1. Absolute residuals should have a half-normal distribution if $X_i \sim N(\mu, \sigma^2)$. Thus, if

$$r_i = |X_i - median(\{X_{i=1}^n\})|$$

then, to derive our constant, if

$$\frac{1}{5} = \mathrm{P}(r_i < q)$$

then

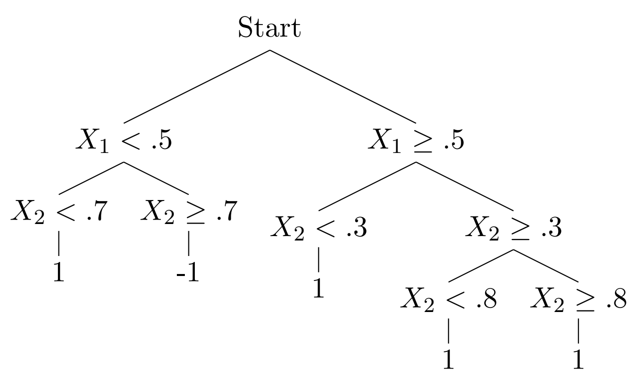
$$\begin{aligned} \frac{6}{10} &= \Phi\left(\frac{q}{\sigma}\right) \\ \Phi^{-1}\left(\frac{6}{10}\right) &= \frac{q}{\sigma} \\ \sigma &= \frac{q}{\Phi^{-1}\left(\frac{6}{10}\right)} \end{aligned}$$

So our constant C should be

$$\frac{1}{\Phi^{-1}\left(\frac{6}{10}\right)} = 3.947$$

This, however, is a worse estimator than MAD, since an estimator for the first quintile of the absolute residuals of a normal distribution will have higher variance than a comparable estimator of the median.

3. (a)



(b) This model incorrectly identifies three points, giving it a training error of 23%.

(c)

$$\hat{f}(.6, .2) = 1, \quad \hat{f}(.3, .9) = -1$$