

COMP 341 – ASSIGNMENT 5

REPORT

Özgür Taylan Özcan

54284

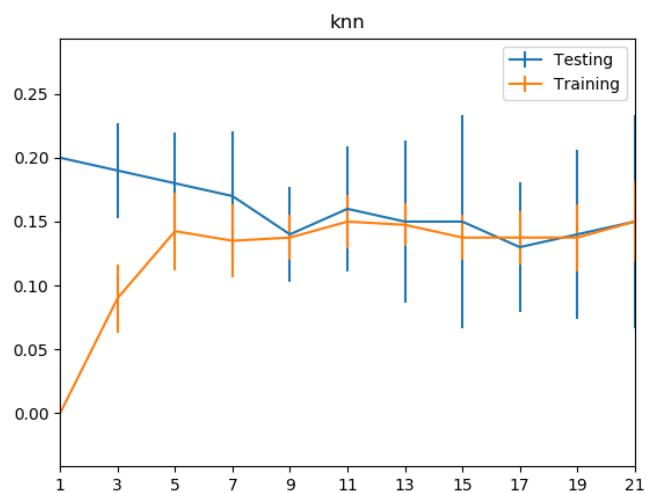
Q1) We should normalize the histogram because the sizes of images are different than each other, which means each image contains different amount of pixels. When we normalize histograms, they become invariant to image sizes. Then, we can use them as features to compare images in different sizes.

Q2) My *knn.png* output is as follows:

The range (1, 5) corresponds to overfitting, because testing error is high while training error is low.

It is not clear, but the range (19, -) may correspond to underfitting, because both training and test errors increase in this range.

$K = 17$ gives the best result. Testing error is at its minimum for this value.

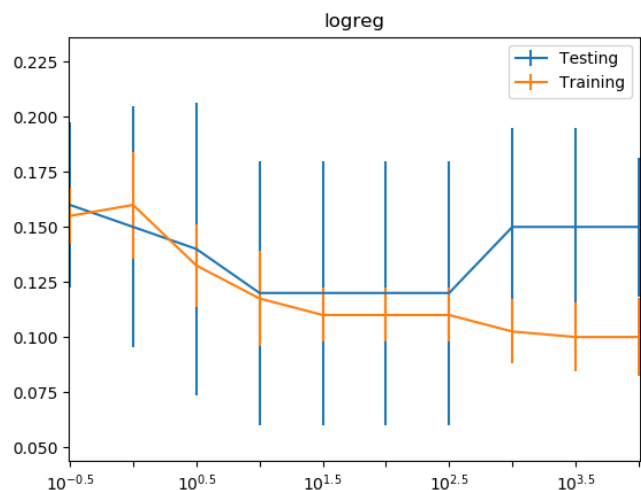


Q3) My *logrec.png* output is as follows:

The range (10^3 , -) corresponds to overfitting, because testing error is high while training error is low.

The range ($10^{-0.5}$, $10^{0.5}$) may correspond to underfitting, because training and test errors are high in this range.

The value $10^{1.0}$ gives the best result. Testing error is at its minimum for this value, and training error is also low.

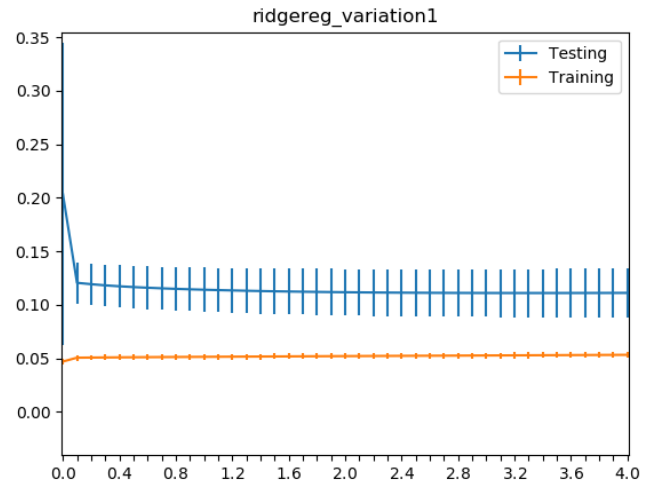


Q4) I would prefer logistic-regression, because it has a slightly lower error rate compared to kNN algorithm.

Q5) My *ridgereg_variation1.png* is as follows:

The range (0.0, 0.1) corresponds to overfitting, because testing error is high while training error is low.

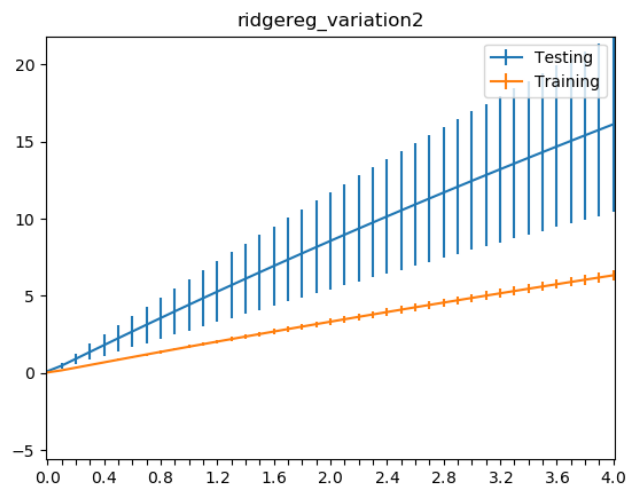
Ridge-regression is slightly better than linear regression, because it produces slightly lower error rates. The outputs on the console also support this claim.



Q6) My *ridgereg_variation2.png* is as follows:

I cannot make inference about overfitting and underfitting on this graph.

Ridge-regression is slightly better than linear regression, because it produces slightly lower error rates. The outputs on the console also support this claim.



Q7) My *ridgereg_airfoil.png* is as follows:

I cannot make inference about overfitting and underfitting on this graph.

Ridge-regression is slightly better than linear regression, because it produces slightly lower error rates. The outputs on the console also support this claim.

