Teorija kodiranja in kriptografija 2013/2014

Shannonov izrek

Arjana Žitnik

Univerza v Ljubljani, Fakulteta za matematiko in fiziko

Ljubljana, 27. 5. 2014

Teorija informacij

Claude Shannon je leta 1948 v članku

”A Mathematical Theory of Communication”

vpeljal koncepte in teoreme, sedaj priznane kot osnova teorije informacij.

Komunikacijski sistem

Predstavil je komunikacijski sistem, ki je sestavljen iz

- oddajnika
- komunikacijskega kanala, na katerega vpliva šum
- sprejemnika, ki mora rekonstruirati poslano sporočilo

Informacija in entropija

Da bi lahko kvantitativno analiziral prenos sporočil čez komunikacijski kanal, je vpeljal mero za količino informacije.

Informacija nastane, ko se zgodi eden od možnih naključnih dogodkov, na primer, pri metu kocke pade 3.

Veliko informacije dobimo, ko se zgodi dogodek, katerega verjetnost je majhna. Ko se zgodi dogodek, katerega verjetnost je velika, potem nismo dosti izvedeli.

Entropija ali mera nedoločenosti slučajne spremenljivke X je enaka povprečni količini informacije, ki nastane, ko X zavzame določeno vrednost.

Kako smiselno definirati entropijo

- $X_1 = \begin{cases} 1; & \text{če jutri na poti v šolo srečamo Billa Gatesa} \\ 0; & \text{sicer} \end{cases}$
- $X_2 = \begin{cases} 1; & \text{če pri metu kovanca pade številka} \\ 0; & \text{če pri metu kovanca pade slika} \end{cases}$
- $X_3 =$ število pik pri metu kocke
- $X_4 =$ številka srečke, ki bo prihodnji teden zadela na loteriji

$$H(X_1) < H(X_2) < H(X_3) < H(X_4).$$

Diskretne slučajne spremenljivke

Naj bo X diskretna slučajna spremenljivka s končno zalogo vrednosti:

$$X : \begin{pmatrix} x_1 & x_2 & \dots & x_n \\ p_1 & p_2 & \dots & p_n \end{pmatrix}$$

Potem označimo

$$H(X) = H(p_1, p_2, \dots, p_n)$$

.

Zahteve za funkcijo H

- $H(p_1, p_2, \ldots, p_n)$ je zvezna funkcija n spremenljivk; vrstni red $p_1, p_2, \ldots, p_n$ ni pomemben
- $H(p_1, p_2, \ldots, p_n) \geq 0$; vrednost nič je dosežena le, če je eden od p_i enak 1
- $H(p_1, p_2, \ldots, p_n, 0) = H(p_1, p_2, \ldots, p_n)$
- $H(p_1, p_2, \ldots, p_n)$ doseže max. vrednost pri enakomerni porazdelitvi ($p_1 = p_2 = \cdots = p_n = 1/n$).
- X in Y neodvisni slučajni spremenljivki:

$$H(X, Y) = H(X) + H(Y)$$

- $H(1/2, 1/2) = 1$

Definicija entropije

Izkaže se, da je funkcija s temi (in še nekaj dodatnimi) zahtevami enolično določena. Definiramo:

$$H(p_1, p_2, \ldots, p_n) = -\sum_{i=1}^{n} p_i \log_2 p_i = \sum_{i=1}^{n} p_i \log_2 \frac{1}{p_i}.$$

Velja:

- $H(\frac{1}{n}, \frac{1}{n}, \ldots, \frac{1}{n}) = \log_2 n$
- $H(\frac{1}{2^n}, \frac{1}{2^n}, \ldots, \frac{1}{2^n}) = n$
- $H(\frac{1}{2}, \frac{1}{2}) = 1$

Pogojna entropija

Naj bosta X in Y diskretni slučajni spremenljivki:

$$X : \begin{pmatrix} x_1 & x_2 & \dots & x_n \\ p_1 & p_2 & \dots & p_n \end{pmatrix} \qquad Y : \begin{pmatrix} y_1 & y_2 & \dots & y_n \\ q_1 & q_2 & \dots & q_n \end{pmatrix}$$

Potem je

$$H(X|Y = y_j) = -\sum_{i=1}^{n} P[X = x_i|Y = y_j] \log_2(P[X = x_i|Y = y_j])$$

pogojna entropija slučajne spremenljivke X glede na dogodek $Y = y_j$.

Pogojna entropija slučajne spremenljivke X glede na slučajno spremenljivko Y je definirana z

$$H(X|Y) = \sum_{j=1}^{n} H(X|Y = y_j) \cdot P[Y = y_j].$$

Velja:

- $H(X, Y) = H(Y) + H(X|Y)$,
- če sta X in Y neodvisni: $H(X|Y) = X$,
- $H(X, X) = 0$.

Informacija

Definicija informacije je usklajena z entropijo.

Naj bo A dogodek z verjetnostjo p. Potem je

$$I(A) = -\log_2 p = \log_2 \frac{1}{p}.$$

Enota informacije: 1 bit
(nastane, če se zgodi dogodek z verjetnostjo 1/2).

Povprečna medsebojna informacija

Povprečna medsebojna informacija slučajnih spremenljivk X in Y je definirana z

$$\begin{aligned} I(X;Y) &= H(X) - H(X|Y) \\ &= \sum_{i=1}^{n}\sum_{j=1}^{n} P[X=x_i, Y=y_j] \log_2 \frac{P[X=x_i|Y=y_j]}{P[X=x_i]}. \end{aligned}$$

- Če sta X in Y neodvisni: $I(X;Y) = 0$.
- Če $X = Y$: $I(X;Y) = H(X)$.
- $I(X;Y) = I(Y;X)$.

Kapaciteta komunikacijskega kanala

Preko komunikacijskega kanala pošiljamo simbole iz abecede Σ_1, prejmemo simbole iz abecede Σ_2.

- Poslani simboli: slučajna spremenljivka X z zalogo vrednosti Σ_1 in verjetnostmi p_i.
- Prejeti simboli: slučajna spremenljivka Y z zalogo vrednosti Σ_2.

Kapaciteta komunikacijskega kanala

$$C = \sup_{(p_i)} I(X; Y)$$

je največja informacija na simbol, ki jo lahko pošljemo skozi komunikacijski kanal.

(Enota: *bps*, kar pomeni bits per symbol)

Informacijska zmogljivost koda

Informacijska zmogljivost (angl. information rate) (n, M, d)-koda $\mathcal{C}$ je

$$r(\mathcal{C}) = \frac{1}{n} \log_2 M.$$

Za linearen $[n, k, d]$-kod nad GF(q) je to enako

$$\frac{k}{n} \log_2 q.$$

Shannonov izrek

Izrek

Naj bo C kapaciteta komunikacijskega kanala in naj bo $0 < R < C$.

Potem za vsak $\epsilon > 0$ obstaja bločni kod z informacijsko zmogljivostjo vsaj R, za katerega je verjetnost napake pri prenosu kodne besede manjša od ϵ.

Obratno, če $R > C$, potem poljubno majhna napaka ni možna.

Opomba: iz dokaza izreka se vidi, da je za majhno napako potrebno vzeti kod z zelo veliko dolžino.

Dvojiški simetrični kanal

Izrek

Kapaciteta dvojiškega simetričnega kanala z verjetnostjo napake p je enaka $C(p) = 1 - H(p, 1-p)$.

Dokaz. Če kanal oddaja simbol 0 z verjetnostjo α in simbol 0 z verjetnostjo $1 - \alpha$, potem velja

$$I(X; Y) = \Omega(\alpha + p - 2\alpha p) - \Omega(p),$$

kjer je

$$\Omega(\alpha) = \alpha \log_2 \frac{1}{\alpha} + (1-\alpha) \log_2 \frac{1}{1-\alpha}$$

binarna entropijska funkcija in

$$C = \sup_{p(x_i)} I(X; Y) = 1 - \Omega(p) = 1 - H(p);$$

supremum je dosežen pri $\alpha = 1/2$.

Primer - kapaciteta dvojiškega simetričnega kanala

Verjetnost napake p	kapaciteta $C(p) = 1 - H(p, 1-p)$.
0.01	0.919207
0.02	0.858559
0.03	0.805608
0.04	0.757708
0.05	0.713603
0.06	0.672555
0.07	0.634076
0.08	0.597821
0.09	0.56353
0.1	0.531004
0.2	0.278072
0.3	0.118709
0.4	0.0290494
0.5	0.