



Heuristické optimalizačné procesy



Pravdepodobnostné algoritmy

Marian.Mach@tuke.sk

<http://neuron.tuke.sk/~machm>

November, 2016

EDA

- Pravdepodobnostný model
 - pravdepodobnostná distribúcia zložiek
 - učenie
 - parametrov modelu
 - štruktúry modelu (ak nie je pevne daná)
- 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: $r \in S$
- $k=0$
- $M = \text{init}()$
- $P_k = \{b_1(k), \dots, b_s(k)\} = \{\text{URP}(), \dots, \text{URP}()\} / \{M(), \dots, M()\}$
- $g(P_k) = \{g(b_1(k)), \dots, g(b_s(k))\}$
- **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
- Typy
 - 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

UMDA

- 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: $r \in S$
- $k=0$
- ~~$M = \text{init}()$~~
- $P_k = \{b_1(k), \dots, b_s(k)\} = \{\text{URP}(), \dots, \text{URP}()\} / \{\text{M}(), \dots, \text{M}()\}$
- $g(P_k) = \{g(b_1(k)), \dots, g(b_s(k))\}$
- **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** term()
- **return** $r = \text{best}$

PBIL

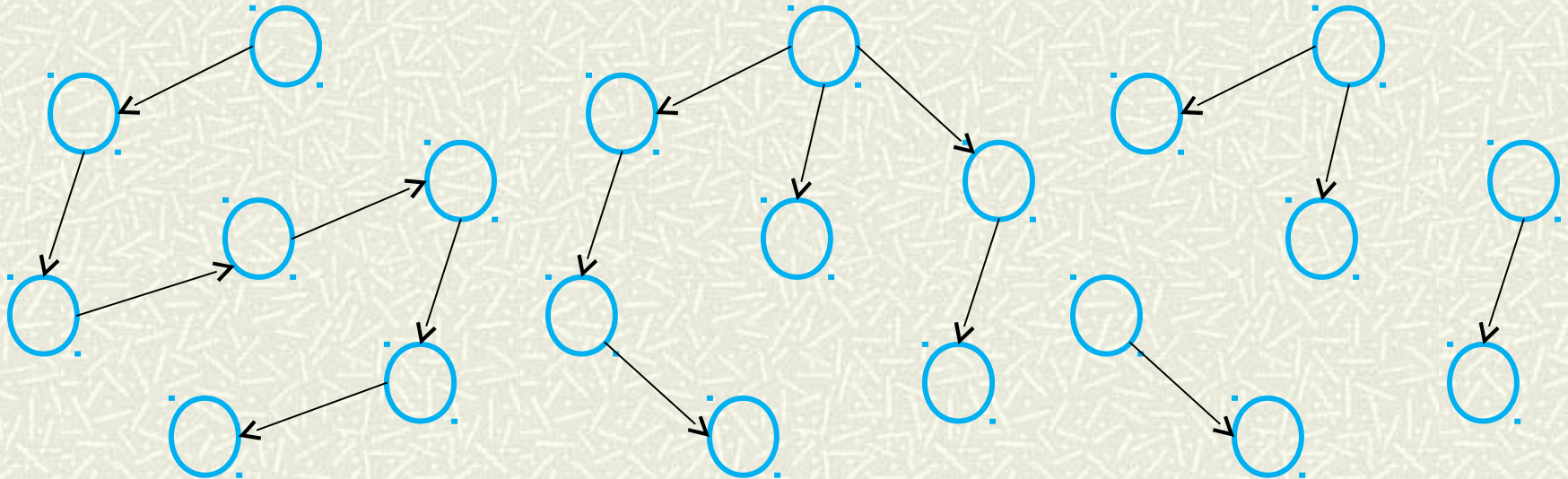
- Population-Based Incremental Learning
- Model
 - inicializovaný na neutrálne hodnoty (0.5)
 - úprava modelu voči najlepšiemu kandidátovi X
 - $M_i = M_i (1 - \alpha) + \alpha X_i$
 - mutácia modelu
 - $M_i = M_i (1 - \beta) + \beta \text{rand}()$
- Populácia
 - kandidát má tvar binárneho reťazca
 - populácia slúži iba na výber jedného kandidáta
 - generačná náhrada

Štruktúra PBIL

- input: π
- output: $r \in S$
- $k=0$
- $M = \text{init}(0.5)$
- $P_k = \{b_1(k), \dots, b_s(k)\} = \{\text{URP}(), \dots, \text{URP}()\} / \{M(), \dots, M()\}$
- $g(P_k) = \{g(b_1(k)), \dots, g(b_s(k))\}$
- **repeat**
- $D' = \text{vyber_best}(P_k, g(P_k))$
- $M = \text{odhadni/uprav}(M, D')$
- **if** $\text{rand}() < p_{\text{mut}}$ **then** $M = \text{mutuj}(M)$
- $D'' = \text{generuj}(M)$
- $g(D'')$
- $P_{k+1} = \text{vyber}(P_k \pm D'')$
- $k = k + 1$
- **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



MIMIC

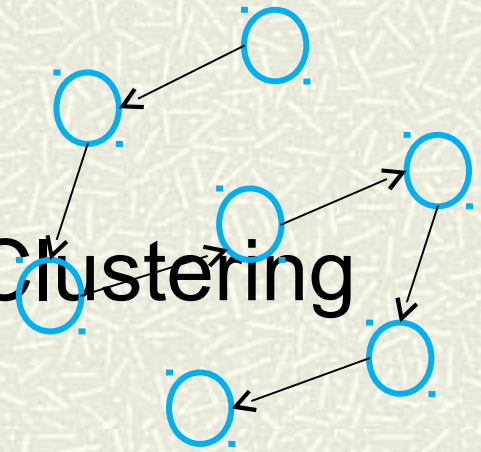
▪ Mutual Inform. Maximizing Input Clustering

▪ 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 minimálnou entropiou
 - predlžovanie reťaze: premenná minimalizujúca podmienenú entropiu

▪ Populácia

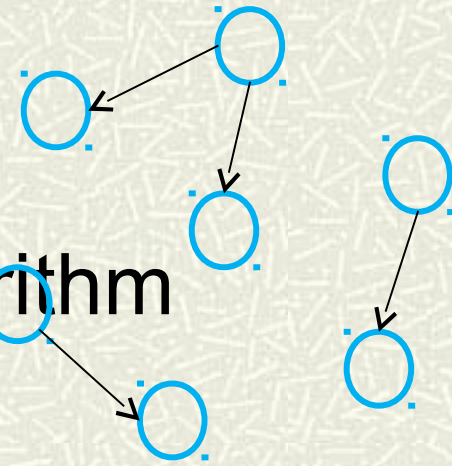
- generovanie premenných v poradí podľa reťaze
- pre učenie sa používa zvyšujúci sa percentil
- noví kandidáti sa pridávajú k už generovaným



Štruktúra MIMIC

- input: π
- output: $r \in S$
- $k=0$
- ~~$M = \text{init}()$~~
- $P_k = \{b_1(k), \dots, b_s(k)\} = \{\text{URP}(), \dots, \text{URP}()\} / \{\text{M}(), \dots, \text{M}()\}$
- $g(P_k) = \{g(b_1(k)), \dots, g(b_s(k))\}$
- **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** term()
- **return** $r = \text{best}$

BMDA



- 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

Štruktúra BMDA

- input: π
- output: $r \in S$
- $k=0$
- ~~$M = \text{init}()$~~
- $P_k = \{b_1(k), \dots, b_s(k)\} = \{\text{URP}(), \dots, \text{URP}()\} / \{\text{M}(), \dots, \text{M}()\}$
- $g(P_k) = \{g(b_1(k)), \dots, g(b_s(k))\}$
- **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** term()
- **return** $r = \text{best}$

BMDA – závislostný graf

- Závislosť a nezávislosť
 - detekcia nezávislosti
 - $p(x_i|x_j) = p(x_i, x_j)/p(x_j) = p(x_i)p(x_j)/p(x_j) = p(x_i)$
 - kritérium
 - $\chi^2 = \sum (\text{pozorované} - \text{očakávané})^2 / \text{očakávané}$
 - $\text{pozorované} = N p(x_i, x_j)$ $\text{očakávané} = N p(x_i)p(x_j)$
- 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