

Métodos Avanzados de Procesamiento y Gestión de Datos

Código: 104377
Créditos ECTS: 6

Titulación	Tipo	Curso	Semestre
2503758 Ingeniería de Datos	OT	4	1

Contacto

Nombre: Daniel Franco Puntos
Correo electrónico: daniel.franco@uab.cat

Uso de idiomas

Lengua vehicular mayoritaria: español (spa)
Algún grupo íntegramente en inglés: No
Algún grupo íntegramente en catalán: No
Algún grupo íntegramente en español: No

Otras observaciones sobre los idiomas

La documentació, enunciats de pràctiques i exàmens poden estar en català, castellà o angles

Equipo docente

Antonio Espinosa Morales

Prerequisitos

Esta asignatura no tiene PRE-REQUISITOS. Se recomienda tener cursadas las Asignaturas de Computación en Entornos a la Nube, Desarrollo de Aplicaciones de Datos Masivas, Redes Neuronales y Aprendizaje Profundo y Visualización de Datos.

Objetivos y contextualización

El objetivo de esta asignatura es conocer los métodos y conceptos avanzados de procesamiento y gestión de datos masivos, tanto desde el punto de vista de generación, adecuación, transmisión y almacenamiento como de procesamiento y análisis para extraer información útil. Asimismo, también es un objetivo utilizar las herramientas adecuadas para llevar a cabo el trabajo con datos masivos, tanto de forma interactiva y local como en batch y remota y en diferido y en tiempo real.

Competencias

- Concebir, diseñar e implementar sistemas de almacenamiento de datos de forma eficiente y segura.
- Prevenir y solucionar problemas, adaptarse a situaciones imprevistas y tomar decisiones.
- Que los estudiantes hayan desarrollado aquellas habilidades de aprendizaje necesarias para emprender estudios posteriores con un alto grado de autonomía.
- 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.

- Trabajar cooperativamente, en entornos complejos o inciertos y con recursos limitados, en un contexto multidisciplinar, asumiendo y respetando el rol de los diferentes miembros del equipo.

Resultados de aprendizaje

1. Estudiar las adaptaciones que se realizan a los algoritmos de análisis y consulta de datos para que preserven la privacidad de los datos de entrada, de los modelos aprendidos o de las salidas de los modelos usados en el ámbito de business intelligence.
2. Prevenir y solucionar problemas, adaptarse a situaciones imprevistas y tomar decisiones.
3. Que los estudiantes hayan desarrollado aquellas habilidades de aprendizaje necesarias para emprender estudios posteriores con un alto grado de autonomía.
4. 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.
5. Trabajar cooperativamente, en entornos complejos o inciertos y con recursos limitados, en un contexto multidisciplinar, asumiendo y respetando el rol de los diferentes miembros del equipo.

Contenido

1. Métodos avanzados de procesamiento y gestión de big data
2. Referencia de arquitecturas en la nube para gestión de datos masivos y diseño de aplicaciones: lotes, streaming, descentralizadas.
3. Análisis de desafíos y casos de estudio con big data orientados a la nube
4. Metodología de MVP para el diseño y desarrollo de soluciones en la nube para administrar datos masivos
5. Herramientas de evaluación para aplicaciones y sistemas de gestión de datos masivos

Metodología

Siguiendo la metodología del aprendizaje por retos, la asignatura se basa en un conjunto coordinado de trabajos prácticos que conducen a la elaboración y presentación de una propuesta de solución técnica del problema planteado por los agentes proponentes del reto, aportando finalmente algunas propuestas solución para resolverlos.

La tarea se realizará en grupos reducidos de estudiantes, que pueden agruparse de forma flexible en función de la dinámica de trabajo. La actividad práctica irá acompañada de un conjunto de sesiones teóricas y metodológicas de soporte, así como de la tutoría de todo el proceso de aprendizaje.

Las etapas de la metodología de aprendizaje por retos son:

1. Descubrimiento

La primera fase supone un doble reconocimiento:

- a) por un lado, del ámbito de aplicación y la problemática de estudio para cada uno de los grupos de trabajo, entrando en contacto con los primeros actores del tema;
- b) por otra, de las herramientas necesarias para desarrollar adecuadamente la investigación posterior: la elaboración de la diagnosis, y el procedimiento para realizar un Plan de Desarrollo.

2. Investigación

La investigación debe seguir el proceso de análisis de las tres partes de un plan de desarrollo: la comprensión de la demanda, de la oferta y de las necesidades. Esta etapa también tiene una doble vertiente:

- a) Por un lado se exploran los documentos de carácter general del reto propuesto.
- b) Por otra parte, se toma contacto con los recursos y servicios necesarios, realizando entrevistas a sus responsables y visitando instalaciones, observando a los usuarios y al entorno.

3. Identificación de necesidades y propuestas de mejora

La tercera fase se caracteriza por el desarrollo de un prototipo para la obtención de resultados, identificando con precisión las carencias en cada uno de los sectores, llegando a diseñar propuestas de mejora en un nivel inicial para los retos analizados por cada grupo.

Estas actividades contarán con el refuerzo de conferencias y talleres orientados a la obtención de resultados creativos.

4. Presentación de resultados

La comunicación de resultados es un paso imprescindible del proceso. Se realizará en tres formatos:

- Memoria de la propuesta
- Pósteres
- Presentaciones orales

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			
Análisis, diseño y desarrollo de prototipos	25	1	1, 2, 3, 4, 5
Clases expositivas sobre tecnología, metodología y sobre el caso de estudio	20	0,8	1, 4
Tipo: Autónomas			
Trabajo autónomo	100	4	1, 2, 3, 4, 5

Evaluación

La evaluación tendrá en cuenta:

- a) Los resultados finales del proceso de elaboración de la Memoria de Diseño de la solución propuesta (Memoria escrita, presentación oral y póster), donde se evaluará el procedimiento y la consecución en la resolución del reto planteado.
- b) El proceso gradual de aprendizaje, a partir de tres informes de seguimiento (MVPs incrementales).

La asignatura sigue un calendario de aprendizaje y de evaluación continua que es necesario ir siguiendo puntualmente. Es necesario respetar las fechas de entrega de las tareas. El retraso en las entregas supondrá una penalización.

La nota de la asignatura será la media de las notas obtenidas en los distintos ítems evaluados. La no realización de alguno de los ítems implica que la asignatura sea "No evaluable". Para poder promediar habrá que haber obtenido al menos un 4 en cada uno de los ítems evaluables.

Recuperación: la recuperación tiene como requisito que se hayan presentado todos los ítems solicitados en la evaluación.

Serán ítems recuperables los ejercicios de seguimiento y el póster. Se podrán recuperar, respectivamente, con otro examen y con la revisión del póster suspendido. Por su naturaleza, no se podrá recuperar la Memoria escrita ni las presentaciones orales.

En caso de que el estudiante realice cualquier irregularidad que pueda conducir a una variación significativa de la calificación de un acto de evaluación, se calificará con 0 este acto de evaluación, con independencia del proceso disciplinario que pueda instruirse. En caso de que se produzcan diversas irregularidades en los actos de evaluación de una misma asignatura, la calificación final de esta asignatura será 0.

Actividades de evaluación

Título	Peso	Horas	ECTS	Resultados de aprendizaje
Memoria final	30	2	0,08	1, 2, 3, 4, 5
Poster i presentación oral	20	1	0,04	1, 2, 3, 4, 5
informes MVP	50	2	0,08	1, 2, 3, 4, 5

Bibliografía

- Dan C. Marinescu. "Cloud Computing. Theory and Practice". Morgan-Kaufmann. 2018.

-AWS Certified Cloud Practitioner Study Guide; Ben Piper, David Clinton; Sybex (14 de junio de 2019); ISBN-10: 1119490707, ISBN-13: 978-1119490708

-Infrastructure as Code; Kief Morris; O'Reilly Media; 1 edition (June 17, 2016); ISBN-10: 1491924357, ISBN-13: 978-1491924358

-Amazon Web Services in Action, 2E; Andreas Wittig, Michael Wittig; Manning Publications; Edición: 2nd edition (30 de septiembre de 2018); ISBN-10: 1617295116, ISBN-13: 978-1617295119

-Microsoft Azure Essentials - Fundamentals of Azure, 2nd Ed; Michael Collier, Robin Shahan; 2016; https://download.microsoft.com/download/6/6/2/662DD05E-BAD7-46EF-9431-135F9BAE6332/9781509302963_I

-Mastering Cloud Computing : Foundations and Applications Programming. Buyya, Rajkumar;Vecchiola, Christian;y más Elsevier Science & Technology 2013. ISBN: ISBN number:9780124114548, ISBN number:9780124095397

This is Service Design Thinking: Basics - Tools - Cases; Marc Stickdorn; BIS Publishers; ISBN 9063692560 (ISBN13: 978906369256) 2012

Design Research: Methods and Perspectives Brenda Laurel; The MIT Press; 2003; ISBN-10 : 0262122634; ISBN-13 : 978-0262122634

Gamestorming: A Playbook for Innovators, Rulebreakers, and Changemakers; Dave Gray, Sunni Brown, James Macanujo; O'Reilly Media; ISBN-10 : 0596804172 ISBN-13 : 978-0596804176; 2010

The Lean Startup: How Today's Entrepreneurs Use Continuous Innovation to Create Radically Successful Businesses; Eric Ries ; Currency; ISBN-10 : 9780307887894, ISBN-13 : 978-0307887894; 2011

Software

Visual Studio Code

Apache Spark

Redis

Power BI

Qlick

Azure Cloud

AWS Cloud

Google Cloud