



EVROPSKÁ UNIE  
Evropské strukturální a investiční fondy  
Operační program Výzkum, vývoj a vzdělávání



Název projektu	Rozvoj vzdělávání na Slezské univerzitě v Opavě
Registrační číslo projektu	CZ.02.2.69/0.0./0.0/16_015/0002400

**Dolování dat**

**Strojové učení**

**Jan Górecki**



**SLEZSKÁ  
UNIVERZITA**  
OBCHODNĚ PODNIKATELSKÁ  
FAKULTA V KARVINĚ

# Obsah přednášky

---

- Co je to učení
- Co je to strojové učení
- Metody učení
- Učení jako aproximace



# Učení

---



- **Je přirozenou charakteristikou živých organismů nezbytnou pro jejich přežití a vývoj (oheň, ...)**
  - **Schopnost učit se bývá často považováno za definici inteligence**
  - **Snaha vybavit touto schopností i stroje**
  - **Často bez toho nelze problém ani řešit (vidění, porozumění textu)**
  - **Prvky učení lze nalézt v: explorační analýza dat (statistika), rozpoznávání obrazů nebo strojové učení (AI), adaptivní a učící se systémy (kybernetika)**
-



Things learn when they change their behavior in a way that makes them perform better in a future.

(Witten, Frank, 1999)

The field of machine learning is concerned with the question of how to construct computer programs that automatically improve with experience.

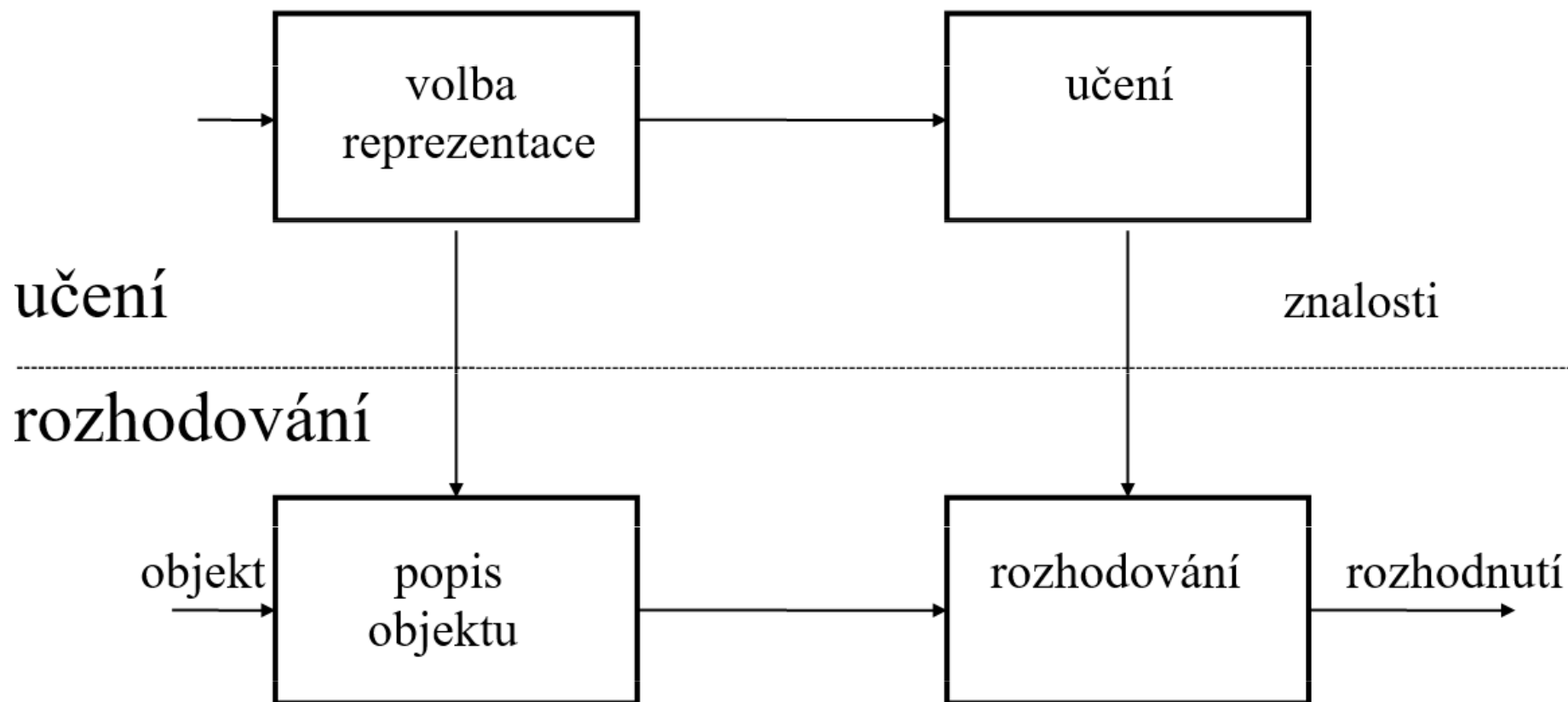
(Mitchell, 1997)

---

# Obecné schéma učícího se systému



SLEZSKÁ  
UNIVERZITA  
OBCHODNĚ PODNIKATELSKÁ  
FAKULTA V KARVINĚ

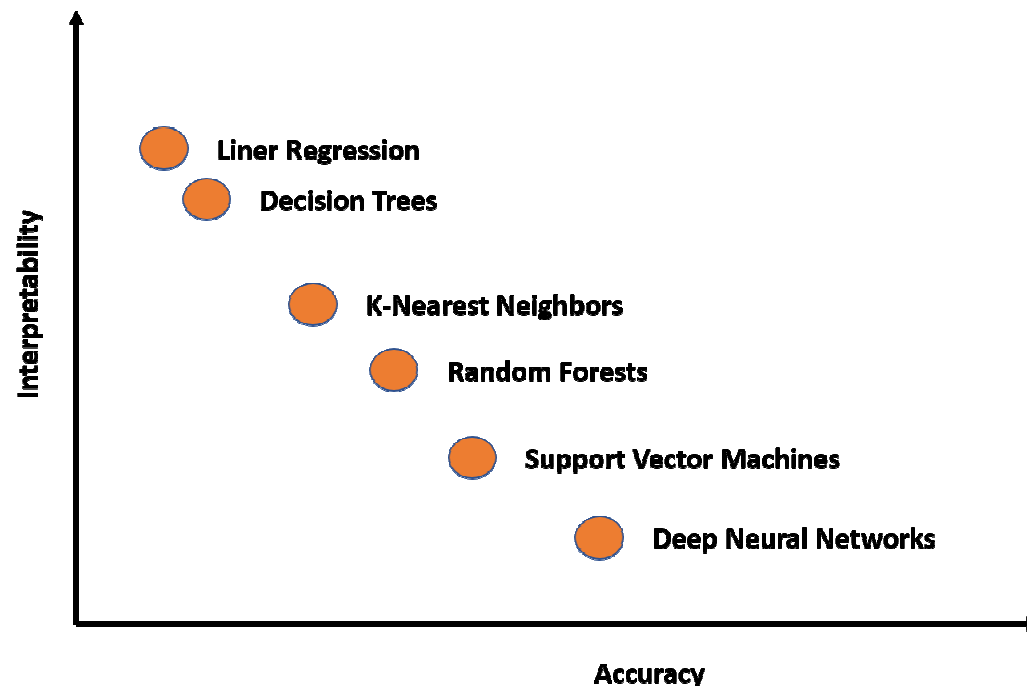


# Metody učení – principy pro získávání znalostí



SLEZSKÁ  
UNIVERZITA  
OBCHODNĚ PODNIKATELSKÁ  
FAKULTA V KARVINĚ

- **statistické metody** – kontingenční tabulky, regresní metody, shluková analýza
- **symbolické metody umělé inteligence** - rozhodovací stromy a pravidla
- **subsymbolické metody umělé inteligence** - neuronové sítě, bayesovské sítě nebo genetické algoritmy.



# Informace o správnosti učení

---



SLEZSKÁ  
UNIVERZITA  
OBCHODNĚ PODNIKATELSKÁ  
FAKULTA V KARVINĚ

Může mít podobu:

1. příkladů zařazených do tříd (**učení s učitelem - supervised learning**)
2. odměny za správné chování a tresty za chování nesprávné (**učení posilováním - reinforcement learning**)
3. žádné (**učení bez učitele - unsupervised learning**)

Příklady:

- Úspěch ChatGPT stojí na správném spojení všech těchto tří typů učení
  - Hide and seek
-

# Principy empirického učení z dat – 1.

---



SLEZSKÁ  
UNIVERZITA  
OBCHODNĚ PODNIKATELSKÁ  
FAKULTA V KARVINĚ

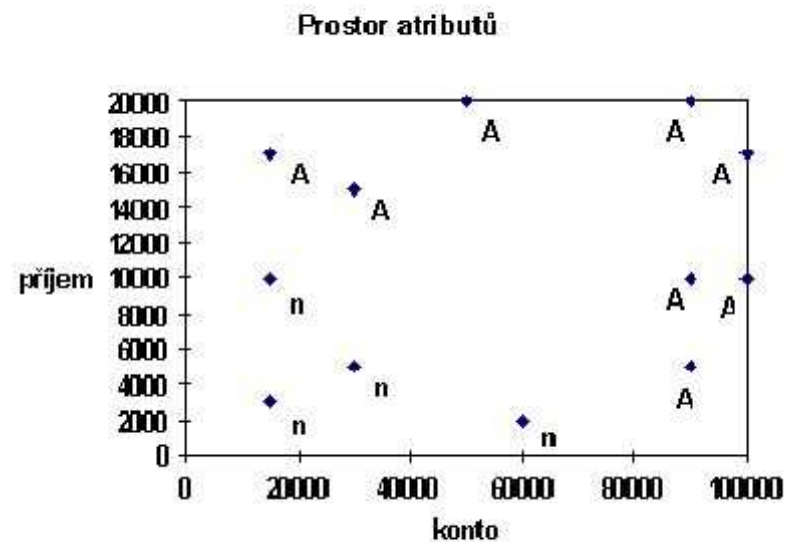
- objekty, patřící do téže třídy mají podobné charakteristiky (**učení na základě podobnosti, similarity-based learning**)
    - příklady téže třídy vytvářejí shluky v prostoru vstupních atributů
    - cílem učení je tyto shluky nalézt a popsat
-



# Principy empirického učení z dat – 2.



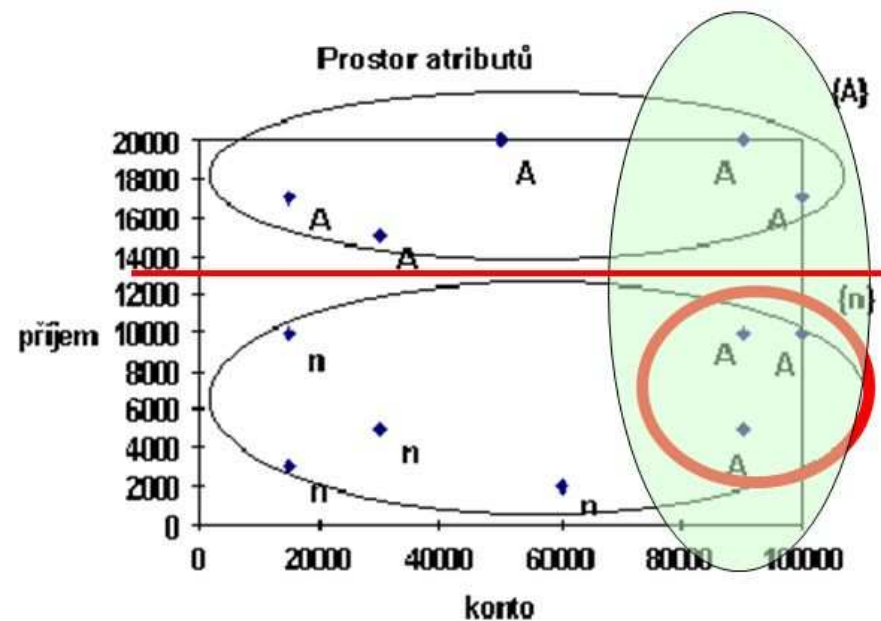
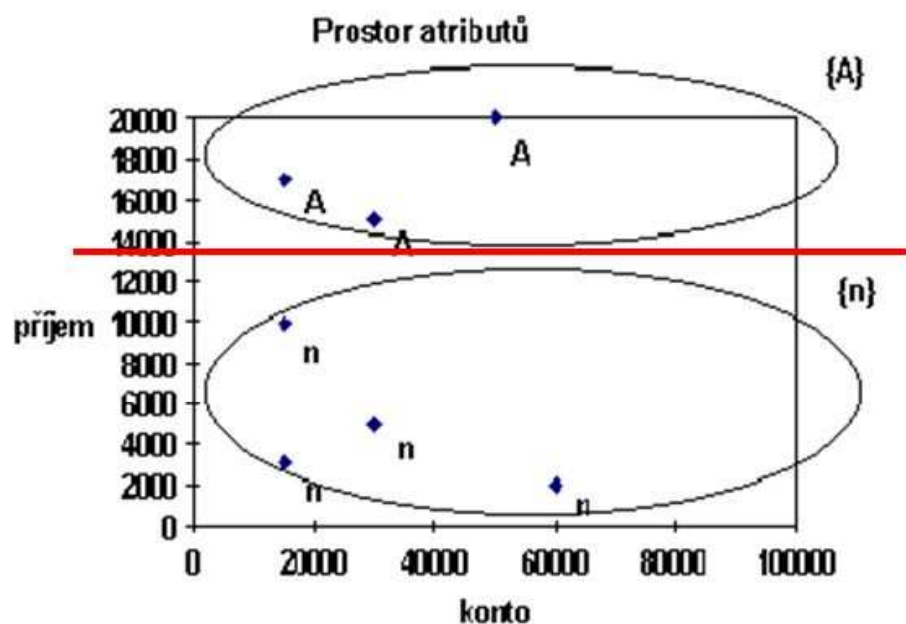
klient	konto	příjem	půjčit
01	15000	3000	ne
02	15000	10000	ne
03	15000	17000	Ano
04	30000	5000	ne
05	30000	15000	Ano
06	50000	20000	Ano
07	60000	2000	ne
08	90000	5000	Ano
09	90000	10000	Ano
10	90000	20000	Ano
11	100000	10000	Ano
12	100000	17000	Ano



# Principy empirického učení z dat – 3.



- z konečného počtu příkladů odvozujeme obecné znalosti (**induktivnost**)



# Obecná definice strojového učení

---



SLEZSKÁ  
UNIVERZITA  
OBCHODNĚ PODNIKATELSKÁ  
FAKULTA V KARVINĚ

Analyzovaná data:

$$D = \begin{bmatrix} X_{11} & X_{12} & \dots & X_{1m} \\ X_{21} & X_{22} & \dots & X_{2m} \\ \vdots & \vdots & & \vdots \\ X_{n1} & X_{n2} & \dots & X_{nm} \end{bmatrix}$$

Řádky tabulky reprezentují sledované **objekty**

Sloupce tabulky odpovídají **atributům**

---

# Obecná definice strojového učení (s učitelem)



SLEZSKÁ  
UNIVERZITA  
OBCHODNĚ PODNIKATELSKÁ  
FAKULTA V KARVINĚ

Přidáme-li cílový atribut do datové tabulky, získáme data vhodná pro použití některé metody učení s učitelem (tzv. **trénovací data**).

$$\mathbf{D}_{\text{TR}} = \begin{bmatrix} x_{11} & x_{12} & \dots & x_{1m} & y_1 \\ x_{21} & x_{22} & \dots & x_{2m} & y_2 \\ \vdots & \vdots & & \vdots & \vdots \\ x_{n1} & x_{n2} & \dots & x_{nm} & y_n \end{bmatrix}$$

Objekt (trénovací příklad) z této tabulky budeme značit

$$\mathbf{o}_i = [ \mathbf{x}_i, y_i ]$$

**Klasifikační úloha:** hledáme **znalosti** (reprezentované rozhodovací funkcí (modelem)  $f$ ), které by umožňovaly k hodnotám vstupních atributů nějakého objektu přiřadit vhodnou hodnotu atributu cílového

$$f: \mathbf{x} \rightarrow y.$$

# Obecná definice strojového učení (s učitelem)



SLEZSKÁ  
UNIVERZITA  
OBCHODNĚ PODNIKATELSKÁ  
FAKULTA V KARVINĚ

V průběhu klasifikace se tedy pro hodnoty vstupních atributů  $\mathbf{x}$  nějakého objektu odvodí hodnota cílového atributu  $\hat{y}$ :

$$\hat{y} = f(\mathbf{x}).$$

Odvozená hodnota  $\hat{y}$  se pro objekty z trénovacích dat **může lišit** od skutečné hodnoty  $y$ . Můžeme tedy pro každý objekt  $\mathbf{o}_i \in D_{\text{TR}}$  vyčíslit *chybu klasifikace*  $Q_f(\mathbf{o}_i, \hat{y}_i)$ :

V případě **numerického** atributu  $C$  může být touto chybou například čtverec rozdílu skutečné a odvozené hodnoty cílového atributu

$$Q_f(\mathbf{o}_i, \hat{y}_i) = (y_i - \hat{y}_i)^2,$$

v případě **kategoriálního** atributu  $C$  může být touto chybou informace o tom že se odvozená a skutečná hodnota vzájemně liší,

$$Q_f(\mathbf{o}_i, \hat{y}_i) = \begin{cases} 1 & \text{pro } y_i \neq \hat{y}_i \\ 0 & \text{pro } y_i = \hat{y}_i \end{cases}$$

# Chyba na trénovacích datech

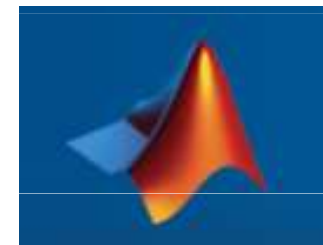
---



SLEZSKÁ  
UNIVERZITA  
OBCHODNĚ PODNIKATELSKÁ  
FAKULTA V KARVINĚ

Pro celou trénovací množinu  $D_{TR}$  pak můžeme vyčíslit souhrnnou chybu  $Err(f, D_{TR})$ , například jako střední chybu

$$Err(f, D_{TR}) = \frac{1}{n} \sum_{i=1}^n Q_f(\mathbf{o}_i, \hat{y}_i)$$



Cílem učení je nalézt takové znalosti  $f^*$ , které by minimalizovaly tuto chybu

$$Err(\underline{f^*}, D_{TR}) = \min_f Err(f, D_{TR}).$$

---

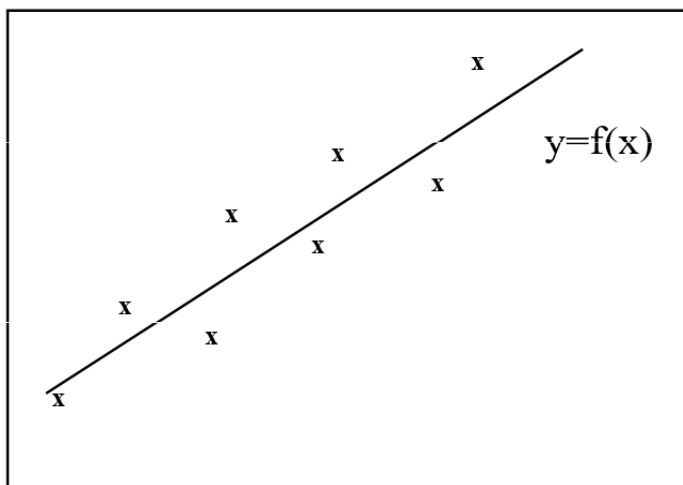
# Učení jako aproximace



- Hledáme „pouze“ parametry modelu

## Příklad:

na základě hodnot funkce v konečném počtu bodů snažíme zrekonstruovat její obecnou podobu



$$f(x) = q_1x + q_0$$

# Metoda nejmenších čtverců

---



SLEZSKÁ  
UNIVERZITA  
OBCHODNĚ PODNIKATELSKÁ  
FAKULTA V KARVINĚ

Hledání minima celkové chyby

$$\min \sum_i (y_i - f(x_i))^2$$

se převádí na řešení rovnice

$$\frac{d}{dq} \sum_{i=1}^n (y_i - f(\mathbf{x}_i))^2 = 0$$

---



# Řešení



## 1) analytické (známe typ funkce)

řešení soustavy rovnic pro parametry funkce

$$\begin{aligned} \min \sum_i (y_i - f(x_i))^2 &\Rightarrow \frac{\partial}{\partial q_0} \sum_{i=1}^n (y_i - (q_1 x_i + q_0))^2 = -2 \sum_{i=1}^n y_i + 2q_1 \sum_{i=1}^n x_i + 2q_0 n \\ f(x) = q_1 x + q_0 &\quad \frac{\partial}{\partial q_1} \sum_{i=1}^n (y_i - (q_1 x_i + q_0))^2 = -2 \sum_{i=1}^n x_i y_i + 2q_1 \sum_{i=1}^n x_i^2 + 2q_0 \sum_{i=1}^n x_i \end{aligned} \Rightarrow \begin{aligned} q_0 &= \frac{(\sum_i y_i)(\sum_i x_i^2) - (\sum_i x_i y_i)(\sum_i x_i)}{n(\sum_i x_i^2) - (\sum_i x_i)^2} \\ q_1 &= \frac{n(\sum_i x_i y_i) - (\sum_i x_i)(\sum_i y_i)}{n(\sum_i x_i^2) - (\sum_i x_i)^2} \end{aligned}$$

## 2) numerické (neznáme typ funkce)

gradientní metody

# Gradientní metody



$$\nabla \text{Err}(\mathbf{q}) = \left[ \frac{\partial \text{Err}}{\partial q_0}, \frac{\partial \text{Err}}{\partial q_1}, \dots, \frac{\partial \text{Err}}{\partial q_Q} \right]$$

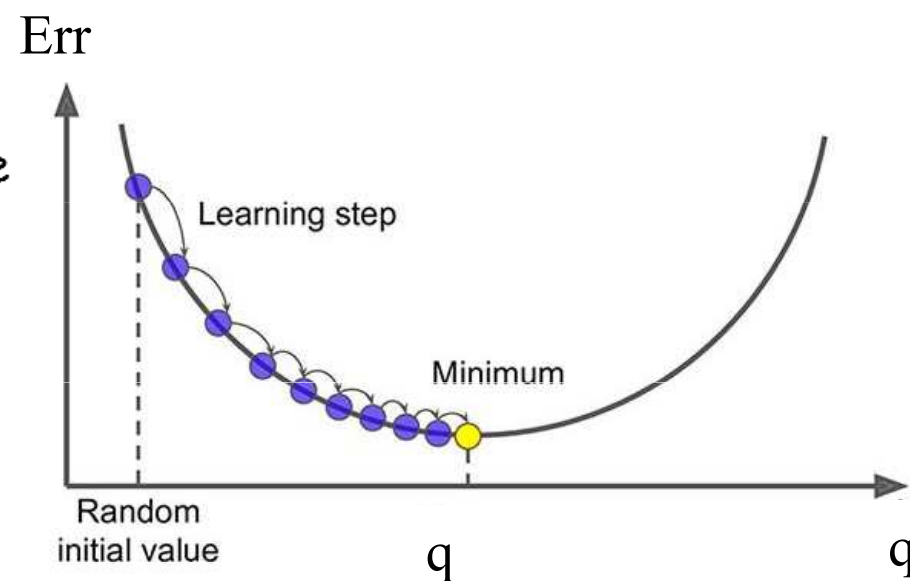
Modifikace znalostí  $\mathbf{q} = [q_0, q_1, \dots, q_Q]$  pak probíhá podle algoritmu

$$q_j \leftarrow q_j + \Delta q_j$$

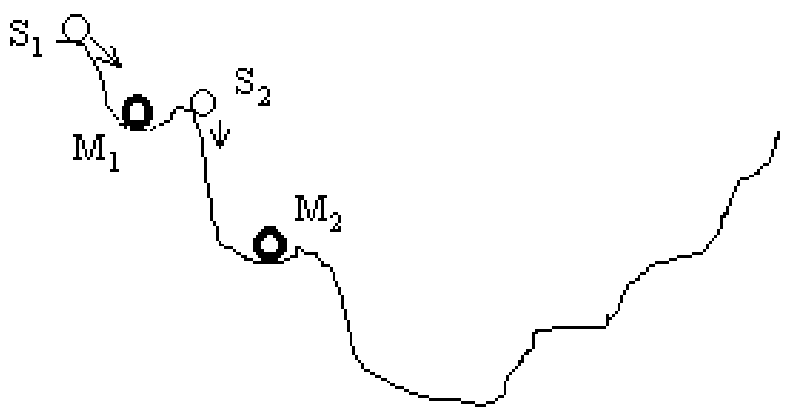
kde

$$\Delta q_j = -\eta \frac{\partial \text{Err}}{\partial q_j}$$

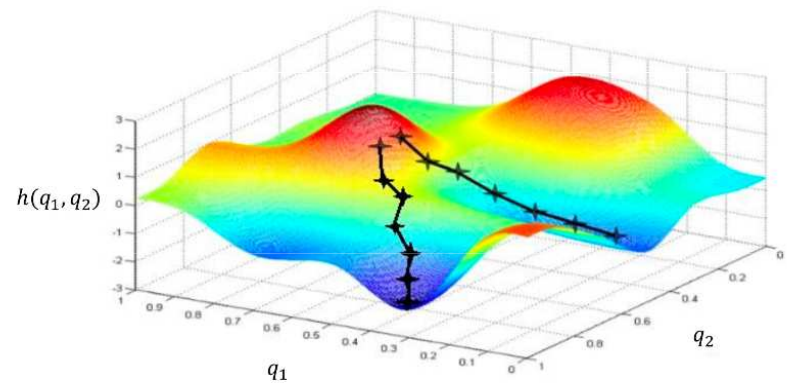
a  $\eta$  je parametr vyjadřující „velikost kroku“ kterým se přibližujeme k minimu funkce  $\text{Err}$ .



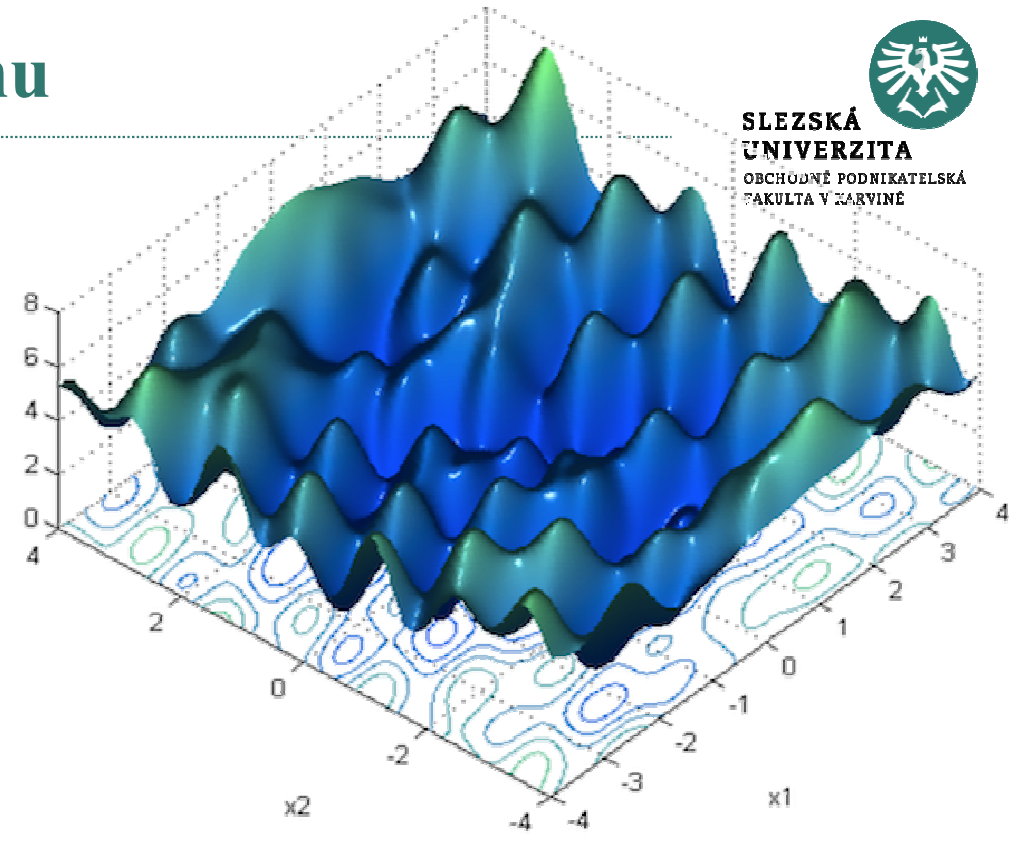
# Problém uváznutí v lokálním minimu



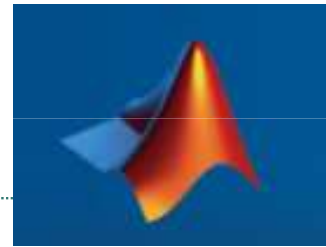
Non-convex Example



Wolfram Global Problem



SLEZSKÁ  
UNIVERZITA  
OBCHODNĚ PODNIKATELSKÁ  
FAKULTA V KARVINĚ



# Děkuji za pozornost

Některé snímky převzaty od:

prof. Ing. Petr Berka, CSc. [berka@vse.cz](mailto:berka@vse.cz)

# Metody učení (dle vynaloženého úsilí)

---



**SLEZSKÁ  
UNIVERZITA**  
OBCHODNĚ PODNIKATELSKÁ  
FAKULTA V KARVINĚ

1. učení zapamatováním (rote learning neboli biflování),
  2. učení se z instrukcí (learning from instruction, learning by being told),
  3. učení se z analogie (learning by analogy, instance-based learning, lazy learning),
  4. učení na základě vysvětlení (explanation-based learning),
  5. učení se z příkladů (learning from examples),
  6. učení se z pozorování a objevování (learning from observation and discovery),
-

## Principy empirického učení z dat – 4.

---



SLEZSKÁ  
UNIVERZITA  
OBCHODNĚ PODNIKATELSKÁ  
FAKULTA V KARVINĚ

- Příklady rozděleny do 2 (někdy 3) množin:
    - **trénovací data** pro vytvoření modelu
    - (**validační data** pro doladění parametrů)
    - **testovací data** pro otestování modelu
-



- **empirické** – vychází se z velkého množství příkladů a žádných (nebo jen mála) počátečních znalostí
  - **analytické** – vychází se z velkého množství počátečních znalostí a jen několika (ilustračních) příkladů
-