

# MT2013122\_ASSIGNMENT4

## Multivariate data visualization using parallel coordinates of bank-note authentication data

### Data Set Information:-

Data were extracted from images that were taken from genuine and forged banknote-like specimens. For digitization, an industrial camera usually used for print inspection was used. The final images have 400x 400 pixels. Due to the object lens and distance to the investigated object gray-scale pictures with a resolution of about 660 dpi were gained. Wavelet Transform tool were used to extract features from images.

Data set Characteristics : - Multivariate

Number of Instances :- 1372

Attribute characteristics :- Real

Number of Attributes :- 5

### Attribute Information: -

1. variance of Wavelet Transformed image (continuous)
2. skewness of Wavelet Transformed image (continuous)
3. curtosis of Wavelet Transformed image (continuous)
4. entropy of image (continuous)
5. class (integer)

### Visualization Algorithm :-

The given data set is multivariate data set. Therefore, parallel coordinates method is used.

### Parallel coordinates method:-

Parallel coordinates methos is generally used to visualize high dimensional geometry or analyzing multi-variate data set. By using parallel axes for dimensions, the parallel coordinates techniques can represent N-dimensional data in a 2-dimensional space.

### Data Processing: -

Data is mapped from values to length of Axis. So the min\_value has lowest postion on the corresponding axis and max\_value has heghest postion on the corresponding axis.

### Implemenation:-

To show the data set of 5 attributes, therefore a backdrop is drwan consisiting of 5 parallel lines, vertically and equally spaced. A point in the n – dimensional is represented as a polyline with vertices on the parallel axes: te position of the vertex on the its axis corresponds to the ith coordinate of the point. The color of lines are determined the range of value in the first- dimension, according to that it gives color value red, green or blue.

**Insights: -**

1. The data in variance of wavelet transformed image and skewness of wavelet transformed atr not related.
2. The data in skewness and curtosis of wavelet transformed image the lines in between are crossing in X-shape, So theses are negative relationship.
3. The data in curtosis of wavelet Transformed image and entropy of image there is negative correlation.
4. The data lines in between the entropy of image and class, aproximetly half of lines skewed toward class0 and half towards class1.
5. The data in between the entory of image and class are not correlated.
6. The all the instances which have high variance of wavelet transformed image values in the class0  
(i.e. The Blue colored lines.)
7. The most of the instances which have low variance of wavelet transformed image values in the class1.(i.e. The red colored lines.)