

Modelos de Inteligencia Artificial

Curso de Especialización de Inteligencia Artificial y Big Data
IES Gran Capitán 2024/25



UNIDAD 3.

Procesamiento del Lenguaje Natural y sus aplicaciones. Potencial y limitaciones.

Índice de contenidos

1. Procesamiento de lenguaje natural
2. Potencial de las técnicas existentes de procesamiento de lenguaje. Limitaciones
 1. Potencial de procesamiento del lenguaje natural
 1. Reconocimiento del habla (ASR, Automatic Speech Recognition)
 2. Síntesis de texto o voz
 3. Detección de entidades nombradas (NER, Named Entity Recognition)
 4. Traducción automática
 5. Similitud de textos
 6. Análisis del sentimiento
 2. Modelos y técnicas existentes
 1. BERT (Bidirectional Encoder Representations from Transformers)
 2. Otros modelos
 3. Limitaciones. La ambigüedad
3. Formación del investigador en PLN.
4. Elaboración de un sistema de procesamiento de lenguaje orientado a una tarea específica



2. Potencial de las técnicas existentes de procesamiento de lenguaje. Limitaciones.

2.1. Potencial de procesamiento del lenguaje natural

Potencial del procesamiento del lenguaje natural

Aplicaciones:

- **Chatbots** → el desarrollo de los asistentes de compra o sistemas conversacionales.
- **Análisis del sentimiento y de la opinión** → detectar sesgos y opiniones sobre un tema (rrss o en encuestas).
- **Clasificación automática de documentos.**
- **Búsqueda de similitudes en textos** → plagios, detección de entidades para contextualizar y clasificar texto, búsqueda de información en múltiples idiomas, traducción automática, reconocimiento del habla...

Reconocimiento del habla (ASR, Automatic Speech Recognition)

Disciplina encargada de convertir los fonemas emitidos por un ser humano en espectros de onda de audio captados mediante un dispositivo de entrada de sonido (ej. micrófono) de una máquina, que tras ser procesados dentro del contexto de un modelo lingüístico, den lugar a las grafías escritas del mismo.



Librería de Python capaz de interactuar con diversas API para el reconocimiento de voz.

*Librería Pyaudio ->> gestionar el micrófono
CMU Sphinx ->> trabajo offline*

...



Ejemplos tutorizados sobre el reconocimiento de voz

```
#Instalar pip install SpeechRecognition
#Instalar pip install pyaudio (puede que pida instalar "brew install
portaudio")
import speech_recognition as sr
recognizer = sr.Recognizer()
mic = sr.Microphone()
with mic as source:
    audio = recognizer.listen(source)
text = recognizer.recognize_google(audio, language = 'ES')
print(f'Has dicho: {text}')
```

[Link a video de youtube](#)

Síntesis de texto a voz

Conocido como TTS (Text to speech), contrario a ASR, a partir de un texto escrito lo reproduce por voz mediante un dispositivo de salida de audio (altavoces).



<https://github.com/NVIDIA/NeMo/tree/main/examples/tts>

<https://github.com/NVIDIA/NeMo/tree/main/tutorials/tts>



Se puede instalar una librería genérica de TTS como **espeak TTS** (offline) o **gtts** (online a través de Google). Para reproducirlo se puede emplear cualquier función de salida de audio que se haya instalado como **playsound** o por medio de una librería **os** → **os.system("mpg321 fichero.mp3")**

ASR y TTS

Bien de manera escrita, o bien mediante el auxiliar de las aplicaciones de reconocimiento de voz (ASR) y de síntesis de texto a voz (TTS), los agentes conversacionales se emplean para multitud de tareas, desde ser:

- Asistentes al usuario → ventas
- Resolutores de problemas postventas → métricas y procedimiento de control de calidad y satisfacción al cliente.



Librería de Phyton para síntesis de texto a voz

Ejemplos tutorizados sobre síntesis de texto a voz

```
"""
```

```
Convertir texto a voz (TTS) con Python usando gTTS
```

```
Ejemplo 3: Escribir welcome en un archivo abierto con open, en idioma inglés
```

```
"""
```

```
from gtts import gTTS
```

```
tts = gTTS('welcome', lang='en')
```

```
tts.save('welcom.mp3')
```

Ejemplos de gtts

Ejemplos tutorizados sobre síntesis de texto a voz

```
"""
```

```
    Convertir texto a voz (TTS) con Python usando gTTS
```

```
    Ejemplo 4: Escribir múltiples idiomas en un archivo
```

```
"""
```

```
from gtts import gTTS
```

```
tts_en = gTTS('hello', lang='en')
```

```
tts_fr = gTTS('bonjour', lang='fr')
```

```
with open('hello_bonjour.mp3', 'wb') as f:
```

```
    tts_en.write_to_fp(f)
```

```
    tts_fr.write_to_fp(f)
```

Ejemplos de gtts

Ejemplos tutorizados sobre síntesis de texto a voz

```
"""
```

```
    Convertir texto a voz (TTS) con Python usando gTTS
```

```
    Ejemplo 5: Escribir hola mundo en un archivo, en idioma español y después  
    reproducirlo
```

```
"""
```

```
from gtts import gTTS
```

```
from playsound import playsound
```

```
NOMBRE_ARCHIVO = "sonido_generado.mp3"
```

```
tts = gTTS('Hola mundo. Estamos convirtiendo texto a voz con Python.' , lang='es-us')
```

```
# Nota: podríamos llamar directamente a save
```

```
with open(NOMBRE_ARCHIVO, "wb") as archivo:
```

```
    tts.write_to_fp(archivo)
```

```
playsound(NOMBRE_ARCHIVO)
```

Ejemplos de gtts