

Introducción a Learning Analytics con  
ejemplos prácticos

# **UD 04. Anexo I - Herramientas para el análisis de texto**

---

Autor: Sergi García Barea

Actualizado Noviembre 2023


## Licencia




**Reconocimiento – NoComercial - CompartirIgual (BY-NC-SA):** No se permite un uso comercial de la obra original ni de las posibles obras derivadas, la distribución de las cuales se debe hacer con una licencia igual a la que regula la obra original.

## Nomenclatura

A lo largo de este tema se utilizarán distintos símbolos para distinguir elementos importantes dentro del contenido. Estos símbolos son:

 **Importante**

 **Atención**

 **Interesante**

## ÍNDICE

<b>1. Introducción.</b>	<b>3</b>
<b>2. Herramientas online destacadas</b>	<b>3</b>
<b>3. Ejemplo en Python proporcionado "analizador.py"</b>	<b>3</b>
<b>4. KNIME y el análisis de textos</b>	<b>4</b>
<b>5. Power BI Desktop y el análisis de textos</b>	<b>4</b>
<b>6. ChatGPT y el análisis de textos</b>	<b>4</b>

## UD04. ANEXO I - HERRAMIENTAS PARA EL ANÁLISIS DE TEXTO

### 1. INTRODUCCIÓN.

En este punto vamos a hablar de herramientas que pueden ayudarnos en el análisis de texto. Enlazaremos herramientas online destacadas. Además, como extra, presentamos un programa en Python realizado para el curso (con un fin didáctico) y enlazaremos información sobre cómo realizar extracción de información en textos usando KNIME.

En algunos momentos nos referiremos al término n-grama. Este término se refiere a conjuntos de palabras seguidas, generalmente referidas al conteo. Por ejemplo, “un resumen” es un 2-grama y “herramienta que permite” un 3-grama.

### 2. HERRAMIENTAS ONLINE DESTACADAS

En este punto describimos herramientas online destacadas para el análisis de textos:

- <https://countwordsfree.com/>
  - Herramienta que permite contar palabras, frases, líneas. Además, permite ver frecuencias tanto de palabras como de n-gramas hasta  $n=4$ . También permite en el análisis obviar “stop words” (palabras vacías) en distintos idiomas.
  - Entre otras funciones, indica el tiempo estimado de un texto y el tiempo estimado de locución del mismo, así como la posibilidad de que la web haga de locutor del mismo.
- <http://guidetodatamining.com/ngramAnalyzer/>
  - Analizador de n-gramas de un texto.
- <https://www.browserling.com/tools/remove-punctuation>
  - Herramienta para eliminar símbolos de puntuación online.
- <https://legible.es/>
  - Herramienta para obtener métricas de legibilidad de un texto.
- <https://niram.org/read/>
  - Indica el tiempo estimado de lectura de un texto.
- <https://tobloef.com/text2mindmap/>
  - Herramienta para crear mapas mentales a partir de texto.

### 3. EJEMPLO EN PYTHON PROPORCIONADO “ANALIZADOR.PY”

Con un fin puramente instructivo y por relacionar la temática con la programación, se ha desarrollado para el curso un pequeño programa en Python 3 para realizar análisis de texto.

Este programa nos proporcionará información sobre el texto, palabras, etc. e incluso intentará realizar un resumen basado en intentar localizar de forma rudimentaria frases con mayor carga de contenido y presentarlas ordenadas.

El programa está disponible en “**AnalizadorTextos.zip**”. Incluye comentarios y requiere la instalación de la biblioteca PDFMiner con “pip install pdfminer” y “pip install pdfminer.six”.

Puede utilizarse para:

- Extraer información de ficheros txt. Se usa ***“python analizador.py fichero.txt”***
- Extraer información de ficheros PDF. Se usa ***“python analiazar.py fichero.pdf pdf”***
  - El último parámetro (“pdf”) es para indicar que es ese tipo de fichero.

#### 4. KNIME Y EL ANÁLISIS DE TEXTOS

KNIME es ampliamente utilizado para el análisis y extracción de información de textos.

Para el uso de las mismas, deben instalarse las siguientes extensiones de KNIME:

- “KNIME textprocessing”
- “KNIME textprocessing Spanish Language Pack”

En esta unidad no hemos tratado el uso de KNIME como herramienta de análisis de textos, pero si quieres saber más al respecto, tienes unos tutoriales aquí:

- <https://www.youtube.com/watch?v=E0p5p90onDA>
- <https://files.knime.com/sites/default/files/Text%20Mining%20Course%20for%20KNIME%20Analytics%20Platform.pdf>

#### 5. POWER BI DESKTOP Y EL ANÁLISIS DE TEXTOS

Power BI Desktop permite realizar análisis de textos avanzados utilizando distintos servicios Azure Cognitive Services. Estos servicios permiten extraer información mediante técnicas de procesamiento del lenguaje natural. Estos servicios generalmente son de pago.

#### 6. CHATGPT Y EL ANÁLISIS DE TEXTOS

ChatGPT es un modelo de lenguaje desarrollado por OpenAI que utiliza la arquitectura GPT-3.5 (aunque de pago, ya existe la versión que utiliza GPT-4.0). Se puede acceder desde <https://chat.openai.com/>

ChatGPT es de sobra conocido, pero vamos a intentar describirlo como una inteligencia artificial avanzada diseñada para comprender y generar texto en lenguaje natural. A través de un extenso entrenamiento con diversos conjuntos de datos, ChatGPT ha aprendido patrones lingüísticos y tiene la capacidad de responder a preguntas, generar texto creativo y participar en conversaciones de manera coherente, analizar otros textos, etc.

En el contexto de Learning Analytics, ChatGPT puede ser utilizado para multitud de análisis, entre ellos el análisis de textos educativos. Entre otras cosas, en el análisis de textos puede ser utilizado para:

- Asistencia Virtual: Implementar un asistente virtual basado en ChatGPT que responda a consultas de estudiantes, proporcionando información adicional y clarificaciones sobre temas específicos.
- Evaluación Automática: Utilizar ChatGPT para analizar respuestas de estudiantes y proporcionar retroalimentación automática. Esto puede facilitar la revisión rápida de tareas y exámenes.
- Análisis de Sentimientos: Emplear ChatGPT para analizar el tono y el sentimiento en las

interacciones escritas, lo que puede ser útil para comprender la participación y el compromiso de los estudiantes.

- **Análisis de Retroalimentación Escrita:** Utilizar ChatGPT para analizar la retroalimentación escrita proporcionada por los docentes a los estudiantes. Esto puede incluir la identificación de patrones comunes en los comentarios para evaluar áreas de mejora y fortalezas en el rendimiento académico.
- **Extracción de Información Clave:** Aplicar ChatGPT para extraer información clave de documentos académicos y textos relacionados con el contenido del curso. Esto facilita la identificación de conceptos fundamentales y la creación de resúmenes automáticos.
- **Detección de Plagio:** Integrar ChatGPT en sistemas de detección de plagio para analizar similitudes y disparidades en el estilo de escritura de los estudiantes. Esto puede ayudar a identificar posibles casos de plagio o colaboración no autorizada.
- **Análisis de Interacciones Estudiante-Contenido:** Emplear ChatGPT para analizar las interacciones escritas entre los estudiantes y el contenido del curso. Esto puede revelar patrones de comprensión, áreas de confusión y niveles de participación, ofreciendo información valiosa para ajustar estrategias de enseñanza.
- **Seguimiento de Progresos a través de Escritos Reflexivos:** Implementar ChatGPT en el análisis de escritos reflexivos de los estudiantes. Esto permite un seguimiento detallado de su progreso, identificando cambios en el pensamiento crítico y la comprensión conceptual a lo largo del tiempo.

ChatGPT ofrece una herramienta muy poderosa para analizar y mejorar la interacción textual en el ámbito educativo, brindando nuevas posibilidades para la personalización del aprendizaje y la automatización de tareas relacionadas con el análisis de textos en Learning Analytics.

Por último, recordad qué es fundamental tener en cuenta las normativas de protección de datos al utilizar servicios como ChatGPT. Se debe respetar lo indicado en la Ley Orgánica de Protección de Datos y Garantía de Derechos Digitales (LOPDGDD) en España y el Reglamento General de Protección de Datos (RGPD) en la Unión Europea para la protección de la privacidad y la gestión de datos personales. Este artículo es ilustrativo de posibles problemas con las inteligencias artificiales generativas <https://www.newtral.es/chatgpt-privacidad-ia-generativas-rgpd/20230326/>