Wednesday, March 24, 2021 10:06 AM

Lecture 16

Decorrelated Classifiers

Suppose we have a binary classifier $\vec{x} \longrightarrow f_{cl}(\vec{x})$

$\vec{x}$: features; $f_{cl}(\vec{x})$: class probability (1 sig, 0 bg)

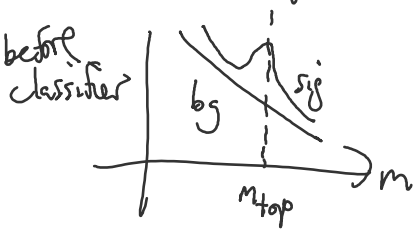
Suppose we have a "protected" feature m that we want f_{cl} to be statistically independent of.

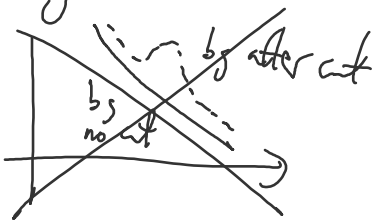
How to accomplish this?

Cannot in general just "exclude" m from training features $\vec{x}$.

$f_{cl}(\vec{x})$ becomes random variable

motivation: want use m for sideband bg. estimation



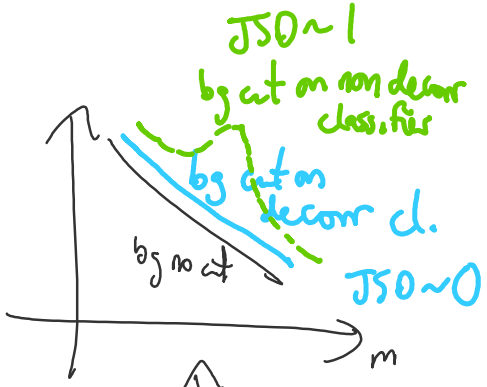


Wednesday, March 24, 2021 10:34 AM

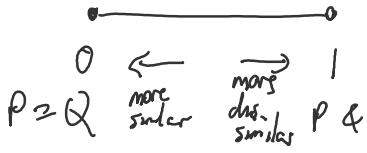
Metrics for decorrelation: $JSD(P,Q) = \frac{1}{2}\left(KL\left(P \,||\, \frac{P+Q}{2}\right) + KL\left(Q \,||\, \frac{P+Q}{2}\right)\right)$

$KL(P||M) = \int dx\, P(x) \log \frac{P(x)}{M(x)}$.

measure of similarity btw two distributions



$0 \le JSD \le 1$ ← if JSD computed w/ $\log_2$



$P = Q$ P & Q no overlap at all

choose a cut → e.g. cut at 50% signal efficiency $JSD50$.

fpr ↓ $\frac{1}{fpr} = R50$