**Assignment #2 Basic Concepts in Information Theory**

yuweic3

2017/9/11

**#1 Entropy**

1. ,
2. Maximum entropy is achieved when all 6 words equally appears in the article, while the minimum entropy occurs when one word is certain and the others are impossible in the article.
3. The maximum value of article is , when , , and article has the same length with article .

**#2 Conditional Entropy and Mutual Information**

1. when and are independent. At this same, would be the same as the original entropy of and knowing does not help at all in resolving the uncertainty of .

**#3 Mutual Information of Words**

1. ,
2. According to the table, we compute probabilities:

,

,

when :

,

when :

, ,

then we compute

Finally,

Similarly, we get

1. My intuition is that “Compared with baseball, program is more related with computer”. And yes, the result confirms my intuition as . The larger the is, the more correlative and are.

**#4 Kullback-Leibler Divergence**

When distributions are equal,

1. Divergence is not symmetric.

In most cases, and is not equal, so If and only if , that .

1. According to the definition of Kullback-Leibler divergence, only if implies for all . In other words, if there is that but , we cannot define regularly. That is, the divergence is infinite.

In such case, we can either use Bayesian prior to smooth the distribution, or practically just give up those values that does not make sense.