

Design

Student ID: 1698774 Name: Qianyi Lu

a)

To create a KNN tree, randomly choose the sum of a row to reorder the data and divide the data in to two parts. Repeating this action until every image is in a leaf. When searching this tree, the function first finds out the best neighbor. Then back up to the left/right node and add the image there to the set. After reordering the set and cut it to length $\leq k$, if the length is less than k or the set has been changed which means found neighbors might not be the best neighbors, trace back to the last layer. In the end, counting the labels of the found k images to get the predict result.

To create a decision tree, reordering and dividing the data 24times by the sum of each row. Add the information rate of two parts of each cutting and compare them to get the least information rate. Repeating this action until all the leave have information rate 0. When searching this tree, use the recorded middle value to find a route. When the present node's information rate is satisfied the requirement, searching ends and return the predict result.

For neural network, create a layer which use 576 pixels as input x and has an output of 2 variables. Comparing the result to modifying the $y = \text{weight} * x + b$ function until it satisfied the test data. I haven't find a way to deal with overlapping.

b)

	Training Accuracy	Testing Accuracy
K = 1	1.0	0.68
K = 2	1.0	0.68
K = 3	0.835	0.655
K = 4	0.89	0.675
K = 5	0.805	0.665
K = 6	0.88	0.65
K = 7	0.77	0.63
K = 8	0.84	0.645
K = 9	0.775	0.665
K = 10	0.84	0.65
DT(IC = 0 bits)	1.0	0.76
DT(IC = 0.1 bits)	1.0	0.76
DT(IC = 0.2 bits)	1.0	0.76
DT(IC = 0.3 bits)	0.995	0.765
DT(IC = 0.4 bits)	0.995	0.765
DT(IC = 0.5 bits)	0.975	0.765
CNN	1.0	1.0

	Training time (s)	Testing time (s)
K = 1	0.3142054080963135	0.38028836250305176
K = 2		0.4312891960144043
K = 3		0.54640793800354
K = 4		0.4643082618713379
K = 5		0.473339319229126
K = 6		0.48636388778686523
K = 7		0.531360387802124
K = 8		0.5534119606018066
K = 9		0.5683860778808594
K = 10		0.5834343433380127
DT(IC = 0 bits)	0.19211506843566895	0.019015073776245117
DT(IC = 0.1 bits)		0.01100921630859375
DT(IC = 0.2 bits)		0.011005163192749023
DT(IC = 0.3 bits)		0.009008169174194336
DT(IC = 0.4 bits)		0.009004354476928711
DT(IC = 0.5 bits)		0.008008241653442383
CNN	14.84309697151184	23.43645453453064