Implement the Support Vector Machine (SVM) algorithm for binary classification for the problem addressed in Project 1 (PLA). Use stochastic gradient descent. **Please do not use an off-the-shelf implementation of SVM from any package**. Use the same data set (as in attachment - Project1\_for reference.zip), i.e., 50 data points for training (25 for each class – separable & non-separable) and 30 for testing. As in the PLA project, consider two separate cases for the training data (the test data set remains the same for the two different training sets): Training data points are linearly separable, and Training data points are not linearly separable.

**Please submit a single doc/docx/pdf/ipynb file** (no other file type, please, with the following exception: **if you are submitting an ipynb file, please ALSO submit the corresponding html)** containing the source code, the training and test data, and a numbered list providing brief notes on:

1. whether the training data points are linearly separable,
2. whether the test points are linearly separable,
3. your initial choice of the weights and constants,
4. the final solution equation of the line (decision boundary),
5. the total number of weight vector updates that your algorithm made,
6. the final misclassification error, if any (expressed as a percentage), on the training data as well as on the test data,
7. the margin width (in both separable and non-separable cases, separately),
8. whether the SVM solution is the same as or different from the PLA solution (i) in the linearly separable case, (ii) in the non-separable case,
9. a brief comparison between the PLA and SVM solutions for the linearly non-separable case (e.g., number of mis-classified points).

Add notes for any other special issues/techniques that you think might be important in your implementation.