Project 5: Distance-Based Learning

Paul Beggs
GitHub Link

November 14, 2025

1 K-Nearest Neighbors Analysis

For Tables 1 through 3 (Appendix A), we can see that the KNN algorithm has perfect accuracy for 'H' for all partitions and values of k. The 'A' has near-perfect accuracy when $k = 3, 5$, but misses one in the first, second, and fourth partition when $k = 7$. I believe this discrepancy is because the 'H's are more distinguishable than the 'A's because when I drew them, there was less variation. That is, the 'A's outer connecting lines can be longer, shorter, wider, or more constricted. Whereas, the 'H's tended to all look more alike, and therefore had less variation. It also seems like when k is smaller, we get a better percentage correct.

Then, for Tables 4 through 6 (Appendix B), we see a similar result to the first 3 tables. 'E' had a perfect record for each value of k, and for each permutation. 'H' had similar results, but slightly fumbled when $k = 5$ for the first and fourth permutation. Finally, 'A' performed well, but still had problems across each value of k. I chalk these results up to a similar conclusion to the previous paragraph: while 'E' does have similar characteristics to 'A' and 'H' (specifically, the middle bridging part and the left straight edge), but because of the other 2 straight lines in the 'E,' it's enough to differentiate it from the other letters.

For the remaining tables (Appendix C), some letters perform significantly better than others. Like, 'T' and 'V' have perfect accuracy, but 'A,' 'H,' 'E,' and 'Q' all perform poorly. It's difficult to nail down if there is any meaningful change to be reported from a change in values of k. However, just from looking, it would appear that when $k = 7$, the algorithm performs the worst, $k = 5$ performs the best, and $k = 3$ is in the middle.

2 Self-Organizing Map Analysis

3 Comparison of Both Distance-Based Methods

Appendix A: KNN With Two Letters

	1st Partition			2nd Partition			3rd Partition			4th Partition		
Lbl.	**Ex.**	**Cor.**	**%**	**Ex.**	**Cor.**	**%**	**Ex.**	**Cor.**	**%**	**Ex.**	**Cor.**	**%**
A	5	5	100%	6	6	100%	5	4	80%	6	6	100%
H	5	5	100%	4	4	100%	5	5	100%	3	3	100%

Table 1: KNN_3 4-Way Cross Validation Results With Two Letters

	1st Partition			2nd Partition			3rd Partition			4th Partition		
Lbl.	**Ex.**	**Cor.**	**%**	**Ex.**	**Cor.**	**%**	**Ex.**	**Cor.**	**%**	**Ex.**	**Cor.**	**%**
A	4	4	100%	6	5	83%	5	5	100%	5	5	100%
H	6	6	100%	4	4	100%	5	5	100%	4	4	100%

Table 2: KNN_5 4-Way Cross Validation Results With Two Letters

	1st Partition			2nd Partition			3rd Partition			4th Partition		
Lbl.	**Ex.**	**Cor.**	**%**	**Ex.**	**Cor.**	**%**	**Ex.**	**Cor.**	**%**	**Ex.**	**Cor.**	**%**
A	5	4	80%	6	5	83%	3	3	100%	5	4	80%
H	5	5	100%	4	4	100%	7	7	100%	4	4	100%

Table 3: KNN_7 4-Way Cross Validation Results With Two Letters

Appendix B: KNN With Three Letters

	1st Partition			2nd Partition			3rd Partition			4th Partition		
Lbl.	**Ex.**	**Cor.**	**%**	**Ex.**	**Cor.**	**%**	**Ex.**	**Cor.**	**%**	**Ex.**	**Cor.**	**%**
A	6	4	67%	4	4	100%	4	4	100%	5	5	100%
E	5	5	100%	7	7	100%	6	6	100%	2	2	100%
H	4	4	100%	4	4	100%	5	5	100%	7	7	100%

Table 4: KNN_3 4-Way Cross Validation Results With Three Letters

	1st Partition			2nd Partition			3rd Partition			4th Partition		
Lbl.	**Ex.**	**Cor.**	**%**	**Ex.**	**Cor.**	**%**	**Ex.**	**Cor.**	**%**	**Ex.**	**Cor.**	**%**
A	4	4	100%	7	7	100%	5	4	80%	3	3	100%
E	4	4	100%	3	3	100%	6	6	100%	5	5	100%
H	7	6	86%	3	3	100%	4	4	100%	6	5	83%

Table 5: KNN_5 4-Way Cross Validation Results With Three Letters

	1st Partition			2nd Partition			3rd Partition			4th Partition		
Lbl.	**Ex.**	**Cor.**	**%**	**Ex.**	**Cor.**	**%**	**Ex.**	**Cor.**	**%**	**Ex.**	**Cor.**	**%**
A	3	3	100%	4	4	100%	6	6	100%	6	4	67%
E	7	7	100%	5	5	100%	4	4	100%	4	4	100%
H	5	5	100%	6	6	100%	5	5	100%	4	4	100%

Table 6: KNN_7 4-Way Cross Validation Results With Three Letters

Appendix C: KNN With Eight Letters

	1st Partition			2nd Partition			3rd Partition			4th Partition		
Lbl.	**Ex.**	**Cor.**	**%**	**Ex.**	**Cor.**	**%**	**Ex.**	**Cor.**	**%**	**Ex.**	**Cor.**	**%**
A	4	4	100%	3	3	100%	7	6	85%	6	4	67%
E	6	3	50%	3	1	33%	7	6	85%	4	3	75%
H	4	3	75%	5	4	80%	4	3	75%	7	1	14%
L	7	7	100%	7	7	100%	1	1	100%	5	5	100%
Q	8	6	75%	1	1	100%	6	4	67%	5	4	80%
T	3	3	100%	5	5	100%	7	7	100%	5	5	100%
V	4	4	100%	5	5	100%	5	5	100%	6	6	100%
Z	4	4	100%	11	9	81%	3	3	100%	2	2	100%

Table 7: KNN_3 4-Way Cross Validation Results With Eight Letters

	1st Partition			2nd Partition			3rd Partition			4th Partition		
Lbl.	**Ex.**	**Cor.**	**%**	**Ex.**	**Cor.**	**%**	**Ex.**	**Cor.**	**%**	**Ex.**	**Cor.**	**%**
A	3	3	100%	5	4	80%	8	6	75%	4	4	100%
E	5	4	80%	3	2	67%	4	3	75%	8	5	63%
H	5	1	20%	4	2	50%	4	3	75%	7	4	57%
L	2	2	100%	3	3	100%	7	4	57%	8	8	100%
Q	9	3	33%	5	4	80%	3	3	100%	3	3	100%
T	5	5	100%	7	7	100%	4	4	100%	4	4	100%
V	4	4	100%	6	6	100%	7	7	100%	3	3	100%
Z	7	7	100%	7	7	100%	3	3	100%	3	3	100%

Table 8: KNN_5 4-Way Cross Validation Results With Eight Letters

	1st Partition			2nd Partition			3rd Partition			4th Partition		
Lbl.	**Ex.**	**Cor.**	**%**	**Ex.**	**Cor.**	**%**	**Ex.**	**Cor.**	**%**	**Ex.**	**Cor.**	**%**
A	3	3	100%	7	5	71%	9	5	56%	1	1	100%
E	2	1	50%	5	3	60%	7	4	57%	6	6	100%
H	8	8	100%	2	2	80%	4	2	50%	7	2	29%
L	7	3	43%	5	5	100%	2	2	100%	5	5	100%
Q	6	4	67%	7	5	71%	2	1	50%	5	4	80%
T	7	7	100%	1	1	100%	7	7	100%	5	5	100%
V	4	4	100%	7	7	100%	2	2	100%	6	6	100%
Z	3	3	100%	6	6	100%	7	6	86%	2	2	100%

Table 9: KNN_7 4-Way Cross Validation Results With Eight Letters

Appendix D: SOM With Two Letters

	1st Partition			2nd Partition			3rd Partition			4th Partition		
Lbl.	**Ex.**	**Cor.**	**%**	**Ex.**	**Cor.**	**%**	**Ex.**	**Cor.**	**%**	**Ex.**	**Cor.**	**%**
A	3	3	100%	7	6	86%	6	4	67%	3	3	100%
H	7	5	71%	3	3	100%	4	4	100%	6	6	100%

Table 10: SOM_6 4-Way Cross Validation Results With Two Letters

	1st Partition			2nd Partition			3rd Partition			4th Partition		
Lbl.	**Ex.**	**Cor.**	**%**	**Ex.**	**Cor.**	**%**	**Ex.**	**Cor.**	**%**	**Ex.**	**Cor.**	**%**
A	7	6	86%	3	2	67%	6	2	33%	3	3	100%
H	3	3	100%	7	7	100%	4	3	75%	6	5	83%

Table 11: SOM_8 4-Way Cross Validation Results With Two Letters

	1st Partition			2nd Partition			3rd Partition			4th Partition		
Lbl.	**Ex.**	**Cor.**	**%**	**Ex.**	**Cor.**	**%**	**Ex.**	**Cor.**	**%**	**Ex.**	**Cor.**	**%**
A	6	2	33%	3	3	100%	6	3	50%	4	4	100%
H	4	4	100%	7	7	100%	4	4	100%	5	5	100%

Table 12: SOM_{10} 4-Way Cross Validation Results With Two Letters

Appendix E: SOM With Three Letters

	1st Partition			2nd Partition			3rd Partition			4th Partition		
Lbl.	**Ex.**	**Cor.**	**%**	**Ex.**	**Cor.**	**%**	**Ex.**	**Cor.**	**%**	**Ex.**	**Cor.**	**%**
A	5	1	20%	6	4	67%	5	0	0%	3	1	33%
E	4	4	100%	4	4	100%	4	4	100%	8	8	100%
H	6	6	100%	5	4	80%	6	6	100%	3	3	100%

Table 13: SOM_6 4-Way Cross Validation Results With Three Letters

	1st Partition			2nd Partition			3rd Partition			4th Partition		
Lbl.	**Ex.**	**Cor.**	**%**	**Ex.**	**Cor.**	**%**	**Ex.**	**Cor.**	**%**	**Ex.**	**Cor.**	**%**
A	6	1	17%	6	5	83%	4	4	100%	3	3	100%
E	5	5	100%	5	5	100%	6	6	100%	4	4	100%
H	4	4	100%	4	4	100%	5	5	100%	7	3	43%

Table 14: SOM_8 4-Way Cross Validation Results With Three Letters

	1st Partition			2nd Partition			3rd Partition			4th Partition		
Lbl.	**Ex.**	**Cor.**	**%**	**Ex.**	**Cor.**	**%**	**Ex.**	**Cor.**	**%**	**Ex.**	**Cor.**	**%**
A	5	3	60%	4	3	75%	3	1	33%	7	1	14%
E	6	6	100%	4	4	100%	8	7	88%	2	2	100%
H	4	2	50%	7	4	57%	4	4	100%	5	4	80%

Table 15: SOM_{10} 4-Way Cross Validation Results With Three Letters

Appendix F: SOM With Eight Letters

	1st Partition			2nd Partition			3rd Partition			4th Partition		
Lbl.	**Ex.**	**Cor.**	**%**	**Ex.**	**Cor.**	**%**	**Ex.**	**Cor.**	**%**	**Ex.**	**Cor.**	**%**
A	4	0	0%	4	1	25%	6	3	50%	6	2	33%
E	4	1	25%	5	2	40%	3	2	67%	8	0	0%
H	6	0	0%	6	0	0%	3	0	0%	5	0	0%
L	4	1	25%	6	6	100%	8	4	50%	2	2	100%
Q	6	3	50%	5	0	0%	4	3	75%	5	2	40%
T	6	6	100%	6	6	100%	4	4	100%	4	4	100%
V	3	3	100%	2	2	100%	8	8	100%	7	7	100%
Z	7	0	0%	6	5	83%	4	4	100%	3	3	100%

Table 16: SOM_6 4-Way Cross Validation Results With Eight Letters

	1st Partition			2nd Partition			3rd Partition			4th Partition		
Lbl.	**Ex.**	**Cor.**	**%**	**Ex.**	**Cor.**	**%**	**Ex.**	**Cor.**	**%**	**Ex.**	**Cor.**	**%**
A	7	0	0%	6	1	17%	2	2	100%	5	0	0%
E	2	1	50%	6	5	83%	9	4	44%	3	2	67%
H	4	2	50%	4	0	0%	6	0	0%	6	1	17%
L	4	4	100%	8	8	100%	4	4	100%	4	4	100%
Q	7	3	43%	3	2	67%	6	2	33%	4	2	50%
T	8	8	100%	3	3	100%	3	3	100%	6	6	100%
V	2	2	100%	5	3	60%	6	5	83%	7	7	100%
Z	6	1	17%	5	4	80%	4	2	50%	5	5	100%

Table 17: SOM_8 4-Way Cross Validation Results With Eight Letters

	1st Partition			2nd Partition			3rd Partition			4th Partition		
Lbl.	**Ex.**	**Cor.**	**%**	**Ex.**	**Cor.**	**%**	**Ex.**	**Cor.**	**%**	**Ex.**	**Cor.**	**%**
A	6	3	50%	3	1	33%	4	4	100%	7	2	29%
E	6	3	50%	4	2	50%	7	2	29%	3	1	33%
H	4	2	50%	8	1	12%	7	0	0%	1	0	0%
L	7	7	100%	3	3	100%	8	8	100%	2	2	100%
Q	3	1	33%	4	0	0%	4	0	0%	9	4	44%
T	5	5	100%	9	9	100%	3	3	100%	3	3	100%
V	5	5	100%	4	4	100%	4	4	100%	7	7	100%
Z	4	3	75%	5	4	80%	3	2	67%	8	4	50%

Table 18: SOM_{10} 4-Way Cross Validation Results With Eight Letters