HW1 CS464-1  
Görkem Kadir Solun 22003214

# 1 Probability Review

## Q1.1

The event that a blue coin is chosen  
 The event that the red coin is chosen  
 The event that the yellow coin is chosen  
The probability that box 1 is chosen  
 Event that box 2 is chosen

What we want  
Event that tail comes twice

The event that blue coin landed tail.

The event that yellow coin landed tail.

The event that red coin landed tail.

## Q1.2

Blue coins are fair, so let's only check blue coins.

## Q1.3

We check red this time.

# 2 Amazon Reviews Classification

## Question 3.1

### 1. What are the percentages of each category in the y\_train.csv y\_test.csv? Draw a pie chart showing percentages.

A pie chart with different colored circles

Description automatically generated

### 2. What is the prior probability of each class? Write your answer to the report.

TODO

### 3. Is the training set balanced or skewed towards one of the classes? Do you think having an imbalanced training set affects your model? If yes, please explain briefly how it can affect the model.

TODO

### 4. How many times do the words ”good” and ”bad” appear in the training documents with the label ”positive”, including multiple occurrences, and what are the log ratio of their occurrences within those documents, i.e, ln(P(good|Y = positive)) and ln(P(bad|Y = positive))?

TODO