

Introducción a Learning Analytics con
ejemplos prácticos

UD 04. Learning Analytics aplicado a análisis de textos (resúmenes, mapas y presentaciones)



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 DE CONTENIDO

Introducción	3
Resúmenes	3
¿En qué consiste?	3
¿Qué objetivos relacionados podemos definir?	3
¿Qué métricas podemos utilizar?	4
¿Qué elementos podemos analizar?	4
¿Cómo podemos presentar la información obtenida?	5
¿Que actuaciones podemos realizar con los alumnos?	5
¿Cómo podemos evaluar si el proceso ha sido correcto?	5
Bibliografía	5
Licencias de elementos utilizados	5
Autores (en orden alfabético)	5

UD04. LEARNING ANALYTICS APLICADO A RESÚMENES Y MAPAS

1. INTRODUCCIÓN

En esta unidad hablaremos del Learning Analytics aplicado al análisis de textos. Para enfocar el tema en elementos concretos, hablaremos de resúmenes, esquemas/mapas y presentaciones.

Los mapas mentales, esquemas y resúmenes son habitualmente utilizados para sintetizar información relacionada con un tema. Además en concreto, en el mundo de la informática es bastante habitual utilizar las llamadas “cheat sheets” que resumen comandos, instrucciones, algoritmos o partes importantes de algún tema, con el fin de ser consultadas de manera rápida mientras se trabaja.

Aunque hay ocasiones en que estos recursos son proporcionados ya realizados a los alumnos, es bastante habitual no proporcionarlos (o no al menos en primera instancia) y pedirle al alumno que realice los suyos propios.

¿Cuál es el objetivo de pedirselo a los alumnos? El objetivo, además de forzar que el alumnado revise a fondo el tema, es trabajar sus habilidades de síntesis. Incluso a veces, se puede pedir un mapa/resumen en los primeros estadios de un tema y otro al finalizar el tema, para que el propio alumno pueda comparar su progresión.

En general, la calidad de un mapa, esquema o resumen de un alumno depende de dos factores:

- Su habilidad para realizar este tipo de recursos.
- Su conocimiento de la materia.

También habitualmente, pedimos a los alumnos que realicen presentaciones. Un error muy común realizado por alumnos (y también por profesores) es realizar presentaciones con cantidades de texto inadecuadas.

En esta unidad, estudiaremos cómo realizar “Learning Analytics” en este tipo de recursos.

2. ¿QUÉ TIPOS DE TÉCNICAS UTILIZAREMOS? ANÁLISIS DE TEXTOS

Para los análisis de esta unidad, utilizaremos técnicas de análisis de texto básicas. No entraremos muy a fondo en ellas, ya que meterse muy a fondo nos llevaría a tocar temas de “Procesamiento del lenguaje natural” y es posible que ese campo se nos escape un poco del objetivo del curso.

Aunque en esta unidad nos centramos en resúmenes, mapas y presentaciones, estas técnicas pueden aplicarse a otro tipo de contenidos que tengan textos. De vosotros depende su buen uso para extraer información.

En el “**Anexo I - Herramientas para el análisis de texto**”, citamos algunas herramientas disponibles online y también describiremos algunas que hemos creado “ad-hoc” en Python para este curso.

3. RESÚMENES

3.1 ¿En qué consiste?

Un resumen consiste en un texto resultado de la transformación de otro texto de mayor envergadura, pero que representa las ideas claves del texto/tema principal.

Generalmente al indicar a un alumno que realice un resumen se le sugiere un umbral de tamaño proporcional a las características (tamaño, complejidad, etc.) del texto original.

3.2 ¿Qué objetivos relacionados podemos definir?

Partiendo de la realización de actividades de resumen, podemos intentar definir unos objetivos tanto individuales como colectivos relacionados con la realización de resúmenes.

! **Atención:** los objetivos propuestos son simplemente ideas para saber qué objetivos podríamos plantearnos. **No es necesario plantearse todos los objetivos, solo aquellos que creamos nos puedan ayudar en un momento dado.**

- **Objetivos (Individuo)**
 - **Sobre un único resumen.**
 - Medir la calidad de un resumen.
 - Realizar sugerencias a un resumen.
 - **Sobre una serie de resúmenes del mismo tema en distintos instantes de tiempo.**
 - Medir incremento de conocimiento del tema en el tiempo.
 - **Sobre resúmenes de distintos temas pero en una misma etapa de trabajo del tema.**
 - Medir incremento de la habilidad de realizar resúmenes, comparando series de resúmenes en un mismo estadio del tema.
- **Objetivos (Colectivos)**
 - **Sobre un único resumen.**
 - Obtener una instantánea del nivel de la clase para un resumen concreto.
 - **Sobre una serie de resúmenes del mismo tema.**
 - Comprobar si la clase está realizando un incremento del conocimiento.
 - **Sobre resúmenes de distintos temas**
 - Medir incremento de la habilidad de realizar resúmenes, comparando series de resúmenes en un mismo estadio del tema.

3.3 ¿Qué métricas podemos utilizar?

Observando los posibles objetivos propuestos, nos planteamos qué datos educativos queremos recolectar y qué métricas podemos usar para ello.

En este modelo que vamos a proponer y con el fin de facilitar la obtención de métrica relacionadas con el resumen, el profesor debe seleccionar palabras y expresiones clave relacionadas con el tema de estudio. Una correcta selección de palabras y expresiones influirá en la calidad de la información obtenida.

! **Atención:** para facilitar la selección palabras y expresiones clave, se puede aplicar sobre el texto original un software que muestre palabras sueltas y expresiones (N-Gramas) ordenados por frecuencia de aparición.

Opcionalmente, se podrá incluir una posible valoración subjetiva del profesor, para el estudio de tendencias.

Procedemos a definir posibles métricas.

- **Métricas extraíbles de un resumen**
 - ¿Su tamaño está dentro del umbral propuesto? [**Verdadero/Falso**]
 - Proporción de palabras/expresiones clave presentes en el resumen. Se extrae de la formula $\frac{PalabrasClavePresentesResumen}{TotalPalabrasClave}$ [**Decimal**]
 - Valoración subjetiva del profesor [**Entero, del 1 al 10**]
 - Palabras/expresiones clave aparecidas de un conjunto [**Conjunto de palabras**]
 - Palabras/expresiones clave no aparecidas de un conjunto [**Conjunto de palabras**]
 - **Importante:** las palabras/expresiones, pueden tanto del contexto del documento, como expresiones útiles para una redacción rica (conectores o similares).
 - Obtener índice de legibilidad (según distintas métricas) del texto.
 - **Importante:** estos índices se basan en general en que cuanto más cortas y simples son sus palabras, más fácil es de leer. Una buena legibilidad, según el nivel, no tiene porque implicar un buen texto.
 - Velocidad de lectura estimada del texto.
- **Para comparar resúmenes**
 - Tendencias en ajuste al umbral propuesto.
 - Diferencias entre la proporción entre palabras clave aparecidas.
 - Diferencia entre valoraciones subjetivas del profesor.
 - Diferencia entre índices de legibilidad.
 - ¿Que palabras/expresiones clave son comunes a resúmenes?
 - ¿Que palabras/expresiones clave no utilizadas son comunes a resúmenes?

3.4 ¿Qué elementos podemos analizar?

Una vez obtenida la información pertinente, podemos plantearnos cuestiones para analizar

A continuación vamos a presentar ejemplos de cuestiones que podrían analizarse:

! **Atención:** las cuestiones para analizar aquí planteadas, son solo sugerencias. No es obligatorio analizar todas, solo a las que por nuestro contexto queremos buscar respuesta.

Individual

- ¿Por qué el alumno ha seleccionado un conjunto de palabras/expresiones clave?
- ¿Por qué motivos el alumnado **no** ha seleccionado un conjunto de palabras/expresiones?
- ¿El alumno sigue la tendencia de ajustarse a los umbrales?
- ¿La legibilidad es la esperada, acorde al nivel del alumno?
- ¿El alumno está progresando en la calidad de sus resúmenes, tanto de forma absoluta como relativa al grupo?

Colectivo

- ¿Los alumnos eligen las palabras/expresiones clave correctamente?
- ¿Los alumnos **no** están seleccionando un conjunto de palabras/expresiones clave?
 - En este y el anterior, puede aplicarse “del contexto” o de “expresiones útiles”.
- ¿Los alumnos se ajustan a los umbrales?
- ¿La legibilidad de los alumnos es la esperada, acorde al nivel del alumno?
- ¿El grupo está progresando en la calidad de los resúmenes?

3.5 ¿Cómo podemos presentar la información obtenida?

Tanto para complementar el análisis de las cuestiones planteadas en el punto anterior como para incluso usarlos como complemento dentro de las actuaciones a realizar, es necesario presentar la información obtenida.

! **Atención:** aquí planteamos algunos ejemplos de representación. Estas representaciones en la práctica deberán adaptarse a las características concretas de quien pueda observarlas para obtener información sobre ellos (analista, profesores, alumnos, etc.)

- **Gráficos circulares**
 - Gráfico circular, indicando a modo de “barra de progreso” el número de palabras clave usadas
- **Mapas de calor**
 - Mapa de calor relacionando alumnos y palabras clave utilizadas por ellos.
- **Resumen**
 - Total palabras, umbral de palabras.
 - Palabras clave usadas/total palabras clave esperadas
 - Listado de palabras clave no utilizadas.

3.6 ¿Qué actuaciones podemos realizar con los alumnos?

Una vez obtenidos los datos, planteadas las preguntas y visualizados los datos, es posible que podamos plantearnos posibles actuaciones.

A continuación, vamos a presentar algunos ejemplos limitados de situaciones genéricas y posibles medidas de actuación:

Individual

- **Al alumno con problemas para ajustarse al umbral propuesto**
 - Reflexionar con el alumno sobre si está ajustándose correctamente o no a los umbrales, así como los motivos de este hecho.
- **Al alumno cuyas palabras clave no son significativas o no se ajustan a las usadas por la mayoría del grupo:**
 - Reflexionar junto a él sobre porque no se han tratado los temas de las palabras clave usadas y no usadas por el alumno.
- Reflexionar sobre el nivel de legibilidad del texto, teniendo en cuenta la complejidad deseada del mismo.

Colectivo

- **Si un número significativo de alumnos no se ajusta al umbral**
 - Reflexionar si las exigencias y orientaciones son adecuadas para que el resumen pueda ajustarse en tamaño al umbral exigido.
- **Sobre palabras clave utilizadas por la mayoría del alumnado**
 - Reflexionar sobre que se ha hecho para que el alumno haya elegido esas palabras clave y sobre si ese era el efecto deseado.
- **Sobre palabras clave no utilizadas por la mayoría del alumnado**
 - Reflexionar sobre las orientaciones dadas al alumno para realizar el resumen o las características del mismo eran adecuadas para incluir los temas no tratados (tamaño sugerido, lenguaje, etc.)

4. MAPAS/ESQUEMAS

Los mapas conceptuales/mentales y los esquemas, al basarse su análisis en la presencia de palabras clave, se puede aplicar todo lo comentado al análisis de resúmenes (detección de palabras clave, análisis de palabras utilizadas, etc.).

Además, el análisis de mapas puede ser complementado con elementos de análisis de grafos. Sin entrar en complejidades relacionadas con análisis de grafos, un dato sencillo y que puede ser útil es la profundidad (es decir, la distancia entre el padre y el hijo más lejano en el mapa conceptual).

Para saber más, posees información en este artículo y en los que son citados dentro del mismo https://www.researchgate.net/publication/312727056_Aplicacion_de_tecnicas_de_mineria_de_grafo_para_el analisis_de_texto

También indicar que hay una potente herramienta de análisis de grafos de código abierto llamada Gephi <https://gephi.org/>

5. PRESENTACIONES

En las presentaciones, debemos tener en cuenta dos tipos de análisis.

El primero, al igual que resúmenes, mapas y esquemas, es el análisis de la presencia de palabras clave, utilizando el proceso de análisis de resúmenes (detección de palabras clave, análisis de palabras utilizadas, etc.).

El segundo, es el tiempo de lectura estimado de cada transparencia. Podemos medirlo por tiempo estimado o por número de palabras. Además, es conveniente, al igual que en resúmenes, fijar un umbral de palabras a usar en cada transparencia.

5.1 Tiempo de lectura en presentaciones

Un error muy común es hacer presentaciones con transparencias que incluyen mucho texto. De hecho, un error muy común en el ámbito educativo es crear “apuntes/transparencias”, que al final son “apuntes copiados en formato transparencia”.

Al final no son útiles como transparencias (en una presentación, exigir una lectura detallada cansa a quien lo lee y evita que se fije en otros detalles) y apenas aportan ventajas a los alumnos respecto a apuntes en otros formatos.

De hecho, por este motivo, tienen en la actualidad un gran éxito a la hora de comunicar formatos como la PechaKucha <https://es.wikipedia.org/wiki/PechaKucha> que consiste en hacer una presentación con 20 diapositivas y dedicando 20 segundos a cada diapositiva.

Herramientas como <https://niram.org/read/> pueden ayudarte a conocer el tiempo de lectura de un texto.

Por ejemplo, en una diapositiva de 20 segundos, si vamos a incluir texto relacionado con los condicionales de programación, analizamos los siguientes posibles textos. Suponemos que al alumno le fijamos un umbral de entre 2 y 15 palabras por transparencia:

- “Los condicionales: if (si en español) es un bloque donde se entra si se cumple una condición. Si no se cumple, va al else (sino en español) ”
 - **Tiempo lectura estimado:** 8 segundos.
 - **Palabras utilizadas:** 25 (no se ajusta al umbral).
 - Se ha perdido la gran mayoría de la transparencia, incluidas explicaciones, posibles elementos visuales o parte del texto.
- Los condicionales: “Condicionales: if se entra si se cumple una condición. Si no se cumple, entra al else”
 - **Tiempo de lectura estimado:** 4 segundos.
 - **Palabras utilizadas:** 16 (no se ajusta al umbral, aunque por poco).
 - Más adecuado que el anterior, permite un mayor foco en lo que cuenta el ponente, aunque en muchos casos merece la pena algo más corto
- “Condicionales: if y else”
 - **Tiempo de lectura estimado:** 1 segundo.
 - **Palabras utilizadas:** 3 (se ajusta acomodadamente al umbral).
 - Acompañado de una explicación de unos 20 segundos, el oyente puede focalizarse en lo que se cuenta.

6. CONCLUSIÓN

Hemos hecho distintas reflexiones sobre objetivos, métricas, cuestiones, visualizaciones e incluso actuaciones que podemos hacer con análisis de textos.

Hemos hablado ampliamente de cómo realizarlo en resúmenes y hemos dado pinceladas de adaptación a realizarlos con mapas conceptuales/mentales y presentaciones, haciendo hincapié en estas últimas en la velocidad de lectura y el número de palabras por transparencia.

7. BIBLIOGRAFÍA

[1] Speech and language processing (Dan Jurafsky) <https://web.stanford.edu/~jurafsky/slp3/>

[2] Handbook of Learning Analytics (Charles Lang, George Siemens, Alyssa Wise, Dragan Gašević) https://www.researchgate.net/publication/324687610_Handbook_of_Learning_Analytics