Q&A

In our classroom, there are only handful amount of female student as compared to male students, do you need to get the same amount of training samples for both male and female students?
Bayes formula (Bayes rule)

\[ p(j | x) = \frac{p(x | j) p(j)}{p(x)} \]

\[ p(x) = \sum_j p(x | j) p(j) \]

\( a - \text{posteriori probability} \) (posterior probability)

\( a - \text{priori probability} \) (prior probability)

Conditional probability density (pdf – probability density function) (likelihood)

From domain knowledge

Normalization constant (evidence)

Bayes decision rule

\[ p(j | x) = \frac{p(x | j) p(j)}{p(x)} \]

\[ p(x | j) p(j) \]

\[ p(x | j) p(j) \]

Maximum \( a - \text{posteriori probability} \) (MAP)

For a given \( x \), if \( p(j_1 | x) > p(j_2 | x) \)

then \( x \) belongs to class \( j_1 \), otherwise, \( j_2 \).
Conditional probability of error

\[ P(\omega_1 | x) = \frac{p(x | \omega_1) P(\omega_1)}{p(x)} \]

\[ P(\omega_2 | x) = \frac{p(x | \omega_2) P(\omega_2)}{p(x)} \]

\[ P(error | x) = \begin{cases} P(\omega_1 | x) & \text{if we decide } \omega_1 \\ P(\omega_2 | x) & \text{if we decide } \omega_2 \end{cases} \]

\[ P(error) = \int_{-\infty}^{\infty} P(error | x) p(x) dx = \min[P(\omega_1 | x), P(\omega_2 | x)] \]

Decision regions

The effect of any decision rule is to partition the feature space into \( c \) decision regions \( \mathcal{R}_1, \mathcal{R}_2, \ldots, \mathcal{R}_c \).

Overall probability of error

\[ P(error) = \int_{\mathcal{R}_1} P(error | x) p(x) dx + \int_{\mathcal{R}_2} P(error | x) p(x) dx \]

\[ P(error) = \int_{\mathcal{R}_1} P(\omega_1 | x) p(x) dx + \int_{\mathcal{R}_2} P(\omega_2 | x) p(x) dx \]

\[ P(error) = \int_{\mathcal{R}_1} P(\omega_1 | x) p(x) dx + \int_{\mathcal{R}_2} P(\omega_2 | x) p(x) dx \]
The conditional risk

Given \( x \), the conditional risk of taking action \( \alpha_i \) is:

\[
R(\alpha_i | x) = \sum \lambda(\alpha_i | \omega_j) P(\omega_j | x)
\]

\( \lambda_{ij} \) is the loss when decide \( x \) belongs to class \( i \) while it should be \( j \).

Likelihood ratio - two category classification

\[
R(\alpha_i | x) = \lambda_{i1} P(\omega_1 | x) + \lambda_{i2} P(\omega_2 | x)
\]

\[
R(\alpha_2 | x) = \lambda_{21} P(\omega_1 | x) + \lambda_{22} P(\omega_2 | x)
\]

If \( R(\alpha_1 | x) < R(\alpha_2 | x) \), then decide \( \omega_2 \)

\[
(\lambda_{21} - \lambda_{11}) P(\omega_1 | x) > (\lambda_{22} - \lambda_{12}) P(\omega_2 | x)
\]

Zero-One loss

\[
\lambda(\alpha_i | \omega_j) = \begin{cases} 0 & i = j \\ 1 & i \neq j \\ \end{cases}
\]

\[
R(\alpha_i | x) = \sum P(\omega_j | x) = 1 - P(\omega_i | x)
\]

\[
p(x | \omega_1) > \frac{\lambda_{22} - \lambda_{12}}{\lambda_{21} - \lambda_{11}} P(\omega_1)
\]

\[
p(x | \omega_2) = \frac{\lambda_{22} - \lambda_{12}}{\lambda_{21} - \lambda_{11}} P(\omega_2)
\]

\[
p(x | \omega_2) = \frac{P(\omega_2)}{P(\omega_1)}
\]
Recap

- Bayes decision rule \(\rightarrow\) maximum a-posteriori probability
- Conditional risk \(\rightarrow\) likelihood ratio
- Decision regions \(\rightarrow\) How to calculate the overall probability of error

For a given \(x\), if \(P(\omega_1 | x) > P(\omega_2 | x)\) then \(x\) belongs to class 1, otherwise, 2.

\[
P(\omega_j | x) = \frac{p(x | \omega_j) P(\omega_j)}{p(x)}
\]

\[
R(\omega_j | x) = \sum_{j=1}^{J} \lambda(\omega_j | \omega_j) P(\omega_j | x)
\]

Overall probability of error

\[
P(\text{error}) = \int_{\Omega_1} P(\omega_1 | x)p(x)dx + \int_{\Omega_2} P(\omega_2 | x)p(x)dx
\]