

נתון dataset של מאמרים מתויגים - המשימה היא לחזות קטגוריה של מאמר בהסתמך על המילים בטקסט

להלן הצורה להוריד את ה dataset הרלוונטי:

```
from sklearn.datasets import fetch_20newsgroups

# Load the 20newsgroups dataset
categories = ['alt.atheism', 'soc.religion.christian', 'comp.graphics',
             'sci.med']
train_data = fetch_20newsgroups(subset='train', categories=categories,
                                shuffle=True, random_state=42)
test_data = fetch_20newsgroups(subset='test', categories=categories,
                                shuffle=True, random_state=42)
len(train_data.data), len(test_data.data)
```