

## Példa 1. Döntési fa számítása/1

1. Legyen a felhasználandó példahalmaz:

Példa sz.	Nagy(x)	Fekete(x)	Ugat(x)	JóKutya(x)
X1	Igen	Igen	Igen	Nem
X2	Igen	Igen	Nem	Igen
X3	Nem	Nem	Igen	Nem
X4	Nem	Igen	Igen	Igen
X5	Nem	Igen	Nem	Nem
X6	Igen	Nem	Nem	Igen

ahol *Nagy*, *Fekete* és *Ugat* az elemi attribútumok, *JóKutya* pedig a besorolás, melynek logikai függvényét tanuljuk.

2. A példahalmazban  $n = 3$  negatív és  $p = 3$  pozitív példa van, azaz a példahalmazt képviselő döntési fa információtartalma:  $I = I(1/2, 1/2) = 1$  bit.

3. A *Nagy* attribútumtesztnek két értéke van.

Az 'Igen' ág mentén:  $n_1 = 1$  és  $p_1 = 2$ , a 'Nem' ág mentén:  $n_2 = 2$  és  $p_2 = 1$  példa van.

Abból kifolyólag a *Nagy* attribútum nyeresége:

$$N_y(Nagy) = I - 1/2 I(1/3, 2/3) - 1/2 I(1/3, 2/3) = 1 - I(1/3, 2/3) = A$$

A *Fekete* attribútumtesztnek szintén két értéke van.

Az 'Igen' ág mentén:  $n_1 = 2$  és  $p_1 = 2$ , a 'Nem' ág mentén:  $n_2 = 1$  és  $p_2 = 1$  példa van.

Abból kifolyólag a *Fekete* attribútum nyeresége:

$$N_y(Fekete) = I - 1/3 I(1/2, 1/2) - 2/3 I(1/2, 1/2) = 1 - I(1/2, 1/2) = 1 - 1 = 0$$

Az *Ugat* attribútumtesztnek szintén két értéke van.

Az 'Igen' ág mentén:  $n_1 = 2$  és  $p_1 = 1$ , a 'Nem' ág mentén:  $n_2 = 1$  és  $p_2 = 2$  példa van.

Abból kifolyólag az *Ugat* attribútum nyeresége ugyanolyan, mint a *Nagy* attribútum nyeresége:

$$N_y(Ugat) = N_y(Nagy) = A.$$

A döntési fa gyökerébe tehát a *Nagy* és az *Ugat* attribútumok valamelyikét kell választani.

Legyen a fa gyökere a *Nagy* teszt.

4. Vizsgáljuk most a *Nagy* = Igen ágat. E ág mentén:

Példa sz.	Fekete(x)	Ugat(x)	JóKutya(x)
X1	Igen	Igen	Nem
X2	Igen	Nem	Igen
X6	Nem	Nem	Igen

példa jön le. Az ág végére levélsomópontot tenni még nem lehet, a teszteket folytatni kell.

A redukált példahalmazban  $n = 1$  negatív és  $p = 2$  pozitív példa van, azaz a példahalmazt képviselő döntési fa információtartalma:  $I' = I(1/3, 2/3)$  bit.

A *Fekete* attribútumtesztnek két értéke van.

Az 'Igen' ág mentén:  $n1 = 1$  és  $p1 = 1$ , a 'Nem' ág mentén:  $n2 = 0$  és  $p2 = 1$  példa van. Abból kifolyólag a *Fekete* attribútum nyeresége:

$$Ny(\text{Fekete}) = I' - 2/3 I(1/2, 1/2) - 1/3 I(0, 1) = I' - 2/3$$

Az *Ugat* attribútumtesztnek szintén két értéke van.

Az 'Igen' ág mentén:  $n1 = 1$  és  $p1 = 0$ , a 'Nem' ág mentén:  $n2 = 0$  és  $p2 = 2$  példa van.

Abból kifolyólag az *Ugat* attribútum nyeresége:

$$Ny(\text{Ugat}) = I' - 1/3 I(0, 1) - 2/3 I(0, 1) = I'$$

A tesztsomópontba tehát az *Ugat* attribútumot választjuk.

Az *Ugat* = Igen ág végén csak az X1 (negatív) példa, az *Ugat* = Nem ág végén csak az X2 és X6 (pozitív) példák találhatók, azaz a tesztek végére levélsomópontokat lehet már tenni és ebbe az irányba a fa építését befejezni.

5. Vizsgáljuk most a *Nagy* = Nem ágat. E ág mentén:

Példa sz.	Fekete(x)	Ugat(x)	JóKutya(x)
X3	Nem	Igen	Nem
X4	Igen	Igen	Igen
X5	Igen	Nem	Nem

példa jön le. Az ág végére levélsomópontot tenni még nem lehet, a tesztek folytatni kell.

A redukált példahalmazban  $n = 2$  negatív és  $p = 1$  pozitív példa van, azaz a példahalmazt képviselő döntési fa információtartalma:  $I' = I(1/3, 2/3)$  bit.

A *Fekete* attribútumtesztnek két értéke van.

Az 'Igen' ág mentén:  $n1 = 1$  és  $p1 = 1$ , a 'Nem' ág mentén:  $n2 = 1$  és  $p2 = 0$  példa van.

Abból kifolyólag a *Fekete* attribútum nyeresége:

$$Ny(\text{Fekete}) = I' - 2/3 I(1/2, 1/2) - 1/3 I(0, 1) = I' - 2/3$$

Az *Ugat* attribútumtesztnek szintén két értéke van.

Az 'Igen' ág mentén:  $n1 = 1$  és  $p1 = 1$ , a 'Nem' ág mentén:  $n2 = 1$  és  $p2 = 0$  példa van.

Abból kifolyólag az *Ugat* attribútum nyeresége:

$$Ny(\text{Ugat}) = I' - 2/3 I(1/2, 1/2) - 1/3 I(0, 1) = I' - 2/3$$

A két érték azonos, a csomópontba tehát a *Fekete* és az *Ugat* attribútumok valamelyikét kell választani. Legyen a megválasztott attribútum a *Fekete*.

A *Fekete* = Igen ág végén az X4 (pozitív) és X5 (negatív) példa, a *Fekete* = Nem ág végén csak az X3 (negatív) példa található, azaz a negatív ágat a Nem levélsomóponttal le lehet már zárni, a pozitív ágat további teszttel folytatni kell.

6. Ezen az ágon csak az *Ugat* maradt megválaszolható tesztnek. Az *Ugat* = Igen ág mentén az X4 (pozitív), az *Ugat* = Nem ág mentén az X5 (negatív) példa maradt. Mindkét ágat tehát alkalmas levélsomóponttal le lehet zárni.

Kifogytunk a példákból, kifogytunk az attribútumokból, a döntési fa elkészült.

Ha gráfszerű alakjából kiemeljük a logikai képletét, azt kapjuk, hogy:

$$\forall x. \text{JóKutya}(x) = (\text{Nagy}(x) \wedge \neg \text{Ugat}(x)) \vee (\neg \text{Nagy}(x) \wedge \neg \text{Fekete}(x) \wedge \text{Ugat}(x))$$

## Példa 2. Döntési fa számítása/2

Egy ágens kellemes házi állatka kiválasztásához annak logikai definícióját szeretne megtanulni döntési fa módszerrel az alábbi 6 példa alapján. Információelméleti mennyiségek mérlegelésével tervezze meg ezt a döntési fát!

	Hörcsög	Tengeri Malac	Kutya	Nagy	Kellemes Állat
x1	I	N	N	N	+
x2	I	N	N	I	-
x3	N	I	N	I	-
x4	N	I	N	N	+
x5	N	N	I	I	-
x6	N	N	I	N	+

A példa halmazt képviselő döntési fának  $I_{fa} = I(1/2, 1/2) = 1$  bit az információ tartalma. Vegyük sorra az egyes attribútum tesztek nyereségét:

$$R(\text{Hörcsög}) = 2/6 I(1/2, 1/2) + 4/6 I(1/2, 1/2) = 1$$

$$R(\text{TengeriMalac}) = 2/6 I(1/2, 1/2) + 4/6 I(1/2, 1/2) = 1$$

$$R(\text{Kutya}) = 2/6 I(1/2, 1/2) + 4/6 I(1/2, 1/2) = 1$$

$$R(\text{Nagy}) = 1/2 I(1, 0) + 1/2 I(0, 1) = 0$$

$$N_y(\text{Hörcsög}) = I_{fa} - R(\text{Hörcsög}) = 0$$

$$N_y(\text{TengeriMalac}) = I_{fa} - R(\text{TengeriMalac}) = 0$$

$$N_y(\text{Kutya}) = I_{fa} - R(\text{Kutya}) = 0$$

$$N_y(\text{Nagy}) = I_{fa} - R(\text{Nagy}) = 1$$

Az első teszt tehát a „Nagy”. A teszt után az „Igen” ágban csúppán negatív példa, a „Nem” ágban csúppán pozitív példa marad, ezek tehát már a „Nem” és az „Igen” levelek és a fát ezzel be is fejeztük.

### Példa 3. Döntési fa számítása/3

Egy ágens kellemes házi állatka kiválasztásához annak logikai definícióját szeretne megtanulni döntési fa módszerrel az alábbi 6 példa alapján. Információelméleti mennyiségek mérlegelésével tervezze meg ezt a döntési fát!

	Hörcsög	Tengeri Malac	Kutya	Nagy	Kellemes Állat
x1	I	N	N	N	+
x2	I	N	N	I	+
x3	N	I	N	I	+
x4	N	I	N	N	+
x5	N	N	I	I	-
x6	N	N	I	N	-

A példa halmazt képviselő döntési fának  $I_{fa} = I(1/3, 2/3) = 0.9183$  bit az információ tartalma. Vegyük sorra az egyes attribútum tesztek nyereségét:

$$R(\text{Hörcsög}) = 2/6 I(1, 0) + 4/6 I(1/2, 1/2) = 2/3$$

$$R(\text{TengeriMalac}) = 2/6 I(1,0) + 4/6 I(1/2, 1/2) = 2/3$$

$$R(\text{Kutya}) = 2/6 I(0, 1) + 4/6 I(1, 0) = 0$$

$$R(\text{Nagy}) = 1/2 I(2/3, 1/3) + 1/2 I(2/3, 1/3) = I(2/3, 1/3) = 0.9183$$

$$N_y(\text{Hörcsög}) = I_{fa} - R(\text{Hörcsög}) = 0.9183 - 2/3$$

$$N_y(\text{TengeriMalac}) = I_{fa} - R(\text{TengeriMalac}) = 0.9183 - 2/3$$

$$N_y(\text{Kutya}) = I_{fa} - R(\text{Kutya}) = 0.9183$$

$$N_y(\text{Nagy}) = I_{fa} - R(\text{Nagy}) = 0$$

Az első teszt tehát a „Kutya”. A teszt után az „Igen” ágban csúpan negatív példa, a „Nem” ágban csúpan pozitív példa marad, ezek tehát már a „Nem” és az „Igen” levelek és a fát ezzel be is fejeztük.

#### Példa 4. Döntési fa számítása/4

Egy ágens kellemes házi állatka kiválasztásához annak logikai definícióját szeretne megtanulni döntési fa módszerrel az alábbi 8 példa alapján. Információelméleti mennyiségek mérlegelésével tervezze meg ezt a döntési fát!

	Hörcsög	Agresszív	Nagy	Kellemes Állat
x1	I	N	N	+
x2	I	N	I	-
x3	I	I	N	+
x4	I	I	I	-
x5	N	N	N	+
x6	N	N	I	+
x7	N	I	N	-
x8	N	I	I	-

$$I(1/2, 1/2) = 1$$

$$N_y(H) = I - (1/2 \times I(1/2, 1/2) + 1/2 \times I(1/2, 1/2)) = 1 - 1 = 0$$

$$N_y(A) = I - (1/2 \times I(1/4, 3/4) + 1/2 \times I(3/4, 1/4)) = 1 - 0.8113 = 0.1887$$

$$N_y(N) = I - (1/2 \times I(1/4, 3/4) + 1/2 \times I(3/4, 1/4)) = 1 - 0.8113 = 0.1887$$

$$I(1/4, 3/4) = 0.8113$$

Az első lehet az **Agresszív**, vagy a **Nagy**. Legyen pl. az **Agresszív**, akkor most **Agresszív = Igen** ágba:

	Hörcsög	Nagy	Kellemes Állat
x3	I	N	+
x4	I	I	-
x7	N	N	-
x8	N	I	-

$$I'(1/4, 3/4) = 0.8113$$

$$N_y(H) = I' - (1/2 \times I(1/2, 1/2) + 1/2 \times I(0, 1)) = 0.8113 - 1/2 = 0.3113$$

$$N_y(N) = I' - (1/2 \times I(1/2, 1/2) + 1/2 \times I(0, 1)) = 0.8113 - 1/2 = 0.3113$$

Bármelyik lehet a második. Legyen a második pl. a **Nagy**, akkor most **Nagy = Igen** ágba:

	Nagy	Kellemes Állat
x4	I	-
x8	I	-

Itt lesz már egy levél (**Nem**), de a:

**Nagy = Nem** ágban:

	Nagy	Kellemes Állat
x3	N	+
x7	N	-

nem lesz levél, mert szükség van még további attribútumra, ami a **Hörsög**:

b. **Agresszív = Nem** ágban:

	Hörsög	Nagy	Kellemes Állat
x1	I	N	+
x2	I	I	-
x5	N	N	+
x6	N	I	+

a helyzet hasonló lesz.

### Példa 5. Döntési fa számítása/5

Egy ágens kellemes házi állatka kiválasztásához annak logikai definícióját szeretne megtanulni döntési fa módszerrel az alábbi 8 példa alapján. Információelméleti mennyiségek mérlegelésével tervezze meg ezt a döntési fát!

	Kutya	Szelíd	Nagy	Kellemes Állat
x1	I	N	N	+
x2	I	N	I	-
x3	I	I	N	+
x4	I	I	I	+
x5	N	N	N	+
x6	N	N	I	-
x7	N	I	N	+
x8	N	I	I	+

$$I = I(1/4, 3/4) = 0.8113$$

$$N_y(K) = I - (1/2 \times I(1/4, 3/4) + 1/2 \times I(3/4, 1/4)) = 0$$

$$N_y(S) = I - (1/2 \times I(0, 1) + 1/2 \times I(1/2, 1/2)) = 0.3113$$

$$N_y(N) = I - (1/2 \times I(1/2, 1/2) + 1/2 \times I(0, 1)) = 0.3113$$

$$I(1/4, 3/4) = 0.8113$$

a. Legyen az első pl. a **Szelid**, akkor most

**Szelid = Igen** ágban:

	Kutya	Nagy	Kellemes Állat
x3	I	N	+
x4	I	I	+
x7	N	N	+
x8	N	I	+

Itt máris egy levelet alakíthatunk (**Igen**)

A **Szelid = Nem** ágban:

	Kutya	Nagy	Kellemes Állat
x1	I	N	+
x2	I	I	-
x5	N	N	+
x6	N	I	-

Itt még tovább kell lépni, a **Kutya** nyeresége kisebb lesz, mint a **Nagy** teszté ( $Ny(K) = 0$ ,  $Ny(N) = 1$ ). A **Nagy** beválasztása után szintén levélhez jutunk el (**Igen, Nem**)

### **Példa 6. Döntési fa számítása/6**

Egy döntési fa tanulásánál a tanító példahalmaz 5 példából áll. A példaattributumok a, b, és c, a d a besorolás oszlopa. Információ mérlegelése alapján döntse el, hogy a fát hogyan építené meg.

	a	b	c	d
x1	I	I	N	N
x2	N	I	N	I
x3	N	I	I	I
x4	N	N	I	N
x5	I	N	N	N

a	I	p/n	0/2
	N	p/n	2/1
b	I	p/n	2/1
	N	p/n	0/2
c	I	p/n	1/1
	N	p/n	1/2

$$I_f = I(2/5, 3/5) = -2/5 \lg_2(2/5) - 3/5 \lg_2(3/5) = 0.971$$

$$N_y(a) = I_f - (2/5 I(0,1) + 3/5 I(2/3, 1/3))$$

$$N_y(b) = I_f - (3/5 I(2/3,1/3) + 2/5 I(0, 1))$$

$$N_y(c) = I_f - (2/5 I(1/2,1/2) + 3/5 I(1/3, 2/3))$$

$$I(2/3, 1/3) = -2/3 \lg_2(2/3) - 1/3 \lg_2(1/3) = 0.9183$$

$$N_y(a) = 0.971 - 3/5 \times 0.9183 = .42$$

$$N_y(b) = 0.971 - 3/5 \times 0.9183 = .42$$

$$N_y(c) = 0.971 - 2/5 - 3/5 \times 0.9183 = .02$$

$$+(x_2 \ x_3) \ -(x_1 \ x_4 \ x_5)$$

teszt1 = a

$$a = I \quad +(-) \ -(x_1 \ x_5) \quad \text{NEM}$$

$$a = N \quad +(x_2 \ x_3) \ -(x_4) \quad \text{tovább}$$

teszt2 = b

$$b = I \quad +(x_2 \ x_3) \ -(-) \quad \text{IGEN}$$

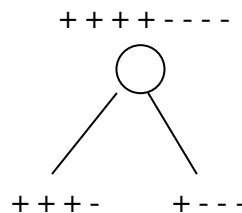
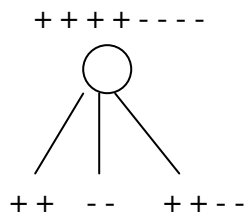
$$b = N \quad +(-) \ -(x_4) \quad \text{NEM}$$

### Példa 7. Döntési fa számítása/7

Egy döntési fa szerkesztésénél a további elágazást az  $A_1$ , ill.  $A_2$  attribútum teszttel lehet folytatni. A + és a - a fa adott helyén értelmezett pozitív és a negatív példát jelenti.

Az  $A_1$  attribútum teszt után

Az  $A_2$  attribútum teszt után



Számítsa ki a  $Nyereség(A_1)$  és a  $Nyereség(A_2)$  értékeit. A számításnál használja  $I(1/2, 1/2)=1$  bit,  $I(0,1)=0$  bit és  $I(1/4, 3/4)=0.8$  bit értékeket. Melyik tesztet építene be a fába és miért?

*Az attribútum nyeresége a fa információ tartalma az attribútumig, levonva az attribútum teszt utáni fa átlagos információ tartalmát (ld. könyvbeli képlet).*

*Itt  $A_1$  esetén:  $I(1/2, 1/2) - (1/4 \times I(0,1) + 1/4 \times I(1,0) + 1/2 \times I(1/2, 1/2)) = 0.5$*

*és*

*$A_2$  esetén:  $I(1/2, 1/2) - (1/2 \times I(3/4, 1/4) + 1/2 \times I(3/4, 1/4)) = 1 - 1/2 \times 0.8 - 1/2 \times 0.8 = 0.2$*

*Azaz az  $A_1$ -et kell választani.*



### Példa 8. A Pillanatnyilag Legjobb Hipotézis számítása/1

Ágens a pillanatnyi legjobb hipotézis módszerével tanul. Alábbi 6 példát kap. A kiinduló hipotézis az első példa logikai kifejezése. Írja le a tanulás folyamatát, azaz milyen a soron lévő példa (jó, HP, HN), mi a művelet (általánosítás, specializálás) és mi a pillanatnyi legjobb hipotézis alakja?

	Hörcsög	Tengeri Malac	Kutya	Nagy	Példa jellege
x1	I	N	N	N	+
x2	I	N	N	I	-
x3	N	I	N	I	-
x4	N	I	N	N	+
x5	N	N	I	I	-
x6	N	N	I	N	+

A kiinduló hipotézis:  $H1(x) = x1 = \text{Hörcsög}(x) \wedge \neg \text{Tengeri Malac}(x) \wedge \neg \text{Kutya}(x) \wedge \neg \text{Nagy}(x)$

*Az első példa tehát már megvan.*

*Az x2 példa egy negatív példa és  $H1(x2) = \text{Hamis}$ , tehát ez is rendben van.*

*Az x3 példa egy negatív példa és  $H1(x3) = \text{Hamis}$ , tehát ez is rendben van.*

*Az x4 példa egy pozitív példa és  $H1(x4) = \text{Hamis}$ , tehát ez egy Hamis Negatív és valamit tenni kell a hipotézissel. A hipotézist általánosítani kell konjunkció elhagyásával, de úgy, hogy az új hipotézis az eddigi példákon jól viselkedjen.*

$$H2(x) = \neg \text{Kutya}(x) \wedge \neg \text{Nagy}(x)$$

*Az x5 példa egy negatív példa és  $H2(x5) = \text{Hamis}$ , tehát ez is rendben van.*

*Az x6 példa egy pozitív példa és  $H2(x6) = \text{Hamis}$ , tehát ez egy Hamis Negatív és valamit tenni kell a hipotézissel. A hipotézist általánosítani kell konjunkció elhagyásával, de úgy, hogy az új hipotézis az eddigi példákon jól viselkedjen.*

$$H3(x) = \neg \text{Nagy}(x)$$

### Példa 8. A Pillanatnyilag Legjobb Hipotézis számítása/1

Ágens a pillanatnyi legjobb hipotézis módszerével tanul. Alábbi 5 példát kap. A kiinduló hipotézis a H1. Írja le a tanulás folyamatát.

	Fekete	Nagy	Kerek	Milyen példa
x1	I	N	I	+
x2	N	N	N	-
x3	I	I	I	+
x4	N	I	N	-
x5	I	N	N	+

A kiinduló hipotézis:  $H1(x) = x1 = \text{Fekete}(x) \wedge \neg \text{Nagy}(x) \wedge \text{Kerek}(x)$

x2 helyes negatív  $H2(x) = H1(x) = \text{Fekete}(x) \wedge \neg \text{Nagy}(x) \wedge \text{Kerek}(x)$

x3 HN általánosítás  $H3(x) = \text{Fekete}(x) \wedge \neg \text{Nagy}(x)$   
nem jó, x3-t nem fogadja

$$H3(x) = \text{Fekete}(x) \wedge \text{Kerek}(x)$$

x4 helyesen negatív  $H4(x) = \text{Fekete}(x) \wedge \text{Kerek}(x)$

x5 HN általánosítás  $H5(x) = \text{Kerek}(x)$

nem jó, x5-t nem fogadja el

$$H5(x) = \text{Fekete}(x)$$

kifogytunk a példákból,  $H5(x)$  minden példát jól sorolja be.