

Análisis de Grafos y Redes

Código: 106567
Créditos ECTS: 6

Titulación	Tipo	Curso	Semestre
2504392 Inteligencia Artificial / Artificial Intelligence	OB	2	2

Contacto

Nombre: Josep Lladós Canet

Correo electrónico: josep.llados@uab.cat

Uso de idiomas

Lengua vehicular mayoritaria: inglés (eng)

Algún grupo íntegramente en inglés: No

Algún grupo íntegramente en catalán: No

Algún grupo íntegramente en español: No

Equipo docente

Cristina Perez Sola

Prerequisitos

Para una mejor seguimiento de la asignatura, es recomendable haber superado la asignatura de segundo semestre de primer año "ingeniería de Datos" y de primer semestre de segundo año "Fundamentos de Aprendizaje Automático". También es recomendable tener competencias básicas de programación adquiridas en las asignaturas previas "Fundamentos de Programación".

Objetivos y contextualización

Las estructuras de grafos son de gran utilidad en representación de datos a partir de sus relaciones. Los grafos son un lenguaje general para describir y analizar entidades con relaciones/interacciones. Son de gran utilidad en la representación y navegación de redes sociales, el modelaje de moléculas, la representación geográfica para software de navegación, biometría, representación de conocimiento, representación de escenas, etc. Además, la teoría de grafos ofrece una metodología robusta i matematicamente fundamentada para caracterizar, recorrer o analizar este tipo de estructuras. En esta asignatura se cubren los siguientes objetivos: Adquirir conocimientos en Teoría básica de grafos, optimización de recorridos, algoritmos sobre grafos, minería de datos para grafos, grafos dinámicos, difusión y propagación de la información en redes, búsqueda en internet, análisis de redes sociales, comunidades y perfiles de usuarios públicos, sistemas de recomendación.

Competencias

- Analizar y resolver problemas de forma efectiva, generando propuestas innovadoras y creativas para alcanzar los objetivos.
- Conocer y utilizar de forma eficiente las técnicas y herramientas de representación, manipulación, análisis y gestión de datos a gran escala.
- Introducir cambios en los métodos y los procesos del ámbito de conocimiento para dar respuestas innovadoras a las necesidades y demandas de la sociedad.

- Que los estudiantes tengan la capacidad de reunir e interpretar datos relevantes (normalmente dentro de su área de estudio) para emitir juicios que incluyan una reflexión sobre temas relevantes de índole social, científica o ética.

Resultados de aprendizaje

1. Analizar y resolver problemas de forma efectiva, generando propuestas innovadoras y creativas para alcanzar los objetivos.
2. Aplicar técnicas de minería de datos específicas para grafos, seleccionando adecuadamente los algoritmos en función del objetivo del análisis, el volumen de los datos a procesar y las capacidades de procesamiento disponibles.
3. Conocer las herramientas básicas de manipulación de diferentes tipos de datos estructurados, semiestructurados y no estructurados.
4. Ponderar los riesgos y las oportunidades de las propuestas de mejora tanto propias como ajenas.
5. Proponer nuevas maneras de medir el éxito o el fracaso de la implementación de propuestas o ideas innovadoras.
6. Que los estudiantes tengan la capacidad de reunir e interpretar datos relevantes (normalmente dentro de su área de estudio) para emitir juicios que incluyan una reflexión sobre temas relevantes de índole social, científica o ética.
7. Utilizar adecuadamente los métodos de visualización de datos.

Contenido

Tema 1. Introducción. Definiciones, conceptos, representaciones.

Tema 2. Métricas en grafos. Algoritmos de centralidad y prestigio (Degree Centrality, Closeness Centrality, Betweenness Centrality).

Tema 3. Generación de grafos sintéticos. Grafos aleatorios. Síntesis de grafos.

Tema 4. Visualización de grafos. Plataforma Gephi.

Tema 5. Búsqueda y recorridos. Graph transversal (Breadth First Search, Depth First Search), Shortest Path, A*, Minimum Spanning Tree, Random Walk.

Tema 6. Búsqueda en internet. Ranking y relevancia para modelos web. Algoritmo PageRank. Crawling.

Tema 7. Algoritmos para detección de comunidades y nodos influyentes. Graph clustering. Subgrupos comunes. Reputación.

Tema 8. Propagación de etiquetas y difusión en grafos. Modelos de difusión. Modelos epidemiológicos. Belief propagation. Message passing.

Tema 9. Principios de Machine Learning basado en grafos. Matching, kernels y embeddings.

Tema 10. Caso práctico.

Metodología

La asignatura consta de 4 horas semanales presenciales (dos sesiones de 2 horas cada una). Las clases se realizarán en un aula con ordenadores (o se recomendará que el estudiante disponga de portátil en el aula). No se distingue entre horarios de teoría, problemas y prácticas de laboratorio en horarios y aula diferenciados, sino que se irán alternando según convenga a la misma aula. De manera general, se dedicará una semana a cada tema (en algunos temas serán dos semanas). Para cada tema se introducirán los conceptos teóricos, partiendo de materiales que se recomendará haber mirado previamente, alternando con actividades más aplicadas (resolución de problemas o seminarios). Se fomentará la participación activa en la resolución de problemas / ejercicios, trabajándolos en el aula, y fomentando la exposición y el debate analíticos de los

resultados. Algunas sesiones, principalmente las últimas del curso, consistirán en seminarios de trabajo más práctico, donde se planteará un problema a resolver de carácter más general, que requiera el diseño e implementación de una solución a partir de la combinación de los conceptos teóricos. Para facilitar el seguimiento, el primer día de curso se distribuirá un calendario detallado de sesiones.

Sesiones de teoría. Consiste en clases magistrales con material multimedia disponible en el Campus Virtual de la UAB. El objetivo principal de estas clases es introducir las nociones básicas que deben permitir al alumno / a coger una visión real de la teoría de grafos y su aplicación en sistemas de búsqueda de información, en particular en internet. Para fomentar sesiones más interactivas, y validar que se adquieren los conocimientos, se distribuirán previamente materiales a través del campus virtual, recomendando que se hayan mirado previamente a la sesión. Durante la sesión de teoría, se introducirán actividades prácticas (resolución de ejercicios o casos prácticos más extensos), y otras actividades de seguimiento (cuestionarios interactivos).

Sesiones de problemas. Durante el curso se alternarán las sesiones de clases magistrales con la resolución de ejercicios en frente del ordenador, con herramientas de apoyo como por ejemplo el entorno *NetworkX*. Serán ejercicios que permitirán monitorizar la comprensión de los contenidos teóricos. Se proporcionará una colección de ejercicios como base de trabajo. Aunque no es una división estricta, de las dos sesiones semanales, una se dedicará más a contenidos teóricos, y otra a trabajar problemas del tema correspondiente. Podrá haber evaluación de la entrega de algún ejercicio (anunciando previamente cuáles son los que los estudiantes deberán entregar).

Seminarios / casos prácticos. Algunas sesiones se dedicarán a resolver un problema más amplio y general, que requiera la combinación de varias herramientas vistas en los temas correspondientes. En los seminarios se promueve fundamentalmente la capacidad de análisis y síntesis, así como el razonamiento crítico y la toma de decisiones del alumno frente a la resolución del problema. Habrá un caso práctico que corresponde al último tema del curso. A medio curso puede haber un caso práctico más curso que cubra los primeros temas.

Nota: se reservarán 15 minutos de una clase dentro del calendario establecido por el centro o por la titulación para que el alumnado rellene las encuestas de evaluación de la actuación del profesorado y de evaluación de la asignatura o módulo.

Actividades

Título	Horas	ECTS	Resultados de aprendizaje
Tipo: Dirigidas			
Clases de problemas	15	0,6	1, 2, 3, 4, 5, 6, 7
Clases de teoría	30	1,2	2, 3, 6, 7
Seminarios y prácticas	15	0,6	1, 2, 3, 4, 5, 7
Tipo: Supervisadas			
Preparación problemas/seminarios	15	0,6	1, 2, 3, 4, 5, 6, 7
Tutorías	15	0,6	1, 2, 3, 4, 5, 6, 7
Tipo: Autónomas			
Estudio / preparación examen	25	1	1, 2, 3, 4, 5, 6, 7
Trabajo personal	30	1,2	1, 2, 3, 4, 5, 6, 7

Evaluación

La evaluación de la asignatura (sobre 10 puntos) constará de las siguientes pruebas evaluatorias:

Exámenes parciales (N1). Se harán dos exámenes parciales durante el curso (E1 y E2). Los exámenes contendrán tanto preguntas de conceptos teóricos como ejercicios de estilo de los que se habrán hecho en clase. La nota mínima para superar cada examen es de 5. En caso de que no se supere uno de los dos parciales, se dispondrá de un examen de recuperación a final de curso de la parte no superada. Corresponde a un 50% de la nota final.

Pruebas basadas en ejercicios en clases de problemas (N2). Se solucionarán cuestionarios, que pueden ser on-line, y / o resolución de problemas adicionales a los vistos en clase. Forma parte de la evaluación continua, y por tanto no es recuperable. Corresponde a un 10% de la nota final.

Entrega de un pequeño proyecto a partir del trabajo de los seminarios. En los seminarios se abordará un caso práctico de forma tutorizada. Se deberá entregar al final el resultado del trabajo. Este trabajo se hará en grupo. Forma parte de la evaluación continua y no tiene recuperación, pero se considerarán mecanismos de recuperación o compensación en determinados casos. Corresponde a un 40% de la nota final.

La nota final de la asignatura se calculará de la siguiente manera:

$$NF = 0,5 * N1 + 0,1 * N2 + 0,4 * N3$$

Para aprobar la asignatura es necesario haber conseguido una puntuación mínima de 5 en todas las calificaciones, así como en las pruebas parciales para liberar materia establecidos a lo largo del curso. A criterio de los profesores / as se podrá, sin embargo, establecer compensaciones entre las notas N1, N2 y N3.

La asignatura será evaluada como No Evaluable sólo en el caso que el alumno no se haya presentado a ninguna de las pruebas de evaluación contempladas ni haya entregado total o parcialmente los trabajos.

En caso de no superar la asignatura debido a que alguna de las actividades de evaluación no alcanza la nota mínima requerida, la nota numérica del expediente será el valor menor entre 4.5 y la media ponderada de las notas. Con las excepciones de que se otorgará la calificación de "no evaluable" a los / as estudiantes que no participen en ninguna de las actividades de evaluación, y de que la nota numérica del expediente será el valor menor entre 3.0 y la media ponderada de las notas en caso de que el estudiante haya cometido irregularidades en un acto de evaluación (y por tanto no será posible el aprobado por compensación).

Se otorgarán las Matrículas de Honor dentro de los máximos admitidos por la normativa de la UAB (en función del número de matriculados) a las notas más altas iguales o superiores a 9.

Para cada actividad de evaluación, se indicará un sitio , fecha y hora de revisión en la que el estudiante podrá revisar la actividad con el profesor / a. En este contexto, se podrán hacer reclamaciones sobre la nota de la actividad, que serán evaluadas por el profesorado responsable de la asignatura. Si el estudiante no se presenta en esta revisión, no se revisará posteriormente esta actividad.

Ver apartado "PLAGIO" sobre medidas en casos de irregularidades por plagio en las actividades de evaluación.

RECUPERACIONES:

Exámenes parciales (N1). Se harán dos exámenes parciales de teoría liberatorios durante horas lectivas. Los alumnos que no superen esta prueba (con nota igual o superior a 5), dispondrán de un examen de recuperación en la fecha de evaluación final programada para la titulación.

Pruebas basadas en ejercicios y en caso práctico (N2 y N3). El trabajo de ejercicios y práctico evalúa en forma de evaluación continuada durante el seguimiento en clase. Por lo tanto no habrá ninguna actividad de recuperación al final del curso.

FECHAS DE EVALUACIÓN:

Las fechas de evaluación y entrega de trabajos se publicarán en el campus virtual y pueden estar sujetos a cambios de programación por motivos de adaptación a posibles incidencias. Siempre se informará en el campus virtual sobre estos cambios ya que se entiende que es el mecanismo habitual de intercambio de información entre profesores y estudiantes.

ALUMNOS REPETIDORES:

No se guardan notas parciales (teoría o prácticas) de un curso para otro. Sin embargo, a criterio del profesor y en función de las evaluaciones de cursos previos, se podrán establecer compensaciones. Esta información se anunciará el día de la presentación de la asignatura, y el campus virtual.

PLAGIO:

Sin perjuicio de otras medidas disciplinarias que se estimen oportunas, y de acuerdo con la normativa académica vigente, las irregularidades cometidas por un estudiante que puedan conducir a una variación de la calificación se calificarán con un cero (0). Las actividades de evaluación calificadas de esta forma y por este procedimiento no serán recuperables. Si es necesario superar cualquiera de estas actividades de evaluación para aprobar la asignatura, esta asignatura quedará suspendida directamente, sin oportunidad de recuperarla en el mismo curso. Estas irregularidades incluyen, entre otros:

- la copia total o parcial de una práctica, informe, o cualquier otra actividad de evaluación;
- dejar copiar;
- presentar un trabajo de grupo no hecho íntegramente por los miembros del grupo;
- presentar como propios materiales elaborados por un tercero, aunque sean traducciones o adaptaciones, y en general trabajos con elementos no originales y exclusivos del estudiante;
- tener dispositivos de comunicación (como teléfonos móviles, smart watches, etc.) accesibles durante las pruebas de evaluación teóricas-prácticas individuales(exámenes).

En resumen: copiar, dejar copiar o plagiar en cualquiera de las actividades de evaluación equivale a un SUSPENSO con nota inferior a 3,0.

ACLARACIÓN FINAL:

Para cualquier duda o discrepancia, prevalecerá la información más actualizada que se comunicará el día de la presentación de la asignatura y que se publicará en el campus virtual.

Actividades de evaluación

Título	Peso	Horas	ECTS	Resultados de aprendizaje
Dos exámenes parciales	50%	4	0,16	1, 2, 3, 4, 5, 6, 7
Presentación trabajos prácticos y seminarios	50%	1	0,04	1, 2, 3, 4, 5, 7

Bibliografía

Los contenidos de la asignatura explicados en clase se extraen de diferentes fuentes. Muchos materiales son en línea, extraídos de materiales multimedia, vídeos, etc. Durante el curso, se proporcionarán los materiales de confección propia necesarios para el seguimiento de la asignatura a través del campus virtual. También se proporcionarán enlaces a documentación en línea, software para los ejercicios prácticos, etc. en formato libre.

Libros en línea de interés:

- Albert-Laszlo Barabási. Network Science. <http://networksciencebook.com/>
- David Easley and Jon Kleinberg. Networks, Crowds, and Markets. <https://www.cs.cornell.edu/home/kleinber/networks-book/>

Software

Para la resolución de problemas y casos prácticos se utiliza software abierto como Gephi, y la librería NetworkX en Python. Algunos ejercicios de clase se distribuirán en la plataforma Google Colab.