UNIVERZITET U NOVOM SADU
FAKULTET TEHNIČKIH NAUKA
NOVI SAD

Ivana Dojić

BAJESOVE MREŽE - MODELIRANJE U NETICI I PRIMER PRIMENE NA TENIS

- master rad -

Mentor:
dr Jelena Ivetić

Novi Sad, 2016.

УНИВЕРЗИТЕТ У НОВОМ САДУ ● **ФАКУЛТЕТ ТЕХНИЧКИХ НАУКА**
21000 НОВИ САД, Трг Доситеја Обрадовића 6

КЉУЧНА ДОКУМЕНТАЦИЈСКА ИНФОРМАЦИЈА

Редни број, **РБР**:		
Идентификациони број, **ИБР**:		
Тип документације, **ТД**:		Монографски рад
Тип записа, **ТЗ**:		Штампа
Врста рада, **ВР**:		Мастер рад
Аутор, **АУ**:		Ивана Дојић
Ментор, **МН**:		др Јелена Иветић
Наслов рада, **НР**:		Бајесове мреже – моделирање у Нетици и пример примене на тенис
Језик публикације, **ЈП**:		Српски
Језик извода, **ЈИ**:		Српски
Земља публиковања, **ЗП**:		Србија
Уже географско подручје, **УГП**:		Војводина
Година, **ГО**:		2016.
Издавач, **ИЗ**:		Ауторски препринт
Место и адреса, **МА**:		Факултет техничких наука, Нови Сад
Физички опис рада, **ФО**: (поглавља/страна/ цитата/табела/слика/графика/прилога)		(6/58/8/3/34/0/0)
Научна област, **НО**:		Математика
Научна дисциплина, **НД**:		Примењена математика
Предметна одредница/Кqучне речи, **ПО**:		Бајесове мреже, Нетика, тенис
УДК		
Чува се, **ЧУ**:		У библиотеци Факултета техничких наука
Важна напомена, **ВН**:		
Извод, **ИЗ**:		Бајесове мреже су снажан математички апарат за анализирање узрочних односа у сложеним моделима. Користе математичке алате: графове, вероватноће и Марковљеве моделе. У овом раду су описане Бајесове мреже и дати су примери који су моделирани у Нетици. Дат је оригиналан пример њихове примене у тенису.
Датум прихватања теме, **ДП**:		
Датум одбране, **ДО**:		24.09.2016.
Чланови комисије, **КО**:	Председник:	др Александар Купусинац
	Члан:	др Ксенија Дорословачки
	Члан, ментор:	др Јелена Иветић

Потпис ментора

Образац **Q2.HA.04-05** - Издање 1

UNIVERSITY OF NOVI SAD ● **FACULTY OF TECHNICAL SCIENCES**
21000 NOVI SAD, Trg Dositeja Obradovića 6

KEY WORDS DOCUMENTATION

Accession number, **ANO**:	
Identification number, **INO**:	
Document type, **DT**:	Monographic
Type of record, **TR**:	Printed
Contents code, **CC**:	Master thesis
Author, **AU**:	Ivana Dojić
Mentor, **MN**:	Jelena Ivetić, PhD
Title, **TI**:	Bayesian networks – Modelling in Netica and example of application in tennis
Language of text, **LT**:	Serbian
Language of abstract, **LA**:	Serbian
Country of publication, **CP**:	Serbia
Locality of publication, **LP**:	Vojvodina
Publication year, **PY**:	2016.
Publisher, **PB**:	Authors reprint
Publication place, **PP**:	Faculty of Technical Sciences, Novi Sad
Physical description, **PD**: (chapters/pages/ref./tables/pictures/graphs/appendixes)	(6/58/8/3/34/0/0)
Scientific field, **SF**:	Mathematics
Scientific discipline, **SD**:	Applied mathematics
Subject/Key words, **S/KW**:	Bayesian networks, Netica, tennis
UC	
Holding data, **HD**:	The library of the Faculty of Technical Sciences
Note, **N**:	
Abstract, **AB**:	The Bayesian networks are a strong mathematical tool for the analysis of causal relations in complex models. They use the following mathematical tools: graphs, probability and Markov's models. This thesis describes the Bayesian networks and gives examples modelled in Netica. There is also an original example of their application in tennis.
Accepted by the Scientific Board on, **ASB**:	
Defended on, **DE**:	September 24th, 2016.
Defended Board, **DB**: President:	Aleksandar Kupusinac, PhD
Member:	Ksenija Doroslovački, PhD
Member, Mentor:	Jelena Ivetić, PhD

Menthor's sign

	УНИВЕРЗИТЕТ У НОВОМ САДУ ● **ФАКУЛТЕТ ТЕХНИЧКИХ НАУКА** 21000 НОВИ САД, Трг Доситеја Обрадовића 6	Број:
	ЗАДАТАК ЗА МАСТЕР РАД	Датум:

(Податке уноси предметни наставник - ментор)

СТУДИЈСКИ ПРОГРАМ:	Математика у техници
РУКОВОДИЛАЦ СТУДИЈСКОГ ПРОГРАМА:	Проф. др Јованка Пантовић

Студент:	Ивана Дојић	Број индекса:	В1 4/2015
Област:	примењена математика		
Ментор:	др Јелена Иветић		

НА ОСНОВУ ПОДНЕТЕ ПРИЈАВЕ, ПРИЛОЖЕНЕ ДОКУМЕНТАЦИЈЕ И ОДРЕДБИ СТАТУТА ФАКУЛТЕТА ИЗДАЈЕ СЕ ЗАДАТАК ЗА МАСТЕР РАД, СА СЛЕДЕЋИМ ЕЛЕМЕНТИМА:
- проблем – тема рада;
- начин решавања проблема и начин практичне провере резултата рада, ако је таква провера неопходна;

НАСЛОВ МАСТЕР РАДА:

Бајесове мреже – моделирање у Нетици и пример примене на тенис

ТЕКСТ ЗАДАТКА:

Задатак мастер рада се састоји из два дела – теоријског и примењеног.

У оквиру теоријског дела задатка студент треба да представи теоријске основе Бајесових мрежа и релевантну математичку позадину потребну за њихово разумевање.

У оквиру примењеног дела, задатак је упознати се са неким од софтверских пакета за рад са Бајесовим мрежама; детаљније проучити неку област примене Бајесових мрежа и коначно, предложити оригинални модел помоћу Бајесових мрежа у изабраној области, коришћењем одабраног софтверског пакета.

Руководилац студијског програма:	Ментор рада:

Примерак за: O - Студента; O - Ментора

Sadržaj

Glava 1

Uvod

Zašto su nam uopšte važni modeli? Pomoću modela možemo da predstavimo opšta stanja pojmova i pojava u svetu. Na primer, posmatrajmo problem vožnje automobila, koji je naveden u [27]. Da bismo naučili da vozimo automobil, potrebno je prvo da naučimo osnovne stvari kao što je pritiskanje papučice za gas, da bi se automobil uopšte kretao, i upravljanje volanom. Upravljanje volanom omogućava pomeranje točkova u željenom pravcu, što zauzvrat određuje smer i pravac kretanja našeg automobila. Ovde je jasno da imamo uzročno-posledičnu vezu. Još od davnina ljudi koriste modele kao sredstva za predstavljanje raznih kompleksnih situacija u svom okruženju. To je zato što modele možemo upotrebljavati i u situacijama koje su slične onim već naučenim. Kako voziti automobil koji je, na primer, limuzina? Razlog zašto smo sposobni da vozimo i limuzinu je taj što je model vožnje automobila veoma snažan i primenljiv je u sličnim situacijma. Kompjuterski programi napravljeni u vidu obuke za vožnju automobila su veoma korisni za ljude koji žele da nauče da voze. Modeli su takođe korisni za shvatanje našeg okruženja, kao i otkrića zakona prirode u naukama kao što su biologija, hemija i fizika. Predmet modelovanja može da bude bilo šta iz našeg okruženja: kuća, automobil, ekosistem, market...

U većini modela postoje zavisnosti i nezavisnosti između osobina modela, koje nazivamo atributi. Međutim, treba naći model koji ne pripada nijednoj krajnosti, tj. da nisu svi atributi međusobno zavisni ili da nisu svi atributi međusobno nezavisni. Dakle, tražimo model u kojem su neki atributi zavisni od drugih, a neki su nezavisni. Kao rezultat ove potrebe prilikom modelovanja nastale su Bajesove mreže. Zato ćemo se u ovom radu i baviti Bajesovim mrežama, koje predstavljaju grafičku strukturu pogodnu za modelovanje znanja u prisustvu neizvesnosti.

Rad se sastoji od 6 glava.

Prva glava je uvodna, daje uvid u celokupan rad i objašnjava motivaciju za bavljenje Bajesovim mrežama.

Druga glava daje osnovne informacije o pojmovima koje ćemo koristiti u radu.

Nju čine 3 poglavlja: teorija verovatnoće, teorija grafova i Markovljev model. U poglavlju 2.1 dati su osnovni pojmovi verovatnoće, kao i uslovne verovatnoće i zajedničke raspodele verovatnoća. Dato je nekoliko teorema od kojih je dokazana Bajesova teorema. Zatim, u poglavlju 2.2 definisani su osnovni pojmovi iz teorije grafova. Definisani su usmereni i neusmereni graf i dat je primer matrice povezanosti. U poglavlju 2.3 govoreno je o Markovljevim modelima. Definisani su Markovljev lanac i Markovljev proces. Navedena je osobina Markovljevog svojstva i na kraju je definisan skriveni Markovljev model, koji je kroz primer [2.3.8] objašnjen.

Treća glava govori o veštačkoj inteligenciji i mašinskom učenju. Sastoji se od 2 celine. U prvoj celini 3.1 data je opšta priča o veštačkoj inteligenciji, objašnjen je Tjuringov test i pokazana je uska povezanost veštačke inteligencije i Bajesovih mreža preko predstavljanja jačine verovanja. Druga celina 3.2 se nadovezuje na prvu, jer je mašinsko učenje oblast veštačke inteligencije. Data je definicija mašinskog učenja i navedeno je nekoliko primera njegove široke primene.

U četvrtoj glavi su predstavljene Bajesove mreže. U poglavlju 4.1 dat je uvod u Bajesove mreže, definisani su osnovni pojmovi: odnosi roditelj-dete, predak-potomak, pokazane su uslovne zavisnosti među promenljivama preko verovatnoća, definisano je Markovljevo svojstvo i pokazano kako utiče na računanje zajednčke raspodele verovatnoća. Sve je to ilustrovano na primeru raka pluća [4.1.5]. Zatim, u poglavlju 4.2 je prikazana Netika i objašnjen je princip rada Netike. Na primeru [4.1.5] je pokazano kako funkcioniše Netika, tj. kako sama računa verovatnoće pod datim uslovima. Zatim, u poglavlju 4.3 su pokazane vrste rezonovanja, uslovne nezavisnosti i d-razdvajanje. Na kraju ovog poglavlja pokazane su vrste Bajesovih mreža od kojih su izdvojene dinamičke Bajesove mreže i o njima je govoreno u poglavlju 4.4.

Peta glava je bazirana na primeni Bajesovih mreža. Pokazana je njihova primena u računarstvu, medicini, ekonomiji, prepoznavanju govora, biologiji i sportu. Primena u sportu je izdvojena kao poglavlje 5.1 i u tom poglavlju je prikazana primena u tenisu. Dat je originalan primer Bajesove mreže na osvojen poen u tenisu [5.1.1] i dat je primer Bajesove mreže koja predstavlja procenu sposobnosti igrača na osnovu servisa i reterna preko MultiSkill i TrueSkill sistema za rangiranje igrača [5.1.5].

Šesta glava je zaključna i u njoj je sumirana predstavljena tema uz date zaključne komentare.

Na kraju rada dati su istorijski podaci o matematičarima koji su spominjani u radu i naveden je spisak korišćene literature koja je sortirana abecedno.

Glava 2

Osnovni pojmovi

U ovoj glavi navešćemo osnovne pojmove iz teorije verovatnoće i teorije grafova i objasnićemo skriveni Markovljev model koji će nam biti potreban u daljem radu. Daćemo nekoliko teorema. Većina dokaza navedenih teorema je izostavljena.

2.1 Osnovni pojmovi teorije verovatnoće

Navodimo osnovne pojmove (verovatnoća, slučajna promenljiva, funkcija raspodele i gustine, matematičko očekivanje i disperzija), kao i nekoliko teorema od kojih ćemo dokazati Bajesovu teoremu. Za literaturu smo koristili [18, 26, 30, 35, 40].

2.1.1 Slučajne promenljive

2.1.1 Definicija *Familija $\mathcal{F}$ podskupova od $\Omega \neq \emptyset$ je σ-**algebra** ako zadovoljava aksiome:*

1. $\Omega \in \mathcal{F}$;
2. *Ako* $A \in \mathcal{F}$ *onda i komplement* $\overline{A} \in \mathcal{F}$;
3. *Ako su* $A_i \in \mathcal{F}$, $i \in \mathbb{N}$, *tada i* $\cup_{i=1}^{\infty} A_i \in \mathcal{F}$.

Elementarni događaj ω je osnovni pojam koji se ne definiše. Neka je skup Ω skup elementarnih događaja i neka je na njemu data σ-algebra $\mathcal{F}$. Skupove iz σ-algebre zovemo *slučajnim događajima*, skup Ω *siguran događaj*, a $\emptyset$ *nemoguć događaj*. Uređen par $(\Omega, \mathcal{F})$ zovemo *merljiv prostor događaja*.

2.1.2 Definicija *Neka je $(\Omega, \mathcal{F})$ merljiv prostor događaja. Skupovna funkcija $P : \mathcal{F} \to [0, 1]$ zove se **verovatnoća** ako ima osobine:*

1. $P(\Omega) = 1$;
2. *Ako su* $A_1, A_2, \ldots$ *disjunktni događaji iz* $\mathcal{F}$ *tada važi* $P\big(\bigcup_{n=1}^{\infty} A_n\big) = \sum_{n=1}^{\infty} P(A_n)$.

Uređena trojka $(\Omega, \mathcal{F}, P)$ *naziva se* ***prostor verovatnoće***.

2.1.3 Definicija *Suprotan događaj događaju* A *obeležava se sa* $\overline{A}$, *gde je* $\overline{A}$ *komplement u odnosu na* Ω.

2.1.4 Definicija *Preslikavanje* $X : \Omega \to \mathbb{R}$ *iz prostora verovatnoće* $(\Omega, \mathcal{F}, P)$ *u merljiv prostor* $(R, \mathcal{B}(\mathbb{R}))$ *naziva se* ***slučajna promenljiva*** *ako inverzna slika svakog Borelovog skupa pripada* σ*-algebri* $\mathcal{F}$, *odnosno za proizvoljno* $B \in \mathcal{B}(\mathbb{R})$ *važi* $X^{-1}(B) \in \mathcal{F}$.

2.1.5 Definicija *Funkciju* $F_X : \mathbb{R} \to [0,1]$ *definisanu sa* $F_X(x) = P\{\omega \in \Omega : X(\omega) \leq x\}$ *nazivamo* ***funkcija raspodele*** *slučajne promenljive* X.

Za naš rad biće prilično važne slučajne promenljive. Razlikujemo slučajne promenljive diskretnog i slučajne promenljive apsolutno neprekidnog tipa. S tim u vezi imamo sledeću definiciju.

2.1.6 Definicija *Slučajna promenljiva* $I_A(\omega) = \begin{cases} 1, & \omega \in A \\ 0, & \omega \notin A \end{cases}$, *za događaj* $A \in \mathcal{F}$ *naziva se indikator događaja* A. *Ako postoji disjunktno razbijanje sigurnog događaja* $\Omega = A_1 \cup A_2 \cup \dots$ *i ako postoji najviše prebrojiv skup* $\{x_1, x_2, \dots\} \subset \mathbb{R}$ *tada slučajnu promenljivu oblika*

$$X(\omega) = \sum_{i=1}^{\infty} x_i I_{A_i}(\omega) \tag{2.1}$$

zovemo ***diskretna slučajna promenljiva***.

2.1.7 Definicija *Slučajna promenljiva* X *je* ***apsolutno neprekidnog tipa*** *ako postoji nenegativna integrabilna funkcija* $\varphi_X(x) : \mathbb{R} \to \mathbb{R}_+$ *takva da za svaki skup* $B \in \mathcal{B}(\mathbb{R})$ *važi*

$$P\{X \in B\} = \int_B \varphi_X(x) dx. \tag{2.2}$$

Funkcija $\varphi_X(x)$ *zove se gustina raspodele slučajne promenljive* X.

U ovom radu baziraćemo se na slučajne promenljive diskretnog tipa. Shodno tome, definisaćemo višedimenzionalne diskretne promenljive i raspodelu višedimenzionalnih diskretnih promenljivih.

2.1.8 Definicija *Neka su* $X_1, X_2, \dots, X_n$ *slučajne promenljive diskretnog tipa definisane na istom prostoru verovatnoće* $(\Omega, \mathcal{F}, P)$. *Uređena n-torka* $(X_1, X_2, \dots, X_n)$ *naziva se* ***višedimenzionalna diskretna slučajna promenljiva***.

2.1.9 Definicija *Funkciju* $F : \mathbb{R}^n \to [0,1]$ *definisanu sa* $F_{(X_1, X_2, \dots, X_n)}(x_1, x_2, \dots, x_n)$- $= P\{\omega \in \Omega : X_1(\omega) \leq x_1 \wedge X_2(\omega) \leq x_2 \wedge \dots \wedge X_n(\omega) \leq x_n\} = P\{X_1 \leq x_1, X_2 \leq x_2, \dots, X_n \leq x_n\}$ *nazivamo* ***funkcija raspodele višedimenzionalne diskretne slučajne promenljive***.

2.1.10 Definicija *Slučajne promenljive* $X_1, \dots, X_n$ *su* ***nezavisne*** *ako za svako* $k \leq n$, *za svaki izbor indeksa* $(i_1, \dots, i_k)$ *i za proizvoljne Borelove skupove* $B_1, \dots, B_k \in \mathcal{B}(\mathbb{R})$ *važi*

$$P\left(\bigcap_{j=1}^{k}\{X_{i_k} \in B_j\}\right) = \prod_{j=1}^{k} P\{X_{i_j} \in B_j\}.$$

2.1.11 Teorema *Slučajne promenljive* $X_1, \dots, X_n$ *na prostoru verovatnoće* $(\Omega, \mathcal{F}, P)$ *su nezavisne ako i samo ako je*

$$F_{(X_1,\dots,X_n)}(x_1, \dots, x_n) = F_{X_1}(x_1) \cdots F_{X_n}(x_n).$$

2.1.12 Definicija *Preslikavanje* $X : \Omega \to \mathbb{R}^n$ *iz prostora verovatnoće* $(\Omega, \mathcal{F}, P)$ *u merljiv prostor* $(\mathbb{R}^n, \mathcal{B}(\mathbb{R}^n))$ *naziva se* ***slučajan vektor*** *(n-dimenzionalna slučajna promenljiva), ako inverzna slika svakog Borelovog skupa pripada σ-algebri* $\mathcal{F}$, *tj. za svako* $B \in \mathcal{B}(\mathbb{R}^n)$ *važi* $X^{-1}(B) \in \mathcal{F}$.

2.1.13 Teorema *Preslikavanje* $X : \Omega \in \mathbb{R}^n$ *čija su komponentna preslikavanja* $X = (X_1, \cdots, X_n)$ *je slučajan vektor ako i samo ako je svako* $X_i, i = 1, \dots, n$ *jednodimenzionalna slučajna promenljiva.*

2.1.2 Numeričke karakteristike slučajnih promenljivih

U ovom radu često će nam biti važni pojedini parametri koji će dobro okarakterisati osobine raspodele verovatnoća slučajne promenljive. Stoga ćemo da definišemo *očekivanje slučajne promenljive* i *disperziju slučajne promenljive*.

2.1.14 Definicija ***Matematičko očekivanje*** *diskretne slučajne promenljive* X *date sa (2.1) definišemo sa*

$$E(X) = \sum_{k=1}^{\infty} x_k p(x_k)$$

gde smo koristili oznaku $p(x_k) = P(A_k) = P\{X = x_k\}, \ k = 1, \dots n$ *i očekivanje postoji ako i samo ako je* $\sum_{k=1}^{\infty} |x_k| p(x_k) < \infty$.

Za apsolutno neprekidnu slučajnu promenljivu X *sa gustinom* $\varphi_X(x)$ *datu sa (2.2) očekivanje je dato sa*

$$E(X) = \int_{-\infty}^{\infty} x\varphi_X(x)dx$$

i ono postoji ako integral apsolutno konvergira, odnosno ako je $\int_{-\infty}^{\infty} |x|\varphi_X(x)dx < \infty$.

2.1.15 Teorema *Osobine matematičkog očekivanja su:*

1. $E(c) = c$ *gde je* $c \in \mathbb{R}$ *proizvoljna konstanta;*

2. $E(cX) = cE(X)$ *gde je* $c \in \mathbb{R}$ *proizvoljna konstanta;*

3. $E(X+Y) = E(X) + E(Y)$.

2.1.16 Definicija *Neka je X slučajna promenljiva.* ***Disperzija*** *(varijansa) slučajne promenljive X u oznaci $D(X)$ definisana je sa*

$$D(X) = E\big((X - E(X))^2\big).$$

2.1.17 Teorema *Osobine disperzije su:*

1. $D(X) = E(X^2) - E^2(X)$;
2. $D(X) \geq 0$;
3. $D(X) = 0$ *ako i samo je* $X = const$ *skoro sigurno;*
4. $D(cX) = c^2 D(X)$ *gde je* $c \in \mathbb{R}$ *proizvoljna konstanta;*
5. $D(X+c) = D(X)$ *gde je* $c \in \mathbb{R}$ *proizvoljna konstanta.*

2.1.3 Uslovna verovatnoća i Bajesova teorema

2.1.18 Definicija *Neka su A i B događaji iz istog prostora verovatnoća $(\Omega, \mathcal{F}, P)$ i neka je $P(B) > 0$. Tada je* ***uslovna verovatnoća*** *događaja A, ako se ostvario događaj B, jednaka*

$$P(A|B) = \frac{P(AB)}{P(B)}. \tag{2.3}$$

Uslovne verovatnoće događaja iz istog prostora verovatnoća, u odnosu na neki događaj iz tog prostora, imaju sve osobine verovatnoće.

Ako posmatramo više događaja iz istog prostora verovatnoća, uslovne verovatnoće nam mogu pomoći pri određivanju verovatnoće njihovog preseka.

2.1.19 Definicija *Presek događaja A i B, u oznaci AB, je događaj koji se ostvaruje kada su se realizovala oba događaja, i A i B.*

2.1.20 Definicija *Neka su $A_1, A_2, \ldots, A_n$ događaji iz istog prostora verovatnoće i neka je presek tih događaja neprazan skup. Tada je*

$$P(A_1 A_2 \ldots A_n) = P(A_1)P(A_2|A_1)P(A_3|A_1A_2)\ldots P(A_n|A_1A_2\ldots A_{n-1}). \tag{2.4}$$

2.1.21 Definicija *Neka su događaji A i B iz istog prostora verovatnoća. Kažemo da su događaji A i B nezavisni ako važi $P(AB) = P(A)P(B)$.*

Iz gore navedenih definicija lako se može zaključiti da za nezavisne događaje A i B važi: $P(A|B) = P(A)$ i $P(B|A) = P(B)$.

2.1.22 Definicija *Neka je* $(\Omega, \mathcal{F}, P)$ *prostor verovatnoća i* $H_1, H_2, \ldots, H_n \in \mathcal{F}$ *zadovoljavaju sledeće uslove:*

1. $P(H_i) > 0, i = 1, 2, \ldots, n,$

2. $H_i \cap H_j = \emptyset, i, j = 1, 2, \ldots, n, i \neq j,$

3. $\sum_{i=1}^{n} H_i = \Omega.$

Događaji $H_1, H_2, \ldots, H_n$ *se nazivaju* ***hipoteze****, a* $\{H_1, H_2, \ldots, H_n\}$ ***potpun sistem događaja****.*

2.1.23 Teorema ***(Formula potpune verovatnoće.)*** *Ako događaji* $H_1, H_2, \ldots, H_n$ *čine potpun sistem događaja u odnosu na događaj* A*, tada je*

$$P(A) = \sum_{k=1}^{n} P(H_k) P(A|H_k).$$

Kažemo da su događaji A i B *nesaglasni* ako je $AB = \emptyset$.

2.1.24 Teorema *(**Bajesova**[1] **formula**.) Ako događaji* $H_1, H_2, \ldots, H_n$ *čine potpun sistem događaja u odnosu na događaj* A*, tada je*

$$P(H_i|A) = \frac{P(H_i)P(A|H_i)}{\sum_{j=1}^{n} P(H_j)P(A|H_j)} (i = 1, 2, \ldots, n), A \in \mathcal{F}$$

Dokaz. Posmatrajmo $P(H_i|A)$ i $P(A|H_i)$. Na osnovu (2.3) imamo da je

$$P(H_i|A) = \frac{P(H_iA)}{P(A)}, \quad P(A|H_i) = \frac{P(H_iA)}{P(H_i)}.$$

Zatim, na osnovu (2.4) u ovoj notaciji imamo

$$P(H_i)P(A|H_i) = P(A)P(H_i|A),$$

odakle je

$$P(H_i|A) = \frac{P(H_i)P(A|H_i)}{P(A)}. \tag{2.5}$$

Pošto je realizovanje događaja A uslovljeno jednim od događaja H_j, verovatnoća od A može da se izrazi kao zbir uslovnih verovatnoća tog događaja

$$P(A) = P(H_1)P(A|H_1) + P(H_2)P(A|H_2) + \cdots + P(H_n)P(A|H_n) =$$

[1]Više o Tomasu Bajesu može se pročitati u poglavlju ,,Istorijski podaci" na strani 52.

$$= \sum_{j=1}^{n} P(H_j)P(A|H_j). \tag{2.6}$$

Zamenom (2.6) u (2.5) dobija se

$$P(H_i|A) = \frac{P(H_i)P(A|H_i)}{\sum_{j=1}^{n} P(H_j)P(A|H_j)} (i = 1, 2, \ldots, n).$$

□

Dokaz smo preuzeli iz [43].

2.2 Teorija grafova

Graf je apstraktni matematički objekat koji služi kao model za opis strukture podataka. Predstavlja se pomoću tačaka, koje predstavljaju podatke, i linija koje ih spajaju. Za ovo poglavlje bazirali smo se na literaturu [7].

2.2.1 Neusmereni graf

2.2.1 Definicija *Neka je V konačan skup i neka je E skup dvoelementnih podskupova (neuređenih parova) od V. (**Neusmereni) graf** G je uređeni par $G = (V, E)$. Elemente skupa V nazivamo **čvorovi grafa** G, a elementi skupa E su **grane grafa** G.*

Neka su $a, b \in V$ dva čvora. Granu između čvorova a i b označavaćemo sa (a, b) i u tom slučaju kažemo da su a i b povezani. Ako je (a, b) isto što i (b, a) onda kažemo da je graf **neusmeren**.

Put između dva čvora u grafu je niz grana koji spaja ova dva čvora, u kome se svaka grana iz grafa pojavljuje najviše jedanput. **Povezan** graf je takav neusmereni graf kod koga su bilo koja dva čvora povezana putem. Ako postoje dva čvora koja se ne mogu povezati, graf je **nepovezan**.

2.2.2 Definicija ***Petlja*** *ili* ***ciklus*** *je put koji počinje i završava se istim čvorom.*

Ako je petlja u čvoru a, obeležavamo je sa (a, a). **Stepen čvora** a u grafu je broj grana grafa koji imaju kraj u tom čvoru i označavamo ga sa d_a. Petlju brojimo dva puta. Ako ne naglasimo drugačije smatraćemo da nemamo petlji, tj. da imamo *prost* graf.

2.2.3 Definicija *Graf koji sadrži barem jedan ciklus naziva se* ***cikličan*** *graf. U suprotnom graf je* ***acikličan****.*

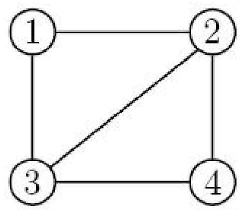
Da bismo predstavili čvorove u grafu, koristimo takozvanu matricu povezanosti.

2.2.4 Definicija *Neka je $G = [V, E]$ neusmereni graf gde $V = \{1, 2, \ldots, n\}$. **Matrica povezanosti** (eng. adjacency matrix) A grafa G je $n \times n$ simetrična matrica definisana sa*

$$A(i,j) = \begin{cases} 1, & ako\ (i,j) \in E\ ili\ (j,i) \in E \\ 0, & inače. \end{cases}$$

Skup čvorova V grafa $G = [V, E]$ je jedinstveno određen pomoću pridružene matrice povezanosti. Možemo pisati $G = [V, A]$ umesto $G = [V, E]$.

2.2.5 Primer U ovom primeru je dat jedan neusmereni graf i njegova matrica povezanosti.



	1	2	3	4
1	0	1	1	0
2	1	0	1	1
3	1	1	0	1
4	0	1	1	0

Slika 1: Primer neusmerenog grafa sa matricom povezanosti.

2.2.2 Usmereni graf

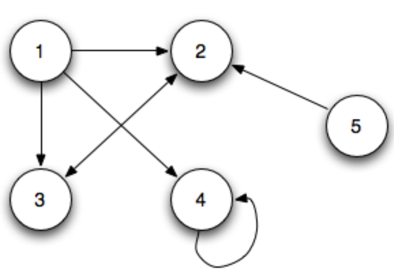
2.2.6 Definicija *Uređen par* $G = [V, E]$ *je* ***usmereni graf (digraf)*** *ako je* V *skup čvorova i* E *je skup uređenih parova čvorova iz* V*.*

2.2.7 Definicija *Neka je* $G = [V, E]$ *digraf gde* $V = \{1, 2, \ldots, n\}$*. Tada je* ***matrica povezanosti*** *grafa* G $n \times n$ *matrica takva da je:*

$$A(i,j) = \begin{cases} 1, & ako\ (i,j) \in E \\ 0, & inače. \end{cases}$$

Kod neusmerenih grafova nismo imali ,,strelice” između čvorova. Suprotno tome, usmereni grafovi imaju ,,strelice” između čvorova. Sada se lako može zaključiti da kod usmerenih grafova ako imamo ,,strelicu” od a do b, ne moramo imati strelicu od b do a. Matrica povezanosti neusmerenog grafa jedinstveno određuje graf, kao što je slučaj i kod usmerenog grafa. Obratimo pažnju da ovde matrica povezanosti može biti asimetrična matrica, pa tako možemo zaključiti da je digraf $G = [V, A]$ simetričan ako je A simetrična matrica. Očigledno, simetričan digraf je ekvivalentan sa neusmerenim grafom, pa je neusmereni graf specijalan slučaj usmerenog grafa.

2.2.8 Primer U ovom primeru je dat jedan usmereni graf i njegova matrica povezanosti.



	1	2	3	4	5
1	0	1	1	1	0
2	0	0	1	0	0
3	0	1	0	0	0
4	0	0	0	1	0
5	0	1	0	0	0

Slika 2: Primer usmerenog grafa sa matricom povezanosti.

2.3 Markovljevi modeli

Statističke metode *Markovljevi modeli* i *Skriveni Markovljevi modeli* prvi put su predstavljene i proučavane krajem šezdesetih i početkom sedamdesetih godina prošlog veka. Kako je struktura ovih modela veoma složena i zahteva jaku računarsku podršku u njihovoj implementaciji, ovi modeli su postali popularni tek poslednjih decenija uporedo sa razvojem softvera. U ovom poglavlju objasnićemo Markovljev i skriveni Markovljev model i daćemo primer radi lakšeg razumevanja. Literatura koju smo koristili za Markovljeve modele je [6, 13, 21, 34, 40].

Markovljev model

Pre nego što definišemo skriveni Markovljev model, definisaćemo slučajni proces i reći ćemo nešto o Markovljevom modelu (procesu, lancu).

2.3.1 Definicija *$\{X_t : t \in [0, T)\}$ je* ***slučajni proces*** *ako je X funkcija $X : \Omega \times [0, T) \to \mathbb{R}$, koja je za svako fiksirano $t \in [0, T)$ slučajna promenljiva $X_t : \Omega \to \mathbb{R}$.*

Parametar t ćemo interpretirati kao vreme (u opštem slučaju ne mora da bude vreme). Napomenimo da je slučajna promenljiva X_t diskretnog tipa, tj.uzima vrednosti iz skupa $\{0, 1, 2, \dots\}$. Ona predstavlja funkciju dve promenljive, zavisi od t i ω. Obično ne pišemo ω.

2.3.2 Definicija *Slučajni proces $\{X_t : t \in [0, T)\}$ je* ***diskretno vredan*** *ako za sve $t \in [0, T)$ važi $X_t : \Omega \to S_x$, gde je S_x najviše prebrojiv skup ($S_x = (x_1, x_2, \dots)$). Skup S_x zovemo* ***skup stanja*** *slučajnog procesa.*

Ako fiksiramo t (tj. $t = t_0$), tada X_{t_0} nazivamo **zasekom** slučajnog procesa za $t = t_0$, a ako fiksiramo ω (tj. $\omega = \omega_0$) onda je $X_t(\omega_0)$ realna funkcija realne promenljive koju nazivamo **trajektorija**. Ako fiksiramo oba parametra (tj. $t = t_0$, $\omega = \omega_0$), onda je $X_{t_0}(\omega_0)$ realan broj. Proces X_t ćemo interpretirati kao stanje sistema u vremenu t. U opštem slučaju skup T je podskup od R, ali u ovom radu mi ćemo za T uzimati diskretan skup ili interval $[0, +\infty)$.

Markovljev model nosi naziv po Andreju Markovu[2]. Teorija Markovljevih modela proučava nizove slučajnih promenljivih kod kojih je skup T diskretan ili $T \in [0, +\infty)$. S tim u zavisnosti razlikujemo slučajne lance i slučajne procese. Ako je T diskretan skup, tj. $T = t_0, t_1, \dots, t_n, \dots$ (neće se izgubiti na opštosti ako zapišemo ovako $T = 0, 1, \dots, n, \dots$), tada slučajni proces X_n, $n \in \mathbb{N}_0$ nazivamo **slučajni lanac**. S druge strane, ako za T uzmemo interval $[0, +\infty)$ tada imamo **slučajni proces sa neprekidnim vremenom**. To znači da sistem X_t posmatramo u neprekidnom vremenu. Shodno tome razlikujemo **Markovljeve lance** i **Markovljeve procese**.

Generalno, Markovljev model (resp. proces i lanac) možemo opisati preko sekvence stanja koja mogu da prelaze jedno u drugo, ili mogu da ostanu u istom

[2]Više o Andreju Markovu može se naći u poglavlju „Istorijski podaci” na strani 52.

stanju. Posmatrajmo sistem koji se sastoji od n unapred poznatih mogućih stanja $\{x_1, \dots, x_n\}$. Broj stanja može biti konačan ili najviše prebrojiv. Jasno je da se u svakom trenutku možemo naći tačno u jednom takvom stanju. Neka se moguće promene stanja sistema vrše u jednakim vremenskim intervalima. Neka važi sledeća relacija:

$$P\{X_t = x_j | X_{t-1} = x_i, X_{t-2} = x_k, \dots\} = P\{X_t = x_j | X_{t-1} = x_i\}$$

gde X_t predstavlja stanje u kojem se sistem nalazi u trenutku t. Ovo nam ustvari govori da to u koje će stanje sistem preći u sledećem trenutku zavisi isključivo od njegovog trenutnog stanja, bez obzira koja stanja su mu prethodila. Sada ćemo definisati verovatnoće prelaska sistema iz jednog stanja u drugo na sledeći način:

$$a_{ij} = P\{X_t = x_j | X_{t-1} = x_i\}, \text{ gde su } 1 \leq i, j \leq n$$

Parametri a_{ij} se nazivaju **verovatnoće prelaza** i moraju da zadovoljavaju sledeće uslove:

$$a_{ij} \geq 0 \quad , \quad \sum_{i,j=1}^{n} a_{ij} = 1$$

Kvadratne matrice čiji su elementi navedene verovatnoće prelaza nazivamo **stohastičke matrice**. Napomenimo da je verovatnoća prelaza nezavisna od vremena. Takođe, verovatnoće prelaza ne moraju nužno biti simetrične ($a_{ij} \neq a_{ji}$) i sistem se u pojedinom stanju može zadržati proizvoljno dugo ($a_{ii} \neq 0$). Vrednost $a_{ij}(t)$ tumačićemo kao verovatnoću da sistem pređe iz stanja i u stanje j za vreme t. Sada ćemo definisati **tranzicionu matricu** (eng. transition probability matrix), odnosno matricu verovatnoće prelaza.

$$A_t = \begin{bmatrix} a_{00}(t) & a_{01}(t) & a_{02}(t) & \dots \\ a_{10}(t) & a_{11}(t) & a_{12}(t) & \dots \\ a_{20}(t) & a_{21}(t) & a_{22}(t) & \dots \\ \vdots & \vdots & \vdots & \ddots \end{bmatrix}$$

Raspodela verovatnoće u trenutku t izgleda ovako:

$$X_t = \begin{pmatrix} 0 & 1 & 2 & \dots \\ a_0(t) & a_1(t) & a_2(t) & \dots \end{pmatrix}$$

gde je $a_k(t) = P\{X_t = x_k\}$. Verovatnosni vektor $a(t) = (a_1(t), a_2(t), \dots)$, nazivamo **vektor stanja**. Vektor stanja predstavlja verovatnoće stanja sistema X_t u trenutku t. Podrazumeva se da o vektoru stanja možemo govoriti samo ako slučajni proces ima konačno mnogo stanja.

2.3.1 Markovljev lanac

2.3.3 Definicija *Slučajni proces* $\{X_t : t \in [0,T)\}$ *je* ***Markovljev lanac*** *ako za sve* $t_1 < t_2 < \cdots < t_n$ *važi* $P\{X_{t_n} = x_n | X_{t_{n-1}} = x_{n-1}, \ldots, X_{t_1} = x_1\} = P\{X_{t_n} = x_n | X_{t_{n-1}} = x_{n-1}\}$. *Skup* T *je diskretan skup.*

2.3.4 Definicija *Slučajni proces* $\{X_t : t \in [0,T)\}$ *je* ***homogeni Markovljev lanac****, ako verovatnoće prelaza ne zavise od* t_n *i* t_{n-1}*, već samo od razlike* $t_n - t_{n-1} = \tau$*, tako da se može zapisati* $P\{X_{t+\tau} = x_i | X_t = x_j\} = a_{ij}(\tau), i,j = 1,2,\ldots$ *(verovatnoća prelaska u* τ *koraka).*

Homogeni Markovljev lanac je u potpunosti određen:

- verovatnoćama stanja $a_i(t)$, gde $i \in \{1,2,\ldots,n\}$, $t \in T$,
- verovatnoćama prelaza $a_{ij}(t)$, gde $i,j \in \{1,2,\ldots,n\}$, $t \in T$.

Verovatnoća prelaska za više koraka se određuje jednačinama Čepmen[3]-Kolmogorov.

2.3.5 Teorema *Za svako* $i,j \in \{1,2,\ldots,m\}$ *i svako* $t,k \in T$ *važi:*

$$a_{ij}(t+k) = \sum_{s=1}^{m} a_{is}(t) a_{sj}(k).$$

U matričnom obliku je $A(t+k) = A(t)A(k)$*. Ovu relaciju nazivamo* ***jednačina Čepmen-Kolmogorova****.*

Uz pomoć ove i drugih teorema, može se zaključiti da je homogeni Markovljev lanac potpuno određen *početnim verovatnoćama stanja* (u trenutku 0) i verovatnoćama prelaza. Ako je $a(t) = a(0)$ za svako $t \in T$, kažemo da je homogeni Markovljev lanac **stacionaran**, u protivnom je **nestacionaran**.

Još ćemo reći nešto o finalnim verovatnoćama. Posmatrajmo (nestacionarne) homogene Markovljeve lance, koji kada $t \to \infty$, teže ka stacionarnom procesu. To ustvari znači da za dovoljno veliko t (što intuitivno označava da sistem dugo radi) verovatnoće stanja $a_j(t)$, $j \in \{1,2,\ldots,m\}$, teže konstanti koju ćemo označavati sa ${a_j}^*$. Konstanta ${a_j}^*$ ne zavisi od t, a verovatnoće prelaza $a_{ij}(t)$, $i,j \in \{1,2,\ldots,m\}$, ne zavise ni od i ni od t. Ovo možemo predstaviti preko graničnih vrednosti:

$$\begin{aligned} \lim_{t\to\infty} a_j(t) &= {a_j}^* \\ \lim_{t\to\infty} a_{ij}(t) &= {a_j}^* \end{aligned}$$

Verovatnoće ${a_j}^*$, $j \in \{1,2,\ldots,m\}$, nazivamo **finalnim verovatnoćama** i interpretiramo ih kao verovatnoće stanja sistema u dalekoj budućnosti.

[3]Više o Sidneju Čepmenu i Andreju Kolmogorovu može se naći u poglavlju ,,Istorijski podaci" na stranama 52-53.

2.3.2 Markovljev proces

2.3.6 Definicija *Slučajni proces* $\{X_t : t \in [0,T)\}$ *je* ***Markovljev proces*** *ako za sve* $t_1 < t_2 < \cdots < t_n$ *važi* $P\{X_{t_n} = x_n | X_{t_{n-1}} = x_{n-1}, \ldots, X_{t_1} = x_1\} = P\{X_{t_n} = x_n | X_{t_{n-1}} = x_{n-1}\}$. *Skup* T *je interval* $[0, +\infty)$.

2.3.7 Definicija *Slučajni proces* $\{X_t, t \in [0, +\infty)\}$ *je* ***homogeni Markovljev proces*** *ako za sve* $0 \leq t_0 < t_1 < \cdots < t < h$ *i sve* $x_{i_0}, x_{i_1}, \ldots, x_i, x_j \in S$ *važi:*

$$P\{X_h = x_j | X_{t_0} = x_{i_0}, X_{t_1} = x_{i_1}, \ldots, X_t = x_i\} =$$

$$= P\{X_h = x_j | X_t = x_i\} = P\{X_{h-t} = x_j | X_0 = x_i\}.$$

U analogiji sa Markovljevim lancima možemo i ovde zaključiti sledeće:

- Homogeni Markovljevi procesi su u potpunosti određeni:
  - početnim verovatnoćama $a(0) = [a_1(0), \ldots, a_n(0)]$,
  - verovatnoćama prelaza $A(t) = [a_{ij}(t)]_{n \times n}, t \in [0, +\infty)$.
- Važi jednačina Čepmen-Kolmogorova;
- Finalne verovatnoće $a^* = [a_1{}^*, a_2{}^*, \ldots, a_n{}^*]$ definišemo ako postoji $t_0 \in [0, +\infty)$ takvo da je $A(t_0) > 0$, tada za svako $j \in \{1, 2, \ldots, n\}$ važi

$$\begin{aligned} \lim_{t \to \infty} a_{ij}(t) &= a_j{}^* \in (0,1) \\ \sum_{j=1}^{n} a_j{}^* &= 1. \end{aligned}$$

Iz prethodnih definicija i definicije Markovljevih lanaca možemo zaključiti da generalno Markovljevi modeli nemaju pamćenje. To ustvari znači da verovatnoća nekog događaja koji će se realizovati u budućnosti (u nekom trenutku t) ne zavisi od prošlosti (od trenutaka $t_1, t_2, \ldots, t_{n-1}$), već samo od sadašnjosti (od trenutka t_n). Drugim rečima, raspodela slučajne promenljive X_t u momentu t_n zavisi samo od vrednosti x_{n-1} procesa u momentu t_{n-1}, a ne zavisi od ostalih vrednosti $x_{n-2}, \ldots, x_1$ procesa u momentima $t_{n-2} > t_{n-3} > \cdots > t_1$. Ovu osobinu Markovljevih procesa nazivamo **Markovljevo svojstvo**.

Ovim su opisani Markovljevi modeli prvog reda, tj. modeli kod kojih verovatnoća stanja sistema u trenutku $t + 1$ zavisi samo od stanja u trenutku t. Postoje i Markovljevi modeli višeg reda, ali oni prevazilaze okvire ovog rada.

2.3.3 Skriveni Markovljev model

Za razliku od Markovljevog modela koji se sastoji od niza stanja koji emituju neki niz opažanja koji nam je uvek poznat, skriveni Markovljev model (eng. hidden Markov model, HMM) sadrži stanja koja su ,,skrivena", kao što i samo ime ovih modela to govori. Ta stanja nam nisu poznata pri emitovanju nekog niza vrednosti, ali nam je poznat taj niz vrednosti i pomoću njega možemo doći do nekih zaključaka o nizu ,,skrivenih" stanja. Literatura na koju smo se bazirali u ovom poglavlju je [6, 14, 15, 21].

Poznata opažana stanja ćemo obeležavati sa Y_t, a skrivena stanja sa X_t. Skriveni Markovljev model ima takav naziv zbog dve stvari. Prvo, zbog stanja koja su ,,nevidljiva" posmatraču, i drugo što stanja moraju zadovoljavati Markovljevo svojstvo.

Napomenimo da su skrivena stanja promenljive diskretnog tipa, dok opažana mogu biti promenljive ili diskretnog ili neprekidnog tipa.

Razvoj HMM-a (u daljem tekstu ćemo pisati HMM) temelji se na Markovljevom lancu prvog reda. Kod HMM nije poznat niz stanja, već samo funkcija koja predstavlja verovatnoću emitovanja svakog simbola u svakom stanju.

HMM možemo opisati kao konačni nedeterministički automat sa konačnim brojem stanja (simbola) , koji kada pređe u novo stanje emituje opažanje.

Sada ćemo definisati parametre koji jednoznačno određuju HMM:

1. $\{\mathbf{X}\} = \{X_1, X_2, \dots\}$ je Markovljev lanac koji ne možemo direktno opažati, takozvani niz simbola,

2. $\{\mathbf{Y}\} = \{Y_1, Y_2, \dots\}$ je niz opažanja,

3. **N** - ćemo definisati kao ukupan broj stanja, $N = N_X + N_Y$, gde N_X predstavlja broj skrivenih stanja, a N_Y broj opažanih stanja,

4. **A** - je tranziciona matrica verovatnoća prelaza stanja koju smo već definisali
$$a_{ij} = P\{X_t = j | X_{t_{-1}} = i\},$$

5. **E** - ćemo definisati kao matricu verovatnoća emitovanja simbola
$$e_{ij} = P\{Y_t = j | X_t = i\},$$

6. a_i - kao što smo već ranije definisali, a_i predstavlja verovatnoće stanja, tj. $a_i = P\{X_t = i\}$.

HMM je generisan trojkom (A, E, a).

Zbog osobine Markovljevog svojstva, možemo dati zajedničku raspodelu verovatnoće niza simbola i opažanja:

$$P\{X_{1:T}, Y_{1:T}\} = P\{X_1\}P\{Y_1|X_1\} \times \prod_{t=2}^{T} P\{X_t|X_{t_{-1}}\}P\{Y_t|X_t\}$$

gde $X_{1:T}$ predstavlja niz $X_1, X_2, \ldots, X_T$. Ovu zajedničku raspodelu verovatnoće možemo predstaviti na grafu, poznatijem kao **Bajesova mreža**.

HMM nam omogućavaju jednostavno modeliranje procesa iz stvarnog sveta, ali najpre je potrebno rešiti tri osnovna problema HMM:

1. **Problem evaluacije.** Rešavanjem ovog problema dobija se verovatnoća niza opažanja ($P(Y)$) u odnosu na neki zadani model. Rešava se *Forward-backward* algoritmom. Više o ovom algoritmu se može naći u [1].

2. **Problem dekodiranja.** Rešavanjem ovog problema dobija se najverovatniji niz simbola (X) za dobijeni niz opažanja (Y). Rešava se *Viterbijevim algoritmom*. Više o ovom algoritmu se može naći u [12].

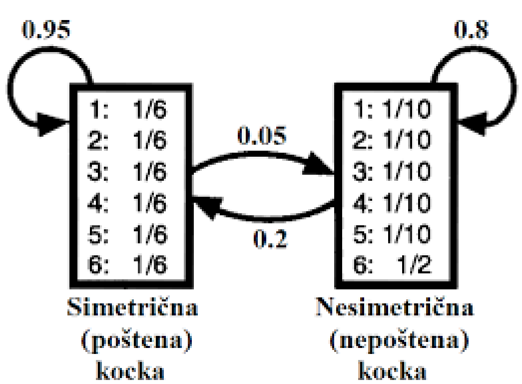
3. **Problem treniranja.** Rešavanjem ovog problema dobijaju se parametri modela: emisione i tranzicione verovatnoće (e_{ij} i a_{ij}). Rešava se *Baum-Welch* algoritmom o kom se više može naći u [2].

Dakle, pomoću ova tri algoritma se računaju parametri modela koji su nam potrebni. Detaljna analiza ovih algoritama prevazilazi okvire ovog rada, tako da ćemo koristiti već izračunate parametre. Kodovi algoritama su detaljno obrađeni u [6, 31, 38].

HMM imaju široku primenu u računarstvu, u prepoznavanju govora, rukopisa, gestikulacije, bioinformatici, ...

Sada ćemo dati primer skrivenog Markovljevog modela radi lakšeg razumevanja. Primer je uzet iz [15].

2.3.8 Primer nepoštene kockarnice.



Slika 3: Primer nepoštene kockarnice.

Primer se interpretira na sledeći način. Kockar baca kockice igajući neku igru i pri tome koristi samo dve kockice. Jedna od tih kockica je ,,poštena" (daje svaki broj sa verovatnoćom jedne šestine), a druga je ,,nepoštena" i daje šesticu sa verovatnoćom jedne polovine, dok ostali brojevi imaju verovatnoću jedne desetine (te verovatnoće možemo videti na dijagramu). S vremena na vreme, da ne bi bio uhvaćen u varanju, kockar menja kockice.

Dva stanja data na dijagramu predstavljaju kockice. Strelice između stanja predstavljaju verovatnoće prelaza iz jednog stanja u drugo. Date su i verovatnoće da se ne menja stanje kockica (to predstavljaju strelice iznad stanja). Unutar stanja su prikazane emisione verovatnoće brojeva pod uslovom da li je korišćena poštena ili nepoštena kockica.

U skladu sa uvedenom terminologijom, primer nepoštene kockarnice predstavićemo na sledeći način:

1. $X = \{p, n\}$ je niz simbola gde p predstavlja poštenu kockicu, a n nepoštenu kockicu;

2. $Y = \{1, 2, 3, 4, 5, 6\}$ je niz opažanja koji predstavlja brojeve na kockici;

3. $N = 8$, jer je $N_X = 2$, a $N_Y = 6$;

4. Tranziciona matrica izgleda ovako:

$$A = \begin{bmatrix} 0,95 & 0,05 \\ 0,2 & 0,8 \end{bmatrix}.$$

   Na primer, ako smo bacili poštenu kockicu verovatnoća da će opet biti bačena poštena je $a_{pp} = 0,95$, dok je verovatnoća da će biti zamenjena nepoštenom $a_{pn} = 0,05$.

5. Emisiona matrica izgleda ovako:

$$E = \begin{bmatrix} \frac{1}{6} & \frac{1}{6} & \frac{1}{6} & \frac{1}{6} & \frac{1}{6} & \frac{1}{6} \\ \frac{1}{10} & \frac{1}{10} & \frac{1}{10} & \frac{1}{10} & \frac{1}{10} & \frac{1}{2} \end{bmatrix}.$$

   Prva vrsta emisione matrice predstavlja verovatnoće kada je korišćena poštena kockica, a druga vrsta kada je korišćena nepoštena kockica. Na primer, $e_{p_3} = P\{Y = 3|X = p\} = \frac{1}{6}$ ili $e_{n_6} = P\{Y = 6|X = n\} = \frac{1}{2}$.

6. Vektor stanja je $a = [0,95 \;\; 0,8]$, gde je $a_p = 0,95$ i $a_n = 0,8$.

Tri osnovna problema HMM u primeru nepoštene kockarnice odredili bi sledeće parametre:

1. *Problem evaluacije* bi odredio verovatnoću niza brojeva, tj. $P(Y)$.

2. *Problem dekodiranja* bi odredio najverovatniji niz kockica za dati niz brojeva. Na primer, ako smo prilikom bacanja kockice dobili sledeći niz $Y = (2\,3\,1\,6\,6\,5)$, problem dekodiranja bi odredio najverovatniji niz X. Mi ne znamo pri kojem bacanju je korišćena poštena, a pri kojem nepoštena kockica. To zna samo kockar. Jedna od mogućnosti niza X za takav niz Y je $X = (n, n, n, p, p, n)$.

3. *Problem treniranja* u ovom slučaju je rešen. Date su nam i emisione i tranzicione verovatnoće.

Kao što smo već videli ovi problemi se rešavaju algoritmima koji prevazilaze okvire ovog rada i više o njima se može naći u $[6, 31, 38]$.

Korišćenjem Bajesovih mreža možemo dati odgovore na pitanja koja ovaj problem donosi.

Glava 3

Veštačka inteligencija i mašinsko učenje

U ovoj glavi ćemo pisati o veštačkoj inteligenciji i mašinskom učenju. Veštačka inteligencija je naučna disciplina koju je teško precizno definisati. Mi ćemo samo dati primere nekih definicija vođenih rezonovanjem i ponašanjem i objasnićemo ispitivanje inteligencije preko Tjuringovog testa. Takođe, objasnićemo usku povezanost veštačke inteligencije i Bajesovih mreža. Definisaćemo mašinsko učenje i navešćemo nekoliko primera njegove primene. Veštačka inteligencija je veoma opširna oblast koja iz dana u dan sve više pomera svoje granice. Mi ćemo se u ovom radu fokusirati samo na osnovne pojmove i ideje, a više o ovoj oblasti se može naći u [13, 20, 23, 29, 36].

3.1 Veštačka inteligencija

Kada govorimo o prirodnoj inteligenciji najčešće mislimo na nadarenost, prirodnu sposobnost pravilnog rasuđivanja, sposobnost snalaženja u situacijama na osnovu prethodno stečenog iskustva, itd. Prirodno se postavlja pitanje: Da li inteligencija predstavlja jednu zasebnu sposobnost ili je ona skup različitih i nepovezanih mogućnosti? Na ovo pitanje se nadovezuju mnogobrojna pitanja, kao npr: Na koji način je znanje smešteno u ljudskom mozgu? Kako se uopšte stiče novo znanje? Na ova pitanja ne može da odgovori samo istraživač, već je potrebno i znanje psihologa, filozofa, neurologa,... Ljudski mozak je veliki neiskorišćen materijal i za sada ga nikakva veštačka inteligencija ne može potisnuti iz upotrebe, ali ga može zameniti na mestima gde čovek nije preko potreban.

Za veštačku inteligenciju su veoma bitni inteligentni sistemi, koje, neformalno, možemo definisati kao entitete kod kojih postoji sposobnost inteligentnog ponašanja, tj. ponašanja koje srećemo kod ljudi.

Veštačka inteligencija (eng. Artificial intelligence, AI) je oblast računarstva koja je u najvećoj meri posvećena inteligentnim sistemima. Drugim rečima, veštačka inteligencija je naučna oblast koja ima za cilj da razvije program (softver), koji će računarima omogućiti da se ponašaju na način koji bi mogao biti okarakterisan kao inteligentan. Izraz veštačka inteligencija koristi se od sredine pedesetih godina i za njegovo uvođenje smatra se najzaslužnijim Džon Makarti[1].

Sada ćemo dati nekoliko neformalnih definicija veštačke inteligencije preuzetih iz [29], koje mogu da se podele u četiri kategorije:

1. Sistemi koji razmišljaju kao ljudi;
2. Sistemi koji se ponašaju kao ljudi;
3. Sistemi koji razmišljaju racionalno;
4. Sistemi koji se ponašaju racionalno.

Za prvu kategoriju imamo sledeće definicije:

,,Uzbudljiv novi pokušaj da se kompjuteri nateraju da razmišljaju...mašine sa umom, u punom i bukvalnom smislu.” (Haugeland, 1985. u [19]).

,,Aktivnosti koje asociramo sa ljudskim razmišljanjem, aktivnosti kao što su donošenje odluka, rešavanje problema, učenje...” (Bellman, 1978. u [3]).

U drugoj kategoriji imamo sledeće definicije:

,,Veština pravljenja mašina koje izvršavaju funkcije koje zahtevaju inteligenciju kada se izvršavaju od strane ljudi.” (Kurzweil, 1990. u [25]).

,,Izučavanje kako naterati računare da rade stvari u kojima su, trenutno, ljudi bolji.” (Rich i Knight, 1991. u [22]).

Za treću kategoriju su date sledeće definicije:

,,Izučavanje mentalnih sposobnosti kroz korišćenje kompjuterskih modela.” (Chamiak i McDermott, 1985. u [10]).

,,Izučavanje računanja koje omogućava opažanje, rasuđivanje i delovanje.” (Winston, 1992. u [41]).

Za poslednju kategoriju imamo sledeće:

,,Računarska inteligencija je nauka o dizajniranju inteligentnih agenata.” (Poole, 1998. u [16]).

,,Veštačka inteligencija se bavi inteligentnim ponašanjem veštačkih naprava.” (Nilsson, 1998. u [33]).

Veštačka inteligencija nema samo za cilj da kreira mašine koje će vršiti funkcije koje zahtevaju inteligenciju, već ona proučava i mentalne modele i aktivnosti kroz korišćenje računara.

[1] Više o Džonu Makartiju može se naći u poglavlju ,,Istorijski podaci” na strani 53.

3.1.1 Tjuringov test

Alan Tjuring[2], po kojem je test dobio ime, je 1950. naterao mnoge na razmišljanje kada se zapitao da li mašine mogu da prođu test inteligentnog ponašanja. Tjuringov test[3] predstavlja igru pogađanja da li je onaj koji odgovara na pitanja nezavisnog ispitivača čovek ili mašina. Mašina će proći test ako ispitivač, posle postavjanja nekoliko pitanja u pisanoj formi, ne može da odredi da li je odgovor dao čovek ili mašina. Mašina koja bi prošla Tjuringov test, morala bi da ima sledeće sposobnosti:

- **Obradu prirodnih jezika** da bi mogla uspešno da komunicira;
- **Reprezentaciju znanja** da bi mogla da skladišti ono što zna i prima na osnovu senzora;
- **Automatsko rezonovanje** da bi mogla da koristi uskladištene informacije za odgovaranje na pitanja i za donošenje novih zaključaka;
- **Mašinsko učenje** da bi mogla da se adaptira na nove okolnosti.

Kako je fizička simulacija čoveka nepotrebna za inteligenciju, Tjuringov test je namerno izbegavao direktnu fizičku interakciju između ispitivača i računara. Međutim, takozvani potpuni Tjuringov test uključuje video signal, tako da ispitivač može da testira percepcijske i motorne sposobnosti ispitanika. Mašina koja bi prošla potpuni Tjuringov test, morala bi još da ovlada i:

- **Računarskom vizijom** u cilju vizuelne percepcije i objekata u njoj;
- **Robotikom** radi kretanja u prostoru i manipulacijom objekata u njemu.

Dominantan deo veštačke inteligencije čine baš ovih 6 sposobnosti, a Tjuringov test u neizmenjenoj formi i danas predstavlja test dostignuća u ovoj oblasti. Još uvek nije zabeležen slučaj da je neka mašina prošla Tjuringov test. Definicije i priču o Tjuringovom testu smo uzeli iz [29].

Postoje dva glavna pravca u razvoju veštačke inteligencije. To su:

- proučavanje prirodne inteligencije (spoznavanje funkcija mozga, modeliranje rada mozga, simuliranje čovekovog ponašanja, reagovanja i rezonovanja);
- postizanje inteligentnog ponašanja primenom drugačijih pristupa, kakvi se ne mogu sresti u prirodnim sredinama.

Da bi veštačka inteligencija mogla da oponaša situacije iz realnog sveta, morala bi da bude osposobljena da koristi verovatnoću, što se zove *Bajesovo rasuđivanje*. Da bi predstavljao stvarne situacije, inteligentan sistem mora da oponaša bar sledeće tri vrste nesigurnosti:

[2] Više o Alanu Tjuringu može se naći u poglavlju ,,Istorijski podaci" na strani 54.

[3] Originalno je Tjuring ovaj test formulisao kao igru pogađanja ko je muškarac, a ko je žena, na osnovu pitanja nezavisnog ispitivača.

- **Neznanje**. U prirodi našeg razuma je da bude nesigurno oko mnogo stvari. Na primer, nesigurni smo da li naš partner u pokeru zaista ima fleš ili nas samo blefira?

- **Neodređenost ili fizička slučajnost**. Ovo objašnjava da smo svi mi u praksi nedeterministi. Zašto je to tako? Uzmimo za primer bacanje novčića. Iako nam je već poznat način na koji se baca novčić i kako mu se daje spin prilikom bacanja, nikada nećemo moći da budemo sigurni da li će pasti glava ili pismo. Deterministi će uvek smatrati da će se detaljnijim istraživanjem moći tvrditi na koju će stranu novčić pasti, ali takav stav je ravan naučnoj fantastici.

- **Neodređenost u širem smislu**. Skoro pa je nemoguće odrediti da li je osoba hrabra ili ne, da li je osoba lepa ili ne, jer ne postoje validni testovi koji to određuju.

Kvantifikaciju našeg ,,nošenja" sa nesigurnostima nazivamo *jačinom verovanja*. *Bajesizam* zastupa filozofiju koja tvrdi sledeće: da bi se shvatilo ljudsko razmišljanje koje je ograničeno neizvesnošću i nesigurnošću, račun verovatnoće je jedan od najvažnijih alata za predstavljanje jačine verovanja. Priča o nesigurnostima i Bajesizmu nalazi se u [23].

Zbog potrebe predstavljanja jačine verovanja koriste se Bajesove mreže. U daljem radu, bavićemo se Bajesovim mrežama, koje objašnjavaju prednosti verovanja kao verovatnoće. Posebno ćemo obrazložiti uslovne verovatnoće i pokazati kako one utiču na jačinu verovanja. Pre nego što pređemo na priču o Bajesovim mrežama, reći ćemo nešto i o mašinskom učenju.

3.2 Mašinsko učenje

Mašinsko učenje je podoblast veštačke inteligencije. Veštačka inteligencija i mašinsko učenje za cilj imaju da se stvori mašina koja će moći da radi ono što joj se naredi. Pored toga veštačka inteligencija uključuje i apstraktno mišljenje, tj. da mašina sama može da razmišlja. Zato je mašinsko učenje usko povezano sa veštačkom inteligencijom. Više o mašinskom učenju se može naći u [20, 28].

Mašinsko učenje predstavlja skup algoritama, teorijskih rezultata i primena iz različitih oblasti veštačke inteligencije, ali i drugih oblasti kao što su matematika, psihologija, filozofija i druge nauke. Mašinsko učenje za cilj ima da se stvore programi koji su osposobljeni da uče. Navešćemo par razloga zbog kojih bi mašine i računari trebali da budu osposobljeni da uče:

- Neki problemi se vrlo teško mogu algoritamski opisati osim primerima u toku praktičnog rada (npr. prepoznavanje govora kod kojeg je neophodno personalizovati bazu znanja o konkretnom glasu ili prepoznavanje lica radi autentifikacije);

- Analiza podataka (eng. Data mining) koja omogućava da se otkriju neočigledni entiteti i relacije među njima u velikoj količini podataka;
- Ponekad su količine podataka i odnosa među njima toliko velike da je ljudima gotovo nemoguće da u celini obuhvate znanje o njima, pa je praktičnije postupno mašinski obuhvatiti takvo znanje;
- U mnogim oblastima se kontinuirano prikupljaju podaci sa ciljem da iz njih izvučemo neki zaključak (npr. u medicini prikupljamo podatke o pacijentima i terapijama);
- Ljudi zaboravljaju, mašine ne!

Ovo su samo neke od prednosti mašinskog učenja.

Mašinsko učenje možemo definisati na različite načine, a mi ćemo ga definisati na način na koji je definisano u literaturi [9].

3.2.1 Definicija *Za program (mašinu) M kaže se da uči iz iskustva E u odnosu na klasu zadataka T i meru performansi P (mera je definisana nad osobinom ili skupom osobina koje takođe moraju biti definisane), ukoliko se njegove mere performansi na klasi zadataka T unapređuju sa iskustvom E.*

Težnja mašinskog učenja je da se što više približi ljudskom učenju po efikasnosti, kao i da ga objasni pomoću teorijskog modela. Ovo je za sada daleko od ostvarenja u pravom smislu, ali su uspesi u rešavanju pojedinih praktičnih problema u mnogim slučajevima impresivni.

3.2.1 Primeri primena mašinskog učenja

Navešćemo samo nekoliko od mnogobrojnih primena mašinskog učenja. Navedeni primeri kao i druge primene mogu se pronaći u [20, 39].

Jedan od najstarijih praktičnih rezultata jeste ALVINN sistem koji je naučen da vozi javnim putem u prisustvu drugih vozila bez ljudske pomoći brzinom od oko 110km/h. Taj sistem je zasnovan na neuronskoj mreži i uspešno je vozio na putu dužine oko 140km. Projekat razvoja autonomnog vozila je sa razvojem dubokih neuronskih mreža dobio teži zadatak, a to je proširenje svojih mogućnosti. U toku je razvoj vozila koje treba da bude u stanju da samostalno učestvuje u gradskoj vožnji, koja je značajno komplikovanija od vožnje na autoputu. Izazovi za tehnike mašinskog učenja koje ovaj problem donosi su prepoznavanje puta i učesnika u saobraćaju, a jedan od najvećih izazova je donošenje odluka.

Mnoge kompanije iz bezbednosnih razloga koriste sistem za prepoznavanje lica. Na osnovu fotografije ili video snimka, sistem treba da prepozna koja je osoba u pitanju. Drugim rečima, sistem treba da klasifikuje osobu (npr. Alisa, Džon, Majkl, ...) ili da je tretira kao nepoznatu osobu. Sličan, ali konceptualno drugačiji problem je verifikacija, gde sistem treba da utvrdi da li je osoba ona koja tvrdi da jeste. Za razliku od prethodnog sistema, ovaj sistem radi na principu da/ne pitanja.

Sistem TD-Gammon za igranje igre Backgammon je igrajući protiv sebe više od milion partija i nastavljajući da uči u igri sa ljudskim igračima, dostigao nivo igre u rangu svetskog šampiona. Na sličnim principima, ali konstruisan je sistem AlfaGo koji je 2015. i 2016. ubedljivo pobedio evropskog, a zatim i svetskog šampiona u igri go. Ova igra je poznata kao jedan od, do sada, najozbiljnijih izazova veštačkoj inteligenciji u domenu igranja igara, pošto po broju mogućih stanja daleko prevazilazi i šah.

Društvena mreža se može predstaviti kao graf čiji čvorovi predstavljaju učesnike mreže, a grane postoje između učesnika koji su povezani u mreži (npr. prijateljstvo na mreži Fejsbuk). Pojam društvene mreže nije ograničen na mreže na internetu, već se odnosi i na bilo kakav vid povezanosti ljudi u relanom životu. Ovaj primer predstavlja jednu od primena metoda mašinskog učenja nad grafovima. U ovom kontekstu metode mašinskog učenja se koriste za predviđanje budućih veza među učesnicima, npr. prilikom preporučivanja učesnicima mreže sa kime se mogu povezati.

Glava 4

Bajesove mreže

U ovoj glavi bavićemo se specijalnim vrstama modela, Bajesovim mrežama. Definisaćemo osnovne pojmove Bajesove mreže s aspekta strukture grafa i s aspekta verovatnoća. Objasnićemo kako funkcioniše softver Netika, softver u kojem se crtaju Bajesove mreže. Zatim, pokazaćemo uslovne nezavisnosti i d-razdvajanje i objasnićemo kako se ,,dekodira" Bajesova mreža, tj. do kojih zaključaka dolazimo. Sve to ćemo ilustrovati na primeru raka pluća [4.1.5]. Na kraju ćemo pokazati koje sve vrste Bajesovih mreža postoje i detaljnije ćemo objasniti dinamičke Bajesove mreže. Može se naći puno literature iz ove oblasti, a mi ćemo se najviše bazirati na [5, 14, 23, 27, 32].

4.1 Osnovni pojmovi Bajesovih mreža

Sa matematičke tačke gledišta, Bajesova mreža je usmereni, aciklični graf koji predstavlja odnose među promenljivama i daje rapodele verovatnoće svake promenljive, dok sa stanovišta primene, Bajesova mreža predstavlja ključnu kompjutersku tehnologiju za rad sa verovatnoćama u veštačkoj inteligenciji. Prvo ćemo predstaviti sintaksu Bajesovih mreža, a kasnije i čitanje interpretiranih informacija (semantiku). Judea Perl[1] je prvi uveo termin Bajesove mreže.

Matematički, za predstavljanje Bajesovih mreža postoje 2 aspekta: struktura koja predstavlja graf i verovatnoće dodeljene čvorovima na grafu.

Prvi aspekt su, dakle, strukture koje predstavljaju graf. Podsetićemo se potrebnih termina iz teorije grafova koje smo objasnili u poglavlju 2.2 i uvešćemo posebnu terminologiju za ove strukture:

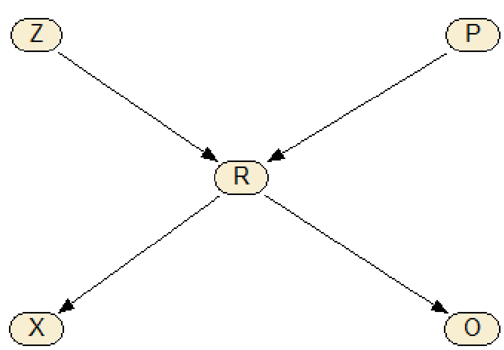
- *Usmereni* graf G je uređen par $[V, E]$, gde je V skup čvorova, a E je skup uređenih parova čvorova iz V;
- *Acikličan* graf je graf koji ne sadrži nijednu petlju;

[1] Više o Judea Perlu može se naći u poglavlju ,,Istorijski podaci" na strani 54.

- Kao što ćemo videti kasnije, čvorovi će biti slučajne promenljive;
- Grane će predstavljati odnose između čvorova koji mogu biti:
  - *Roditelj*. Ako imamo npr. dva čvora X_i i X_j i granu od X_i do X_j, tada je čvor X_i roditelj čvoru X_j. Funkciju roditelja ćemo obeležavati sa $rod(X)$.
  - *Dete*. Shodno definisanom odnosu roditelja, čvor X_j je dete čvora X_i.
  - *Predak*. Ako imamo npr. čvorove X_i, X_j i X_k, i grane od X_i do X_j i od X_j do X_k, tada su preci čvora X_k čvorovi X_i i X_j, a predak čvora X_j je čvor X_i (čvor X_i nema pretke).
  - *Potomak*. Shodno definisanim precima, potomci čvora X_i su čvorovi X_j i X_k, a potomak čvora X_j je čvor X_k (čvor X_k nema potomke).
  - *Nepotomci*. Shodno definisanim precima i potomcima, nepotomci nekog čvora X_i su svi njegovi preci i sam čvor X_i.
  - *Koren*. Koren je čvor koji nema roditelje. U primeru za pretke i potomke čvor X_i je koren. Koreni čvorovi predstavljaju glavne uzroke nekog problema koji modelujemo.
  - *List*. List je čvor koji nema decu. U primeru za pretke i potomke cvor X_k je list. Lisni čvorovi predstavljaju posledice nekog problema koji modelujemo.

Napomenimo da postoji grana iz svakog elementa $rod(X)$ u X. Sada ćemo na jednom primeru ilustrovati navedene odnose.

4.1.1 Primer Primer jednog acikličnog usmerenog grafa.



Slika 4: Primer usmerenog acikličnog grafa.

U skladu sa uvedenom terminologijom, odredićemo odnose na datom primeru:

- Čvorovi su $\{Z, P, R, X, O\}$;
- Roditelji datih čvorova su:
  - $rod(R) = \{Z, P\}$,
  - $rod(X) = \{R\}$,
  - $rod(O) = \{R\}$,
  - Čvorovi Z i P nemaju roditelje.
- Shodno navedenim roditeljima, deca datih čvorova su:
  - dete čvora Z je čvor R,
  - dete čvora P je čvor R,
  - deca čvora R su čvorovi X i O.
  - čvorovi X i O nemaju decu.
- Preci čvorova su:
  - preci čvora R su Z i P,
  - preci čvorova X i O su čvorovi: R, Z i P,
  - čvorovi Z i P nemaju pretke.
- Shodno navedenim precima, potomci su:
  - potomci čvorova Z i P su: R, X i O,
  - potomci čvora R su X i O,
  - čvorovi X i O nemaju potomke.
- Nepotomci čvorova su:
  - nepotomak čvora Z je sam čvor Z. Isto važi i za čvor P (sam je svoj nepotomak),
  - nepotomci čvora R su čvorovi: R, Z i P,
  - nepotomci čvora X su čvorovi: X, R, Z i P,
  - nepotomci čvora O su čvorovi: O, R, Z i P.
- Koreni su čvorovi Z i P.
- Listovi su čvorovi X i O.

Ovim smo objasnili prvi aspekt Bajesovih mreža.

Drugi aspekt su verovatnoće. Pomoću odgovarajućih statističkih metoda možemo proceniti uslovne zavisnosti na grafu. Stoga, Bajesove mreže su poznate i kao *mreže poverenja* (eng. belief networks).

U prethodno definisanoj grafičkoj strukturi:

- čvorovi predstavljaju slučajne promenljive, koje predstavljaju karakteristike datog modela;
- grane predstavljaju zavisnosti između slučajnih promenljivih.

Verovatnoća nad svim slučajnim promenljivama, $P(X_1, X_2, \ldots, X_n)$, naziva se **zajednička raspodela verovatnoće**. Povezan sa Bajesovim mrežama je i skup uslovnih raspodela verovatnoća, odnosno uslovna verovatnoća svake slučajne promenljive s obzirom na svoje roditelje (koji uključuje i prethodne verovatnoće tih promenljivih bez roditelja).

Da bismo uveli svu potrebnu terminologiju Bajesovih mreža počinjemo sa skupom slučajnih promenljivih koje predstavljaju odabrane karakteristike našeg proizvoljnog modela. Napomenimo da radimo sa slučajnim promenljivama diskretnog tipa. Neka su $X_1, X_2, \ldots, X_n$ slučajne promenljive. Prvo odredimo totalno uređenje tih promenljivih, $X_1, X_2, \ldots, X_n$. Pomoću pravila lanca (eng. chain rule), kojim možemo da izračunamo zajedničku raspodelu skupa slučajnih promenljivih koristeći samo uslovne verovatnoće, pretvorićemo konjunkciju u uslovnu verovatnoću:

$$P(X_1 = x_1 \wedge X_2 = x_2 \wedge \cdots \wedge X_n = x_n) = \prod_{i=1}^{n} P\left(X_i = x_i | X_1 = x_1 \wedge \cdots \wedge X_{i-1} = x_{i-1}\right).$$

Ili, u smislu višedimenzionalne slučajne promenljive i distribucije verovatnoće imamo sledeće:

$$P(X_1, X_2, \ldots, X_n) = P(X_1)P(X_2|X_1)P(X_3|X_1, X_2)\ldots P(X_n|X_1, X_2, \ldots, X_{n-1}) = \\ = \prod_{i=1}^{n} P\left(X_i|X_1, \ldots, X_{i-1}\right).$$

U slučaju kada imamo dato totalno uređenje slučajnih promenljivih $X_1, X_2, \ldots, X_n$, roditelj slučajne promenljive X_i je ustvari minimalan skup prethodnika od X_i u totalnom uređenju takvih da su drugi prethodnici od X_i uslovno nezavisni od roditelja od X_i. Dakle, $rod(X_i) \subseteq \{X_1, X_2, \ldots, X_{i-1}\}$ pa možemo zaključiti

$$P(X_i|X_1, \ldots, X_{i-1}) = P\left(X_i|rod(X_i)\right).$$

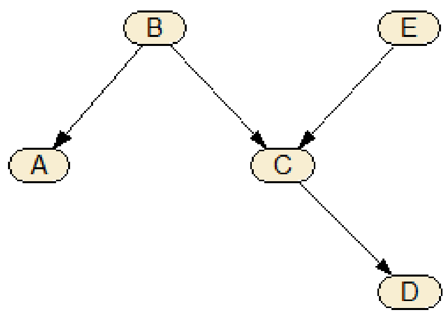
Ako postoji više od jednog minimalnog skupa, bilo koji minimalan skup može biti izabran da bude roditelj. Može biti više od jednog minimalnog skupa samo u slučaju kada su neki od prethodnika determinističke funkcije drugih. Sada, primenom pravila lanca i definicije roditelja dobijamo

$$P(X_1, X_2, \ldots, X_n) = \prod_{i=1}^{n} P\left(X_i|rod(X_i)\right).$$

Još ćemo da definišemo Markovljevo svojstvo za čvorove u grafu.

4.1.2 Definicija *Pretpostavimo da imamo zajedničku raspodelu verovatnoća slučajnih promenljivih P iz nekog određenog skupa V i imamo usmereni aciklični graf $G = (V, E)$. Kažemo da (G, P) zadovoljava Markovljevo svojstvo ako za svaku promenljivu $X \in V$, X je uslovno nezavisna od skupa svih njenih nepotomaka s obzirom na skup svih njenih roditelja. Kada (G, P) zadovoljava Markovljevo svojstvo, kažemo da G i P međusobno zadovoljavaju Markovljevo svojstvo.*

4.1.3 Primer U ovom primeru cemo pokazati kako se računa zajednička raspodela verovatnoća primenom pravila lanca proizvoda i pokazaćemo kako se računa zajednička raspodela pod uslovom da važi Markovljevo svojstvo.



Slika 5: Primer Bajesove mreže sa 5 čvorova.

Neka je totalno uređenje promenljivih za ovaj primer: (D, A, C, B, E). Zajednička raspodela verovatnoća promenljivih A, B, C, D i E datih ovom Bajesovom mrežom je sledeća:

$$P(D, A, C, B, E) = P(D|A, C, B, E)P(A|C, B, E)P(C|B, E)P(B|E)P(E).$$

Raspodela pod uslovom da važi Markovljevo svojstvo bi bila:

$$P(D, A, C, B, E) = P(D|C)P(A|B)P(C|B, E)P(B)P(E).$$

Obratimo pažnju da se zahvaljujući Markovljevom svojstvu redukuje broj parametara našeg modela. Primetimo da su B i D koreni, pa su zbog toga njihove verovatnoće samo $P(B)$ i $P(D)$, tj. nemamo uslovne verovatnoće.

Sada možemo da damo definiciju Bajesovih mreža.

4.1.4 Definicija *Neka je P zajednička raspodela verovatnoće slučajnih promenljivih iz skupa V, i $G = (V, E)$ aciklični usmereni graf. Kažemo da je (G, P)* ***Bajesova mreža*** *ako (G, P) zadovoljava Markovljevo svojstvo.*

Na osnovu svega navedenog, možemo da zaključimo da se Bajesove mreže sastoje od:

- Acikličnog usmerenog grafa, gde je svaki čvor slučajna promenljiva X_i;
- Domena svake slučajne promenljive (pod domenom podrazumevamo skup vrednosti slučajne promenljive);
- Skupa uslovnih raspodela verovatnoća datih sa $P(X_i|rod(X_i))$ za svaku promenljivu X_i.

Napomenuli smo da slučajne promenljive moraju biti diskretnog tipa, što znači da njihove vrednosti moraju biti jedan od sledećih tipova vrednosti:

- Bulove vrednosti - binarne vrednosti, tačno ($\top$) ili netačno ($\perp$);
- Uređene vrednosti koje mogu biti:
  - kategoričke - npr. { nisko, srednje, visoko};
  - numeričke - npr. ocena koja može biti { 1, 2, 3, 4, 5 };
- Realne vrednosti - npr. kada imamo čvor *Visina* koji predstavlja visinu osobe, tada su moguće vrednosti od npr. 50 do 250 cm.

Obratimo pažnju na sledeće bitne napomene u vezi Bajesovih mreža:

1. Moraju biti aciklične. Struktura acikličnog grafa nam garantuje da ne postoji čvor koji može biti sopstveni predak ili sopstveni potomak. Ovaj uslov je značajan i za faktorizaciju zajedničke raspodele verovatnoće.

2. Slučajne promenljive su date u fiksnom redosledu koji je određen pravilom lanca. Iste slučajne promenljive dekomponovane na različit način rezultiraju različitim Bajesovim mrežama.

3. Markovljevo svojstvo znatno olakšava računanje zajedničke raspodele verovatnoće, jer redukuje broj parametara datog modela.

4. Radi lakšeg razumevanja Bajesovih struktura, po konvenciji, obično su mreže predstavljene u vidu (naopakog) stabla, gde grane idu od vrha ka dole. Tada su koreni na vrhu, a listovi na dnu tog stabla.

5. Ako čvor ima mnogo roditelja ili ako se roditeljima može dodeliti veliki broj vrednosti, onda je tabela uslovnih verovatnoća veoma velika. Specijalno, ako imamo Bulove mreže gde promenljiva ima n roditelja, onda uslovnih verovatnoća ima 2^{n+1}. Zbog toga, za iole veće modele gotovo je nemoguće izračunati ,,peške" sve verovatnoće, pa je potrebna softverska podrška. Neki od softvera koji se koriste za Bajesove mreže su: BayesiaLab, GeNIe, Netica, Hugin, . . . Mi ćemo se za potrebe ovog rada poslužiti Netikom, koju ćemo detaljno objasniti u poglavlju 4.2.

Sada ćemo semantikom i verovatnoćama proširiti primer [4.1.1] preuzet iz [23].

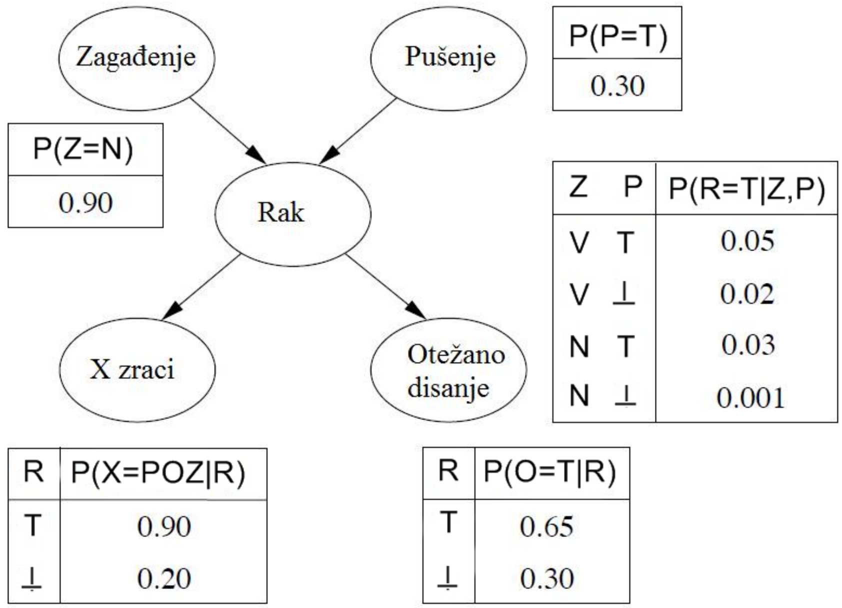
4.1.5 Primer Rak pluća. Pacijent pati od otežanog disanja i zbog toga posećuje lekara, jer sumnja da njegovi simptomi ukazuju na to da možda ima rak pluća. Lekar je svestan da isti simptomi mogu da ukazuju i na tuberkulozu i bronhitis, ali i na rak pluća. On zna da su mogući uzročnici tih simptoma sredina u kojoj boravi (zagađenje vazduha) i da li je pacijent pušac ili ne (što povećava šanse za rak pluća i bronhitis). Lekar šalje pacijenta na rendgen (X-zračenje), ako je rendgen pozitivan to ukazuje na tuberkulozu ili rak pluća.

Pogledajmo sada tipove i vrednosti slučajnih promenljivih za ovaj primer:

Ime čvora (oznaka)	**Tip**	**Vrednosti**
Zagađenje (Z)	Uređena	{ nisko, visoko }
Pušenje (P)	Bulova	$\{ \top, \bot \}$
Rak pluća (R)	Bulova	$\{ \top, \bot \}$
Otežano disanje (O)	Bulova	$\{ \top, \bot \}$
X-zraci (X)	Uređena	{ pozitivna, negativna }

Tabela 1: Tabela tipova i vrednosti slučajnih promenljivih za primer [4.1.5].

Ovakvim izborom tipova i vrednosti slučajnih promenljivih već imamo određeno šta će nam biti predstavljeno u mreži. Primetimo da svi čvorovi imaju samo dve vrednosti, to smo uzeli radi jednostavnijeg modela. Generalno, ne postoji ograničenje za diskretne vrednosti čvorova.



Slika 6: Bajesova mreža za primer [4.1.5].

Kao što smo već spomenuli u radu, prvo za svaki čvor tražimo sve moguće kombinacije vrednosti roditelja tih čvorova. Na primer, posmatrajmo čvor R. Njegovi roditelji su Z i P i oni mogu uzeti sledeće vrednosti $\{< V, \top >, < V, \bot >, < N, \top >, < N, \bot >\}$ (gde je V oznaka za visoko, a N oznaka za nisko zagađenje). Odnosno, kada pročitamo iz tabele što se nalazi pored čvora R to su sledeće verovatnoće: $< 0.05, 0.02, 0.03, 0.001 >$. Naravno, ovo su verovatnoće kada čvor R uzima vrednost tačno ($\top$). Verovatnoća da R uzima netačnu vrednost ($\bot$) pod uslovima Z i P je sledeća: $< 0.95, 0.98, 0.97, 0.999 >$ (dobijamo kao verovatnoću suprotnog događaja).

Još ćemo na ovom primeru izračunati zajedničku raspodelu verovatnoća za sledeće uzete vrednosti:

$$
\begin{aligned}
& P(X = poz \land O = \top \land R = \top \land Z = n \land P = \bot) = \\
& = P(X = poz | O = \top, R = \top, Z = n, P = \bot) \\
& \times P(O = \top | R = \top, Z = n, P = \bot) \\
& \times P(R = \top | Z = n, P = \bot) \\
& \times P(Z = n | P = \bot) \times P(P = \bot) = \\
& = P(X = poz | R = \top) \times P(O = \top | R = \top) \times P(R = \top | Z = n, P = \bot) \\
& \times P(Z = n) \times P(P = \bot) = \\
& = 0,9 \times 0,65 \times 0,001 \times 0,9 \times 0,7 = \\
& = 0,00036855.
\end{aligned}
$$

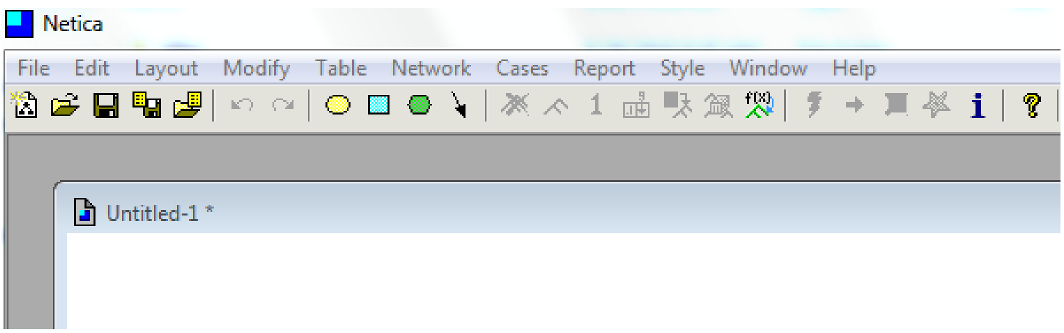
Time smo izračunali verovatnoću da će pacijentu rendgen biti pozitivan pod uslovima: da mu je ustanovljen rak pluća, da otežano diše, da nije pušač i da nije bio izložen visokom stepenu zagađenja.

Uočimo da je zahvaljujući Markovljevom svojstvu računanje zajedničke raspodele verovatnoće mnogo jednostavnije.

4.2 Netika

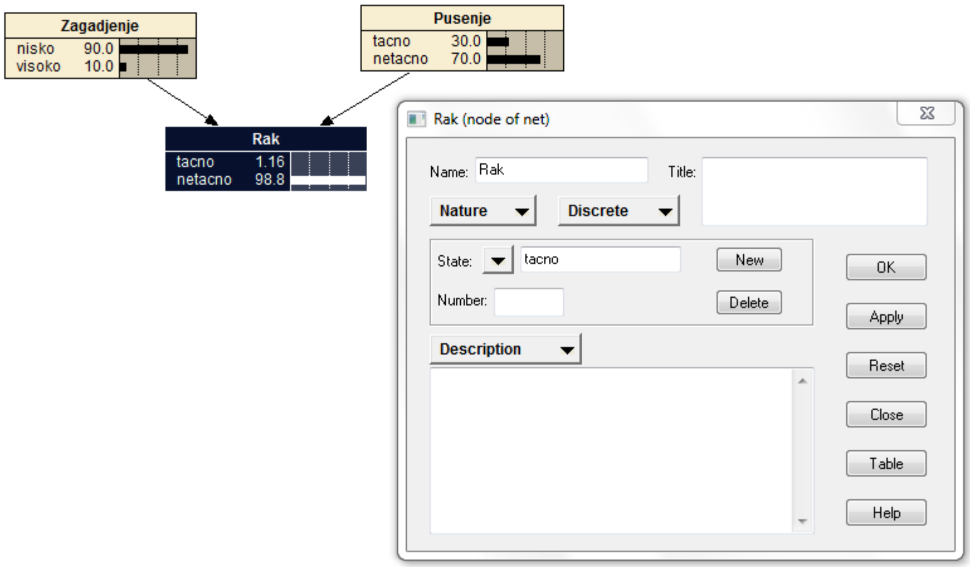
Netika (Netica) je moćan softver koji se koristi za rad sa Bajesovim mrežama. Produkt je Norsysa. Dizajniran je da bude jednostavan, pouzdan i lak za upotrebu zahvaljujući svom interfejsu za crtanje mreže. Za upravljanje nesigurnostima u poslovanju, inženjerstvu, medicini ili ekologiji, Netika je alat kojim se koriste mnoge od vodećih svetskih kompanija. Za opštu priču o Netici koristili smo [44].

Sada ćemo ukratko objasniti princip rada Netike (kako se crtaju čvorovi i zavisnosti među njima, kako se unose verovatnoće i kako na kraju dobijemo mrežu).



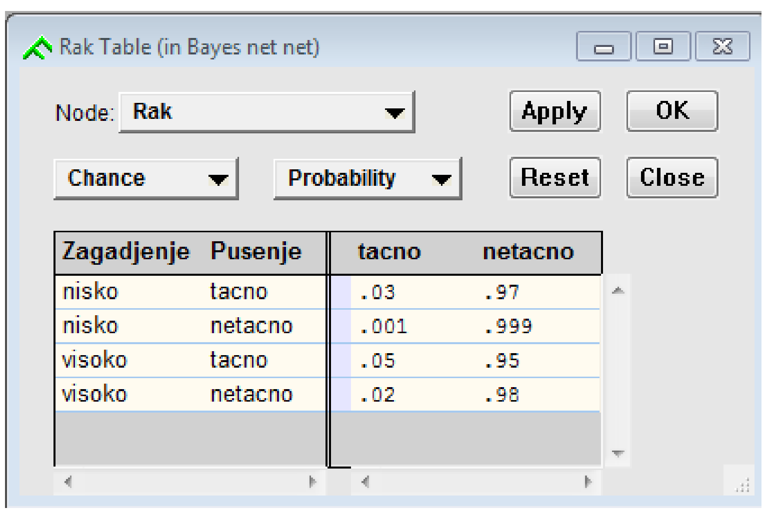
Slika 7: Početni prozor Netike.

Ovo je početni prozor Netike. Primetimo da imamo 3 mogućnosti za crtanje čvorova, poređanih respektivno u obliku: kruga, kvadrata i šestougla. To su, respektivno: čvorovi ,,prirode", čvorovi ,,odluke" i čvorovi ,,vrednosti". Čvorovi prirode predstavljaju promenljive na koje se može uticati od strane donosioca odluke. Čvorovi prirode se koriste za predstavljanje empirijskih ili izračunatih parametara i verovatnoća nastanka različitih stanja. Mi ćemo u Netici crtati mreže preko čvorova prirode. Čvorovi odluke predstavljaju upravljačke promenljive koje mogu direktno biti implementirane od strane donosioca odluke. Ovi čvorovi predstavljaju skup raspoloživih upravljačkih akcija. Usko su povezani sa čvorovima vrednosti, koji se koriste za procenu pravila optimalne odluke na mreži koja će maksimizirati sumu očekivanih vrednosti čvorova vrednosti. Pored toga nalazi se strelica pomoću koje se crtaju zavisnosti između čvorova. Odnosi između čvorova se mogu uneti kao pojedinačne verovatnoće, u obliku jednačina, ili se mogu preuzeti iz neke datoteke sa podacima. Mi ćemo odnose unositi kao verovatnoće. Dvoklikom na čvor otvori se prozor u kojem možemo da: promenimo ime čvora, odredimo da li će biti diskretnog ili neprekidnog tipa, unesemo verovatnoće, pogledamo tabelu unetih verovatnoća, ...



Slika 8: Prozor čvora mreže u Netici.

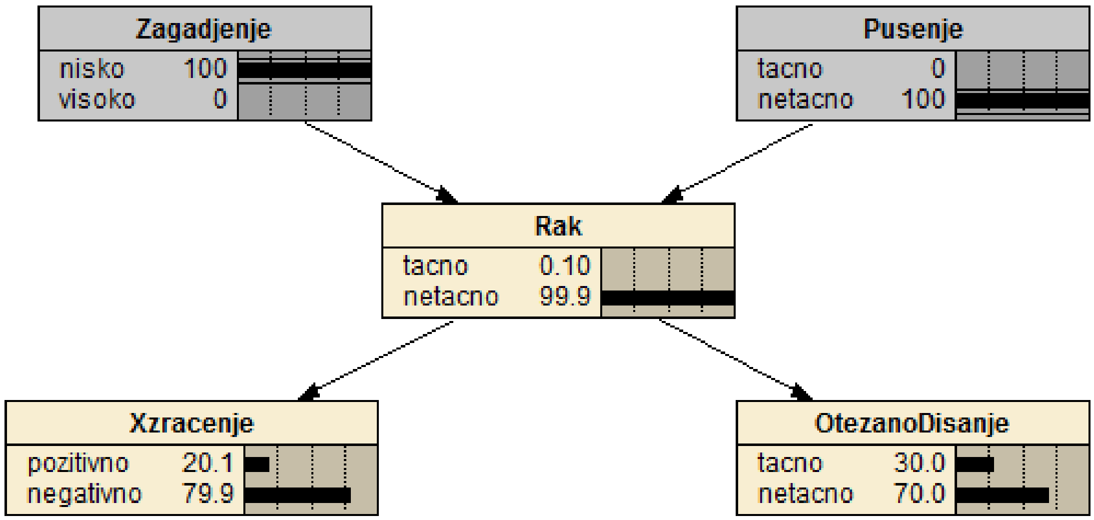
Kada unesemo sve čvorove i odnose među njima, i kada unesemo verovatnoće, tada pritisnemo Compile net (liči na munju) i imamo gotovu mrežu. Napomenimo da Netika sama izračuna uslovne verovatnoće pomoću unetih verovatnoća.



Zagadjenje	Pusenje	tacno	netacno
nisko	tacno	.03	.97
nisko	netacno	.001	.999
visoko	tacno	.05	.95
visoko	netacno	.02	.98

Slika 9: Prozor tabele sa verovatnoćama u Netici.

Kada smo dobili gotovu mrežu, možemo na nekom čvoru da menjamo verovatnoće, a program će sam menjati verovatnoće ostalih čvorova koji zavise od tog čvora. Promenu verovatnoća ćemo demonstrirati na primeru raka pluća, pod pretpostavkom da pacijent nije pušač i da nije bio izložen zagađenju. Primetimo da smo ovu verovatnoću već izračunali, kao jedan od činilaca kada smo računali zajedničku raspodelu verovatnoće na primeru [4.1.5]. Primetimo da u prozoru za verovatnoće zadajemo verovatnoće za zagađenje i pušenje, a u ostalim čvorovima ne menjamo verovatnoće, već Netika sama promeni verovatnoće pod novim zadatim uslovima.



Slika 10: Primer [4.1.5] nacrtan u Netici sa zadatim uslovima.

Netika sadrži kao opciju mogućnost da znanje koje ima o jednoj mreži prenosi na druge mreže sečenjem i lepljenjem, ili da sačuva u modularnoj formi stvaranjem biblioteke čvorova. Naravno, mreže i biblioteke mogu da se sačuvaju u fajlovima ili se mogu štampati.

Netika može da koristi mreže za obavljanje različitih vrsta zaključivanja koristeći savremene algoritme. Netika će naći odgovarajuće vrednosti ili verovatnoće za sve nepoznate promenljive. Ove vrednosti ili verovatnoće mogu biti prikazane na više različitih načina. Slučajevi mogu biti sačuvani u datoteci, a kasnije se mogu vratiti u mrežu (ili drugu mrežu) za dalju pretragu, ili mogu da uzmu u obzir nove informacije o slučaju. Netika može koristeći dijagrame da pronađe optimalne vrednosti koje uvećavaju očekivane vrednosti navedenih promenljivih. Takođe, može izgraditi uslovne planove, jer odluke u budućnosti mogu zavisiti od uočenih zapažanja.

4.3 Zaključivanje pomoću Bajesovih mreža

U ovom poglavlju ćemo pokazati kako predstavljene informacije u Bajesovoj mreži ,,dekodiramo”, odnosno do kojih zaključaka dolazimo. Za literaturu smo koristili [23].

4.3.1 Razumevanje domena preko Bajesovih mreži

Već smo objasnili kako se domen i verovatnoće mogu predstaviti preko Bajesovih mreža. Sada nam je zadatak da shvatimo kako pomoću Bajesovih mreža možemo doći do zaključaka o domenu. Kada posmatramo vrednost neke promenljive, mi želimo da je uslovimo na novu informaciju, tj. da zaključimo nešto o verovatnoći te vrednosti na osnovu poznatih vrednosti drugih promenljivih u mreži. Takav proces uslovljavanja (zaključak) se vrši preko niza informacija u mreži.

Vrste čvorova

Čvorovi mogu biti *upitni* (eng. query) i *dokazani* (eng. evidence) čvorovi. Mi pomoću datih vrednosti za neke dokazane (posmatrane) čvorove, tražimo verovatnoću za upitne čvorove.

Za dokazane čvorove uzimamo one čvorove koji su određeni (imaju određenu vrednost), npr. čvor X ima vrednost x ($X = x$). Ovo nazivamo i *konkretni dokazi*. Npr. pretpostavimo da smo otkrili da je pacijent pušač, tada je čvor P ($P = \top$) konkretan dokaz.

Ponekad postoje dokazi koji su dvosmisleni. Npr. neka čvor X može da ima vrednosti: x_1, x_2, x_3 i x_4. Znamo da x_3 i x_4 sigurno nisu vrednosti, ali ne znamo da li je x_1 ili x_2. Takvi čvorovi se nazivaju *negativni dokazi*.

Vrste rezonovanja

Pomoću kretanja kroz mrežu vidimo kako se menjaju verovatnoće dodeljene pojedinim promenljivama, što nazivamo *rezonovanje*. U zavisnosti od načina kretanja kroz mrežu razlikujemo 4 vrste rezonovanja. To su: *intuitivno rezonovanje, dijagnostičko rezonovanje, uzročno rezonovanje* i *kombinovano rezonovanje*. Objasnićemo ih kroz primer raka pluća [4.1.5].

- *Dijagnostičko rezonovanje* se javlja kada su dokazani čvorovi potomci, a upitni čvorovi preci. Odnosno, koristeći medicinsku terminologiju, kada hoćemo da damo dijagnozu za date simptome. Npr. kada pacijent ima simptom otežanog disanja (potomak), lekar tada daje pretpostavku dijagnoze o raku pluća (predak) i da li je pacijent pušač (predak).

- *Intuitivno rezonovanje* predstavlja oblik zaključivanja kada na osnovu dokazanih čvorova (predaka) želimo da zaključimo o upitnim čvorovima (potomcima). Npr. kada pacijent kaže lekaru da je pušac (roditelj), a ne spominje nijedan simptom, lekar zna da to povećava šanse da pacijent ima rak pluća (dete).

- *Uzročno rezonovanje* daje obrazloženje o uzajamnim uzrocima zajedničkog uticaja, tj. o 2 roditelja jednog deteta. Pretpostavimo da postoje tačno dva moguća uzroka određenog uticaja. Oni su predstavljeni kao v-struktura u Bajesovoj mreži. U primeru uzroci su pušenje i zagađenje, a njihov zajednički uticaj je rak. Na prvi pogled, mi zaključujemo da su ta dva uzroka međusobno nezavisna. Međutim, uočimo njihovu zavisnost kroz sledeće 2 situacije:

  - U prvoj situaciji, ukoliko je dokazani čvor dete, ta činjenica povećava verovatnoću za oba roditelja. Npr. kada pacijent ima rak pluća, lekar zna da to povećava šanse da je pacijent pušač i da je bio izložen zagađenju.

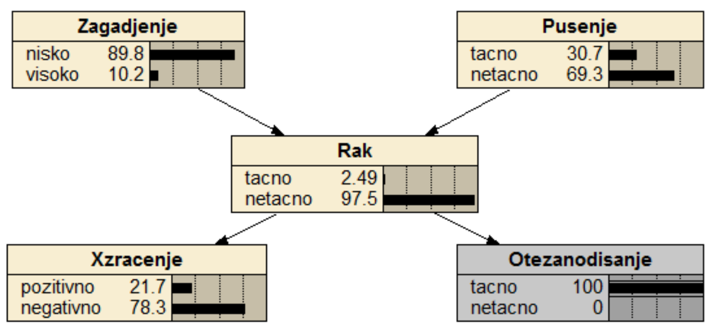
- U drugoj situaciji, ukoliko ima 2 dokazana čvora (dete i jedan roditelj), ta činjenica utiče na verovatnoću drugog roditelja. Npr. kada je pacijent pušač i ima rak pluća, to zauzvrat smanjuje verovatnoću da je pacijent bio izložen visokom stepenu zagađenja. Dakle, iako su ova dva uzroka na početku bila nezavisna, sa znanjem njihovog zajedničkog uticaja prisustvo jednog uzroka smanjuje verovatnoću drugog uzroka.

- *Kombinovano rezonovanje* obuhvata situacije koje nisu uključene u prethodne 3 vrste rezonovanja. Npr. kada pacijent kaže lekaru da je pušac (predak) i da ima simptom otežanog disanja (potomak), lekar zna da su tada velike šanse da pacijent ima rak pluća.

Sada ćemo ilustrovati ove 4 vrste rezonovanja u Netici. Dokazani čvorovi će biti obeleženi sivom bojom, a upitni narandžastom bojom.

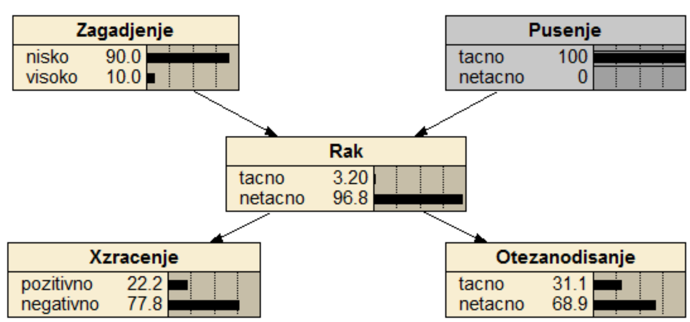
4.3.1 Primer

a)



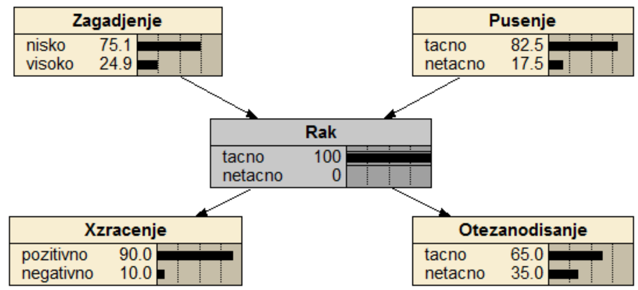
Slika 11: Primer dijagnostičkog rezonovanja u kojem su upitni čvorovi rak i pušenje.

b)



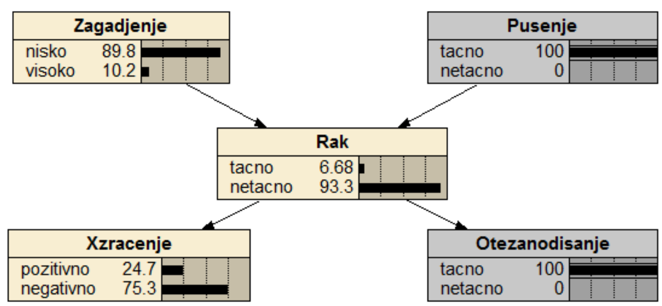
Slika 12: Primer intuitivnog rezonovanja u kojem je upitni čvor rak.

c)



Slika 13: Primer uzročnog rezonovanja u kojem su upitni čvorovi zagađenje i pušenje.

d)

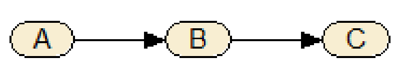


Slika 14: Primer kombinovanog rezonovanja u kojem je upitni čvor rak.

4.3.2 Uslovne nezavisnosti i d-razdvajanje

U Bajesovoj mreži čvor time što postaje dokazan, nekad može izazvati uslovnu nezavisnost između pojedinih čvorova. Uslovne nezavisnosti mogu biti interpretirane u vidu uzročnog lanca, zajedničkog uzroka i zajedničkog uticaja.

Uzročni lanci



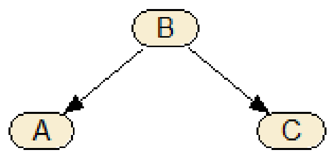
Slika 15: Uzročni lanac.

Na slici 15 je predstavljen uzročni lanac od tri čvora, gde A uzrokuje B, a B uzrokuje C. Uzročni lanac nam daje uslovnu nezavisnost, jer se verovatnoća čvora C računa na sledeći način:

$$P(C|A \wedge B) = P(C|B).$$

To u stvari znači da je verovatnoća čvora C određena čvorom B ista kao i verovatnoća čvora C određena čvorovima A i B. Odnosno, ako je B dokazan čvor, to znači da A ne utiče na naše zaključivanje o C. Ovakvu uslovnu nezavisnost obeležavamo $A \perp C|B$ i čitamo kao A i C su uslovno nezavisni u odnosu na B.

Zajednički uzrok

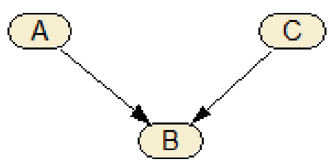


Slika 16: Zajednički uzrok.

Zajednički uzroci (zajednički preci), u slučaju da su dokazani čvorovi, daju nam uslovnu nezavisnost između čvorova dece. Slika 16 nam pokazuje da promenljive A i C imaju zajednički uzrok B. Zajednički uzroci daju nam istu uslovnu nezavisnost kao i uzročni lanci:

$$P(C|A \wedge B) = P(C|B) \equiv A \perp C|B.$$

Zajednički uticaj



Slika 17: Zajednički uticaj.

Zajednički uticaj predstavlja strukturu u kojoj dete ima dva roditelja. Slika 17 prikazuje situaciju gde čvor ima dva uzroka. Zajednički uticaji ili njihovi potomci se ponašaju suprotno u odnosu na uzročne lance i zajedničke uzroke i stvaraju uslovno zavisnu strukturu. Roditelji koji su po osnovnoj strukturi nezavisni ($A \perp C$), kada je zajedničko dete dokazani čvor, postaju *uslovno zavisni*:

$$P(A|C \wedge B) \neq P(A|B) \equiv A \not\perp C|B.$$

Ovo nam govori da ako posmatramo zajednički uticaj (dete), i ako je jedan od uzroka (roditelja) odsutan, time se povećava verovatnoća drugog uzroka.

Ovim smo objasnili kako Bajesove mreže pokazuju uslovne nezavisnosti, ali i kako te nezavisnosti utiču na promenu jačine verovanja.

Uslovna nezavisnost $A \perp C|B$ pokazuje da ako znamo vrednost od B, onda B *blokira* informaciju o C bitnu za A, i obrnuto. Osim što je primenljivo između parova čvorova, blokiranje pomoću uslovne nezavisnosti primenjuje se i između skupova čvorova. Uopštenije, ako znamo da važi Markovljevo svojstvo, možemo utvrditi da li je skup čvorova X nezavisan od skupa čvorova Y u odnosu na skup dokazanih čvorova E. Sada ćemo definisati put, blokiranje i d-razdvajanje.

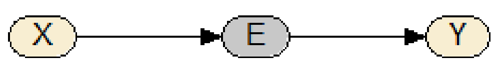
4.3.2 Definicija ***Put*** *između dva skupa čvorova X i Y je bilo koji niz čvorova između člana skupa X i člana skupa Y takav da je svaki susedni par čvorova povezan granom (bez obzira na usmerenje) i ne postoji čvor koji se u nizu pojavljuje dva puta.*

4.3.3 Definicija *Neka su X, Y i E tri disjunktna podskupa čvorova u usmerenom acikličnom grafu, i neka je p put od čvora iz X do čvora u Y. Tada kažemo da E* ***blokira*** *put p ako postoji čvor Z na putu p koji zadovoljava neki od sledećih uslova:*

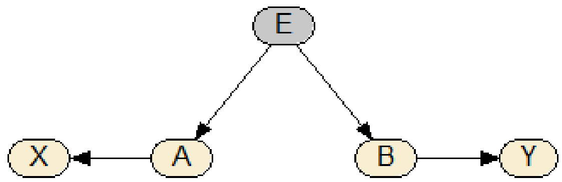
- *Z je u E i jedna grana vodi u Z, a jedna iz Z (lanac);*
- *Z je u E i obe grane vode iz Z (zajednički uzrok);*
- *Ni Z ni bilo koji njegov potomak nije u E, i obe grane vode u Z (zajednički uticaj).*

4.3.4 Definicija *Skup čvorova E* ***d-razdvaja*** *skupove čvorova X i Y ($X \perp Y|E$), ako je svaki put od čvora iz X do čvora iz Y blokiran sa E.*

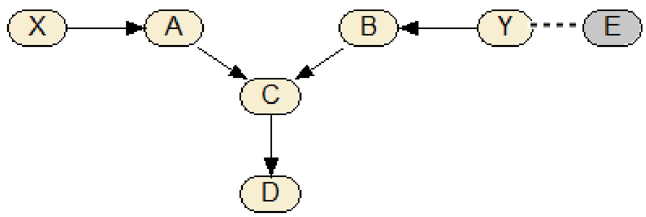
Sada ćemo ilustrativno prikazati gore navedene situacije blokiranja.



Slika 18: Prvi uslov blokiranja.



Slika 19: Drugi uslov blokiranja.



Slika 20: Treći uslov blokiranja.

Primenićemo još d-razdvajanje na primer [4.1.5]. Pretpostavimo da je čvor R dokaz čvor. Tada:

1. Z je d-razdvojeno od X i O (prvi uslov blokiranja);

2. X je d-razdvojeno od O (drugi uslov blokiranja);

3. Ako za dokaze ne uzmemo R, X i O, onda je P d-razdvojeno od Z (treći uslov blokiranja).

Još ćemo reći nešto o vrstama Bajesovih mreža [11]. U mnogim praktičnim primerima struktura Bajesovih mreža je nepoznata i zahteva učenje iz podataka. Ovaj problem je poznat kao problem *Bajesovog učenja*. Bajesovo učenje je vezano za algoritme učenja koji koriste verovatnoću i statistiku kao model. Problem Bajesovog učenja smatra se težim od učenja parametara Bajesovih mreža. U specijalne vrste Bajesovih mreža spadaju i skriveni Markovljevi modeli, neuronske mreže i Kalmanovi filteri. Takođe, specifične vrste Bajesovih mreža su *dinamičke Bajesove mreže*, koje se baziraju na stohastičkim procesima i HMM modelima, i *funkcionalne Bajesove mreže*. Detaljnija priča o svakoj od vrsta prevazilazi okvire ovog rada, ali mi ćemo u narednom poglavlju reći nešto o dinamičkim Bajesovim mrežama.

4.4 Dinamičke Bajesove mreže

Pre nego što pređemo na primenu Bajesovih mreža, reći ćemo nešto o dinamičkim Bajesovim mrežama. Više se može naći u [5, 8, 27, 31, 32]. Dinamičke Bajesove mreže predstavljaju proširenje Bajesovih mreža na modelovanje dinamičkih sistema. Dakle, izraz ,,dinamičke" znači samo da radimo u dinamičkim sistemima, a ne da se mreža menja kroz vreme. Napomenimo da radimo samo sa diskretnim stohastičkim procesima, tako da svaki put kad povećamo t za jedan imamo novo posmatranje. Neka imamo slučajne promenljive $\{Z_1, Z_2, \ldots, Z_n\}$. U dinamičkim Bajesovim mrežama, stanje u trenutku t je predstavljeno preko skupa slučajnih promenljivih $Z_t = (Z_{1,t}, \ldots, Z_{n,t})$. Stanje u vremenu t je zavisno od stanja u prethodnim vremenskim intervalima. Pretpostavimo da svako stanje neposredno zavisi samo od prethodnog stanja (Markovljev lanac prvog reda), i tako treba da predstavimo prelazne verovatnoće $P(Z_t|Z_{t-1})$. To možemo da uradimo pomoću dva vremenska dela Bajesove mreže, koja sadrže promenljive iz Z_t čiji su roditelji promenljive iz Z_{t-1} i/ili Z_t i promenljive iz Z_{t-1} bez svojih roditelja.

Dinamička Bajesova mreža je par $(B_1, B_{\rightarrow})$, gde je B_1 Bajesova mreža koja definiše verovatnoću $P(Z_1^{1:N})$, a $(B_{\rightarrow})$ predstavlja Bajesovu mrežu sa dva vremenska dela koja je definisana sa $P(Z_t|Z_{t-1})$ što u direktnom acikličnom grafu znači sledeće:

$$P(Z_t|Z_{t-1}) = \prod_{i=1}^{N} P(Z_t^i|Rod(Z_t^i))$$

gde je Z_t^i i-ti čvor u delu t (koji može biti skriveni ili opažajući, stoga je $N = N_s + N_o$) dok su $Rod(Z_t^i)$ čvora Z_t^i u grafu. Čvorovi u prvom delu Bajesove mreže nemaju nikakve parametre, ali svaki čvor u drugom delu ima svoju uslovnu raspodelu verovatnoće koju definiše $P(Z_t^i|Rod(Z_t^i))$ za svako $t > 1$. Roditelji čvora, $Rod(Z_t^i)$, mogu biti u istom vremenskom delu ili u prethodnom vremenskom delu. Lukovi između delova su od leva na desno i predstavljaju uzročni protok vremena. Ako postoji grana od Z_{t-1}^i do Z_t^i, onda se čvor Z_t^i naziva *uporan čvor*. Grane između delova su proizvoljne, sve dok je dinamička Bajesova mreža direktni aciklični graf. Intuitivno, usmerene grane unutar delova predstavljaju trenutne uzročnosti. Ako koristimo neusmerene grane unutar delova onda imamo model ograničenja, a ne model uzroka. Takav model se naziva *dinamički lanac grafa*.

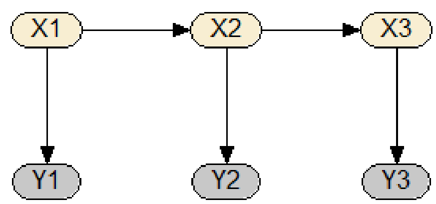
Pretpostavimo da su parametri uslovnih raspodela verovatnoće vremenski invarijantni. Ako se parametri mogu promeniti, možemo ih dodati na stanja i tretirati ih kao slučajne promenljive. Ako postoji ograničen skup mogućih vrednosti parametara, možemo dodati skrivenu promenljivu koja određuje koji skup parametara se koristi.

Zajednička raspodela za niz dužine T može se dobiti ,,razmotavanjem" mreže koja ima T delova, a zatim se pomnožiti sa uslovnim raspodelama verovatnoće:

$$P(Z_{1:T}^{1:N}) = \prod_{i=1}^{N} P_{B_1}(Z_1^i|Rod(Z_t^i)) \times \prod_{t=2}^{T}\prod_{i=1}^{N} P_{B_{\rightarrow}}(Z_t^i|Rod(Z_t^i)).$$

Sada ćemo, radi lakšeg razumevanja, dati primer Bajesove mreže sa dva vremenska dela za standardni HMM za dužinu niza $T = 3$.

4.4.1 Primer

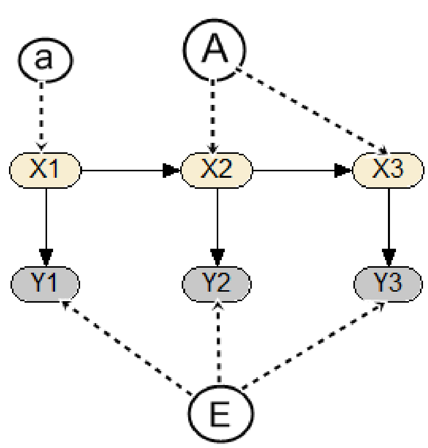


Slika 21: Primer dinamičke Bajesove mrerže.

Iz ovog primera vidimo da HMM možemo predstaviti kao dinamičku Bajesovu mrežu. Čvorovi koji su osenčeni su posmatrane promenljive, dok su neosenčeni čvorovi skrivene promenljive. Ovaj graf predstavlja sledeću uslovno nezavisnu pretpostavku: $X_{t+1} \perp X_{t-1} | X_t$ (Markovljevo svojstvo) i $Y_t \perp Y_{t'} | X_t$, za $t \neq t'$.

Razlika između HMM i dinamičkih Bajesovih mreža je ta što dinamičke Bajesove mreže predstavljaju skrivena stanja u smislu skupa slučajnih promenljivih $X_t^1, \ldots, X_t^{N_s}$. U suprotnom, u HMM, stanja su predstavljena kao jedna slučajna promenljiva X_t.

Sada ćemo isti taj primer proširiti tako što ćemo na grafu docrtati parametre, koji su spojeni isprekidanim granama.



Slika 22: Dopunjen primer [4.4.1].

Parametri su $P(X_1 = i) = \pi(i) = a_i$, $P(X_t = j|X_{t-1} = i) = A(i,j)$ i $P(Y_t = j|X_t = i) = E(i,j)$.

Glava 5

Primena Bajesovih mreža

U ovoj glavi ćemo navesti nekoliko od mnogobrojnih primera primena Bajesovih mreža. Bajesove mreže imaju veliku primenu u računarstvu, medicini, sportu, ekonomiji, kao i u prepoznavanju govora, kontroli vozila, ali i u naukama kao što su biologija, psihologija, meteorologija,... Primeri se mogu naći u [4, 32]. Detaljnije ćemo obraditi primenu u sportu, konkretno u tenisu. Daćemo originalan primer primene Bajesovih mreža na osvojen poen u tenisu i daćemo objašnjenje primera preuzetog iz [24].

- **računarstvo**
  - *kompjuterska vizija*: Pham i saradnici su 2002. razvili sistem za detekciju lica.
  - *kompjuterske igrice*: Valadares je 2002. razvio računarsku igricu koja je model evolucije simuliranog sveta.
  - *softver*: Microsoft je otkrio veliki broj aplikacija. Od 1995. Microsoft Office je koristio Bajesove mreže za odabir pomoći na osnovu upita.
- **medicina**
  - Galan je 2002. kreirao takozvani NasoNet, koji predstavlja sistem koji vrši prognozu i dijagnozu nazofaringealnog karcinoma. Širenje kancera je nedeterministički dinamičan proces, pa shodno tome dizajn modela za dijagnozu i prognozu stepena raka treba da se zasniva na metodi koju čine neizvesnost i vreme. Mreža verovatnosnih događaja u diskretnom vremenu je dinamička Bajesova mreža, koja kaže da su mehanizmi modela uzročno povezani sa vremenom evolucije procesa. Model je dovoljno snažan da može da se primeni na bilo koju drugu vrstu kancera.
  - Sakellaropoulos sa svojim saradnicima je 1999. razvio sistem za prognozu povreda glave.
  - Mani i saradnici su 1997. razvili takozvani MENTOR sistem. Sistem predviđa mentalnu obolelost kod novorođenčadi.

  - Ogunyemi i njegovi saradnici su 2002. razvili TraumaSCAN, što je prototip računarskog sistema koji vrši procenu efekata prodiranja traume u grudima i stomaku.

- **ekonomija**
  - Lažno finansijsko izveštavanje obuhvata: manipulacije, falsifikovanje ili izmenu računovodstvene evidencije i prateće dokumentacije. Lažno finansijsko izveštavanje vodi povećanju informacionog rizika pod kojim se podrazumeva da finansijske informacije koje se koriste za donošenje poslovnih odluka mogu da budu netačne i nepouzdane. Bajesova mreža se pokazala uspešnom za otkrivanje lažnog finansijskog izveštavanja. Ona identifikuje lažne finansijske izveštaje sa verovatnoćom od 0,9.

- **prepoznavanje govora**
  - Bilmes je 2000. primenio dinamičke Bajesove mreže na prepoznavanje govora.
  - Nefian je sa svojim saradnicima 2002. razvio sistem za audio-vizuelno prepoznavanje govora.

- **biologija**
  - Friedman je 2000. razvio tehniku za učenje uzročno-posledične veze među genima, analizirajući podatke ekspresije gena, koja predstavlja proces kojim se informacija gena koristi za sintezu funkcionalnog genskog produkta. Ova tehnika je rezultat projekta u kojem su primenjene dinamičke Bajesove mreže na analizu ekspresije gena.
  - Friedman je 2002. razvio metod za rekonstrukciju filogenetskih drveća. Taj metod se koristi u SEMPHY, što predstavlja sredstvo za računanje maksimalne verovatnoće filogenetskih rekonstrukcija.

Primeni Bajesovih mreža u sportu ćemo se posvetiti u poglavlju 5.1 u kojem ćemo dati konkretne primere primene na tenis.

5.1 Primena Bajesovih mreža na tenis

U ovom poglavlju ćemo dati 2 primera primene Bajesovih mreža na poene u tenisu. Oba primera predstavljaju primenu na servis i retern. Prvi naveden primer je originalan, dok je drugi preuzet iz [24] i koristi MuktiSkill i TrueSkil sisteme koje ćemo malo detaljnije objasniti.

5.1.1 Primer Poen se odigrava između igrača 1 i igrača 2. Posmatramo poene samo u kojima servira igrač 1. Za ovu Bajesovu mrežu slučajne promenljive su: *prvi servis, retern prvog servisa, osvojen poen pri prvom servisu, drugi servis, retern drugog servisa* i *osvojen poen pri drugom servisu*. Oznake koje ćemo koristiti u tabeli verovatnoća za ove slučajne promenljive su respektivno: $S1, R1, P1, S2, R2$ i $P2$. Napomenimo da su sve promenljive binarne. Za potrebe ovog primera verovatnoće su uzete iz statistike meča Huan Martin Del Potro-Novak Đoković odigranog na Olimpijskim igrama u Riu 2016. godine, gde je na servisu Del Potro.

S1	**R1**	**P1**
P(S1=u)=0,58	P(R1=p|S1=u)=0,84	P(P1=⊤$\|S1 = u, P1 = p) = 0,86$
P(S1=nu)=0,42	P(R1=p|S1=nu)=0	P(P1=⊤$\|S1 = u, R1 = np) = 1$

Tabela 2: Tabela verovatnoća za primer [5.1.1].

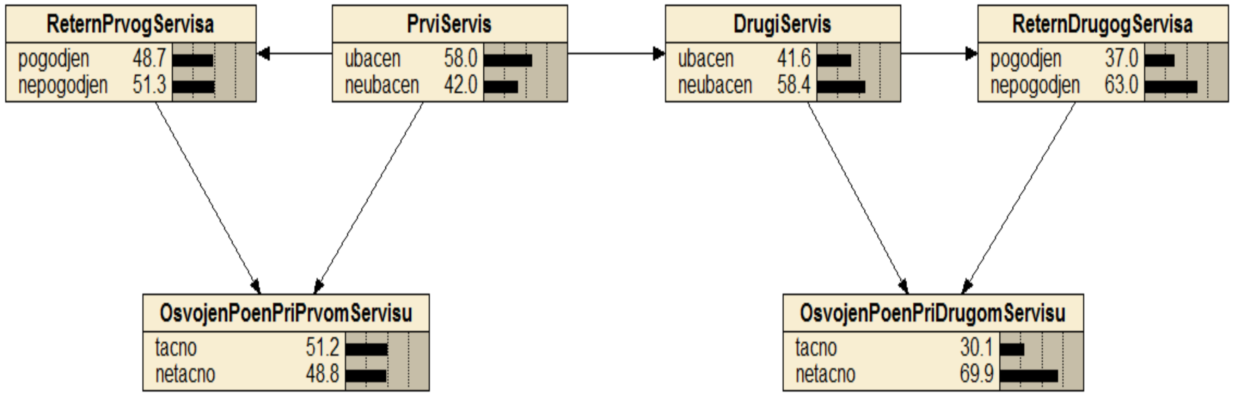
S2	**R2**	**P2**
P(S2=u|S1=u)=0	P(R2=p|S2=u)=0,89	P(P2=⊤$\|S2 = u, R2 = p) = 0,69$
P(S2=u|S1=nu)=0.99	P(R2=p|S2=nu)=0	P(P2=⊤$\|S2 = u, R2 = nu) = 1$

Tabela 3: Tabela verovatnoća za primer [5.1.1].

Napomenimo da smo za vrednosti slučajnih promenljivih koristili skraćene oznake: u-ubačen, nu-neubačen, p-pogođen i np-nepogođen.

Primetimo da drugi servis zavisi od prvog servisa i da retern zavisi od servisa. Osvojen poen pri prvom servisu zavisi od prvog servisa i od reterna prvog servisa, analogno za osvojen poen pri drugom servisu. Primetimo da čvor osvojen poen pri prvom servisu pokazuje s kojom verovatnoćom je Del Potro osvojio poen pod uslovom da je ubacio prvi servis. Analogno za osvojen poen pri drugom servisu.

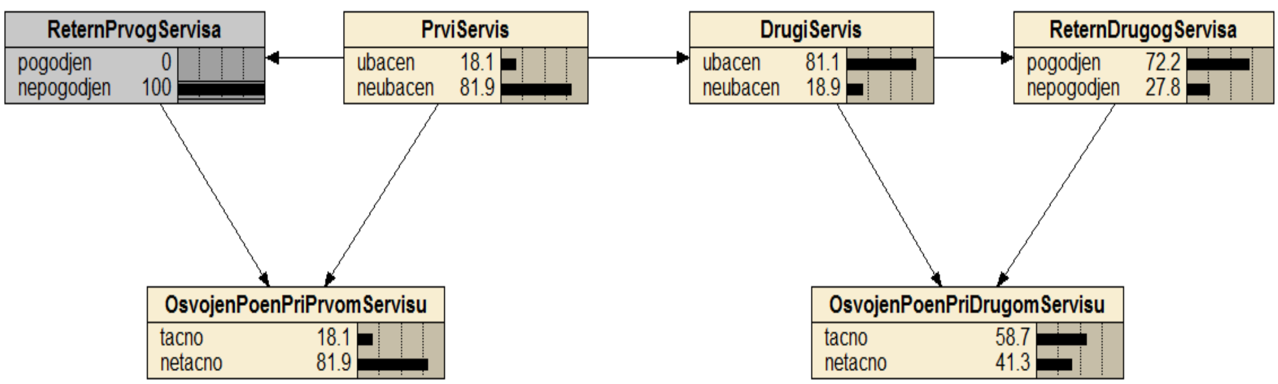
Mreža za ovaj primer nacrtana u Netici izgleda ovako:



Slika 23: Bajesova mreža koja predstavlja poen u tenisu.

Sada pomoću Netike lako možemo da izračunamo verovatnoće koje želimo. Ilustrovaćemo na par primera.

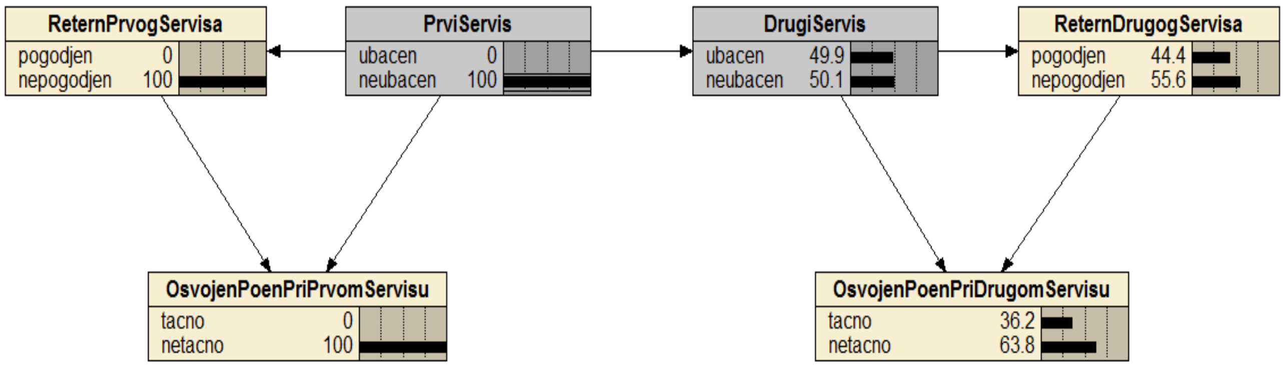
5.1.2 Primer Pitamo se kolika je verovatnoća da Đoković osvoji poen ako znamo da nije pogodio retern prvog servisa?



Slika 24: Bajesova mreža u kojoj je dokazani čvor retern prvog servisa.

Pogledajmo tako dobijene verovatnoće. S obzirom da Đoković nije pogodio retern prvog servisa, verovatnoća da je Del Potro uopšte ubacio prvi servis se smanjila na 0,18 (što predstavlja i verovatnoću da je Del Potro osvojio poen pod ovim uslovom). Pogledajmo šta se dešava sa verovatnoćom drugog servisa. Verovatnoća da je Del Potro ubacio drugi servis sada iznosi 0,81, a verovatnoća da je osvojio taj poen je 0,59. Dakle, odgovor na naše pitanje je sledeći: verovatnoća da je Đoković osvojio poen pod uslovom da nije pogodio retern prvog servisa je 0,41.

5.1.3 Primer Kolika je verovatnoća da je Del Potro osvojio poen pri drugom servisu, pod uslovima da nije ubacio prvi servis, a da je drugi servis ubacio sa verovatnoćom od 0,5?

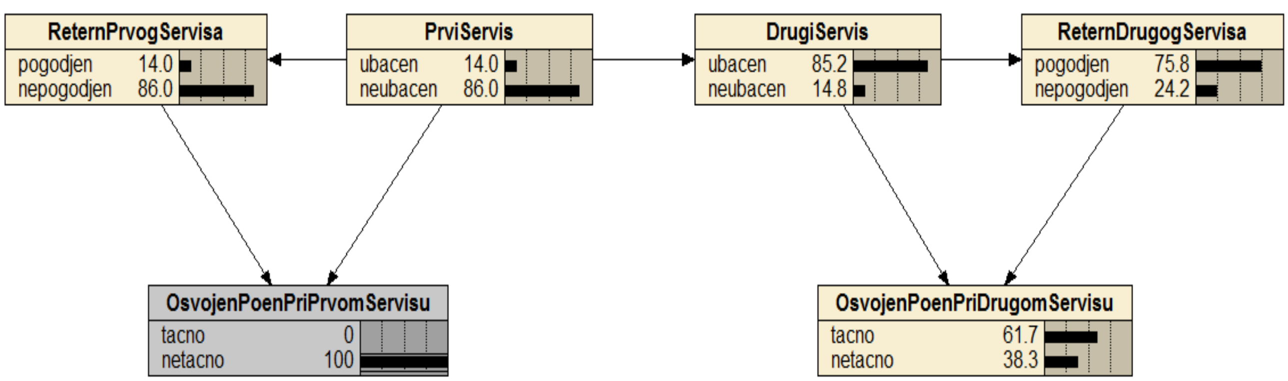


Slika 25: Bajesova mreža u kojoj su dokazani čvorovi prvi i drugi servis.

Iz mreže vidimo da je verovatnoća da je Đoković pogodio retern drugog servisa pod datim uslovom 0,45 i vidimo da je verovatnoća da je Del Potro osvojio taj poen oko 0,36.

Hajde još da vidimo šta je sa verovatnoćama ako znamo da Del Potro nije osvojio poen pri prvom servisu.

5.1.4 Primer Pitamo se kolika je verovatnoća da je Del Potro osvojio poen pri drugom servisu ako sigurno znamo da nije osvojio poen pri prvom servisu?



Slika 26: Bajesova mreža u kojoj je dokazani čvor osvojen poen pri prvom servisu.

Kao što možemo da vidimo iz dobijene mreže, verovatnoća da je Del Potro uopšte ubacio prvi servis se pod ovim uslovom smanjila na 0,14, dok verovatnoća da je ubacio drugi servis sada iznosi oko 0,85. Del Potro će pod ovim uslovom osvojiti poen s verovatnoćom oko 0,62.

Sada ćemo dati primer Bajesove mreže koja predstavlja poen u tenisu *MultiSkill sistema*. MultiSkill je sistem za procenu sposobnosti igrača u tenisu baziran na dinamičkim Bajesovim mrežama. MultiSkill sistem se bazira na *TrueSkill sistemu*, Bajesovoj mreži koja predstavlja rejting igrača preko jedne sposobnosti. Upravo to predstavlja razliku između MultiSkill i TrueSkill sistema. U MultiSkill sistemu svaki igrač ima više sposobnosti koje variraju u zavisnosti od uslova određene igre. Na primer, teniseri pokazuju različite sposobnosti u pogledu na njihove veštine za servis, retern i aseve. Ove sposobnosti jednog igrača mogu se intuitivno smatrati kao kombinacija pojedinačnih sposobnosti više pojedinih igrača koji čine ekipu u TrueSkill sistemu.

Svrha sistema za rangiranje je da prepozna i procenjuje sposobnosti igrača svaki put kada igrač odigra utakmicu. Na osnovu sposobnosti igrača sistem stalno menja percepciju. TrueSkill se koristi samo u konačnom plasmanu svih timova u igri, kako bi se ažurirale procene sposobnosti za sve igrače koji učestvuju u toj igri. Ovakvi sistemi za rejting mogu da se koriste u mnogim sportovima. Literatura koju smo koristili je [17, 24].

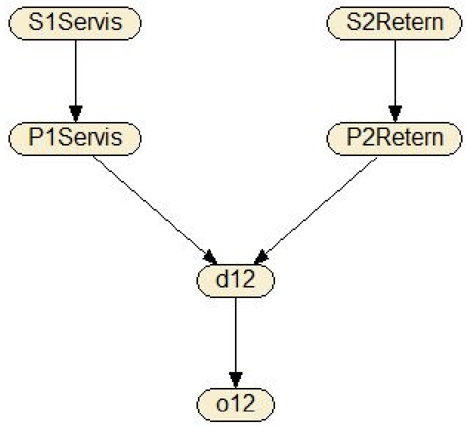
Česta promena neizvesnosti omogućava sistemu da rano pravi velike promene u procenama sposobnosti. Shodno s tim, TrueSkill može da proceni sposobnosti pojedinih igrača iz vrlo malog broja igara. U našem slučaju igra će predstavljati jedan poen. Za razliku od TrueSkill sistema, MultiSkill pravi promene posle svakog odigranog poena.

MultiSkill sistem za procenu primenjuje tri glavna pristupa:

- Koristi promenljive koje imaju normalnu raspodelu za model razvoja sposobnosti igrača tokom vremena;
- Predstavlja igru preko Bajesove mreže;
- Koristi očekivano proširenje zaključaka za promenu sposobnosti datih posmatranom igrom.

Poen se odigrava između igrača 1 i 2 pod uslovom da igrač 1 servira. Napomenimo da se čvorovi sa indeksom 1 odnose na igrača 1, dok čvorovi sa indeksom 2 na igrača 2. Ovaj primer predstavlja Bajesovu mrežu koja računa rejting igrača pomoću MultiSkill i TrueSkill sistema.

5.1.5 Primer



Slika 27: Bajesova mreža koja predstavlja poen u tenisu preko MultiSkill i TrueSkill sistema.

U primeru čvorovi predstavljaju sledeće:

- S predstavlja *sposobnost* igrača;
- P predstavlja *performans* igrača;
- d predstavlja razliku performansa.

Sposobnost igrača predstavlja talenat i umeće igrača, a performans je predstavljanje sposobnosti u datom trenutku. Zbog toga imamo takve zavisnosti u mreži, tj. sposobnosti utiču na performanse. Parametar d predstavlja razliku performansa $P1$ i $P2$. Ako je d veće od nule, poen je osvojio igrač 1, u suprotnom poen je osvojio igrač 2. TrueSkill sistem koristi parametre za sposobnost S, performans P i razliku performansa d. Napomenimo još da ti parametri predstavljaju brojeve koji su dobijeni normalnom raspodelom. Detaljna analiza ovih sistema prevazilazi okvire ovog rada, tako da nećemo objašnjavati parametre. Primetimo još da ishod poena zavisi od razlike performansa.

Glava 6

Zaključak

Bajesove mreže su snažan matematički aparat za analiziranje uzročnih odnosa u složenim modelima poput dijagnostikovanja bolesti ili opisivanja modela za donošenje odluka. Koriste raznorazne matematičke alate: grafove, verovatnoće i Markovljeve modele. U ovom radu smo predstavili osnovne pojmove Bajesovih mreža, kao grafičke strukture i u smislu verovatnoća. Pokazali smo inspirisanost veštačkom inteligencijom i mašinskim učenjem. Zatim, pokazali smo njihovu mogućnost rezonovanja i naveli koje vrste Bajesovih mreža postoje. Od vrsta smo definisali dinamičke Bajesove mreže. Zbog kompleksnosti računanja verovatnoća koristili smo softver Netiku. Pokazali smo široku primenu Bajesovih mreža, a najviše smo se posvetili primeni u tenisu. Predstavili smo originalan model za procenu verovatnoće osvojenog poena. Pomoću Netike smo modelirali primere i računali željene verovatnoće.

Istorijski podaci

U ovom poglavlju reći ćemo nešto o matematičarima koje smo spominjali u radu. Literatura koju smo koristili je [45].

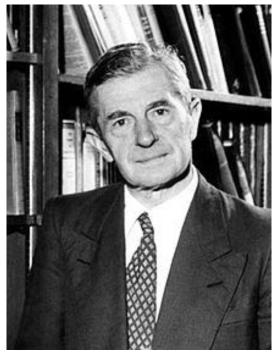
Tomas Bajes (1702-1761) bio je engleski matematičar rođen u Londonu. Po profesiji je bio sveštenik. Kao matematičar bio je član kraljevskog društva u Londonu. Po njemu je dobilo ime bajesovo pravilo koje je od velikog značaja u verovatnoći. Osim toga, ustanovio je i poseban pristup u teoriji statističkog zaključivanja.

Slika 28: Tomas Bajes.

Andrej Markov (1856-1922) bio je ruski matematičar rođen u Rjazanju. Studirao je na univerzitetu u Sankt Peterburgu i bio je član Ruske akademije nauka. Najviše se bavio stohastičkim procesima, a njegova istraživanja su poznata kao Markovljevi procesi.

Slika 29: Andrej Markov.

Sidni Čepmen (1888-1970) bio je britanski matematičar i geofizičar rođen u Mančesteru. Bio je profesor matematike u Londonu i Oksfordu, a predavao je i na mnogim univerzitetima u svetu. Istraživao je kinetičku teoriju plinova, geomagnetizam i atmosfersku fiziku. Po njemu je nazvan krater na mesecu.

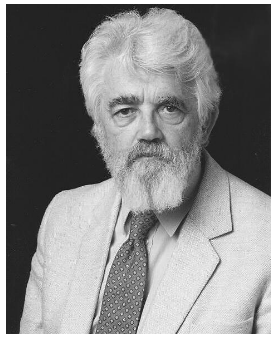


Slika 30: Sidni Čepmen.

Andrej Kolmogorov (1903-1987) bio je ruski matematičar rođen u Tambovu. Studirao je na Moskovskom državnom univerzitetu. Bavio se teorijom verovatnoće, topologijom, mehanikom, matematičkom analizom... Dobio je brojne nagrade od kojih ćemo izdvojiti poslednju, nagradu Lobačevskog.

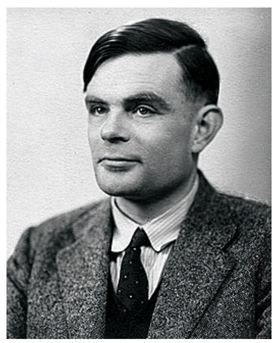
Slika 31: Andrej Kolmogorov.

Džon Makarti (1927-2011) bio je američki naučnik rođen u Bostonu. Zbog velikog doprinosa na polju veštačke inteligencije dobio je Tjuringovu nagradu. Zaslužan je za stvaranje samog pojma "veštačka inteligencija". Projektovao je programski jezik Lisp.



Slika 32: Džon Makarti.

Alan Tjuring (1912-1954) bio je engleski matematičar, logičar i kriptograf rođen u Londonu. Radio je u Nacionalnoj fizičkoj laboratoriji, a tokom Drugog svetskog rata radio je u Blečli parku. Razvio je tehnike za razbijanje šifara, uključujući metod bombe, elektromehaničke mašine. Smatra se ocem modernog računarstva. Poznat je po svom Tjuringovom testu koji je dao značajan doprinos veštačkoj inteligenciji.



Slika 33: Alan Tjuring.

Judea Perl je izraelski kompjuterski naučnik i filozof rođen u Tel Avivu 1936. Zaslužan je za razvoj Bajesovih mreža i prvi je uveo taj termin. Takođe, zaslužan je za razvoj teorije uzročnosti i dao je doprinos u oblasti veštačke inteligencije kroz razvoj računara za verovatnoće. Dobitnik je Tjuringove nagrade.

Slika 34: Judea Perl.

Bibliografija

[1] Austin, S., Placeway, P., Schwartz, R. The forward-backward search algorithm, Acoustics, Speech, and Signal Processing, 1991. ICASSP-91., 1991 International Conference on, 1991.

[2] Baldi, P., Chauvin, Y. Smooth on-line learning algorithms for hidden Markov models, Neural Computation 6.2 (1994): 307-318, 1994.

[3] Bellman, R. An introduction to artificial intelligence: Can computers think?, Thomson Course Technology, 1978.

[4] Bešlić, I., Bešlić, D., Zakić, V. Data mining tehnike za otkrivanje lažnog finansijskog izveštavanja, Univerzitet Singidunum, 2014.

[5] Bilmes, J. Learning Bayesian Multinets, Proceedings of the Sixteenth conference on Uncertainty in artificial intelligence, Morgan Kaufmann Publishers Inc., 2000.

[6] Blunsom, P. Hidden Markov Models. Lecture notes, 2004.

[7] Bondy, J., Murty, U. *Graph Theory*, Springer, 2008.

[8] Boudali, H., Dugan, J. A discrete-time Bayesian network reliability modeling and analysis framework, Reliability Engineering and System Safety, 2005.

[9] Carbonell, J., Michalski, R., Mitchell, T. Machine learning: An artificial intelligence approach, Springer science, 2013.

[10] Charniak, E., McDermott, D. Introduction to Artificial intelligence, Reading, Addison-Wesley, 1985.

[11] Faltin, F., Kenett, R., Ruggeri, F. Encyclopedia of statistics in quality and reliability, Wiley Chichester, England, 2007.

[12] Forney, D. The viterbi algorithm, Proceedings of the IEEE 61.3 (1973): 268-278, 1973.

[13] Ghahramani, Z. International Journal of Pattern Recognition and Artificial Intelligence. University College London, 2001.

[14] Ghahramani, Z. An introduction to hidden Markov models and Bayesian networks. International Journal of Pattern Recognition and Artificial Intelligence, World Scientific, 2001.

[15] Glušica, M. Skriveni Markovljevi modeli i primena u bioinformatici, Magistarski rad, Prirodno-matematički fakultet, Beograd, 2013.

[16] Goebel, R., Mackworth, A., Poole, D. Computational intelligence: a logical approach, Oxford University Press New York, 1998.

[17] Graepel, T., Herbrich, R., Minka, T. $TrueSkill^{TM}$: A Bayesian skill rating system, Advances in neural information processing systems, 2006.

[18] Hadžić, O. *Numeričke i statističke metode u obradi eksperimentalnih podataka I deo*, drugo izdanje, Institut za matematiku, Novi Sad, 1992.

[19] Haugeland, J. Artificial intelligence: The Very Idea, Cambridge, 1985.

[20] Janičić, P., Nikolić, M. Veštačka inteligencija, Matematički fakultet univerziteta u Beogradu, 2016.

[21] Karuza, M. Izvedba funkcija za skrivene Markovljeve modele. Sveučilište u Zagrebu, 2011.

[22] Knight, K., Rich, E. Artificial intelligence, McGraw-Hill, 1991.

[23] Korb, B., Nicholson, E. *Bayesian Artificial Intelligence*, 2th Edition, CRC Press, Boca Raton FL USA, 2011.

[24] Korzekwa, D. The MultiSkill Tennis Model, Machine Learning Summer School, Germany, 2013.

[25] Kurzweil, R., Richter, R., Schneider, M. The age of intelligent machines, MIT press Cambridge, 1990.

[26] Lozanov Crvenković, Z. *Statistika*, Prirodno-matematički fakultet, Novi Sad, 2012.

[27] Margaritis, D. Learning Bayesian Network Model Structure from Data. Ph.D. Thesis, Department of Computer Science, Carnegie Mellon University, 2003.

[28] Mašinsko učenje, inteligentni sistemi, seminarski rad. Visoka poslovna škola strukovnih studija, Čačak.

[29] Milosavljević, M. Veštačka inteligencija, prvo izdanje. Univerzitet Singidunum, Beograd, 2015.

[30] Mladenović, P. *Verovatnoća i statistika*, VESTA - Matematički fakultet, Beograd, 1995.

[31] Murphy, K. Dynamic Bayesian Networks: representation, inference and learning, University of California, Berkeley, 2002.

[32] Neapolitan, R. Learning Bayesian Networks, Pearson Prentice Hall Upper Saddle River, 2004.

[33] Nilsson, N. Artificial intelligence: a new synthesis, Elsevier, 1998.

[34] Rajter Ćirić, D. *Stohastička analiza - skripta* Prirodno-matematički fakultet, Novi Sad, 2005.

[35] Rajter Ćirić, D. *Verovatnoća*, treće izdanje, Prirodno-matematički fakultet, Novi Sad, 2013.

[36] Rakić, S. Veštačka inteligencija: ekspertni sistemi i neuronske mreže, e-knjiga Fakultet za menadžment, Novi Sad, 2003.

[37] Reynolds, D. Gaussian mixture models, Encyclopedia of biometrics. Springer, 2015.

[38] Rudman, M. Kompleksnost skrivenih Markovljevih modela, diplomski rad. Sveučilište u Zagrebu, 2015.

[39] Smola, A., Vishwanathan, S. Introduction to Machine learning, Cambridge University, 2008.

[40] Stojaković, M. Verovatnoća i slučajni procesi, Fakultet tehničkih nauka, Novi Sad, 2013.

[41] Winston, P. Learning by building identification trees, Boston: Addison-Wesley Publishing Company, 1992.

[42] http://artint.info/html/ArtInt$_1$48.*html*, *avgust*2016.

[43] https://es.scribd.com/doc/307680679/500-Osnovi-Teorije-Verovatnoce, 2016.

[44] https://www.norsys.com/netica.html, avgust 2016.

[45] https://www.wikipedia.org/, avgust 2016.

Biografija

Ivana Dojić rođena je u Somboru 5.12.1992. godine. Završila je osnovnu školu ,,Ivo Lola Ribar" 2007. godine kao nosilac Vukove diplome. Zatim je upisala Gimnaziju ,,Veljko Petrović" društveni smer. Istu je završila 2011. godine nakon čega je upisala osnovne studije matematike na Departmanu za matematiku i informatiku Prirodno-matematičkog fakulteta Univerziteta u Novom Sadu. Ivana je nacionalni teniski sudija od 2011. godine i time se aktivno bavi sve do danas. Studije je završila zaključno sa junskim rokom 2015. godine sa prosečnom ocenom 8.75. Master studije upisala je iste godine na Fakultetu tehničkih nauka u Novom Sadu, smer Matematika u tehnici. Sve ispite položila je u junskom roku 2016. godine sa prosečnom ocenom 9.38 i time stekla uslov za odbranu ovog master rada.

Novi Sad, septembar 2016. Ivana Dojić