**Script 3 help that I have taken from AI**

**You said:**

hello chat gpt

**ChatGPT said:**

Hey Naveer, good to see you! What are you working on today? Need help with a project, concept, or something else entirely?

**You said:**

I am working on a project in my digital humanities class where i am working on al Jazeera articles and analysing them, i need to now filter out some stops words,but i couldn't recall the code or library, can you please give me some help?

**ChatGPT said:**

Absolutely. To filter out stop words from text in Python, the most commonly used libraries are **NLTK** and **spaCy**. Here's a quick rundown for both:

**✅ Using NLTK (Natural Language Toolkit):**

python

CopyEdit

import nltk

from nltk.corpus import stopwords

from nltk.tokenize import word\_tokenize

# Download stopwords if you haven't already

nltk.download('stopwords')

nltk.download('punkt')

# Sample text

text = "This is a sample article from Al Jazeera about the Gaza conflict."

# Tokenize text

words = word\_tokenize(text)

# Remove stopwords

filtered\_words = [word for word in words if word.lower() not in stopwords.words('english')]

print(filtered\_words)

**✅ Using spaCy:**

python

CopyEdit

import spacy

# Load English tokenizer, POS tagger, etc.

nlp = spacy.load("en\_core\_web\_sm")

# Sample text

text = "This is a sample article from Al Jazeera about the Gaza conflict."

# Process the text

doc = nlp(text)

# Filter stop words

filtered\_words = [token.text for token in doc if not token.is\_stop]

print(filtered\_words)

Let me know if you're using a specific tokenizer (like from Stanza) or need to do this on a large dataset. I can help you write a custom stopword filter, batch processor, or integrate it into your existing pipeline.

**You said:**

okay i will be using the first one. thank u