



Ingeniería de datos

Syllabus

Equipo Docente

Domagoj Vrgoč

Domagoj Vrgoč obtuvo su grado de doctorado en la Escuela de Ingeniería de la Universidad de Edimburgo en el año 2014. Sus áreas de interés son bases de datos, la Web semántica, y teoría de computación. Su tesis de doctorado fue destacado en el concurso de British Computing Society en el año 2015 como una de las mejores de su generación. Domagoj ha servido como miembro del comité del programa en las mejores conferencias en su área como PODS, IJCAI, o ISWC.

Descripción del curso

Este curso está diseñado hacia la comprensión de modelos de datos. Un tema imprescindible en el mundo de hoy. Con esto en mente el curso tiene una orientación hacia la aplicación de conceptos en el mundo real.

La ingeniería de datos se encuentra hoy en día en varias encrucijadas y al mismo tiempo en un boom expansivo que la ha catapultado hacia un lugar fundamental tanto en el desarrollo de la industria como en la toma de decisiones financieras o políticas sociales, entre muchas otras aplicaciones.

En este curso, los estudiantes analizarán, en una primera parte, el modelo relacional de bases de datos y su lenguaje de consultas SQL. Con esto, los alumnos analizarán las ventajas y desventajas de este modelo y tendrán, por añadidura, una opinión más formada respecto de las potencialidades que la ingeniería de datos comprende y al mismo tiempo tendrán una relación experimental con las herramientas utilizadas en la disciplina.

Luego de completar esta primera parte, los estudiantes comprenderán el modelo de datos NoSQL. Como todo el curso, la orientación en los contenidos y análisis de casos tienen una es de carácter práctico y están dirigidos hacia su aplicación dentro de contextos reales. El curso cuenta con una serie de actividades formativas (clases de cátedra, videos, talleres), evaluaciones y un proyecto final.

Resultados de Aprendizaje

1. Elaborar modelos de datos para implementarlos sobre un motor relacional aplicando conceptos de dependencias funcionales y normalización .
2. Aplicar el lenguaje de consultas SQL a un nivel intermedio para realizar consultas sobre bases de datos.
3. Identificar las fortalezas y debilidades de las bases de datos NoSQL para implementarlas en el contexto adecuado.
4. Elaborar consultas sencillas en bases de datos de documentos para hacer búsquedas básicas de información.
5. Aplicar las bases de datos en procesos de análisis de datos e identificar problemas éticos y de privacidad en el manejo de datos

La metodología de enseñanza comprende:

- Exposiciones a través de clases sincrónicas
- Análisis de casos
- Test basados en la interacción con bases de datos
- Contenidos entregados por las video clases
- Lecturas y reflexión individual a través de foros.

Estructura del curso

El curso está estructurado de la siguiente forma:

1. Módulo 1

- 1.1. Conceptos Fundamentales de bases de datos
- 1.2. Modelo Relacional de datos
- 1.3. Introducción al lenguaje de consulta SQ

2. Módulo 2

- 2.1. SQL avanzado (agregación, consultas anidadas)
- 2.2. Diseño de una BD Relacional

3. Módulo 3

- 3.1. Rol de SQL en Ciencia de Datos
- 3.2. Bases de Datos NoSQL

El requisito académico comprende cuestionarios, tareas, foros de discusión y un trabajo final. También, el curso comprende actividades formativas sin evaluación. El porcentaje de aprobación de las evaluaciones es de 50%.

Actividad	Evaluación
Cuestionarios	20% nota final
Tareas	45% nota final
Foros de discusión	20% nota final
Trabajo Final	15% nota final

Información General

Sigla: MDS3000

Créditos: 5 UC

Requisitos: No tiene

Restricciones: No tiene

Carácter: Mínimo

Tipo: Cátedra

Calificación: Estándar

Horas: 90 horas total (24 directas y 66 indirectas)

Integridad Académica

La Pontificia Universidad Católica de Chile aspira a lograr la excelencia en la creación y transferencia de conocimiento y en la formación de las personas, inspirada en una concepción católica y siempre al servicio de la iglesia y de la sociedad.

A través de la Educación Continua, se espera aportar al cumplimiento de los ejes de desarrollo de la universidad, innovando en la docencia, fortaleciendo la interdisciplina y profundizando el compromiso público.

Asimismo la Dirección de Educación Continua, se compromete a una constante búsqueda de la satisfacción de las expectativas de los estudiantes, a fortalecer las competencias de los funcionarios y profesionales, a mejorar continuamente el sistema de gestión de la calidad y a mantenerse siempre dentro del marco legal vigente.

Bibliografía

Mínima:

- Ramakrishnan, R., Gehrke, J., "Database Management Systems (3rd Ed.)", McGraw-Hill, 2003.

Complementaria:

- Elmasri, R., Navathe, S. "Fundamentals of Database Systems (7th Ed.)", Pearson, 2015.
- Walmsley, P., "XQuery", 2nd Ed. O'Reilly, 2016.
- Redmont E. and Wilson, J. "Seven Databases in Seven Weeks: A Guide to Modern Databases and NoSQL" Pragmatic Bookshelf 2012
- CouchDB Reference Manual, Nov 2015.
- Kemper, C. "Beginning Neo4J" Apress 2015
- Ponniah P. "DW fundamentals for IT Professionals" 2nd Ed Wiley 2010
- DuCharme B. "Learning SparQL" 2nd Ed O'Reilly 2013
- Walmsley P. "XQuery" 2nd Ed O'Reilly 2016