Mutual Information I(X;Y): Measuring Entropy Reduction

How much reduction in the entropy of X can we obtain by knowing Y?

Mutual Information: $\mathbf{I(X; Y)= H(X) – H(X|Y) = H(Y)-H(Y|X)}$

Properties:

- Non-negative: $I(X;Y)\geq 0$
- Symmetric: $I(X;Y)=I(Y;X)$
- $I(X;Y)=0$ iff X & Y are independent

When we fix X to rank different Ys, I(X;Y) and H(X|Y) give the same order but I(X;Y) allows us to compare different (X,Y) pairs.