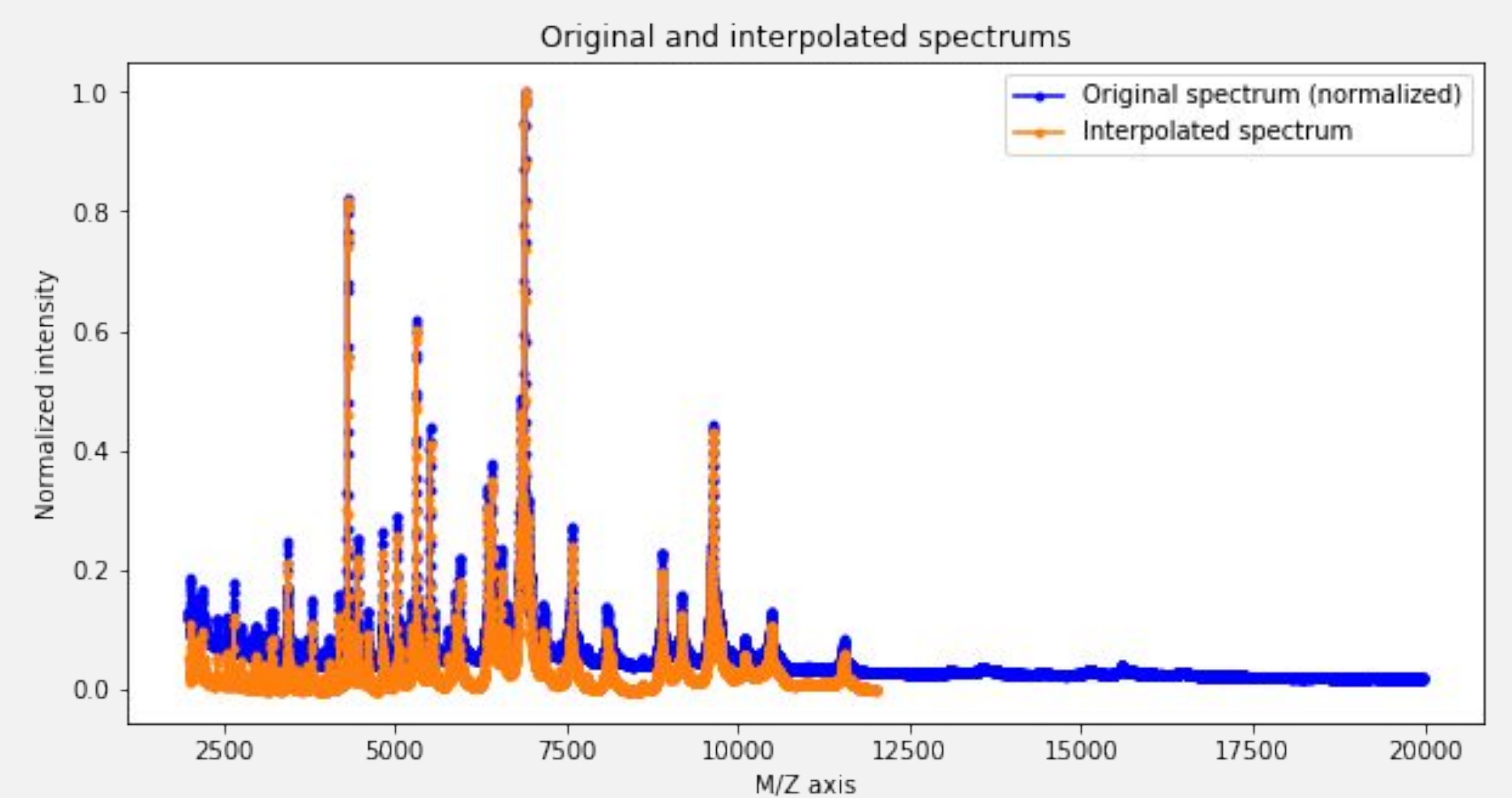


Kaggle challenge

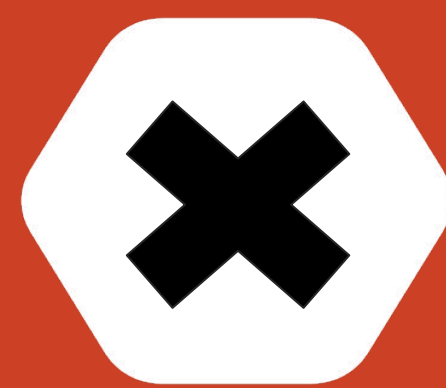
S. Aureus antibiotic resistance

PREPROCESSING

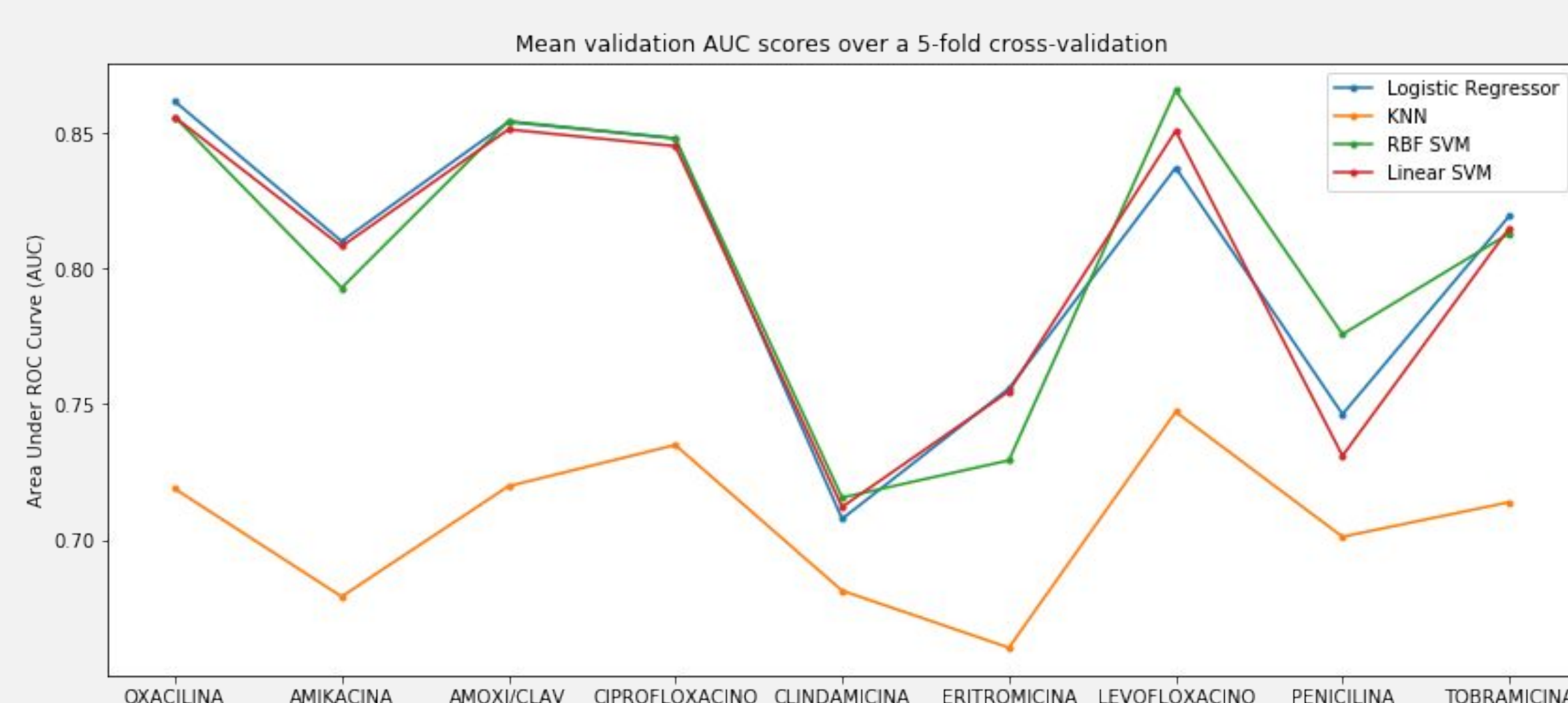
Remove noisy, low variance samples
Interpolation to standard MZ axis
Data pruning above 12.000 MZ
Baseline removal to avoid anomalies
Normalization by largest peak



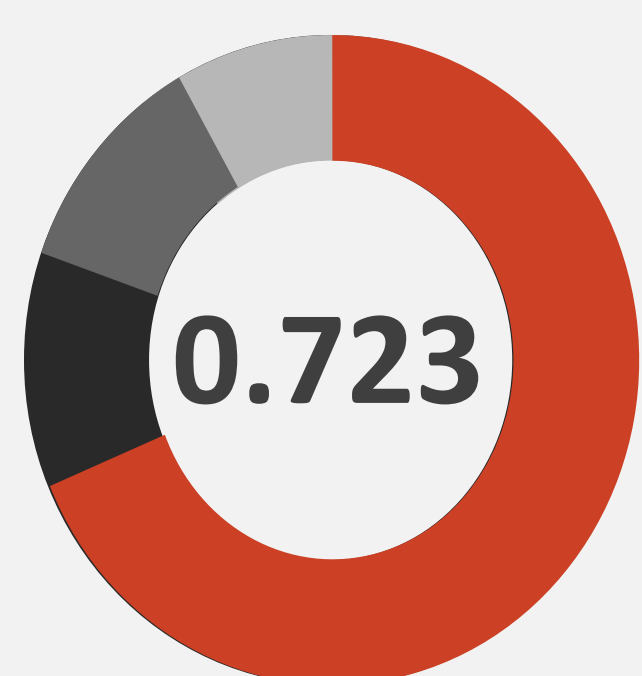
Ada Boost
Classification Trees
Linear/Poly Regression



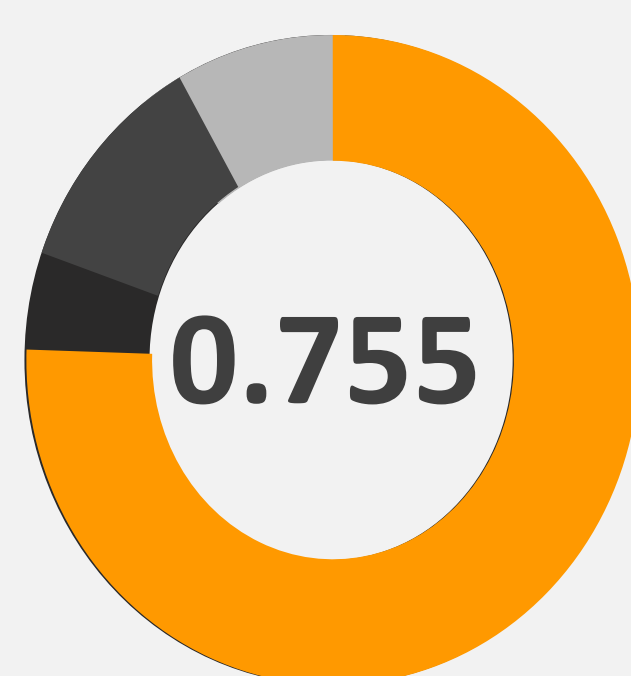
Low MZ magnification
Gaussian Mixture Model
Threshold peak detection



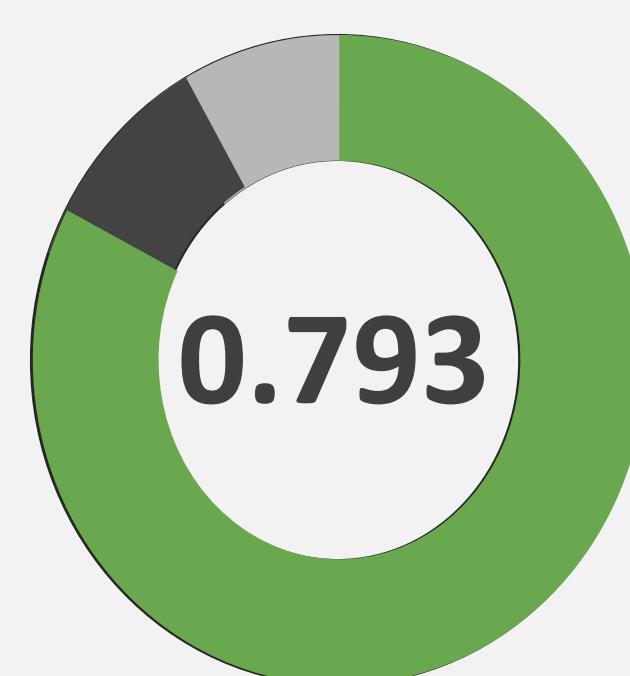
Neural Network (MLP)
OneVsRest Classifier
KNN Classifier
Gaussian Process



Gaussian Process
with peak detection



Logistic Regression



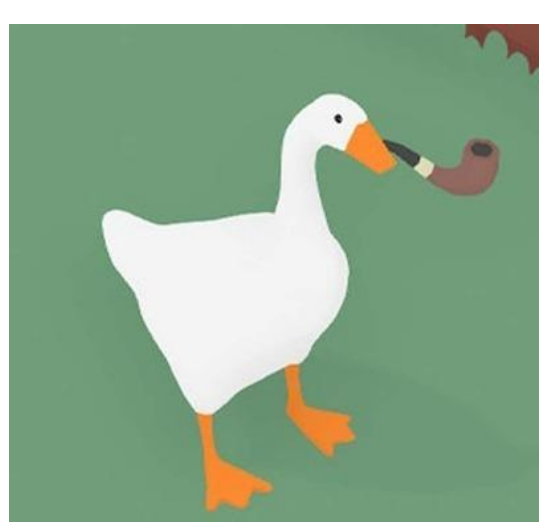
Linear SVC

KISS: THE LINEAR SVC



KISS: Keep It Short and Simple
No further data processing
SVC: Winner of the Cross-Validation Wars
C value = 1 with a Linear Kernel

Plechawska-Wojcik, M. (2012). A Comprehensive Analysis of MALDI-TOF Spectrometry Data. *Medical Informatics*, 69.



Jesús Herrera López
100330730



Javier García Ordóñez
100345713

