



Stuart
Russell
Peter
Norvig



Umjetna inteligencija

Moderan pristup

Preveo Boris Debić

4. izdanje

Sadržaj

I Umjetna inteligencija

1	Uvod	1
1.1	Što je UI?	1
1.2	Temelji umjetne inteligencije	6
1.3	Povijest umjetne inteligencije	17
1.4	Najnovija postignuća	28
1.5	Rizici i dobrobiti umjetne inteligencije	31
	Sažetak	35
	Bibliografske i povijesne bilješke	36
2	Inteligentni agenti	37
2.1	Agenti i okruženja	37
2.2	Ispravno ponašanje: koncept racionalnosti	40
2.3	Priroda okruženja	43
2.4	Struktura agenata	48
	Sažetak	61
	Bibliografske i povijesne bilješke	62
II	Rješavanje problema	
3	Rješavanje problema pretraživanjem	64
3.1	Agenti rješavači problema	64
3.2	Primjeri problema	68
3.3	Algoritmi pretraživanja	72
3.4	Strategije neinformiranog pretraživanja	78
3.5	Informirane (heurističke) strategije pretraživanja	87
3.6	Heurističke funkcije	100
	Sažetak	108
	Bibliografske i povijesne bilješke	109
4	Pretraživanje u složenim okruženjima	113
4.1	Lokalno pretraživanje i optimizacijski problemi	113
4.2	Lokalno pretraživanje u kontinuiranim prostorima	123
4.3	Pretraživanje s nedeterminističkim djelovanjima	125
4.4	Pretraživanje u djelomično sagledljivim okruženjima	129
4.5	Agenti sekvencijalnog pretraživanja i okruženja	138
	Sažetak	144
	Bibliografske i povijesne bilješke	145
5	Suparničko pretraživanje i igre	150
5.1	Teorija igara	150
5.2	Optimalne odluke u igrama	152

5.3	Heurističko alfa–beta pretraživanje stabla	160
5.4	Monte Carlo pretraživanje stabla	165
5.5	Stohastičke igre	169
5.6	Djelomično sagledljive igre	172
5.7	Ograničenja algoritama za pretraživanje u igrama	177
	Sažetak	178
	Bibliografske i povijesne bilješke	179
6	Problemi zadovoljenja ograničenja	185
6.1	Definiranje problema zadovoljenja ograničenja	185
6.2	Propagacija ograničenja: zaključivanje u PZO-ovima	191
6.3	Pretraživanje vraćanjem unatrag za probleme zadovoljenja ograničenja	197
6.4	Lokalno pretraživanje za PZO-ove	203
6.5	Struktura problema	205
	Sažetak	209
	Bibliografske i povijesne bilješke	210
III	Znanje, zaključivanje, i planiranje	
7	Logički agenti	214
7.1	Agenti temeljeni na znanju	215
7.2	Wumpusov svijet	216
7.3	Logika	219
7.4	Propozicijska logika: vrlo jednostavna logika	223
7.5	Propozicijsko dokazivanje teorema	228
7.6	Učinkovita provjera propozicijskog modela	238
7.7	Agenti temeljeni na propozicijskoj logici	243
	Sažetak	253
	Bibliografske i povijesne bilješke	254
8	Logika prvog reda	257
8.1	Ponovno o reprezentaciji	257
8.2	Sintaksa i semantika logike prvog reda	262
8.3	Korištenje logike prvog reda	271
8.4	Inženjerstvo znanja u logici prvog reda	278
	Sažetak	283
	Bibliografske i povijesne bilješke	284
9	Zaključivanje u logici prvog reda	286
9.1	Propozicijsko nasuprot zaključivanju prvog reda	286
9.2	Unifikacija i zaključivanje prvog reda	288
9.3	Ulančavanje unaprijed	292
9.4	Ulančavanje unatrag	299
9.5	Rezolucija	305
	Sažetak	316
	Bibliografske i povijesne bilješke	317

10	Reprezentacija znanja	321
10.1	Ontološko inženjerstvo	321
10.2	Kategorije i objekti	324
10.3	Događaji	329
10.4	Mentalni objekti i modalna logika	333
10.5	Sustavi rasuđivanja za kategorije	336
10.6	Rasuđivanje sa zadanim informacijama	340
	Sažetak	345
	Bibliografske i povijesne bilješke	345
11	Automatizirano planiranje	351
11.1	Definicija klasičnog planiranja	351
11.2	Algoritmi za klasično planiranje	355
11.3	Heuristike za planiranje	360
11.4	Hijerarhijsko planiranje	364
11.5	Planiranje i djelovanje u nedeterminističkim domenama	373
11.6	Vrijeme, rasporedi i resursi	382
11.7	Analiza pristupa planiranju	386
	Sažetak	387
	Bibliografske i povijesne bilješke	388
IV	Neizvjesno znanje i rasuđivanje	
12	Kvantificiranje neizvjesnosti	393
12.1	Djelovanje uz neizvjesnost	393
12.2	Osnovna notacija vjerojatnosti	396
12.3	Zaključivanje korištenjem potpunih združenih razdioba	403
12.4	Neovisnost	406
12.5	Bayesovo pravilo i njegova primjena	407
12.6	Naivni Bayesovi modeli	410
12.7	Preispitivanje wumpusovog svijeta	412
	Sažetak	415
	Bibliografske i povijesne bilješke	416
13	Vjerojatnosno rasuđivanje	420
13.1	Reprezentacija znanja u neizvjesnim domenama	420
13.2	Semantika Bayesovih mreža	422
13.3	Egzaktno zaključivanje u Bayesovim mrežama	435
13.4	Približno zaključivanje za Bayesove mreže	443
13.5	Uzročne mreže	458
	Sažetak	462
	Bibliografske i povijesne bilješke	462
14	Vjerojatnosno rasuđivanje u vremenu	469
14.1	Vrijeme i neizvjesnost	469
14.2	Zaključivanje u vremenskim modelima	473

14.3	Skriveni Markovljevi modeli	481
14.4	Kalmanovi filtri	487
14.5	Dinamičke Bayesove mreže	493
	Sažetak	505
	Bibliografske i povijesne bilješke	505
15	Vjerojatnosno programiranje	509
15.1	Relacijski vjerojatnosni modeli	510
15.2	Modeli vjerojatnosti otvorenog svijeta	516
15.3	Praćenje složenog svijeta	523
15.4	Programi kao vjerojatnosni modeli	528
	Sažetak	533
	Bibliografske i povijesne bilješke	533
16	Donošenje jednostavnih odluka	538
16.1	Kombiniranje vjerovanja i želja pod neizvjesnošću	538
16.2	Osnove teorije korisnosti	539
16.3	Funkcije korisnosti	542
16.4	Višesvojstvena funkcija korisnosti	550
16.5	Mreže odlučivanja	555
16.6	Vrijednost informacije	557
16.7	Nepoznate preferencije	563
	Sažetak	567
	Bibliografske i povijesne bilješke	568
17	Donošenje složenih odluka	572
17.1	Problemi sekvencijalnog odlučivanja	572
17.2	Algoritmi za MDP-ove	582
17.3	Problemi razbojnika	591
17.4	Djelomično sagledljivi MDP-ovi	598
17.5	Algoritmi za rješavanje POMDP-ova	600
	Sažetak	605
	Bibliografske i povijesne bilješke	606
18	Višeagentno donošenje odluka	610
18.1	Svojstva višeagentnih okruženja	610
18.2	Teorija nekooperativnih igara	616
18.3	Teorija kooperativnih igara	637
18.4	Donošenje kolektivnih odluka	643
	Sažetak	656
	Bibliografske i povijesne bilješke	657
V	Strojno učenje	
19	Učenje iz primjera	662
19.1	Oblici učenja	662

19.2	Nadzirano učenje	664
19.3	Učenje stabala odlučivanja	668
19.4	Odabir i optimizacija modela	676
19.5	Teorija učenja	683
19.6	Linearna regresija i klasifikacija	687
19.7	Neparametarski modeli	697
19.8	Ansambl učenje	707
19.9	Razvoj sustava strojnog učenja	716
	Sažetak	726
	Bibliografske i povijesne bilješke	727
20	Učenje vjerojatnosnih modela	733
20.1	Statističko učenje	733
20.2	Učenje s potpunim podacima	736
20.3	Učenje sa skrivenim varijablama: EM algoritam	749
	Sažetak	758
	Bibliografske i povijesne bilješke	759
21	Duboko učenje	762
21.1	Jednostavne jednosmjerne mreže	763
21.2	Izračunski grafovi za duboko učenje	768
21.3	Konvolucijske mreže	772
21.4	Algoritmi učenja	777
21.5	Generalizacija	780
21.6	Povratne neuronske mreže	785
21.7	Nenadzirano učenje i učenje prijenosom	788
21.8	Primjene	794
	Sažetak	797
	Bibliografske i povijesne bilješke	797
22	Podržano učenje	802
22.1	Učenje iz nagrada	802
22.2	Pasivno podržano učenje	804
22.3	Aktivno podržano učenje	810
22.4	Generalizacija u podržanom učenju	816
22.5	Traženje politike	823
22.6	Učenje naukovanjem i inverzno podržano učenje	825
22.7	Primjene podržanog učenja	828
	Sažetak	831
	Bibliografske i povijesne bilješke	832
VI	Komuniciranje, opažanje, i djelovanje	
23	Obrada prirodnog jezika	837
23.1	Jezični modeli	837
23.2	Gramatika	847

23.3	Raščlanjivanje	850
23.4	Proširene gramatike	856
23.5	Komplikacije pravog prirodnog jezika	860
23.6	Zadaci prirodnog jezika	863
	Sažetak	865
	Bibliografske i povijesne bilješke	866
24	Duboko učenje za obradu prirodnog jezika	871
24.1	Vektorska reprezentacija riječi	871
24.2	Povratne neuronske mreže za NLP	875
24.3	Modeli niza u niz	879
24.4	Arhitektura pretvarača	884
24.5	Predobuka i učenje prijenosom	886
24.6	Najsuvremenija dostignuća	890
	Sažetak	893
	Bibliografske i povijesne bilješke	894
25	Računalni vid	897
25.1	Uvod	897
25.2	Formiranje slike	898
25.3	Jednostavne značajke slike	904
25.4	Klasifikacija slika	911
25.5	Detektiranje objekata	915
25.6	3D svijet	917
25.7	Upotreba računalnog vida	922
	Sažetak	934
	Bibliografske i povijesne bilješke	935
26	Robotika	941
26.1	Roboti	941
26.2	Robotsko sklopovlje	942
26.3	Kakav problem rješava robotika?	946
26.4	Robotska percepcija	947
26.5	Planiranje i upravljanje	954
26.6	Planiranje neizvjesnih pokreta	972
26.7	Podržano učenje u robotici	975
26.8	Ljudi i roboti	977
26.9	Alternativni robotski sustavi	985
26.10	Domene primjene	987
	Sažetak	991
	Bibliografske i povijesne bilješke	992
VII	Zaključna poglavlja	
27	Filozofija, etika i sigurnost UI	998
27.1	Ograničenja UI	998

27.2	Mogu li strojevi doista misliti?	1001
27.3	Etika UI	1003
	Sažetak	1023
	Bibliografske i povijesne bilješke	1023
28	Budućnost UI	1029
28.1	Komponente UI	1029
28.2	Arhitekture UI	1035
A	Matematička podloga	1041
A.1	Analiza složenosti i $O()$ notacija	1041
A.2	Vektori, matrice i linearna algebra	1043
A.3	Razdiobe vjerojatnosti	1045
	Bibliografske i povijesne bilješke	1047
B	Natuknice o jezicima i algoritmima	1048
B.1	Definiranje jezika Backus-Naurovim oblikom (BNF)	1048
B.2	Opisivanje algoritama pseudokodom	1049
B.3	Dodatni materijali	1050
	Bibliografija	1051
	Kazalo	1087