Department of Computer Science

California State University, Channel Islands

MATHCOMPPH-546: Pattern Recognition

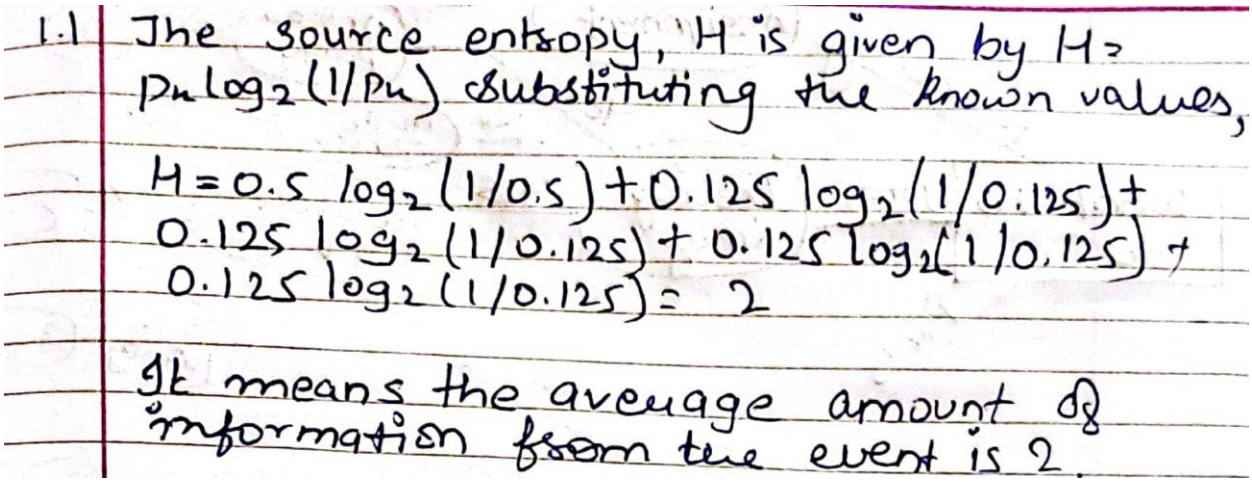
Lesson 3 phys546 Non-Metric Methods HW_3

Student Name: Sandipta Subir Khare

Student Major: Computer Science

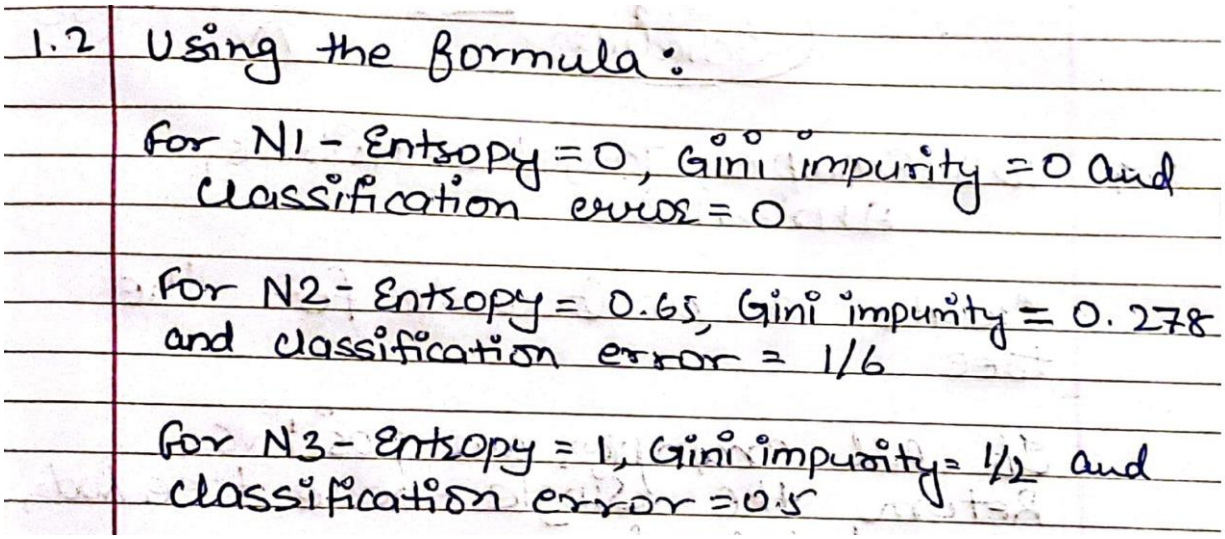
1.1 Suppose that the probability of five events are P(1) ¼ 0.5, and P(2) ¼ P(3) ¼ P (4)¼ P(5) ¼ 0.125. Calculate the entropy. Explain in words what this means.

Ans.



1.2 Three binary nodes, N1, N2, and N3, split examples into (0, 6), (1,5), and (3,3), respectively. For each node, calculate its entropy, Gini impurity, and classification error.

Ans.

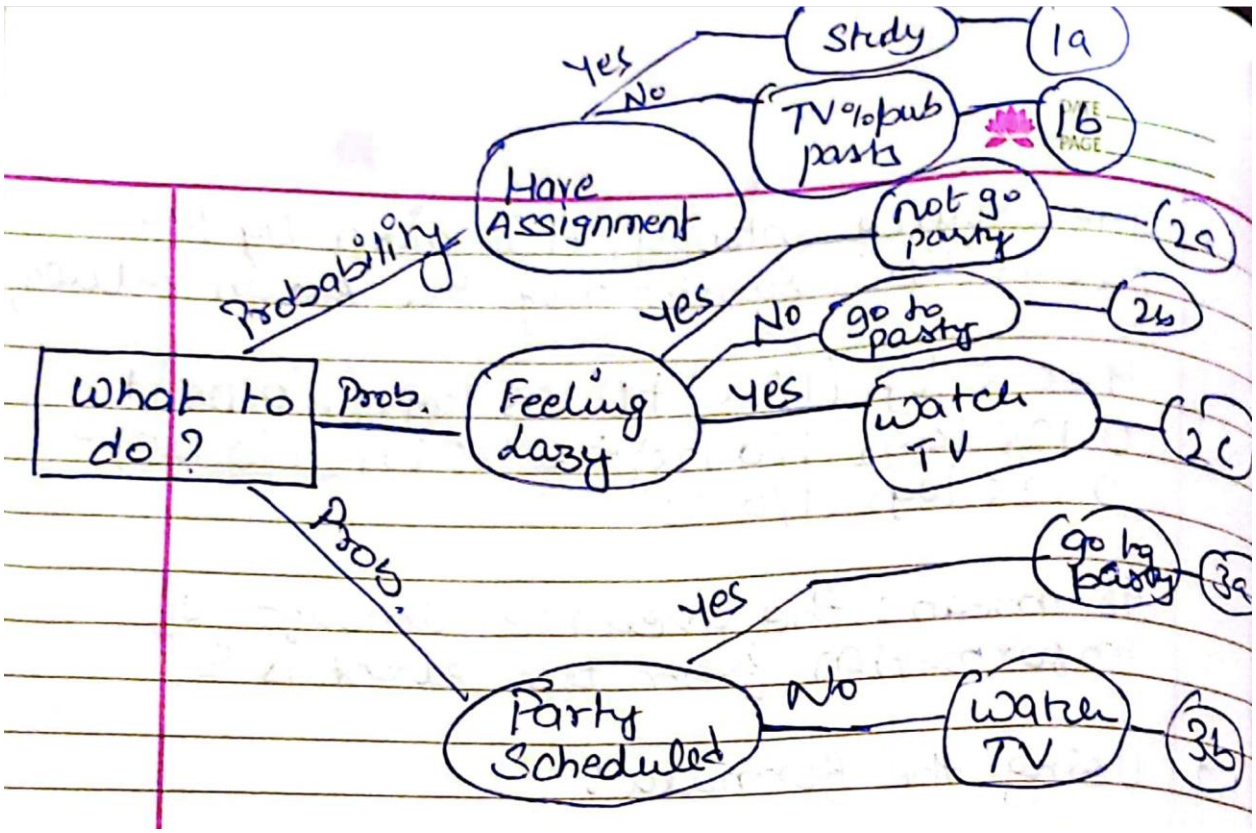


1.3 Build a decision tree that computes the logical AND function.

Ans.

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 Imagine that there are four things that you like to do in the evening: going to a pub, watching TV, going to a party, or studying (!). The choice is sometimes made for you—if you have an assignment due the next day, you need to study, if you're feeling lazy then the pub isn't for you, and if there isn't a party then you can't got it. You are looking for a decision tree which will help you decide what to do each evening. Here is a list of everything you've done in the past 10 days.

Deadline?	Is there a party?	Lazy?	Activity
Urgent	Yes	Yes	Party
Urgent	No	Yes	Study
Near	Yes	Yes	Party
None	Yes	No	Party
None	No	Yes	Pub
None	Yes	No	Party
Near	No	No	Study
Near	No	Yes	TV
Near	Yes	Yes	Party
Urgent	No	No	Study

(The first thing to do is to work out which feature to use as the starting (root) node. For this you need to compute the entropy, and then find out which feature has the maximal information gain).

Ans.

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