4th Place - Stacking Ensemble using XGB only

Thanks Kaggle for this fun playground competition! As my first playground competition, this experience has been incredibly valuable. I've learned numerous ML techniques from the generous Kaggle community members who shared their approaches.

Final Submissions

- **Best CV (0.3842 CV score)**: 50-model ensemble using XGB

- **Best LB**: (0.38384 public score and 0.38454 private score)

Ensemble Architecture

The winning approach was an ensemble with XGBoost only and it contained the following:

1. TabTransformer Models (3 total)

- 2 TabTransformers with feature engineering

- 1 TabTransformer without feature engineering
  Reference: TabTransformer

2. Gradient Boosting Models

- **YDF**: 2 models (no FE)

- **CatBoost**: 3 models (2 with FE, 1 with feature selection)

- **LightGBM**: ~10 models including:

  o Standard LGBM with/without FE

  o LGBM GOSS

  o Different feature combinations

- **XGBoost**: Multiple configurations:

  o Original data only

  o Augmented data

  o Repeated KFold

  o Diverse feature sets

3. Neural Networks (6 total)

- Various architectures with different feature engineering
  Reference: [NN](#)

4. Other Models

- 2 AutoGluon models with FE

- 2 Random Forests (no FE)

- Linear models:

  - 1 Linear Regression

  - 1 SGD model
    Reference: [Linear Regression](#)