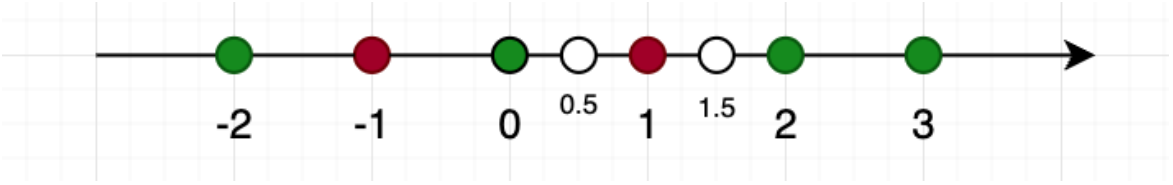
I239 - Assignment 3: SVM

Part 1: Report

a. Using SVM to classify the following training data (all the color points):

- 3 positive points (green points): [-2, 0, 2, 3]
- 2 negative points (red points): [-1, 1]



Note:

- *each point contains 1 dimension.*
- *draw the graph of your solution (your hyper-plane or line you found to separate data into 2 classes).*

Tips:

- *Best solution of this problem can separate all the training points by 1 hyper-plane (or line).*

Require:

- What points are the support vectors (support points) that you choose? why?
- Calculate the SVM weights of hyper-plane (or line) to classify.
- Calculate the hard margin.
- Estimate class (positive or negative) for a new point A=(0.5), B=(1.5).

Part 2: Programming

Implement the SVM algorithm (the sklearn library is allowed) in python language using data from in this colab notebook **(**https://drive.google.com/file/d/1P1-zLbHJ2_yMHMn-Uadcai5p5wFHInVH/view?usp=sharing**)** (the training data is the samples id from 1 - 10, test data is samples id from 11 - 14). The implementation must include two basic functions: train (fit) and predict.

Reference: https://scikit-learn.org/stable/modules/svm.html#classification.
Tips: To mapping all string values or "ranges of numeric values" into numeric values, check the colab notebook for example. SVM only works on the numeric data.
Output requirements:

- Show the support vectors found by your code.
- Show the SVM vector found by your code.
- Show the number of support vectors for each class.
- Calculate Precision, Recall and F1 score of your model.

Notes

***For the report, you must submit the pdf file.**
For programming exercise, you must submit file .py or .ipynb *(file .ipynb downloaded from google colab is recommended)

All submit in a folder with the name:
"StudentID_DT_SVM.zip"
(e.g: 1920xxx.zip)
All wrong format submissions will be rejected.