

## Cartas Descriptivas

Datos de identificación				
Unidad Académica	<b>Facultad de Ciencias Administrativas y Sociales (Ensenada). Facultad de Ciencias Administrativas (Mexicali). Facultad de Contaduría y Administración (Tijuana).</b>			
Programa	<b>Maestría en Gestión de Tecnologías de la Información y la Comunicación</b>			
Nombre de la asignatura	<b>Big Data</b>			
Tipo de Asignatura	<b>Optativa</b>			
Clave (Posgrado e Investigación)	<b>4054</b>			
Horas teoría	<b>2</b>	Horas laboratorio	<b>0</b>	Créditos Totales
Horas taller	<b>2</b>	Horas prácticas de campo	<b>0</b>	<b>6</b>
Perfil de egreso del programa				
El egresado de la Maestría en Gestión de Tecnologías de la Información y la Comunicación tendrá la capacidad de planear, organizar, desarrollar, dirigir, controlar, proponer e implementar proyectos tecnológicos innovadores, en las organizaciones, generando soluciones que contribuyan a la competitividad de las mismas, mediante la aplicación de metodologías y técnicas vanguardistas en el ámbito de las TIC dentro de un marco social ético, responsable y sostenible.				
Definiciones generales de la asignatura				
<b>Aportación de esta materia al perfil de egreso del estudiante.</b>	Diseñar, organizar, desarrollar, proponer e implementar proyectos tecnológicos innovadores altamente escalables para aceptar, almacenar y analizar grandes volúmenes de datos no estructurados en modo batch y / o tiempo real, dentro de las organizaciones, generando soluciones que contribuyan a la competitividad de las mismas, mediante la aplicación de metodologías y técnicas vanguardistas en el ámbito de las TIC.			
<b>Descripción de la orientación de la asignatura en coherencia con el perfil de egreso.</b>	La asignatura optativa se orienta a fortalecer los conocimientos de manipulación, almacenamiento y análisis de grandes cantidades de datos, con el propósito de que tengan los elementos necesarios para proponer sistemas altamente escalables en beneficio de la organización beneficiaria.			
<b>Cobertura de la asignatura.</b>	Se revisarán las técnicas de MapReduce para el procesamiento paralelo y Hadoop, un framework de código abierto que permite implementar de forma barata y eficiente MapReduce en problemas de escala en internet. Se consideran herramientas que proporcionan un acceso similar a SQL para datos no estructurados: Pig y Hive. Se analizarán las soluciones de almacenamiento NoSQL.			
<b>Profundidad de la asignatura.</b>	La asignatura va desde la investigación documental para conocer los fundamentos y teoría de Big Data, hasta el desarrollo de una aplicación sencilla, pasando por el desarrollo de prácticas en la misma universidad en donde experimenten los conceptos vistos en clase.			

<b>Temario</b>			
<b>Unidad</b>	<b>Objetivo</b>	<b>Tema</b>	<b>Producto a evaluar (evidencia de aprendizaje)</b>
I. Conceptualización de Big Data	Identificar los conceptos de Big Data a través de una investigación bibliográfica para que el estudiante se familiarice con el tema.	1.1. Introducción a Biga Data 1.2. Conceptualización de Big Data 1.3. Porqué utilizar Big Data 1.4. Impactos del uso de Biga Data	Examen escrito  Reporte de investigación
II. Metodologías de Big Data	Examinar con creatividad las metodologías más comunes en el análisis de datos a través de diferentes metodologías para reconocer cuándo utilizar una en particular.	2.1. Clasificación de datos. 2.2. Métodos Estadísticos. 2.3. Árboles de decisión. 2.4. Regresión. 2.5. Dimensionalidad y selección de características	Reporte de prácticas
III. Técnicas de depuración de datos	Seleccionar la técnica de limpieza de datos más conveniente mediante un estudio comparativo para la manipulación de los mismos.	3.1. Calidad de los datos 3.2. Limpieza de datos. 3.3. Métodos más utilizados para la limpieza de datos 3.4. Preparación de los datos	Reporte de prácticas
IV. Técnica MapReduce	Analizar las técnicas MapReduce mediante una investigación documental para su aplicación en la manipulación de datos.	4.1. Concepto de procesamiento en paralelo 4.2. Técnicas MapReduce 4.3 Función Map() 4.4 Función Reduce() 4.5 Arquitectura del MapReduce	Reporte de prácticas
V. Procesamiento Hadoop	Implementar el procesamiento en paralelo de datos a través de un marco de trabajo abierto para mejorar la productividad del desarrollador.	5.1. Introducción a Hadoop 5.2. Procesamiento por lotes con Hive (HiveQL) y Pig Latin. 5.3. HBase y sistemas NoSQL. 5.4. Análisis de grandes volúmenes de datos con Mahout	Reporte de prácticas  Proyecto final
<b>Estrategias de aprendizaje utilizadas:</b> <ul style="list-style-type: none"> <li>• Exposición por parte del instructor.</li> <li>• Participación de los estudiantes en los temas expuestos por parte del instructor.</li> <li>• Reporte de las prácticas desarrolladas por los estudiantes.</li> <li>• Desarrollo de un proyecto sencillo por parte de los estudiantes.</li> </ul>			

## Métodos y estrategias de evaluación:

- Exámenes parciales 10%
- Participación en clase 10%
- Ensayos y prácticas 20%
- Proyecto final 60%

## Bibliografía:

### Básica:

Data Science and Big Data Analytics: Discovering, Analyzing, Visualizing and Presenting Data  
EMC Education Services  
Wiley; 1 edition, 2015  
ISBN-13: 978-1118876138 ISBN-10: 111887613X

Instalar Big Data paso a paso: Hadoop, Hive y Spark con seguridad Kerberos SSL y LDAP  
Victor J. Isla  
Amazon Digital Services LLC, 2016  
ASIN: B01CEIQOY

Hadoop: The Definitive Guide  
Tom White  
O'Reilly Media; 3 Edition, 2012  
ISBN13: 9781449389734

Mining the Social Web: Data Mining Facebook, Twitter, LinkedIn, Google+, GitHub, and More  
Matthew A. Russell  
O'Reilly Media; Second Edition edition, 2013  
ISBN-13: 978-1449367619 ISBN-10: 1449367615

Big Data: Principles and best practices of scalable realtime data systems  
Nathan Marz, James Warren  
Manning Publications; 1 edition, 2015  
ISBN-13: 978-1617290343 ISBN-10: 1617290343

Data Smart: Using Data Science to Transform Information into Insight  
John W. Foreman  
Wiley; 1 edition, 2013  
ISBN-13: 978-1118661468 ISBN-10: 111866146X

Analítica predictiva: Predecir El futuro utilizando big data  
Eric Siegel, Vicente Gonzalez Leon  
Anaya Multimedia, 2013  
ISBN-13: 978-8441534421, 844153442X

## **Complementaria**

Big Data Revolution: What farmers, doctors and insurance agents teach us about discovering big data patterns

Rob Thomas, Patrick McSharry

Wiley; 1 edition, 2015

ISBN-13: 978-1118943717 ISBN-10: 1118943716

Una introducción al Big Data en Cirugía

del Pozo Gil de Pareja Pablo, Campillo Soto Álvaro, López Oliva Ana Isabel

Publicia, 2015

ISBN-13: 978-3639550382, 3639550382

Modelado Predictivo para la Inteligencia de Negocios

Libros Científicos

CreateSpace Independent Publishing Platform, 2015

ISBN-13: 978-1514797907, 1514797909

Data Science for Business: What You Need to Know about Data Mining and Data-Analytic Thinking

Foster Provost, Tom Fawcett

O'Reilly Media; 1 edition (August 19, 2013)

ISBN-13: 978-1449361327, 1449361323

Data Driven: How Performance Analytics Delivers Extraordinary Sales Results

Jenny Dearborn

Wiley; 1 edition, 2015

ISBN-13: 978-1119043126 ISBN-10: 1119043123

Learning Spark: Lightning-Fast Big Data Analysis

Holden Karau, Andy Konwinski, Patrick Wendell, Matei Zaharia

O'Reilly Media; 1 edition, 2015

ISBN-13: 978-1449358624 ISBN-10: 1449358624

Nombre y firma de quién diseñó carta descriptiva:

1. **José Manuel Valencia Moreno**<sup>1</sup>
2. **César Carlo López Portillo**
3. **Omar Álvarez Xochihua**<sup>1</sup>
4. **José Ángel González Fraga**<sup>1</sup>
5. **Nicolasa Valenciana Moreno**<sup>2</sup>

<sup>1</sup>Cuerpo Académico de Tecnologías de Información y Visualización

<sup>2</sup>Cuerpo Académico de Investigación en la Adopción de las TIC'S y la Competitividad de las Mipymes

Nombre y firma de quién autorizó carta descriptiva:

- **Dr. Sergio Octavio Vázquez Núñez**  
Director de la Facultad de Contaduría y Administración – Tijuana
- **Dr. Raúl González Núñez**  
Director de la Facultad de Ciencias Administrativas – Mexicali
- **Dra. Mónica Lacavex Berumen**  
Director de la Facultad de Ciencias Administrativas y Sociales – Ensenada

Nombre(s) y firma(s) de quién(es) evaluó/revisó(evaluaron/ revisaron) la carta descriptiva:

- **Dra. Margarita Ramírez Ramírez**  
Coordinadora de Posgrado de la Facultad de Contaduría y Administración-Tijuana
- **Dr. Manuel Alejandro Ibarra Cisneros**  
Coordinador de Posgrado de la Facultad de Ciencias Administrativas-Mexicali
- **Dr. Ariel Moctezuma Hernández**  
Coordinador de Posgrado de la Facultad de Ciencias Administrativas y Sociales-Ensenada