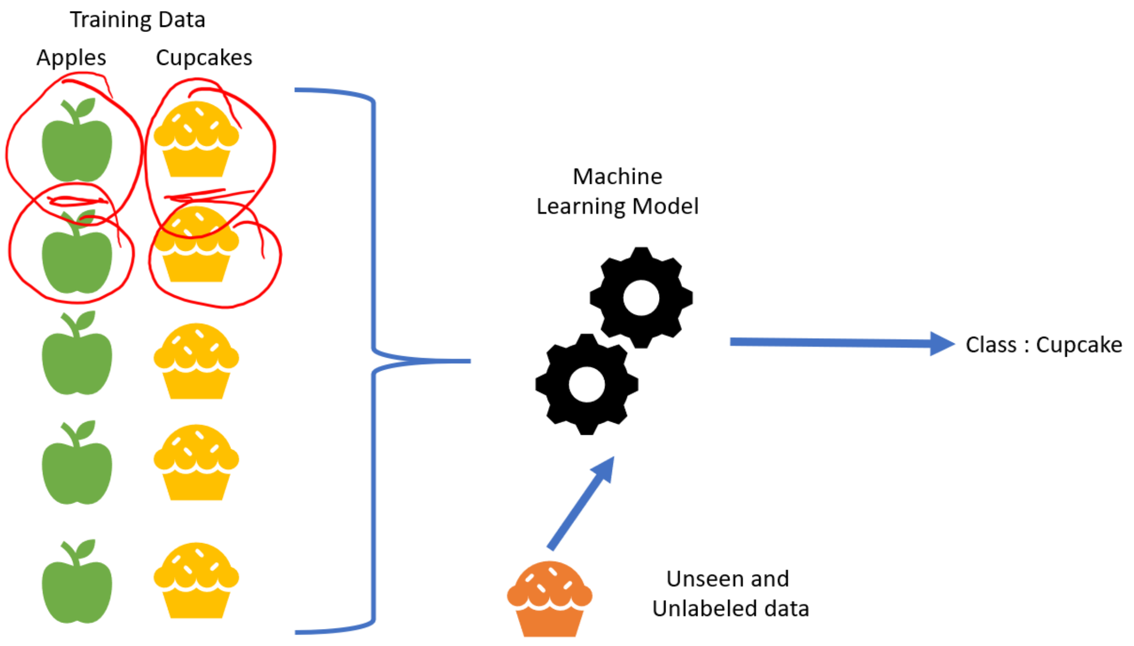
# Machine Learning

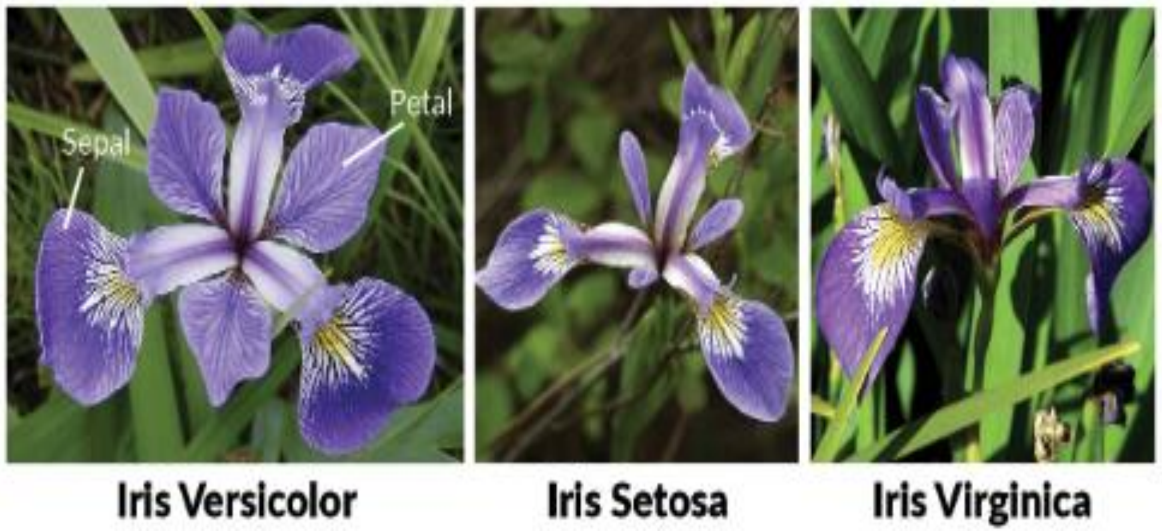
Dr. Muhammad Adeel Nisar

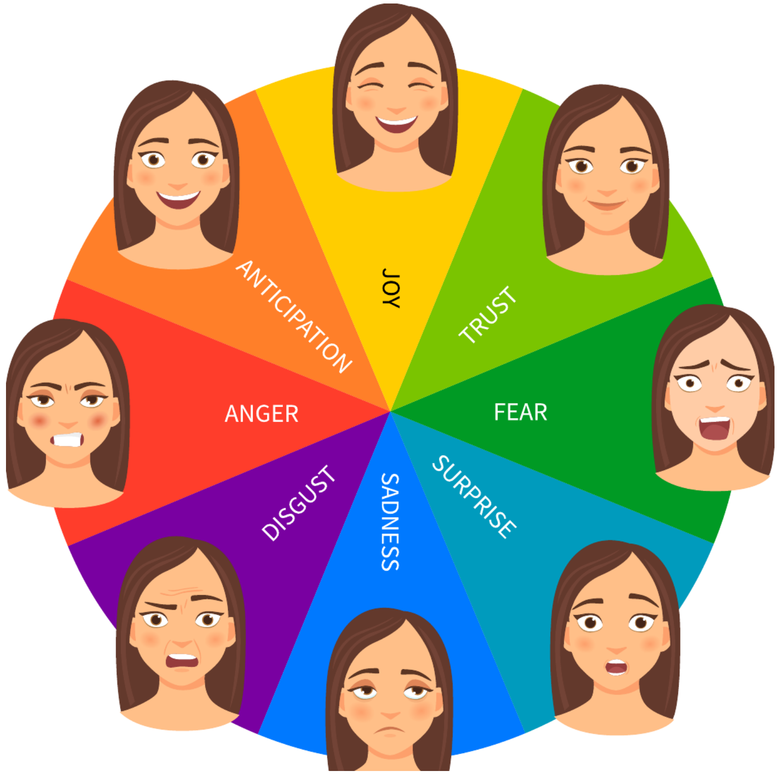
Assistant Professor – Department of IT,
Faculty of Computing and Information Technology,
University of the Punjab, Lahore

# Classification

# Classification Examples



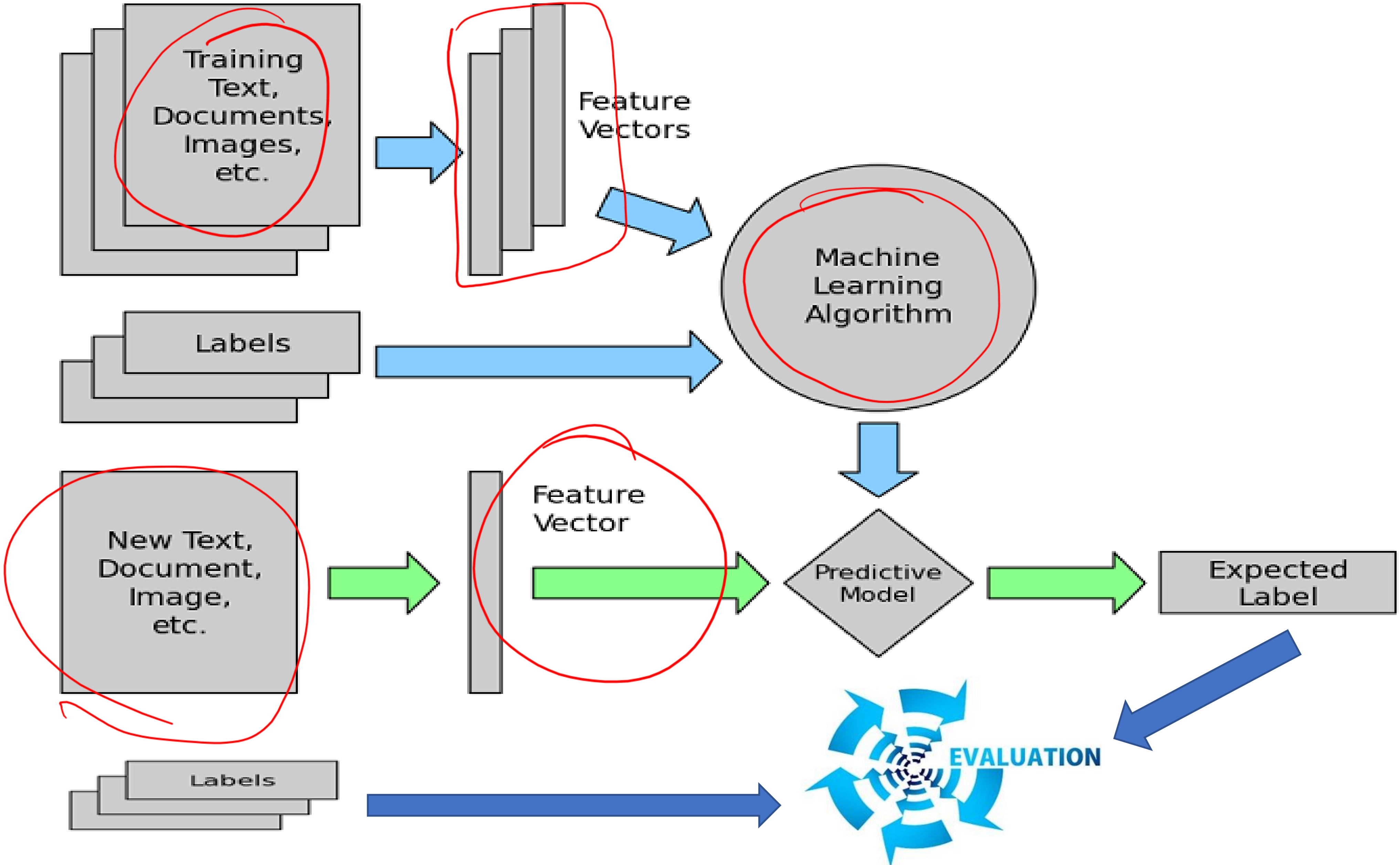
# Classification Algorithms

- Classification algorithms are used when the output variable is categorical, which means there are two or more classes.
- Algorithms
  - Naïve Bayes
  - Logistic Regression
  - Support vector Machines
  - Random Forest
  - Decision Trees

# How to Perform Classification

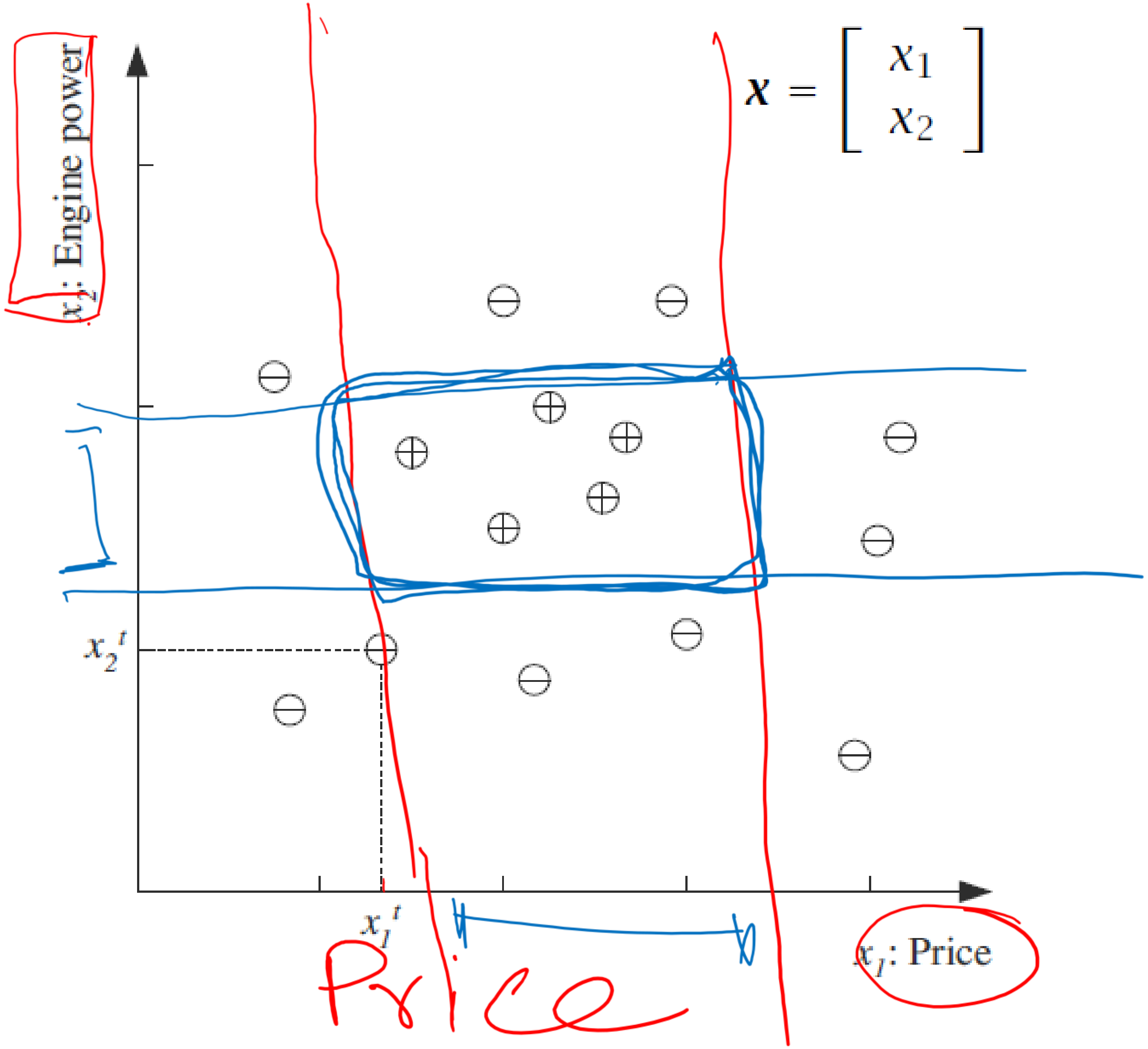
Training Data

Testing Data

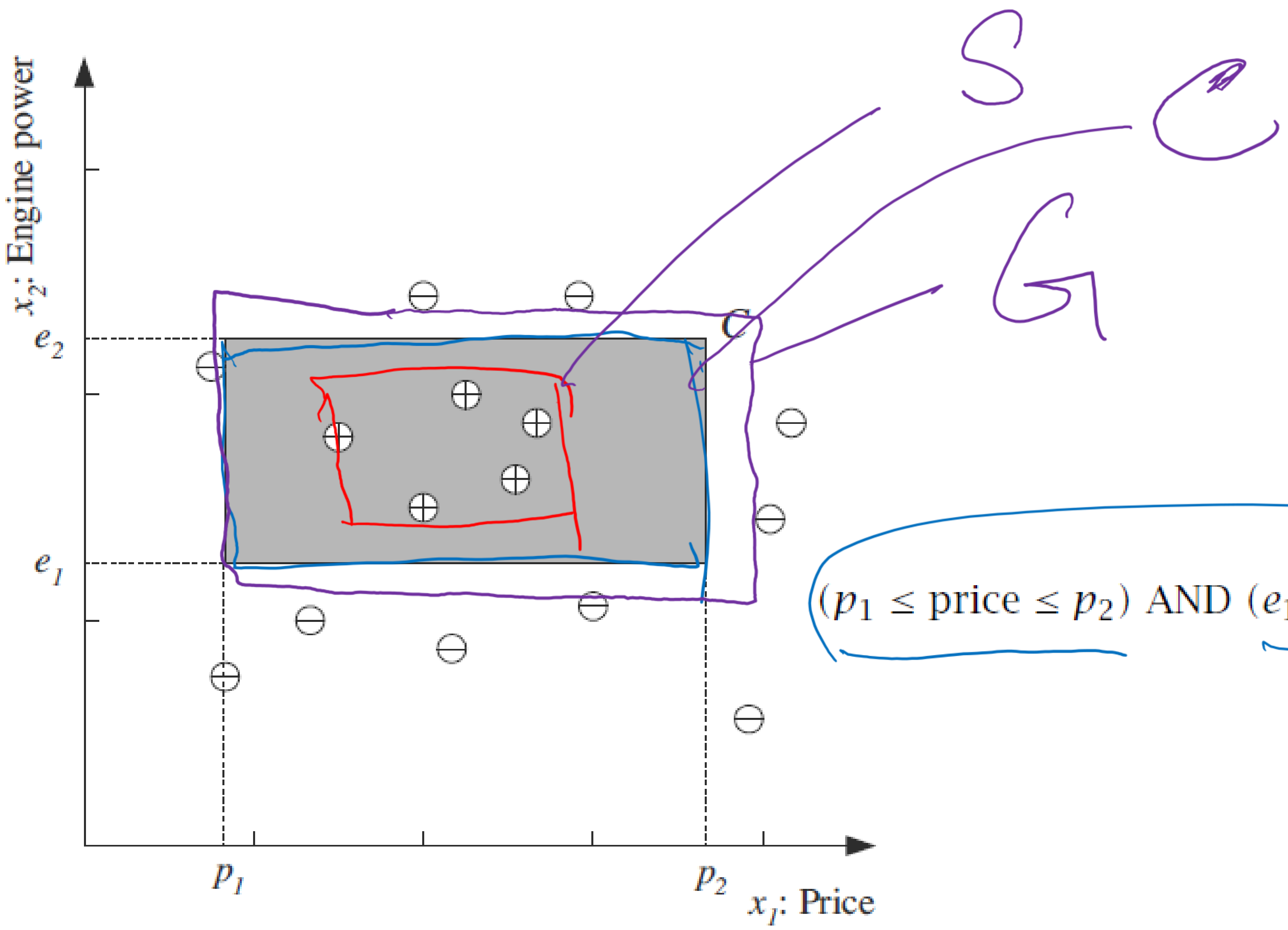


# Classification of a Family Car

Engine Power

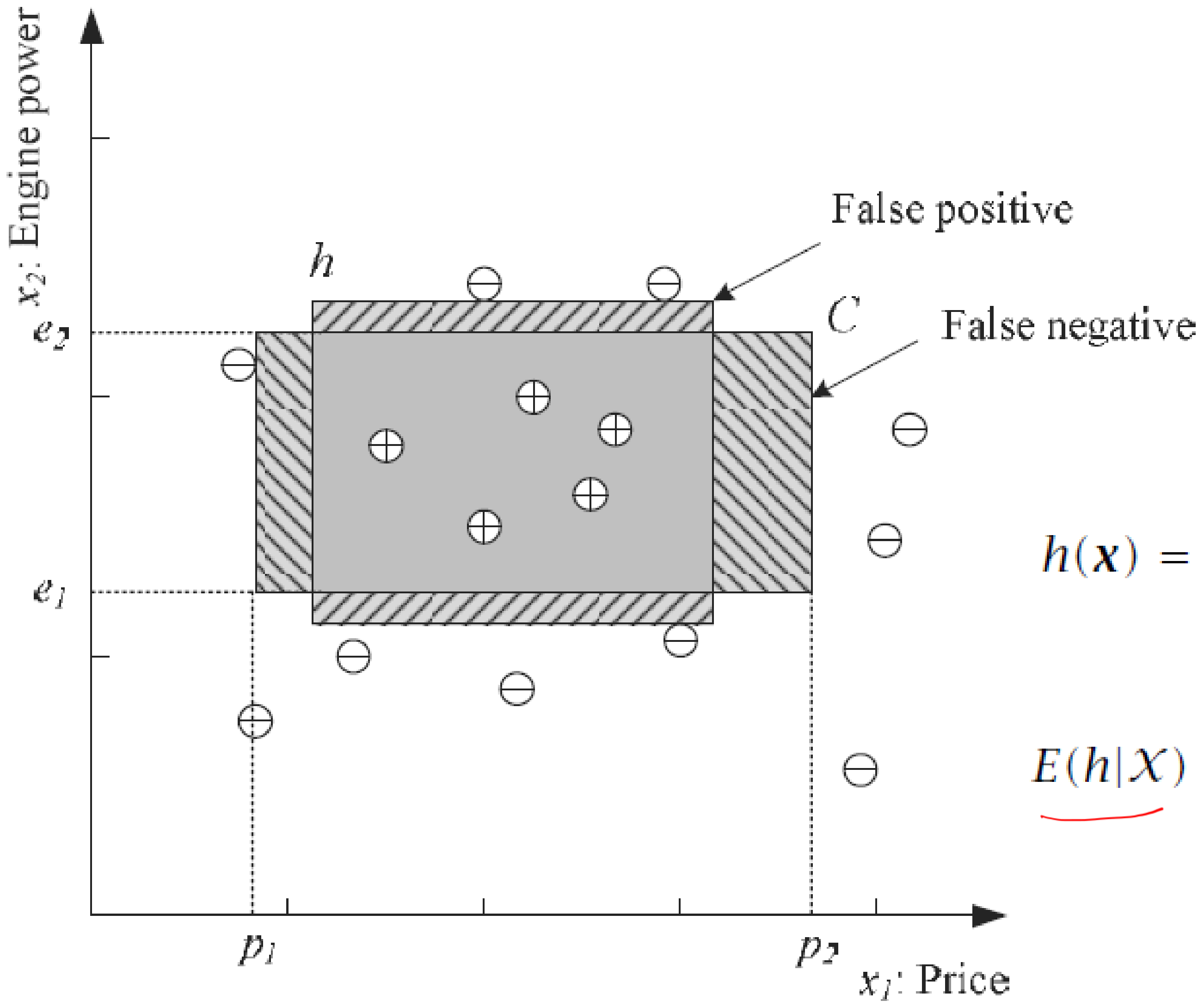


# Classification of a Family Car




$(p_1 \leq \text{price} \leq p_2)$ AND $(e_1 \leq \text{engine power} \leq e_2)$

# Classification of a Family Car



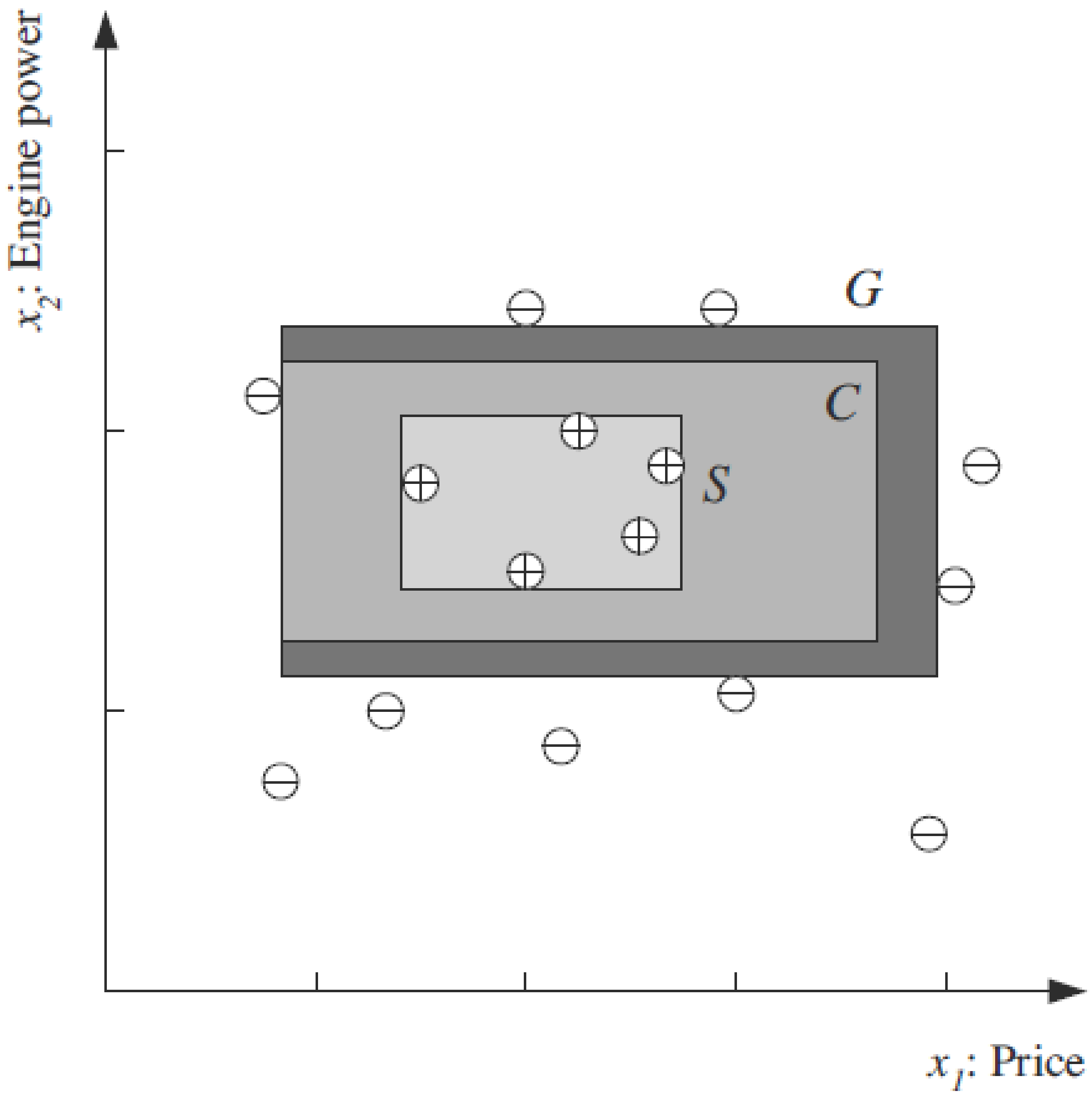

$$r = \begin{cases} 1 & \text{if } \boldsymbol{x} \text{ is a positive example} \\ 0 & \text{if } \boldsymbol{x} \text{ is a negative example} \end{cases}$$

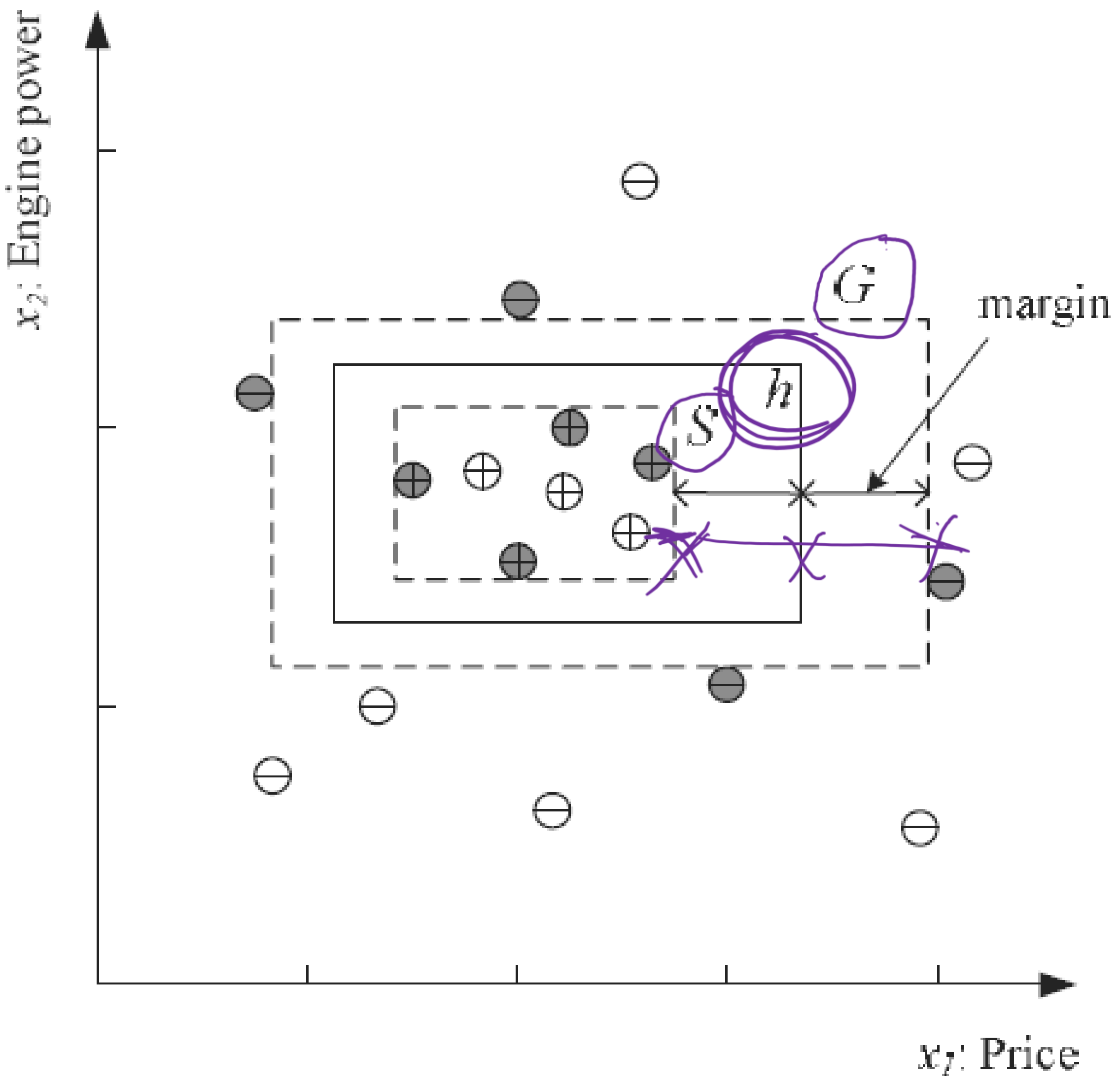
$$\mathcal{X} = \{\mathbf{x}^t, r^t\}_{t=1}^N$$

$$h(\boldsymbol{x}) = \begin{cases} 1 & \text{if } h \text{ classifies } \boldsymbol{x} \text{ as a positive example} \\ 0 & \text{if } h \text{ classifies } \boldsymbol{x} \text{ as a negative example} \end{cases}$$

$$E(h|\mathcal{X}) = \sum_{t=1}^{N} 1(h(\mathbf{x}^t) \neq r^t)$$

# Classification of a Family Car



# Evaluation of a Classifier

| Car | Actual | Predicted |
|---|---|---|
| **1** | FC | FC |
| **2** | FC | NFC |
| **3** | NFC | FC |
| **4** | FC | FC |
| **5** | NFC | FC |
| **6** | NFC | NFC |
| **7** | FC | NFC |
| **8** | FC | FC |
| **9** | NFC | FC |
| **10** | FC | FC |
| **11** | NFC | NFC |
| **12** | NFC | NFC |

FC : Family Car
NFC: Not a Family Car

# Model Evaluation Measures

- **True positives (TP):** data points predicted/labeled as positive that are actually positive
- **False positives (FP):** data points predicted/labeled as positive that are actually negative
- **True negatives (TN):** data points predicted/labeled as negative that are actually negative
- **False negatives (TN):** data points predicted/labeled as negative that are actually positive

# Model Evaluation Measures

**Confusion Matrix**

| | | Actual | |
|---|---|---|---|
| | | **Positive** (FC) | **Negative** (NFC) |
| **Predicted** | **Positive** (FC) | 5 TP | FP 2 |
| | **Negative** (NFC) | 2 FN | 3 TN |

| Car | Actual | Predicted |
|---|---|---|
| 1 | FC | FC |
| 2 | FC | FC |
| 3 | NFC | FC |
| 4 | FC | FC |
| 5 | NFC | FC |
| 6 | NFC | NFC |
| 7 | FC | NFC |
| 8 | FC | FC |
| 9 | FC | FC |
| 10 | FC | NFC |
| 11 | NFC | NFC |
| 12 | NFC | NFC |

# Model Evaluation Measures

| Confusion Matrix | | Actual | |
|---|---|---|---|
| | | Positive (FC) | Negative (NFC) |
| Predicted | Positive (FC) | 5 | 2 |
| | Negative (NFC) | 2 | 3 |

| Car | Actual | Predicted |
|---|---|---|
| 1 | FC | FC |
| 2 | FC | FC |
| 3 | NFC | FC |
| 4 | FC | FC |
| 5 | NFC | FC |
| 6 | NFC | NFC |
| 7 | FC | NFC |
| 8 | FC | FC |
| 9 | FC | FC |
| 10 | FC | NFC |
| 11 | NFC | NFC |
| 12 | NFC | NFC |

# Model Evaluation Measures

- **Accuracy:** Closeness of a measured value to a standard or known value ((TP + TN)/(TP+TN+FN+FP)) = **8/12** = **0.67**
- **Recall:** Ability of a classification model to identify all relevant instances (TP / TP + FN) = **5/7**
- **Precision:** Ability of a classification model to return only relevant instances (TP / TP + FP) = **5/7**
- **F1 score:** A single metric that combines recall and precision using the harmonic mean ( 2 (Precision * Recall)/(Precision + Recall)) **= ?**

# Example 2

| Confusion Matrix | | Actual | |
|---|---|---|---|
| | | Positive | Negative |
| Predicted | Positive | | |
| | Negative | | |

- Accuracy = ?
- Recall = ?
- Precision = ?
- F1-Score = ?

| Images | Actual Label | Predicted Label |
|---|---|---|
| Image1 | Happy | Sad |
| Image2 | Sad | Sad |
| Image3 | Sad | Happy |
| Image4 | Happy | Happy |
| Image5 | Happy | Happy |
| Image6 | Sad | Happy |
| Image7 | Sad | Happy |
| Image8 | Happy | Happy |

# Example 3

| | | Actual | |
|---|---|---|---|
| Confusion Matrix | | Positive | Negative |
| Predicted | Positive | | |
| | Negative | | |

- Accuracy = ?
- Recall = ?
- Precision = ?
- F1-Score = ?

| Images | Actual Label | Model 1 | Model 2 |
|---|---|---|---|
| Patient1 | Not Corona | Not Corona | Not Corona |
| Patient2 | Not Corona | Not Corona | Not Corona |
| Patient3 | Not Corona | Not Corona | Not Corona |
| Patient4 | Not Corona | Not Corona | Not Corona |
| Patient5 | Not Corona | Not Corona | Not Corona |
| Patient6 | Not Corona | Not Corona | Corona |
| Patient7 | Not Corona | Not Corona | Corona |
| Patient8 | Not Corona | Not Corona | Corona |
| Patient9 | Corona | Not Corona | Corona |
| Patient10 | Corona | Not Corona | Corona |

# Probabilities (recap)

Slides Courtesy: Dr. Kamran Malik

# Probabilities (Recap)

- $P(X = 0) =$
- $P(Y = 3) =$
- $P(X = 1, Y = 2) =$
- $P(Y = 2, X = 1) =$
- $P(X = 1 \mid Y = 2) =$
- $P(Y = 2 \mid X = 1) =$

| X | Y |
|---|---|
| 0 | 0 |
| 0 | 1 |
| 1 | 0 |
| 1 | 2 |
| 2 | 3 |
| 2 | 0 |
| 2 | 3 |
| 1 | 3 |
| 1 | 2 |
| 0 | 3 |
| 0 | 2 |
| 0 | 0 |

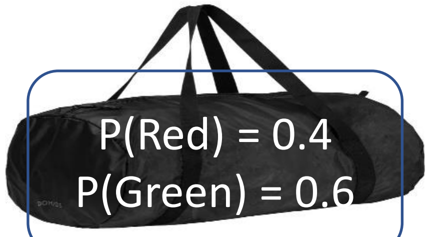
# Probabilities (Recap)

- $P(X = 0) = \mathbf{5/12}$
- $P(Y = 3) = \mathbf{4/12}$
- $P(X = 1, Y = 2) = \mathbf{2/12}$
- $P(Y = 2, X = 1) = \mathbf{2/12}$
- $P(X = 1 \mid Y = 2) = \mathbf{2/3}$
- $P(Y = 2 \mid X = 1) = \mathbf{2/4}$

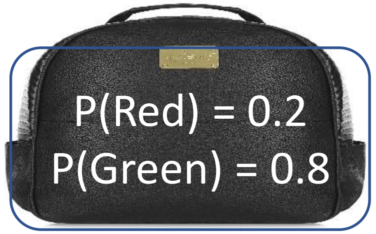
| X | Y |
|---|---|
| 0 | 0 |
| 0 | 1 |
| 1 | 0 |
| 1 | 2 |
| 2 | 3 |
| 2 | 0 |
| 2 | 3 |
| 1 | 3 |
| 1 | 2 |
| 0 | 3 |
| 0 | 2 |
| 0 | 0 |

# Probabilities (Recap)

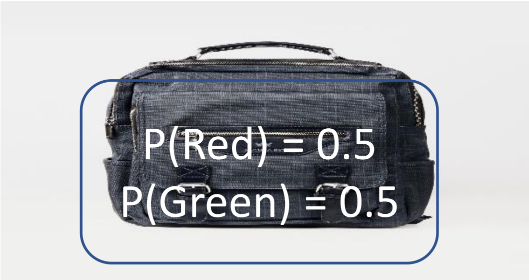
P(bag1) = 0.3 P(bag2) = 0.5 P(bag3)=0.2



P(Red) = 0.4
P(Green) = 0.6

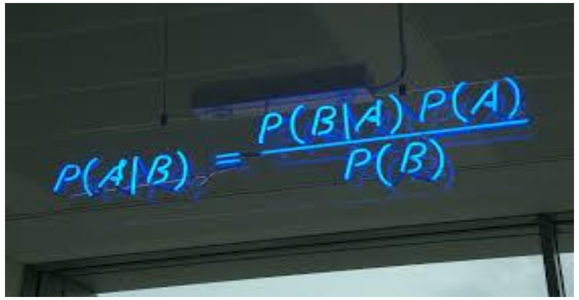


P(Red) = 0.2
P(Green) = 0.8



P(Red) = 0.5
P(Green) = 0.5

- P(bag2) =
- P(Red|bag1) =
- P (Red) =
- P(bag1|Red) =



# Probabilities (Recap)

- $P(A, B)$ or $P(A \cap B) = ?$
- $P(A, B) = P(A \mid B) * P(B) = P(B \mid A) * P(A)$
- $P(A, B \mid C) = P(A \mid B, C) * P(B \mid C) = P(B \mid A, C) * P(A \mid C)$
- $P(A \cap B \cap C) = P(A, B, C)$
- $P(A \cap B \cap C \cap D) = ?$

# Probabilities (Recap)

if all events are independent

- $P(A, B) = P(A) * P(B)$
- $P(A, B, C) =$
- $P(A, B, C, D) =$
- $P(A1, A2, . . ., An) =$

Conditionally independent

- $P(A, B \mid C) = P(A \mid C) * P(B \mid C)$
- $P(A, B, C \mid D) =$
- $P(A1, A2, . . ., An \mid Z) =$

# Homework

- Book 1: 2.1, 2.2, 2.3, 5.6
- Book 2: Start Reading Chapter 3
- Book 3: Chapter 2 (2.1 to 2.5)