

CODE

```
from sklearn.datasets import fetch_20newsgroups
import numpy as np
from sklearn.metrics import confusion_matrix
from sklearn.metrics import classification_report

cat = ['alt.atheism', 'soc.religion.christian', 'comp.graphics', 'sci.med']
twenty_train = fetch_20newsgroups(subset = 'train', categories = cat,
shuffle = True)
twenty_test = fetch_20newsgroups(subset = 'test', categories = cat, shuffle
= True)

from sklearn.feature_extraction.text import CountVectorizer
count_vect = CountVectorizer()
X_train_tf = count_vect.fit_transform(twenty_train.data)

from sklearn.feature_extraction.text import TfidfTransformer
tfidf_transformer = TfidfTransformer()
X_train_tfidf = tfidf_transformer.fit_transform(X_train_tf)
# X_train_tfidf.shape

from sklearn.naive_bayes import MultinomialNB
from sklearn.metrics import accuracy_score
from sklearn import metrics
mod = MultinomialNB()
mod.fit(X_train_tfidf, twenty_train.target)
X_test_tf = count_vect.transform(twenty_test.data)
X_test_tfidf = tfidf_transformer.transform(X_test_tf)
predicted = mod.predict(X_test_tfidf)

print("Accuracy: ",accuracy_score(twenty_test.target,predicted))
print(classification_report(twenty_test.target,predicted,target_names=twenty_test.target_names))
print(metrics.confusion_matrix(twenty_test.target,predicted))
```

OUTPUT

```
('Accuracy: ', 0.83488681757656458)
      precision  recall f1-score  support

alt.atheism      0.97    0.60    0.74    319
comp.graphics    0.96    0.89    0.92    389
sci.med          0.97    0.81    0.88    396
soc.religion.christian  0.65    0.99    0.78    398

micro avg       0.83    0.83    0.83   1502
macro avg       0.89    0.82    0.83   1502
weighted avg    0.88    0.83    0.84   1502

[[192  2  6 119]
 [ 2 347  4  36]
 [ 2 11 322  61]
 [ 2  2  1 393]]
```