



Recuperación de Información Multimedia

Introducción MIR

CC5213 – Recuperación de Información Multimedia

Departamento de Ciencias de la Computación

Universidad de Chile

Juan Manuel Barrios – <https://juan.cl/mir/> – 2020



¿Qué es la Recuperación de Información?

- **Information Retrieval (IR)** estudia cómo representar, organizar, almacenar y acceder a **información** existente en documentos
 - Recuperar (*to retrieve*) documentos relevantes a la necesidad de información del usuario
 - Documentos: página web, texto de emails, conjunto de tags, fichas de libros, etc.
 - Consulta: frase, keywords, preguntas

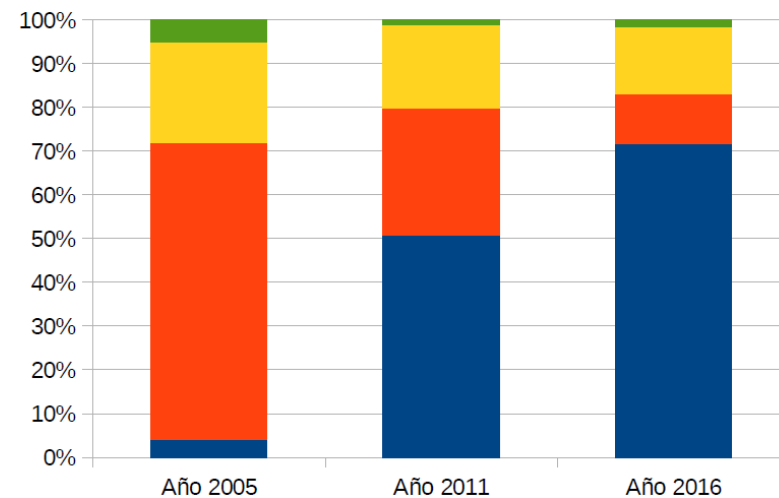
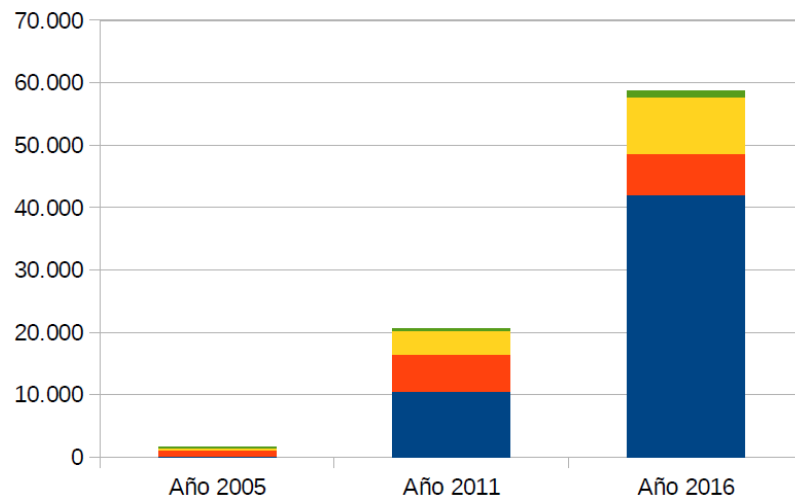


Relevancia y Ranking

- **Relevancia:** Cuantificar cuan útil es un documento en satisfacer la necesidad de información del usuario
- **Ranking:** Ordenar documentos según relevancia a la consulta
- Bases de Datos vs IR:
 - Bases de Datos: obtener eficientemente ***todos*** los elementos que cumplen cierta condición, transacciones
 - Recuperación de Información: obtener eficientemente ***los mejores*** elementos que satisfacen una necesidad

Evolución del uso de Internet

Tráfico de datos de usuarios de Internet [PB mensual]



- Otros
- Web, email, datos
- P2P
- Video

En 2016 más del 70% del tráfico son videos:

- Sin embargo audio, imágenes y videos son una “caja negra” usando solo sus metadatos
- Se necesitan métodos para poder buscar dentro de ellos...

Fuente:

Global IP Traffic Forecast and Methodology, 2006-2011.
Cisco Visual Networking Index: Forecast and Methodology, 2011-2016
Cisco Visual Networking Index: Forecast and Methodology, 2016-2021



¿Qué es la Recuperación de Información Multimedia?

■ Multimedia Information Retrieval (MIR)

- Recuperar (*to retrieve*) **archivos multimedia** relevantes a la necesidad de información del usuario
- Documentos: Audio, Imagen, Video, Objetos 3D, etc.
 - Ej: fotos personales, películas, grabación de cámaras de seguridad, canciones, etc.
- La búsqueda requiere analizar **contenido multimedia**, es decir, revisar los píxeles, samples de audio, frames de video
- La búsqueda **NO requiere metadatos** (no es necesario usar etiquetas, EXIF, ID3, ni textos)
- Consulta: keywords, “by-example” (buscar algo parecido a un documento modelo), “by-sketch” (buscar usando bosquejos)



Áreas Involucradas en MIR

- **Análisis de Contenido Multimedia**
 - Imágenes, Audio, Videos y Texto
- **Estructuras de Datos**
 - Algoritmos eficientes, Métodos de búsqueda, Índices
- **Ciencia de Datos / Inteligencia Artificial**
 - Data Mining, Machine Learning, Deep Learning, Visualización



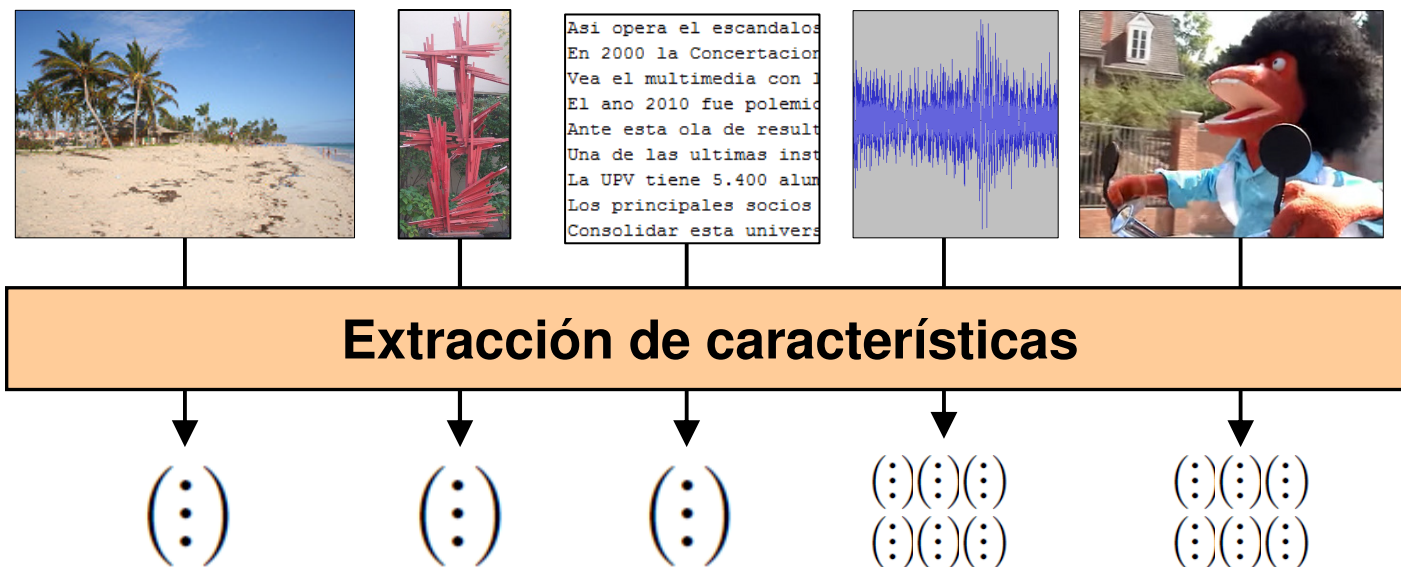
Temas a estudiar durante el semestre

A continuación se describen las unidades o partes de este curso

Parte 1

■ Descripción de Contenido Multimedia

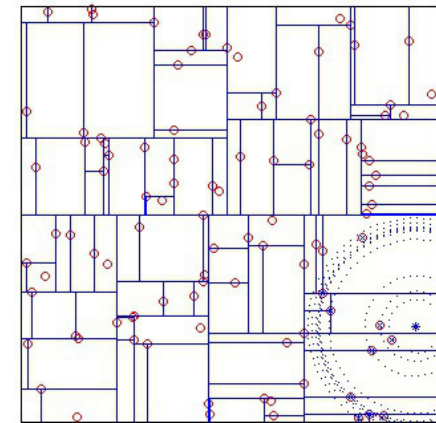
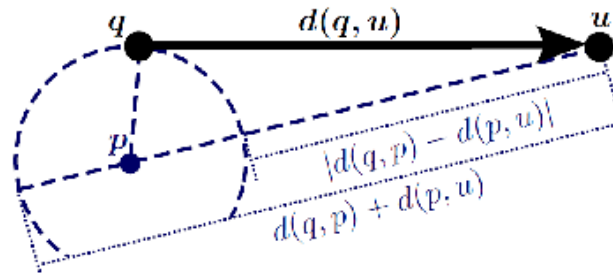
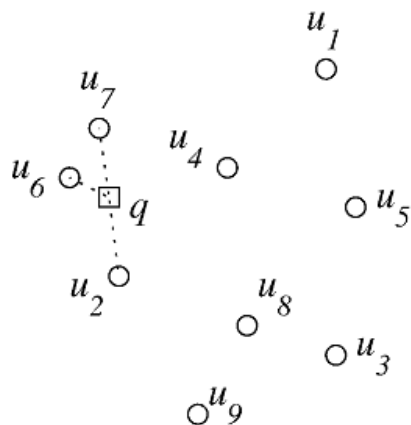
- Métodos para analizar imágenes, audio, videos y texto
- El contenido de cada documento multimedia (foto, canción, video, email, etc.) se representa por uno o más vectores



Parte 2

■ Búsqueda por Similitud

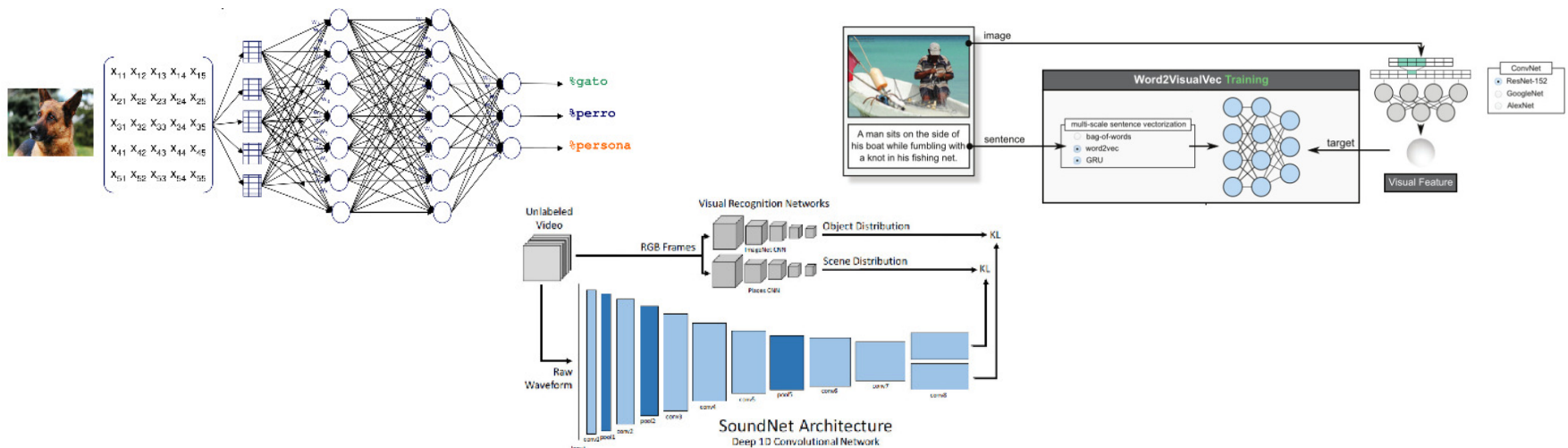
- Métodos para resolver eficientemente búsquedas k-NN en espacios vectoriales y métricos
- Resolver búsquedas en millones de vectores
- Índices para espacios vectoriales y métricos



Parte 3

■ Métodos Avanzados e Investigación

- Técnicas actuales de descripción y búsqueda usando Codebooks y Deep Learning
- Uso de redes neuronales (convolucionales, recurrentes, auto-encoders) para descripción y búsqueda





Casos de Estudio

Distintos problemas que motivan los
contenidos del curso

Casos de Estudio #1

- Dada una imagen de consulta, buscar otras imágenes parecidas

Imagen de
consulta



Casos de Estudio #1

- Buscar imágenes parecidas según algún criterio (color, forma, etc.)

Resultados



Ej: Dada la imagen de una playa encontrar otras imágenes parecidas.

Casos de Estudio #2

- Dado un catálogo de objetos, reconocer la aparición de un objeto conocido en una foto



Imagen de consulta

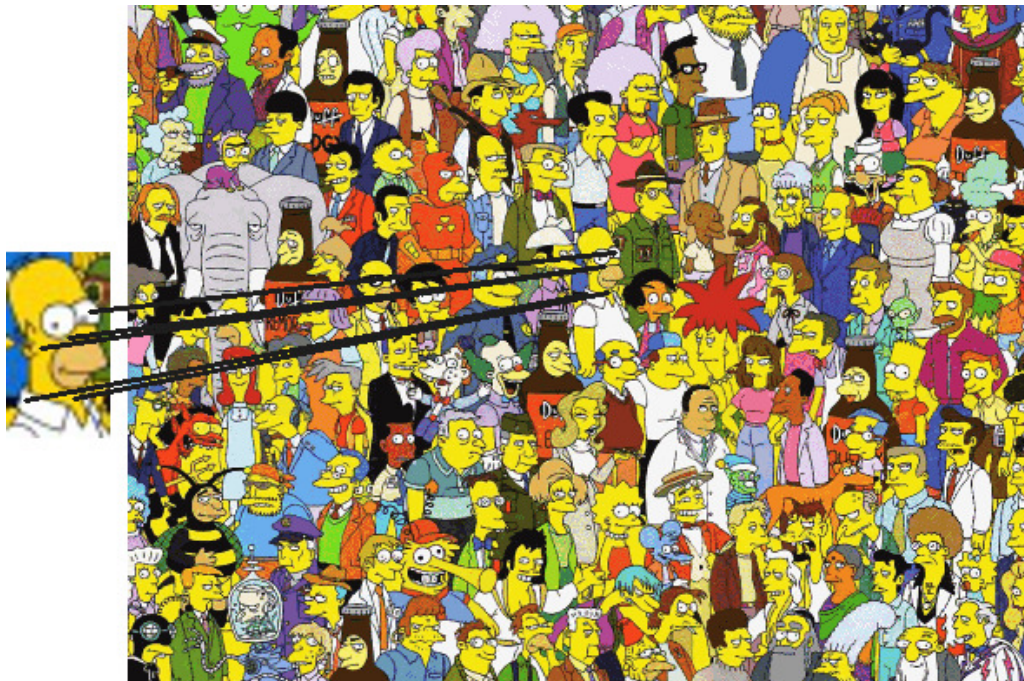


Ver <https://impresee.com/>



Casos de Estudio #2

- Además de la aparición del objeto, determinar la ubicación y pose del objeto



Casos de Estudio #3

- Dado un video de Internet determinar la escena original de la que proviene
- Reconocer una canción según un trozo de audio (“Shazam”)

youtube



Cap.5



Cap.10



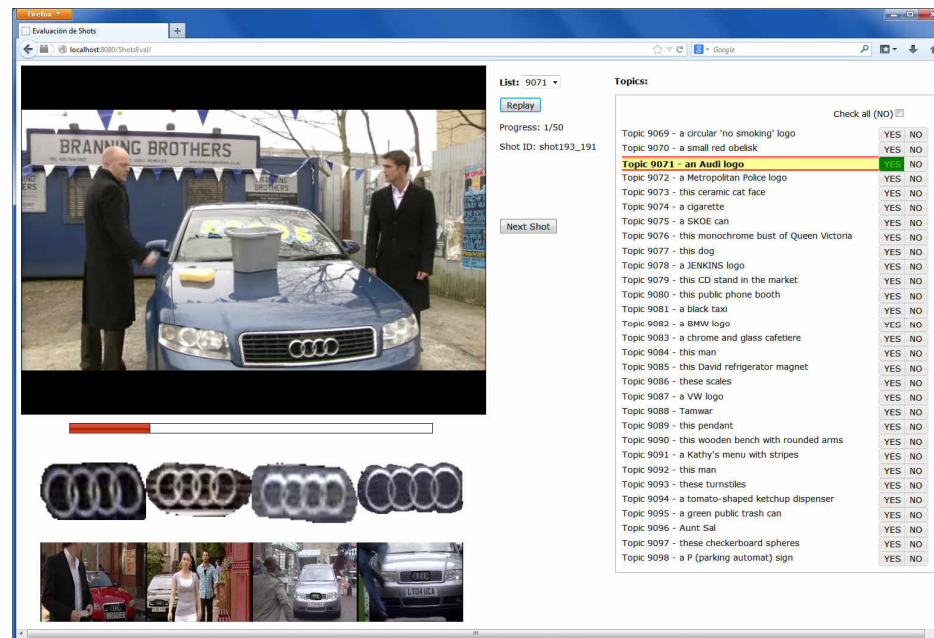
Cap.23



Ver <https://sourceforge.net/projects/p-vcd/>

Casos de Estudio #4

- Buscar la aparición de algún producto, logo, etc. en televisión



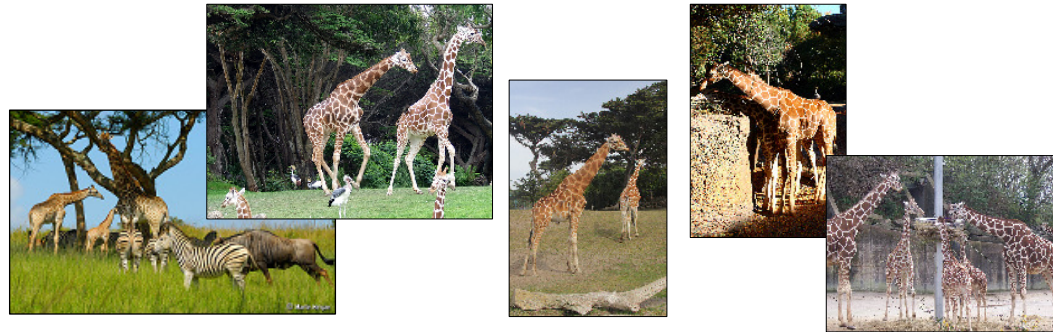
Ej.: Buscar las apariciones
del logo Audi
© BBC EastEnders

Ver <https://www.robots.ox.ac.uk/~vgg/research/vgoogle/>

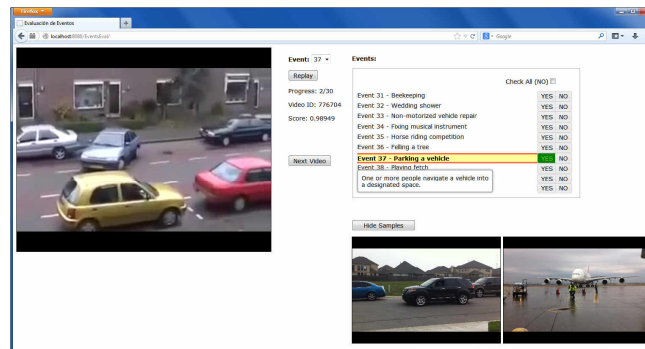
Casos de Estudio #5

- Buscar por texto en imágenes o videos “sin etiquetar”

Ej: Buscar fotos con “jirafas comiendo”



Ej: Buscar videos con “un vehículo estacionando”



Muchas Aplicaciones Novedosas y Problemas Abiertos

Google: Flujo de videos en YouTube hace imposible eliminar todo el contenido terrorista

La compañía justificó que la gran cantidad de videos no se puede detectar toda la propaganda

29 de Enero de 2015 | 09:26 | AP

Google sigue los pasos de Facebook e implementará inteligencia artificial contra contenidos terroristas

La empresa enfocará sus esfuerzos a través de YouTube, un

19 de Junio de 2017 | 16:39 | D

Liga inglesa demanda a YouTube por violar derechos de autor

La asociación que maneja el campeonato inglés inició la acción legal en Nueva York, ante los numerosos videos con imágenes de la Premier League que están disponibles en el sitio web.

La aplicación móvil Shazam permitirá a sus usuarios identificar objetos

La herramienta que reconoce la música permite a los usuarios identificar productos en un supermercado

05 de Marzo de 2015 | 16:42 | Reuters

Apple confirma la adquisición de Shazam y aventura "interesantes planes" para el futuro

Las compañías deberán esperar hasta el próximo año para comenzar a integrar ambos servicios

11 de Diciembre de 2017 | 15:12 | Emol

Facebook reconocerá automáticamente canciones que suenan de fondo

La red social también podrá identificar los programas de televisión que ven sus usuarios mientras escriben mensajes.

Aplicación chilena encuentra productos con solo dibujarlos o sacarles una foto

• Creada por investigadores universitarios, la herramienta analiza imágenes y busca los objetos en



De la idea al prototipo

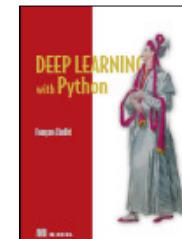
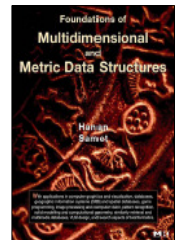
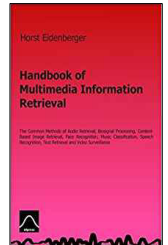
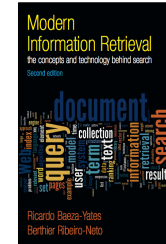
Spotify amplía su alcance más allá de la música, sumando videos y podcasts a su servicio

La compañía de Daniel Ek estrenó hoy un nuevo modelo que busca ayudar al descubrimiento de contenido en la app. También sumó nuevas funciones para escuchar música al hacer deporte.

20 de Mayo de 2015 | 13:33 | Emol

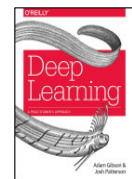
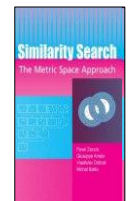
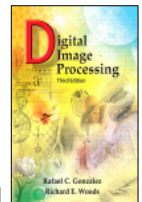
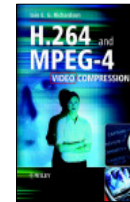
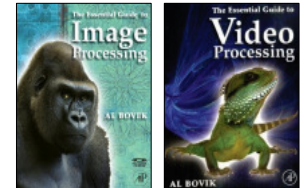
Bibliografía

- **Modern Information Retrieval.** Baeza-Yates, Ribeiro-Neto, 2011.
- **Handbook of Multimedia Information Retrieval.** Eidenberger. 2012.
- **Multimedia Retrieval.** Blanken, de Vries, Blok, Feng. 2007.
- **Foundations of Multidimensional and Metric Data Structures.** Samet. 2006.
- **Deep Learning with Python.** Chollet. 2018.



Bibliografía

- **The Essential Guide to Image/Video Processing.** Bovik. 2009.
- **H.264 and MPEG-4 Video Compression.** Richardson. 2003.
- **Digital Image Processing.** Gonzalez, Woods. 2008.
- **Data Mining: The Textbook.** Aggarwal. 2015.
- **Similarity Search The Metric Space Approach.** Zezula, Amato, Dohnal, Batko. 2006.
- **Computer Vision. Algorithms and Applications.** Szeliski. 2011.
- **Deep Learning: A Practitioner's Approach.** Patterson, Gibson. 2017.





CC5213

Recuperación de Información Multimedia

Departamento de Ciencias de la Computación (DCC)
Facultad de Ciencias Físicas y Matemáticas (FCFM)
Universidad de Chile



Datos del Curso (Otoño 2020)

- **CC5213–Recuperación de Información Multimedia**


- ☐ Departamento de Ciencias de la Computación (DCC), FCFM, Universidad de Chile
- ☐ Pregrado (ingeniería civil en computación) y Posgrado (magister y doctorado en ciencias mención computación)

- **6 Créditos**

- **Miércoles y Viernes 8:30 – 9:45**

- **Requisito (basta uno):**

- ☐ CC5206 – Introducción a la Minería de Datos
- ☐ EL4106 – Inteligencia Computacional
- ☐ MA5204 – Aprendizaje de Máquinas
- ☐ CC4102 – Diseño y Análisis de algoritmos



Evaluación CC5213 (Otoño 2020)

- Seis mini-controles
 - **NC** = $(C1 + C2 + C3 + C4 + C5 + C6) / 6$
- Cinco tareas de programación
 - **NT** = $(T1 + T2 + T3 + T4 + T5) / 5$
- Dos presentaciones de lecturas
 - **NL** = $(L1 + L2) / 2$
- Proyecto final (T=trabajo realizado, P=presentación final)
 - **NP** = $(T + P) / 2$
- Para aprobar:
NC \geq 4.0, **NL** \geq 4.0, **NT** \geq 4.0, **NP** \geq 4.0
- Nota Final:
(NC + NL + NT + NP) / 4



Mini-Controles (NC)

- Una pregunta escrita al inicio de clases
- Individual
- 15 minutos máximo
- Se permiten apuntes en papel, slides impresas, pero nada digital
- Se avisarán con una clase de anticipación
- Si **NC** ≥ 4.0 es posible volver a rendir uno o más al final del semestre pero sin apuntes



Tareas de Programación (NT)

- Resolver problemas usando Python 3 o C++ 11
- Individual
- Se evaluará:
 - La calidad de la solución
 - La claridad del código fuente
 - Explicación de cómo usar la tarea (readme.txt)
- Las tareas pueden tener actividades opcionales para lograr bonus en nota
- Si **NT** ≥ 4.0 es posible volver a enviar una o más tareas al finalizar el semestre



Presentación de Lecturas (NL)

- Exponer publicaciones científicas
 - Tiempo máximo de 10 minutos
 - Se publicará una lista papers para elegir
 - Individual o grupos de 2 estudiantes
-
- Fechas de presentaciones en semana 11 y semana 15



Proyecto de Síntesis (NP)

- Resolver algún problema real a elección que aplique contenidos del curso
- Tarea con “tema libre” que debe incluir:
 - Descripción de contenido multimedia
 - Búsqueda por similitud
- Individual o grupos de 2 estudiantes
- Presentación de resultados al finalizar el semestre (semana 17)



Clases

- Las clases inician a las **8:30 en punto** y duran **75 minutos**
 - Mini-controles inician a las 8:30 y terminan a las 8:45
- Se publicará material docente en U-Cursos con anticipación a cada clase
 - **Recomendación:** imprimirlo, leerlo con anticipación y tomar apuntes sobre él durante la clase
- Las clases se enfocan en explicar la intuición detrás de los distintos temas, mostrar ejemplos y discutir ventajas/desventajas de las distintas técnicas
 - **Recomendación:** asistir a todas las clases y participar
- No se permite el uso de computador durante la clase



No se permite el uso de computador durante la clase

- “As expected, the multitasking students learned less than those focused on the lecture, scoring about 11 percent lower on a test. What is more surprising: **the learning of students near the multitaskers also suffered**. Students who could see the screen of a multitasker’s laptop (but were not multitasking themselves) scored 17 percent lower on comprehension than those who had no distracting view. It’s hard to stay focused when a field of laptops open to Facebook, Snapchat, and email lies between you and the lecturer.”
 - <https://www.brookings.edu/research/for-better-learning-in-college-lectures-lay-down-the-laptop-and-pick-up-a-pen/>
- Más Información:
 - <https://doi.org/10.1016/j.compedu.2012.10.003>
 - <https://www.sciencedirect.com/science/article/abs/pii/S0272775716303454>
 - <https://www.theatlantic.com/technology/archive/2014/05/to-remember-a-lecture-better-take-notes-by-hand/361478/>
 - <https://www.nytimes.com/2017/11/22/business/laptops-not-during-lecture-or-meeting.html>

Planificación (Otoño 2020)

Semana	Miércoles	Tema	Viernes	Tema
1	11 Mar	Introducción Motivación	13 Mar	C++ y OpenCV Procesamiento De Imágenes
2	18 Mar	Detección de Bordes Detección de Líneas	20 Mar	Descriptores Descriptores globales (parte 1)
3	25 Mar	Mini-Control 1 Dominio de Frecuencias	27 Mar	FERIADO
4	1 Abr	Descriptores globales (parte 2)	3 Abr	Descriptores Audio
5	8 Abr	Descriptores Color	10 Abr	Mini-Control 2 Procesamiento de Videos
6	15 Abr	Búsquedas por Similitud	17 Abr	Repaso de Estructuras de Datos
7	22 Abr	Índices Multidimensionales (parte 1)	24 Abr	Mini-Control 3 Índices Multidimensionales (parte 2)
8	29 Abr	Índices Multidimensionales (parte 3)	1 May	FERIADO
9	6 May	Maldición de la Dimensionalidad PCA	8 May	Índices Métricos
10	13 May	Mini-Control 4 Descriptores Locales	15 May	Codebooks
	20 May	Vacaciones	22 May	Vacaciones
11	27 May	Presentación 1	29 May	Presentación 1
12	3 Jun	Métricas de Evaluación	5 Jun	Mini-Control 5 Descriptores de Texto
13	10 Jun	Repaso de Machine Learning	12 Jun	Deep Learning (imágenes)
14	17 Jun	Mini-Control 6 Deep Learning (texto)	19 Jun	Deep Learning (audio, multimodal)
15	24 Jun	Presentación 2	26 Jun	Presentación 2
16				
17		Presentación Proyectos		



Reglas ante situaciones especiales

- Si no se realiza una clase (e.g. paro de un día) esa clase **se pierde** y no se recupera:
 - La siguiente clase se hace un resumen y continúa la planificación
 - Si había un mini-control fijado ese día, se pierde
 - Si había presentación de papers, se debe subir un video con la presentación (el plazo será la siguiente clase)
- Si se suspenden semanas completas (e.g. paros continuados por semanas) usualmente se replanifican las semanas
- Si se acorta el semestre, se pierden esas clases y se debe subir un video de cada presentación faltante



Recursos

- Sitio Web del Curso:

<https://juan.cl/mir/>

- Repositorio GitHub:

<https://github.com/juanbarrios/Curso-Recuperacion-de-Informacion-Multimedia>