



**RESUMEN – TESIS DE GRADO**

AUTORES: LINA MARCELA VARGAS TORRES  
MIGUEL EDUARDO BECERRA NAVARRO

FACULTAD: INGENIERIA

PLAN DE ESTUDIOS: INGENIERIA DE SISTEMAS

DIRECTOR: CARMEN JANETH PARADA

TITULO DE LA TESIS: ANALISIS Y CLASIFICACIÓN DE LOS ALGORITMOS DE LA  
TÉCNICA KNOWLEDGE DISCOVERY FROM DATABASES EN EL PROCESO DE  
MINERIA DE DATOS

RESUMEN:

Se identificaron y analizaron las diferentes técnicas empleadas en el proceso KDD (descubrimiento de conocimiento a partir de bases de datos), lo cual permitió determinar el comportamiento de cada algoritmo en cuanto al costo computacional, tiempo de respuesta, eficiencia y demás factores que caracterizan el funcionamiento de cada uno de ellos. Igualmente se clasificaron los algoritmos dentro de los tipos de tarea (clasificación, segmentación/clustering y análisis de dependencia) de acuerdo a los objetivos o metas que se desean lograr. Por último se compararon de acuerdo a su funcionamiento, metas, tipo de tarea y conocimiento novedoso, que puedan aportar a través del uso de Data Mining, orientadas al análisis inteligente de los datos.

CARACTERÍSTICAS:

PAGINAS: 100

PLANOS:

ILUSTRACIONES:

CD-ROM: 1

ANALISIS Y CLASIFICACIÓN DE LOS ALGORITMOS DE LA TÉCNICA  
KNOWLEDGE DISCOVERY FROM DATABASES EN EL PROCESO DE MINERIA  
DE DATOS

LINA MARCELA VARGAS TORRES  
MIGUEL EDUARDO BECERRA NAVARRO

UNIVERSIDAD FRANCISCO DE PAULA SANTANDER  
FACULTAD DE INGENIERIA  
PLAN DE ESTUDIOS DE INGENIERIA DE SISTEMAS  
SAN JOSÉ DE CUCUTA  
2008

ANALISIS Y CLASIFICACIÓN DE LOS ALGORITMOS DE LA TÉCNICA  
KNOWLEDGE DISCOVERY FROM DATABASES EN EL PROCESO DE MINERIA  
DE DATOS

LINA MARCELA VARGAS TORRES  
MIGUEL EDUARDO BECERRA NAVARRO

Trabajo de grado presentado como requisito para optar al título de  
Ingeniero de Sistemas

Director  
CARMEN JANETH PARADA  
Ingeniero de Sistemas

UNIVERSIDAD FRANCISCO DE PAULA SANTANDER  
FACULTAD DE INGENIERIA  
PLAN DE ESTUDIOS DE INGENIERIA DE SISTEMAS  
SAN JOSÉ DE CUCUTA  
2008



## ACTA DE SUSTENTACION DE UN TRABAJO DE GRADO

FECHA: 29 DE ABRIL DE 2008 HORA: 4:00 p. m.

LUGAR : AUDITORIO "J. J. MALDONADO"- UFPS

PLAN DE ESTUDIOS: INGENIERIA DE SISTEMAS

TITULO DE LA TESIS: "ANALISIS Y CLASIFICACION DE LOS ALGORITMOS DE LA TECNICA KNOWLEDGE DISCOVERY FROM DATABASES EN EL PROCESO DE MINERIA DE DATOS".

JURADOS: EDWARD GILBERTO PUERTO CUADROS  
MILTON JESUS VERA CONTRERAS  
NELSON BELTRAN GALVIS

DIRECTOR : INGENIERA CARMEN JANETH PARADA.

NOMBRE DE LOS ESTUDIANTES:	CODIGO	CALIFICACION	
		NUMERO	LETRA
LINA MARCELA VARGAS TORRES	0152348	3,5	TRES, CINCO
MIGUEL EDUARDO BECERRA NAVARRO	0152252	3,5	TRES, CINCO

# APROBADA

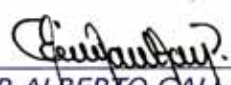
FIRMA DE LOS JURADOS

  
EDWARD GILBERTO PUERTO CUADROS

\_\_\_\_\_  
MILTON JESUS VERA CONTRERAS

  
NELSON BELTRÁN GALVIS

Vo.Bo. \_\_\_\_\_

  
OSCAR ALBERTO GALLARDO PEREZ  
Coordinador Comité Curricular

Betty M.

## CONTENIDO

	Pág.
INTRODUCCION	12
1. MINERIA DE DATOS	15
1.1 DEFINICIÓN	15
1.2 TÉCNICAS DE MINERÍA DE DATOS	16
1.3 EL APRENDIZAJE	18
1.3.1 Aprendizaje supervisado (predicción)	20
1.3.2 Aprendizaje no supervisado (descripción)	20
2. EXTRACCIÓN DE CONOCIMIENTO EN BASES DE DATOS (KDD)	22
2.1 DEFINICIÓN	22
2.2 EL PROCESO	22
2.3 TIPOS DE CONOCIMIENTO	24
2.4 ESTRUCTURA DEL PROCESO DE OBTENCIÓN DEL CONOCIMIENTO, KDD	25
2.4.1 Etapa 1	26

2.4.2 Etapa 2	27
2.4.3 Etapa 3	28
2.4.4 Etapa 4	30
2.4.5 Etapa 5	31
2.4.6 Etapa 6	31
3. ANÁLISIS DE LOS ALGORITMOS DE MINERÍA EN EL PROCESO DE DESCUBRIMIENTO DE CONOCIMIENTO, KDD	33
3.1 ALGORITMO	33
3.1.1 Definición	33
3.1.2 Características	33
3.2 ANÁLISIS	34
3.2.1 Enfoques en el análisis	34
3.2.2 Eficiencia	35
3.2.3 Cálculo de costos de algoritmos	36
3.2.4 Órdenes de complejidad computacional	36
4. CLASIFICACIÓN DE LAS TECNICAS DE DATA MINING	39

4.1 LOS TIPOS DE TAREA	39
4.1.1 Clasificación	39
4.1.2 Regresión	40
4.1.3 Asociación	40
4.1.4 Agrupamiento (clustering) / segmentación	41
4.2 METAS	41
4.2.1 Descriptivas	41
4.2.2 Predictivas	42
4.3 PRODUCTOS MINERÍA DE DATOS COMERCIALES Y NO COMERCIALES	43
5. CONCLUSIONES	45
6. RECOMENDACIONES	47
BIBLIOGRAFÍA	48
ANEXOS	49