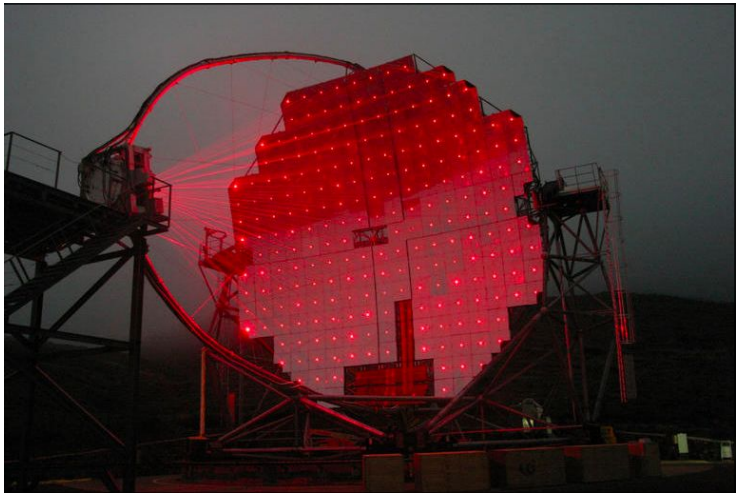
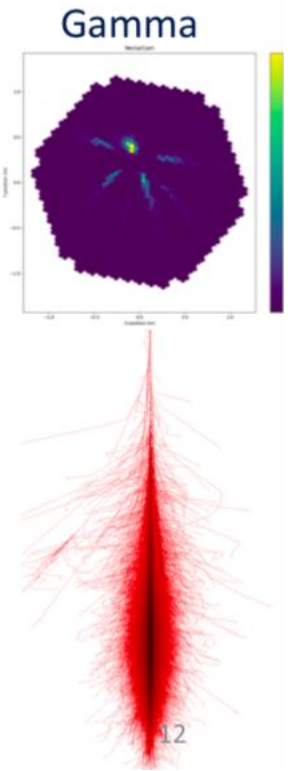
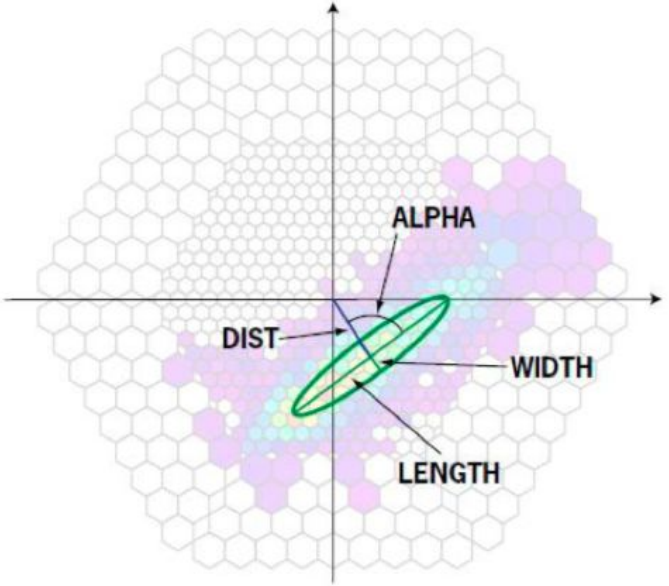
Gamma | Hadron Classification

Max Gebhard

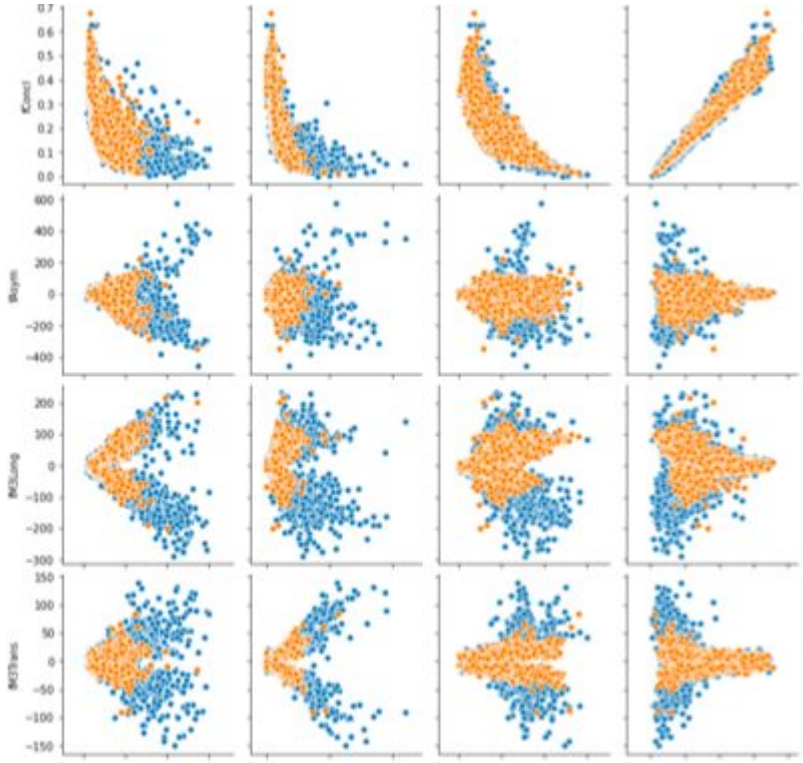
Imaging

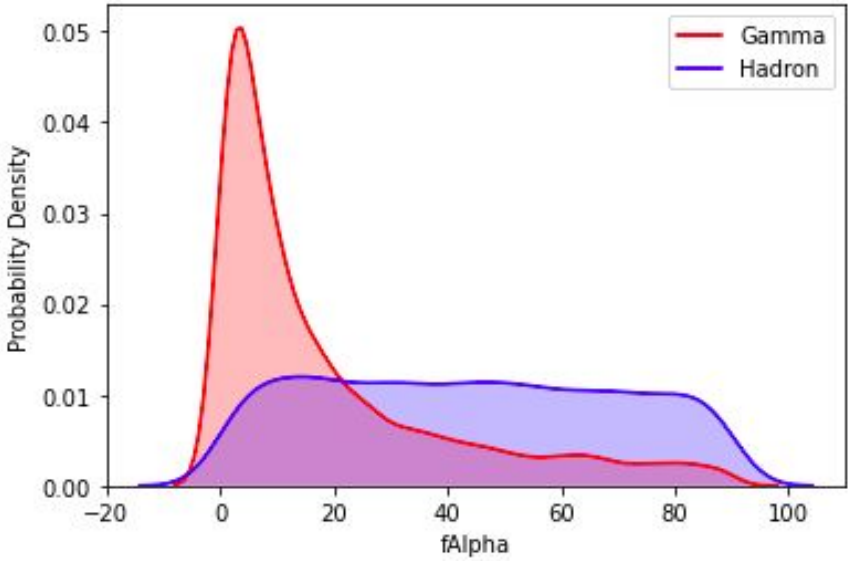






SNS Plots

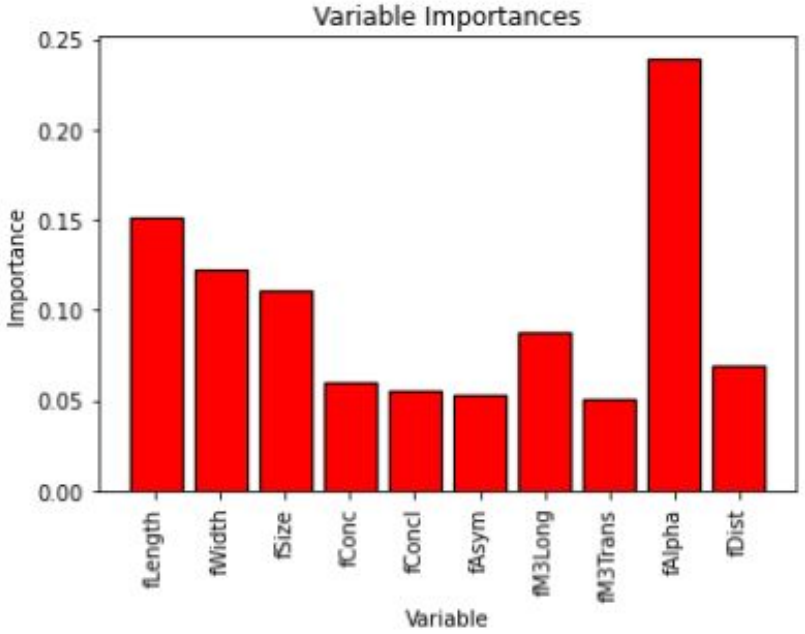



Gamma
Hadron
Probability Density
0.05
0.04
0.03
0.02
0.01
0.00
−20
0
20
40
60
80
100
fAlpha

Initial Models by F1 Balance

	Model	Accuracy	Precision	Recall	F1	CV_Precision
2	Random Forest	0.876972	0.873819	0.945239	0.908127	0.861921
3	SVM	0.823081	0.812522	0.947177	0.874697	0.791886
1	Decision Tree	0.811777	0.823077	0.906048	0.862572	0.825102
4	kNN	0.789695	0.805455	0.893145	0.847036	0.786804
0	Logistic	0.778128	0.795948	0.887097	0.839054	0.790059
5	Bayes	0.728181	0.734893	0.912097	0.813962	0.786804

Feature Importance and Initial Random Forest




```
n_estimators = 50, random_state = 3, n_jobs = -1
              precision    recall  f1-score   support

           0       0.87      0.76      0.81      1357
           1       0.87      0.94      0.90      2447

    accuracy                           0.87      3804
   macro avg       0.87      0.85      0.86      3804
weighted avg       0.87      0.87      0.87      3804

0.8649704107553504
```

Feature Engineering

1 $alpha < 7.8$

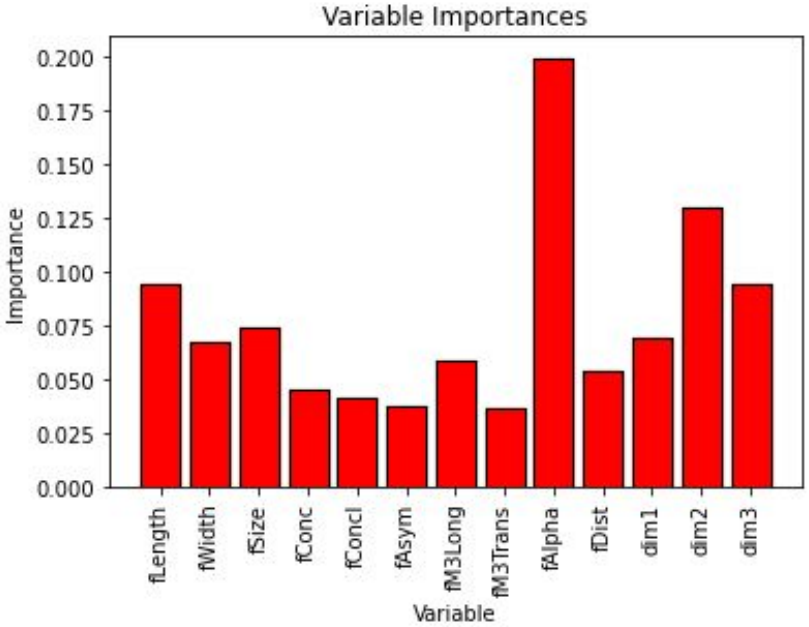
2 $\boxed{length/(-67. + 40. * size)} < 1.35$

3 $\boxed{width/(27.9 - 22.5 * size + 6.7 * size^2)} < 1.12$

4 $size > 3.2$

5 $\boxed{\text{dens} = \frac{\log_{10}(\text{size})}{\text{length} \times \text{width}}}$

No Cut Feature Importance and Second RF



n_estimators = 50, n_jobs = -1

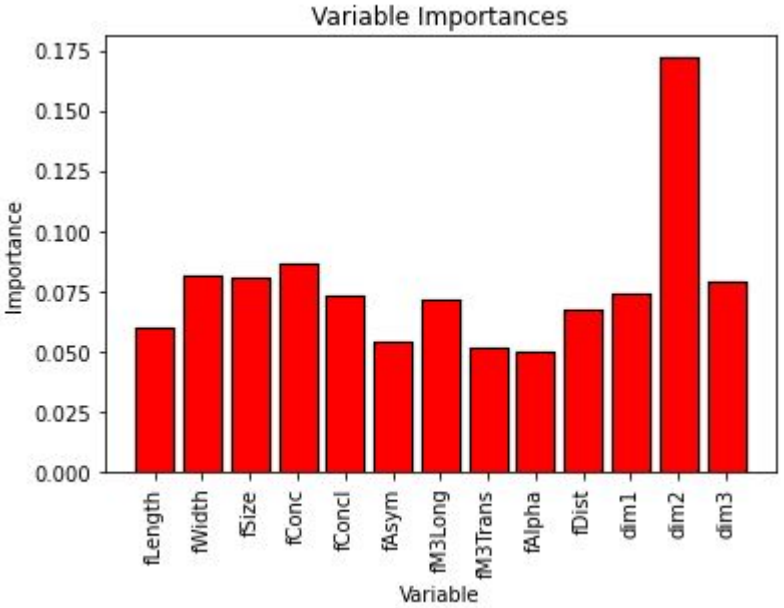
	precision	recall	f1-score	support
0	0.88	0.81	0.84	1318
1	0.90	0.94	0.92	2486
accuracy			0.90	3804
macro avg	0.89	0.88	0.88	3804
weighted avg	0.89	0.90	0.89	3804

0.8888344079006407

Feature Engineering

1	$alpha < 7.8$
2	$length/(-67. + 40. * size) < 1.35$
3	$width/(27.9 - 22.5 * size + 6.7 * size^2) < 1.12$
4	$size > 3.2$
5	$\text{dens} = \frac{\log_{10}(\text{size})}{\text{length} \times \text{width}}$

Full Cut Feature Importance and Modified RF



n_estimators = 50, n_jobs = -1

	precision	recall	f1-score	support
0	0.85	0.39	0.53	72
1	0.96	0.99	0.97	948
accuracy			0.95	1020
macro avg	0.90	0.69	0.75	1020
weighted avg	0.95	0.95	0.94	1020

0.9477007407452296


Thanks for listening!

Extra: Optimizing Hyperparameters for XGBoost