Decision Tree

Geometric Intuition

- classification
  - Decision Tree can be trained to a pure classification with 100% accuracy
  - Decision Trees are hyperparameters and can be tuned to achieve better results and avoid overfitting, similar to other models
- regressionusing splitting
  - Decision Trees can also be used for Regression problems with the same approach
  - Regression Trees are hyperparameters
- Terminology
  - Decision Trees can also be used for Regression problems with the same approach
  - Regression Trees are hyperparameters
- Pros
  - Table data
  - Decision Tree in R plot
  - Decision Trees
  - Simple data
  - Decision Trees
  - Categorical features
  - How to decide which column should be considered for root node
- Disadvantages
  - How to select optimal decision nodes
  - How to decide stopping criteria of tree or pruning
  - Hierarchical primitives to understand complex decision boundaries
- Hyperparameters
  - Hyperparameters of the decision tree in Sklearn
- Multi-dim and continuous features
- Feature importance
  - Information Gain
  - Gini Impurity
- Entropy
  - Entropy is a measure for indeterminate behaviour
  - The measure of disorder
  - Entropy
  - Information gain entropy = sum of entropies
  - Entropy = -p1log(p1) - p2log(p2)
  - Example 1
  - Example 1 uncertainty = 1 (maximum)
  - Very Knowledgeable = low entropy
  - How to calculate Entropy
  - Example 1
  - Entropy
  - Entropy = -1/6log(1/6) - 5/6log(5/6)
  - Entropy = 0.65
  - Example 2
  - Entropy
  - Entropy = 0
  - Example 3
  - Entropy
  - Entropy = 1
  - Entropy for continuous variables
  - Entropy for continuous variables
  - For entropy, I let entropy because here the data is not pure and there is some uncertainty. I have set entropy stop criterion as 0.4
  - Information Gain
  - Information gain = 1 - Entropy
  - Entropy for Probability
  - Information Gain
  - Gini Impurity used to train decision tree
  - Gini Impurity is a measure of how often a randomly chosen element would be incorrectly labeled
  - The information gain is based on the decrease in entropy after a dataset is split on an attribute
  - Constructing a decision tree is all about finding attribute that returns the highest information gain
  - Information Gain
  - Information Gain
  - Formula
  - Example
  - Data set
  - Outlook
  - Temp
  - Humidity
  - Windy
  - Play
  - Calculate the entropy of column
  - Calculate the entropy of column temp
  - Calculate the entropy of column humidity
  - Calculate the entropy of column windy
  - Calculate the information gain
  - Calculate the information gain
  - Calculate the information gain
  - Calculate the information gain
  - Calculate the information gain
  - Calculate the information gain
  - Entropy calculating the child entropies
- Gini Impurity
  - Gini Impurity is the probability of a randomly chosen element being incorrectly labeled if it was randomly labeled according to the distribution of labels in the subset
  - Gini Impurity
  - Gini Impurity = 1 - (p1^2 + p2^2)
  - Gini Impurity = 1 - (0.36 + 0.64)
  - Gini Impurity = 0.48
  - Gini Impurity for Probability
- Entropy vs Gini Impurity
  - Entropy vs Gini Impurity
  - Gini Impurity is used to train decision trees
  - Gini Impurity is a measure of how often a randomly chosen element would be incorrectly labeled
  - The information gain is based on the decrease in entropy after a dataset is split on an attribute
  - Constructing a decision tree is all about finding attribute that returns the highest information gain
  - Information Gain
  - Information Gain
  - Formula
  - Example
  - Data set
  - Outlook
  - Temp
  - Humidity
  - Windy
  - Play
  - Calculate the entropy of column
  - Calculate the entropy of column temp
  - Calculate the entropy of column humidity
  - Calculate the entropy of column windy
  - Calculate the information gain
  - Calculate the information gain
  - Calculate the information gain
  - Calculate the information gain
  - Calculate the information gain
  - Calculate the information gain
  - Entropy calculating the child entropies
- Information Gain
  - Information gain
  - Information gain
  - Information