

Description

In this experiment, 10 perceptrons were created to work together to predict values from a mnist dataset, training and testing. The perceptrons predicted values for the training data, its accuracy computed, then it predicted the values for the test data, computing the accuracy as well. The perceptrons were then trained, which this was repeated for 50 epochs. After the 50 epochs were done, a confusion matrix was created for the perceptrons running on the test data after training was complete.

Note:

For the confusion matrices:

Predictions that match the actual are in **bold**

False predictions greater than true predictions are highlighted with **red**

False predictions less than true predictions but close are highlight **yellow**

Experiment 1: learning rate = 0.001

Learning Rate = 0.001

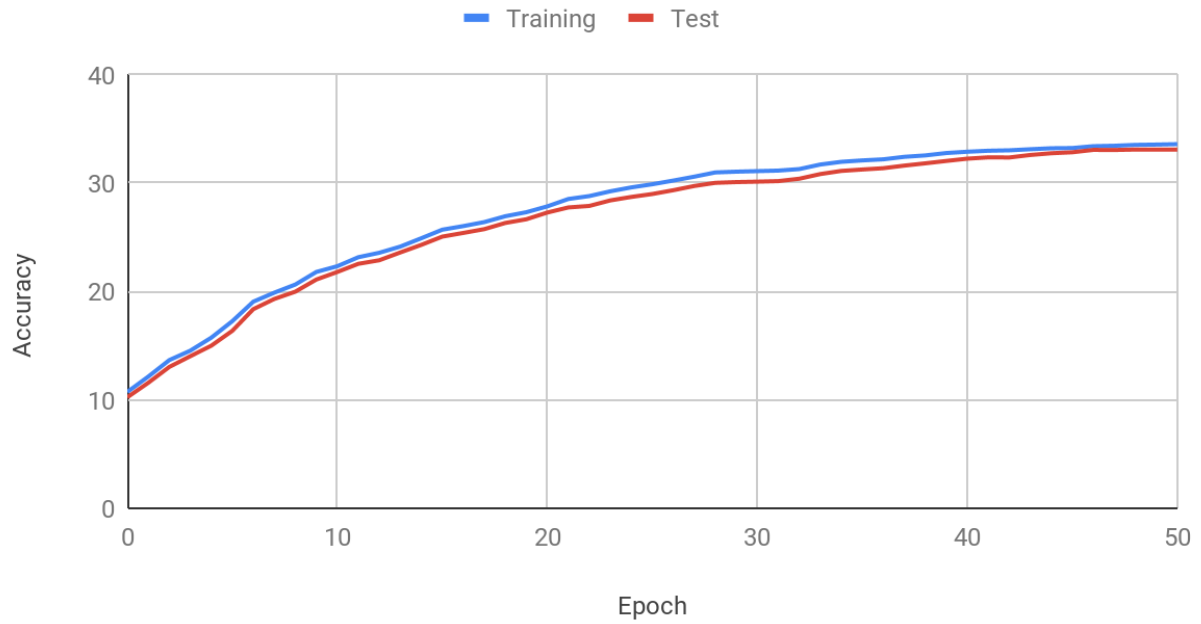


Figure 1: Perceptron chart with a 0.001 learning rate

Looking at the chart above, there seems to be no overfitting of data whatsoever, and no noticeable oscillations and seems to converge at 33 to 34% accuracy, which must be determined with more epochs.

	0	1	2	3	4	5	6	7	8	9
0	437	34	86	88	43	51	88	66	64	23
1	135	556	39	82	14	38	100	50	93	28
2	195	128	206	110	70	29	153	72	34	35
3	165	103	59	347	36	65	54	82	64	35
4	154	90	33	59	240	67	63	93	115	68
5	182	78	51	78	82	144	68	90	80	39
6	162	53	84	33	44	38	434	54	12	66
7	127	28	37	60	110	44	51	355	32	184
8	163	126	43	119	62	74	47	55	226	59
9	128	34	14	67	125	36	32	173	36	364

In the confusion matrix, for most of the predicted values, they tend match rather well with the actual values, except for the actual value of 5, which is somehow confused with the number 0. The value 2 is at the brink of being confused with 2, but all other values are predicting rather well with relatively high prediction values.

Experiment 2: learning rate = 0.01

Learning Rate = 0.01

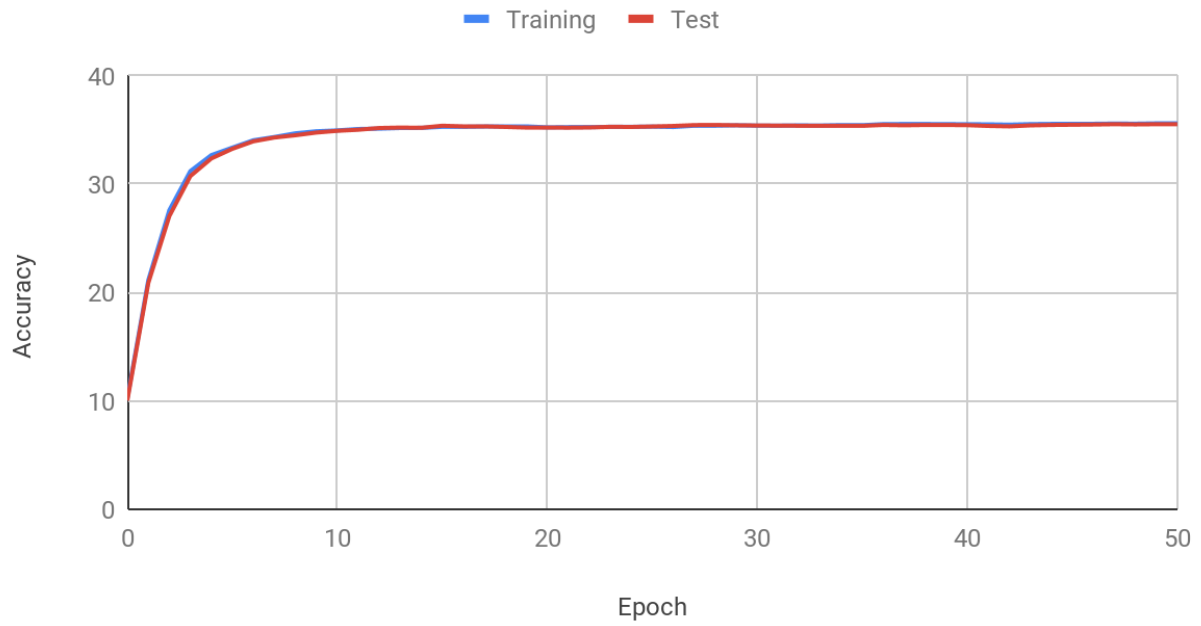


Figure 2: Perceptron chart with a 0.01 learning rate

Looking at the graph above, there seems to be no overfitting but there is an oscillation between 34 and 35%, however, the accuracy overall might increase, even if it's slightly, if more epochs occurred.

	0	1	2	3	4	5	6	7	8	9
0	456	33	69	92	27	27	118	43	75	40
1	142	581	22	67	7	31	100	44	108	33
2	180	139	209	106	37	15	173	58	50	65
3	173	115	51	395	20	41	59	44	58	54
4	156	91	26	70	184	70	76	78	102	129
5	166	79	41	85	68	150	2	73	92	56
6	155	68	48	23	25	24	498	45	19	53
7	123	38	15	97	47	26	61	344	35	242
8	141	126	39	126	45	68	56	38	256	79
9	133	37	11	89	67	24	35	99	35	479

In the confusion matrix above, the actual value 5 is being confused with 0, since there are more occurrences of a 0 prediction with a 5 as the actual. No other false prediction is over or really close to the true predictions.

Experiment 3: learning rate = 0.1

Learning Rate = 0.1

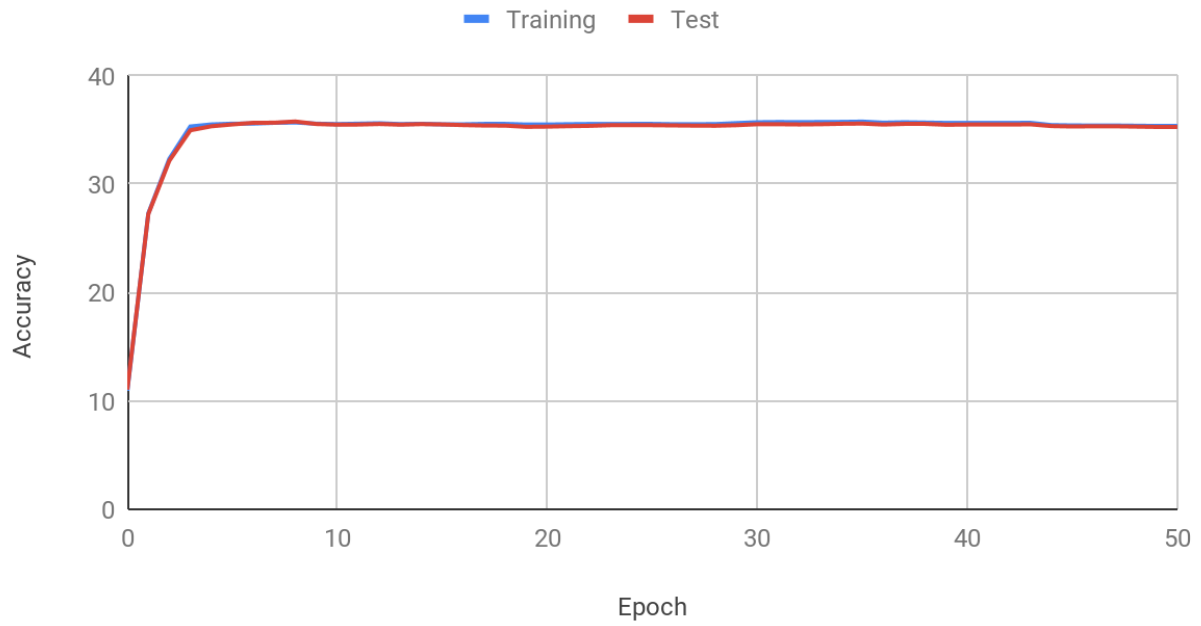


Figure 3: Perceptron chart with a 0.1 learning rate

Looking at the chart above, there seems to be an oscillation between 34 and 35% and the accuracy seems to decrease, but this might not be the case if more epochs occurred, since the learning rate is somewhat high.

	0	1	2	3	4	5	6	7	8	9
0	436	25	104	89	23	54	10	42	71	30
1	116	562	62	58	8	47	93	40	114	35
2	170	133	246	103	44	31	158	56	49	42
3	153	102	72	384	15	64	51	57	64	48
4	164	88	24	67	177	89	67	80	99	127
5	170	78	52	74	64	176	69	76	89	44
6	165	61	60	22	20	36	485	45	21	43
7	124	30	26	97	53	26	55	365	38	214
8	139	122	51	122	42	94	45	44	244	71
9	134	38	15	85	67	31	34	120	33	452

For all true predictions, they are the most predicted out of their respected predictions in their row. However, the 0 prediction for 4 and 5 seem to be close their respected true prediction. In all experiments 5 has been giving the perceptrons a problem in all experiments, however, the true prediction is highest in its row in this experiment.

Overall, In each learning, the value 5 gives the perceptrons the most problem, even though it is not a full confusion when the learning rate is 0.1. The 0.1 learning rate tends to work quite well but it learns too quickly. Maybe if the learning rate reaches a specific accuracy, the learning rate should decrease or increase depending on the accuracy. This is discussion for another time, however it might be beneficial to look into. The plot charts for the 0.01 and 0.1 learning rate learn very quickly but oscillate around their peak, which is in the mid 34-36% respectively. However, the 0.001 learning is a gradually ascending chart, meaning it will eventually reach a peak, or even exceed the 35% peak due to it's more precise learning rate.