Learning Reflection Week 4

● Graded

Student

Piyush Acharya

Total Points

2 / 2 pts

Question 1

Learning Reflection

2 / 2 pts

✔ **+ 2 pts** Learning reflection contains all required sections

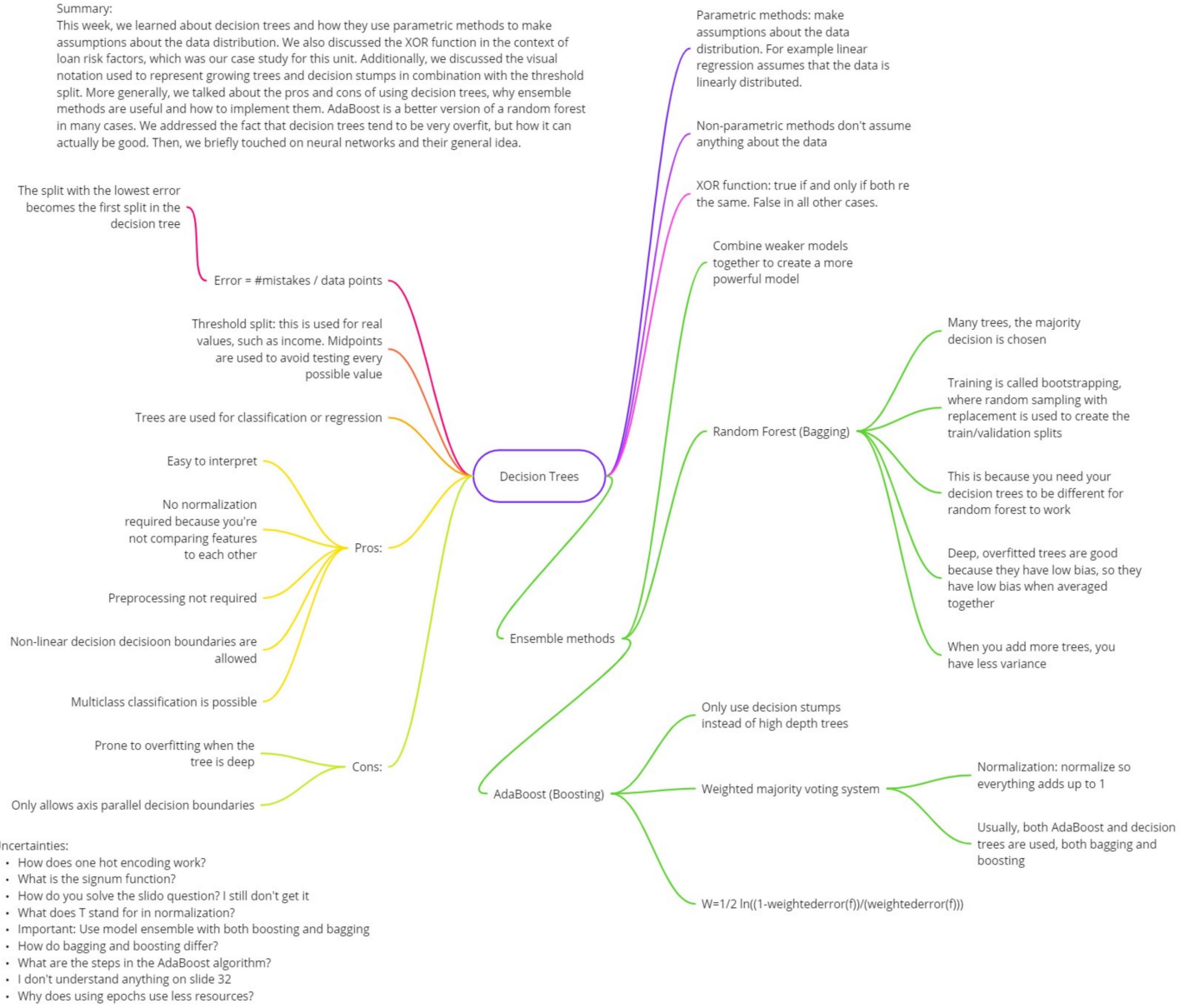
+ 1 pt Either learning reflection is missing one requested section **or** submitted late

+ 0.5 pts Learning reflection is missing one requested section **and** submitted late

+ 0 pts Learning reflection not submitted or missing substantial portions

Question assigned to the following page: 1

Summary:
This week, we learned about decision trees and how they use parametric methods to make assumptions about the data distribution. We also discussed the XOR function in the context of loan risk factors, which was our case study for this unit. Additionally, we discussed the visual notation used to represent growing trees and decision stumps in combination with the threshold split. More generally, we talked about the pros and cons of using decision trees, why ensemble methods are useful and how to implement them. AdaBoost is a better version of a random forest in many cases. We addressed the fact that decision trees tend to be very overfit, but how it can actually be good. Then, we briefly touched on neural networks and their general idea.



Uncertainties:

- How does one hot encoding work?
- What is the signum function?
- How do you solve the slido question? I still don't get it
- What does T stand for in normalization?
- Important: Use model ensemble with both boosting and bagging
- How do bagging and boosting differ?
- What are the steps in the AdaBoost algorithm?
- I don't understand anything on slide 32
- Why does using epochs use less resources?