

■ Day 3 – Data Science Math Gym

Entropy & Jensen's Inequality (Feynman Style + Visuals)

■ Concept 1: Entropy

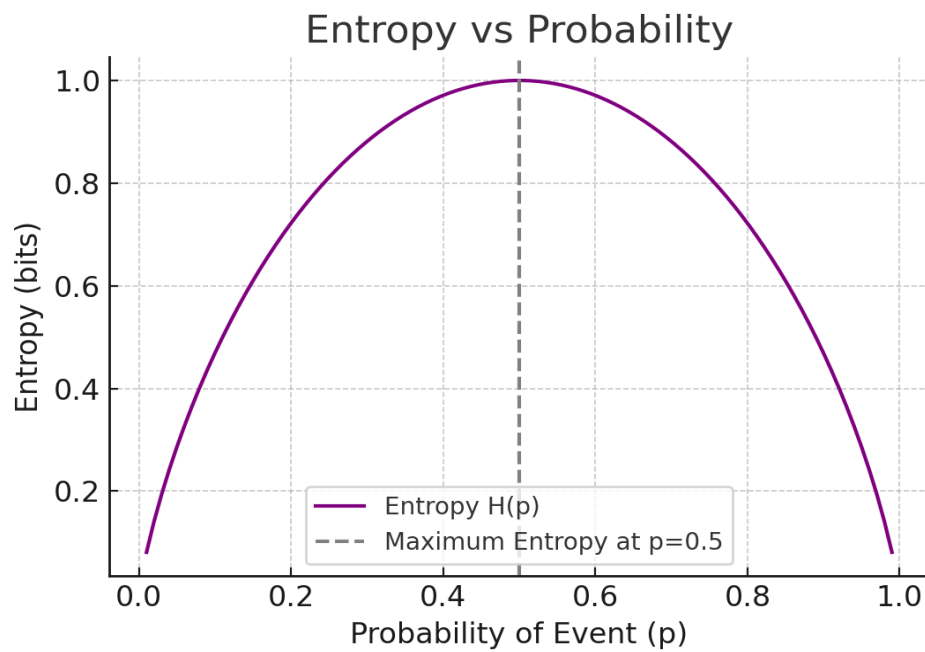
Entropy is the amount of uncertainty or "surprise" in a system. If something is very predictable (like a coin always landing heads), entropy is low. If it's unpredictable (like a fair coin flip), entropy is high. Entropy is used in decision trees, cross-entropy loss, and measures how messy or uncertain a dataset is. Formula: $H(P) = -\sum p(x) \log(p(x))$

■ Concept 2: Jensen's Inequality

Jensen's Inequality says that when you apply a convex function to an average, it will always be less than (or equal to) the average of the function applied to individual values. It's a way of saying "convex curves bend upwards" — and that has major consequences in optimization, machine learning loss functions, and variational inference. Formula: $f(E[X]) \leq E[f(X)]$ for convex functions f .

■ Visual Examples

■ Entropy vs. Probability



■ Jensen's Inequality

