

Suicide Squad's Machine Learning Project

How we arrived at a score of 0.11599 and ranked top 8.7% in the public leaderboard

The Team

Members

Ansel Santos

Sal Lascano

Moon Kang

Yicong Xu

Strengths

- We have a strong coder in the team
- Members have background in project management and in working with big teams
- The team dynamics is strong, with the members sharing the same vision on how the project should be done

Team Strategy

- To concentrate in delivering the project requirements
- Improve on what we were able to build by learning more about machine learning code and functions

Summary of the Machine Learning Process

Exploratory Data Analysis

Getting to know the dataset

We did a scatter plot of the target variable and features which have a high correlation to the target. We also did a correlation heatmap.

Cleaning and Transforming

Preparing the data

- Dealing with missingness
- Transforming the features






Modeling, Grid Search and CV

Balancing between accuracy and variance

We identified the models which produced the best score and then did grid search and stacking to further improve our score

Result

Score of 0.11610 ranked in the
top 8.7% in the public
leaderboard

368	▼ 44	Liuxinran		0.11598	50	11h
369	new	xudongmei		0.11598	2	1d
370	new	Suicide Squad		0.11599	41	5m
Your Best Entry ⬆ Your submission scored 0.11599, which is not an improvement of your best score. Keep trying!						
371	▼ 46	Sharath Kothapalli		0.11600	4	2mo
372	▼ 46	Charu		0.11610	10	1mo

Other models we tried

Why we ended up with Stacking
Lasso and Gradient Boosting
despite trying out other models

- Ridge Regression
 - XGBoost
 - Light GBM
 - Random Forest
 - Multiple Linear Regression
-

Closing Notes

Wrapping Things Up

Being humble and keeping our feet grounded worked well for the team. This gave the team room for each other's ideas, which when combined created a better result.

Closing Notes

Wrapping Things Up

We'd also like to note that keeping the team's bond strong was very important. Having beer together after a long day's work made the team stronger. This made us move forward from problems without clashing with each other, which was a big factor in making us deliver the requirements effectively.
