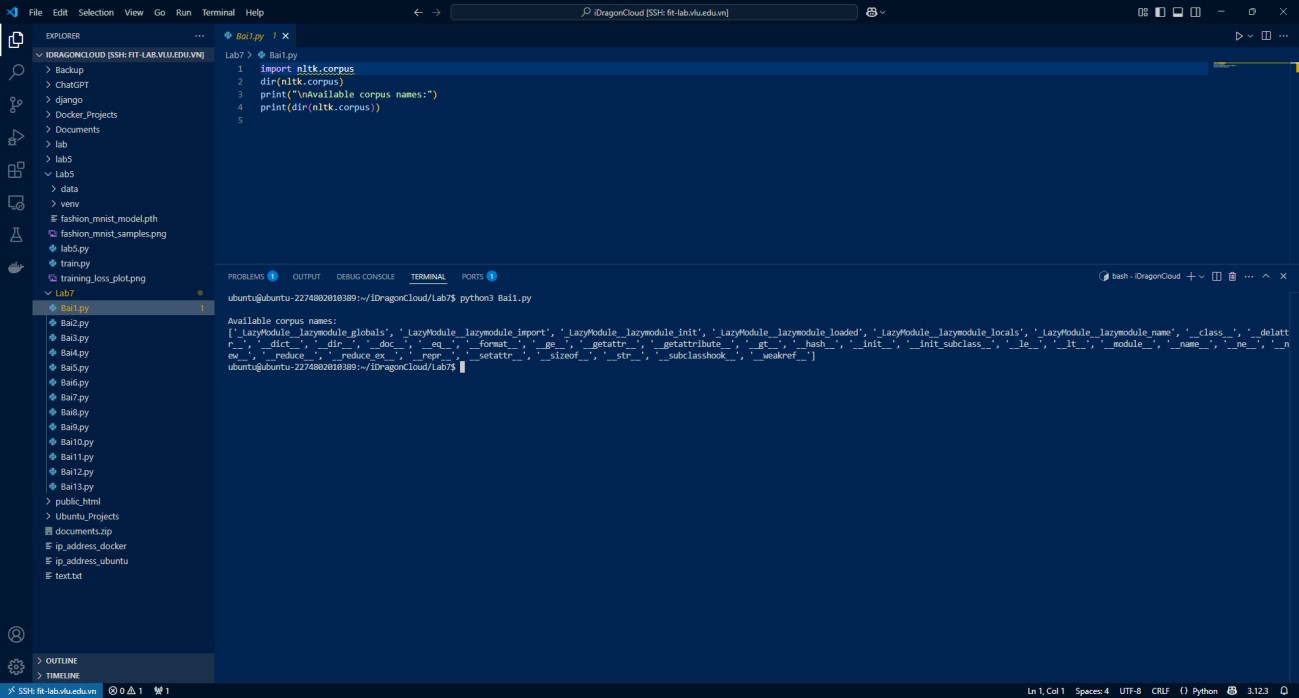
Họ tên: Bùi Bá Minh Đăng

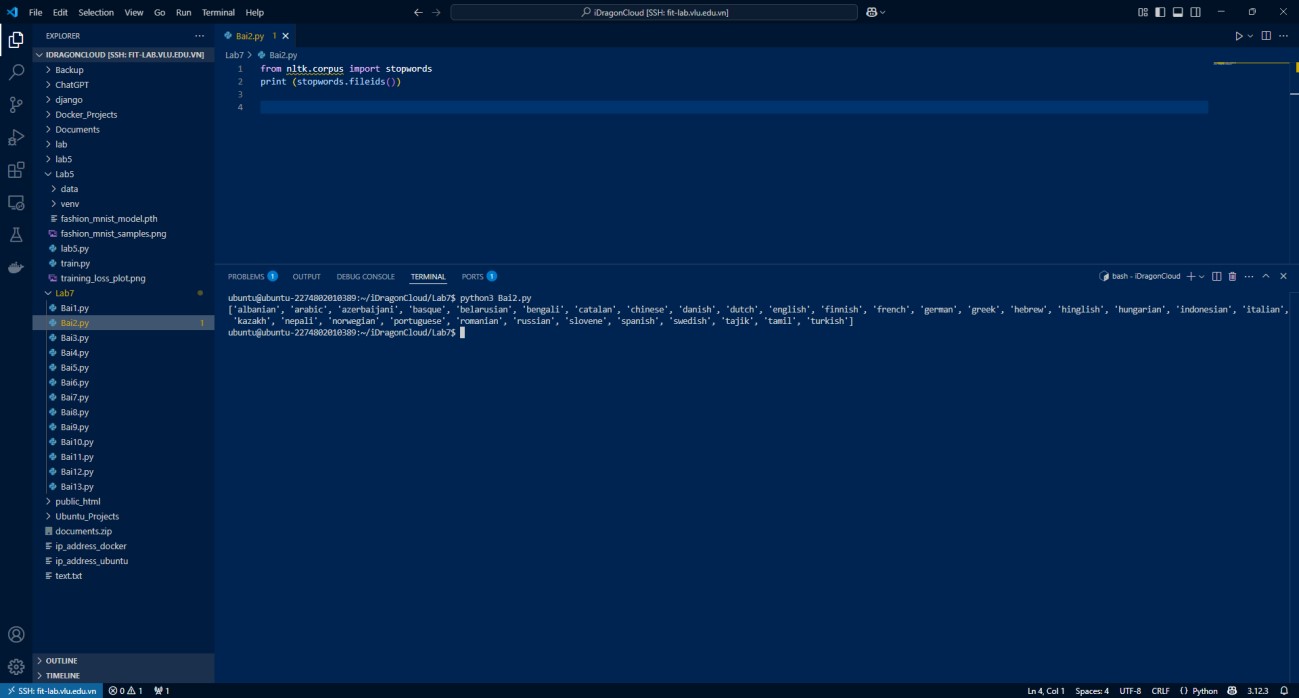
Mssv: 2274802010174

# Lab 7 PHÂN TÍCH DỮ LIỆU DẠNG VĂN BẢN VỚI NLTK

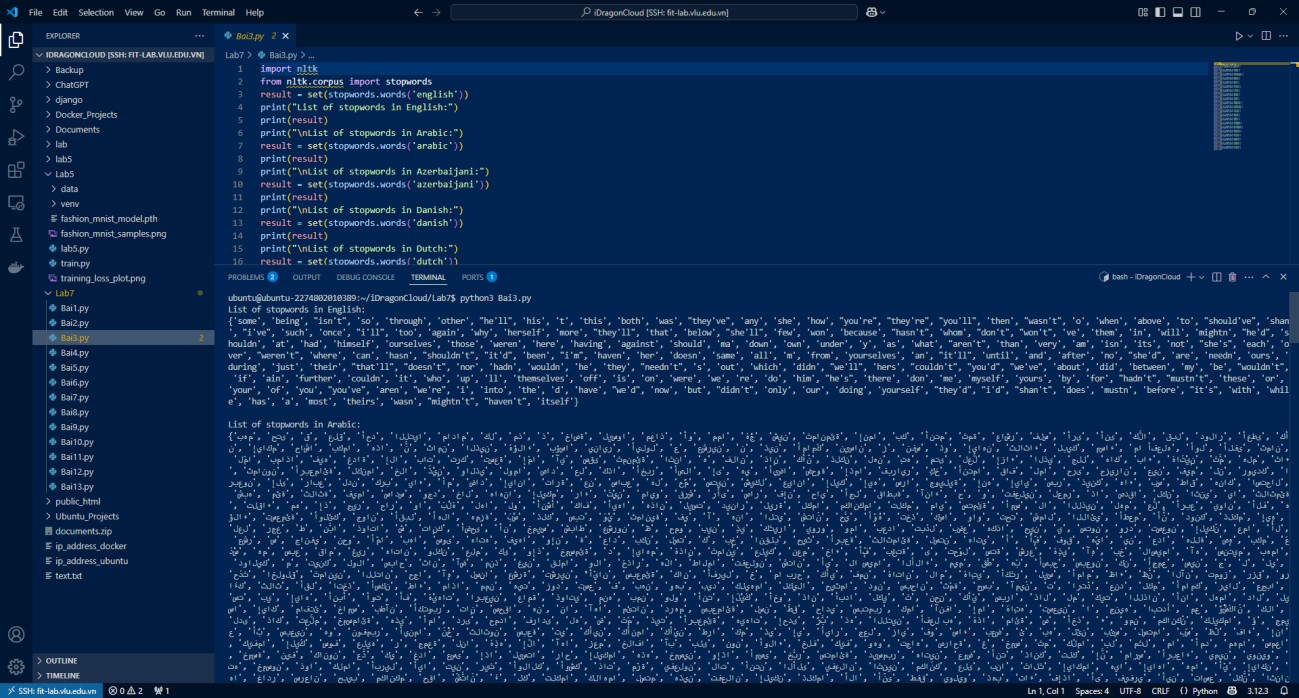
1. Viết chương trình Python với thư viện NLTK để liệt kê các tên của copus.



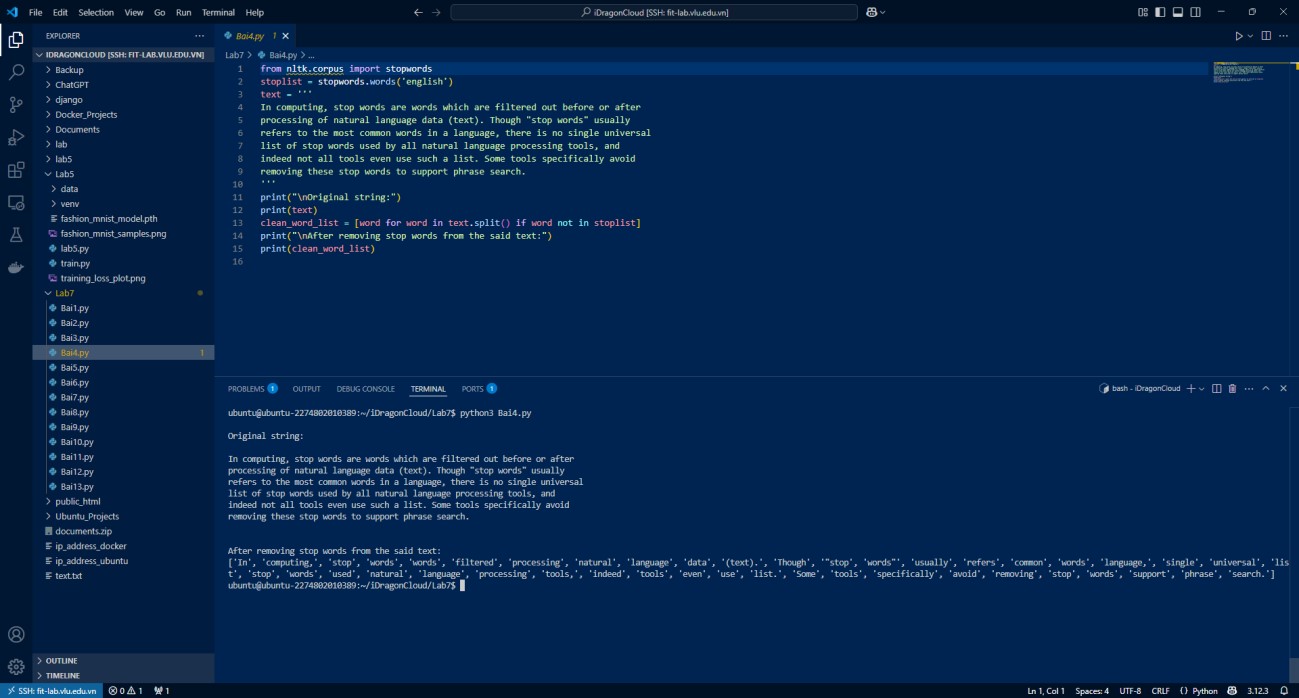
1. Viết chương trình Python với thư viện NLTK để liệt kê danh sách các stopword bằng các ngôn ngữ khác nhau.



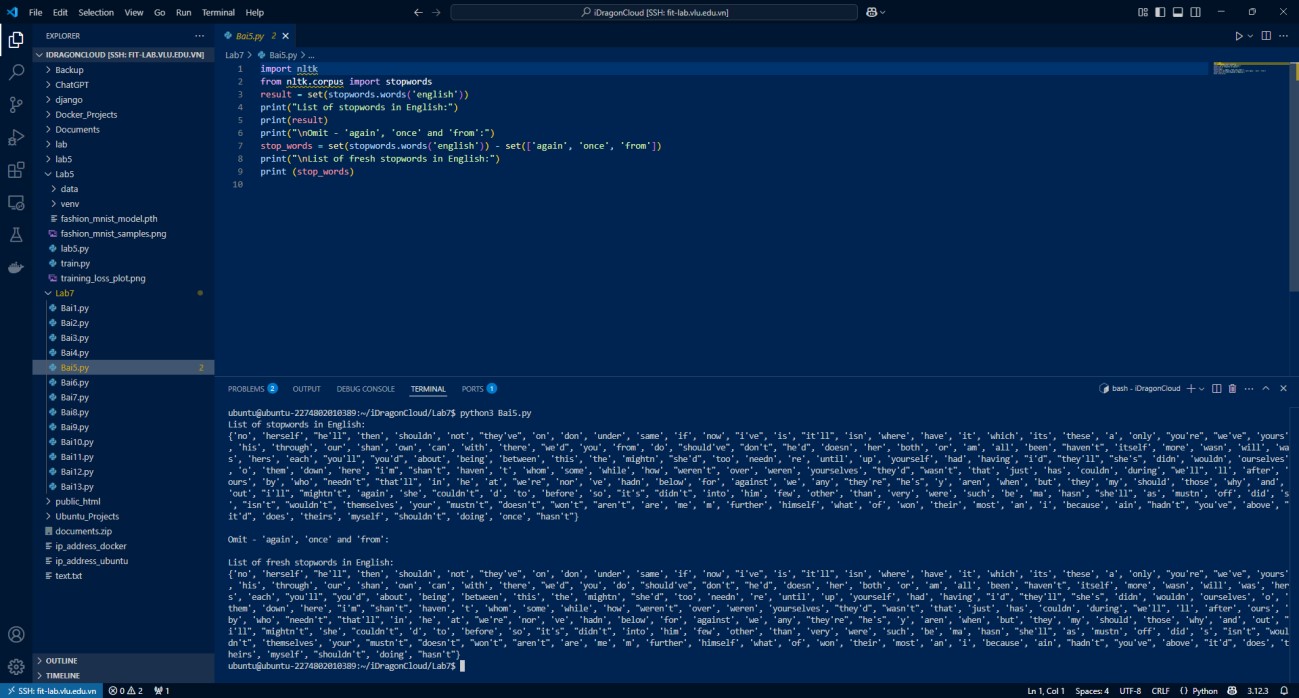
1. Viết chương trình Python với thư viện NLTK để kiểm tra danh sách các stopword bằng các ngôn ngữ khác nhau.



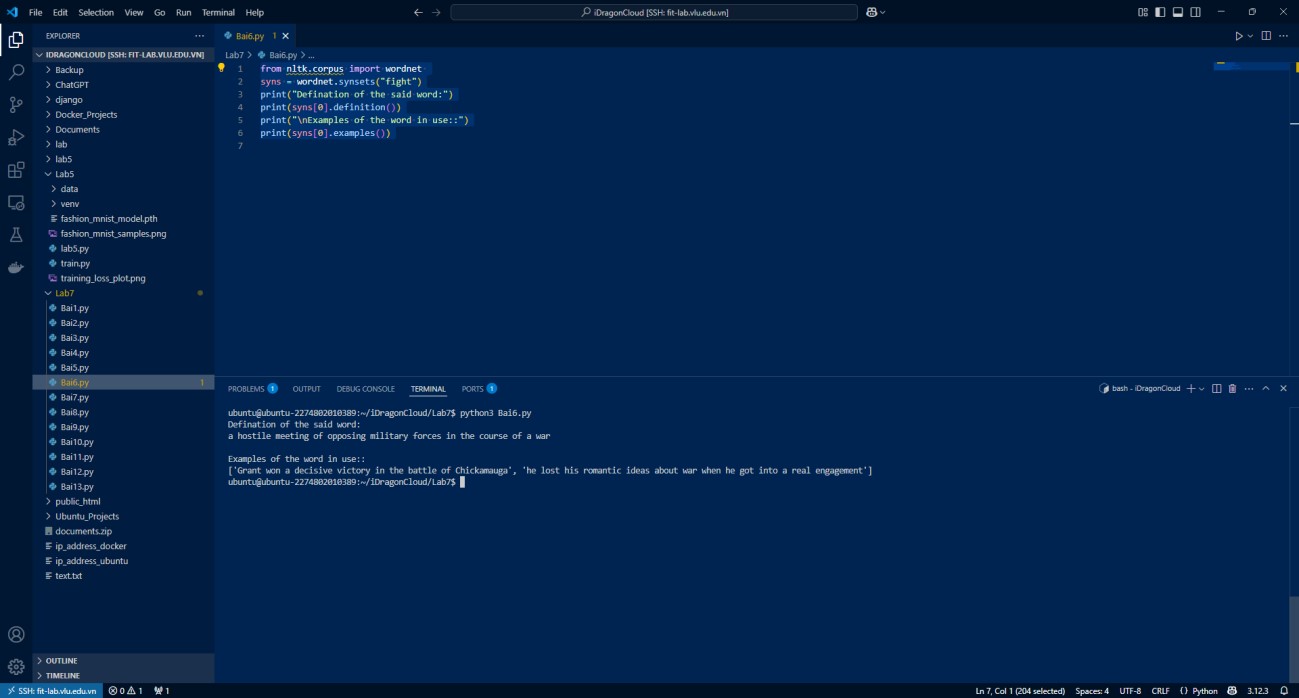
1. Viết chương trình Python với thư viện NLTK để loại bỏ các stopword từ một văn bản đã cho.



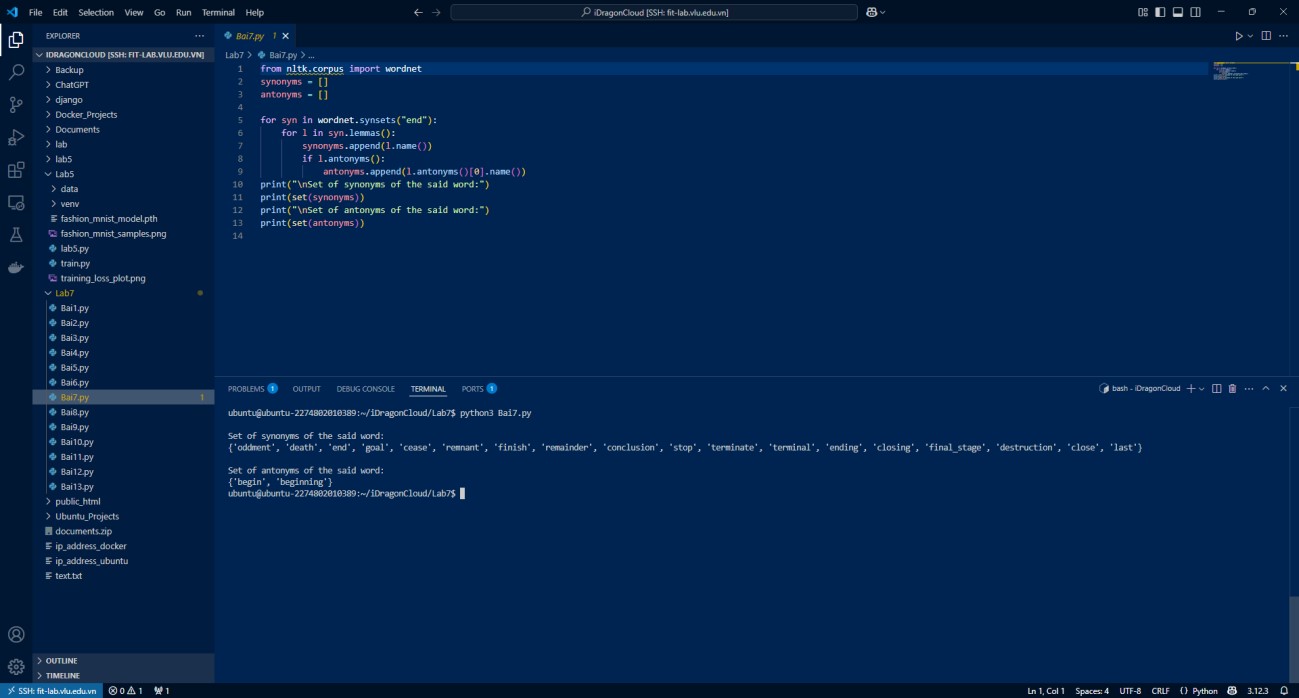
1. Viết chương trình Python với thư viện NLTK bỏ qua các stopword từ danh sách các stopword.



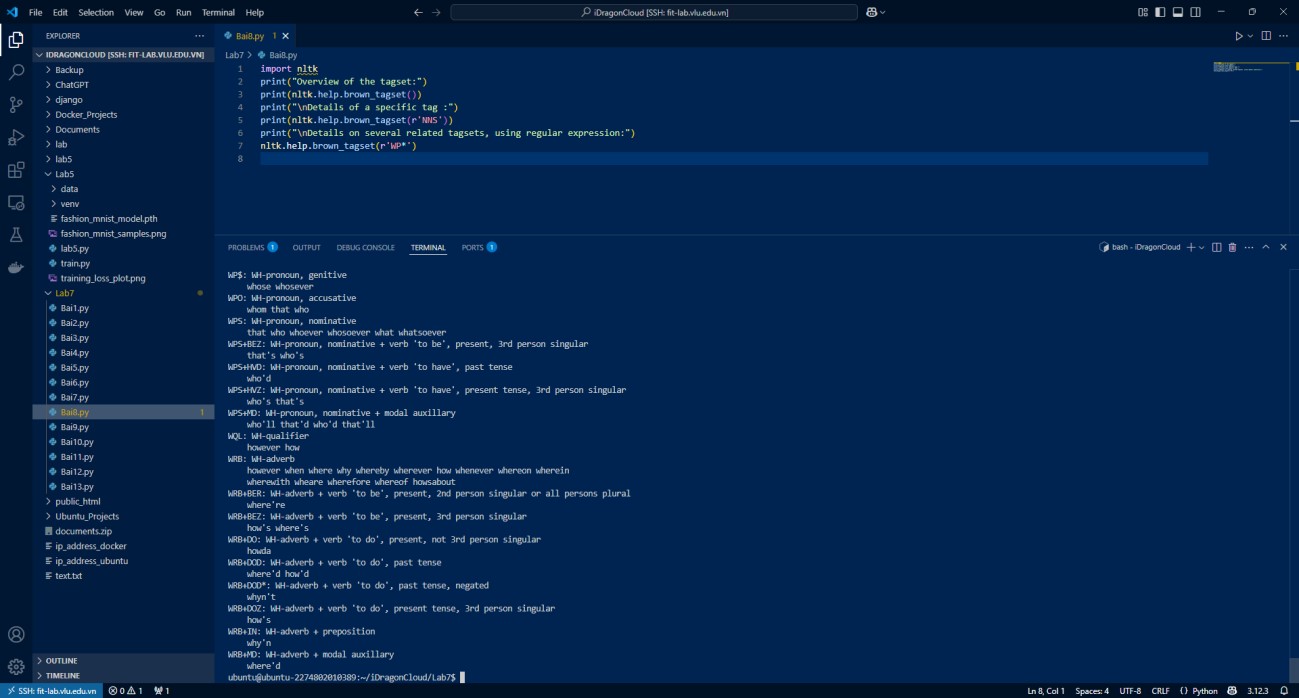
1. Viết một chương trình Python với thư viện NLTK để tìm định nghĩa và ví dụ của một từ đã cho bằng WordNet từ Wikipedia,



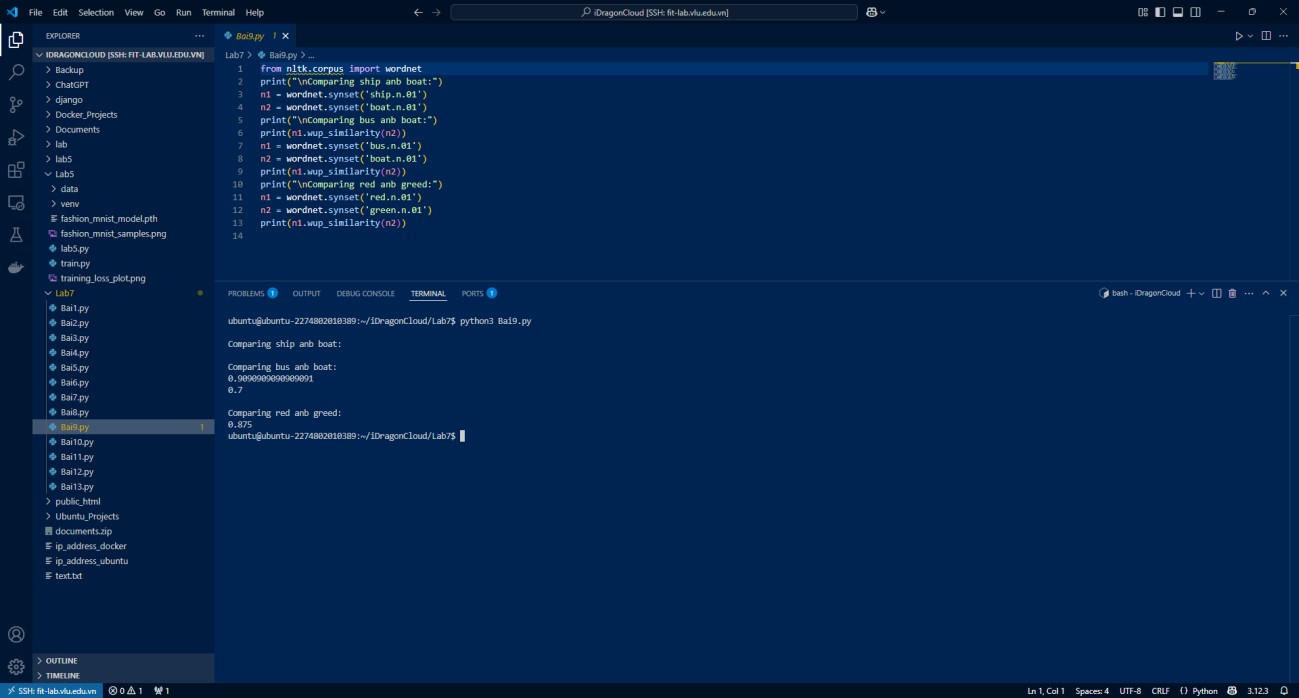
1. Viết chương trình Python với thư viện NLTK để tìm tập hợp các từ đồng nghĩa và trái nghĩa của một từ nào đó.



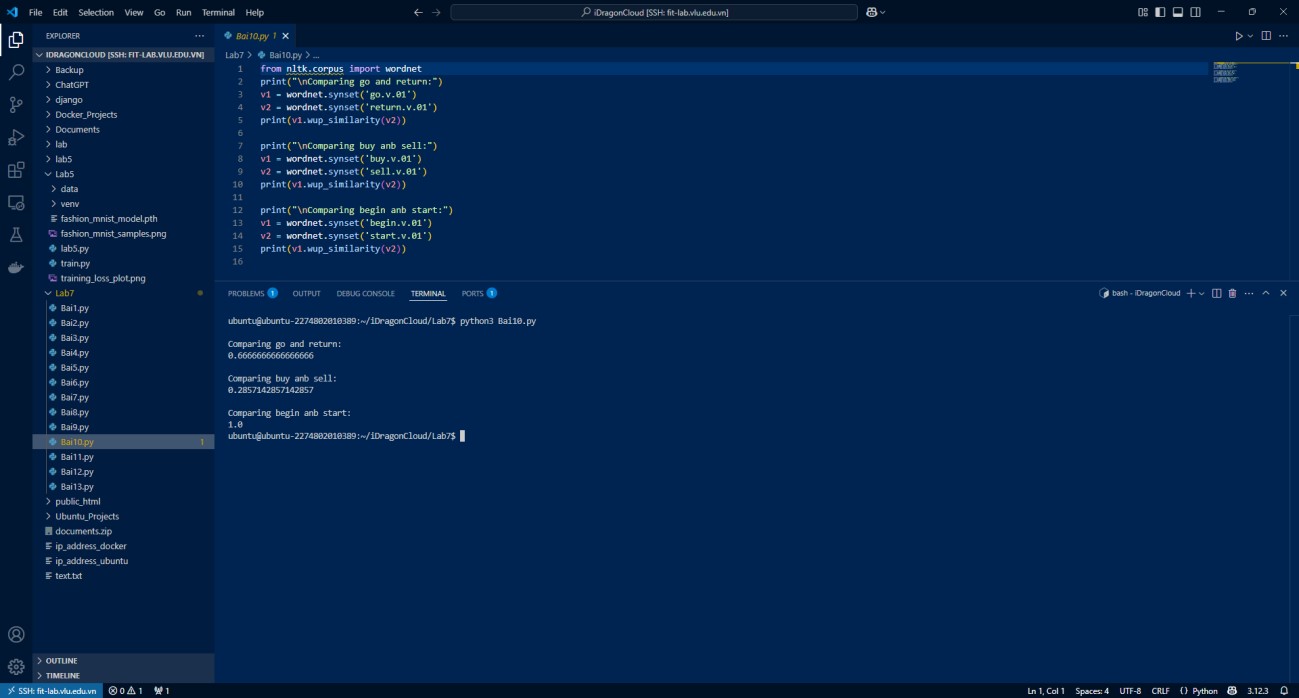
1. Viết chương trình Python với thư viện NLTK để có cái nhìn tổng quan về bộ tag, chi tiết của một tag cụ thể trong bộ tag và chi tiết về một số bộ tag liên quan, sử dụng biểu thức chính quy.



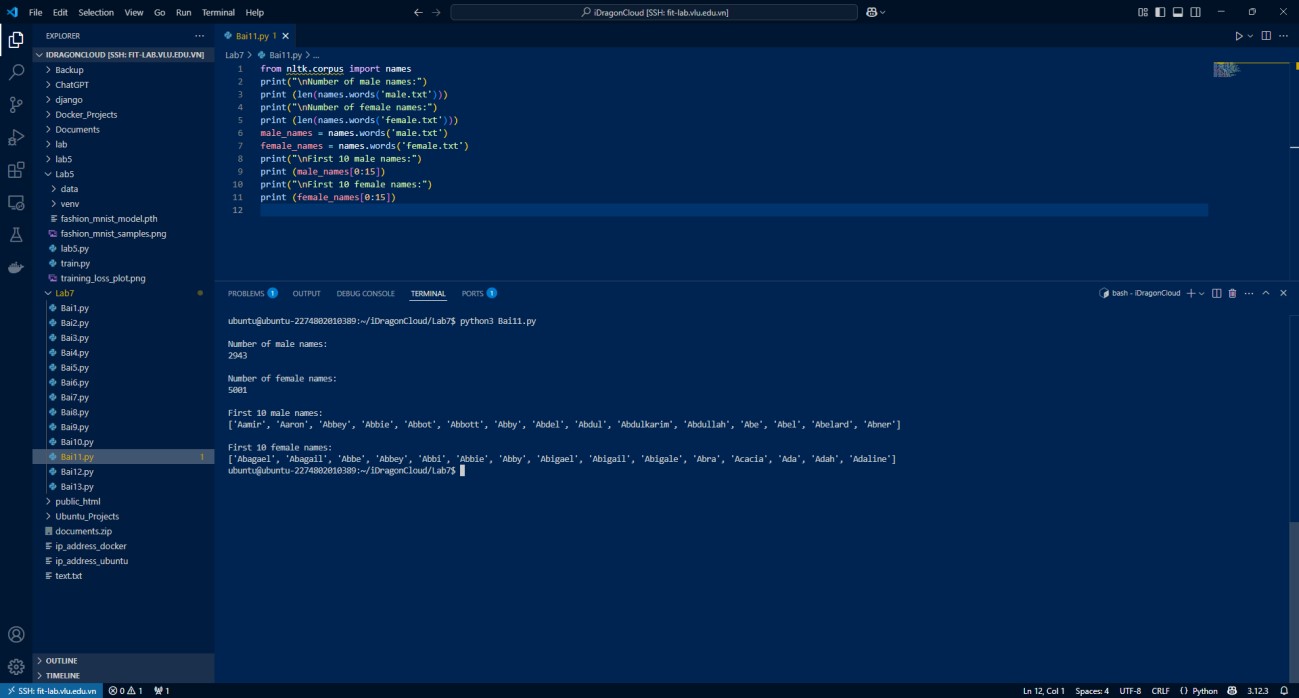
1. Viết chương trình Python với thư viện NLTK để so sánh sự giống nhau của hai danh từ đã cho.



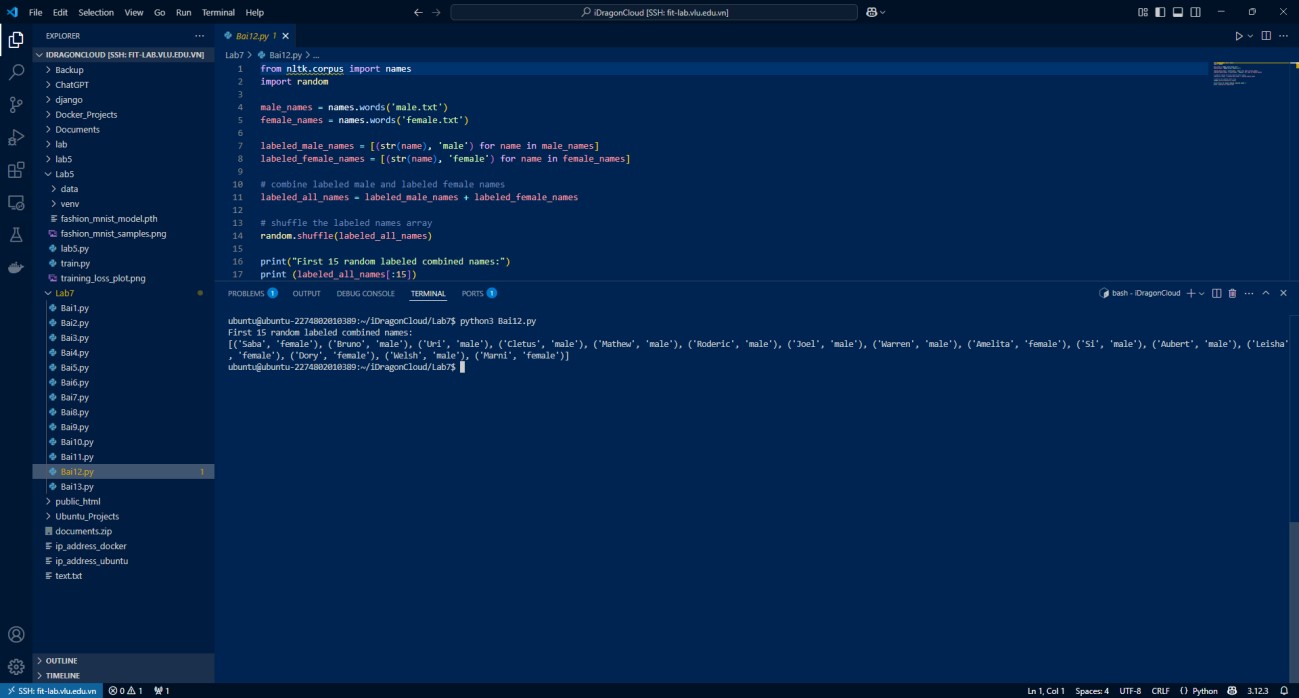
1. Viết chương trình Python với thư viện NLTK để so sánh sự giống nhau của hai động từ đã cho.



1. Viết chương trình Python với thư viện NLTK để tìm số lượng tên nam và nữ trong các tên kho ngữ liệu. In tên 10 nam và nữ đầu tiên. Lưu ý: Kho văn bản tên chứa tổng cộng khoảng 2943 nam (male.txt) và 5001 nữ (Female.txt) tên. Kho được biên soạn bởi Kantrowitz, Ross.



1. Viết chương trình Python với thư viện NLTK để in 15 kết hợp ngẫu nhiên đầu tiên được gắn nhãn nam và được gắn nhãn tên nữ từ kho tên.



1. Viết chương trình Python với thư viện NLTK để trích xuất ký tự cuối cùng của tất cả các tên được gắn nhãn và tạo mảng mới với chữ cái cuối cùng của mỗi tên và nhãn được liên kết.

