

PRÁCTICA 2. HIPERPLANOS

PARTE OBTATIVA

JAVIER RIVILLA ARREDONDO

Práctica 2. Hiperplanos (Parte opcional)

Al igual que en los clasificadores basados en la práctica anterior vamos a utilizar ahora hiperplanos y comparar entre ambos los resultados obtenidos.

Empezaré explicando el número de iteraciones realizadas para ello usaremos los siguientes datos a completar:

En común, todos estos hiperplanos han sido probados con 8 iteraciones y un 70/30 (Aprendizaje/test).

- 100 hiperplanos.

Iteración	Aciertos (Aprendizaje)	Errores (Aprendizaje)	F.P. (Aprendizaje)	Aciertos (Test)	Errores (Test)	F.P. (Test)	V.P (Test)
1	0,584	28	110	0,563	21	110	0,3
2	0,596	29	121	0,537	18	121	0,4
3	0,639	40	108	0,59	15	108	0,5
4	0,624	40	103	0,603	16	103	0,467
5	0,64	37	109	0,583	16	109	0,467
6	0,637	32	113	0,57	16	113	0,467
7	0,631	34	112	0,57	17	112	0,433
8	0,656	35	109	0,577	18	109	0,4

Resumen de las estadísticas:

Tiempo requerido para el aprendizaje	1.841 s
Clasificadores utilizados	800(8*100)
Tasa de aciertos (APRENDIZAJE)	65.57143%
Tasa de aciertos (TEST)	57.666664%
Imágenes: 1000	1000
Imágenes para aprender	700(70.0%)
Imágenes de test	300(30.0%)
Comentarios:	Debemos saber qué estamos realizando simplemente 8 iteraciones, como en el caso de los clasificadores daban los porcentajes mayores.

- 2000 hiperplanos.

Iteración	Aciertos (Aprendizaje)	Errores (Aprendizaje)	F.P. (Aprendizaje)	Aciertos (Test)	Errores (Test)	F.P. (Test)	V.P (Test)
1	0,614	36	109	0,59	14	109	0,576
2	0,647	45	107	0,573	21	107	0,364
3	0,684	46	99	0,6	21	99	0,364
4	0,706	43	86	0,657	17	86	0,485
5	0,719	46	84	0,653	20	84	0,394
6	0,736	45	81	0,667	19	81	0,424
7	0,741	43	92	0,63	19	92	0,424
8	0,75	45	90	0,637	19	90	0,424

Tiempo requerido para el aprendizaje	31.808 s
Clasificadores utilizados	16000(8*2000)
Tasa de aciertos (APRENDIZAJE)	75.0%
Tasa de aciertos (TEST)	63.666664%
Imágenes: 1000	1000
Imágenes para aprender	700(70.0%)
Imágenes de test	300(30.0%)
Comentarios:	Esta prueba es muy interesante, si te fijas bien al aumentar el número de hiperplanos el tiempo aumenta considerablemente, pero obtenemos mejores resultados. Por esta razón estoy utilizando 8 iteraciones porque si uso más puede tardar más.

- 3000 hiperplanos.

Iteración	Aciertos (Aprendizaje)	Errores (Aprendizaje)	F.P. (Aprendizaje)	Aciertos (Test)	Errores (Test)	F.P. (Test)	V.P (Test)
1	0,596	32	131	0,537	8	131	0,636
2	0,614	29	102	0,623	11	102	0,5
3	0,637	38	107	0,59	16	107	0,273
4	0,651	40	93	0,64	15	93	0,318
5	0,657	40	97	0,627	15	97	0,318
6	0,684	37	93	0,647	13	93	0,409
7	0,699	35	84	0,667	16	84	0,273
8	0,719	39	83	0,663	18	83	0,182

Tiempo requerido para el aprendizaje: 47.674 s
Clasificadores utilizados: 24000(8*3000)
Tasa de aciertos (APRENDIZAJE): 71.85714%
Tasa de aciertos (TEST): 66.333336%
Imágenes: 1000
Imágenes para aprender: 700(70.0%)
Imágenes de test: 300(30.0%)

- 4000 hiperplanos.

Iteración	Aciertos (Aprendizaje)	Errores (Aprendizaje)	F.P. (Aprendizaje)	Aciertos (Test)	Errores (Test)	F.P. (Test)	V.P (Test)
1	0,607	47	101	0,617	14	101	0,481
2	0,627	50	100	0,617	15	100	0,444
3	0,659	51	88	0,65	17	88	0,37
4	0,677	52	93	0,63	18	93	0,333
5	0,71	49	80	0,673	18	80	0,333
6	0,721	49	81	0,657	22	81	0,185
7	0,746	47	75	0,683	20	75	0,259
8	0,747	46	77	0,673	21	77	0,222

Tiempo requerido para el aprendizaje: 63.211 s
Clasificadores utilizados: 32000(8*4000)
Tasa de aciertos (APRENDIZAJE): 74.71429%
Tasa de aciertos (TEST): 67.333336%
Imágenes: 1000
Imágenes para aprender: 700(70.0%)
Imágenes de test: 300(30.0%)

- 5000 hiperplanos.

Iteración	Aciertos (Aprendizaje)	Errores (Aprendizaje)	F.P. (Aprendizaje)	Aciertos (Test)	Errores (Test)	F.P. (Test)	V.P (Test)
1	0,599	37	118	0,553	16	118	0,407
2	0,644	40	105	0,597	16	105	0,407
3	0,669	51	87	0,65	18	87	0,333
4	0,683	48	82	0,663	19	82	0,296
5	0,679	47	91	0,633	19	91	0,296
6	0,701	45	83	0,657	20	83	0,259
7	0,723	46	84	0,647	22	84	0,185
8	0,74	45	77	0,677	20	77	0,259

Tiempo requerido para el aprendizaje: 79.123 s
Clasificadores utilizados: 40000(8*5000)
Tasa de aciertos (APRENDIZAJE): 74.0%
Tasa de aciertos (TEST): 67.666664%
Imágenes: 1000
Imágenes para aprender: 700(70.0%)
Imágenes de test: 300(30.0%)

- 6000 hiperplanos.

Iteración	Aciertos (Aprendizaje)	Errores (Aprendizaje)	F.P. (Aprendizaje)	Aciertos (Test)	Errores (Test)	F.P. (Test)	V.P (Test)
1	0,611	36	114	0,56	18	114	0,4
2	0,661	29	106	0,587	18	106	0,4
3	0,689	35	99	0,6	21	99	0,3
4	0,71	33	84	0,65	21	84	0,3
5	0,713	38	79	0,673	19	79	0,367
6	0,749	36	71	0,697	20	71	0,333
7	0,756	33	68	0,707	20	68	0,333
8	0,743	33	55	0,757	18	55	0,4

Tiempo requerido para el aprendizaje: 57.86 s
Clasificadores utilizados: 48000(8*6000)
Tasa de aciertos (APRENDIZAJE): 74.28571%
Tasa de aciertos (TEST): 75.666664%
Imágenes: 1000
Imágenes para aprender: 700(70.0%)
Imágenes de test: 300(30.0%)

- 7000 hiperplanos.

Iteración	Aciertos (Aprendizaje)	Errores (Aprendizaje)	F.P. (Aprendizaje)	Aciertos (Test)	Errores (Test)	F.P. (Test)	V.P (Test)
1	0,621	44	101	0,597	20	101	0,394
2	0,623	50	101	0,613	15	101	0,545
3	0,67	42	91	0,633	19	91	0,424
4	0,667	48	79	0,677	18	79	0,455
5	0,703	43	81	0,677	16	81	0,515
6	0,694	43	84	0,66	18	84	0,455
7	0,704	46	76	0,69	17	76	0,485
8	0,731	46	76	0,68	20	76	0,394

Tiempo requerido para el aprendizaje: 66.861 s
Clasificadores utilizados: 56000(8*7000)
Tasa de aciertos (APRENDIZAJE): 73.14285%
Tasa de aciertos (TEST): 68.0%
Imágenes: 1000
Imágenes para aprender: 700(70.0%)
Imágenes de test: 300(30.0%)

En conclusión, podemos observar qué cuanto más aumentamos el número de hiperplanos más tiempo tarda, pero los resultados que obtenemos son mucho mejores que los anteriores.

Ahora bien, vamos a comprobar como irá en caso de hacer mas de 8 iteraciones, vamos a utilizar 2000 hiperplanos y 25 iteraciones.

Iteración	Aciertos (Aprendizaje)	Errores (Aprendizaje)	F.P. (Aprendizaje)	Aciertos (Test)	Errores (Test)	F.P. (Test)	V.P. (Test)
1	0,594	32	120	0,53	21	120	0,382
2	0,633	32	95	0,617	20	95	0,412
3	0,639	34	96	0,623	17	96	0,5
4	0,657	33	98	0,6	22	98	0,353
5	0,674	31	90	0,617	25	90	0,265
6	0,686	31	88	0,62	26	88	0,235
7	0,703	34	88	0,623	25	88	0,265
8	0,719	35	83	0,643	24	83	0,294
9	0,729	38	75	0,667	25	75	0,265
10	0,739	36	74	0,673	24	74	0,294
11	0,754	37	76	0,657	27	76	0,206
12	0,757	37	65	0,693	27	65	0,206
13	0,773	38	65	0,697	26	65	0,235
14	0,781	36	69	0,68	27	69	0,206
15	0,771	39	64	0,69	29	64	0,147
16	0,779	38	56	0,717	29	56	0,147
17	0,811	40	54	0,723	29	54	0,147
18	0,806	43	50	0,737	29	50	0,147
19	0,809	43	48	0,743	29	48	0,147
20	0,824	42	49	0,733	31	49	0,088
21	0,833	37	45	0,747	31	45	0,088
22	0,843	41	44	0,75	31	44	0,088
23	0,836	44	47	0,743	30	47	0,118
24	0,837	44	44	0,753	30	44	0,118
25	0,837	42	42	0,76	30	42	0,118

Como podemos observar a simple vista en los datos, los falsos positivos decremantan por cada iteración por lo que afecta positivamente a nuestro rendimiento. Además, a continuación, vamos a poner una comparación entre tiempos de ambos, tanto hiperplanos como clasificadores con umbrales.

Los tiempos que hemos obtenido anteriormente más los que obtuvimos en la práctica anterior son los siguientes:

Hiper/Clasificadores	Hiperplanos(t)	Clasificadores(t)
100	1,82 s	1,02 s
2000	31,8 s	10,62s
3000	47,8 s	11,35 s
4000	63,5 s	15,44s
5000	78,7 s	17,9s

6000	96,5 s	21,45 s
7000	110,1 s	23,52 s

Como podemos comprobar los resultados obtenidos por hiperplanos son mucho más tardíos que los obtenidos por clasificadores, no obstante, los resultados obtenidos por hiperplanos son mucho mejores que los obtenidos por clasificadores débiles.

La conclusión a la que yo llego es que como el tiempo realmente se consume principalmente enfocado al entrenamiento, será mejor utilizar en este caso hiperplanos ya que realmente vamos a adquirir mejores resultados a la hora de entrenar que haciéndolo por clasificadores débiles.