

LTAT.02.006 Andmeteaduse meetodid

8. praktikum

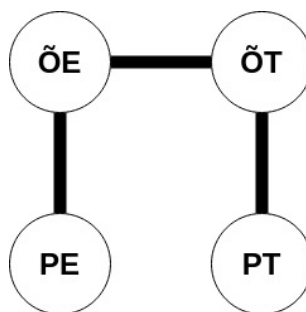
2. - 3. november 2021

1. Koostada eelmise praktikumi viimases ülesandes vaadeldud Bayesi võrguga ekvivalentne Markovi juhuslik väli. Allpool on toodud uuesti ära need tabelid, mis selle vaadeldud Bayesi võrgu ära määravad. Markovi juhuslikku välja koostades võtta tõenäosused $P(\tilde{\text{ohurõhkEile}})$ ja $P(\tilde{\text{ohurõhkTäna}} | \tilde{\text{ohurõhkEile}})$ kokku üheks faktoriks. Arvutada välja ka normaliseeriv konstant Z .

$P(\tilde{\text{ohurõhkEile}} = \text{madal})$	$\tilde{\text{ohurõhkEile}}$	$P(\tilde{\text{ohurõhkTäna}} = \text{madal} \tilde{\text{ohurõhkEile}})$
0.4	madal	0.3
	kõrge	0.2

$\tilde{\text{ohurõhkEile}}$	$P(\text{PilvesEile} = \text{jah} \tilde{\text{ohurõhkEile}})$
madal	0.6
kõrge	0.4

$\tilde{\text{ohurõhkTäna}}$	$P(\text{PilvesTäna} = \text{jah} \tilde{\text{ohurõhkTäna}})$
madal	0.6
kõrge	0.4



$$\underbrace{P(\tilde{\text{OE}})P(\tilde{\text{OT}}|\tilde{\text{OE}})}_{\phi_1} \quad \underbrace{P(\text{PE}|\tilde{\text{OE}})}_{\phi_2} \quad \underbrace{P(\text{PT}|\tilde{\text{OT}})}_{\phi_3}$$

$$\begin{aligned}
 \phi_1(\tilde{\text{OE}}=\text{m}, \tilde{\text{OT}}=\text{m}) &= P(\tilde{\text{OE}}=\text{m}) \cdot P(\tilde{\text{OT}}=\text{m}|\tilde{\text{OE}}=\text{m}) = 0.4 \cdot 0.3 = 0.12 \\
 \phi_1(\tilde{\text{OE}}=\text{m}, \tilde{\text{OT}}=\text{k}) &= P(\tilde{\text{OE}}=\text{m}) \cdot \underbrace{P(\tilde{\text{OT}}=\text{k}|\tilde{\text{OE}}=\text{m})}_{1 - P(\tilde{\text{OT}}=\text{m}|\tilde{\text{OE}}=\text{m})} = 0.4 \cdot (1 - 0.3) = 0.28 \\
 \phi_1(\tilde{\text{OE}}=\text{k}, \tilde{\text{OT}}=\text{m}) &= P(\tilde{\text{OE}}=\text{k}) \cdot P(\tilde{\text{OT}}=\text{m}|\tilde{\text{OE}}=\text{k}) = 0.6 \cdot 0.2 = 0.12 \\
 \phi_1(\tilde{\text{OE}}=\text{k}, \tilde{\text{OT}}=\text{k}) &= P(\tilde{\text{OE}}=\text{k}) \cdot \underbrace{P(\tilde{\text{OT}}=\text{k}|\tilde{\text{OE}}=\text{k})}_{1 - P(\tilde{\text{OT}}=\text{m}|\tilde{\text{OE}}=\text{k})} = 0.6 \cdot (1 - 0.2) = 0.48
 \end{aligned}$$

ϕ_1		
$\tilde{O}E$	$\tilde{O}T$	$P(\tilde{O}E, \tilde{O}T)$
m	m	0.12
m	k	0.28
k	m	0.12
k	k	0.48

ϕ_2		
$\tilde{O}E$	PE	$P(PE \tilde{O}E)$
m	e	0.4
m	j	0.6
k	e	0.6
k	j	0.4

ϕ_3		
$\tilde{O}T$	PT	$P(PT \tilde{O}T)$
m	e	0.4
m	j	0.6
k	e	0.6
k	j	0.4

$\tilde{O}E$	$\tilde{O}T$	PE	PT	$\phi_1 \cdot \phi_2 \cdot \phi_3$
m	m	e	e	$0.12 \cdot 0.4 \cdot 0.4$
				+
m	m	e	j	$0.12 \cdot 0.4 \cdot 0.6$
				+
			...	
				+
k	k	j	j	$0.48 \cdot 0.4 \cdot 0.4$
$Z=1$				

2. Teisendada eelmises ülesandes leitud Markovi juhusliku välja faktoreid sellisel, et need sisaldaksid vaid täisarve, aga Markovi juhuslik väli jääks algsega samaväärseks. Saadud Markovi juhusliku välja abil vastata järgmistele küsimustele ilma normaliseerivat konstanti Z kasutamata:

(a) Leida kumb järgnevatest tõenäosustest on suurem:

$P(\tilde{O}hur\tilde{o}hkEile = \text{madal}, \tilde{O}hur\tilde{o}hkT\tilde{a}na = \text{k\ddot{o}rge}, \text{PilvesEile} = \text{jah}, \text{PilvesT\ddot{a}na} = \text{ei})$

$P(\tilde{O}hur\tilde{o}hkEile = \text{k\ddot{o}rge}, \tilde{O}hur\tilde{o}hkT\tilde{a}na = \text{k\ddot{o}rge}, \text{PilvesEile} = \text{jah}, \text{PilvesT\ddot{a}na} = \text{ei})$

(b) Milline on kõige tõenäolisem tunnuste väärtuste kombinatsioon?

(c) Milline on kõige vähem tõenäolisem tunnuste väärtuste kombinatsioon ja mitmekordne on tõenäosuse erinevus kõige tõenäolisema kombinatsiooniga?

$\tilde{O}E$	$\tilde{O}T$	ϕ_1
m	m	12
m	k	28
k	m	12
k	k	48

$\tilde{O}E$	PE	ϕ_2
m	e	40
m	j	60
k	e	60
k	j	40

$\tilde{O}T$	PT	ϕ_3
m	e	40
m	j	60
k	e	60
k	j	40

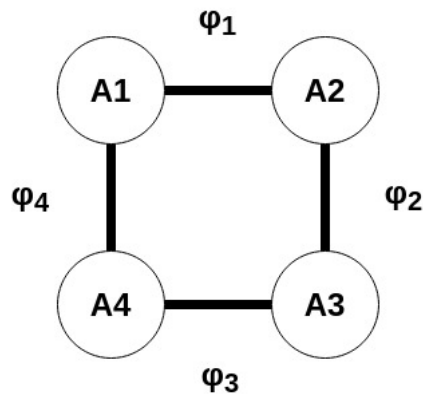
$\tilde{O}E$	$\tilde{O}T$	PE	PT	$\phi_1 \cdot \phi_2 \cdot \phi_3$
m	k	j	e	$28 \cdot 60 \cdot 60 = 100800$
k	k	j	e	$48 \cdot 40 \cdot 60 = 115200$

$\tilde{O}E$	$\tilde{O}T$	PE	PT	$\phi_1 \cdot \phi_2 \cdot \phi_3$
k	k	e	e	$48 \cdot 60 \cdot 60 = 172800$
m	m	e	e	$12 \cdot 40 \cdot 40 = 19200$
k	m	j	e	$12 \cdot 40 \cdot 40 = 19200$
				$172000/19200 = 9x$

Leida normaliseeriva konstandi Z väärtus ning leida kõige tõenäolisema ja kõige vähem tõenäolisema tunnuste väärtuste kombinatsioonide tõenäosused.

ÕE	ÕT	PE	PT	$\phi_1 \cdot \phi_2 \cdot \phi_3$
m	m	e	e	$12 \cdot 40 \cdot 40$
				+
m	m	e	j	$12 \cdot 40 \cdot 60$
				+
		...		
				+
k	k	j	j	$48 \cdot 40 \cdot 40$
				$Z=1e6$
				$172800/1e6 = 0.1728$
				$19200/1e6 = 0.0192$

3. Vaatleme 2x2 tipuga ruutvõret, kus tipud on koordinaatidel (0,0), (0,1), (1,0), (1,1) ning servad on tippude vahel, mille kaugus on 1. Olgu igas tipus binaarne tunnus väärtustega +1 ja -1 ning kõik neli faktorit olgu sellised, kus tunnuste võrdsete väärtuste korral on faktori väärtus 2 ja vastasel juhul 1. Leida tõenäosus, et kõik tunnused on võrdsed väärtusega +1.



A_1	A_2	ϕ_1
+	+	2
+	-	1
-	+	1
-	-	2

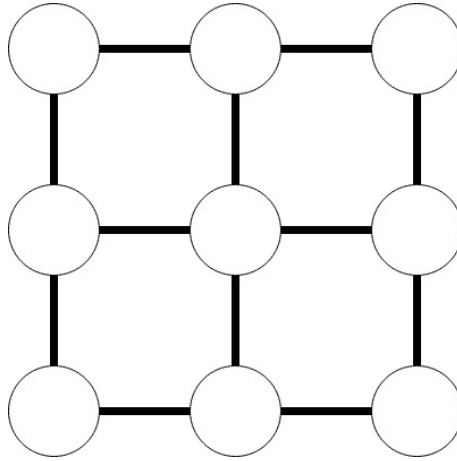
A_2	A_3	ϕ_2
+	+	2
+	-	1
-	+	1
-	-	2

A_3	A_4	ϕ_3
+	+	2
+	-	1
-	+	1
-	-	2

A_4	A_1	ϕ_4
+	+	2
+	-	1
-	+	1
-	-	2

A1	A2	A3	A4	$\phi_1 \cdot \phi_2 \cdot \phi_3 \cdot \phi_4$
+	+	+	+	16
+	+	+	-	4
+	+	-	+	4
+	+	-	-	4
+	-	+	+	4
+	-	+	-	1
+	-	-	+	4
+	-	-	-	4
-	+	+	+	4
-	+	+	-	4
-	+	-	+	1
-	+	-	-	4
-	-	+	+	4
-	-	+	-	4
-	-	-	+	4
-	-	-	-	16
				$Z=16 \cdot 2 + 1 \cdot 2 + 4 \cdot 12 = 82$
				$16/82 = 0.195$

4. Vaatleme samasugust ruutvõret nagu eelmises ülesandes, aga nüüd 3x3 tipuga koordinaatidel $(0, 0), (0, 1), (0, 2), (1, 0), (1, 1), (1, 2), (2, 0), (2, 1), (2, 2)$. Oletame, et vaatlusena teame, et punktis $(0, 1)$ on väärtus $+1$ ja punktis $(0, 0)$ on väärtus -1 . Mis on kõige tõenäolisem kombinatsioon kõikide tunnuste väärtustest?



$$P \begin{pmatrix} - & + & ? \\ ? & ? & ? \\ ? & ? & ? \end{pmatrix}$$

$$P \begin{pmatrix} + & + & + \\ + & + & + \\ + & + & + \end{pmatrix} = 2^{12}/z \quad P \begin{pmatrix} + & - & + \\ - & + & - \\ + & - & + \end{pmatrix} = 1/z$$

$$P \begin{pmatrix} - & + & + \\ + & + & + \\ + & + & + \end{pmatrix} = 1^2 \cdot 2^{10}/z$$