

# Proyecto 1 - Codificador de fuente para una fuente discreta sin memoria

Sistemas de Comunicaciones Digitales I - 2024

## Objetivos

1. Investigar en la bibliografía lo siguiente:
  1. Conceptos de fuente de información discreta y fuente discreta sin memoria (DMS)
  2. Concepto de entropía. Entropía de una DMS
  3. Concepto de Codificación de Fuente. El código Unicode y la codificación UTF-8
  4. Codificación óptima para una DMS. Teorema de codificación de fuente
  5. Codificación Huffman como código óptimo para una DMS
  6. La ley de los grandes números y su aplicación a la estimación de probabilidades de ocurrencia de símbolos de una fuente discreta sin memoria en base a una muestra de su salida
2. Utilizando un lenguaje de programación (se recomienda Python)
  1. Realizar una rutina que, tomando como entrada un archivo de texto utf-8, cree una tabla de frecuencias de ocurrencia de caracteres en dicho archivo.
  2. Realizar una clase (o equivalente) que simule una fuente discreta sin memoria en base a una tabla de símbolos con sus probabilidades, permitiendo producir cadenas de símbolos
  3. Realizar una rutina que genere un codificador de Huffman para una fuente discreta sin memoria dada su tabla de probabilidad de ocurrencia de símbolos
  4. Obtener de [Proyecto Gutenberg](#) la versión en texto plano utf-8 de algún libro.
    1. Crear a partir de dicho libro una simulación de fuente discreta sin memoria y un codificador Huffman.
    2. Generar cadenas de símbolos de la fuente discreta sin memoria y codificarlas con el codificador de Huffman.
    3. Evaluar el rendimiento del codificador. ¿Cúmple con el Teorema de codificación de fuente?

## Entregables

Todos los entregables estarán contenidos en un repositorio git público alojado en GitHub. La entrega se hará mediante un enlace al repositorio. Los entregables son:

- Un informe conteniendo los resultados de la investigación bibliográfica y las prácticas realizadas.

- El código desarrollado durante las prácticas, que permite reconstruir los resultados presentados.

A continuación se indica la estructura del informe, el contenido de cada sección y los criterios de evaluación que se aplicarán para su calificación.

- *Título.* El título será *Codificador de fuente para una fuente discreta sin memoria*
- *Autor.* Debe indicar el nombre completo
- *Resumen.* (1 punto)
  - Describir brevemente el contenido del informe
  - Criterio de evaluación: Explica de forma clara de que trata el informe. Esto es, sin leer el resto del informe y solo en base al resumen un lector puede comprender que se investigó, que prácticas se llevaron a cabo y una idea general de sus resultados.
- *Introducción.* (2 puntos)
  - Exponer los resultados de la investigación del objetivo 1.
  - Plantear los objetivos para el resto del proyecto (sub-objetivos del objetivo 2)
  - Criterios de evaluación.
    - \* Expone de forma clara los temas 1.1 a 1.6
    - \* Incluye figuras y tablas necesarias
    - \* Cita las fuentes consultadas según normas APA. Consultar (Sánchez, 2019) para más detalles
- *Desarrollo.* (2,5 puntos)
  - Explicar el desarrollo de cada punto del objetivo 2
  - Incluir las figuras, tablas, fórmulas y pseudocódigo necesarios para ilustrar lo presentado.
  - Criterios de evaluación.
    - \* Cita las fuentes consultadas según normas APA (Sánchez, 2019).
    - \* Explica en alto nivel el diseño del software realizado
    - \* Toda figura tiene un epígrafe y número de figura, toda tabla un título y número de tabla, todo listado de pseudocódigo un título y número de listado.
    - \* Toda figura, tabla y listado son referenciados en el texto por sus correspondientes números (ej. “el Listado 1...”)
- *Resultados y discusión* (2,5 puntos)
  - Iniciar con un resumen de los principales resultados obtenidos
  - Exponer la interfaz de programación de aplicaciones (API) del software desarrollado
  - Presentar los resultados del objetivo 2.4
  - Contrastar los resultados obtenidos con la teoría
  - Criterios de evaluación.

- \* De leer el resumen es posible conocer en general los resultados obtenidos y si están de acuerdo a lo esperado
- \* La API expuesta permite hacer uso del software realizado
- \* Utiliza tablas, listados y figuras según sea necesario para mostrar los resultados obtenidos
- \* Toda tabla, figura y listado que se incluya contiene información relevante, es numerada y referenciada correctamente en el texto
- \* Se discuten los resultados obtenidos indicando si están de acuerdo con lo esperado
- *Conclusiones* (1 punto)
  - Explica brevemente lo aprendido durante la realización del proyecto
  - Criterios:
    - \* Indica los conceptos explorados durante la realización del proyecto
    - \* Contrasta los resultados obtenidos con los objetivos y lo esperado del desarrollo teórico
- *Referencias* (1 punto)
  - Lista las referencias utilizadas, utiliza formato APA (Sánchez, 2020). No es necesario respetar exactamente las medidas indicadas en la norma, solo el estilo general y el contenido requerido.
  - Criterios:
    - \* Todas las referencias corresponden a citas en el texto
    - \* Respetar el formato APA

## Referencias

- Sánchez, C. (08 de febrero de 2019). Cita de parafraseo. Normas APA (7ma edición). <https://normas-apa.org/citas/cita-de-parafraseo/>
- Sánchez, C. (24 de enero de 2020). Referencias APA. Normas APA (7ma edición). <https://normas-apa.org/referencias/>