

Suppose you are trying to decide among a few different choices of kernel and are also choosing parameters such as  $C$ ,  $\sigma^2$ , etc. How should you make the choice?

Choose whatever performs best on the training data.

Choose whatever performs best on the cross-validation data.

Correct

Choose whatever performs best on the test data.

Choose whatever gives the largest SVM margin.