Entropija

Nenad Markuš

7. listopada 2014.

Model izvora informacije <u>bez memorije</u>

Model izvora informacije bez memorije

- Slučajna varijabla X poprima vrijednosti iz skupa $\mathbb{X} = \{x_1, x_2, \ldots, x_n\}$ s vjerojatnostima p_i:

$$X \sim \begin{pmatrix} x_1 & x_2 & \ldots & x_n \\ p_1 & p_2 & \ldots & p_n \end{pmatrix}$$

Model izvora informacije bez memorije

- Slučajna varijabla X poprima vrijednosti iz skupa $\mathbb{X} = \{x_1, x_2, \ldots, x_n\}$ s vjerojatnostima p_i:

$$X \sim \begin{pmatrix} x_1 & x_2 & \ldots & x_n \\ p_1 & p_2 & \ldots & p_n \end{pmatrix}$$

- Primjer: bacanje novčića

Model izvora informacije bez memorije

- Slučajna varijabla X poprima vrijednosti iz skupa $\mathbb{X} = \{x_1, x_2, \ldots, x_n\}$ s vjerojatnostima p_i:

$$X \sim \begin{pmatrix} x_1 & x_2 & \ldots & x_n \\ p_1 & p_2 & \ldots & p_n \end{pmatrix}$$

- Primjer: bacanje novčića
- **Poruka** duljine N je element skupa $\mathbb{X}^N$ nastao realizacijama N nezavisnih kopija varijable X

Neka razumna razmatranja o informaciji

Neka razumna razmatranja o informaciji

- Ako je $p_1 = 1$, a za sve ostale $p_i = 0$, tada nema iznenađenja, jer unaprijed znamo ishod prijenosa. Intuitivno, u ovom slučaju ne dolazi do prijenosa informacije.

Neka razumna razmatranja o informaciji

- Ako je $p_1 = 1$, a za sve ostale $p_i = 0$, tada nema iznenađenja, jer unaprijed znamo ishod prijenosa. Intuitivno, u ovom slučaju ne dolazi do prijenosa informacije.
- Informacija je obrnuto proporcionalna vjerojatnosti pojave simbola. Ako primimo manje vjerojatan simbol, iznenađenje je veće.

Neka razumna razmatranja o informaciji

- Ako je $p_1 = 1$, a za sve ostale $p_i = 0$, tada nema iznenađenja, jer unaprijed znamo ishod prijenosa. Intuitivno, u ovom slučaju ne dolazi do prijenosa informacije.
- Informacija je obrnuto proporcionalna vjerojatnosti pojave simbola. Ako primimo manje vjerojatan simbol, iznenađenje je veće.
- Neodređenost, iznenađenje i informacija su vezani:
    1. Prije nekog događaja (eksperiment, prijem poruke, i sl.) postoji određena količina neodređenosti
    2. Kad se događaj zbije, postoji određena količina iznenađenja
    3. Nakon događaja nastala je određena količina informacije

Neka razumna razmatranja o informaciji

- Ako je $p_1 = 1$, a za sve ostale $p_i = 0$, tada nema iznenađenja, jer unaprijed znamo ishod prijenosa. Intuitivno, u ovom slučaju ne dolazi do prijenosa informacije.
- Informacija je obrnuto proporcionalna vjerojatnosti pojave simbola. Ako primimo manje vjerojatan simbol, iznenađenje je veće.
- Neodređenost, iznenađenje i informacija su vezani:
  1. Prije nekog događaja (eksperiment, prijem poruke, i sl.) postoji određena količina neodređenosti
  2. Kad se događaj zbije, postoji određena količina iznenađenja
  3. Nakon događaja nastala je određena količina informacije
- Ako uparimo dva nezavisna izvora, X i Y, koliko se tada generira informacije?

Mjera količine informacije

Mjera količine informacije

- Nyquist (1924.) i Hartley (1928.) zaključuju da mjera za količinu informacije mora imati logaritamski karakter.

Mjera količine informacije

- Nyquist (1924.) i Hartley (1928.) zaključuju da mjera za količinu informacije mora imati logaritamski karakter.
    1. $\log n > 0$, za svaki $n \in \mathbb{N}$

Mjera količine informacije

- Nyquist (1924.) i Hartley (1928.) zaključuju da mjera za količinu informacije mora imati logaritamski karakter.
    1. $\log n > 0$, za svaki $n \in \mathbb{N}$
    2. $m < n \implies \log m < \log n$

Mjera količine informacije

- Nyquist (1924.) i Hartley (1928.) zaključuju da mjera za količinu informacije mora imati logaritamski karakter.
  1. $\log n > 0$, za svaki $n \in \mathbb{N}$
  2. $m < n \implies \log m < \log n$
  3. $\log(m \cdot n) = \log m + \log n$

Mjera količine informacije

- Nyquist (1924.) i Hartley (1928.) zaključuju da mjera za količinu informacije mora imati logaritamski karakter.
  1. $\log n > 0$, za svaki $n \in \mathbb{N}$
  2. $m < n \implies \log m < \log n$
  3. $\log(m \cdot n) = \log m + \log n$
- Shannon (1948.) predlaže mjeru koja uzima u obzir različite vjerojatnosti pojavljivanja pojedinih simbola.

Entropija

$$H = -\sum_{i=1}^{n} p_i \log p_i$$

Primjeri

Primjeri

- Bacanje novčića

Primjeri

- Bacanje novčića
- Binarna varijabla

$$B \sim \begin{pmatrix} 0 & 1 \\ p & (1-p) \end{pmatrix}$$

Primjeri

- Bacanje novčića
- Binarna varijabla

$$B \sim \begin{pmatrix} 0 & 1 \\ p & (1-p) \end{pmatrix}$$

- $H(X) = -\mathbb{E}[\log \mathbb{P}(X)]$

Primjeri

- Bacanje novčića
- Binarna varijabla

$$B \sim \begin{pmatrix} 0 & 1 \\ p & (1-p) \end{pmatrix}$$

- $H(X) = -\mathbb{E}[\log \mathbb{P}(X)]$
- Geometrijska razdioba

$$G \sim \begin{pmatrix} 1 & 2 & \dots & g & \dots \\ p & p(1-p) & \dots & p(1-p)^g & \dots \end{pmatrix}$$

Svojstva entropije

Svojstva entropije

1. $H \geq 0$

Svojstva entropije

1. $H \geq 0$
2. $H \leq \log n$ ako izvor generira n simbola

Svojstva entropije

1. $H \geq 0$
2. $H \leq \log n$ ako izvor generira n simbola
3. $H(X, Y) = H(X) + H(Y)$ ako su X i Y statistički nezavisni izvori

Entropija i kompresija podataka

Entropija i kompresija podataka

- Koliko je, u prosjeku, binarnih pitanja potrebno postaviti da bi se otkrila varijabla X?

$$X \sim \begin{pmatrix} a & b & c & d \\ \frac{1}{2} & \frac{1}{4} & \frac{1}{8} & \frac{1}{8} \end{pmatrix}$$

Entropija i kompresija podataka

- Koliko je, u prosjeku, binarnih pitanja potrebno postaviti da bi se otkrila varijabla X?

$$X \sim \begin{pmatrix} a & b & c & d \\ \frac{1}{2} & \frac{1}{4} & \frac{1}{8} & \frac{1}{8} \end{pmatrix}$$

- $\mathbb{P}(Q = 1) = 0.5$, $\mathbb{P}(Q = 2) = 0.25$, $\mathbb{P}(Q = 3) = 0.25$

Entropija i kompresija podataka

- Koliko je, u prosjeku, binarnih pitanja potrebno postaviti da bi se otkrila varijabla X?

$$X \sim \begin{pmatrix} a & b & c & d \\ \frac{1}{2} & \frac{1}{4} & \frac{1}{8} & \frac{1}{8} \end{pmatrix}$$

- $\mathbb{P}(Q = 1) = 0.5$, $\mathbb{P}(Q = 2) = 0.25$, $\mathbb{P}(Q = 3) = 0.25$
- $\mathbb{E}(Q) = 1.75$

Entropija i kompresija podataka

- Koliko je, u prosjeku, binarnih pitanja potrebno postaviti da bi se otkrila varijabla X?

$$X \sim \begin{pmatrix} a & b & c & d \\ \frac{1}{2} & \frac{1}{4} & \frac{1}{8} & \frac{1}{8} \end{pmatrix}$$

- $\mathbb{P}(Q = 1) = 0.5$, $\mathbb{P}(Q = 2) = 0.25$, $\mathbb{P}(Q = 3) = 0.25$
- $\mathbb{E}(Q) = 1.75$
- Kolika je entropija varijable X?

Entropija i kompresija podataka

- Koliko je, u prosjeku, binarnih pitanja potrebno postaviti da bi se otkrila varijabla X?

$$X \sim \begin{pmatrix} a & b & c & d \\ \frac{1}{2} & \frac{1}{4} & \frac{1}{8} & \frac{1}{8} \end{pmatrix}$$

- $\mathbb{P}(Q = 1) = 0.5$, $\mathbb{P}(Q = 2) = 0.25$, $\mathbb{P}(Q = 3) = 0.25$
- $\mathbb{E}(Q) = 1.75$
- Kolika je entropija varijable X?
- Ako izvor entropije H generira niz od N simbola, taj će se niz moći jednoznačno predstaviti u prosjeku s $N \cdot H$ bitova.

Uvjetna entropija

Uvjetna entropija

- Poznavanjem Y možemo reći nešto o X, u općem slučaju.

Uvjetna entropija

- Poznavanjem Y možemo reći nešto o X, u općem slučaju.
- Uvjetna entropija:

$$H(X|Y=y) = -\sum_{x \in \mathbb{X}} \mathbb{P}(X=x|Y=y) \log \mathbb{P}(X=x|Y=y)$$

Uvjetna entropija

- Poznavanjem Y možemo reći nešto o X, u općem slučaju.
- Uvjetna entropija:

$$H(X|Y=y) = -\sum_{x\in\mathbb{X}} \mathbb{P}(X=x|Y=y)\log\mathbb{P}(X=x|Y=y)$$

- Prosječna uvjetna entropija:

$$H(X|Y) = \sum_{y\in\mathbb{Y}} \mathbb{P}(Y=y)H(X|Y=y)$$

Svojstva uvjetne entropije

Svojstva uvjetne entropije

1. $H(X|Y) \geq 0$

Svojstva uvjetne entropije

1. $H(X|Y) \geq 0$
2. $H(X|Y) \neq H(Y|X)$

Svojstva uvjetne entropije

1. $H(X|Y) \geq 0$
2. $H(X|Y) \neq H(Y|X)$
3. $H(X,Y) = H(X) + H(Y|X) = H(Y) + H(X|Y)$

Obrada podataka

Obrada podataka

- Može li obrada podataka proizvesti informaciju?

Obrada podataka

- Može li obrada podataka proizvesti informaciju?
- **Ne:** $H(V) \geq H(f(V))$

Obrada podataka

- Može li obrada podataka proizvesti informaciju?
- **Ne:** $H(V) \geq H(f(V))$
- Ipak, nije sva informacija korisna:

  > *We are drowning in information and starving for knowledge.*

Veza s termodinamikom

Veza s termodinamikom

- Entropija fizičkog sustava:

$$S = k \log W$$

Veza s termodinamikom

- Entropija fizičkog sustava:

$$S = k \log W$$

-