**Analysis of k-nearest-neighbors**

The k-nearest neighbors has 3 parameters that affect the outcome. The actual value of k that we need to choose. The number of test cases that we compare or find our distance from and third the distance function that we use.

The table below shows how changing the two factors of k and no of test cases considered affects the Run-time and accuracy of the results.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| k | No of test cases considered | Accuracy | Correct Classification | Run-Time( in Secs) |
| 2 | 200 | 55.77942736 | 526 | 13.38 |
| 2 | 1000 | 57.58218452 | 543 | 40.8 |
| 3 | 200 | 58.6426299 | 553 | 13.92 |
| 3 | 1000 | 61.08165429 | 576 | 45.02 |
| 4 | 200 | 60.12725345 | 567 | 13.84 |
| 4 | 1000 | 61.08165429 | 576 | 41.8 |
| 5 | 200 | 62.67232238 | 591 | 13.79 |
| 5 | 1000 | 62.99045599 | 594 | 42.93 |
| 6 |  | 62.99045599 | 594 |  |
| 6 | 1000 | 62.99045599 | 594 |  |
| 10 | 200 | 62.77836691 | 592 | 13.18 |
| 10 | 1000 | 64.26299046 | 606 | 41.34 |
| 10 | 2000 | 66.91410392 | 631 | 76.68 |
| 10 | 10000 | 69.56521739 | 656 | 362.97 |
| 10 | 20000 | 68.82290562 | 649 | 1459.99 |
| 15 | 1000 | 65.32343584 | 616 | 42.67 |
| 50 | 1000 | 67.86850477 | 640 | 41.89 |
| 100 | 1000 | 67.126193 | 633 | 67.12 |
| 100 | 10000 | 70.51961824 | 665 | 928.4 |
| 1000 | 1000 | 26.61717922 | 251 | 47.14 |

As we can see from the result above that the run-time of the algorithm is dependent on the number of test cases we consider. The higher the number the more run-time for the classifier. We also see that the accuracy is also dependent on test cases. Mostly, having a big number of test cases improves the accuracy of the classifier as we can see from the table. We also note that the accuracy also dependent on number of neighbors that we consider. A too small number will always result in a lower accuracy and a very large number may also decrease the accuracy as we may be dependent on the outliers.

So, basically it is a trade-off between the number of test cases, run-time that we need and number of neighbors(k).

However, to improve the accuracy we can use technique such as instead of having a large number of test cases. Having a set of test cases that cover all types and orientation of images thus that would improve accuracy. An improved data-set would be more accurate. We can also change the distance function, to improve accuracy.

**Analysis Neural Networks**:-

There were a number of parameters to choose from to test the performance of a neural net.

The table shows success percentage for different values of learning rate alpha and number of neurons in hidden layer.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Alpha/Neurons | 8 | 16 | 32 | 48 | 64 |
| 0.10 | 57.6349 | 62.2677 | 68.2217 | 68.7349 | 69.2356 |
| 0.20 | 60.6791 | 64.9219 | 69.3875 | **70.8127** | 70.4821 |
| 0.30 | 61.9362 | 65.3752 | 69.4938 | 69.4637 | 69.2273 |
| 0.40 | 60.2193 | 64.7013 | 68.6512 | 69.1082 | 69.2512 |
| 0.50 | 60.0012 | 63.8211 | 68.0068 | 68.4359 | 68.7172 |

The best result is obtained when alpha is set to 0.20 and there are 48 neurons in the hidden layer.

The training set is being iterated for 10 epochs only due to time consideration. The result can be improved by learning the training set better by increasing the number of epochs.