

Email Spam Detection with Machine Learning

```
import pandas as pd

from sklearn.feature_extraction.text import CountVectorizer

from sklearn.model_selection import train_test_split

from sklearn.naive_bayes import MultinomialNB

from sklearn.metrics import accuracy_score

data = pd.read_csv("/content/spam.csv", encoding="iso-8859-1")

vectorizer = CountVectorizer(stop_words="english")

X = vectorizer.fit_transform(data["v2"])

y = data["v1"]

X_train, X_test, y_train, y_test = train_test_split(X, y, test_size=0.2)

model = MultinomialNB()

model.fit(X_train, y_train)



▾ MultinomialNB  

    MultinomialNB()



y_pred = model.predict(X_test)

accuracy = accuracy_score(y_test, y_pred)

print("Accuracy:", accuracy)

Accuracy: 0.9811659192825112

new_email = "SMS. ac Sptv: The New Jersey Devils and the Detroit Red Wings play Ice Hockey. Correct or Incorrect? End? Reply END SPTV"

new_email_vector = vectorizer.transform([new_email])

prediction = model.predict(new_email_vector)[0]

print("Prediction:", prediction)

Prediction: spam
```

