co-operators

Investing in your future. Together.

McMaster Workshop 2024

January 2024

Introduction

XYZ insurance is interested in improving the pricing of their private passenger automobile product. Currently, for every vehicle, they estimate the frequency of loss, and multiply this by the average severity of all claims, to arrive at a pure premium for each vehicle.

$$\boldsymbol{Pure\ Premium} = \underbrace{\boldsymbol{E[Number\ of\ Claims]}}_{\textbf{Frequency}} * \underbrace{\boldsymbol{E[Cost\ of\ Claim \mid There\ was\ a\ Claim]}}_{\textbf{Severity}}$$

Then they load this for expenses and a profit provision to arrive at an indicated rate, or premium:

$$\boldsymbol{Indicated\ Average\ Rate} = \frac{\boldsymbol{Pure\ Premium\ (including LAE) + Fixed\ UW\ Expense\ Per\ Exposure}}{\boldsymbol{1.0 - Variable\ Expense\%} - \boldsymbol{Target\ UW\ Profit\%}}$$

The Chief Actuary has stated that the easiest way to improve their pricing is to update the severity model. Currently, the estimate is the same for everyone. This initiative aims to use modeling to give each vehicle a different severity estimate. Frequency, expenses, and target UW profit don't change.

Furthermore, after finding the new premium, management is expected to present to leadership and inform them on the expected dislocation, or the change in premiums, that policyholders can expect.

The Data

All the data can be found ___PUT LOCATION___

Data Dictionary – Definitions for the fields in the data.

Claims_years_1_to_3.csv – Dataset to create new severity model using this dataset.

Submission_Data.csv – Dataset to predict on using new severity model.

Dislocation_dataset.csv – Dataset to calculate premium & dislocation using new severity model.

Qualification Round

Your task is to build a severity model using **Claims_years_1_to_3**. The model will be evaluated using **Submission_Data** – the actual values are withheld. Your submission file must include 2 fields: id_policy and your predictions, labelled "Prediction". The RMSE will be used to measure accuracy of your predictions against actual values. Top teams will qualify through to the final round on March 9th, 2024 – held in person at McMaster University!

The Starter Kit provides a skeleton for participants to get started. You are encouraged to build from it. Using other tools is also permitted and encouraged.

Participants may find it useful to read section 2.7.1 and 4.3 of the following article: GLM Monograph

We invite you to explore any technique or model when it comes to modelling severity.