

## Homework 4: Decision Trees

AUTHORS: Jed Pulley

---

**DO NOT POLLUTE!** AVOID PRINTING, OR PRINT 2-SIDED MULTIPAGE.

### Problem 4.1

**Passenger Class:** First class is 1 and other classes are 0.

**Sex:** I left sex alone, since it was already binary.

**Age:** I split age based on the median. If you were above the median age, you get 0, otherwise 1.

**Siblings/Spouse:** I chose 1 for no siblings/spouses and 0 for any number of them.

**Parents/Children:** Similarly, I chose 1 for no parents/children and 0 for any

**Fare:** I split fare based on the median, if you are above, you get 1, otherwise 0.

### Problem 4.2

See *decision\_trees.ipynb* under section 4.2 for code

### Problem 4.3

See *decision\_trees.ipynb* under section 4.3 for code

Two variables, *max\_depth* and *min\_samples\_split*, are defined as my stopping conditions. If a node exceeds the max depth, then stop. Otherwise, if a node is about to be split more times than our minimum sample split, then it will stop as well. This prevents the tree from getting too deep or wide.

### Problem 4.4

See *decision\_trees.ipynb* under section 4.4 for code

**Problem 4.5****Problem 4.6****Problem 4.7**

(a)

(b)

(c)

**Problem 4.8**

(a)

(b)

(c)

**Problem 4.9****Problem 4.10**