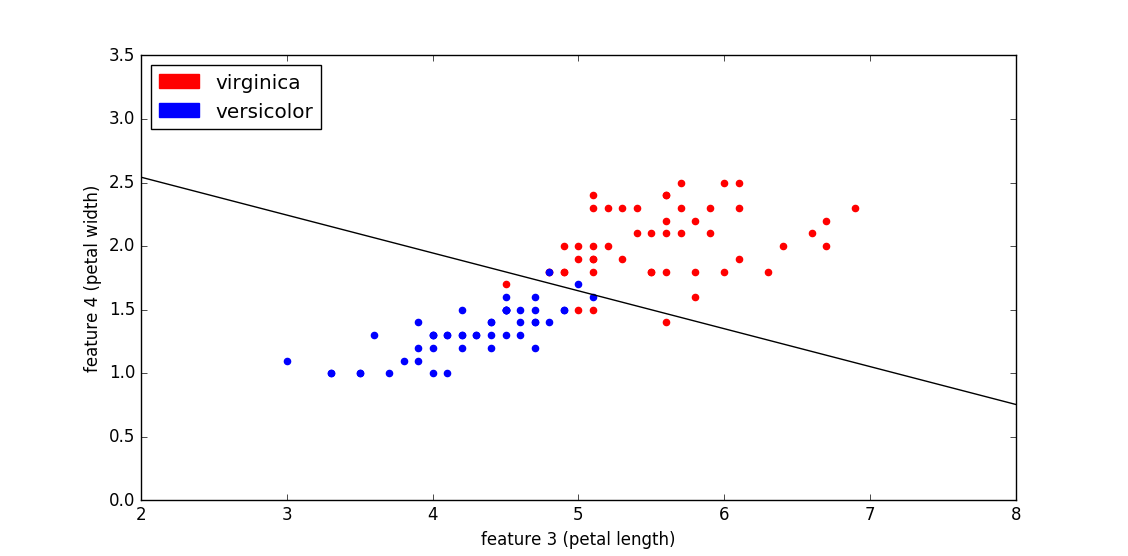
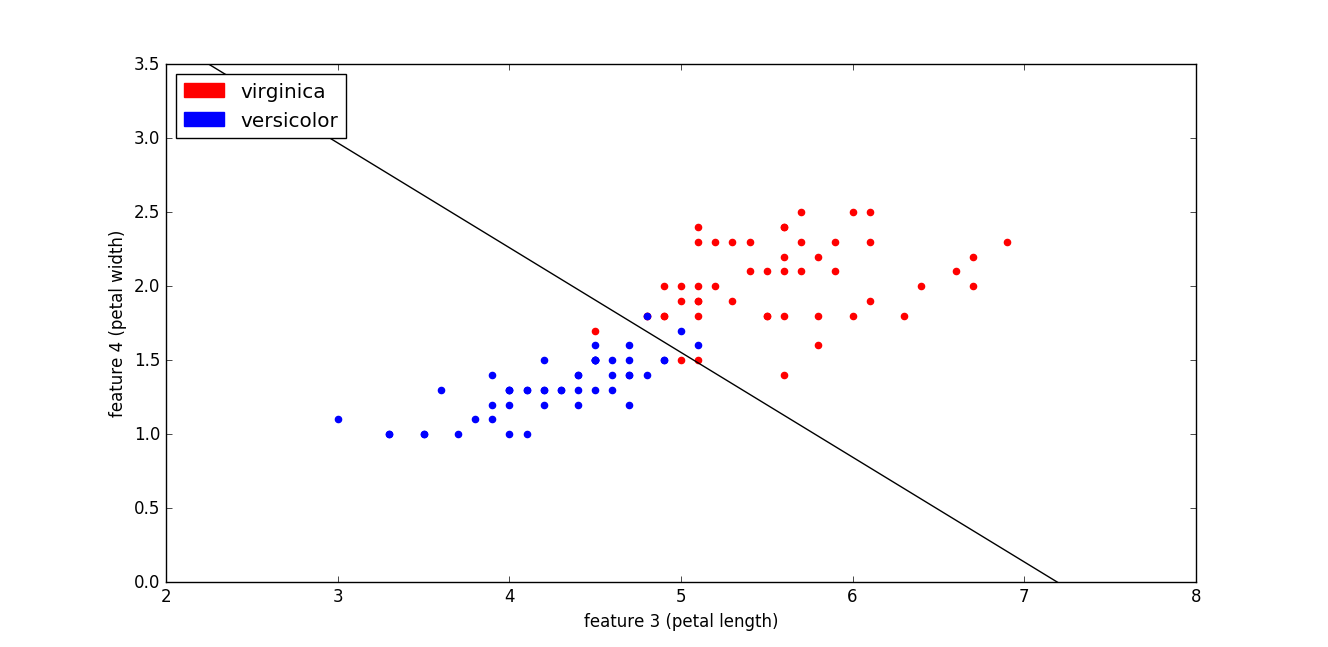
HW8 Answers

# Q1 Part A



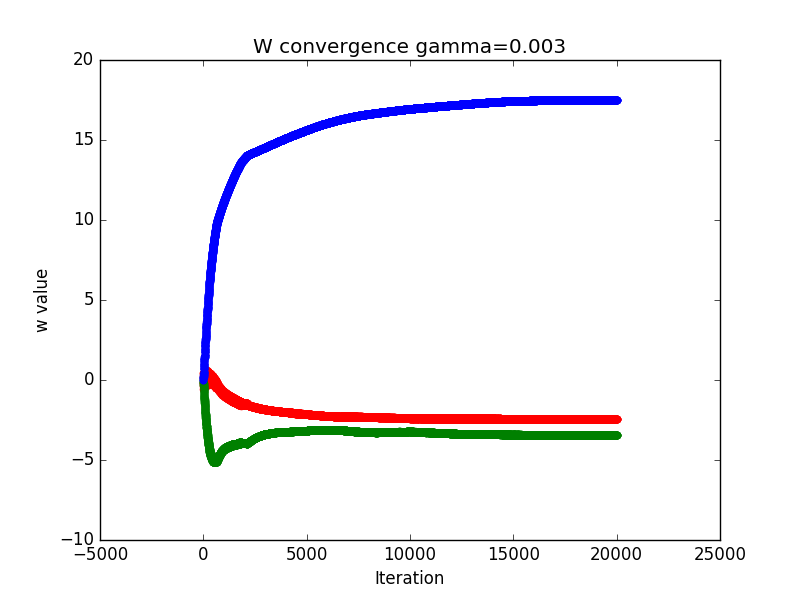
# Q1 Part B

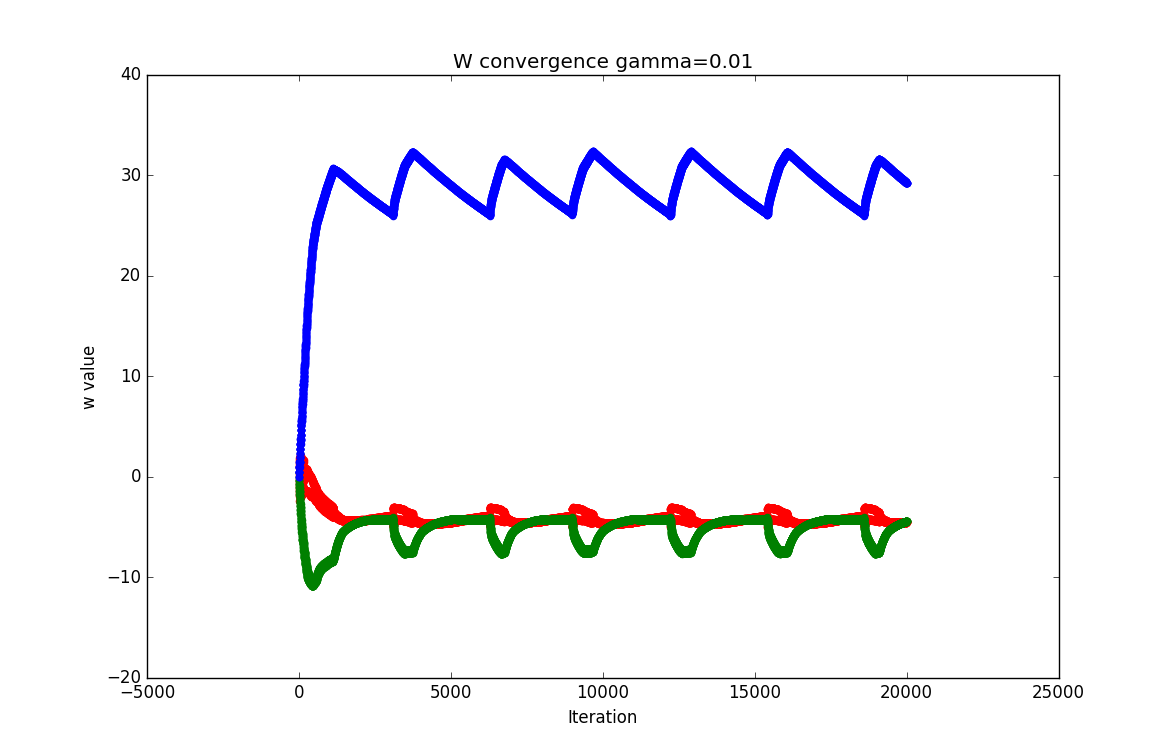


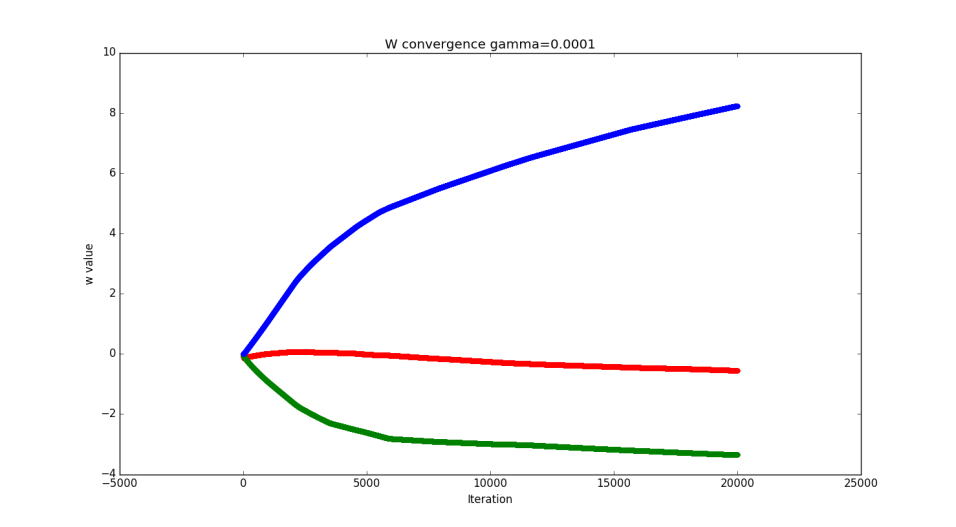
It can be seen that the SVM decision boundary more evenly separates classes by creating a line that is positioned more evenly between the class distributions. The LS line, while close to the boundary between classes, does not separate classes near the boundary as evenly.

# Q1 Part C

**Gamma = 0.003**

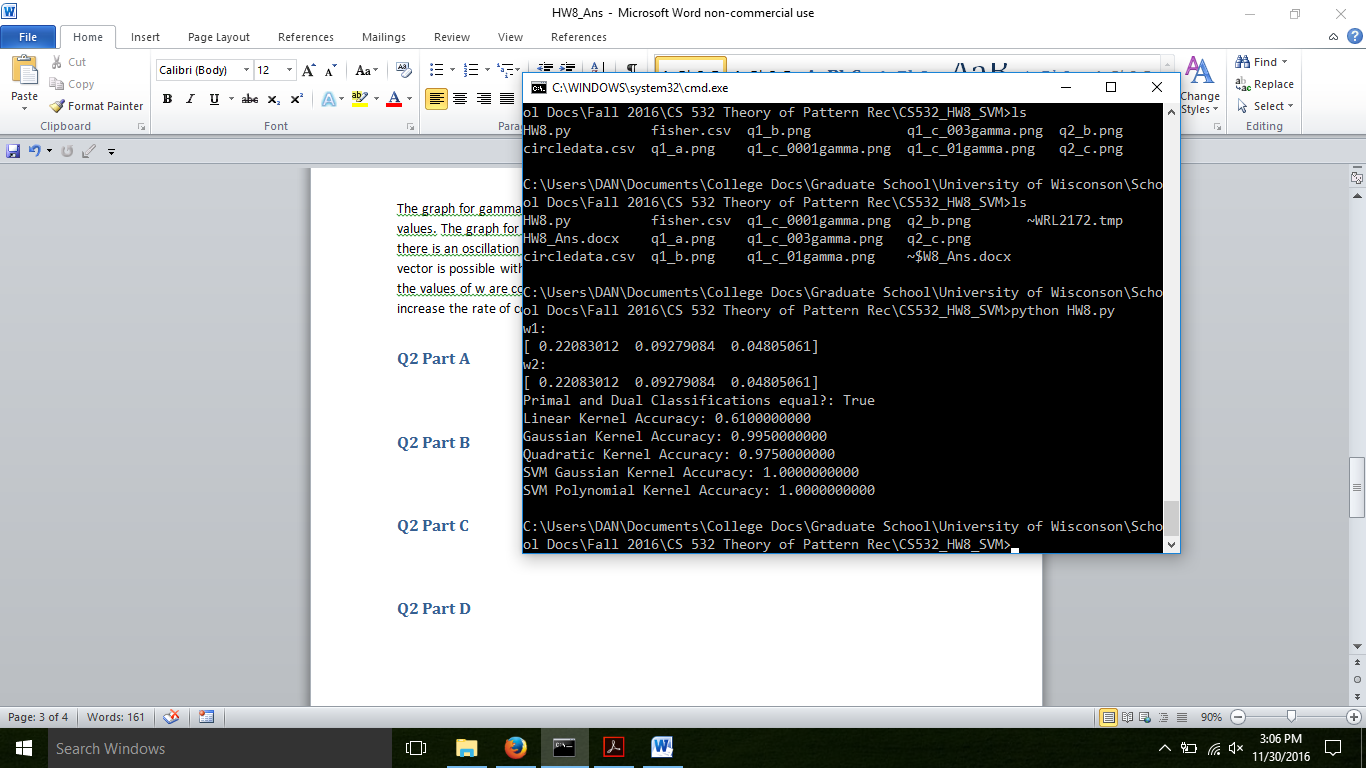
****

**Gamma = 0.01**

**Gamma = 0.0001**

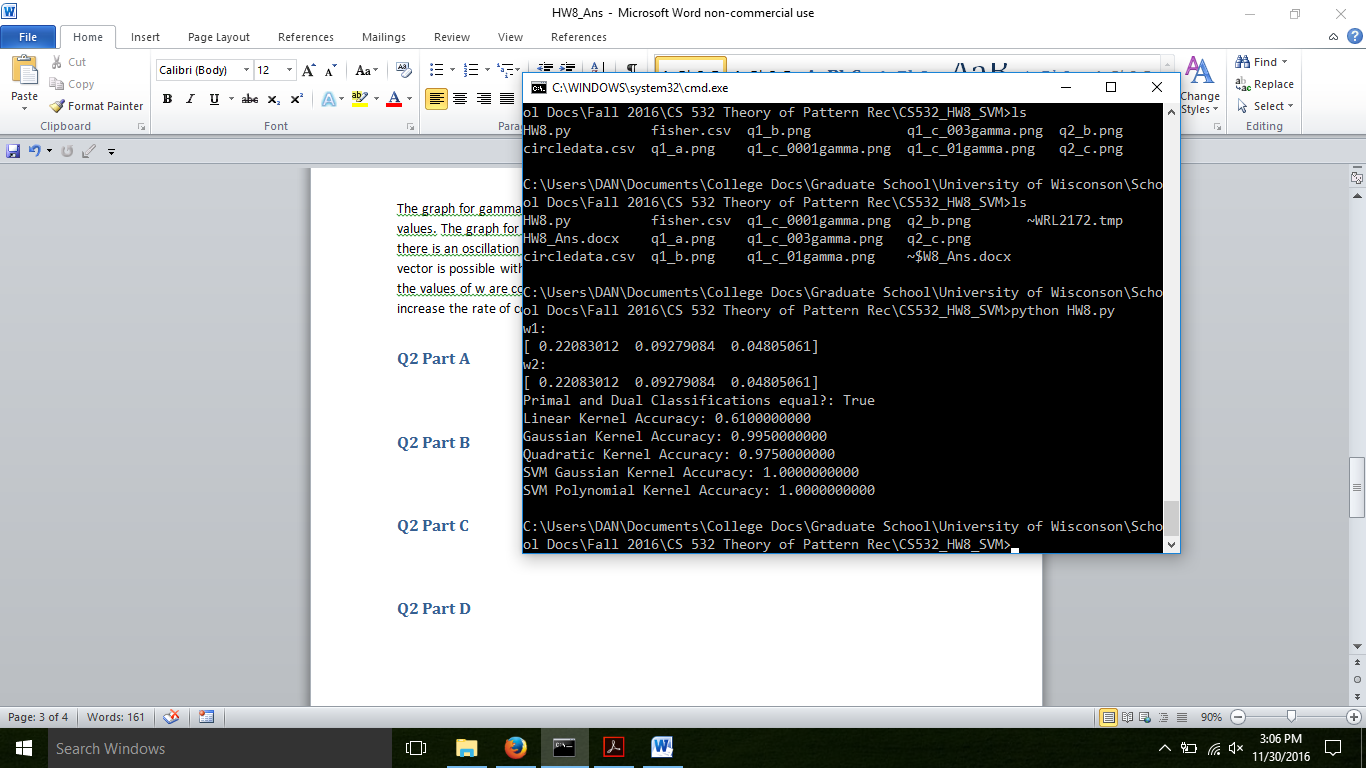
The graph for gamma equaling 0.003 shows that the values of w are converging to steady values. The graph for gamma equaling 0.01 shows that, while convergence happens quickly, there is an oscillation in the w values. This means that convergence to a better w parameter vector is possible with a lower gamma value. The graph for gamma equaling 0.0001 shows that the values of w are converging, but very slowly. This indicates that gamma can be increased to increase the rate of convergence.

# Q2 Part A



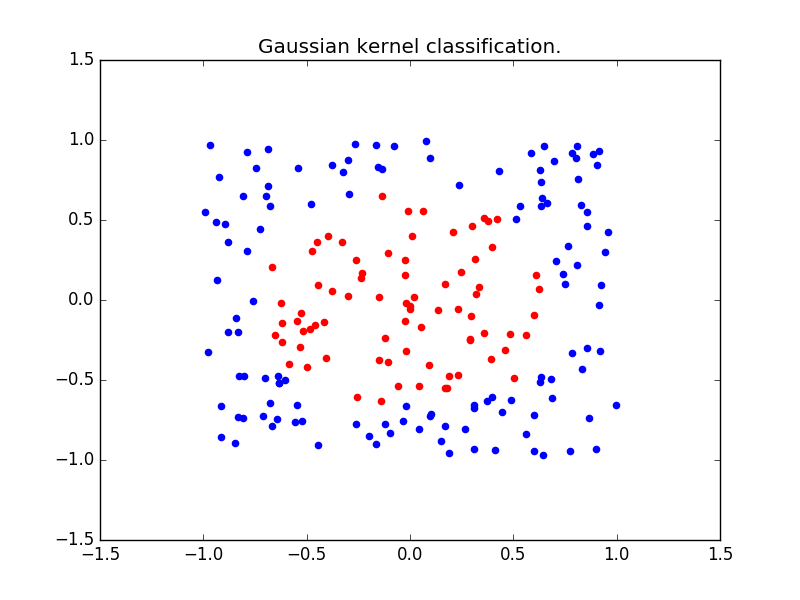
Show above is the output of the program for the primal and dual LS formulations. Both have the same w values, and both produce the same classification lists. Finally, the linear kernel accuracy is shown.

# Q2 Part B

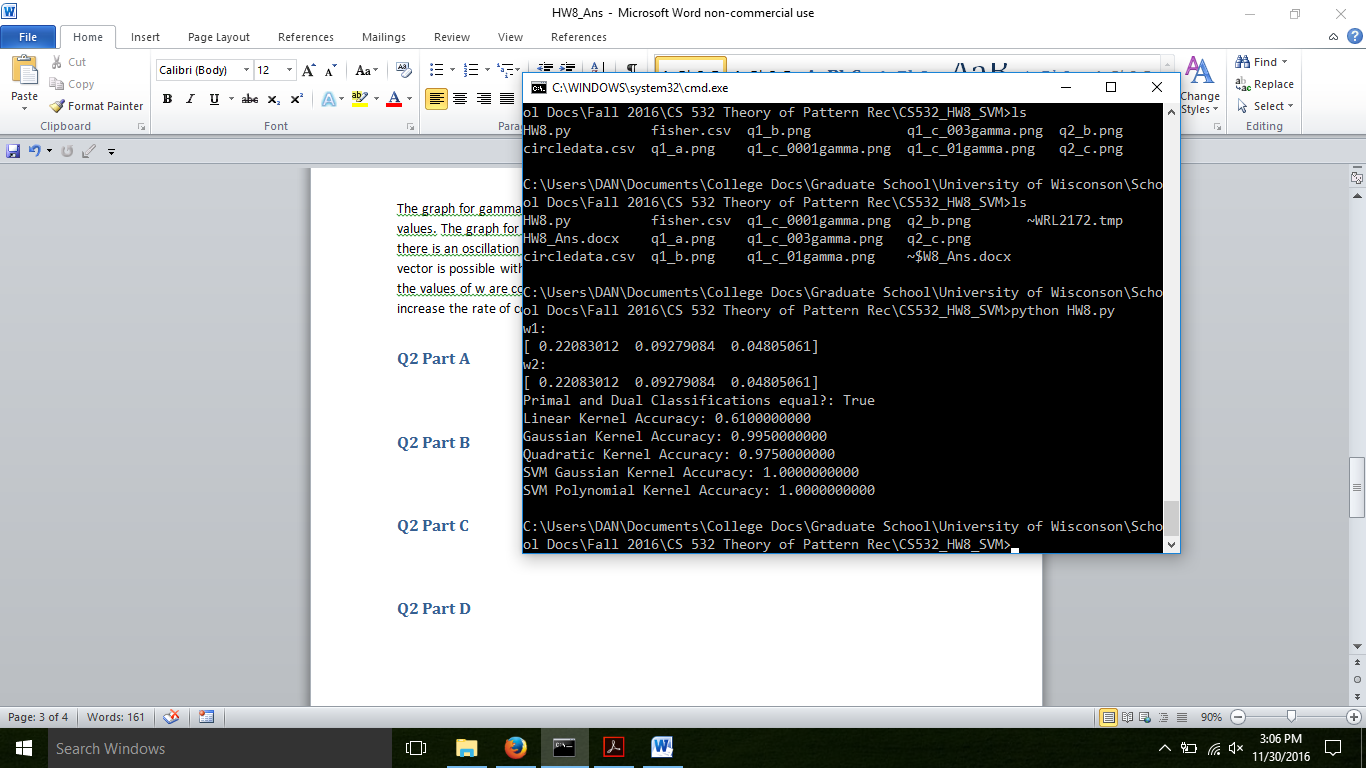


Shown above is the program output for a gaussian kernel LS problem. Its accuracy as compared with the primal linear LS solution is much higher. This is expected since the class distribution is roughly concentric circles.

**Gaussian Kernel Classification Plot**

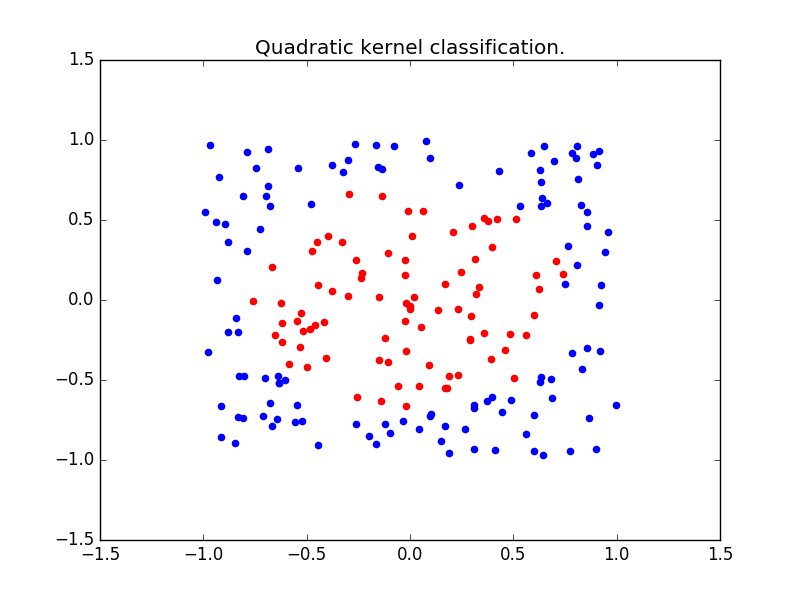
****

# Q2 Part C

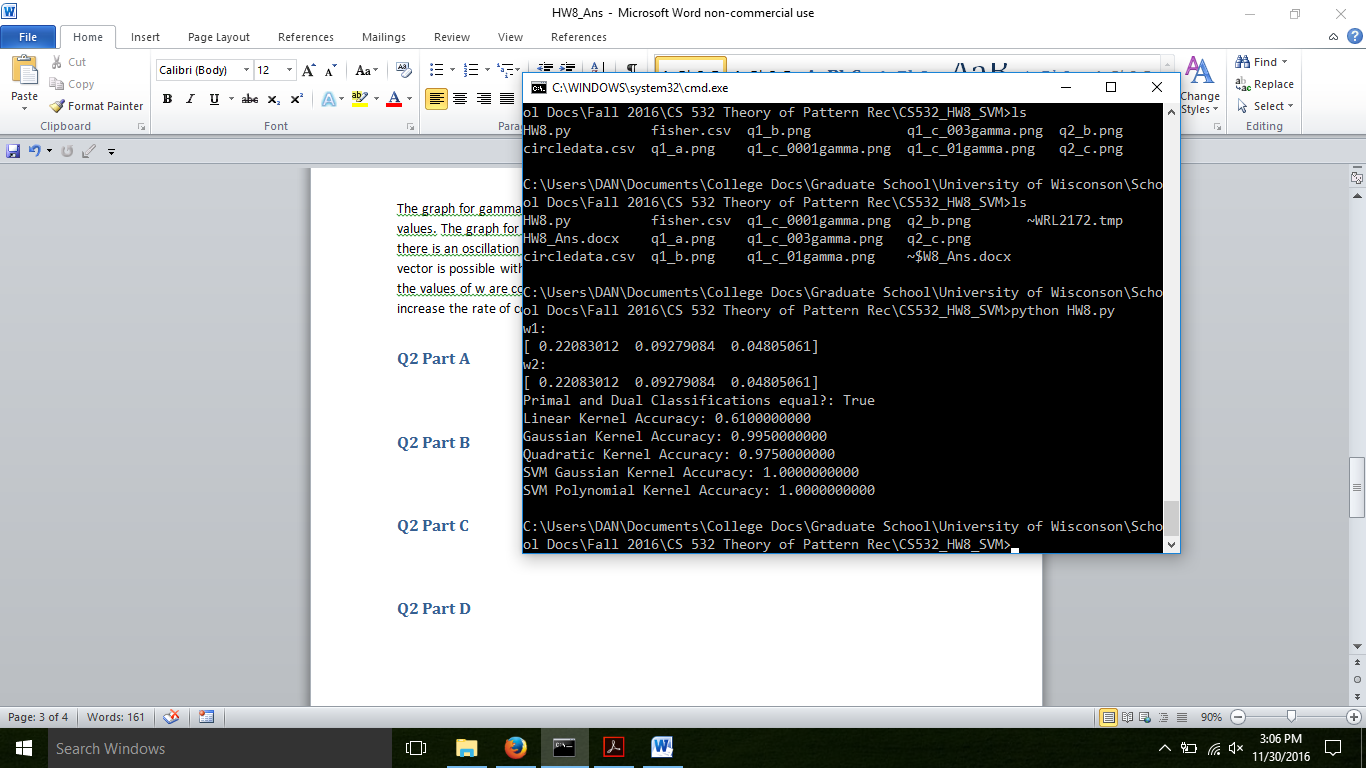


Shown above is the program output for a gaussian kernel LS problem. Its accuracy as compared with the primal linear LS solution is much higher. This is expected since the class distribution is roughly concentric circles. The polynomial kernel however is not as accurate as the Gaussian kernel, indicating that the data distribution more closely a Gaussian distribution, rather than a polynomial one.

**Polynomial Kernel Classification Plot**

****

# Q2 Part D



Shown above is the program output for both a gaussian kernel, and a polynomial kernel LS problem. Scikit-Learn’s svm.SVC class was used to generate a model to fit the data with the previously stated kernels. The resulting accuracy as compared with the LS Gaussian and polynomial kernel solution is slightly higher. This is likely due to the fact that the SVM not only tries to find a decision boundary, but tries to find the boundary which best separates the class distributions. The resulting higher performance means that the decision boundary found due to the hinge loss function can better discriminate between classes than the least squares loss function.