



1. DATOS GENERALES DE LA UNIDAD DE APRENDIZAJE (UA) O ASIGNATURA			
Nombre de la Unidad de Aprendizaje (UA) o Asignatura			Clave de la UA
MINERIA DE DATOS			I5911
Modalidad de la UA	Tipo de UA	Área de formación	Valor en créditos
PRESENCIAL	CURSO	BÁSICO PARTICULAR	8
UA de pre-requisito	UA simultaneo	UA posteriores	
NINGUNO	NINGUNO	NINGUNO	
Horas totales de teoría	Horas totales de práctica	Horas totales del curso	
51	17	68	
Licenciatura(s) en que se imparte		Módulo al que pertenece	
INGENIERIA INFORMÁTICA		MODULO 4 CÓMPUTO FLEXIBLE (SOFT COMPUTING)	
Departamento		Academia a la que pertenece	
CIENCIAS COMPUTACIONALES		TÉCNICAS MODERNAS DE PROGRAMACIÓN	
Elaboró o revisó		Fecha de elaboración o revisión	
ISRAEL ROMÁN GODÍNEZ LUIS ANTONIO MEDELLIN SERNA		23 DE JULIO DE 2018	



2. DESCRIPCIÓN DE LA UA O ASIGNATURA

Presentación

El uso cotidiano de la computadora en nuestra sociedad ha mejorado substancialmente la capacidad de generar y recolectar grandes volúmenes de datos a partir de fuentes muy variadas. Este crecimiento desmesurado de datos ha generado la necesidad, urgente, de nuevas técnicas y herramientas automáticas que puedan ayudarnos a transformar ese conjunto de datos en información y conocimiento. Es a partir de esto último que se ha desarrollado una, relativamente, nueva rama de las ciencias computacionales conocida como Minería de Datos. Ésta se define como la extracción automática de patrones representativos de conocimiento que están, implícitamente, almacenados en grandes volúmenes de información, almacenes de datos, la Web, y otros repositorios de información. (Han, J., Pei, J., & Kamber, M. (2011). *Data mining: concepts and techniques*. Elsevier.)

Relación con el perfil

Modular

Desarrollar, en el alumno, las habilidades necesarias para proponer e implementar soluciones a problemas que involucren descubrir conjuntos de patrones, con el fin de otorgar información relevante de un volumen de datos, aplicando algoritmos de minería de datos. Dichas habilidades, en conjunto con otras desarrolladas en otras unidades de aprendizaje como programación e ingeniería del software, contribuye de manera fundamental en el perfil del egresado de la Ingeniería en Informática dotándolo de conocimientos necesarios en el ámbito de las tecnologías de la información y comunicación (TIC), para desempeñarse en el ámbito laboral de la región o a nivel nacional.

De Egreso

El egresado de Ingeniería Informática, habrá adquirido competencias profesionales e intelectuales en los campos de Sistemas de Información, gestión de las tecnologías de la información, sistemas robustos, paralelos, distribuidos y computación flexible (Centro Universitario de Ciencias Exactas e Ingenierías - UDG (2016). Ingeniería Informática [en línea]. Disponible en: <http://www.cucei.udg.mx/es/oferta-academica/licenciaturas/licenciatura-en-ingenieria-informatica> [Recuperado el 9 de Junio de 2017]).



UNIVERSIDAD DE GUADALAJARA

Competencias a desarrollar en la UA o Asignatura		
Transversales	Genéricas	Profesionales
<p>Desarrolla la capacidad de abstracción de problemas, no solo aquellos relacionados con el análisis de conjuntos de datos, si no también aquellos problemas que requieran de un análisis y solución generalizada.</p> <p>El alumno es capaz de aplicar las herramientas brindadas en esta unidad de aprendizaje tanto en entornos comunes como poco conocidos.</p> <p>Mediante trabajos colaborativos, el alumno aprenderá a resolver problemas en equipo.</p>	<p>Capacidad de abstracción, análisis y síntesis.</p> <p>Capacidad de aplicar los conocimientos en la práctica.</p> <p>Capacidad para organizar y planificar el tiempo</p> <p>Responsabilidad social y compromiso ciudadano</p> <p>Capacidad de comunicación oral y escrita</p> <p>Capacidad de comunicación en un segundo idioma</p> <p>Habilidades para buscar, procesar y analizar información procedente de fuentes diversas</p> <p>Capacidad crítica y autocrítica</p> <p>Capacidad de trabajo en equipo</p> <p>Habilidad para trabajar en forma autónoma</p> <p>Compromiso ético</p> <p>Compromiso con la calidad (Proyecto Tuning (2004-2007).</p> <p>Competencias genéricas de América Latina [en línea]. Disponible en: http://www.tuningal.org/es/competencias/geologia [Recuperado el 20 de Julio de 2016])</p>	<p>El egresado podrá desarrollarse en el sector público y privado, o emprender su propia empresa en los campos relacionados con el uso eficiente de las Tecnologías de la Información y Comunicación (TIC); el diseño de software que resuelva diversas aplicaciones complejas, involucrándose en las diferentes ramas de la ingeniería; el diseño y mantenimiento de sistemas de información, mediante los cuales se guía la operación cotidiana de las empresas; y en el campo de la docencia e investigación.</p> <p>(DIVEC-CUCEI-UDG (2015). Ingeniería Informática, Plan Modular. [en línea]. Disponible en: http://www.cucei.udg.mx/sites/default/files/oferta_academica/licenciaturas/folletos/lic_info.pdf [Recuperado el 20 de Julio de 2016]).</p>
Saberes involucrados en la UA o Asignatura		
Saber (conocimientos)	Saber hacer (habilidades)	Saber ser (actitudes y valores)



UNIVERSIDAD DE GUADALAJARA

<ul style="list-style-type: none">- Define los términos datos, información y conocimiento.- Diferencia concepto, instancia y atributo.- Describe las técnicas de pre-procesamiento de datos- Identifica las medidas de similitud y de diferencia- Interpreta las diferentes representaciones de salidas de datos- Distingue entre los diferentes algoritmos de aprendizaje máquina- Contrasta las diferentes técnicas de evaluación de credibilidad de los algoritmos	<ul style="list-style-type: none">- Comunicación oral y escrita- Capacidad de abstracción- Manejo de razonamiento lógico- Capacidad para analizar y resolver problemas- Capacidad de aprender por cuenta propia- Destreza para el uso de la computadora	<ul style="list-style-type: none">- Llega dentro del tiempo establecido a la clase- Guarda silencio mientras otras personas están hablando- No toma el trabajo ajeno como si fuera propio- Trabaja de forma autónoma en las actividades dentro y fuera del aula- Colabora con sus compañeros en las actividades dentro y fuera del aula- Entrega los productos de las actividades en la fecha que se le indica- Participa activamente durante la clase
---	--	--

Producto Integrador Final de la UA o Asignatura

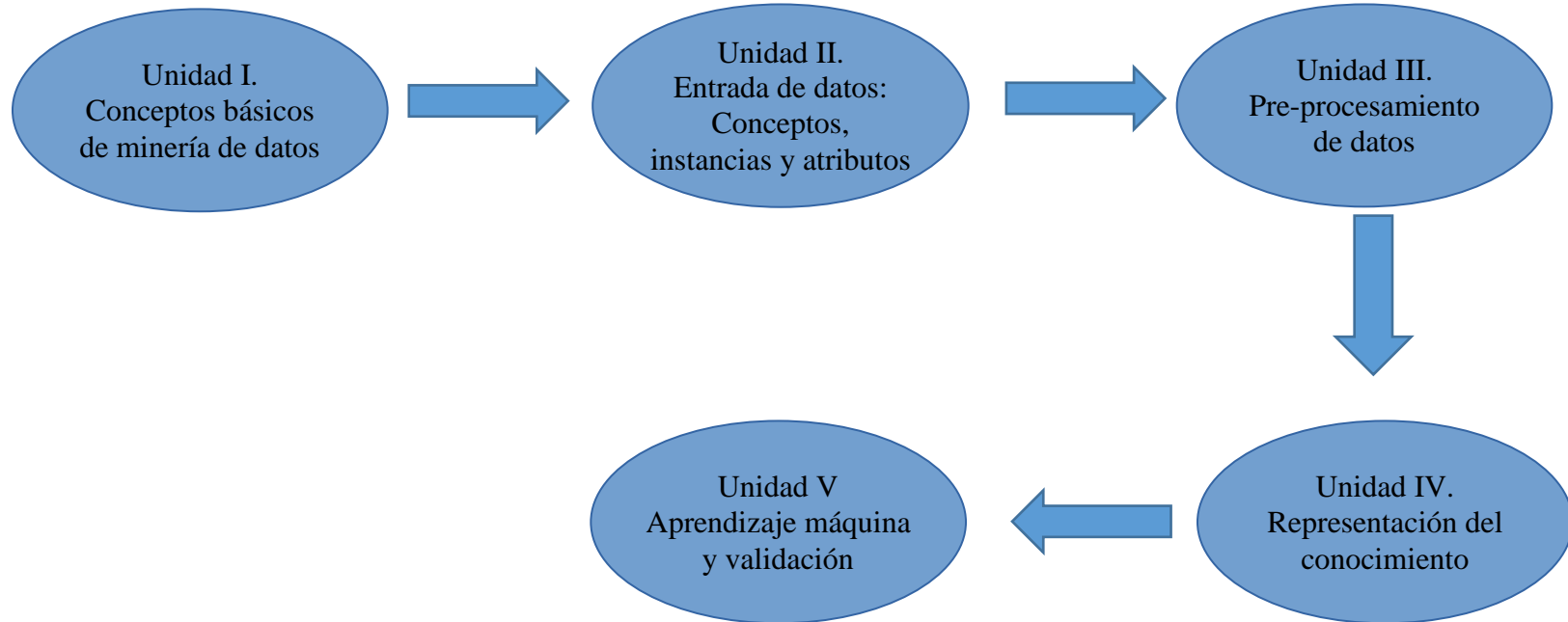
Título del Producto: Descubriendo patrones.

Objetivo: Implementar la metodología de descubrimiento de datos en un conjunto de datos.

Descripción: Utilizando las diferentes herramientas de minería de datos, el estudiante deberá proponer un modelo que puede ser de clasificación, regresión, agrupamiento o asociación represente la problemática establecida. Deberá verificar que el modelado propuesto cumpla con alguna de las técnicas de validación de modelos. Posteriormente, dicho modelo deberá ser implementado utilizando algún lenguaje de programación (establecido por el profesor o el alumno). El análisis deberá proponer un conjunto de reglas que permita tomar decisiones sobre conjuntos de datos nuevos (desconocidos) y proponer algunas estrategias de mejoras para el cliente, utilizando dichas reglas.



3. ORGANIZADOR GRÁFICO DE LOS CONTENIDOS DE LA UA O ASIGNATURA





4. SECUENCIA DEL CURSO POR UNIDADES TEMÁTICAS

Unidad temática 1

Objetivo de la unidad temática

Los alumnos deberán ser capaces de:

- Reconocer los problemas que puedan ser solucionados con el uso de algún modelo minería de datos
- Listar las características de los conjuntos de entradas para los algoritmos de minería de datos
- Identificar los tipos de modelos de minería de datos y que tipos de problemas resuelven
- Clasificar a los diferentes actores que participan en los procesos de minería de datos

Introducción:

La es parte del proceso de descubrimiento de conocimiento. Es necesario identificar las diferentes etapas del descubrimiento de conocimiento e identificar los tipos de problemas que se pueden resolver a través de esta técnica (clasificación, asociación, agrupamiento, regresión).

Contenido temático	Saberes involucrados	Producto de la unidad temática
<p>1. CONCEPTOS BÁSICOS DE MINERÍA DE DATOS</p> <p>1.1 Introducción</p> <p>1.2 Problemas que resuelve la minería de datos.</p> <p>1.3 Alcances y limitaciones de la minería de datos.</p>	<ul style="list-style-type: none">• Conocer las características principales de los procesos de descubrimiento de datos.• Reconocer los principales actores que intervienen en el descubrimiento de datos.	<ul style="list-style-type: none">• Documento electrónico que contiene las respuestas de un cuestionario de conceptos básicos de minería de datos.• Elaborar un organizador gráfico (cuadro o tabla comparativa) de las características principales del proceso de descubrimiento de datos.• Incluir la reflexión o conclusión del estudiante acerca de lo investigado.



UNIVERSIDAD DE GUADALAJARA

Actividades del docente	Actividades del estudiante	Evidencia de la actividad	Recursos materiales	y	Tiempo destinado
El docente solicitará a los alumnos la realización de las actividades correspondientes a los temas de la UA.	Trabajo en equipo en la realización de actividades.	Se entregarán un documento electrónicos (uno por cada actividad realizada en el aula), que muestre el trabajo en equipo.	Computadora Libro		4 Horas
Entregado el documento al profesor, éste emitirá retroalimentación como parte de la heteroevaluación.	Presentación de temas vistos en las actividades	Entrega de las diapositivas	Diapositivas Cañón Pintarron		4 Horas
El docente identificará a los estudiantes que parezcan tener una menor comprensión del tema e intentará aclarar sus dudas.	Responder a las diferentes preguntas elaboradas, tanto por compañeros como por el docente.	N/A	Pintarron		2 Horas

Unidad temática 2:

Objetivo de la unidad temática

Los alumnos deberán ser capaces de identificar los elementos principales en un conjunto de entrada de datos :

- Instancias
- Atributos
- Tipos de datos
- Similitud entre tipos de datos
- Similitud entre tipos de datos mezclados

Introducción:

Para poder encontrar patrones es necesario poner, en términos computacionales, las necesidades de alguna problemática del mundo real. Para ello es necesario conocer las diferentes formas de representación de los conjuntos de datos, los tipos de estos y como medir similitudes entre elementos que comparten los mismo o diferentes tipos de datos

Contenido temático	Saberes involucrados	Producto de la unidad temática
2. Tema 2.1 Instancias 2.2 Atributos 2.3 Tipos de datos 2.4 Similitud entre tipos de datos	<ul style="list-style-type: none"> • Ser capaz de construir un conjunto de datos que se ajuste a las necesidades de los algoritmos de minería de datos • Identificar los diferentes tipos de datos que existen. 	<ul style="list-style-type: none"> • Documento donde demuestre el estudiante la correcta construcción de conjuntos de datos. • También describa e



UNIVERSIDAD DE GUADALAJARA

	<ul style="list-style-type: none"> • Ser capas de medir la similitud y disimilitud entre dos instancias, con los mismos o diferentes tipos de datos. 	<p>identifique los diferentes tipos de datos que se encuentran en los conjuntos.</p> <ul style="list-style-type: none"> • Por último, identifique, describa y aplique las diferentes medidas de similitud y disimilitud de datos. • Una vez terminado el documento se debe compartir con al menos dos compañeros para coevaluación.
--	---	---

Actividades del docente	Actividades del estudiante	Evidencia de la actividad	Recursos materiales y	Tiempo destinado
El docente solicitará a los alumnos la realización de las actividades correspondientes a los temas de la UA (Una por cada tema).	Trabajo en equipo en la realización de actividades.	Se entregarán un documento electrónicos (uno por cada actividad realizada en el aula), que muestre el trabajo en equipo.	Computadora Libro	5 HORAS
Entregado el documento al profesor, éste emitirá retroalimentación como parte de la heteroevaluación.	Presentación de temas vistos en las actividades	Entrega de las diapositivas	Diapositivas Cañón Pintarron	5 HORAS
El docente identificará a los estudiantes que parezcan tener una menor comprensión del tema e intentará aclarar sus dudas.	Responder a las diferentes preguntas elaboradas, tanto por compañeros como por el docente.	N/A	Pintarron	5 HORAS

Unidad temática 3:



Objetivo de la unidad temática

Los alumnos deberán ser capaces de:

- Identificar los pasos del pre-procesamiento de datos
- Conocer diferentes técnicas de análisis estadístico de los conjuntos de datos
- Identificar cuál es la mejor técnica de análisis estadístico para cada conjunto de datos
- Utilizar las diferentes estrategias de sustitución de valores faltantes.
- Análisis de correlación entre los atributos de los diferentes conjuntos de datos

Introducción:

Dar a conocer al alumno las diferentes estrategias de pre-procesamiento de datos, desde sus proceso hasta las diferentes estrategias estadísticas que te permiten ajustar tu conjunto de datos para poder crear modelos que representen de mejor forma un problema del mundo real.

Contenido temático		Saberes involucrados		Producto de la unidad temática	
3. Tema 3.1 Preprocesamiento de datos 3.2 Limpieza de datos 3.3 Análisis estadístico univariable 3.4 Análisis estadístico bivariable		<ul style="list-style-type: none"> • Uso de algunas estadísticas para análisis univariable. • Uso de algunas estadísticas para análisis bivariable. • Conocimiento de las principales herramientas para la limpieza de datos 		<ul style="list-style-type: none"> • Documento que describa e identifique el pre-procesamiento de datos. • Análisis que contenga las principales estrategias de limpieza de datos • Análisis de los atributos de un conjunto de datos • Análisis bivariable de un conjunto de datos. 	
Actividades del docente	Actividades del estudiante	Evidencia de la actividad	Recursos materiales y	Tiempo destinado	
El docente solicitará a los alumnos la realización de las actividades correspondientes a los temas de la UA (Una por cada tema).	Trabajo en equipo en la realización de actividades.	Se entregarán un documento electrónicos (uno por cada actividad realizada en el aula), que muestre el trabajo en equipo.	Computadora Libro	4 HORAS	
Entregado el documento al	Presentación de temas vistos en las	Entrega de las	Diapositivas	4 HORAS	



UNIVERSIDAD DE GUADALAJARA

profesor, éste emitirá retroalimentación como parte de la heteroevaluación.	actividades	diapositivas	Cañón Pintarron	
El docente identificará a los estudiantes que parezcan tener una menor comprensión del tema e intentará aclarar sus dudas.	Responder a las diferentes preguntas elaboradas, tanto por compañeros como por el docente.	N/A	Pintarron	2 HORAS
Unidad temática 4:				
<p>Objetivo de la unidad temática Los alumnos deberán ser capaces de: - Identificar las diferentes formas de representación de los modelos de aprendizaje. - Obtener los patrones de un conjunto de datos que represente un problema en particular</p> <p>Introducción: Los diferentes modelos de aprendizaje entregan diferentes representaciones de salida de información. Por lo cual es necesario que el alumno conozca e identifique cada una de ellas, con el fin de poder interpretarlas.</p>				
Contenido temático	Saberes involucrados	Producto de la unidad temática		
<p>4. Tema 4.1 Representación de salida de modelos de clasificación. 4.2 Representación de salida de modelos de asociación 4.3 Representación de salida de modelos de agrupamiento. 4.4 Representación de salida de modelos de regresión</p>	<ul style="list-style-type: none"> • Uso de las herramientas para el aprendizaje de los diferentes modelos. 	<ul style="list-style-type: none"> • Documento que describa e identifique las diferentes representaciones por modelo. • Incluir una solución a un conjunto de casos de estudio a través de la elaboración de diferentes modelos. • Una vez entregado el documento al profesor, emitirá retroalimentación de la actividad como parte de la heteroevaluación. 		
Actividades del docente	Actividades del estudiante	Evidencia de la actividad	Recursos materiales y	Tiempo destinado



UNIVERSIDAD DE GUADALAJARA

El docente solicitará a los alumnos la realización de las actividades correspondientes a los temas de la UA (Una por cada tema).	Trabajo en equipo en la realización de actividades.	Se entregarán un documento electrónicos (uno por cada actividad realizada en el aula), que muestre el trabajo en equipo.	Computadora Libro	5 HORAS
Entregado el documento al profesor, éste emitirá retroalimentación como parte de la heteroevaluación.	Presentación de temas vistos en las actividades	Entrega de las diapositivas	Diapositivas Cañón Pintarron	4 HORAS
El docente identificará a los estudiantes que parezcan tener una menor comprensión del tema e intentará aclarar sus dudas.	Responder a las diferentes preguntas elaboradas, tanto por compañeros como por el docente.	N/A	Pintarron	4 HORAS

Unidad temática 5:

Objetivo de la unidad temática

Los alumnos deberán ser capaces de:

- Conocer algunos modelos de clasificación.
- Conocer algunos modelos de regresión
- Conocer algunos modelos de asociación
- Conocer algunos modelos de agrupamiento

Introducción:

La minería de datos se basa en la exploración de un conjunto de datos usando un conjunto diferente de herramientas computacionales que estan dentro del área de aprendizaje automático. En esta unidad de aprendizaje se presentan al menos un algoritmo para cada uno de los modelos de aprendizaje.

Contenido temático	Saberes involucrados	Producto de la unidad temática
5. Tema 5.1 Algoritmo de clasificación: ZeroR, OneR, J48, CART. 5.2 Algoritmos de regresión: regresión lineal. 5.3 Algoritmos de asociación: A priori 5.4 Algoritmos de agrupamiento: k-means. 5.5 K-fold Cross-Validation, Hold-out,	Conocer el funcionamiento de los diferentes algoritmos. Construir un programa que realice las tareas de cada uno de los modelos. Validar los modelos construidos utilizando alguna	<ul style="list-style-type: none"> • Documento que describa y muestre el resultado de aplicar cada uno de los modelos con un conjunto de datos. • Programar cada uno de los modelos de



UNIVERSIDAD DE GUADALAJARA

Percentage Split		de las técnicas de validación		aprendizaje. <ul style="list-style-type: none"> • Provar diferentes set de datos en los programas contruidos utilizando las técnicas de validación • Una vez terminado el documento se debe compartir con al menos dos compañeros para coevaluación. 	
Actividades del docente	Actividad del estudiante	Evidencia de la actividad	Recursos materiales y	Tiempo destinado	
El docente solicitará a los alumnos la realización de las actividades correspondientes a los temas de la UA (Una por cada tema).	Trabajo en equipo en la realización de actividades.	Se entregarán un documento electrónicos (uno por cada actividad realizada en el aula), que muestre el trabajo en equipo.	Computadora Libro	15 HORAS	
Entregado el documento al profesor, éste emitirá retroalimentación como parte de la heteroevaluación.	Presentación de temas vistos en las actividades	Entrega de las diapositivas	Diapositivas Cañón Pintarron	3 HORAS	
El docente identificará a los estudiantes que parezcan tener una menor comprensión del tema e intentará aclarar sus dudas.	Responder a las diferentes preguntas elaboradas, tanto por compañeros como por el docente.	N/A	Pintarron	2 HORAS	

5. EVALUACIÓN Y CALIFICACIÓN

Requerimientos de acreditación:

Ordinaria

De acuerdo con el Reglamento General de Evaluación y Promoción de Alumnos de la Universidad de Guadalajara:

Artículo 20. Para que el alumno tenga derecho al registro del resultado final de la evaluación en el periodo ordinario, establecido en el calendario escolar aprobado por el H. Consejo General Universitario, se requiere:

- I. Estar inscrito en el plan de estudios y curso correspondiente, y
- II. Tener un mínimo de asistencia del 80% a clases y actividades registradas durante el curso.

(<http://www.secgral.udg.mx/sites/archivos/normatividad/general/ReglamentoGralEPAAlumnos.pdf>) *Extraordinaria*

De acuerdo con el Reglamento General de Evaluación y Promoción de Alumnos:



UNIVERSIDAD DE GUADALAJARA

Artículo 25. La evaluación en periodo extraordinario se calificará atendiendo a los siguientes criterios:

- I. La calificación obtenida en periodo extraordinario, tendrá una ponderación del 80% para la calificación final;
- II. La calificación obtenida por el alumno durante el periodo ordinario, tendrá una ponderación del 40% para la calificación del periodo extraordinario, y
- III. La calificación final para la evaluación en periodo extraordinario será la que resulte de la suma de los puntos obtenidos en las fracciones anteriores.

Artículo 27. Para que el alumno tenga derecho al registro de la calificación en el periodo extraordinario, se requiere:

- I. Estar inscrito en el plan de estudios y curso correspondientes.
- II. Haber pagado el arancel y presentar el comprobante correspondiente.
- III. Tener un mínimo de asistencia del 65% a clases de actividades registradas durante el curso.

(<http://www.secgral.udg.mx/sites/archivos/normatividad/general/ReglamentoGralEPAlumnos.pdf>)

Criterios generales de evaluación:

Entregar los productos en la plataforma electrónica elegida por el profesor y en el día y hora establecidos, cumpliendo con los criterios de cada una de las actividades.

Evidencias o Productos

Evidencia o producto	Competencias y saberes involucrados	Contenidos temáticos	Ponderación
Actividades de Aprendizaje	Deberá reconocer las necesidades de una problemática de descubrimiento de patrones y aplicar los recursos necesarios para solucionarla de manera eficiente.	Consiste en las evidencias, producto del trabajo en clases, tareas y actividades; entregadas de manera presencial o a través de algún medio electrónico de acuerdo con la selección del profesor.	30%
Exámenes Parciales	Demostrar los conocimientos adquiridos a través de una serie de preguntas formuladas por el profesor.	Aplicación de dos exámenes durante el curso con valor de 25 puntos cada uno, en los cuales de evaluarán los aspectos del Saber y el Saber Hacer descritos en el documento, los cuales podrán ser aplicados de manera escrita o a	50%



UNIVERSIDAD DE GUADALAJARA

		través del apoyo de algún entorno de aprendizaje o plataforma virtual de acuerdo con la selección del profesor	
Proyecto final	Capacidad de aplicar conocimientos	Aplicación del proceso de descubrimiento de datos en un conjunto de datos presentado por el profesor. Se entregará un documento que muestre las conclusiones encontradas.	20%

Producto final

Descripción		Evaluación	
Título: Análisis de un conjunto de datos		Criterios de fondo: <ul style="list-style-type: none"> - Lista de requerimientos del problema - Análisis estadístico de los datos. - Preprocesamiento de datos - Minería de datos - Interpretación del modelo aprendido - Elaboración de una serie de conclusiones sobre el conjunto de patrones encontrados. Criterios de forma: <ul style="list-style-type: none"> - Hoja de presentación 	Ponderación
Objetivo: Que el alumno aplique el proceso de descubrimiento de datos en un conjunto de datos.			20%
Caracterización El proyecto debe dar una solución a un problema de identificación de patrones utilizando alguna técnica de aprendizaje automático.			



UNIVERSIDAD DE GUADALAJARA

	<ul style="list-style-type: none">- Descripción del problema.- Técnicas de limpieza de datos aplicada.- Aplicación del conjunto de datos sobre el modelo de aprendizaje seleccionado.	
Otros criterios		
Criterio	Descripción	Ponderación



6. REFERENCIAS Y APOYOS

Referencias bibliográficas

BASICA:

Jiawei, H., & Kamber, M. (2001). *Data mining: concepts and techniques*. San Francisco, CA, itd: Morgan Kaufmann, 5.

Witten, I. H., & Frank, E. (2005). *Data Mining: Practical machine learning tools and techniques*. Morgan Kaufmann.

COMPLEMENTARIA:

Hernández Orallo, J., Ramírez Quintana, M. J., & Ferri Ramírez, C. (2004). *Introducción a la Minería de Datos*. Pearson Prentice Hall.

Apoys (videos, presentaciones, bibliografía recomendada para el estudiante)

Unidad temática 1:

Unidad temática 2:

Unidad temática 3:

Unidad temática 4:

Unidad temática 5: