

## REPORT 1

Steven Guo

cs 4300

University of Missouri-St. Louis

### Abstract

I chose to work on this project because it would be helpful in deciding on what car to buy in the future.

Car Evaluation Database was derived from a simple hierarchical decision model originally developed for the demonstration of DEX, M. Bohanec, V. Rajkovic: Expert system for decision making. Sistemica 1(1), pp. 145-157, 1990.). The model evaluates cars according to the following concept structure:

- CAR car acceptability
  - . PRICE overall price
    - . . buying buying price
    - . . main price of the maintenance
  - . TECH technical characteristics
    - . . COMFORT comfort
      - . . . doors number of doors
      - . . . persons capacity in terms of persons to carry
      - . . . lug boot the size of luggage boot
    - . . safety estimated safety of the car

Input attributes are printed in lowercase. Besides the target concept (CAR), the model includes three intermediate concepts: PRICE, TECH, COMFORT. Every concept is in the original model related to its lower level descendants by a set of examples (for these examples sets see [Web Link]).

The Car Evaluation Database contains examples with the structural information removed, i.e., directly relates CAR to the six input attributes: buying, main, doors, persons, lug boot, safety.

Because of known underlying concept structure, this database may be particularly useful for testing constructive induction and structure discovery methods.

Attribute Information:

Class Values:

unacc, acc, good, vgood

Attributes:

buying price: vhigh, high, med, low.

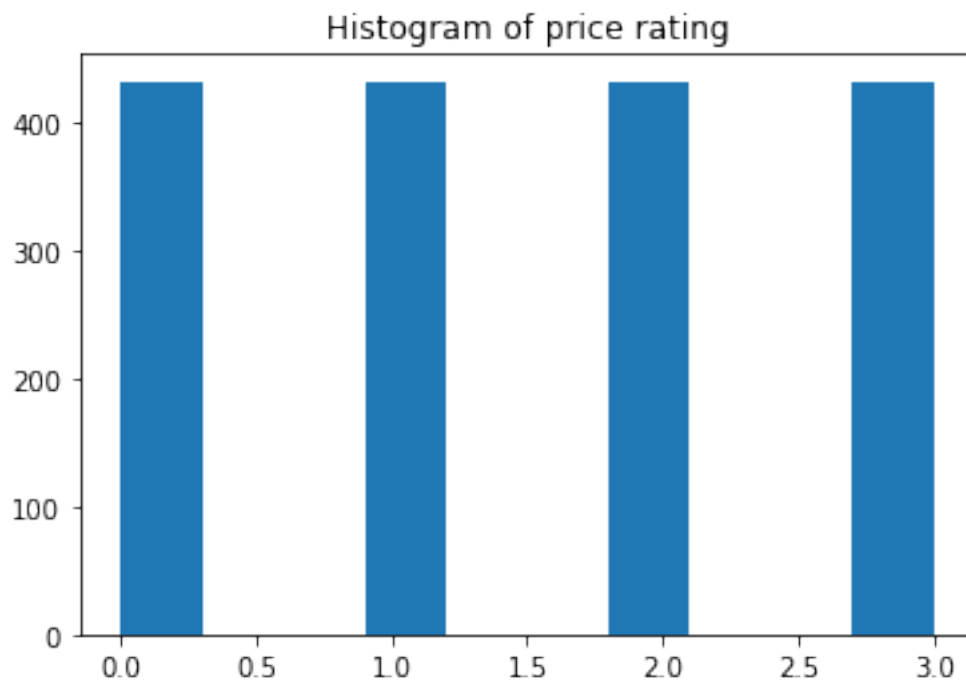
maintenance: vhigh, high, med, low.

doors: 2, 3, 4, 5more.

persons: 2, 4, more.

lug boot: small, med, big.

safety: low, med, high.



categories are: vhigh, high, med, low.

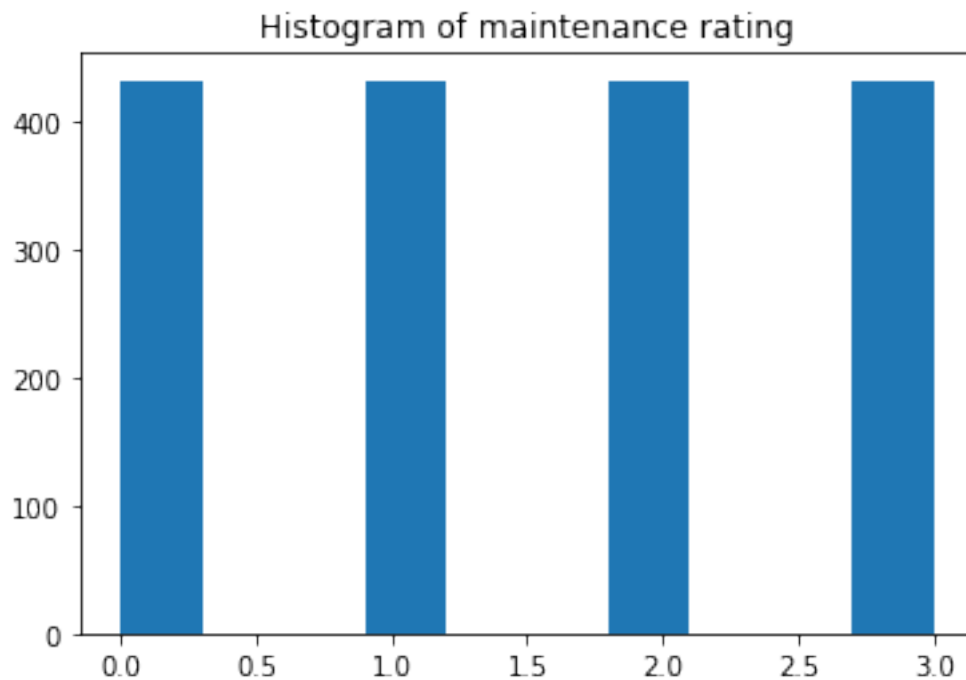
In relation to the graph vhigh = 3, high = 2, med = 1, low = 0.

mean= 1.5

median= 1.5

max= 3

min= 0



categories are: vhigh, high, med, low.

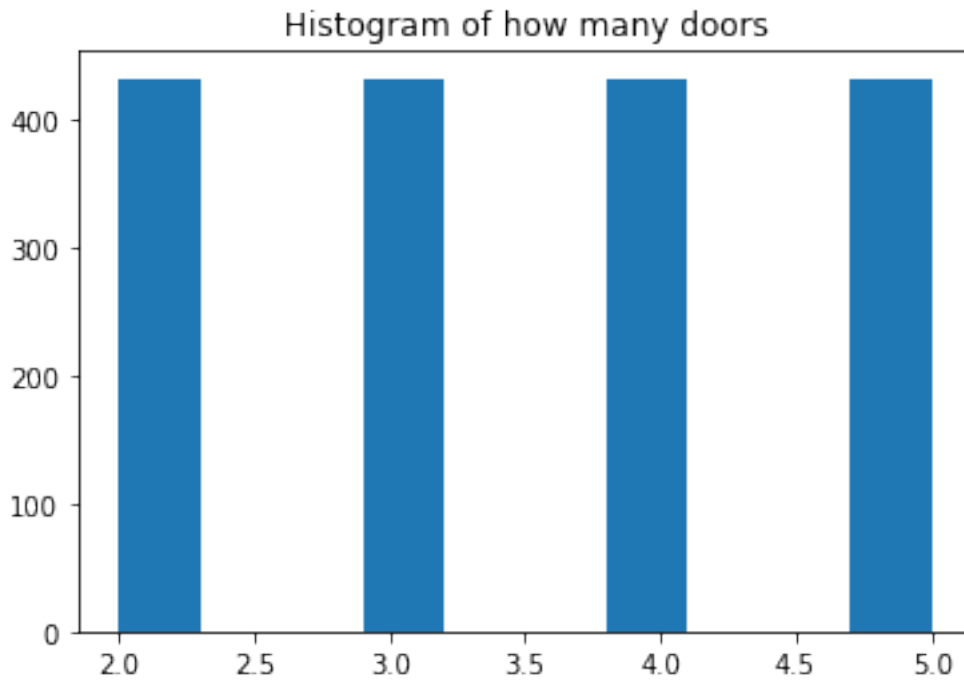
In relation to the graph vhigh = 3, high = 2, med = 1, low = 0.

mean= 1.5

median= 1.5

max= 3

min= 0



categories are: 2, 3, 4, 5more.

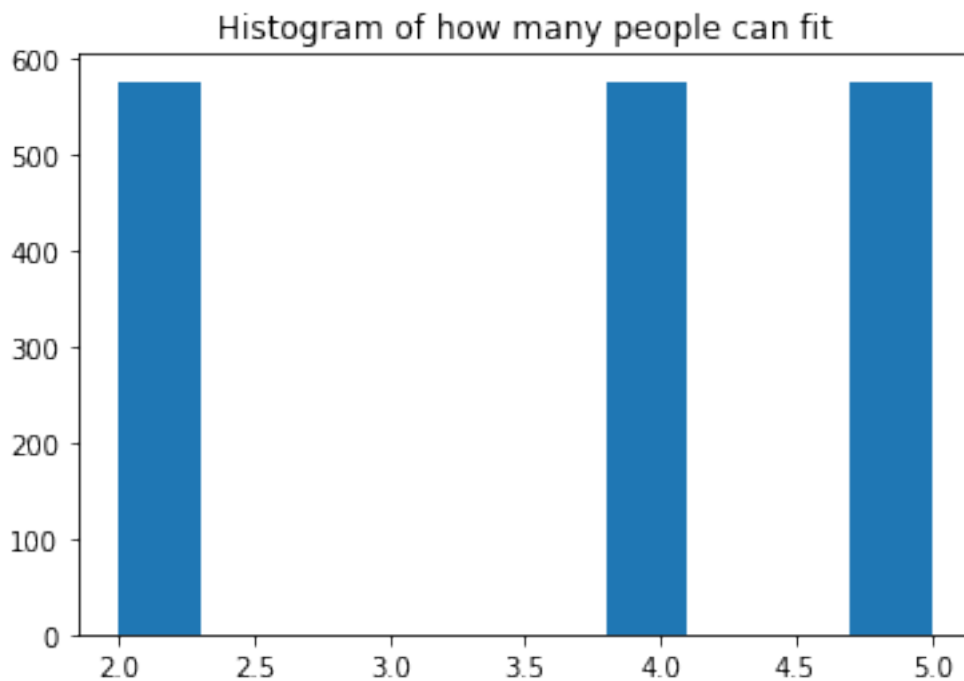
only thing that changed is 5more which is five or more. I just changed it to 5 to represent 5 or more

mean= 3.5

median= 3.5

max= 5

min= 2



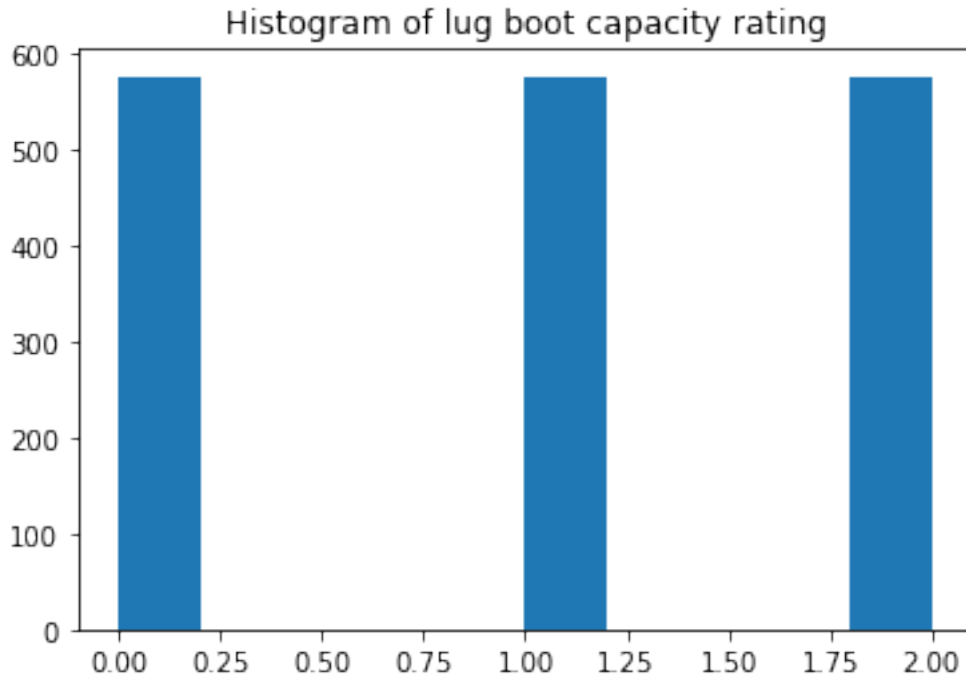
categories are: 2, 4, more.

only thing that changed is more which is more than 4 people. I just changed it to 5 mean= 3.6

median= 4

max= 5

min= 2



categories are: small, med, big.

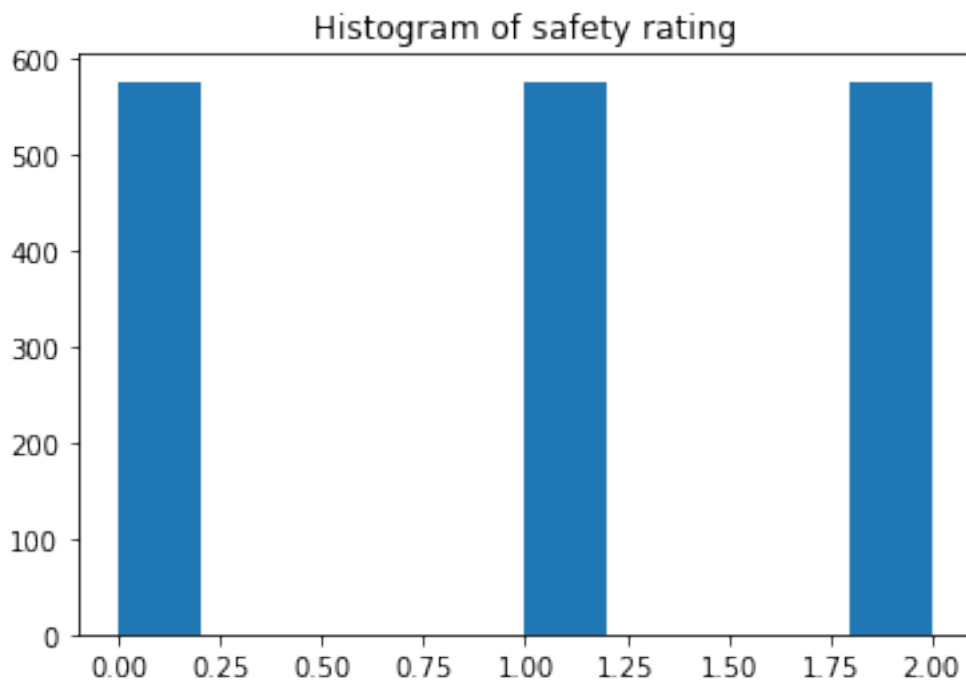
In relation to the graph small = 0, med = 1, big = 2.

mean= 1

median= 1

max= 2

min= 0



categories are: low, med, high.

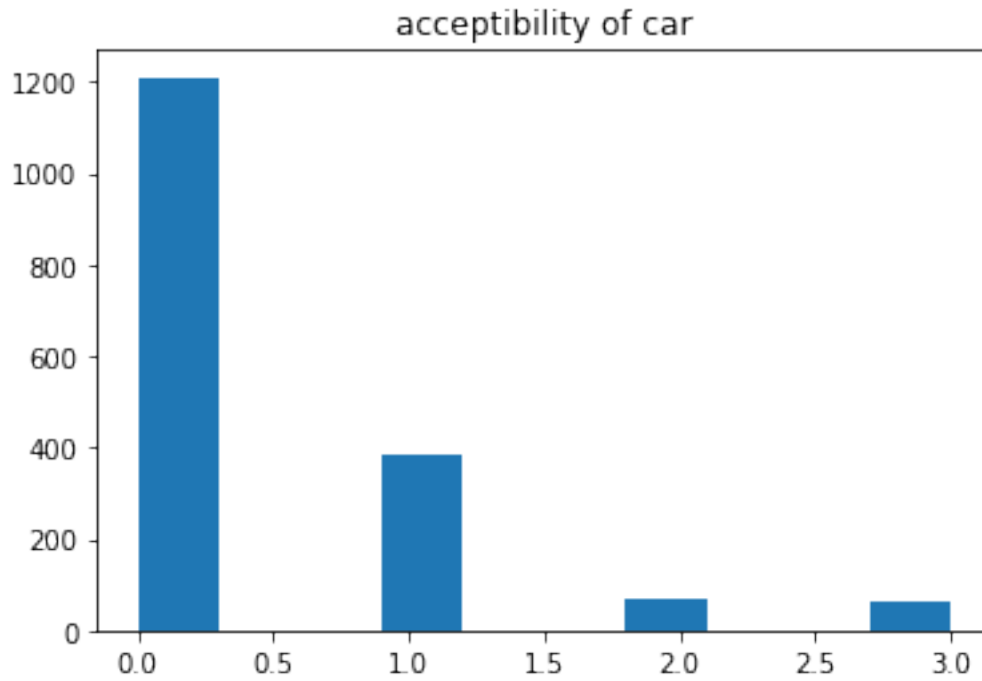
In relation to the graph low = 0, med = 1, high = 2.

mean= 1

median= 1

max= 2

min= 0



categories are: unacc, acc, good, vgood In relation to the graph unacc = 0, acc = 1, good = 2, vgood = 3

mean= 0.41

median= 0

max= 3

min= 0

0- 1210

1- 384

2- 69

3- 65

I would classify the car as acceptable(1) or not acceptable(0), therefore the amount of acceptable cars is 522. the amount of acceptable cars is = 30% and 70% for unacceptable cars.