

Solemne 1 - Sistemas Inteligentes (Pauta)

miércoles 27 de abril de 2015

Profesor: Alejandro Figueroa

Ayudante: Alexander Espina

- Está prohibido el uso de teléfonos celulares durante el desarrollo de la prueba.
- La prueba debe responderse con un lápiz de tinta indeleble, de lo contrario no hay opción a correcciones.
- Cualquier alumno que sea sorprendido intentando copiar será calificado con una nota 1.
- Está prohibido conversar durante la prueba. Recuerde que su compañero puede estar concentrado y el ruido puede perturbarlo en el desarrollo de su prueba.
- Utilice sólo las hojas entregadas para escribir sus respuestas.

Pregunta 1 (60 puntos)

Considera la siguiente salida de tres clasificadores:

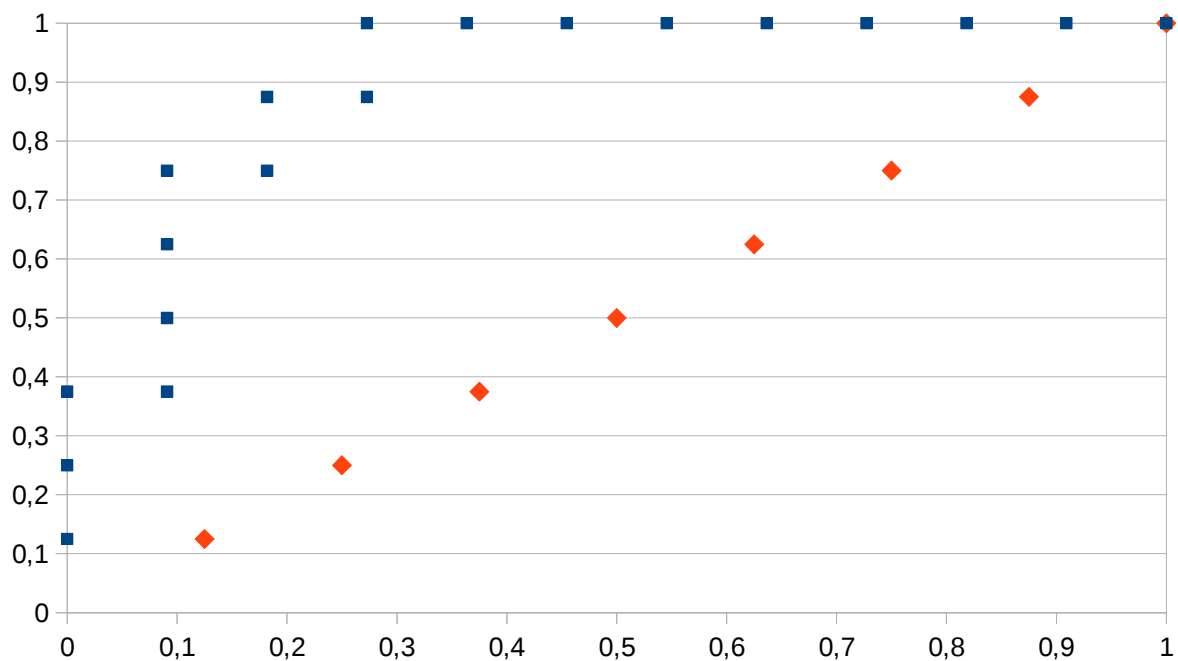
Etiqueta	Clasificador 1			Clasificador 2			Clasificador 3		
1	0,5247	0,0032	0,4721	0,5913	0,1199	0,2887	0,558	0,06155	0,3804
1	0,6109	0,2898	0,0993	0,7214	0,1998	0,0788	0,66615	0,2448	0,08905
2	0,0779	0,8945	0,0276	0,0389	0,959	0,0021	0,0584	0,92675	0,01485
2	0,2745	0,7095	0,016	0,2958	0,7	0,0041	0,28515	0,70475	0,01005
1	0,9563	0,0425	0,0011	0,8407	0,0798	0,0795	0,8985	0,06115	0,0403
1	0,6132	0,2568	0,1301	0,4292	0,492	0,0788	0,5212	0,3744	0,10445
1	0,8093	0,169	0,0217	0,66	0,1944	0,1456	0,73465	0,1817	0,08365
2	0,0035	0,9953	0,0012	0,0389	0,959	0,0021	0,0212	0,97715	0,00165
2	0,8687	0,0915	0,0398	0,8991	0,0556	0,0453	0,8839	0,07355	0,04255
2	0,3671	0,4741	0,1588	0,1743	0,8127	0,013	0,2707	0,6434	0,0859
1	0,9917	0,0022	0,006	0,9917	0,0012	0,0071	0,9917	0,0017	0,00655
2	0,412	0,4235	0,1645	0,2601	0,6592	0,0807	0,33605	0,54135	0,1226
3	0,5134	0,0126	0,4739	0,7069	0,0967	0,1964	0,61015	0,05465	0,33515
3	0,0181	0,0062	0,9757	0,2494	0,2085	0,5422	0,13375	0,10735	0,75895
1	0,8783	0,0496	0,072	0,9329	0,0311	0,036	0,9056	0,04035	0,054
1	0,8064	0,1671	0,0265	0,8751	0,0747	0,0502	0,84075	0,1209	0,03835
2	0,6169	0,1806	0,2025	0,4609	0,1907	0,3484	0,5389	0,18565	0,27545
2	0,0779	0,8945	0,0276	0,0389	0,959	0,0021	0,0584	0,92675	0,01485
2	0,0035	0,9953	0,0012	0,0389	0,959	0,0021	0,0212	0,97715	0,00165

Considerando que los puntajes entregados por estos clasificadores son una distribución de probabilidad. Calcule:

- el MRR para cada uno de ellos. (10 puntos)
- el accuracy para cada uno de ellos. (10 puntos)
- La Precisión, el Recall y el F-Score considerando interpretando cada una de las clases como positiva. (15 puntos, 5 c/u)
- Grafique la Curva ROC considerando la clase 1 como positiva y el tercer clasificador. (10 puntos)
- Para el punto d), calcule el AUC. (10 puntos)
- ¿Qué observa del clasificador 3? (5 puntos)

	Etiqueta	Clasificador 1			Clasificador 2			Clasificador 3					
Cant clase 1	1	0,5247	0,0032	0,4721	0,5913	0,1199	0,2887	0,558	0,06155	0,3804			
	1	0,6109	0,2898	0,0993	0,7214	0,1998	0,0788	0,66615	0,2448	0,08905			
	2	0,0779	0,8945	0,0276	0,0389	0,959	0,0021	0,0584	0,92675	0,01485			
	2	0,2745	0,7095	0,016	0,2958	0,7	0,0041	0,28515	0,70475	0,01005			
	1	0,9563	0,0425	0,0011	0,8407	0,0798	0,0795	0,8985	0,06115	0,0403			
	1	0,6132	0,2568	0,1301	0,4292	0,492	0,0788	0,5212	0,3744	0,10445			
	1	0,8093	0,169	0,0217	0,66	0,1944	0,1456	0,73465	0,1817	0,08365			
	2	0,0035	0,9953	0,0012	0,0389	0,959	0,0021	0,0212	0,97715	0,00165			
	2	0,8687	0,0915	0,0398	0,8991	0,0556	0,0453	0,8839	0,07355	0,04255			
	2	0,3671	0,4741	0,1588	0,1743	0,8127	0,013	0,2707	0,6434	0,0859			
	1	0,9917	0,0022	0,006	0,9917	0,0012	0,0071	0,9917	0,0017	0,00655			
	2	0,412	0,4235	0,1645	0,2601	0,6592	0,0807	0,33605	0,54135	0,1226			
	3	0,5134	0,0126	0,4739	0,7069	0,0967	0,1964	0,61015	0,05465	0,33515			
	3	0,0181	0,0062	0,9757	0,2494	0,2085	0,5422	0,13375	0,10735	0,75895			
	1	0,8783	0,0496	0,072	0,9329	0,0311	0,036	0,9056	0,04035	0,054			
	1	0,8064	0,1671	0,0265	0,8751	0,0747	0,0502	0,84075	0,1209	0,03835			
	2	0,6169	0,1806	0,2025	0,4609	0,1907	0,3484	0,5389	0,18565	0,27545			
	2	0,0779	0,8945	0,0276	0,0389	0,959	0,0021	0,0584	0,92675	0,01485			
	2	0,0035	0,9953	0,0012	0,0389	0,959	0,0021	0,0212	0,97715	0,00165			
	Cant clase 1	8											
Cant clase 2	9												
Cant clase 3	2												
		C1	C2	C3	PC1	PC2	PC3	RR1	RR2	RR3	ACC1	ACC2	ACC3
		1	1	1	1	1	1	1	1	1	0	0	0
		1	1	1	1	1	1	1	1	1	0	0	0
		2	2	2	1	1	1	1	1	1	0	0	0
		2	2	2	1	1	1	1	1	1	0	0	0
		1	1	1	1	1	1	1	1	1	0	0	0
		1	2	1	1	2	1	1	0,5	1	0	1	0
		1	1	1	1	1	1	1	1	1	0	0	0
		2	2	2	1	1	1	1	1	1	0	0	0
		1	1	1	2	2	2	0,5	0,5	0,5	1	1	1
		2	2	2	1	1	1	1	1	1	0	0	0
		1	1	1	1	1	1	1	1	1	0	0	0
		2	2	2	1	1	1	1	1	1	0	0	0
		1	1	1	2	2	2	0,5	0,5	0,5	1	1	1
		3	3	3	1	1	1	1	1	1	0	0	0
		1	1	1	1	1	1	1	1	1	0	0	0
		1	1	1	1	1	1	1	1	1	0	0	0
		1	1	1	3	3	3	0,33	0,333	0,333	1	1	1
		2	2	2	1	1	1	1	1	1	0	0	0
		2	2	2	1	1	1	1	1	1	0	0	0
Cant clase 1		11	10	11			MRR	0,91	0,886	0,912	0,8421	0,7895	0,8421
Cant clase 2		7	8	7									
Cant clase 3		1	1	1									
Precision c1		0,727	0,7	0,727									
Precision c2		1	0,875	1									
Precision c3		1	1	1									
Recall c1		1	0,875	1									
Recall c2		0,778	0,778	0,778									
Recall c3		0,5	0,5	0,5									
Fscore c1		0,842	0,778	0,842									
Fscore c2		0,875	0,824	0,875									
Fscore c3		0,667	0,667	0,667									

		TPR	FNR
1	0,9917	0,125	0
1	0,9056	0,25	0
1	0,8985	0,375	0
0	0,8839	0,375	0,09091
1	0,84075	0,5	0,09091
1	0,73465	0,625	0,09091
1	0,66615	0,75	0,09091
0	0,61015	0,75	0,18182
1	0,558	0,875	0,18182
0	0,5389	0,875	0,27273
1	0,5212	1	0,27273
0	0,33605	1	0,36364
0	0,28515	1	0,45455
0	0,2707	1	0,54545
0	0,13375	1	0,63636
0	0,0584	1	0,72727
0	0,0584	1	0,81818
0	0,0212	1	0,90909
0	0,0212	1	1



$AUC = 1 - 0,09091 * (1 - 0,375) - (0,18182 - 0,09091) * (1 - 0,75) - (0,27273 - 0,18182) * (1 - 0,875) = 0,90909$

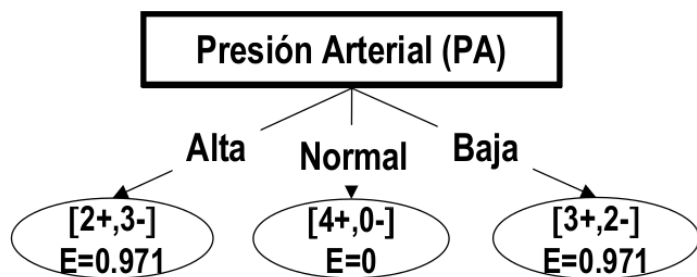
El tercer clasificador es una combinación lineal de los dos anteriores.

Pregunta 2 (40 puntos)

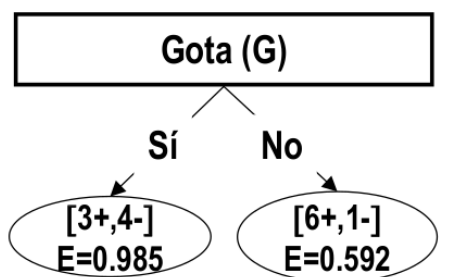
De acuerdo a lo visto en clases y ayudantía, construya el árbol de decisión para el siguiente problema.

Paciente	Presión Arterial	Urea en sangre	Gota	Hipotiroidismo	Administrar Tratamiento
1	Alta	Alta	Sí	No	No
2	Alta	Alta	Sí	Sí	No
3	Normal	Alta	Sí	No	Sí
4	Baja	Normal	Sí	No	Sí
5	Baja	Baja	No	No	Sí
6	Baja	Baja	No	Sí	No
7	Normal	Baja	No	Sí	Sí
8	Alta	Normal	Sí	No	No
9	Alta	Baja	No	No	Sí
10	Baja	Normal	No	No	Sí
11	Alta	Normal	No	Sí	Sí
12	Normal	Normal	Sí	Sí	Sí
13	Normal	Alta	No	No	Sí
14	Baja	Normal	Sí	Sí	No

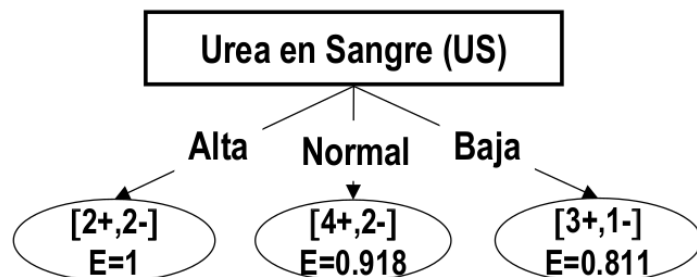
Fuente: https://www.nebrija.es/~cmalagon/inco/apuntes_mios/ejemplo_ID3_clase.pdf



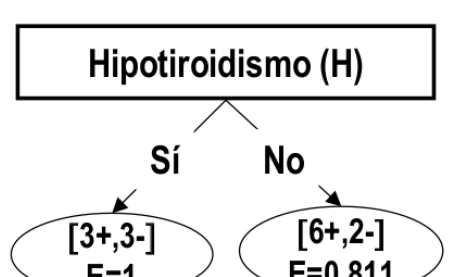
$$G(S, "PA") = 0.940 - (5/14) \cdot 0.971 - (4/14) \cdot 0 - (5/14) \cdot 0.971 = 0.246$$



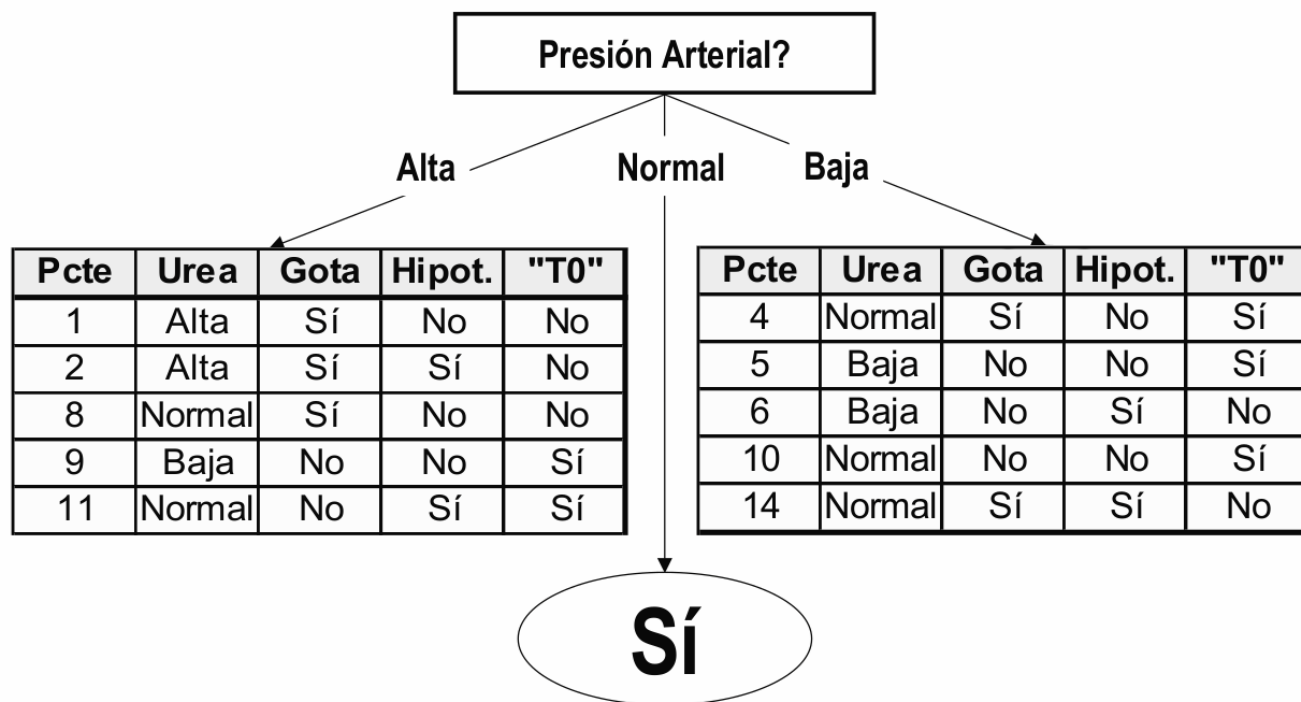
$$G(S, "G") = 0.940 - (7/14) \cdot 0.985 - (7/14) \cdot 0.592 = 0.151$$



$$G(S, "US") = 0.940 - (4/14) \cdot 1 - (6/14) \cdot 0.918 - (4/14) \cdot 0.811 = 0.029$$



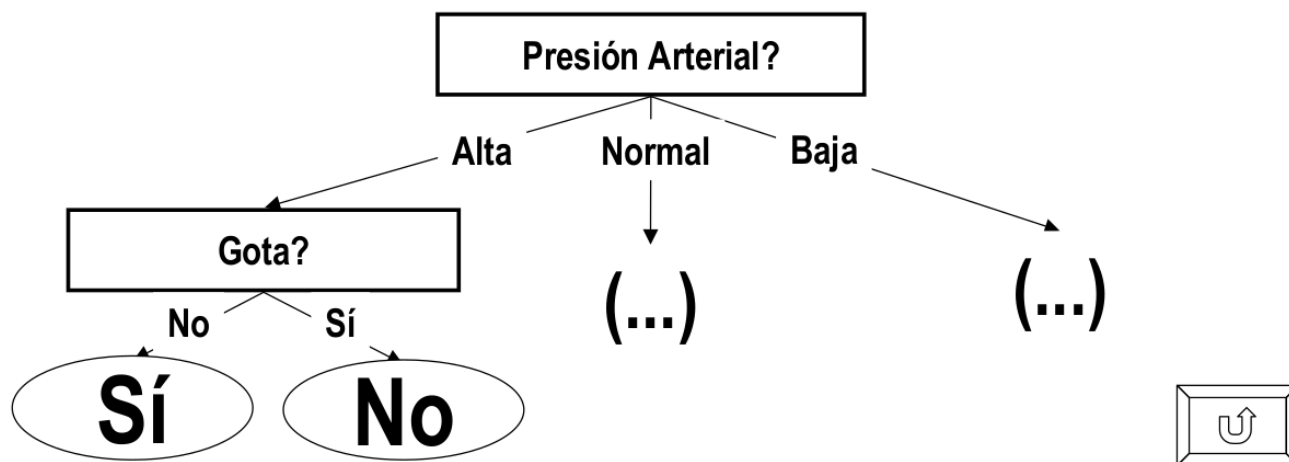
$$G(S, "H") = 0.940 - (6/14) \cdot 1 - (8/14) \cdot 0.811 = 0.048$$



$$G(S, "U") = E[(2+, 3-)] - \frac{2}{5}E[(0+, 2-)] - \frac{2}{5}E[(1+, 1-)] - \frac{1}{5}E[(1+, 0-)] = 0.571$$

$$G(S, "G") = E[(2+, 3-)] - \frac{3}{5}E[(0+, 3-)] - \frac{2}{5}E[(2+, 0-)] = 0.971 \quad \leftarrow$$

$$G(S, "H") = E[(2+, 3-)] - \frac{2}{5}E[(1+, 1-)] - \frac{3}{5}E[(1+, 2-)] = 0.020$$



$$G(S, "U") = E[(3+, 2-)] - \frac{3}{5}E[(2+, 1-)] - \frac{2}{5}E[(1+, 1-)] = 0.020$$

$$G(S, "G") = E[(3+, 2-)] - \frac{2}{5}E[(1+, 1-)] - \frac{3}{5}E[(2+, 1-)] = 0.020$$

$$G(S, "H") = E[(3+, 2-)] - \frac{2}{5}E[(0+, 2-)] - \frac{3}{5}E[(3+, 0-)] = 0.971 \quad \leftarrow$$

