Stat 348: Allstate Claims Severity

Nathan Chantry

Problem Background

Sponsored by Allstate (are you in good hands?)

Predict cost of claims

Evaluated on MAE (mean absolute error)

131 Categories, 188318/125546 Observations (Train/Test)

Description of Feature Engineering

```
my_recipe <- recipe(loss ~ ., data = train) |>
  step_rm(id) |>
  step_other(all_nominal_predictors(), threshold = .001) |>
  step_lencode_glm(all_nominal_predictors(), outcome = vars(loss)) |>
  step_corr(all_numeric_predictors(), threshold = 0.6) |>
  step_normalize(all_numeric_predictors())|>
  step_zv(all_predictors())
```

Model Comparison

Boosted Trees: 1188.76692

BART: 1214.47657

Random Forest: 3018.06585

Details of Best Model/Final Score

Trees: 1000

Tree Depth: 8

Learn Rate: 0.1