

ONLINE MASTERS IN **DATA SCIENCE**

DSC 255 - MACHINE LEARNING FUNDAMENTALS

AN SVM EXAMPLE

SANJOY DASGUPTA, PROFESSOR

UC San Diego

COMPUTER SCIENCE & ENGINEERING
HALICIOĞLU DATA SCIENCE INSTITUTE



The (Soft-Margin) SVM

Allow each data point $x^{(i)}$ some **slack** ξ_i .

$$\begin{aligned} \min_{w \in \mathbb{R}^d, b \in \mathbb{R}, \xi \in \mathbb{R}^n} \quad & \|w\|^2 + C \sum_{i=1}^n \xi_i \\ \text{s.t.: } & y^{(i)}(w \cdot x^{(i)} + b) \geq 1 - \xi_i \quad \text{for all } i = 1, 2, \dots, n \\ & \xi \geq 0 \end{aligned}$$

Parameter C manages the tradeoff between margin and slack.

Sentiment Data

Sentences from reviews on Amazon, Yelp, IMDB, each labeled as positive or negative.

- Needless to say, I wasted my money.
- He was very impressed when going from the original battery to the extended battery.
- I have to jiggle the plug to get it to line up right to get decent volume.
- Will order from them again!

Data details:

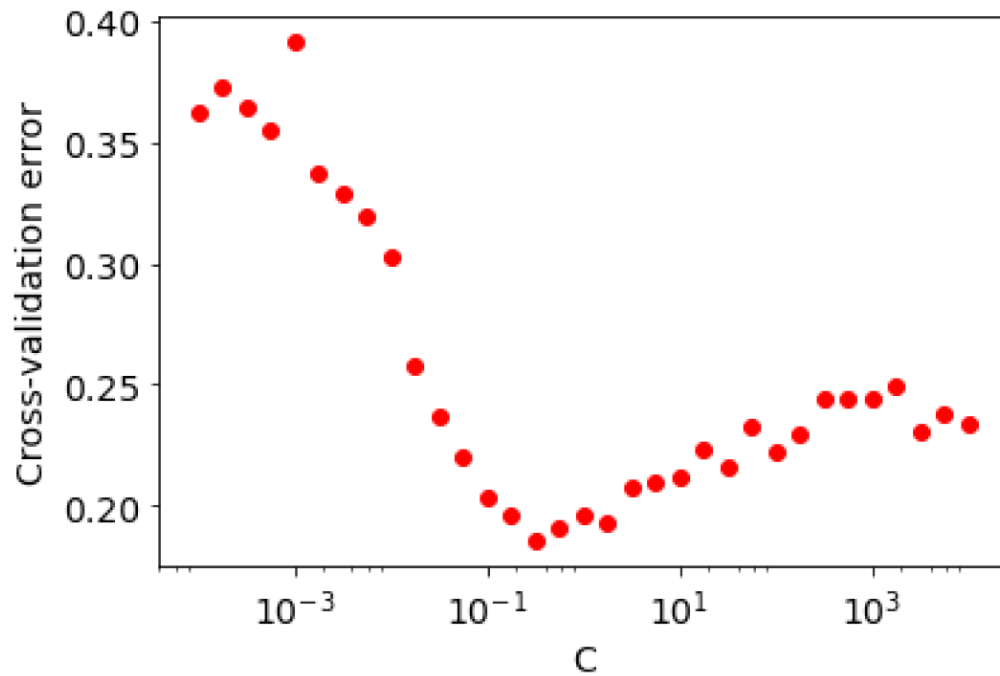
- Bag-of-words representation using a vocabulary of size 4500
- 2500 training sentences, 500 test sentences

What C to use?

C	training error (%)	test error (%)	# support vectors
0.01	23.72	28.4	2294
0.1	7.88	18.4	1766
1	1.12	16.8	1306
10	0.16	19.4	1105
100	0.08	19.4	1035
1000	0.08	19.4	950

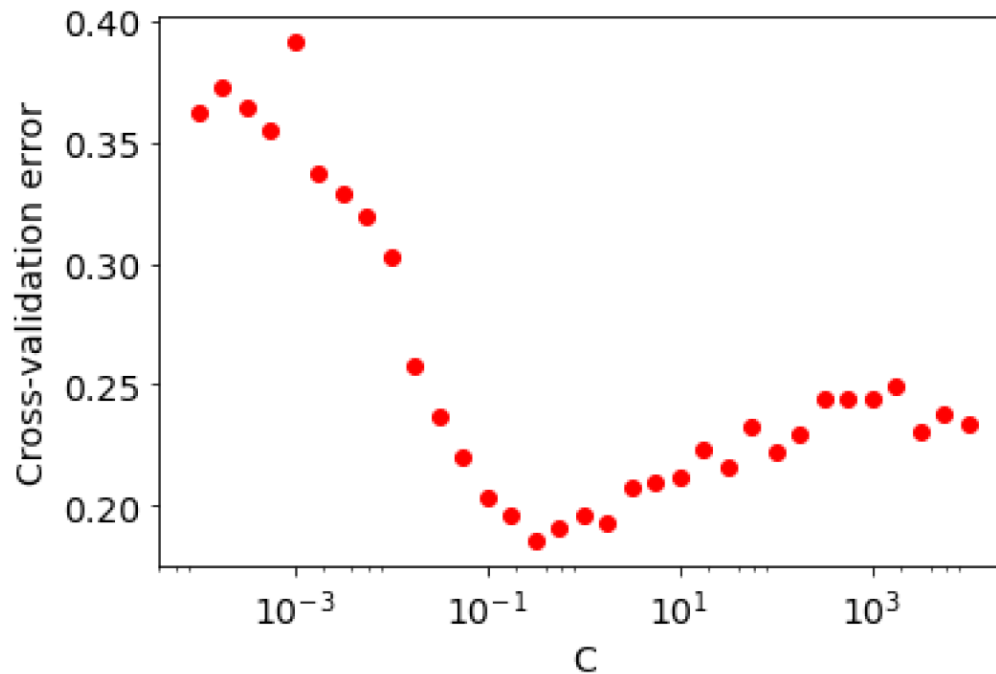
Cross-Validation

Results of 5-fold cross-validation:



Cross-Validation

Results of 5-fold cross-validation:



Chose $C = 0.32$. Test error: 15.6%