Aamleen Ahmed
2020002

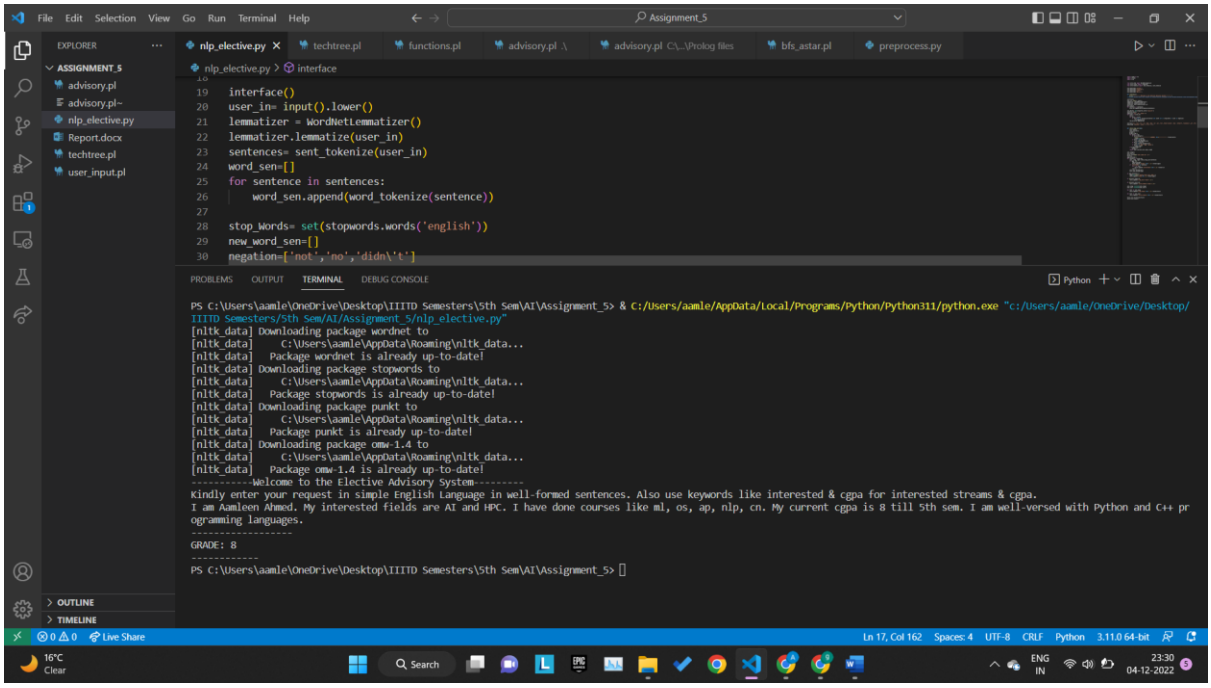
Assignment 5 – NLP + Prolog

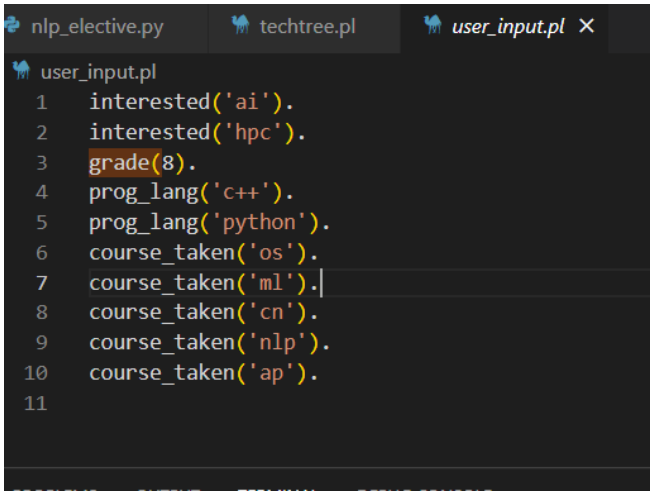
Artificial Intelligence

Instructions to work

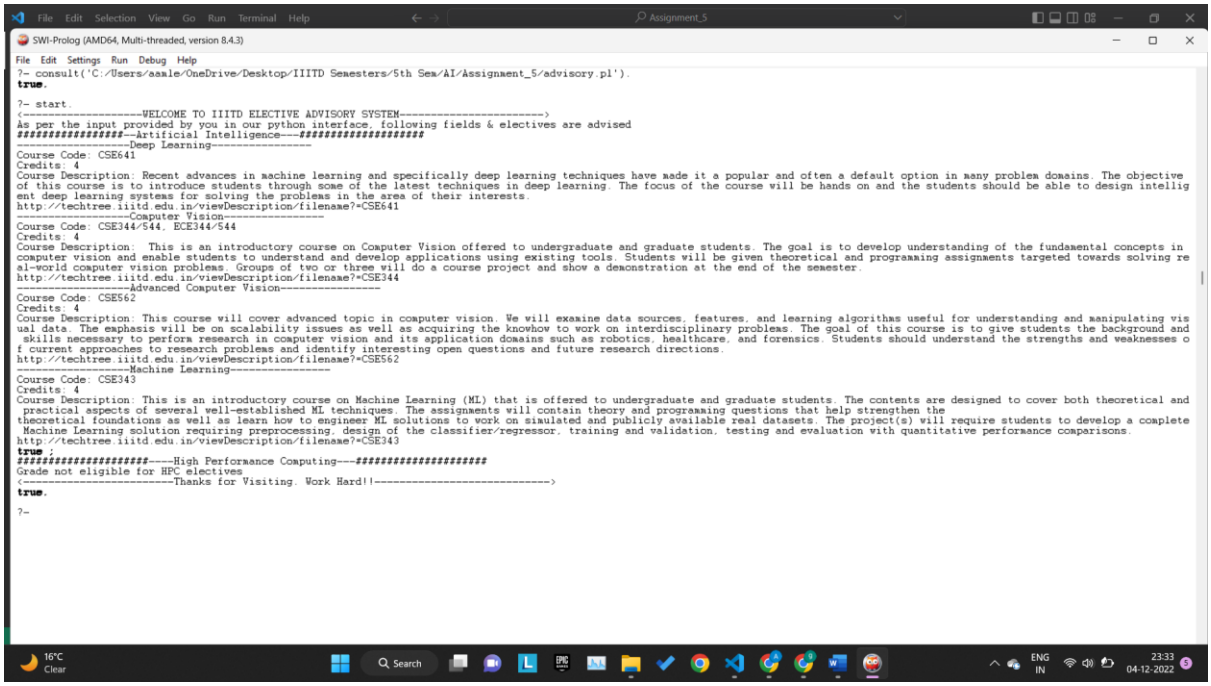
Run python file first (nlp_electives.py). It will ask for user input, which you can provide in your simple english language. You just need to mention your interested stream in 1 sentence, and cgpa in 1 sentence. The highlighted words are keywords, along with any of their related words, so use them for the program to detect accordingly. Rest courses provided will be prerequisites.

Now once python file is completed, open prolog and consult and run the advisory.pl file. Change the path of other consulted files first. Then just write start. to start the elective advisory. Keep pressing ';' after outputs to continue for next interested fields, unless Thank you message is printed.





Facts file generated by NLP program



Coding Procedure/Planning

Step 1: Started by taking input from the user.

Step 2: Preprocessing → Using various NLP libraries, preprocessed the data to extract meaningful tokens. In detail steps:

a. Downloaded words, stoping words (general connecting words and all), and punctuation marks.
b. Using lemmatizer, simplified the input to club multiple words of the same origin (root word) to be grouped under single root word. This will help us in reducing the data & just giving the context by the root words.
c. Now tokenizing the input into sentences, and then further extracting each word as token from each sentence. Words of same sentence are stored in one list, and together all sentences form nested list.
d. Next, removed the stopping words to simplify the data and remove any useless words. Avoided negation statements for proper context.
e. Now using our predefined courses & programming languages, extracted the course choice and interested stream of user. Also took input of cgpa in 1 of the sentences. If not provided, cgpa input is taken separately.
f. Now just writing all of this data as facts:
    a. Interested() : contains the fields which interests user
    b. Course_taken() : specifies the courses done by user already
    c. Prog_lang() : specifies the user's programming languages known.
    d. Grade() : specifies the user grade between 1-10, integer only.
g. Now output file (user_input.pl) is ready

Step 3: Prolog-> Now consulting the user_input.pl file for input facts made using NLP, and also techtree.pl that contains all the course details.

Step 4: In prolog, checking for the interests of user, then checking for prereq courses done, and then grade requirement. If all requirements are met, then electives to be taken are shown else appropriate message is shown.
