CSD316 - Assignment 3

Pranav Dahiya

1710110249

The dual coordinate descent algorithm was used to train the SVM. Refer to [1] for more details. The results for a linear SVM are shown in table 1 below. The regularization parameter was tuned to 500 and each SVM was trained for 1000 epochs.

Table 1: Results for SVM without kernel

	TP	FP	TN	FN
1	43	3	238	5
2	35	0	241	13
3	45	2	239	3
4	40	3	238	8
5	42	0	241	6
6	45	0	242	3
7	36	1	240	12
8	34	2	239	14
9	42	10	231	6
10	42	1	241	7
average	40.40	2.20	239.00	7.70
standard deviation	4.03	2.97	3.13	4.00

The following kernels were used in the kernelized SVMs:

1. Linear:
$$K(x, y) = x^T y + 1$$
2. Polynomial:
$$K(x, y) = (x^T y + 1)^3$$
3. Gaussian:
$$K(x, y) = e^{-\|x-y\|^2}$$

The results for the kernelized SVMs are shown in tables 2, 3 and 4.

Table 2: Results for SVM with linear kernel

	TP	FP	TN	FN
1	45	2	239	3
2	42	1	241	7
3	43	3	238	5
4	46	4	237	2
5	45	3	238	3
6	36	0	241	12
7	45	7	234	3
8	45	0	242	3
9	43	0	241	5
10	43	10	231	5
average	43.30	3.00	238.20	4.80
standard deviation	2.87	3.30	3.49	2.94

Table 3: Results for SVM with polynomial kernel

	TP	FP	TN	FN
1	48	4	237	0
2	47	1	241	1
3	47	9	232	1
4	48	10	231	0
5	47	1	241	2
6	45	2	239	3
7	48	8	233	0
8	47	17	224	1
9	48	0	241	0
10	47	1	240	1
average	47.20	5.30	235.90	0.90
standard deviation	0.92	5.54	5.72	0.99

Table 4: Results for SVM with gaussian kernel

	TP	FP	TN	FN
1	46	0	242	2
2	45	1	240	3
3	48	3	238	0
4	48	5	236	0
5	44	0	241	4
6	47	5	236	1
7	47	7	234	1
8	45	2	239	3
9	44	11	230	4
10	46	1	241	3
average	46.00	3.50	237.70	2.10
standard deviation	1.49	3.54	3.74	1.52

Table lists the performance metrics for all of the SVMs, the linear discriminator in assignment 2 and the naive Bayes classifier in assignment 1.

Table 5: Comparison of performance of classifiers

Metric	SVM		Linear SVM		Poly SVM		Gauss SVM		LMS		Naive Bayes	
	μ	σ	μ	σ	μ	σ	μ	σ	μ	σ	μ	σ
TPR(Recall)	0.84	0.08	0.90	0.06	0.98	0.02	0.96	0.03	0.89	0.03	0.20	0.09
FPR	0.01	0.01	0.01	0.01	0.02	0.02	0.01	0.01	0.04	0.01	0.00	0.00
TNR	0.99	0.01	0.99	0.01	0.98	0.02	0.99	0.01	0.96	0.01	1.00	0.00
Precision	0.95	0.06	0.94	0.06	0.91	0.09	0.93	0.06	0.84	0.05	1.00	0.00
Accuracy	0.97	0.02	0.97	0.01	0.98	0.02	0.98	0.01	0.95	0.01	0.87	0.01
F1	0.89	0.05	0.92	0.04	0.94	0.05	0.94	0.03	0.86	0.01	0.33	0.12

References

[1] Cho-Jui Hsieh et al. “A Dual Coordinate Descent Method for Large-scale Linear SVM”. In: *Proceedings of the 25th International Conference on Machine Learning*. ICML ’08. Helsinki, Finland: ACM, 2008, pp. 408–415. ISBN: 978-1-60558-205-4. DOI: 10.1145/1390156.1390208. URL: http://doi.acm.org/10.1145/1390156.1390208.