ДОМАШНА ЗАДАЧА 2

ГРУПА 3

ВЕШТАЧКА ИНТЕЛИГЕНЦИЈА

БАЕСОВИ МРЕЖИ И МАШИНСКО УЧЕЊЕ

Изработила: Дабеска Мила 193263

1. БАЕСОВИ МРЕЖИ

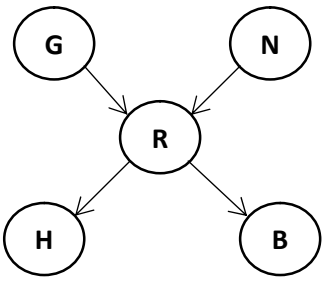
Климатски промени

(а) Да се претстават случајните променливи со кои го моделирате проблемот користејќи Баесова мрежа. Образложете го вашиот избор!

- **Променливи во овој проблем ќе бидат:**

1. **Регулатива (R)** – логичка променлива која се справува со проблемите врзани со климатските промени
2. **Државните институции (G)** - логичка променлива која се справува со решавање на проблемот
3. **Буџет (B)** - логичка променлива која се занимава со инвестиции и трошоци во државата
4. **Здравје (H)** - логичка променлива која се занимава со здравјето на населението
5. **Невладиниот сектор (N)** - логичка променлива која се занимава со намалување на последиците од климатските промени и групите кои се посветени за подобрување на здравјето на луѓето во државата

(б) Да се даде графички приказ на Баесовата мрежа!



(в) Колку вкупно параметри (поединечни вредности на веројатностите) се потребни за да се дефинираат условните веројатносни распределби (во табелите покрај секој јазол) за оваа Баесова мрежа?

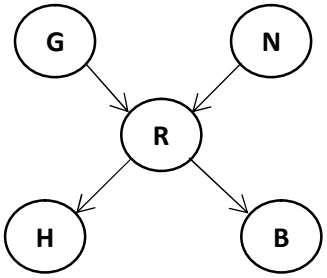
- Вкупниот број на поединечни вредности на веројатностите кои се потребни за случајните променливи во овој проблем се 20.

1. По 2 параметри за маргиналните веројатности на G и N, вкупно 4 параметри
2. По 4 параметри за условните веројатности на R за дадени H и B, вкупно 8 параметри.
3. По 8 параметри за условната веројатност на G и N за дадено R.

(г) Да се прикажат маргиналните и условни веројатносни табели користејќи ги податоците од описот на проблемот. За вредностите за кои немате податоци изберете вредности според ваш избор.

G	P(G)
0	0.84
1	0.16

N	P(N)
0	0.72
1	0.28



R	H	P(H|R)
0	0	0.35
0	1	0.2
1	0	0.3
1	1	0.15

R	B	P(B|R)
0	0	0.2
0	1	0.36
1	0	0.05
1	1	0.39

G	N	R	P(R|G,N)
0	0	0	0.09
0	0	1	0.07
0	1	0	0.03
0	1	1	0.05
1	0	0	0.3
1	0	1	0.15
1	1	0	0.25
1	1	1	0.06

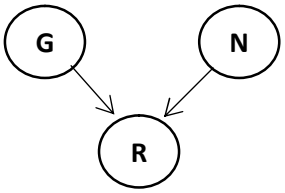
(д) Напишете го изразот за тотална здружена веројатност за случајните променливи кои се дефинирани во Баесовата мрежа. Користете го правилото на синџир според подредување.

$$P(G,N,R,H,B) = P(G) * P(N) * P(R \mid G,N) * P(H \mid R) * P(B \mid R)$$

(ѓ) Наведете ги сите независности (апсолутни и условни) помеѓу променливите кои произлегуваат од структурата на Баесовата мрежа. Образложете!

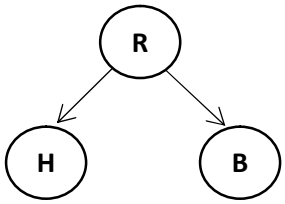
-**Елементарни тројки:**

1. **Заедничка последница (Common effect)**



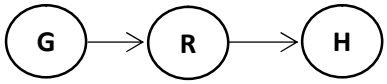
Доколку вредноста на **R** ни е непозната само тогаш, **G и N** се апсолутно независни.

2. **Заедничка причина (Common cause)**



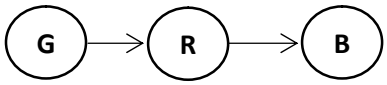
Доколку вредноста на **R** ни е позната само тогаш, **H и B** се условно независни за дадено **R**

3. **Каузален синџир**



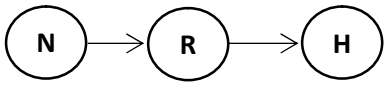
Доколку вредноста на **R** ни е позната само тогаш, **G и H** се условно независни за дадено **R**. Односно доколку **R** ни е непознато тогаш протекува влијание помеѓу **G и H.**

Каузален синџир



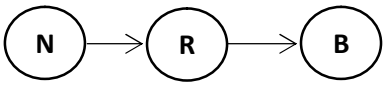
Доколку вредноста на **R** ни е позната само тогаш, **G и B** се условно независни за дадено **R**. Односно доколку **R** ни е непознато тогаш протекува влијание помеѓу **G и B**.

Каузален синџир



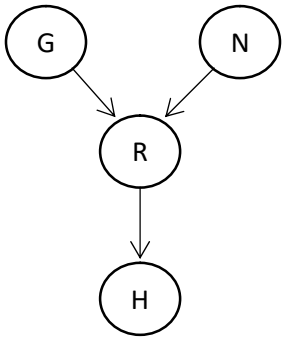
Доколку вредноста на **R** ни е позната само тогаш, **N и H** се условно независни за дадено **R**. Односно доколку **R** ни е непознато тогаш протекува влијание помеѓу **N и H.**

Каузален синџир



Доколку вредноста на **R** ни е позната само тогаш, **N и B** се условно независни за дадено **R**. Односно доколку **R** ни е непознато тогаш протекува влијание помеѓу **N и B.**

(е) Напишете го изразот за веројатност дека постои позитивен тренд во подобрување на здравјето на населението (пр. намален број на болни од рак или намален број на белодробни заболувања). Да се пресмета нумеричката вредност на оваа веројатност, користејќи ги вредностите од веројатносните табели.



-Здружена веројатност

$$P(G,N,R,H) = P(G) * P(N) * P(R \mid G,N) * P(H \mid R)$$

- Прошируваме со веројатноста на R

$$P(H = 1) * P(R) = P(H = 1, R = 0) * P(G) + P(H = 1, R = 1) * P(G)$$

- Прошируваме со веројатноста на G

$$P(H = 1, R = 0, G = 0) * P(N) + P(H = 1, R = 0, G = 1) * P(N) + P(H = 1, R = 1, G = 0) * P(N) + P(H = 1, R = 1, G = 1) * P(N)$$

- **Пресметување веројатност:**

- $P(H = 1, R = 0, G = 0, N = 0) = P(G = 0) * P(N = 0) * P(R = 0 \mid G = 0, N = 0) * P(H = 1 \mid R = 0) = 0.84 * 0.72 * 0.09 * 0.2 = 0.01$

- $P(H = 1, R = 0, G = 0, N = 1) = P(G = 0) * P(N = 1) * P(R = 0 \mid G = 0, N = 1) * P(H = 1 \mid R = 0) = 0.84 * 0.28 * 0.03 * 0.2 = 0.001$

- $P(H = 1, R = 0, G = 1, N = 0) = P(G = 1) * P(N = 0) * P(R = 0 \mid G = 1, N = 0) * P(H = 1 \mid R = 0) = 0.16 * 0.72 * 0.3 * 0.2 = 0.007$

- $P(H=1,R=0,G=1,N=1) = P(G=1) * P(N=1) * P(R=0 \mid G=1,N=1) * P(H=1 \mid R=0) = 0.16 * 0.28 * 0.25 * 0.2 = 0.002$

- $P(H=1,R=1,G=0,N=0) = P(G=0) * P(N=0) * P(R=1 \mid G=0,N=0) * P(H=1 \mid R=1) = 0.84 * 0.72 * 0.07 * 0.15 = 0.006$

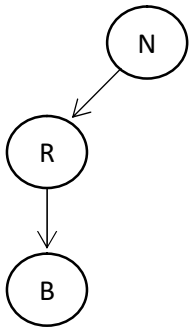
- $P(H=1,R=1,G=0,N=1) = P(G=0) * P(N=1) * P(R=1 \mid G=0,N=0) * P(H=1 \mid R=1) = 0.84 * 0.28 * 0.07 * 0.15 = 0.002$

- $P(H=1,R=1,G=1,N=0) = P(G=1) * P(N=0) * P(R=1 \mid G=1,N=0) * P(H=1 \mid R=1) = 0.16 * 0.72 * 0.15 * 0.15 = 0.003$

- $P(H=1,R=1,G=1,N=1) = P(G=1) * P(N=1) * P(R=1 \mid G=1,N=1) * P(H=1 \mid R=1) = 0.16 * 0.28 * 0.06 * 0.15 = 0.0004$

- $P(H=1) = 0.01 + 0.001 + 0.007 + 0.002 + 0.006 + 0.002 + 0.003 + 0.0004 = \mathbf{0.0314}$

(ж) Напишете го изразот за веројатност дека постоела зголемена активност на невладиниот сектор, доколку буџетот не претрпел измени. Да се пресмета нумеричката вредност на веројатноста, користејќи ги вредностите од веројатносните табели.



- **Баесова теорема**

$$P(N=1|B=0) = \frac{P(B=0|N=1) * P(N=1)}{P(B=0)}$$

$$= \frac{(P(B=0|N=1,R=0) + P(B=0|N=1,\ R=1)) * P(N=1)}{P(B=0,R=0) + P(B=0,\ R=1)} =$$

$$\frac{(P(B=0|R=0) + P(B=0|R=1)) * P(N=1)}{P(B=0|R=0) + P(B=0|R=1)}$$

$$P(N=1|B=0) = \frac{P(B=0|N=1) * P(N=1)}{P(B=0)} = \frac{0.2 + 0.05 + 0.28}{0.2 + 0.05} = \frac{0.53}{0.25} = 2.12$$

2. МАШИНСКО УЧЕЊЕ

#ден	Температура	Ветер	Притисок	Загадено?
1	висока	силна	висок	**НЕ**
2	висока	силна	нормален	**НЕ**
3	ниска	силна	висок	**ДА**
4	ниска	слаба	нормален	**НЕ**
5	нормална	слаба	нормален	**ДА**
6	висока	силна	нормален	**НЕ**
7	ниска	слаба	нормален	**ДА**
8	нормална	силна	висок	**ДА**
9	нормална	слаба	висок	**НЕ**
10	нормална	слаба	нормален	**ДА**

А. Наивен Баесов класификатор

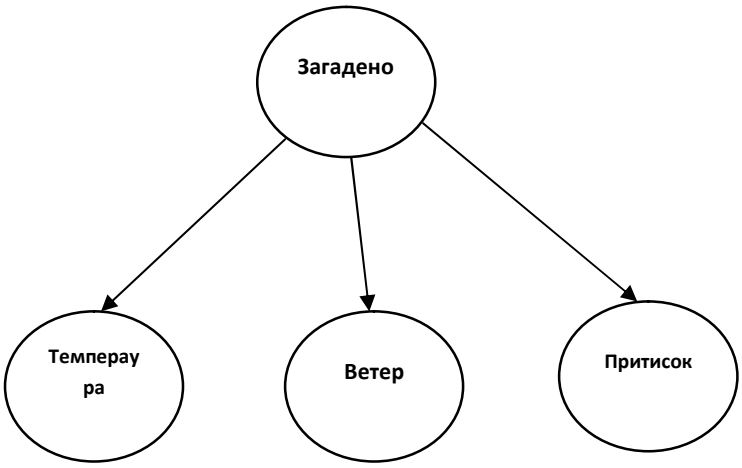
(а) Нацртајте го моделот на Наивниот Баесов класификатор и прикажете ги сите негови параметри.

- Карактеристики:

- Температура
- Ветер
- Притисок

- **Излезна лабела(родител(причина) на сите влезни карактеристики(последици)):**

- Загадено



-Тотална здружена веројатност:

$$P(\text{Загадено}, \text{Температура}, \text{Ветер}, \text{Притисок}) = P(\text{Загадено}) * P(\text{Температура}|\text{Загадено}) * P(\text{Ветер}|\text{Загадено}) * P(\text{Притисок}|\text{Загадено})$$

-Параметри:

Загадено		P(Да)/P(Не)
ДА	5	5/10
НЕ	5	5/10
Вкупно	10	100%

Притисок	ДА	НЕ	P(Да)	P(Не)
Висок	2	2	2/5	2/5
Нормален	3	3	3/5	3/5
Вкупно	5	5	100%	100%

Ветер	ДА	НЕ	P(Да)	P(Не)
Силна	2	3	2/5	3/5
Слаба	3	2	3/5	2/5
Вкупно	5	5	100%	100%

Температура	ДА	НЕ	P(Да)	P(Не)
Висока	0	3	0/5	3/5
Ниска	2	1	2/5	1/5
Нормална	3	1	3/5	1/5
Вкупно	5	5	100%	100%

- Со користење на **Лапласово порамнување со k=2** виртуелни примероци ќе биде:

$$\text{Plap}, k(Y) = \frac{c(Y) + k}{N + k|Y|}$$

Загадено		P(Да)/P(Не)
ДА	7	7/14
НЕ	7	7/14

Притисок	ДА	НЕ	P(Да)	P(Не)
Висок	4	4	4/7	4/7
Нормален	5	5	5/7	5/7

Ветер	ДА	НЕ	P(Да)	P(Не)
Силна	4	5	4/7	5/7
Слаба	5	4	5/7	4/7

Температура	ДА	НЕ	P(Да)	P(Не)
Висока	2	5	2/7	5/7
Ниска	4	3	4/7	3/7
Нормална	5	3	5/7	3/7

(б) Како ќе биде класифициран ден која има ниска температура на воздух, ветерот е силен и притисокот на воздухот е нормален?

- **Загадено=да:**

$$\begin{aligned}&\mathrm{P}(\text{Загадено} = \text{да}, \text{Температура} = \text{ниска}, \text{Ветер} = \text{силен}, \text{Притисок} = \text{нормален})\\&= \mathrm{P}(\text{Загадено} = \text{да}) * \mathrm{P}(\text{Температура} = \text{ниска}|\text{Загадено} = \text{да})\\&* \mathrm{P}(\text{Ветер} = \text{силен}|\text{Загадено} = \text{да})\\&* \mathrm{P}(\text{Притисок} = \text{нормален}|\text{Загадено} = \text{да}) = \frac{7}{14} * \frac{4}{7} * \frac{4}{7} * \frac{5}{7}\\&= 0{,}1153395 \sim 0{,}11\end{aligned}$$

- **Загадено=не:**

$$\begin{aligned}&\mathrm{P}(\text{Загадено} = \text{не}, \text{Температура} = \text{ниска}, \text{Ветер} = \text{силен}, \text{Притисок} = \text{нормален})\\&= \mathrm{P}(\text{Загадено} = \text{не}) * \mathrm{P}(\text{Температура} = \text{ниска}|\text{Загадено} = \text{не})\\&* \mathrm{P}(\text{Ветер} = \text{силен}|\text{Загадено} = \text{не})\\&* \mathrm{P}(\text{Притисок} = \text{нормален}|\text{Загадено} = \text{не}) = \frac{7}{14} * \frac{3}{7} * \frac{5}{7} * \frac{5}{7}\\&= 0{,}1083815 \sim 0{,}11\end{aligned}$$

#ден	**Температура**	**Ветер**	**Притисок**	**Загадено?**
11	ниска	силен	нормален	да

(в) Колкава е веројатноста воздухот да биде загаден во ден во кој притисок на воздухот е висок и температурата е нормална?

$$\begin{aligned}&\mathrm{P}(\text{Загадено}, \text{Температура}, \text{Ветер}, \text{Притисок})\\&= \mathrm{P}(\text{Загадено}) * \mathrm{P}(\text{Температура}|\text{Загадено}) * \mathrm{P}(\text{Ветер}|\text{Загадено})\\&* \mathrm{P}(\text{Притисок}|\text{Загадено})\end{aligned}$$

$$\mathrm{P}(\text{Загадено} = \text{да}, \text{Температура} = \text{нормална}, \text{Притисок} = \text{висок}) = \frac{5}{10} * \frac{5}{7} * \frac{4}{7} = 0.20$$

Б. Перцептрон

(а) Предложете соодветен модел на перцептрон. Поставете иницијални вредности за параметрите на перцептронот различни од 0 (нула). Каков е ефектот на параметри еднакви на 0 (нула) во однос на тоа како перцептронот предвидува соодветна класа? Каков ќе биде ефектот на вредностите кои вие ги поставивте?

- Вектор на тежини:

- **Да -> позитивна класа , Не -> негативна класа**

$$W = [BIAS, Wтемпература, Wветер, Wпритисок] = [1,1,1,1]$$

Загадено	BIAS
Да	1
Не	-1

- **Одредување нумерички вредности:**

Ветер	Нумеричка
Слаб	1
Силен	2

Температура	Нумеричка
Ниска	1
Нормална	2
Висока	3

Притисок	Нумеричка
Нормален	1
Висок	2

Загадено	Нумеричка
Не	-1
Да	1

- **Заедно:**

# ден	Температура	Ветер	Притисок	Загадено?
1	3	2	2	-1
2	3	2	1	-1
3	1	2	2	1
4	1	1	1	-1
5	2	1	1	1
6	3	2	1	-1
7	1	1	1	1
8	2	2	2	1
9	2	1	2	-1
10	2	1	1	1

- **Кога параметрите се 0:**

$$W = [\text{BIAS}, \text{Wтемпература}, \text{Wветер}, \text{Wпритисок}] = [1,0,0,0]$$

Примерок 1: $f(x) = [\text{BIAS}, \text{fтемпература}, \text{fветер}, \text{fпритисок}] = [-1,3,2,2]$

Точна класа: $-$ (Не) $\rightarrow y * = -1$

Предвидена класа: $W * f(x)$

$$= \text{BIAS} * \text{BIAS} + \text{Wтемпература} * \text{fтемпература} + \text{Wветер} * \text{fветер} + \text{Wпритисок} * \text{fпритисок} = -1 > 0 \rightarrow y = -1$$

- **Кога параметрите се 1:**

$$W = [\text{BIAS}, \text{Wтемпература}, \text{Wветер}, \text{Wпритисок}] = [1,1,1,1]$$

Примерок 1: $f(x) = [\text{BIAS}, \text{fтемпература}, \text{fветер}, \text{fпритисок}] = [-1,3,2,2]$

Точна класа: $-$ (Не) $\rightarrow y * = -1$

Предвидена класа: $W * f(x)$

$$= \text{BIAS} * \text{BIAS} + \text{Wтемпература} * \text{fтемпература} + \text{Wветер} * \text{fветер} + \text{Wпритисок} * \text{fпритисок} = 8 > 0 \rightarrow y = 1$$

- **Предвидената класа не се совпаѓа со точната и треба да се ажурира векторот на тежини:**

$w \leftarrow w + y * * f = w - f = [1,1,1,1] - [-1,3,2,2] = [0,-2,-1,-1]$

Предвидена класа: $W * f(x)$

$$= \text{BIAS} * \text{BIAS} + \text{Wтемпература} * \text{fтемпература} + \text{Wветер} * \text{fветер} + \text{Wпритисок} * \text{fпритисок} = -10 > 0 \rightarrow -(\text{Не}) \rightarrow y * = -1$$

(б) Покажете една епоха од учењето на перцептронот.

$$W = [\text{BIAS}, \text{Wтемпература}, \text{Wветер}, \text{Wпритисок}] = [0,-2,-1,-1]$$

- $W * f(x) = \text{BIAS} * \text{BIAS} + \text{Wтемпература} * \text{fтемпература} + \text{Wветер} * \text{fветер} + \text{Wпритисок} * \text{fпритисок} = 0 * -1 + (-2 * 3) + (-1 * 2) + (-1 * 2) = -10 > 0 \rightarrow -(\text{Не}) \rightarrow y * = -1$
- $W * f(x) = \text{BIAS} * \text{BIAS} + \text{Wтемпература} * \text{fтемпература} + \text{Wветер} * \text{fветер} + \text{Wпритисок} * \text{fпритисок} = 0 * -1 + (-2 * 3) + (-1 * 2) + (-1 * 1) = -9 > 0 \rightarrow -(\text{Не}) \rightarrow y * = -1$
- $W * f(x) = \text{BIAS} * \text{BIAS} + \text{Wтемпература} * \text{fтемпература} + \text{Wветер} * \text{fветер} + \text{Wпритисок} * \text{fпритисок} = 0 * 1 + (-2 * 1) + (-1 * 2) + (-1 * 2) = -6 > 0 \rightarrow -(\text{Не}) \rightarrow y * = -1$
- $W * f(x) = \text{BIAS} * \text{BIAS} + \text{Wтемпература} * \text{fтемпература} + \text{Wветер} * \text{fветер} + \text{Wпритисок} * \text{fпритисок} = 0 * -1 + (-2 * 1) + (-1 * 1) + (-1 * 1) = -4 > 0 \rightarrow -(\text{Не}) \rightarrow y * = -1$
- $W * f(x) = \text{BIAS} * \text{BIAS} + \text{Wтемпература} * \text{fтемпература} + \text{Wветер} * \text{fветер} + \text{Wпритисок} * \text{fпритисок} = 0 * 1 + (-2 * 2) + (-1 * 1) + (-1 * 1) = -6 > 0 \rightarrow -(\text{Не}) \rightarrow y * = -1$

- $W * f(x) = BIAS * BIAS + Wтемпература * fтемпература + Wветер * fветер + Wпритисок * fпритисок = 0 * -1 + (-2 * 3) + (-1 * 2) + (-1 * 1) = -9 > 0 \rightarrow -(Не) \rightarrow y *= -1$
- $W * f(x) = BIAS * BIAS + Wтемпература * fтемпература + Wветер * fветер + Wпритисок * fпритисок = 0 * 1 + (-2 * 1) + (-1 * 1) + (-1 * 1) = -4 > 0 \rightarrow -(Не) \rightarrow y *= -1$
- $W * f(x) = BIAS * BIAS + Wтемпература * fтемпература + Wветер * fветер + Wпритисок * fпритисок = 0 * 1 + (-2 * 2) + (-1 * 2) + (-1 * 2) = -8 > 0 \rightarrow -(Не) \rightarrow y *= -1$
- $W * f(x) = BIAS * BIAS + Wтемпература * fтемпература + Wветер * fветер + Wпритисок * fпритисок = 0 * -1 + (-2 * 2) + (-1 * 1) + (-1 * 2) = -7 > 0 \rightarrow -(Не) \rightarrow y *= -1$
- $W * f(x) = BIAS * BIAS + Wтемпература * fтемпература + Wветер * fветер + Wпритисок * fпритисок = 0 * 1 + (-2 * 2) + (-1 * 1) + (-1 * 1) = -6 > 0 \rightarrow -(Не) \rightarrow y *= -1$

(в) Дали постои гаранција дека алгоритамот за учење на перцептронот ќе конвергира?

Алгоритамот за учење на перцептронот ќе конвергира доколку дадениот перцептрон во една епоха нема да има никакво ажурирање на векторот на тежини.Со тоа во оваа задача покажав дека нема потреба од ажурирање на векторот на тежини.

В. Дрво на одлучување

(а) Покажете ја целосната постапка (вклучително и сите пресметки) за избор на најдобар атрибут за поделба во коренот на дрвото на одлучување

- S=[Да:5 , Не:5]

- Ентропија

$$H = -\frac{Не}{n} * \log2 * \frac{Не}{n} - \frac{Да}{n} * \log2 * \frac{Да}{n}$$

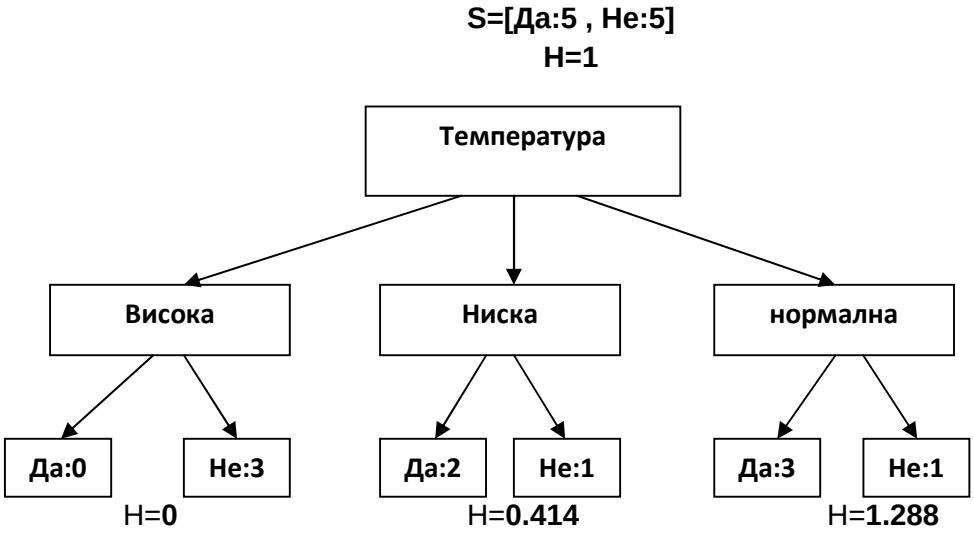
(n – вкупен број на податоци)

$$H = -\frac{5}{10} * \log2 * \frac{5}{10} - \frac{5}{10} * \log2 * \frac{5}{10} = 1$$

- Информациска добивка

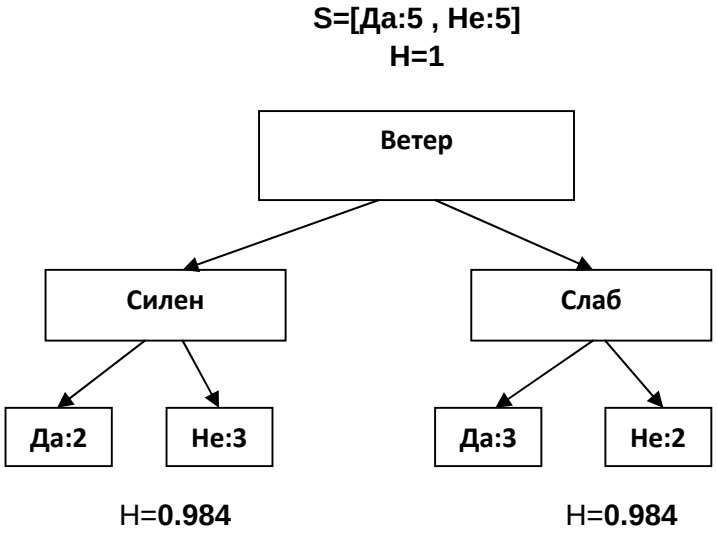
$$IG = H(родител) - [average\ H(дете)]$$

1. *Прво дрво*



$$\text{Gain}(S, \text{Температура}) = 1 - \frac{3}{10} * 0 - \frac{3}{10} * 0.414 - \frac{4}{10} * 1.288 = 0.3606$$

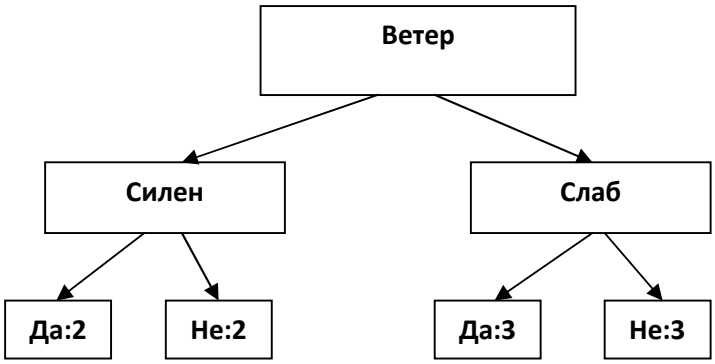
2. *Второ дрво*



$$\text{Gain}(S, \text{Ветер}) = 1 - \frac{5}{10} * 0.984 - \frac{5}{10} * 0.984 = 0{,}016$$

3. Трето дрво

S=[Да:5 , Не:5]
H=1

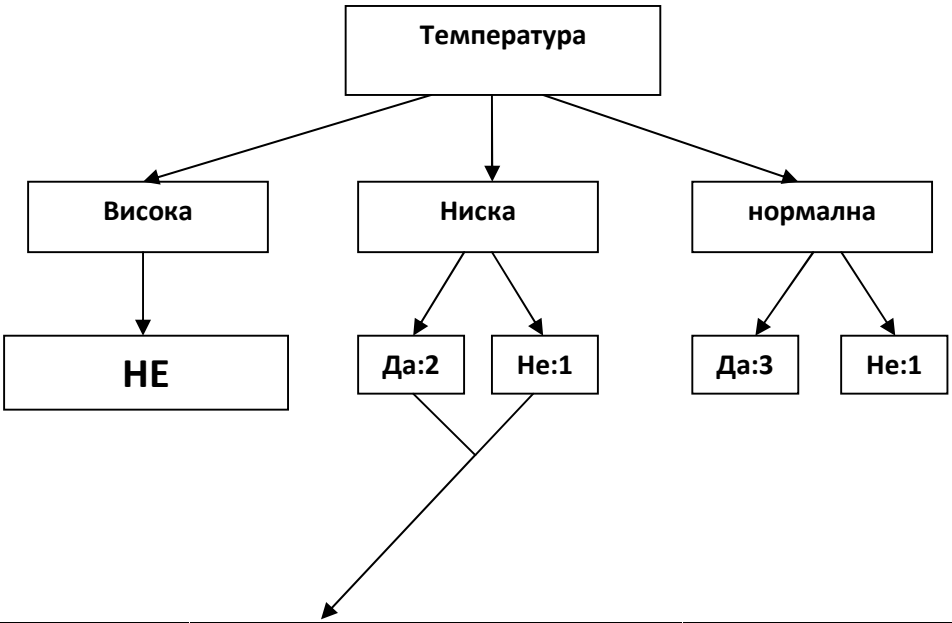


H=1 H=1

$$\text{Gain}(\text{S}, \text{Ветер}) = 1 - \frac{4}{10} * 1 - \frac{6}{10} * 1 = 0$$

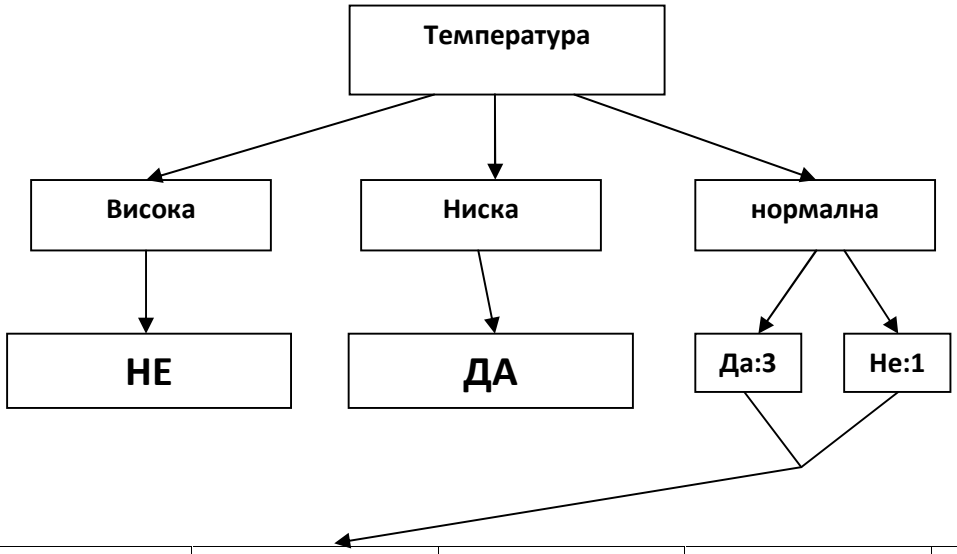
- За корен на дрвото го земаме **Првото дрво** односно атрибутот **Температура**

(б) Покажете го целосното дрво на одлучување без да ги покажувате пресметките. За секој лист во дрвото на одлучување дајте образложение зошто ја предвидува класата која сте ја доделиле.



#ден	Температура	Ветер	Притисок	Загадено?
3	ниска	силна	висок	**ДА**
4	ниска	слаба	нормален	**НЕ**
7	ниска	силна	нормален	**ДА**

- Вредноста **ДА** се појавува 2 пати,а вредноста **НЕ** се појавува еднаш ја доделуваме вредноста **ДА** како лист бидејќи се јавува почесто.



#ден	Температура	Ветер	Притисок	Загадено?
5	нормална	слаба	нормален	**ДА**
8	нормална	силна	висок	**ДА**
9	нормална	слаба	висок	**НЕ**
10	нормална	слаба	нормален	**ДА**

- Вредноста **ДА** се појавува три пати,а вредноста **НЕ** се појавува еднаш ја доделуваме вредноста **ДА** како лист бидејќи се јавува почесто.

✓ **ФИНАЛНО**