PDV - 20.11.2026

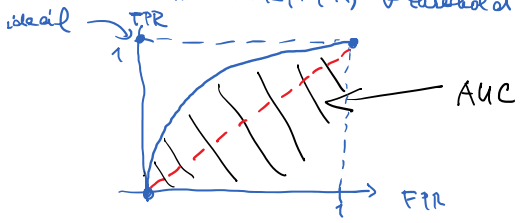
piatok 20. novembra 2026 10:59

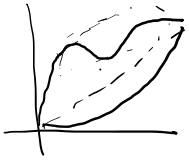
Minule: Logist. reg.: $y_i \sim Alt(\pi_i)$

$$\pi_i = \underbrace{\frac{e^{x_i^T\beta}}{1+e^{x_i^T\beta}}}_{g(x_i^T\beta)}$$

Kvalita modelu: accuracy, ...

ROC krivka: (TPR, FPR) $\forall$ threshold





PC cvič.: 27.11., 4.12.

- 24.11.: skončíme okolo 18:00
- 1.12.: skončíme okolo 18:00

- DÚ
- Píš ?

jedno číslo: AUC (Area under curve): čím vyššie, tým lepšie
→ kvalita cez všetky thresholdy

- $y = x$: klasifikátor, čo nič nevie ($\hat{\pi}$ je rovnomerne náhodné bez ohľadu na dáta)
  → AUC = 1/2
- AUC $\in [0, 1]$, ideálne aspoň $[1/2, 1]$
- toto celé: pre ľub. klasif. model

Viactriedová klasifikácia

- $\underline{C}$ rôznych tried
- namiesto $P(y_i = 1) = \pi_i$ vs. $P(y_i = 0) = 1 - \pi_i$

teraz: $P(y_i = 1) = \pi_{1,i}$
$\vdots$
$P(y_i = C) = \pi_{C,i}$

$$x \Rightarrow \begin{matrix} \hat{\pi}_{1,i} \\ \vdots \\ \hat{\pi}_{C,i} \end{matrix} \Rightarrow \hat{y}_i: \text{trieda s najv. pp.} \in \operatorname{argmax}_j \hat{\pi}_{j,i}$$

(napr. 50%, 30%, 20%)
↳ $\hat{y}_i = 1$

- napr. Multinomická reg.

Neurónové siete

⟨Strojové uč., AZKD, Neurónové siete, Moderné techniky strojového učenia⟩

(1.) Predpis modelu: závisí od parametrov θ
: $x \longmapsto \hat{y} = f_\theta(x)$

- lin. reg.: $\hat{y} = \beta_0 + \beta_1 x_1 + \dots + \beta_{p-1} x_{p-1} = f_\theta(x)$ (pre $\theta = \beta$)

(2.) $\hat{\theta}$: dáta $x_1, \dots, x_n \in \mathbb{R}^p$
$y_1, \dots, y_n \in \{1, \dots, C\}$
$\in \mathbb{R}$
$\in \{0, 1\}$
...

: minimalizujeme chybu, napr.

$$\sum_{i=1}^{n} (y_i - f_\theta(x_i))^2 \to \min_\theta \Rightarrow \hat{\theta}$$

lin. reg.: $\sum_{i=1}^{n} (y_i - (\beta_0 + \dots + \beta_{p-1} x_{i,p-1}))^2 \to \min_\beta$

Neur. siete: to isté, akurát (1.) je veľmi komplikovaný

logist. reg.: $x \to \beta_0 + \underbrace{\beta_1 x_1 + \dots + \beta_p x_p}_{x^T\beta} \to \pi = g(x^T\beta + \beta_0)$ $\quad g(t) = \frac{e^t}{1+e^t} = \frac{1}{1+e^{-t}}$

pre viac tried: $x \nearrow \theta_{0,1} + \theta_{1,1}x_1 + \theta_{2,1}x_2 + \dots + \theta_{p,1}x_p \to \pi_1 = g(\theta_{0,1} + x^T\theta_1)$ $\quad \theta_j = \begin{pmatrix} \theta_{1,j} \\ \vdots \end{pmatrix}$

pre viac tried :

$$x \to \begin{cases} \theta_{0,1} + \theta_{1,1}x_1 + \theta_{2,1}x_2 + \dots + \theta_{p,1}x_p \to \pi_1 = g(\theta_{0,1} + x^T\theta_1) \\ \theta_{0,2} + \theta_{1,2}x_1 + \theta_{2,2}x_2 + \dots + \theta_{p,2}x_p \to \pi_2 = g(\theta_{0,2} + x^T\theta_2) \\ \vdots \\ \theta_{0,c} + \theta_{1,c}x_1 + \theta_{2,c}x_2 + \dots + \theta_{p,c}x_p \to \pi_c = g(\theta_{0,c} + x^T\theta_c) \end{cases} \qquad \theta_j = \begin{pmatrix} \theta_{1,j} \\ \vdots \\ \theta_{p,j} \end{pmatrix}$$

parametre : $\theta_{0,1}, \dots, \theta_{0,c} \in \mathbb{R}$, $\theta_1, \dots, \theta_c \in \mathbb{R}^p$

- π_j by mali byť pp. : $\pi_j \geq 0 \leftarrow \checkmark$ $\quad \pi_j = g(\dots) \geq 0$

  $\sum_{j=1}^{c} \pi_j = 1 \leftarrow \times$: nenasčítajú sa

  $\Rightarrow$ zachováme $e^{(\,)}$, ale predelíme súčtom všetkých :

$$\pi = \begin{pmatrix} \dfrac{e^{z_1}}{\sum_{j=1}^{c} e^{z_j}} \\ \vdots \\ \dfrac{e^{z_c}}{\sum_{j=1}^{c} e^{z_j}} \end{pmatrix} \begin{matrix} \geq 0 \\ \\ \geq 0 \end{matrix} \quad , \sum_{j=1}^{c} \pi_j = 1$$

softmax

- ozn. : $\theta_{0,1} + x^T\theta_1 = z_1$

  $\vdots$

  $\theta_{0,c} + x^T\theta_c = z_c$

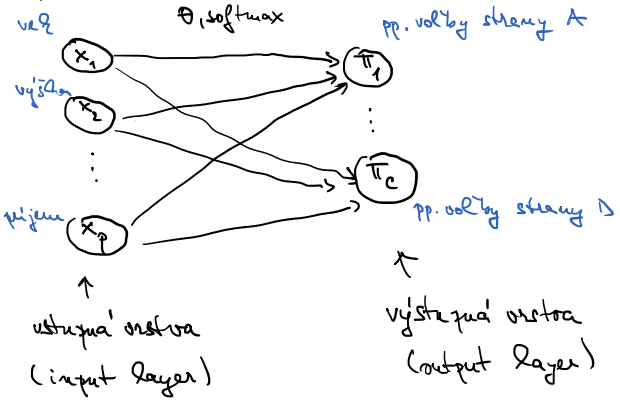
$\Rightarrow \pi = \text{softmax}(z)$

$z = (1, 10)$ $\quad \Rightarrow \text{softmax}(z) = \left(\frac{2{,}6}{22\,028{,}6}, \frac{22\,026}{22\,028{,}6}\right) \doteq (0, 1)$

$e^z = (e^1, e^{10}) = (2{,}6 ; 22\,026)$ $\quad \Rightarrow$ "skoro" maximum : softmax

$\Rightarrow$ zvýrazňuje rozdiely, výsledná predikcia bude jednoznačnejšia

zápis tohto :



: $z_1 = \theta_{0,1} + x^T\theta_1$

$\vdots$

$z_c = \theta_{0,c} + x^T\theta_c$

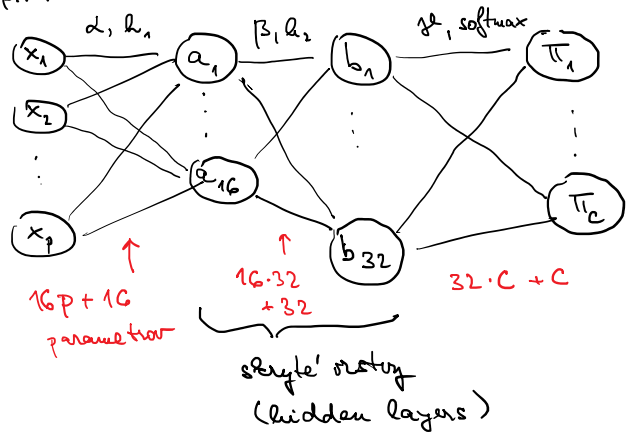
bias

$$, \begin{pmatrix} \pi_1 \\ \vdots \\ \pi_c \end{pmatrix} = \text{softmax}\begin{pmatrix} z_1 \\ \vdots \\ z_c \end{pmatrix}$$

θ : váhy, weights

zložitejšie : viac vrstiev : hlboké učenie (deep learning) :

Napr. :



$a_1 = h_1(\alpha_{0,1} + x^T\alpha_1)$ $\qquad \alpha_i \in \mathbb{R}^p$

$\vdots$

$a_{16} = h_1(\alpha_{0,16} + x^T\alpha_{16})$

$b_1 = h_2(\beta_{0,1} + a^T\cdot\beta_1)$ $\qquad \beta_i \in \mathbb{R}^{16}$

$\vdots$

$b_{32} = h_2(\beta_{0,32} + a^T\cdot\beta_{32})$

$z_1 = \gamma_{0,1} + b^T\gamma_1$

$\vdots$

$z_c = \gamma_{0,c} + b^T\gamma_c$

$$, \begin{pmatrix} \pi_1 \\ \vdots \\ \pi_c \end{pmatrix} = \text{softmax}\begin{pmatrix} z_1 \\ \vdots \\ z_c \end{pmatrix}$$

h_1, h_2 : aktivačné fcie (activation functions)

$\bigcirc$: neuróny (uzly)

c.h. : 16 · 16 + 16·32 + 32 + 32C + C parametrov

spolu: $16p + 16 + 16 \cdot 3C + 32 + 3C$ = parametrov

$= 560 + 16p + 33C$ parametrov

príklady aktivačných fcií: $h(t) = \frac{e^t}{1+e^t}$: logistická, sigmoid



$h(t) = \max(0,t)$: ReLU

Rectified Linear Unit



→ štandardne: nelineárne

- toto celé: (1.) $x \mapsto f_\theta(x)$

(2.) parametre = ?

maximalizácia $\ell(\theta) = \sum_{i=1}^{n} \left[\sum_{j=1}^{C} y_{ij} \ln(\pi_j(x_i)) \right] \longrightarrow \max_\theta$

j-ta výstupná pp. pre i-ty objekt (pre $x_i \in \mathbb{R}^p$)

$y_i = \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} y_{i1} \\ y_{i2} \\ \vdots \\ y_{i6} \end{pmatrix}$ ..., $y_{i3} = 1$, ostatné $y_{ij} = 0$

volí stranu 3

$(C=6)$

log-likelihood multinomického rozd.,

cross-entropy