Then possibly biased on val set

→ final metrics on another held out dataset

"test set"

Example 1: Gaussian toy model (10d)

$$\begin{cases} b_{ij}\ (0) & p_{b_{ij}}(x) = \mathcal{N}(0,1)^{10} \\ s_{ij}\ (1) & p_{s_{ij}}(x) = \mathcal{N}(0.5,1)^{10} \end{cases}$$

- can compare w/ NP lemma

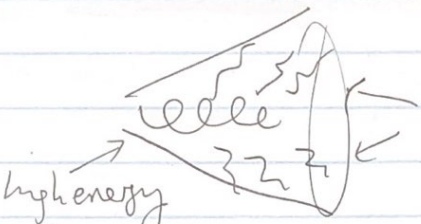
$$P_{NP}(s_{ij}|x) = \frac{p_{s_{ij}}(x)}{p_{b_{ij}}(x) + p_{s_{ij}}(x)}$$

- train simple DNN on samples
  - plot losses
  - ROC curve
  - AUC
  - classifier output

- Jupyter nb demo

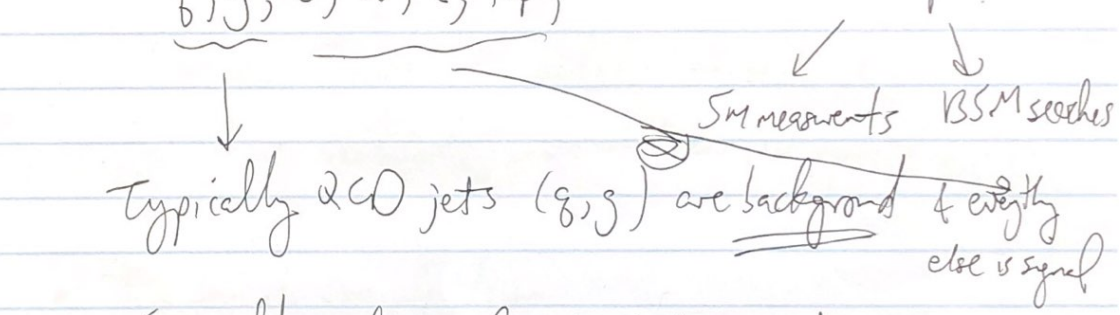
Example 2: Jet classification @ LHC

- Jets are essential objects @ LHC
  Product of strong interaction (QCD) → showers of hadrons

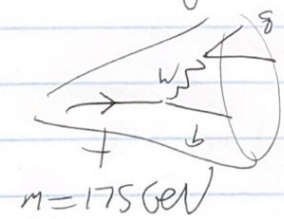


↓ fills out a cone of energy "jet"

- Many different flavors of jets — important to tell them apart!
  $q, g, b, W, Z, top, \ldots$



- Especially rich example: boosted top jets



$m = 175$ GeV
Heaviest ptcl in SM

- Decays to bW
  W can decay to quarks
- If top is very energetic → boosted decay products collimated "fat jet"

- Top jet has substructure → expect 3 "prongs" of energy $\approx b, q, q'$
  - → expect $m_{jet} \approx m_{top}$
  - → expect $q q'$ mass $\approx m_W$.
  - → expect b jet (won't use here) or its own w/ special properties

Note: QCD xsec thousands times more than top (or more) dependson energy (~1000x @ $m_{jj} \sim 1$ TeV)

- Meanwhile QCD jets are relatively structureless.

→ can use these properties to classify tops vs QCD!

(unlike Gaussian toy model, here we do not know true likelihoods, so do not know NP optimal classifier)

- Jupyter nb example dems
  - (m_j, m_W, τ_{32}) → cuts
    → DNN
  - jet constituents → DNN
  - ⋮

Metric: R30
1/fpr @ tpr = 30%
more informative than AUC/ACC