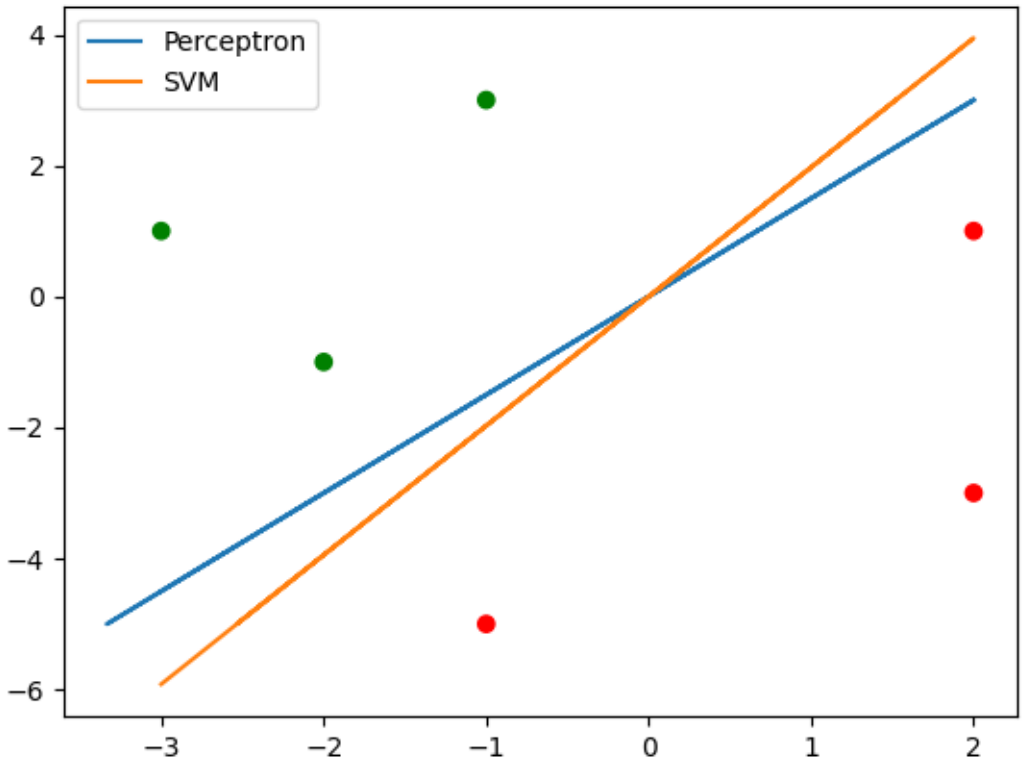
Exercise 1

See Referencesappendix for code.



(a) Perceptron Learning:

Updating vector $w = (0, 0)$ using $(x, y) = ((2, 1), -1)$
$w = (-2, -1) \rightarrow w = (0, 0) + y = -1 * x = (2, 1)$

Updating vector $w = (-2, -1)$ using $(x, y) = ((-1, 3), 1)$
$w = (-3, 2) \rightarrow w = (-2, -1) + y = 1 * x = (-1, 3)$

Margin: 0.3076923076923077

(b) SVM Learning:

0.0 w^* : $(-0.6074814098863559, 0.30790724037405226)$
Margin: 1.9555332420486629

Exercise 2

(a) $\hat{w} = \mathbf{1} = (1, \ldots, 1) \in \{1\}^n$ is a suitable weight vector, since $\langle \hat{w}, x \rangle$ is only positive, iff x contains more 1's than -1's.

(b) $\lambda = n$, since $\|x\|$ is maximum when x consists of either only 1's or only -1's.
$\gamma = \frac{1}{n}$ since the margin is minimal for a x which consists of an by one number off amount of 1's and $-1'$. Thus, $\frac{|\langle w,x \rangle|}{\|w\|} = \frac{1}{n}$
Using Theorem 1.13 we can derive that the perceptron algorithm finds a linear separator after at most $\left(\frac{\lambda}{\gamma}\right)^2 = \left(\frac{n}{\frac{1}{n}}\right)^2 = n^4$ updates.

Appendix

Code for Exercise 1

```python
from matplotlib import pyplot
from sklearn.svm import LinearSVC

S = [
    ((2, 1), -1),
    ((-1, 3), 1),
    ((-3, 1), 1),
    ((-2, -1), 1),
    ((-1, -5), -1),
    ((2, -3), -1),
]


def plot(S, w_perc, w_svm):
    # scatter points
    x_values = [s[0][0] for s in S]
    y_values = [s[0][1] for s in S]
    colors = ['green' if s[1] == 1 else 'red' for s in S]

    pyplot.scatter(x_values, y_values, c=colors)

    # plot linear seperators
    x_min = min(x_values)
    x_max = max(x_values)
    y_min = min(y_values)
    y_max = max(y_values)

    for w in [w_perc, w_svm]:
        ortho_w = (-w[1], w[0])

        p_1 = (x_min, ortho_w[1] * (x_min / ortho_w[0]))
        p_2 = (x_max, ortho_w[1] * (x_max / ortho_w[0]))
        p_3 = (ortho_w[0] * (y_min / ortho_w[1]), y_min)
        p_4 = (ortho_w[0] * (y_max / ortho_w[1]), y_max)

        p_x_values = (p_1[0], p_2[0], p_3[0], p_4[0])
        p_y_values = (p_1[1], p_2[1], p_3[1], p_4[1])

        pyplot.plot(p_x_values, p_y_values, label=('Perceptron' if w==w_perc else 'SVM'))#

    pyplot.legend()
```

```
42
43      # save to file
44      pyplot.savefig(f'exercise_01.png')
45
46
47  def sgn(value) -> int:
48      if value > 0:
49          return 1
50      elif value == 0:
51          return 0
52      else:
53          return -1
54
55
56  def dot_product(a, b) -> int:
57      return a[0] * b[0] + a[1] * b[1]
58
59
60  def check_consistency(S, w) -> bool:
61      for s in S:
62          if sgn(dot_product(s[0], w)) != s[1]:
63              return False
64      return True
65
66
67  def perceptron(S) -> tuple:
68      w = (0, 0)
69      while not check_consistency(S, w):
70          for s in S:
71              if sgn(dot_product(s[0], w)) != s[1]:
72                  w_old = w
73                  # w <- w + yx
74                  w_x = w[0] + s[1] * s[0][0]
75                  w_y = w[1] + s[1] * s[0][1]
76                  w = (w_x, w_y)
77                  # printing formatted for latex. Just copy and paste
78                  print(f'Updating vector $w={w_old}$ using $(x,y)={s}$ \\\\')
79                  print(
80                      f'$w={w} \\rightarrow w={w_old} + y={s[1]} * x={s[0]}$
    \\\\ \n\\bigskip \n')
81      return w
82
83
84  def margin(S, w) -> float:
85      distances = [abs(dot_product(w, s[0]))/(dot_product(w, w)) for s in S]
86      distances = sorted(distances)
87      return distances[0]
88
89
90  def svm(S) -> tuple:
91      classifier = LinearSVC(fit_intercept=False) # force heterogenous (
    fit_intercept=False)
92      classifier.fit([[s[0][0], s[0][1]] for s in S], [s[1] for s in S])
93      print(classifier.intercept_)
94      return (classifier.coef_[0][0], classifier.coef_[0][1])
95
96
97  if __name__ == '__main__':
98      print(f'Perceptron Learning: \\\\ \n\\bigskip \n')
99      w_perc = perceptron(S)
```

```
100     print(f'Margin: ${margin(S,w_perc)}$')
101
102     print(f'SVM Learning: \\\\ \n\\bigskip \n')
103     w_svm = svm(S)
104     print(f'$w^*: {w_svm}$ \\\\')
105     print(f'Margin: ${margin(S, w_svm)}$')
106
107     plot(S, w_perc, w_svm)
```