

Danh sách nhóm:

1. Lê Tấn Lộc - 20521546
2. Nguyễn Hồng Anh Thư - 20520313
3. Bùi Quốc Thịnh - 20520934

Mô tả bài toán

- **Bài toán:** Phát hiện ngôn ngữ được sử dụng trong văn bản.
- **Input:** Văn bản bằng ngôn ngữ tự nhiên.
- **Output:** Đưa ra dự đoán mã loại ngôn ngữ được dùng trong văn bản cùng tỷ lệ phần trăm của ngôn ngữ đó

Text

xin chào

Language:

{'vi': '0.9999997381609155'}

Detect

- **Ý nghĩa:** Phát hiện ngôn ngữ rất hữu ích trong việc xử lý ngôn ngữ tự nhiên và là tiền đề cho các mô hình dịch máy. Nó hỗ trợ máy tính trong việc nhận biết loại ngôn ngữ để từ đó đưa ra những hướng xử lý tiếp theo

Mô tả API được cung cấp, cách sử dụng

- **Mô tả:** API Text Analysis hỗ trợ thực hiện các tác vụ như tóm tắt văn bản, phát hiện ngôn ngữ, phân tích cảm xúc trong văn bản, trích xuất văn bản từ tập tin, nhận dạng vật thể được đặt tên, trích xuất bài báo, trích xuất trang web.
 - + Trích xuất bài báo: Trích xuất dữ liệu quan trọng từ các nguồn tin tức như tiêu đề, văn bản, tóm tắt, từ khóa, tác giả, hình ảnh chính, tất cả hình ảnh, liên kết, v.v.
 - + Trích xuất văn bản từ tập tin: Trích xuất văn bản từ các tệp như "csv, doc, docx, eml, epub, json, html, htm, msg, odt, pdf, pptx, ps, rtf, txt, xls, xlsx, gif, jpg, jpeg, png, tiff, tif, mp3, ogg, wav".
 - + Phát hiện ngôn ngữ: Phát hiện ngôn ngữ trong đoạn văn bản.
 - + Nhận dạng thực thể được đặt tên: Nhận dạng thực thể được đặt tên là nhiệm vụ khai thác thông tin nhằm xác định vị trí và phân loại các thực thể được đặt tên được đề cập trong văn bản phi cấu trúc thành các danh mục được xác định trước như tên người, tổ chức, địa điểm, biểu thức thời gian, số lượng, giá trị tiền tệ, tỷ lệ phần trăm, ...
 - + Phân tích cảm xúc: Phân tích cảm xúc là việc sử dụng xử lý ngôn ngữ tự nhiên để xác định và trích xuất thông tin chủ quan một cách có hệ thống. Phân tích tình cảm được áp dụng rộng rãi cho các bài đánh giá phim và phản hồi khảo sát.

- + Tóm tắt văn bản: Tóm tắt đoạn văn bản dài hoặc câu chuyện.
- + Trích xuất trang web: Trích xuất nội dung từ các trang web chung chung.

Tuy nhiên, ở demo này chúng em chỉ sử dụng tác vụ phát hiện ngôn ngữ của API

- **Cách sử dụng:** Sử dụng module requests trong Python với phương thức request. Cú pháp: `requests.request(method, url, args)` hỗ trợ gửi yêu cầu truy cập API đến URL được chỉ định.
 - + URL của API: <https://text-analysis12.p.rapidapi.com/language-detection/api/v1.1>

Code ứng dụng demo:

[LAB05.ipynb - Colaboratory](#)

```
import gradio as gr
import requests
import json

def Language_detection(text):
    url = "https://text-analysis12.p.rapidapi.com/language-detection/api/v1.1"

    payload = {"text": text}
    headers = {
        "content-type": "application/json",
        "X-RapidAPI-Key": "a7833d2622mshd4e0dce16cf7abcp138a03jsnba1cb6751b51",
        "X-RapidAPI-Host": "text-analysis12.p.rapidapi.com"
    }

    response = requests.request("POST", url, json=payload, headers=headers)
    data = json.loads(response.text)
    return(data['language_probability'])

with gr.Blocks() as demo:
    text = gr.Textbox(label="Text")
    lang = gr.Textbox(label="Language: ")
    greet_btn = gr.Button("Detect")
    greet_btn.click(fn=Language_detection, inputs=text, outputs=lang)

demo.launch()
```

Link demo:

<https://28309.gradio.app>