

**NAME: SABIKA ZAHRA**

**ROLL NUMBER: 074**

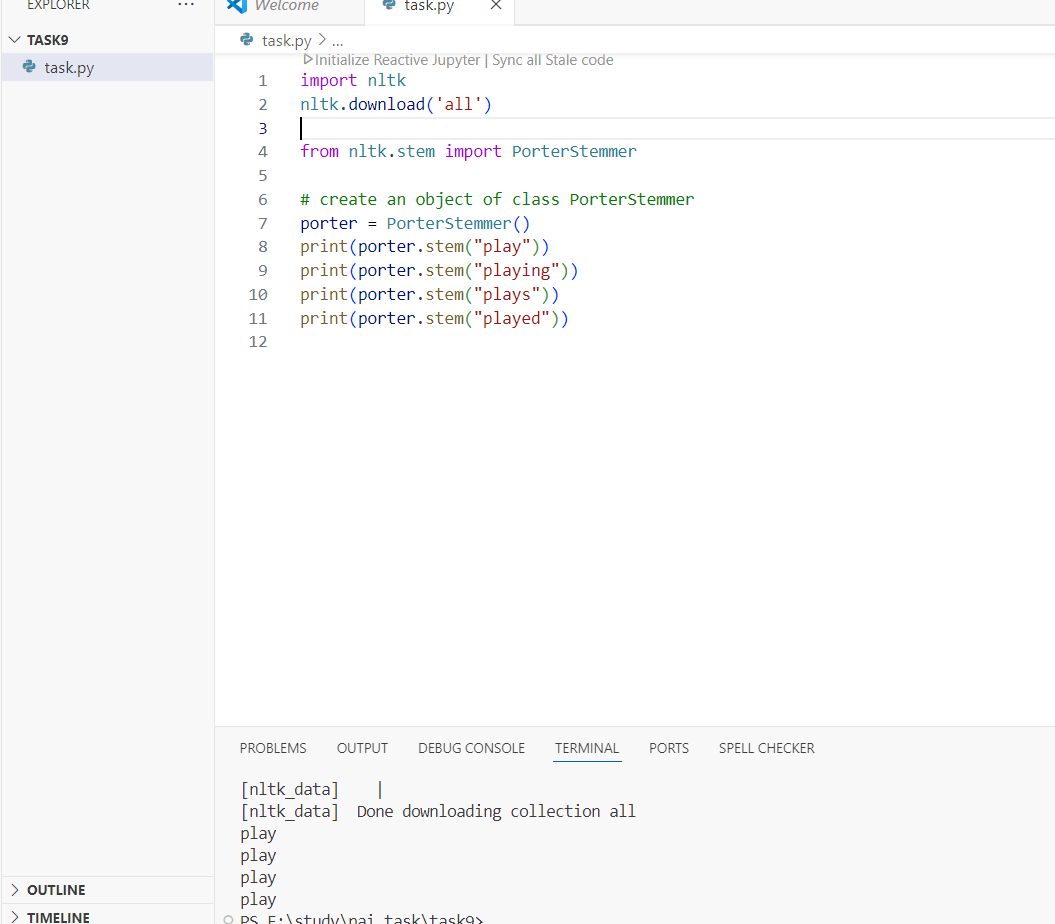
**CLASS: BSAI**

**SECTION: ALPHA**

**SEMESTER: 4**

**SUBMITTED TO: SIR MUNEEB SALEEM**

**DATE: 2,20,2025**



* Imports the nltk library and downloads all its datasets (e.g., stopwords, corpora).
* Lemmatization uses vocabulary to return the base/dictionary form (e.g., "better" → "good").
* All variants ("playing", "plays", "played") are stemmed to the root **"play"**.