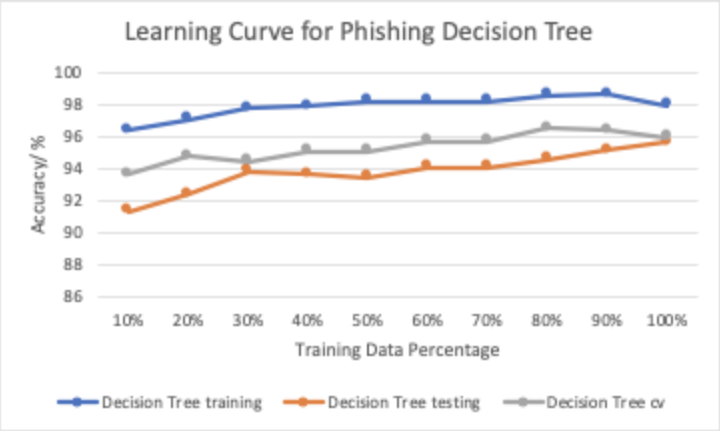
Supervised Learning Models For Classification

Zhang, Shuai
GT ID: zhang429

Phishing Decision Tree

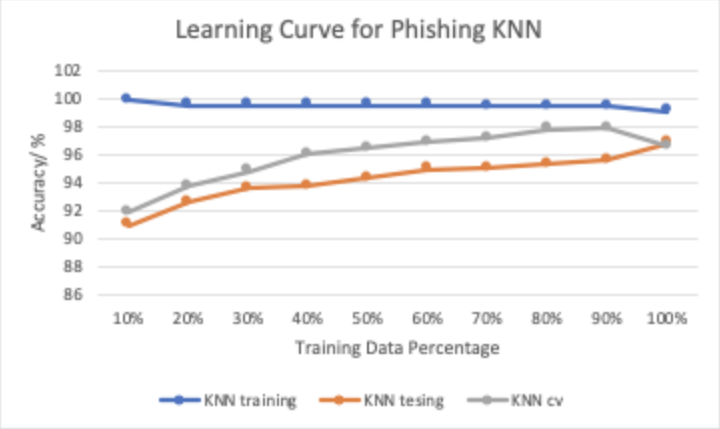
confidenceFactor	pruning	minNumObj	Training accuracy	Testing accuracy	CV accuracy
0.1	Y	2	96.1489	94.8447	95.0762
0.2	Y	2	97.1698	95.4175	95.6319
0.25	Y	2	97.3249	95.5381	95.6061
0.30	Y	2	97.5058	95.4778	95.7483
0.35	Y	2	97.5446	95.4778	95.7870
0.40	Y	2	97.9064	95.6587	95.8129
0.45	**Y**	**2**	**97.9581**	**95.6587**	**95.9421**
0.50	Y	2	98.0098	95.6587	95.8775
0.55	Y	2	98.3717	95.8999	95.7483
0.65	Y	2	98.3717	95.8999	95.7483
0.80	Y	2	98.3717	95.8999	95.7483
0.90	Y	2	98.3717	95.8999	95.7483
N/A	N	2	98.3717	95.9602	95.6707



Phishing KNN

K	Training accuracy	Testing accuracy	CV accuracy

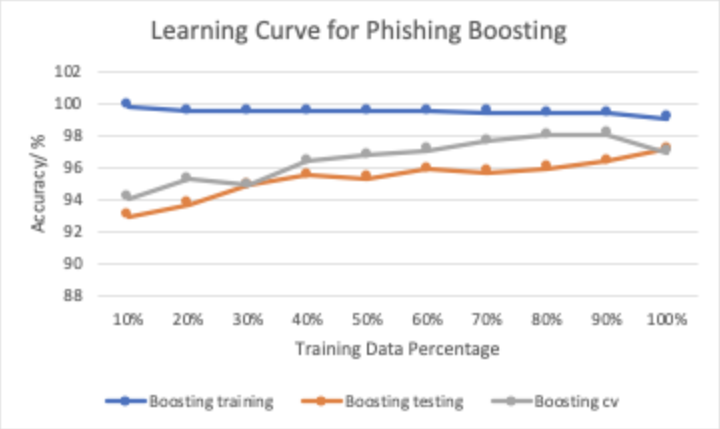
1	**99.0825**	**96.8345**	**96.5753**
3	97.1827	95.5683	94.8695
5	96.0326	95.2970	94.4559
10	94.9082	94.694	94.2621
20	94.2621	93.7594	93.9002
30	93.9778	93.669	93.5255
50	93.1895	92.6138	92.9956
80	92.7113	92.1917	92.5691
100	92.5562	92.0711	92.2978



Phishing Boosting

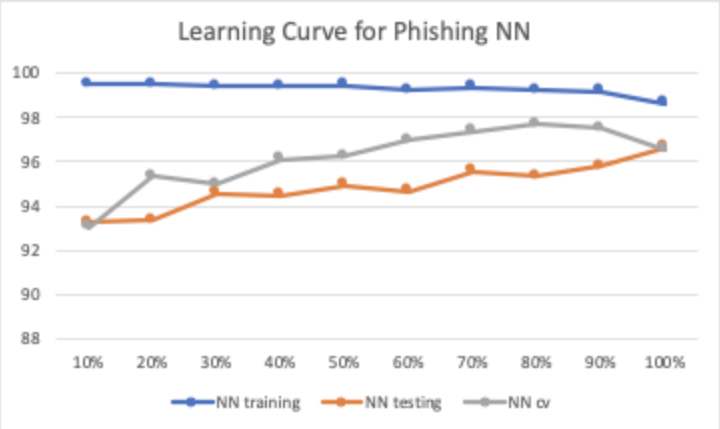
numIterations	Decision Tree confidenceFactor	Training accuracy	Testing accuracy	CV accuracy
10	0.25	99.0825	96.9249	96.7433
10	0.45	99.0825	96.9852	96.8338
20	0.25	99.0825	97.136	96.8726
20	0.45	99.0825	96.8646	96.8597
30	**0.25**	**99.0825**	**97.136**	**96.9889**
30	0.45	99.0825	96.9249	96.8855
40	0.25	99.0825	97.136	96.8726
40	0.45	99.0825	96.9249	96.808

50	0.25	99.0825	97.1058	96.8984
50	0.45	99.0825	96.8948	96.8338



Phishing Neural Networks

learningRate	momentum	hiddenLayers	Training accuracy	Testing accuracy	CV accuracy
0.3	0.2	16	98.824	96.2918	96.1101
0.2	0.1	16	98.8498	96.5632	96.2781
0.1	0.1	16	98.9661	96.7139	96.123
0.2	0.2	16	98.9015	95.9301	96.2135
0.4	**0.3**	**16**	**98.6172**	**96.6235**	**96.5366**
0.5	0.4	16	98.7723	96.744	96.3298
0.5	0.5	16	98.6818	96.7139	96.1747



Phishing SVM

kernelType	Training accuracy	Testing accuracy	CV accuracy
linear	94.2492	93.8499	94.0165
RBF, gamma = 1.0	99.0178	96.9551	95.7999
RBF, gamma = 0.5	**98.6948**	**97.1661**	**96.7046**
RBF, gamma = 0.25	97.8547	96.6235	96.3169
Polynomial, gamma = 0.03, degree = 3	94.3267	93.6087	93.9778
Polynomial, gamma = 0.01, degree = 3	91.419	90.5336	90.8245
Polynomial, gamma = 0.10, degree = 3	97.3378	96.1109	95.9809
Polynomial, gamma = 0.20, degree = 3	98.4492	96.5029	96.472
Polynomial, gamma = 0.40, degree = 3	98.9274	96.5933	96.2264
Polynomial, gamma = 0.70, degree = 3	99.0695	95.9602	96.0972
Polynomial, gamma = 1.0, degree = 3	99.0825	95.7793	96.1489

https://github.com/renatopp/arff-datasets/blob/master/classification/letter.arff

Letter SVM

kernelType	Training accuracy	Testing accuracy	CV accuracy
linear	87.4286	84.1	85.1
RBF, gamma = 0	**99.4429**	**96.7667**	**96.6357**
RBF, gamma = 0.25	99.9857	92.4	92.1786
RBF, gamma = 0.5	100	79.6	79.0929

RBF, gamma = 1.0	100	48.9333	48.05
Polynomial, gamma = 0.03, degree = 3	99.9071	94.8	94.3
Polynomial, gamma = 0.01, degree = 3	98.5929	94.6333	94.5071
Polynomial, gamma = 0.10, degree = 3	100	94.7	94.2286
Polynomial, gamma = 0.20, degree = 3	100	94.7	94.2286
Polynomial, gamma = 0.40, degree = 3	100	94.7	94.2286
Polynomial, gamma = 1.0, degree = 3	100	94.7	94.2286

Letter Neural Networks

learningRate	momentum	hiddenLayers	Training accuracy	Testing accuracy	CV accuracy
0.3	0.2	21	84.1286	81.5333	82.0714
0.2	0.1	21	84.9786	82.6333	81.8429
0.1	**0.1**	**21**	**85.5643**	**82.8167**	**82.2857**
0.4	0.3	21	83.5857	81.1167	82.0786
0.5	0.5	21	82.7857	80.1000	80.3357

Letter Boosting

numIterations	Decision Tree confidenceFactor	Training accuracy	Testing accuracy	CV accuracy
10	0.25			
10	0.45			
20	0.25			
20	0.45			
30	0.25			

30	0.45			
40	0.25			
40	0.45			
50	0.25			
50	0.45			

Letter Decision Tree

confidenceFactor	pruning	minNumObj	Training accuracy	Testing accuracy	CV accuracy
0.1	Y	2			
0.2	Y	2			
0.25	Y	2			
0.30	Y	2			
0.35	Y	2			
0.40	Y	2			
0.45	Y	2			
0.50	Y	2			
0.55	Y	2			
0.65	Y	2			
0.80	Y	2			
0.90	Y	2			
N/A	N	2			

Letter KNN

K	Training accuracy	Testing accuracy	CV accuracy
1			
3			
5			

10			
20			
30			
50			
80			
100			