ANSWER 3

	A_1	A_2	A_3	output y
x_1	1	0	0	0
x_2	1	0	1	0
x_3	0	1	0	0
x_4	1	1	1	1
x_5	1	1	0	1

We want to choose the attribute to split in such a way that remaining entropy is minimized.

i.e. we compute Remainder (A_i) for each A_i at split and select one that gives least value.

$$\Rightarrow \text{Remainder}(A_i) = \sum_{k=1}^{2} \frac{p_k + n_k}{p+n} B\left(\frac{p_k}{p_k+n_k}\right)$$

$k=1$ corresponds to $A_i = 0$, $k=2$ to $A_i = 1$

where $B(q) = -(q \log_2 q + (1-q)\log_2(1-q))$

p_k is no. of positive egs and n_k is no. of neg egs for k^{th}

first split

$$\text{Remainder}(A_1) = \frac{4}{5}\left(-\frac{2}{4}\log_2\left(\frac{2}{4}\right) - \frac{2}{4}\log_2\left(\frac{2}{4}\right)\right) + \frac{1}{5}\left(-0 - \frac{1}{1}\log\left(\frac{1}{1}\right)\right) =$$

$$\frac{4}{5}\left(\frac{1}{2} + \frac{1}{2}\right) + 0 = 0.8.$$

$$\text{Remainder}(A_2) = \frac{3}{5}\left(-\frac{2}{3}\log_2\left(\frac{2}{3}\right) - \frac{1}{3}\log_2\left(\frac{1}{3}\right)\right) + \frac{2}{5}\left(-0 - \frac{2}{2}\log_2\left(\frac{2}{2}\right)\right)$$

$$= \frac{3}{5}\frac{1}{3}\left(2\log_2\left(\frac{3}{2}\right) + \log_2(3)\right) = 0.55.$$

$$\text{Remainder}(A_3) = \frac{2}{5}\left(-\frac{1}{2}\log_2\left(\frac{1}{2}\right) - \frac{1}{2}\log_2\left(\frac{1}{2}\right)\right)$$

$$+ \frac{3}{5}\left(-\frac{1}{3}\log_2\left(\frac{1}{3}\right) - \frac{2}{3}\log_2\left(\frac{2}{3}\right)\right)$$

$$= \frac{1}{5}(1+1) + \frac{1}{5}\left(\log_2 3 + 2\log_2 \frac{3}{2}\right)$$

$$= 0.9509 = 0.95$$

Hence, A_2 minimizes remaing entropy.
choose A_2 to split.

All with $A_2 = 0$ are output $y = 0$.
so, we are left with $\{x_3, x_4, x_5\}$.
and A_1 and A_3.

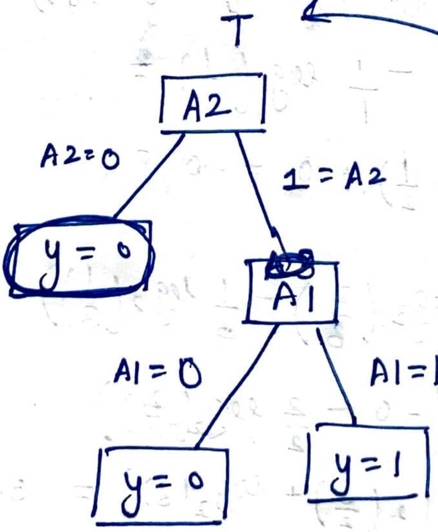
$$\text{Remainder}(A_1) = \frac{2}{3}\left(-\frac{2}{2}\log_2\left(\frac{2}{2}\right) - 0\right)$$

$$+ \frac{1}{3}\left(-0 - \frac{1}{1}\log_2\left(\frac{1}{1}\right)\right) = 0$$

$$\text{Remainder}(A_3) = \frac{1}{3}\left(-\frac{1}{1}\log_2\left(\frac{1}{1}\right)\right) + \frac{2}{3}\left(\frac{-1}{2}\log_2\left(\frac{1}{2}\right) - \frac{1}{2}\log_2\left(\frac{1}{2}\right)\right)$$

$$= \frac{1}{3}(0) + \frac{2}{3}\left(\frac{1}{2} + \frac{1}{2}\right) = \frac{2}{3} = 0.67$$

Hence, A_1 minimizes remaining entropy to 0.
i.e. we split wrt A_3 now.



final decision tree