

## Algoritmo de aprendizaje para árbol de clasificación

**Miguel Lentisco Ballesteros**

Aconsejamos las ramas a cada alumno reflejada en la siguiente tabla:

Caso	Gusta Mates	Quiere trabajar	Nota media	Gusta hardware	Como Trabaja	Gusta programar	Rama(s) aconsejadas
Alumno 1	Si	Docencia	Alta	No	Mucho	Si	CSI
Alumno 2	No	Empresa Pública	Media	No	Normal	No	SI
Alumno 3	Si	Empresa Privada	Media	Si	Normal	Si	CSI, IC
Alumno 4	No	Empresa Privada	Baja	No	Poco	Si	TI
Alumno 5	No	Empresa Privada	Alta	Si	Mucho	Si	IC
Alumno 6	Si	Docencia	Media	No	Poco	Si	IS
Alumno 7	No	Docencia	Alta	No	Normal	Si	IS
Alumno 8	No	Empresa Pública	Baja	Si	Normal	No	IC
Alumno 9	Si	Empresa Privada	Alta	No	Normal	Si	CSI
Alumno 10	No	Empresa Pública	Baja	No	Poco	Si	TI
Alumno 11	Si	Empresa Privada	Alta	Si	Normal	No	CSI, IC
Alumno 12	No	Empresa Privada	Baja	Si	Poco	Si	IC
Alumno 13	No	Empresa Privada	Alta	Si	Normal	Si	IC
Alumno 14	Si	Docencia	Media	No	Mucho	Si	CSI
Alumno 15	No	Docencia	Media	No	Normal	No	TI
Alumno 16	No	Empresa Pública	Media	Si	Poco	No	SI

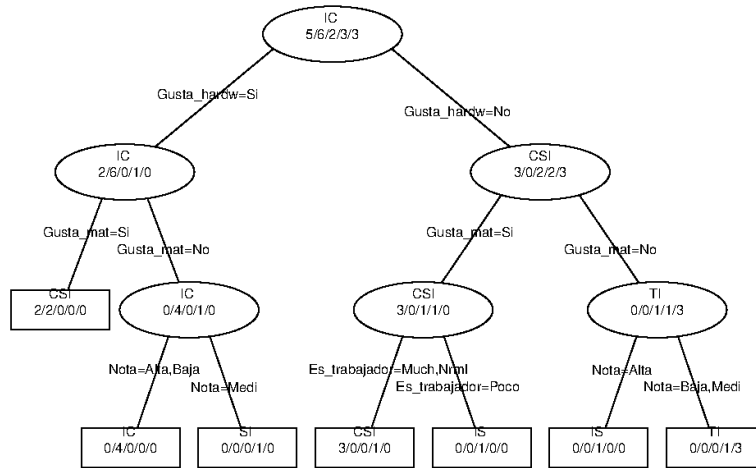
Consideramos con varias recomendaciones al mismo alumno como ejemplos

separados, y agrupamos según la frecuencia de cada atributo en cada rama:

<b>Atributos</b>	<b>Valores</b>	<b>CSI</b>	<b>IC</b>	<b>IS</b>	<b>SI</b>	<b>TI</b>
<i>Gusta matemáticas</i>	Sí	5	2	1	0	0
	No	0	4	1	2	3
<i>Quiere trabajar</i>	Docencia	2	0	2	0	1
	Pública	0	1	0	2	1
	Privada	3	5	0	0	1
<i>Nota media</i>	Alta	3	3	1	0	0
	Media	2	1	1	2	1
	Baja	0	2	0	0	2
<i>Gusta hardware</i>	Sí	2	6	0	1	0
	No	3	0	2	1	3
<i>Es trabajador</i>	Mucho	2	2	0	0	0
	Normal	3	3	1	1	1
	Poco	0	1	1	1	2
<i>Gusta programar</i>	Sí	5	4	2	0	2
	No	0	2	0	2	1

Ahora usamos un algoritmo de aprendizaje basado en entropía (ID3) de manera que obtenemos el siguiente árbol de decisión:

### Árbol de Clasificación



Finalmente obtenemos las reglas:

- IF (Gusta\_hardw=Si AND Gusta\_mat=SI) OR (Gusta\_hardw=No AND Gusta\_mat=Si AND (Es\_trabajador=Much OR Es\_trabajador=Nrml)) THEN **CSI**
- IF Gusta\_hardw=Si AND Gusta\_mat=No AND (Nota=Alta OR Nota=Baja) THEN **IC**
- IF Gusta\_hardw=No AND ((Gusta\_mat=Si AND Es\_trabajador=Poco) OR (Gusta\_mat=No AND Nota=Alta)) THEN **IS**
- IF Gusta\_hardw=Si AND Gusta\_mat=No AND Nota=Medi THEN **SI**
- IF Gusta\_hardw=No AND Gusta\_mat=No AND (Nota=Baja OR Nota=Medi) THEN **TI**

Que son válidas.