papergrid

Date: / /

# COM SCI 161: HOMEWORK 9.

## QUESTION 1.

Step 1. To find the attribute that minimizes the conditional entropy of D:

A: $ENT(D|A) = \sum_a Pr(a)\, ENT(D|a)$

$= Pr(a)\, ENT(D|a) + Pr(\bar{a})\, ENT(D|\bar{a})$

$= \frac{11}{22} ENT(D|a) + \frac{11}{22} ENT(D|\bar{a})$

$= \frac{1}{2} ENT(D|a) + \frac{1}{2} ENT(D|\bar{a})$

$= \frac{1}{2}\left[-\left[Pr(d|a)\log_2(Pr(d|a)) + Pr(\bar{d}|a)\log_2(Pr(\bar{d}|a))\right]\right]$

$+ \frac{1}{2}\left[-\left[Pr(d|\bar{a})\log_2(Pr(d|\bar{a})) + Pr(\bar{d}|\bar{a})\log_2(Pr(\bar{d}|\bar{a}))\right]\right]$

$= \frac{1}{2}\left[-\left[(7/11)\log_2(7/11) + (4/11)\log_2(4/11)\right]\right]$

$+ \frac{1}{2}\left[-\left[3/11\log_2(3/11) + (8/11)\log_2(8/11)\right]\right].$

$= 0.47283 + 0.42268$

$= 0.8955$

B: $ENT(D|B) = \sum_b Pr(b)\, ENT(D|b)$

$= \frac{14}{22} ENT(D|b) + \frac{8}{22} ENT(D|\bar{b})$

$= \frac{14}{22}\left[-\left[Pr(d|b)\log_2(Pr(d|b)) + Pr(\bar{d}|b)\log_2(Pr(\bar{d}|b))\right]\right]$

$+ \frac{8}{22}\left[-\left[Pr(d|\bar{b})\log_2(Pr(d|\bar{b})) + Pr(\bar{d}|\bar{b})\log_2(Pr(\bar{d}|\bar{b}))\right]\right]$

$= 14/22\left[-\left[8/14\log_2(8/14) + 6/14\log_2(6/14)\right]\right]$

$+ 8/22\left[-\left[2/8\log_2(2/8) + 6/8\log_2(6/8)\right]\right]$

$= 0.62696 + 0.29501$

$= 0.92197$

C: $ENT(C|A) = \sum_c Pr(c)\, ENT(D|c)$

$= 7/22\, ENT(D|c) + 15/22\, ENT(D|\bar{c})$

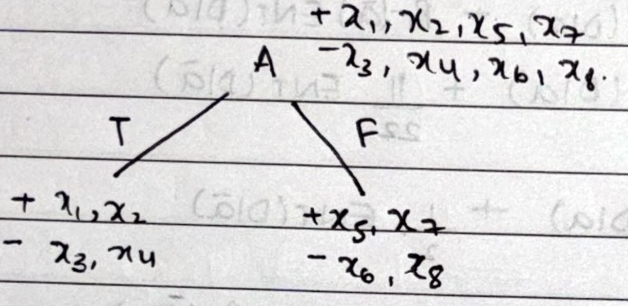
$= 7/22\left[-\left[Pr(d|c)\log_2(Pr(d|c)) + Pr(\bar{d}|c)\log_2(Pr(\bar{d}|c))\right]\right]$

$+ 15/22\left[-\left[Pr(d|\bar{c})\log_2(Pr(d|\bar{c})) + Pr(\bar{d}|\bar{c})\log_2(Pr(\bar{d}|\bar{c}))\right]\right]$

$= 7/22 [-[4/7 \log_2 4/7 + 3/7 \log_2 3/7]] + 15/22 [-[6/15 \log_2 6/15$
$+ 9/15 \log_2 9/15]]$
$= 0.31348 + 0.66201$
$= 0.97549.$

Here A has the lowest conditional entropy, so we split on A.

STEP2:




CASE 1: branch A=T $\{x_1, x_2, x_3, x_4\}$

$ENT(D|B) = \sum Pr(b)\, ENT(D|b)$
$= 7/11\, ENT(D|b) + 4/11\, ENT(D|\bar{b})$
$= 7/11 [-[Pr(d|b) \log_2 (Pr(d|b)) + Pr(\bar{d}|b) \log_2 (Pr(\bar{d}|b))]]$
$+ 4/11 [-[Pr(d|\bar{b}) \log_2 (Pr(d|\bar{b})) + Pr(\bar{d}|\bar{b}) \log_2 (Pr(\bar{d}|\bar{b}))]]$
$= 7/11 [-[7/7 \log_2 (7/7) + 0 \log_2 (0)]] + 4/11 [-[0 \log_2 (0)$
$+ 4/4 \log_2 (4/4)]]$
$= 0$

Since $ENT(D|B) = 0$, smallest conditional entropy possible, we split on B.

CASE 2: branch A=F $\{x_5, x_7, x_6, x_8\}$.

$ENT(D|B) = 7/11\, ENT(D|b) + 4/11\, ENT(D|b)$
$= 7/11 [-[Pr(d|b) \log_2 (Pr(d|b)) + Pr(\bar{d}|b) \log_2 (Pr(d|b))]]$
$+ 4/11 [-[Pr(d|\bar{b}) \log_2 (Pr(d|\bar{b})) + Pr(\bar{d}|\bar{b}) \log_2 (Pr(\bar{d}|\bar{b}))]]$
$= 7/11 [-[1/7 \log_2 1/7 + 6/7 \log_2 6/7]] + 4/11 [-[2/4 \log_2 2/4$
$+ 2/4 \log_2 2/4]]$
$= 0.37652 + 0.36363$
$= 0.74016.$

$ENT(D|C) = \frac{3}{11} ENT(D|C) + \frac{8}{11} ENT(D|C)$

$= \frac{3}{11}\left[-\left[Pr(d|c)\log_2(Pr(d|c)) + Pr(\bar{d}|c)\log_2(Pr(\bar{d}|c))\right]\right]$

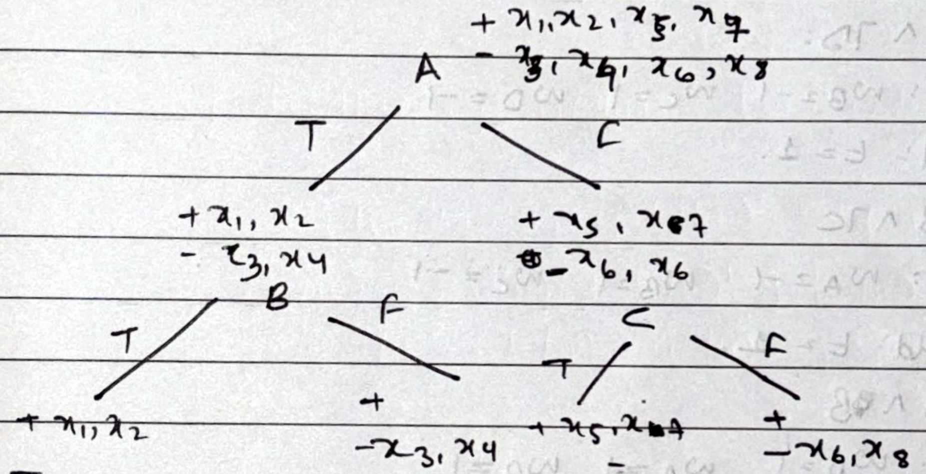
$+ \frac{8}{11}\left[-\left[Pr(d|\bar{c})\log_2(Pr(d|\bar{c})) + Pr(\bar{d}|\bar{c})\log_2(Pr(\bar{d}|\bar{c}))\right]\right]$

$= \frac{3}{11}\left[-\left[\frac{3}{3}\log_2\frac{3}{3} + 0\log_2 0\right]\right] + \frac{8}{11}\left[-\left[0\log_2 0 + \frac{8}{8}\log_2\frac{8}{8}\right]\right]$

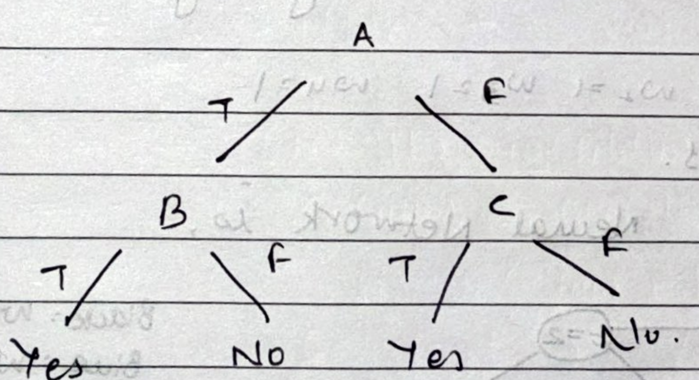
$= 0$

As $ENT(D|C) = 0$ gives the smallest possible conditional entropy possible, we split C.

STEP 3: We split on B for the branch A = T and we split on C when A = F.




The final decision tree is,



QUESTION 2.

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

$= ((A \lor \neg B) \land \neg(\neg C \lor D)) \lor (\neg(A \lor \neg B) \land (\neg C \lor D))$

$= ((A \lor \neg B) \land (C \land \neg D)) \lor ((\neg A \land B) \land (\neg C \lor D))$

$= (A \land C \land \neg D) \lor (\neg B \land C \land \neg D) \lor (\neg A \land B \land \neg C) \lor (D \land \neg A \land B)$

Weights & Thresholds:

Node 1: $A \land C \land \neg D$

Weights: $w_A = 1$ $w_C = 1$ $w_D = -1$

Threshold: $t = 2$

Node 2: $\neg B \land C \land \neg D$.

weights: $w_B = -1$ $w_C = 1$ $w_D = -1$

Threshold: $t = 1$.

Node 3: $\neg A \land B \land \neg C$

weights: $w_A = -1$ $w_B = 1$ $w_C = -1$

Threshold: $t = 1$.

Node 4: $D \land \neg A \land B$

weights: $w_D = 1$ $w_A = -1$ $w_B = 1$

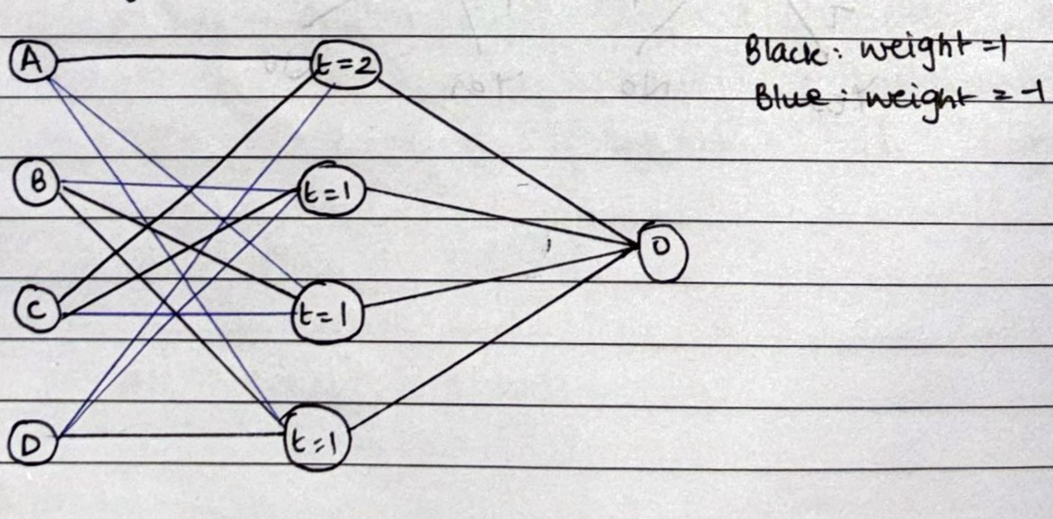
Threshold: $t = 1$.

OUTPUT: The output node will activate if any intermediate nodes are active.

weights: $w_1 = 1$ $w_2 = 1$ $w_3 = 1$ $w_4 = 1$

Threshold: $t = 1$.

Thus, the final Neural Network is,



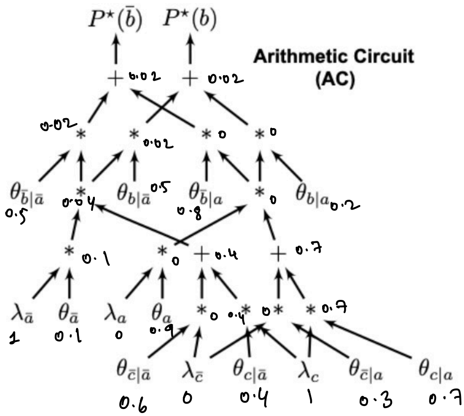
# Question 3

(a). For $e_1 = \bar{a}, c$

$\lambda_{\bar{a}} = 1 \quad \lambda_a = 0 \quad \lambda_c = 1 \quad \lambda_{\bar{c}} = 0$

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

$P^*(b) = 0.02$

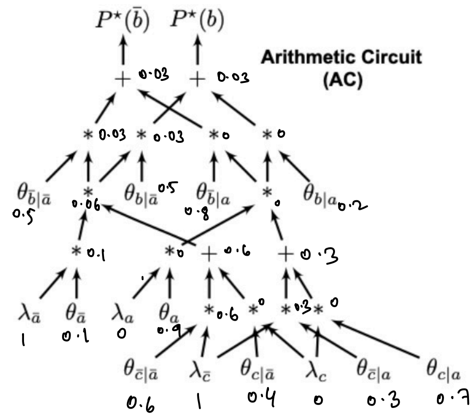



For $e_2 = \bar{a}, \bar{c}$

$\lambda_{\bar{a}} = 1 \quad \lambda_{\bar{c}} = 1 \quad \lambda_a = 0 \quad \lambda_c = 0$

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

$P^*(b) = 0.03$

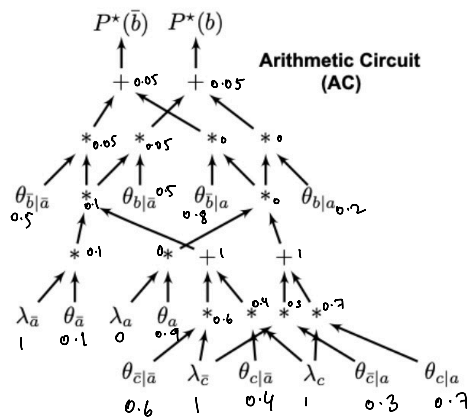



For $e_3 = \bar{a}$

$\lambda_{\bar{a}} = 1 \quad \lambda_{\bar{c}} = 1 \quad \lambda_a = 0 \quad \lambda_c = 1$

$P^*(\bar{b}) = 0.05$

$P^*(b) = 0.05$



(b) $P^*(\bar{b})$ : unnormalized probability of $\bar{b}$ given the evidence
$P^*(b)$ : unnormalized probability of $b$ given the evidence.

(c) $e_1$ : $P(\bar{b}|e_1) = \dfrac{P^*(\bar{b})}{P^*(\bar{b}) + P^*(b)} = \dfrac{0.02}{0.02+0.02} = 0.5$

$e_2$ : $P(\bar{b}|e_2) = \dfrac{P^*(\bar{b})}{P^*(\bar{b}) + P^*(b)} = \dfrac{0.03}{0.03+0.03} = 0.5$

$e_3$ : $P(\bar{b}|e_3) = \dfrac{P^*(\bar{b})}{P^*(\bar{b}) + P^*(b)} = \dfrac{0.05}{0.05+0.05} = 0.5$