```
#!/usr/bin/env python
# coding: utf-8

# In[9]:


import pandas as pd
import os
import
nltk
from nltk.corpus import stopwords
from sklearn.feature_extraction.text import
TfidfVectorizer
from sklearn.model_selection import train_test_split
from sklearn.naive_bayes
import MultinomialNB
from sklearn.metrics import accuracy_score


#input the directory of data
set.
os.chdir(r'C:\Users\SATHWIK\Desktop\New folder')
# Load the dataset
df =
pd.read_csv('BBCNews.csv')
df.head()


# In[11]:


# Preprocessing: Tokenization, Stopword
removal, and Vectorization
stop_words = stopwords.words('english')
vectorizer =
TfidfVectorizer(stop_words=stop_words)
corpus = df['Text'].tolist()
X =
vectorizer.fit_transform(corpus)
y = df['Category'].values

# Split the dataset into train and
test sets
X_train, X_test, y_train, y_test = train_test_split(X, y, test_size=0.2,
random_state=42)

# Train a Naive Bayes classifier on the training set
clf =
MultinomialNB()
clf.fit(X_train, y_train)

# Evaluate the accuracy of the classifier on the
test set
y_pred = clf.predict(X_test)
accuracy = accuracy_score(y_test,
y_pred)
print('Accuracy:', accuracy)


# In[13]:




# In[ ]:
```