APR: Práctica sobre Mixtura de gaussianas

Roselló Beneitez, N. U.; Roselló Oviedo, M.

6 de Enero de 2020

Índice

Índice de figuras

1. Descripción de la práctica

En esta práctica se ha realizado una clasificación para la base de datos de dígitos *MNIST* mediante diversas técnicas: un clasificador multinomial, un clasificador gaussiano y un clasificador basado en mixtura de gaussianas.

Tanto el clasificador multinomial como el clasificador gaussiano se nos proporcionaban ya implementados, por lo que solo ha bastado ejecutar estos programas para obtener los resultados (véase “Resultados obtenidos”).

En el modelo de clasificador basado en mixtura de gaussianas, se ha empleado el algoritmo *EM* (Esperanza - Maximización) para el aprendizaje de las muestras por cada componente. Para ello se ha tenido que implementar únicamente el paso *M*, puesto que el paso E ya se daba completo.

2. Un pequeño ejercicio completo de aprendizaje

2.1. Ejercicio 4.1

Cuando el valor del suavizado (α) se mantiene constante e igual a 1, los resultados son los siguientes:

PCA	% error	IC (95 %)
10	10.82	±0.6088%
20	4.96	±0.4255%
30	3.91	±0.3799%
40	3.75	±0.3724%
50	3.64	±0.3671%
60	3.88	±0.3785%
70	3.85	±0.3771%
80	4	±0.3841%
90	4.1	±0.3886%
100	4.3	±0.3976%

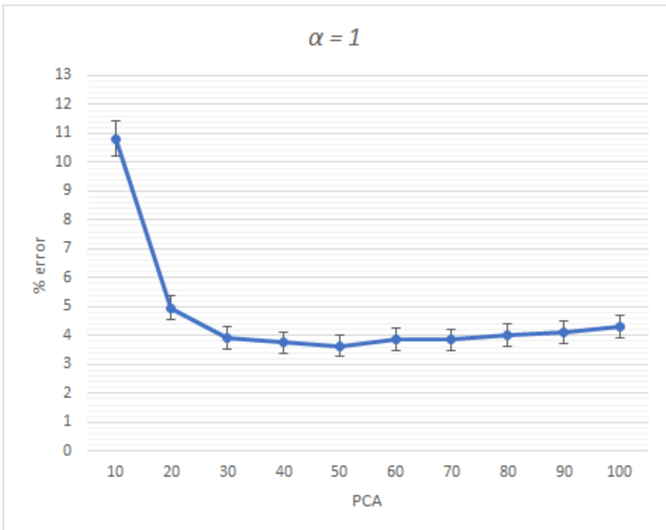

Figura 1: Error e IC 95 % en función de PCA para $\alpha = 1$

2.2. Ejercicio 4.2

Se ha experimentado con valores del suavizado en el conjunto 0.001, 0.01, 0.1, 0.2, 0.5, 0.9, 0.95, 0.99. Los resultados se muestran a continuación:

% error		α								
		0.001	0.01	0.1	0.2	0.5	0.9	0.95	0.99	1
PCA	10	11.77	11.92	11.73	11.49	10.97	10.82	10.82	10.81	10.82
	20	6.12	6.52	6.29	5.99	5.26	4.97	4.98	4.98	4.96
	30	5.26	5.89	5.64	5.10	4.51	3.99	3.99	3.92	3.91
	40	5.16	5.91	5.54	5.18	4.41	3.85	3.78	3.77	3.75
	50	5.05	5.95	5.59	5.15	4.34	3.70	3.70	3.63	3.64
	60	5.33	6.40	6.12	5.53	4.64	3.95	3.92	3.90	3.88
	70	5.31	6.50	6.17	5.60	4.59	3.92	3.88	3.86	3.85
	80	5.31	6.58	6.26	5.72	4.82	4.10	4.03	3.99	4.00
	90	5.36	6.77	6.84	6.11	4.87	4.12	4.07	4.09	4.10
	100	5.42	6.96	6.80	6.12	4.98	4.42	4.34	4.29	4.30

IC 95%		α								
		0.001	0.01	0.1	0.2	0.5	0.9	0.95	0.99	1
PCA	10	±0.6316%	±0.6351%	±0.6307%	±0.6250%	±0.6125%	±0.6088%	±0.6088%	±0.6086%	±0.6088%
	20	±0.4698%	±0.4839%	±0.4759%	±0.4651%	±0.4375%	±0.4260%	±0.4264%	±0.4264%	±0.4255%
	30	±0.4375%	±0.4615%	±0.4522%	±0.4312%	±0.4067%	±0.3836%	±0.3836%	±0.3804%	±0.3799%
	40	±0.4336%	±0.4622%	±0.4484%	±0.4344%	±0.4024%	±0.3771%	±0.3738%	±0.3733%	±0.3724%
	50	±0.4292%	±0.4637%	±0.4503%	±0.4332%	±0.3994%	±0.3700%	±0.3700%	±0.3666%	±0.3671%
	60	±0.4403%	±0.4797%	±0.4698%	±0.4480%	±0.4123%	±0.3818%	±0.3804%	±0.3794%	±0.3785%
	70	±0.4395%	±0.4832%	±0.4716%	±0.4506%	±0.4102%	±0.3804%	±0.3785%	±0.3776%	±0.3771%
	80	±0.4395%	±0.4859%	±0.4748%	±0.4552%	±0.4198%	±0.3886%	±0.3855%	±0.3836%	±0.3841%
	90	±0.4414%	±0.4924%	±0.4948%	±0.4694%	±0.4219%	±0.3896%	±0.3873%	±0.3882%	±0.3886%
	100	±0.4438%	±0.4988%	±0.4934%	±0.4698%	±0.4264%	±0.4029%	±0.3994%	±0.3972%	±0.3976%

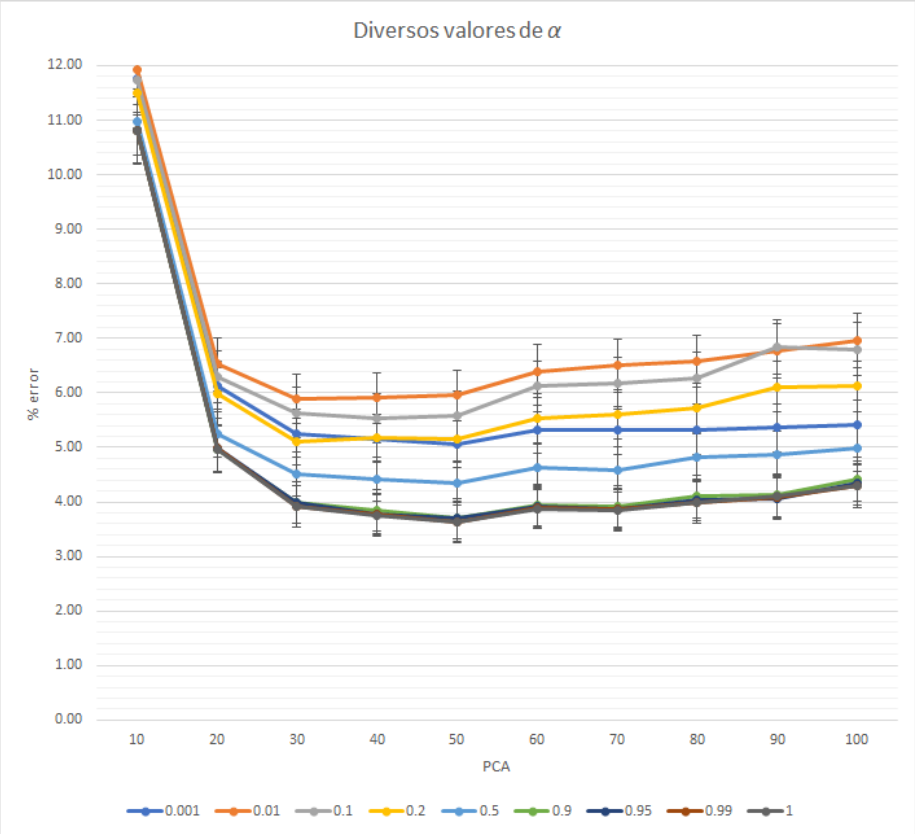

Figura 2: Error e IC 95 % en función de PCA para diversos valores de α

3. Aplicación de SVM a MNIST

3.1. Ejericio 5.1

El código adjunto muestra el paso M. Las ecuaciones están debidamente comentadas como el número de ecuación que implementan.

3.2. Ejericio 5.2

Puesto que se ha observado que aplicar un suavizado menor a 0.5 carece de sentido, y con el fin de ahorrar ejecuciones, los valores de α han cambiado: ahora están en el conjunto $\{0.5, 0.7, 0.9, 0.95, 0.99\}$.

k=1	10	20	30	40	50	60	70	80	90	100
α = 0.5	10.97	5.26	4.51	4.41	4.34	4.64	4.59	4.82	4.87	4.98
α = 0.7	10.83	5.07	4.16	4.05	3.96	4.24	4.21	4.43	4.45	4.68
α = 0.9	10.82	4.97	3.99	3.85	3.70	3.95	3.92	4.10	4.12	4.42
α = 0.95	10.82	4.98	3.99	3.78	3.70	3.92	3.88	4.03	4.07	4.34
α = 0.99	10.81	4.98	3.92	3.77	3.63	3.90	3.86	3.99	4.09	4.29

k=1	10	20	30	40	50	60	70	80	90	100
α = 0.5	±0.6125%	±0.4375%	±0.4067%	±0.4024%	±0.3994%	±0.4123%	±0.4102%	±0.4198%	±0.4219%	±0.4264%
α = 0.7	±0.6091%	±0.4300%	±0.3914%	±0.3864%	±0.3822%	±0.3949%	±0.3936%	±0.4033%	±0.4042%	±0.4140%
α = 0.9	±0.6088%	±0.4260%	±0.3836%	±0.3771%	±0.3700%	±0.3818%	±0.3804%	±0.3886%	±0.3896%	±0.4029%
α = 0.95	±0.6088%	±0.4264%	±0.3836%	±0.3738%	±0.3700%	±0.3804%	±0.3785%	±0.3855%	±0.3873%	±0.3994%
α = 0.99	±0.6086%	±0.4264%	±0.3804%	±0.3733%	±0.3666%	±0.3794%	±0.3776%	±0.3836%	±0.3882%	±0.3972%

Figura 3: Error e IC 95 % en función de PCA para diversos α con $k = 1$

El resultado en forma de gráfica se presenta en la siguiente página.

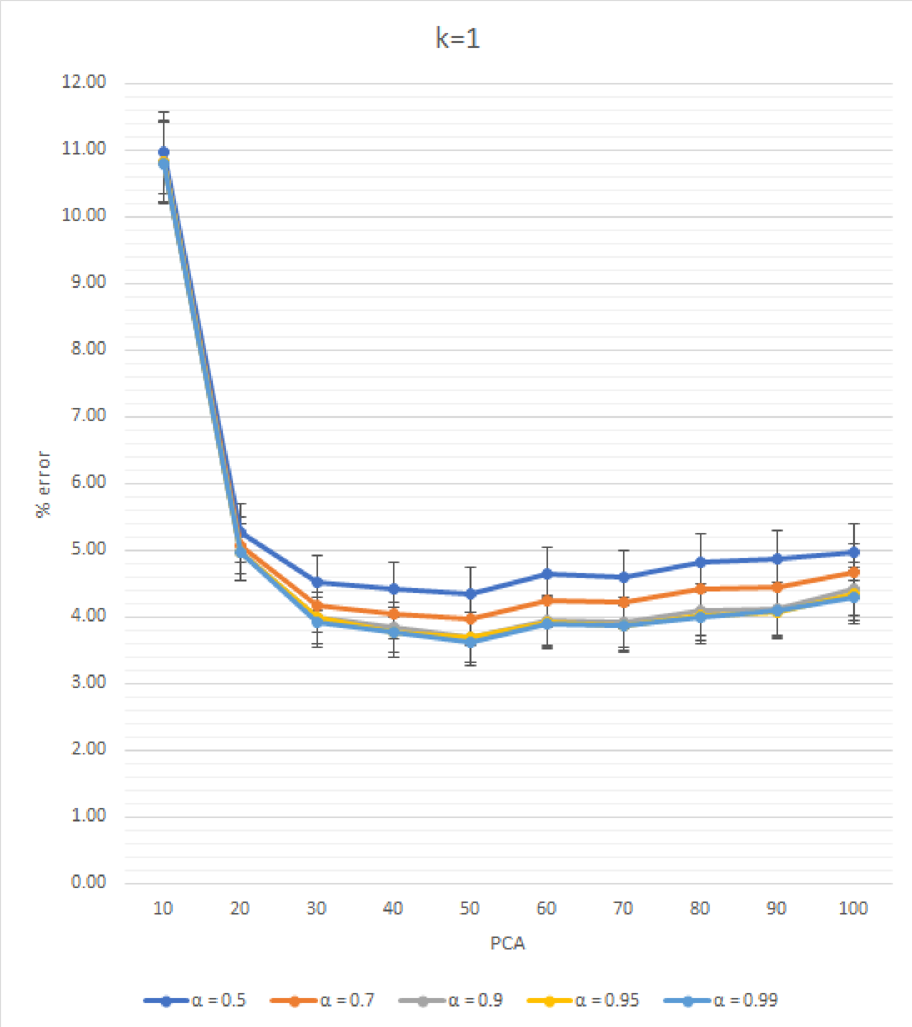

Figura 4: Gráfica del error en función de PCA para diversos α con $k = 1$

3.3. Ejericio 5.3

Los valores del suavizado son los mismos que en el ejercicio anterior.

Las figuras de las siguientes páginas representan el resultado según las dimensiones reducidas de PCA en función del número de componentes.

Se han vuelto a incluir los datos para $k = 1$ con el fin de dar una visión más completa de los resultados. El mejor error es de un 2.20 %, obtenido para $k = 9$ con 40 dimensiones de PCA y $\alpha = 0.9$ (marcado en amarillo).

α = 0.5	PCA 10	PCA 20	PCA 30	PCA 40	PCA 50	PCA 60	PCA 70	PCA 80	PCA 90	PCA 100
K = 1	10.97	5.26	4.51	4.41	4.34	4.64	4.59	4.82	4.87	4.98
K = 2	10.24	4.80	3.99	4.09	3.89	4.08	4.12	4.37	4.90	5.42
K = 3	9.14	4.65	3.58	3.64	3.60	4.21	4.31	4.72	4.83	4.98
K = 4	9.30	4.12	4.16	3.58	3.74	4.19	3.95	4.34	4.84	5.34
K = 5	9.30	3.81	3.68	3.28	3.43	3.55	4.16	4.10	4.55	4.90
K = 6	8.82	4.28	3.18	3.19	2.84	3.42	3.77	3.91	4.21	4.74
K = 7	8.92	4.21	3.54	3.12	3.31	3.68	4.15	4.24	3.97	4.54
K = 8	8.13	3.80	2.87	2.89	2.89	3.28	3.59	4.12	4.77	4.61
K = 9	8.40	3.64	2.98	2.98	3.35	3.51	4.06	4.24	4.59	4.81
K = 10	8.47	3.72	2.93	3.41	3.45	3.31	3.70	4.34	4.28	4.31

α = 0.7	PCA 10	PCA 20	PCA 30	PCA 40	PCA 50	PCA 60	PCA 70	PCA 80	PCA 90	PCA 100
K = 1	10.83	5.07	4.16	4.05	3.96	4.24	4.21	4.43	4.45	4.68
K = 2	10.08	4.82	3.87	3.69	3.64	3.87	4.07	4.21	4.82	5.08
K = 3	8.63	4.38	3.52	3.26	3.37	3.70	3.96	4.10	4.18	4.78
K = 4	8.81	3.91	3.52	3.34	3.24	3.49	3.66	4.30	4.54	5.17
K = 5	8.94	3.38	3.33	3.00	3.11	3.14	3.73	3.97	4.20	4.63
K = 6	8.13	3.82	2.97	2.73	2.77	3.34	3.19	3.56	3.65	4.22
K = 7	8.33	3.69	2.81	2.93	2.86	3.08	3.55	3.58	3.73	3.97
K = 8	7.63	3.59	2.64	2.59	2.76	3.00	3.23	3.75	4.42	4.41
K = 9	7.68	3.39	2.82	2.53	2.84	3.36	3.08	3.29	3.99	4.22
K = 10	7.80	3.30	2.81	2.59	2.93	3.00	3.38	3.47	3.90	3.95

α = 0.9	PCA 10	PCA 20	PCA 30	PCA 40	PCA 50	PCA 60	PCA 70	PCA 80	PCA 90	PCA 100
K = 1	10.82	4.97	3.99	3.85	3.70	3.95	3.92	4.10	4.12	4.42
K = 2	9.70	4.63	3.72	3.34	3.46	3.64	3.99	4.10	4.97	5.18
K = 3	8.62	4.28	3.30	2.97	3.23	3.38	3.65	3.74	4.06	4.84
K = 4	8.68	3.67	3.27	3.00	3.06	3.39	3.51	3.85	4.49	5.54
K = 5	8.04	3.44	2.95	2.91	2.85	3.06	3.37	3.71	4.21	4.83
K = 6	7.68	3.60	2.66	2.50	2.37	3.04	3.06	3.69	3.80	4.30
K = 7	7.87	3.29	2.75	2.64	2.78	2.75	3.05	3.75	3.77	4.07
K = 8	7.29	3.20	2.36	2.30	2.54	2.70	2.73	3.49	4.12	4.20
K = 9	7.07	3.30	2.64	2.20	2.56	3.13	3.00	3.01	3.51	3.76
K = 10	7.16	2.96	2.75	2.45	2.65	2.85	3.29	3.33	3.66	3.73

α = 0.95	PCA 10	PCA 20	PCA 30	PCA 40	PCA 50	PCA 60	PCA 70	PCA 80	PCA 90	PCA 100
K = 1	10.82	4.98	3.99	3.78	3.70	3.92	3.88	4.03	4.07	4.34
K = 2	9.61	4.60	3.76	3.22	3.41	3.67	3.96	4.08	5.04	5.34
K = 3	8.54	4.20	3.15	2.97	3.26	3.37	3.60	3.73	4.04	4.82
K = 4	8.44	3.61	3.19	2.90	2.88	3.25	3.40	3.77	4.67	5.62
K = 5	7.86	3.46	2.82	2.88	2.84	3.13	3.36	3.70	4.17	4.93
K = 6	7.61	3.29	2.53	2.46	2.29	2.93	3.02	3.79	3.75	4.24
K = 7	7.73	3.19	2.65	2.45	2.74	2.81	3.03	3.66	3.91	4.14
K = 8	7.45	3.07	2.36	2.43	2.52	2.74	2.66	3.53	3.99	4.28
K = 9	6.92	3.31	2.58	2.29	2.40	2.98	2.88	2.98	3.39	3.73
K = 10	6.89	3.00	2.60	2.40	2.65	2.94	3.29	3.15	3.54	3.80

α = 0.99	PCA 10	PCA 20	PCA 30	PCA 40	PCA 50	PCA 60	PCA 70	PCA 80	PCA 90	PCA 100
K = 1	10.81	4.98	3.92	3.77	3.63	3.90	3.86	3.99	4.09	4.29
K = 2	9.69	4.57	3.75	3.20	3.34	3.68	3.96	4.11	5.03	5.39
K = 3	8.62	4.19	3.09	3.06	3.21	3.29	3.54	3.76	4.14	4.86
K = 4	8.33	3.55	3.14	2.83	2.74	3.24	3.34	3.81	4.74	5.55
K = 5	7.77	3.27	2.74	2.77	2.90	3.08	3.40	3.61	4.19	5.03
K = 6	7.59	3.16	2.56	2.44	2.37	2.87	3.13	3.75	3.82	4.36
K = 7	7.61	3.28	2.65	2.40	2.70	2.78	3.14	3.57	3.88	4.17
K = 8	7.27	2.99	2.41	2.23	2.40	2.63	2.62	3.39	3.99	4.16
K = 9	6.84	3.18	2.51	2.23	2.41	3.00	2.83	3.14	3.39	3.76
K = 10	6.85	3.00	2.52	2.38	2.67	2.93	3.18	3.08	3.57	3.90

Figura 5: Error según componentes y PCA para diversos suavizados en MNIST

α = 0.5	PCA 10	PCA 20	PCA 30	PCA 40	PCA 50	PCA 60	PCA 70	PCA 80	PCA 90	PCA 100
K = 1	±0.6125%	±0.4375%	±0.4067%	±0.4024%	±0.3994%	±0.4123%	±0.4102%	±0.4198%	±0.4219%	±0.4264%
K = 2	±0.5942%	±0.4190%	±0.3836%	±0.3882%	±0.3790%	±0.3877%	±0.3896%	±0.4007%	±0.4231%	±0.4438%
K = 3	±0.5648%	±0.4127%	±0.3642%	±0.3671%	±0.3651%	±0.3936%	±0.3980%	±0.4157%	±0.4202%	±0.4264%
K = 4	±0.5692%	±0.3896%	±0.3914%	±0.3642%	±0.3719%	±0.3927%	±0.3818%	±0.3994%	±0.4206%	±0.4407%
K = 5	±0.5692%	±0.3752%	±0.3690%	±0.3491%	±0.3567%	±0.3627%	±0.3914%	±0.3886%	±0.4085%	±0.4231%
K = 6	±0.5558%	±0.3967%	±0.3439%	±0.3444%	±0.3256%	±0.3562%	±0.3733%	±0.3799%	±0.3936%	±0.4165%
K = 7	±0.5587%	±0.3936%	±0.3622%	±0.3408%	±0.3506%	±0.3690%	±0.3909%	±0.3949%	±0.3827%	±0.4080%
K = 8	±0.5357%	±0.3747%	±0.3272%	±0.3283%	±0.3283%	±0.3491%	±0.3646%	±0.3896%	±0.4177%	±0.4110%
K = 9	±0.5437%	±0.3671%	±0.3333%	±0.3333%	±0.3527%	±0.3607%	±0.3868%	±0.3949%	±0.4102%	±0.4194%
K = 10	±0.5457%	±0.3709%	±0.3305%	±0.3557%	±0.3577%	±0.3506%	±0.3700%	±0.3994%	±0.3967%	±0.3980%

α = 0.7	PCA 10	PCA 20	PCA 30	PCA 40	PCA 50	PCA 60	PCA 70	PCA 80	PCA 90	PCA 100
K = 1	±0.6091%	±0.4300%	±0.3914%	±0.3864%	±0.3822%	±0.3949%	±0.3936%	±0.4033%	±0.4042%	±0.4140%
K = 2	±0.5901%	±0.4198%	±0.3780%	±0.3695%	±0.3671%	±0.3780%	±0.3873%	±0.3936%	±0.4198%	±0.4304%
K = 3	±0.5504%	±0.4011%	±0.3612%	±0.3481%	±0.3537%	±0.3700%	±0.3822%	±0.3886%	±0.3923%	±0.4182%
K = 4	±0.5555%	±0.3799%	±0.3612%	±0.3522%	±0.3470%	±0.3597%	±0.3680%	±0.3976%	±0.4080%	±0.4340%
K = 5	±0.5592%	±0.3542%	±0.3517%	±0.3344%	±0.3402%	±0.3418%	±0.3714%	±0.3827%	±0.3932%	±0.4119%
K = 6	±0.5357%	±0.3757%	±0.3327%	±0.3194%	±0.3217%	±0.3522%	±0.3444%	±0.3632%	±0.3676%	±0.3940%
K = 7	±0.5416%	±0.3695%	±0.3239%	±0.3305%	±0.3267%	±0.3386%	±0.3627%	±0.3642%	±0.3714%	±0.3827%
K = 8	±0.5203%	±0.3646%	±0.3142%	±0.3113%	±0.3211%	±0.3344%	±0.3465%	±0.3724%	±0.4029%	±0.4024%
K = 9	±0.5219%	±0.3547%	±0.3245%	±0.3078%	±0.3256%	±0.3532%	±0.3386%	±0.3496%	±0.3836%	±0.3940%
K = 10	±0.5256%	±0.3501%	±0.3239%	±0.3113%	±0.3305%	±0.3344%	±0.3542%	±0.3587%	±0.3794%	±0.3818%

α = 0.9	PCA 10	PCA 20	PCA 30	PCA 40	PCA 50	PCA 60	PCA 70	PCA 80	PCA 90	PCA 100
K = 1	±0.6088%	±0.4260%	±0.3836%	±0.3771%	±0.3700%	±0.3818%	±0.3804%	±0.3886%	±0.3896%	±0.4029%
K = 2	±0.5801%	±0.4119%	±0.3709%	±0.3522%	±0.3582%	±0.3671%	±0.3836%	±0.3886%	±0.4260%	±0.4344%
K = 3	±0.5501%	±0.3967%	±0.3501%	±0.3327%	±0.3465%	±0.3542%	±0.3676%	±0.3719%	±0.3868%	±0.4206%
K = 4	±0.5518%	±0.3685%	±0.3486%	±0.3344%	±0.3376%	±0.3547%	±0.3607%	±0.3771%	±0.4059%	±0.4484%
K = 5	±0.5329%	±0.3572%	±0.3316%	±0.3295%	±0.3261%	±0.3376%	±0.3537%	±0.3705%	±0.3936%	±0.4202%
K = 6	±0.5219%	±0.3651%	±0.3154%	±0.3060%	±0.2981%	±0.3365%	±0.3376%	±0.3695%	±0.3747%	±0.3976%
K = 7	±0.5278%	±0.3496%	±0.3205%	±0.3142%	±0.3222%	±0.3205%	±0.3370%	±0.3724%	±0.3733%	±0.3873%
K = 8	±0.5095%	±0.3450%	±0.2975%	±0.2938%	±0.3084%	±0.3177%	±0.3194%	±0.3597%	±0.3896%	±0.3932%
K = 9	±0.5024%	±0.3501%	±0.3142%	±0.2875%	±0.3096%	±0.3413%	±0.3344%	±0.3349%	±0.3607%	±0.3728%
K = 10	±0.5053%	±0.3322%	±0.3205%	±0.3030%	±0.3148%	±0.3261%	±0.3496%	±0.3517%	±0.3680%	±0.3714%

α = 0.95	PCA 10	PCA 20	PCA 30	PCA 40	PCA 50	PCA 60	PCA 70	PCA 80	PCA 90	PCA 100
K = 1	±0.6088%	±0.4264%	±0.3836%	±0.3738%	±0.3700%	±0.3804%	±0.3785%	±0.3855%	±0.3873%	±0.3994%
K = 2	±0.5777%	±0.4106%	±0.3728%	±0.3460%	±0.3557%	±0.3685%	±0.3822%	±0.3877%	±0.4288%	±0.4407%
K = 3	±0.5478%	±0.3932%	±0.3423%	±0.3327%	±0.3481%	±0.3537%	±0.3651%	±0.3714%	±0.3859%	±0.4198%
K = 4	±0.5449%	±0.3656%	±0.3444%	±0.3289%	±0.3278%	±0.3476%	±0.3552%	±0.3733%	±0.4136%	±0.4514%
K = 5	±0.5275%	±0.3582%	±0.3245%	±0.3278%	±0.3256%	±0.3413%	±0.3532%	±0.3700%	±0.3918%	±0.4243%
K = 6	±0.5197%	±0.3496%	±0.3078%	±0.3036%	±0.2932%	±0.3305%	±0.3354%	±0.3743%	±0.3724%	±0.3949%
K = 7	±0.5235%	±0.3444%	±0.3148%	±0.3030%	±0.3200%	±0.3239%	±0.3360%	±0.3680%	±0.3799%	±0.3905%
K = 8	±0.5147%	±0.3381%	±0.2975%	±0.3018%	±0.3072%	±0.3200%	±0.3154%	±0.3617%	±0.3836%	±0.3967%
K = 9	±0.4974%	±0.3506%	±0.3107%	±0.2932%	±0.3000%	±0.3333%	±0.3278%	±0.3333%	±0.3547%	±0.3714%
K = 10	±0.4964%	±0.3344%	±0.3119%	±0.3000%	±0.3148%	±0.3311%	±0.3496%	±0.3423%	±0.3622%	±0.3747%

α = 0.99	PCA 10	PCA 20	PCA 30	PCA 40	PCA 50	PCA 60	PCA 70	PCA 80	PCA 90	PCA 100
K = 1	±0.6086%	±0.4264%	±0.3804%	±0.3733%	±0.3666%	±0.3794%	±0.3776%	±0.3836%	±0.3882%	±0.3972%
K = 2	±0.5798%	±0.4093%	±0.3724%	±0.3450%	±0.3522%	±0.3690%	±0.3822%	±0.3891%	±0.4284%	±0.4426%
K = 3	±0.5501%	±0.3927%	±0.3392%	±0.3376%	±0.3455%	±0.3496%	±0.3622%	±0.3728%	±0.3905%	±0.4215%
K = 4	±0.5416%	±0.3627%	±0.3418%	±0.3250%	±0.3200%	±0.3470%	±0.3522%	±0.3752%	±0.4165%	±0.4487%
K = 5	±0.5247%	±0.3486%	±0.3200%	±0.3217%	±0.3289%	±0.3386%	±0.3552%	±0.3656%	±0.3927%	±0.4284%
K = 6	±0.5191%	±0.3429%	±0.3096%	±0.3024%	±0.2981%	±0.3272%	±0.3413%	±0.3724%	±0.3757%	±0.4002%
K = 7	±0.5197%	±0.3491%	±0.3148%	±0.3000%	±0.3177%	±0.3222%	±0.3418%	±0.3637%	±0.3785%	±0.3918%
K = 8	±0.5089%	±0.3338%	±0.3006%	±0.2894%	±0.3000%	±0.3137%	±0.3131%	±0.3547%	±0.3836%	±0.3914%
K = 9	±0.4948%	±0.3439%	±0.3066%	±0.2894%	±0.3006%	±0.3344%	±0.3250%	±0.3418%	±0.3547%	±0.3728%
K = 10	±0.4951%	±0.3344%	±0.3072%	±0.2988%	±0.3160%	±0.3305%	±0.3439%	±0.3386%	±0.3637%	±0.3794%

Figura 6: Intervalos de confianza al 95 % del error según componentes y PCA para diversos suavizados en MNIST

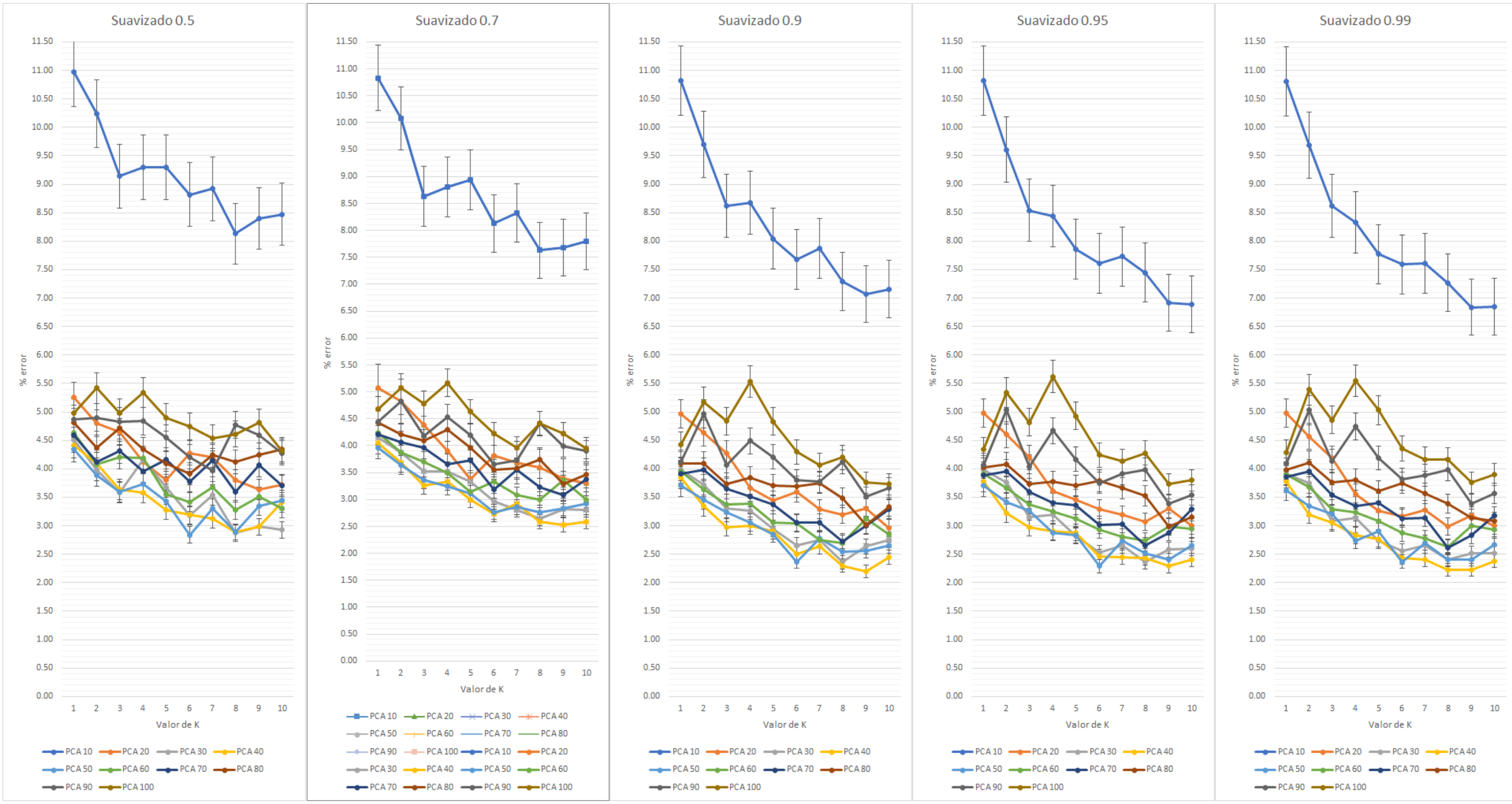

Figura 7: Gráfica del error según componentes y PCA para diversos suavizados en MNIST

4. Conclusiones

El valor del suavizado es importante a la hora de entrenar un modelo de mixtura de gaussianas. Un suavizado pequeño incrementa ligeramente la tasa de error con respecto a otros valores, mientras que un suavizado grande altera mucho el error tanto para entrenamiento como para test (véase $\alpha = 0.1$ en los resultados). De hecho, con un suavizado elevado, se puede comprobar en el entrenamiento cómo la mixtura de gaussianas "desaprende": debido al poco peso que tiene la matriz original con respecto a la matriz identidad, el error en cada ejecución aumenta. Esto a su vez se debe a que, si el valor de α es pequeño, se asume que los datos tienen poca covarianza ya que se le está dando más valor a la matriz identidad (únicamente varianza intraclase) que a la matriz calculada en la obtención de la matriz de covarianza (tanto varianza intraclase como varianza interclase), lo cual no tiene por qué ser cierto. Lo óptimo es que el valor del suavizado se encuentre entre 0.9 y 0.95, tal y como se ha podido comprobar en las tablas anteriores.

En cuanto al error, el menor valor ha sido encontrado con la proyección a 40 dimensiones PCA, 9 componentes de la mixtura gaussiana y un valor del suavizado igual a 0.9. En general, pasadas unas dimensiones de proyección determinadas de PCA (entre 60 y 70 dimensiones), el error tiende a aumentar, así como si incrementamos en exceso las componentes de la mixtura. Los mejores errores se encuentran con entre 8 y 9 componentes de la mixtura gaussiana; con 10 componentes el error comienza a aumentar de nuevo. Esto no sigue del todo el esquema lógico a priori: 10 componentes deberían de ser el número ideal para clasificar 10 clases (en nuestro caso, 10 dígitos). Sin embargo, es posible que las representaciones de algunos números compartan ciertas características y una clase pueda estar entre varias distribuciones solapadas (componentes de la mixtura); permitiendo, por ejemplo, clasificar dos grupos de números mediante una sola distribución de la mixtura, o clasificar tres grupos de números mediante dos distribuciones, lo que permitiría "ahorrarse" componentes de la mixtura gaussiana final. Cabría la posibilidad de que, al buscar 10 componentes, exista un conflicto con dichos solapamientos y esto de lugar a cierto grado de sobreentrenamiento al tratar de diferenciar entre dos distribuciones muy similares.