Assignment 6

**Analysis of KNN:**

Table 1:

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **DISTANCE FUNCTION** | **NO. OF ‘K’ NEAREST NEIGHBOURS** | **FRACTION OF TRAINING SET** | **ACCURACY**  **(in Percentage)** | **TIME**  **(in Second)** |
| MANHATTAN | 20 | 1 | 70.02 | 4528.12 |
| 1/10 | 68.29 | 499.34 |
| 1/20 | 66.70 | 236.59 |
| 1/40 | 65.64 | 116.82 |
| 40 | 1 | 70.62 | 5438.02 |
| 1/10 | 69.41 | 412.85 |
| 1/20 | 68.39 | 260.69 |
| 1/40 | 66.06 | 119.15 |
| EUCLIDEAN | 20 | 1 | 70.41 | 6963.37 |
| 1/10 | 68.92 | 618.179 |
| 1/20 | 66.38 | 308.10 |
| 1/40 | 65.64 | 157.77 |
| 40 | 1 | 71.89 | 7994.40 |
| 1/10 | 69.67 | 628.90 |
| 1/20 | 68.39 | 315.44 |
| 1/40 | 66.80 | 149.83 |

**PARAMETER’S INFLUENCE:**

* KNN depends on k value chosen, distance function and number of training data.
* KNN gives best results when manhattan distance is chosen with accuracy same as euclidean but with lesser time.
* K value when given in the range of 30-40 KNN gives better results for both the distance function analyzed.
* KNN takes longer time when entire training data set is chosen.

**Analysis of Neural Net:**

**Table 2**:

|  |  |  |  |
| --- | --- | --- | --- |
| **LEARNING RATE** | **NO. OF NODES IN HIDDEN LAYER** | **ACCURACY PERCENTAGE (%)** | **TIME**  **(in sec)** |
| 0.001 | 3 | 23.75 | 0.37 |
| 4 | 25.87 | 0.41 |
| 5 | 23.75 | 0.46 |
| 6 | 35.94 | 0.50 |
| 0.1 | 3 | 64.88 | 0.39 |
| 4 | 67.23 | 0.43 |
| 5 | 23.75 | 0.48 |
| 6 | 25.87 | 0.53 |
| 0.5 | 3 | 67.13 | 0.35 |
| 4 | 65.53 | 0.43 |
| 5 | 70.83 | 0.45 |
| 6 | 70.73 | 0.52 |
| 0.7 | 3 | 66.38 | 0.42 |
| 4 | 70.51 | 0.42 |
| 5 | 67.65 | 0.48 |
| 6 | 69.45 | 0.50 |

**PARAMETER’S INFLUENCE:**

* Neural net performance is affected by learning rate and nodes in hidden layer.
* Neural net performs well with a learning rate of 0.5 and 0.7 and hidden node 5 or 6 giving an accuracy of 70%.
* We have given the neural net performance in the Table 2 with the parameters learning rate and number of hidden nodes and output efficiency and time in seconds.
* On varying the train\_data size for training the neural net, we observe the following results:

1.Learning rate:0.5,train\_data=train\_data/2

Efficiency=23.57% Time=0.4882s

2.Learning rate:0.7,train\_data=train\_data/2

Efficiency=71.04% Time=0.5452s

**OBSERVATION:**

Based on our observation, Neural network gives better accuracy within lesser time. So, we would recommend neural network as classifier with the following parameters:

Learning rate=0.5

Number of nodes in hidden layer=5

###Neural net behaves erratic due to the random initial weights chosen, on rerun it gives considerably good efficieny###

