

Data

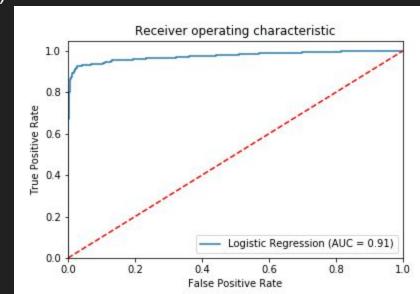
University of Manchester

- 1. Mean of the integrated profile
- 2. Excess kurtosis of the integrated profile
- 3. Standard deviation of the integrated profil
- 4. Skewness of the integrated profile
- 5. Mean of the DM-SNR curve
- 6. Excess kurtosis of the DM-SNR curve
- 7. Standard deviation of the DM-SNR curve
- 8. Skewness of the DM-SNR curve

17,900 data points

Exploration

- Data cleaned, normalized
- Skewed target variable
 - Tried oversampling, undersampling, and SMOTE
- Created base model (logistic regression)



Results

Model (grid search)	AUC
SVM (SGD)	0.969
SVM (LinearSVC)	0.964
Logistic Regression	0.960
Random Forest	0.957
KNN	0.953
ADABoost	0.951
Decision Tree	0.941

Recommendations

