



Data Mining y Big Data: un enfoque integrado

Hoy en día casi cualquier empresa u organismo tiene datos almacenados sobre su actividad, clientes o proveedores, y se puede extraer información útil analizando esos datos. En particular, si se almacenan datos sobre clientes, se puede determinar cómo servirles mejor, cómo tenerlos más satisfechos, conseguir más negocio o aumentar su fidelidad hacia nuestra marca. Simplemente entendiendo qué quieren, qué cosas les resultan importantes y cuáles no, qué les gusta y qué les irrita en su relación con la empresa.

Data Mining y Big Data, son dos términos que están muy relacionados entre sí y tienen fines muy parecidos, pero no exactamente iguales. De forma resumida, se puede decir que Data Mining consiste en el conjunto de técnicas para la extracción de la información y que Big Data es la tecnología capaz de capturar, gestionar y procesar en un tiempo razonable y de forma veraz estos datos, es decir que hay una integración mutua entre técnica y herramienta.

Objetivo: introducir una mirada diferente e integradora al Análisis Inteligente de Datos con énfasis en el procesamiento de Datos Masivos o Big Data.

Requisito de admisión: conocimientos básicos de bases de datos.

Cupo: cupo limitado.

Aranceles:

- Público en general: \$6.700.
- Docentes FACET: \$3.350.
- Estudiantes FACET: \$1.500.

Contenidos mínimos:

- **Bases de Datos no Relaciones:** Introducción y evolución. Propiedades BASE. Comparación con BD relacionales. Tipos de Bases de datos NoSQL. Escalabilidad – Desafíos. Transformación y migración SQL a NoSQL.
- **Big Data:** Fundamentos de Big Data. Definición y dimensiones en Big Data. Relación entre Big Data y Científico de Datos/Ingeniero de Datos. Aplicaciones de Big Data. Modelos de datos y modelos de procesamiento en Big Data. Fundamentos de Map Reduce. Ética, seguridad, privacidad en Big Data. Casos de uso - IoT.
- **Data Warehousing:** Concepto de Data Warehouse y características. Concepto de Data Marts y características. Proceso de creación de Data Warehouse y Data Marts.
- **Minería de Datos:** Introducción. Concepto de minería de datos. Terminología relacionada. Tipos de datos. Tipos de modelos. Aplicaciones. El Proceso de KDD y sus fases. Preparación



de los datos. Técnicas y algoritmos de minería de datos. Evaluación, interpretación y difusión de la información.

Fecha de inicio: viernes 23 de agosto de 2019.

Modalidad de dictado: 2 clases teórico-prácticas semanales.

Carga horaria: 40 horas de reloj.

Lugar de realización: Laboratorio de redes de computadoras, Block 1 – 3º piso (DEEC – FACET – UNT).

Forma de aprobación: 80% de asistencia a clases y aprobación de un examen final con nota igual o superior a 6/10.

Expositores:

- Mag. Lic. Pablo J. Thomas (FI – UNLP). Se adjunta Curriculum Vitae resumido.
- Mag. Ing. Carlos Albaca Paraván (FACET – UNT).

Bibliografía:

- Introducción a las Bases de Datos. Fundamentos y Diseño. Rodolfo Bertone y Pablo Thomas. Prentice Hall – Person Education, 2011. ISBN 978-987-615-136-8
- Fundamentos de Bases de Datos. Silberschatz y Korth. McGraw Hill. ISBN 9788448136543
- The Data Warehouse Toolkit: The Complete Guide to Dimensional Modeling. Ralph Kimball y Margy Ross. Wiley Edition. ISBN: 978-0471200246
- NOSQL PROBLEM LITERATURE REVIEW, Jarosław Kurpanik, Małgorzata Pańkowska. ISSN 2083-8611. Nr 234. 2015.
- Big Data: A Brief investigation on NoSQL Databases, Roshni Bajpayee, Sonali Priya Sinha, Vinod Kumar, , International Journal of Innovations & Advancement in Computer Science, IJIACS ISSN 2347 –8616, Volume 4, Issue 1 January 2015
- Big Data for Dummies. Judith Hurwitz, Alan Nugent, Dr. Fern Halper and Marcia Kaufman. Published by John Wiley & Sons, Inc. ISBN 978-1-118-64401-0
- Next-Generation Big Data Analytics: State of the Art, Challenges, and Future Research Topics, Zhihan Lv, Houbing Song, Senior Member, IEEE, Pablo Basanta-Val, Anthony Steed, and Minh Jo, Senior Member, IEEE, IEEE Transactions On Industrial Informatics, Vol. 13, No. 4, August 2017.
- Introducción a la minería de datos. José Hernández Orallo, María José Ramírez Quintana y César Ferri Ramírez. Pearson Educación. 2004. ISBN: 84-205-4091-9
- Data Mining: practical machine learning tolos and techniques (4th edition). Ian H. Witten, Eibe Frank, Mark A. Hall and Christopher J. Pal. Elsevier. 2017. ISBN: 978-0-12-804291-5.