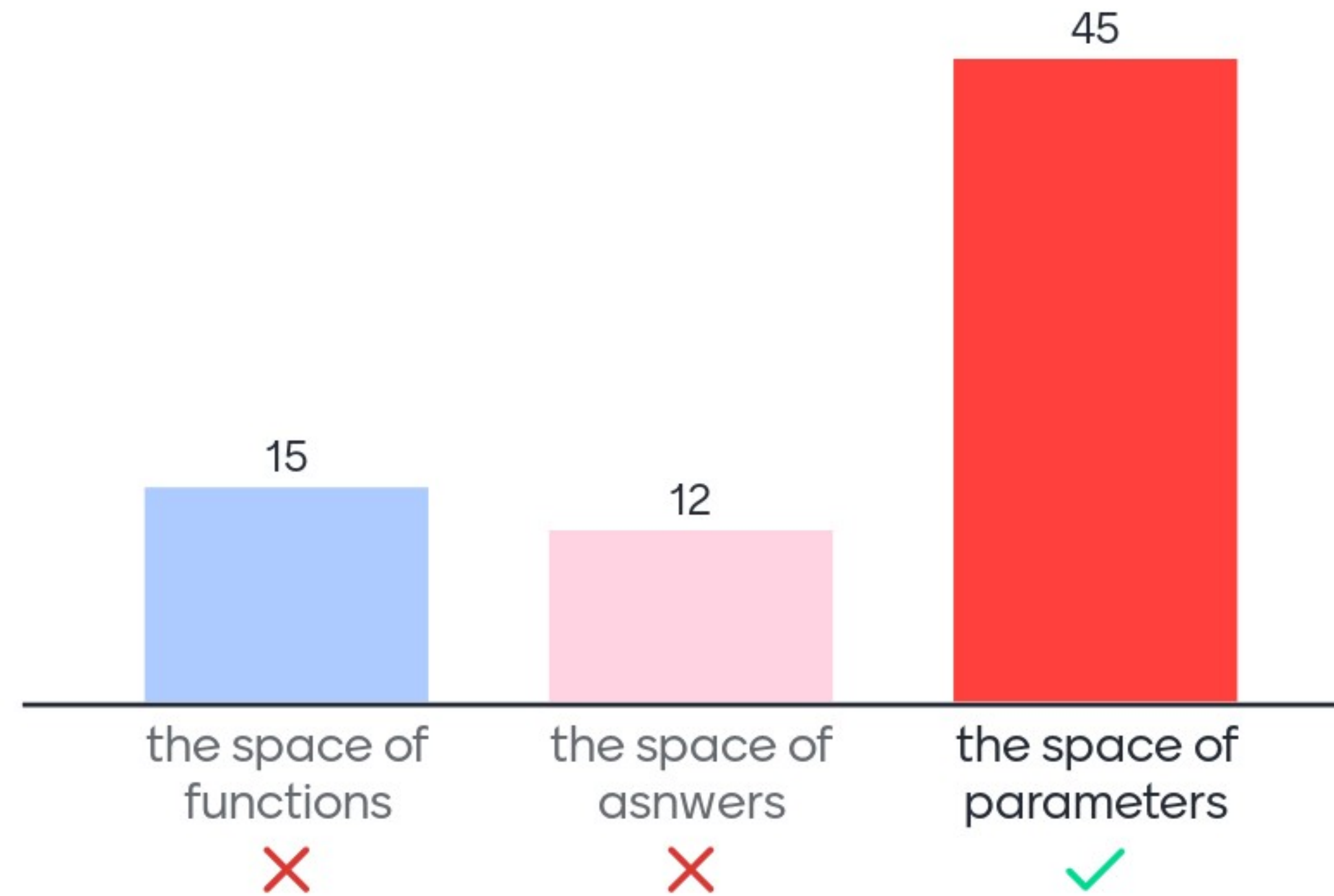


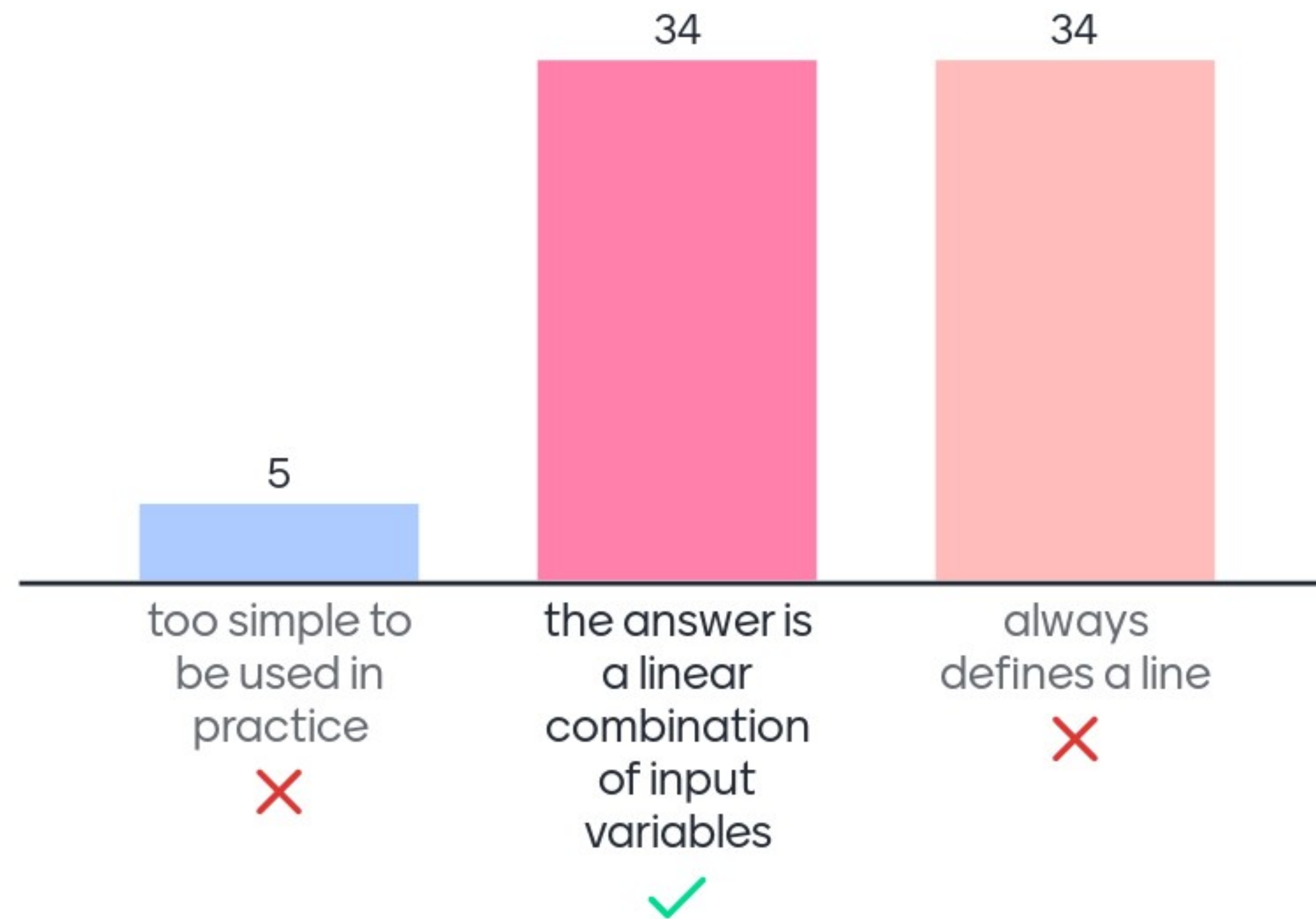
IN3050/IN4050 *week 7*

Classification and Logistic Regression

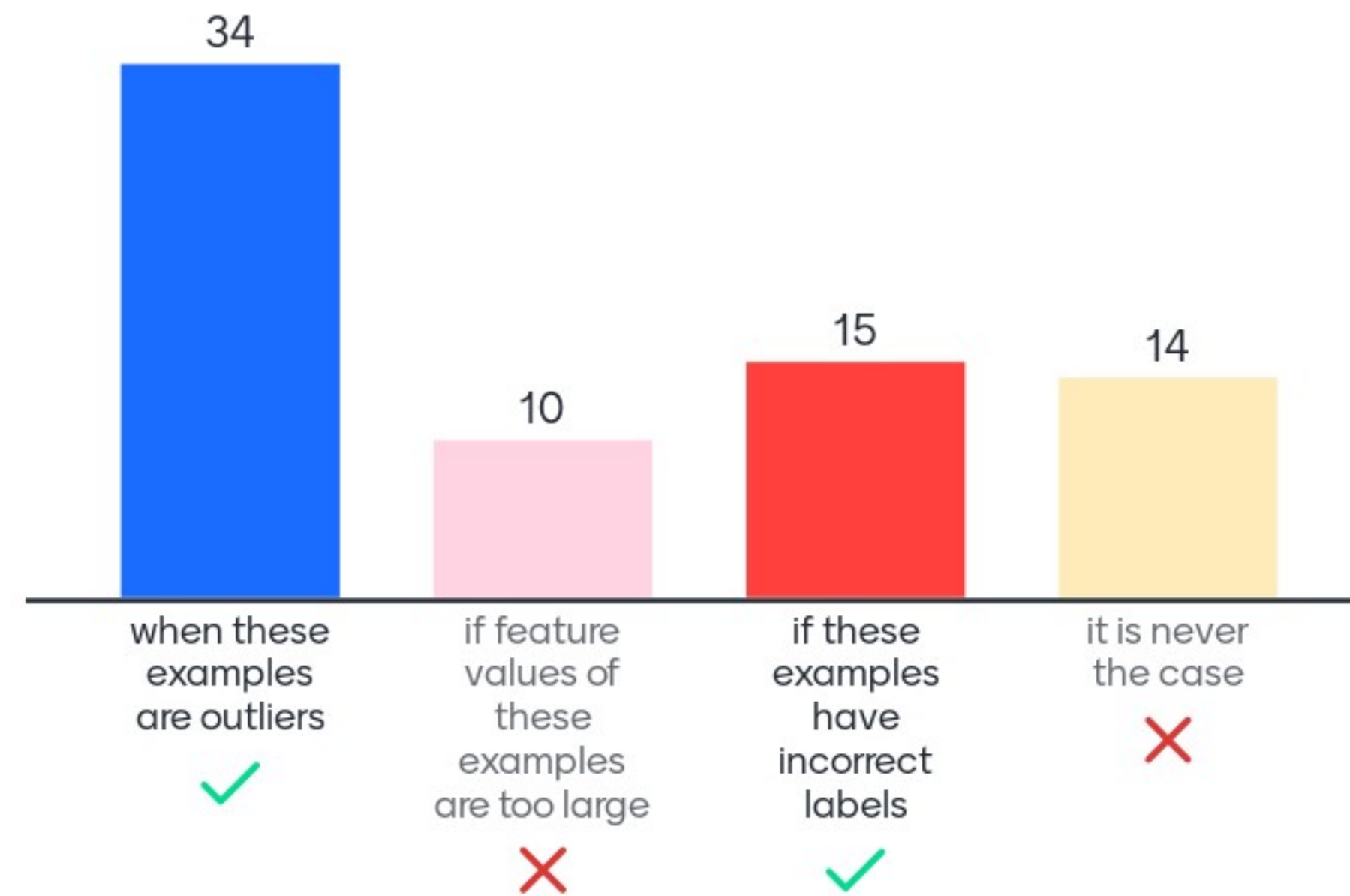
When training a supervised classifier, we are searching over:



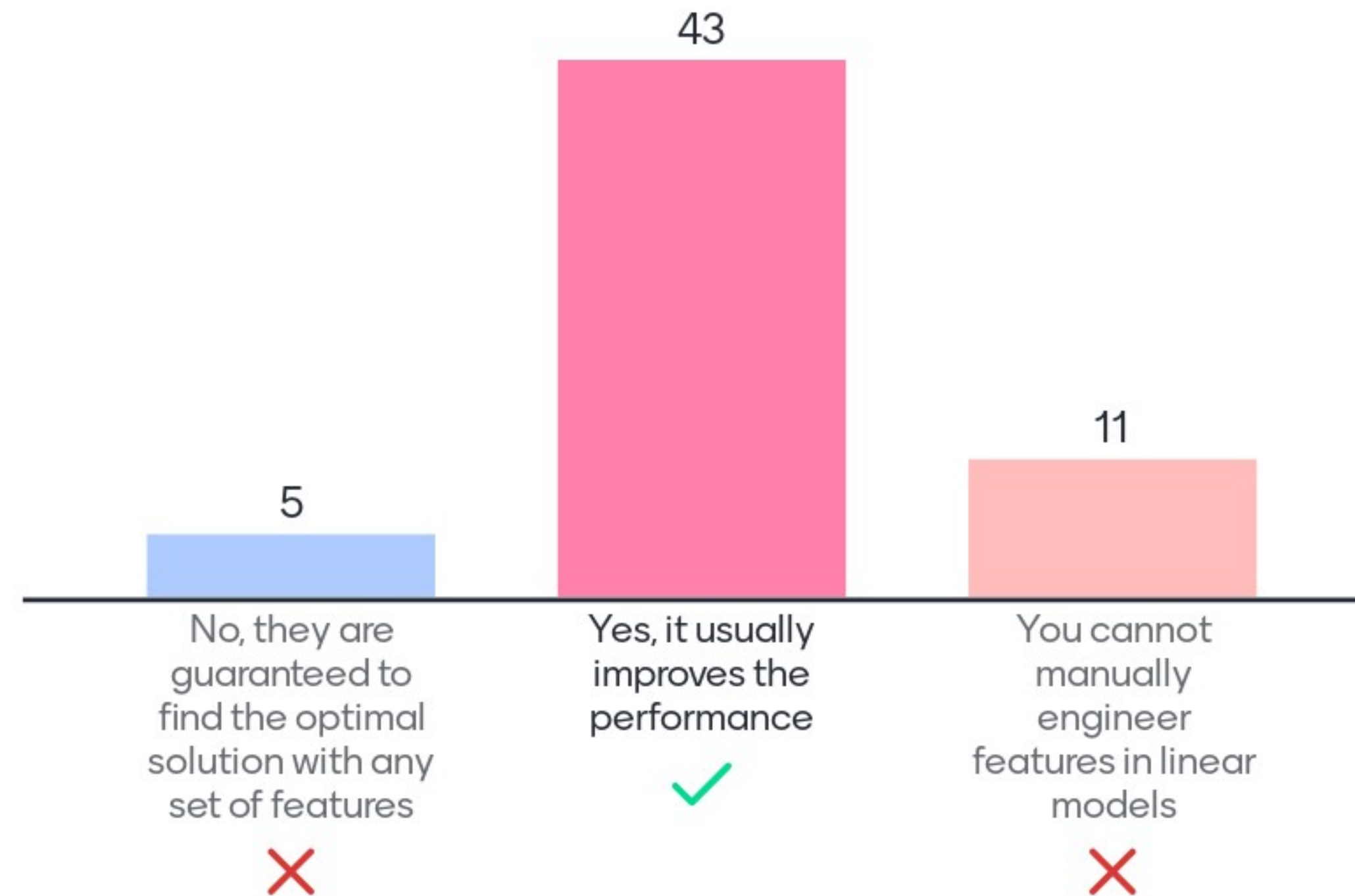
What does it mean for a function to be *linear*?



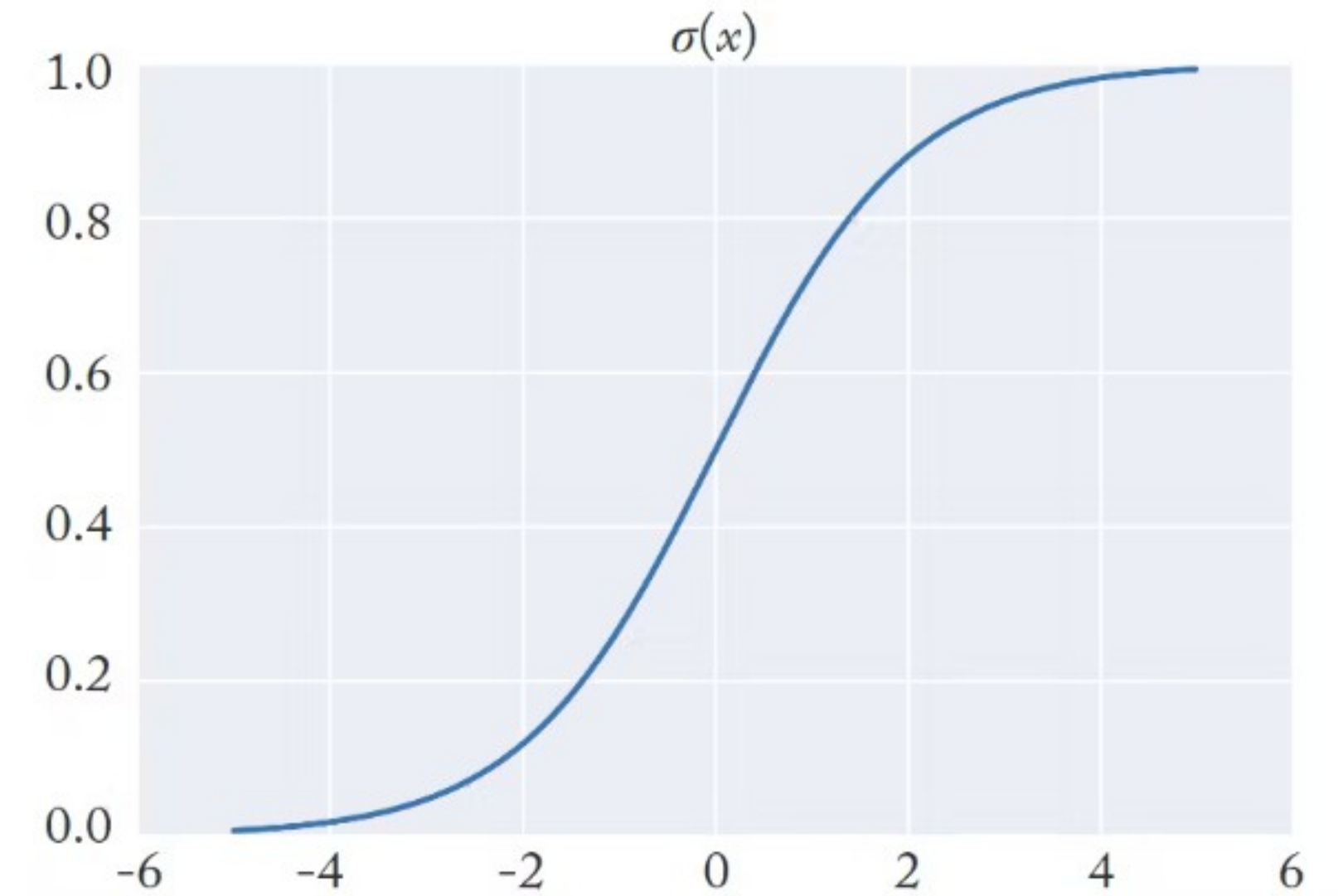
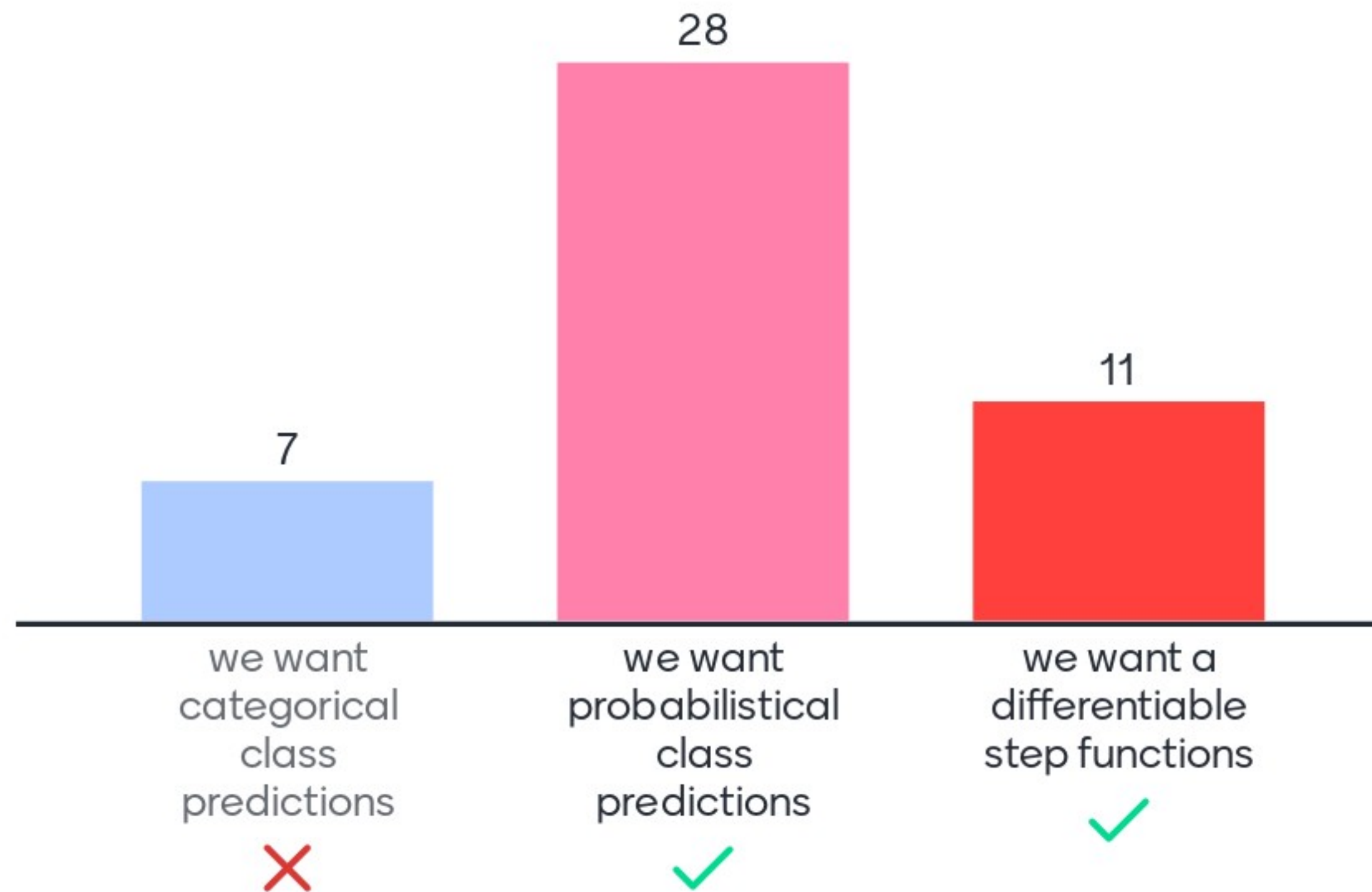
In what cases it is beneficial that the trained model classifies incorrectly some of the examples in the train set?



Do we need to carefully chooses (feature engineering) for linear models?



What is the purpose of applying a squashing function (like sigmoid) to the output of our classifier?



YOU TRAINED A 4-CLASS LOGISTIC CLASSIFIER.

For an example with the correct class of "2", it predicted $[0.2, 0.7, 0.05, 0.05]$. Compute categorical cross-entropy loss (times 10).

