

INSTITUTO POLITÉCNICO NACIONAL
SECRETARÍA ACADÉMICA
DIRECCIÓN DE EDUCACIÓN SUPERIOR



PROGRAMA SINTÉTICO

UNIDAD ACADÉMICA: UNIDAD PROFESIONAL INTERDISCIPLINARIA EN INGENIERÍA CAMPUS COAHUILA; (UPIIC), ESCUELA SUPERIOR DE COMPUTO (ESCOM), UNIDAD PROFESIONAL INTERDISCIPLINARIA EN INGENIERÍA CAMPUS TLAXCALA (UPIIT)	
PROGRAMA ACADÉMICO: Ingeniería en Inteligencia Artificial	
UNIDAD DE APRENDIZAJE: Big Data	SEMESTRE: VI, VII

PROPÓSITO DE LA UNIDAD DE APRENDIZAJE:				
Manipula bases de datos de gran tamaño a partir de algoritmos de minado de datos y herramientas de procesamiento.				
CONTENIDOS:	I. Modelo de Big Data II. Preprocesamiento de Big Data III. Analítica de Big Data IV. Flujo continuo (streaming) de Big Data V. Visualización de Big Data			
ORIENTACIÓN DIDÁCTICA:	Métodos de enseñanza		Estrategias de aprendizaje	
	a) Inductivo	X	a) Estudio de casos	
	b) Deductivo	X	b) Aprendizaje basado en problemas	
	c) Analógico	X	c) Aprendizaje orientado proyectos	X
	d) Heurístico	X	d)	
EVALUACIÓN Y ACREDITACIÓN:	Diagnóstica	X	Saberes Previamente Adquiridos	X
	Solución de casos	X	Organizadores gráficos	X
	Problemas resueltos		Problemarios	
	Reporte de proyectos	X	Exposiciones	X
	Reportes de indagación		Otras evidencias a evaluar: Presentación electrónica y reporte de exposición.	
	Reportes de prácticas	X		
	Evaluaciones escritas			
BIBLIOGRAFÍA BÁSICA:	Autor(es)	Año	Título del documento	Editorial / ISBN
	*Alla, S.	2016	<i>Big Data Analytics with Hadoop 3.</i>	Packt Publisher Inc. / 9781788628846
	*Azarmi, B.	2016	<i>Scalable Big Data Architecture: A Practitioner's Guide to Choosing Relevant Big Data Architecture.</i>	Apress / 9781484213261
	Dasgupta, N.	2018	<i>Practical Big Data Analytics.</i>	Packt Publisher Inc. / 9781783554393
	Luengo, J., García, D., Ramírez, S., García, S. & Herrera, F.	2020	<i>Big Data Preprocessing: Enabling Smart Data.</i>	Springer / 9783030391041
	Mehta, R.	2017	<i>Big Data Analytics with Java.</i>	

clásica



INSTITUTO POLITÉCNICO NACIONAL
SECRETARÍA ACADÉMICA
DIRECCIÓN DE EDUCACIÓN SUPERIOR






PROGRAMA DE ESTUDIOS

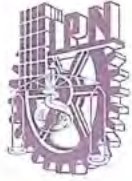
UNIDAD DE APRENDIZAJE: Big Data

HOJA 2 DE 7

UNIDAD ACADÉMICA: UNIDAD PROFESIONAL INTERDISCIPLINARIA EN INGENIERÍA CAMPUS COAHUILA; (UPIIC), ESCUELA SUPERIOR DE COMPUTO (ESCOM), UNIDAD PROFESIONAL INTERDISCIPLINARIA EN INGENIERÍA CAMPUS TLAXCALA (UPIIT)		
PROGRAMA ACADÉMICO: Ingeniería en Inteligencia Artificial		
SEMESTRE: VI, VII	ÁREA DE FORMACIÓN: Formación Profesional	MODALIDAD: Escolarizada
TIPO DE UNIDAD DE APRENDIZAJE: Teórica-práctica/Optativa		
VIGENTE A PARTIR DE: Agosto 2022	CRÉDITOS	
	TEPIC: 7.5	SATCA: 6.3
INTENCIÓN EDUCATIVA		
La unidad contribuye al perfil de egreso de la Ingeniería en Inteligencia Artificial con las habilidades de análisis, diseño, implementación, validación e implantación de sistemas inteligentes para manipular bases de datos de gran tamaño a partir de algoritmos de minado de datos y herramientas de procesamiento. Así mismo fomenta el trabajo en equipo, pensamiento lateral y el pensamiento crítico; asumiendo una actitud de responsabilidad y ética en su desempeño profesional y personal.		
Esta unidad se relaciona de manera antecedente con Minería de datos y de manera lateral con Aplicaciones de sistemas multiagentes.		
PROPÓSITO DE LA UNIDAD DE APRENDIZAJE		
Manipula bases de datos de gran tamaño a partir de algoritmos de minado de datos y herramientas de procesamiento.		

TIEMPOS ASIGNADOS HORAS TEORÍA/SEMANA: 3.0 HORAS PRÁCTICA/SEMANA: 1.5 HORAS TEORÍA/SEMESTRE: 54.0 HORAS PRÁCTICA/SEMESTRE: 27.0 HORAS APRENDIZAJE	UNIDAD DE APRENDIZAJE DISEÑADA POR: Comisión de Diseño del Programa Académico.	AUTORIZADO Y VALIDADO POR: 
	APROBADO POR: Comisión de Programas Académicos del H. Consejo General Consultivo del IPN.	
	21/06/2022	

DIRECCIÓN DE EDUCACIÓN SUPERIOR



INSTITUTO POLITÉCNICO NACIONAL
SECRETARÍA ACADÉMICA
DIRECCIÓN DE EDUCACIÓN SUPERIOR

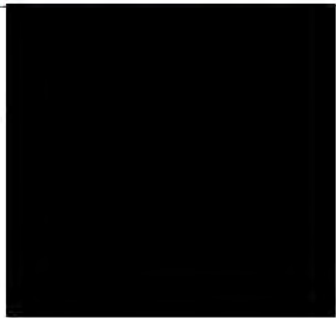


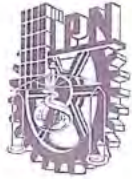
UNIDAD DE APRENDIZAJE: Big Data

HOJA 3 DE 7

UNIDAD TEMÁTICA I Modelo de Big Data	CONTENIDO	HORAS CON DOCENTE		HRS AA
		T	P	
UNIDAD DE COMPETENCIA Describe el procesamiento de Big Data con base en sus herramientas, problemática, las fases del procesamiento y modelo.	1.1 El modelo de Big Data 1.1.1 De las 3 Vs a más de 6 Vs 1.1.2 El porqué de grandes bases de datos 1.1.3 Más de 4 Vs	1.5		1.0
	1.2 Fases del Procesamiento de Big Data 1.2.1 Preprocesamiento 1.2.2 Analítica 1.2.3 Flujo continuo (streaming) 1.2.4 Visualización	1.5		1.0
	1.3 Tipos y fuentes de Big Data 1.3.1 Tipos 1.3.2 Fuentes 1.3.3 Conjuntos	1.5	1.5	1.0
	1.4 Administración e indexación de memoria 1.4.1 Indexación 1.4.2 Modelo vectorial	1.5	1.5	1.0
	1.5 Herramientas para Big Data 1.5.1 Ecosistema Hadoop 1.5.2 Ecosistema Spark 1.5.3 Ecosistema en la nube EC2	3.0	1.5	2.0
	Subtotal	9.0	4.5	6.0

UNIDAD TEMÁTICA II Preprocesamiento en Big data	CONTENIDO	HORAS CON DOCENTE		HRS AA
		T	P	
UNIDAD DE COMPETENCIA Emplea el paradigma de reducción en la manipulación de datos a partir de la dimensionalidad y el tamaño de los datos.	2.1 Manipulación de datos	1.5		
	2.2 Reducción de la dimensionalidad	3.0	1.5	1.0
	2.3 Reducción de datos	3.0	1.5	1.0
	2.4 Paradigma de reducción 2.4.1 El marco MapReduce	3.0	1.5	1.0
	Subtotal	10.5	4.5	3.0





INSTITUTO POLITÉCNICO NACIONAL
SECRETARÍA ACADÉMICA
DIRECCIÓN DE EDUCACIÓN SUPERIOR



UNIDAD DE APRENDIZAJE: Big Data

HOJA 4 DE 7

UNIDAD TEMÁTICA III Analítica de Big Data	CONTENIDO	HORAS CON DOCENTE		HRS AA
		T	P	
UNIDAD DE COMPETENCIA Aplica algoritmos de minado de datos a grandes tablas de acuerdo con la estructura y las herramientas.	3.1 Tablas NoSQL 3.1.1 Estructura 3.1.2 Operaciones	3.0	1.5	1.0
	3.2 Algoritmos de minado de grandes tablas 3.2.1 Page rank 3.2.2 Similaridades 3.2.3 Agrupación y clasificación	4.5	3.0	2.5
	3.3 Herramientas de minado en Hadoop 3.3.1 Mahout 3.3.2 MLib 3.3.3 Minehashing	4.5	1.5	2.5
	Subtotal	12.0	6.0	6.0

UNIDAD TEMÁTICA IV Flujo continuo (streaming) de Big Data	CONTENIDO	HORAS CON DOCENTE		HRS AA
		T	P	
UNIDAD DE COMPETENCIA Demuestra el minado de datos de flujos continuos a partir de su estructura y las herramientas.	4.1 Redes sociales 4.1.1 Estructura de las redes sociales	3.0	1.5	1.0
	4.2 Minado de datos en flujos continuos	6.0	3.0	2.0
	4.3 Herramientas 4.3.1 Flink 4.3.2 FlinkML 4.3.3 BigDaPFlink	6.0	1.5	2.0
	Subtotal	15.0	6.0	5.0

UNIDAD TEMÁTICA V Visualización de Big Data	CONTENIDO	HORAS CON DOCENTE		HRS AA
		T	P	
UNIDAD DE COMPETENCIA Construye tableros de control a partir de las herramientas de visualización.	5.1 Visualización de datos 5.1.1 Objetivo de la visualización 5.1.2 Tableros de control	1.5		
	5.2 Visualización usando lenguaje R	1.5	1.5	1.0
	5.3 Visualización usando Google Charts	1.5	1.5	1.0
	5.4 Visualización usando Python	1.5		
	5.5 Visualización usando Tableau	1.5		
Subtotal		7.5		



INSTITUTO POLITÉCNICO NACIONAL
SECRETARÍA ACADÉMICA
DIRECCIÓN DE EDUCACIÓN SUPERIOR

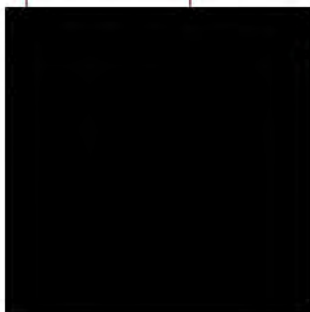


UNIDAD DE APRENDIZAJE: Big Data

HOJA 5 DE 7

ESTRATEGIAS DE APRENDIZAJE	EVALUACIÓN DE LOS APRENDIZAJES
<p>Estrategia de Aprendizaje Orientado a Proyectos</p> <p>El estudiante desarrollará las siguientes actividades:</p> <ol style="list-style-type: none"> 1. Indagación documental de diferentes temas del programa con lo que elaborarán un mapa conceptual o mental. 2. Análisis de casos específicos de los temas vistos. 3. Desarrollo de proyectos donde se compruebe el aprendizaje de la unidad <ol style="list-style-type: none"> I) Aplicación del Preprocesamiento en Big data II) Minado a grandes Tablas III) Minado de flujos continuos IV) Tableros para la visualización de datos 4. Exposición de casos prácticos. 5. Exposición de proyecto. 6. Realización de prácticas. 	<p>Evaluación diagnóstica.</p> <p>Portafolio de evidencias:</p> <ol style="list-style-type: none"> 1. Mapa mental o conceptual. 2. Solución de casos. 3. Reportes del proyecto. Entrega final del proyecto. 4. Presentación. 5. Reporte de exposición del proyecto. 6. Reporte de prácticas.

RELACIÓN DE PRÁCTICAS			
PRÁCTICA No.	NOMBRE DE LA PRÁCTICA	UNIDADES TEMÁTICAS	LUGAR DE REALIZACIÓN
1	Instalación del ecosistema Hadoop	I	Laboratorio de Cómputo
2	MapReduce y su aplicación	II	
3	Operaciones con tablas NoSQL	III	
4	Minado en big data: encontrando similitudes	III	
5	Minado en big data: clasificación o agrupación	III	
6	Minado en redes sociales	IV	
7	R para visualización de datos	V	
8	Google Charts en la visualización de datos	V	
9	Python para visualizar datos	V	
10	El Desarrollo de tableros	V	
		TOTAL DE HORAS:	27.0





INSTITUTO POLITÉCNICO NACIONAL
SECRETARÍA ACADÉMICA
DIRECCIÓN DE EDUCACIÓN SUPERIOR



UNIDAD DE APRENDIZAJE: Big Data

HOJA 6 DE 7

Bibliografía							
Tipo	Autor(es)	Año	Título del documento	Editorial/ISBN	Documento		
					Libro	Antología	Otros
B	*Alla, S.	2016	<i>Big Data Analytics with Hadoop 3.</i>	Packt Publisher Inc. / 9781788628846	X		
B	*Azarmi, B.	2016	<i>Scalable Big Data Architecture: A Practitioner's Guide to Choosing Relevant Big Data Architecture.</i>	Apress / 9781484213261	X		
C	Caballero, R., Martín, E. & Riesco, A.	2018	<i>BIG DATA con PYTHON. Recolección, almacenamiento y proceso.</i>	RC Libros / 9788494897207	X		
C	Damji, S., Wenig, B., Das, T. & Lee, D.	2020	<i>Learning Spark.</i>	O'Reilly Media, Inc. / 9781492050049	X		
B	Dasgupta, N.	2018	<i>Practical Big Data Analytics.</i>	Packt Publisher Inc. / 9781783554393	X		
C	Deitel, J. & Deitel, M.	2019	<i>Intro to Python for Computer Science and Data Science: Learning to Program with AI, Big Data and The Cloud.</i>	Pearson / 9780135404676	X		
C	Elahi, I.	2019	<i>Scale Programming for Big Data Analytics : Get Started With Big Data Analytics Using Apache Spark</i>	Apress / 9781484248102	X		
C	Lakshmanan, V. & Tigani, J.	2019	<i>Google BigQuery: The Definitive Guide.</i>	O'Reilly Media Inc. / 9781492044468	X		
B	Luengo, J., García, D., Ramírez, S., García, S. & Herrera, F.	2020	<i>Big Data Preprocessing: Enabling Smart Data.</i>	Springer / 9783030391041	X		
B	Mehta, R.	2017	<i>Big Data Analytics with Java.</i>	Packt Publisher Inc. / 9781787288980	X		
C	Ryzko, D.	2020	<i>Modern Big Data Architectures: a multi-Agent Systems Perspective.</i>	Wiley & sons / 9781119597933	X		
C	Weissman, B. & Van de Laar, E.	2020	<i>SQL Server Big Data Clusters: Data Virtualization, Data Lake, and AI Platform.</i>	Apress / 9781484259849	X		

Recursos digitales

Autor, año, título y Dirección Electrónica		Texto	Simulador	Animación	Video	Otros
Amazon. AWS (s/a). Elastic compute cloud (EC2) de capacidad modificable en la nube. https://aws.amazon.com/es/ec2/						
Anaconda. https://www.anaconda.com/products/individual .						
Foundation. (2006). Apache Hadoop. https://hadoop.apache.org/						
. (s/a). Google Charts. https://developers.google.com/chart						
Foundation. (2001). Welcome to python.Org. https://www.python.org						
(s/a). R: The R Project for Statistical Computing. https://www.r-						X
udio. https://rstudio.com						X



INSTITUTO POLITÉCNICO NACIONAL
SECRETARÍA ACADÉMICA
DIRECCIÓN DE EDUCACIÓN SUPERIOR



UNIDAD DE APRENDIZAJE: Big Data

HOJA 7 DE 7

PERFIL DOCENTE: Ingeniería en Computación o áreas afines preferentemente con grado de Maestría y/o Doctorado en áreas afines a la Inteligencia Artificial.

EXPERIENCIA PROFESIONAL	CONOCIMIENTOS	HABILIDADES DIDÁCTICAS	ACTITUDES
Preferentemente dos años en el sector productivo. Al menos dos años de docencia a Nivel Superior y/o posgrado.	En Inteligencia Artificial, Big Data, minería de datos, Hadoop, Spark. Del Modelo Educativo Institucional (MEI).	Comunicación efectiva Capacidad de transmitir conocimientos Capacidad de organización y planificación Liderazgo Capacidad para el manejo de grupos	Ética Respeto Responsabilidad Honestidad Empatía Tolerancia Compromiso social e institucional Disponibilidad para trabajar en equipo

ELABORÓ

REVISÓ

AUTORIZÓ

Dr. Eric Manuel Rosales Peña Alfaro
Coordinador

Mr. en C Ivan Giovanny Irujo
García
Subdirector Académico

Ing. Enrique Lima Morales
Subdirector Académico UP