

Asignatura: CI5128 – Ingeniería del conocimiento**Cuatrimestre: 2º****Examen: Parcial****Convocatoria: Ordinaria****Grupo: SIT1****Curso: 2006/2007****Fecha: 15-3-2007**

Una compañía de seguros posee la siguiente tabla en la que se relaciona el riesgo de accidente grave con las características de diferentes modelos. Nos han contratado para construir un modelo que prediga si se debe hacer o no la póliza a un cliente dado (según el coche que posea)

Marca	Combustible	Aspiración	Puertas	Tracción	Precio	Riesgo
Peugeot	gasolina	normal	4	trasera	medio	alto
Peugeot	diesel	turbo	4	trasera	medio	medio
Peugeot	gasolina	normal	4	trasera	medio	medio
Renault	gasolina	normal	4	trasera	alto	alto
Renault	diesel	turbo	4	trasera	alto	bajo
Renault	gasolina	normal	4	delantera	bajo	medio
Renault	gasolina	normal	2	delantera	bajo	bajo
Volkswagen	gasolina	normal	4	delantera	bajo	alto
Volkswagen	gasolina	normal	4	delantera	bajo	medio
Volkswagen	diesel	turbo	4	delantera	bajo	medio
Volkswagen	gasolina	normal	4	delantera	bajo	medio
Volkswagen	gasolina	normal	2	delantera	medio	bajo

1. **(4 puntos)** Aplicar el algoritmos de eliminación de candidatos (o el espacio de versiones) a los ejemplos 1, 5, 7 y 12. Considera sólo para este caso que los coches que tiene riesgo alto son ejemplos negativos y que los de medio y bajo son positivos. ¿Qué concepto se ha aprendido?
2. **(2 puntos)** ¿Cual sería el **error estimado** de un algoritmo genérico que, sobre un conjunto de test de 100 ejemplos, comete 18 errores (con un 95 % de confianza, siendo para este intervalo $z=1,96$)?
3. **(4 puntos)** Obtener **solamente** el primer nodo del árbol mediante el algoritmo C4.5.