STRENGTH AND WEAKNESS

SYRACUSE UNIVERSITY
School of Information Studies

EXTEND BINARY CLASSIFICATION TO MULTICLASS

Given n classes, e.g.

Sentiment = {positive, negative, neutral, no opinion}

One-versus-one (pairwise) strategy:

Create n(n – 1)/2 classifiers: pos|neg, pos|neu, pos|np, neg|neu, neg|np, neu|np

Pick the most confident prediction.

One-versus-all strategy:

Create n classifiers: positive or not, negative or not, neutral or not, np or not

Pick the most confident prediction.

SVMS' STRENGTH

High tolerance to noisy data

Flexibility in data representation: Well suited for continuous- or discrete-valued inputs and outputs

Probabilistic prediction result

Scalability: Successful on extremely large problems

Successful on a wide array of real-world data

PROBABILISTIC OUTPUT OF SVMS

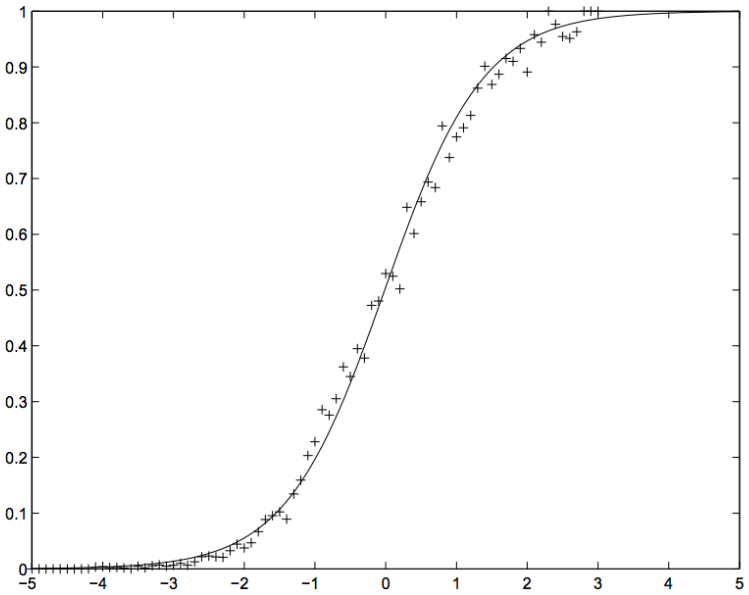

Figure 2: The fit of the sigmoid to the data for a linear SVM on the Adult data set (as in Figure 1). Each plus mark is the posterior probability computed for all examples falling into a bin of width 0.1. The solid line is the best-fit sigmoid to the posterior, using the algorithm described in this chapter.

Platt, J. (1999). Probabilistic outputs for support vector machines and comparisons to regularized likelihood methods. *Advances in Large Margin Classifiers, 10*(3), 61– 74.

SVMS' WEAKNESS

Require a number of parameters for each kernel type

Interpretability

- Easy interpretation for linear kernel
- Difficult to interpret the model generated by nonlinear kernels