
 <b>UNIVERSIDAD DISTRITAL</b> <b>FRANCISCO JOSÉ DE CALDAS</b>	<b>FORMATO DE SYLLABUS</b>		Código: AA-FR-003	 <small>Sistema Integrado de Gestión</small>
	Macroproceso: Direccionamiento Estratégico		Versión: 01	
	Proceso: Autoevaluación y Acreditación		Fecha de Aprobación: 27/07/2023	

<b>FACULTAD:</b>	<b>Ingeniería</b>		
<b>PROYECTO CURRICULAR:</b>	<b>Ingeniería de Sistemas</b>	<b>CÓDIGO PLAN DE ESTUDIOS:</b>	

I. IDENTIFICACIÓN DEL ESPACIO ACADÉMICO

NOMBRE DEL ESPACIO ACADÉMICO: **Minería de datos**

Código del espacio académico:		Número de créditos académicos:				3	
Distribución horas de trabajo:	HTD	4	HTC	2	HTA	3	
Tipo de espacio académico:	Asignatura	x	Cátedra				

NATURALEZA DEL ESPACIO ACADÉMICO:

Obligatorio Básico		Obligatorio Complementario	x	Electivo Intrínseco		Electivo Extrínseco	
--------------------	--	----------------------------	---	---------------------	--	---------------------	--

CARÁCTER DEL ESPACIO ACADÉMICO:

Teórico		Práctico		Teórico-Práctico	x	Otros:		Cuál: _____
---------	--	----------	--	------------------	---	--------	--	-------------

MODALIDAD DE OFERTA DEL ESPACIO ACADÉMICO:

Presencial	x	Presencial con incorporación de TIC		Virtual		Otros:		Cuál: _____
------------	---	-------------------------------------	--	---------	--	--------	--	-------------

II. SUGERENCIAS DE SABERES Y CONOCIMIENTOS PREVIOS

Probabilidad y estadística  
Conocimientos en programación de computadores

III. JUSTIFICACIÓN DEL ESPACIO ACADÉMICO

En el contexto actual, la minería de datos se constituye como una disciplina esencial para la interpretación y el entendimiento de los grandes volumen de datos provenientes de todas las áreas del conocimiento y las interacciones humanas. Esta asignatura tiene una alta relevancia práctica ya que se alinea con las tendencias actuales y futuras del mercado laboral.

IV. OBJETIVOS DEL ESPACIO ACADÉMICO (GENERAL Y ESPECÍFICOS)

**Objetivo General:**  
Entender los conceptos básicos de la minería de datos: aprender sobre la preparación y limpieza de datos previa a su análisis, aplicar técnicas de minería de datos para la identificación de patrones, utilizar herramientas de software para la minería de datos y evaluar los resultados obtenidos de la aplicación de técnicas de minería de datos.

**Objetivos específicos:**  
\* Capacidad de análisis y para interpretar volúmenes de datos.  
\* Capacidad para aplicar la minería de datos en la resolución de problemas.  
\* Habilidad para presentar resultados del análisis.  
\* Capacidad para trabajar con datos provenientes de diferentes fuentes y en múltiples formatos."  
\* Habilidad para evaluar modelos de minería de datos

V. COMPETENCIAS Y RESULTADOS DE APRENDIZAJE

Competencias	Dominio-Nivel	RA	Resultados de Aprendizaje
Capacidad de análisis y para interpretar volúmenes de datos.	Cognitivo-Intelectual	1	Aplicar técnicas de limpieza y preparación datos.
	Cognitivo-Intelectual	2	Realizar e interpretar análisis exploratorio.
	Cognitivo-Intelectual	3	Comprender los conceptos de la minería de datos.
Capacidad para aplicar la minería de datos en la resolución de problemas.	Cognitivo-Intelectual	4	Identificar cuando usar clasificación, agrupación, o regresión
	Cognitivo-Intelectual	5	Definir el proceso de aplicación de la minería de datos a un problema específico
	Cognitivo-Intelectual	6	Aplicar la reducción de dimensionalidad en problemas con muchas variables

Habilidad para presentar resultados del análisis.	Cognitivo-Intelectual	7	Utilizar herramientas de visualización para identificar patrones y tendencias
	Cognitivo-Intelectual	8	Construir informes finales resultantes del análisis de datos
	Cognitivo-Intelectual	9	Analizar y presentar resultados de procesos de minería
Capacidad para trabajar con datos provenientes de diferentes fuentes y en múltiples formatos.	Cognitivo-Intelectual	10	Obtener e integrar datos de diferentes fuentes
	Cognitivo-Intelectual	11	Obtener e integrar datos en diferentes formatos
	Cognitivo-Intelectual	12	Incorporar textos como fuentes de datos
Habilidad para evaluar modelos de minería de datos	Cognitivo-Intelectual	13	Evaluar clasificadores a partir de métricas.
	Cognitivo-Intelectual	14	Evaluar agrupamientos mediante índices de evaluación.
	Cognitivo-Intelectual	15	Identificar y prevenir overfitting y underfitting.

VI. CONTENIDOS TEMÁTICOS

Revisión histórica y conceptual  
 -Aplicaciones y casos de uso  
 -Evolución de la minería de datos  
 -Revisión conceptual  
 -Proceso de minería: CRISP-DM  
 Recolección y preparación de datos:  
 -Fuentes y métodos  
 -Eliminación de ruido y datos duplicados  
 -Imputación de datos faltantes  
 -Normalización y estandarización  
 Introducción a la clasificación  
 -Árboles de decisión  
 -Clasificación por vecindad  
 -Naive Bayes  
 -Máquinas de soporte vectorial (SVM)  
 -Métricas de evaluación de la clasificación: Matriz de confusión, precisión, recall, F1-score  
 Introducción a la regresión  
 -Regresión lineal  
 -Regresión logística  
 -Regularización: Ridge y Lasso  
 Introducción al agrupamiento  
 -Agrupación por K medias  
 -Algoritmo de agrupamiento jerárquico  
 -Agrupación por DBSCAN  
 -Métricas de evaluación de agrupamientos: Índice de Silhouette y coeficiente de Dunn  
 Introducción a las reglas de asociación  
 -Algoritmo A priori  
 -Algoritmo Eclat  
 -Métricas de evaluación: Confianza, lift y leverage  
 Reducción de dimensionalidad  
 -Análisis de Componentes Principales  
 -Análisis de Correspondencia  
 -Selección de características

VII. ESTRATEGIAS DE ENSEÑANZA QUE FAVORECEN EL APRENDIZAJE

- Asistencia a clases expositivas y de discusión
- Trabajo en Equipo
- Talleres Investigativos: Aprendizaje por descubrimiento (Heurístico)
- Talleres Prácticos: Mapas conceptuales y simulaciones
- Talleres Teóricos: Elaboración y lectura de artículos “científicos” (documentación).

VIII. PROPUESTA DE EVALUACION

Resultados de aprendizaje (RA) a ser evaluados:	Resultados de aprendizaje asociados a las evaluaciones (T: Teórico / P: Práctico)					
	Actividades Entregables	Talleres	Parciales	Informes de proyecto final	Proyecto final	Exposiciones
RA1	X	X				
RA2					X	
RA3			X			

RA4		X			X	
RA5			X	X	X	
RA6		X				
RA7	X					
RA8				X		
RA9				X		
RA10	X	X				
RA11		X				
RA12		X				
RA13			X	X		
RA14	X			X	X	
RA15			X		X	
Tipo de evaluación						
Porcentaje de evaluación (%)	25	20	30	5	25	
Trabajo Individual (I) o Grupal (G)	I,G	G	I	G	G	
Tipo de nota	0-5	0-5	0-5	0-5	0-5	0-5
IX. MEDIOS Y RECURSOS EDUCATIVOS						
<ul style="list-style-type: none"><li>• Salón normal con pizarrón para sesiones de cátedra y para sesiones de discusión.</li><li>• Acceso a Videobeam.</li><li>• Página web para publicar material didáctico, guías de trabajo, talleres, etc.</li><li>• Videos didácticos alrededor de los temas de la asignatura.</li><li>• Ttalleres investigativos y prácticos.</li><li>• Acceso al material bibliográfico recomendado</li></ul>						
X. PRÁCTICAS ACADÉMICAS - SALIDAS DE CAMPO						
No aplica						
XI. BIBLIOGRAFÍA						
<b>Básicas:</b> Data Mining: Concepts and Techniques by Jiawei Han, Micheline Kamber, and Jian Pei Pattern Recognition and Machine Learning by Christopher M. Bishop						
<b>Complementarias:</b> The Elements of Statistical Learning: Data Mining, Inference, and Prediction by Trevor Hastie, Robert Tibshirani, and Jerome Friedman Python Data Science Handbook by Jake VanderPlas Introduction to Machine Learning with Python: A Guide for Data Scientists by Andreas C. Müller and Sarah Guido						
<b>Páginas web</b>						
XII. SEGUIMIENTO Y ACTUALIZACIÓN DEL SYLLABUS						
Fecha revisión por Consejo Curricular:						
Fecha aprobación por Consejo Curricular:				Número de acta:		