Solutions to exercises: week 1

Exercise 1.1
(c) The quadratic, linear and Fisher discriminant are not affected, the nearest mean and the k-NN are. It is because in the first classifiers a covariance matrix $\hat{\Sigma}$ is estimated. Therefore automatically the scale of the feature values is estimated in these classifiers.
(d) In many cases it is an advantage, because it means that you don't need to optimize the scaling of the features. The disadvantage is that if you know that one of the features is more informative, it is not clear how you can bias your classifier to use that feature more .