1.1 The source entropy, H is given by $H = P_n \log_2(1/P_n)$ Substituting the known values,

$H = 0.5 \log_2(1/0.5) + 0.125 \log_2(1/0.125) + 0.125 \log_2(1/0.125) + 0.125 \log_2(1/0.125) + 0.125 \log_2(1/0.125) = 2$

It means the average amount of information from the event is 2.

1.2 Using the formula:

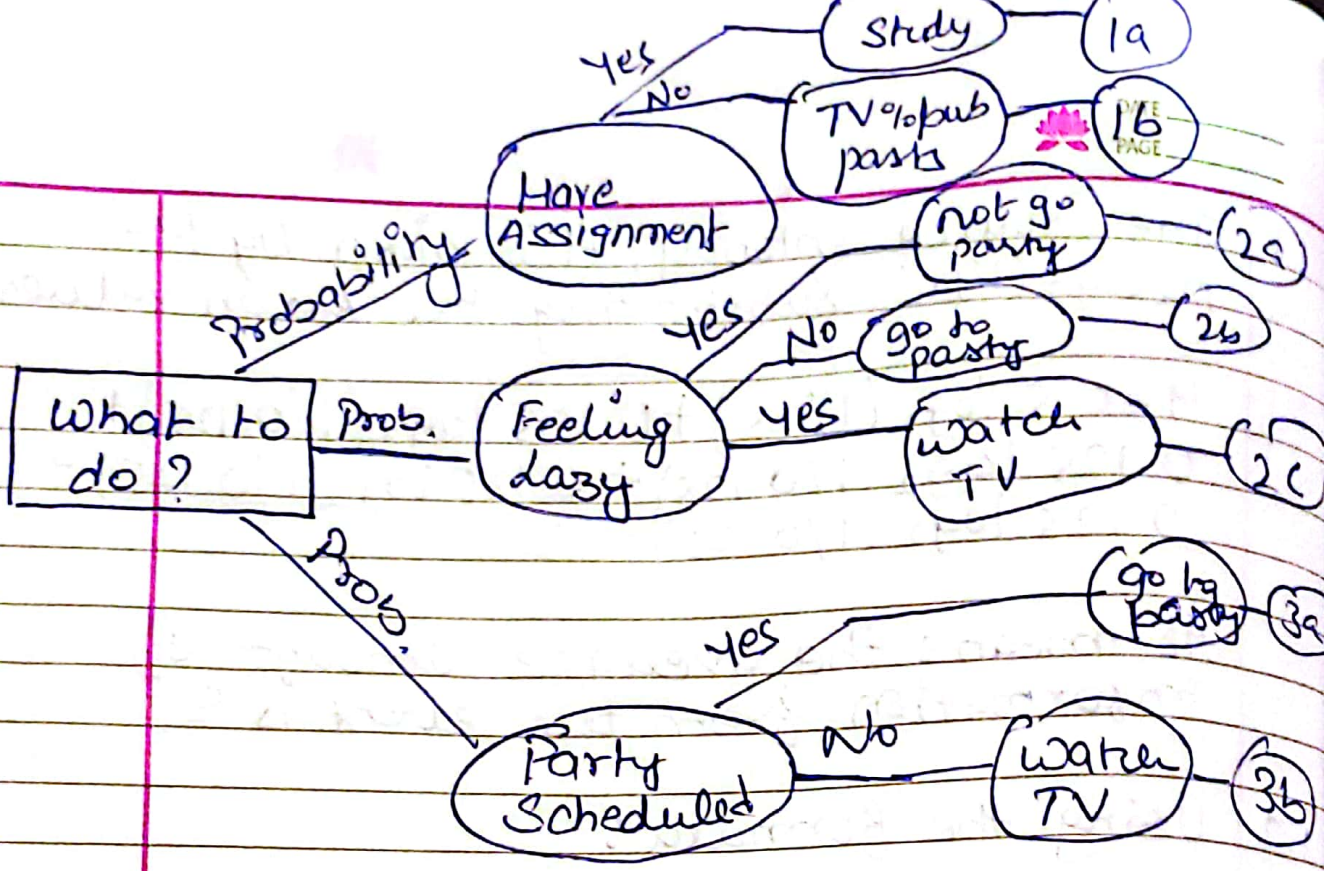
For N1 - Entropy = 0, Gini impurity = 0 and classification error = 0

For N2 - Entropy = 0.65, Gini impurity = 0.278 and classification error = 1/6

For N3 - Entropy = 1, Gini impurity = 1/2 and classification error = 0.5

1.3
- Going to Pub
- Watching TV
- Going to Party
- Studying

1. If assignment next day - Study
2. Feeling lazy - No Pub.
3. No Party scheduled - cannot go to party



1.4 ID3 (Examples, Target_Attribute, Attributes).

Create a root node for the tree

If all examples are positive, Return the single-node tree Root, with Label = +.

If all examples are negative. Return the single-node tree Root with Label = -.

If number of Predicting attributes is empty, then Return the single node tree Root,

with Label = most common value of the target attribute in the examples

Otherwise Begin

A ← The attribute that best classifies examples.

Decision Tree attribute for root = A

For each possible value, v_i, of A,
Add a new tree branch below root, corresponding to the test $A = v_i$.
Let Examples(v_i) be the subset of examples that have the value v_i for A
If examples(v_i) is empty
Then below this new branch add a leaf node with label = most common target values in the examples
Else, below this new branch add the subtree ID3(Example(v_i), target - attribute, attributes - {A})

End
Return Root