



# ANALÝZA A KLASIFIKACE DAT



prof. Ing. Jiří Holčík, CSc.



INVESTICE DO ROZVOJE VZDĚLÁVÁNÍ

# III. BAYESŮV KLASIFIKÁTOR

# ZÁKLADNÍ POJMY A PŘEDPOKLADY

- ✓ při řešení praktických úloh je třeba předpokládat, že obrazy signálů jsou ovlivněny víceméně náhodnými fluktuacemi zdroje signálu, v přenosové cestě, při předzpracování i analýze, které se nepodaří zcela eliminovat.
- ✓ ztrátová funkce  $\lambda(\omega_r|\omega_s)$  udává ztrátu při chybné klasifikaci obrazu ze třídy  $\omega_s$  do třídy  $\omega_r$ .
- ✓ matice ztrátových funkcí

$$\mathbf{\lambda} = \begin{bmatrix} \lambda(\omega_1|\omega_1) & \lambda(\omega_1|\omega_2) & \cdots & \lambda(\omega_1|\omega_R) \\ \lambda(\omega_2|\omega_1) & \lambda(\omega_2|\omega_2) & \cdots & \lambda(\omega_2|\omega_R) \\ \vdots & \vdots & \ddots & \vdots \\ \lambda(\omega_R|\omega_1) & \lambda(\omega_R|\omega_2) & \cdots & \lambda(\omega_R|\omega_R) \end{bmatrix}$$

- ✓ střední ztráta  $J(\mathbf{a})$  udává průměrnou ztrátu při chybné klasifikaci obrazu  $\mathbf{x}$

# KRITÉRIUM MINIMÁLNÍ STŘEDNÍ ZTRÁTY

- ☑ pokud se soustředíme na obrazy pouze ze třídy  $\omega_s$ , je střední ztráta dána průměrnou hodnotou z  $\lambda(d(\mathbf{x}, \mathbf{a})|\omega_s)$  vzhledem ke všem obrazům ze třídy  $\omega_s$ , tj.

$$J_s(\mathbf{a}) = \int \lambda(d(\mathbf{x}, \mathbf{a})|\omega_s) \cdot p(\mathbf{x}|\omega_s) d\mathbf{x}$$

kde  $p(\mathbf{x}|\omega_s)$  je podmíněná hustota  
pravděpodobnosti výskytu obrazu  $\mathbf{x}$  ve třídě  $\omega_s$

# KRITÉRIUM MINIMÁLNÍ STŘEDNÍ ZTRÁTY

- ☑ Celková střední ztráta  $J(\mathbf{a})$  je průměrná hodnota ze ztrát  $J_s(\mathbf{a})$

$$J(\mathbf{a}) = \sum_{s=1}^R J_s(\mathbf{a}) \cdot P(\omega_s) = \int \sum_{s=1}^R \lambda(d(\mathbf{x}, \mathbf{a}) | \omega_s) \cdot p(\mathbf{x} | \omega_s) \cdot P(\omega_s) d\mathbf{x}$$

- ☑ nebo podle Bayesova vzorce (  $P(\omega_s | \mathbf{x}) \cdot p(\mathbf{x}) = p(\mathbf{x} | \omega_s) \cdot P(\omega_s)$  )

$$J(\mathbf{a}) = \int \sum_{s=1}^R \lambda(d(\mathbf{x}, \mathbf{a}) | \omega_s) \cdot p(\mathbf{x}) \cdot P(\omega_s | \mathbf{x}) d\mathbf{x}$$

kde  $p(\mathbf{x})$  je hustota pravděpodobnosti výskytu obrazu  $\mathbf{x}$  v celém obrazovém prostoru a  $P(\omega_s | \mathbf{x})$  je podmíněná pravděpodobnost, že daný obraz patří do třídy  $\omega_s$  (tzv. aposteriorní pravděpodobnost třídy  $\omega_s$ ).

# KRITÉRIUM MINIMÁLNÍ STŘEDNÍ ZTRÁTY

- ☑ Návrh optimálního klasifikátoru, který by minimalizoval střední ztrátu, spočívá v nalezení takové množiny parametrů rozhodovacího pravidla  $\mathbf{a}^*$ , že platí

$$J(\mathbf{a}^*) = \min_{\mathbf{a}} J(\mathbf{a})$$

- ☑ Dosadíme-li za  $J(\mathbf{a})$  z předchozího vztahu, je

$$J(\mathbf{a}^*) = \min_{\mathbf{a}} \int_{\mathcal{X}} \sum_{s=1}^R \lambda(d(\mathbf{x}, \mathbf{a}) | \omega_s) \cdot p(\mathbf{x} | \omega_s) \cdot P(\omega_s) d\mathbf{x}$$

- ☑ Je-li ztrátová funkce  $\lambda(\omega_r | \omega_s)$  konstantní pro všechny obrazy z  $\omega_s$ , je dále

$$J(\mathbf{a}^*) = \int_{\mathcal{X}} \min_r \sum_{s=1}^R \lambda(\omega_r | \omega_s) \cdot p(\mathbf{x} | \omega_s) \cdot P(\omega_s) d\mathbf{x}$$

# KRITÉRIUM MINIMÁLNÍ STŘEDNÍ ZTRÁTY

- ☑ Označíme-li ztrátu při klasifikaci obrazu  $\mathbf{x}$  do třídy  $\omega_r$

$$L_x(\omega_r) = \sum_{s=1}^R \lambda(\omega_r | \omega_s) \cdot p(\mathbf{x} | \omega_s) \cdot P(\omega_s)$$

tak po dosazení dostaneme

$$J(\mathbf{a}^*) = \int_x \min_r L_x(\omega_r) d\mathbf{x}$$

Úloha nalezení minima celkové střední ztráty se tak převedla na minimalizaci funkce  $L_x(\omega_r)$ . Optimální rozhodovací pravidlo  $d(\mathbf{x}, \mathbf{a}^*)$  podle kritéria minimální celkové střední ztráty je

$$L_x(d_{ME}(\mathbf{x}, \mathbf{a}^*)) = \min_r L_x(\omega_r)$$

# KRITÉRIUM MINIMÁLNÍ STŘEDNÍ ZTRÁTY

- ☑ Chceme-li využít principu diskriminačních funkcí

$$\min L_x(\omega_r) = \max(-L_x(\omega_r))$$

- ☑ Diskriminační funkci optimálního klasifikátoru podle kritéria minimální chyby pak definujeme

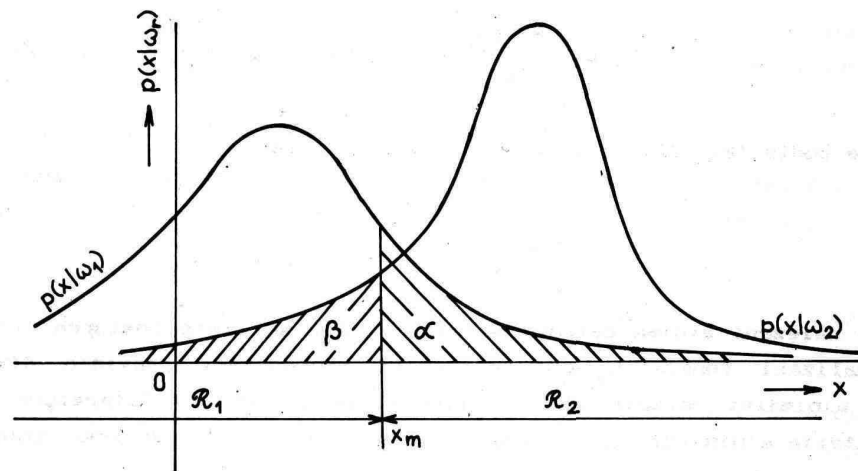
$$g_r(\mathbf{x}) = -L_x(\omega_r) = -\sum_{s=1}^R \lambda(\omega_r | \omega_s) \cdot p(\mathbf{x} | \omega_s) \cdot P(\omega_s)$$



# KRITÉRIUM MINIMÁLNÍ STŘEDNÍ ZTRÁTY DICHOTOMICKÝ KLASIFIKÁTOR

Celková střední ztráta v případě dvou tříd je

$$\begin{aligned} J(\mathbf{a}) &= \int_{\mathcal{R}_1} \sum_{s=1}^2 \lambda(\omega_1|\omega_s) \cdot p(\mathbf{x}|\omega_s) \cdot P(\omega_s) d\mathbf{x} + \int_{\mathcal{R}_2} \sum_{s=1}^2 \lambda(\omega_2|\omega_s) \cdot p(\mathbf{x}|\omega_s) \cdot P(\omega_s) d\mathbf{x} = \\ &= \lambda(\omega_1|\omega_1) \cdot P(\omega_1) \int_{\mathcal{R}_1} p(\mathbf{x}|\omega_1) \cdot d\mathbf{x} + \lambda(\omega_1|\omega_2) \cdot P(\omega_2) \int_{\mathcal{R}_1} p(\mathbf{x}|\omega_2) \cdot d\mathbf{x} + \\ &+ \lambda(\omega_2|\omega_1) \cdot P(\omega_1) \int_{\mathcal{R}_2} p(\mathbf{x}|\omega_1) \cdot d\mathbf{x} + \lambda(\omega_2|\omega_2) \cdot P(\omega_2) \int_{\mathcal{R}_2} p(\mathbf{x}|\omega_2) \cdot d\mathbf{x} = \\ &= \lambda(\omega_1|\omega_1) \cdot P(\omega_1) \cdot (1 - \alpha) + \lambda(\omega_1|\omega_2) \cdot P(\omega_2) \cdot \beta + \lambda(\omega_2|\omega_1) \cdot P(\omega_1) \cdot \alpha + \lambda(\omega_2|\omega_2) \cdot P(\omega_2) \cdot (1 - \beta) \end{aligned}$$



# KRITÉRIUM MINIMÁLNÍ STŘEDNÍ ZTRÁTY DICHOTOMICKÝ KLASIFIKÁTOR

Diskriminační funkce pro dichotomický klasifikátor bude

$$\begin{aligned}g(\mathbf{x}) &= g_1(\mathbf{x}) - g_2(\mathbf{x}) = -L_{\mathbf{x}}(\omega_1) + L_{\mathbf{x}}(\omega_2) = \\&= -\lambda(\omega_1|\omega_1) \cdot p(\mathbf{x}|\omega_1) \cdot P(\omega_1) - \lambda(\omega_1|\omega_2) \cdot p(\mathbf{x}|\omega_2) \cdot P(\omega_2) + \\&\quad + \lambda(\omega_2|\omega_1) \cdot p(\mathbf{x}|\omega_1) \cdot P(\omega_1) + \lambda(\omega_2|\omega_2) \cdot p(\mathbf{x}|\omega_2) \cdot P(\omega_2) = \\&= (\lambda(\omega_2|\omega_1) - \lambda(\omega_1|\omega_1)) \cdot p(\mathbf{x}|\omega_1) \cdot P(\omega_1) + (\lambda(\omega_2|\omega_2) - \lambda(\omega_1|\omega_2)) \cdot p(\mathbf{x}|\omega_2) \cdot P(\omega_2)\end{aligned}$$

Položíme-li tento výraz nule dostaneme vztah pro hraniční plochu dichotomického klasifikátoru, ze kterého můžeme určit poměr hustot pravděpodobnosti výskytu obrazu  $\mathbf{x}$  v každé z obou klasifikačních tříd - **věrohodnostní poměr**

$$\Lambda_{12} = \frac{p(\mathbf{x}|\omega_1)}{p(\mathbf{x}|\omega_2)} = \frac{(\lambda(\omega_1|\omega_2) - \lambda(\omega_2|\omega_2)) \cdot P(\omega_2)}{(\lambda(\omega_2|\omega_1) - \lambda(\omega_1|\omega_1)) \cdot P(\omega_1)}$$

Obraz  $\mathbf{x}$  zařadíme do třídy  $\omega_1$ , když je věrohodnostní poměr větší než výraz na pravé straně, je-li menší pak obraz  $\mathbf{x}$  zařadíme do třídy  $\omega_2$ .

# VĚROHODNOSTNÍ POMĚR I.

- ☑ Sumarizuje veškerou informaci získanou experimentem.
- ☑ Pravděpodobnost, že jev (data) nastane za daných podmínek (hypotéza) děleno pravděpodobností, že stejný jev nastane za jiných podmínek. Podmínky jsou vzájemně se vylučující.

# VĚROHODNOSTNÍ POMĚR II.

*Věrohodnostní poměr (likelihood ratio) LR* udává podíl pravděpodobnosti, že se vyskytne nějaký jev  $A$  za určité podmínky (jev  $B$ ), k pravděpodobnosti, že se jev  $A$  vyskytne, když podmínka neplatí (jev non $B$ ). Má-li například pacient náhlou ztrátu paměti (jev  $A$ ), chceme znát věrohodnostní poměr výskytu jevu  $A$  v případě, že má mozkový nádor (jev  $B$ ), tj. podíl pravděpodobnosti, s jakou ztráta paměti vzniká při nádoru mozku, k pravděpodobnosti, s jakou vzniká v ostatních případech. Věrohodnostní poměr je tedy podíl podmíněných pravděpodobností

$$LR = \frac{P(A|B)}{P(A|nonB)}$$

# KRITÉRIUM MINIMÁLNÍ PRAVDĚPODOBNOTI CHYBNÉHO ROZHODNUTÍ

Díky obtížnému stanovení hodnot ztrátových funkcí  $\lambda(\omega_r|\omega_s)$  se kritérium minimální chyby zjednodušuje použitím jednotkových ztrátových funkcí definovaných

$$\lambda(\omega_r|\omega_s) = \begin{cases} 0 & \text{pro } r = s \\ 1 & \text{pro } r \neq s \end{cases}$$

Matrice jednotkových ztrátových funkcí má pak tvar

$$\lambda = \begin{bmatrix} 0 & 1 & \dots & 1 \\ 1 & 0 & \dots & 1 \\ \vdots & \vdots & \ddots & \vdots \\ 1 & 1 & \dots & 0 \end{bmatrix}$$

a celková ztráta je

$$J(\mathbf{a}) = \sum_{\substack{s=1 \\ s \neq r}}^R \int_{X-\mathcal{R}_s} p(\mathbf{x}|\omega_s) \cdot P(\omega_s) d\mathbf{x}$$

což je hodnota pravděpodobnosti chybného rozhodnutí.

# KRITÉRIUM MINIMÁLNÍ PRAVDĚPODOBNOTI CHYBNÉHO ROZHODNUTÍ

Dosadíme-li hodnoty jednotkových ztrátových funkcí do vztahu pro ztrátu při klasifikaci obrazu do chybné třídy

$$L_{\mathbf{x}}(\omega_r) = \sum_{\substack{s=1 \\ s \neq r}}^R p(\mathbf{x}|\omega_r) \cdot P(\omega_s) = \sum_{s=1}^R p(\mathbf{x}|\omega_s) \cdot P(\omega_s) - p(\mathbf{x}|\omega_r) \cdot P(\omega_r)$$

a s využitím Bayesova vztahu

$$L_{\mathbf{x}}(\omega_r) = p(\mathbf{x}) \sum_{s=1}^R P(\omega_s|\mathbf{x}) - p(\mathbf{x}|\omega_r) \cdot P(\omega_r) = p(\mathbf{x}) - p(\mathbf{x}|\omega_r) \cdot P(\omega_r)$$

$p(\mathbf{x})$  nezávisí na klasifikační třídě a tedy neovlivňuje výběr minima.

Diskriminační funkci tedy můžeme určit jako

$$g(\mathbf{x}) = p(\mathbf{x}|\omega_r) \cdot P(\omega_r)$$

# KRITÉRIUM MINIMÁLNÍ PRAVDĚPODOBNOTI CHYBNÉHO ROZHODNUTÍ

V případě dichotomického klasifikátoru je diskriminační funkce

$$g(\mathbf{x}) = p(\mathbf{x}|\omega_1) \cdot P(\omega_1) - p(\mathbf{x}|\omega_2) \cdot P(\omega_2)$$

A věrohodnostní poměr je potom

$$\Lambda_{12} = \frac{p(\mathbf{x}|\omega_1)}{p(\mathbf{x}|\omega_2)} = \frac{P(\omega_2)}{P(\omega_1)}$$

# KRITÉRIUM MAXIMÁLNÍ APOSTERIORNÍ PRAVDĚPODOBNOSTI

- ✓ Modifikujeme-li vztah pro ztrátu při chybné klasifikaci obrazu podle Bayesova vztahu (  $P(\omega_s|\mathbf{x}).p(\mathbf{x}) = p(\mathbf{x}|\omega_s).P(\omega_s)$  ) platí

$$L_x(\omega_r) = \sum_{s=1}^R \lambda(\omega_r|\omega_s).p(\mathbf{x}).P(\omega_s|\mathbf{x}) = p(\mathbf{x}) \sum_{s=1}^R \lambda(\omega_r|\omega_s).P(\omega_s|\mathbf{x})$$

- ✓ Hustota pravděpodobnosti  $p(\mathbf{x})$  nezávisí na klasifikační třídě a tedy místo  $L_x(\omega_r)$  lze použít

$$L'_x(\omega_r) = \frac{L_x(\omega_r)}{p(\mathbf{x})} = \sum_{s=1}^R \lambda(\omega_r|\omega_s).P(\omega_s|\mathbf{x})$$

a s jednotkovými ztrátovými funkcemi je

$$L'_x(\omega_r) = \sum_{\substack{s=1 \\ s \neq r}}^R P(\omega_s|\mathbf{x}) = \sum_{s=1}^R P(\omega_s|\mathbf{x}) - P(\omega_r|\mathbf{x}) = 1 - P(\omega_r|\mathbf{x})$$



# KRITÉRIUM MAXIMÁLNÍ APOSTERIORNÍ PRAVDĚPODOBNOSTI

- ☑ Minimum ztráty  $L'_x(\omega_r)$  je právě tehdy, když  $P(\omega_r|\mathbf{x})$  je maximální. Tzn. že jako diskriminační funkci můžeme zvolit právě hodnotu aposteriori pravděpodobnosti třídy  $\omega_r$ , tj.

$$g_r(\mathbf{x}) = P(\omega_r|\mathbf{x})$$

- ☑ Pro případ dichotomického klasifikátoru je diskriminační funkce

$$g(\mathbf{x}) = P(\omega_1|\mathbf{x}) - P(\omega_2|\mathbf{x}) = 0.$$

Z toho plyne, že hranicí mezi třídami určuje vztah

$$P(\omega_1|\mathbf{x}) = P(\omega_2|\mathbf{x})$$

nebo

$$\frac{P(\omega_1|\mathbf{x})}{P(\omega_2|\mathbf{x})} = 1$$

Podle tohoto kritéria zařídíme obraz do té třídy, jejíž aposteriori pravděpodobnost je při výskytu obrazu  $\mathbf{x}$  větší.

# KRITÉRIUM MAXIMÁLNÍ PRAVDĚPODOBNOTI (MINIMAX)

Neznáme-li apriorní pravděpodobnosti všech tříd, předpokládáme rovnoměrné rozložení (pravděpodobnost všech tříd je táž ( $P(\omega_s) = P(\omega) = 1/R$ ). Potom celková střední ztráta

$$J(\mathbf{a}) = \frac{1}{R} \sum_{s=1}^R \int_{\mathcal{X}} \lambda(\omega_r | \omega_s) \cdot p(\mathbf{x} | \omega_s) d\mathbf{x}$$

dosáhne minima, když

$$J(\mathbf{a}^*) = \frac{1}{R} \min_{\forall \mathbf{a}} \int_{\mathcal{X}} \sum_{s=1}^R \lambda(\omega_r | \omega_s) \cdot p(\mathbf{x} | \omega_s) d\mathbf{x}$$

Diskriminační funkci lze jako v předchozích případech definovat jako

$$g_r(\mathbf{x}) = -L_{\mathbf{x}}(\omega_r) = - \sum_{s=1}^R \lambda(\omega_r | \omega_s) \cdot p(\mathbf{x} | \omega_s)$$

# KRITÉRIUM MAXIMÁLNÍ PRAVDĚPODOBNOSTI (MINIMAX)

- ✓ V případě dichotomie je věrohodnostní poměr

$$\Lambda_{12} = \frac{p(\mathbf{x}|\omega_1)}{p(\mathbf{x}|\omega_2)} = \frac{(\lambda(\omega_1|\omega_2) - \lambda(\omega_2|\omega_2))}{(\lambda(\omega_2|\omega_1) - \lambda(\omega_1|\omega_1))}$$

- ✓ Pokud jsou ceny správného rozhodnutí nulové, tj.  $\lambda(\omega_1|\omega_1) = \lambda(\omega_2|\omega_2) = 0$ , je

$$\Lambda_{12} = \frac{p(\mathbf{x}|\omega_1)}{p(\mathbf{x}|\omega_2)} = \frac{(\lambda(\omega_1|\omega_2))}{(\lambda(\omega_2|\omega_1))}$$

- ✓ Obraz je zařazen do třídy  $\omega_1$ , když je věrohodnostní poměr než poměr cen ztrát chybných zatřídění. Jsou-li obě ceny stejné, je obraz zařazen do té třídy, pro kterou je hodnota  $p(\mathbf{x}|\omega_s)$  větší.

# KRITÉRIUM MAXIMÁLNÍ PRAVDĚPODOBNOTI (MINIMAX)

