

Índice general

1. Introducción	1
1.1. Descripción de la problemática	2
1.2. Objetivos del trabajo	2
1.2.1. Objetivo general	2
1.2.2. Objetivos específicos	2
1.3. Contribución de la tesis	2
2. Artículo	
Análisis comparativo de técnicas de machine learning para predecir la deserción de estudiantes en varios niveles de estudios.	4
2.1. Introducción	5
2.1.1. Relevancia para la gestión de operaciones	7
2.2. Marco teórico	9
2.3. Metodología	12
2.3.1. Metodología <i>Data Mining (DM)</i>	12
2.3.2. Descripción de los datos	13
2.3.3. Medición del desempeño	17
2.3.4. Resumen de medidas de desempeño	18
2.4. Resultados	19
2.4.1. Information gain	19
2.4.2. Resultados modelos	20
2.4.3. Discusión de resultados	23
2.5. Conclusiones	26
Bibliografía	28
A. Código fuente del algoritmo utilizado para conformar el <i>Data Warehouse</i>	32

Índice de figuras

2.1. Modelos propuestos para predecir deserción.	6
2.2. Modelo propuesto de deserción por niveles y <i>Dashboard</i> propuesto.	7
2.3. Proceso <i>KDD</i>	12
2.4. Ejemplo Validación Cruzada con 10 particiones.	13
2.5. Dashboard - Matrículas de la Universidad Católica del Maule	14

Índice de tablas

2.1. Resumen de literatura consultada.	11
2.2. Atributos más representativos del <i>Data WareHouse (DW)</i>	13
2.3. Atributos seleccionados de las bases de datos.	16
2.4. Matriz de confusión y las medidas de desempeño más comunes a partir de ella.	18
2.5. Ranking de atributos utilizando Information Gain para los datasets dados.	20
2.6. Resultados para el modelo global previo al balance de clases.	21
2.7. Resultados para el modelo global con clases balanceadas.	21
2.8. Resultados modelo primer nivel.	22
2.9. Resultados modelo segundo nivel.	22
2.10. Resultados modelo tercer nivel.	23