Lecture 7: Decision Trees

Aru Gyani

Based on Prof. Rishabh Iyer's slides.

Part I

Probability, Random Variables, and Entropy

1 Probability

The following content is sourced from the following:

- seeing-theory
- Khan Academy
- Professor Rishabh Iyer's class notes

1.1 Discrete Probability

Definition 1.1 Sample space

Sample space specifies the set of possible outcomes.

For example, $\Omega = \{H, T\}$ would be the set of possible outcomes of a coin flip. Simply put, we are saying a coin flip could be either *heads* or *tails*.

Definition 1.2 Probability

Probability is simply how likely something is to happen

For each element ω inside of our sample space Ω, there is a number $p(\omega) \in [0, 1]$ called a probability. This represents how likely the event ω is to happen.

The total probability of all possible outcomes (each outcome being an element ω) within the sample space Ω adds up to 1. This means that one of the outcomes in the sample space is certain to occur.

$$\sum_{\omega \in \Omega} p(\omega) = 1$$

For exaxmple, a biased coin might have $p(H) = .6$ and $p(T) = .4$.

Note: $[0, 1]$ is a range, not a set. This number is always between 0 and 1, where 0 indicates impossibility and 1 indicates certainty.

Definition 1.3 Event

An **event** is a subset of the sample space Ω, e.g.

- Let $\Omega = \{1, 2, 3, 4, 5, 6\}$ be the 6 possible outcomes of a dice roll.
- $A = \{1, 5, 6\} \subseteq \Omega$ would be the event that the dice roll comes up as a one, five, or six.

The probability of an event is just the sum of all the outcomes that it contains.

$$p(A) = p(1) + p(5) + p(6)$$

Note: $A \subseteq \ \Omega$ means that A is a subset of Ω. Below, is a brief recap of what a subset is.

Definition 1.4 Subset

A set A is a **subset** of another set B if all of elements of the set A are elements of the set B. In other words, the set A is contained inside the set B.

1.2 Conditional Probability

1.2.1 Independence

We say two events are independent if knowing one event occurred doesn't change the probability of the other event. For example, the probability that a fair coin shows "heads" after being flipped is $\frac{1}{2}$. What if we knew the day was Tuesday? Does this change the probability of getting "heads"? Of course not. The probability of getting "heads", given that it's a Tuesday, is still $\frac{1}{2}$. Thus, the result of a coin flip and the day being Tuesday are independent events. More reading .

Definition 1.5 Independence

Two events A and B are independent if

$$p(A \cap B) = p(A)P(B)$$

Informally, $p(A \cap B)$ indicates the probability of A and B (aka probability of A intersection B). This means that it indicates the likelihood of both events happening.

Let's suppose that we have a fair die: $p(1) = \cdots = p(6) = \frac{1}{6}$ (the probability of rolling any number on the dice is equal). If $A = \{1, 2, 5\}$ and $B = \{3, 4, 6\}$ are A and B independent?

2 Discrete Random Variables

3 Entropy

3.1 Conditional Entropy

3.2 Information Gain

Part II
Decision Trees

Definitions