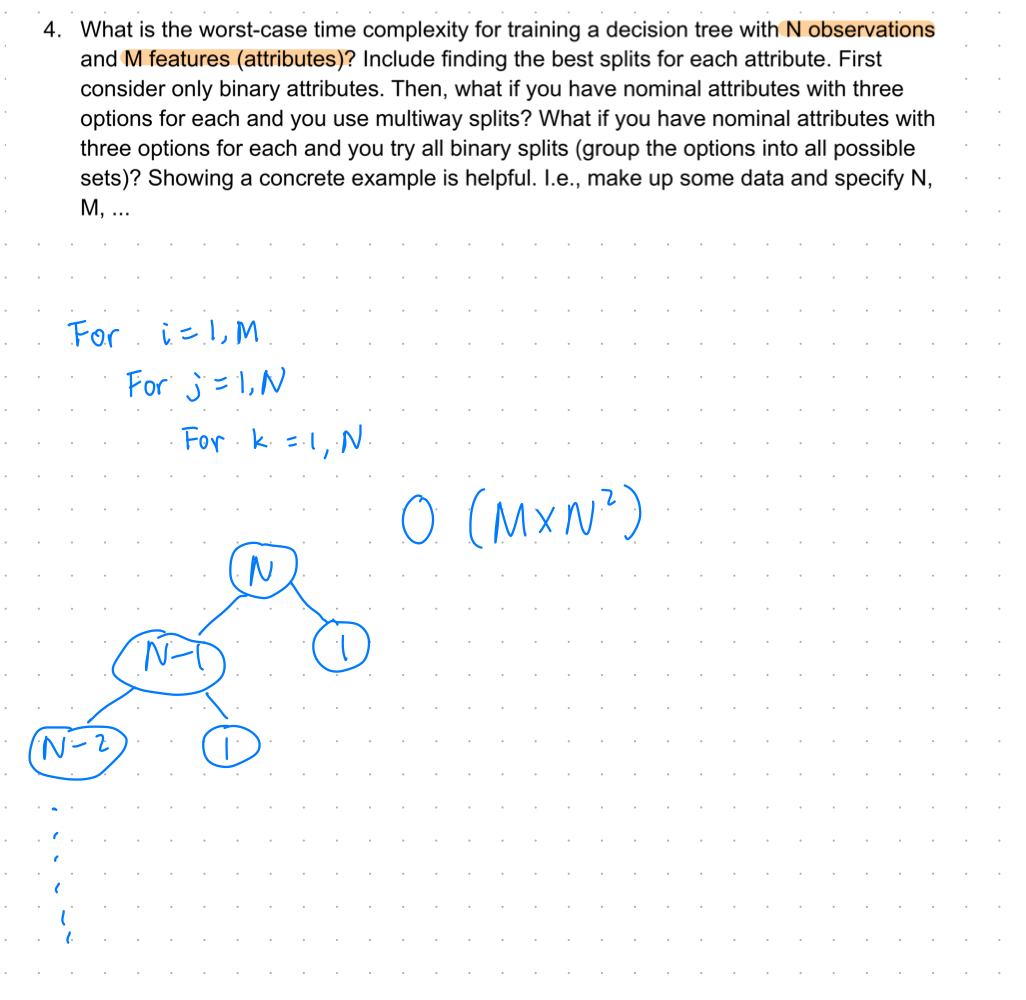
Chapter 3:

**Theory**

****

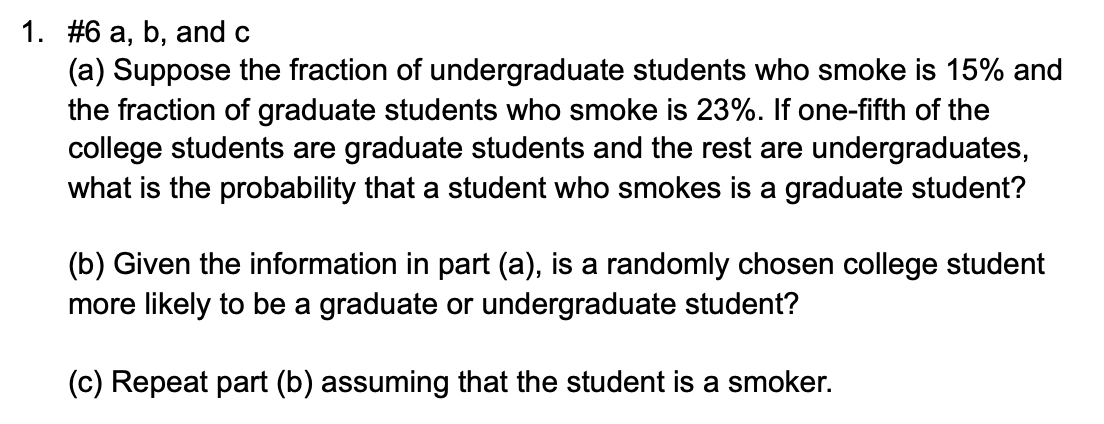
Update: For binary split, we would have M choices for picking the first splitting criteria. After picking the splitting features, N splitting thresholds need to be considered. For the worst-case scenario, we need to repeat this process N times since the tree can be split very unevenly. Therefore, the Big-O will be O(M\*N\*N).

If our features are nominal and assume binary split, the splitting thresholds are easier to find. Thus, the big-O would be O(M\*(number of categories)\*N).

5. Discuss normalizing numeric features for building a decision tree. Is this necessary? Does it affect the decision tree that is built?

I think normalization would not have an impact on the decision tree split. That is because scaling will not affect the distribution of points so the number of data points below and above the splitting threshold will stay the same. **Therefore, I think scaling is not necessary for building a decision tree.**

Chapter 4:



My answer:

P(smoke|undergrad) = 0.15

P(smoke|grad) = 0.23

P(undergrad) = 0.8

P(grad) = 0.2

(a)

P(grad|smake) = (0.23\*0.2)/(0.15\*0.8 + 0.23\*0.2) = 0.28

(b)

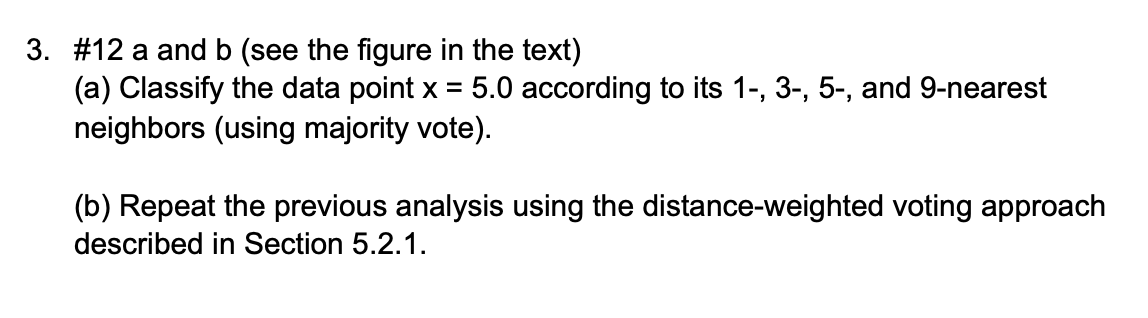
Undergraduate student since P(undergrad) > P(grad)

(c)

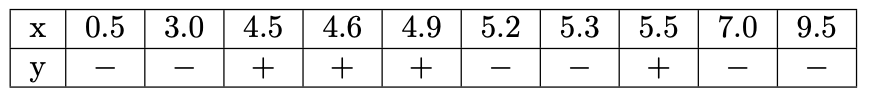
P(grad|smake) = (0.23\*0.2)/(0.15\*0.8 + 0.23\*0.2) = 0.28

P(undergrad|smake) = (0.15\*0.8)/(0.15\*0.8 + 0.23\*0.2) = 0.72

Since P(undergrad|smake) > P(grad|smake), that person is more likely to be an undergraduate student.



My answer:



(a)

+

-

+

-

(a) Rewrite:

When K=1, the closest point to X=5 is X=4.9. The class of X=4.9 is positive so we believe the predicted class of X=5 is also positive.

When K=3, the closest points to X=5 are X=4.9,5.2,5.3. We take the majority vote so we believe the predicted class of X=5 is negative.

When K=5, the closest points to X=5 are X=4.6,4.9,5.2,5.3,5.5. We take the majority vote so we believe the predicted class of X=5 is positive.

When K=9, the closest points to X=5 are X=3.0...9.5. We take the majority vote so we believe the predicted class of X=5 is negative.

(b) Rewrite:

When K=1, the closest point to X=5 is X=4.9. The class of X=4.9 is positive so we believe the predicted class of X=5 is also positive. (the same as unweighted KNN)

When K=3, the closest points to X=5 are X=4.9,5.2,5.3. We take the majority vote and weighted them so we believe the predicted class of X=5 is positive.

(1/0.1\*1)+(1/0.2\*-1)+(1/0.3\*-1)>0, so positive

When K=5, the closest points to X=5 are X=4.6,4.9,5.2,5.3,5.5. We take the majority vote and weighted them so we believe the predicted class of X=5 is positive.

(1/0.4\*1)+(1/0.1\*1)+(1/0.2\*-1)+(1/0.3\*-1)+(1/0.5\*1)>0, so positive

When K=9, the closest points to X=5 are X=3.0...9.5. We take the majority vote and weighted them so we believe the predicted class of X=5 is positive.