BDT's: A Beginner's Perspective

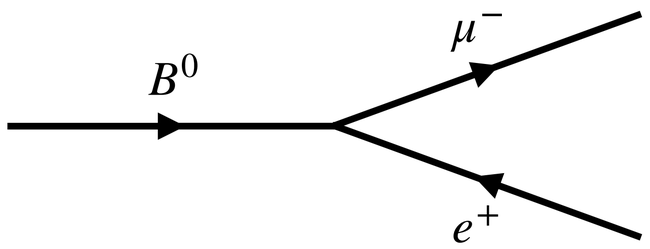
Santi Naylor

VIRGINIA TECH

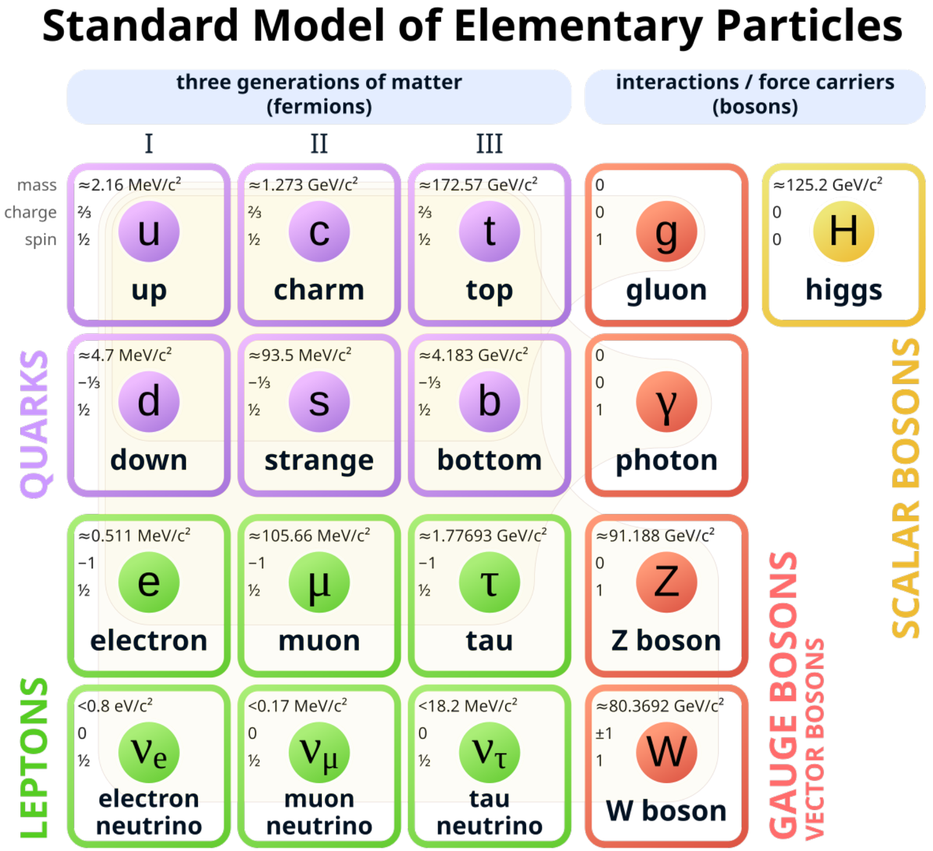
Belle II

$B^0 \to \ell\ell$

- $\ell \in \{e, \mu\}$
- Lepton flavor violating/heavily suppressed decay modes
  - $\mathcal{B}\left(B^0 \to e^+e^-\right) \lesssim 10^{-9}$
  - $\mathcal{B}\left(B^0 \to \mu^+\mu^-\right) \lesssim 10^{-10}$
  - $\mathcal{B}\left(B^0 \to e^\pm\mu^\mp\right) \lesssim 10^{-9}$



	L_e	L_μ
e^-	$+1$	0
e^+	-1	0
μ^-	0	$+1$
μ^+	0	-1



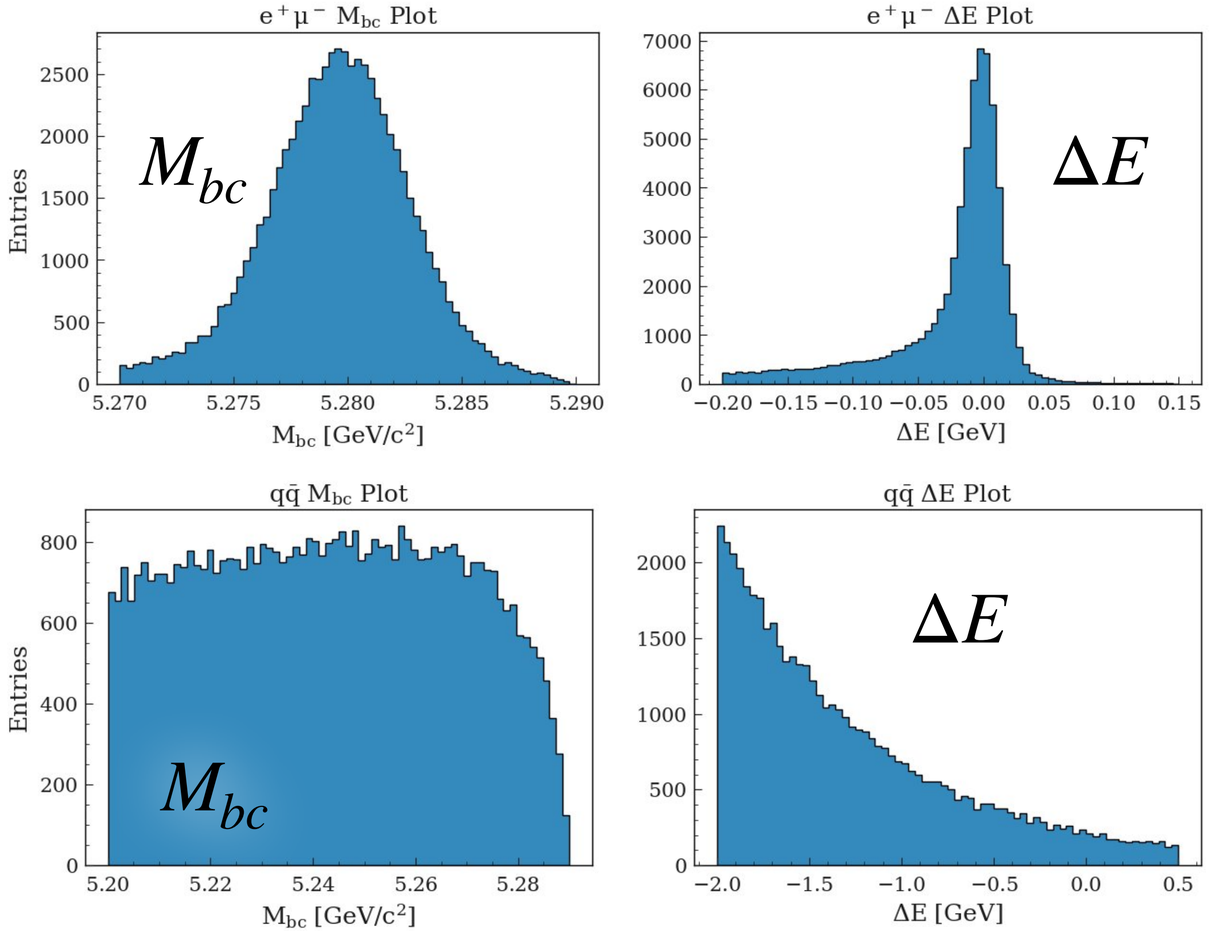
Signal vs. Background

Signal Modes

- $B^0 \to e^{\pm}\mu^{\mp}$
- $B^0 \to e^+e^-$
- $B^0 \to \mu^+\mu^-$

Background Modes

- $B^0 \to K^+K^-$
- $B^0 \to K^{\pm}\pi^{\mp}$
- $B^0 \to \pi^+\pi^-$
- $e^+e^- \to q\bar{q}: q \in \{u, d, s, c\}$
- $e^+e^- \to \tau^+\tau^-$
- $e^+e^- \to e^+e^-X$
- Other Rare B^0 decays



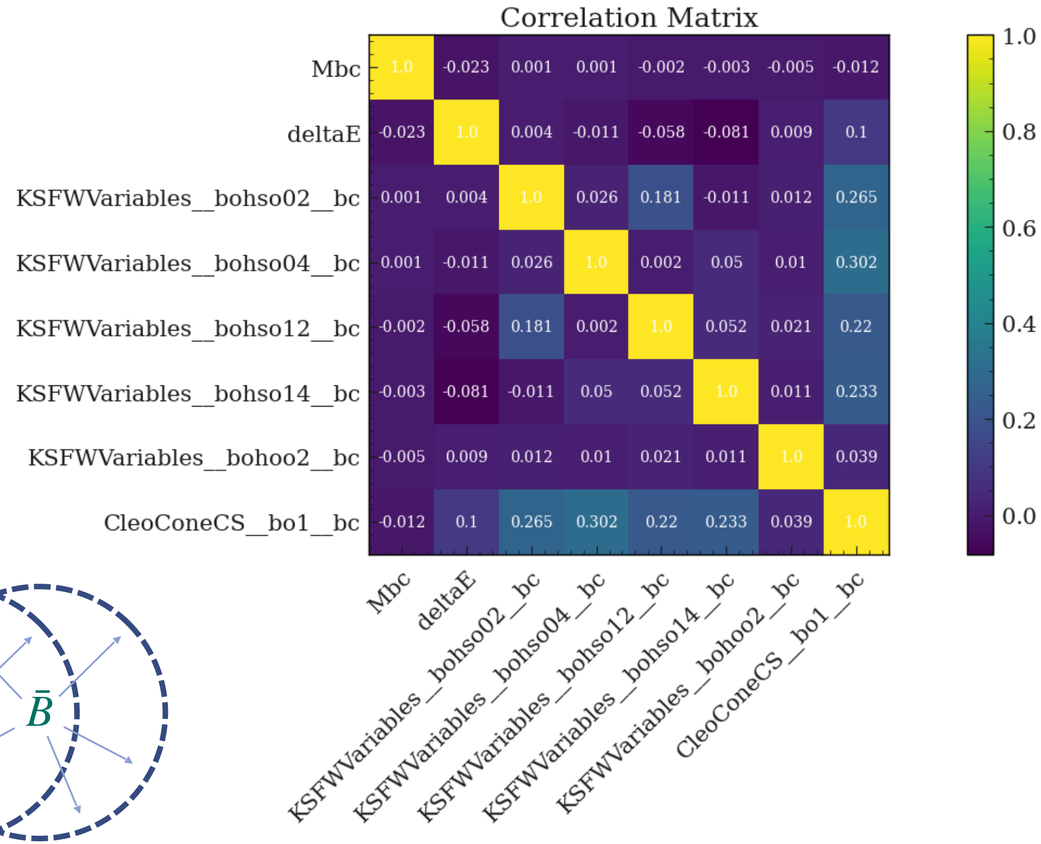
$$M_{bc} = \sqrt{(E_{beam})^2 - (p_B)^2}$$

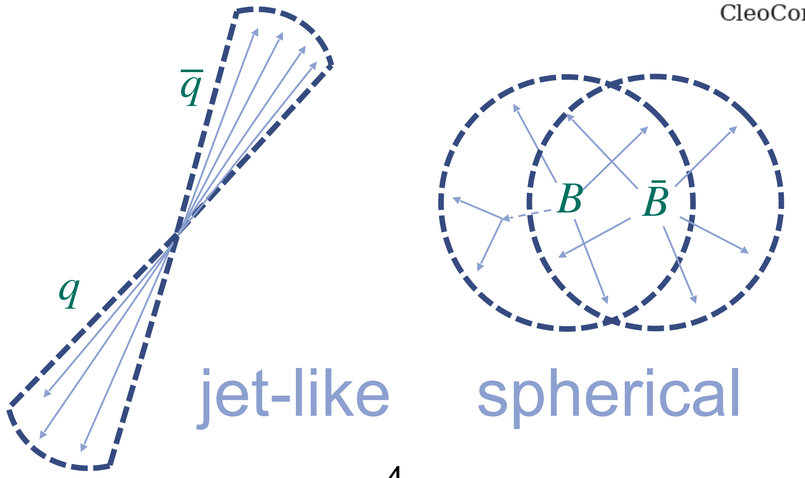
$$\Delta E = E_B - \frac{E_{beam}}{2}$$

Continuum Suppression

$e^+e^- \rightarrow q\bar{q}$

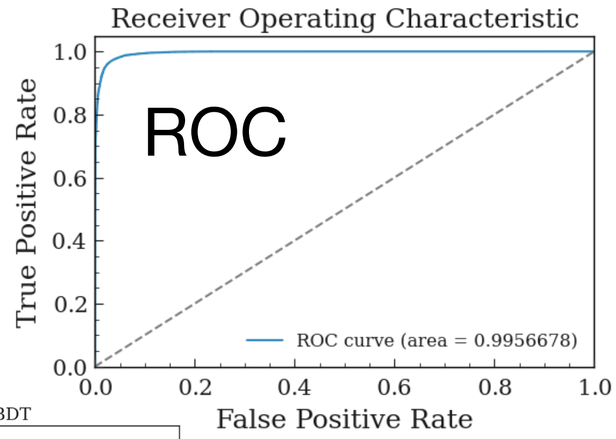
- Boosted Decision Trees (BDT's)
  - Scikit-Learn
  - Signal = 1 and Background = 0
- Choose distinctive variables
  - Event Shape
    - KSFW and Cleo Cones
  - Check correlations

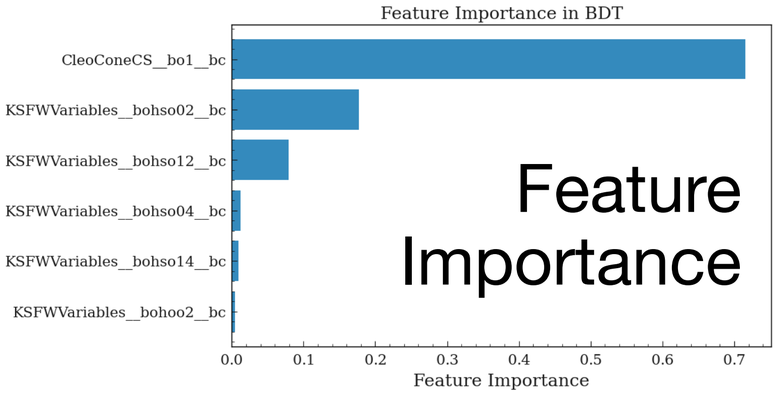


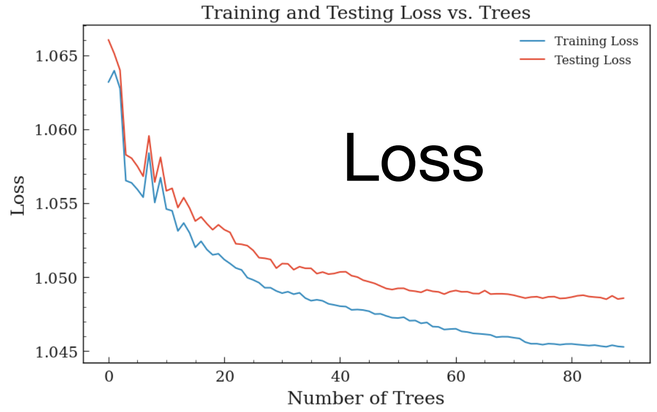


Training and Testing

- Input variables
  - Do not train on fitting variables
- Receiver Operating Characteristic (ROC)
  - Area Under Curve (AUC) shows performance
- Feature Importance
  - Shows how relevant the variables are in the model
  - Cut variables with small feature importance
- Loss
  - Difference between predicted values and true values
  - Minimize



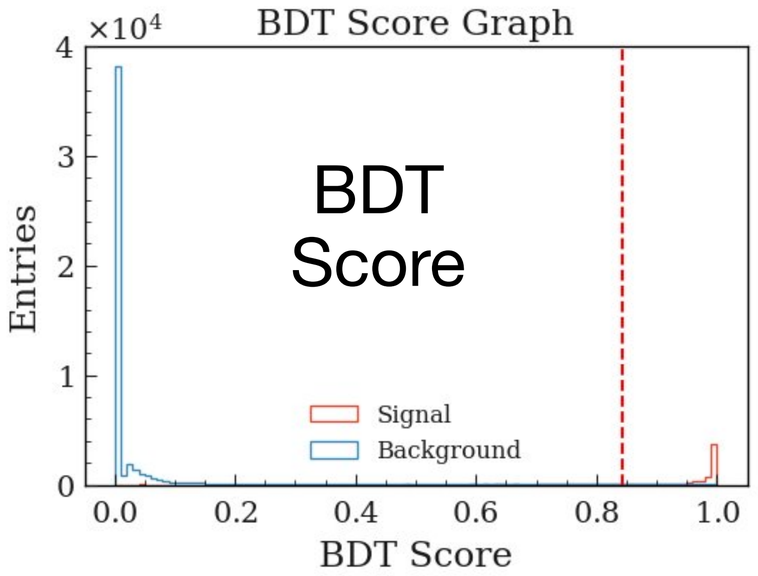


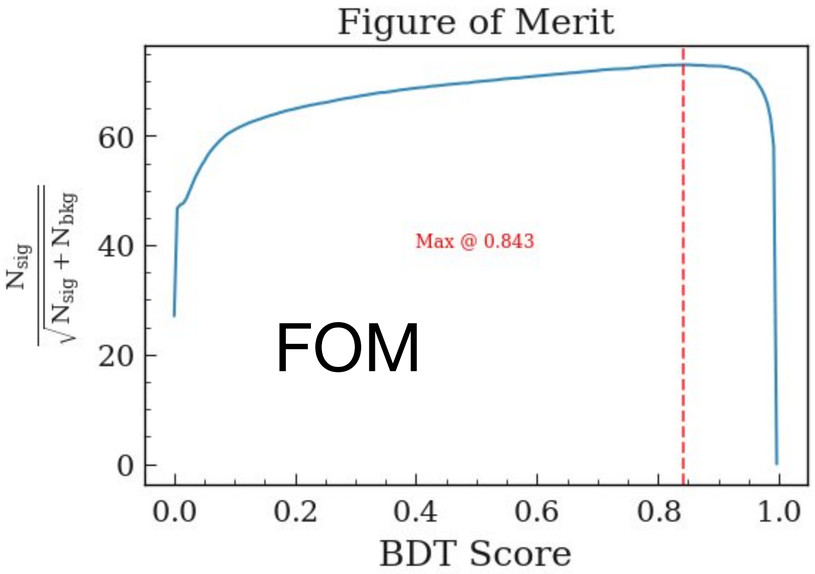


Scoring

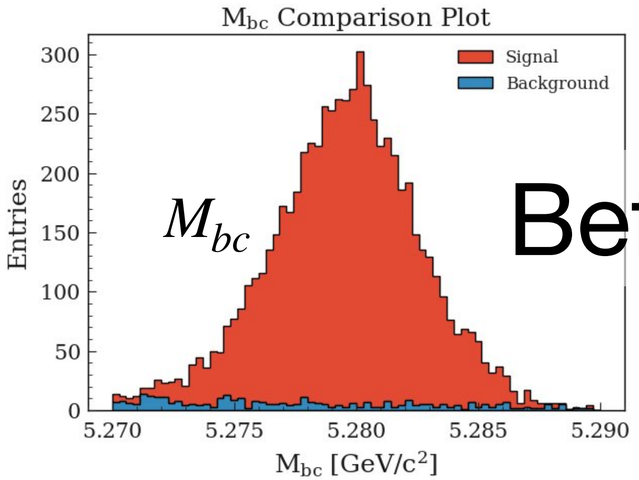
- BDT Score
  - Two Distributions around 0 and 1
- Optimal BDT Score Cut
  - Figure of Merit

$$\frac{N_{sig}}{\sqrt{N_{sig} + N_{bkg}}}$$

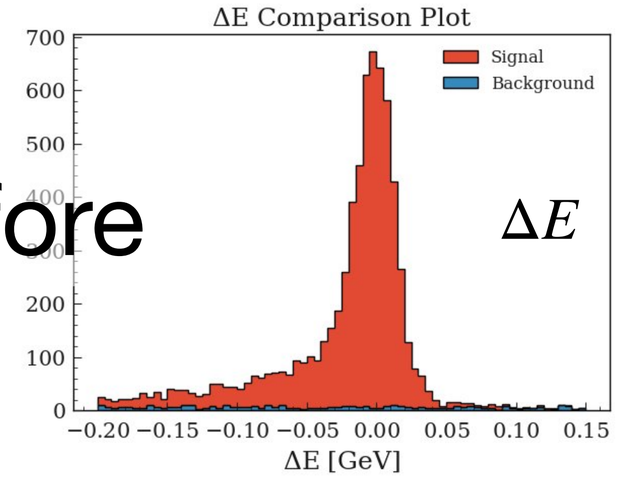


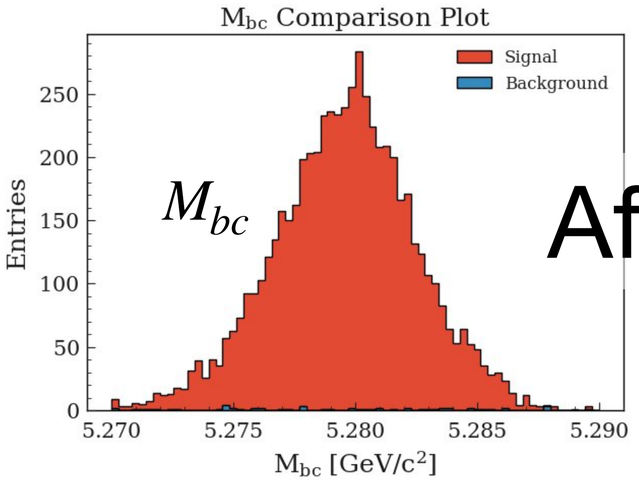


Results

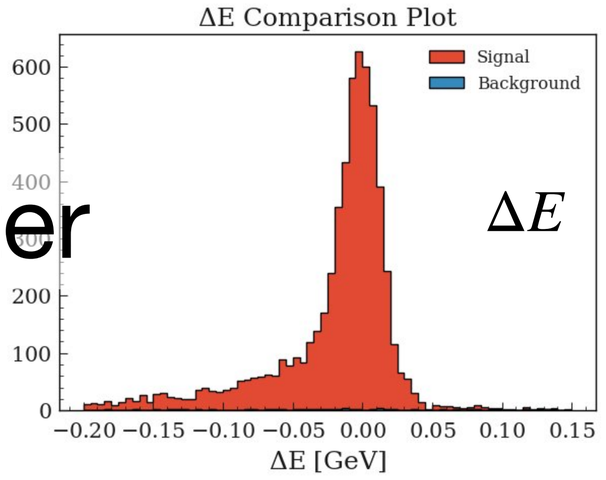


Before





After



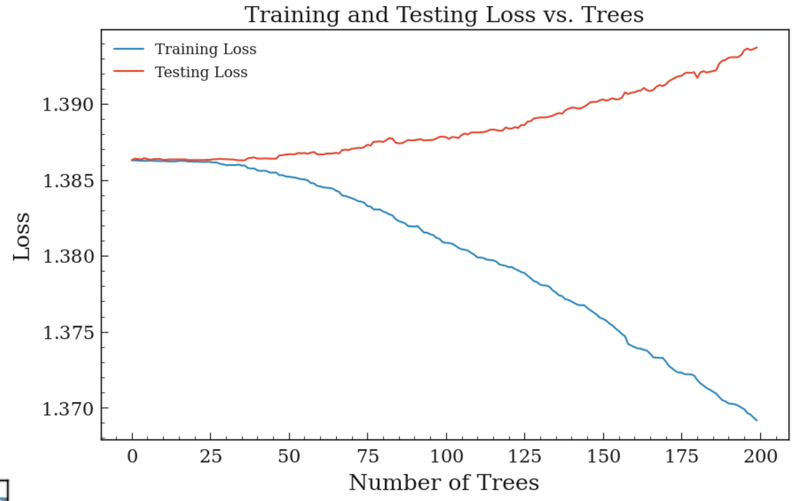
	Before	After	%
Signal	6440	5904	91.7
Background	368	55	14.9

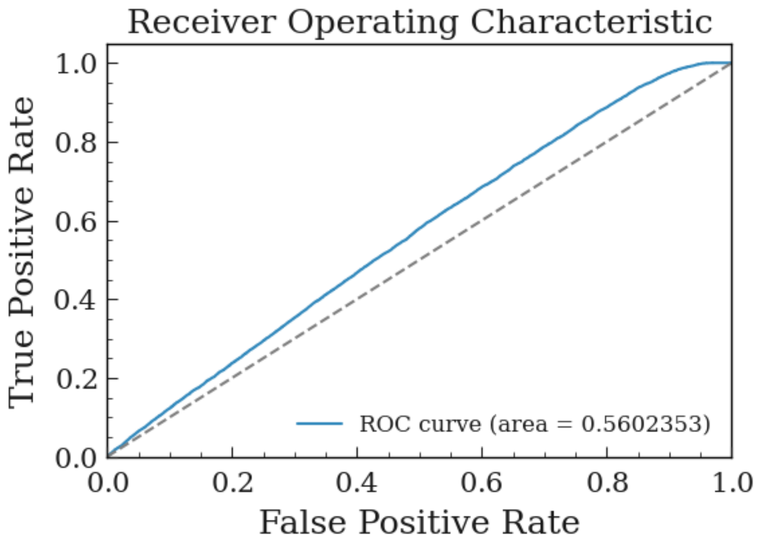
VIRGINIA TECH

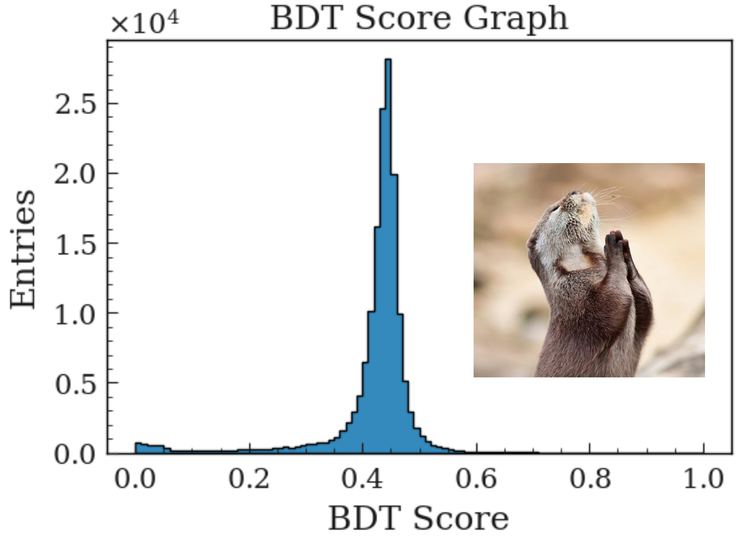
*Dataset not realistically weighted

Limitations

- Bad input = Bad output
- Overfitting
- Example using hadronic background modes:
  $B^0 \to K^+K^-,\ B^0 \to K^\pm\pi^\mp,\ B^0 \to \pi^+\pi^-$
  - Can't use event shape
  - Try other kinds of variables (e.g. kinematic)
  - Better to not use BDT







Thank You

VIRGINIA TECH

Belle II