

Starts from Entropy

Ruichen Wang

November 9, 2018

Abstract

Mainly about entropy and a brief introduction of its related algorithms.
Including logistic regression, variational auto encoder ...

Contents

1 What is entropy?

1.1 Def of information

$$I(x) = -\log P(x)$$

Given x is some event, $P(x)$ is probability which event x happens.

1.2 Why using $-\log P(x)$?

The reason we pick $-\log$ as the transform function can be explained as follows:

- Lower probability, higher information
- Higher probability, lower information
- Multi-event happens, the probability is multiplied. the information is summed

Mathematics Explanation

As we know, $P(x) \in [0, 1]$, and larger $P(x)$ should have smaller information.

$$P(x_1, x_2) = P(x_1) * P(x_2)$$

$$\log P(x_1, x_2) = \log P(x_1) + \log P(x_2)$$

2 Average of Information(Entropy)