

## Thư viện Python:

Thư viện	Mục đích chính
nltk	Xử lý ngôn ngữ tự nhiên (tokenization, POS tagging, WordNet)
networkx	Xây dựng và xử lý đồ thị (graph), dùng để biểu diễn quan hệ từ đồng nghĩa
matplotlib	Vẽ đồ thị trực quan hóa bằng biểu đồ, node, edges
rich	In giao diện đẹp trong terminal (bảng, màu sắc, hiệu ứng)
sentence-transformers	So sánh ngữ nghĩa giữa các câu (semantic similarity) bằng embedding

## Corpus NLTK:

Dữ liệu	Vai trò
punkt	Dùng để chia câu và từ (tokenizer)
averaged_perceptron_tagger	Gán loại từ (noun, verb, adj, ...)
wordnet	Cơ sở dữ liệu từ đồng nghĩa, dùng để tra từ và xây đồ thị

```
#!/bin/bash

echo "[*] BẮT ĐẦU CÀI ĐẶT MÔI TRƯỜNG LAB..."
sleep 1

# Cập nhật pip và cài đặt các thư viện cần thiết
pip install --upgrade pip setuptools wheel

echo "[*] Cài đặt thư viện Python cần thiết..."
pip install nltk networkx matplotlib rich sentence-transformers

echo "[*] Tải dữ liệu NLTK cần thiết..."
python3 -m nltk.downloader punkt averaged_perceptron_tagger wordnet
```

Bước này ta sẽ Sử dụng NetworkX để xây dựng một biểu đồ không mong muốn của các mối quan hệ đồng nghĩa. Mỗi từ là một nút. Thêm một cạnh giữa hai từ nếu chúng là từ đồng nghĩa.

Ta sử dụng lệnh sau để bắt đầu phần này:

*encode\_vertex\_2bits*