

# Information Theory: Our Learning Map



What will we learn to measure?

Aarav

**Probability** is a number from 0 to 1 that expresses a chance. 0.5 means 50%.

Uncertainty, and the quality of probability predictions.

Maya



## 1 Uncertainty of a fair coin

We will measure how uncertain a result is.

Before the coin lands



Heads  
50%

Tails  
50%



Which side will land?

We cannot know this result in advance. We will measure this uncertainty.



## 2 Scoring a probability prediction

We will score the probability given to the answer that happened.

After we know the answer



Observed answer  
Cat

Model A

90%  
for cat

Model B

60%  
for cat



Both choose cat. We will also score the chance each gave to the observed answer.

## 3 Compare two probability reports

We will measure the extra average prediction loss from this mismatch.

Reference chances

A  75%  
B  25%

Our reported chances

A  50%  
B  50%



Loss is a number used to score prediction error. We will measure the extra average loss from this mismatch.

## 4 Use a helpful clue

We will measure how knowing the weather reduces uncertainty on average.

Weather helps predict umbrella



On rainy days:  
umbrella on  
3 of 4 days



Across all days:  
umbrella on  
4 of 8 days

Uncertainty about umbrella

**Before seeing weather**

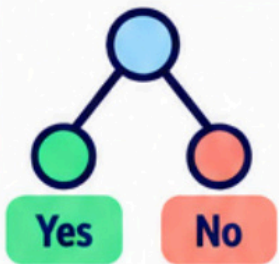
We are uncertain about the umbrella choice.

**After seeing weather**

We are less uncertain about the umbrella choice.

## Real-world applications

Decision tree splits



Choose a useful yes/no question that reduces uncertainty.

Predict the next word



Score the chance given to the next word.

Image reconstruction



Train a model to rebuild an image.