Georgia Institute of Technology

ECE 6254 - Assignment 3

Fall 2020 - v1.0

- There are 2 problems over 2 pages (including the cover page)
- The problems are not necessarily in order of difficulty.
- All problems are assigned the same weight in the overall grade.
- Each question is graded out of two points (0 for no meaningful work, 1 for partial work, 2 if correct)
- Unless otherwise specified, you should concisely indicate your reasoning and show all relevant work.
- The grade on each problem is based on our judgment of your level of understanding as reflected by what you have written. If we can't read it, we can't grade it.
- Please use a pen and not a pencil if you handwrite your solution.
- **You must submit your exam on Gradescope. Make sure you allocate time for the submission**.

Problem 1: Synthetic Bayes' classifier

Consider a binary classification problem involving a single (scalar) feature x and suppose that $X|Y=0$ and $X|Y=1$ are continuous random variables with densities given by

$$p_{X|Y}(x|0) = \begin{cases} e^{-x} \text{ if } x > 0 \\ 0 \text{ else,} \end{cases} \qquad p_{X|Y}(x|1) = \begin{cases} 1 \text{ if } x \in [a; a+1] \\ 0 \text{ else,} \end{cases}$$

for some $a \geqslant 0$. Further assume that $\mathbb{P}(Y=0) = \mathbb{P}(Y=1) = \frac{1}{2}$ and that we are using the 0-1 loss function.

[Q1] Sketch $p_{X|Y}(x|1)$ and $p_{X|Y}(x|0)$.

[Q2] Derive the expression of the optimal Bayes estimator and simplify your expression as much as possible.

[Q3] Calculate the Bayes risk as a function of the parameter a and simplify your answer as much as possible.

[Q4] Suppose now that we do not know the true value of parameter a but have an estimate $\hat{a}$ obtained at training. Determine the risk for the classifier that results from replacing a with $\hat{a}$ in the classification rule you derived earlier in part **[Q2]**, i.e., assume that the data is still generated according to the distributions $p_{X|Y=0}$ and $p_{X|Y=1}$ but we use the classification rule based on $\hat{a}$. (Your answer should be in terms of a and the estimate $\hat{a}$, and you may assume that $\hat{a} \geqslant 0$.)

[Q5] Given $\hat{a}$, we now treat a as a random variable and suppose that we are given a guarantee that $\mathbb{P}(|\hat{a} - a| \geqslant \epsilon) \leqslant \delta$ for some $\epsilon \geqslant 0$ and $\delta \in (0,1)$. Give an upper bound on the worst-case possible risk $\mathbb{P}\left(h^{\hat{a}}(x) \neq y\right)$ in terms of *only* the quantities we know (i.e., $\widehat{a}$, ϵ, and δ, but not a.) (*Hint:* Break up the analysis of $\mathbb{P}\left(h^{\hat{a}}(x) \neq y\right)$ into two cases depending on $|a - \hat{a}|$ using the law of total probability and calculate the worst-case risk under either scenario.)

Problem 2: Bayes for Gaussian distributions

Consider the regular 0/1 loss and assume that data is generated according to $P(y=0) = P(y=1) = 1/2$. Also assume that the class-conditional densities are Gaussian with mean μ_0 and covariance Σ_0 under class 0, and mean μ_1 and covariance Σ_1 under class 1. For the following two cases, you are asked to i) Sketch contours of the level sets of the class conditional densities and label them with $p(x|y=0)$ and $p(x|y=1)$ with $x \in \mathbb{R}^2$; ii) Draw the decision boundaries obtained using the Bayes optimal classifier in each case and indicate the regions where the classifier will predict class 0 and where it will predict class 1. *You must justify your answer to explain your sketch in enough details to claim credit. It is possible to do some quantitative analysis and be very precise.*

[Q1] **Case 1:** $\mu_0 = -\mu_1$ and

$$\Sigma_0 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \quad \Sigma_1 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

[Q2] **Case 2:** $\mu_0 = \mu_1$ and

$$\Sigma_0 = \begin{bmatrix} 1 & 0 \\ 0 & 4 \end{bmatrix} \quad \Sigma_1 = \begin{bmatrix} 4 & 0 \\ 0 & 1 \end{bmatrix}$$