

Rapport de laboratoire 3

GTI770

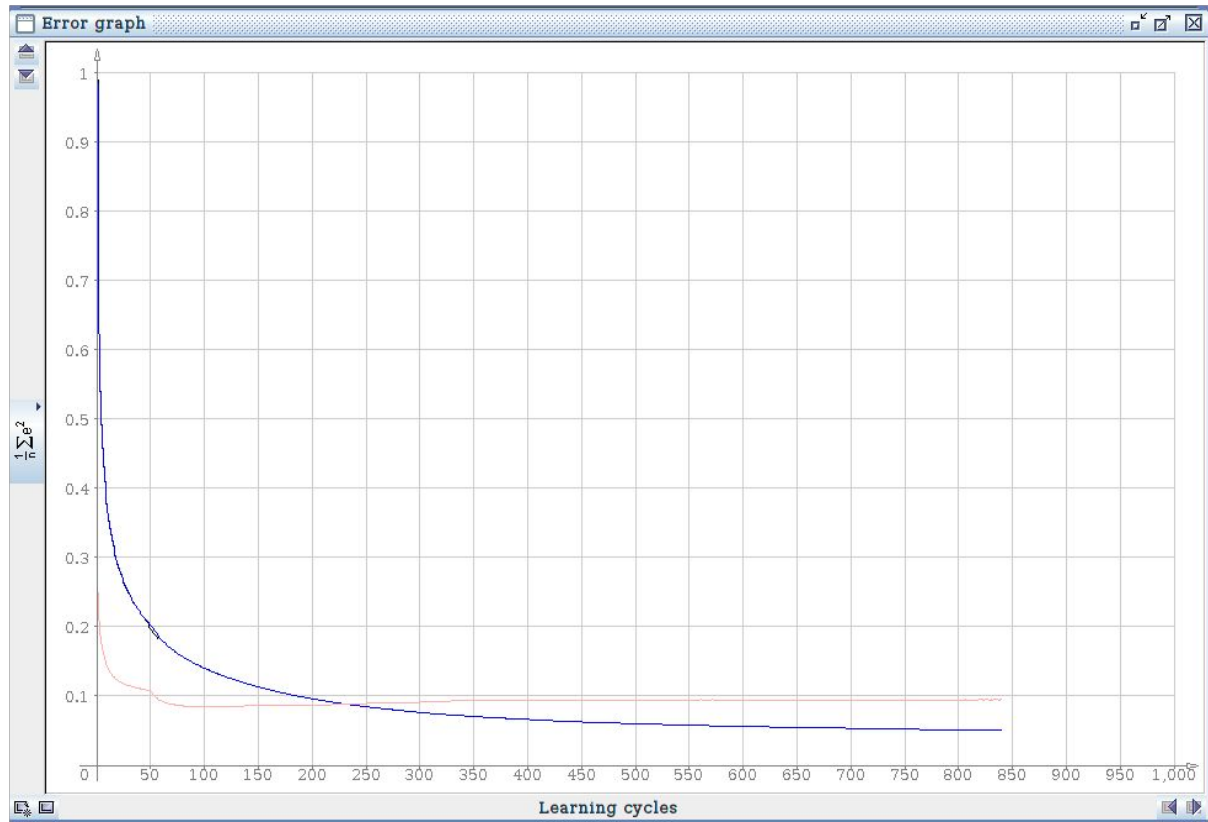
par

Olivier Arteau
Simon Grondin

#2

Le temps d'apprentissage est d'environ 1h30 pour les 2000 itérations.

#3



a) On peut observer qu'à partir d'environ 225, le taux d'erreur sur l'ensemble de test stagne. Étant donné que le nombre d'itération est à 2000 ce qui largement au-dessus, on peut donc affirmer sans problème que l'on est dans une situation de surapprentissage.

b) La valeur idéal est autour de 225 puisque c'est à ce point qu'on obtient la meilleur performance sur l'ensemble de test.

#4

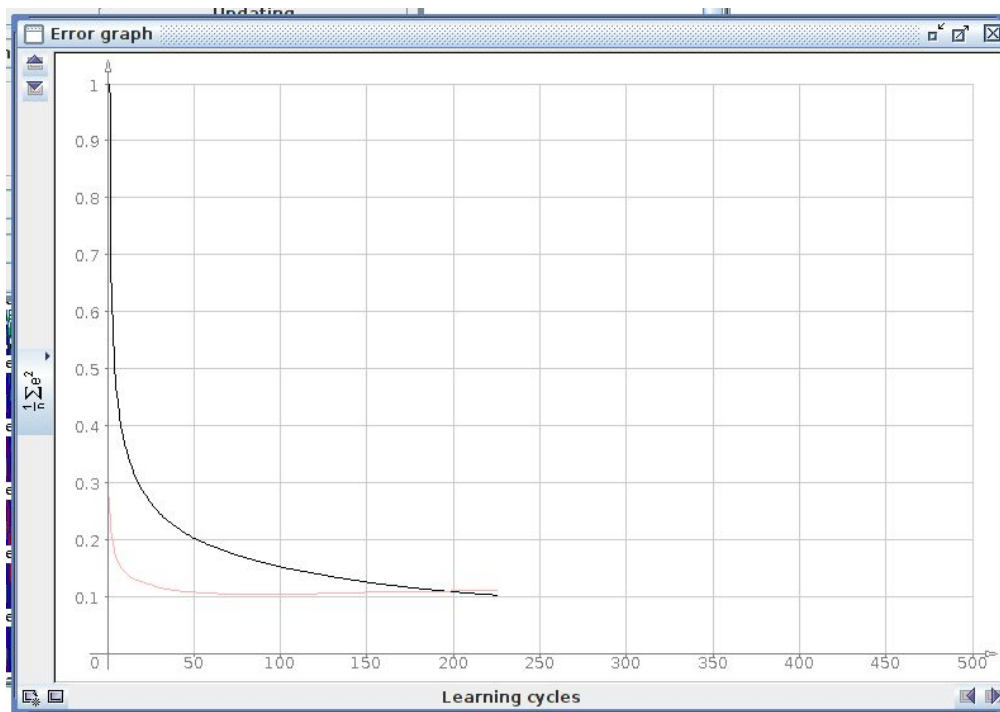
Ensemble de données	Exemples testés	Exemples correctement classifiés	Taux de succès
T2 (Apprentissage)	37440	36993	98.81%
T1 (Test)	12092	11560	95.60%
T3 (Validation)	11941	11111	93.05%

Matrice de confusion pour l'ensemble T3

[illegible][illegible]

#5

a)



Ensemble de données	Exemples testés	Exemples correctement classifiés	Taux de succès
T2 (Apprentissage)	37440	36556	97.64%
T1 (Test)	12092	11273	93.23%
T3 (Validation)	11941	10960	91.78%

Matrice de confusion pour l'ensemble T3

class	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17
0	423	1	0	0	0	3	2	5	2	1	3	0	4	0	0	0	1	6
1	2	401	0	6	3	2	5	2	0	0	2	0	0	0	6	0	1	2
2	0	0	506	0	2	1	3	0	0	0	0	3	0	0	2	0	0	0
3	1	5	0	358	0	0	0	0	0	3	0	0	2	0	17	5	1	0
4	0	1	2	0	348	6	1	0	0	1	2	0	0	0	0	0	0	1
5	0	0	0	0	1	394	0	1	0	1	1	0	0	0	0	9	0	2
6	0	2	2	0	0	1	368	2	0	0	1	2	0	0	4	0	6	0
7	3	0	0	0	0	0	2	387	0	0	3	0	1	1	0	0	0	0
8	0	3	7	6	31	2	1	0	563	98	0	7	0	0	0	2	5	0
9	0	0	0	2	0	0	1	0	0	411	0	0	0	0	1	0	0	0
10	0	1	2	0	2	2	1	7	0	0	336	9	0	0	0	0	0	7
11	0	0	2	0	1	0	0	0	0	0	0	489	0	0	0	0	0	0
12	1	0	0	0	0	0	0	34	0	0	2	0	401	10	0	2	0	2
13	0	0	0	0	0	0	0	9	1	0	1	0	4	405	0	0	0	0
14	1	2	3	32	0	0	1	0	1	0	0	0	3	0	386	1	17	0
15	1	1	0	5	0	3	0	0	0	1	0	0	0	0	0	449	1	0
16	1	2	1	1	0	1	3	1	0	0	0	0	1	0	6	3	417	3
17	2	3	4	0	0	2	0	1	0	0	11	2	0	0	0	3	2	413
18	0	1	0	1	0	1	1	0	0	12	1	0	0	0	0	0	0	0
19	1	0	0	0	0	3	0	1	0	6	1	0	1	0	0	3	0	0
20	0	0	0	0	0	0	0	4	0	3	3	1	0	3	0	0	1	1
21	0	0	0	0	0	0	0	1	0	0	0	0	0	2	0	0	0	0
22	0	0	1	1	0	0	0	1	1	0	0	0	4	4	0	0	0	0
23	2	0	0	0	0	0	0	0	1	0	18	1	0	0	0	0	0	1
24	0	0	0	0	0	0	0	3	1	1	1	0	0	0	0	0	0	0
25	0	3	1	1	8	0	0	0	0	1	0	1	0	0	0	0	0	4
noclass	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0

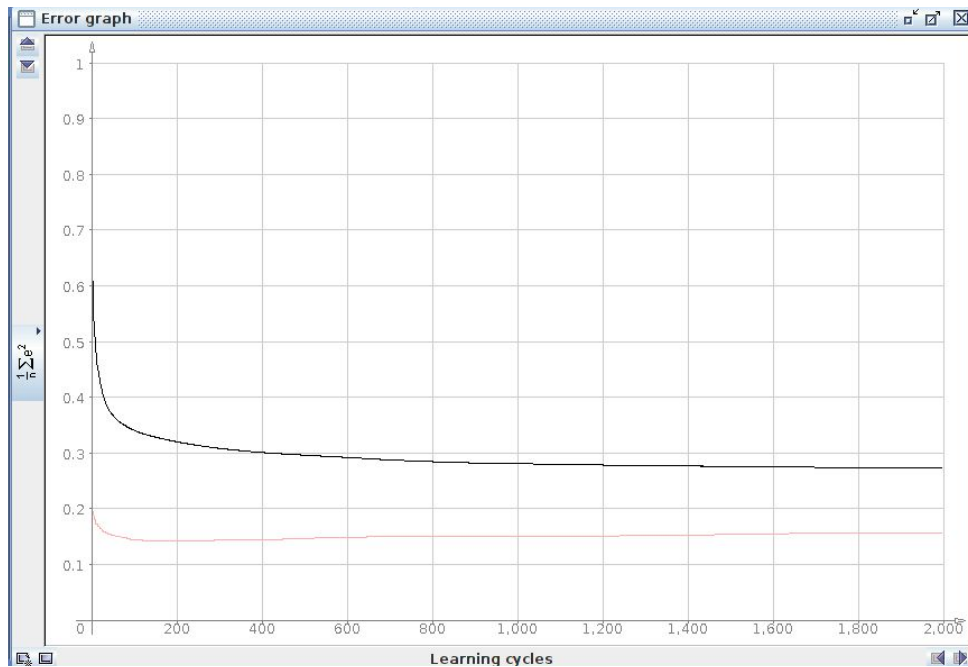
18	19	20	21	22	23	24	25	unknown
0	0	0	1	1	6	0	0	0
2	0	0	1	0	0	0	0	0
0	0	0	1	0	0	0	0	0
0	0	1	2	1	0	0	0	0
0	0	0	0	0	0	0	3	0
0	8	0	0	0	0	0	2	0
0	0	0	1	0	0	0	0	0
0	1	0	2	0	1	1	0	0
8	18	0	0	0	5	3	56	0
8	1	0	0	0	0	0	2	0
1	0	1	1	0	7	0	0	0
0	0	2	2	0	0	0	0	0
0	0	2	1	3	2	0	0	0
0	0	0	1	17	1	0	0	0
0	0	11	0	1	0	0	0	0
0	0	0	4	0	0	2	0	0
0	0	7	1	4	0	0	0	0
0	0	1	0	1	1	0	0	0
422	0	0	0	0	0	4	2	0
0	452	0	0	0	0	0	1	0
0	0	405	30	6	0	1	0	0
0	0	11	463	4	0	1	0	0
0	0	5	0	457	0	1	0	0
0	0	0	1	1	443	4	0	0
0	4	1	18	1	8	415	0	0
0	0	0	0	0	0	0	448	0
0	0	0	0	0	0	0	0	0

b)

La principale différence entre les deux résultats, est que les résultats sur l'ensemble d'apprentissage sont beaucoup plus près que celui de l'ensemble de test. On n'est donc pas en train de faire du surapprentissage. Par contre, les résultats de l'ensemble de test et de validation sont à peu près équivalents.

#6

L'analyse était assez rapide, mais les taux d'erreurs sont plats. Il n'y a pas d'apprentissage détectable selon le nombre de cycles



#7

L'analyse a pris un temps incroyablement long dû au grand nombre de neurones dans la couche cachée. Les résultats stagnent à partir de 50 cycles. Il y a donc beaucoup de surapprentissage pour 2000 itérations. L'erreur sur les données d'entraînement est d'environ 20% et l'erreur sur les données de test était d'environ 10%.

#8

Architecture du réseau	# de cycle d'apprentissage	Temps d'apprentissage	Erreur MSE	Erreur MSE		
				T1	T2	T3
108-67-26	2000	1h30	10%	95.6	98.8	93.1
108-67-26	225	0h15	10%	93.2	97.6	91.8
108-13-26	2000	0h10	15%	90.1	93.7	86.8
108-335-26	50	2h00	10%	96.2	97.8	94.2

Meilleur matrice de confusion pour 108-335-26.

class	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17
0	1426	0	0	0	1	1	0	3	0	0	3	1	1	0	0	3	1	0
1	9	1407	1	8	2	0	1	0	0	0	0	0	0	0	5	0	1	3
2	0	0	1419	0	0	0	2	0	1	0	0	9	0	0	0	0	1	1
3	3	0	0	1378	0	0	1	0	0	0	0	1	1	0	50	0	4	0
4	0	4	8	0	1401	6	4	0	1	0	1	5	0	0	0	0	0	2
5	4	0	0	0	1	1423	0	0	0	0	0	0	0	0	0	5	0	0
6	3	10	4	1	1	6	1397	3	0	0	0	0	0	0	1	1	7	0
7	19	1	0	0	0	1	0	1375	0	2	3	0	4	16	0	1	0	0
8	0	0	0	0	0	1	0	0	1414	4	0	6	0	0	2	2	0	0
9	1	0	0	1	0	0	0	0	2	1413	0	0	0	0	0	0	0	0
10	4	0	0	0	0	1	0	2	1	0	1418	6	0	1	0	1	0	1
11	0	0	1	0	1	2	0	0	0	0	0	1427	0	0	0	0	0	1
12	1	0	0	0	0	0	0	2	0	0	1	0	1429	4	1	0	0	0
13	2	0	0	0	0	0	0	2	0	1	2	0	3	1418	1	1	0	0
14	2	2	2	17	0	1	0	0	1	0	0	0	0	0	1400	0	10	0
15	4	0	0	10	0	3	0	0	1	1	0	0	0	0	0	1407	4	1
16	3	3	1	2	1	0	3	0	0	0	0	0	0	0	12	4	1409	0
17	20	1	0	0	1	2	0	0	0	0	4	1	1	1	0	2	0	1403
18	0	0	1	0	1	2	0	1	0	3	0	0	0	0	1	0	0	0
19	3	0	0	0	1	5	0	0	0	0	0	0	0	0	0	1	0	0
20	0	1	1	1	1	0	0	0	0	0	1	0	2	1	1	0	0	0
21	0	0	0	0	1	0	0	0	1	1	0	0	0	1	0	0	0	0
22	1	0	0	0	0	0	0	5	0	0	0	1	4	30	0	0	1	0
23	5	0	0	1	0	2	0	0	0	0	10	1	0	1	0	0	0	0
24	0	0	0	2	0	3	0	0	0	4	0	0	1	1	0	1	0	0
25	0	1	0	0	0	1	0	0	0	1	0	1	0	0	0	0	0	0
noiclass	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0

18	19	20	21	22	23	24	25	unknown
0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	3	0
4	0	1	2	0	0	0	0	0
0	0	0	1	1	0	0	0	0
5	0	0	0	0	1	2	0	0
2	0	0	0	0	0	2	1	0
2	0	1	0	0	0	3	0	0
0	1	5	2	3	3	4	0	0
1	1	0	4	0	0	2	3	0
6	9	3	3	0	0	2	0	0
0	0	2	1	0	1	1	0	0
0	0	0	0	0	2	1	5	0
0	0	1	0	1	0	0	0	0
1	0	1	2	4	1	0	1	0
0	1	3	1	0	0	0	0	0
1	3	1	1	0	0	3	0	0
0	0	1	0	0	0	1	0	0
0	1	0	0	0	0	2	1	0
1431	0	0	0	0	0	0	0	0
0	1428	0	0	0	0	2	0	0
2	2	1376	48	2	0	0	2	0
2	1	27	1399	0	0	7	0	0
0	0	5	9	1384	0	0	0	0
0	0	1	2	0	1402	10	5	0
0	2	2	13	0	2	1409	0	0
0	1	0	0	0	1	0	1434	0
0	0	0	0	0	0	0	0	0

On remarque que les paramètres qui donnent la meilleur précision sont ceux qui ont le plus de neurones dans la couche cachée. Par contre ces paramètres étaient excessivement lent à entraîner et l'amélioration était relativement faible (1% à 8%) d'amélioration. Dans un contexte où la rapidité d'apprentissage est importante, le compromis avec 67 neurones est très intéressant en terme de précision et de gain en vitesse.