

Guía docente de la asignatura

Big Data I

Fecha última actualización: 27/07/2021
Fecha de aprobación por la Comisión Académica: 27/07/2021

Máster

Máster Universitario en Ciencia de Datos e Ingeniería de Computadores

MÓDULO

Módulo Big Data y Cloud Computing

RAMA

Ingeniería y Arquitectura

CENTRO RESPONSABLE DEL TÍTULO

Escuela Internacional de Posgrado

Semestre

Segundo

Créditos

4

Tipo

Optativa

Tipo de enseñanza

Presencial

PRERREQUISITOS Y/O RECOMENDACIONES

- Conocimiento básico del S.O. Linux: uso y acceso remoto a servidores (ssh).
- Conocimientos de programación de ordenadores.
- Conocimientos básicos de bases de datos.

BREVE DESCRIPCIÓN DE CONTENIDOS (Según memoria de verificación del Máster)

CloudComputing. Tecnologías para el procesamiento de datos en la nube.
Almacenamiento Masivo de Datos: Datawarehouse, NoSQL, bases de datos no estructuradas.

COMPETENCIAS

COMPETENCIAS BÁSICAS

- CB6 - Poseer y comprender conocimientos que aporten una base u oportunidad de ser originales en desarrollo y/o aplicación de ideas, a menudo en un contexto de investigación.
- CB7 - Que los estudiantes sepan aplicar los conocimientos adquiridos y su capacidad de resolución de problemas en entornos nuevos o poco conocidos dentro de contextos más amplios (o multidisciplinares) relacionados con su área de estudio.



- CB8 - Que los estudiantes sean capaces de integrar conocimientos y enfrentarse a la complejidad de formular juicios a partir de una información que, siendo incompleta o limitada, incluya reflexiones sobre las responsabilidades sociales y éticas vinculadas a la aplicación de sus conocimientos y juicios.
- CB9 - Que los estudiantes sepan comunicar sus conclusiones y los conocimientos y razones últimas que las sustentan a públicos especializados y no especializados de un modo claro y sin ambigüedades.
- CB10 - Que los estudiantes posean las habilidades de aprendizaje que les permitan continuar estudiando de un modo que habrá de ser en gran medida autodirigido o autónomo.

COMPETENCIAS GENERALES

- CG01 - Capacidad de acceso y gestión de la información
- CG02 - Capacidad de análisis y síntesis
- CG03 - Capacidad de organización y planificación
- CG04 - Capacidad emprendedora
- CG05 - Capacidad para tomar decisiones de forma autónoma
- CG07 - Motivación por la calidad
- CG08 - Capacidad para trabajar en equipo

COMPETENCIAS ESPECÍFICAS

- CE03 - Capacidad para la aplicación de técnicas y metodologías que permitan abordar desde nuevas perspectivas los problemas de interés, gracias a la disponibilidad de las plataformas de computación y comunicación con altos niveles de prestaciones.
- CE04 - Capacidad de análisis de aplicaciones en ámbitos de biomedicina y bioinformática, optimización y predicción, control avanzado, y robótica bioinspirada, tanto desde el punto de vista de los requisitos para una implementación eficaz de los algoritmos y las técnicas de computación que se usan para abordarlas, como de las características deseables en las arquitecturas donde se ejecutan
- CE05 - Capacidad para modelar y resolver problemas reales o académicos mediante técnicas de ciencia de datos
- CE06 - Capacidad para modelar y resolver problemas reales o académicos mediante tecnologías inteligentes o de inteligencia computacional.
- CE08 - Conocer algunas de las principales aplicaciones de la ciencia de datos y de las tecnologías inteligentes

COMPETENCIAS TRANSVERSALES

- CT01 - Ser consciente de la importancia del desarrollo sostenible y demostrar sensibilidad medioambiental.
- CT02 - Ser consciente del derecho a la no discriminación y al acceso universal al conocimiento de las personas con discapacidad.

RESULTADOS DE APRENDIZAJE (Objetivos)

El alumno aprenderá a manejar las tecnologías de cloud computing, y almacenamiento y procesamiento de grandes volúmenes de datos bajo las tecnologías actuales de Big Data tales como Hadoop y Spark, así como su procesamiento en bases de datos. En particular, el alumno:



- Conocerá y manejará la instalación de sistemas de cloud computing, servidores de Big Data bajo el ecosistema Hadoop, así como el diseño de servicios
- Aprenderá a manejar bases de datos NoSQL en la nube.
- Sabrá aplicar las técnicas anteriormente mencionadas para modelar y resolver problemas reales o académicos de big data.
- Conocerá las principales herramientas software existentes para el manejo de información en la nube y la tecnología big data: HIVE, PIG ...

PROGRAMA DE CONTENIDOS TEÓRICOS Y PRÁCTICOS

TEÓRICO

1. Introducción a Cloud Computing
 - Concepto de computación en la nube.
 - Niveles de servicios cloud.
 - Introducción a Big Data.
2. Big Data: Almacenamiento de datos masivos
 - Introducción a las bases de datos NoSQL.
3. Procesamiento y almacenamiento de datos.
 - Ecosistema Hadoop.

PRÁCTICO

1. Introducción a Cloud Computing
 - Despliegue de máquinas virtuales y contenedores.
 - Aplicaciones de cloud computing: Big Data con Spark.
2. Big Data: Almacenamiento de datos masivos
 - Uso y aplicación de herramientas de bases de datos: MongoDB, Base, Cassandra
3. Big Data: Procesamiento.
 - Aplicación de Apache Impala.

BIBLIOGRAFÍA

BIBLIOGRAFÍA FUNDAMENTAL

- R. Buyya, J. Broberg, A. Goscinski, (ed.) “Cloud Computing. Principles and Paradigms”, Wiley, 2011
- K. Chodorow, “MongoDB. The Definitive Guide”, O’Reilly, 2013
- J. Hurwitz, M. Kaufman, F. Halper, R. Bloor, “Cloud Computing for Dummies”, Wiley, 2010
- N. Marz, J. Warren, “Big Data: Principles and best practices of scalable realtime data systems”, Manning, 2014
- M. Miller, “Cloud Computing. Web-based applications that change the way you work and collaborate online”, Que, 2009.
- T. White, “Hadoop: The Definitive Guide”, O’Reilly, 3rd Ed. 2012

BIBLIOGRAFÍA COMPLEMENTARIA



- S.P. Kane, "Docker Up & Running: Shipping Reliable Containers in Production", O'Reilly, 2018.
- H. Karau, et al. "Learning Spark: Lightning-Fast Big Data Analysis", O'Reilly, 2015.
- I. Mell, "Docker in Practice", Manning, 2015.
- K. Morris, "Infrastructure as Code: Managing Servers in the Cloud", O'Reilly, 2015.

ENLACES RECOMENDADOS

- spark.apache.org
- hadoop.apache.org
- aws.amazon.com
- www.azure.com
- www.docker.com
- <https://m.oursky.com/saas-paas-and-iaas-explained-in-one-graphic-d56c3e6f4606>
- <https://www.mongodb.com/>

METODOLOGÍA DOCENTE

- MD01 Lección magistral/expositiva
- MD02 Resolución de problemas y estudio de casos prácticos
- MD03 Prácticas de laboratorio
- MD04 Seminarios
- MD05 Análisis de fuentes y documentos
- MD06 Realización de trabajos en grupo
- MD07 Realización de trabajos individuales
- MD08 Seguimiento del TFM

EVALUACIÓN (instrumentos de evaluación, criterios de evaluación y porcentaje sobre la calificación final)

EVALUACIÓN ORDINARIA

El método de evaluación consistirá en:

- La parte de Cloud Computing supondrá un 50% de la calificación global de la asignatura y se evaluará mediante un examen escrito con múltiples preguntas sobre la materia y la realización de un trabajo práctico sobre el uso de tecnologías cloud.
- La parte de Almacenamiento de Datos Masivos supondrá un 40% de la calificación global de la asignatura y se evaluará mediante trabajos prácticos. Se le entregará a cada alumno un conjunto de bases de datos y se pedirá que aplique las herramientas estudiadas en MongoDB para responder a una serie de cuestiones plantadas: consultas para MongoDB y bases de datos NoSQL.
- La parte de procesamiento mediante Apache Impala supondrá un 10% de la calificación final. Se evaluará mediante un trabajo práctico que consistirá en:
 - la selección de un conjunto de datos experimental
 - la realización de transformaciones mediante ImpalaQL sobre el conjunto



seleccionado.

Para la evaluación, se hará uso de las plataformas docentes y herramientas de trabajo colaborativo proporcionadas por la Universidad de Granada.

Para obtener una calificación en la asignatura, será obligatorio realizar la evaluación de todas las partes de la asignatura y que la calificación global iguale o supere la calificación de 5 sobre 10. En caso de no evaluar todas las partes de la asignatura, la calificación obtenida no podrá superar el 4,5 sobre 10.

Si no se supera la asignatura en la Convocatoria Ordinaria, el estudiante o la estudiante podrán optar por conservar la calificación de aquella parte o aquellas partes que superen la calificación de 5 sobre 10 para la Convocatoria Extraordinaria.

EVALUACIÓN EXTRAORDINARIA

El método de evaluación consistirá en:

- La parte de Cloud Computing supondrá un 50% de la calificación global de la asignatura y se evaluará mediante un examen escrito con múltiples preguntas sobre la materia y la resolución de ejercicios prácticos usando herramientas de tecnología cloud.
- La parte de Almacenamiento de Datos Masivos supondrá un 40% de la calificación global de la asignatura y se evaluará mediante trabajos prácticos. Se le entregará a cada alumno un conjunto de bases de datos y se pedirá que aplique las herramientas estudiadas en MongoDB para responder a una serie de cuestiones plantadas: consultas para MongoDB y bases de datos NoSQL.
- La parte de procesamiento mediante Apache Impala supondrá un 10% de la calificación final y se evaluará mediante un trabajo práctico que consistirá en:
 - la selección de un conjunto de datos experimental
 - la realización de transformaciones mediante ImpalaQL sobre el conjunto seleccionado.

Aquel estudiante que hubiera obtenido una calificación igual o superior a 5 sobre 10 en la evaluación de alguna de las partes referidas anteriormente en la convocatoria ordinaria no tendrá obligación de volver ser evaluado de esa parte en la convocatoria extraordinaria, pudiendo conservar la obtenida en la convocatoria ordinaria.

Para obtener una calificación en la asignatura, será obligatorio obtener calificación en todas las partes de la asignatura y que la calificación global iguale o supere la calificación de 5 sobre 10. En caso de no evaluar todas las partes de la asignatura, la calificación obtenida no podrá superar el 4,5 sobre 10.

EVALUACIÓN ÚNICA FINAL

Alternativamente a la evaluación continua, el alumno puede optar por la evaluación única final según lo dispuesto en la normativa de evaluación y de calificación de los estudiantes de la Universidad de Granada (<http://secretariageneral.ugr.es/pages/normativa/ugr/ncg7121>). Para acogerse a la evaluación única final, el estudiante, en las dos primeras semanas de impartición de la asignatura, lo solicitará al Coordinador del Máster, quien dará traslado al profesorado correspondiente, alegando y acreditando las razones que le asisten para no poder seguir el sistema de evaluación continua.

La evaluación única final se realizará en un solo acto académico. Dicha prueba (evaluada de 0 a





10) incluirá pruebas tanto de tipo teórico como práctico que garanticen que el alumno ha adquirido la totalidad de las competencias descritas en esta guía docente.

