

```
1 package edu.caltech.cs2.lab03;
2 import java.util.*;
3
4 public class DecisionTree {
5     private final DecisionTreeNode root;
6
7     public DecisionTree(DecisionTreeNode root) {
8         this.root = root;
9     }
10
11     public String predict(Dataset.Datapoint point) {
12         return this.predictHelper(point, this.root);
13     }
14
15     private String predictHelper(Dataset.Datapoint point, DecisionTreeNode curr) {
16         // Node is an outcome node
17         if (curr.isLeaf()) {
18             return ((OutcomeNode) curr).outcome;
19         }
20
21         // Node is an attribute node
22         String treeAttribute = ((AttributeNode) curr).attribute;
23         String pointFeature = point.attributes.get(treeAttribute);
24         curr = ((AttributeNode) curr).children.get(pointFeature);
25
26         return predictHelper(point, curr);
27     }
28
29     public static DecisionTree id3(Dataset dataset, List<String> attributes) {
30         return new DecisionTree(id3helper(dataset, attributes));
31     }
32
33     private static DecisionTreeNode id3helper(Dataset data, List<String> attributes) {
34
35         // If all remaining data points have same outcome,
36         // return OutcomeNode with this outcome
37         String outcome = data.pointsHaveSameOutcome();
38         if (outcome.length() > 0) {
39             return new OutcomeNode(outcome);
40         }
41
42         // If no attributes left, return OutcomeNode with most common outcome
43         if (attributes.size() == 0) {
44             String mostCommonOutcome = data.getMostCommonOutcome();
45             return new OutcomeNode(mostCommonOutcome);
46         }
47
48         // Find the lowest entropy attribute, a.
49         String a = data.getAttributeWithMinEntropy(attributes);
50
51
52         // Create new list with all attributes, except a
53         List<String> newAttributes = new ArrayList<>(attributes);
```

```

54     newAttributes.remove(a);
55
56     // For each feature, f, of a:
57     List<String> lowestEntropyFeatures = data.getFeaturesForAttribute(a);
58     Map<String, DecisionTreeNode> attributeToNode = new HashMap<>();
59     for (String f : lowestEntropyFeatures) {
60
61         // Find all the data points that have the feature f
62         Dataset newDSet = data.getPointsWithFeature(f);
63
64         // If there are none
65         if (newDSet.isEmpty()) {
66
67             // Guess! Make an outcome child with the most common outcome.
68             String mostCommonOutcome = newDSet.getMostCommonOutcome();
69             OutcomeNode newNode = new OutcomeNode(mostCommonOutcome);
70             attributeToNode.put(f, newNode);
71
72             // Otherwise:
73             } else {
74
75                 // Use the data! Recursively make a child out of the remaining data points,
76                 // using all attributes except for a.
77                 attributeToNode.put(f, id3helper(newDSet, newAttributes));
78             }
79         }
80
81         // Return the attribute node with the children generated above.
82         return new AttributeNode(a, attributeToNode);
83     }
84 }

```