16th Aug — PS625 notes

- NN is a system of nodes and edges. In training, the error difference changes the weights of edges till it converges.

- Supervised ML
  1. Classification — labels are discrete
  2. Regression — output is a point on a spectrum

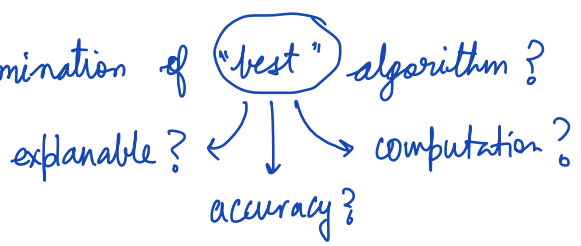
⇒ How are annotations obtained for training dataset?

1. Manually labelled (need multiple annotators for subjective labels)

   " Annotator agreement is necessary, but not sufficient for a good dataset "

2. Natural sources, eg - temperature

⇒ Determination of "best" algorithm?

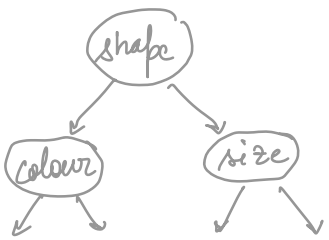


⇒ Bias-variance tradeoff — enough variance to make model "not brittle"

Classification

① Decision trees : find the feature which divides the dataset closest to the actual answer

ass^{m}: features don't affect eachother



decreasing order of usefulness of a feature

② Bayesian classifier : ass^{m}: linearity precondition, ie, all features return points on a linear scale

↑

true for decision trees, naive-bayes, KNN

③ K-nearest neighbours

Regression

① Linear

② Logistical

Unsupervised ML

- Clustering – find internal patterns & natural groups
  user input reqd: how many clusters?
- Chinese restaurant / Indian buffet — ways to find distribution over data
- Topic modelling
- NNs in unsupervised ML → Generative adversarial network (GAN)