

# Control 2

CC5213 – Recuperación de Información Multimedia

Departamento de Ciencias de la Computación

Universidad de Chile

Profesor: Juan Manuel Barrios

Fecha entrega: viernes 16 de mayo de 2025

## Semana 05

1. Suponga que descarga la canción “Despacito”, de 3 minutos y 49 segundos de duración, y guarda el audio en un archivo llamado “A.raw” en formato PCM de **16 bits** (s16le) y sample rate de **22.050**.
  - a. (0.25 puntos) Señale el tamaño en bytes de “A.raw” y la frecuencia máxima audible al reproducirlo.
  - b. (0.25 puntos) Suponga que descarga la misma canción, pero guarda el audio en un archivo llamado “B.raw” en formato PCM de **32 bits** (32f) y sample rate **11.025**. ¿Al reproducir “B.raw” oirá un sonido de mejor o peor calidad que “A.raw”? Justifique en máximo dos líneas.
2. Suponga que tiene un audio de **48.000** samples por segundo al que le desea realizar un análisis de espectro con una Transformada de Fourier usando ventanas de tamaño 40 ms.
  - a. (0.25 puntos) ¿Cuántos coeficientes tendrá la transformada de Fourier de cada ventana?
  - b. (0.25 puntos) Suponga que la Transformada de Fourier de una ventana obtiene alta energía para una señal de 40 ciclos ¿Qué frecuencia de sonido es audible en esa ventana?

3. De acuerdo al video “05.5-Descriptores de audio y MFCC” publicado en la sección de Enlaces de U-Cursos ([ver acá](#)) responda las siguientes preguntas:
- (0.25 puntos) En el minuto 19:35 se da una razón de por qué es mala idea buscar audios similares comparando sus coeficientes de la transformada de Fourier. Explique en máximo dos líneas cómo la escala Mel evita el problema mencionado.
  - (0.25 puntos) En el minuto 25:20 se menciona qué significa o qué representa el primer coeficiente del descriptor MFCC. Señale si para el caso de la tarea 2 es buena idea o mala idea descartar el primer coeficiente de MFCC. Justifique en máximo dos líneas.
4. En un concurso de TV, dos participantes **A** y **B** deben imitar lo mejor posible cierto sonido **M**. Cada participante registra su intento en un archivo de audio. Para definir al ganador, cada grabación se divide en ventanas de audio, para cada ventana se calcula un descriptor adhoc al concurso. Los descriptores de cada concursante se comparan con los descriptores del audio modelo mediante una función de distancia adhoc, obteniendo las siguientes matrices de distancias:

Matriz de distancia entre descriptores del concursante A y audio M:

	<b>M<sub>1</sub></b>	<b>M<sub>2</sub></b>	<b>M<sub>3</sub></b>	<b>M<sub>4</sub></b>
<b>A<sub>1</sub></b>	2	10	2	3
<b>A<sub>2</sub></b>	11	3	11	8
<b>A<sub>3</sub></b>	8	2	8	7
<b>A<sub>4</sub></b>	1	9	3	2
<b>A<sub>5</sub></b>	3	7	3	2

Matriz de distancia entre descriptores del concursante B y audio M:

	<b>M<sub>1</sub></b>	<b>M<sub>2</sub></b>	<b>M<sub>3</sub></b>	<b>M<sub>4</sub></b>
<b>B<sub>1</sub></b>	2	8	6	1
<b>B<sub>2</sub></b>	2	12	2	5
<b>B<sub>3</sub></b>	13	3	13	10
<b>B<sub>4</sub></b>	2	10	2	3
<b>B<sub>5</sub></b>	2	8	4	1
<b>B<sub>6</sub></b>	2	8	6	1

- a. (0.3 puntos) Calcule la distancia **DTW** entre **A** y **M**. Señale la alineación óptima entre las ventanas de **A** y de **M**, y el costo total de esa alineación.
- b. (0.3 puntos) Calcule la distancia **DTW** entre **B** y **M**. Señale la alineación óptima entre las ventanas de **B** y de **M**, y el costo total de esa alineación.
- c. (0.2 puntos) Señale el ganador del concurso. Justifique en no más de dos líneas por qué es el ganador.

Hint: Revise el ejemplo de cálculo de DTW publicado en "[Slides 05.3-Distancia DTW.pdf](#)" en la sección Material Docente de U-Cursos ([ver acá](#)).

## Semana 06

1. (0.6 puntos) Para las siguientes resoluciones de video señale la cantidad de bytes que se necesitan para guardar un segundo de video sin compresión. Escriba cualquier supuesto que necesite realizar.
- a. 1920 x 1080, 60 fps, 4:2:0.
  - b. 1280 x 720, 24 fps, 4:4:4.
  - c. 3180 x 2160, 30 fps, 4:2:0.
  - d. 1920 x 1440, 30 fps, 4:4:4.
2. (0.8 puntos) Se tiene un video compuesto de 28 cuadros, donde un encoder MPEG-1 decidió que el tipo de cada cuadro (I, P o B) será de acuerdo a la siguiente tabla:

Número de cuadro	Tipo de cuadro	Número de cuadro	Tipo de cuadro
1	I	15	B
2	B	16	I
3	B	17	B
4	P	18	B
5	B	19	P
6	P	20	B
7	B	21	P
8	B	22	B
9	I	23	I
10	B	24	B
11	B	25	P
12	P	26	B
13	B	27	B
14	P	28	I

Suponga que durante la reproducción del video anterior, el decoder MPEG-1 tuvo un error al decodificar cierto cuadro. Señale todos los cuadros donde se verán errores cuando el cuadro erróneo es:

- a. El cuadro 6
- b. El cuadro 9
- c. El cuadro 18
- d. El cuadro 19

## Semana 08

1. Una empresa necesita un buscador de imágenes para que dada una foto de consulta recupere la mayor cantidad de imágenes similares dentro de la primera página de resultados (que es de tamaño 8).

Para esto se implementaron dos sistemas con algoritmos distintos y se desea comparar la efectividad de cada uno. Se elaboraron dos casos de prueba **Q1** y **Q2** y se determinó que las imágenes relevantes para **Q1** son {**d37, d06, d22, d80**} y para **Q2** son {**d48, d71, d14**}.

La respuesta de cada sistema con los primeros 12 resultados fue:

Sistema 1		
Lugar	Q1	Q2
1	d48	d95
2	<b>d80</b>	d83
3	d59	<b>d48</b>
4	<b>d22</b>	d05
5	d24	d10
6	<b>d06</b>	<b>d14</b>
7	d54	d06
8	<b>d37</b>	<b>d71</b>
9	d13	d66
10	d47	d74
11	d42	d25
12	d71	d53

Sistema 2		
Lugar	Q1	Q2
1	<b>d22</b>	<b>d14</b>
2	<b>d80</b>	d25
3	d75	d21
4	d95	d92
5	d10	d29
6	d87	d11
7	d55	d80
8	d29	d96
9	d20	d95
10	d62	<b>d48</b>
11	d71	d42
12	<b>d37</b>	d65

- a. (0.4 puntos) Calcule **Mean Average Precision (MAP)** de ambos sistemas.
- b. (0.4 puntos) Calcule **Mean Reciprocal Rank (MRR)**, **precision@1**, **recall@12** y **R-Precision** de ambos sistemas.
- c. (0.3 puntos) Señale cuál sistema logra mejor efectividad. Justifique con métricas.

2. Una estudiante de computación para su trabajo de memoria desarrolló un buscador de fotografías “near duplicates” cuyo objetivo es encontrar la mayor cantidad de imágenes duplicadas de una consulta.

La estudiante implementó tres algoritmos distintos de descriptores visuales y en su trabajo desea determinar el descriptor que logra mayor efectividad. Para la evaluación, la estudiante construyó un conjunto de 20 imágenes de consulta y para cada consulta determinó las respuestas relevantes dentro de un dataset de imágenes.

Para cada uno de los tres algoritmos, la estudiante calculó el Average Precision (AP) de cada consulta, y a esos 20 valores les calculó el promedio ( $\mu$ ) y la desviación estándar ( $\sigma$ ) obteniendo los siguientes resultados:

	Algoritmo 1	Algoritmo 2	Algoritmo 3
<b>Promedio AP (<math>\mu</math>)</b>	0.671	0.811	0.744
<b>Desv.estándar AP (<math>\sigma</math>)</b>	0.216	0.165	0.017

- a. (0.3 puntos) Calcule el MAP de cada algoritmo y calcule el intervalo de confianza de cada uno (utilice un p-value de 5%). Considerando los intervalos de confianza obtenidos, señale el mejor y el peor algoritmo.
- b. (0.3 puntos) La estudiante construyó un conjunto más grande de consultas, ahora con 200 imágenes de consulta y calculó el AP de las nuevas consultas. Suponga que obtuvo el mismo promedio AP ( $\mu$ ) y desviación estándar ( $\sigma$ ) de la tabla anterior, pero con la diferencia que ahora es el resultado de evaluar 200 consultas en vez de 20. Calcule los nuevos intervalos de confianza para este caso. ¿Hay alguna diferencia en el algoritmo que resulta ser el mejor y el peor?

3. La empresa RIM-BET necesita un software que prediga los resultados de la fase de grupos de la Copa Mundial de Fútbol, y para esto tiene dos proveedores que le ofrecen el mejor servicio de predicción del mercado.

Al ingeniero jefe de RIM-BET se le ocurre la idea de evaluar ambos servicios ingresando los datos del mundial Sudáfrica 2010 y comparar las predicciones que entregue cada servicio con lo que realmente sucedió. En particular, al evaluar el caso del Grupo H, se ingresaron los datos de los equipos y ambos servicios entregaron las siguientes predicciones:

Respuesta Servicio 1		Respuesta Servicio 2	
Lugar	Equipo	Lugar	Equipo
1°	Suiza	1°	Chile
2°	Honduras	2°	Honduras
3°	España	3°	Suiza
4°	Chile	4°	España

El resultado real para el Grupo H fue:

Resultado Real	
Lugar	Equipo
1°	España
2°	Chile
3°	Suiza
4°	Honduras

- a) (0.3 puntos) Calcule el Coeficiente de Spearman para ambos servicios y señale el que se acerca más a la respuesta real.
- b) (0.3 puntos) Calcule el Coeficiente de Kendall-Tau para ambos servicios y señale el que se acerca más a la respuesta real.

## Entrega

- Puede desarrollarlo en papel y enviar una foto (.jpg, .png), o puede desarrollarlo en formato digital en una planilla (.xlsx .ods), un documento (.docx) u otro formato exportado a .pdf.
- El plazo máximo de entrega es el **viernes 16 de mayo de 2025** hasta las 23:59. Existirá una segunda fecha (por definir) para entregar su respuesta sin descuentos en la nota.

**El control es \*individual\* y debe ser de su autoría, es decir, no pueden ser resueltos por otro estudiante, no se pueden copiar respuestas de Internet, no se permite usar ChatGPT ni similares. En caso de detectar copia o plagio se asignará nota 1.0 a las o los estudiantes involucrados.**