**Exercise 1a - Support Vector Machines (SVMs)**

|  |  |  |
| --- | --- | --- |
| **Task** | **Maximum Points** | **Student Score** |
| Create a personal google account if you don't have one and go to <https://colab.research.google.com/>  You can use either Google Colab or anaconda/Spyder for the first 3 practice exercises and Assignments, but you must use Google Colab for the last two Assignments.  To install anaconda go to: <https://docs.anaconda.com/anaconda/install/windows/>  And follow the instructions there.  Run the anaconda prompt (Start->Anaconda Prompt)  And type the following commands one by one [by typing the command and pressing enter] and follow the instructions to install the required packages and dependencies:  conda install numpy  conda install scikit-learn  conda install spyder  [It’s fine if they are already installed] | 5 |  |
| Download SVM\_440.ipynb and in Colab do File->Open Notebook->Upload and drag in SVM\_440.ipynb  Or run spyder and load svm\_440.py using File->Open | 5 |  |
| Run the notebook or python module and provide a screenshot proving that you were successful | 10 |  |
| Look up the details of the miceprotein data set and provide an overview.   This is a dataset used for classification.   Pay attention to the target variable and be aware that there are 8 possible target classes. | 20 |  |
| Write a description of what the code does. | 20 |  |
| Look up what the 'ovo' parameter does in  svm.SVC(decision\_function\_shape='ovo').  Try running with 'ovr' instead--do you see any differences? Explain. | 10 |  |
| Explain what the confusion matrix is and compare the confusion matrices you see in the training and test sets. | 15 |  |
| Explain what precision,  recall, f1-score and support are and compare these statistics between the train and test cases. | 15 |  |
| **Total:** | 100 |  |