1. Có cần phải scale features trước khi huấn luyện các SVM models không? Tại sao?

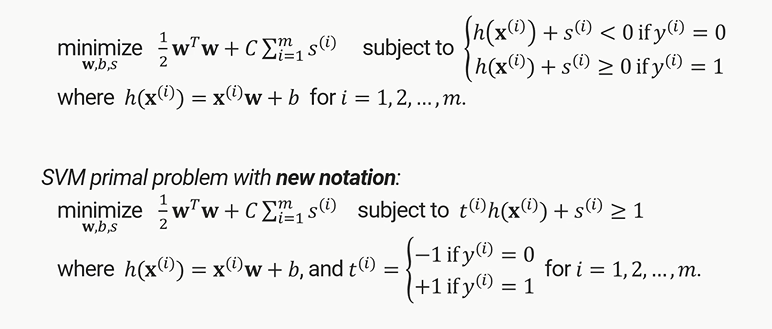
Có cần thiết. Phải scaling features trước khi huấn luyện các model SVM. Vì các thuật toán SVM nhạy cảm với features scaling vì nếu các features khoảng giá trị không bằng nhau thì gây ra margin hẹp. Và nếu khoảng giá trị các thuộc tính bằng nhau thì margin sẽ lớn hơn lúc đó model như thế sẽ tốt hơn.

1. Primal và dual problem là gì? Tại sao cần giải dual problem mà không giải trực tiếp primal problem?

Primal problem là bài toán gốc

Dual problem là bài toán gần với bài toán gốc (solution là chặn dưới của primal problem)

Cần giải dual problem mà không giải trực tiếp primal problem vì khi giải được dual problem thì mình sẽ biết được chặn dưới của primal problem nên mình sẽ giải được primal problem và giải dual problem nó đơn giản nhanh chóng hơn.



1. Khi một SVM classifier với Gaussian RBF kernel bị underfitting thì chúng ta nên tăng hay giảm tham số γ (gamma) và tham số C? Tại sao?

Khi một SVM classifier với Gaussian RBF kernel bị underfitting thì chúng ta nên giảm gamma và tham số C. Vì C càng nhỏ thì regularization càng mạnh càng ít ngăn overfitting nên nó tránh được underfitting nhưng phải giảm để gamma và C hợp lí để tránh trường hợp overfitting.