Overview

Notes

Boosted Decision Trees

Definition (Von Neuman Entropy).

A measure of how much uncertainty there is in ρ regarding which pure state that we have. It also can be used as a measure of entanglement between two particles.

$$S(\rho) = \sum_i^N \lambda_i \log_2 \frac{1}{\lambda i}$$

The goal is to cut the data in such a way to classify, where we maximize the information gained through aking a cut. More specifically, you're maximizing information gain

Selecting Models

Models have inherent inductive bias, but we use that to make any decisions.

Is there an appropriate analogue to the AuC and RoC for when you're doing continuous classification?

You can look at distribution of variables, and then go off of that along with MSE. In reality though, continuous classification doesn't have a meaningful analogue to the RoC

It might be a good idea to tweak the loss function to hate false positives more

Neural Nets

The idea is simple. By Introducing a nonlinearity to the network, we can construct more complex decision boundaries. Where a BDT can only make 1D slices, a neural network can make classification boundaries non-linear and multiple dimensions

Definition 1 (Latent Space).

Starting with one space which describes your data, a latent space is a new representation of the data which having used weights and features distilled has reduced dimensions

In particle physics, is there a good way to interpret the weights on the kernel to better understand what our model is learning off of?

I could try using the lab CNN, on this new data? It should go really fast, given the small dataset

Learning rate tends to matter the most

Graph Neural Nets

How would you train something that looks like a molecule?
One of the inductive biases is that nodes nearby.
GNNs have an attention matrix, to tell us at a given node which one to pay the most attention to

Transformers