In [1]:
```
import pandas as pd
import numpy as np
from sklearn.feature_extraction.text import CountVectorizer, TfidfVectorizer
from sklearn.naive_bayes import BernoulliNB, MultinomialNB
from sklearn.metrics import accuracy_score, classification_report
```

In [3]:
```
train_df = pd.read_csv('train.csv')
test_df = pd.read_csv('test.csv')
```

In [4]:
```
train_df.head()
```

Out[4]:

	text	label
0	Unfortunately, the frustration of being Dr. Go...	0
1	Been going to Dr. Goldberg for over 10 years. ...	1
2	I don't know what Dr. Goldberg was like before...	0
3	I'm writing this review to give you a heads up...	0
4	All the food is great here. But the best thing...	1

In [6]:
```
test_df.head()
```

Out[6]:

	text	label
0	Contrary to other reviews, I have zero complai...	1
1	Last summer I had an appointment to get new ti...	0
2	Friendly staff, same starbucks fair you get an...	1
3	The food is good. Unfortunately the service is...	0
4	Even when we didn't have a car Filene's Baseme...	1

In [7]:
```
X_train = train_df['text']
y_train = train_df['label']
X_test = test_df['text']
y_test = test_df['label']
```

In [8]:
```
#Vectorization
vectorizer_binary = CountVectorizer(binary=True)
X_train_binary = vectorizer_binary.fit_transform(X_train)
X_test_binary = vectorizer_binary.transform(X_test)

vectorizer_count = CountVectorizer(binary=False)
X_train_count = vectorizer_count.fit_transform(X_train)
X_test_count = vectorizer_count.transform(X_test)
```

In [9]:
```
# TF-IDF Vectorizer
vectorizer_tfidf = TfidfVectorizer()
X_train_tfidf = vectorizer_tfidf.fit_transform(X_train)
X_test_tfidf = vectorizer_tfidf.transform(X_test)
```

In [10]:
```
bnb = BernoulliNB()
mnb = MultinomialNB()
```

In [11]:
```
bnb.fit(X_train_binary, y_train)
predictions_binary = bnb.predict(X_test_binary)
accuracy_binary = accuracy_score(y_test, predictions_binary)
```

In [12]:
```
mnb.fit(X_train_count, y_train)
predictions_count = mnb.predict(X_test_count)
accuracy_count = accuracy_score(y_test, predictions_count)

mnb.fit(X_train_tfidf, y_train)
predictions_tfidf = mnb.predict(X_test_tfidf)
accuracy_tfidf = accuracy_score(y_test, predictions_tfidf)
```

In [13]:
```
print(f'Accuracy with BernoulliNB and Binary Count Vectorizer: {accuracy_binary
print(f'Accuracy with MultinomialNB and Count Vectorizer: {accuracy_count:.3f}'
print(f'Accuracy with MultinomialNB and TF-IDF Vectorizer: {accuracy_tfidf:.3f}
```

```
Accuracy with BernoulliNB and Binary Count Vectorizer: 0.760
Accuracy with MultinomialNB and Count Vectorizer: 0.868
Accuracy with MultinomialNB and TF-IDF Vectorizer: 0.884
```

In [15]:
```
# Detailed Classification Report for the best model
print("\nClassification Report for MultinomialNB with TF-IDF Vectorizer:")
print(classification_report(y_test, predictions_tfidf))
```

```
Classification Report for MultinomialNB with TF-IDF Vectorizer:
              precision    recall  f1-score   support

           0       0.88      0.89      0.89     19000
           1       0.89      0.87      0.88     19000

    accuracy                           0.88     38000
   macro avg       0.88      0.88      0.88     38000
weighted avg       0.88      0.88      0.88     38000
```

In [ ]:
```

```