**Bước 1**: Tải và cài đặt python

- Tải python tại trang: <https://www.python.org/downloads/windows/>

- Tải bản Download Windows installer (64-bit)

- Mở file python-3.11.5-amd64.exe đã tải về, tick chọn 2 options và next đến khi cài đặt xong

**Bước 2**: Kiểm tra máy đã cài đặt python thành công

- Gõ Window + R -> Nhập cmd -> Ok

- Gõ python --version -> Hiện phiên bản là thành công

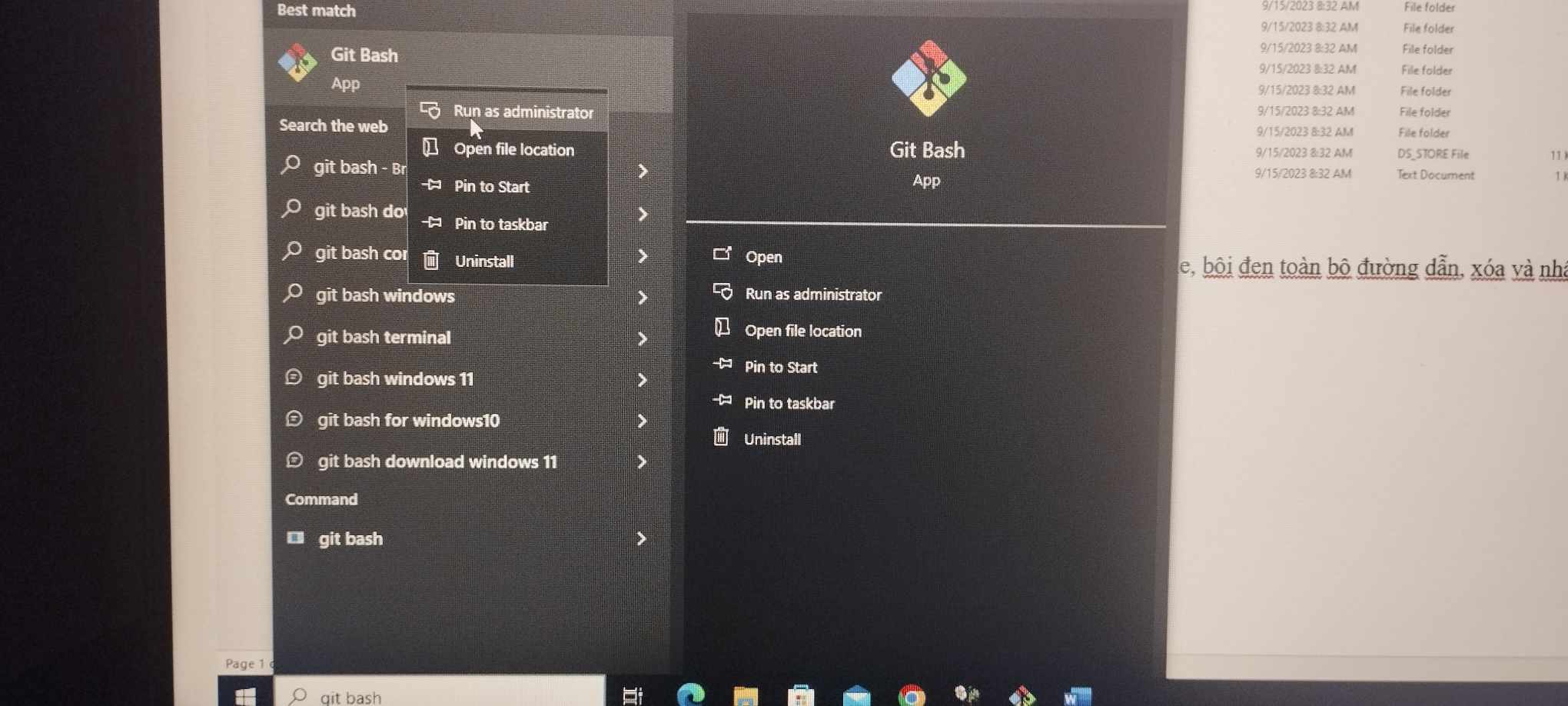
A black background with white text

Description automatically generated

**Bước 3**: Cài đặt git và git bash:

* Làm theo hướng dẫn ở link sau: <https://funix.edu.vn/chia-se-kien-thuc/cai-dat-git-va-git-bash-windows/>

**Bước 4**: Mở git bash bằng quyền Administrator:



* Kết quả như sau:

A black screen with purple text

Description automatically generated

**Bước 5**: Điều hướng tới thư mục tools (sau khi tải về, thư mục tools ở đây mặc định là ở thư mục Download):

* Gõ câu lệnh: **cd Downloads**
* Gõ câu lệnh: **cd tools**

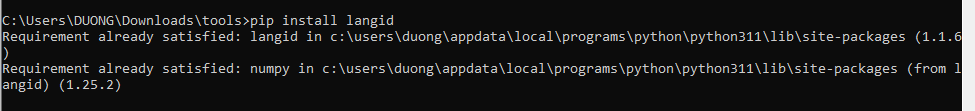
A screenshot of a computer program

Description automatically generated

**Bước 6**: Cài đặt thư viện:

Gõ những dòng lệnh sau một cách lần lượt: (gõ tất cả các lệnh)

* **pip install langid**



* **pip install langdetect**

A screen shot of a computer

Description automatically generated

* **pip install docx2txt**

A computer screen shot of text

Description automatically generated

* **pip install nltk**

**Bước 7**. Chạy tools:

**Một số lưu ý:**

- Đặt các file đầu vào vào thư mục data. File đầu vào có định dạng docx với cấu trúc sau:

+ Nếu là 2 file song ngữ: **<filename>(EN).docx và <filename>(VN).docx**

+ Nếu là file bilingual: **<tên file>(EN-VN)**

- Chạy bằng câu lệnh: **bash file\_name.sh**

- Folder output sẽ chứa các file đầu ra. Định dạng file đầu ra: out\_en\_filename.txt (ví dụ với file ifrs.docx đầu vào thì 2 file xuất ra là out\_en\_ifrs.txt và out\_vi\_ifrs.docx)

**Tool 1: File British:**

* Chuẩn bị dữ liệu: Copy các file đầu vào vào thư mục data trong folder British

A screenshot of a computer

Description automatically generated

* Bước 1: Vào thư mục British bằng git bash:
* Gõ câu lệnh: **cd British**

A screenshot of a computer

Description automatically generated

* Chạy file british.sh: **bash british.sh**

A screenshot of a computer program

Description automatically generated

* File đầu ra trong thư mục output:

A screenshot of a computer

Description automatically generated

**Tool 2: File IFRS:**

* Chuẩn bị dữ liệu: Copy các file đầu vào vào thư mục data trong folder IFRS (tương tự với thư mục British)
* Bước 1: Quay về thư mục tools ( nếu đang ở trong một thư mục con trong thư mục tools )
* Gõ câu lệnh: **cd ..**

A black screen with yellow and purple text

Description automatically generated

* Bước 2: Vào thư mục IFRS bằng git bash:
* Gõ câu lệnh: **cd IFRS**

A black screen with colorful text

Description automatically generated

* Chạy file ifrs.sh: **bash ifrs.sh**

A screenshot of a computer

Description automatically generated

* File đầu ra trong thư mục output

**Tool 3: File PBV:**

* Chuẩn bị dữ liệu: Copy các file đầu vào vào thư mục data trong folder PBV (tương tự với thư mục British)
* Bước 1: Quay về thư mục tools ( nếu đang ở trong một thư mục con trong thư mục tools )
* Gõ câu lệnh: **cd ..**

A black screen with yellow and purple text

Description automatically generated

* Bước 2: Vào thư mục PBV bằng git bash:
* Gõ câu lệnh: **cd PBV**

A black screen with text on it

Description automatically generated

* Chạy file pbv.sh: **bash pbv.sh**

A computer screen with text

Description automatically generated

* File đầu ra trong thư mục output

**Tool 4: File ND01:**

* Chuẩn bị dữ liệu: Copy các file đầu vào vào thư mục data trong folder ND01 (tương tự với thư mục British)
* Bước 1: Quay về thư mục tools ( nếu đang ở trong một thư mục con trong thư mục tools )
* Gõ câu lệnh: **cd ..**

A black screen with yellow and purple text

Description automatically generated

* Bước 2: Vào thư mục ND01 bằng git bash:
* Gõ câu lệnh: **cd ND01**

A screenshot of a computer program

Description automatically generated

* Chạy file nd01.sh: **bash nd01.sh**

A screen shot of a computer

Description automatically generated

* File đầu ra trong thư mục output

**Tool 5: File QD02:**

* Chuẩn bị dữ liệu: Copy các file đầu vào vào thư mục data trong folder QD02 (tương tự với thư mục British)
* Bước 1: Quay về thư mục tools ( nếu đang ở trong một thư mục con trong thư mục tools )
* Gõ câu lệnh: **cd ..**

A black screen with yellow and purple text

Description automatically generated

* Bước 2: Vào thư mục **QD02** bằng git bash:
* Gõ câu lệnh: **cd QD02**

A screenshot of a computer

Description automatically generated

* Chạy file qd02.sh: **bash qd02.sh**

A screenshot of a computer

Description automatically generated

* File đầu ra trong thư mục output

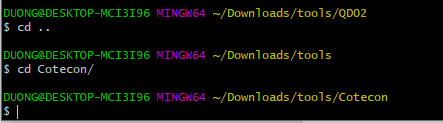
**Tool 6: File Cotecon:**

* Chuẩn bị dữ liệu: Copy các file đầu vào vào thư mục data trong folder Cotecons (tương tự với thư mục British)
* Bước 1: Quay về thư mục tools ( nếu đang ở trong một thư mục con trong thư mục tools )
* Gõ câu lệnh: **cd ..**

A black screen with yellow and purple text

Description automatically generated

* Bước 2: Vào thư mục **Cotecon** bằng git bash:
* Gõ câu lệnh: **cd Cotecon**



* Chạy file cotecon.sh: **bash cotecon.sh**

A screenshot of a computer program

Description automatically generated

* File đầu ra trong thư mục output

**Tool 7: File Woori:**

* Chuẩn bị dữ liệu: Copy các file đầu vào vào thư mục data trong folder Woori (tương tự với thư mục British)
* Bước 1: Quay về thư mục tools ( nếu đang ở trong một thư mục con trong thư mục tools )
* Gõ câu lệnh: **cd ..**

A black screen with yellow and purple text

Description automatically generated

* Bước 2: Vào thư mục **Woori** bằng git bash:
* Gõ câu lệnh: **cd Woori**

A screenshot of a computer

Description automatically generated

* Chạy file woori.sh: **bash woori.sh**

A screenshot of a computer program

Description automatically generated

* File đầu ra trong thư mục output