Confusion Table 2

Robert Kramer
Perceptron Hw 1
18 JAN 2016

Experiment:

I used python to train perceptrons using the all-pairs method to classify letters from the UCI machine-learning repository. I began exploring the data by importing into R because of familiarity. I found the data was roughly uniformly distributed by letter. I used this information to split the data in half by picking the first 10,000 rows for training and the next 10000 for testing. I switched to python and made use of the pandas, numpy, and itertools libraries. See the attached code for details. After running the all-pairs algorithm I obtained an overall accuracy of 55.78% on the test data. Confusion Table 1 shows the results of this first run. My algorithm is very bad at finding H and G. I tried running with more epochs and varying the eta's, but overall I didn't get qualitatively different answers. I randomly reordered the data and ran the program on a loop to see how much accuracy changed. The accuracy varies between 50% and 64%. When I looked at the confusion tables, the letters completely misclassified changed. I named the Network dataframe with the perceptron name, accuracy, and number of epochs before it stopped improving. Network Figure 1 shows a sample for clarification. When looking at each individual pairwise training set, I found that training accuracy was generally over 95%, but a few were as low as 52%. In the example, it is easy to see the results of the poor performance of the GS perceptron. G was incorrectly classified as S 141 times. These low accuracy perceptrons only ran a few epochs. With this information, I ran the experiment with each perceptron training for 20 epochs and got overall accuracies in the 70%'s as expected. Confusion Table 2 shows my final attempt to maximize accuracy. I ran each perceptron for 20 epochs and shuffled the training data until an overall accuracy on the training data was above 76%. This lead to an overall accuracy of 77.35% on the test data. Confusion Table 2 shows the results.

Confusion Table 1

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z
A	323	6	0	0	0	0	0	1	0	0	2	8	3	0	0	0	7	0	16	8	12	0	1	9	0	0
B	0	291	0	0	0	19	0	3	0	0	0	0	1	0	0	1	0	0	54	0	0	0	0	3	0	0
C	0	2	213	0	9	1	0	0	0	0	23	11	3	0	3	0	3	0	35	47	8	0	0	0	0	0
D	0	201	0	116	0	9	0	4	0	0	0	0	6	0	10	4	0	0	39	16	2	0	0	11	0	0
E	0	9	17	0	108	2	0	1	0	0	1	7	0	0	0	0	11	3	164	40	0	0	0	7	0	0
F	0	9	0	1	1	258	0	0	0	0	1	0	0	0	0	6	0	0	20	97	0	0	2	1	0	0
G	0	35	89	0	0	5	1	15	0	0	4	3	5	0	0	7	70	3	141	11	6	5	3	3	0	0
H	0	78	9	10	0	9	0	77	0	0	5	0	8	30	0	1	4	3	19	32	21	0	0	21	0	0
I	0	4	0	2	2	21	0	0	273	4	0	3	0	0	0	0	0	0	64	3	0	0	0	15	0	0
J	3	5	6	3	0	17	0	1	24	199	0	0	0	0	24	0	1	0	60	2	4	0	0	6	0	0
K	0	37	51	4	13	0	0	2	0	0	143	14	2	0	0	0	2	12	24	9	12	0	0	44	0	0
L	0	4	2	9	5	1	1	2	0	0	2	262	0	0	0	0	4	1	86	1	1	0	0	5	0	0
M	3	10	1	0	0	0	0	11	0	0	9	0	327	3	0	0	0	0	0	0	13	0	5	0	0	0
N	0	6	1	25	0	2	0	40	0	0	5	1	26	269	7	2	0	0	0	2	17	2	2	2	0	0
O	0	44	18	14	0	3	0	23	0	0	0	0	5	0	179	3	18	0	21	14	16	0	14	1	0	0
P	0	11	0	3	0	183	0	3	0	0	1	0	0	1	0	165	0	0	7	11	1	0	2	0	6	0
Q	0	30	2	1	0	2	0	5	0	0	0	3	0	0	4	2	271	0	68	8	3	9	5	0	0	0
R	0	186	2	2	0	0	0	6	0	0	12	6	7	0	3	0	0	143	10	0	0	0	0	17	0	0
S	0	12	1	0	0	16	0	0	0	0	0	2	0	0	0	0	3	0	340	19	0	0	0	0	0	0
T	0	3	1	0	2	8	0	0	0	0	6	0	1	0	0	8	0	0	11	328	1	0	0	0	0	0
U	0	2	0	0	0	0	0	1	0	0	3	1	15	0	1	0	0	1	0	24	348	0	11	0	0	0
V	0	8	1	0	0	5	0	5	0	0	0	0	3	0	0	15	0	0	1	19	25	282	12	0	6	0
W	0	4	0	0	0	0	0	4	0	0	0	0	36	1	0	1	0	0	0	0	5	6	341	0	0	0
X	0	9	0	0	8	3	0	2	0	0	0	4	0	0	0	0	12	0	147	31	1	0	0	171	0	0
Y	0	1	0	0	0	15	0	6	0	0	0	3	1	0	0	16	2	0	21	180	9	62	1	1	60	0
Z	0	3	0	0	17	5	0	0	0	3	0	0	0	0	0	0	0	0	226	32	0	0	0	0	0	90

Confusion Table 2

Network Table 1 with pairwise perceptron name, accuracy of pair prediction, and number of epochs.

GS 0.54 2	GT 0.91 2	GU 0.97 3	GV 0.97 6	GW 0.98 2	GX 0.96 3
-0.00747396801932	0.117650028934	0.114445259908	0.135995754692	-0.563158592391	-0.97870853406
-0.108915357056	0.402999401915	-0.112370760567	0.162416206815	-1.02838824802	0.636161683716
-0.813467691625	-0.410909634351	-0.207861869999	0.133892786002	0.722848860803	0.483470815993
0.271631205579	0.0580062003307	-0.62171314126	0.118671756388	0.111720102458	0.37674573966
0.336832832139	-0.345709741267	-0.101027527428	0.59692631826	-0.0734007045608	-0.83205733432
-0.571059430401	-0.298064265061	-0.335990056921	-1.38553518389	-0.0021394262604	0.704800982553
-1.90194998459	-0.0471448156566	-0.535238840841	0.0553460784725	0.377001070514	-1.26490319441
-0.577557644262	-1.71789103672	-0.959102103158	-1.2839360107	-0.775300929239	-1.33680071035
1.93581634206	1.28589997128	-0.423204533963	1.45341384882	0.445460854801	3.38159596187
-1.57736062577	-1.60083286992	-0.548858753066	-0.489069164791	0.558717911086	-1.50670091939
-0.951986818749	-0.00519244150905	0.168342318621	0.259993793315	0.644950592483	0.669671079139
-0.323707729755	-1.4376461305	-1.48227761499	-1.86783018094	-0.655122296414	-0.492521626068
1.91519205069	1.15268078938	0.736498841082	1.9592673289	0.897105173709	0.591401613035
1.62686708776	-0.0441541851053	-1.58974900544	-1.67232262244	-2.46386499546	-0.333365044025
0.0213717524855	0.10727330365	-0.476814119792	-0.714524666796	0.235022663576	-0.0393943071245
-1.8742507693	-0.399916623917	2.53600305353	1.47013594461	1.74714597218	0.408338456889
0.812582187992	1.50986747721	2.09247954615	1.52966897971	0.595420878314	1.51017587998

Confusion Table 2

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z
A	338	0	0	5	0	0	3	2	0	11	6	0	0	0	3	0	2	7	5	1	5	2	0	1	4	1
B	0	202	0	24	3	0	0	4	30	7	0	0	0	0	0	0	4	39	54	0	0	2	0	2	1	0
C	1	0	199	1	39	0	62	0	1	0	19	10	1	0	10	0	0	0	4	4	7	0	0	0	0	0
D	1	10	0	375	0	0	0	3	0	4	0	1	0	1	4	1	0	4	1	7	5	0	0	1	0	0
E	0	0	1	0	292	0	13	0	2	1	4	8	0	0	0	0	2	13	15	6	0	0	0	10	0	3
F	0	2	4	5	11	218	2	14	7	29	1	1	0	1	0	29	0	6	23	36	0	1	1	4	1	0
G	0	0	8	13	6	0	265	7	1	0	3	10	3	0	2	3	41	16	18	0	1	5	4	0	0	0
H	0	2	0	45	2	0	10	124	0	13	14	1	2	7	48	0	3	31	0	3	8	4	0	8	2	0
I	0	0	1	8	0	4	0	0	324	23	0	2	0	0	1	0	0	0	17	0	0	0	0	10	0	1
J	0	0	0	7	0	0	0	2	22	311	0	0	0	0	5	0	1	0	4	0	0	0	0	2	0	1
K	0	1	1	10	3	0	18	1	0	0	262	12	0	0	5	0	0	42	0	0	3	0	0	11	0	0
L	0	0	0	5	15	0	7	0	1	0	0	337	0	0	0	0	6	4	4	0	0	0	0	7	0	0
M	4	4	0	0	0	0	1	4	0	0	0	0	323	10	0	0	0	8	0	0	16	0	12	0	0	0
N	1	1	0	40	0	0	0	22	0	2	10	0	5	301	2	0	0	4	0	0	8	9	1	0	3	0
O	1	0	0	42	0	0	8	7	0	2	0	1	4	0	275	2	7	6	0	4	1	0	13	0	0	0
P	0	1	0	11	2	17	6	3	3	9	0	1	0	0	8	314	2	1	5	1	0	0	0	0	10	0
Q	1	0	0	5	0	0	18	0	3	9	0	11	0	0	19	3	289	5	37	0	0	3	3	1	4	2
R	2	14	0	13	1	0	1	8	0	0	11	4	1	2	6	0	0	326	1	1	0	2	0	1	0	0
S	0	4	1	1	10	0	3	0	21	12	0	15	0	0	1	0	5	7	271	7	0	0	0	5	2	28
T	0	1	0	1	2	0	4	0	2	0	4	3	0	0	0	2	0	4	4	332	0	0	0	5	1	4
U	2	0	0	8	0	0	2	5	0	1	5	0	4	0	12	0	0	1	0	55	300	1	10	0	1	0
V	0	2	0	0	0	0	2	5	0	1	0	1	1	0	1	1	0	6	0	3	1	342	8	0	8	0
W	0	0	0	0	0	0	1	7	0	0	0	0	9	11	3	0	0	5	0	0	7	7	348	0	0	0
X	0	1	0	10	7	0	4	0	7	7	15	2	0	0	1	0	1	2	5	1	1	0	0	322	2	0
Y	1	0	0	3	0	8	0	0	0	4	0	0	1	0	0	0	6	0	0	31	1	33	1	6	283	0
Z	0	0	0	0	11	0	0	0	5	35	0	2	0	0	0	0	2	0	47	5	0	0	0	5	0	264