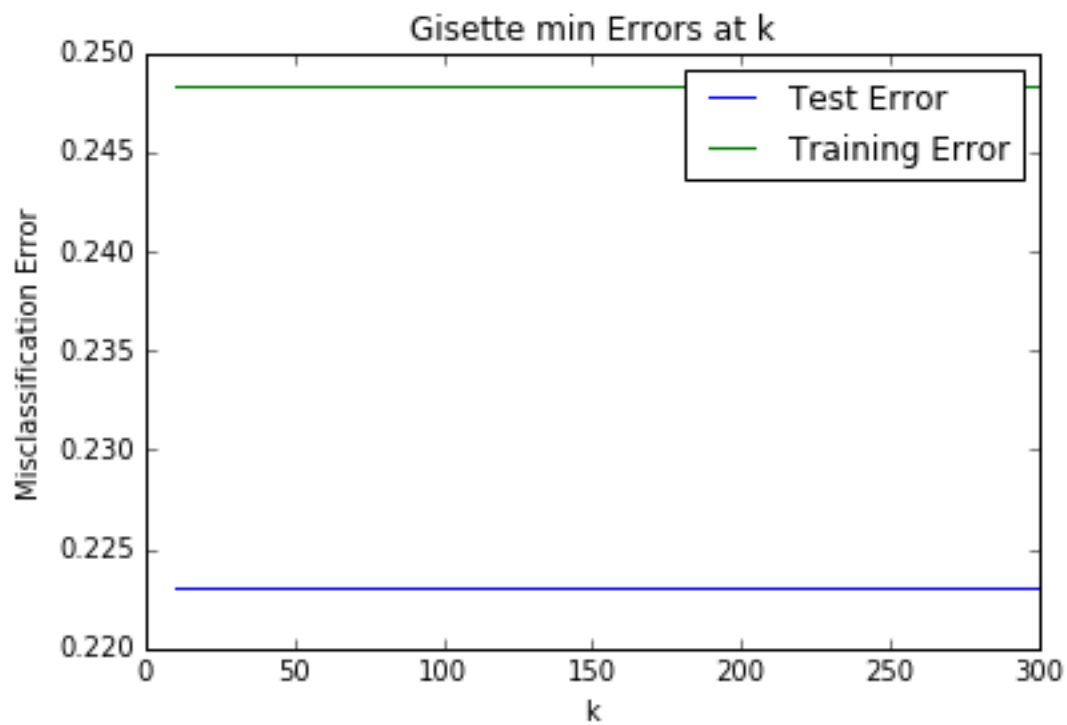
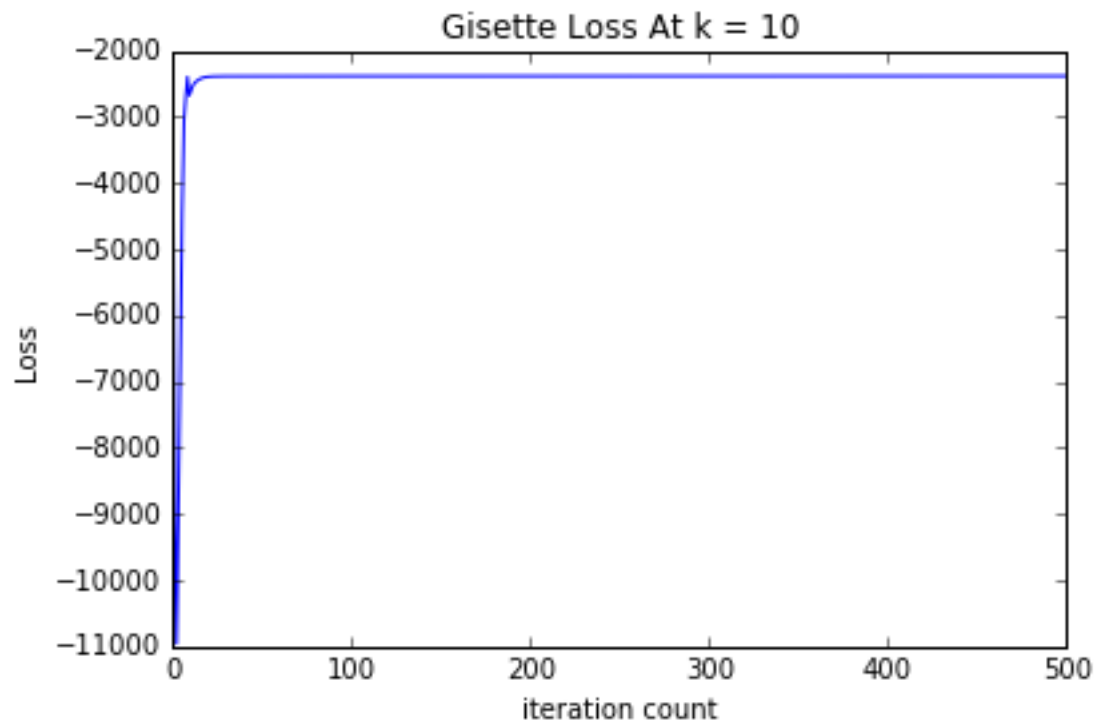
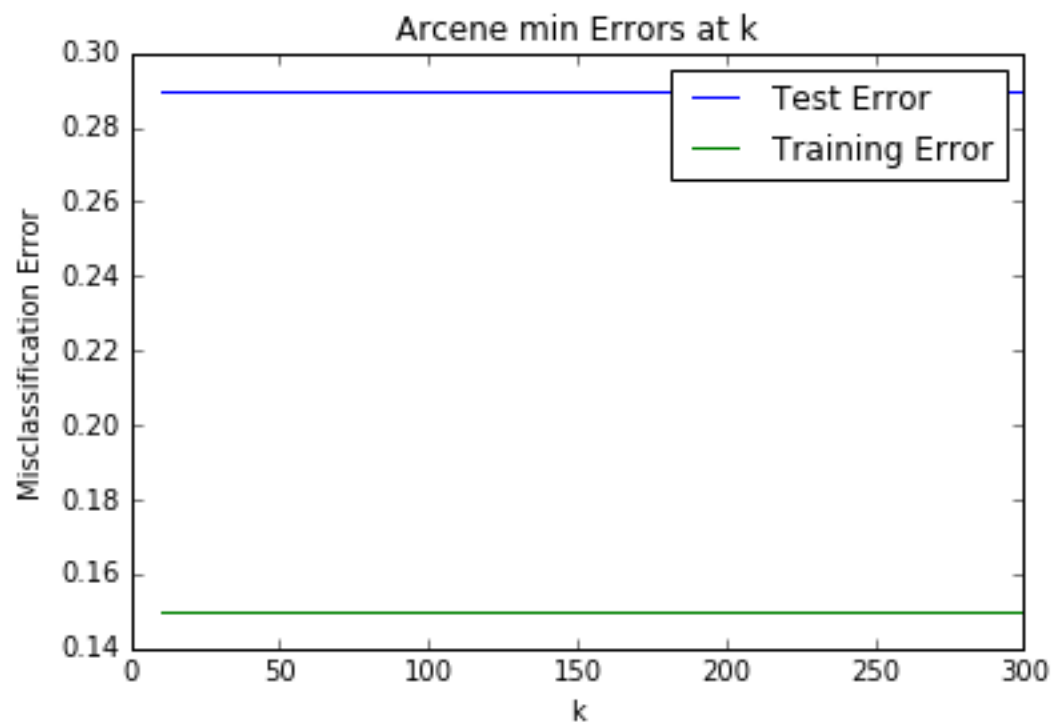
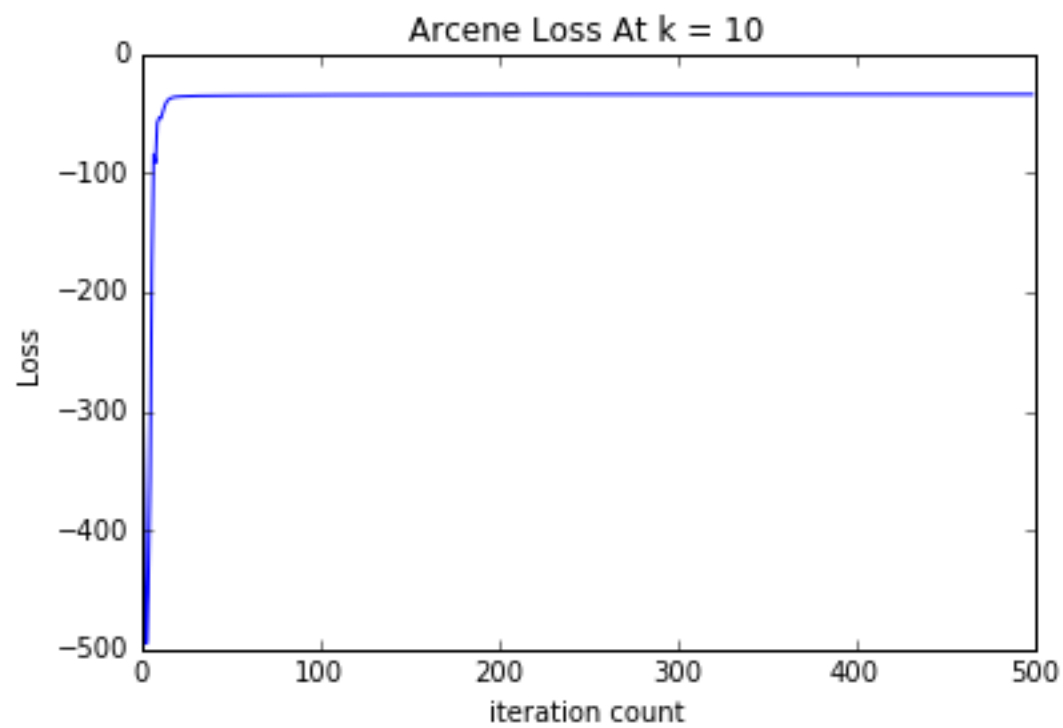


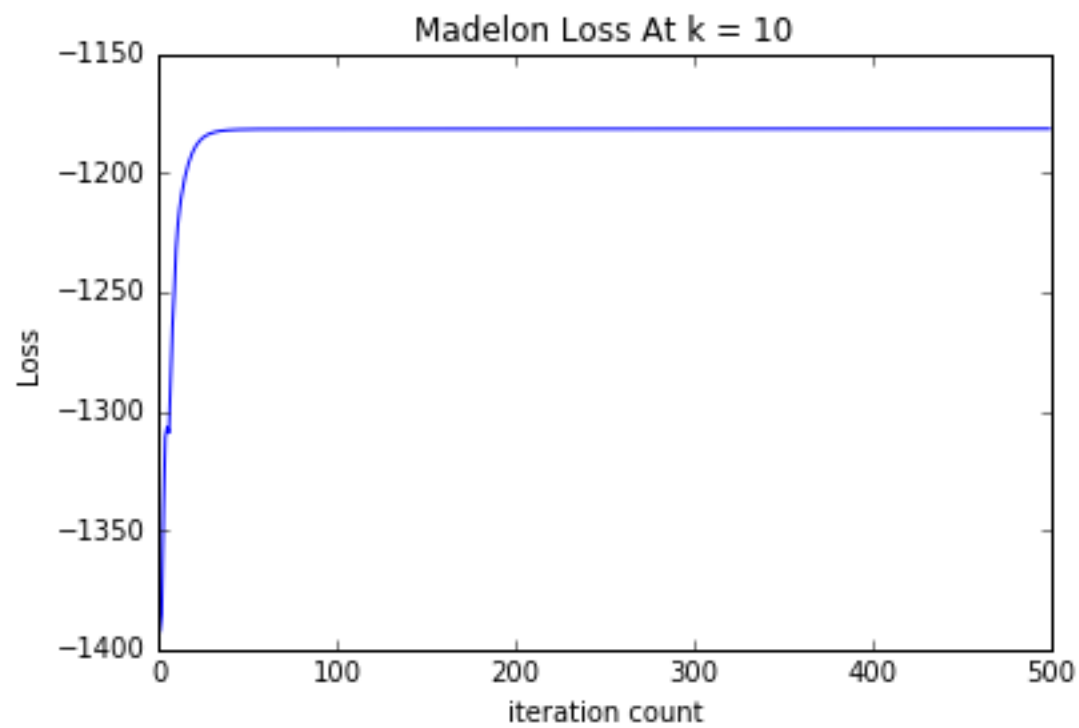
Gisette



Arcene



Madelon



Misclassification Errors

Dataset (k)	Test Error	Train Error
Gisette (10)	0.223	0.2483
Gisette (30)	0.223	0.2483
Gisette (100)	0.223	0.2483
Gisette (300)	0.223	0.2483
Arcene (10)	0.29	0.15
Arcene (30)	0.29	0.15
Arcene (100)	0.29	0.15
Arcene (300)	0.29	0.15
Madelon (10)	0.395	0.3805
Madelon (30)	0.395	0.3805
Madelon (100)	0.395	0.3805
Madelon (300)	0.395	0.3805

Code

Included is only the training and testing portions of our code. For full code see https://github.com/johansent/ML_Fall2017/tree/master/hw6.

```
def updateWeights(X,Y,w, Ncol,Nrow, learningRate, s):
    summation = [0] * Ncol
    lorenz = []
    for k, Row in X.iterrows():
        row = np.array(Row)
        product = Y[k] * np.dot(w,row)
        if(product <= 1):
            divisor = 2 + product*product - (2 * product)
            numerator = 2 * (product - 1)
            value = (numerator / divisor) * row * Y[k]
            summation = summation + value

    lorenz.append(np.log(1 + (product - 1)**2))
```

```

    else:
        lorenz.append(0)

    derivative = summation + (s * w)
    w = w - (learningRate * derivative)
    loss = -(sum(lorenz) + s * np.dot(w, w))
    return w, loss

def Test(w, X, Y):
    Y = [0 if y <= 0 else 1 for y in np.array(Y)]
    wx = np.dot(X,w)
    Ypredict = [0 if x < 0 else 1 for x in wx]
    results = np.array(Y) - np.array(Ypredict)
    return sum(abs(results))/len(Y)

```

Bibliography

1. Stéfan van der Walt, S. Chris Colbert and Gaël Varoquaux. The NumPy Array: A Structure for Efficient Numerical Computation, Computing in Science & Engineering, 13, 22-30 (2011), DOI:10.1109/MCSE.2011.37
2. John D. Hunter. Matplotlib: A 2D Graphics Environment, Computing in Science & Engineering, 9, 90-95 (2007), DOI:10.1109/MCSE.2007.55
3. Wes McKinney. Data Structures for Statistical Computing in Python, Proceedings of the 9th Python in Science Conference, 51-56 (2010)
4. Fabian Pedregosa, Gaël Varoquaux, Alexandre Gramfort, Vincent Michel, Bertrand Thirion, Olivier Grisel, Mathieu Blondel, Peter Prettenhofer, Ron Weiss, Vincent Dubourg, Jake Vanderplas, Alexandre Passos, David Cournapeau, Matthieu Brucher, Matthieu Perrot, Édouard Duchesnay. Scikit-learn: Machine Learning in Python, Journal of Machine Learning Research, 12, 2825-2830 (2011)
5. conner.xyz at <https://stackoverflow.com/questions/20517650/how-to-delete-the-last-column-of-data-of-a-pandas-dataframe>