

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:

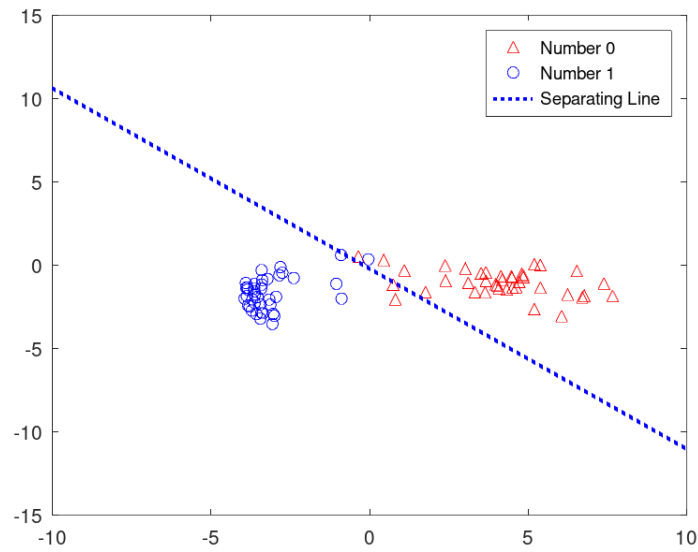


Figure 1: Perceptron(fixed) in action

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

Note: Error rates below are calculated as $(FP + FN) / TP + FP + FN$

2.1 Original Data Set

Training

OVO40 Training Confusion Matrix

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

Class	0	1	2	3	4	5	6	7	8	9
FP	114	118	364	490	288	529	220	321	468	432
FN	229	255	572	666	448	719	338	471	753	748
Error Rate	0.0579	0.0576	0.1474	0.1757	0.1204	0.2101	0.0900	0.1201	0.1905	0.1841

Correct Classifications	Errors	Rejections
0.913350	0.055733	0.030917

Testing

OVO40 Testing Confusion Matrix

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

Class	0	1	2	3	4	5	6	7	8	9
FP	27	17	50	97	61	74	31	52	78	57
FN	31	36	86	90	69	123	53	90	122	129
Error Rate	0.0577	0.0468	0.1258	0.1688	0.1242	0.2033	0.0848	0.1322	0.1908	0.1743

Correct Classifications	Errors	Rejections
0.917100	0.055200	0.027700

2.2 Extended Data Set

Training

OVO860 Training Confusion Matrix

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

Class	0	1	2	3	4	5	6	7	8	9
FP	9	26	5	4	27	12	9	31	7	27
FN	10	45	14	17	35	10	16	25	17	43
Error Rate	0.0032	0.0106	0.0032	0.0034	0.0106	0.0041	0.0042	0.0089	0.0041	0.0118

Correct Classifications	Errors	Rejections
0.996130	0.002617	0.001250

Testing

OVO860 Testing Confusion Matrix

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

Class	0	1	2	3	4	5	6	7	8	9
FP	16	9	18	31	18	18	20	17	24	32
FN	14	17	34	39	35	42	28	48	44	38
Error Rate	0.0306	0.0229	0.0498	0.0677	0.0529	0.0658	0.0484	0.0628	0.0688	0.0678

Correct Classifications	Errors	Rejections
0.966100	0.020500	0.013400

3 One versus Rest

3.1 Original Data Set

Training

OVR40 Training Confusion Matrix

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

Class	0	1	2	3	4	5	6	7	8	9
FP	108	120	195	231	181	688	142	157	338	529
FN	1205	945	1510	2344	1692	1862	936	1767	2236	1443
Error Rate	0.2191	0.1571	0.2766	0.4057	0.3115	0.4177	0.1771	0.2990	0.4157	0.3041

Correct Classifications	Errors	Rejections
0.734333	0.044817	0.220850

Testing

OVR40 Testing Confusion Matrix

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

Class	0	1	2	3	4	5	6	7	8	9
FP	26	10	30	34	27	94	33	29	65	66
FN	174	155	282	364	269	298	155	311	370	265
Error Rate	0.1985	0.1431	0.2933	0.3817	0.2932	0.3943	0.1906	0.3212	0.4182	0.3086

Correct Classifications	Errors	Rejections
0.735700	0.041400	0.222900

3.2 Extended Data Set

Training

OVR860 Training Confusion Matrix

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

Class	0	1	2	3	4	5	6	7	8	9
FP	17	20	42	31	23	45	17	52	22	55
FN	129	150	216	266	229	260	134	226	361	274
Error Rate	0.0249	0.0256	0.0447	0.0487	0.0430	0.0559	0.0259	0.0439	0.0654	0.0559

Correct Classifications	Errors	Rejections
0.962580	0.005400	0.032017

Testing

OVR860 Testing Confusion Matrix

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

Class	0	1	2	3	4	5	6	7	8	9
FP	8	6	7	11	11	5	9	7	7	17
FN	33	29	59	63	57	70	51	71	85	66
Error Rate	0.0414	0.0309	0.0632	0.0729	0.0681	0.0837	0.0611	0.0758	0.0934	0.0814

Correct Classifications	Errors	Rejections
0.941600	0.008700	0.049700

4 One versus Rest - Balanced

4.1 Original Data Set

Training

OVR40 - Balanced Training Confusion Matrix

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

Class	0	1	2	3	4	5	6	7	8	9
FP	108	120	195	231	181	688	142	157	338	529
FN	1205	945	1510	2344	1692	1862	936	1767	2236	1443
Error Rate	0.2191	0.1571	0.2766	0.4057	0.3115	0.4177	0.1771	0.2990	0.4157	0.3041

Correct Classifications	Errors	Rejections
0.731987	0.043682	0.224331

Testing

OVR40 - Balanced Testing Confusion Matrix

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

Class	0	1	2	3	4	5	6	7	8	9
FP	108	120	195	231	181	688	142	157	338	529
FN	1205	945	1510	2344	1692	1862	936	1767	2236	1443
Error Rate	0.2191	0.1571	0.2766	0.4057	0.3115	0.4177	0.1771	0.2990	0.4157	0.3041

Correct Classifications	Errors	Rejections
0.735500	0.042500	0.222000

4.2 Extended Data Set

Training

OVO860 - Balanced Training Confusion Matrix

Class	0	1	2	3	4	5	6	7	8	9	Rejected
0	5413	0	0	0	0	0	1	0	0	0	7
1	0	5403	1	0	1	0	1	6	0	1	8
2	0	1	5414	0	0	0	1	4	0	0	1
3	0	0	0	5417	0	0	0	0	0	1	3
4	2	3	0	0	5406	0	2	1	0	6	1
5	2	0	0	1	0	5412	1	0	1	0	4
6	3	2	1	0	2	1	5409	0	0	0	3
7	0	3	7	0	0	0	0	5360	1	24	26
8	0	2	0	1	0	3	0	1	5412	1	1
9	0	2	0	1	28	0	0	3	1	5375	11

Class	0	1	2	3	4	5	6	7	8	9
FP	7	13	8	4	29	4	6	15	3	33
FN	8	18	7	4	15	9	12	61	9	46
Error Rate	0.0028	0.0057	0.0028	0.0015	0.0081	0.0024	0.0033	0.0141	0.0022	0.0147

Correct Classifications	Errors	Rejections
0.996510	0.002287	0.001199

Testing

OVO860 - Balanced Testing Confusion Matrix

Class	0	1	2	3	4	5	6	7	8	9	Rejected
0	5413	0	0	0	0	0	1	0	0	0	7
1	0	5403	1	0	1	0	1	6	0	1	8
2	0	1	5414	0	0	0	1	4	0	0	1
3	0	0	0	5417	0	0	0	0	0	1	3
4	2	3	0	0	5406	0	2	1	0	6	1
5	2	0	0	1	0	5412	1	0	1	0	4
6	3	2	1	0	2	1	5409	0	0	0	3
7	0	3	7	0	0	0	0	5360	1	24	26
8	0	2	0	1	0	3	0	1	5412	1	1
9	0	2	0	1	28	0	0	3	1	5375	11

Class	0	1	2	3	4	5	6	7	8	9
FP	108	120	195	231	181	688	142	157	338	529
FN	1205	945	1510	2344	1692	1862	936	1767	2236	1443
Error Rate	0.2191	0.1571	0.2766	0.4057	0.3115	0.4177	0.1771	0.2990	0.4157	0.3041

Correct Classifications	Errors	Rejections
0.963700	0.022700	0.013600

5 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:

OVO1275 Training Confusion Matrix

Class	0	1	2	3	4	5	6	7	8	9	Rejected
0	962	0	0	0	0	0	3	1	1	1	12
1	0	1121	1	1	1	0	2	0	3	0	6
2	2	0	1004	2	1	0	1	5	6	1	10
3	0	0	1	980	0	4	0	2	6	2	15
4	2	0	3	0	951	0	4	1	0	10	11
5	2	0	1	8	0	850	4	1	6	5	15
6	4	2	2	1	7	4	924	0	3	1	10
7	0	4	5	4	1	0	0	988	3	8	15
8	1	0	2	9	1	3	2	3	933	3	17
9	1	2	1	3	6	4	2	3	1	973	13

Direct performance comparison:

OVO1275

Class	0	1	2	3	4	5	6	7	8	9
FP	12	8	16	28	17	15	18	16	29	31
FN	23	13	25	40	22	26	26	28	45	44
Error Rate	0.0349	0.0186	0.0398	0.0653	0.0399	0.0451	0.0451	0.0428	0.0732	0.0718

Versus OVO860

Class	0	1	2	3	4	5	6	7	8	9
FP	16	9	18	31	18	18	20	17	24	32
FN	14	17	34	39	35	42	28	48	44	38
Error Rate	0.0306	0.0229	0.0498	0.0677	0.0529	0.0658	0.0484	0.0628	0.0688	0.0678

OVO1275

Correct Classifications	Errors	Rejections
0.968600	0.019000	0.012400

Versus OVO860

Correct Classifications	Errors	Rejections
0.966100	0.020500	0.013400

While the improvement is not significant, both of these methods are already very accurate. There are small improvements in correct classifications and rejections and even the error rate is slightly lower. Similar, but even greater 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.

6 Conclusion

This mini-project/laboratory work has shown that choosing the right linear classification ensemble method comes with different trade-offs. One versus Rest can be trained faster, but its performance is lacking. One versus One is more accurate, but it takes longer to train. Feature extension has shown significant improvements in classification accuracy, for OvO and OvR methods, however, the training time is increased greatly. Giving a balanced training set to One versus Rest method has improved its performance. However, it still lags behind One versus One method.