EPART Lab 3 Report

Rustenis Tolpeznikas

December 14, 2024

1 Implementing Perceptron

The perceptron was implemented as followed in the lab. Below is a figure of the perceptron in action:

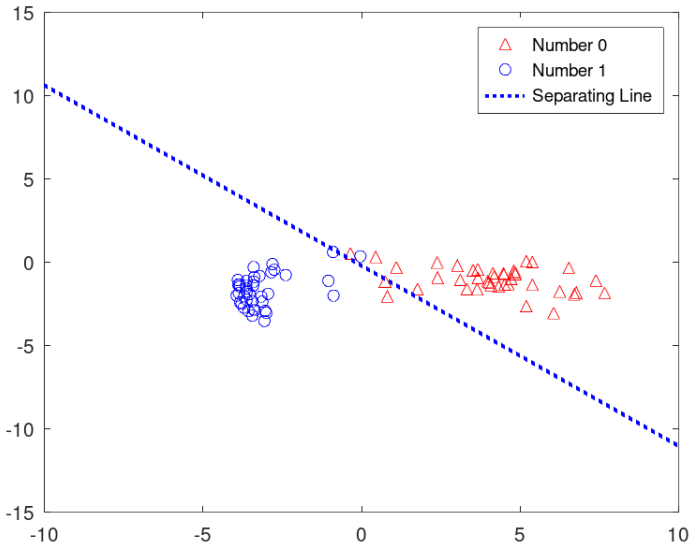

Figure 1: Perceptron(fixed) in action

This report was performed with a fixed perceptron learning rate assumed to be 1, rather than decaying one.

1.1 Data Set

The data set consists of 10 classes, each representing a digit from 0 to 9. In this report, class 1 corresponds to digit 0, as the data set is shifted by one.

Type	**0**	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**
Training	5923	6742	5958	6131	5842	5421	5918	6265	5851	5949
Testing	980	1135	1032	1010	982	892	958	1028	974	1009

2 One versus One

2.1 Original Data Set

Training

Class	**0**	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**Rejected**
0	5694	1	21	11	5	54	22	3	16	2	94
1	0	6487	28	15	2	8	2	14	72	6	108
2	23	16	5386	58	38	17	61	32	83	10	234
3	10	14	69	5465	1	197	9	28	101	27	210
4	8	8	28	3	5394	5	28	22	16	183	147
5	24	5	24	170	11	4702	66	10	104	32	273
6	19	2	65	2	28	86	5580	1	22	0	113
7	5	6	44	35	24	12	1	5794	13	127	204
8	14	54	61	122	7	120	31	16	5098	45	283
9	11	12	24	74	172	30	0	195	41	5201	189

Correct Classifications	Errors	Rejections
0.913350	0.055733	0.030917

Testing

Class	0	1	2	3	4	5	6	7	8	9	Rejected
0	949	0	2	2	0	6	4	1	0	1	15
1	0	1099	6	1	0	1	0	0	13	0	15
2	4	4	946	7	5	3	8	7	11	1	36
3	0	1	10	920	2	23	0	4	12	3	35
4	1	0	4	1	913	1	6	6	3	24	23
5	8	1	4	40	4	769	7	0	19	4	36
6	6	2	13	1	5	11	905	0	3	0	12
7	0	3	13	9	5	0	0	938	6	21	33
8	4	2	5	27	4	25	6	4	852	3	42
9	4	4	1	9	36	4	0	30	11	880	30

2.2 Extended Data Set

Training

Class	0	1	2	3	4	5	6	7	8	9	Rejected
0	5913	0	0	0	1	1	2	0	1	0	5
1	0	6697	3	2	3	0	2	7	5	2	21
2	0	3	5944	0	0	0	1	7	0	0	3
3	0	1	0	6114	0	6	0	3	0	1	6
4	2	5	0	0	5807	0	2	2	0	18	6
5	2	0	0	2	0	5411	1	0	0	0	5
6	5	2	1	0	1	2	5902	0	0	0	5
7	0	5	1	0	4	0	0	6240	1	4	10
8	0	6	0	0	0	3	1	1	5834	2	4
9	0	4	0	0	18	0	0	11	0	5906	10

Testing

Class	0	1	2	3	4	5	6	7	8	9	Rejected
0	966	0	0	0	0	0	4	1	0	1	8
1	0	1118	1	1	1	0	3	0	3	0	8
2	5	1	998	3	1	0	3	6	5	1	9
3	0	0	2	971	0	6	0	3	7	3	18
4	1	0	3	0	947	0	4	2	1	13	11
5	3	0	0	11	1	850	3	0	5	2	17
6	5	2	1	0	6	5	930	0	2	0	7
7	0	4	6	1	3	0	0	980	3	10	21
8	2	0	3	11	1	4	2	3	930	2	16
9	0	2	2	4	5	3	1	2	0	971	19

3 One versus Rest

3.1 Original Data Set

Training

Class	0	1	2	3	4	5	6	7	8	9	Rejected
0	4718	0	7	4	3	47	13	3	2	3	1123
1	1	5797	21	9	4	29	2	8	77	11	783
2	8	19	4448	38	36	21	40	46	73	25	1204
3	14	5	54	3787	1	244	16	18	51	61	1880
4	7	8	12	2	4150	15	11	6	39	196	1396
5	28	9	11	60	39	3559	38	9	52	27	1589
6	8	7	7	1	13	81	4982	0	13	0	806
7	11	8	49	8	14	17	1	4498	13	152	1494
8	19	58	17	72	1	169	21	8	3615	54	1817
9	12	6	17	37	70	65	0	59	18	4506	1159

Testing

Class	0	1	2	3	4	5	6	7	8	9	Rejected
0	806	0	0	1	0	6	5	1	0	1	160
1	0	980	0	2	0	1	3	1	10	0	138
2	1	3	750	7	3	3	7	11	24	3	220
3	0	0	4	646	0	39	2	3	9	4	303
4	1	0	0	1	713	2	3	1	6	29	226
5	7	1	1	8	4	594	9	2	8	4	254
6	5	0	1	1	1	11	803	0	4	0	132
7	3	3	18	2	3	1	0	717	2	19	260
8	7	2	4	10	4	21	4	3	604	6	309
9	2	1	2	2	12	10	0	7	2	744	227

3.2 Extended Data Set

Training

Class	0	1	2	3	4	5	6	7	8	9	Rejected
0	5794	1	3	1	0	1	1	1	0	2	119
1	0	6592	9	2	1	0	0	10	3	2	123
2	3	5	5742	4	4	0	1	8	4	4	183
3	0	0	10	5865	0	21	0	6	3	6	220
4	0	1	1	0	5613	0	3	3	1	15	205
5	4	2	2	9	0	5161	6	3	5	8	221
6	4	1	2	0	4	5	5784	0	2	0	116
7	1	2	10	3	2	2	0	6039	2	12	192
8	3	7	5	7	1	10	6	3	5490	6	313
9	2	1	0	5	11	6	0	18	2	5675	229

Testing

Class	0	1	2	3	4	5	6	7	8	9	Rejected
0	947	0	0	0	0	0	0	0	0	1	32
1	0	1106	1	1	0	0	2	0	1	0	24
2	2	0	973	0	1	0	0	3	1	0	52
3	0	0	0	947	0	1	0	0	1	2	59
4	0	1	1	0	925	0	2	1	0	4	48
5	2	0	0	4	1	822	2	0	3	1	57
6	3	2	0	0	2	3	907	0	0	0	41
7	0	2	3	1	1	0	0	957	0	6	58
8	0	0	2	3	1	1	2	2	889	3	71
9	1	1	0	2	4	0	1	1	1	943	55

4 Further Improvements

For further classification improvements, I began by thinking big. I spent few hours trying to implement LDA transformation, but it proved too much of a hassle, with many code changes required. Then I tried normalizing the data before PCA transformation. That not only did not improve the results, but made the extended data performance significantly worse (as low as 20% accuracy). Lastly, I have decided to change a single parameter named *comp_count* from 40 to 50.
For the comparison, I have chosen the One versus One method with extended data set. Classification accuracy is the metric of choice. As now there are 50 primary components, extension results in 1275 total features. Thus, the new ensemble is shortly named *OVO1275* Here are the results: S

Similar improvements were observed in all the other ensemble methods, for both original and extended data sets. However, as this is a brute force method, the training time is even longer than before.

5 Conclusion