

PLAN DE ESTUDIOS (PE): Licenciatura en Ingeniería en Tecnologías de la Información

AREA: Tecnologías de la Información

ASIGNATURA: Minería de Datos

CÓDIGO: ITIM-261

CRÉDITOS: 5

FECHA: 15 de Julio de 2013



1. DATOS GENERALES

Nivel Educativo:	Licenciatura
Nombre del Plan de Estudios:	Licenciatura en Ingeniería en Tecnologías de la Información
Modalidad Académica:	Presencial
Nombre de la Asignatura:	Minería de Datos
Ubicación:	Nivel Formativo
Correlación:	
Asignaturas Precedentes:	Administración de Bases de Datos
Asignaturas Consecuentes:	Inteligencia de Negocios, Tratamiento de información y Teoría de decisiones
Conocimientos, habilidades, actitudes y valores previos:	<p>Conocimientos:</p> <ul style="list-style-type: none"> • Conocer, diseñar y saber manejar bases de datos. • Conocer y saber manejar lenguajes de consulta de bases de datos. • Conocer y aplicar técnicas estadísticas en el análisis de tendencias de datos. • Conocer y aplicar un lenguaje de programación de propósito general. <p>Habilidades:</p> <ul style="list-style-type: none"> • Reconocer escenarios en donde tiene aplicación la abstracción de información. • Habilidad para el manejo de la información en áreas interdisciplinarias. • Seleccionar las herramientas adecuadas para la solución a problemas relacionados con datos. • Capacidad de expresión oral y escrita. • Capacidad de abstraer conocimiento a partir de lecturas. • Trabajo colaborativo. <p>Actitudes:</p> <ul style="list-style-type: none"> • Ser flexible para la implementación de técnicas y aplicaciones en diferentes plataformas.



	<ul style="list-style-type: none"> • Proponer la pertinencia de la aplicación de soluciones de Middleware a la solución de problemas. • Responsabilidad. <p>Valores:</p> <ul style="list-style-type: none"> • Interés por conocer los principios del manejo de información digitalizada. • Preocuparse por estar actualizado con las tendencias y herramientas. • Deseo por aprender a resolver problemáticas que involucran abstracción de información. • Ética profesional en el desarrollo de sus actividades.
--	--

2. CARGA HORARIA DEL ESTUDIANTE (Ver matriz 1)

Concepto	Horas por periodo		Total de horas por periodo	Número de créditos
	Teoría	Práctica		
Horas teoría y práctica (16 horas = 1 crédito)	48	32	80	5
Total	48	32	80	5

3. REVISIONES Y ACTUALIZACIONES

Autores:	Ambrosio Vázquez Alma Delia Archundia Sierra Etelvina Boone Rojas María del Rocío González Calleros Juan Manuel Guerrero García Josefina Olmos Pineda Iván Pérez de Celis Herrero María Concepción Somodevilla García María Josefa
Fecha de diseño:	15 de Julio de 2013
Fecha de aprobación por parte de la academia de área	15 de Noviembre de 2013
Fecha de aprobación por parte de CDESC-UA	13 de Diciembre de 2013
Fecha de revisión del Secretario Académico	20 de Enero de 2013



4. PERFIL DESEABLE DEL PROFESOR (A) PARA IMPARTIR LA ASIGNATURA:

Disciplina profesional:	Ciencias de la Computación, Tecnologías de la información, Informática o Áreas Afines
Nivel académico:	Maestría
Experiencia docente:	Mínima 2 años
Experiencia profesional:	Mínima 2 años

5. OBJETIVOS:

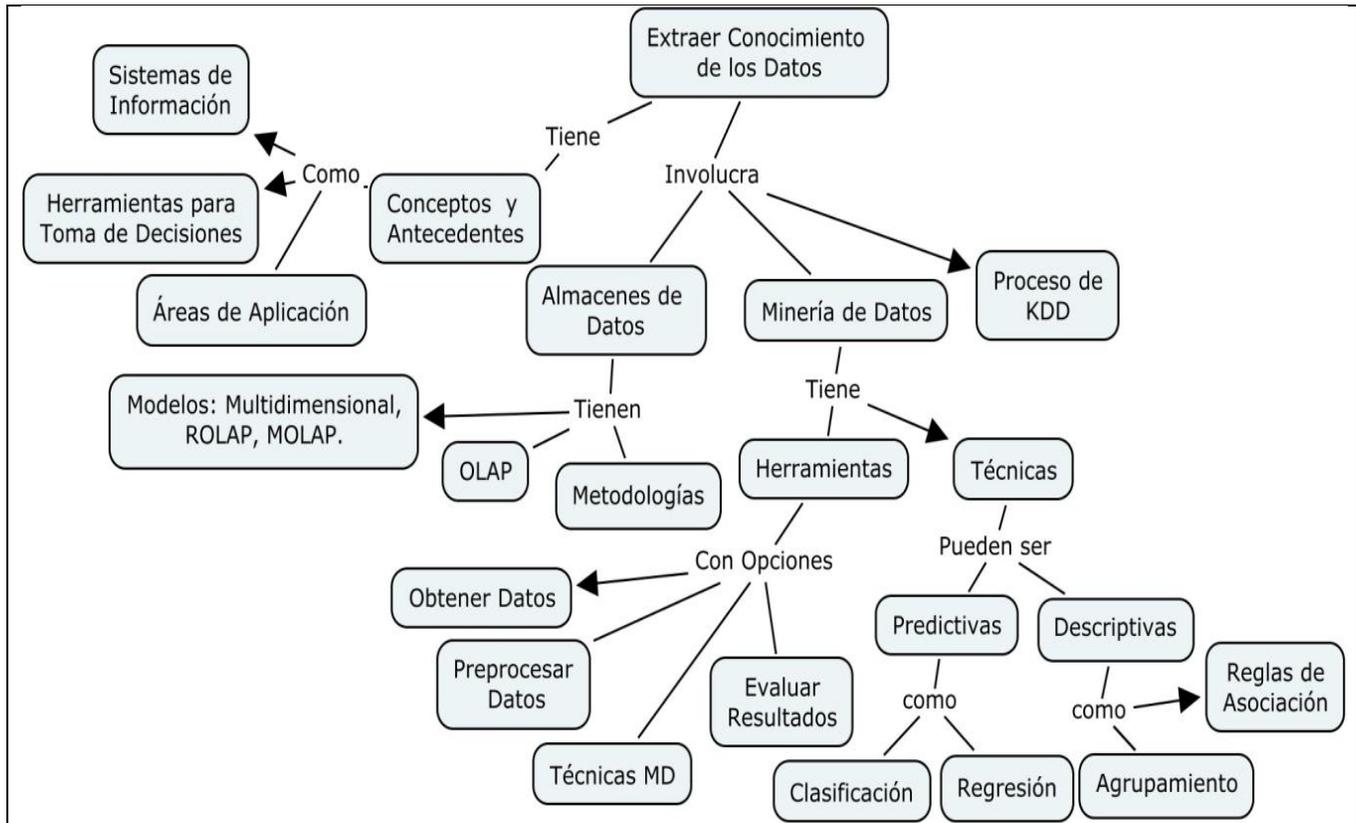
5.1 General: Al finalizar el curso, el alumno conocerá el proceso de descubrir, extraer y almacenar información relevante de amplias bases de datos a través de programas de búsqueda e identificación de patrones.

5.2 Específicos: El alumno será capaz de:

- Conocer las ventajas y casos donde es aconsejable recopilar información interna y externa en un Almacén de Datos, así como distinguir las propiedades y finalidades de las diferentes herramientas de soporte a la toma de decisiones.
- Conocer el modelo multidimensional de los almacenes de datos y los operadores de refinamiento asociados: drill, roll, slice & dice, pivot; así como conocer la arquitectura y diferentes implementaciones de almacenes de datos.
- Reconocer la problemática del análisis de grandes volúmenes de datos y de los beneficios de su uso sistemático para la obtención de modelos y patrones predictivos o descriptivo para elegir, en un problema concreto, qué técnicas de minería de datos son más apropiadas.
- Generar los modelos y patrones elegidos utilizando una herramienta de minería de datos para aplicarlos en la solución de un problema específico.

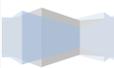


6. REPRESENTACIÓN GRÁFICA DE LA ASIGNATURA:



7. CONTENIDO

Unidad 1	Objetivo Específico	Contenido Temático/Actividades de aprendizaje	Bibliografía	
			Básica	Complementaria
Introducción	Conocer las ventajas y casos donde es aconsejable recopilar información interna y externa en un Almacén de Datos, así como distinguir las	1.1. Finalidades y Evolución de los Sistemas de Información. 1.2. Herramientas para la Toma de Decisiones. 1.3. Interrelación entre Almacenes de Datos, OLAP y Minería de Datos. 1.4. Áreas de Aplicación de los Sistemas de Soporte a la Toma de Decisiones.	Pérez López, C., Santin González, D. (2007) Minería de Datos: Técnicas y Herramientas Ediciones España: Paraninfo, S.A. Rajaraman, A.,	Hernández Orallo, J., Ramírez Quintana, M.J., Ferri Ramírez, C. (2004). Introducción a la Minería de Datos. España: Pearson Prentice Hall Tan, P.N., Steinbach, M., Kumar, V. (2006).



Unidad 1	Objetivo Específico	Contenido Temático/Actividades de aprendizaje	Bibliografía	
			Básica	Complementaria
	propiedades y finalidades de las diferentes herramientas de soporte a la toma de decisiones.		<p>Leskovec, J., Ullman, J.D. (2011). Mining of Massive Datasets. Cambridge University Press.</p> <p>Witten, I.H., Frank, E., Hall, M.A. (2011). Data Mining: Practical Machine Learning Tools and Techniques 3th ed.). USA: Morgan Kaufman Publishers</p> <p>Hastie, T., Tibshirani, R., Friedman, J. (2009). The Elements of Statistical Learning: Data Mining, Inference, and Prediction. Springer</p>	<p>Introduction to Data Mining. Addison-Wesley</p> <p>Han, J., Kamber, M. (2006) Data Mining Concepts and Techniques. (2nd ed.). USA: Morgan Kaufman Publishers</p>

Unidad 2	Objetivo Específico	Contenido Temático/Actividades de aprendizaje	Bibliografía	
			Básica	Complementaria
Almacenes de datos	Conocer el modelo multidimensional de los almacenes	2.1. Introducción a los Almacenes de Datos. 2.2. Arquitectura de un Sistema de Almacén de	Rajaraman, A., Leskovec, J., Ullman, J.D. (2011). Mining	Hernández Orallo, J., Ramírez Quintana, M.J., Ferri Ramírez, C.

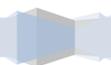


Unidad 2	Objetivo Específico	Contenido Temático/Actividades de aprendizaje	Bibliografía	
			Básica	Complementaria
	de datos y los operadores de refinamiento asociados: drill, roll, slice & dice, pivot; así como conocer la arquitectura y diferentes implementaciones de almacenes de datos	<p>Datos.</p> <p>2.3. Explotación de un Almacén de Datos utilizando OLAP.</p> <p>2.4. Sistemas ROLAP y MOLAP.</p> <p>2.5. Carga y Mantenimiento de un Almacén de Datos.</p> <p>2.6. Metodología de Diseño de un Almacén de Datos.</p> <p>2.7. Nuevas tendencias en los almacenes de datos</p> <p>2.7.1. Bases de Datos Espaciales</p> <p>2.7.2. Bases de Datos Multimedia (audio/video)</p> <p>2.7.3. Bases de Datos Estructuradas</p> <p>2.7.4. Bases de Datos Visuales (imágenes)</p>	<p>of Massive Datasets. Cambridge University Press.</p> <p>Witten, I.H., Frank, E., Hall, M.A. (2011). Data Mining: Practical Machine Learning Tools and Techniques 3th ed.). USA: Morgan Kaufman Publishers</p> <p>Hastie, T., Tibshirani, R., Friedman, J. (2009). The Elements of Statistical Learning: Data Mining, Inference, and Prediction. Springer</p>	<p>(2004). Introducción a la Minería de Datos. España: Pearson Prentice Hall</p> <p>Tan, P.N., Steinbach, M., Kumar, V. (2006). Introduction to Data Mining. Addison-Wesley</p> <p>Han, J., Kamber, M. (2006) Data Mining Concepts and Techniques. (2nd ed.). USA: Morgan Kaufman Publishers</p> <p>Pérez López, C., Santin González, D. (2007) Minería de Datos: Técnicas y Herramientas Ediciones España: Paraninfo, S.A.</p>

Unidad 3	Objetivo Específico	Contenido Temático/Actividades de aprendizaje	Bibliografía	
			Básica	Complementaria
Minería de datos	Reconocer la problemática del análisis de grandes volúmenes de datos y de los beneficios de su uso	<p>3.1. Introducción a la Minería de Datos.</p> <p>3.2. El proceso KDD</p> <p>3.3. Técnicas para el preprocesamiento de los datos</p> <p>3.3.1. Data Summarizacion</p> <p>3.3.2. Limpieza de datos:</p>	<p>Rajaraman, A., Leskovec, J., Ullman, J.D. (2011). Mining of Massive Datasets. Cambridge University</p>	<p>Hernández Orallo, J., Ramírez Quintana, M.J., Ferri Ramírez, C. (2004). Introducción a la Minería de Datos. España: Pearson</p>



Unidad 3	Objetivo Específico	Contenido Temático/Actividades de aprendizaje	Bibliografía	
			Básica	Complementaria
	sistemático para la obtención de modelos y patrones predictivos o descriptivo para elegir, en un problema concreto, qué técnicas de minería de datos son más apropiadas.	binning, regresión 3.3.3. Reducción de dimensionalidad: wavelets, PCA, LDA 3.3.4. Integración y transformación de datos: análisis de normalización, de correlaciones 3.4. Técnicas de Minería de Datos 3.4.1 Técnicas predictivas 3.4.2. Técnicas Descriptivas 3.5. Introducción a las técnicas de clasificación 3.5.1. Conceptos básicos: definición de clases, precisión en los clasificadores 3.5.2. Estimación de error de los clasificadores 3.5.2.1. Validación cruzada 3.5.2.2. Recuerdo 3.5.2.3. Precisión 3.5.2.4. F-measure 3.5.2.5. Curvas ROC 3.6. Técnicas de clasificación 3.6.1. Clasificación basada en árboles 3.6.2. Clasificación probabilística 3.6.3. Clasificación basada en modelos conexionistas 3.7. Técnicas de asociación de datos 3.7.1. Funciones de distancia 3.7.2. Algoritmos de asociación de datos 3.8. Evaluación de Hipótesis	Press. Witten, I.H., Frank, E., Hall, M.A. (2011). Data Mining: Practical Machine Learning Tools and Techniques 3th ed.). USA: Morgan Kaufman Publishers	Prentice Hall Tan, P.N., Steinbach, M., Kumar, V. (2006). Introduction to Data Mining. Addison-Wesley Han, J., Kamber, M. (2006) Data Mining Concepts and Techniques. (2 nd ed.). USA: Morgan Kaufman Publishers Manning, C.D., Raghavan, P, Schütze, H. (2008). Introduction to Information Retrieval. Cambridge University Press



Unidad 4	Objetivo Específico	Contenido Temático/Actividades de aprendizaje	Bibliografía	
			Básica	Complementaria
Desarrollo de proyectos utilizando una herramienta	Generar los modelos y patrones elegidos utilizando una herramienta de minería de datos para aplicarlos en la solución de un problema específico.	4.1 Generación y obtención del conjunto de datos de aprendizaje. 4.2 Conexión con bases de datos. 4.3 Preprocesamiento aplicando filtros. 4.4 Aplicación de técnicas predictivas y descriptivas. 4.5 Evaluación de resultados.	Rajaraman, A., Leskovec, J., Ullman, J.D. (2011). Mining of Massive Datasets. Cambridge University Press. Witten, I.H., Frank, E., Hall, M.A. (2011). Data Mining: Practical Machine Learning Tools and Techniques 3th ed.). USA: Morgan Kaufman Publishers Hastie, T., Tibshirani, R., Friedman, J. (2009). The Elements of Statistical Learning: Data Mining, Inference, and Prediction. Springer	Hernández Orallo, J., Ramírez Quintana, M.J., Ferri Ramírez, C. (2004). Introducción a la Minería de Datos. España: Pearson Prentice Hall Tan, P.N., Steinbach, M., Kumar, V. (2006). Introduction to Data Mining. Addison-Wesley Han, J., Kamber, M. (2006) Data Mining Concepts and Techniques. (2 nd ed.). USA: Morgan Kaufman Publishers



8. CONTRIBUCIÓN DEL PROGRAMA DE ASIGNATURA AL PERFIL DE EGRESO

Asignatura	Perfil de egreso (anotar en las siguientes tres columnas, cómo contribuye la asignatura al perfil de egreso)		
	Conocimientos	Habilidades	Actitudes y valores
La minería de datos pretende que el alumno aprenda a reconocer la problemática del análisis de grandes volúmenes de datos y aplicar técnicas de extracción de datos apropiadas para transformarlos en una estructura comprensible para su uso posterior.	Conocer y comprender el proceso de la Minería de Datos para extraer conocimiento desde base de datos y la aplicación de estas técnicas a la solución de problemas de reconocimiento de patrones, clasificación y pronóstico.	Capacidades para resolver problemas relacionados con el manejo de grandes volúmenes de datos, trabajo en equipo, toma de decisiones asertivas.	Entender los conflictos éticos y profesionales importantes el diseño de bases de datos y la seguridad en torno a ellas. Creatividad para resolver problemas relacionados con el manejo y visualización de grandes volúmenes de datos.

9. Describa cómo el eje o los ejes transversales contribuyen al desarrollo de la asignatura

Eje (s) transversales	Contribución con la asignatura
Formación Humana y Social	El alumno mostrará ante el grupo sus proyectos, y sus compañeros aprenderán a respetar el trabajo expuesto aportando críticas constructivas.
Desarrollo de Habilidades en el uso de las Tecnologías de la Información y la Comunicación	El alumno deberá hacer uso de las TIC para obtener información y software que le sirva de guía en la el curso de minería de datos mediante ejemplos. La búsqueda de información deberá ser de fuentes confiables electrónicas.
Desarrollo de Habilidades del Pensamiento Complejo	A partir del análisis de problemas que requieren utilizar minería de datos, los alumnos proponen soluciones adecuadas de acuerdo al contexto de implementación computacional.
Lengua Extranjera	El alumno deberá utilizar bibliografía de cada unidad en inglés con el fin de enterarse de los conceptos básicos y avances en el tema de minería de datos.
Innovación y Talento Universitario	Aplicación de sus talentos para innovar en el diseño de proyectos con calidad que resuelvan problemas de la vida real utilizando las bases de datos.
Educación para la Investigación	Se fomenta el hábito de investigar e implementar procesos de análisis y diseño de sistemas, así como los cambios continuos de la tecnología invita al alumno a estar en una constante investigación de los avances de desarrollo de bases de datos.

10. ORIENTACIÓN DIDÁCTICO-PEDAGÓGICA.

Estrategias y Técnicas de aprendizaje-enseñanza	Recursos didácticos
<p>Estrategias de aprendizaje:</p> <ul style="list-style-type: none"> • De las lecturas sugeridas por el docente, realizar: <ul style="list-style-type: none"> - Comprensión - Comparaciones - Análisis - Síntesis • Desarrollar proyectos de manera colaborativa. <p>Estrategia de enseñanza:</p> <ul style="list-style-type: none"> • Solucionar problemas reales mediante la aplicación de los conceptos planteados en clase. • Evaluación mediante rubricas de las actividades colaborativas. <p>Ambientes de aprendizaje:</p> <ul style="list-style-type: none"> • Aula • Laboratorio • Lenguajes de especificación y modelado. <p>Técnicas</p> <ul style="list-style-type: none"> • grupales • de debate • del diálogo • de estudio de casos • cuadros sinópticos • mapas conceptuales • comparación • lluvia de ideas 	<ul style="list-style-type: none"> • Libros del área de Bases de Datos y Minería de Datos • Software especializado para manejo de bases de datos y minería de datos • Pizarrón • Plataforma educativa Moodle/Blackboard • Materiales audiovisuales: extractos de conferencias especializadas en el área • Servicios telemáticos: páginas Web, Weblog, tours virtuales, webquest, correo electrónico, chats, foros, unidades didácticas y cursos on-line



11. CRITERIOS DE EVALUACIÓN

Criterios	Porcentaje
▪ Exámenes	25
▪ Tareas	15
▪ Prácticas de laboratorio	20
▪ Proyectos intermedios	15
▪ Proyecto final	25
Total	100%

12. REQUISITOS DE ACREDITACIÓN

Estar inscrito como alumno en la Unidad Académica en la BUAP
Asistir como mínimo al 80% de las sesiones
La calificación mínima para considerar un curso acreditado será de 6
Cumplir con las actividades académicas y cargas de estudio asignadas que señale el PE

13. Anexar (copia del acta de la Academia y de la CDESC- UA con el Vo. Bo. del Secretario Académico)

