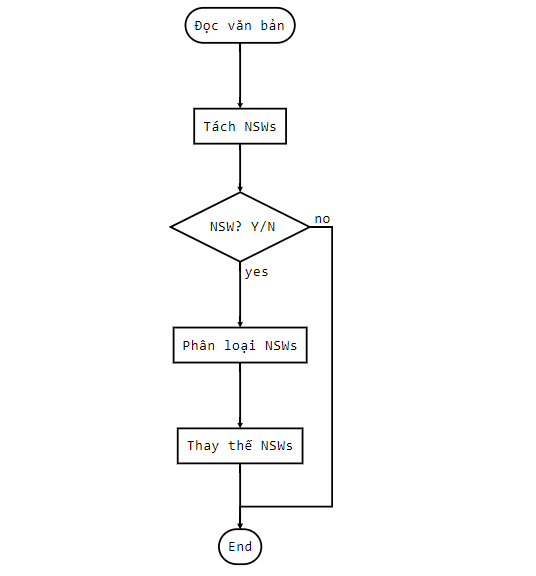
**Normalize-text for TTS**

**1. Active Flow**



**2. Requirements**

- num2words

- beautifulsoup4

- lxml

- pronouncing

- tqdm

- unicodedata

- flair

**3. Function**

- config: chứa dict sử dụng cho từng loại NSW (Các từ điển được lấy từ folder dict)

- build\_en2vi\_dict: tạo thư viện đọc từ tiếng anh từ bộ từ điển của thư viện pronouncing

- en2vi: hàm xây dựng âm tiết tiếng anh theo chuẩn ARPABET

- letter2vi: hàm phân loại NSWs dạng chữ

- numb2vi: hàm phân loại NSWs dạng số

- ultis: chứa các hàm đọc các NSWs đã được phân loại

- utils/tokenizer: hàm tách câu và tách từ

- utils/normalize\_utils: hàm chuẩn hóa câu

**4. Maintain**

- Xử lý lỗi sai sót của hàm en2vi

- Xử lý các trường hợp đặc biệt

- Cải thiện mô hình phân loại NSWs