



Máquinas de Vetores de Suporte (SVM)

Exercício

No exercício proposto, foram gerados dados em formato de espiral pertencentes às duas classes distintas, conforme mostra a figura 1, obtidos a partir da biblioteca *mlbench* do R.

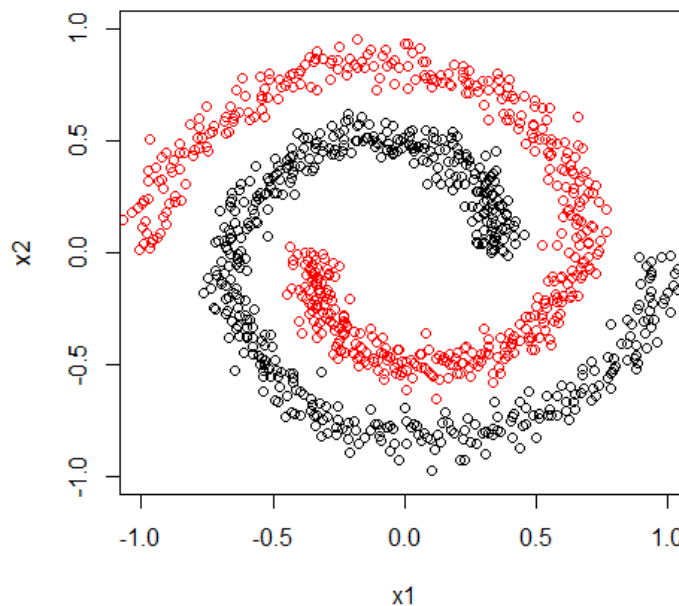


Figura 1: Dados de entrada para o problema.

Os dados foram divididos em 10 *folds* para a técnica de validação cruzada, sendo que a cada iteração, um grupo funcionava como dado de teste e os demais como dados de entrada.

Para cada *fold* criado, os dados de entrada foram utilizados para o treinamento de uma rede SVM. O parâmetro $C = 20$ foi assim definido para que a margem seja pequena e evite um *overfitting*, uma vez que a variância dos dados gerados já era baixa suficiente. O parâmetro $\sigma = 0,8$ foi utilizado para que menos vetores de suporte fossem utilizados, uma vez que o tamanho de cada gaussiana era maior. Os vetores de suporte são mostrados na figura 2.

Esta combinação levou a uma acurácia média de 100%, ou seja, para todos os *folds* criados foram feitas classificações perfeitas dos dados de teste. O hiperplano de separação é mostrado na figura 3, plotado utilizando a função *contour* do R.

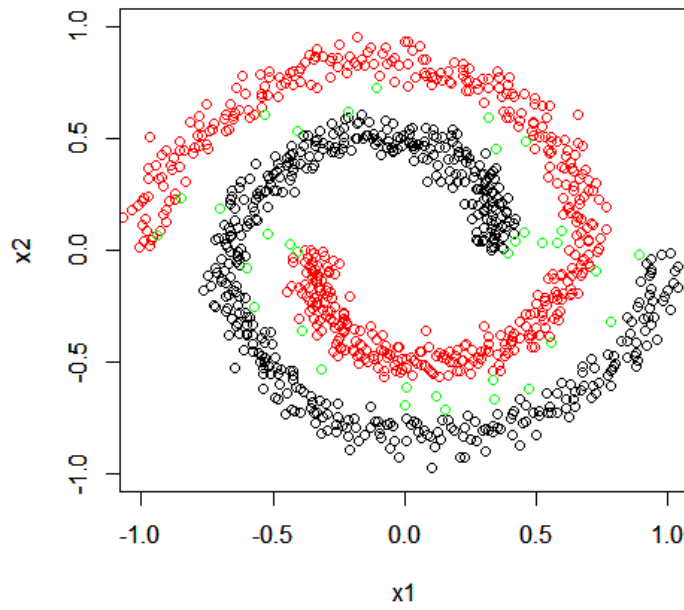


Figura 2: Pontos de Suporte do VSM em verde.

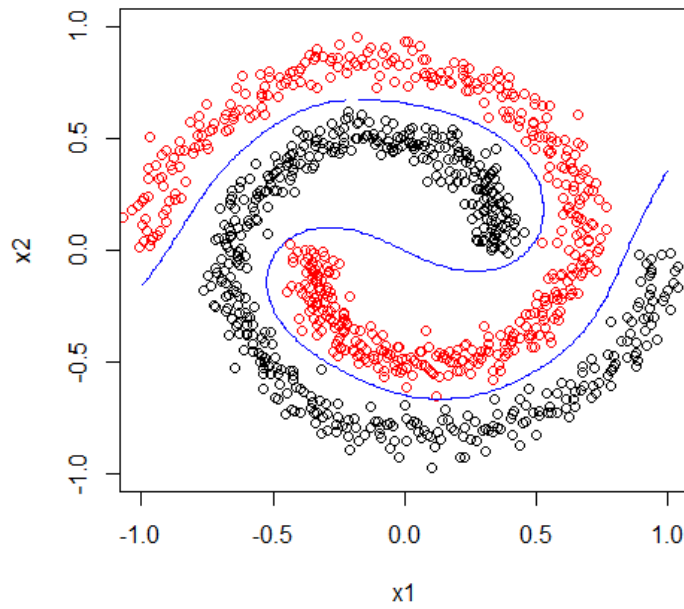


Figura 3: Superfície de Separação em duas dimensões.

Por fim, o hiperplano de separação em três direções, feito com a função *persp3D*, é mostrado na figura 4.

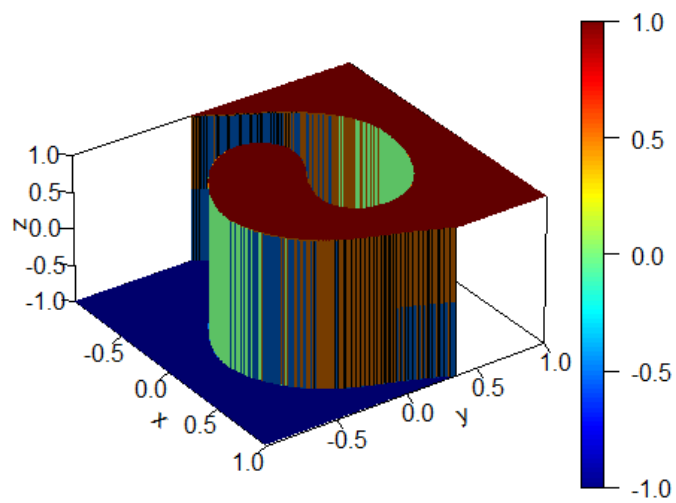


Figura 4: Superfície de Separação em três dimensões.