

Formulario de aprobación de curso de posgrado/educación permanente

Asignatura: Sistemas de Información para el Análisis de Grandes Volúmenes de Datos

(Si el nombre contiene siglas deberán ser aclaradas)

Modalidad:

(posgrado, educación permanente o ambas)

Posgrado

Educación permanente

Profesor de la asignatura ¹: Dr. Adriana Marotta, grado 4, Instituto de Computación
(título, nombre, grado o cargo, instituto o institución)

Profesor Responsable Local ¹:

(título, nombre, grado, instituto)

Otros docentes de la Facultad: MSc. Flavia Serra, grado 2, Instituto de Computación
(título, nombre, grado, instituto)

Docentes fuera de Facultad:

(título, nombre, cargo, institución, país)

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)

[Si es curso de posgrado]

Programa(s) de posgrado: Maestría en Informática (Peduciba), Doctorado en Informática (Peduciba)

Instituto o unidad: Instituto de Computación

Departamento o área: Grupo Concepción de Sistemas de Información

Horas Presenciales: 70

(se deberán discriminar las horas en el ítem Metodología de enseñanza)

Nº de Créditos: 10

[Exclusivamente para curso de posgrado]

(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 enseñanza)

Público objetivo: Estudiantes de posgrado interesados en el área de modelos, procesamiento y análisis de datos multidimensionales.

Cupos: No tiene cupo

(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, hasta completar el cupo asignado)

Objetivos: Los sistemas de soporte a la toma de decisiones proveen asistencia a distintos niveles gerenciales de una organización para analizar información estratégica. Estos sistemas se apoyan en Inteligencia de Negocio (Business Intelligence), la cual consiste en metodologías, arquitecturas y tecnologías que transforman datos crudos en información significativa y útil para la toma de decisiones. El modelo de datos multidimensional es el modelo típicamente utilizado para el procesamiento analítico en línea (OLAP), mecanismo utilizado para la explotación de los datos contenidos en este tipo de sistemas. Los objetivos principales de este curso son los siguientes: (i) brindar al estudiante conocimientos teóricos básicos sobre los sistemas de análisis multidimensional de datos, su diseño y su construcción, teniendo en cuenta los nuevos modelos y tecnologías para manejo de grandes volúmenes de datos, (ii) lograr que el estudiante adquiera cierta práctica en el desarrollo de estos sistemas, (iii) estudiar los principales avances en el área de investigación de este tema.

Conocimientos previos exigidos: Bases de Datos y Programación

Conocimientos previos recomendados:

Metodología de enseñanza:

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

Descripción de la metodología: El curso se desarrollará en base a: clases teóricas de exposición de los distintos temas, trabajos en grupo de profundización en algunos temas en base al estudio y presentación de artículos científicos, y desarrollo de un proyecto de aplicación de los conocimientos que se van adquiriendo.

[Obligatorio]

Detalle de horas:

- Horas de clase (teórico): 30
- Horas de clase (práctico): 10
- Horas de clase (laboratorio): 20
- Horas de consulta: 6
- Horas de evaluación: 4
 - Subtotal de horas presenciales: 70
- Horas de estudio: 30
- Horas de resolución de ejercicios/prácticos:
- Horas proyecto final/monografía: 50
 - Total de horas de dedicación del estudiante: 150

Forma de evaluación: Los estudiantes serán evaluados a través de la realización de un proyecto y una prueba final individual.

Temario:

1. Introducción general del curso
 - Motivaciones
 - Sistemas de información para análisis multidimensional de datos. Data Warehouse y OLAP.
 - Sistemas de información para análisis de grandes volúmenes de datos (Big Data)
2. Modelo Multidimensional y Diseño Conceptual
 - Modelos Multidimensionales
 - Problemas y pautas en Diseño Conceptual
 - Diseño Conceptual a partir de requerimientos y a partir de fuentes de datos
3. Diseño Lógico
 - Diseño Lógico Relacional
 - Diseño Lógico a partir del esquema conceptual
 - Problemas y pautas en Diseño Lógico
4. Big Data en Sistemas OLAP
 - Introducción a bases de datos No-SQL
 - Arquitecturas posibles
 - OLAP sobre bases de datos No-SQL
 - Inteligencia de Negocio incluyendo datos de redes sociales (Social BI)
5. Proceso de Carga y Actualización
 - Tipos de operaciones, integración y calidad de datos
 - ETL tradicional
 - ETL en Tiempo Real (ELT)
6. Tecnologías
 - Herramientas de carga, manipulación de datos y análisis.

Bibliografía:

(título del libro-nombre del autor-editorial-ISBN-fecha de edición)

- Vaisman, A.; Zimányi, E. (2014). Data Warehouse Systems Design and Implementation. Springer Series: Data-Centric Systems and Applications 2014, XXVI, 603 p. ISBN 978-3-642-54654-9
- Kimball, R. (2002). The Data Warehouse Toolkit. John Wiley & Sons, 2002. ISBN: 0-471-20024-7
- Chevalier, M.; El Malki, M.; Koplaku, A.; Teste, O.; Tournier, R. Implementation of multidimensional databases in column-oriented NoSQL systems. 19th East-European Conference on Advances in Databases and Information Systems (ADBIS), Sep 2015, Poitiers, France. pp. 79-91.
- Chevalier, M.; El Malki, M.; Koplaku, A.; Teste, O.; Tournier, R. Document-Oriented Models for Data Warehouses. (2016) In: 18th International Conference on Enterprise Information Systems (ICEIS 2016), 25 April 2016 - 28 April 2016 (Rome, Italy).
- Gómez, L.; Kuijpers, B.; Vaisman, A. Performing OLAP over Graph Data: Query Language, Implementation, and a Case Study. En Proceedings of the Int. Workshop on Real-Time Business Intelligence and Analytics, BIRTE, Munich, Germany, August 28, 2017, 2017, pp. 6:1–6:8.