PROBABILITY : The Math of Machine Learning

Math of Machine Learning & Data Analytics

Here I'll begin covering the math of machine learning. Rather then cover machine learning by simply showing you some examples using some popular libraries, I will cover it in depth so that you have a complete understanding of the topic.

Probability

Probability is the branch of mathematics that focuses on how likely an event is to occur. With probability we'll use 1 as the example that we are certain something will occur. We use 0 to mean we are certain something won't occur. Decimals between those 2 values are used for all other probabilities.

Probability is defined as P(A) where P is probability and A is the event we are checking. To find the probability of an even we divide the preferred event by the sum of all other possible events.

A coin flip represents a clear probability. It is either heads or tails and the probability is .5. Here is the code to simulate that.

CODE

```
# Create the list
flip_list = []

# Populate the list with 100 Hs and Ts
# Trick : random.choice() returns a random value from the list
for i in range(1, 101):
    flip_list += random.choice(['H', 'T'])

# Output results
print("Heads : ", flip_list.count('H'))
print("Tails : ", flip_list.count('T'))
```

A single die roll is equally simple with a probability of 1/6 or .167.

A double dice roll is where things start to get interesting. Here we can see that the odds of getting a 2 is 1 in 36, while a 7 is 6 in 36. This occurs because there are just more ways to roll a 7 with 2 dice. This is known as a Probability Frequency Distribution Table. It is a collection of all probabilities.

Let's see how this translates into code.

CODE

```
dice_list = []
for i in range(1, 501):
    dice_list.append((randrange(1, 7) + randrange(1, 7)))
for i in range(2, 13):
    print(f"{i} : {dice_list.count(i)} : {dice_list.count(i) / 500:.2f}")
```

A Complement of an Event is everything it is not. It is defined with the symbol A'. The sum of your event probabilities must equal 1 So the probability that you won't roll a 1 is the sum of all other probabilities. With dice the probability that you will roll something other than a 1 is 5/6th.

Logical Operations with Events

We can combine events using logical operators to analyze other events. 2 AND 1 gives a probability of 2/36th or 0.06.

2 OR 1 gives a probability of 20/36 or .56. There is a formula for calculating OR conditions. P(A OR B) = P(A) + P(B) - P(A AND B). So to calculate P(2 OR 1)

1. P(2 OR 1) = P(2) + P(1) - P(2 AND 1)
2. P(2) = 11/36
3. P(1) = 11/36
4. P(2 AND 1) = 2/36
5. P(2 OR 1) = 11/36 + 11/36 - 2/36 = 20/26

And, NOT 2 provides a probability of 25/36 or 0.69. Sometimes we can calculate nots by subtracting from 1. Suppose we have P(NOT 6 AND 6). We could find that probability this way.

1. P(6 AND 6) = 1/36
2. P(NOT 6 AND 6) = 36/36 - 1/36 = 0.97

Combinatorics

Combinators is concerned with counting the number of ways items can be ordered. There are 3 main parts of combinatorics. Permutations define the number of ways we can arrange a set of elements. Variations focuses on arranging a number of elements into a limited number of spaces. And, Combinations focuses on the number of different ways we can pick certain elements. Through examples these will make sense.

Permutations

Permutations are concerned with how many ways we can arrange a set of items. This value is the factorial of the number of items. The factorial of n = 1 * 2 * 3 ... * n. So the total number of arrangement options for 2 items is 2. The total for 3 is 1 * 2 * 3 or 6.

Factorial Trick

To save time when calculating factorials if the numerator is greater than the denominator you can cancel out multiplying like values in each.