

Lab 1 Machine Learning

Anton Stråhle & Jan Alexandersson

25 januari 2020

Assignment 0

Assignment 1

Dataset	Entropy
MONK-1	1.0
MONK-2	0.9571174283
MONK-3	0.9998061328

Assignment 2

Assignment 3

Dataset	A1	A2	A3	A4	A5	A6
MONK-1	0.07527	0.00584	0.00471	0.02631	0.28703	0.00076
MONK-2	0.00376	0.00246	0.00106	0.01566	0.01728	0.00625
MONK-3	0.00712	0.29374	0.00083	0.00289	0.25591	0.00708

Assignment 4

Assignment 5

Run the code below for MONK-1 and MONK-3 and their respective testdata to obtain a sample which we can compute the mean and variance on.

```
def prune(data, test):  
    pruned_trees_fraction = []  
    for frac in fraction:  
        train, val = partition(data, frac)
```

```

t = d.buildTree(train, m.attributes)
all_pruned = d.allPruned(t)

#I suppose t is included in all_pruned
best_tree_perf = 0
for t in all_pruned:
    candidate_perf = d.check(t, val)
    if best_tree_perf < candidate_perf:
        best_tree_perf = candidate_perf
        best_tree = t

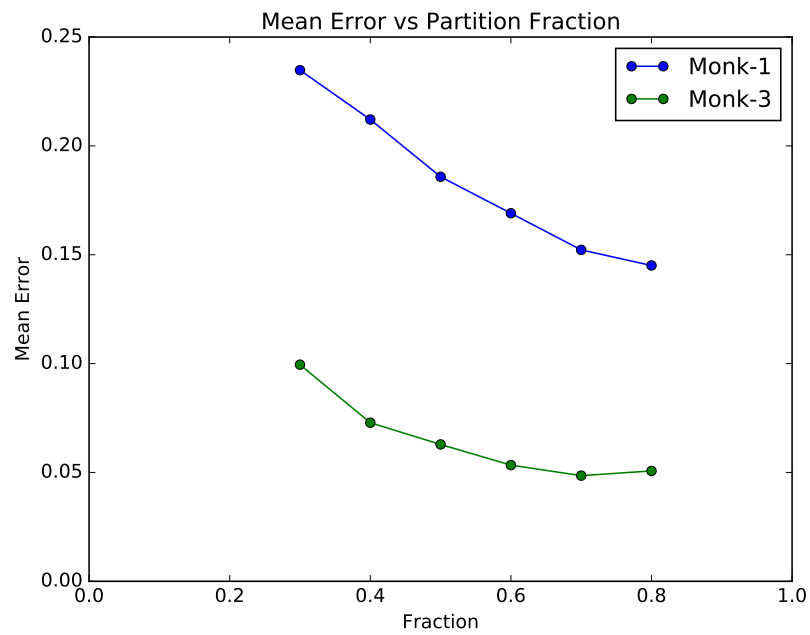
pruned_trees_fraction.append(1-d.check(best_tree, test))
return pruned_trees_fraction

```

Dataset	Error(Train)	Error(Test)
MONK-1	0.0	0.1713
MONK-2	0.0	0.30787
MONK-3	0.0	0.05556

Assignment 6

Assignment 7



Fraction =	0.3	0.4	0.5	0.6	0.7	0.8
MONK-1 Sd	0.0436	0.0405	0.0409	0.0432	0.0415	0.0383
MONK-3 Sd	0.0578	0.0421	0.0378	0.0322	0.0284	0.0298