Big Data. Técnicas de análisis de datos

Curso Académico: (2020 / 2021) **Fecha de revisión: 31/07/2020 11:37:03**

Departamento asignado a la asignatura: Departamento de Biblioteconomía y Documentación

Coordinador/a: GARCIA ZORITA, JOSE CARLOS

Tipo: Optativa Créditos ECTS : 3.0

Curso : 1 Cuatrimestre : 0

REQUISITOS (ASIGNATURAS O MATERIAS CUYO CONOCIMIENTO SE PRESUPONE)

- Bases de Datos (340-17452).
- Análisis Estadístico Avanzado (340-17467).
- Visualización de Datos (340-17468).
- Técnicas Métricas de Información (340-17461).
- Ciencia de Datos (340-17477).
- Visualización de información (335-17283).

OBJETIVOS

COMPETENCIAS BÁSICAS DE LA MATERIA

CB8 Aprender a diseñar estrategias para el análisis y explotación de grandes volúmenes de datos.
CB9 Realizar propuestas de gestión de datos adaptadas a distintos contextos y organizaciones.

COMPETENCIAS GENERALES DE LA MATERIA

CG2 Aprender a identificar posibles líneas de trabajo e investigación relacionadas con el análisis, depuración y explotación de datos.
CG6 Conocer los distintos modelos de negocio relacionados con el Big data.
CG8 Aprender a identificar el valor y uso potencial de los datos.
CG9 Aprender a adaptar y desarrollar los métodos y técnicas de la documentación a los nuevos conocimientos y aptitudes que requiere la gestión y análisis de datos.

RESULTADOS DEL APRENDIZAJE

- El alumno adquirirá conocimientos sobre las técnicas y métodos necesarios para diseñar y evaluar estrategias para la gestión de datos.

El estudiante tras la superación de la materia deberá:

- Aplicar técnicas para elaborar estudios e informes que permitan analizar y evaluar la gestión de datos en organizaciones.
- Conocer y comprender conceptos y términos relacionados con big data.
- Diseñar, planificar e implementar procesos de depuración y análisis de datos.
- Aprender a identificar los datos con valor estratégico y de alto impacto para la organización.
- Conocer los distintos métodos y técnicas de extracción, depuración y enlazado de datos.
- Adquirir un dominio suficiente de distintas herramientas tecnológicas para realizar procesos de depuración, extracción y enlazado de datos.

El estudiante, tras la superación de la asignatura, deberá:

- Conocer las limitaciones y capacidades técnicas requeridas en Big Data.
- Conocer y comprender el valor de los datos procedentes de distintas fuentes.
- Conocer herramientas que permitan diseñar, preparar, analizar y manejar grandes volúmenes de

información, estructurada o no estructurada.
- Usar técnicas de análisis de datos que permitan obtener conclusiones válidas para la toma de decisiones.
- Normalizar, relacionar y enriquecer la información procedente de distintos conjuntos de datos.
- Utilizar distintas técnicas y herramientas para el análisis, depuración y enlazado de datos.

DESCRIPCIÓN DE CONTENIDOS: PROGRAMA

Módulo 1.- Introducción. Fundamentos básicos. Conceptos
1.1. Conceptos principales. El Big data en el contexto de la gestión de documentos
1.2. Nuevas tendencias en la gestión de datos. Del Big data al Smart data
1.3. El Big data en las estrategias de las organizaciones públicas y privadas
1.4. Principales iniciativas y sectores de aplicación

Módulo 2.- Técnicas de análisis y depuración de datos.
2.1. Auditoría y análisis de la calidad de los datos
2.2. Depuración de datos no estructurados. Principales herramientas. Introducción a Open Refine.
2.3. Minería de datos no estructurados
2.4. Definición de modelos predictivos con técnicas de data mining

Módulo 3.- Herramientas de procesamiento y enlazado de datos
3.1. Hacia la web semántica. Interoperabilidad y open linked data
3.2. Técnicas de Linked data. Utilización de web services para relacionar bases de datos
3.3. Análisis y extracción de información. Soluciones open source
3.4. Desarrollo de un proyecto de procesamiento de datos. Estudio de casos prácticos

ACTIVIDADES FORMATIVAS, METODOLOGÍA A UTILIZAR Y RÉGIMEN DE TUTORÍAS

ACTIVIDADES FORMATIVAS DEL PLAN DE ESTUDIOS REFERIDAS A MATERIAS

AF1 Trabajo individual para el estudio de materiales teóricos y prácticos.
AF2 Ejercicios para la resolución de casos prácticos.
AF3 Clases presenciales teórico-prácticas.
AF4 Tutorías.
AF5 Trabajo final.
AF6 Participación activa en los foros de la asignatura.

METODOLOGÍA DOCENTE

MD1 Exposiciones orales que describen los conceptos principales de la materia.
MD2 Lectura crítica de textos recomendados.
MD3 Resolución de casos prácticos y problemas individualmente o en grupo
MD4 Exposición y discusión en clase de temas relacionados con el contenido de la materia, así como de casos prácticos.
MD5 Elaboración de trabajos e informes de manera individual o en grupo
MD6 Lectura de materiales docentes teóricos y prácticos

RÉGIMEN DE TUTORÍAS

Los horarios de las tutorías se podrán consultar en Aula Global. Además de las tutorías oficiales, los alumnos pueden solicitar y concertar con el profesor tutorías adicionales fuera de esos horarios.

SISTEMA DE EVALUACIÓN

Peso porcentual del Examen/Prueba Final:	50
Peso porcentual del resto de la evaluación:	50

SISTEMA DE EVALUACIÓN

Test y ejercicios prácticos	50 %
Examen final (*)	50%

Peso porcentual del Examen/Prueba Final:	50
Peso porcentual del resto de la evaluación:	50

(*) El examen final se realizará en modalidad presencial en la Universidad Carlos III de Madrid, y se deberá superar para aprobar la asignatura.

CONVOCATORIA EXTRAORDINARIA

En esta asignatura, en ausencia de normativa de la Universidad específica para estudios de posgrado, si el alumno no ha seguido la evaluación continua, tendrá derecho a la realización de un examen que le permita obtener el 60% de la calificación máxima final en la convocatoria ordinaria. En la extraordinaria, el examen le permitirá obtener hasta el 75% de la calificación máxima final.

BIBLIOGRAFÍA BÁSICA

- CABALLERO, Rafael y MARTIN, Enrique Las bases del Big Data, Madrid: CATARATA, 2017

- GÓMEZ GARCÍA, José Luis Introducción al big data, Barcelona: UOC, 2015

- JOYANES AGUILAR, Luis Big data: análisis de grandes volúmenes de datos en organizaciones, Barcelona: Marcombo, 2013

- LARA TORRALBO, Juan Alfonso Minería de datos, Madrid: CEF, 2014

- MARR, Bernard Big Data. La utilización del Big Data, el análisis y los parámetros SMART para tomar mejores decisiones y aumentar el rendimiento, Madrid: TEEL, 2016

- MARR, Bernard Big data en la práctica, Madrid: TEEL, 2017

- MAYER-SCHÖNBERGER, Viktor Big data: la revolución de los datos masivos, Madrid: Turner, 2013

- NETTLETON, David F Data mining: fundamentos y metodologías, Barcelona: UOC, 2007

- SCHMARZO, Bill Big data: el poder de los datos, Madrid: Anaya Multimedia, 2014

- SIEGEL, Eric Analítica predictiva: predecir el futuro utilizando Big Data, Madrid: Anaya Multimedia, 2013