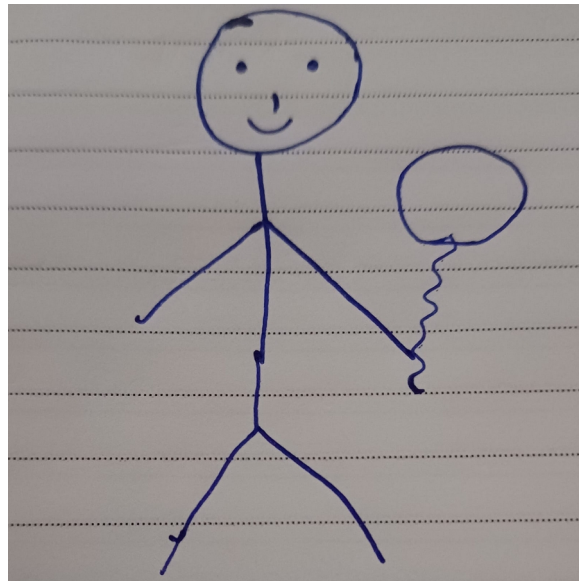


DA5401 - Data Analytics Lab

Assignment – 1

1.1 Data Acquisition

I drew a diagram of a man holding a balloon. The diagram is shown below.

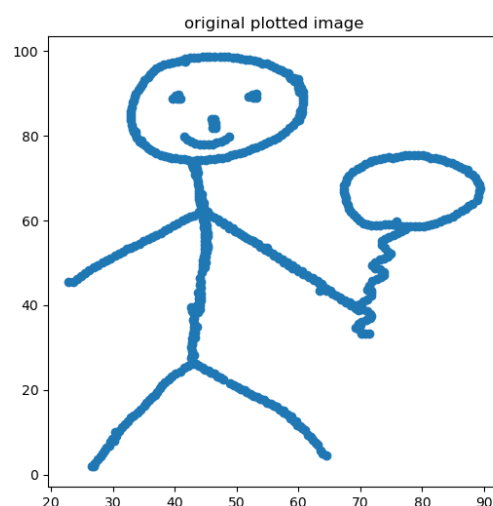


Then, I uploaded this image to <https://wpd.starrydata2.org/> and generated the data points. I chose the x1 and x2 points as 0 and 100, and y1, y2 points as 0,100 generating a 100 x 100 grid.

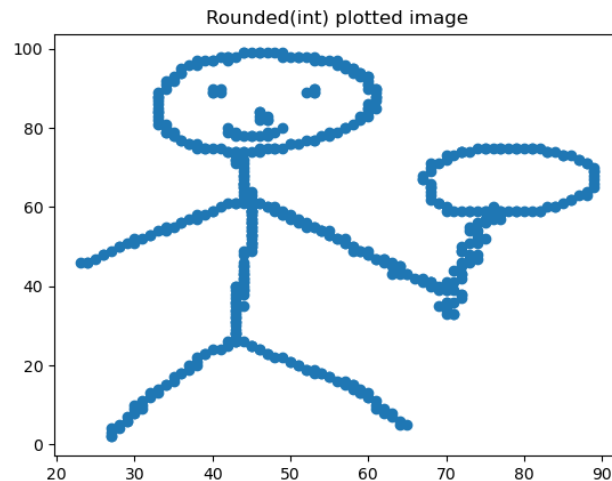
The default operation resulted in points including the ruled lines from the notebook. Then I used the pen tool to carefully highlight the drawing, then downloaded the points as .csv file and stored it as dataset.csv.

1.2 Data Cleansing and Loading

Now, we load the .csv file to a dataframe using pandas library. Then the scatter plot is plotted using matplotlib library, which is shown below.



Here the data points are in the float datatype with values after decimal. Now, I converted them into integer format, so that we can store the values in a 100 x 100, 2 dimensional matrix. The image after converting the float points into integers looks like



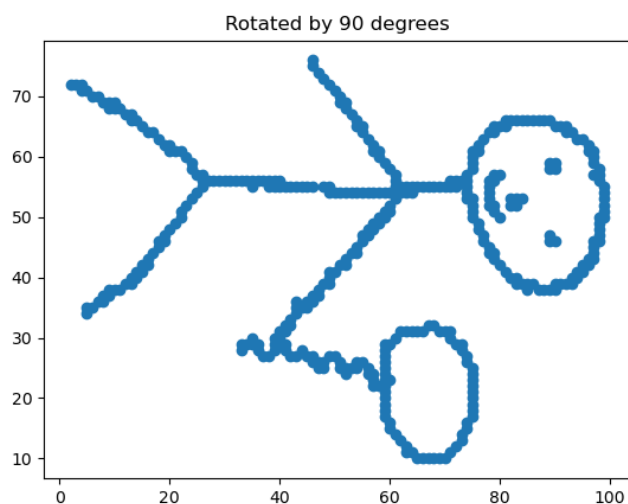
After converting into integers, the value of `matrix[x][y]` is set as 1 for every x,y in our dataframe. The rest of the values remain as 0.

1.3 Transformation

1.3.1 Rotating by 90 degrees

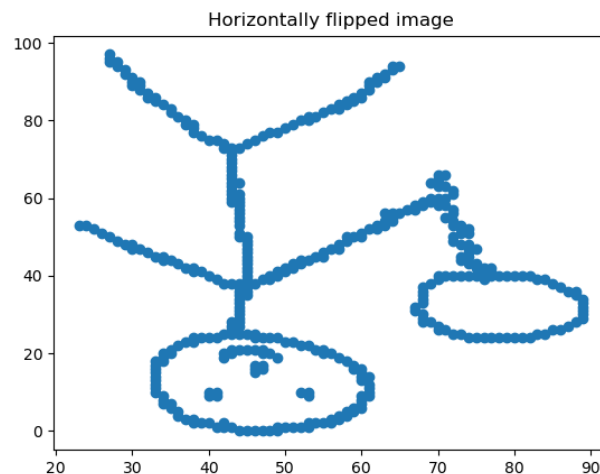
We use the `rot90` from numpy library to rotate the matrix than takes 2 arguments: the name of the matrix and the value for k, -1 representing clockwise direction.

After rotating, we converted back the matrix into dataframe using pandas library and plotted the scatterplot which looks like



1.3.2 Horizontal Flip

To horizontally flip the image, we *flip/r* function from numpy library that takes the matrix as input, and returns the horizontally flipped matrix as output. This matrix is again converted to dataframe and a scatterplot is plotted.



1.4 Visualization

The horizontally flipped version and 90° rotated image is plotted and attached in submission file. Running the main.py file will also display the transformed images that are attached in this document.

References

- <https://numpy.org/doc/>
- <https://pandas.pydata.org/docs/>
- <https://matplotlib.org/stable/index.html>
- <https://www.youtube.com/watch?v=P7GbGdMvopU>