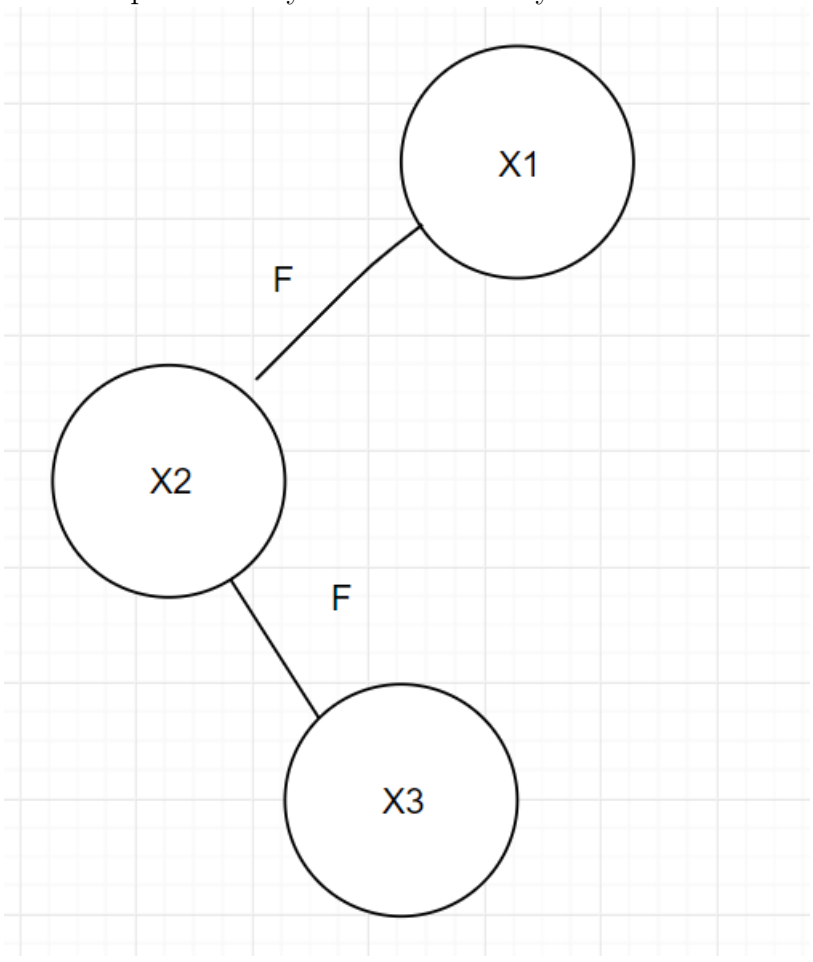
# Machine Learning
# Homework 1

Aaron Templeton

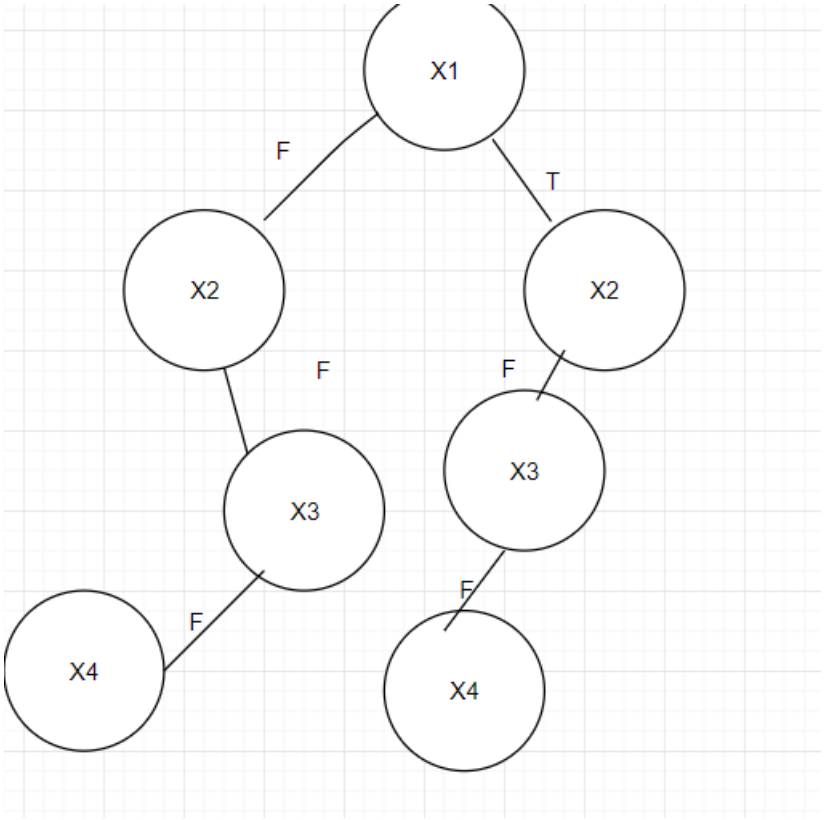
September 11, 2019

## Decision Trees

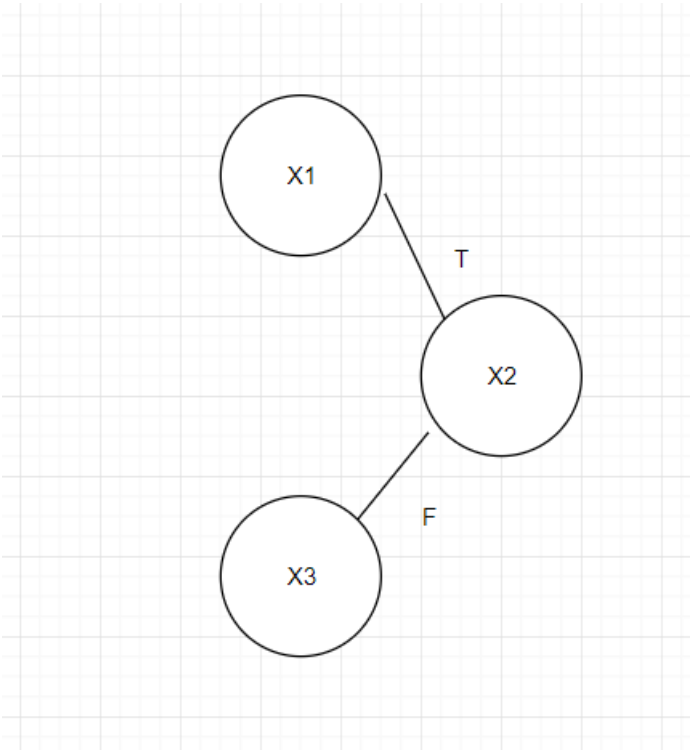
(1) (a) $\neg x_1 \wedge (x_2 \; xor \; x_3)$
can be represented by decision tree only




(b) $(x_1 \; nor \; x_2) \vee (x_3 \wedge (\neg x_4))$
can be represented as decision tree only




(c) $x_1 \wedge \neg x_2 \wedge x_3$
represent as tree and linear classifier




linear classifier: $x_1 - x_2 + x_3 \geq 2$

(2) (a) there are $2^4$ possible functions that map the four features to a boolean decision.

all functions are consistent with dataset

(b) entropy is

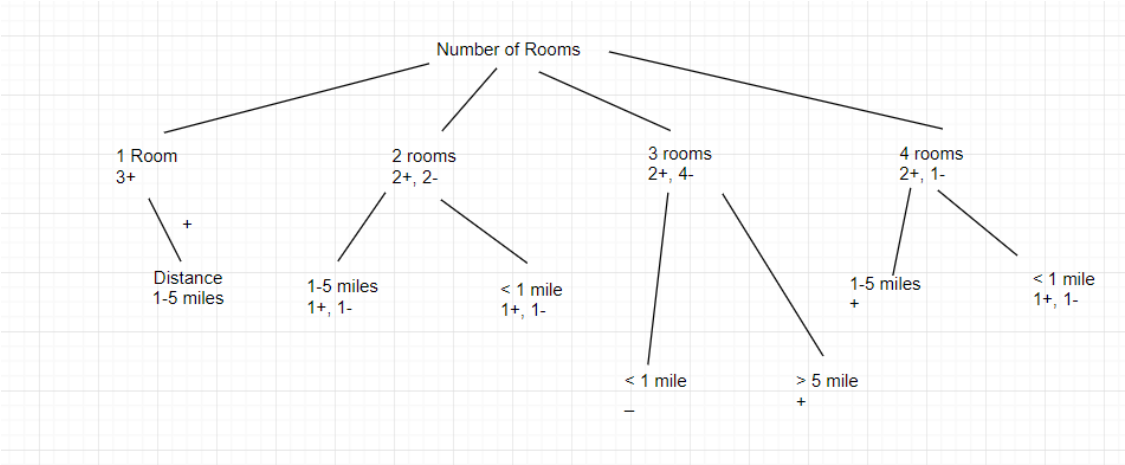
$$-(\frac{9}{16})log_2(\frac{9}{16}) - (\frac{7}{16})log_2(\frac{7}{16}) = 0.987$$

(c) answers in table

| Feature | Information Gain |
|---|---|
| Number of rooms | 0.223 |
| Apartment condition | 0.013 |
| Distance | 0.196 |
| Price | 0.185 |

(d) the feature with the highest information gain which is Number of rooms

(e) tree below




(f) predictions 1-10 below

1. Like
2. Like
3. Like
4. Like
5. Dislike
6. Like
7. Like
8. Like
9. Like
10. Like

it is almost completely accurate. differences on 2 predictions

(3) (a) $1 - ((9/16)^2 + (7/16)^2) = 0.492$

| Feature | Information Gain Gini Index |
|---|---|
| Number of rooms | 0.219 |
| Apartment condition | 0.235 |
| Distance | 0.125 |
| Price | 0.078 |

(b) table below

(c) Root should be Apartment condition. The tree is different

# Experiments

(1) (a) root feature is

(b) information gain is 0.975

(c) max depth = 2

(d) error on training = 0

(e) error on test = 0