**Support Vector Machines**

**Data Exploration Questions:**

1. What is the number of attributes in each Dataset?
   1. Optdigits Dataset: 64 input+1 class attribute
   2. Amazon Reviews Dataset: 1 input(review)+1 class attribute(rating)
2. What is the number of observations?
   1. Optdigits Dataset:
      1. Training = 3823
      2. Test = 1797
   2. Amazon Reviews Dataset:
      1. Training = 146824
      2. Test = 36707
3. What is the mean and standard deviation of each attribute?
   1. Optdigits Dataset:

****

* 1. Amazon Reviews Dataset:
     1. Training(Rating): Mean = 4.121798 SD = 1.284965
     2. Test(Rating): Mean = 4.115046 SD = 1.285226

1. What is the distribution of the different classes in each of the datasets?
   1. Optdigits Dataset:
      1. Class: No of examples in training set
      2. 0: 376
      3. 1: 389
      4. 2: 380
      5. 3: 389
      6. 4: 387
      7. 5: 376
      8. 6: 377
      9. 7: 387
      10. 8: 380
      11. 9: 382
      12. Class: No of examples in testing set
      13. 0: 178
      14. 1: 182
      15. 2: 177
      16. 3: 183
      17. 4: 181
      18. 5: 182
      19. 6: 181
      20. 7: 179
      21. 8: 174
      22. 9: 180
   2. Amazon Reviews Dataset:
      1. Training:
         1. 12146
         2. 9040
         3. 13364
         4. 26509
         5. 85765
      2. Test:
         1. 3037
         2. 2270
         3. 3415
         4. 6696
         5. 21289

**Results – Optdigits Dataset**

* Result for running Optdigits dataset on SMO (Implements John Platt's **sequential minimal optimization** algorithm for training a support vector classifier.) on WEKA.

**Parameters:**

* For Normalized Poly Kernel
  + Exponent Value = 2.0
  + C = 5.0
  + Cross Validation Applied
  + Standardization of Data



* For Normalized Poly Kernel
  + Exponent Value = 2.0
  + C = 1.0
  + Standardization of Data



* For Normalized Poly Kernel
  + Exponent Value = 2.0
  + C = 3.0
  + Standardization of Data



* For Normalized Poly Kernel
  + Exponent Value = 2.0
  + C = 5.0
  + Standardization of Data



* For Poly Kernel
  + Exponent Value = 2.0
  + C = 5.0
  + Cross Validation Applied
  + Standardization of Data



* For Poly Kernel
  + Exponent Value = 2.0
  + C = 1.0
  + Standardization of Data
  + 
* For Poly Kernel
  + Exponent Value = 2.0
  + C = 2.0
  + Standardization of Data
  + 
* For Poly Kernel
  + Exponent Value = 2.0
  + C = 3.0
  + Standardization of Data



* For Poly Kernel
  + Exponent Value = 2.0
  + C = 5.0
  + Standardization of Data



* For PUK Kernel
  + Exponent Value = 2.0
  + C = 5.0
  + Standardization of Data
  + Cross Validation Data



* For PUK Kernel
  + Exponent Value = 2.0
  + C = 1.0
  + Standardization of Data



* For PUK Kernel
  + Exponent Value = 2.0
  + C = 10.0
  + Standardization of Data



* For PUK Kernel
  + Exponent Value = 2.0
  + C = 100.0
  + Standardization of Data



* For RBF Kernel
  + Exponent Value = 2.0
  + C = 5.0
  + Standardization of Data
  + Cross Validation



* For RBF Kernel
  + Exponent Value = 2.0
  + C = 1.0
  + Standardization of Data



* For RBF Kernel
  + Exponent Value = 2.0
  + C = 5.0
  + Standardization of Data



* For RBF Kernel
  + Exponent Value = 2.0
  + C = 10.0
  + Normalization of Data



* For RBF Kernel
  + Exponent Value = 2.0
  + C = 5.0
  + No Standardization or Normalization of Data



**Amazon Reviews Dataset – Results**

* Results for running **SupportVectorClassifier** Perceptron by scikit-learn of Python, using Bag of Words with TF-IDF (Short Term Frequency – Inverse Document Frequency) on Amazon Reviews Dataset.
* Number of features from Tf-IDF = 2500 from a dummy dataset of containing Train = 8000 instances, Test = 2000 instances.
* SVC Parameters and Results:

1. Poly Kernel with a degree of 10 using balanced weighting for the class.
   1. 
2. RBF Kernel with C=50 and using default gamma value for scikit learn which is auto.
   1. 
3. RBF Kernel with C=100 and using default gamma value for scikit learn which is auto.
   1. 
4. RBF Kernel with C=300 and using gamma value = 0.1.
   1. 
5. RBF Kernel with C=400 and using gamma value = 0.1.
   1. 
6. RBF Kernel with C=300 and using gamma value = 1.0
   1. 
7. RBF Kernel with C=1 and using gamma value = 1.0
   1. 
8. RBF Kernel with C=10 and using gamma value = 1.0
   1. 
9. RBF Kernel with C=100 and using gamma value = 1.0
   1. 
10. Linear Kernel with C=10 and using gamma value = 1.0
    1. 
11. Linear Kernel with C=100 and using gamma value = 1.0
    1. 
12. Linear Kernel with C=100 and using gamma value = 0.1
    1. 
13. Sigmoid Kernel with C=1 and using gamma value = 0.1
    1. 
14. Sigmoid Kernel with C=10 and using gamma value = 0.1
    1. 
15. Sigmoid Kernel with C=100 and using gamma value = 0.1
    1. 
16. Sigmoid Kernel with C=100 and using gamma value=1.0
    1. 