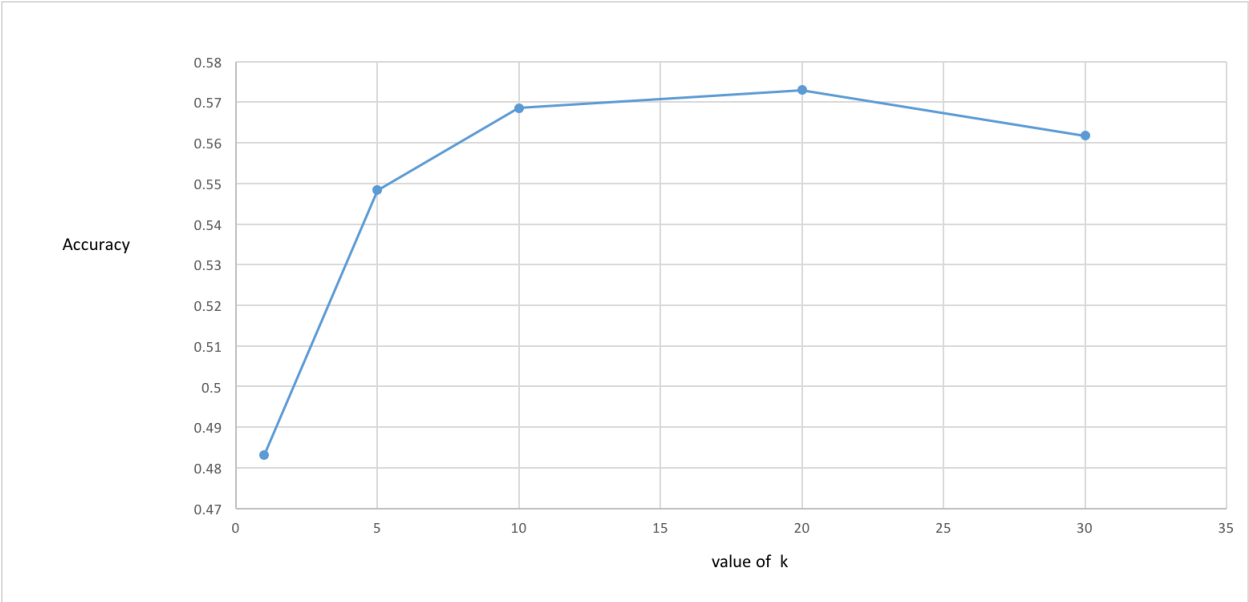
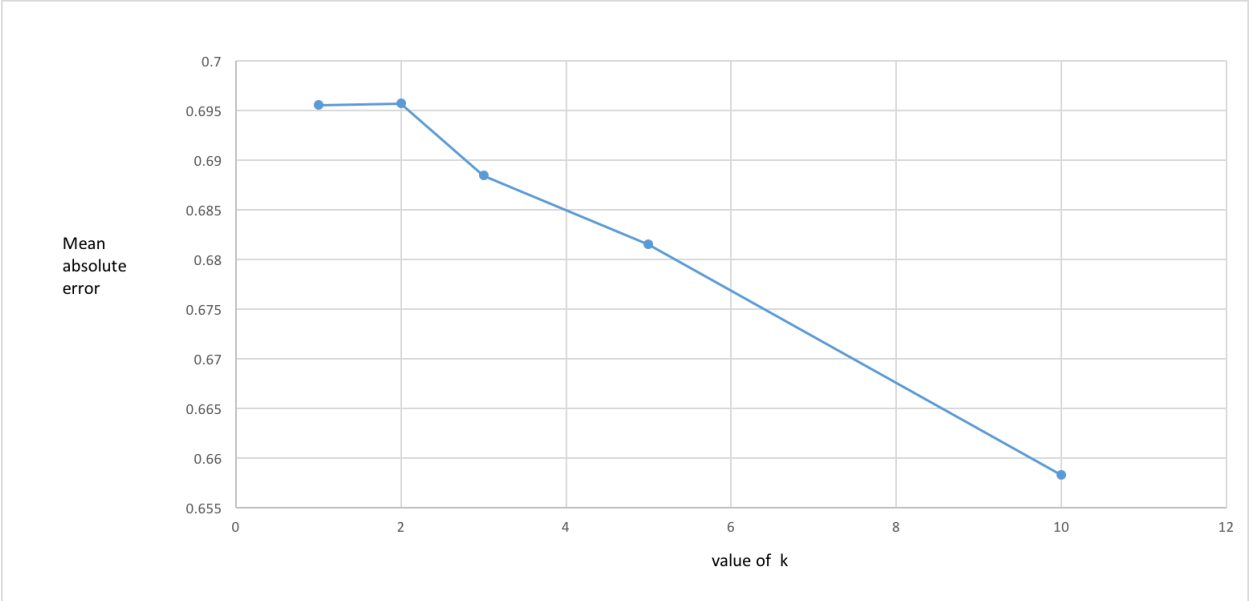
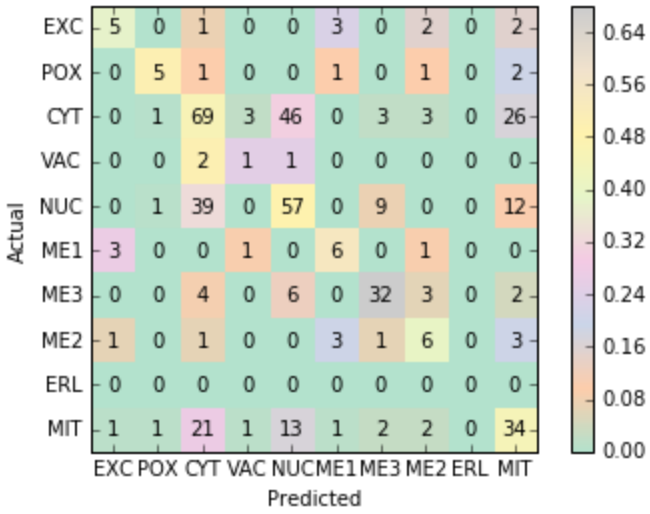
2)



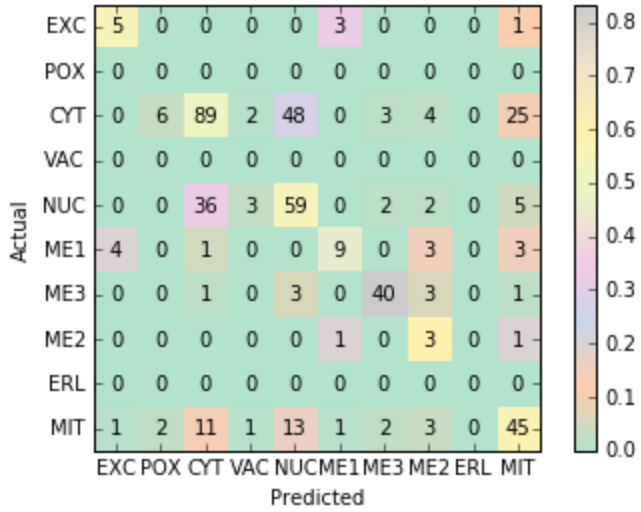

Accuracy as a function of k for the yeast dataset




Mean absolute error as a function of k for the wine dataset




Confusion matrix for Yeast k=1




Confusion matrix for Yeast k=30

The larger the k the predictor becomes better at classifying things. However, too large of a K tends to overfit the data. The accuracy of the classifier thus starts to decline.

3)

| | Euclidean Distance | Best Distance | Best Node | Priority Queue |
|---|---|---|---|---|
| | | ∞ | | (f, 0) |

| | | | | |
|---|---|---|---|---|
| Pop f | 5 √2 | 5 √2 | f | (h, 0), (c, 1) |
| Pop h | 5 √2 | 5 √2 | f | (i, 0), (c, 1), (g, 5) |
| Pop i | 3 | 3 | i | (c, 1),(j, 3) (g, 5) |
| Pop c | 2 | 2 | c | (e, 0),(b, 0),(j, 3) (g, 5) |
| Pop e | √61 | 2 | c | (b, 0),(d, 0),(j, 3) (g, 5) |
| Pop b | 2√5 | 2 | c | (d, 0),(j, 3),(a, 4) (g, 5) |
| Pop d | √29 | 2 | c | (j, 3),(a, 4) (g, 5) |
| Pop j<br>Return c | | | | |

The nearest neighbor is c