Základy matematiky a statistiky pro humanitní obory II

Pavel Rychlý Vojtěch Kovář

Fakulta informatiky, Masarykova univerzita
Botanická 68a, 602 00 Brno, Czech Republic
{pary, xkovar3}@fi.muni.cz

část 8

Obsah přednášky

Entropie náhodné veličiny

- Míra informace náhodné veličiny
  - kolik informace získáme, když se dozvíme hodnotu náhodné veličiny
  - „hodnota informace", kterou nám veličina dává
  - měří se v bitech
  - nulová entropie = jsme schopni určit hodnotu veličiny se 100% jistotou
- Počátky
  - 40. léta (Shannon)
  - potřeba přenést informaci co nejmenší možnou zprávou

Entropie

- Vzorec
  - $H(p) = H(X) = -\sum_{x \in X} p(x) log_2 p(x)$
  - X = množina možných hodnot
  - p = pravděpodobnostní rozložení
- Příklad – hod dvěma mincemi, počítáme panny
  - p(0) = 1/4, p(1) = 1/2, p(2) = 1/4
  - $H(p) = -(1/4 log_2(1/4) + 1/2 log_2(1/2)) + 1/4 log_2(1/4)) = -(-2/4 - 1/2 - 2/4) = 1.5$ *bitu*
- Pokud budou na obou mincích padat pouze panny
  - p(0) = 0, p(1) = 0, p(2) = 1
  - $H(p) = -(log_2(1)) = -(0) = 0$
  - $\rightarrow$ nemusíme předávat žádnou informaci, abychom zjistili, že padly dvě panny

Podmíněná entropie

- Podobně jako podmíněná pravděpodobnost
  - H(X|Y) – entropie veličiny X za předpokladu, že známe hodnoty veličiny Y
  - $H(p) = H(X|Y) = \sum_{x \in X} p(x) H(Y|X = x)$
- Řetízkové pravidlo (chain rule)
  - H(X, Y) = H(X) + H(Y|X)

Mutual information (vzájemná informace)

- Míra informace, kterou jedna náhodná proměnná říká o jiné
  - vzorec: MI(X; Y) = H(X) - H(X|Y) = H(Y) - H(Y|X)
  - 0, pokud jsou veličiny nezávislé
  - čím vyšší, tím více hodnoty jedné vlastnosti určují hodnoty druhé vlastnosti
- Příklad použití – kolokace
  - X: výskyt slova *a* (např. „základní") v textu
  - Y: výskyt slova *b* (např. „škola") v textu
  - MI je měřítkem „síly" kolokace těchto dvou slov
  - je tím vyšší, čím vyšší je počet souvýskytů slov a tím nižší, čím jsou slova častější