Amy Seo
505 328 863

1.

Example	Input Attributes			Class	#
	A	B	C	D	
$\mathbf{x}_1$	t	t	t	Yes	1
$\mathbf{x}_2$	t	t	f	Yes	6
$\mathbf{x}_3$	t	f	t	No	3
$\mathbf{x}_4$	t	f	f	No	1
$\mathbf{x}_5$	f	t	t	Yes	1
$\mathbf{x}_6$	f	t	f	No	6
$\mathbf{x}_7$	f	f	t	Yes	2
$\mathbf{x}_8$	f	f	f	No	2

- initial entropy:

$$ENT(D) = -\sum_x Pr(x)\log_2(x)$$
$$= -\left(\frac{10}{22}\log_2\left(\frac{10}{22}\right) + \frac{12}{22}\log_2\left(\frac{12}{22}\right)\right)$$
$$= 0.994$$

- choosing first attribute to split:
  - attribute A:

$$ENT(D|A) = \sum_a Pr(a)\,ENT(D|a)$$
$$= \left(\frac{11}{22}\,ENT(D|a)\right) + \left(\frac{11}{22}\,ENT(D|\bar{a})\right)$$
$$= \underbrace{\left(\frac{11}{22}\times\left(-\left(\frac{7}{11}\log_2\frac{7}{11} + \frac{4}{11}\log_2\frac{4}{11}\right)\right)\right)}_{A=t} + \underbrace{\left(\frac{11}{22}\times\left(-\left(\frac{3}{11}\log_2\frac{3}{11} + \frac{8}{11}\log_2\frac{8}{11}\right)\right)\right)}_{A=f}$$
$$\approx 0.8955$$

  - attribute B:

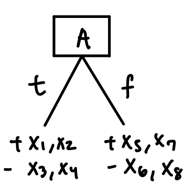
$$ENT(D|B) = \sum_b Pr(b)\,ENT(D|b)$$
$$= \left(\frac{14}{22}\,ENT(D|b)\right) + \left(\frac{8}{22}\,ENT(D|\bar{b})\right)$$
$$= \underbrace{\left(\frac{14}{22}\times\left(-\left(\frac{8}{14}\log_2\frac{8}{14} + \frac{6}{14}\log_2\frac{6}{14}\right)\right)\right)}_{B=t} + \underbrace{\left(\frac{8}{22}\times\left(-\left(\frac{2}{8}\log_2\frac{2}{8} + \frac{6}{8}\log_2\frac{6}{8}\right)\right)\right)}_{B=f}$$
$$\approx 0.9220$$

  - attribute C:

$$ENT(D|C) = \sum_c Pr(c)\,ENT(D|c)$$
$$= \left(\frac{7}{22}\,ENT(D|c)\right) + \left(\frac{15}{22}\,ENT(D|c)\right)$$
$$= \underbrace{\left(\frac{7}{22}\times\left(-\left(\frac{4}{7}\log_2\frac{4}{7} + \frac{3}{7}\log_2\frac{3}{7}\right)\right)\right)}_{C=t} + \underbrace{\left(\frac{15}{22}\times\left(-\left(\frac{6}{15}\log_2\frac{6}{15} + \frac{9}{15}\log_2\frac{9}{15}\right)\right)\right)}_{C=f}$$
$$\approx 0.9755$$

based on conditional entropies, attribute A is the most discriminating since its value was the lowest.

- choosing second attribute to split:



- for $A = t$: $\{x_1, x_2, x_3, x_4\}$

  → attribute B:

$$ENT(D|B) = \sum_b Pr(b)\, ENT(D|b)$$
$$= \frac{7}{11} ENT(D|b) + \frac{4}{11} ENT(D|\bar{b})$$
$$= \left(\frac{7}{11} \times -(1 \log 1)\right) + \left(\frac{4}{11} \times -(1 \log 1)\right)$$
$$= 0$$

  → attribute C:

$$ENT(D|C) = \sum_c Pr(c)\, ENT(D|c)$$
$$= \frac{4}{11} ENT(D|c) + \frac{7}{11} ENT(D|\bar{c})$$
$$= \left(\frac{4}{11} \times -\left(\frac{1}{4}\log_2\frac{1}{4} + \frac{3}{4}\log_2\frac{3}{4}\right)\right) + \left(\frac{7}{11} \times -\left(\frac{6}{7}\log_2\frac{6}{7} + \frac{1}{7}\log_2\frac{1}{7}\right)\right)$$
$$= 0.6715$$

B is the next most discriminating attribute as its conditional entropy is 0.

- for $A = f$: $\{x_5, x_6, x_7, x_8\}$

  → attribute B:

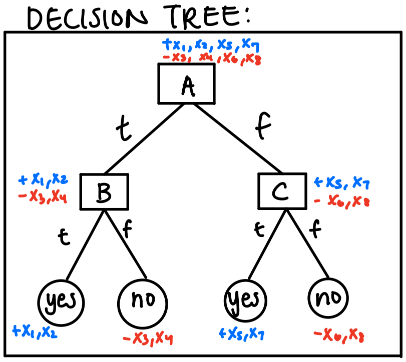
$$ENT(D|B) = \sum_b Pr(b)\, ENT(D|b)$$
$$= \frac{7}{11} ENT(D|b) + \frac{4}{11} ENT(D|\bar{b})$$
$$= \left(\frac{7}{11} \times -\left(\frac{1}{7}\log_2\frac{1}{7} + \frac{6}{7}\log_2\frac{6}{7}\right)\right) + \left(\frac{4}{11} \times -\left(\frac{2}{4}\log_2\frac{2}{4} + \frac{2}{4}\log_2\frac{2}{4}\right)\right)$$
$$= 0.7402$$

  → attribute C:

$$ENT(D|C) = \sum_c Pr(c)\, ENT(D|c)$$
$$= \frac{3}{11} ENT(D|c) + \frac{8}{11} ENT(D|\bar{c})$$
$$= \left(\cancel{\frac{3}{11} \times -}\underbrace{(1 \log 1)}_{0}\right) + \left(\cancel{\frac{8}{11} \times -}\underbrace{(1 \log 1)}_{0}\right)$$
$$= 0$$

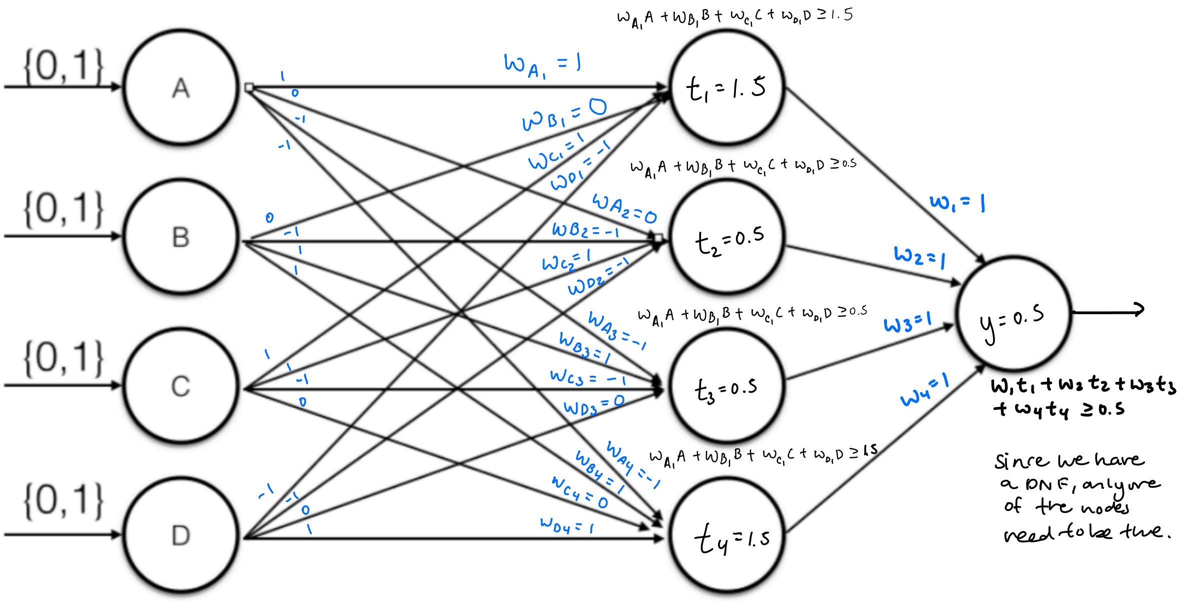
C is the next most discriminating attribute as its conditional entropy is 0.

DECISION TREE:



2. $(A \lor \neg B) \oplus (\neg C \lor D)$

$p \oplus q$ can be rewritten as $(p \land \neg q) \lor (\neg p \land q)$
letting $p = A \lor \neg B$ and $q = \neg C \lor D$, we can rewrite the original expression as:
$((A \lor \neg B) \land (\neg(\neg C \lor D))) \lor ((\neg(A \lor \neg B)) \land (\neg C \lor D))$
$((A \lor \neg B) \land (C \land \neg D)) \lor ((\neg A \land B) \land (\neg C \lor D))$
$(A \land C \land \neg D) \lor (\neg B \land C \land \neg D) \lor (\neg A \land B \land \neg C) \lor (\neg A \land B \land D)$

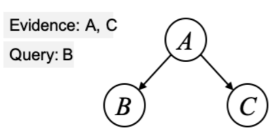


$w_1 t_1 + w_2 t_2 + w_3 t_3 + w_4 t_4 \geq 0.5$

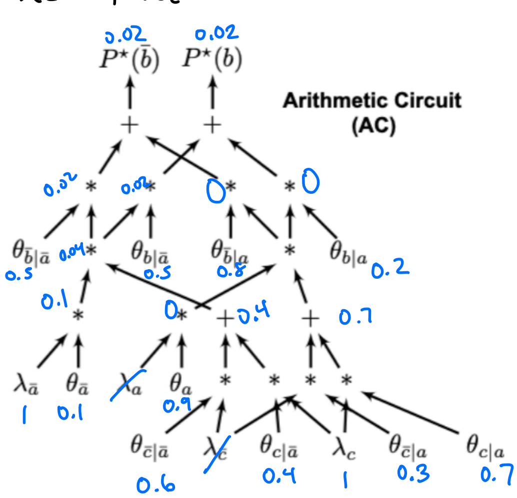
Since we have a DNF, only one of the nodes need to be true.

3. $\theta_a = 0.9$
$\theta_{b|a} = 0.2$
$\theta_{b|\bar{a}} = 0.5$
$\theta_{c|a} = 0.7$
$\theta_{c|\bar{a}} = 0.4$

evidence:
$e_1 = \bar{a}, c$
$e_2 = \bar{a}, \bar{c}$
$e_3 = \bar{a}$



a) $e_1 = \bar{a}, c$
$\lambda_a = 0, \lambda_{\bar{a}} = 1$
$\lambda_c = 1, \lambda_{\bar{c}} = 0$

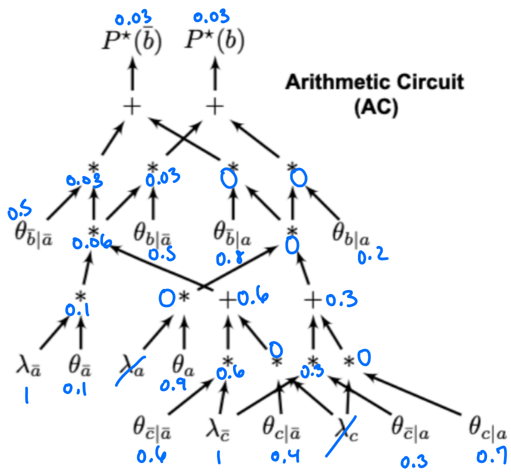


$P^*(\bar{b}) = 0.02$
$P^*(b) = 0.02$

$e_2 = \bar{a}, \bar{c}$

$\lambda_a = 0, \lambda_{\bar{a}} = 1$

$\lambda_c = 0, \lambda_{\bar{c}} = 1$



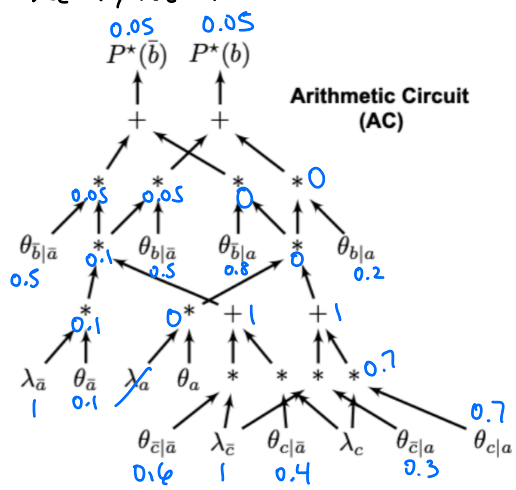
$P^*(\bar{b}) = 0.03$

$P^*(b) = 0.03$

$e_3 = \bar{a}$

$\lambda_a = 0, \lambda_{\bar{a}} = 1$

$\lambda_c = 1, \lambda_{\bar{c}} = 1$



b) The 2 circuit outputs $P^*(b)$ and $P^*(\bar{b})$ represent the probability of b and evidence e_i happening, and the probability of not b and evidence e_i happening respectively:

$P^*(b)$ means $P(b, e)$

$P^*(\bar{b})$ means $P(\bar{b}, e)$

c) $Pr(\bar{b} \mid e_1) = \frac{Pr(\bar{b}, e_1)}{Pr(e_1)} = \frac{Pr(\bar{b}, e_1)}{Pr(b, e_1) + Pr(\bar{b}, e_1)} = \frac{0.02}{0.02+0.02} = \boxed{\frac{1}{2}}$

$Pr(\bar{b} \mid e_2) = \frac{Pr(\bar{b}, e_2)}{Pr(e_2)} = \frac{Pr(\bar{b}, e_2)}{Pr(b, e_2) + Pr(\bar{b}, e_2)} = \frac{0.03}{0.03+0.03} = \boxed{\frac{1}{2}}$

$Pr(\bar{b} \mid e_3) = \frac{Pr(\bar{b}, e_3)}{Pr(e_3)} = \frac{Pr(\bar{b}, e_3)}{Pr(b, e_3) + Pr(\bar{b}, e_3)} = \frac{0.05}{0.05+0.05} = \boxed{\frac{1}{2}}$