Introduction to Information Theory

Complexity Explorer 2017

Unit 10.1 Homework Solutions

1. To compute the information theoretic quantities, we first compute the marginal and joint probabilities using the data in the CSV file.

   In terms of marginal probabilities, out of 31 days:
   There were 18 days when it was hot, so $P(X=x_{hot}) = 18/31$ and $P(X=x_{not\text{-}hot}) = 13/31$.
   There were 12 days when it was rainy, so $P(Y=y_{rain}) = 12/31$ and $P(Y=y_{no\text{-}rain}) = 19/31$.
   (August is the monsoon season in Santa Fe, NM, so it is actually the rainiest time of the year).

   In terms of joint probabilities, out of 31 days:
   There were 7 days when it was both hot and rainy, so $P(X=x_{hot}, Y=y_{rain}) = 7/31$.
   There were 11 days when it was hot and not rainy, so $P(X=x_{hot}, Y=y_{no\text{-}rain}) = 11/31$.
   There were 5 days when it was not hot and not rainy, so $P(X=x_{not\text{-}hot}, Y=y_{no\text{-}rain}) = 5/31$.
   There were 8 days when it was not hot and rainy, so $P(X=x_{not\text{-}hot}, Y=y_{rain}) = 8/31$.

   Finally, we are ready to compute our information theoretic quantities. The information in X and Y are respectively:
   $I(X) = -(18/31) \log_2 (18/31) - (13/31) \log_2 (13/31) \approx 0.981$ bits
   $I(Y) = -(12/31) \log_2 (12/31) - (19/31) \log_2 (19/31) \approx 0.963$ bits

   The joint information:
   $I(XY) = -(7/31) \log_2 (7/31) - (11/31) \log_2 (11/31) - (5/31) \log_2 (5/31) - (8/31) \log_2 (8/31) \approx 1.944$ bits

   The conditional information quantities are
   $I(X|Y) = I(XY) - I(Y) \approx 1.944 - 0.963 = 0.981$ bits
   $I(Y|X) = I(XY) - I(X) \approx 1.944 - 0.981 = 0.963$ bits

   The mutual information is
   $I(X{:}Y) = I(X) + I(Y) - I(XY) \approx 0.98 + 0.96 - 1.94 = 0$ bits

Let X indicate the state of the two binary inputs, and Y indicate the binary output. There are four possible input values: X=(0,0), X=(0,1), X=(1,0), and X=(1,1), which we assume are distributed uniformly.

Our goal is to calculate the mutual information between X and Y. We use the formula $I(X{:}Y) = I(X) + I(Y) - I(XY)$.

I(X) is the information in the inputs. Since there are 4 possible input values, and they are uniformly distributed, we can compute

$I(X) = -(1/4) \log_2(1/4) -(1/4) \log_2(1/4) -(1/4) \log_2(1/4) -(1/4) \log_2(1/4) = 2$ bits

I(XY) is the joint information in the input and output values. To calculate these, we need to compute the joint probabilities. Since Y is equal to 1 only when X=(1,1), and is 0 otherwise, we can write

P(X=(0,0), Y=0) = 1/4
P(X=(0,0), Y=1) = 0 (remember, it is only possible for Y to equal 1 when X=(1,1))
P(X=(0,1), Y=0) = 1/4
P(X=(0,1), Y=1) = 0
P(X=(1,0), Y=0) = 1/4
P(X=(1,0), Y=1) = 0
P(X=(1,1), Y=0) = 0
P(X=(1,1), Y=1) = 1/4

The joint information is then

$I(XY) = -(1/4) \log_2 (1/4) -(1/4) \log_2 (1/4) - (1/4) \log_2 (1/4) - (1/4) \log_2 (1/4) = 2$ bits

(for simplicity, here we have dropped terms involved 0 probabilities from the calculation.)

Finally, we compute the marginal probabilities of Y,

P(Y=0) = P(X=(0,0), Y=0) + P(X=(1,0), Y=0) + P(X=(0,1), Y=0) + P(X=(1,1), Y=0) = 3/4
P(Y=1) = P(X=(0,0), Y=1) + P(X=(1,0), Y=1) + P(X=(0,1), Y=1) + P(X=(1,1), Y=1) = 1/4

This gives the information in Y as

$I(Y) = -(3/4) \log_2 (3/4) -(1/4) \log_2 (1/4) \approx 0.81$

Combining gives

$I(X:Y) = I(X) + I(Y) - I(XY) = 2 + 0.81 - 2 = 0.81$