Assignment 3
Group size: 1 individual max
Deadline: 25th December 2020

Statement

In this assignment you have to implement a clustering technique. Format of the dataset is as follows:

```
214
11

1    1.52101    13.64    4.49    1.10    71.78    0.06    8.75    0.00    0.00    1
2    1.51761    13.89    3.60    1.36    72.73    0.48    7.83    0.00    0.00    1
3    1.51618    13.53    3.55    1.54    72.99    0.39    7.78    0.00    0.00    1
4    1.51766    13.21    3.69    1.29    72.61    0.57    8.22    0.00    0.00    1
5    1.51742    13.27    3.62    1.24    73.08    0.55    8.07    0.00    0.00    1
6    1.51596    12.79    3.61    1.62    72.97    0.64    8.07    0.00    0.26    1
7    1.51743    13.30    3.60    1.14    73.09    0.58    8.17    0.00    0.00    1
8    1.51756    13.15    3.61    1.05    73.24    0.57    8.24    0.00    0.00    1
9    1.51918    14.04    3.58    1.37    72.08    0.56    8.30    0.00    0.00    1
```

- The digit "214" in first row is the number of rows
- Digit "11" in second row is the number of columns
- Third row is an empty one
- Rest is a grid of data.

Input data sets:

Download the following dataset and transform it in the above mentioned format.

- http://archive.ics.uci.edu/ml/datasets/Iris

Write an application, in any programming language, to apply the following tasks to the input datasets.

Tasks to do:

Calculation Correlation Matrix:

- Create a correlation matrix from the data matrix using Pearson's correlation coefficient
- The correlation matrix will be a *NXN* matrix (where *N* is number of records in your input dataset) containing Pearson's correlation coefficient between each of the row in data matrix
- Pearson's correlation coefficient formula:

$$\sum \left(x - \bar{x}\right)\left(y - \bar{y}\right) / \sqrt{\sum \left(x - \bar{x}\right)^2 \sum \left(y - \bar{y}\right)^2}$$

Discretize:

- Calculate median/mean of each column of the correlation matrix and set all the values in that column that are above the calculated median/mean to 1 and rest to 0

Visualize:

- Convert the discretized matrix into bitmap. Sample image follow.