



Programa de Recuperación de Información y Recomendaciones en la Web (WIR)

1. NOMBRE DE LA UNIDAD CURRICULAR

Recuperación de Información y Recomendaciones en la Web (WIR)

2. CRÉDITOS

10 créditos

3. OBJETIVOS DE LA UNIDAD CURRICULAR

La recuperación de la información no comenzó con Internet, pero es hoy el motor de innovación en el área, debido principalmente a explosión de publicaciones y a la enorme cantidad de creadores de información que hay en ella. Toda esta información no podría seguir creciendo si no se pudiera encontrar, indexar y analizar para que cada usuario pueda obtener de manera rápida la información que es pertinente a sus necesidades.

El objetivo principal de este curso es enseñar a los estudiantes los fundamentos básicos de la recuperación de la información, en particular para Internet y luego profundizar en temas avanzados.

La primera parte del curso presenta los temas principales relacionados con la recuperación de la información, en particular relacionados con el funcionamiento de los motores de búsqueda de Internet. En la segunda parte del curso se discuten temas avanzados, como por ejemplo la optimización de la recuperación de información mediante “relevance feedback” o expansión de consultas. Finalmente se tratan los temas de filtrado colaborativo y el análisis de enlaces, para estudiar como el comportamiento colectivo ayuda en la recuperación de la información en Internet.

4. METODOLOGÍA DE ENSEÑANZA

El curso es de carácter teórico y de tipo proyecto-presentación. Durante 13 semanas el estudiante tendrá en promedio 4 horas semanales de clases teóricas y de

monitoreos/consultas para el apoyo a la realización de un proyecto. Se estima unas 2,5 horas semanales de estudio adicional y 5,5 horas semanales para la realización del proyecto.

En resumen la dedicación esperada por parte de los estudiantes es:

- 52 hs. Clases
- 70 hs. realización de proyectos
- 30hs. estudio individual

5. TEMARIO

1. Recuperación booleana
2. Vocabularios y listas de ocurrencias de palabras
3. Diccionarios y recuperación tolerante
4. Construcción y compresión de índices
5. Puntaje y asignación de pesos en el modelo de espacio vectorial
6. Calculo de puntajes en un sistema completo de búsqueda
7. Evaluación de la recuperación de la información
8. Relevance feedback y expansión de consultas
9. Recuperación de la información probabilística
10. Clasificación de texto mediante regla de Bayes y modelo de espacio vectorial
11. Filtrado colaborativo
12. Análisis de enlaces

6. BIBLIOGRAFÍA

6.1 Básica

Christopher D. Manning, Prabhakar Raghavan and Hinrich Schütze, Introduction to Information Retrieval, Cambridge University Press. 2008. <http://nlp.stanford.edu/IR-book/>
Para todos los temas.

7. CONOCIMIENTOS PREVIOS EXIGIDOS Y RECOMENDADOS

7.1 Conocimientos Previos Exigidos: Conocimientos básicos de Probabilidad y Estadística, Bases de Datos, Análisis de algoritmos, Algoritmos y estructura de datos, Grafos

7.2 Conocimientos Previos Recomendados: Probabilidad y Estadística avanzada, Investigación Operativa

ANEXO A

Para todas las Carreras

A1) INSTITUTO

Instituto de Computación

A2) CRONOGRAMA TENTATIVO

Semana 1	Recuperación booleana, vocabularios y listas de ocurrencias de palabras
Semana 2	Diccionarios y recuperación tolerante
Semana 3	Construcción y compresión de índices
Semana 4	Monitoreos/consultas
Semana 5	Puntaje y asignación de pesos en el modelo de espacio vectorial
Semana 6	Calculo de puntajes en un sistema completo de búsqueda
Semana 7	Evaluación de la recuperación de la información
Semana 8	Monitoreos/consultas
Semana 9	Relevance feedback y expansión de consultas
Semana 10	Recuperación de la información probabilística
Semana 11	Clasificación de texto con la regla de Bayes y mediante el modelo de espacio vectorial
Semana 12	Filtrado colaborativo
Semana 13	Análisis de enlaces

A3) MODALIDAD DEL CURSO Y PROCEDIMIENTO DE EVALUACIÓN

Clases teóricas y de monitoreos/consultas

Para abordar los temas propuestos se realizarán dos clases semanales, algunas semanas teóricas y otras de monitoreos/consultas para el apoyo a la realización de un proyecto.

Proyecto

Los estudiantes deberán realizar un proyecto que puede constar esencialmente de tres actividades: investigación sobre un tema relacionado con el curso, diseño de una solución a un problema (usando lo aprendido) y una posible implementación de la solución. El tema específico deberá ser validado por el docente previamente. El estudiante deberá escribir un informe sobre el trabajo realizado y hacer una presentación oral del trabajo (max. 15 minutos).

Evaluación

El curso se evaluará a partir de:

- La asistencia a un mínimo del 60% de las clases, requerimiento eliminatorio que no lleva puntos

- Informe escrito sobre el proyecto realizado (60%)
- Presentación oral del proyecto (40%)

El curso se aprueba con el 60% de los puntos, teniendo un mínimo del 60% en cada actividad.

A4) CALIDAD DE LIBRE

Esta unidad curricular no adhiere a la resolución de Calidad de Libre.

A5) CUPOS DE LA UNIDAD CURRICULAR

Cupo máximo: 50 estudiantes

ANEXO B para la carrera Ingeniería en Computación (plan 97)

B1) ÁREA DE FORMACIÓN

Bases de Datos y Sistemas de Información.

B2) UNIDADES CURRICULARES PREVIAS

Para el Curso: Exámenes aprobados de:
Programación 3 y
Fundamentos de bases de datos

Para el Examen: No aplica

ANEXO B para la carrera Ingeniería en Computación (plan 87)

B1) ÁREA DE FORMACIÓN

No corresponde

B2) UNIDADES CURRICULARES PREVIAS

Para el Curso: Previas comunes a las electivas y
Exámenes aprobados de:
Programación III y
Bases de datos

Para el Examen: No aplica

Observación: Esta unidad curricular se corresponde con una electiva