1 STROINO UCENIE

1.1 Problemski prostor, ocenjevanje znanja

1.2 Evalviranje hipotez

Pomembni kriteriji:

- · konsistentnost hipotez z primeri (ucnimi)
- · splosnost (tocnost za nevidene primere)
- · razumljivost hipotez

TP=true positive, TN-true negative, FP-false positive (napaka 1. tipa), FNfalse negative (napaka 2. tipa)

Klasifikacijska tocnost = $\frac{TP+TN}{TP+TN+FP+FN} = \frac{TP+TN}{N}$ Obcutljivost/senzitivnost = $TPR = \frac{TP}{TP+FN}$

1.3 GRADNIA ODI OCITVENIH DREVES

Za koliko se entropija zmanjsa po delitvi z Atributom A:

Informacijski prispevek (najbolj informativni atribut maksimizira informacijski prispevek minimizira Ires:

Gain(A) =
$$H(A) - H_{res}(A)$$

 $H_{res}(A) = -\sum_{a_i \in A} p(A = a_i)$

$$H_{\text{res}}(A) = -\sum_{a_i \in A} p(A=a_i) \sum_{c_i \in C} p(C=c_i|A=a_i) \log_2 p(C=c_i|A=a_i)$$

Razmerje inofrmacijskega prispevka atributa A:

$$IGR(A) = \frac{Gain(A)}{H(A)}$$

1.3.1 TDIDT (Top down induction decision tree) algoritem

Pozresen algoritem, ki lokalno izbira najbolsi atribut.

- kratkoviden algoritem

1.3.2 BINARIZACIJA ATRIBUTOV

Aleternativa za resevanje problematike z vecvrednostnimi atributi:

Strategije (za primer $B = \{Y, G, R, B\}$):

- [{Y},{R,G,B}] (one-vs-all)
- [{Y,R},{G,B}]
- · vpeljava bianrnih atributov za vsako barvo

Primer B = {Y, G, R}, konstruiramo 3 nove binarne atribute:

barva	Y	G	R	
Y	1	0	0	Prednost: manjse vejanje drevesa.
$G \rightarrow$	0	1	0	rrednost: manjse vejanje drevesa.
R	0	0	1	

1.4 Ucenje iz sumnih podatkov (rezanje)

tocnost t...verjetnost pravilnosti klasifikacije

napaka e ... 1-t

relativna frekvenca
$$p = \frac{n}{N}$$

m-ocena
$$p = \frac{n + p_a * m}{N + m}$$

 $m\dots$ koliko zaupam apriorni verjetnosti

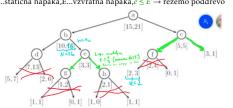
pa apriorna verjetnost (pove domenski ekspert) (v nasem primeru relativna frekvenca)

Laplacova ocena verjetnosti $p = \frac{n+1}{N+k}$

k...stevilo vseh moznih razredov

1.4.1 MEP (MINIMAL ERROR PRUNNING)

e...staticna napaka,E...vzvratna napaka, $e \le E \rightarrow$ rezemo poddrevo



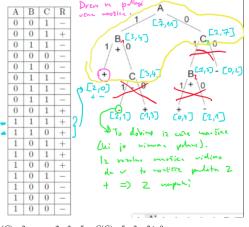
(Laplace)

$$\begin{array}{l} e_L(d) = 1 - t = 1 - \frac{13 + 1}{20 + 2} = 0.363 \\ E_L(d) = 12/20 \cdot e_L(d_l) + 8/20 \cdot e_L(d_d) = \frac{12}{20} \cdot (1 - \frac{7 + 1}{12 + 2}) + \frac{8}{20} (1 - \frac{13 + 1}{20 + 2}) \end{array}$$

1.4.2 REP (REDUCED ERROR PRUNNING)

Ucna mnozica: 70% za gradnjo, 30% za rezanje (z rezanjem odstranimo poddrevesa, ki niso kriticna in so redundantna tako zmansamo velikost drevesa) G(v)=st. napacnih klasifikacij v poddrevesu - st. napacnih klasifikacij v korenu poddrevesa

 $G(v) \ge 0 \Rightarrow \text{rezemo podrevo}$



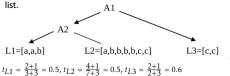
$$e(C) = 3$$
, $e_T = 2 + 3 = 5$, $G(C) = 5 - 3 = 2 \ge 0 \rightarrow \text{rezemo}$

1.5 Ocenievanie uspesnosti modelov

tocnost t ... verjetnost pravilnosti klasifikacije

Laplacova ocena verjetnosti $p = \frac{n+1}{N+k}$

k...stevilo vseh moznih razredov





$$e = 1 - (P(B = 0)P(R = 0|B = 0) + P(B = 1)P(R = 1|B = 1))$$

1.6 OBRAVNANVA MANKAJOCIH ATRIBUTOV, NAVINI BAYESOV KLASIFIKATOR

1.6.1 NAIVNI BAYES

Ce poznamo razred, kam klasificiramo ce nepoznamo atributov:

Klasifikator:
$$\operatorname{argmax}_{c \in C} P(c) \prod_{i=1}^{n} P(x_i|c)$$

c...razred, $x_i...$ atributi

Verietnost::

$$P(C = c | x_1, ..., x_n) = \frac{P(C = c)P(X_1 = x_i | C = c)P(X_2 = x_j | C = c)...}{P(X_1 = x_i)P(X_2 = x_j)...}$$

Primer pritisk, vreme, razred glavobol

$X \backslash Y$	ne	da
P(G)	$P(G = da) = \frac{5}{9}$	$P(G = ne) = \frac{4}{9}$
P(P=srednji)	$P(P = sre. G = ne) = \frac{1}{5}$	$P(P = sre. G = da) = \frac{2}{4}$
P(V=dezevno)	$P(V = dez G = ne) = \frac{1}{5}$	$P(V = dez G = da) = \frac{2}{4}$
$P(y) \prod_{i=1}^{n} P(x_i y)$	$\frac{5}{9} \cdot \frac{1}{5} \cdot \frac{1}{5}$	$\frac{4}{9} \cdot \frac{2}{4} \cdot \frac{2}{4}$

1.6.2 Nomogragmi

Ciljni razred $C = c_7$

 $X_{X_i = x_j} = \ln\left(\frac{P(X_i = x_j | C = c_T)}{P(X_i = x_j | C = \overline{c_T})}\right)$

1.7 K-najblizjih sosedov



1.8 NADZOROVANO UCENJE (SUPERVISED LEARNING)

Ucni primeri so podani/oznaceni kot vrednosti vhodov in izhodov.

$$(\vec{x}_1, \vec{y}_1), (\vec{x}_2, \vec{y}_2), \dots, (\vec{x}_N, \vec{y}_N)$$

 \vec{x}_i ... atributi, \vec{y}_i ... ciljna spremenljivka

Locimo dve vrsti problemov:

- 1. Klasifikacijski problemi yi diskretna
- 2. Regresijski problemi y i zvezna

1.8.1 Lokalno utezena regresija

$$u(\vec{x}_{?}) = \frac{\sum_{i=1}^{K} w_{i} \cdot f(\vec{x}_{i})}{\sum_{i=1}^{K} w_{i}}, w_{i}(d)...utez$$

A	В	С	dolžina	d(xiixi)	Wi	We-f(-xi	2
0	0	0	9	4	115	9/5	
0	0	0	10	4	115	2	
0	1	1	9	2_	113	3	1. () S. w f(x;)
0	2	0	12	2	113	4	h(x,)= \$ w. f(x,)
0	2	1	12	1	112	6	4
1	0	0	12	3	1/4	3	20
1	0	0	15	3	114	15/4	$= \frac{20}{45} = 11.359$
1	1	1	11	1	112	MIL	15
1	1	1	15	1	112	15/2	
1	1	1	9	1	112	912	
1	2	0	9	1	112	512	
1	2	1	12	0	1	12	
					76	1111	

- Manhattansko razdaljo za m jedrno funkcijo $w_i = \frac{1}{1+d_{i1}}$

1.8.2 Regresijska drevesa

Linearna regresija je poseben primer regresijskega drevesa.

V listih regresijskega drevesa vcasih napovemo kar povprecno vrednost.

1.9 Nenadzorovano ucenje (unsupervised learning)

Ucni primeri niso oznaceni (nimajo ciljne spremenljivke), ucimo se vzorcev v podatkih, (npr. grucenje)

1.9.1 HIERARHICNO GRUCENJE

Poveze po podobnosti med primeri, primer zacne kot samostojna gruca, na koncu vsi primeri pripadajo eni gruci

Dendrogram: drevo, ki predstavlja grucenje.

Single-linkage: povezava med grucami je najkrajsa razdalje med primeroma iz razlicnih gruc.

Complete-linkage: najdaljsa razdalja med primeroma iz razlicnih gruc. Average-linkage: povprecna razdalja med primeroma iz razlicnih gruc.

Tocke A(3,1),B(1,2),C(3,4),D(5,2),E(1,1), manhattan, complete linkage:

		A	В	C	D	E			A	BE	С	D
ı	A	0	3	3	3	2	_		Α	DE		D
l		"		4	4	-		A	0	3	3	3
l	В		0	4	4	1	\rightarrow	BE		0	5	5
l	C			0	4	5				Ü		
l	D				0	5		C			0	4
ı					U	5		D				0
l	E					0			ı			

		ABE	C	D			ABE	CD
_	ABE	0	5	5			ADE	<u>CD</u>
\rightarrow			0	4	\rightarrow	ABE	0	5
	C		U	4		CD		0
	D			0			I	

1.9.2 K-MEANS

- 1. V prostor dodamo k centroidov, ki predstavljajo gruce.
- 2. Izracunamo ketri centroid je najblizji vsakemu primeru.
- 3. Izracunamo nove centre gruc = $\frac{1}{|G|} \sum_{i \in G} x_i$
- 4. Ponovimo korake 2 in 3 dokler se centri ne premaknejo.

V mnozici tock A(3,1),B(1,2),C(3,4),D(5,2),E(1), manhattanska razdalja, zacetni vrednosti centroidov C1(4,4) in C2(5,4).

Tocka	d(X,C1)	d(X,C2)	Gruca
A	4	5	C1
В	5	6	C1
С	1	2	C1
D	3	2	C2
E	6	7	C1

V naslednji iteraciji sta koordinati centroidov:

$$C1 = (\frac{3+1+3+1}{4}, \frac{1+2+4+1}{4}) = (2,2)$$
 in $C2 = D = (5,2) \dots$

1.10 Spodbujevalno ucenje - reinforcement learning

Inteligentni agent se uci iz zaporedja nagrad in kazni

1.11 Ocenjevanje ucenja

k-fold, celo ucno mnozico razbij na k disjunktnih podmnozic za vsako od k podmnozic uporabi mnozico kot testno mnozico, preostalih k-1 mnozic kot ucno mnozico.

2 Preiskovanie

2.1 Neinformirani preiskovalni algoritmi

2.1.1 ISKANJE V SIRINO

2.1.2 ISKANJE V GLOBINO

Izboljsave (Iskanje s sestopanjem, iterativno poglabljanje)

2.1.3 ITERATIVNO POGLABLIANIE

problem gobinsko omejenega iskanja -> nastavitev meje l Mejo l postopoma povecujemo za 1, dokler ne najdemo resitve.

- popolnost: Da
- · optimalnost: Da
- casovna zahtevnost O(b^d)
- prostorska zahtevnost O(bd)

Boljse od iskanja v globino/sirino

2.1.4 DVOSMERNO ISKANJE

Pozenemo vzporedni iskanji od zacetka do cilja in od cilja do zacetka.

Implemenatcija dvosmernega iskanja:

- · ciljno vozlisce mora biti znano
- originalni problemski prostor preslikamo v dvosmerni prosto stanj E1, E2 dosegljiv iz E in S1,S2,S3 dosegljiv iz S (S,E) -> (S1, E1), (S1,E2), (S2, E1) (S2, E2)... Vozlisce (Si, Ei) je v dvosmernem prostur ciljo vozlisce ce velja E=S (soda dolzina na isto mesto pridemo iz obeh strani) ali S->E (liha pot sosednja)

2.2 Informirani preiskovalni algoritmi

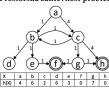
Ideja: preiskovanje usmerjamo z dodatnim znanjem hevristiko (ocenitvena funkcija za obetavnost vozlisca)

- optimisticna/dopustna: $\forall n : h(n) \le h^*(n)$ (h^* je optimalna ocena)

optimalna: $h(n) = h^*(n)$ pesimisticna: $h(n) \ge h^*(n)$

A* is informed version of dijkstra (uses heuristics and pq), ce h(dopustna)=popolna in optimalna

Casovna zahtevnost odvisna od hevristike: $E = (h^* - h)/h^*$, $O(b^E \cdot d)$, b-stopnja vejanja, d-globina optimalne resitve



f(n) = g(n) + h(n), g(n) cena do vozlisca, h(n) hevristika

Razvijamo dokler ne pridemo do ciljnega vozlisca

Razvijano	Generirana	Priority Queue
/	a(4)	[]
a	b(7) c(6)	[c(6), b(7)]
c	b'(11) g(12) h(8)	[b(7),h(8),b'(11),g(12)]
b	c'(4) d(8) e(5)	[c'(4),e(5),h(8),d(8),b'(11),g(12)]
f		

2.2.2 IDA^* (Iterative deepening A^*)

f(n) = g(n) + h(n), g(n)=cena poti do n

Meja	Razvijano	Generirana	DFS (list)
0	/	s(7)	/
7	/	s(7)	s
	s	a(8) b(7) c(7)	b, c
	ь	f(6) h(5)	f h c
	f	g(7) h(9) i(11)	g h c
	<u>g</u>		

2.2.3 Kakovost hevristicnih funkcij

Kakovost h ocenimo z stevilom generiranih vozlisc ter efektivnim faktorjem vejanja (N vozlisc je algoritem generiral da je na globini d nasel resitev) Hocemo imeti dopustne hevristike s cim visjimi vrednostmi in sprejelmjivo ceno (casom izracuna)

Ce $h_2(n) \ge h_1(n)$, $\forall n$ potem h_2 **dominira** h_1

2.3 preiskovanje grafov AND/OR, nedeterministicno okolje

Pomagajo resevati probleme z **dekompozicijo na manjse probleme** Uporabnost:

- · princip deli in vladaj
- · iskanje v nedeterministicnih okoljih
- igre med dvema nasprotnikoma s popolno informacijo (sah, dama)
- ekspertno resevanje problem

2.3.1 AO*

- posplositev A* na grafe AND/OR
- popoln in optimalen ⇔ h(n) ne precenjuje dejanske cene do cilja

 ${\cal F}(N)$ ocena za usmerjanje preiskovanja, ${\cal H}(N)$ dinamicna hevristicna ocena Postopek:

- 1. Razvij najcenejse vozlisce
- ce list in koncno (oznaci), preveri 3. korak, nadaljuj v 1.
- ce list in ni koncno (oznaci) vrednost vozlisca = ∞
- 2. Posodobi vse predhodnike
- v AND starsih, cena starsa = ∑ sinov + povezava v
- v OR starsih, cena starsa = min(sinovi) + povezava v
- Koncaj ko obstaja pot od zacetnega vozlisca, po kateri v AND vozliscih po vseh sinovih prides do cilja, v OR vozliscih v vsaj enem

2.3.2 ALGORITEM MINIMAX

 $O(n^d)$

2.3.3 REZANIE ALFA-BETA

V najblsem primeru zmansa iz $O(b^{m \cdot d})$ na $O(b^{m \cdot d/2})$

2.4 Lokalno preiskovalni algoritmi

plezanje na hrib, simulirano ohlajanje, gen.algoritmi...

2.4.1 Lokalno iskanje v snopu

generiraj k nakljucnih zacetnih stanj

iz vsakegea generiraj sosede

izberi k najboljsih naslednikov

ponavljaj (iz maksimum stohasticno iskanje -> 1-verjetnost/sumvseh)

3 PLANIRANJE

plan zaporedje akcij, ki pripelje od zacetnega do koncnega stanja

3.1 Planiranje s sredstvi in cilji (STRIPS)

Agentu opisemo svet in postavimo fizikalne omejitve.

Ne zagotovalja optimalne resitve, obravnavamo le en cilj naenkrat (ko ga dosezemo, se lahko ostali izgubijo) = Sussmanova anomalija

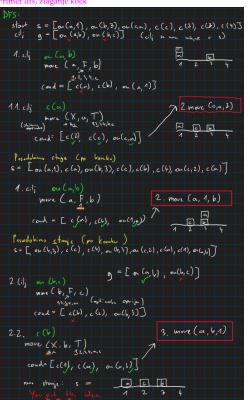
Akcija move(X, From, To)

- pogoj: $cond=[clr(X), on(X,F), clr(T)] \rightarrow pogoji za izvajanje akcije,$
- poz. ucinki: adds=[on(X, T), clr(F)] → nova stanja,
- neg. ucinki: dels=[on(X, F), clr(T)] → izbrisana stanja,
- omejitve: constr=[F ≠ T, X≠ F, X≠ T, block(X)] → omejitve akcij (fizikalne omejitve),

Algoritem:

- 1. Izberi se neresen cilj iz mnozice CILJEV
- 2. Izberi akcijo, ki izbrani cilj doda v stanje
- 3. Omogoci izbrano akcijo (izpolni pogoje)
- 4. Izvedi akcijo (ki izopolni najvec pogojev)
- Ce obstajajo nereseni cilji ⇒ 1.

Primer dfs, zlaganje kod



3.2 Planiranje z regresiranjem ciljev (STRIPS)

Resitev za sussmanovo anomalijo

Zacnemo v ciljih, regresiramo do zacetka ($G_i \subset S_0$):

1. $G_{i+1} = G_i \cup \operatorname{cond}(A) - \operatorname{adds}(A)$

- 2. **POGOJ**: $G_i \cap \text{dels}(A) = \emptyset$
- 3. Preveri da ni protislovja (npr. $G_{i+1} = [on(b,c),...,c(c)...]$)
- \rightarrow zactno stanje = [on(a,1), on(b,a), c(b), on(c,3), c(c)]
- \rightarrow hocemo da zacetno_stanje $\subset G_i$
- 1. $G_0 = [on(a,b), on(b,c)]$
 - on(a,b): $A_0 = move(a, From, b)$
 - From = 1
 - POGOJ: $G_0 \cap dels(A_0) = \emptyset \checkmark$
 - $G_1 = [on(a,b), on(b,c), c(a), c(b), on(a,1)] [c(1), on(a,b)] \checkmark$
- 2. $G_1 = [on(b,c),c(a),c(b),on(a,1)]$
 - c(a): $A_1 = move(X, a, To)$
 - X = c, To = 2
 - POGOJ: $G_1 \cap dels(A_1) = \emptyset \checkmark$
 - $G_2 = [\underline{\text{on}(b,c)},c(a),c(b),\text{on}(a,1),\underline{c(c)},c(2),\text{on}(c,a)]$ -[c(a), on(c,2)] X(protislovje)
 - on(b,c): $A_2 = move(b, From, c)$
 - From = 3
 - POGOJ: $G_2 \cap dels(A_2) = \emptyset \checkmark$
 - $G_2 = [on(b,c),c(a),c(b),on(a,1),c(c),c(b),on(b,3)] \checkmark$
- 3. $G_2 = ...$

3.3 RAZPOREJANJE OPRAVIL PDDL)

Razsirimo lahko notacijo (PDDL):

Akcija1 < Akcija2: Akcija1 se mora zgoditi pred Akcijo2

Resources podajo stevila razpolozljivih resursov

DURATION opredejljuje trajanje posamezne akcije

CONSUME opredeljuje (trajno) porabo dolocene kolicine resursov

USE opredeljuje (zacasno) zasedenost kolicine resursov med izvajanjem akcije

Metoda kriticne poti

kriticna pot: pot, ki je **najdaljsa** in doloca dolzino trajanja celotnega plana vsaki akciji priredimo par [**ES**, **LS**]

- ES: najbolj zgodnji mozen zacetek (Earliest Start)
- $\quad ES(start) = 0, \quad ES(B) = \max_{A < B} [ES(A) + Duration(A)]$
- LS: najbolj pozen mozen zacetek (Latest Start)
- $\quad \mathsf{LS}(\mathsf{Finish}) = \mathsf{ES}(\mathsf{Finish}), \, LS(A) = \min_{A < B} \left[LS(B) Duration(A) \right]$

rezerva(slack)=LS-ES (casovna rezerva) Algoritem po hevristiki minimum slack → na vsaki iteraciji ima prednost akcija ki ima izpolnjene predhodnike in najnizji slack, nato posodobi [ES in LS] za celotni graf in ponovi.

3.4 Preiskovanje brez informacije o stanju

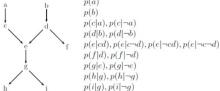
belief states - graf, zacetna stanja = potencna mnozica vseh stanj, konec samo koncno stanje

4 SKLEPANJE

4.1 Bayesovske mreze

Baye. mreza = Usmerjen graf, kjer so podane zahtevane verjetnosti:

- Za vozlisca brez starsev verjetnosti P(v_i)
- Za vozlisca z starsi pogojne verjetnosti vseh kombinacij starsev



Pravila ver-

jetnostnega sklepanja:

1. **Konjunkcija**: $P(X_1 X_2 | C) = P(X_1 | C)P(X_2 | X_1 C)$

•
$$P(A_1 \cap \dots \cap A_n) = \prod_{k=1}^n P\left(A_k \mid \bigcap_{j=1}^{k-1} A_j\right)$$

- 2. Gotov dogodek: $P(X \mid ...X \cdot ...) = 1$
- 3. Nemogoc dogodek: $P(X \mid ... \overline{X}...) = 0$
- 4. Negacija: $P(\overline{X} \mid C) = 1 P(X \mid C)$
- 5. Ce je Y naslednik od X in je Y vsebovan v pogojnem delu: $P(X \mid YC) = P(X \mid C) \cdot \frac{P(Y \mid XC)}{P(Y \mid C)}$
- 6. Ce pogojni del ne vsebuje naslednika od X:
- (a) ce X **nima** starsev: $P(X \mid C) = P(X)$, P(X) je podan
- (b) ce ima X starse S: $P(X \mid C) = \sum_{S \in P_X} P(X \mid S)P(S \mid C)$
- 7. Iz 6b zgoraj: P(i | gc) = P(i | g)

4.2 Ovojnica Markova

X je **neodvisno** od vseh ostalih \Leftrightarrow podani **starsi**, **otroci** in **starsi otrok**



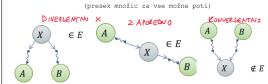
4.3 D-LOCEVANIE

A in B v mrezi sta **neodvisni** \Leftrightarrow obstaja mnozica vozlisc E, ki d-locuje A in B, potem sledi: $\{P(A|EB) = P(A|E)\}$

za vsako neusmerjeno pot P med A in B v bayesovski mreži:
za vsako vozlišče X na poti P:
analiziraj pogoj za pripadnost X množici E glede na tip:

divergentno ali zaporedno vozlišče: X∈E
konvergentno vozlišče: Xin nasledniki∉E
S_X = množice vozlišč, ki ustrezajo pogoju za X
S_D = U_X(S_Y) // množice, ki d-ločujejo samo na poti P

(unija množic za vozlišča na poti) $\mathbf{E} = \bigcap_{P} S_{P} \qquad // \text{ množice, ki d-ločujejo v celi mreži}$



! pri konvergentnem izlocimo tudi vse naslednike X

Primer d-locevanje vozlisc c in d

P(d|cbe)=P(d|be).P(d|cabe)=P(d|abe)



 $\rightarrow P(d|ca) = P(d|a), P(d|cb) = P(d|b), P(d|cab) = P(d|ab),$

4.4 I-EKVIVALENTNOST

Mrezi sta I-ekvivalentni ce imate **enako strukturo** (ob ignoriranju usmerjenosti povezav) in **ista konvergentna vozlisca**:

