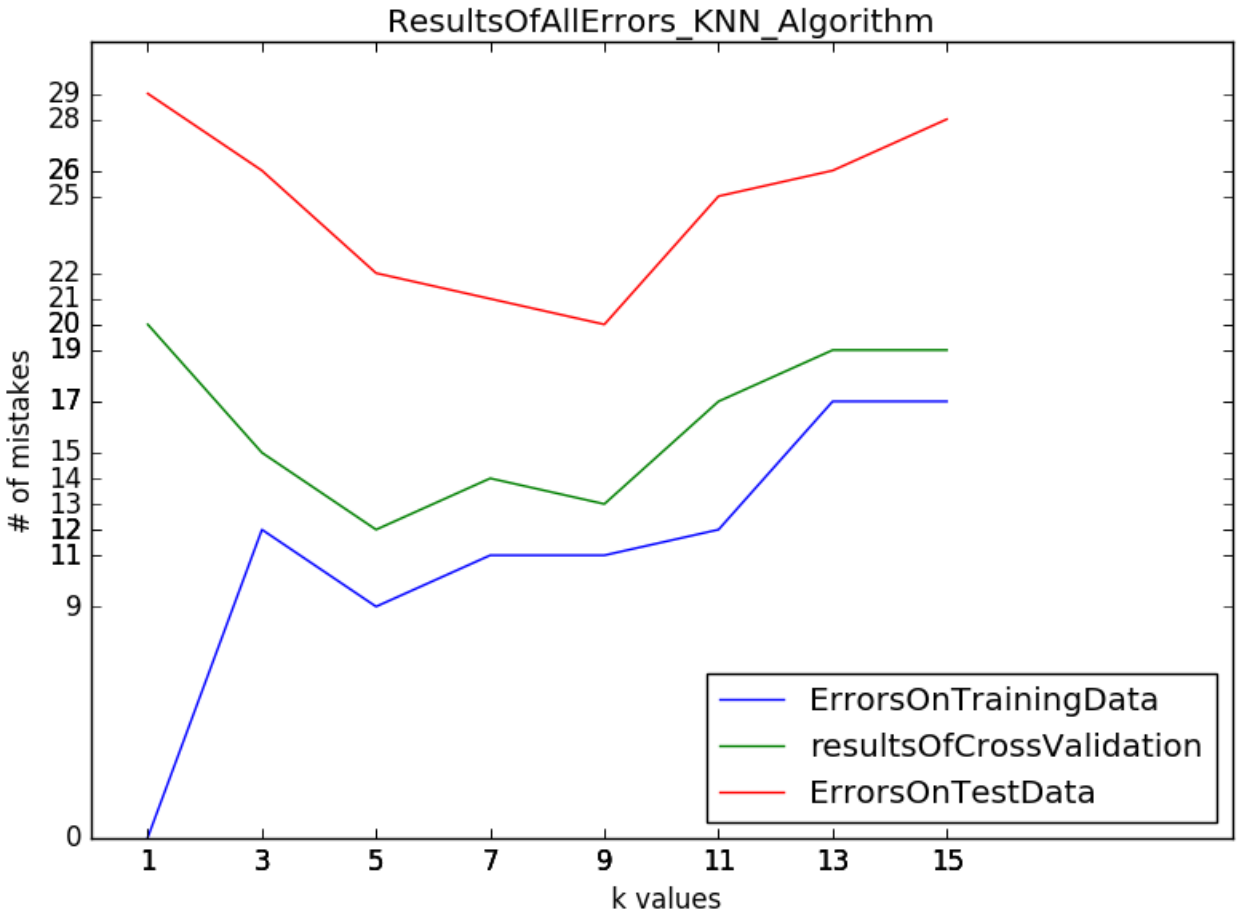
CS434: Machine Learning & Data Mining (Spring 2016)
Implementation Assignment 2
Muratcan CICEK
(No Group)

NOTES:

- I used Python for the implementation of this assignment and I printed some of my experiments, I also plotted some results. They are below.

Part I:

- I have performed model selection using leave-one-out cross-validation to select the best K for the given learning task using the provided training data.
- I have considered the following K values plotted the results as below



- As you can see above 5 and 9 seem the optimal k values. Smaller k values cause over fitting and larger k values cause under fitting.

Part II:

- I have considered all the following tests and from this experience, I have observed that checking x6 == 2 has maximum information gain, `0.498363754696`. In this stump left child has 68 negative instances which belong to class label y = 0. The right child has all of 39 positive instances (y=1) and 17 negative instances which are errors of my decision stump on training data.

    Test: Is x1 == 1? Information Gain = 0.00753540706527
    Test: Is x1 == 2? Information Gain = 0.0006038106728
    Test: Is x1 == 3? Information Gain = 0.0143465953988
    Test: Is x2 == 1? Information Gain = 6.98614835981e-08
    Test: Is x2 == 2? Information Gain = 0.00143423326152
    Test: Is x2 == 3? Information Gain = 0.0013638906969
    Test: Is x3 == 1? Information Gain = 0.00572757919074
    Test: Is x3 == 2? Information Gain = 0.00572757919074
    Test: Is x4 == 1? Information Gain = 0.00482682747333
    Test: Is x4 == 2? Information Gain = 0.00741166801119
    Test: Is x4 == 3? Information Gain = 0.000264861406531
    Test: Is x5 == 1? Information Gain = 0.00975442583748
    Test: Is x5 == 2? Information Gain = 0.00915430152135
    Test: Is x5 == 3? Information Gain = 0.0003754516682
    Test: Is x5 == 4? Information Gain = 0.000529988102218
    Test: Is x6 == 1? Information Gain = 0.498363754696
    Test: Is x6 == 2? Information Gain = 0.498363754696

- The same decision stump (x6 == 2) on test data also has only 216 negative instances on its left child and on right child, it has all 144 positive and 72 negative instances. So I have 72 errors.
-
-
-