## Write a Python script that:

1. Use Genism to preprocess data from a sample text file, follow basic procedures like tokenization, stemming, lemmatization etc.

```
In [1]:

★ from gensim.parsing.preprocessing import preprocess_string, strip_punctuat

        from gensim.parsing.porter import PorterStemmer
        def preprocess_text(text):
            filters = [lambda x: x.lower(), strip_punctuation, strip_numeric, remo
            return preprocess_string(text, filters)
        sample text = "This is a sample text file for preprocessing. It contains m
        processed_data = preprocess_text(sample_text)
        stemmer = PorterStemmer()
        stemmed_data = [stemmer.stem(word) for word in processed_data]
        print("Lemmatized Data:", processed_data)
        print("Stemmed Data:", stemmed_data)
        Lemmatized Data: ['sample', 'text', 'file', 'preprocessing', 'contains',
        'multiple', 'words', 'sentences']
        Stemmed Data: ['sampl', 'text', 'file', 'preprocess', 'contain', 'multipl
         ', 'word', 'sentenc']
```

1 of 1 25-01-2025, 14:54