



GENERALITAT
VALENCIANA
Conselleria de Educació,
Cultura i Esport

TOTS
A UNA
VEU



Carrer d'Illueca, 28
03206 Elx, Espanya
03013224.secret@gva.es
(+34) 966.912.260



Fondo Social Europeo
El FSE invierte en tu futuro
UNIÓN EUROPEA

ENUNCIADO

En esta práctica vamos a ampliar las funcionalidades de la aplicación realizada en la práctica anterior. En concreto, vamos a llevar a cabo, además, un análisis de sentimiento del texto introducido por el usuario y un despliegue de la aplicación para que esté disponible online.

Parte 1 (5 puntos)

Para ello, vamos a utilizar dos librerías diferentes, una para el idioma inglés y otra para castellano:

- Inglés: NLTK Vader (**2 puntos**)
- Castellano: spanish_sentiment_analysis (**2 puntos**)

Apóyate en este [vídeo](#) y modifica el código existente.

Como resultado del análisis de sentimiento, muestra por pantalla también el valor de la polaridad calculado, según el idioma introducido (**1 punto**).

Adjunta a la entrega el enlace al repositorio de Github.

Parte 2 (5 puntos)

Basándote en este [enlace](#), haz el despliegue de tu aplicación en Flask a la plataforma Heroku. Adjunta a la entrega:

- Enlace a tu aplicación web.
- Un pequeño documento donde se demuestre cómo funciona.

Parte 1:

Repositorio GitHub: <https://github.com/Zurichk/NLP-Sentimientos-spaCy>

Comparando las 2 librerías:

(Nota, para la parte en español, no he podido utilizar la librería `spanish_sentiment_analysis`, ya que he tratado de instalarla de diferentes formas, pero siempre obtenía errores, y he utilizado la librería `pysentimiento` <https://github.com/pysentimiento/pysentimiento/>.)

```
#Análisis de sentimientos
#nltk.download('vader_lexicon')
from nltk.sentiment.vader import SentimentIntensityAnalyzer #Inglés
sid = SentimentIntensityAnalyzer()

#https://github.com/pysentimiento/pysentimiento/
from pysentimiento import analyzer #Español
sentiment_es = analyzer.SentimentAnalyzer(lang='es')
#emotion_es = analyzer.EmotionAnalyzer(lang='es')
#from classifier import SentimentClassifier #Español
#clf = SentimentClassifier()
```

La librería Vader obtiene mejores resultados con frases cortas, realizamos varias pruebas con los mismos textos en inglés y español.

Texto de prueba Español:

Adrián García estuvo en Madrid y visitó la tienda Apple, el pasado jueves 02 de Marzo de 2022, y fue atendido por Charles, que hablaba en inglés, durante 2 horas, y se gastó unos 600 euros en un Iphone 10.

Por la tarde, se fue a pasear por el parque del Retiro y estuvo con Miguel Pérez durante 3 horas. Los 2 estaban muy felices.

Test text English:

Adrián García was in Madrid and visited the Apple store, last Thursday, March 02, 2022, and was attended by Charles, who spoke in english, for 2 hours, and spent about 600 euros on an Iphone 10. In the afternoon, He went for a walk in the Retiro park and was with Miguel Pérez for 3 hours.

Resultado en Español:

Resultado

Número de registros: 9

- ('PER', 'Adrián García')
- ('LOC', 'Madrid')
- ('ORG', 'Apple')
- ('PER', 'Charles')
- ('MISC', 'inglés')
- ('MISC', 'Iphone 10')
- ('LOC', 'parque del Retiro')
- ('PER', 'Miguel Pérez')
- ('MISC', 'Los 2')

Análisis de Sentimientos:

- SentimentOutput(output=POS, probas=(POS: 0.997 NEU: 0.002, NEG: 0.001))

Resultado en Ingles:

Resultado

Número de registros: 12

- ('PERSON', 'Adrián García')
- ('GPE', 'Madrid')
- ('ORG', 'Apple')
- ('DATE', 'last Thursday, March 02, 2022')
- ('PERSON', 'Charles')
- ('LANGUAGE', 'english')
- ('TIME', '2 hours')
- ('MONEY', 'about 600 euros')
- ('TIME', 'the afternoon')
- ('GPE', 'Retiro')
- ('PERSON', 'Miguel Pérez')
- ('TIME', '3 hours')

Análisis de Sentimientos:

- {'neg': 0.0, 'neu': 1.0, 'pos': 0.0, 'compound': 0.0}

Vemos que nos devuelve un resultado totalmente neutral.

```
Idioma Español: cargamos es_core_news_lg
SentimentOutput(output=POS, probas={POS: 0.997, NEU: 0.002, NEG: 0.001})
127.0.0.1 - - [12/Apr/2022 11:49:24] "POST /process HTTP/1.1" 200 -
127.0.0.1 - - [12/Apr/2022 11:49:24] "GET /main.css HTTP/1.1" 404 -
127.0.0.1 - - [12/Apr/2022 11:49:24] "GET /main.js HTTP/1.1" 404 -
Idioma Ingles: cargamos en_core_web_md
{'neg': 0.0, 'neu': 1.0, 'pos': 0.0, 'compound': 0.0}
127.0.0.1 - - [12/Apr/2022 11:51:16] "POST /process HTTP/1.1" 200 -
127.0.0.1 - - [12/Apr/2022 11:51:16] "GET /main.css HTTP/1.1" 404 -
127.0.0.1 - - [12/Apr/2022 11:51:16] "GET /main.js HTTP/1.1" 404 -
```

Ahora vamos a analizar frases un poco más cortas.

Hola me llamo Adrian, hoy está lloviendo, y estoy muy triste.

Resultado

Número de registros: 1

- ('PER', 'Adrian')

Análisis de Sentimientos:

- SentimentOutput(output=NEG, probas={NEG: 0.999, NEU: 0.000, POS: 0.000})

```
Idioma Español: cargamos es_core_news_lg
SentimentOutput(output=NEG, probas={NEG: 0.997, NEU: 0.002, POS: 0.000})
127.0.0.1 - - [12/Apr/2022 12:16:25] "POST /process HTTP/1.1" 200 -
127.0.0.1 - - [12/Apr/2022 12:16:25] "GET /main.css HTTP/1.1" 404 -
127.0.0.1 - - [12/Apr/2022 12:16:25] "GET /main.js HTTP/1.1" 404 -
```

Hello, my name is Adrian, today it is raining, and I am very sad.

Resultado

Número de registros: 2

- ('PERSON', 'Adrian')
- ('DATE', 'today')

Análisis de Sentimientos:

- {'neg': 0.22, 'neu': 0.78, 'pos': 0.0, 'compound': -0.5256}

```
Idioma Ingles: cargamos en_core_web_md
{'neg': 0.22, 'neu': 0.78, 'pos': 0.0, 'compound': -0.5256}
127.0.0.1 - - [12/Apr/2022 12:17:14] "POST /process HTTP/1.1" 200 -
127.0.0.1 - - [12/Apr/2022 12:17:14] "GET /main.css HTTP/1.1" 404 -
127.0.0.1 - - [12/Apr/2022 12:17:14] "GET /main.js HTTP/1.1" 404 -
```

El compound (la suma de los 3 valores) nos devuelve un valor negativo.

Otro dato a tener en cuenta es que Vader nos devuelve valores entre -1 y 1, mientras que pySentimiento devuelve valores entre 0 y 1, por eso a la hora de obtener los resultados, esta última nos devuelve valores más cercanos a 1.

Vamos a realizar otro ejemplo con una frase con 2 sentimientos encontrados.

Cuando sale el sol estoy contento y cuando llueve estoy triste.

Obtenemos un resultado Neutral:

Análisis de Sentimientos:

- SentimentOutput(output=NEU, probas={NEU: 0.984, NEG: 0.015, POS: 0.001})

When the sun rises I am happy and when it rains I am sad.

Análisis de Sentimientos:

- {'neg': 0.185, 'neu': 0.595, 'pos': 0.22, 'compound': 0.1531}

Si esta misma frase cambiamos la palabra SAD a mayúsculas, también cambia la polaridad a negativo, dando mayor importancia a la palabra en mayúsculas.

When the sun rises I am happy and when it rains I am SAD.

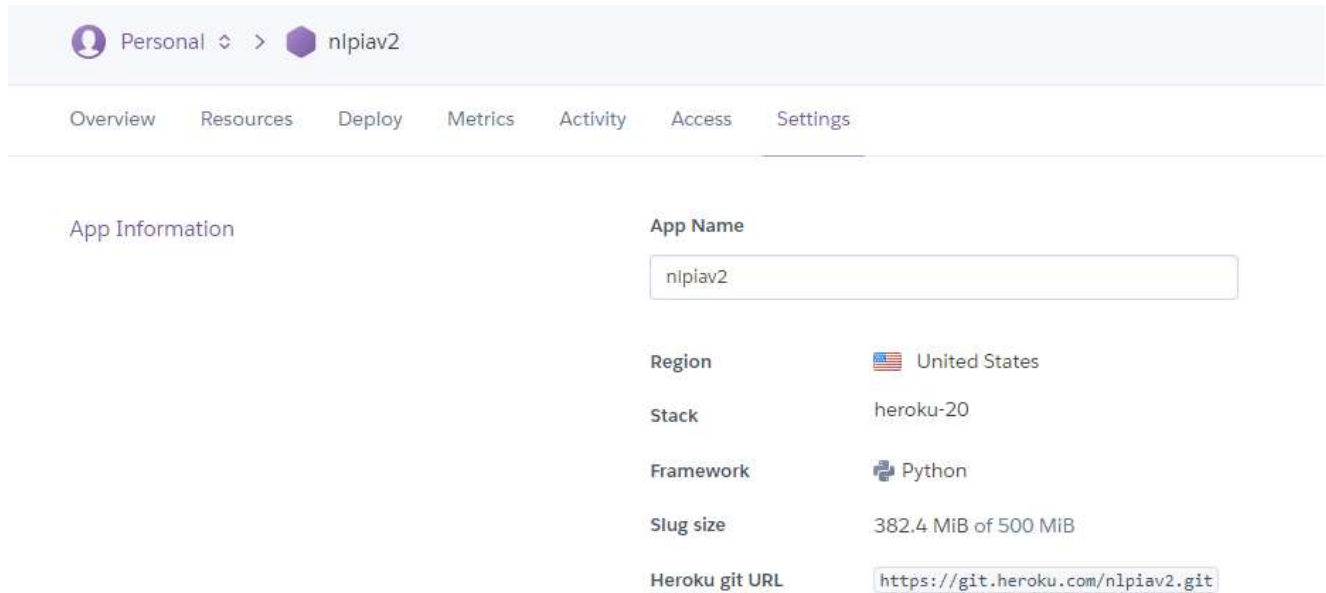
Análisis de Sentimientos:

- {'neg': 0.219, 'neu': 0.57, 'pos': 0.211, 'compound': -0.0343}

Parte 2

Haz el despliegue de tu aplicación en Flask a la plataforma Heroku.

Para esta parte he tenido que quitar la parte en español ya que la aplicación con Español ocupaba 1.1GB comprimido, y Heroku solo permite 500Mb.



Link: <https://nlpiav2.herokuapp.com/>

Comandos utilizados para desplegar la app.

Generamos el requirements:

```
pip freeze > requirements.txt
```

Una vez creada la cuenta en heroku. Desde nuestro entorno virtual(cmd):

```
heroku login (para loguearnos)
```

Generamos o creamos el archivo Procfile, indicando el directorio donde estará nuestra app

```
echo web: gunicorn --chdir code app:app > Procfile
```

Iniciamos git:

```
git init
git add Procfile requirements.txt
git commit -m "Agregar archivos de implementación de Heroku"
```

Creamos la app en heroku y subimos los cambios.

```
heroku create nlpiav2
git push heroku master
```

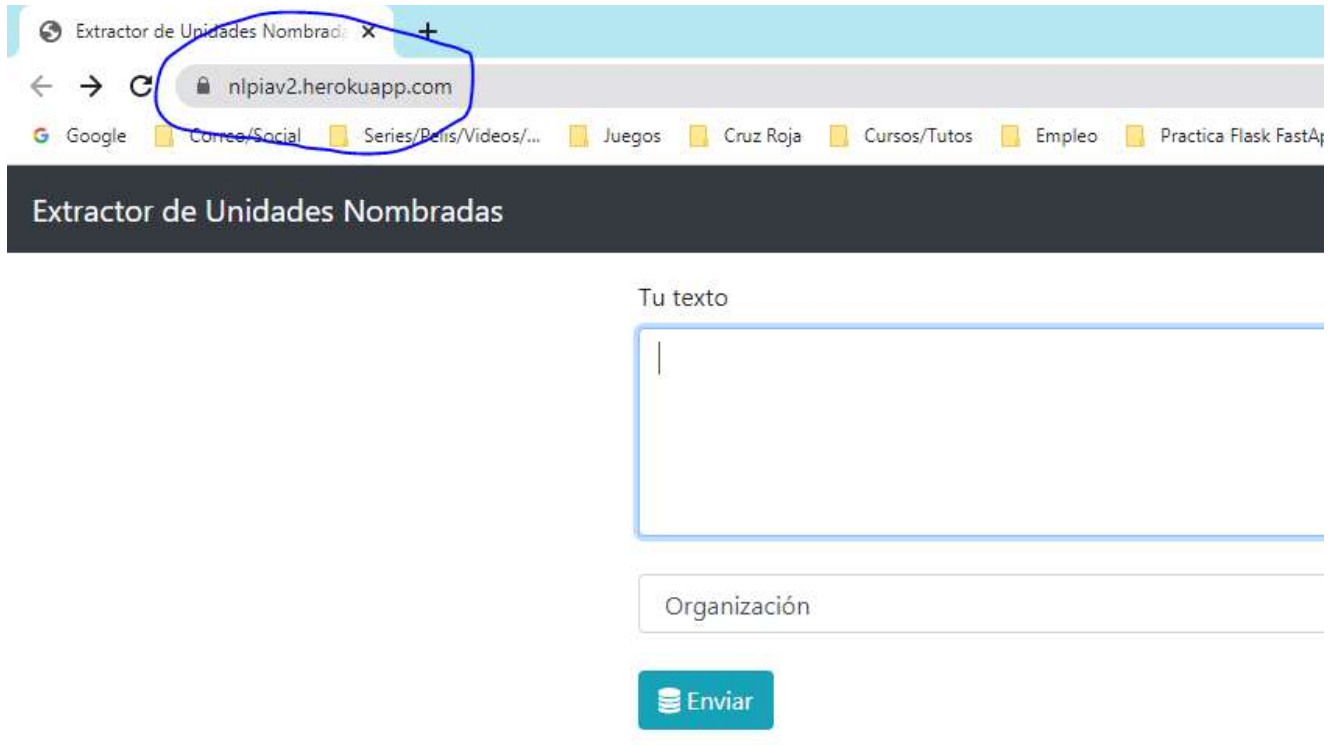
Una vez se han instalado todas las librerías incluidas en el requirements, comprobamos que todo está correcto.

Y subimos la carpeta code con la app.py y templates

```
git add code
git commit -m "subo code"
git push heroku master
```

En otra ventana de cmd, nos conectamos a la consola de heroku, para cargar en_core_web_md:
heroku run bash
(e instalamos el paquete de idioma) python -m spacy download en_core_web_md

Lanzamos la aplicación:
heroku open



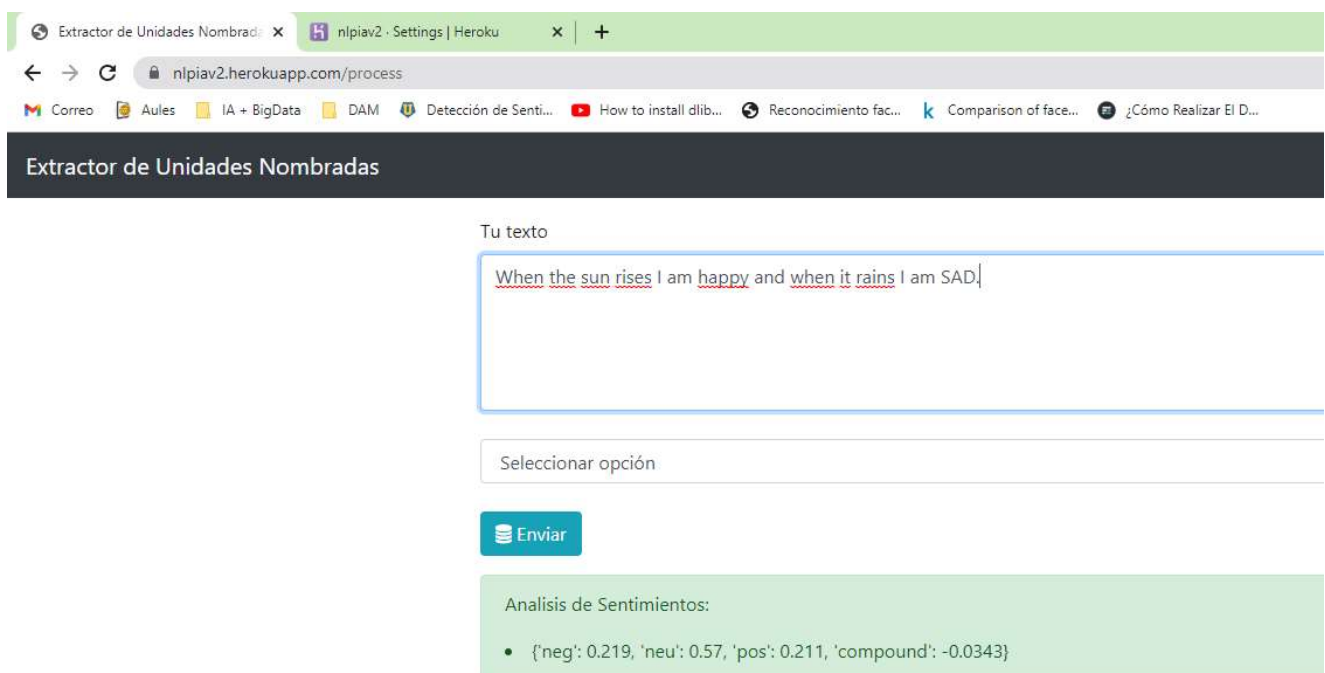
Extractor de Unidades Nombradas

Tu texto

Organización

Enviar

Comprobamos que funciona correctamente, con alguno de los textos utilizados anteriormente.



Extractor de Unidades Nombradas

Tu texto

Seleccionar opción

Enviar

Análisis de Sentimientos:

- {'neg': 0.219, 'neu': 0.57, 'pos': 0.211, 'compound': -0.0343}

Tu texto

Hello, my name is Adrian, today it is raining, and I am very sad.

Seleccionar opción

Enviar

Resultado

Número de registros: 2

- ('PERSON', 'Adrian')
- ('DATE', 'today')

Análisis de Sentimientos:

- {'neg': 0.22, 'neu': 0.78, 'pos': 0.0, 'compound': -0.5256}