ML Model Guide

5 Steps to a Great Model

01

Pre-Modeling
- Import libraries
- Build preprocessor

02

Bare Minimum
- Instantiate
- Put into a pipeline with preprocessor
- Fit on training data
- Evaluate with appropriate metric

03

Choose Best Metric
- Regression metrics: R2, MAE (MAPE), MSE (RMSE)
- Classification metrics: Recall, Precision, Accuracy, F1, confusion matrices, ROC AUC

04

Balance Classes (Classifiers Only)
- class_weights = 'balanced'
- Undersampling
- Oversampling
- SMOTE

05

Tune Hyperparameters
- Choose 1-2 most important params
- Use GridSearchCV (cross-validates)
- Pass GridSearchCV the appropriate scoring strategy