Gradient descent

Data size = 30			Data size = 80		
Learning rate	Train accuracy	Test accuracy	Learning rate	Train accuracy	Test accuracy
0.010	0.983	0.643	0.0005	0.555	0.423
0.015	0.990	0.687	0.0010	0.779	0.628
0.020	0.983	0.703	0.0050	0.984	0.808
0.025	0.960	0.700	0.0100	0.955	0.795
0.030	0.927	0.673	0.0150	0.856	0.748

Random hill climbing

Data size = 30			Data size = 80		
Learning rate	Train accuracy	Test accuracy	Learning rate	Train accuracy	Test accuracy
0.05	0.54	0.3	0.05	0.451	0.354
0.1	0.92	0.527	0.1	0.825	0.6525
0.15	1	0.63	0.15	0.9575	0.77125
0.2	1	0.657	0.2	0.99375	0.785
0.25	1	0.737	0.25	1	0.78375
0.3	1	0.673	0.3	1	0.80125
0.35	1	0.697	0.35	1	0.775

Simulated annealing

Data size = 30			Data size = 80		
Learning rate	Train accuracy	Test accuracy	Learning rate	Train accuracy	Test accuracy
0.05	0.747	0.37	0.050	0.655	0.480
0.1	1	0.63	0.100	0.961	0.751
0.15	1	0.717	0.150	0.997	0.832
0.2	1	0.717	0.200	1.000	0.840
0.25	1	0.727	0.250	1.000	0.840
0.3	1	0.75	0.300	1.000	0.840
0.35	1	0.74	0.350	1.000	0.833

Genetic algorithm

Data size = 30			Data size = 80		
(Popu, mutP)	Train accuracy	Test accuracy	(Popu, mutP)	Train accuracy	Test accuracy
(20, 0.5)	0.2	0.13	(20, 0.5)	0.204	0.138
(50, 0.1)	0.225	0.172	(50, 0.1)	0.231	0.174
(100, 0.05)	0.23	0.163	(100, 0.05)	0.235	0.171
(200, 0.1)	0.267	0.23	(200, 0.1)	0.268	0.240
(1000, 0.1)	0.291	0.284	(1000, 0.1)	0.295	0.289