



<b>UNIVERSIDAD NACIONAL DE CATAMARCA</b>				
<b>FACULTAD DE TECNOLOGÍA Y CIENCIAS APLICADAS</b>				
DEPARTAMENTO DE INFORMATICA				
<b>PROGRAMA DE:</b>			<b>Código:</b> 7CIII	
<b>MINERÍA DE DATOS y DATA WAREHOUSE</b>			<b>Área:</b> ELECTIVA- SISTEMAS GERENCIALES	
			<b>Curso:</b> 5to año	
<b>Carga horaria Total: 75</b>			<b>Plan:</b> 2011	
<b>Horas</b>			<b>Régimen:</b> Cuatrimestral	
			<b>Cuatrimestre:</b> 1er	
			<b>Cuerpo Docente</b>	
<b>Teórico</b>	<b>Actividad Práctica</b>			Profesor: Dr. Hernán César Ahumada Profesor Adjto Semiexclusivo Lic. Claudia Alejandra Bazán JTP Simple
	FE <sup>1</sup>	RPI <sup>2</sup>	ADyP <sup>3</sup>	
40	20	15	--	
<b>Correlativas</b>				
<b>C2 – Inteligencia Empresarial</b>				
<b>OBJETIVOS:</b>				
<b>Objetivos:</b>				
<ul style="list-style-type: none"><li>• Que el alumno logre familiarizarse con los conceptos básicos y fases del proceso de extracción de conocimiento</li><li>• Que el alumno conozca los fundamentos y el alcance de las principales técnicas de minería de datos.</li><li>• Que los alumnos sean capaces de identificar las técnicas y métodos que sean necesarios para la resolución inteligente de problemas de procesamiento de información que lo requieran.</li><li>• Que los alumnos adquieran los elementos conceptuales necesarios para diseñar y conducir el proceso de implementación de sistemas de extracción de conocimiento.</li><li>• Que el alumno experimente el uso de herramientas de extracción del conocimiento basadas en sistemas inteligentes para la aplicación en proyectos de minería de datos.</li></ul>				

<sup>1</sup> FE: Horas dedicadas a la Formación Experimental

<sup>2</sup> RPI: Horas dedicadas a la Resolución de Problemas de Ingeniería

<sup>3</sup> ADyP: Horas dedicadas a las Actividades de Diseño y Proyecto



### **CONTENIDOS MINIMOS:**

Fundamentos de Minería de Datos. Características. Objetivos. Alcances. Fases de un Proyecto de Minería de Datos. Extensiones de la Minería de Datos: Web Mining y Text Mining. Estructura y Diseño de un Data Warehouse. Data Marts. Bases de Datos y Warehouse. Middleware de Data Warehouse. Data Warehouse e Inteligencia de Negocios: uso de Minería de Datos y OLAP.

### **PROGRAMA ANALITICO:**

#### **Unidad I: Introducción a la minería de datos.**

Pirámide Dato-Información-Conocimiento-Sabiduría. Sistemas de Información: transaccionales y analíticos. Inteligencia de Negocios: definición, tipos. Datawarehouse: Bases de Datos y Warehouse. Data Warehouse e Inteligencia de Negocios: uso de Minería de Datos y OLAP.

Fundamentos de Minería de Datos. Características. Objetivos. Alcances. Relación con otras disciplinas. Aplicaciones. Sistemas y herramientas de minería de datos. Etapas del proceso de extracción de conocimiento. Fases del proceso de extracción de conocimiento: fase de integración y recopilación, fase de selección, limpieza y transformación, fase de minería de datos, fase de evaluación e interpretación, fase de difusión, uso y monitorización.

#### **Unidad II: Introducción al Aprendizaje Automatizado**

Aprendizaje automatizado. Concepto. Métodos predictivos y descriptivos. Ejemplos. Campos de aplicación. Aprendizaje de concepto. Ejemplos y atributos. Espacio de hipótesis. Sesgo y varianza. Extensiones de la Minería de Datos: Web Mining y Text Mining

#### **Unidad III: Aprendizaje Supervisado**

Características. Conceptos básicos. Tipos: clasificación, regresión. Funciones de pérdida. Medidas de desempeño. Técnicas de evaluación: conjuntos de entrenamiento y test, validación cruzada, bagging. Métodos de aprendizaje supervisado: método de Bayes, árboles de decisión, métodos basados en reglas, redes neuronales, máquinas de vectores soporte. Técnicas de comparación de métodos.

#### **Unidad IV: Aprendizaje No Supervisado**

Características. Conceptos básicos. Medidas de similaridad y disimilaridad. Agrupamiento. Métodos particionales. Métodos jerárquicos. Métodos basados en densidad. Métodos basados en grafos.

### **ACTIVIDADES PRÁCTICAS:**

1. Introducción a R. (Semana 3)
2. Codificación de algoritmo del vecino más cercano en lenguaje R. (Semana 5)
3. Implementación de árboles de decisión usando R. (Semana 7)
4. Codificación de algoritmo Naive Bayes en lenguaje R. (Semana 9)
5. Algoritmo SVM en R. (Semana 11)
6. Aplicación de técnicas de agrupamiento (clustering) en R. (Semana 13)



Modalidad de la Actividad Práctica	Horas Totales
Formación Experimental (simulación, otros)	20
Resolución de Problemas de Ingeniería	15
Actividades Proyecto y Diseño	--

## **BIBLIOGRAFÍA:**

Título	Autores	Editorial	Año Edic.	Cant. Disp
Data Mining: Practical Machine Learning Tools and Techniques, 4 <sup>th</sup> Ed.	Witten, Ian et al.	Morgan Kaufmann	2017	
Minería de datos Modelos y algoritmos	Gironés Roig, Jordi et. al.	Universitat Oberta de Catalunya (UOC)	2017	
Data Mining: Concepts and Techniques, 3 <sup>rd</sup> . Ed.	Jiawei Han Micheline Kamber Jian Pei	Morgan Kaufmann	2011	
Machine Learning with R, 2nd. Ed.	Brett Lantz	Packt Publishing	2015	
Introduction to Machine Learning, 3 <sup>rd</sup> . ed.	Alpaydin, Ethem	MIT Press	2014	
An Introduction to Statistical Learning with Applications in R	James, G., Witten, D., Hastie, T., Tibshirani, R.	Springer	2013	
Introducción a la minería de datos	Hernández Orallo, J. et al	Pearson	2007	
Machine Learning	Mitchell, Tom	McGraw-Hill	1997	

### **CONDICIONES PARA REGULARIZAR LA ASIGNATURA:**

El alumno deberá cumplimentar no menos del ochenta por ciento (80%) de asistencia a las clases teórico-prácticas. – (según reglamento, alumno, art 19, c).

El alumno deberá aprobar cada uno de los trabajos prácticos. Los trabajos prácticos pueden rehacerse tantas veces como se estime oportuno hasta que el alumno obtenga los resultados deseados.

**EXAMEN FINAL :** se efectuará sobre el programa analítico completo desarrollado por la cátedra en el periodo lectivo correspondiente al de la obtención de la regularidad en la asignatura. ( según reglamento alumno art 28)

**EXAMEN LIBRE:** el alumno deberá inscribirse hasta con diez (10) días hábiles de anticipación a la fecha prevista para el examen y deberá constar la autorización correspondiente de la cátedra. Se establecerá que en fecha previa a la de examen, el alumno asista a la Facultad para la realización de ejercicios relacionados a las actividades prácticas. Obtenido como nota seis (6) o más en el examen práctico, el alumno tendrá el derecho para rendir la parte oral hasta en dos turnos ordinarios consecutivos siguientes