

Job-a-thon Report

A brief on the approach, which I have used to solve the problem.

Step-1. Visualizing the data to understand the pattern.

Step-2. Data cleaning-Imputation with standard deviation.

Step-3. Outlier Treatment- I have treated the outlier with median.

Step-4. Data Preparation-

- Imbalance Treatment
- Standardization

Step-5. Training Machine Learning models without performing hyperparameter tuning

Step-6. Training Machine Learning models without performing hyperparameter tuning

What data-pre-processing / feature engineering ideas really worked? How did you discover them?

I followed an under-sampling technique and I have taken this technique after implementing various techniques. Of all it worked well. And next to it I have standardized the data using standardisation method.

What does your final model look like? How did you reach it?

After implementing all the models without hyperparameter tuning and with hyperparameter tuning. Finally, I got an accuracy of 57.4 percent with Random Forest and 57.8 percent with XGboost.