

### Hoja de Trabajo # 3

#### Problema # 1

##### Partes de una oración

La oración consta de 9 partes, de las cuales 5 son variables y 4 son invariables.

- Sustantivos: nombran personas, lugares o cosas. Pueden ser el sujeto de la oración.
  - Comunes: nombran a cualquier persona, lugar o cosa. Ejemplo: mujer, gato, país.
  - Propios: nombran a una persona, lugar o cosa distinguiéndolas de su clase. Ejemplo: Pablo, Guatemala, Nilo.
  - Concretos: nombran entidades que son perceptibles por los sentidos.
  - Abstractos: nombran ideas o sentimientos.
  - Individuales: nombran a una única persona, lugar o cosa.
  - Colectivos: nombran a un conjunto de seres u objetos.
- Adjetivo: parte de la oración que acompaña al sustantivo para determinarlo o expresar características del mismo.
  - Posesivos: indican pertenencia. Ejemplo: mío, tuyo, suyo.
  - Demostrativos: indican ubicación espacial. Ejemplo: este, aquella, ese
  - Calificativo: señalan alguna cualidad del sustantivo. Ejemplo: lindo, feo
  - Numeral: indican un número determinado.
  - Indefinido: se refieren al número o cantidad de forma imprecisa. Ejemplo: mucho, poco, bastante
- Artículo: se antepone al sustantivo. Pueden ser definidos o indefinidos.
- Pronombre: señalan o representan a personas u objetos ya conocidos por hablante y oyente.
- Verbo: indica acción, estado de ánimo y acontecimientos
- Adverbio: modifica el significado del verbo u otras palabras.
- Preposición: relaciona los elementos de una oración. Pueden indicar procedencia, destino, dirección, lugar, etc.
- Conjunción: parte invariable que se utiliza para enlazar oraciones y establecer relaciones de jerarquía entre ellas.
- Interjección: palabras únicas que no tienen variaciones y se emplean en forma aislada, por lo cual funcionan como oración completa. Ejemplo: ¡uy!

#### Problema #2: Natural Language Processing

NLP es una subrama de lingüística y computer science que estudia la interacción entre computadoras y lenguajes naturales (como el idioma español). Su enfoque principal consiste en el desarrollo de cómo las computadoras pueden procesar grandes cantidades de datos de lenguaje natural.

Dentro de sus principales componentes se encuentran análisis más específicos tales como:

- Análisis morfológico: extrae raíces y unidades léxicas complejas. Una misma palabra puede tener varios significados y esto se puede deducir por el contexto de la oración.
- Análisis sintáctico: analiza la estructura de la frase mediante la gramática de la lengua
- Análisis semántico: extracción del significado de la frase y resolución de ambigüedades estructurales.
- Análisis pragmático: análisis del texto más allá del límite de la frase. Una oración puede significar lo que no está diciendo por elementos tales como la ironía y el sarcasmo.

### Problema # 3

1. `[aA-zZ0-9.0-]+@[a-z]+.[a-z]+`
2. `http[s]*://www.[a-z0-9-]+.(com)?(edu)?(net)?` Para los dominios especificados.  
`http[s]*://www.[a-z0-9-]+.[a-z]+` para todos los dominios
3. `[0-9a-fA-F]+`
4. `\b(25[0-5]|2[0-4][0-9]|1[0-9][0-9])?`
5. `[0-3]|3[0-1]\d|/[0-1]|1[0-2]/1[0-9]|0-9+`

### Problema # 4

Ver notebook y pdf adjunto en carpeta

```
In [211]: df.describe()
```

```
Out[211]:
```

	day	month	year
count	21000.000000	21000.000000	21000.000000
mean	15.624762	6.466476	2016.686905
std	8.782500	3.468418	1.441601
min	1.000000	1.000000	2014.000000
25%	8.000000	3.000000	2015.000000
50%	16.000000	6.000000	2017.000000
75%	23.000000	10.000000	2018.000000
max	31.000000	12.000000	2019.000000

```
In [212]: df.mean()
```

```
Out[212]: day      15.624762  
month      6.466476  
year      2016.686905  
dtype: float64
```