



Heuristické optimalizačné procesy

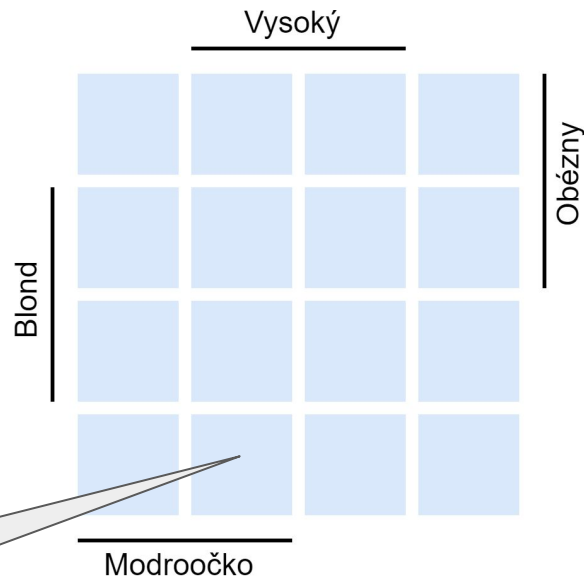
Pravdepodobnostné algoritmy

prednáška 9

Ing. Ján Magyar, PhD.
ak. rok. 2023/2024 ZS

Určenie pravdepodobnosti

Vysoký	Blond	Modroočko	Obézny
A	A	A	A
N	N	A	N
N	N	N	A
...



Počet(Vysoký=A,Blond=N,Modroočko=A,Obézny=N)

Počet(všetky osoby)

Združená distribúcia pravdepodobnosti

základná distribúcia

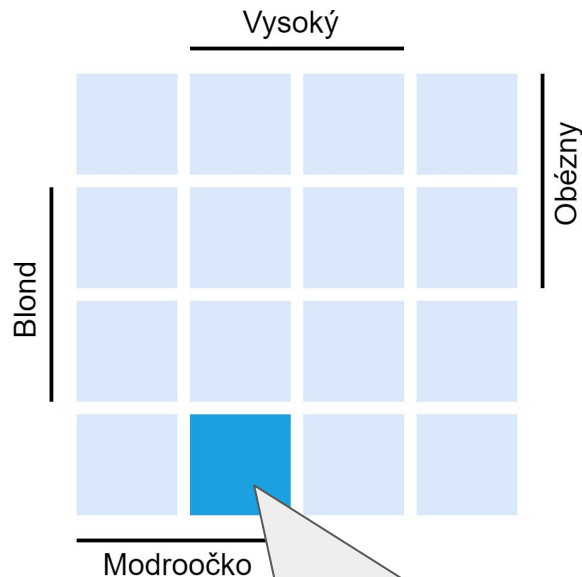
ostatné sa z nej odvodzajú

uvažujú sa jednotlivé polia
(všetky premenné)

$$p(V=.,B=.,M=.,O=.)$$

počet prvkov distribúcie: $2^4 = 16$

$$\text{platí } \sum_V \sum_B \sum_M \sum_O p(V,B,M,O) = 1$$



$$p(\text{Vysoký}=A, \text{Blond}=N, \text{Modroočko}=A, \text{Obézny}=N)$$

Marginálna distribúcia pravdepodobnosti

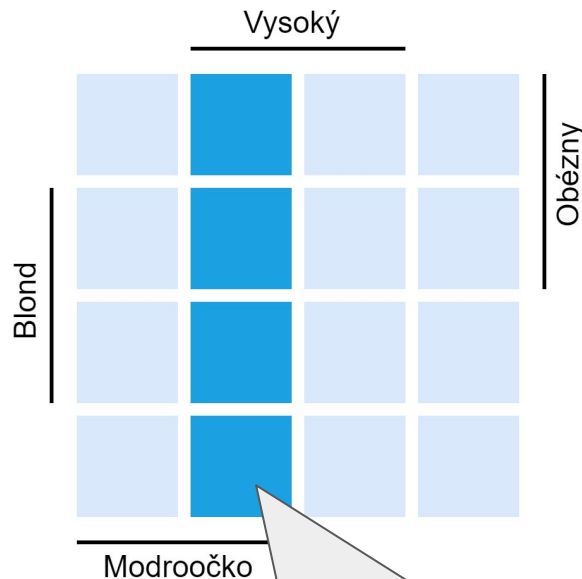
uvažujú sa oblasti polí
(bez niektorých premenných)

marginalizácia

$$p(V=.,M=.) = \sum_B \sum_O p(V=.,B,M=.,O)$$

marginálna distribúcia má
znížený počet prvkov

$$\text{platí } \sum_V \sum_M p(V,M) = 1$$



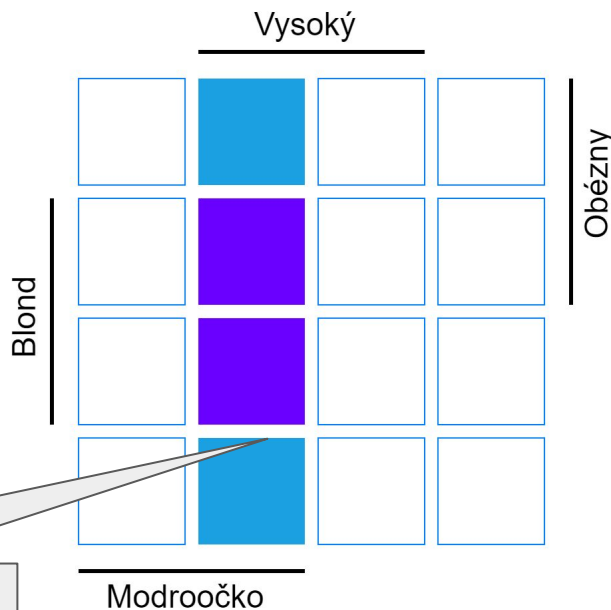
Podmienená distribúcia pravdepodobnosti

uvažujú sa oblasti polí
v oblastiach polí

marginalizácia + normalizácia

$$\begin{aligned} p(B=.|V=.,M=.) \\ &= p(V=.,B=.,M=.) / p(V=., M=.) \\ &= \Sigma_O p(V=.,B=.,M=.,O) / \\ &\quad \Sigma_B \Sigma_O p(V=.,B,M=.,O) \end{aligned}$$

$$p(\text{Blond}=A | \text{Vysoký}=A, \text{Modroočko}=A)$$



platí $\Sigma_B p(B|V=.,M=.) = 1$

$$p(V=.,B=.,M=.) = p(B=.|V=.,M=.) p(V=.,M=.)$$

Nezávislosť premenných

nezávislosť M a V od B a O

- distribúcia $p(M,V)$ nezávisí od hodnôt B a O
- rovnaká pre každú kombináciu hodnôt B a O

platí $p(M=., V=.|B=., O=.) = p(M=., V=.)$

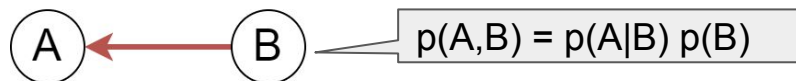
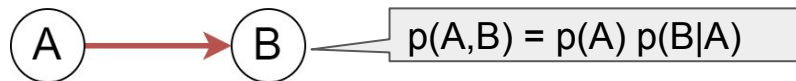
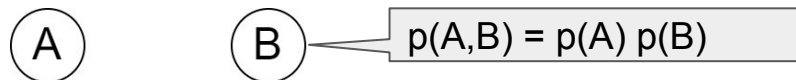
dôsledok $p(A,B) = p(A|B) p(B) = p(B|A) p(A) = p(A) p(B)$

		Vysoký				Obézny
Blond						
		Modroočko				

Faktorizácia združenej distribúcie

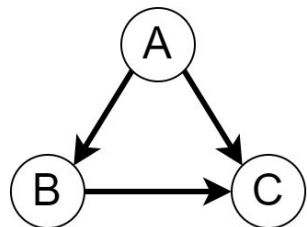
rozloženie na menšie distribúcie

späťne získaná združená distribúcia násobením menších distribúcií

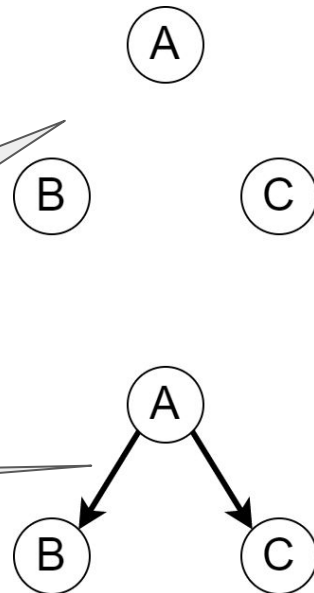


$p(A,B,C) = p(A) p(B) p(C)$

$p(A,B,C) = p(A) p(B|A) p(C|A)$

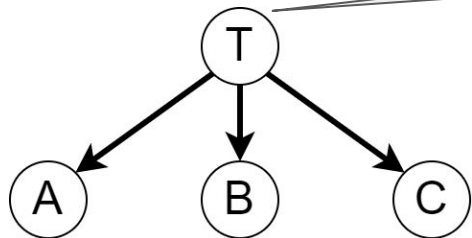


$p(A,B,C) = p(A) p(B|A) p(C|A,B)$



Naivný Bayesov klasifikátor

$$p(A,B,C,T) = p(A|T) p(B|T) p(C|T) p(T)$$



$$p(T|A,B,C) = p(A,B,C,T) / p(A,B,C) =$$

$$\frac{p(A|T) p(B|T) p(C|T) p(T)}{\sum_T p(A,B,C,T)}$$

=

$$\frac{\prod_{X=A,B,C} p(X|T)}{\sum_T p(T) \prod_{X=A,B,C} p(X|T)} p(T)$$

$$\operatorname{argmax}_T p(T) \prod_{X=A,B,C} p(X|T)$$

Estimation of distribution algorithms

pravdepodobnostný model

- pravdepodobnostná distribúcia zložiek

- učenie parametrov (a štruktúry) modelu

populačný algoritmus: *populácia vs. model*

- populácia

 - náhodné generovanie na základe modelu

 - vzorkovanie priestoru kandidátov

- použitie populácie ako učiacej množiny

 - tvorba / vylepšenie modelu

Štruktúra EDA

input: π

output: $r \in S$

$k = 0$

$M = \text{init}()$

$P_k = \{b_1(k), \dots, b_s(k)\} = \{URP(), \dots, URP()\} / \{M(), \dots, M()\}$

$g(P_k) = \{g(b_1(k)), \dots, g(b_s(k))\}$

repeat

$D' = \text{vyber}(P_k, g(P_k))$

$M = \text{odhadni/uprav}(M, D')$

$D'' = \text{generuj}(M)$

$g(D'')$

$P_{k+1} = \text{vyber}(P_k + D'')$

$k = k + 1$

until $\text{term}()$

return $r = \text{best}$

Typy modelov

nemôže byť priamo združená pravdepodobnostná distribúcia

univariantné

žiadne vzájomné závislosti medzi premennými

$$p(x_1, \dots, x_n) = \prod_{i=1, \dots, n} p(x_i)$$

bivariantné

vzájomné závislosti existujú iba medzi dvojicami premenných

multivariantné

vzájomné závislosti existujú medzi skupinami premenných

Univariate Marginal Distribution Algorithm

model

- v každej generácii sa vytvára nový model

- iniciálny model sa nevytvára

- marginálne pravdepodobnosti odhadované z frekvencií

populácia

- generačná náhrada

- D' vyberaná metódami z evolučných algoritmov

Štruktúra UMDA

input: π

output: $r \in S$

$k = 0$

~~$M = \text{init}()$~~

$P_k = \{b_1(k), \dots, b_s(k)\} = \{URP(), \dots, URP()\}$ ~~$\{M(), \dots, M()\}$~~

$g(P_k) = \{g(b_1(k)), \dots, g(b_s(k))\}$

repeat

$D' = \text{vyber}(P_k, g(P_k))$

$M = \text{odhadni/uprav}(M, D')$

$D'' = \text{generuj}(M)$

$g(D'')$

$P_{k+1} = \text{vyber}(P_k, D'')$

$k = k + 1$

until $\text{term}()$

return $r = \text{best}$

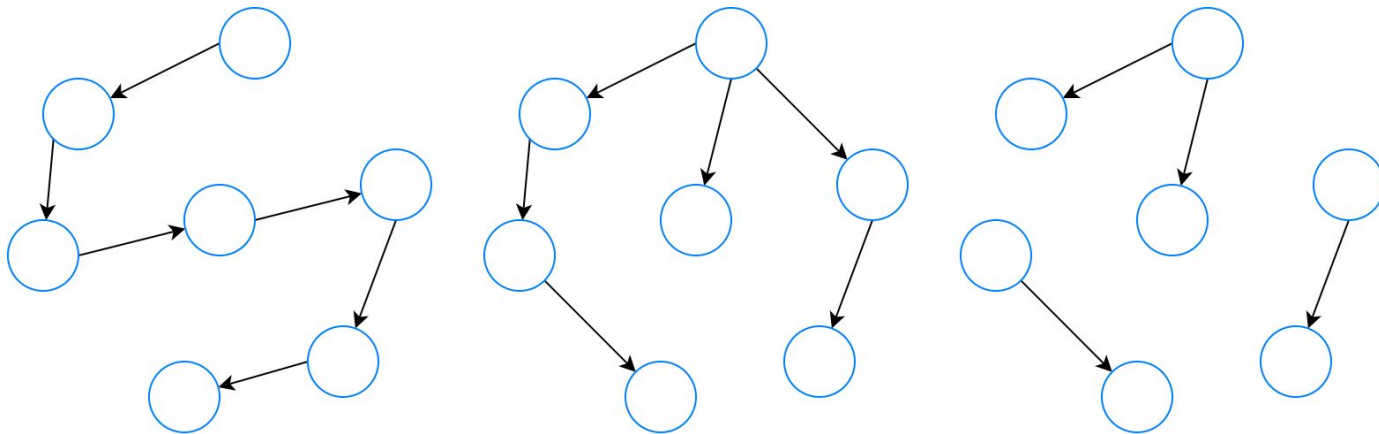
Bivariantné modely

špeciálne tvary závislostí

rôzne úrovne všeobecnosti

reťazec je špeciálnym prípadom stromu

strom je špeciálnym prípadom lesa stromov



Mutual-Information-Maximizing Input Clustering

model

združená distribúcia $p(x_1, \dots, x_n) = p(x_1) \prod_{i=2, \dots, n} p(x_i | x_{i-1})$

učenie štruktúry - greedy prístup

koreň: premenná s maximálnou entropiou

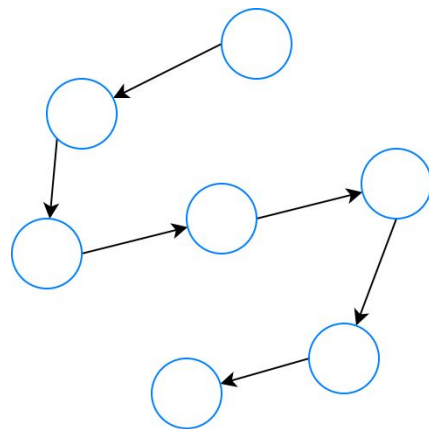
predlžovanie reťazce: premenná minimalizujúca podmienenú entropiu

populácia

generovanie premenných v poradí podľa reťazce

pre učenie sa používa zvyšujúci sa percentil

noví kandidáti sa pridávajú k už generovaným



Štruktúra MIMIC

input: π

output: $r \in S$

$k = 0$

~~$M = \text{init}()$~~

$P_k = \{b_1(k), \dots, b_s(k)\} = \{URP(), \dots, URP()\}$ ~~$\{M(), \dots, M()\}$~~

$g(P_k) = \{g(b_1(k)), \dots, g(b_s(k))\}$

repeat

$D' = \text{vyber}(P_k, g(P_k))$

$M = \text{odhadni/uprav}(M, D')$

$D'' = \text{generuj}(M)$

$g(D'')$

$P_{k+1} = \text{vyber}(P_k + D'')$

$k = k + 1$

until $\text{term}()$

return $r = \text{best}$

Bivariate Marginal Distribution Algorithm

model

združená distribúcia $p(x_1, \dots, x_n) = \prod_{k \in R} p(x_k) \prod_{j \in V \setminus R} p(x_j | x_{r(j)})$

učenie štruktúry $\{V, E, R\}$ - greedy prístup

na základe univariantných a bivariantných frekvencií

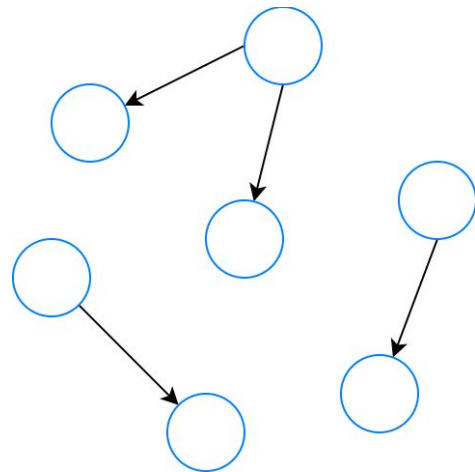
učenie parametrov

populácia

generovanie premenných v poradí podľa závislostí

pre učenie sa používa lepšia časť populácie

noví kandidáti nahrádzajú časť populácie



BMDA - závislostný graf

závislosť a nezávislosť

detekcia nezávislosti: $p(\mathbf{x}_i | \mathbf{x}_j) = p(x_i, x_j) / p(x_j) = p(x_i) \quad p(x_j) / p(x_j) = p(\mathbf{x}_j)$

kritérium

$$X^2 = \sum (\text{pozorované} - \text{očakávané})^2 / \text{očakávané}$$

$$\text{pozorované} = N p(x_i, x_j) \quad \text{očakávané} = N p(x_i) p(x_j)$$

tvorba závislostného grafu:

1. určenie kritéria a nezávislosti dvojíc premenných
2. náhodný výber koreňa stromu
3. iteračné vybudovanie stromu
4. ak ešte nespracovaná premenná, krok 2

Štruktúra BMDA

input: π

output: $r \in S$

$k = 0$

~~$M = \text{init}()$~~

$P_k = \{b_1(k), \dots, b_s(k)\} = \{URP(), \dots, URP()\}$ ~~$\{M(), \dots, M()\}$~~

$g(P_k) = \{g(b_1(k)), \dots, g(b_s(k))\}$

repeat

$D' = \text{vyber}(P_k, g(P_k))$

$M = \text{odhadni/uprav}(M, D')$

$D'' = \text{generuj}(M)$

$g(D'')$

$P_{k+1} = \text{vyber}(P_k) + D''$

$k = k + 1$

until $\text{term}()$

return $r = \text{best}$

otázky?