. Combined with manufacturing principles, a simulation model is run to determine the optimum number of labor and machines required based on an **Exponential distribution** of time between vehicles entering each bay.