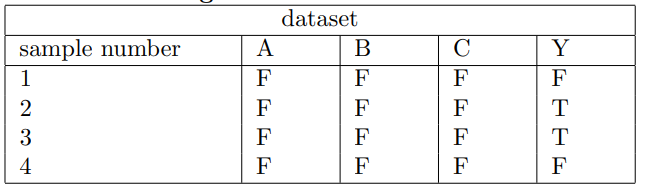
**Theoretical Part**

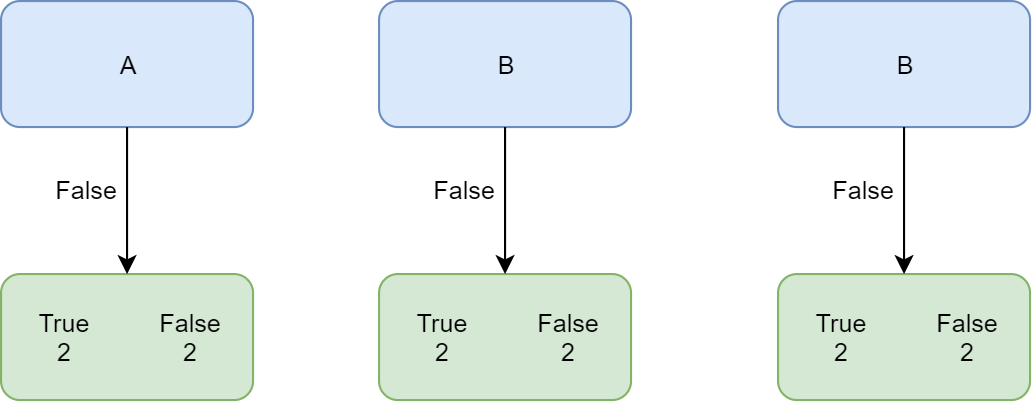
Daniil Rolnik 334018009

1. Decision Tree Algorithm:



1. *Using the dataset above, we want to build a decision tree which classifies Y as True (T) or False (F) given the binary variables A, B, C. Draw the tree that would be learned by the greedy algorithm with zero training error. You do need to show computations.*

Let’s build decision tree from given Boolean table. We will use Gini algorithm.

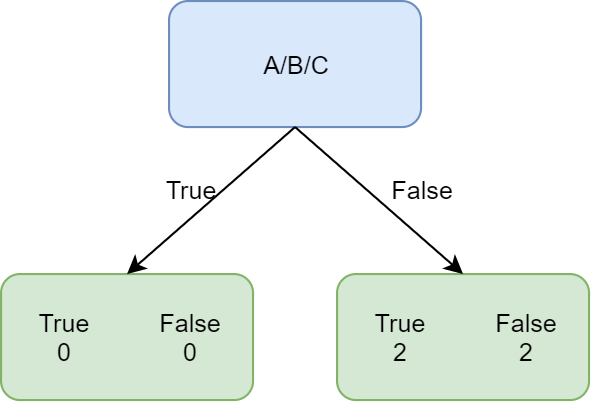


As we can see all three inputs (entered as false) can lead us to both to True and False classification. Let’s check Gini impurity for every input:

**Gini formula** = 1 – (the probability of “true”)^2 – (the probability of “false”)^2

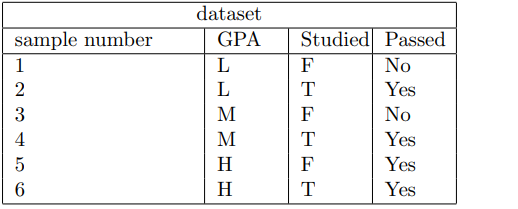
Gini for A = Gini for B = Gini for C = 1 – (2/4)^2 – (2/4)^2 = 0.5

All three Ginis are equal so it makes no difference which node we will take as root node for our decision tree. Let’s take A as root node. If we’ll continue with Gini impurity, we’ll see that all leaves will have the same results, so final result is:



1. *Is this tree optimal (i.e., does it get zero training error with minimal depth)? Explain in less than two sentences. If it is not optimal, draw the optimal tree as well.*

Because we have error equal to 50% in leaves in the best case, the tree is not optimal.



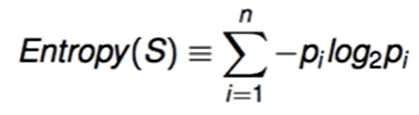
* 1. *We will use the dataset below to learn a decision tree which predicts if people pass machine learning (Yes or No), based on their previous GPA (High, Medium, or Low) and whether or not they studied. On the following calculations use log on base of 2.*

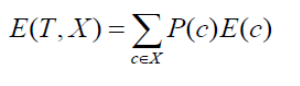
*What is the entropy of H(Passed) ?*

*What is the entropy of H(Passed | GPA)?*

*What is the entropy of H(Passed | Studied)?*

Formulas for entropy:





|  |  |  |
| --- | --- | --- |
| **Entropy** | **Calculation** | **Answer** |
| H(Passed) | (-4/6)\*log2(4/6) + (-2/6)\*log2(2/6) | 0.9183 |
| H(Passed | GPA) | P(L)\*E(1,1) + P(M)\*E(1,1) + P(H)\*E(2,0) =  (2/6)\*((-1/6)\*log2(1/6)+(-1/6)\*log2(1/6)) +  (2/6)\*((-1/6)\*log2(1/6)+(-1/6)\*log2(1/6)) +  (2/6)\*((-2/6)\*log2(2/6) + 0) = 0.2872 + 0.2872 + 0.1761 | 0.75 |
| H(Passed | Studied) | P(T)\*E(3,0) + P(F)\*E(1, 2) =  (3/6)\*((-3/6)\*log2(3/6)+0) + (3/6)\*((-1/6)\*log2(1/6) + (-2/6)\*log2(2/6)) =  0.25 + 0.48 | 0.73 |

7. *Draw the full decision tree, that would be learned for this dataset. You do not need to show any calculations.*

