Empowering Sales Management with Potential Detection And Conversion Analysis

MBa team: Hazel Sima, Peijun Xu | MIT Faculty: Dimitris Bertsimas
Thermo Fisher Scientific: Benjamin Preston, Fabio Piccione, David Giela, Tuan Ngo, Varun Shikumar

Project Overview

Division Level: Transplant Diagnostic Division
Project I: Potential Detection
Objective: Identify potential sales quantity for two major products L & S by state
Goal: Quantify the sales opportunities to fine tune the sales structure and synergize sales forces in regions with high potential

Division Level: Transplant Diagnostic Division
Project II: Conversion Analysis
Objective: Segment Customers into sales buckets based on conversion behaviors and provide customized marketing strategy recommendations
Goal: Generate product profile for each customer and put them into predictive buckets, suggesting future actions to research on how customers react to product updates

Group Level: Special Diagnostic Group
Project III: Sales Forecast
Objective: Build Top-down Forecast model pipeline to predict future sales value on a monthly basis for 27 product groups
Goal: Provide accurate future revenue prediction and credibility to change sales actions with appropriate attention on accelerated markets

Sales Team Management

Leverage three perspectives to generate confidence levels for prediction and synergizes sales forces in regions with high potential

Challenges

- The quarterly and monthly transplant data from UNOS public database
- The death rate and retransmission rate for each transplant center in order to assess potential and optimize diagnostic portfolio
- Number of donors in each medical center
- Competitor’s sales revenue and quantity information
- Current companies’ Salesforce distribution
- Specific features of customers: operations or researches that they purchased the old products
- Number of donors in each medical center
- Limited Timespan
- The customer names from the public database and the company’s transactional dataset is different

COVID Effect

COVID has a strong impact on the revenue and product sales portfolio on mid-2020 and therefore the sales information on 2020 is not an accurate historic indication for future product potential

Solution

- Eliminate the entire 2020’s data and assume that 2021 happens right after 2019.
- Imputation
  - Impute using SARIMA time series prediction
  - Impute using the monthly 1 on 1 ratio with the month in first quarter: assume that the first quarter 2020 is normal
  - Impute using the monthly percentage of each month

Data Quality

- More than 90% of 0 value for some individual products sales revenue and some even include negative values
- Observation Match
- Due to product updates and new product release, we can only use sales revenue from 2018
- Solution: Harden the parameter tuning range based on the seasonality pattern we observed in the aggregated level

Reproducible Tools

- End-to-end machine learning pipeline that connects company’s ETL database and generates sales forecasting for business intelligence dashboard with practical imputation for period affected by Covid-19

Data

- UNOS Transplant Data From 1988 to 2021
- Transplant cases performed yearly in each state
- Transplant waitlist added yearly in each state
- Company Transactional Data
  - Quarterly revenue collected from each center from 2010 to 2021

Modeling

- We applied linear regression models to predict three quantities shown below. Sales prediction is the main indicator that quantifies accurate historic indication for future product potential

Results

- We identified four groups of customers with distinct conversion behaviors
  - 11% Fast Adopters (Orange)
  - 14.5% Adopters (Red)
  - 58.8% Lazy Adopters (Blue)
  - 17.1% Churned Customer (Green)

- We also provide categorized buckets for the customers who only purchased the old products

Business Impacts

- Detected estimated over 0.7M dollars of potential sales opportunities across 28 states in the U.S. for two core products
- Discovered and defined 4 distinct conversion behaviors and estimated the average conversion time of 3 years without promotion
- Estimated a customer churn rate of 17.1% and a conversion rate of 25.3% without promotion
- Improved the accuracy from the baseline prediction by 30% to help the management team better allocate salesforce

Synergize salesforce to the accelerated markets detected with abundant potential sales opportunities and high confidence

- Track current old-product-only customers and flag the customers once they react to product updates
- Update the resource data for the sales forecasting model frequently to obtain accurate historic indication for future product potential

Cases Potential

- Identified MidWest, Northeast and Southwest as the top 3 regions with a lot of potential and high confidence

Sales Forecast

- Monthly average revenue data for 103 products from 2018/1-2021/5
  - We grouped non-core products by divisions to generate 27 product groups
  - We imputed the second and third quarters of 2020 to reduce covid effects

Modeling

- We used time series predictive model, SARIMA, for sales forecasting
  - 1. Able to incorporate seasonality
  - 2. Interpretable parameters and results so that it’s easy to incorporate with the domain knowledge from the sales team

Advantages of SARIMA model:

- Interpretable parameters and results

Results

- We improved the accuracy from the baseline prediction by 20%