Machine Learning in HEP

Finding Higgs Boson

Yusheng Zhao

October 14, 2020

Outline

Naively

```
while(1){
  stick_your_hand_in_the_haystack();

  vehemently_move_your_hand();
  if(needle_stick_in_hand){
    cout << "hurray I found the needle" << endl;
    return 0;
  }

 }
```

A better way?

```
get_a_magent();
use_the_magnet();
```

Magnet is Machine Learning

- Is it really a black box?
- But surely you don't know why it works.
- Not a rigorous algorithm
- There needs to be an invariant over some state that is related to our problem link to where I got this
- Boring, but that's what's awesome about Machine Learning

The most intriguing pattern finding puzzle

Large Amount of Data [1]

- It is costly to develop new algorithm
- Over abundance of data
- Hard and costly for humans to find the pattern

[1] [3]

Higgs Boson [2]

- Last elementary particle in the SM to be observed
- Mechanicsm that gives mass to massive elementary particles predicts its existence.
-

[2] [2]

What more do we know about the Higgs Boson [3]

- Know the rough range of mass
- General Consideration is for it to be smaller than ~ 1 TeV
- Electroweak Measurement says it is < 152 GeV
- LEP Collider says it shoud be > 114.4 GeV

[3] [4]

What haystack?

- what are the back ground noise?
- What are the information that we can get?

Challanges

- Why is it difficults

- describe how the experiment works and diverge to the machine learning part when we hit the obstacle of trying to find what we want to see

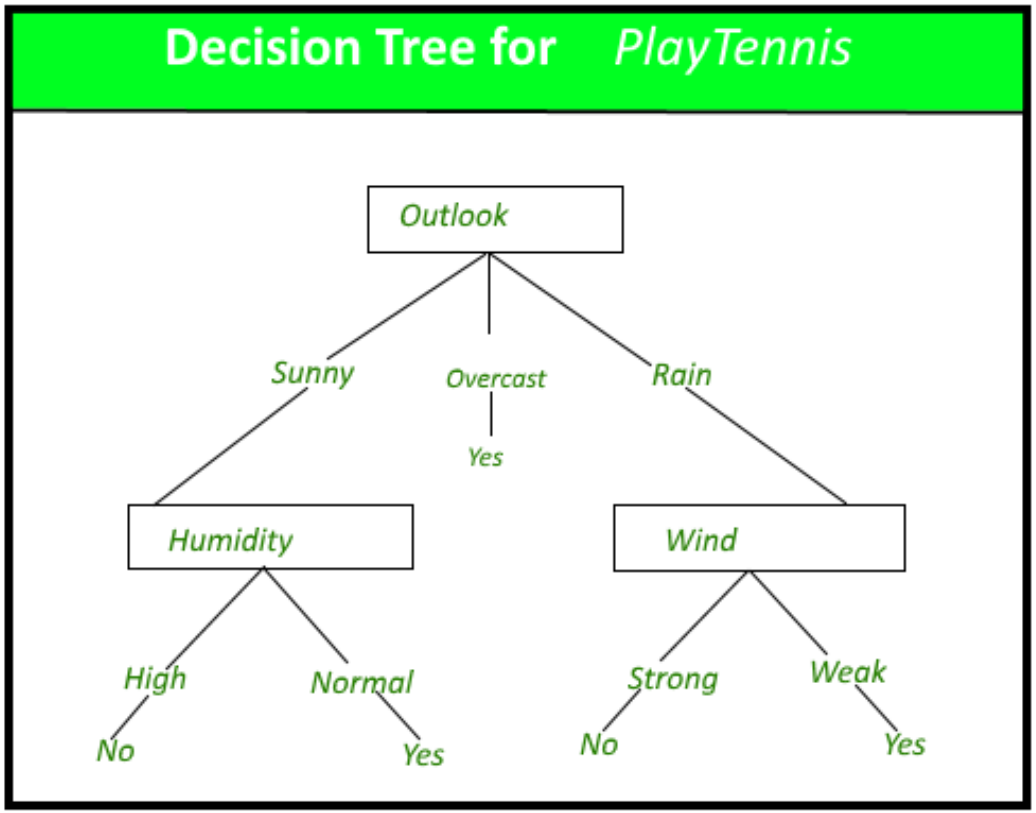
-mix the information about machine learning here.

BDT

Artificial Neural Networks

Decision Trees [1]

- Can be used to assign data to class



Boosted Decision Tree (BDT) [3]

- Convert weak to strong learner ?

Artificial Neural Networks

- Layers of nodes
- Weighted inputs and nonlinear transformation
- rectified linear unit (ReLU)
- Hidden Layers : Deep Learning

Optimizing the model

- Evaluating a cost for model
- Minimize the cost through evolution of model
- Backpropogation (chain rules)
- Stochastic Gradient Descent

- Higgs Boson 2012
- boosted decision tree
- small signals (invariant mass peaks) over large smoothly falling backgrounds

Little bit of physics behind it

-then the Higgs decays and couplings to the heavy W and Z gauge bosons, as well as the heavy third generation quarks (bottom and top) and tau leptons, have been observed by both ATLAS and CMS, and are consistent with the predictions of the SM at the current level of precision

What to look for

- observing Higgs decays and measuring its couplings to fermions outside the third generation
- decays to a pair of muons with oppotiste change (μ^+,μ^-)
- But this only occurs with small probablility 0.02% (other possibilities are Drell-Yan, top quark or W boson pairs production)
- dimuon invariant mass peak near 125 GeV, only a few GeV wide, determined by

the experimental muon momentum resolution. In contrast, the background events exhibit a smoothly falling mass spectrum in the search region from 110 to 160 GeV

Mentioned in paper

-The large amounts of data collected at colliders like the Large Electron-Positron collider (LEP) or the LHC, and at the intensity frontier, mean that the statistical errors on the collected data samples tend to get quite small, and often the systematic effects become important and even limiting. Experience shows that a large, often dominating amount of time in data analysis is spent on estimating and handling the systematic errors, after the express production of first, exploratory, results.

- scultping of variables , what does it mean?

Bibliography


Decision tree, Apr 2019.


G. Aad and et al.
Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC.
Technical Report 1, 2012.


D. Bourilkov.
Machine and deep learning applications in particle physics.
Technical Report 35, 2019.


S. Chatrchyan and et al.
Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC.
Technical Report 1, 2012.