1 Q1

1.1 1 (7 points)

+7 if numbers are right

1.2 2 (23 points)

+7 plot looks correct (circles tend towards the boundary)

+5 say that thre error is a random variable

+6 errors of AC are reasonable, and train ¿ test

+6 mention that AC points tend to be chosen around boundary

+3 mention that low-density points help little (and/or that reducing the grid bounds would help)

+3 mention other thoughts that appear in the solutions

2 Q2

2.1 1 (25 points)

+7 use the hint i.e. write

$$E_{p(x,y)}\left[L(y,\hat{y}(x))\right] = \ldots = \int_x p(x) \sum_y \left(1 - \delta_{y,\hat{y}(x)}\right) p(y|x) dx \qquad (1)$$

(+ 3 if start right but then get confused)

+8 show (convincingly) it is enough to minimize cost for every x

+4 only say it is enough to minimize for every x (not convincing they understand why it's enough)

+10 derive thre rule from (1)

2.2 2 (10 points)

+6 correct rule (in general form)

+4 explain how to derive

+2 show simplification for binary case

do not deduct points fro same mistakes as in 1!

3 Q3

3.1 1 (10 points)

+2 right numbers

+2 right plots

+4 say that as $n \to \infty$ boundar approaches the optimal boundary

+2 only say that boundary becomes linear (without saying that optimal boundary is linear)

+2 explain it is difficult to estimate Gaussians with few points

-2 say the boundary for `XYlarge` **is** linear

3.2 2 (25 points)

+2 right numbers

+2 right plots

+3 say that the two models have similar generalization error when trained on `XYlarge`.

+5 explain disatvantage of generative models with few points (diificult to estimate Gaussians)

+15 explain that boundaries approach each other, and the optimal boundary (only +7 if claims are correct but not explained)

+7 only say boundaries both approach a linear boundary (but do not say that's the optimal boundary)

-5 say one of the methods (or both) can not produce linear boundary (or have different number of degrees of freedom)

-3 say log. reg. is much worse because the boundary looks worse