CO3

Practical 7

1. Questions related to linear transformation.

**Pre-lab.**

1. Explain what is a random variable?
2. Explain linear transformation of random variable.
3. What is Min-Max normalization.

**In-lab.**

Data-set

The Iris flower data set or Fisher's Iris data set is a multivariate data set introduced by the British statistician, eugenicist, and biologist Ronald Fisher in his 1936. The data set consists of 50 samples from each of three species of Iris (Iris setosa, Iris virginica and Iris versicolor). Four features were measured from each sample: the length and the width of the sepals and petals, in centimetres. Based on the combination of these four features, Fisher developed a linear dis-criminant model to distinguish the species from each other. The dataset is available in following link.

<https://www.kaggle.com/arshid/iris-flower-dataset>

1. Read data from the .CSV file. Get the basic statistics like mean, median, variance and standard deviation of the attributes sepal\_length, sepal\_width, petal\_length, petal\_width of the data set.
2. Find the co-relation between the attributes (sepal\_length, sepal\_width) and (petal\_length, petal\_with)
3. Make a linear transformation of the attribute sepal\_length by adding 1.5 to it. Similarly make a linear transformation of the attribute petal\_length by multiplying two.
4. Now find the co-relation between the attributes (sepal\_length, sepal\_width) and (petal\_length, petal\_with). Then analyze it and draw conclusion.
5. Make a linear transformation of the attribute sepal\_length by subtracting 1.5 from it. Similarly make a linear transformation of the attribute petal\_length by dividing two.
6. Now find the co-relation between the attributes (sepal\_length, sepal\_width) and (petal\_length, petal\_with). Then analyze it and draw conclusion.

**Post-lab.**

DATA - SET

A super market XXXX selling the products such as Health and beauty, accessories, Electronic ,Home and lifestyle, Food and beverages, Sports. You can get the data-set in the following link.

<https://www.kaggle.com/aungpyaeap/supermarket-sales>

1. Find the co-relation between the attributes unit-price and Total. Then transform the attribute ‘Total’ using min-max normalization. Then find the co-relation between unit-price and Total and draw conclusion.