

32-SVM-RMarkdown

Le Nhat Tung

1 Giới thiệu về SVM

Support Vector Machine (SVM) là một trong những thuật toán học máy có giám sát (supervised learning) phổ biến nhất. SVM được sử dụng rộng rãi trong các bài toán phân loại (classification) và hồi quy (regression).

1.1 Ý tưởng chính của SVM là gì?

Hãy tưởng tượng bạn có một tập dữ liệu gồm nhiều điểm thuộc hai lớp khác nhau (ví dụ: hoa màu đỏ và hoa màu xanh). SVM cố gắng tìm một đường thẳng (trong không gian 2D) hoặc một siêu phẳng (hyperplane) (trong không gian nhiều chiều) tốt nhất để phân tách hai lớp dữ liệu này.

Đường thẳng “tốt nhất” là đường có **lề (margin) lớn nhất** - tức là khoảng cách từ đường đến các điểm dữ liệu gần nhất của mỗi lớp là lớn nhất.

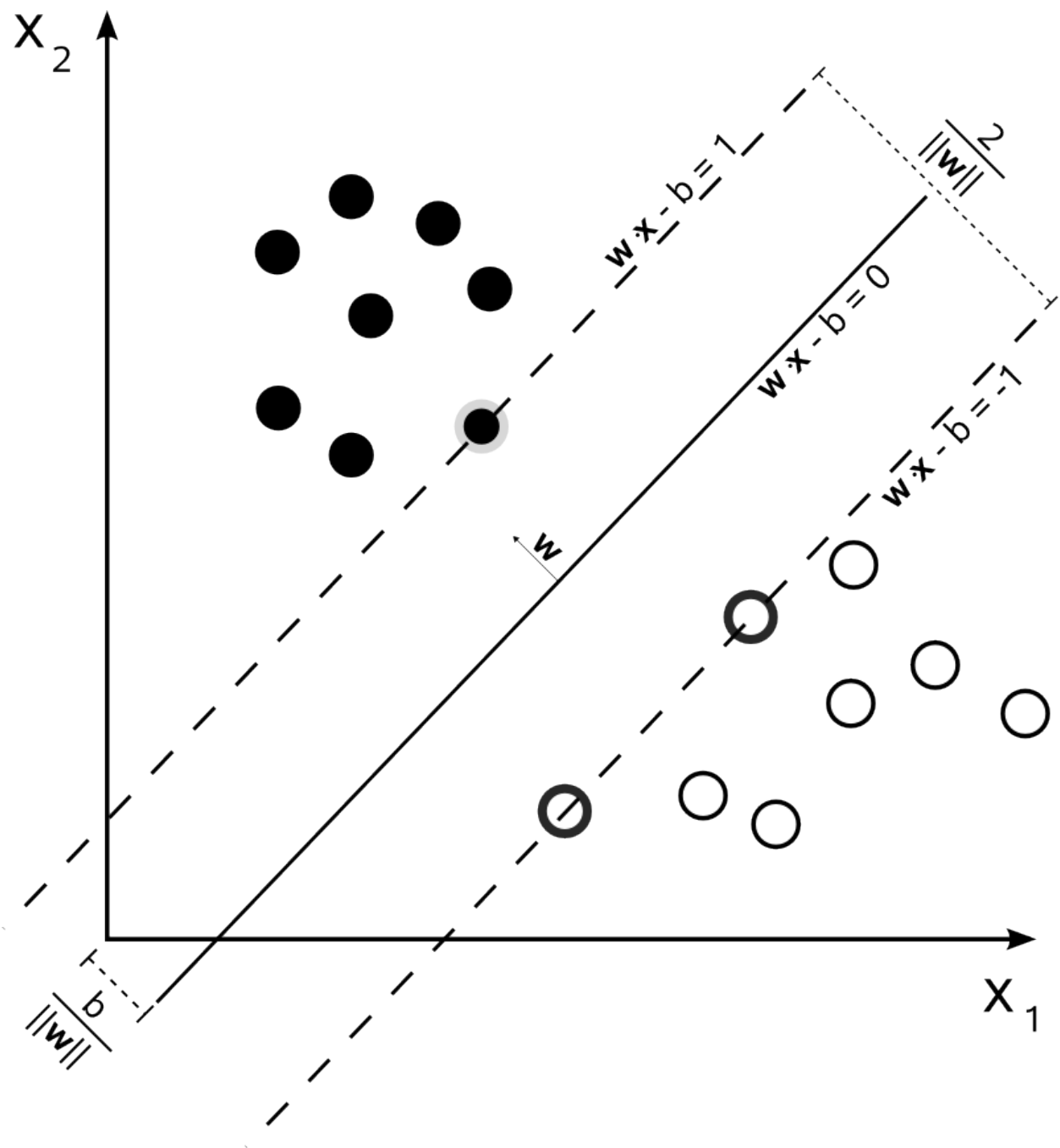


Figure 1: “Mo ta SVM”