

**UNIVERSIDAD AUTÓNOMA DE BAJA CALIFORNIA**  
**COORDINACIÓN DE FORMACIÓN BÁSICA**  
**COORDINACIÓN DE FORMACIÓN PROFESIONAL Y VINCULACIÓN UNIVERSITARIA**  
**PROGRAMA DE UNIDAD DE APRENDIZAJE**

**I. DATOS DE IDENTIFICACIÓN**

1. Unidad Académica: Facultad de Ciencias

2. Programa (s) de estudio: Nivel: Licenciatura en Ciencias Computacionales      3. Vigencia del plan:

4. Nombre de la Unidad de aprendizaje: Minería de Datos      5. Clave:

6. HC: 1   HL: 2   HT: 2   HPC:       HCL:       HE: 1   CR: 6

7. Etapa de formación a la que pertenece: Disciplinaria

8. Carácter de la Unidad de aprendizaje:      Obligatoria   X        Optativa

9. Requisitos para cursar la unidad de aprendizaje:

Formuló: Dr. Everardo Gutiérrez López

Vo.Bo. Dr. Alberto Leopoldo Morán y Solares

Fecha: Septiembre de 2016

Cargo: Subdirector

## **II. PROPÓSITO GENERAL DE LA UNIDAD DE APRENDIZAJE**

El propósito de esta asignatura es introducir a los alumnos en el área de investigación de las ciencias computacionales conocida como Minería de Datos, la cual se dedica a estudiar los modelos y técnicas involucrados en el análisis computacional de datos.

En particular el contenido de esta unidad de aprendizaje se enfoca al estudio de las técnicas de pre-procesamiento, almacenamiento y análisis de datos, así como de los modelos tanto descriptivos como predictivos utilizados en las tareas de minería, terminando con la revisión de las nuevas tendencias en esta área de estudio. Todo esto ejemplificado a través de su aplicación en diversos ámbitos de interés para buscar alternativas viables a las diversas y crecientes problemáticas que requieren soluciones computacionales.

La revisión y análisis del contenido de esta unidad de aprendizaje permitirá al alumno la adquisición de los conocimientos necesarios para su posterior uso en el desarrollo de soluciones en el procesamiento y análisis de datos para su utilización en la solución de los problemas prácticos que se le presenten durante el desarrollo de su actividad profesional.

Esta asignatura es obligatoria y se encuentra en la etapa disciplinaria de la Licenciatura en Ciencias Computacionales.

## **III. COMPETENCIA DE LA UNIDAD DE APRENDIZAJE**

Construir modelos de minería de datos a través de la utilización de las técnicas fundamentales involucradas en el análisis computacional de datos para su aplicación en la solución de problemáticas prácticas, con una actitud propositiva, innovadora, honesta y de responsabilidad.

## **V. EVIDENCIA (S) DE DESEMPEÑO**

Construye un modelo de minería de datos como proyecto final que utilice las técnicas fundamentales involucradas en el análisis computacional de datos para solucionar alguna de las problemáticas prácticas especificadas por el docente. Presenta y reporta los resultados obtenidos tanto de forma oral como escrita.

## V. DESARROLLO POR UNIDADES

### **UNIDAD I. Introducción a la Minería de Datos**

**Competencia:** Analizar los antecedentes y conceptos generales de la minería de datos, sus paradigmas y dominios de aplicación, a partir del estudio de los conceptos básicos y la resolución de problemáticas de ejemplo, para examinar las diferentes etapas que se siguen en el proceso de análisis computacional de datos, con una actitud crítica, sistemática y propositiva.

#### **Contenido**

**Duración: 3 horas**

I. Introducción a la Minería de Datos.

- I.1. Definiciones e importancia de la minería de datos.
- I.2. Tecnologías utilizadas en la Minería de Datos.
- I.3. Descripción estadística de los datos.
- I.4. Visualización de la información.
- I.5. Medidas de similitud y diferencia entre datos.
- I.6. Aplicaciones y retos.

### **UNIDAD II. Técnicas de Preprocesamiento, Almacenamiento y Análisis de Datos**

**Competencia:** Analizar los principales conceptos involucrados en el pre-procesamiento, almacenamiento y análisis de datos, mediante la revisión de las técnicas involucradas en estas estas tareas así como las herramientas computacionales disponibles para realizarlas, con el fin de detectar sus alcances y diseñar alternativas de solución a problemáticas de aplicación propuestas por el docente, con una actitud crítica, propositiva e innovadora.

#### **Contenido**

**Duración: 4 horas**

II. Técnicas de Preprocesamiento, Almacenamiento y Análisis de Datos

- II.1. Importancia y retos del preprocesamiento de datos.
- II.2. Limpieza, Integración y Reducción.
- II.3. Transformación y Discretización..
- II.4. Almacenes de datos.
- II.5. Herramientas de análisis de datos.

**UNIDAD III. Modelos Descriptivos de Minería de Datos.**

**Competencia:** Examinar los principales modelos descriptivos de minería de datos para la construcción de estrategias de caracterización de la información por medio de la implementación de sus algoritmos básicos y el análisis comparativo de su comportamiento en diferentes aplicaciones, con una actitud crítica, propositiva e innovadora.

**Contenido****Duración: 4 horas**

## III. Modelos Descriptivos de Minería de Datos.

- III.1. Análisis de correlación de datos.
- III.2. Descubrimiento de Patrones.
- III.3. Análisis de conglomerados.
- III.4. Aprendizaje de máquina no supervisado.

**UNIDAD IV. Modelos Predictivos de Minería de Datos**

**Competencia:** Concebir la predicción de información mediante la utilización de modelos predictivos de minería de datos para el diseño de soluciones algorítmicas en sistemas computacionales, a partir del análisis de dichos modelos y el diseño de soluciones para casos de aplicación, con una actitud crítica, propositiva e innovadora.

**Contenido****Duración: 3 horas**

## IV. Modelos Predictivos de Minería de Datos

- IV.1. Técnicas de regresión de datos.
- IV.2. Técnicas de clasificación.
- IV.3. Aprendizaje de máquina supervisado.

## **UNIDAD V. Tópicos avanzados de Minería de Datos.**

**Competencia:** Relacionar el área de estudio de la minería de datos, sus métodos y técnicas, con los nuevos retos y áreas de oportunidad en el contexto actual de los sistemas computacionales, para el diseño de propuestas de solución de análisis de datos aplicadas a casos de estudio actuales, así como la exposición del proceso y resultados en formato de presentación y reporte final de la unidad de aprendizaje, con una actitud crítica, sistemática y siguiendo los estándares de redacción de trabajos científicos.

### **Contenido**

**Duración: 2 horas**

#### V. Tópicos avanzados de Minería de Datos.

- V.1. Modelos Prescriptivos.
- V.2. Técnicas de Visualización.
- V.3. Manejo de grandes volúmenes de datos.

## VI. ESTRUCTURA DE LAS PRÁCTICAS

No. de Práctica	Competencia(s)	Descripción	Material de Apoyo	Duración
1	<p>Comparar las diferentes etapas que se siguen en el proceso de análisis computacional de datos para la discriminación de los conceptos generales de la minería de datos, sus paradigmas y dominios de aplicación, a partir del estudio de los conceptos básicos y la resolución de problemáticas de ejemplo, con una actitud crítica, sistemática y propositiva.</p>	<p>Revisión bibliográfica histórica de la minería de datos. Comparación y clasificación de las etapas que se siguen en el proceso de análisis computacional de datos. Programación para la resolución de ejemplos.</p>	<p>Bibliografía, Computadora, Lenguaje Compilador.</p>	<p>8 HL 8 HT</p>
2	<p>Contrastar los principales conceptos involucrados en el pre-procesamiento, almacenamiento y análisis de datos, mediante la revisión de las técnicas de minería de datos involucradas en estas tareas así como las herramientas computacionales disponibles para realizarlas, con el fin de detectar sus alcances y diseñar alternativas de solución a problemáticas de aplicación propuestas por el docente, con una actitud crítica, propositiva e innovadora.</p>	<p>Revisión bibliográfica histórica de la minería de datos. Comparación y clasificación de las alternativas para el pre-procesamiento y análisis de datos . Programación de alternativas de solución para la resolución de problemáticas de aplicación.</p>	<p>Bibliografía, Computadora, Lenguaje Compilador.</p>	<p>8 HL 8 HT</p>
3	<p>Examinar modelos descriptivos de minería de datos para la construcción</p>	<p>Revisión bibliográfica. Comparación de las principales</p>	<p>Bibliografía, Computadora,</p>	<p>8 HL 8 HT</p>

	de estrategias de caracterización de la información por medio de la implementación de sus algoritmos y el análisis comparativo de su comportamiento en aplicaciones, con una actitud crítica, propositiva e innovadora.	estrategias de caracterización de la información y los algoritmos correspondientes. Programación de alternativas de solución para la resolución de problemáticas de aplicación.	Lenguaje Compilador.	
4	Concebir la predicción de información mediante la utilización de modelos predictivos de minería de datos para el diseño de soluciones algorítmicas, a partir del análisis de dichos modelos y el diseño de soluciones para casos de aplicación, con una actitud crítica, propositiva e innovadora.	Revisión bibliográfica. Comparación de las principales estrategias predicción de la información y los algoritmos correspondientes. Programación de alternativas de solución para la resolución de problemáticas de aplicación.	Bibliografía, Computadora, Lenguaje Compilador.	8 HL 8 HT

## VII. METODOLOGÍA DE TRABAJO

El docente motivará el análisis y discriminación de las diversas aproximaciones que siguen los métodos y técnicas de minería de datos con el objetivo de determinar la mejor alternativa para obtener la información de las fuentes de datos a considerar. El proceso deberá llevarlo a cabo el alumno utilizando el pensamiento lógico, para poder proponer soluciones algorítmicas a problemas computacionales de minería de datos, tanto para solucionar problemas teórico-prácticos como su utilización en programas computacionales escritos en un lenguaje moderno de alto nivel.

El docente incentivará la participación del alumno mediante discusiones críticas, intercambio de ideas, dinámicas de grupo, asignaciones extras que representen un reto adicional, entre otros. También se sugiere solicitar investigación en diferentes fuentes bibliográficas sobre temas de actualidad o temáticas que serán discutidos posteriormente en clase para fomentar el autoaprendizaje y la investigación en medios electrónicos (Internet), libros, y revistas sobre temas del área.

El maestro deberá aplicar al menos 2 exámenes de conocimientos durante el periodo, que permitan identificar la obtención de competencias de los estudiantes. Los exámenes podrán ser de varios tipos, tales como: de preguntas abiertas, opción múltiple y solicitud de programas.

## VIII. CRITERIOS DE EVALUACIÓN

### **Requisitos de Acreditación**

Acreditar los parciales en tiempo y forma, y entregar los ejercicios prácticos correspondientes.

Programar el proyecto final, presentarlo frente al grupo y elaborar un reporte del proyecto en formato de artículo.

Cumplir con los requisitos establecidos en el estatuto escolar vigente.

### **Criterios de Evaluación**

El proyecto final se evaluará considerando tres aspectos: resultados prácticos obtenidos con la implementación, presentación del proyecto ante el grupo y elaboración del reporte final. En la evaluación de los exámenes se considerará la capacidad del alumno para discriminar los métodos y técnicas de minería de datos que le permitan resolver los ejercicios que se le planteen.

Para la acreditación de la unidad de aprendizaje se atenderá al Estatuto Escolar Vigente, artículos 70-71, por lo que el estudiante deberá contar un mínimo de 80% de asistencias en el periodo. Tener un mínimo aprobatorio de 60 en su calificación final.

### **Calificación**

Al menos tres exámenes parciales: 45%

Ejercicios de programación: 15%

Ejercicios prácticos: 10%

Proyecto final: 30%.



## IX. BIBLIOGRAFÍA

### Básica

- Peter Flach. Machine Learning: The Art and Science of Algorithms that Make Sense of Data. Cambridge University Press, 409p. 2012.
- Jiawei Han, Micheline Kamber. Data Mining: Concepts and Techniques, Third Edition. Morgan Kaufmann, 744p. 2011.
- Mohammed J. Zaki, Wagner Meira Jr. Data Mining and Analysis: Fundamental Concepts and Algorithms. Cambridge University Press, 562p. 2014
- ACM Special Interest Group on Knowledge Discovery and Data Mining. <http://www.kdd.org/>

### Complementaria

- David J. Hand, Heikki Mannila, Padhraic Smyth. Principles of Data Mining. A Bradford Book, 584p. 2001. [Clásico]
- Christopher D. Manning, Prabhakar Raghavan, Hinrich Schütze. Introduction to Information Retrieval. Cambridge University Press, 496p. 2008. [Clásico]
- Ian H. Witten, Eibe Frank, Mark A. Hall. Data Mining: Practical Machine Learning Tools and Techniques, Third Edition. Morgan Kaufmann, 664p. 2011.
- ACM Special Interest Group on Algorithms and Computation Theory. <http://www.sigact.org/>
- ACM Special Interest Group on Management of Data. <http://www.sigmod.org/>

## X. Perfil Docente Deseable

El docente deberá ser profesionalista en ciencias computacionales o áreas afines, además deberá tener conocimiento de las principales técnicas de análisis de datos, así como experiencia en la resolución de problemáticas que requieran dicho análisis.