
Formulario de Aprobación Curso de Posgrado 2013

Asignatura: Seminario webir (Seminario de Recuperación de Información y Recomendaciones en la Web)

Profesor de la asignatura ¹: Dra. Libertad Tansini, grado 3, Inco
(título, nombre, grado o cargo, Instituto o Institución)

Profesor Responsable Local ¹: Dra. Libertad Tansini, grado 3, Inco
(título, nombre, grado, Instituto)

Otros docentes de la Facultad:
(título, nombre, grado, Instituto)

Docentes fuera de Facultad:
(título, nombre, cargo, Institución, país)

Instituto ó Unidad: Inco
Departamento ó Area: Bases de Datos y Sistemas de Información

¹ Agregar CV si el curso se dicta por primera vez.
(Si el profesor de la asignatura no es docente de la Facultad se deberá designar un responsable local)

Fecha de inicio y finalización: 23/9/2013 hasta 31/12/2013
Horario y Salón: A confirmar

Horas Presenciales: 56
(se deberán discriminar las mismas en el ítem Metodología de enseñanza)

Nº de Créditos: 6
(de acuerdo a la definición de la UdeLaR, un crédito equivale a 15 horas de dedicación del estudiante según se detalla en el ítem metodología de la enseñanza)

Público objetivo y Cupos: Hay restricción de cupo de 20 estudiantes (ver nota adjunta).

(si corresponde, se indicará el número de plazas, mínimo y máximo y los criterios de selección. Asimismo, se adjuntará en nota aparte los fundamentos de los cupos propuestos. Si no existe indicación particular para el cupo máximo, el criterio general será el orden de inscripción en el Depto. de Posgrado, hasta completar el cupo asignado)

Objetivos:

La recuperación de la información ha tenido como motor de innovación, en las últimas décadas, las necesidades surgidas a partir del extenso uso de Internet. 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 del curso es presentarle a los estudiantes los fundamentos básicos de la recuperación de la información, en particular para Internet y luego profundizar en temas avanzados discutiendo por ejemplo la optimización de la recuperación de información mediante "relevante feedback", expansión de consultas, filtrado colaborativo y el análisis de enlaces.

Se pretende además que los estudiantes estudien y presenten en formato de seminarios aspectos específicos del área temática del curso.

Conocimientos previos exigidos: Bases de Datos y Programación 3

Conocimientos previos recomendados: Bases de Datos y Programación

Metodología de enseñanza:

El curso es de carácter teórico y de tipo seminario-presentación. El estudiante tendrá 4 horas semanales de clases teóricas (durante 10 semanas). Se estima unas 2 horas semanales de estudio adicional y 15 horas para la profundización y preparación de un seminario. Los estudiantes deberán asistir a los seminarios como parte del curso (cada estudiante debe asistir a los seminarios de los demás estudiantes, aprox. 16 horas).

(comprende una descripción de las horas dedicadas por el estudiante a la asignatura y su distribución en horas presenciales -de clase práctica, teórico, laboratorio, consulta, etc.- y no presenciales de trabajo personal del estudiante)

- Horas clase (teórico): 40
- Horas clase (práctico): 16
- Horas clase (laboratorio):
- Horas consulta:
- Horas evaluación:
 - Subtotal horas presenciales: 56
- Horas estudio: 20
- Horas resolución ejercicios/prácticos:
- Horas proyecto final/monografía: 15
 - Total de horas de dedicación del estudiante: 91

Forma de evaluación:

- Asistencia, de la menos 80% de las clases y de los seminarios
- Presentación oral de un tema en formato seminario

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. Modelos de lenguajes para la recuperación de la información
11. Clasificación de texto con la regla de Bayes y mediante el modelo de espacio vectorial
12. Support vector machines y aprendizaje automático para documentos
16. Filtrado colaborativo
17. Análisis de enlaces



Facultad de Ingeniería Comisión Académica de Posgrado

Bibliografía:

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

Además utilizaremos artículos para profundizar en aspectos específicos del área temática que serán indicados en la página web del curso.
