COMP8118-A5

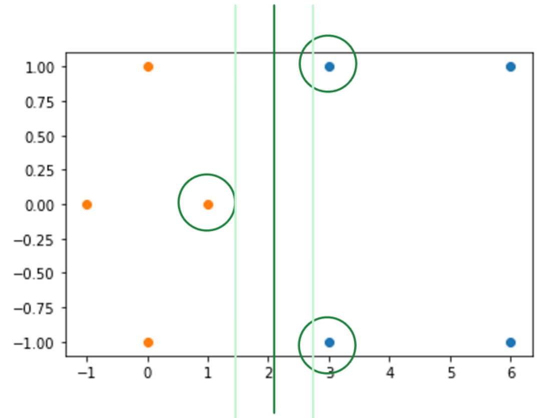
UNIVERSITY OF MEMPHIS

S. Parisa Daj. U00743495

Data Mining Assignment

10/3/2022

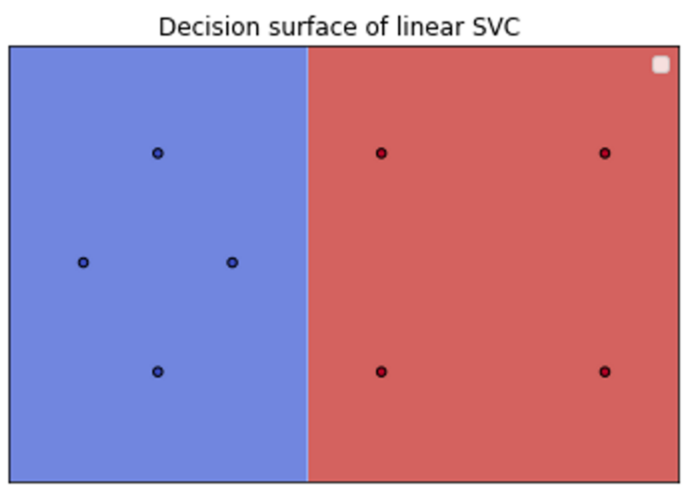
ID	x1	x2	y	
0	1	3	1	1
1	2	3	-1	1
2	3	6	1	1
3	4	6	-1	1
4	5	1	0	-1
5	6	0	1	-1
6	7	0	-1	-1
7	8	-1	0	-1



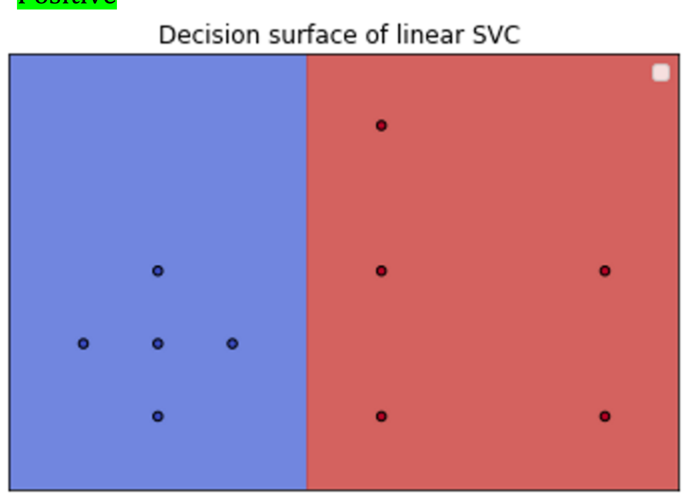
From <https://9z83mig9uzk-496ff2e9c6d22116-0-colab.googleusercontent.com/outputframe.html?vrz=colab-20220930-060045-RC01_478073889>

Q1.

a. Linearly Seperable? Yes
b. Support Vectors: [1., 0.], [3., 1.], [3., -1.]
c. Decision Boundaries:



d. Predict(0, 0) = negative
Predict(3, 3) = Positive



Q2.

Timestamps	x_1	x_2	y
t_1	0.3	0.6	0.2
t_2	0.1	1.0	0.4

a. s0 = y0 = 0

Step 1: Input Forward Propagation

1. *Forget gate variable:* $f_t = \sigma\left(W_f, [x_t \;\; y_{t-1}] + b_f\right)$

$$f_1 = \sigma\left(W_f.[x_1,\, y_0] + b_f\right) = \sigma\left(\begin{pmatrix}0.7\\0.4\\0.1\end{pmatrix}.\begin{pmatrix}0.3\\0.6\\0\end{pmatrix} + 0.1\right)$$
$$= \sigma(0.55) = 0.6341$$

$$f_2 = \sigma\left(W_f.[x_2,\, y_1] + b_f\right) = \sigma\left(\begin{pmatrix}0.7\\0.4\\0.1\end{pmatrix}.\begin{pmatrix}0.1\\1.0\\0.2\end{pmatrix} + 0.1\right)$$
$$= \sigma(0.59) = 0.6434$$

2. *Input gate variable:* $i_t = \sigma(W_i, [x_t \;\; y_{t-1}] + b_i)$

$$i_1 = \sigma(W_i.[x_1,\, y_0] + b_i) = \sigma\left(\begin{pmatrix}0.2\\0.6\\0.7\end{pmatrix}.\begin{pmatrix}0.3\\0.6\\0\end{pmatrix} + 0.4\right)$$
$$= \sigma(0.82) = 0.6942$$

$$i_2 = \sigma(W_i.[x_2,\, y_1] + b_i) = \sigma\left(\begin{pmatrix}0.2\\0.6\\0.7\end{pmatrix}.\begin{pmatrix}0.1\\1.0\\0.2\end{pmatrix} + 0.4\right)$$
$$= \sigma(1.16) = 0.7613$$

3. *Input activation variable:* $a_t = tanh(W_a.[x_t,\, y_{t-1}] + b_a)$

$$a_1 = tanh(W_a.[x_1,\, y_0] + b_a) = tanh\left(\begin{pmatrix}0.3\\0.2\\0.1\end{pmatrix}.\begin{pmatrix}0.3\\0.6\\0\end{pmatrix} + 0.3\right)$$
$$= tanh(0.51) = 0.4699$$

$$a_2 = tanh(W_a.[x_2,\, y_1] + b_a) = tanh\left(\begin{pmatrix}0.3\\0.2\\0.1\end{pmatrix}.\begin{pmatrix}0.1\\1.0\\0.2\end{pmatrix} + 0.3\right)$$
$$= tanh(0.55) = 0.5005$$

4. *Internal state variable:* $s_t = f_t.s_{t-1} + i_t.a_t$

$s_1 = f_1.s_0 + i_1.a_1 = 0.6341 * 0 + 0.6942 * 0.4699 = 0.3262$

$s_2 = f_2.s_1 + i_2.a_2 = 0.6434 * 0.3262 + 0.7613 * 0.5005 = 0.5909$

5. *Output gate variable: o_t=σ(W_o.[x_t, y_(t−1)]+b_o)*

$$o_1 = \sigma(W_o.[x_1,\ y_0] + b_o) = \sigma\left(\begin{pmatrix}0.6\\0.3\\0.1\end{pmatrix}.\begin{pmatrix}0.3\\0.6\\0\end{pmatrix} + 0.2\right)$$
$$= \sigma(0.56) = 0.6364$$
$$o_2 = \sigma(W_o.[x_2,\ y_1] + b_o) = \sigma\left(\begin{pmatrix}0.6\\0.3\\0.1\end{pmatrix}.\begin{pmatrix}0.1\\1.0\\0.2\end{pmatrix} + 0.2\right)$$
$$= \sigma(0.58) = 0.6411$$

6. Final output variable: $y_t = o_t.\tanh(s_t)$

$y_1 = o_1, \tanh(s_1) = 0.6364 * \tanh(0.3262) = 0.2005$

$y_2 = o_2, \tanh(s_2) = 0.6411 * \tanh(0.5909) = 0.3401$

b. Errors

$e_1 = y_1 - y_{1\_old} =$ 0.2005 - 0.2 = 0.0005

$e_2 = y_2 - y_{2\_old} =$ 0.3401 - 0.4 = -0.0599