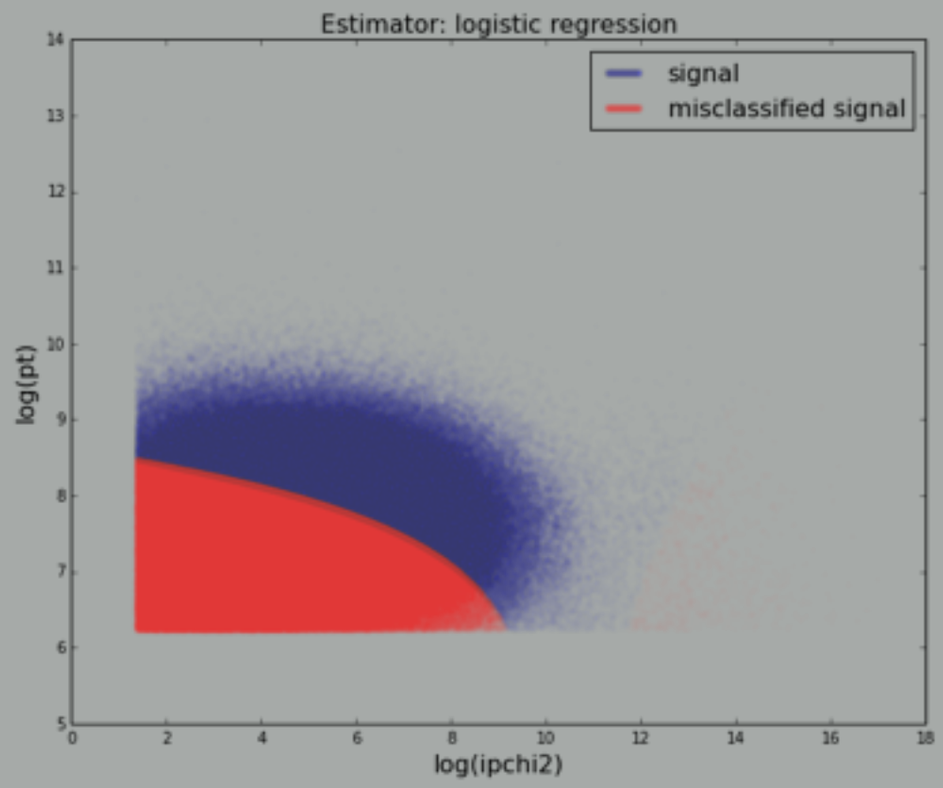
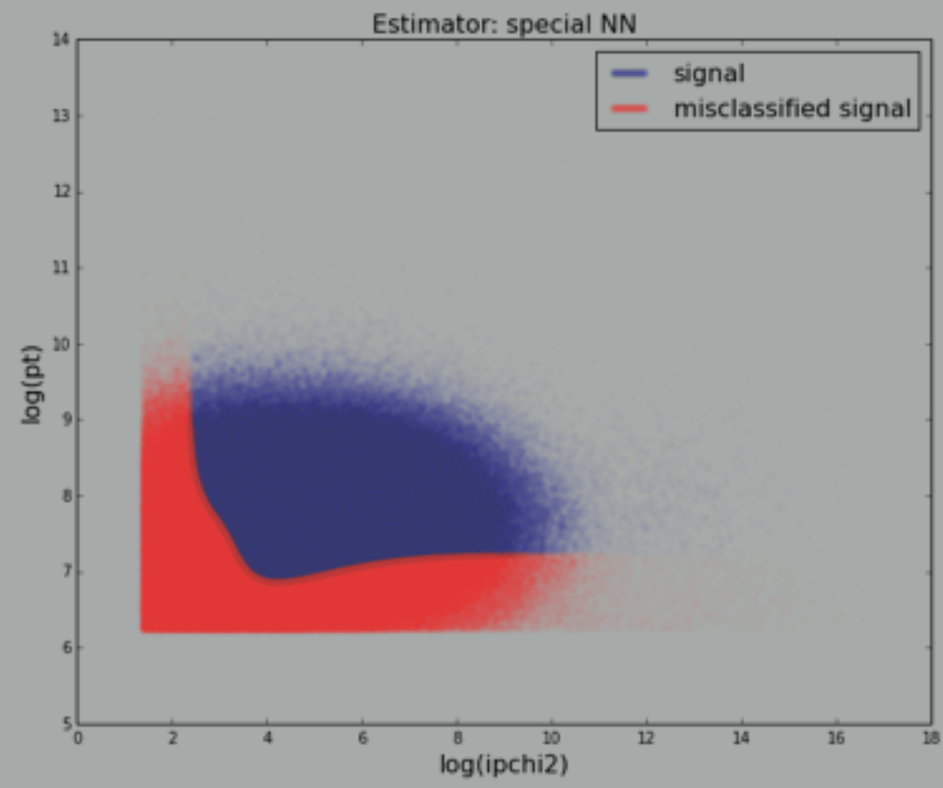
Machine learning in HEP

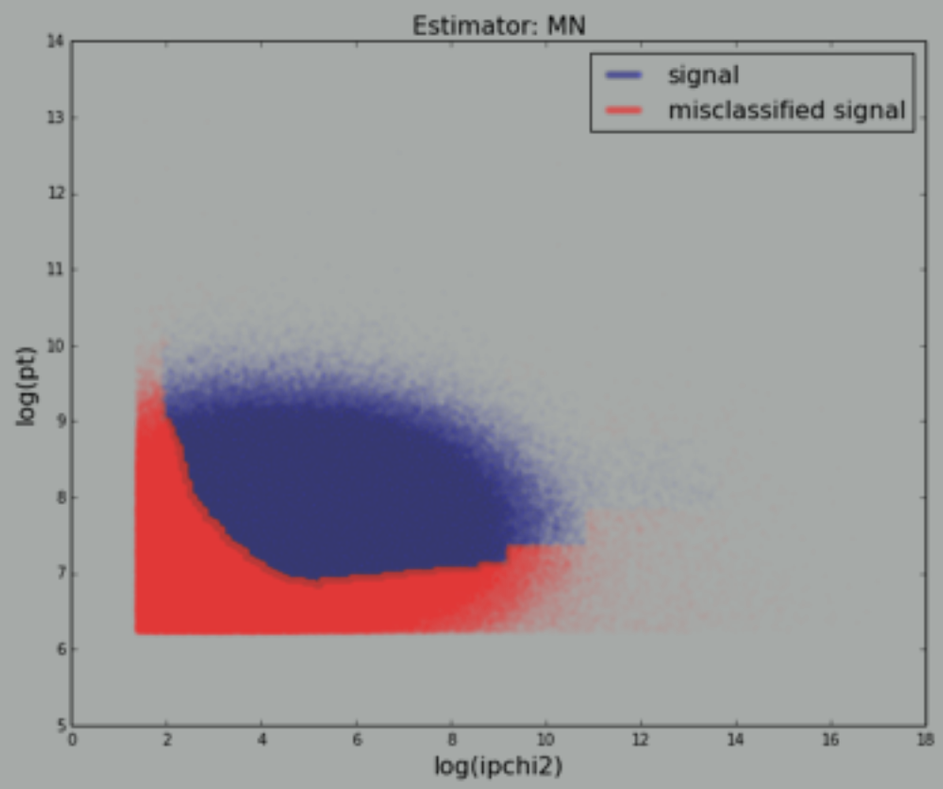
Likhomanenko Tatiana

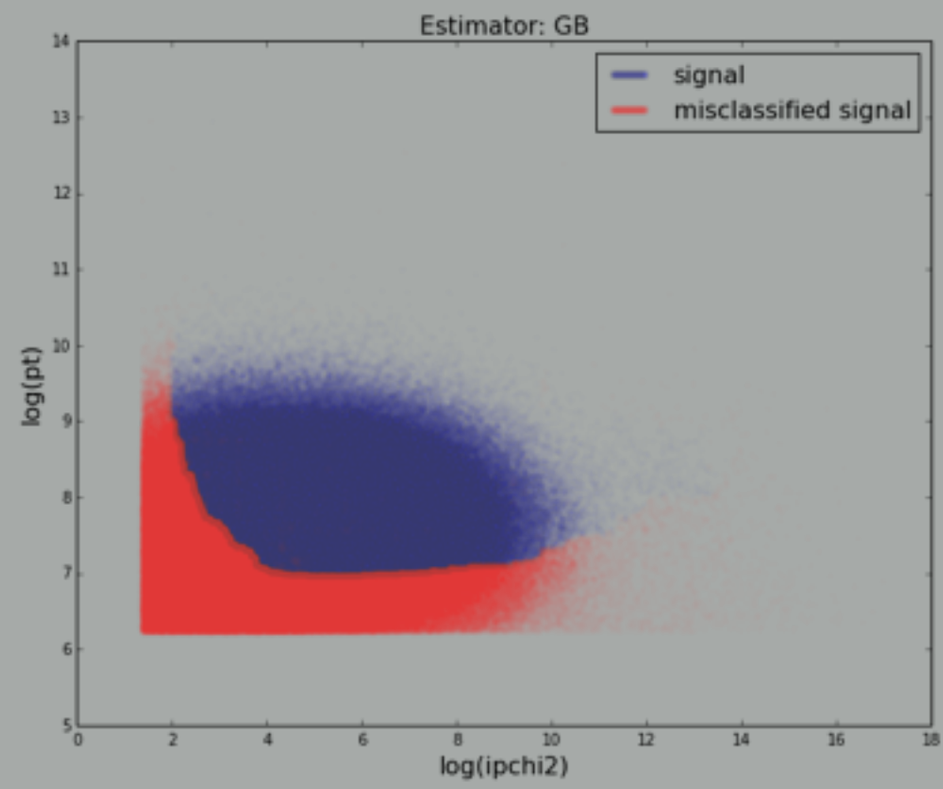
Summer school on Machine Learning in High Energy Physics

Track trigger:
Neural Net vs another models

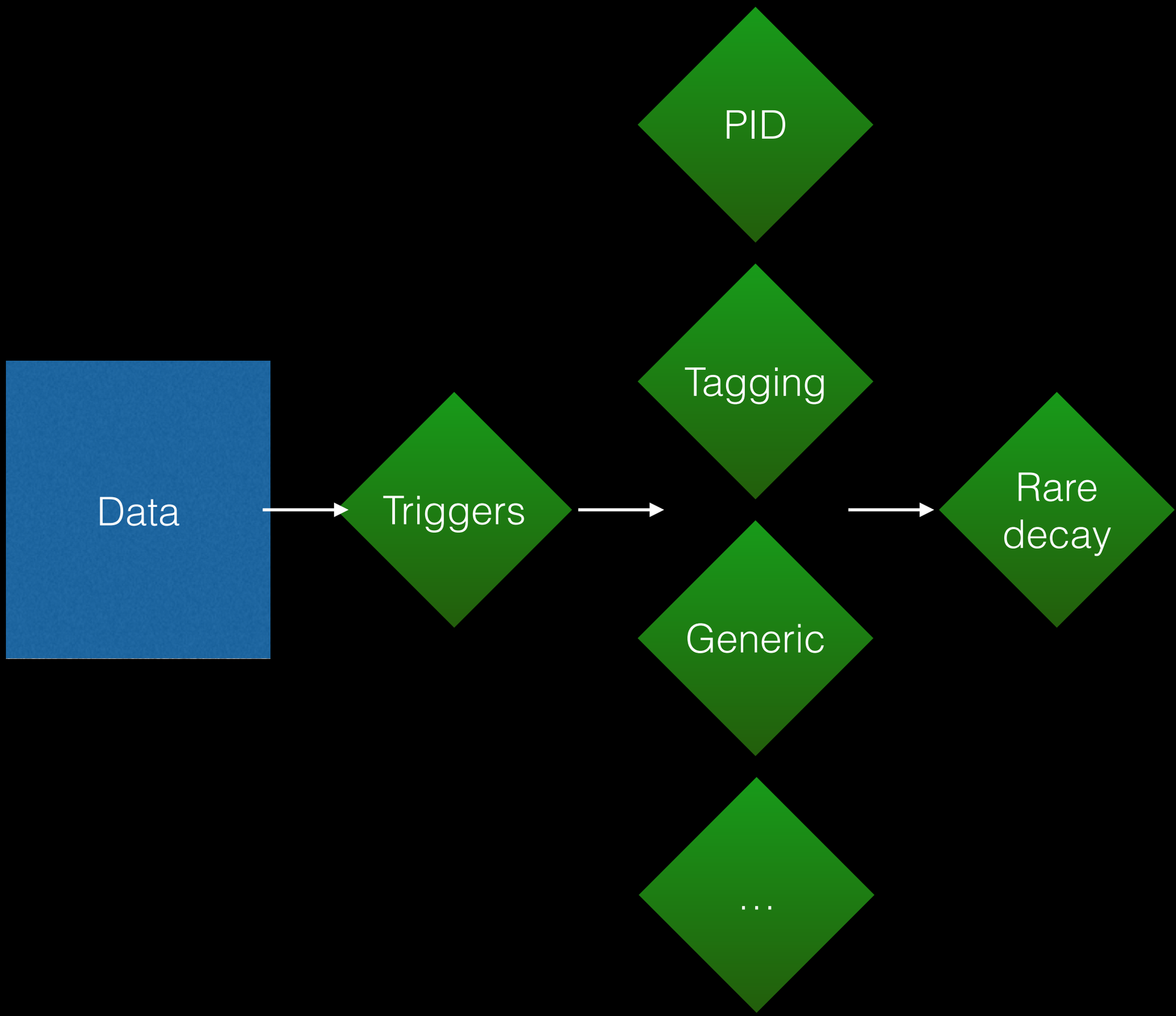








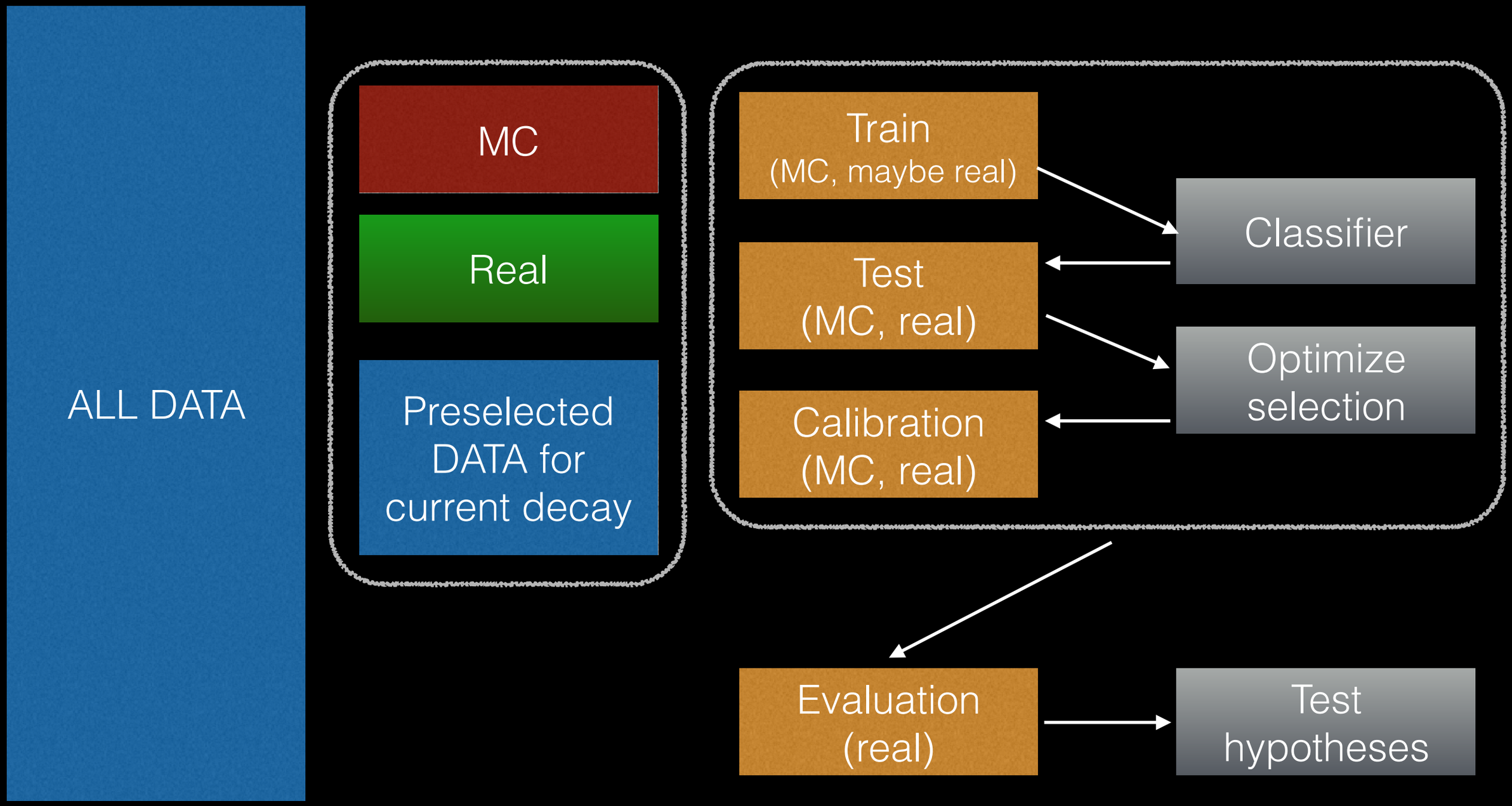
Summary


PID
Data
Triggers
Tagging
Rare
decay
Generic
...

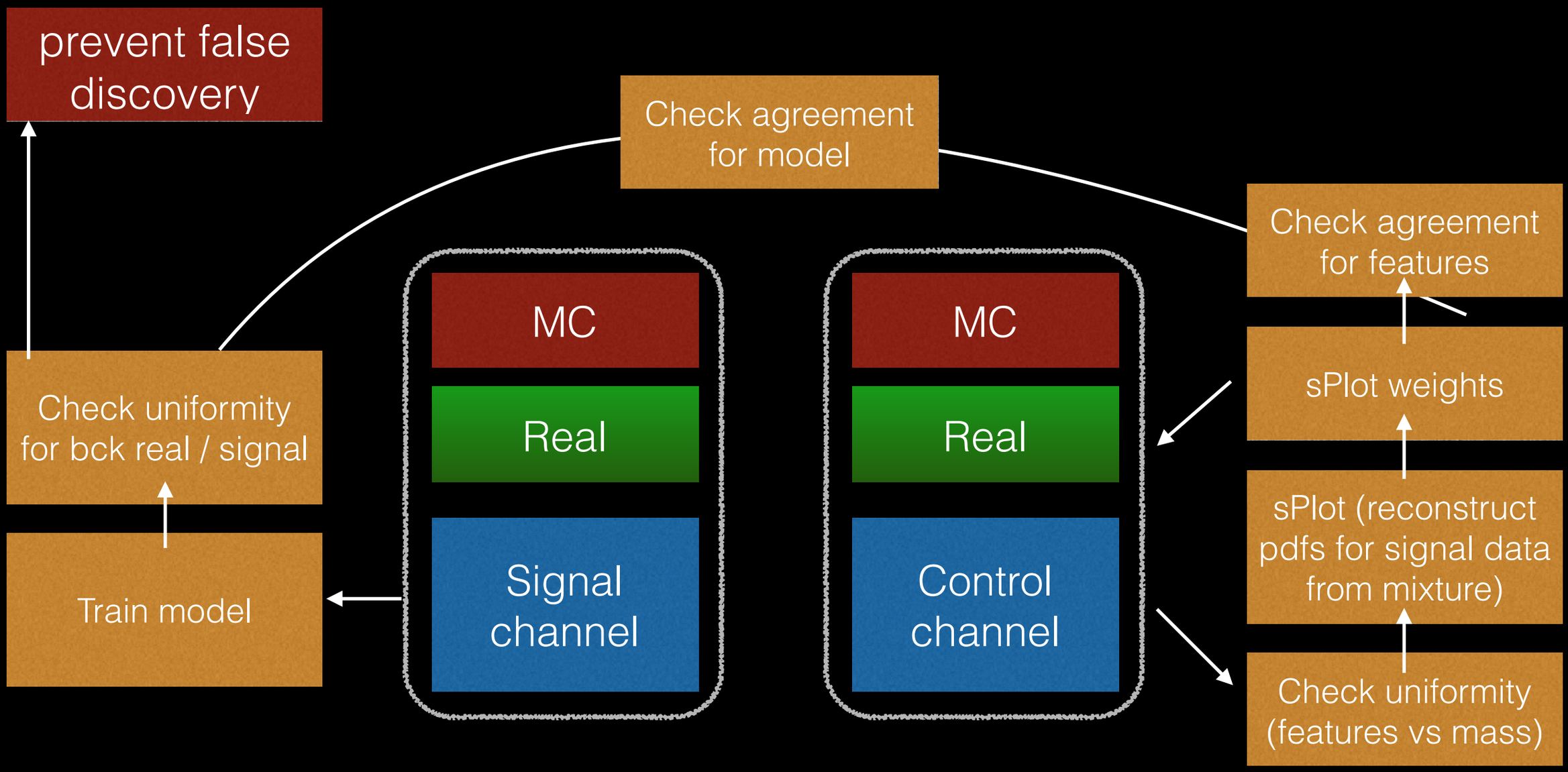
Predict the whole event

- Work with the whole event
- Training sample is set of SVs/tracks
- Noisy signal data
- Apply random forest to clean signal data
- Compare models using the ROC curves and predictions for the whole event

Rare decay analysis



Rare decay analysis: Training



Rare decay analysis: metrics

- Non-uniformity metric: CvM (generate pdf => p-value)
- Disagreement metric: KS (generate pdf => p-value, in 1D without weights KS statistic works)
- Compare distributions using U-test (in ND case we train model and compare distributions for model output)

Rare decay analysis: problems-solutions

- Non-uniformity with mass => throw features / UGB
- Disagreement => throw features / reweighting (bin / GB) / Iterative learning
- Comparison of ND pdfs => apply ML
- Overfitting => check this using only learning curves
- Save data for evaluation stage => Folding scheme

Real data analysis (without MC)

- Don’t have MC data
- Data is unlabeled
- sPlot weights and probabilities to be signal/bck event
- Different strategies to train on sPlot data