Decision Tree Review Session

First, you work on these questions by yourself.
Second, you discuss these questions with your peers in a group (2-3).
Third, reflect and revise your answers.

1. **Decision Tree Fundamentals**
    1. What are the key components of a decision tree? Describe what each internal node, branch, and leaf represents.
    2. What are the advantages and disadvantages of decision trees compared to linear regression models?
    3. Explain how the entropy of a dataset reflects uncertainty. Given a dataset where 6 samples are positive and 2 are negative, compute the entropy of the dataset.
    4. How does information gain (IG) guide the selection of an attribute during tree construction?
    5. What are the base cases for stopping recursion when building a decision tree?
2. **Building Decision Trees**
    1. Describe the greedy heuristic used in decision tree algorithms.
    2. What is the difference between Information gain and Gini index as splitting criteria?
    3. Using a decision tree to make a prediction on a new example, we basicly traverse the tree from root to leaf. (True of False)
    4. What does the term overfitting mean in the context of decision trees?
    5. How do decision trees handle continuous (real-valued) features differently from discrete features?
3. **Ensemble Methods**
    1. What is an ensemble classifier, and what are its main advantages and disadvantages?
    2. Explain how bagging (bootstrap aggregation) works.
    3. Why is bagging particularly effective with decision trees but not with k-nearest neighbors?
    4. Describe how random forests combine multiple trees to make a final prediction.