Bayesian Learning

20CP401T

Himanshu K. Gajera
Department of Computer Science & Engineering
Pandit Deendayal Energy University, Gandhinagar

Feature Space Representation

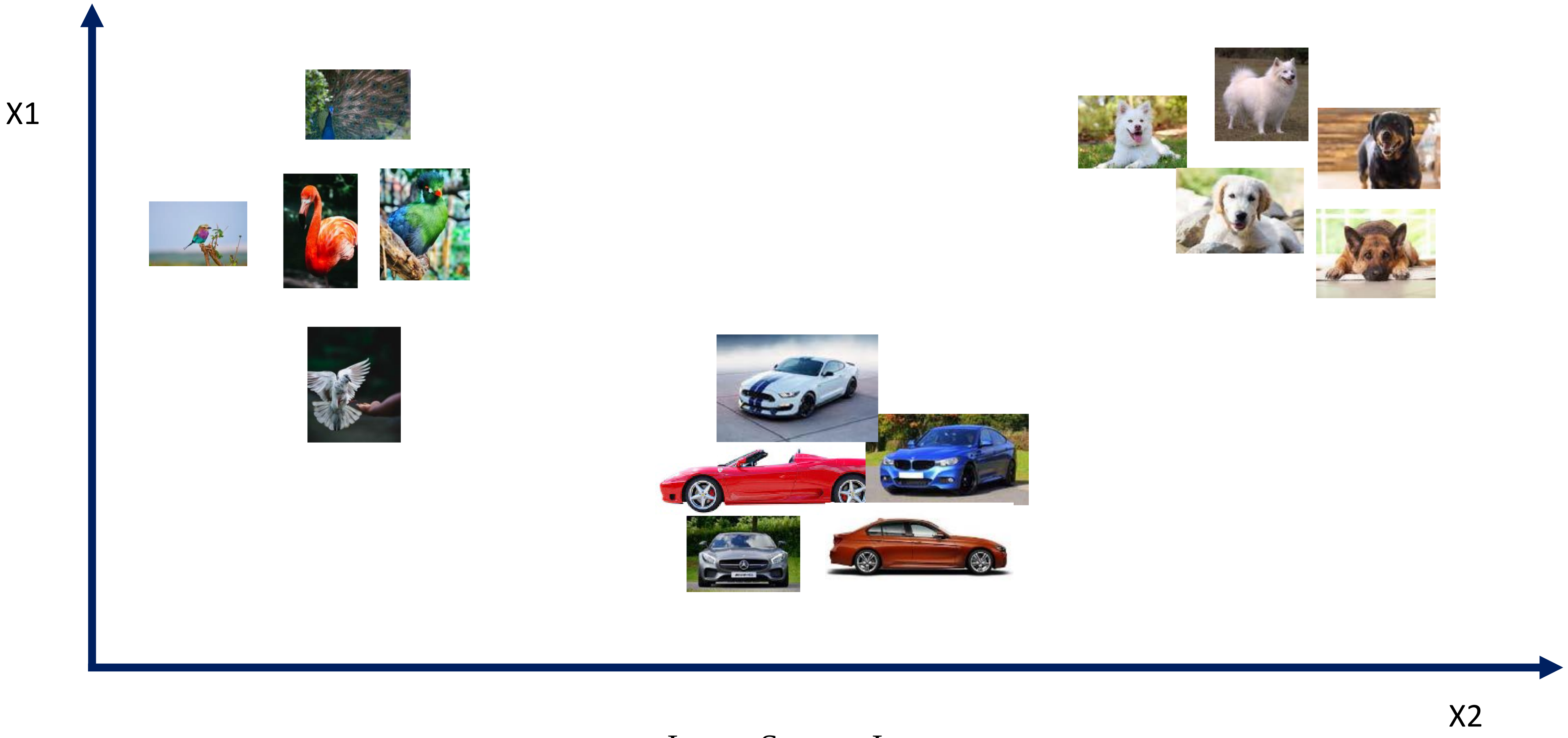
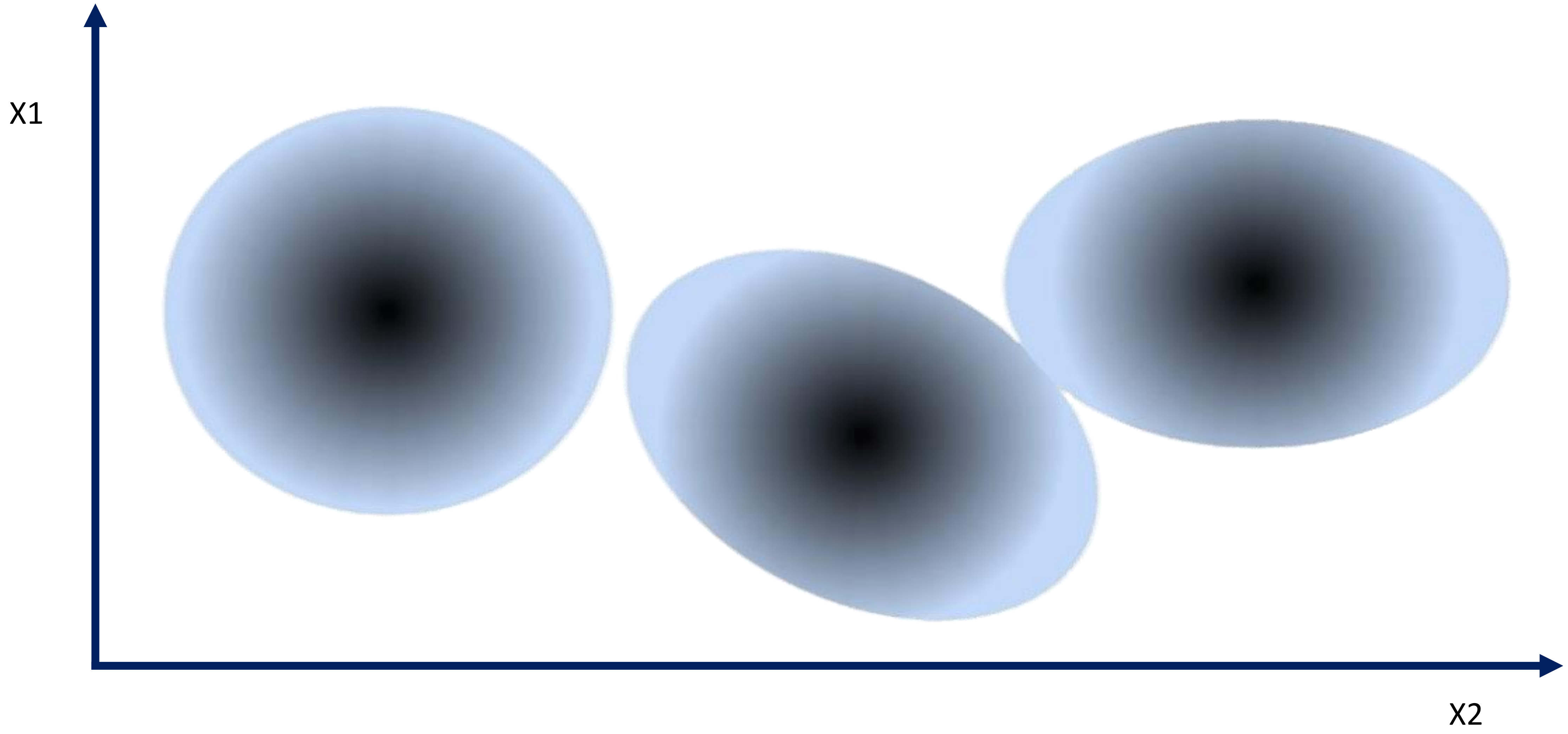
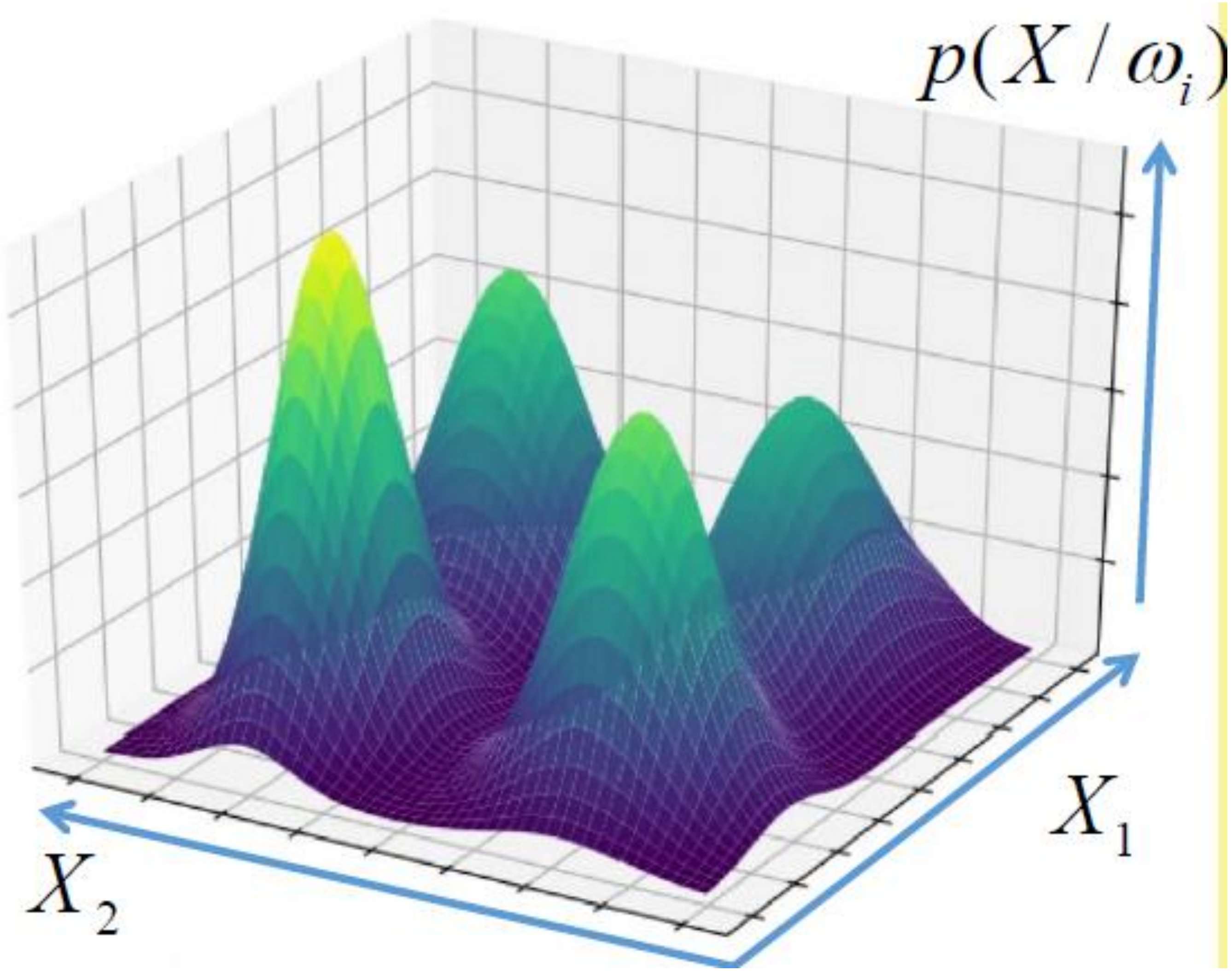


Image Source: Internet

Feature Space Representation



Feature Space Representation



Feature Space Representation

- Design classifiers to recommend decisions that minimize some total expected ”risk”.

  - The simplest risk is the classification error (i.e., costs are equal).

  - Typically, the risk includes the cost associated with different decisions.

Terminology

- State of nature ω ***(random variable)***:
  - e.g., $\omega 1$ for car, $\omega 2$ for bird

- Probabilities $P(\omega 1)$ and $P(\omega 2)$ ***(priors)***:
  - e.g., prior knowledge of how likely is to get a bird or a car

- Probability density function $p(x)$ ***(evidence)***:
  - e.g., how frequently we will measure a pattern with feature value x (e.g., x corresponds to lightness)

Terminology

- Conditional probability density $p(x/\omega j)$ (**likelihood**) :
  - e.g., how frequently we will measure a pattern with feature value x given that the pattern belongs to class ωj

- Conditional probability $P(\omega j /x)$ (**posterior**) :
  - e.g., the probability that the fish belongs to class ωj given measurement x.

Terminology

Likelihood   Class Prior Probability

$$P(c \mid x) = \frac{P(x \mid c)\,P(c)}{P(x)}$$

Posterior Probability   Predictor Prior Probability

$$P(c \mid \mathrm{X}) = P(x_1 \mid c) \times P(x_2 \mid c) \times \cdots \times P(x_n \mid c) \times P(c)$$

Decision Rule Using Prior Probabilities

Decide $\omega 1$ if $P(\omega 1) > P(\omega 2);$ otherwise **decide** $\omega 2$

$$P(error) = \begin{cases} P(\omega_1) & if\ we\ decide\ \omega_2 \\ P(\omega_2) & if\ we\ decide\ \omega_1 \end{cases}$$

or $P(error) = min[P(\omega_1), P(\omega_2)]$

Favors the most likely class.

➢ This rule will be making the same decision all times.

i.e., optimum if no other information is available

Decision Rule Using Conditional Probabilities

Using Bayes' rule, the posterior probability of category ωj given measurement x is given by:

$$P(\omega_j / x) = \frac{p(x/\omega_j)P(\omega_j)}{p(x)} = \frac{likelihood \times prior}{evidence}$$

where $p(x) = \sum_{j=1}^{2} p(x/\omega_j)P(\omega_j)$ (i.e., scale factor – sum of probs = 1)

Decide ω_1 if $P(\omega_1 /x) > P(\omega_2/x)$; otherwise **decide** ω_2

or

Decide ω_1 if $p(x/\omega_1)P(\omega_1)>p(x/\omega_2)P(\omega_2)$ otherwise **decide** ω_2