ASSIGNMENT-17

1) Implementing bag of words and tf-idf using naive Bayes algorithm and check the accuracy

```
import numpy as np
from sklearn.datasets import fetch_20newsgroups
from sklearn.feature_extraction.text import CountVectorizer, TfidfTransformer
from sklearn.naive_bayes import MultinomialNB
from sklearn.pipeline import make_pipeline
from sklearn.model_selection import train_test_split
from sklearn.metrics import accuracy_score

dataset = fetch_20newsgroups(subset='all')
X, y = dataset.data, dataset.target

X_train, X_test, y_train, y_test = train_test_split(X, y, test_size=0.2, random_state=42)

# Define a model using Multinomial Naive Bayes and feature extraction with CountVectorizer and TfidfTransformer
model = make_pipeline(CountVectorizer(), TfidfTransformer(), MultinomialNB())

model.fit(X_train, y_train)
y_pred = model.predict(X_test)

# Calculate the accuracy of the model
accuracy = accuracy_score(y_test, y_pred)
print('Accuracy: %.2f' % (accuracy*100))
```