

<p style="text-align: center;">UNIVERSIDAD AUTÓNOMA DE CHIHUAHUA</p>  <p style="text-align: center;">UNIVERSIDAD AUTÓNOMA DE CHIHUAHUA</p> <p style="text-align: center;">UNIDAD ACADÉMICA INGENIERÍA</p> <p style="text-align: center;">PROGRAMA DE LA UNIDAD DE APRENDIZAJE:</p> <p style="text-align: center;">Big Data Analysis</p>	DES:	INGENIERÍA
	Programa académico	Maestría en Ingeniería en Computación
	Tipo de materia (Obli/Opta):	Optativa
	Clave de la materia:	MICOP217
	Semestre:	2, 3, 4
	Área en plan de estudios (G, E):	G,E
	Total de horas por semana:	4
	<i>Teoría: Presencial o Virtual</i>	2
	<i>Laboratorio o Taller:</i>	0
	<i>Prácticas:</i>	2
	<i>Trabajo extra-clase:</i>	2
	Créditos Totales:	6
	Total de horas semestre (x 16 sem):	64
	<i>Fecha de actualización:</i>	16 mayo 2019
	<i>Prerrequisito (s):</i>	Ninguno
<i>Realizado por:</i>	M.I. Ricardo Carlos	

DESCRIPCIÓN:

El curso introduce al estudiante a la adquisición, transformación, almacenamiento, y procesamiento de grandes volúmenes de datos. Así mismo, le proporciona técnicas estadísticas y computacionales para la visualización, exploración, y análisis de estos.

COMPETENCIAS A DESARROLLAR:

Competencias Genéricas:

Gestión del Conocimiento

Competencias Específicas:

Aplicación de Ciencias de la Computación

DOMINIOS (Se toman de las competencias)	OBJETOS DE ESTUDIO (Contenidos necesarios para desarrollar cada uno de los dominios)	RESULTADOS DE APRENDIZAJE (Se plantean de los dominios y contenidos)	METODOLOGÍA (Estrategias, secuencias, recursos didácticos)	EVIDENCIAS (Productos tangibles que permiten valorar los resultados de aprendizaje)
Soluciona problemas en diversas áreas del conocimiento aplicando las ciencias	<p>I Introducción</p> <p>Definición de big data (BD).</p> <p>Contextualización con áreas de estudio relacionadas: aprendizaje computacional, ciencia de</p>	<p>Define elementos básicos del área.</p> <p>Conoce la problemática de BD.</p>	<p>Exposición del docente.</p> <p>Lecturas individuales.</p> <p>Discusión grupal.</p>	<p>Notas de clase.</p> <p>Examen escrito.</p>

<p>computacionales.</p> <p>Construye prototipos innovadores utilizando herramientas de software y hardware adecuadas a la solución de problemas en cualquier ámbito de desempeño.</p>	<p>datos, sistemas distribuidos.</p> <p>Características de BD.</p> <p>Aplicaciones y retos actuales de BD.</p>	<p>Identifica posibles aplicaciones de técnicas de BD.</p>		
<p>Comprueba los resultados obtenidos del prototipo contra las investigaciones recientes a fin de identificar nuevas contribuciones.</p>	<p>II Adquisición, almacenamiento y procesamiento de datos</p> <p>Tipos de datos y sus problemáticas: estructurados, semi-estructurados, no-estructurados.</p> <p>Data scraping y data wrangling para documentos de texto, sitios y aplicaciones web, imágenes, señales.</p> <p>Tipos de almacenamiento: relacional, llave-valor, jerárquico, de grafos, orientado a documentos.</p> <p>Procesamiento por lotes.</p> <p>Procesamiento en paralelo.</p> <p>Almacenamiento y procesamiento distribuido.</p> <p>Modelo MapReduce.</p> <p>Proyecto de unidad.</p>	<p>Realiza adquisición de datos.</p> <p>Extrae y almacena datos de diversos tipos.</p> <p>Utiliza diferentes tecnologías para almacenar y procesar datos.</p>	<p>Exposición del docente.</p> <p>Lecturas individuales.</p> <p>Discusión grupal.</p> <p>Realización de prácticas individuales.</p>	<p>Código fuente.</p> <p>Reportes escritos.</p> <p>Exposiciones.</p>
	<p>III Exploración de datos</p> <p>Tipos de muestreo y sus implicaciones.</p> <p>Estadísticos descriptivos y gráficos: histogramas, diagramas de dispersión, diagrama de caja.</p> <p>Pruebas de normalidad.</p> <p>Correlación y covarianza.</p> <p>Normalización de variables.</p> <p>Criterios de selección de características: calidad de</p>	<p>Aplica técnicas estadísticas para la exploración y limpieza de datos.</p> <p>Aplica técnicas para reducción de dimensionalidad.</p>	<p>Exposición del docente.</p> <p>Lecturas individuales.</p> <p>Discusión grupal.</p> <p>Realización de prácticas individuales.</p>	<p>Código fuente.</p> <p>Reportes escritos.</p> <p>Exposiciones.</p>

	<p>datos, varianza, correlación, entropía, impureza de Gini, aglomeración, otras; selección hacia adelante y hacia atrás.</p> <p>Reducción de dimensionalidad: SVD, ICA, FA, ISOMAP, otros.</p> <p>Visualización de datos e interpretación de relación entre variables.</p> <p>Proyecto de unidad.</p>	<p>Aplica técnicas de análisis multivariante y de agrupación.</p>		
	<p>IV Análisis de asociación y detección de anomalías</p> <p>Análisis de correspondencias.</p> <p>Reglas de asociación.</p> <p>Métricas: soporte, confianza, interés, coeficiente phi.</p> <p>Algoritmos APRIORI, ECLAT, FP-growth.</p> <p>Algoritmos de agrupación: c-means, GMM, HCA.</p> <p>Detección de anomalías: por proximidad, por densidad, por ordenamiento, por técnicas estadísticas. Local outlier factor, isolation forest, redes bayesianas, elliptic envelope, otros.</p> <p>Proyecto de unidad.</p>	<p>Analiza conjuntos de datos para encontrar relaciones de interés entre entidades y variables.</p> <p>Propone y evalúa modelos para la predicción de eventos y resultados futuros basándose en el análisis de datos.</p>	<p>Exposición del docente.</p> <p>Lecturas individuales.</p> <p>Discusión grupal.</p> <p>Realización de prácticas individuales.</p>	<p>Código fuente.</p> <p>Reportes escritos.</p> <p>Exposiciones.</p>

FUENTES DE INFORMACIÓN (Bibliografía, direcciones electrónicas)	EVALUACIÓN DE LOS APRENDIZAJES (Criterios, ponderación e instrumentos)
<p>Buyya, R., Calheiros, R. N., & Dastjerdi, A. V. (Eds.). (2016). Big Data: Principles and paradigms. MK - Morgan Kaufmann.</p>	

Mistrík, I., Bahsoon, R., Ali, N., Heisel, M., & Maxim, B. (2017). *Software Architecture for Big Data and the Cloud*. Elsevier Science.

Pérez, M. (2015). *Big data: Técnicas, herramientas y aplicaciones*. Alfaomega.

Dietrich, D., Heller, B., & Yang, B. (2015). *Data science & big data analytics: Discovering, analyzing, visualizing and presenting data*. Wiley.

Kotu, V., & Deshpande, B. (2019). *Data science: Concepts and practice*. Morgan Kaufmann.

Husson, F., Lê, S., & Pagès, J. (2017). *Exploratory multivariate analysis by example using R*. CRC Press Taylor & Francis Group.

Nisbet, R., Miner, G., Yale, K., Elder, J. F., & Peterson, A. F. (2018). *Handbook of statistical analysis and data mining applications (2nd ed.)*. Academic Press.

Mehrotra, K. G., Mohan, C. K., & Huang, H. (2018). *Anomaly detection principles and algorithms*. Springer.

Han, J., Kamber, M., & Pei, J. (2012). *Data mining: Concepts and techniques (3rd ed.)*. Elsevier/Morgan Kaufmann.

CRONOGRAMA

Objetos de estudio	Semanas															
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
Introducción																
Adquisición, almacenamiento y procesamiento de datos																
Exploración de datos																
Análisis de asociación y detección de anomalías																