



MÁY HỌC NÂNG CAO (Máy học véc-to hỗ trợ - SVM)

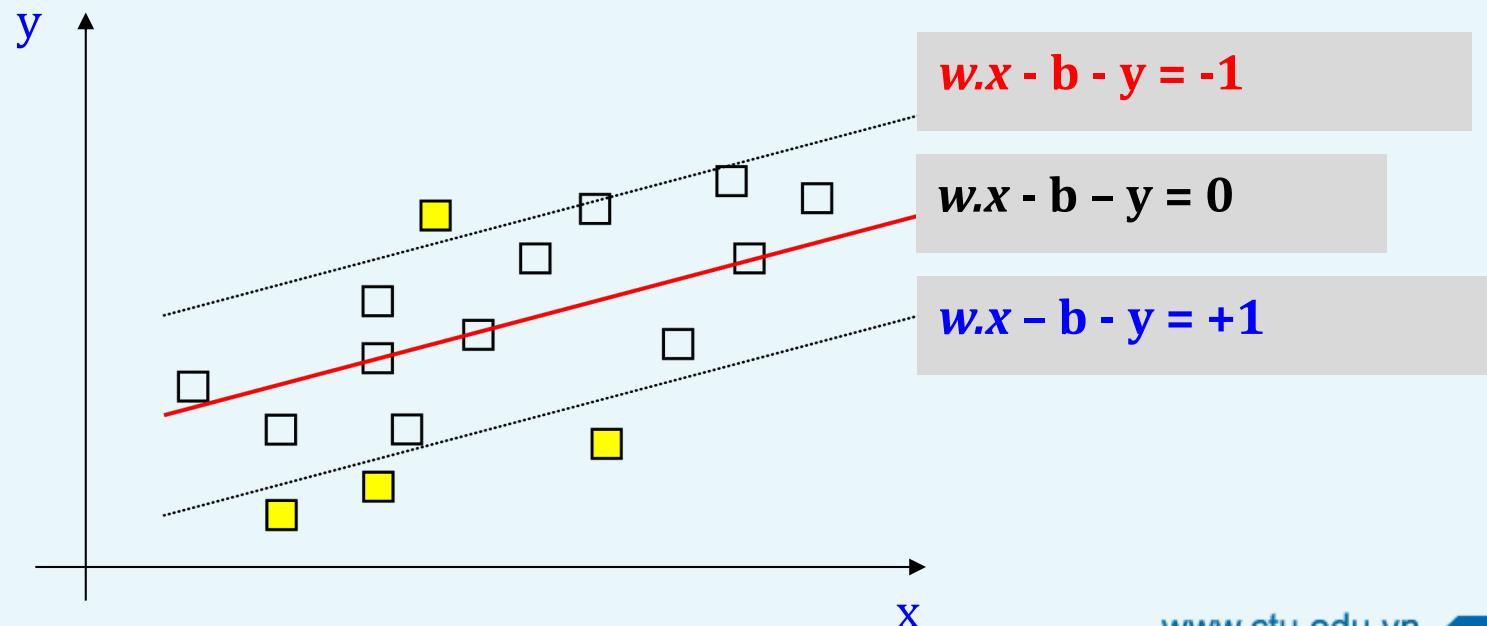
Phạm Nguyên Khang
pnkhang@cit.ctu.edu.vn



CANTHO UNIVERSITY

SVM cho bài toán hồi quy

- SVM cho bài toán hồi quy
 - Tìm đường thẳng (siêu phẳng) “đi qua” dữ liệu
 - $y = wx - b$
 - Sự chênh lệch giá trị thật (y) đến giá trị dự báo (đường thẳng hay $wx - b$) không quá ϵ

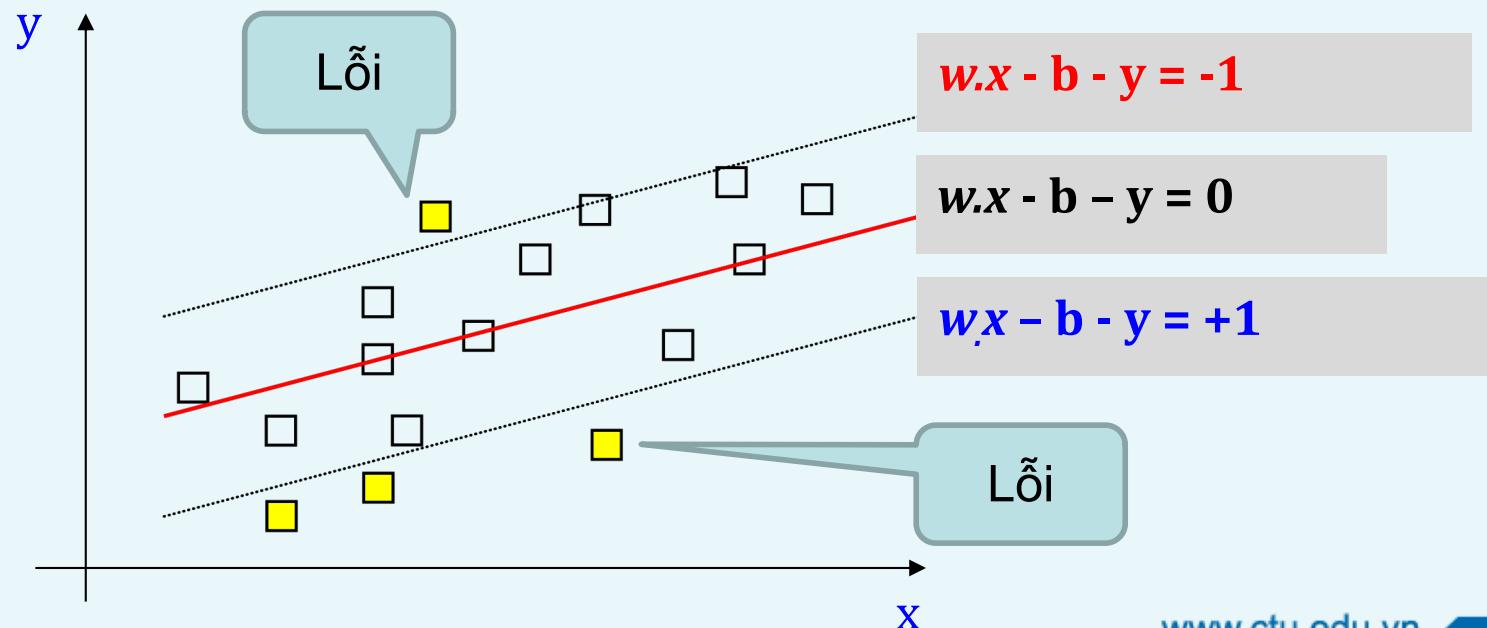




CANTHO UNIVERSITY

SVM cho bài toán hồi quy

- SVM cho bài toán hồi quy
 - Tìm đường thẳng (siêu phẳng) “đi qua” dữ liệu
 - $y = w \cdot x - b$
 - Sự chênh lệch giá trị thật (y) đến giá trị dự báo (đường thẳng hay $wx - b$) không quá ε





CANTHO UNIVERSITY

SVM cho bài toán hồi quy

- SVM cho bài toán hồi quy
 - Tìm đường thẳng (siêu phẳng) “đi qua” dữ liệu
 - $y = wx - b$
 - Khoảng cách từ các điểm dữ liệu đến đường thẳng không quá ε
$$-\varepsilon \leq w.x_i - b - y_i \leq \varepsilon$$
 - Lỗi: những phần tử nằm ngoài phạm vi
 - Bài toán tối ưu
 - Tìm w, b sao cho: $\frac{1}{2} |w|^2 + c \sum_i (z_i^* + z_i) \Rightarrow \min$
 - Với ràng buộc:
$$w.x_i - b - y_i - z_i^* \leq \varepsilon$$
$$w.x_i - b - y_i + z_i \geq -\varepsilon$$
$$z_i^*, z_i \geq 0 \quad (i=1, 2, \dots, m)$$



CANTHO UNIVERSITY

SVM cho bài toán hồi quy

- Sử dụng SVM cho bài toán hồi quy
 - Sau khi tìm được w , b , ta có phương trình đường thẳng hồi quy:

$$y = w \cdot x - b$$

- Dự báo giá trị cho phần tử mới:

$$\text{predict}(x_{\text{mới}}) = w \cdot x_{\text{mới}} - b$$



CANTHO UNIVERSITY

SVM cho bài toán hồi quy

- SVM cho bài toán hồi quy
 - Giải bài toán hồi quy bằng python (xem bảng)

$$\frac{1}{2} |w|^2 + c \sum_i (z_i^* + z_i) \Rightarrow \min$$

$$\begin{aligned} w \cdot x_i - b - y_i - z_i^* &\leq \varepsilon \\ w \cdot x_i - b - y_i + z_i &\geq -\varepsilon \\ z_i^*, z_i &\geq 0 \quad (i=1, 2, \dots, m) \end{aligned}$$



CANTHO UNIVERSITY

SVM cho bài toán hồi quy

- Bài toán đối ngẫu: tuyến tính

$$\begin{aligned} \min \Psi(\alpha, \alpha^*) &= (1/2) \sum_{i=1}^m \sum_{j=1}^m ((\alpha_i - \alpha_i^*)(\alpha_j - \alpha_j^*)(x_i \cdot x_j)) \\ &\quad - \sum_{i=1}^m (\alpha_i - \alpha_i^*)y_i + \varepsilon \sum_{i=1}^m (\alpha_i + \alpha_i^*) \end{aligned}$$

s.t.

$$\sum_{i=1}^m (\alpha_i - \alpha_i^*) = 0$$

$$c \geq \alpha_i, \alpha_i^* \geq 0$$

hằng $c > 0$ sử dụng để chỉnh độ rộng lề và lối



CANTHO UNIVERSITY

SVM cho bài toán hồi quy

- Bài toán đối ngẫu: dự báo

$$predict(x) = \sum_{i=1}^{\#SV} (\alpha_i - \alpha_i^*)(x_i \cdot x) - b$$



CANTHO UNIVERSITY

SVM cho bài toán hồi quy

- Bài toán đối ngẫu: phi tuyến

$$\begin{aligned} \min \Psi(\alpha, \alpha^*) = & (1/2) \sum_{i=1}^m \sum_{j=1}^m ((\alpha_i - \alpha_i^*)(\alpha_j - \alpha_j^*) K(x_i, x_j)) \\ & - \sum_{i=1}^m (\alpha_i - \alpha_i^*) y_i + \varepsilon \sum_{i=1}^m (\alpha_i + \alpha_i^*) \end{aligned}$$

s.t.

$$\sum_{i=1}^m (\alpha_i - \alpha_i^*) = 0$$

$$c \geq \alpha_i, \alpha_i^* \geq 0$$

hàng $c > 0$ sử dụng để chỉnh độ rộng lề và lỗi



CANTHO UNIVERSITY

SVM cho bài toán hồi quy

- Bài toán đối ngẫu: dự báo phi tuyến

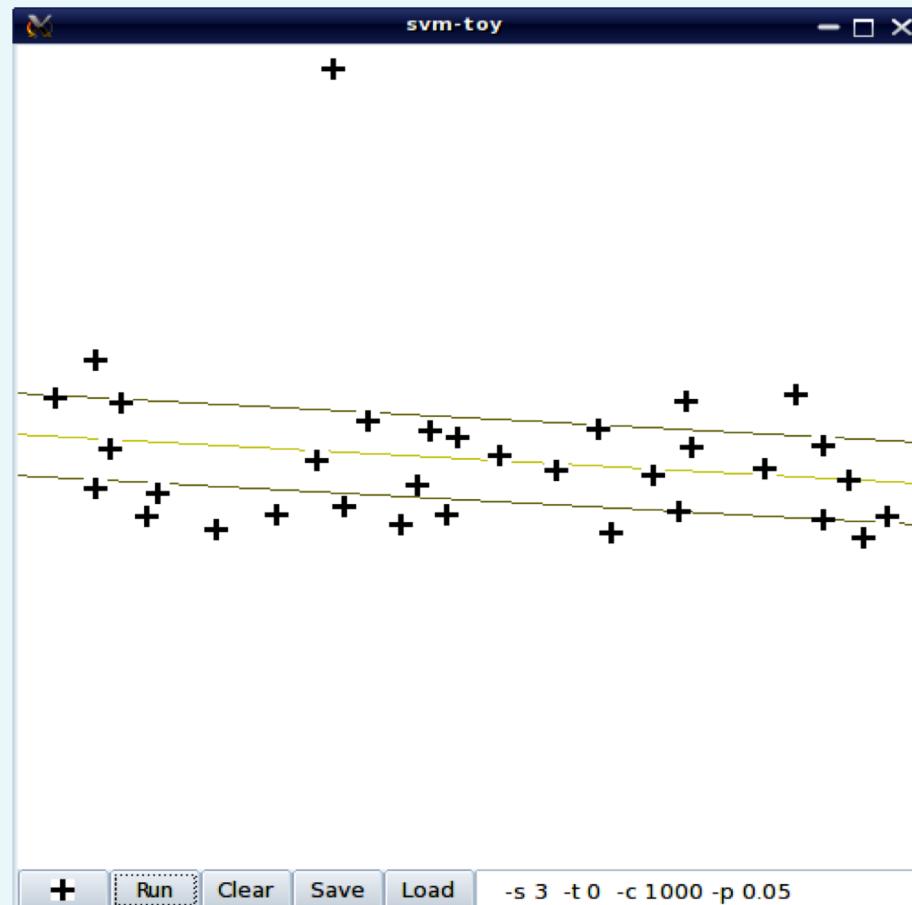
$$predict(x) = \sum_{i=1}^{\#SV} (\alpha_i - \alpha_i^*) K(x_i, x) - b$$



CANTHO UNIVERSITY

SVM cho bài toán hồi quy

- Ví dụ

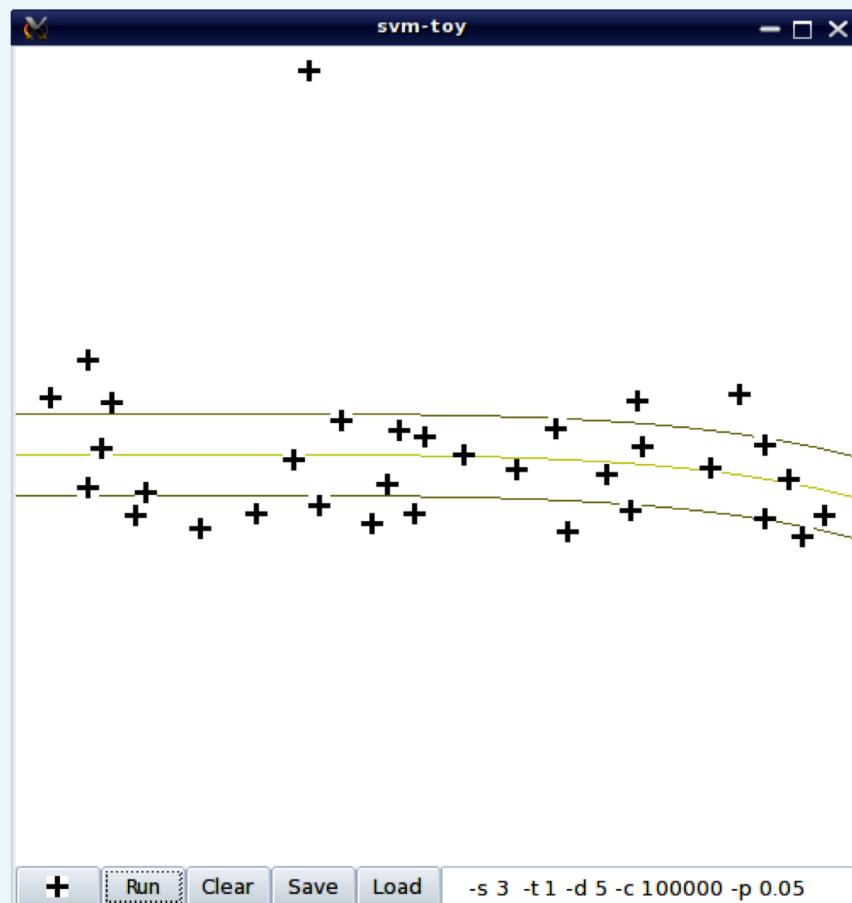




CANTHO UNIVERSITY

SVM cho bài toán hồi quy

- Ví dụ





CANTHO UNIVERSITY

SVM cho bài toán hồi quy

- Ví dụ

