CSE 151A - Discussion 02

Quick Review

Univariate Bayes Classifier : predict class y_i that maximizes $P(Y = y_i|X = x)$ for some input x

1. Estimate $P(Y = y_i|X = x)$ directly with Nearest Neighbor classifier
2. Estimate $P(Y = y_i|X = x)$ indirectly using Bayes Rule
   (a) Estimate $P(X = x|Y = y_i)$ using histograms (non-parametric)
   (b) Estimate $P(X = x|Y = y_i)$ using Gaussians (parametric, μ and σ)

Note : $\text{argmax}_{y_i} P(Y = y_i|X = x)$ is equivalent to $\text{argmax}_{y_i} P(X = x|Y = y_i)P(Y = y_i)$

Problem 1.

Suppose you find yourself in the middle of Oklahoma. You are suddenly approached by a local sheriff who needs help solving a case involving an animal that has been reported missing.
The only evidence procured so far is from a nearby animal park run by a man named Joe.
It reads as follows:

Animal Type	Number in park	μ weight (kg)	σ weight (kg)
Tiger	157	220	10
Lion	83	190	8
Bear	41	210	13

The sheriff tells you that the missing animal weighs exactly **212 kg**.

Using a Bayes Classifier, determine what is most likely to be the type of animal that is missing.
[HINT : recall the formula of the univariate Gaussian $f(x; \mu, \sigma) = \frac{1}{\sqrt{(2\pi\sigma^2)}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$]

Solution: Tiger
We will compute $\text{argmax}_{y_i} P(X = x|Y = y_i)P(Y = y_i)$ by estimation using Gaussians.

Total animals = 157 + 83 + 41 = 281

Priors: $P(Y = \text{Tiger}) = \frac{157}{281} = 0.5587$ $P(Y = \text{Lion}) = \frac{83}{281} = 0.2954$ $P(Y = \text{Bear}) = \frac{41}{281} = 0.1459$

Conditionals:

$$P(X = 212\text{kg}|Y = \text{Tiger}) = \frac{1}{\sqrt{2\pi(10)^2}} e^{-\frac{(212-220)^2}{2(10)^2}} = 0.0290$$

$$P(X = 212\text{kg}|Y = \text{Lion}) = \frac{1}{\sqrt{2\pi(8)^2}} e^{-\frac{(212-190)^2}{2(8)^2}} = 0.0011$$

$$P(X = 212\text{kg}|Y = \text{Bear}) = \frac{1}{\sqrt{2\pi(13)^2}} e^{-\frac{(212-210)^2}{2(13)^2}} = 0.0303$$

$P(X = 212\text{kg}|Y = \text{Tiger}) * P(Y = \text{Tiger}) = (0.0290)(0.5587) = \mathbf{0.0162} \leftarrow$ max value
$P(X = 212\text{kg}|Y = \text{Lion}) * P(Y = \text{Lion}) = (0.0011)(0.2954) = 0.00032$
$P(X = 212\text{kg}|Y = \text{Bear}) * P(Y = \text{Bear}) = (0.0303)(0.1459) = 0.0044$

We predict **Tiger** because the maximum value of $P(Y = y_i|X = 212\text{kg})$ was achieved when $y_i = \text{Tiger}$.

Problem 2.

You then overhear the sheriff speaking on the phone to a very upset woman named Carole. She has tracked one of Joe's trailers and is confident that he is currently on the road transporting as many as 130 tigers to various mall parking lots across the country.
This means that there could only be 27 tigers at the animal park capable of escaping.

Does this new information change what type of animal is most likely to be missing?
Explain your answer. (Assume that the parameters of each Gaussian remain the same.)

Solution: Bear
We will use the same approach as we did in Problem 1.

Total animals = 27 + 83 + 41 = 151

Priors: $P(Y = \text{Tiger}) = \frac{27}{151} = 0.1788 \quad P(Y = \text{Lion}) = \frac{83}{151} = 0.5497 \quad P(Y = \text{Bear}) = \frac{41}{151} = 0.2715$

Conditionals: Same as Problem 1 because of assumption.

$P(X = 212\text{kg}|Y = \text{Tiger}) * P(Y = \text{Tiger}) = (0.0290)(0.1788) = 0.0052$
$P(X = 212\text{kg}|Y = \text{Lion}) * P(Y = \text{Lion}) = (0.0011)(0.5497) = 0.0006$
$P(X = 212\text{kg}|Y = \text{Bear}) * P(Y = \text{Bear}) = (0.0303)(0.2715) = \mathbf{0.0082} \leftarrow$ max value

We predict **Bear** because the maximum value of $P(Y = y_i|X = 212\text{kg})$ was achieved when $y_i = \text{Bear}$.

Problem 3.

Consider the following multivariate Gaussians that are parameterized by some μ and C (covariance matrix). For each, sketch the approximate contour shape (from a top-down view) and classify it as either (i) Spherical, (ii) Diagonal, or (iii) General. Also comment on the name of the contour shape and any assumptions that were made about the data.

a) $\mu = \begin{bmatrix} -1 \\ -1 \end{bmatrix}$ $C = \begin{bmatrix} 1 & 0 \\ 0 & 4 \end{bmatrix}$

b) $\mu = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$ $C = \begin{bmatrix} 1 & -1.5 \\ -1.5 & 4 \end{bmatrix}$

c) $\mu = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$ $C = \begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix}$

d) $\mu = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$ $C = \begin{bmatrix} 4 & 1.5 \\ 1.5 & 1 \end{bmatrix}$

Solution:
a. Diagonal. Axis-aligned ellipse. Data points are independent.
b. General. Non axis-aligned ellipse. Data points are not independent.
c. Spherical. Circular. Data points are independent and the variance is the same in both dimensions.
d. General. Non axis-aligned ellipse. Data points are not independent.
*See follow-up post on Campuswire or view the discussion video for sketches.