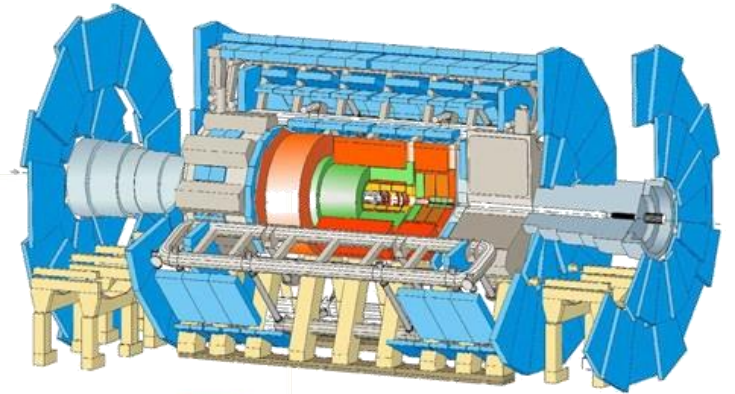


Machine learning para descubrir el bosón de Higgs

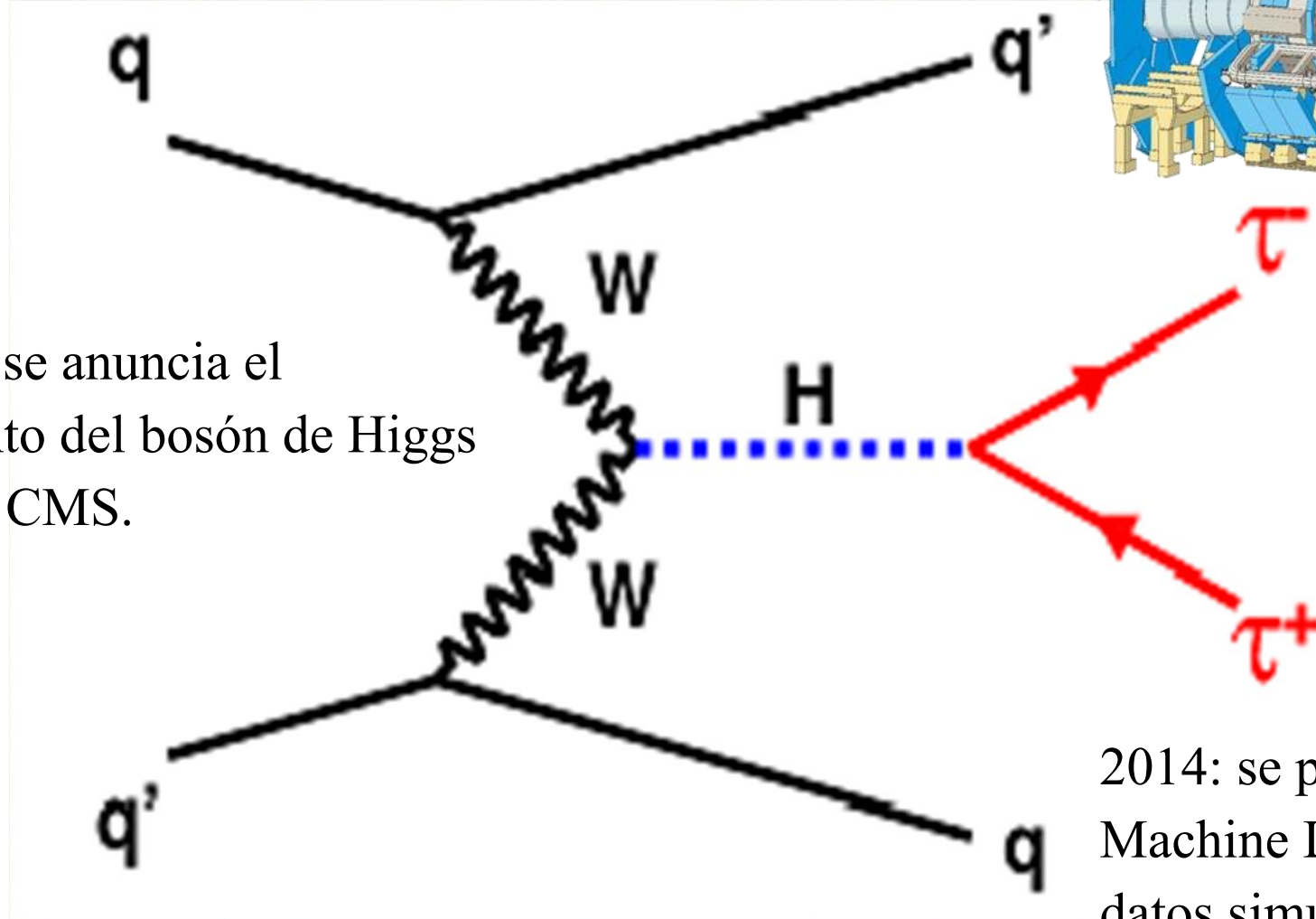
IFIC Summer Student 20/07/2018

Sergio Ferrando y Ámbar Pérez

El bosón de Higgs



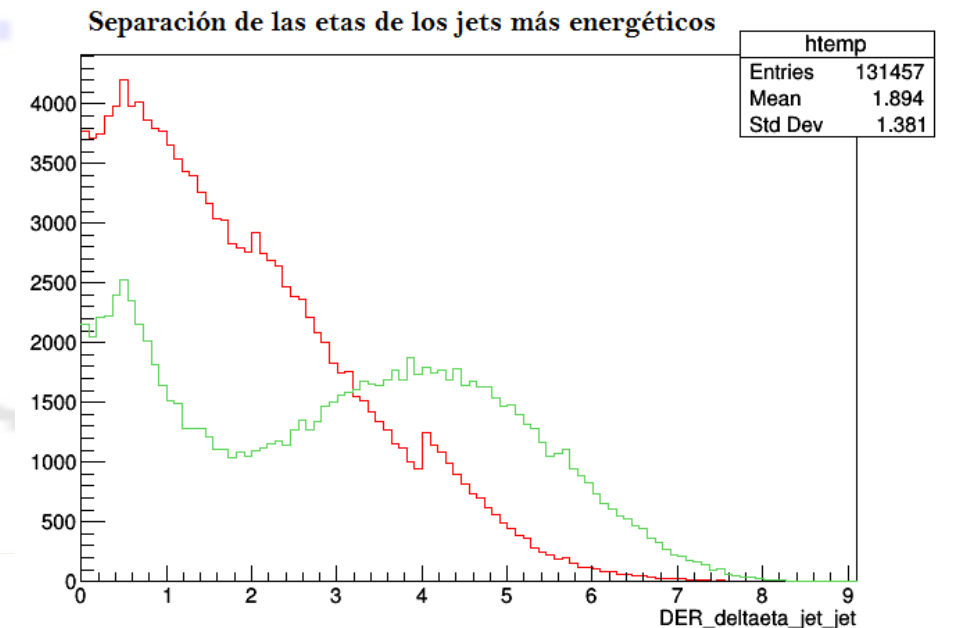
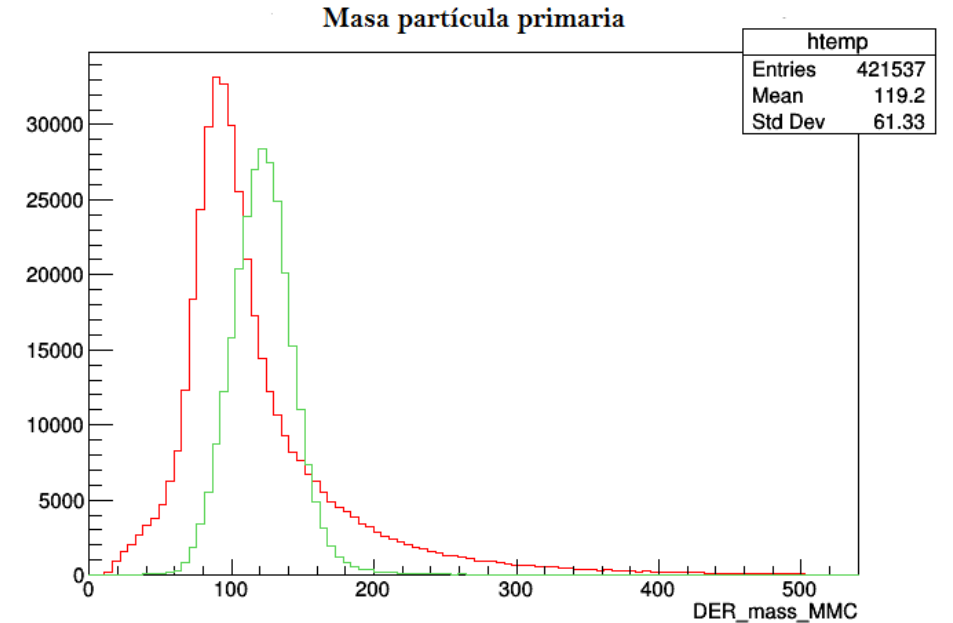
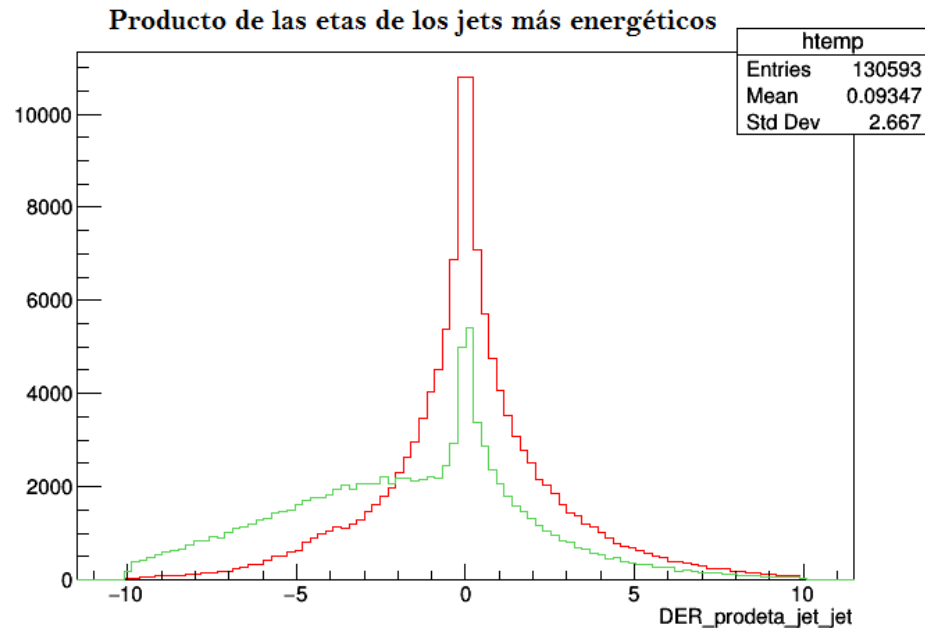
4 julio 2012: se anuncia el descubrimiento del bosón de Higgs en ATLAS y CMS.



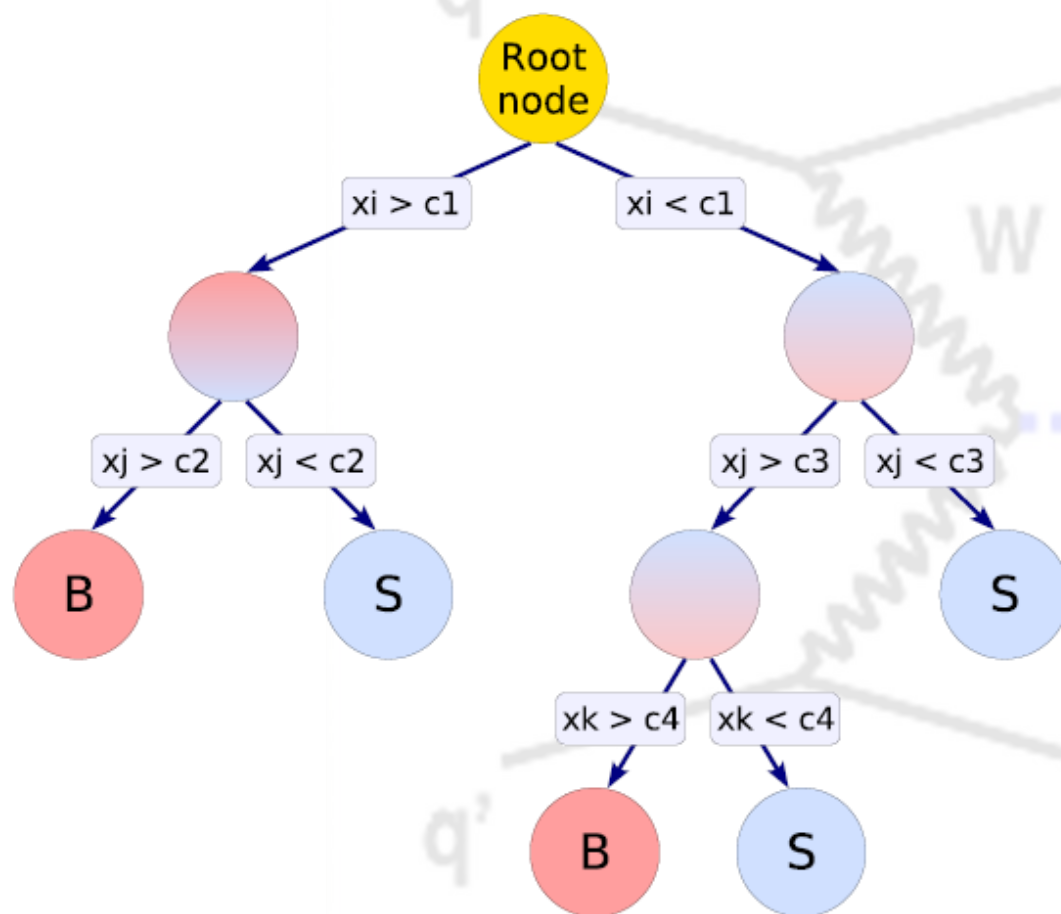
2014: se propone Higgs Boson Machine Learning Challenge con datos simulados de ATLAS.

Variables

Se han utilizando datos simulados con Montecarlo, en los que la proporción señal-background no se corresponde con la proporción real.



Boosted Decision Trees (BDT)



Estructura básica:

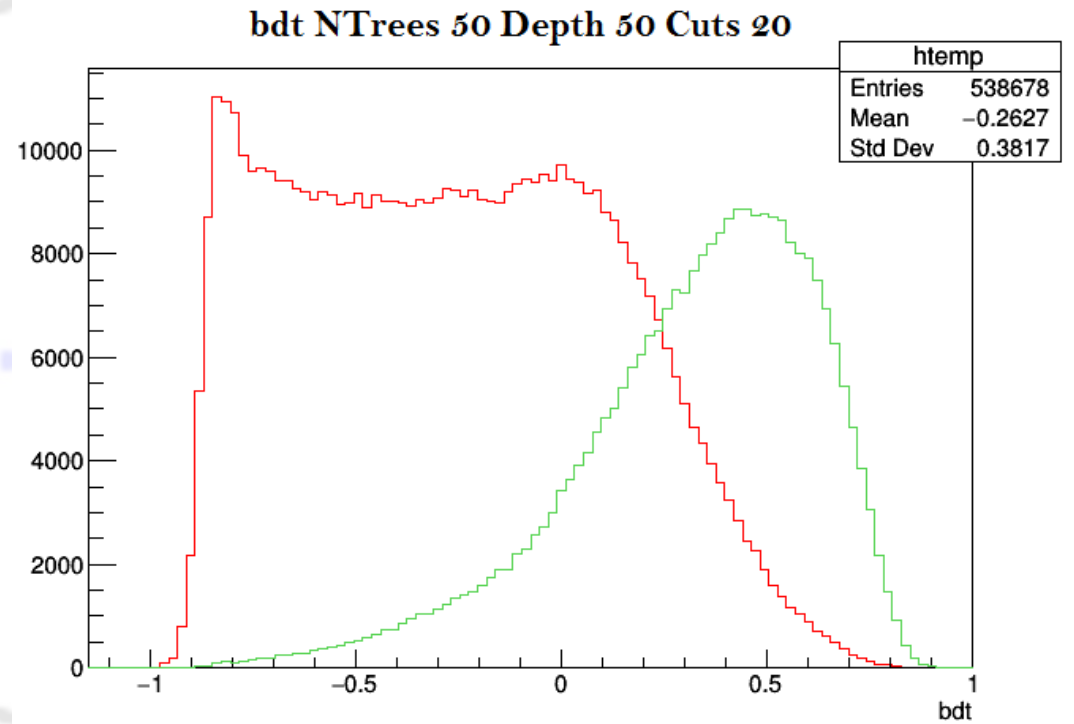
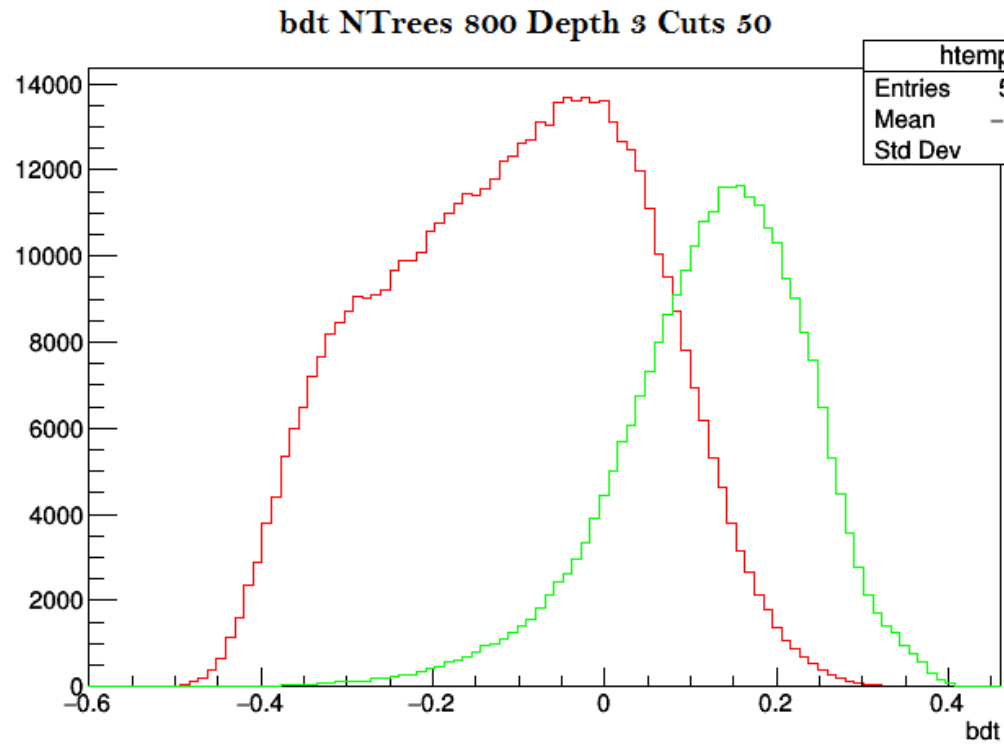
 Tree

 Depth

 Cuts

El BDT utiliza varios árboles a los que atribuye pesos según su eficacia en la fase de entrenamiento.

Variable bdt



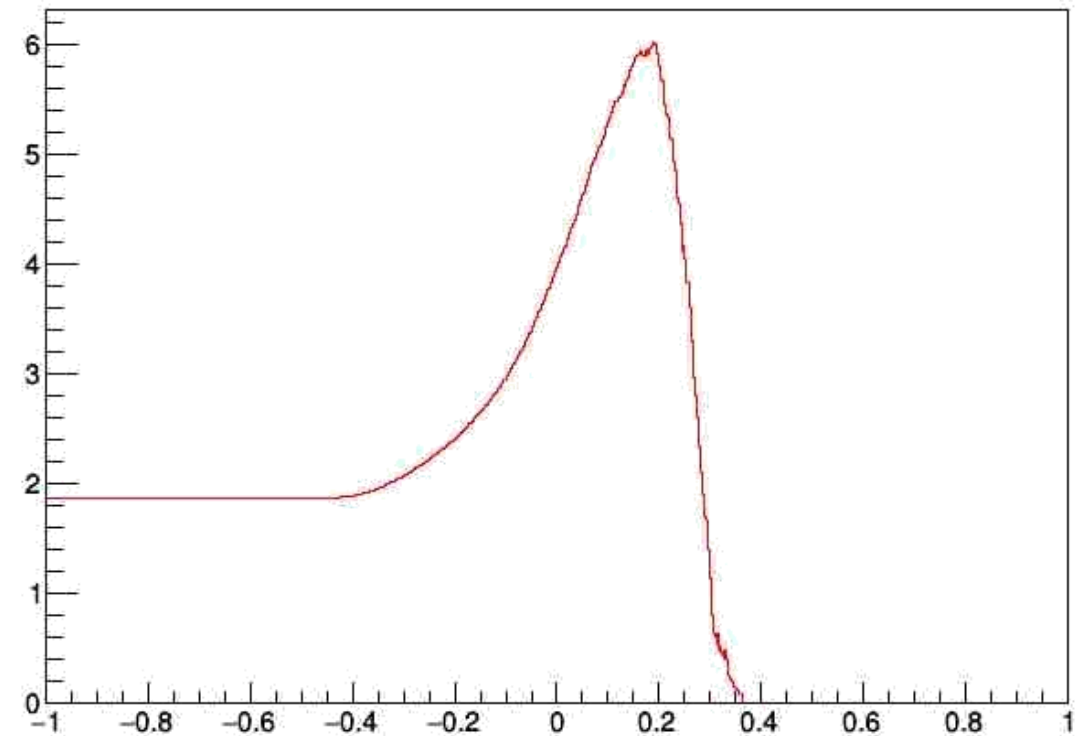
Significancia

$$\frac{S}{\sqrt{b}}$$

Es un criterio para evaluar si ha habido un indicio o un descubrimiento (>5).

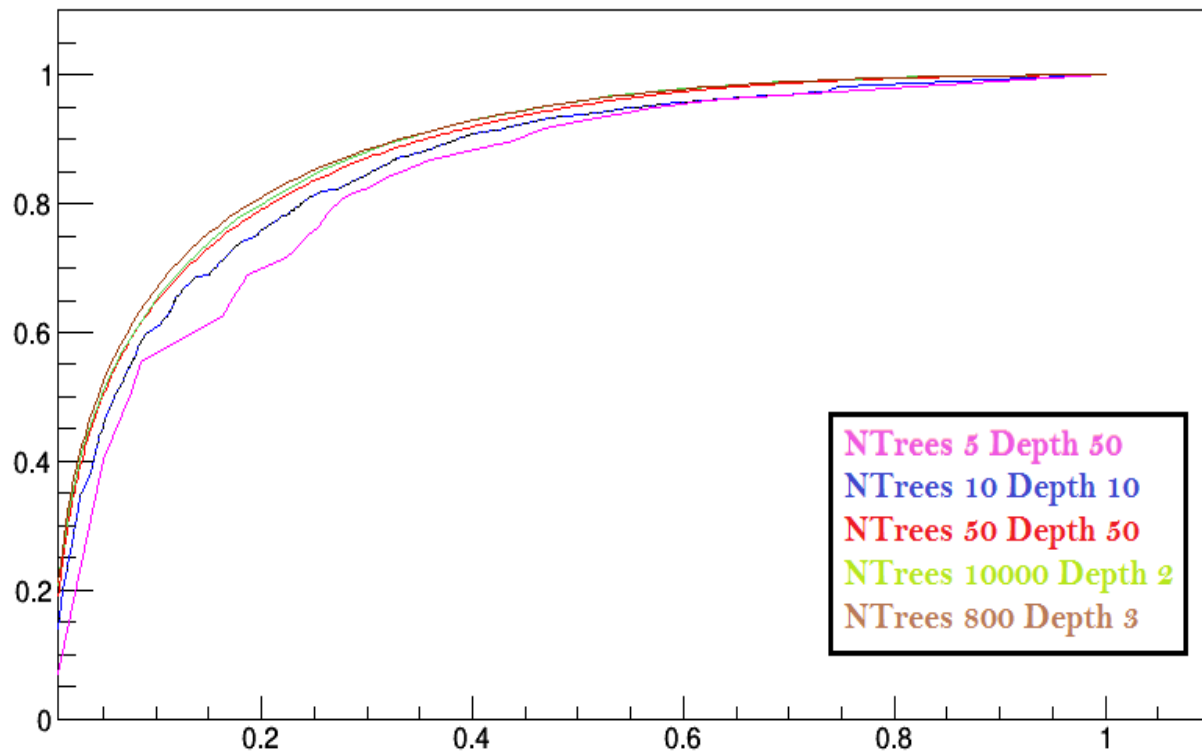
El valor de corte de bdt es el que maximiza la significancia.

Significancia vs bdt

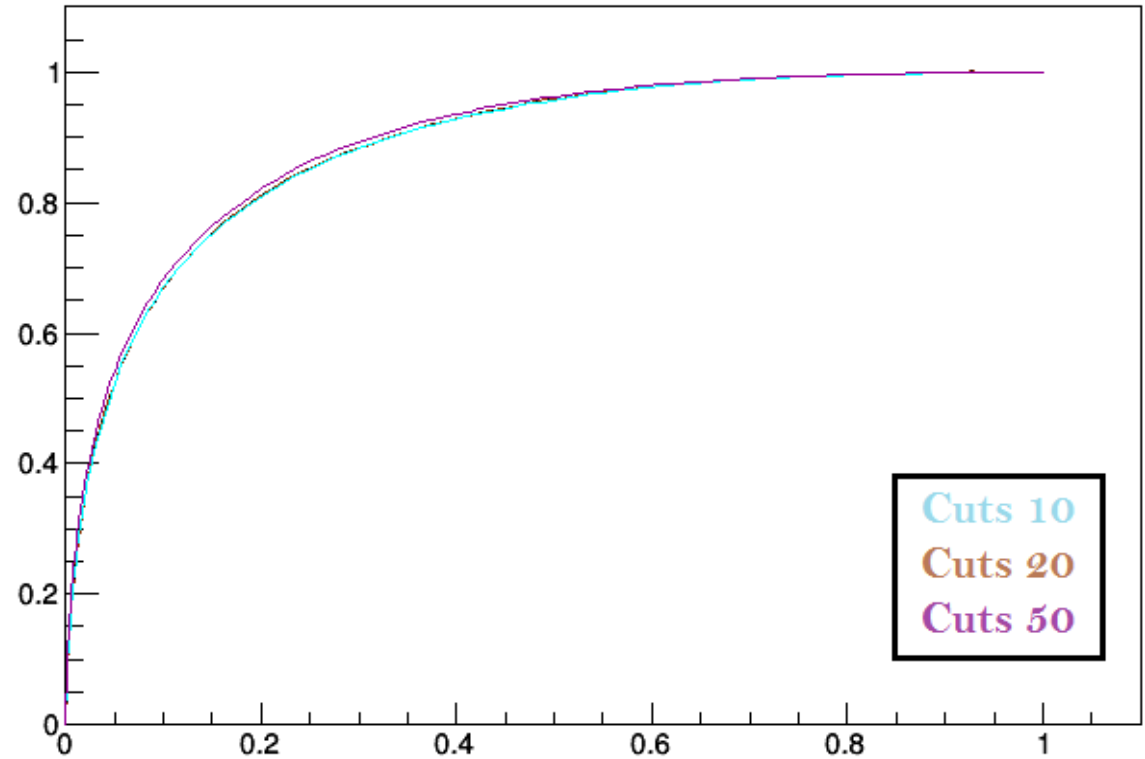


Curvas de eficiencia

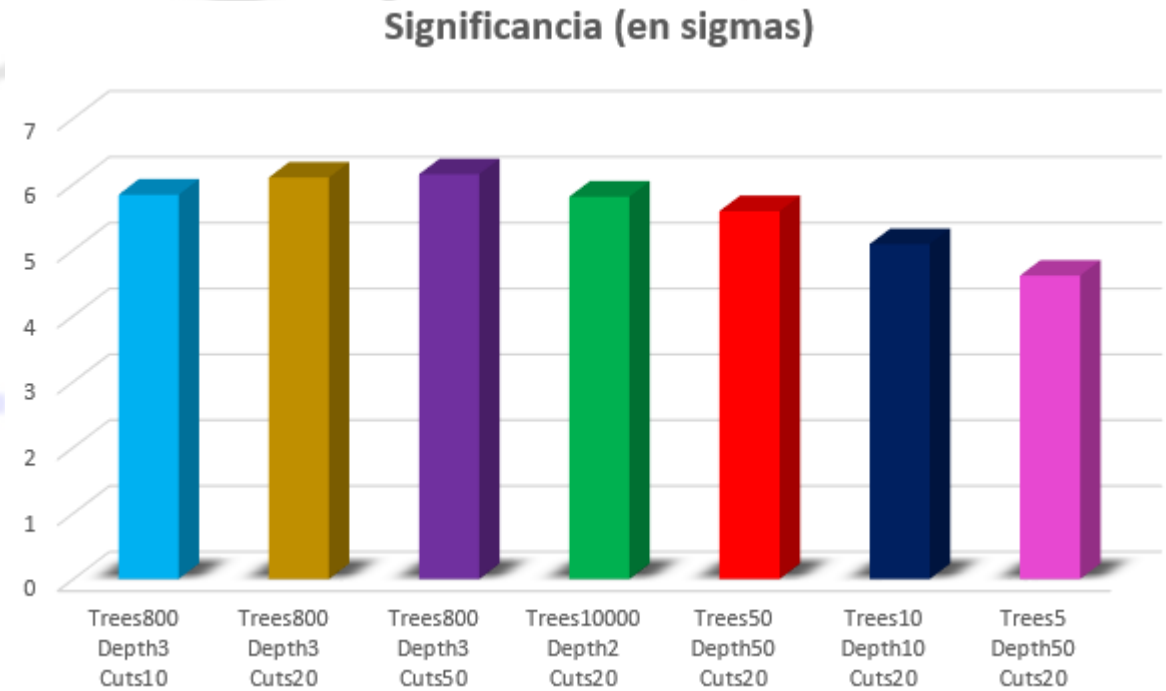
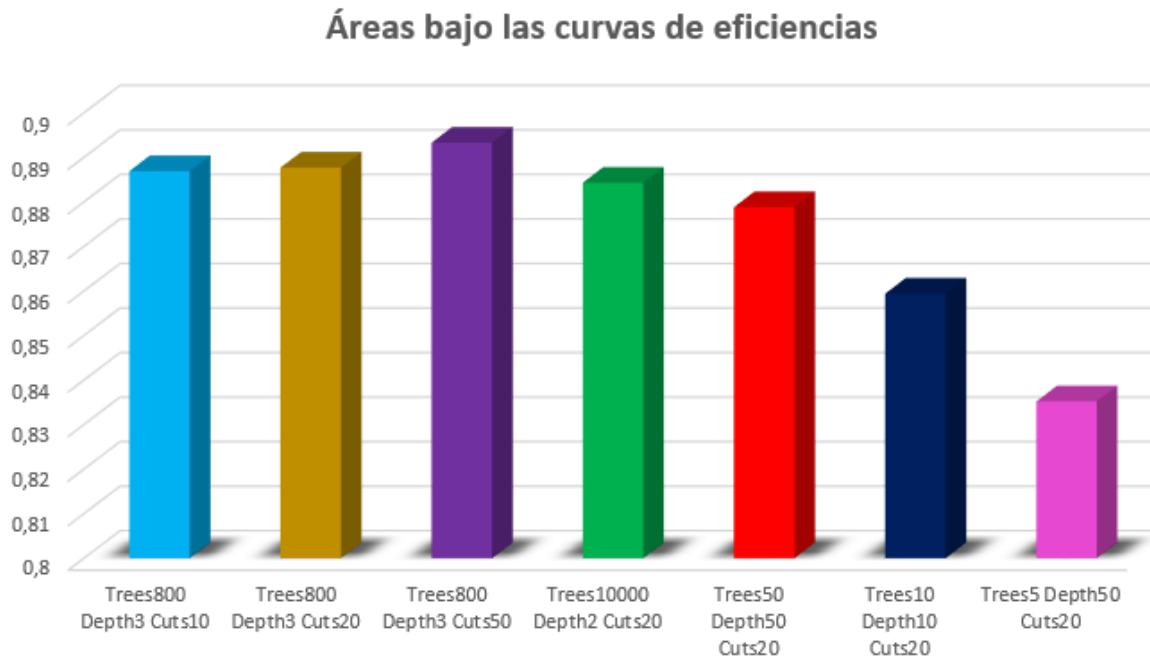
Eficiencias cortes iguales (Cuts 20)



Eficiencias cortes distintos (NTrees 800 Depth 3)



Área bajo la curva de eficiencia y significancia



Conclusiones



Aprendizaje:

- ✱ Detector ATLAS
- ✱ Herramientas de análisis de física experimental
- ✱ Método de análisis multivariado (BDT)
- ✱ Conceptos de estadística



Resultados:

- ✱ Estudio de las variables con mayor poder de separación
- ✱ BDT optimizado para muchos árboles, pocos nodos y número elevado de cortes
- ✱ Descubrir el bosón de Higgs



Mejoras:

- ✱ Comparar con otros métodos de machine learning