10-301/601: Introduction to Machine Learning Decision Tree Pseudocode

Henry Chai & Matt Gormley

9/8/22

Decision Tree: Pseudocode

```
def train(𝒟):
    store root = tree_recurse(𝒟)
def tree_recurse(𝒟′):
    q = new node()
    base case – if (SOME CONDITION):
    recursion – else:
        q.type = internal
        find best attribute to split on, x_d
        q.split = x_d
        for v in V(x_d), all possible values of x_d:
            𝒟_v = {(x^(n), y^(n)) ∈ 𝒟 | x_d^(n) = v}
            q.children(v) = tree_recurse(𝒟_v)
    return q
```

Decision Tree: Pseudocode

```
def train(𝒟):
    store root = tree_recurse(𝒟)
def tree_recurse(𝒟′):
    q = new node()
    base case – if (𝒟′ is empty OR
        all labels in 𝒟′ are the same OR
        all features in 𝒟′ are identical OR
        some other stopping criterion):
            q.type = leaf
            q.label = majority_vote(labels in 𝒟′)

    recursion – else:
    return q
```