# Naive Bayes Spam Classification Training Process

## 1. Data Collection

Gather a dataset of labeled emails. For example:

- Email 1 (Spam): "Get rich quick! Buy our product now!"
- Email 2 (Ham): "Hey friend, want to grab coffee tomorrow?"
- Email 3 (Spam): "Congratulations! You've won a free iPhone!"
- Email 4 (Ham): "Please find attached the report for our meeting."
- Email 5 (Spam): "Enlarge your profits with our business opportunity!"

## 2. Feature Extraction

Select key words as features:

["rich", "buy", "friend", "coffee", "won", "free", "report", "meeting", "business"]

## 3. Training Data Preparation

Create a table showing word occurrence in emails:

| Email | Class | rich | buy | friend | coffee | won | free | report | meeting | business |
|-------|-------|------|-----|--------|--------|-----|------|--------|---------|----------|
| 1 | Spam | 1 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 2 | Ham | 0 | 0 | 1 | 1 | 0 | 0 | 0 | 0 | 0 |
| 3 | Spam | 0 | 0 | 0 | 0 | 1 | 1 | 0 | 0 | 0 |
| 4 | Ham | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 1 | 0 |
| 5 | Spam | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 |

## 4. Calculating Probabilities

### a) Prior Probabilities:

- P(Spam) = 3/5 = 0.6
- P(Ham) = 2/5 = 0.4

**b) Likelihood Probabilities:**

Calculate P(word|Spam) and P(word|Ham) for each word.

Example calculations:

- P("rich"|Spam) = 1/3
- P("rich"|Ham) = 0/2 = 0
- P("friend"|Spam) = 0/3 = 0
- P("friend"|Ham) = 1/2 = 0.5

Apply Laplace smoothing to avoid zero probabilities:

- P("rich"|Spam) = (1+1) / (3+2) = 0.4
- P("rich"|Ham) = (0+1) / (2+2) = 0.25

# 5. Storing the Model

The trained model consists of:

- Prior probabilities: P(Spam) and P(Ham)
- Likelihood probabilities: P(word|Spam) and P(word|Ham) for each word

# 6. Classification Step

To classify a new email:

1. Extract features (words) from the new email.
2. Calculate P(Spam|email) and P(Ham|email) using the stored probabilities:
   P(Spam|email) ∝ P(Spam) * ∏ P(word|Spam) for each word in the email
   P(Ham|email) ∝ P(Ham) * ∏ P(word|Ham) for each word in the email
3. Compare the two probabilities:
   - If P(Spam|email) > P(Ham|email), classify as Spam
   - Otherwise, classify as Ham

Example: New email: "Quick cash opportunity"

P(Spam|email) ∝ 0.6 * P("quick"|Spam) * P("cash"|Spam) * P("opportunity"|Spam)
P(Ham|email) ∝ 0.4 * P("quick"|Ham) * P("cash"|Ham) * P("opportunity"|Ham)

Calculate these values using the stored probabilities and compare them to determine the classification.

Note: In practice, we often work with log probabilities to avoid underflow issues with very small numbers.