



## Sistemas de Ayuda a la Decisión

### Ejercicios sobre el preprocesado

Nombre y Apellidos:

---

## Índice

<b>1. Enunciado</b>	<b>1</b>
<b>2. Recursos de Ayuda</b>	<b>1</b>
<b>3. Anexo I</b>	<b>2</b>
<b>4. Anexo II</b>	<b>2</b>
<b>5. Anexo II</b>	<b>2</b>

---

## 1. Enunciado

1. Suponiendo una tarea de NER (Named Entity Recognition), donde hay que identificar si el texto contiene nombres propios y los que no. Explica brevemente los preprocesos que aplicarías, como quedarían los pequeños textos tras esos preprocesos y calcula BOW en sus versiones one-hot y frecuencia con identificador Id 1 de datos del texto que encontrarás en el Anexo (nota: el diccionario/Vector también está en el Anexo I).
2. Suponiendo una tarea de clasificación de sentimientos en mensajes neutros o que contienen sentimientos, explica brevemente los preprocesos que aplicarías, como quedaría el mensaje tras esos preprocesos y calcula BOW en sus versiones one-hot y frecuencia con identificador Id 1 de datos de mensajes que encontrarás en el Anexo (nota: el diccionario/Vector también está en el Anexo II).
3. Suponiendo una tarea de clasificación de SPAMS, explica brevemente los preprocesos que aplicarías, como quedaría el mensaje tras esos preprocesos y calcula el tf-idf del SMS con identificador Id 1 de datos de SMS-s que encontrarás en el Anexo (nota: el diccionario/Vector también está en el Anexo III).

## 2. Recursos de Ayuda

Existen librerías de Procesamiento del Language Natural (NLP) que son muy útiles para realizar los preprocesos necesarios y para convertir los textos en

vectores numéricos (tf-idf, sent2vec, word2vec...). Aquí se presentan tutoriales sobre su uso.

1. <https://www.nltk.org/>
2. [https://radimrehurek.com/gensim/auto\\_examples/index.html](https://radimrehurek.com/gensim/auto_examples/index.html)
3. <https://spacy.io/usage/spacy-101>
4. <https://pypi.org/project/emosent-py/>
5. <https://pypi.org/project/sent2vec/>

### 3. Anexo I

Id	Text	Clase Real	Clase Pred.
1	<b>Bilbao, well it is a great city.</b>	0	1
2	I love Bilbao because I was born and I live there.	0	1
3	Do you think that this city is nice?	0	0
4	I do not agree, you shouldn't say that.	0	0
5	If I was living in Madrid I would miss the sea.	0	1
6	The sea is great in the summer	0	0

Cuadro 1: Ejemplos de un Conjunto de textos

### 4. Anexo II

Id	Text	Clase Real	Clase Pred.
1	<b>Sorry!, I did not mean to hurt you.</b>	0	1
2	Good Luck! I hope you have a nice future in front of you in the university.	0	1
3	You are an adult now.	0	0
4	Women, they life longer than men.	0	0
5	Oh, I am so happy for you!!!! I love you!!!	0	1
6	Are women happier than men?	0	0

Cuadro 2: Ejemplos de un Conjunto de mensajes

### 5. Anexo II

<b>Id</b>	<b>SMS</b>	<b>Clase Real</b>	<b>Clase Pred.</b>
1	<b>Sorry!, I will call later</b>	0	0
2	Urgent! Finished class where are you.	0	1
3	Congratulations! One year older! I am calling U...	0	1
4	Remember U owe me 20 pounds	0	1
5	Sorry, K. Did you call me just now ah?	0	0
6	Ok i am on the way to home. Do you offer me dinner tonight?	0	1

Cuadro 3: Ejemplos de un Conjunto de SMS