```
#excller day-7!pip install gensim nltk spacy
import re
import gensim
from nltk.stem.porter import PorterStemmer
from nltk.corpus import stopwords
import spacy
import nltk

# Download NLTK stopwords
nltk.download('stopwords')

# Load required modules and data
nlp = spacy.load("en_core_web_sm")
porter_stemmer = PorterStemmer()
stop_words = set(stopwords.words('english'))

def preprocess_text(text):
    # Remove punctuation and convert to lowercase
    text = re.sub(r'[^\w\s]', '', text.lower())

    # Tokenization and removing stopwords
    tokens = [word for word in gensim.utils.simple_preprocess(text) if word not in stop_words]

    # Stemming the tokens
    stemmed_tokens = [porter_stemmer.stem(token) for token in tokens]

    # Lemmatizing using spaCy
    doc = nlp(' '.join(stemmed_tokens))
    lemmatized_tokens = [token.lemma_ for token in doc]

    return lemmatized_tokens

# Sample text for processing
text_content = """
Write a Python script that uses Gensim to preprocess data from a sample text
file. Follow basic procedures like tokenization, stemming, and lemmatization.
Print the final output to verify the preprocessing steps.
"""

# Preprocess the text and print the result
processed_text = preprocess_text(text_content)
print(processed_text)
```

```
[nltk_data] Downloading package stopwords to /root/nltk_data...
[nltk_data]   Package stopwords is already up-to-date!
['write', 'python', 'script', 'use', 'gensim', 'preprocess', 'data', 'sampl', 'text', 'file', 'follow', 'basic', 'procedur', 'like',
```

Start coding or generate with AI.