Machine Learning with Python

KMeans Clustering

Project-1 : Iris Flower species classifications

Use iris flower dataset from 'sklearn' library and try to form clusters of flowers. Figure out if any preprocessing such as scaling would help here.

Iris Versicolor **Iris Setosa** **Iris Virginica**

Iris Flower species classifications

Hints :

1. Load the dataset : .

```
In [2]:  1  from sklearn.datasets import load_iris
         2  iris = load_iris(as_frame=True)
         3  type(iris)

Out[2]:  sklearn.utils.Bunch
```

```
In [4]:  1  df=iris.frame
         2  type(df)

Out[4]:  pandas.core.frame.DataFrame
```

2. Use the below pattern to convert the target and predicted value to species name.

```
1  df['species'] = df['target'].replace(to_replace=[0,1,2], value=['Iris-setosa','Iris Versicolour','Iris Virginica'])
```

3. Plot the elbow chart to find the optimum number of cluster.

4. Predict the accuracy by comparing given 'species' and 'predicted species'.

5. Find out the best the features for the better accuracy ?

6. Try 'MinMaxscaler' & Check out .. Is it affecting the accuracy or not?

7. Use the attractive graph to show the species cluster and centroid points.

Project-2 : Indian Stock Market

Task:

- Firstly, screen the companies and visualise them according to sub-sectors using a pie chart or a bar charts.

- Secondly, visualize the companies by segregating them according to Market Cap in three categories: Large Cap(>20,000 crore), Mid Cap (5,000 to 20,000 crore) & Small Cap(<5,000).

- Thirdly, pick 10 random companies from the entire group and visualize the following through a line chart Find the Intrinsic Value of the company based on 3 cases of growth (g): Assume 3 Cases for g (Growth) : Good (15% Growth) ; Bad (-5% Growth); Best (25%Growth)

- Visualise these 10 companies on a line chart for all 3 cases of growth going forward against its current market cap.

Indian Stock Market

💡 V = (EPS * (8.5 + 2g) * 6) / Y

V   : Intrinsic Value

EPS : The Company's last 12 month earnings per share

8.5   : The constant represents the appropriate P-E ratio for a no-growth
       company as proposed by Graham.

g   : The company's long-term (five years) earnings growth estimate

6   : The average Return of FDs (6%)

Y   : The current yield on AAA corporate bonds.