Meghnath Reddy Challa

Portfolio of Projects - Data Scientist

Customer Defection – Auto New Business Churn Model

- ***Business Problem -*** High Churn Rate within first 90 days.
- ***Phase 1*** – Work with call representative team to retain the customers who have high probability to churn .
- ***Deliverables***
  - A spreadsheet with list of policies and necessary data for reps to use to make a call.
  - Looker Studio Dashboard to monitor the progress of the experiment.
- ***Training set*** - Historical data from 2018-2020 which includes Policy, Driver & Vehicle related features.
- ***Evaluation*** - An XGBoost Classification Model with AUC of 0.71.
- ***Model deployment*** - Using Vertex AI, with bi-weekly batch runs.
- ***Results*** - “Test Group” had a lower churn rate of 3.14% compared to the “Control Group" of 4.61%, saving about ~$500K in lifetime premium.

Phase 2 – High Value/High Risk Customers

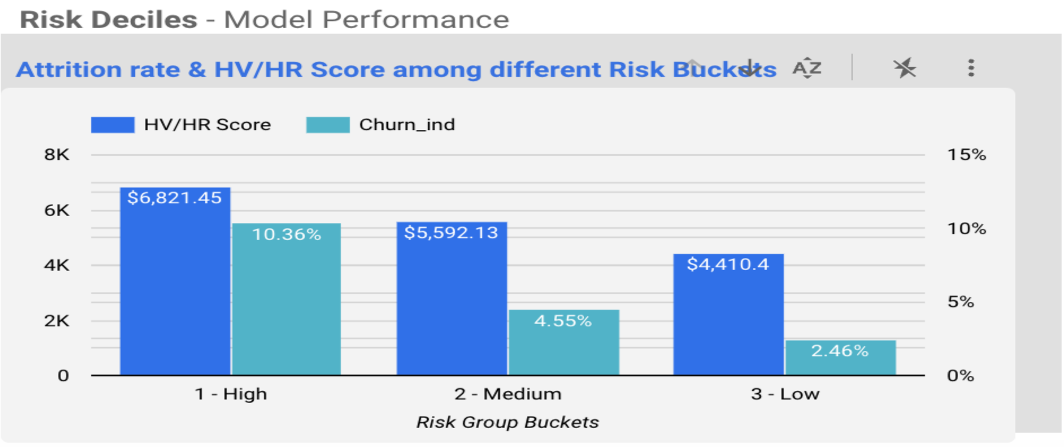
Problem

- How to avoid Unprofitable business ?
  - ❖ Solution - Use Customer LifeTime Value Model
- How to reduce operational cost?
  - ❖ Solution – Introduce Email experiment

Results

By targeting, only the top 20% risk group policies, we saved around $250K and cut operational costs by 80% compared to the first phase.

Model Performance



KPI Forecasting – Quotes, Sales, Responses, WP

- ***Business Problem*** **–** Workforce and Marketing team needed daily level forecasts to allocate resources more effectively compared to previously used monthly forecasts.
- ***Deliverables***
  - A spreadsheet with yearly forecast at a daily level.
  - Collaborate with the Build team in order to develop a Power BI dashboard that allows the business to monitor and compare the Planned and Actual forecast data.
- ***Training set*** - Historical data ranging from 2019 to 2021, which includes actual KPIs, lag values as well as other independent features such as "day_of_week," "month," "is_public_holiday" etc., to aid identifying patterns and trends over time.
- ***Evaluation*** – Models such as XGBoost Regressor, Facebook Prophet, and LSTM Model are used. However, the LSTM Model outperformed the others in relation to MAPE.
- ***Final* Adjusted Model Forecasts** – To match with the monthly planned values.
- ***Results*** – The overall trend and seasonality were accurately captured, although unusual instances of historical spikes were also factored into the forecasting process. For instance, if there was a spike in numbers on Jan 15, 2021, as a result of a special campaign, the model might incorrectly predict a similar trend for the following year.
- ***Next Release*** **–** Started incorporating campaign data to track and analyze spikes more effectively, resulting in improved forecasting.

Other Projects

- ***Business Problem***

How to identify & prioritize Home policies that are subject to external inspection and are at high risk of non-renewal?

Solution : Used SHAP values and identified features like “Age of Home”, “exposed wiring”, “roof rotting”, “siding missing”, “roof type” had a significant impact on non-renewal. Based on these results, recommendations were provided to the underwriting team to assist them in prioritizing the list of policies that required immediate attention.

- Analyzed the frequency and severity of claims within different QI (Quality Improvement) score groups.
- Part of team that was responsible for creation of dataset for Acxiom Daily Refresh to support BI Team.
- Built a dashboard to display the customer base according to census tracts.

Below are some In-flight Projects

- How can we determine whether the potential losses for a given policy in its upcoming term will exceed the pre-defined threshold limit?

  Solution : Built a Loss Ratio model by using policy term level & claims information. However, model still requires refinement and tuning to enhance the evaluation metrics.

- Part of team which focuses on modifying queries that were originally designed for on-premises systems so that they could now reference the point-in-time tables in GCP.
- Think from the perspective of an Underwriter and identify potential factors that may result in an auto policy being flagged for non-renewal. Develop a machine learning model to predict non-renewal.