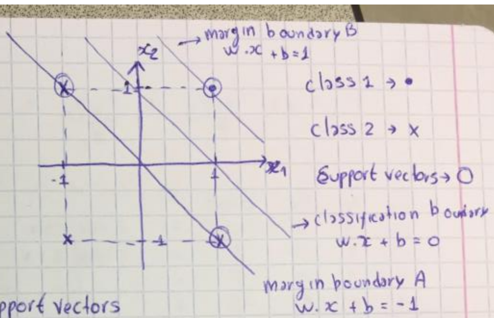
Aprendizagem Automática
Laboratório 5

2.

2.1.

x_1	x_2	dAND
-1	-1	-1
-1	1	-1
1	-1	-1
1	1	1



Support vectors
(1, 1) → 1
(-1, 1) → -1
(1, -1) → -1

$(w.x^s+b)d^s=1$, para todos os support vectors

$$\begin{cases} w_1+w_2+b=1 \\ -w_1+w_2+b=-1 \\ w_1-w_2+b=-1 \end{cases} (\ldots) \Leftrightarrow \begin{cases} b=-1 \\ w_1=w_2=1 \end{cases} \Leftrightarrow w=[1,1]$$

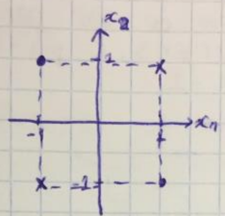
$x_1+x_2-1=0 \Leftrightarrow x_2=-x_1+1$ → hiper plano de decisão

$$\left.\begin{array}{l} x_1+x_2-1=1 \Leftrightarrow \overbrace{x_2=-x_1+2}^{\text{margem B}} \\ x_1+x_2-1=-1 \Leftrightarrow \underbrace{x_2=-x_1}_{\text{margem A}} \end{array}\right\} \text{hiper plano de margem}$$

margem A

2.2.

x_1	x_2	dxOR	
-1	-1	-1	x
-1	1	1	•
1	-1	1	•
1	1	-1	x



$\tilde{x}=f(x)=(x_1,x_2,x_1x_2)^T$

$f(y)=(y_1,y_2,y_1y_2)^T$

$K(x,y)=f(x).f(y) \Leftrightarrow$

$\Leftrightarrow K(x,y)=x_1y_1+x_2y_2+x_1x_2y_1y_2$

Classe 1 → •
Classe -1 → x

Não pode ser utilizada uma classificação linear no ambiente de "input" visto que é impossivel encontrar uma linha recta que separe completamente as duas classes

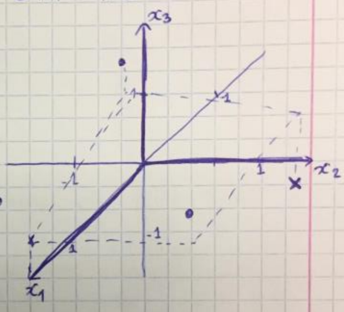
2.3.

$$K(x,y) = \underbrace{x_1 y_1}_{\tilde{x}_1} + \underbrace{x_2 y_2}_{\tilde{x}_2} + \underbrace{x_1 x_2 y_1 y_2}_{\tilde{x}_3}$$

$\tilde{x}_1$	$\tilde{x}_2$	$\tilde{x}_3$	d_{XOR}	
-1	-1	1	-1	x
-1	1	-1	1	∘
1	-1	-1	1	∘
1	1	1	-1	x

classe 1 → ∘
classe -1 → x

GRAFICAMENTE



Pela equação $(\tilde{w} \cdot \tilde{x}^s + b)\, d^s = 1$

$$\begin{cases} -w_1 - w_2 + w_3 + b = -1 \\ -w_1 + w_2 - w_3 + b = 1 \\ w_1 - w_2 - w_3 + b = 1 \\ w_1 + w_2 + w_3 + b = -1 \end{cases} \Leftrightarrow \begin{cases} w_1 = w_2 = 0 \\ w_3 = -1 \\ b = 0 \end{cases}$$

Observando o resultado define-se:

hiperplano de decisão: $\tilde{w} \cdot \tilde{x} + b = 0 \Leftrightarrow \tilde{x}_3 = 0$

hiperplano de margens: $\tilde{w} \cdot \tilde{x} + b = \pm 1 \Leftrightarrow \tilde{x}_3 = \pm 1$

Observando a figura, define-se como plano de decisão $\tilde{x}_3 = 0$
como as margens têm de ser paralelas ao plano de decisão,
então definem-se como margens, o plano $\tilde{x}_3 = -1$ para a classe 1
e $\tilde{x}_3 = 1$ para a classe -1.

2.4.

$$x = \varphi(x) = (\underbrace{x_1}_{x_1}, \underbrace{x_2}_{x_2}, \underbrace{x_1 x_2}_{x_3})$$

Classification boundary $\rightarrow \tilde{x}_3 = 0 \rightarrow x_1 x_2 = 0 \Leftrightarrow x_1 = 0$
$\downarrow$
$x_2 = 0$

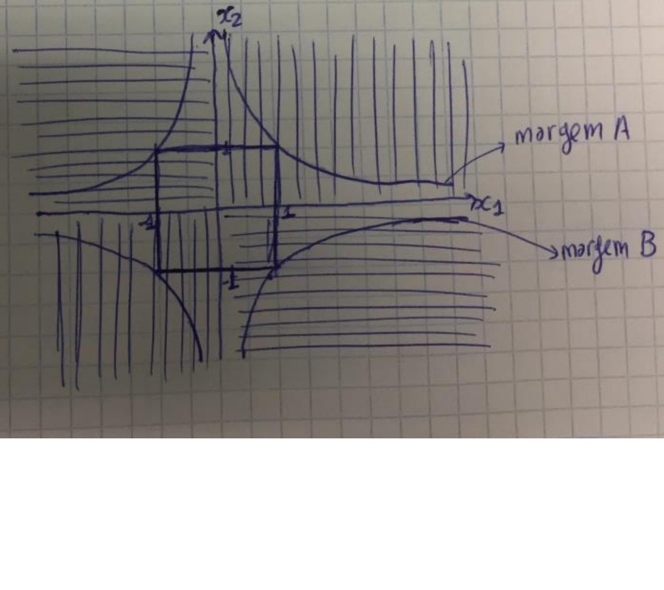
Margin A $\rightarrow \tilde{x}_3 = 1 \rightarrow x_1 x_2 = 1 \Leftrightarrow x_2 = \frac{1}{x_1}$

Margin B $\rightarrow \tilde{x}_3 = -1 \rightarrow x_1 x_2 = -1 \Leftrightarrow x_2 = -\frac{1}{x_1}$

Graficamente

classe 1

classe -1



2.5. Para se obter um output 1:

$$(\tilde{w}\tilde{x}+b)d>0 \xrightarrow{d=1} \tilde{w}\tilde{x}+b=0 \xrightarrow[\tilde{w}_1 - \tilde{w}_2=0]{b=0} \tilde{w}_3\tilde{x}_3>0 \xrightarrow{\tilde{w}_3=-1} -\tilde{x}_3>0$$

Dado que $\tilde{x}_3 = x_1 x_2 \Leftrightarrow -x_1 x_2 > 0 \Leftrightarrow (x_1>0 \wedge x_2<0)$

$$\vee$$

$$(x_1<0 \wedge x_2>0)$$

3.1.(c).

Analisando os resultados obtidos verifica-se que com o aumento do grau do polinómio, valor de p, o erro diminui. Isto é deve-se ao facto de este se adaptar melhor aos dados nas devidas classes. O erro atinge o valor mínimo de 0.36 ,quando o polinómio é de grau 10, e para este mesmo grau atinge-se o numero de 37 “support vectors”. Quando grau do polinómio é superior a grau 10, o número de “support vectors” diminui e tende para 0. Nota-se também um aumento do erro para valores de p superiores a 10. Face aos resultados obtidos e analisando as várias figuras obtidas para os vários valores de p, considera-se como solução ótima o polinómio de grau 10. O resultado obtido esta representado na Figura 1.

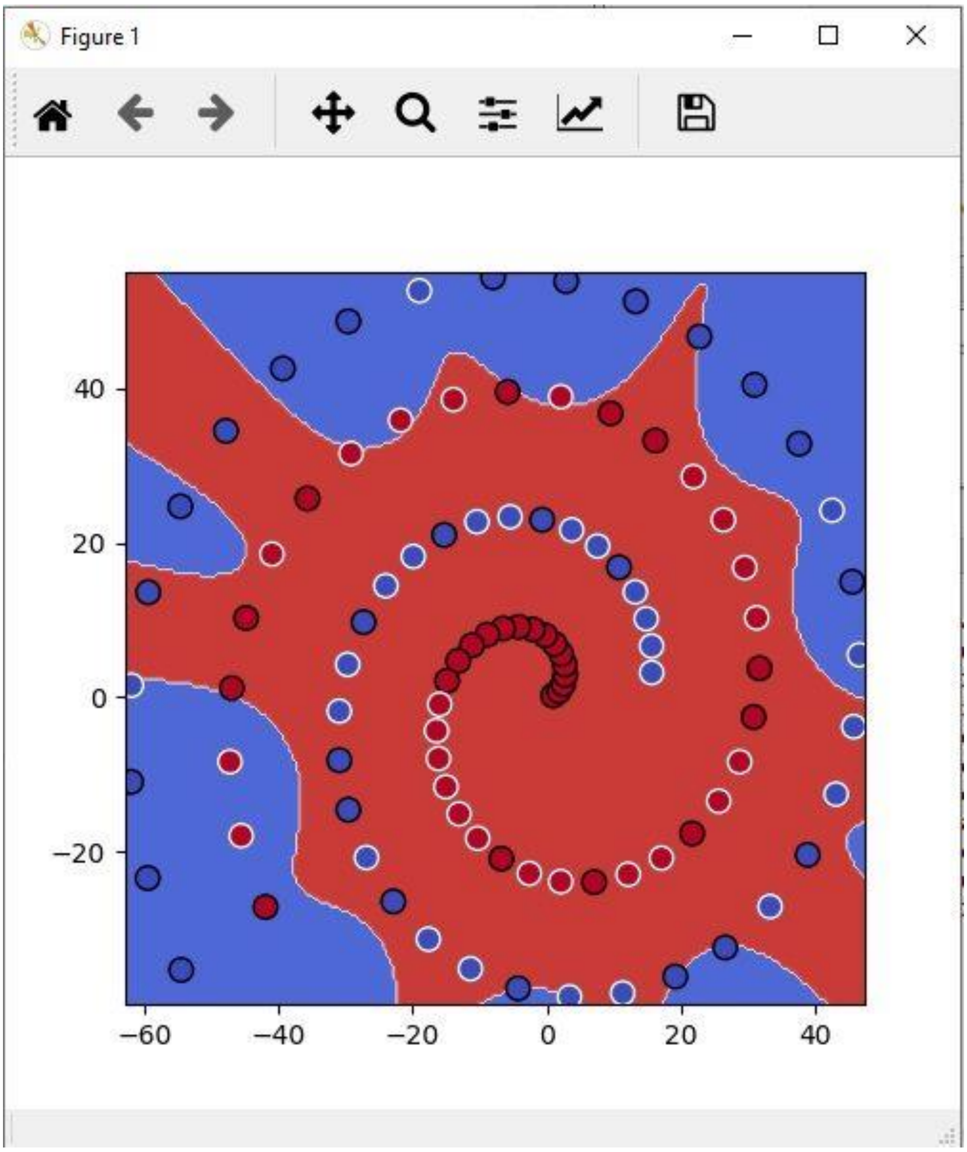


Figura 1

3.1.(d).

Analisando os gráficos, verifica-se que para gamma=0.01, se obtém um erro nulo e o menor número de "support vectors", neste caso 90. Assim sendo escolhemos como solução ótima o valor de gamma atrás definido. Quando comparando com a questão 3.1.(c), apesar de o número de "support vectors" ser superior, o resultado é mais robusto nesta questão. Tal facto deve-se ao aumento das margens face ao gráfico da questão anterior. Acrescenta-se ainda que a utilização de um valor de gamma mais elevado, origina maior erro e consequentemente a necessidade de um maior número "support vectors". O resultado obtido para os resultados escolhidos está representado na Figura 2.

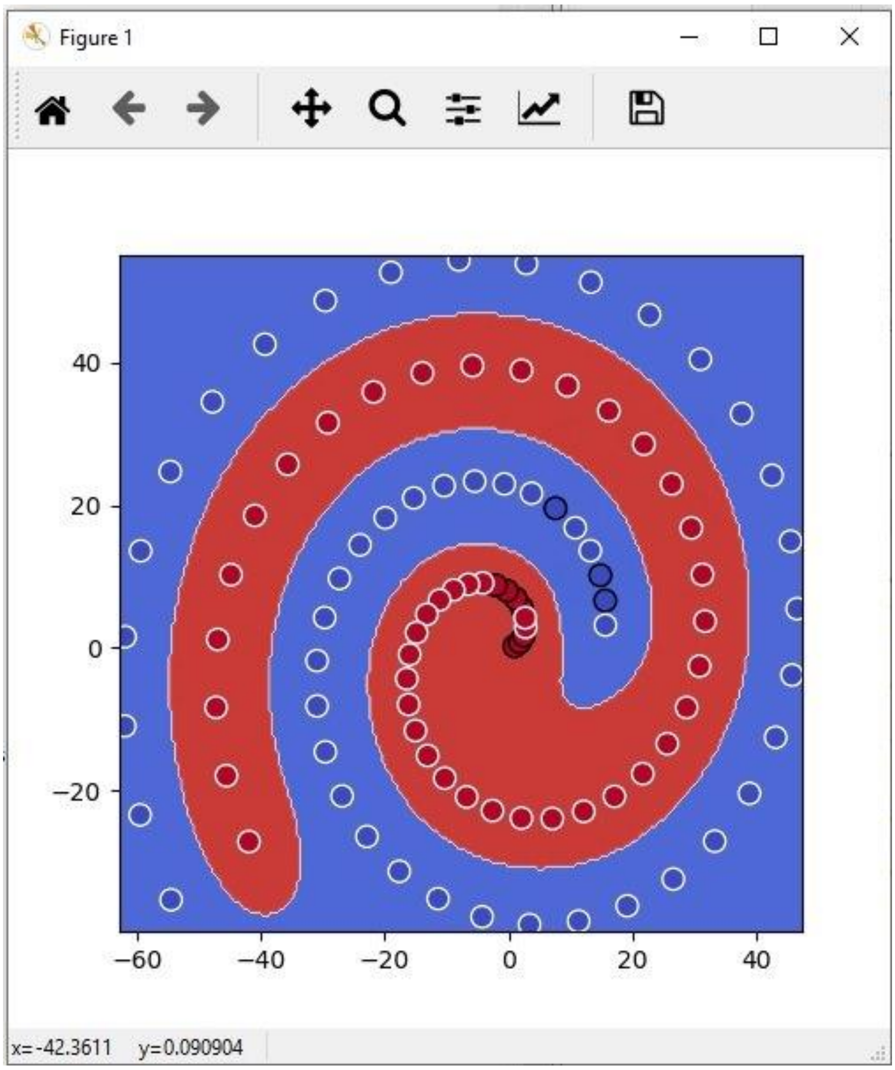


Figura 2

3.2.(c).De acordo com os resultados obtidos, concluímos que , apesar de o erro ser 0 para a maior parte dos valores testados, o valor de gamma que apresenta um numero menor de “support vectors”, é o valor gamma=0.001, que apresenta 11 “support vectors”. O resultado obtido esta representado na Figura 3.

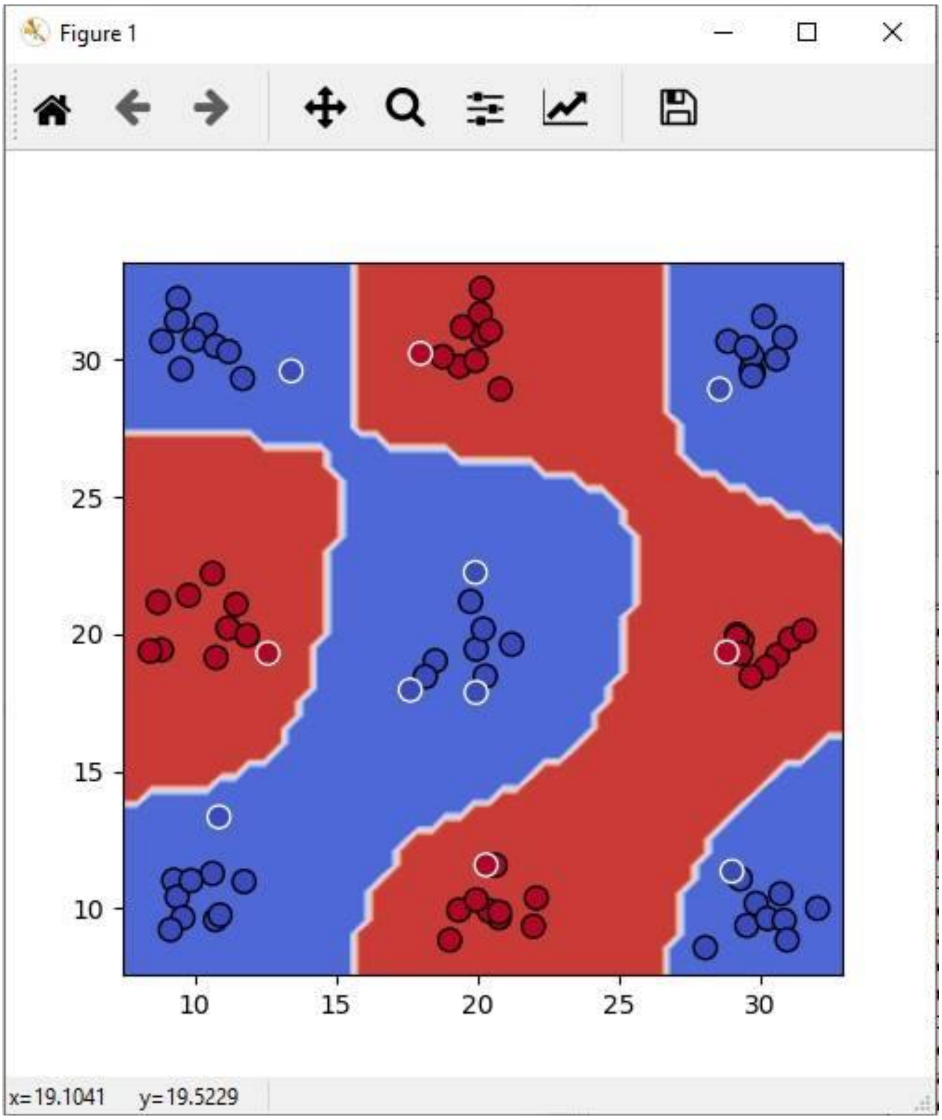


Figura 3

3.3.(c).Usando o mesmo valor de gamma da pergunta anterior (0.001), foi diminuído o valor de C de modo a testar qual o valor que apresenta menor erro. Depois de testados vários valores, concluímos que o erro diminui atingindo um valor mínimo para C=100000, apresentando neste caso 15 "support Vectors". Quando o valor de C continua a diminuir, o erro aumenta assim como o numero de "support vectors". Assim usamos C=100000 e o resultado obtido está representado na Figura 4.

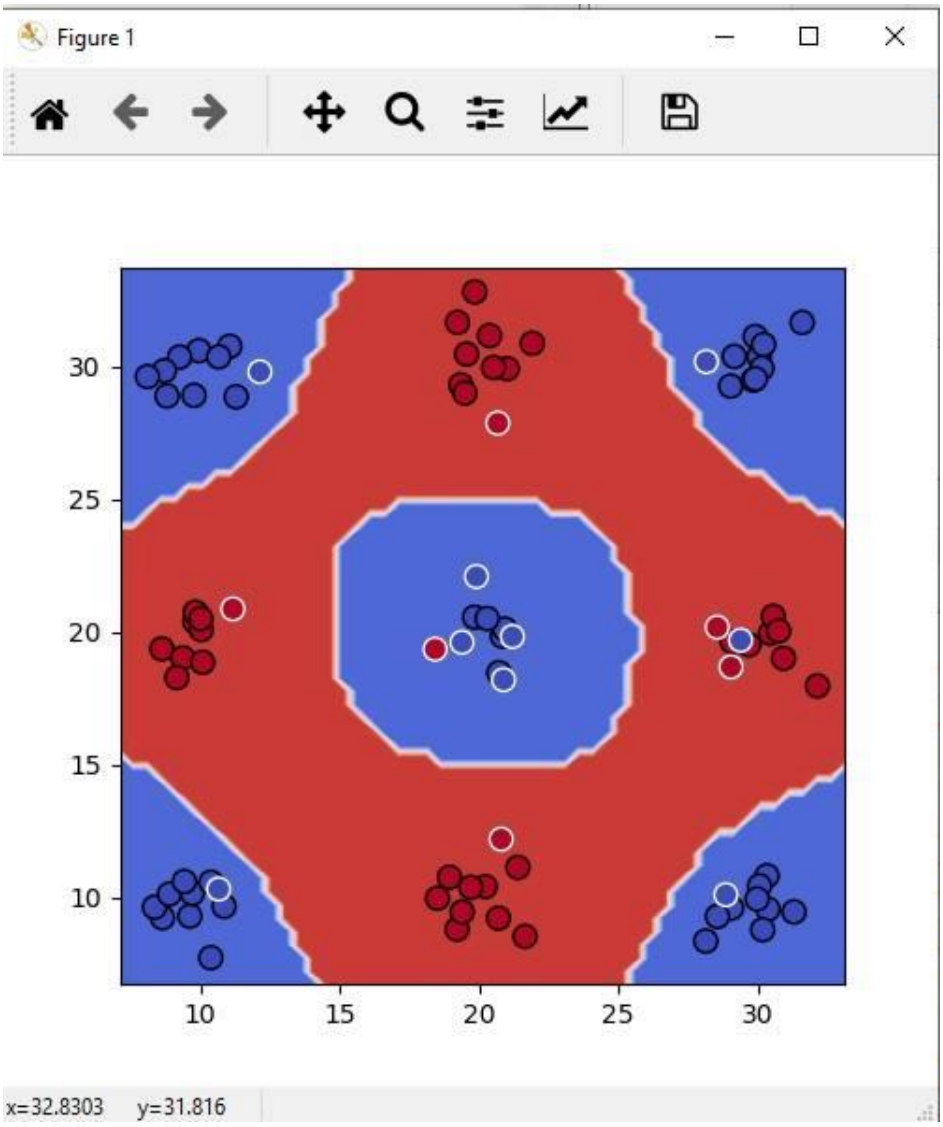


Figura 4

João Morais n83916
Francisco Rabaça n97236