

Universidad TecMilenio

Nombre del alumno: Alexis Leal Mata

Máster en Inteligencia Artificial

Actividad 1

Explorando datos de corpus con Python

# Archivo de código



# Preguntas

1. ¿Cuántas palabras hay en cada categoría?

En la categoría positivas hay 832564

En la categoría negativa hay 751256

1. ¿Cuántas palabras diferentes (únicas) hay en cada categoría?

Positiva 29750

Negativa 27841

1. ¿Cuáles son las palabras con frecuencia 1 (hápax legomenón) en cada categoría? Muestra las primeras 100 palabras si es que las hay.

Positiva 1055

Negativa 11197

1. ¿Cuáles son las palabras más frecuentes en cada categoría? Mostrar las primeras 70 palabras si es que las hay, ¿en qué difieren en cada categoría?

Es muestra en la imagen del pdf

# Enlace del repositorio

<https://github.com/AlexisLeal/NLP-ComputerVision.git>

# Conclusión

Esta actividad fue demasiado interesante porque vimos cómo usar funciones que nos puede de hacer mucha ayuda para analizar un texto en este caso fue un texto de prueba, pero podríamos realizar con otro texto aparte que me obligo usar conocimientos de programación para realizar los ejercicios espero que las demás actividades sean igual de interesante