# **Part 1**

k-nearest neighbor

For k-nearest neighbor I use **k = 19** and get **Error = 0.422472**

|  |  |
| --- | --- |
| PREDICTIONS | ACTUAL RESULT |
| CYT | CYT |
| NUC | CYT |
| CYT | NUC |
| CYT | NUC |
| CYT | NUC |
| EXC | CYT |
| CYT | CYT |
| NUC | CYT |
| CYT | MIT |
| ME1 | ME1 |
| EXC | ME1 |
| ME3 | MIT |
| CYT | MIT |
| ME1 | ME2 |
| CYT | CYT |
| CYT | CYT |
| CYT | CYT |
| MIT | MIT |
| MIT | MIT |
| NUC | NUC |
| CYT | CYT |
| CYT | CYT |
| CYT | CYT |
| ME1 | EXC |
| MIT | MIT |
| CYT | MIT |
| MIT | MIT |
| POX | MIT |
| ME3 | ME3 |
| NUC | CYT |
| ME1 | MIT |
| ME3 | ME3 |
| CYT | CYT |
| CYT | CYT |
| CYT | CYT |
| NUC | POX |
| NUC | NUC |
| MIT | MIT |
| NUC | MIT |
| MIT | MIT |
| ME3 | ME3 |
| CYT | NUC |
| NUC | CYT |
| CYT | CYT |
| NUC | CYT |
| CYT | CYT |
| MIT | NUC |
| NUC | NUC |
| CYT | CYT |
| NUC | NUC |
| NUC | NUC |
| NUC | CYT |
| NUC | NUC |
| MIT | NUC |
| NUC | CYT |
| MIT | MIT |
| CYT | POX |
| NUC | NUC |
| CYT | CYT |
| ME1 | ME1 |
| MIT | MIT |
| MIT | MIT |
| MIT | MIT |
| ME3 | MIT |
| MIT | MIT |
| MIT | MIT |
| MIT | MIT |
| EXC | MIT |
| CYT | CYT |
| CYT | CYT |
| EXC | EXC |
| CYT | VAC |
| CYT | NUC |
| ME2 | ME2 |
| CYT | NUC |
| NUC | NUC |
| EXC | EXC |
| EXC | EXC |
| CYT | MIT |
| CYT | MIT |
| CYT | MIT |
| CYT | MIT |
| ME3 | ME3 |
| POX | POX |
| CYT | NUC |
| NUC | NUC |
| CYT | CYT |
| NUC | NUC |
| MIT | MIT |
| CYT | NUC |
| CYT | ME2 |
| CYT | CYT |
| NUC | ME3 |
| CYT | CYT |
| ME3 | ME3 |
| NUC | CYT |
| CYT | CYT |
| ME2 | ME2 |
| CYT | CYT |
| CYT | CYT |
| CYT | ME2 |
| CYT | ME3 |
| NUC | NUC |
| EXC | EXC |
| CYT | CYT |
| NUC | CYT |
| CYT | CYT |
| ME3 | ME3 |
| ME3 | ME3 |
| CYT | POX |
| ME3 | ME3 |
| ME1 | ME1 |
| ME1 | ME2 |
| CYT | CYT |
| ME3 | ME3 |
| ME3 | ME3 |
| ME1 | ME1 |
| CYT | NUC |
| CYT | NUC |
| CYT | NUC |
| CYT | NUC |
| ME3 | ME3 |
| NUC | NUC |
| CYT | CYT |
| ME3 | CYT |
| CYT | NUC |
| NUC | NUC |
| CYT | NUC |
| CYT | NUC |
| CYT | CYT |
| CYT | CYT |
| NUC | CYT |
| CYT | CYT |
| CYT | ME2 |
| NUC | NUC |
| NUC | NUC |
| MIT | MIT |
| NUC | NUC |
| CYT | NUC |
| CYT | NUC |
| NUC | CYT |
| MIT | MIT |
| NUC | CYT |
| CYT | CYT |
| NUC | NUC |
| NUC | NUC |
| CYT | NUC |
| CYT | NUC |
| ME3 | ME3 |
| NUC | NUC |
| CYT | CYT |
| NUC | NUC |
| NUC | MIT |
| MIT | MIT |
| MIT | NUC |
| NUC | NUC |
| CYT | CYT |
| ME3 | ME3 |
| ME3 | ME3 |
| CYT | POX |
| MIT | MIT |
| MIT | MIT |
| NUC | CYT |
| MIT | MIT |
| CYT | MIT |
| NUC | MIT |
| ME1 | ME1 |
| ME3 | ME3 |
| NUC | CYT |
| NUC | MIT |
| ME1 | EXC |
| NUC | ME3 |
| ERL | CYT |
| MIT | CYT |
| ME3 | ME3 |
| ME2 | ME2 |
| ME2 | ME2 |
| CYT | VAC |
| CYT | CYT |
| NUC | NUC |
| ME3 | ME3 |
| ME3 | ME3 |
| NUC | ME2 |
| CYT | CYT |
| NUC | CYT |
| NUC | NUC |
| ME3 | ME3 |
| CYT | MIT |
| CYT | MIT |
| CYT | MIT |
| CYT | CYT |
| MIT | MIT |
| MIT | ME3 |
| CYT | CYT |
| CYT | NUC |
| MIT | MIT |
| ME3 | ME3 |
| NUC | NUC |
| NUC | CYT |
| MIT | MIT |
| CYT | NUC |
| NUC | NUC |
| NUC | NUC |
| NUC | MIT |
| MIT | MIT |
| MIT | MIT |
| CYT | MIT |
| MIT | MIT |
| CYT | MIT |
| NUC | MIT |
| MIT | MIT |
| MIT | MIT |
| MIT | MIT |
| NUC | NUC |
| MIT | MIT |
| CYT | NUC |
| CYT | MIT |
| MIT | MIT |
| CYT | MIT |
| CYT | EXC |
| ME3 | ME3 |
| NUC | NUC |
| CYT | CYT |
| CYT | ME2 |
| ME3 | NUC |
| MIT | NUC |
| NUC | NUC |
| ME3 | CYT |
| CYT | CYT |
| MIT | NUC |
| NUC | CYT |
| CYT | NUC |
| CYT | NUC |
| NUC | NUC |
| MIT | ME2 |
| ME3 | ME2 |
| ME1 | MIT |
| CYT | MIT |
| CYT | MIT |
| CYT | POX |
| MIT | POX |
| CYT | CYT |
| CYT | CYT |
| CYT | CYT |
| NUC | NUC |
| ME3 | ME3 |
| NUC | NUC |
| CYT | VAC |
| NUC | VAC |
| MIT | MIT |
| NUC | MIT |
| CYT | MIT |
| MIT | MIT |
| CYT | CYT |
| EXC | EXC |
| CYT | CYT |
| NUC | NUC |
| NUC | NUC |
| CYT | ME3 |
| EXC | ME1 |
| ME1 | ME2 |
| ME3 | ME3 |
| NUC | NUC |
| NUC | NUC |
| CYT | NUC |
| CYT | NUC |
| ME3 | NUC |
| CYT | NUC |
| ME2 | ME3 |
| MIT | MIT |
| CYT | CYT |
| NUC | NUC |
| CYT | CYT |
| MIT | NUC |
| NUC | NUC |
| CYT | NUC |
| NUC | NUC |
| NUC | NUC |
| CYT | CYT |
| CYT | CYT |
| MIT | CYT |
| NUC | NUC |
| NUC | CYT |
| CYT | CYT |
| NUC | NUC |
| CYT | MIT |
| CYT | MIT |
| CYT | NUC |
| NUC | NUC |
| CYT | NUC |
| CYT | NUC |
| NUC | NUC |
| NUC | NUC |
| MIT | MIT |
| CYT | NUC |
| NUC | NUC |
| ME3 | NUC |
| NUC | NUC |
| CYT | NUC |
| CYT | CYT |
| CYT | CYT |
| CYT | CYT |
| CYT | CYT |
| CYT | CYT |
| NUC | NUC |
| CYT | NUC |
| CYT | NUC |
| CYT | NUC |
| CYT | NUC |
| NUC | NUC |
| CYT | CYT |
| CYT | CYT |
| CYT | CYT |
| CYT | CYT |
| CYT | CYT |
| CYT | CYT |
| CYT | CYT |
| CYT | CYT |
| CYT | CYT |
| CYT | CYT |
| MIT | CYT |
| NUC | CYT |
| CYT | NUC |
| CYT | CYT |
| CYT | CYT |
| CYT | CYT |
| MIT | CYT |
| MIT | CYT |
| CYT | CYT |
| MIT | MIT |
| MIT | MIT |
| MIT | MIT |
| EXC | ME1 |
| CYT | ME2 |
| NUC | CYT |
| CYT | CYT |
| CYT | CYT |
| CYT | CYT |
| NUC | CYT |
| ME3 | ME3 |
| ME3 | ME2 |
| CYT | MIT |
| EXC | ME1 |
| NUC | CYT |
| MIT | MIT |
| MIT | CYT |
| NUC | NUC |
| CYT | NUC |
| CYT | CYT |
| ME3 | ME3 |
| NUC | NUC |
| NUC | NUC |
| CYT | CYT |
| ME3 | ME3 |
| MIT | NUC |
| NUC | NUC |
| MIT | CYT |
| ME3 | ME3 |
| CYT | CYT |
| CYT | NUC |
| ME3 | ME3 |
| MIT | MIT |
| NUC | NUC |
| NUC | NUC |
| CYT | NUC |
| CYT | NUC |
| CYT | NUC |
| CYT | NUC |
| CYT | NUC |
| NUC | CYT |
| CYT | CYT |
| MIT | MIT |
| NUC | CYT |
| CYT | NUC |
| CYT | ME3 |
| CYT | ME3 |
| EXC | EXC |
| ME3 | ME3 |
| ME1 | ME1 |
| CYT | CYT |
| MIT | MIT |
| NUC | NUC |
| CYT | CYT |
| ME1 | EXC |
| NUC | NUC |
| CYT | CYT |
| CYT | CYT |
| NUC | CYT |
| CYT | CYT |
| CYT | CYT |
| CYT | CYT |
| CYT | CYT |
| NUC | CYT |
| NUC | NUC |
| NUC | NUC |
| CYT | CYT |
| MIT | CYT |
| NUC | CYT |
| NUC | CYT |
| NUC | CYT |
| NUC | CYT |
| CYT | CYT |
| ME1 | ME1 |
| ME1 | ME1 |
| CYT | NUC |
| NUC | NUC |
| CYT | NUC |
| CYT | CYT |
| CYT | CYT |
| CYT | CYT |
| CYT | NUC |
| CYT | CYT |
| CYT | CYT |
| CYT | NUC |
| CYT | CYT |
| CYT | CYT |
| ME3 | ME3 |
| CYT | CYT |
| NUC | NUC |
| CYT | MIT |
| ME1 | VAC |
| CYT | VAC |
| ME1 | ME1 |
| NUC | MIT |
| ME1 | ME1 |
| CYT | CYT |
| ME3 | ME3 |
| POX | POX |
| ME3 | ME3 |
| ME3 | ME3 |
| ME3 | ME3 |
| ME3 | ME3 |
| ME3 | ME3 |
| ME3 | ME3 |
| ME3 | ME3 |
| MIT | ME2 |
| CYT | NUC |
| ME3 | ME3 |
| MIT | CYT |
| NUC | NUC |
| CYT | CYT |
| CYT | MIT |
| NUC | MIT |
| CYT | MIT |
| ME3 | ME2 |

**The number of correctly classified test instance = 257**

**Total number of instances in the test set = 445**

K-fold cross validation

knn =01 cross-validation error = 0.504589

knn =02 cross-validation error = 0.516208

knn =03 cross-validation error = 0.460575

knn =04 cross-validation error = 0.458729

knn =05 cross-validation error = 0.442154

knn =06 cross-validation error = 0.452961

knn =07 cross-validation error = 0.438263

knn =08 cross-validation error = 0.437336

knn =09 cross-validation error = 0.425518

knn =10 cross-validation error = 0.428759

knn =11 cross-validation error = 0.416286

knn =12 cross-validation error = 0.418863

knn =13 cross-validation error = 0.415443

knn =14 cross-validation error = 0.410382

knn =15 cross-validation error = 0.406568

knn =16 cross-validation error = 0.406406

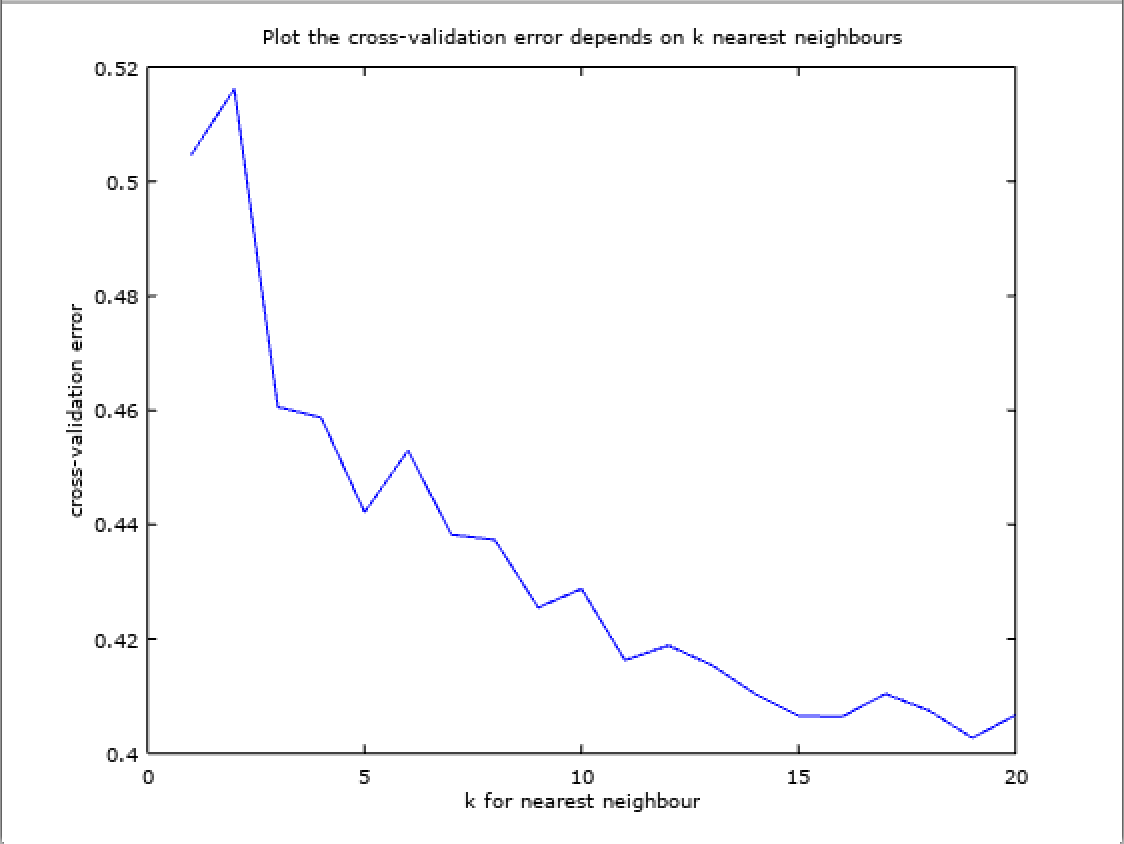
knn =17 cross-validation error = 0.410392

knn =18 cross-validation error = 0.407519

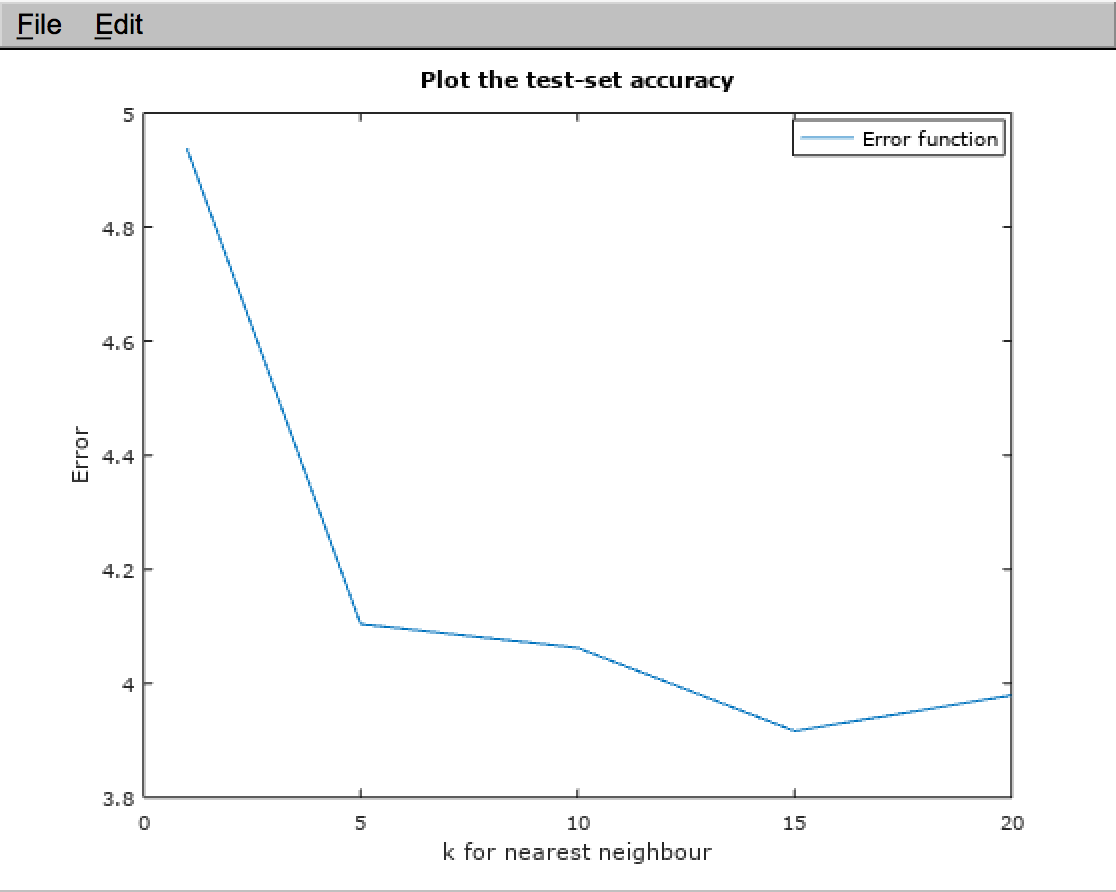
**knn =19 cross-validation error = 0.402720**

knn =20 cross-validation error = 0.406658

**Min cross-validation error = 0.399660** for k-nearest neighbors k **= 19**.



# **Part 2**



|  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| k=1 |  | Predicted | | | | | | | | | |  |
|  | **n = 445** | **CYT** | **ERL** | EXC | ME1 | ME2 | ME3 | MIT | NUC | POX | VAC | **TOTAL** |
| Actual | **CYT** | 68 | 0 | 1 | 0 | 1 | 5 | 15 | 43 | 1 | 4 | **138** |
| **ERL** | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | **0** |
| **EXC** | 0 | 0 | 5 | 3 | 1 | 0 | 1 | 0 | 0 | 0 | **10** |
| **ME1** | 0 | 0 | 3 | 6 | 3 | 1 | 0 | 0 | 1 |  | **14** |
| **ME2** | 3 | 0 | 3 | 1 | 4 | 2 | 1 | 0 | 1 | 3 | **18** |
| **ME3** | 3 | 0 | 0 | 0 | 1 | 32 | 2 | 8 |  | 1 | **47** |
| **MIT** | 23 | 0 | 2 | 0 | 3 | 2 | 36 | 11 | 3 | 1 | **81** |
| **NUC** | 51 | 0 | 1 | 0 | 1 | 4 | 14 | 50 | 0 | 2 | **123** |
| **POX** | 1 | 0 | 0 | 0 | 0 | 0 | 1 | 1 | 5 | 0 | **8** |
| **VAC** | 0 | 0 | 0 | 1 | 1 | 0 | 1 | 0 | 1 | 2 | **6** |
|  | **TOTAL** | **149** | **0** | **15** | **11** | **15** | **46** | **71** | **113** | **12** | **13** |  |
|  |  |  |  |  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |  |  |  |
| k=15 |  | Predicted | | | | | | | | | |  |
|  | **n = 445** | **CYT** | **ERL** | EXC | ME1 | ME2 | ME3 | MIT | NUC | POX | VAC | **TOTAL** |
| Actual | **CYT** | 91 | 2 | 1 | 0 | 0 | 2 | 9 | 33 | 0 | 0 | **138** |
| **ERL** | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | **0** |
| **EXC** | 1 | 0 | 7 | 2 | 0 | 0 | 0 | 0 | 0 | 0 | **10** |
| **ME1** | 0 | 0 | 4 | 9 | 0 | 0 | 1 | 0 | 0 | 0 | **14** |
| **ME2** | 4 | 0 | 1 | 2 | 4 | 4 | 2 | 1 | 0 | 0 | **18** |
| **ME3** | 4 | 0 | 0 | 0 | 1 | 39 | 1 | 2 | 0 | 0 | **47** |
| **MIT** | 26 | 0 | 1 | 2 | 0 | 1 | 42 | 8 | 1 | 0 | **81** |
| **NUC** | 51 | 0 | 0 | 0 | 0 | 4 | 7 | 61 | 0 | 0 | **123** |
| **POX** | 2 | 0 | 0 | 0 | 0 | 0 | 1 | 1 | 4 | 0 | **8** |
| **VAC** | 5 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | **6** |
|  | **TOTAL** | **184** | **2** | **14** | **16** | **5** | **50** | **63** | **106** | **5** | **0** |  |

As we can notice there are amount of correct predicted class on the diagonal of confusion matrixes. For k=15 our program classifies better than for k=1. Except only classes “POX” and “VAC”.

If we divide each diagonal element by relevant row total, we will get an accuracy per test class.

|  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | CYT | ERL | EXC | ME1 | ME2 | ME3 | MIT | NUC | POX | VAC | Total Accuracy |
| K=1 | 0.49 | 0.00 | 0.50 | 0.43 | 0.22 | 0.68 | 0.44 | 0.41 | 0.63 | 0.33 | 0.46742 |
| K=15 | 0.66 | 0.00 | 0.70 | 0.64 | 0.22 | 0.83 | 0.52 | 0.50 | 0.50 | 0.00 | 0.57753 |