Introduction to Decision Trees

Alice Gao

Lecture 18-2

Based on work by K. Leyton-Brown, K. Larson, and P. van Beek

Outline

Learning Goals

By the end of the lecture, you should be able to

- Describe the components of a decision tree.
- Construct a decision tree given an order of testing the features.
- Determine the prediction accuracy of a decision tree on a test set.
- Describe/trace/implement the ID3 algorithm.

What is a decision tree

- Book a flight
- Handle late night cravings
- Decide how to spend Friday evening

Jeeves the valet - training set

Day	Outlook	Temp	Humidity	Wind	Tennis?
1	Sunny	Hot	High	Weak	No
2	Sunny	Hot	High	Strong	No
3	Overcast	Hot	High	Weak	Yes
4	Rain	Mild	High	Weak	Yes
5	Rain	Cool	Normal	Weak	Yes
6	Rain	Cool	Normal	Strong	No
7	Overcast	Cool	Normal	Strong	Yes
8	Sunny	Mild	High	Weak	No
9	Sunny	Cool	Normal	Weak	Yes
10	Rain	Mild	Normal	Weak	Yes
11	Sunny	Mild	Normal	Strong	Yes
12	Overcast	Mild	High	Strong	Yes
13	Overcast	Hot	Normal	Weak	Yes
14	Rain	Mild	High	Strong	No

Jeeves the valet - the test set

Day	Outlook	Temp	Humidity	Wind	Tennis?
1	Sunny	Mild	High	Strong	No
2	Rain	Hot	Normal	Strong	No
3	Rain	Cool	High	Strong	No
4	Overcast	Hot	High	Strong	Yes
5	Overcast	Cool	Normal	Weak	Yes
6	Rain	Hot	High	Weak	Yes
7	Overcast	Mild	Normal	Weak	Yes
8	Overcast	Cool	High	Weak	Yes
9	Rain	Cool	High	Weak	Yes
10	Rain	Mild	Normal	Strong	No
11	Overcast	Mild	High	Weak	Yes
12	Sunny	Mild	Normal	Weak	Yes
13	Sunny	Cool	High	Strong	No
14	Sunny	Cool	High	Weak	No

What is a decision tree

A decision tree performs a sequence of tests in the input features.

- Each node performs a test on one input feature.
- Each arc is labeled with a value of the feature.
- Each leaf node specifies an output value.

How to construct a decision tree

- At each node, choose a remaining feature.
- Create arcs, one for each value of the feature.
- Each child is a subtree with all the remaining examples such that the feature takes the value on the arc.

Example: Decision tree

Construct a decision tree using the following order of testing features:

- Test Outlook first.
- For Outlook=Sunny, test Humidity.
- For Outlook=Rain, test Wind.

Example: Decision tree

Construct a decision tree using the following order of testing features:

- Test Temp first.
- For Temp = Cool or Temp = Mild, test Outlook.
- For Temp = Hot, test Wind.

Observations

- The order of testing features affects the size of the decision tree.
- Many decision trees are consistent with the training examples.

Which decision tree would you prefer?

CQ: Test set accuracy

CQ: Which decision tree has better accuracy on predicting the decisions in the test set?

(A) Decision tree A (the simpler one)

(B) Decision tree B (the more complicated one)

(C) They have the same accuracy.

How many functions can we encode with decision trees?

Each decision tree encodes a propositional formula.

If we have n features, how many different functions can we encode with decisions trees?

Constructing the "best" decision tree

We want a decision tree to be

- Consistent with all the training examples and
- As small (shallow) as possible.

Unfortunately, it is intractable to find the smallest consistent decision tree.
Thus, we will use heuristics to find a small consistent tree.

The decision tree learning algorithm

- A greedy divide-and-conquer approach.
- Test the most important feature at each step.
- Solve the sub-problems recursively.

The ID3 algorithm

Algorithm 1 ID3 Algorithm (Features, Examples)

1: If all examples are positive, return a leaf node with decision yes.
2: If all examples are negative, return a leaf node with decision no.
3: If no features left, return a leaf node with the majority decision of the examples.
4: If no examples left, return a leaf node with the majority decision of the examples in the parent.
5: else
6: choose the most important feature f
7: **for** each value v of feature f **do**
8: add arc with label v
9: add subtree $ID3(F - f, s \in S | f(s) = v)$
10: **end for**

Base cases of the ID3 algorithm

No features left:

No examples left:

Revisiting the Learning Goals

By the end of the lecture, you should be able to

- Describe the components of a decision tree.
- Construct a decision tree given an order of testing the features.
- Determine the prediction accuracy of a decision tree on a test set.
- Describe/trace/implement the ID3 algorithm.