In [1]:
```
#task1
from sklearn.model_selection import train_test_split
from sklearn.feature_extraction.text import TfidfVectorizer
from sklearn.naive_bayes import MultinomialNB
from sklearn.metrics import accuracy_score

# Expanded dataset for better splitting
data = {'label': ['ham', 'spam', 'ham', 'spam', 'ham', 'spam', 'ham', 'spam', 'ham', '
        'text': ['Hello', 'Buy now', 'How are you?', 'Free entry!', 'Call me', 'Winner
                 'See you', 'Exclusive deal']}

labels = [1 if label == 'spam' else 0 for label in data['label']]
X_train, X_test, y_train, y_test = train_test_split(data['text'], labels, test_size=0.
vectorizer = TfidfVectorizer()
X_train_vec = vectorizer.fit_transform(X_train)
X_test_vec = vectorizer.transform(X_test)
model = MultinomialNB()
model.fit(X_train_vec, y_train)
y_pred = model.predict(X_test_vec)

print(f'Accuracy: {accuracy_score(y_test, y_pred)}')

print("Training set:")
for text, label in zip(X_train, y_train):
    print(f"Text: {text}, Label: {label}")

print("\nTesting set:")
for text, label in zip(X_test, y_test):
    print(f"Text: {text}, Label: {label}")
```

```
Accuracy: 0.4
Training set:
Text: How are you?, Label: 0
Text: Exclusive deal, Label: 1
Text: Call me, Label: 0
Text: Free entry!, Label: 1
Text: Good day, Label: 0

Testing set:
Text: See you, Label: 0
Text: Buy now, Label: 1
Text: Winner!, Label: 1
Text: Hello, Label: 0
Text: Click here, Label: 1
```

In [ ]: