



FACULTY OF MATHEMATICS,
PHYSICS AND INFORMATICS

Comenius University
Bratislava

Úvod do Umelej Inteligencie

Rozhodovacie stromy

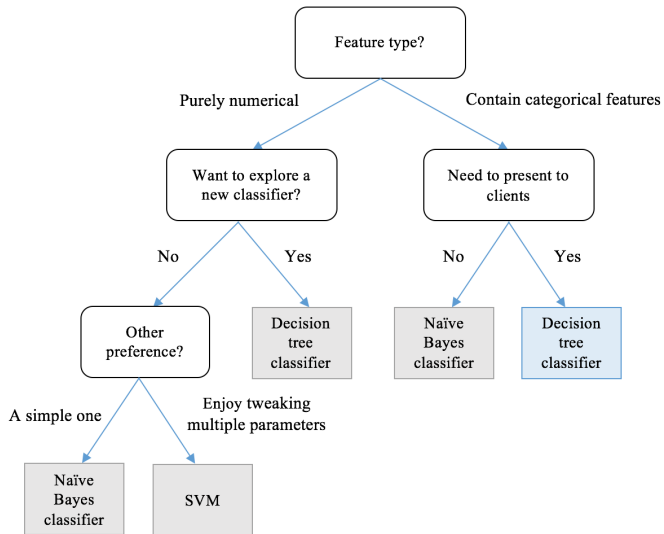
Ing. Viktor Kocur, PhD.

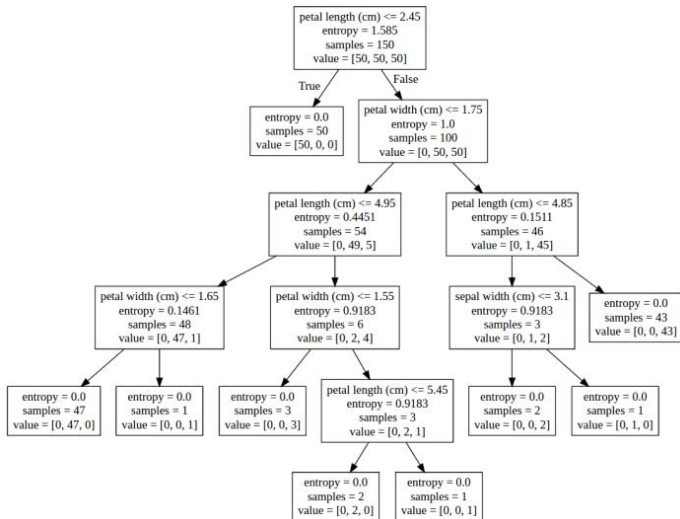
6.12.2023



Klasifikácia: Pod pojmom klasifikácia rozumieme rozdelovanie objektov, alebo javov do konečnej množiny tried. Rozdelovanie realizujeme na základe príznakov.

Príznamy: V strojovom učení popisujeme objekty a javy pomocou tzv. príznakov. Príznak je individuálna merateľná vlastnosť, alebo charakteristika. Príznamy môžu byť väčšinou numerické, alebo kategorické/nominálne.







Strom konštruujeme, tak že vyberáme príznak a jeho hodnotu na základe ktorého rozdelíme množinu prvkov na dve časti. Tento postup opakujeme s oboma podmnožinami až kým nieje splnené ukončujúce kritérium.

Existujú rôzne postupy. Napr.:

- ID3
- C4.5
- CART



Je možné používať rôzne kritéria:

- Všetky objekty v skúmanej množine patria do jednej klasifikačnej triedy
- Strom dosiahol maximálne stanovenú hĺbku
- Počet objektov klasifikovaných v danom uzle je menší ako stanovený prah
- Ohodnotenie najlepšieho príznaku je nižšie ako stanovený prah



Day	Outlook	Temp.	Humidity	Wind	Decision
1	Sunny	Hot	High	Weak	No
2	Sunny	Hot	High	Strong	No
3	Overcast	Hot	High	Weak	Yes
4	Rain	Mild	High	Weak	Yes
5	Rain	Cool	Normal	Weak	Yes
6	Rain	Cool	Normal	Strong	No
7	Overcast	Cool	Normal	Strong	Yes
8	Sunny	Mild	High	Weak	No
9	Sunny	Cool	Normal	Weak	Yes
10	Rain	Mild	Normal	Weak	Yes
11	Sunny	Mild	Normal	Strong	Yes
12	Overcast	Mild	High	Strong	Yes
13	Overcast	Hot	Normal	Weak	Yes
14	Rain	Mild	High	Strong	No



Pri každom kroku vyberáme príznak pre ktorý bude entropia minimálna, teda taký pre ktorý je informačný prínos najväčší (vzájomná informácia s triedami je najväčšia).

Definujme si množinu tried ω . Napríklad $\omega = \{\text{hráme tenis, nehráme tenis}\}$.

Potom máme rôzne príznaky. Napríklad pre príznak oblačnosť môžeme mať hodnoty: slnečno, oblačno, dážď. Pre príznak vietor môžeme mať áno, nie.



$$H(Y) = \sum_{y \in \omega} -P(Y = y) \cdot \log_2(P(Y = y)), \quad (1)$$

kde $P(Y = y)$ označuje pravdepodobnosť danej triedy. V tomto prípade ju vyjadríme ako jednoduchý pomer prvkov z danej triedy voči všetkým prvkom trénovacej množiny.



Budeme tiež používať špecifickú podmienenú entropiu:

$$H(Y|X = v) = H(Y), \text{ len pre hodnoty } Y, \text{ kde } X = v \quad (2)$$

Teda entropia ak sa obmedzíme len na prvky pre ktoré príznak X má hodnoty v .

Vďaka tomu môžeme vypočítať vzájomnú informáciu:

$$I(Y; X) = H(Y) - H(Y|X) = H(Y) - \sum_{x \in \omega} P(X = x) \cdot H(Y|X = x) \quad (3)$$



$$H(Y) = - \sum_{y \in \omega} P(Y = y) \cdot \log_2(P(Y = y))$$

X	Y
Matematika	Áno
História	Nie
Informatika	Áno
Matematika	Nie
Matematika	Nie
Informatika	Áno
História	Nie
Matematika	Áno



$$H(Y) = - \sum_{y \in \omega} P(Y = y) \cdot \log_2(P(Y = y))$$

■ $H(X) = 1.5$

X	Y
Matematika	Áno
História	Nie
Informatika	Áno
Matematika	Nie
Matematika	Nie
Informatika	Áno
História	Nie
Matematika	Áno



$$H(Y) = - \sum_{y \in \omega} P(Y = y) \cdot \log_2(P(Y = y))$$

X	Y
Matematika	Áno
História	Nie
Informatika	Áno
Matematika	Nie
Matematika	Nie
Informatika	Áno
História	Nie
Matematika	Áno

■ $H(X) = 1.5$

■ $H(Y) = 1$



$$H(Y|X = v) = H(Y), \text{ len pre hodnoty } Y, \text{ kde } X = x$$

X	Y
Matematika	Áno
História	Nie
Informatika	Áno
Matematika	Nie
Matematika	Nie
Informatika	Áno
História	Nie
Matematika	Áno



$H(Y|X = v) = H(Y)$, len pre hodnoty Y , kde $X = x$

X	Y
Matematika	Áno
História	Nie
Informatika	Áno
Matematika	Nie
Matematika	Nie
Informatika	Áno
História	Nie
Matematika	Áno

■ $H(Y|X = M) = 1$



$H(Y|X = v) = H(Y)$, len pre hodnoty Y , kde $X = x$

X	Y
Matematika	Áno
História	Nie
Informatika	Áno
Matematika	Nie
Matematika	Nie
Informatika	Áno
História	Nie
Matematika	Áno

■ $H(Y|X = M) = 1$

■ $H(Y|X = H) = 0$



$H(Y|X = v) = H(Y)$, len pre hodnoty Y , kde $X = x$

X	Y
Matematika	Áno
História	Nie
Informatika	Áno
Matematika	Nie
Matematika	Nie
Informatika	Áno
História	Nie
Matematika	Áno

■ $H(Y|X = M) = 1$

■ $H(Y|X = H) = 0$

■ $H(Y|X = I) = 0$



$$H(Y|X) = \sum_{x \in \Omega_x} P(X = x) \cdot H(Y|X = x)$$

X	Y
Matematika	Áno
História	Nie
Informatika	Áno
Matematika	Nie
Matematika	Nie
Informatika	Áno
História	Nie
Matematika	Áno



$$H(Y|X) = \sum_{x \in \Omega_x} P(X = x) \cdot H(Y|X = x)$$

■ $H(Y|X) = 0.5$

X	Y
Matematika	Áno
História	Nie
Informatika	Áno
Matematika	Nie
Matematika	Nie
Informatika	Áno
História	Nie
Matematika	Áno



ID3

`https://sefiks.com/2017/11/20/a-step-by-step-id3-decision-tree-example/`

C4.5

`https://sefiks.com/2018/05/13/a-step-by-step-c4-5-decision-tree-example/`