**Machine Learning Notes**

**Information Entropy:** Basic Quantity in info theory associated to any random variable.

* Entropy is a measure of disorder
* Entropy measures between 0 and 1

**Information Gain:**  Used to measure the reduction of this disorder in our target variable given additional info about it.

IG (Y, X) = E(Y)- E(Y|X)

* The greater the reduction in this uncertainty, the more info is gained about Y form X.
* *It is mainly used to determine which feature to split their nodes on to get closer to predicting the target variable with each split and also to determine when to stop splitting the tree.*

**Occam’s razor:** When presented with competing hypotheses that make the same predictions, one should select the solution with the fewest assumptions.