

Finding lowest RMSE in a Regression

SUBMITTED BY

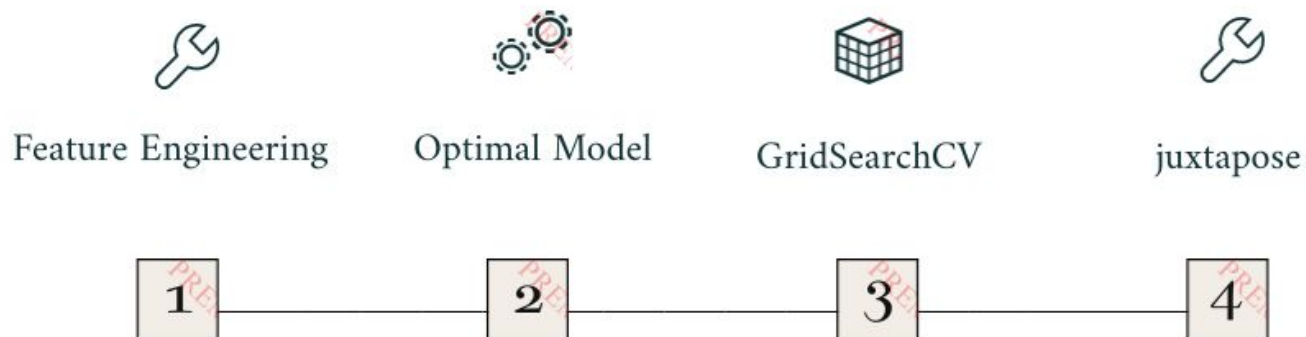
**PRAMOSH
SHRESTHA**

**3rd of July
2021**

Submitted to LeapFrog Tech.

■ PROCESS

Pipeline for the Regression



- The best approach would be to build a Pipeline.
- It's provided by sklearn
- It's a linear process.

Feature Engineering

1

- Check the distribution of data
- Check for null and impute
- Perform Scaling if necessary
- PCA is not necessary for two feature sets.

2

Model

- Time to pick a model
- Neural Networks, Decision Trees, SVR, etc are viable options.
- Choose SVR because it's fast and easily regularised.

3

GridSearchCV

- Research about the model's hyperparameters
- Perform GridSearchCV
- Observer the best parameters
- Select the best model

4

Juxtapose

- Fit and Predict the engineered DataSet on the best estimator
- The RMSE **1.43**
- Fit and Predict on the unengineered DataSet on the unoptimized model
- The RMSE is **3.70**



**The most precious gift we can offer anyone is our attention.
When mindfulness embraces those we love, they will bloom
like flowers**



Thich Nhat Hanh

Thank you for your attention