



INSTITUTO POLITÉCNICO NACIONAL
SECRETARÍA ACADÉMICA
DIRECCIÓN DE EDUCACIÓN SUPERIOR



PROGRAMA SINTÉTICO

UNIDAD ACADÉMICA: ESCUELA SUPERIOR DE CÓMPUTO (ESCOM), UNIDAD PROFESIONAL INTERDISCIPLINARIA DE INGENIERÍA, CAMPUS ZACATECAS (UPIIZ)	
PROGRAMA ACADÉMICO: Ingeniería en Sistemas Computacionales	
UNIDAD DE APRENDIZAJE: Data mining	SEMESTRE: VII PLAN DE ESTUDIOS: 2020

PROPÓSITO DE LA UNIDAD DE APRENDIZAJE			
Desarrolla proyectos estructurados a partir de las técnicas de minería de datos CRISP-DM o SEMMA.			
CONTENIDOS:	<ul style="list-style-type: none"> I Gestión de cubos OLAP II Fundamentos y técnicas de minería de datos estadísticas III Fundamentos estadísticos IV Clasificación V Agrupamiento (Clúster) 		
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		c) Aprendizaje Orientado a Proyectos
EVALUACIÓN Y ACREDITACIÓN:	d) Analítico	X	d)
	Diagnóstica	X	Saberes Previamente Adquiridos
	Solución de casos	X	Organizadores gráficos
	Problemas resueltos	X	Problemarios
	Reporte de proyectos	X	Exposiciones
	Reportes de indagación		Otras evidencias a evaluar:
	Reportes de prácticas	X	
Evaluación escrita			
BIBLIOGRAFÍA BÁSICA:	Autor(es)	Año	Título del documento
	*Hernández, J., Ramírez, M., F, M. & Ferri, C.	2004	Introducción a la Minería de datos
	Larose, D. y Larose, C.	2015	Data Mining and Predictive Analytics
	Pang-Ning, T., Steinbach, M., y Kumar, V	2005	Introduction to data mining
	Ratner, B.	2011	Statistical and Machine-Learning Data Mining Techniques for Better Predictive Modeling and Analysis of Big Data
	Rokach, L. & Maimon, O.	2015	Data Mining with decision trees. Theory and Applications
			Editorial / ISBN
			Pearson/ 84-205-4091-9
			Wiley/ 978-1-118-11619-7
			Pearson/ 978-1-292-02615-2
			CRC Press Taylor & Francis Group/ 978-1-4398-6092-2
			World Scientific Publishing Co. Pte. Ltd/ 978-9814590075



INSTITUTO POLITÉCNICO NACIONAL
SECRETARÍA ACADÉMICA
DIRECCIÓN DE EDUCACIÓN SUPERIOR



PROGRAMA DE ESTUDIOS

UNIDAD DE APRENDIZAJE: Data mining

HOJA 2 DE 9

UNIDAD ACADÉMICA: ESCUELA SUPERIOR DE CÓMPUTO (ESCOM), UNIDAD PROFESIONAL INTERDISCIPLINARIA DE INGENIERÍA, CAMPUS ZACATECAS (UPIIZ)		
PROGRAMA ACADÉMICO: Ingeniería en Sistemas Computacionales		
SEMESTRE: VII PLAN DE ESTUDIOS: 2020	ÁREA DE FORMACIÓN: Profesional	MODALIDAD: Escolarizada
TIPO DE UNIDAD DE APRENDIZAJE: Teórica- práctica/ Optativa		
VIGENTE A PARTIR DE: Enero 2023	CRÉDITOS: TEPIC: 7.5 SATCA: 6.3	
INTENCIÓN EDUCATIVA La unidad de aprendizaje contribuye al perfil de egreso de la Ingeniería en Sistemas Computacionales con el desarrollo de habilidades que le permitan analizar datos descriptivos y de representación visual, aplicar diversas técnicas de tratamiento de datos y el reconocimiento de la naturaleza de las técnicas de minería de datos a aplicar en la solución de problemas. Asimismo, desarrolla habilidades transversales como pensamiento crítico, trabajo en equipo, comunicación efectiva, ética, resolución de problemas, creatividad e ingenio. Esta unidad de aprendizaje se relaciona de manera antecedente con Statistical tools for data analytics y Bases de datos; de manera lateral con Trabajo terminal I; y de forma consecuente con Trabajo terminal II.		
PROPÓSITO DE LA UNIDAD DE APRENDIZAJE Desarrolla proyectos estructurados a partir de las técnicas de minería de datos CRISP-DM o SEMMA.		

TIEMPOS ASIGNADOS HORAS TEORÍA/SEMANA: 3.0 HORAS PRÁCTICA/SEMANA: 1.5 HORAS TEORÍA/SEMESTRE: 54.0 HORAS PRÁCTICA/SEMESTRE: 27.0 HORAS APRENDIZAJE AUTÓNOMO: 24.0 HORAS TOTALES/SEMESTRE: 81.0	UNIDAD DE APRENDIZAJE REDISEÑADA POR: Academia de Ingeniería de Software REVISADA POR: <hr/> M. en C. Iván Giovanni Mosso García Subdirector Académico ESCOM <hr/> M. en C. Roberto Oswaldo Cruz Lejía Subdirector Académico UPIIZ APROBADA POR: Consejo Técnico Consultivo Escolar <hr/> M. en C. Andrés Ortigoza Campos Presidente ESCOM 22/11/2022 <hr/> Dr. Fernando Flores Mejía Presidente del CTCE de UPIIZ 27/06/2022	APROBADO POR: Comisión de Programas Académicos del Consejo General Consultivo del IPN. <p align="center">24/11/2022</p>
	AUTORIZADO Y VALIDADO POR: <hr/> Mtro. Mauricio Igor Jasso Zaranda Secretario Académico	



INSTITUTO POLITÉCNICO NACIONAL
SECRETARÍA ACADÉMICA
DIRECCIÓN DE EDUCACIÓN SUPERIOR



UNIDAD DE APRENDIZAJE: Data mining

HOJA 3 DE 9

UNIDAD TEMÁTICA I Gestión de cubos OLAP	CONTENIDO	HORAS CON DOCENTE		HRS AA
		T	P	
UNIDAD DE COMPETENCIA Resuelve problemas de analítica de datos con base en el análisis del almacenamiento de datos.	1.1 Decision Support System (DSS)	1.5	1.5	1.0
	1.2 Modelo multidimensional	1.5		
	1.3 Administración de cubos	1.5		
	1.3.1 Agregado de funcionalidades			
	1.3.2 Importación de datos y uso de conectores			
	1.3.3 Configuraciones adicionales			
	1.4 Operación de cubos	3.0		1.0
1.4.1 Búsqueda de Información				
1.4.2 Administración de incidentes y problemas				
1.4.3 Informes y análisis de almacenamiento de datos				
1.5 Uso de cubos OLAP para análisis avanzado y solución de problemas	1.5	1.5	1.0	
	Subtotal	9.0	3.0	3.0

UNIDAD TEMÁTICA II Fundamentos y técnicas de minería de datos estadísticas	CONTENIDO	HORAS CON DOCENTE		HRS AA
		T	P	
UNIDAD DE COMPETENCIA Describe las acciones en las etapas del KDD y la minería de datos a partir de la limpieza, discretización y recategorización.	2.1 Fundamentos de minería de datos	1.5	1.5	1.0
	2.1.1 Tipos de datos			
	2.1.2 Clasificación de las técnicas de minería de datos			
	2.1.3 Tipos de modelos o tareas de la minería de datos: clasificación, agrupamiento, regresión, reglas de asociación, series de tiempo y predicción			
	2.2 Knowledge Discovery in databases (KDD)	1.5		1.0
	2.3 Métodos de minería de datos: CRISP-DM (Cross Industry Standard Process), SEMMA (Muestrear, explorar, modificar, modelar y evaluar)	4.5		
	2.4 Limpieza y tratamiento de datos	3.0	3.0	1.5
	2.4.1 Técnicas de imputación			
	2.4.2 Técnicas de normalización			
	2.4.3 Técnicas de transformación y contenedores (binners)			
2.4.4 Discretización y recategorización				
2.5 Análisis estadístico descriptivo de datos: tendencia central, dispersión, tablas de frecuencias y gráficos de representación según los tipos de datos			1.0	
	Subtotal	10.5	4.5	4.5



INSTITUTO POLITÉCNICO NACIONAL
SECRETARÍA ACADÉMICA
DIRECCIÓN DE EDUCACIÓN SUPERIOR



UNIDAD DE APRENDIZAJE: Data mining

HOJA 4 DE 9

UNIDAD TEMÁTICA III Fundamentos estadísticos	CONTENIDO	HORAS CON DOCENTE		HRS AA
		T	P	
UNIDAD DE COMPETENCIA Soluciona problemas de predicción y reducción de dimensionalidad en conjuntos de datos con base en la correlación, estadística inferencial y reducción dimensional.	3.1 Correlación	1.5		1.0
	3.2 Estadística inferencial	3.0	3.0	1.5
	3.2.1 Regresión lineal			
	3.2.2 Regresión múltiple			
	3.2.3 Coeficiente de determinación			
	3.3 Reducción de dimensionalidad	3.0	3.0	1.5
	3.3.1 Análisis de Componentes Principales (ACP)			
	3.3.2 Análisis Factorial Exploratorio y Confirmatorio (AFE y AFC)			
	Subtotal	7.5	6.0	4.0

UNIDAD TEMÁTICA IV Clasificación	CONTENIDO	HORAS CON DOCENTE		HRS AA
		T	P	
UNIDAD DE COMPETENCIA Clasifica conjuntos de datos de acuerdo con técnicas, árboles de decisión, regresión logística y redes neuronales artificiales.	4.1 Significado de la clasificación	3.0		1.5
	4.1.1 El proceso para construir modelos de clasificación			
	4.1.2 La matriz de confusión			
	4.1.3 Métricas de evaluación de la clasificación: exactitud, precisión, tasa de error, sensibilidad (recall), especificidad y comprensibilidad			
	4.1.4 Evaluación del desempeño: método de retención, validación cruzada, submuestreo aleatorio y Bootstrap			
	4.2 Técnicas de clasificación	3.0	1.0	
	4.3 Árboles de decisión	3.0		1.5
	4.3.1 Naturaleza de los árboles de decisión			
	4.3.2 Tipos de árboles: ID3, C4.5, CART y CHAID			
	4.3.3 Criterios de particionamiento: Entropía, índice de ganancia, índice de Gini e impureza			
	4.3.4 Subajuste y sobreajuste			
	4.3.5 Bagging y Boosting			
	4.4 Regresión logística como modelo probabilístico de clasificación	3.0	2.0	2.0
	4.4.1 Naturaleza del clasificador binario			
	4.4.2 Componentes de la función logit			
4.4.3 Función sigmoidea				
4.5 Redes neuronales artificiales	3.0	3.0	1.5	
4.5.1 Red neuronal de una sola capa (ADaptive Llinear NEuron (Adaline))				
	Subtotal	15.0	6.0	6.5



INSTITUTO POLITÉCNICO NACIONAL
SECRETARÍA ACADÉMICA
DIRECCIÓN DE EDUCACIÓN SUPERIOR



UNIDAD DE APRENDIZAJE: Data mining

HOJA 5 DE 9

UNIDAD TEMÁTICA V Agrupamiento (Clúster)	CONTENIDO	HORAS CON DOCENTE		HRS AA
		T	P	
UNIDAD DE COMPETENCIA Crea grupos de ítems característicos a partir de las medidas, similitudes y las métricas de bondad.	5.1 Naturaleza del agrupamiento	4.5		2.0
	5.1.1 Característica y fundamentos teóricos			
	5.1.2 Medidas de similitud y distancia: enlace simple, enlace completo, enlace promedio, centroide y medoide			
	5.1.3 Algoritmos jerárquicos: el dendograma, aglomerativos y divisivo			
	5.1.4 Algoritmos de particionamiento: k-medias y algoritmo del vecino más próximo (KNN)			
	5.1.5 Medida de bondad del clúster: Silueta y estadística Seudo-F			
	5.1.6 Determinación del número de grupos: elbow			
	5.2 Agrupamiento con datos categóricos	3.0	4.5	2.0
	5.3 Naturaleza de las reglas de asociación	4.5	3.0	2.0
	5.3.1 Algoritmo a priori para identificar patrones frecuentes			
	5.3.2 Creación de las reglas de asociación			
	5.3.3 Evaluación de las reglas de asociación			
	5.3.4 Importancia de la evaluación subjetiva de las reglas de asociación			
	5.3.5 Reglas de asociación y el aprendizaje supervisado			
		Subtotal	12.0	7.5



INSTITUTO POLITÉCNICO NACIONAL
SECRETARÍA ACADÉMICA
DIRECCIÓN DE EDUCACIÓN SUPERIOR



UNIDAD DE APRENDIZAJE: Data mining

HOJA: 6 DE 9

ESTRATEGIAS DE APRENDIZAJE	EVALUACIÓN DE LOS APRENDIZAJES
<p>Estrategia de aprendizaje orientado a proyectos</p> <p>El alumno desarrollará las siguientes actividades:</p> <ol style="list-style-type: none">1. Resolución de problemas2. Diseño e implementación de proyecto: Fase 1. Método KDD Fase 2. Métricas de evaluación de resultados Fase 3. Naturaleza de los datos, método KDD y CRISP-DM o SEMMA. Fase 4. Justificación e interpretación de resultados de evaluación3. Análisis de casos de estudio4. Realización de prácticas	<p>Evaluación diagnóstica</p> <p>Portafolio de evidencias:</p> <ol style="list-style-type: none">1. Problemas resueltos2. Reporte de proyecto e implementación3. Solución de casos de datos nativos4. Reporte de prácticas

RELACIÓN DE PRÁCTICAS			
PRÁCTICA No.	NOMBRE DE LA PRÁCTICA	UNIDADES TEMÁTICAS	LUGAR DE REALIZACIÓN
1	Visualización y estructuras multidimensionales	I	Laboratorio de cómputo
2	Etapas del KDD y la relación que guarda con los métodos de minería CRISP-DM y SEMMA	II	
3	Tratamiento de datos	III	
4	Modelos de predicción y reducción de dimensionalidad	IV	
5	Modelos de clasificación mediante árboles, regresión logística y redes neuronales	V	
6	Grupos (clúster) basados en medidas y similitudes	V	
7	Reglas de asociación en la canasta de mercado	V	
		TOTAL DE HORAS:	27.0



INSTITUTO POLITÉCNICO NACIONAL
SECRETARÍA ACADÉMICA
DIRECCIÓN DE EDUCACIÓN SUPERIOR



UNIDAD DE APRENDIZAJE: Data mining

HOJA: 7 DE 9

Bibliografía							
Tipo	Autor(es)	Año	Título del documento	Editorial	Documento		
					Libro	Antología	Otros
C	Anderson, D., Sweeney, D. y Williams, T.	2008	Estadística para administración y economía	Cengage Learning/ 978-607-481-319-7	X		
C	Bennet, J.	2011	Razonamiento estadístico	Pearson/ 978-607-32-0759-1	X		
C	Dunham, M. H.	2002	Data mining: introductory and advanced topics	Prentice Hall/ 0130888923	X		
B*	Hernández, O., Ramírez, J. y Ramírez, F.	2004	Introducción a la Minería de datos	Pearson/ 84-205-4091-9	X		
B	Larose, D. & Larose, C.	2015	Data Mining and Predictive Analytics. Second Edition.	Wiley/ 978-1-118-11619-7	X		
C	Levín, R.	2004	Estadística para administración y economía	Prentice-Hall/ 970-26-0497-4	X		
B*	Pang-Ning, T., Steinbach, M., y Kumar, V.	2005	Introduction to data mining	Pearson/ 978-1-292-02615-2	X		
B	Ratner, B.	2011	Statistical and Machine-Learning Data Mining Techniques for Better Predictive Modeling and Analysis of Big Data	CRC Press Taylor & Francis Group/ 978-1-4398-6092-2	X		
B	Rokach, L. & Maimon, O.	2015	Data Mining with decision trees. Theory and Applications	World Scientific Publishing Co. Pte. Ltd/ 978-9814590075	X		



INSTITUTO POLITÉCNICO NACIONAL
SECRETARÍA ACADÉMICA
DIRECCIÓN DE EDUCACIÓN SUPERIOR



UNIDAD DE APRENDIZAJE: Data mining

HOJA: 8 DE 9

Recursos digitales								
Autor, año, título y Dirección Electrónica	Texto	Simulador	Imagen	Tutorial	Video	Presentación	Diccionario	Otro
Suárez-Ibujes, M. (2008). Conceptos básicos de probabilidades y estadística inferencial. Recuperado el 12 de mayo de 2022, de: https://es.scribd.com/doc/129480693/Conceptos-basicos-de-Probabilidades-y-Estadistica-Inferencial	x							
(s/a). (2009). Análisis de clúster y árboles de clasificación. Recuperado el 12 de mayo de 2022, de: http://halweb.uc3m.es/esp/Personal/personas/jmmarin/esp/DM/tema6dm.pdf	x							
Gutiérrez, A., García, M. & Martínez, J. (2012). Algoritmo de agrupamiento basado en patrones utilizando árboles de decisión no supervisados. Recuperado el 12 de mayo de 2022, de: https://ccc.inaoep.mx/portalfiles/CCC-11-002.pdf	x							
Cabena, P. et al. (1998). Discovering Data Mining: From Concept to Implementation. Recuperado el 12 de mayo de 2022, de: https://www.zanasi-alessandro.eu/publications/zanasi-a-1998-competitive-intelligence-through-data-mining-public-sources/	x							
Bhumika, G. et al. (2017). Analysis of Various Decision Tree Algorithms for Classification in Data Mining. International. Journal of Computer Applications. Recuperado el 18 de octubre de 2022, de: https://www.ijcaonline.org/archives/volume163/number8/27414-2017913660	X							



INSTITUTO POLITÉCNICO NACIONAL
SECRETARÍA ACADÉMICA
DIRECCIÓN DE EDUCACIÓN SUPERIOR



UNIDAD DE APRENDIZAJE: Data mining

HOJA: 9 DE 9

PERFIL DOCENTE: Ingeniería en Computación o áreas afines, preferentemente con grado de Maestría o Doctorado en áreas afines a la Inteligencia Artificial, máquinas de aprendizaje y minería de datos

EXPERIENCIA PROFESIONAL	CONOCIMIENTOS	HABILIDADES DIDÁCTICAS	ACTITUDES
Preferentemente dos años en el sector productivo realizando análisis de datos Mínima de dos años en docencia a Nivel Superior	En máquinas de Aprendizaje Técnicas de minado de datos Modelos de almacenes de datos y procesamiento analítico de datos Del Modelo Educativo Institucional (MEI)	Coordinar grupos de aprendizaje Organizar equipos de aprendizaje Planificación de la enseñanza Manejo de estrategias didácticas centradas en el aprendizaje Manejo de TIC en la enseñanza y para el aprendizaje Comunicación multidireccional Habilidad numérica Toma de decisiones	Compromiso con la enseñanza Congruencia Empatía Honestidad Proactividad Respeto Responsabilidad Tolerancia Liderazgo Ético

ELABORÓ

REVISÓ

AUTORIZÓ

Ing. Tania Rodríguez Sarabia
Coordinadora

M.I.S. Julia Elena Hernández Ríos
Coordinadora

M. en C. Laura Méndez Segundo
Participante

M. en C. Roberto Oswaldo Cruz Lejía
Subdirector Académico UPIIZ

M. en C. Andrés Ortigoza Campos
Director ESCOM

Dra. Fabiola Ocampo Botello
Participante

M. en Ed. Karina Rodríguez Mejía
Participante

M. en C. Iván Giovanni Mosso García
Subdirector Académico ESCOM

Dr. Fernando Flores Mejía
Director UPIIZ