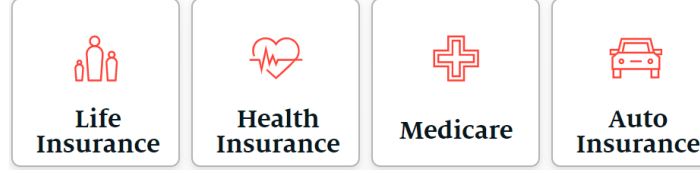


Assurance is a consumer platform that helps shoppers buy insurance and financial products. Their online experience is centralized around the use of data science to help shoppers find the right product.



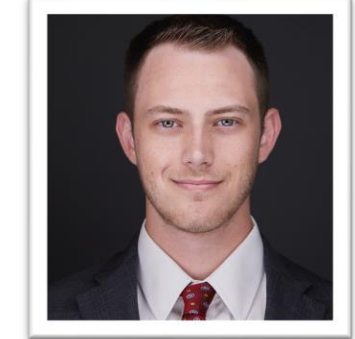
Problem Statement

Will a shopper be Approved or Declined for a life insurance product?

- Predicting the approval status of a shopper is critical in recommending them the correct insurance policy
- If the wrong policy is recommended and a shopper is declined, their status will carry with them and likely prevent them from being able to get life insurance anywhere they go



Kevin Lin, MIT MBAn 2020



Joshua McKenney, MIT MBAn 2020

Allison Arzeno, Assurance Chief Data Scientist
 Nick Howard, Assurance VP Data Science
 Megan Dixon, Assurance Principal Data Scientist
 Professor Dimitris Bertsimas, Project Advisor

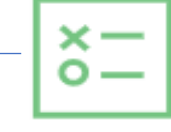
Shopper Profile

- Mostly 50-70 years old
- Lower to middle income
- Not as healthy

Scope

- Looked at policy applications from 2018-current
- Data from more than 100k submissions

Data



Questionnaire

Dynamic questions asked to the shopper (ex tobacco usage, medical conditions, age)



Session Info

Marketing channel, number of visits, etc.



Census Info

Avg income, age/gender breakdown, etc.



Third Party Data

Financial and phone data

Life Policy Types

TERM

Only covers a term period

Simplified TERM (SITERM)

Only covers a term period (with instant decision, easier to qualify)

Whole Life (WL)

Lasts whole life, covers large amounts

Simplified Final Expense (SIFE)

Lasts whole life, covers small amount

Modeling

- We built separate models for each life insurance product to predict probabilities of Instant Approval and Decline
- Logic was built and implemented using the models to to:
 1. Skim off Instant Approvals (recommend the product if we think the shopper is likely to be Instantly Approved)
 2. Filter our Instant Declines (do not recommend a product we think the shopper will be Instantly Declined for)



Model Types

XGBoost
 Logistic Regression



Model Feature Reduction

Feature importance from tree-based models and lasso with cross validation

30k daily shoppers



Individual Data
 Questionnaire
 Session Info
 Census Info
 Third Party Data



TERM + WL Approval Models

only TERM approve

only WL approve

both approve

both no

Recommend TERM policy

Recommend WL

Recommend highest net benefit between TERM and WL

only TERM decline

only WL decline

both decline

no criteria met

Recommend next highest net benefit permanent policy

Recommend next highest net benefit fixed-duration policy

Recommend SIFE/FE

Do not change recommendation

Permanent Policy
 Last entire life – SIFE & WL
Fixed Duration Policy
 Lasts a fixed duration – TERM & SITERM
Highest Net Benefit
 Determined from existing Assurance models



Model Features

Individual Data
 Engineered Features



Model Users

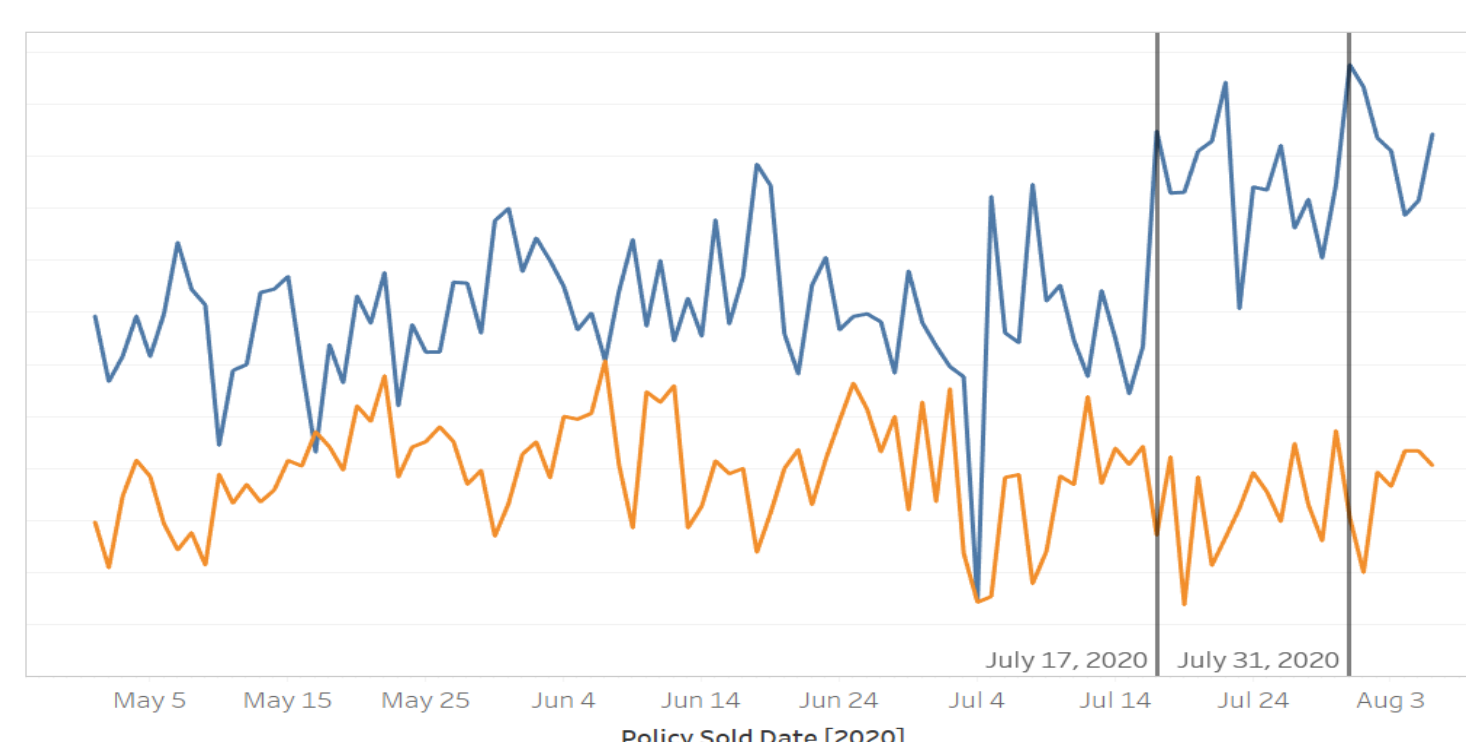
Life insurance shoppers
 30k per day
 From all walks of life

Business Impact

Each of our models is currently making **30k predictions a day** and collectively recommending a different product **10% of the time**. Instant Approval and Instant Decline rates can be seen in the chart below, the black lines denote v1 and v2 of our recommendation changes

Results since July 17th

- instant approvals **↑ 45%**
- instant declines **↓ 10%**
- revenue per shopper **↑ 30%**



Implementation

Pipelined models and data transformations on AWS server for Engineering team to implement



Monitoring Performance

Self updating dashboard tracking daily performance and probability distributions

Other initiatives were going on during the time of the logic implementation, we do not take full credit for the performance increase but believe our project contributed to the growing success of Assurance