Mohammad Mojtaba Roshani (*programmer*) note:

--1. See the highlights please.

--2. See Persian explanation on page 2.

This database contains 13 attributes (which have been extracted from

a larger set of 75)

Attribute Information:

------------------------

-- 1. age

-- 2. sex

-- 3. chest pain type (4 values)

-- 4. resting blood pressure

-- 5. serum cholestoral in mg/dl

-- 6. fasting blood sugar > 120 mg/dl

-- 7. resting electrocardiographic results (values 0,1,2)

-- 8. maximum heart rate achieved

-- 9. exercise induced angina

-- 10. oldpeak = ST depression induced by exercise relative to rest

-- 11. the slope of the peak exercise ST segment

-- 12. number of major vessels (0-3) colored by flourosopy

-- 13. thal: 3 = normal; 6 = fixed defect; 7 = reversable defect

Attributes types

-----------------

Real: 1,4,5,8,10,12

Ordered:11,

Binary: 2,6,9 => I used these indexes and filtered plus convert file to csv

Nominal:7,3,13

Variable to be predicted

------------------------

Absence (1) or presence (2) of heart disease

Cost Matrix

abse pres

absence 0 1

presence 5 0

where the rows represent the true values and the columns the predicted.

No missing values.

270 observations

Dataset link provided by <https://archive.ics.uci.edu/> :

<https://archive.ics.uci.edu/ml/datasets/Statlog+%28Heart%29>

GitHub link (code + edited dataset)provided by Mohammad Mojtaba Roshani:

<https://github.com/MohammadMojtabaRoshani-TOMaaR/ID3-with-gini-index>

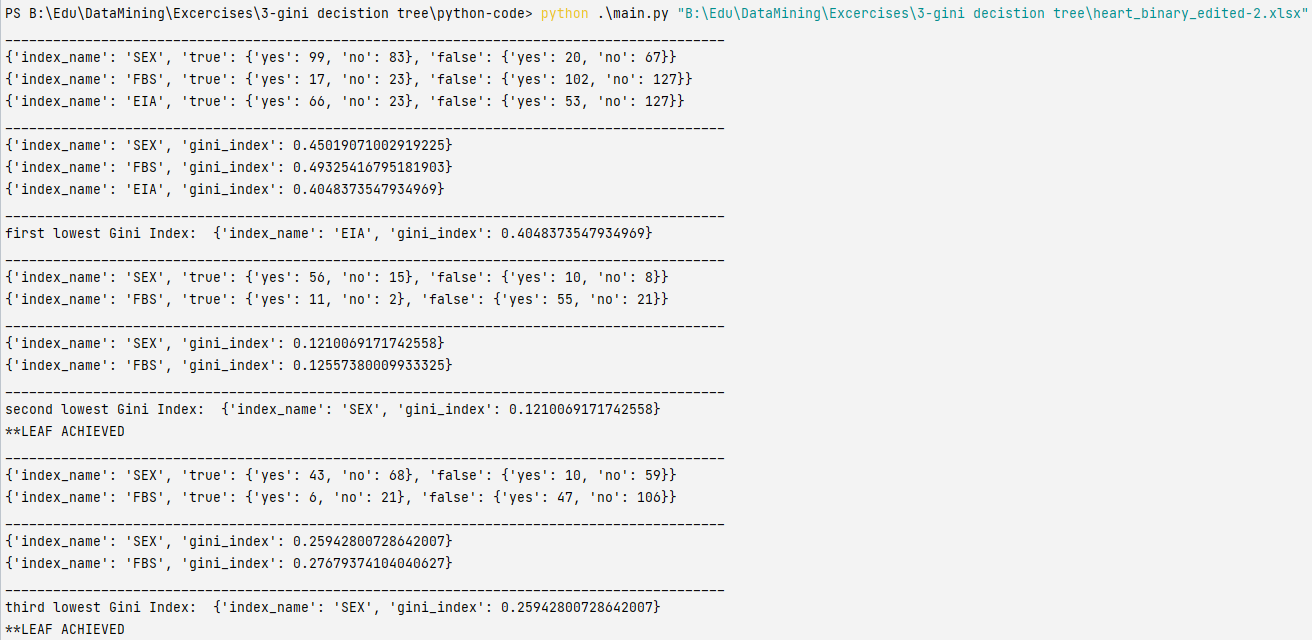
* Essential python library can be found and install in "requirement.txt"

$ pip install -r requirement.txt

* There a \*.xslx dataset in the current directory of program, to run:

python .\main.py "path to \*.xlsx dataset"

note: .dat and .CSV format of dataset are provided

* The numerical output is like this:

* The visual output is like this:

