# 

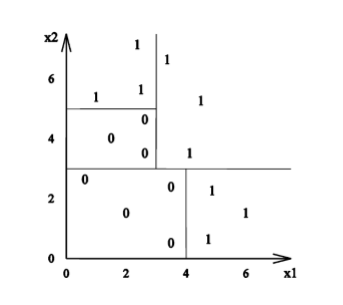
# I526 Applied Machine Learning

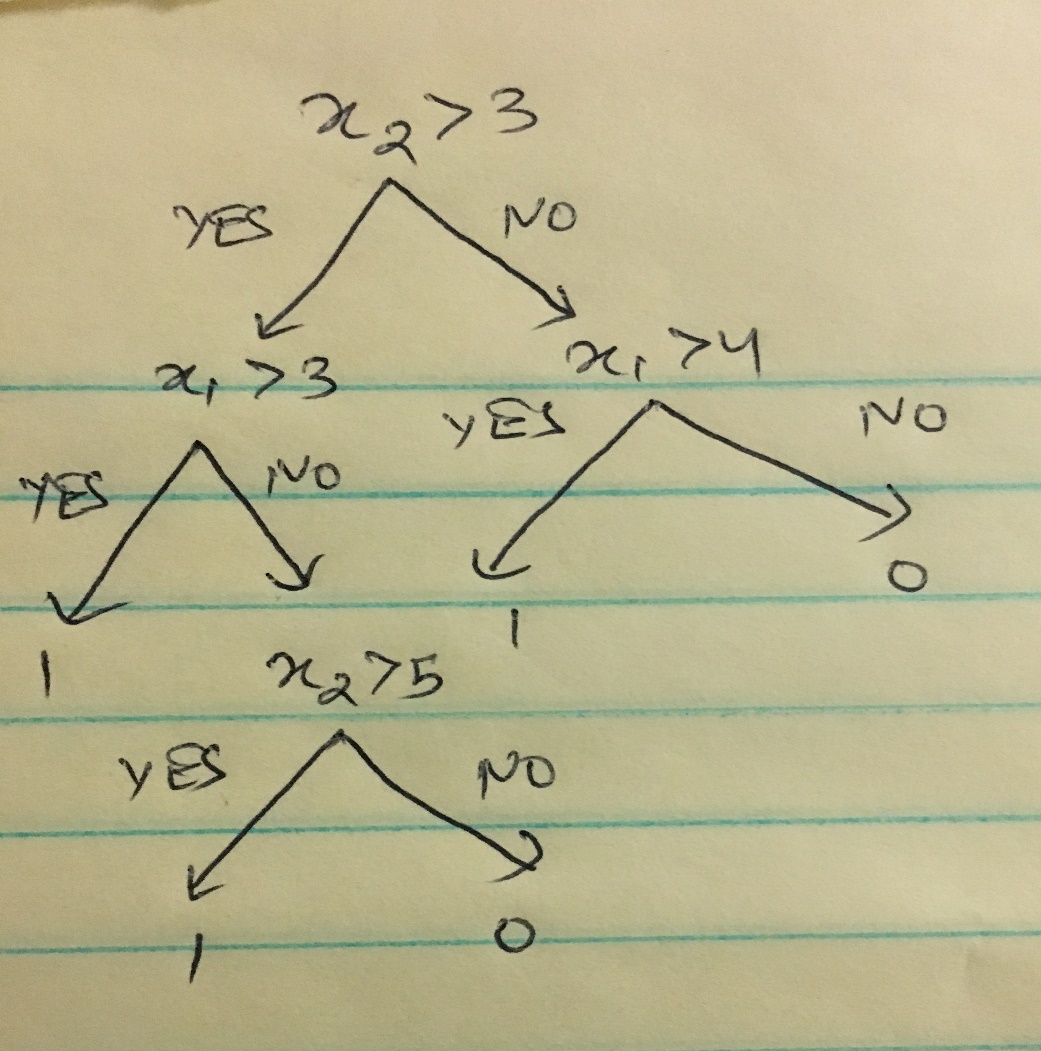
## Homework 2

### Created by:

### Himanshu Goyal (hgoyal)

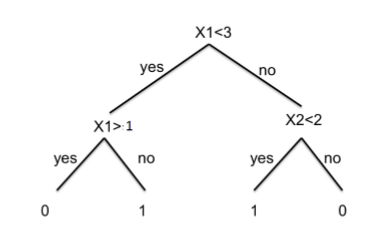
1. *Consider the following decision boundary and draw the equivalent decision-tree*

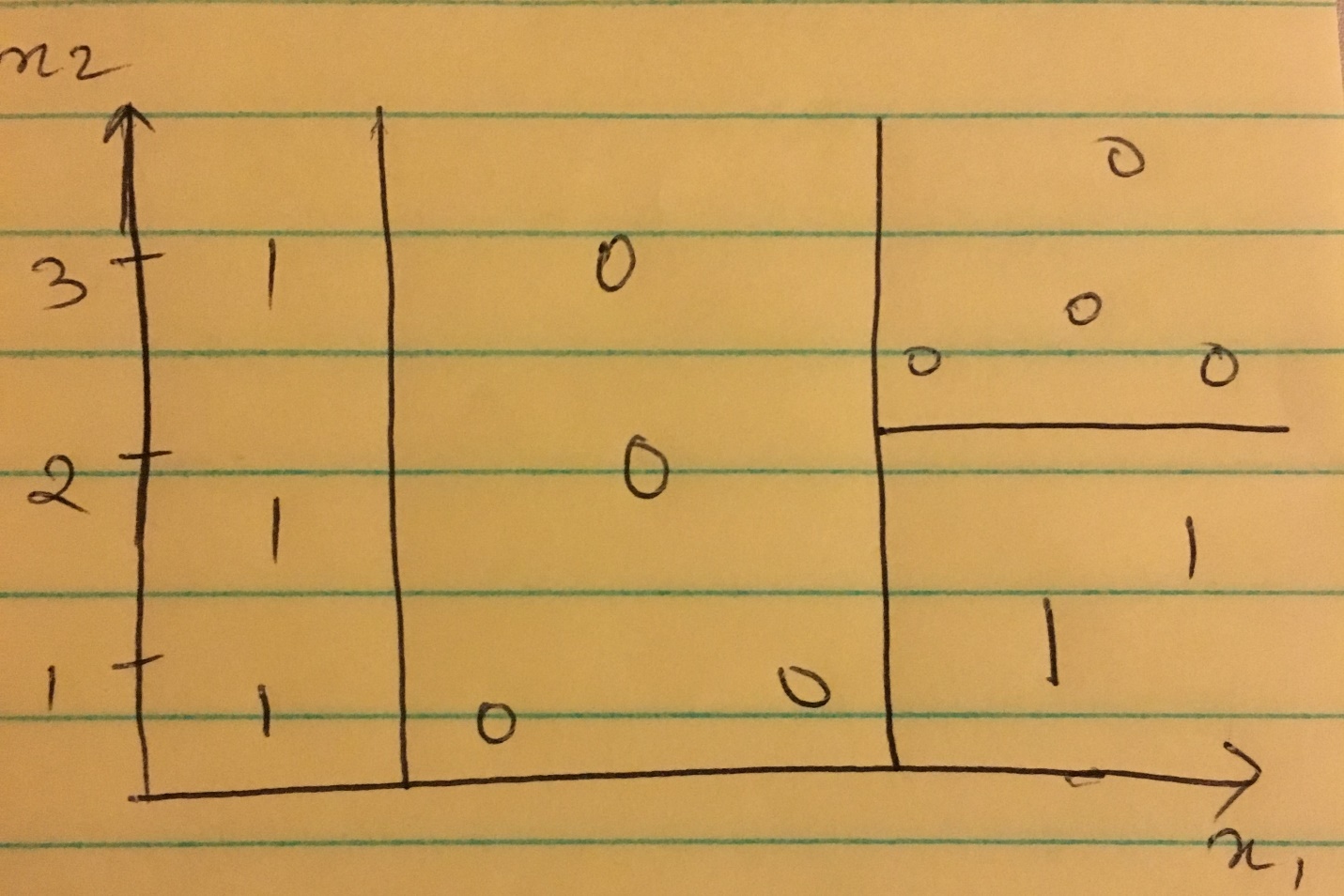
Solution*:*



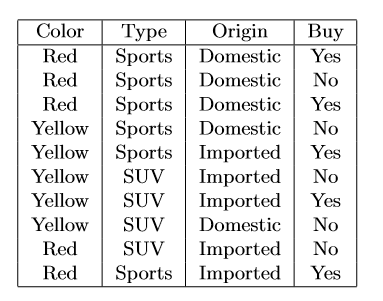
1. Consider the decision-tree presented below. Draw the decision boundary

for this ﬁgure.

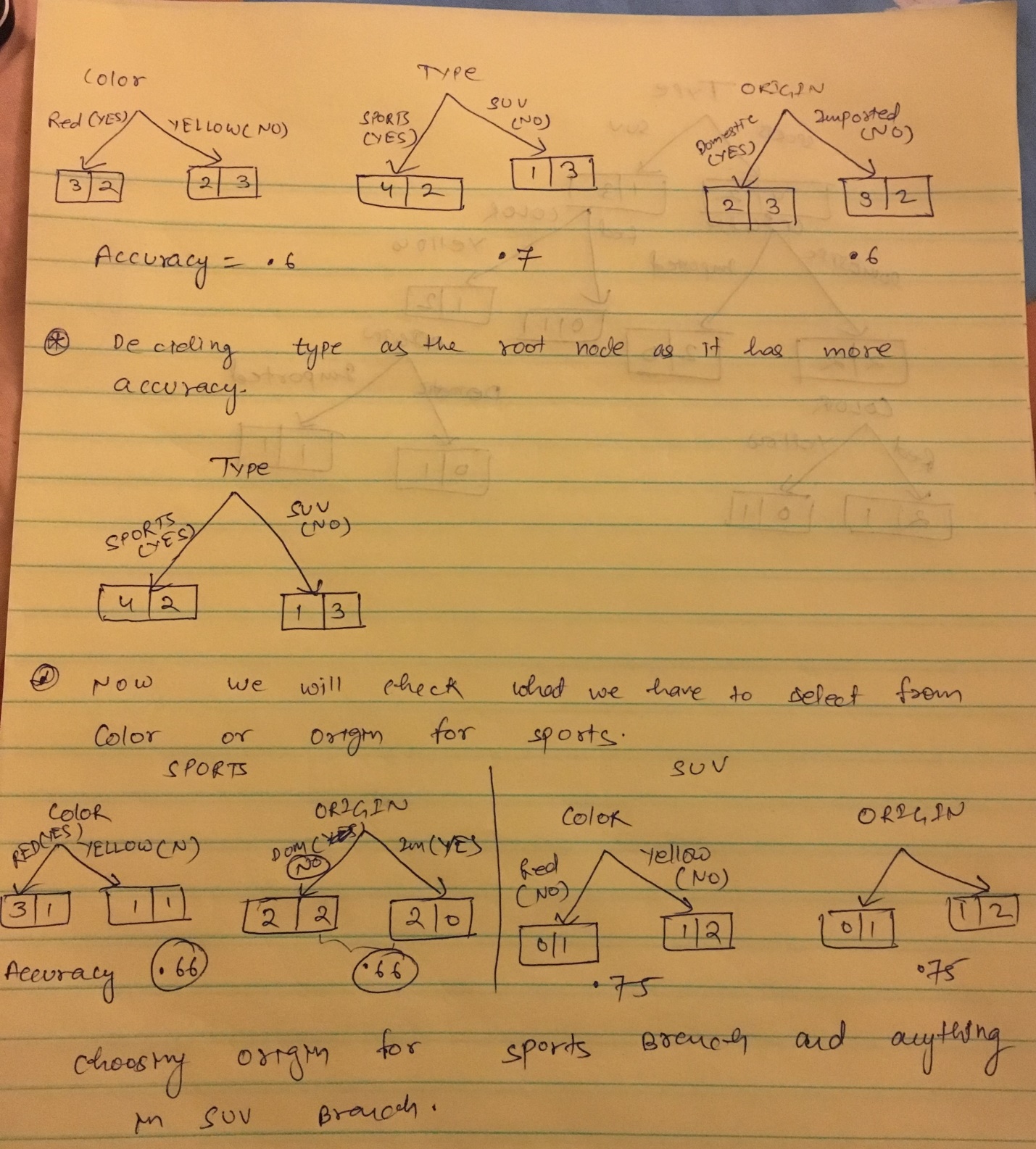
Solution:

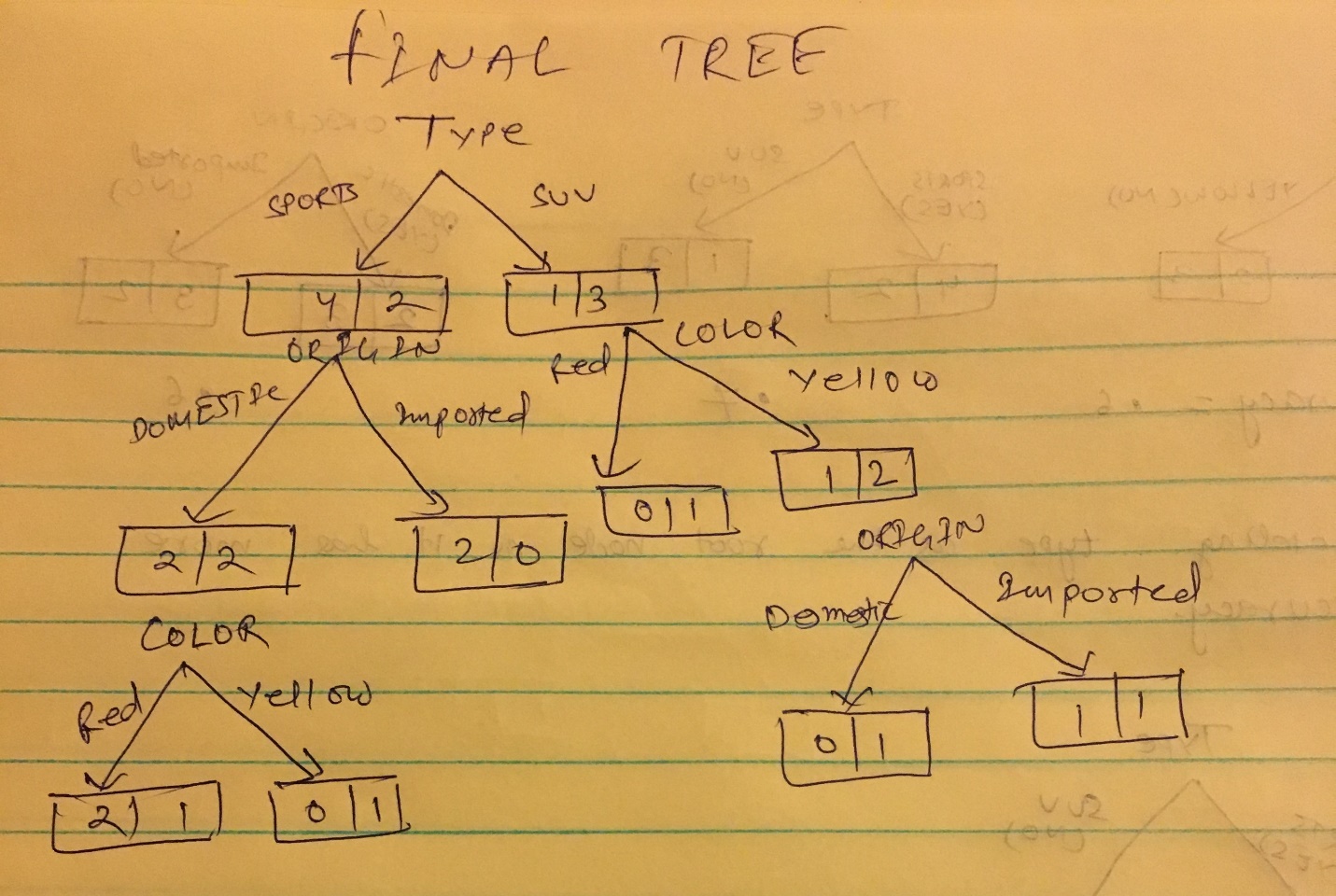
2

1. Consider the data set presented in Table 1. Let the problem be to predict if a car is going to be bought. Draw the decision tree that will be learned if you use accuracy as the splitting criterion. Analyze the overﬁtting nature of this tree.



Solution:





Summary: I am choosing type as root node as it is having 70% as accuracy. I am choosing origin as it provides accuracy of 66% in sports branch. In SUV branch, I can choose anything as the accuracy for both (Color or origin) is coming as 75%. I am choosing color in SUV branch as that provide more results.

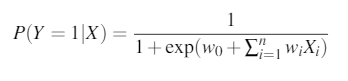
Overfitting: Accuracy of nodes reduces as the depth is increasing. If we allow to grow the tree more, we are going to classify everything correctly. Which will result into overfitting. We should stop increasing the tree as the accuracy is increasing after level 1. As we see most of the leaf nodes are having only one value, this is going to overfit the data.

1. Prove that Logistic Regression learns a linear model.

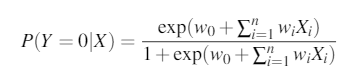
Solution:

If the expected Loss of predicting 0 given that class was 1 is more than loss of predicting 1 then the probability of predicting 1 is greater than that of predicting 0. Taking ratio of probability of predicting 1 and probability of predicting 0 then would be greater than 1. Taking natural log would yield that ratio is greater than 0. If we apply GNB or parameterized form and expand the formula, we would see that logistic regression learns a linear model. Proof based on parameterized form can be seen below.

Logistic regression assumes a parametric model which can be written as follows.

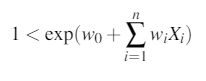
 eq: A

Now we can write P(Y=0|x) as follows after subtracting the above value from 1.

 eq: B

Now, to classify for any givne x we need to maximize P(Y = yk|X), or we can say that we the label Y = 0 if the following condition holds:



Replacing the values from eq A and B this become as follows. 

If we take the log of above condition, we can classify the linear classification rule as follows which is directly dependent on X.

