

Writing Scientific Papers and Software

Bharath Narayanan, Sandra Siby, Ouiame Talbi

Abstract—

I. INITIAL DATA ANALYSIS

- Plots with Seaborn (correlation)
- PRI vs DER features
- Split by jet number
- Histograms
- DER_mass feature
- Outlier analysis
- Some stats on outliers per column?

II. DATA PRE-PROCESSING

- Split by jet number
- Standardisation
- Outlier removal

III. CLASSIFICATION MODELS

- Why we choose logistic regression
- Some description of our final method
- Cross validation

IV. RESULTS

- Logistic regression vs linear regression
- Logistic regression - PRI and DER vs both
- Logistic regression - all data vs split data by jet number
- Improvements (der mass results?)
- Cross validation results
- Changes with hyperparameters
- Final method results

V. SUMMARY