

# Programming Assignment 5 - Part 2

## Section 1

Team Members: Oleg Tielushko (U48136789), Tyler Brown.

Then it should include the results that WEKA provides when you use data set 1 to train the network selecting the “percentage split” option in the “test options” window and setting the percentage split to 80.

The following items are to be included and clearly described in this section 1:

a. The overall accuracy achieved by the trained network.

84.85%

=== Summary ===

Correctly Classified Instances	13576	84.85	%
Incorrectly Classified Instances	2424	15.15	%
Kappa statistic	0.8424		
Mean absolute error	0.0138		
Root mean squared error	0.1		
Relative absolute error	18.7112	%	
Root relative squared error	51.9927	%	
Total Number of Instances	16000		

b. The accuracy by which each letter was classified. Use a format similar to what you used in Part 1.

Letter	Percentage correct from Training Data
A	90.00%
B	94.40%
C	89.20%
D	87.20%
E	80.40%

F	85.00%
G	66.70%
H	73.60%
I	86.30%
J	78.40%
K	83.40%
L	82.90%
M	95.00%
N	86.50%
O	80.80%
P	88.40%
Q	85.80%
R	83.80%
S	70.20%
T	78.40%
U	91.20%
V	90.70%
W	92.90%
X	89.00%
Y	91.30%
Z	81.70%
Average	84.74%

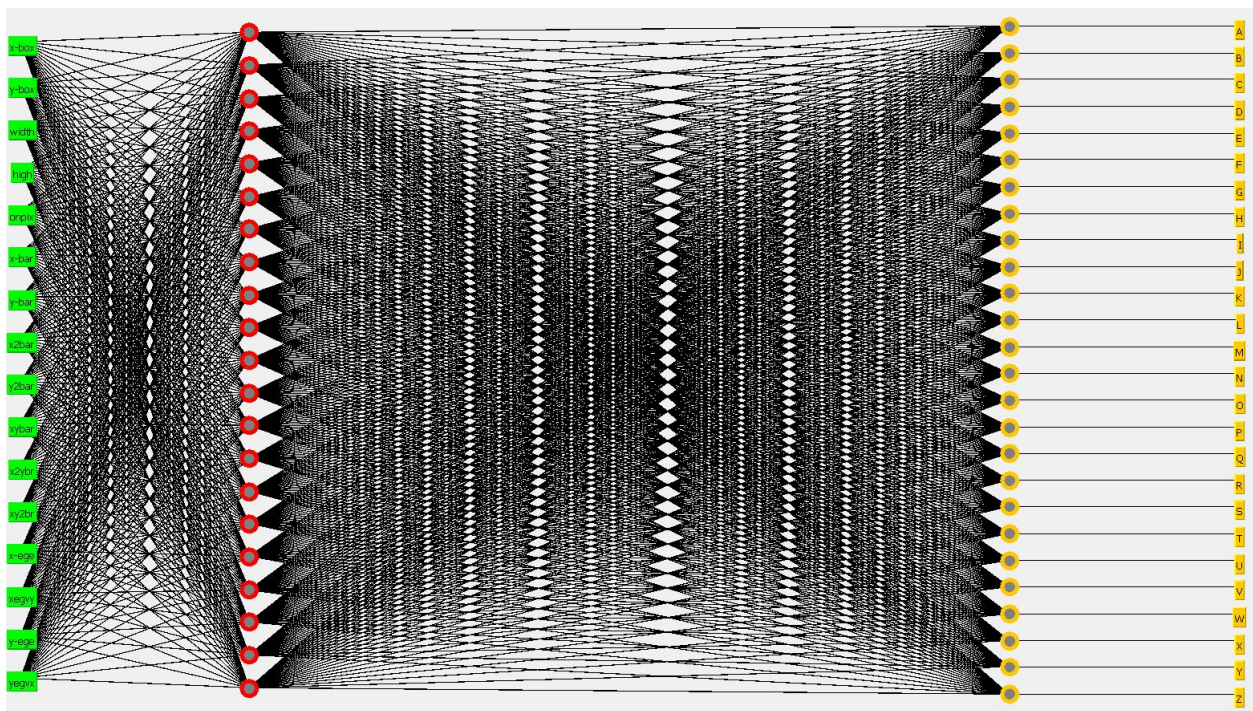
Here is the detailed matrix of classification

	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	<-- classified as	
561	2	2	2	0	0	0	0	0	0	0	0	4	11	0	0	1	3	1	7	0	5	2	3	0	12	7		a = A
0	572	0	5	4	2	1	2	0	0	1	0	0	0	0	1	0	1	8	3	0	0	1	0	3	0	2		b = B
0	0	521	0	5	0	11	3	1	0	10	1	4	0	12	0	9	2	2	0	3	0	0	0	0	0	0		c = C
2	15	0	570	0	0	0	6	4	0	3	1	7	2	8	0	0	17	7	0	4	0	0	2	2	4		d = D	
2	14	18	1	488	18	2	0	0	1	1	0	0	0	0	1	22	3	7	0	0	0	0	12	0	17		e = E	
0	16	1	4	5	525	0	2	2	0	0	0	0	0	0	27	1	0	8	12	0	0	4	7	0	4		f = F	
25	13	10	10	3	8	402	1	1	0	11	5	2	0	4	15	48	20	10	0	2	3	4	4	2	0		g = G	
2	19	9	17	0	5	2	424	0	1	5	0	2	1	18	1	15	31	0	0	3	14	2	0	5	0		h = H	
3	4	1	11	2	2	0	0	512	0	0	6	0	0	0	6	0	0	17	6	0	0	0	17	6	0		i = I	
13	1	2	7	1	29	0	7	20	469	0	0	0	0	0	12	3	2	14	2	1	0	0	7	8	0		j = J	
3	7	5	2	2	1	2	15	1	0	491	2	2	5	0	0	6	23	0	0	3	0	1	7	10	1		k = K	
0	4	10	2	6	0	9	0	0	0	3	505	0	1	2	0	36	7	5	0	0	0	0	16	2	1		l = L	
2	8	2	5	0	1	0	0	0	0	0	0	621	0	0	2	0	10	0	0	0	0	3	0	0	0		m = M	
1	9	1	7	0	0	0	5	0	0	0	0	12	533	4	1	6	11	0	1	5	1	13	0	6	0		n = N	
0	3	9	15	0	1	2	5	0	1	0	3	0	1	485	1	46	6	0	0	7	2	13	0	0	0		o = O	
1	18	0	5	1	21	0	0	4	0	0	0	0	0	4	582	10	1	1	0	1	1	2	0	6	0		p = P	
2	7	5	8	22	0	2	1	5	0	0	2	0	0	12	0	563	6	4	0	0	1	0	2	0	14		q = Q	
7	36	0	15	1	1	1	6	0	0	8	0	2	0	1	3	11	501	0	0	0	0	0	4	1	0		r = R	
12	17	4	21	15	10	0	12	14	5	2	12	0	0	2	5	7	11	425	0	0	2	0	6	5	18		s = S	
0	6	1	13	14	21	4	15	4	1	4	1	0	0	3	0	11	1	4	504	2	2	0	6	11	15		t = T	
5	3	5	1	0	1	0	10	0	0	1	2	11	0	2	0	5	6	0	0	593	1	1	2	1	0		u = U	
0	17	1	2	0	2	0	1	0	0	0	0	1	2	0	6	4	4	0	0	2	555	9	1	5	0		v = V	
7	14	0	0	0	0	0	0	0	0	0	0	10	0	0	1	0	8	0	0	2	2	573	0	0	0		w = W	
0	2	6	4	3	0	0	1	3	4	17	1	0	0	2	4	0	3	6	2	4	0	0	549	3	3		x = X	
0	2	1	1	0	9	0	1	0	0	0	3	1	0	0	2	8	0	0	3	3	17	1	3	575	0		y = Y	
1	9	3	0	12	11	0	0	0	0	0	5	0	0	2	0	9	3	46	0	0	0	0	0	6	477		z = Z	

c. The number of layers in the network and the number of neurons in each layer.

The Neural Network consists of the:

- Input Layer (Attributes) (16 nodes due to 16 attributes of each letter).
- Hidden Layer: Node 26 through Node 46 (21 nodes total)
- Output Layer: Node 0 through Node 25 (26 nodes total)



d. Save the network in case this is the best network you will find.

DONE!

e. A paragraph discussing the results obtained—compare results obtained with this method vs those obtained in PART 1—for example “is the architecture similar?”, “accuracy” “best and worse accuracies on letters” etc.

WEKA constructed the network with a similar architecture: the number of neurons and layers was the same. After analyzing the results obtained from Part 1 with 85% data set dedicated to training, and Part 2 with 80% data set dedicated to training, it is easy to point out that the percentages of correctly identified items are approximately the same. The difference in most cases does not come over the 5% between the letters in sets. A big thing to point out too, is the fact that the most recognized and the least recognized letters are the same in both of the data sets. For example the difference between the most recognized letters A and M in our sets differs by around 1-2% and the difference between least recognized letter G is 2%. This experiment shows us that the percent of the data we devote to training at some point will stop producing better results for the network. Therefore, there needs to be an optimal amount of data records that will allow us to produce the best result in overall accuracy.

## Section 2

Section 2 should include the results that WEKA provides when data set 2 records are tested on the network described in section 1 and saved. The following results are to be included and clearly described in this section:

f. The overall accuracy achieved by the network on this test file  
83.56%

```
=== Summary ===
```

Correctly Classified Instances	3341	83.525	%
Incorrectly Classified Instances	659	16.475	%
Kappa statistic	0.8286		
Mean absolute error	0.0148		
Root mean squared error	0.1047		
Total Number of Instances	4000		

g. The accuracy by which each letter was classified. Use a format similar to what you used in sections 1.

Letter	Percentage correct from Testing Data
A	89.2%
B	95.0%
C	82.9%
D	84.1%
E	74.5%
F	79.0%
G	79.4%
H	67.1%
I	85.8%
J	74.5%
K	78.0%
L	91.4%
M	94.9%
N	81.4%
O	78.4%
P	86.2%
Q	84.3%
R	85.0%
S	62.9%
T	79.1%
U	89.6%
V	91.4%
W	93.3%
X	88.8%
Y	92.3%
Z	83.3%
Average	83.5%

=== Confusion Matrix ===

	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	<-- classified as	
148	0	1	1	0	0	0	0	0	0	0	0	1	4	0	0	0	1	0	3	0	1	0	0	0	4	2		a = A
0	152	0	5	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	2	0	0	0	0	0	0	0		b = B
1	0	126	0	3	0	4	0	0	0	4	0	4	0	2	0	5	0	0	0	3	0	0	0	0	0	0		c = C
0	8	0	127	0	1	0	1	1	1	0	1	2	0	4	0	0	0	1	0	1	0	1	0	1	0	2		d = D
0	4	8	0	120	2	1	0	0	0	2	0	0	0	0	0	0	7	3	1	0	0	0	0	5	0	8		e = E
1	3	1	0	3	124	0	0	2	0	0	0	0	1	0	10	0	0	1	5	0	0	3	2	1	0	0		f = F
4	3	4	3	0	0	135	0	0	0	1	2	0	0	1	1	12	1	1	0	0	1	1	0	0	0	0		g = G
1	5	3	5	1	9	0	106	0	0	3	1	0	0	4	1	5	6	0	0	1	3	0	1	3	0	0		h = H
0	2	0	6	0	1	0	0	139	1	1	2	0	0	0	0	2	0	1	1	0	0	0	3	3	0	0		i = I
1	0	0	2	0	13	0	2	5	111	0	0	0	0	0	0	2	0	0	3	0	0	0	0	6	3	1		j = J
0	5	0	0	0	0	3	6	0	0	117	0	0	1	2	0	2	6	0	0	4	0	0	3	1	0	0		k = K
0	2	0	0	2	0	0	0	0	0	1	139	0	0	1	0	4	0	1	0	1	0	0	0	1	0	0		l = L
2	3	0	1	0	0	0	0	0	0	0	0	131	0	0	0	0	1	0	0	0	0	0	0	0	0	0		m = M
0	3	0	3	0	0	0	1	0	0	1	0	5	136	3	0	1	4	0	0	2	0	7	0	1	0	0		n = N
0	1	6	5	0	0	1	1	0	0	0	1	0	1	120	0	7	2	0	0	3	2	3	0	0	0	0		o = O
0	3	0	0	0	4	0	1	0	0	1	0	0	0	1	125	3	0	0	0	0	1	3	0	3	0	0		p = P
0	5	0	0	4	0	1	0	0	0	1	1	0	0	5	0	107	0	0	0	0	0	0	1	0	2	0		q = Q
1	10	0	3	0	0	0	0	0	0	0	0	1	1	1	1	5	136	0	0	0	0	0	1	0	0	0		r = R
4	6	0	4	5	3	0	2	2	5	0	6	0	0	1	1	0	5	90	2	0	0	0	1	1	5	0		s = S
0	1	0	2	3	6	1	1	0	0	0	0	0	0	0	0	1	2	0	121	0	3	0	3	6	3	0		t = T
2	1	1	0	0	0	0	3	0	0	0	0	6	0	2	0	1	0	0	0	146	0	0	1	0	0	0		u = U
1	3	0	2	0	0	0	1	0	0	1	0	0	0	0	0	0	1	0	0	1	139	2	1	0	0	0		v = V
1	5	0	0	0	0	0	1	0	0	0	0	2	0	0	0	0	0	0	0	0	0	0	126	0	0	0		w = W
0	0	0	2	0	0	0	0	1	1	7	1	0	0	0	0	0	1	1	2	0	0	0	151	2	1	0		x = X
0	0	0	0	0	3	0	0	0	0	0	1	0	0	0	0	3	0	0	2	0	3	0	0	144	0	0		y = Y
0	0	0	0	3	4	0	0	1	0	0	2	0	0	0	0	2	1	9	0	0	0	0	0	3	125	0		z = Z

h. A paragraph discussing the results obtained -- compare results obtained with this method vs those obtained in PART 1.

Surprisingly enough, a lot of the letters do not have the similar percent accuracy. A drastic difference can be seen in letters B, E, Q which have over 10% accuracy differences. Other than that, it is also important to point out that some of the most and least recognized letters are similar in both Part 1 and Part 2.