

CHYBOVÉ FUNKCE

Autor textu:
Ing. Petr Honzík, Ph.D.

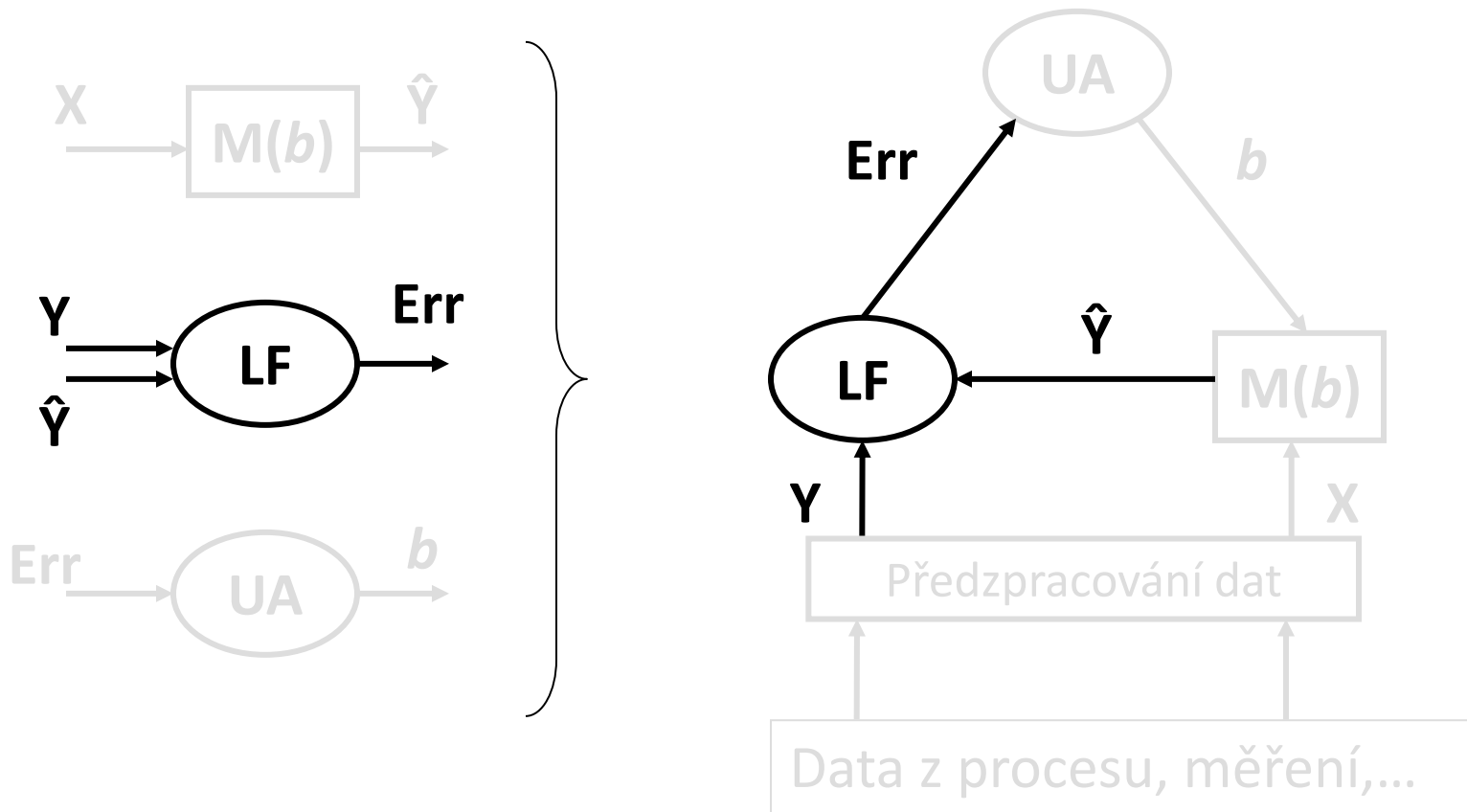
Květen 2014

Komplexní inovace studijních programů a zvyšování kvality výuky na FEKT VUT v Brně
OP VK CZ.1.07/2.2.00/28.0193



INVESTICE DO ROZVOJE VZDĚLÁVÁNÍ

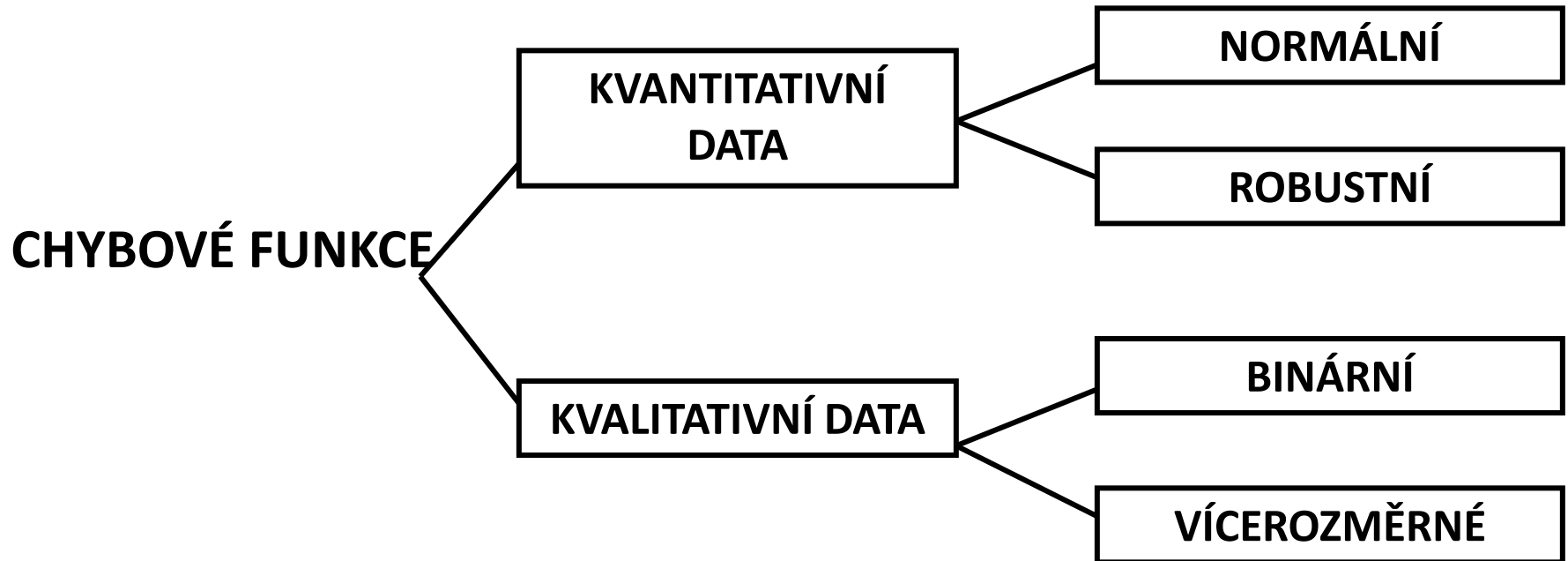
Chybová funkce v učení s učitelem



Chybová funkce (LF) – definice

- $(x, y, f(x)) \in X \times Y \times Y$ je uspořádaná trojice, kde
 - x je vstupní hodnota
 - y je výstupní hodnota
 - $f(x)$ je predikovaná výstupní veličina
- funkce $LF: X \times Y \times Y \rightarrow \langle 0; \infty \rangle$, pro kterou platí, že pro $\forall x \in X$ a $\forall y \in Y$ je $LF(x, y, y) = 0$, je označována jako **chybová funkce**
- chybová funkce (*Loss Function*) \Rightarrow **LF**

Chybové funkce – dělení



LF – kvantitativní

- Nejmenší čtverec vzdálenosti (MNČ)

$$LF = [y_i - f(x_i, \mathbf{b})]^2, \quad Err = \frac{1}{N} \sum_1^N [y_i - f(x_i, \mathbf{b})]^2$$

- Váhová MNČ

$$LF = [y_i - f(x_i, \mathbf{b})]^2 \cdot f(x_i), \quad Err = \frac{1}{N} \sum_{i=1}^N [y_i - f(x_i, \mathbf{b})]^2 \cdot f(x_i)$$

$$LF = [y_i - f(x_i, \mathbf{b})]^2 \cdot \frac{1}{x_i^2}, \quad Err = \frac{1}{N} \sum_{i=1}^N [y_i - f(x_i, \mathbf{b})]^2 \cdot \frac{1}{x_i^2}$$

- Absolutní chyba

$$LF = |y_i - f(x_i, \mathbf{b})|, \quad Err = \frac{1}{N} \sum_1^N |y_i - f(x_i, \mathbf{b})|$$

- Polynomická chyba

$$LF = \frac{1}{d} |y_i - f(x_i, \mathbf{b})|^d$$

LF – kvantitativní – robustní

- absolutní chyba + pásmo necitlivosti ε

$$LF(x_i, y_i, f(x_i, \mathbf{b})) = \max(|y_i - f(x_i, \mathbf{b})| - \varepsilon, 0)$$

- Huberova robustní chyba

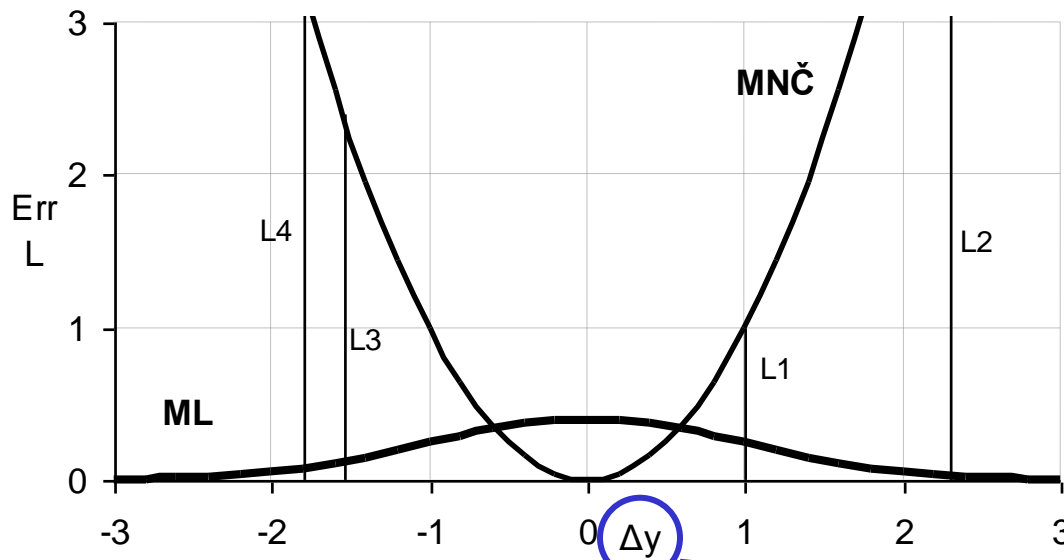
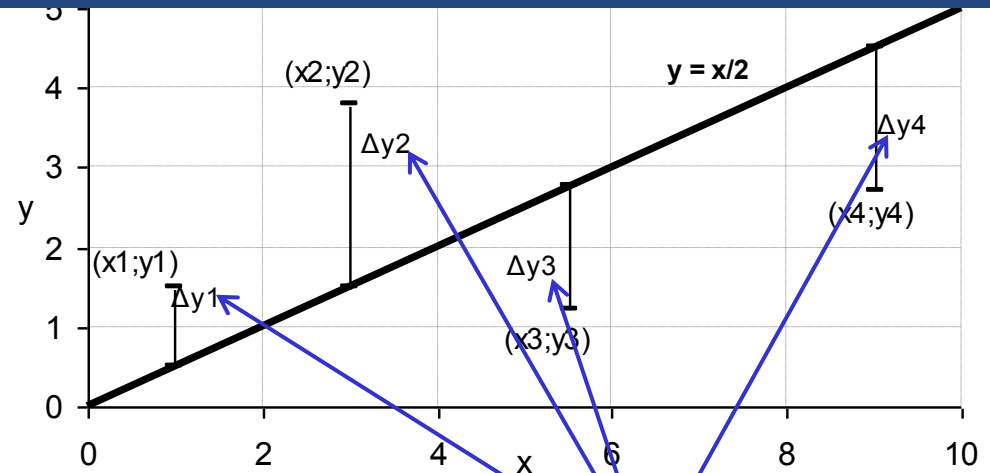
$$LF(x_i, y_i, f(x_i, \mathbf{b})) = \begin{cases} \frac{1}{2\sigma} [y_i - f(x_i, \mathbf{b})]^2 & , \text{pokud } |y_i - f(x_i, \mathbf{b})| \leq \sigma \\ |y_i - f(x_i, \mathbf{b})| - \frac{\sigma}{2} & , \text{jinak} \end{cases}$$

- **Vždy se jako vstupní informace objevuje** $[y_i - f(x_i, \mathbf{b})] = \Delta y_i$

Chybové funkce: MNČ vs. ML

$$L = \prod h(\Delta y_i)$$

$$Err = \sum \Delta y_i^2$$



Chybové funkce: MNČ vs. ML

- **rozdíl v typu transformační funkce** (transformace absolutní hodnoty vstupní difference Δy je v případě MNČ funkce rostoucí, v případě MLE funkce klesající)
- **velikost optimalizační funkce** (u MNČ je funkce označována za chybu a je minimalizována, v případě MLE se hovoří o věrohodnosti a její velikost je maximalizována)
- **způsob výpočtu souhrnné chyby modelu** (u MNČ získána součtem všech dílčích chyb, u MLE je použit součin jednotlivých věrohodností)
- **informace obsažená v celkové chybě modelu** (u MLE má význam absolutní hodnota chyby, která vyjadřuje podmíněnou pravděpodobnost za předpokladu předem zvoleného typu rozložení chyby)

LF – kvalitativní – binární

model \ realita	+, P	-, N
+, \hat{P}	TP – true positive	FP – false positive
-, \hat{N}	FN – false negative	TN – true negative

- $P = TP + FN$ (skutečně pozitivních), $\hat{P} = TP + FP$ (predikovaných), N, \hat{N}
- $T = TP + TN$ (správně klasifikovaných), $T = FP + FN$ (chybně klas.)
- celková správnost (*accuracy*) $Acc = T / (T+F)$
- chyba (*error*) $Err = F / (T+F)$
- senzitivita (úplnost, *recall*) $Sens = TP / (TP+FN) = TP / P$
- specificita $Spec = TN / (TN+FP) = TN / N$
- negativní prediktivní hodnota $NPV = TN / \hat{N}$
- pozitivní prediktivní hodnota (přesnost, *precision*) $PPV = TP / \hat{P}$
- F míra $F = \frac{2 \cdot PPV \cdot Sens}{PPV + Sens} = \frac{2 \cdot TP}{\hat{P} + P}$

Příklad: accuracy vs ROC vs F-míra

Případ 1

m \ r	+	-
+	81	5
-	9	5

V obou případech

$$p(m=+ | r=+) = 0,9$$

$$p(m=- | r=-) = 0,5$$

Případ 2

m \ r	+	-
+	45	25
-	5	25

	1. případ	2. případ
• $Acc = T / (T+F)$	0,86	0,7
• $Err = F / (T+F)$	0,14	0,3
• $Sens = TP / (TP+FN)$	0,9	0,9
• $Spec = TN / (TN+FP)$	0,5	0,5
• $NPV = TN / N$	0,36	0,83
• $PPV = TP / P$	0,83	0,64
• $F\ míra$	0,92	0,75

Často se používá **plocha pod ROC křivkou (AUC)**, která vzniká vynesemím bodů senzitivity a specificity do 2D grafu. Více v přednášce o předzpracování dat.

LF – kvalitativní – vícerozměrné

- větší váha opakováním daných prvků
- nákladová matice (cost matrix)

	realita	A	B	C
model				
A		0	4	10
B		1	0	8
C		1	1	0

? Co je to nákladová matice (cost matrix)?