

ML4 Clustering Analysis

October 27, 2023

```
[1]: import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
from sklearn.cluster import KMeans
import warnings
warnings.filterwarnings('ignore')

data = pd.read_csv("C:/Users/hp/Downloads/Practical_Data/Iris.csv")
data.head()
```

```
[1]:
```

| | Id | SepalLengthCm | SepalWidthCm | PetalLengthCm | PetalWidthCm | Species |
|---|----|---------------|--------------|---------------|--------------|-------------|
| 0 | 1 | 5.1 | 3.5 | 1.4 | 0.2 | Iris-setosa |
| 1 | 2 | 4.9 | 3.0 | 1.4 | 0.2 | Iris-setosa |
| 2 | 3 | 4.7 | 3.2 | 1.3 | 0.2 | Iris-setosa |
| 3 | 4 | 4.6 | 3.1 | 1.5 | 0.2 | Iris-setosa |
| 4 | 5 | 5.0 | 3.6 | 1.4 | 0.2 | Iris-setosa |

```
[2]: pip install --upgrade threadpoolctl
```

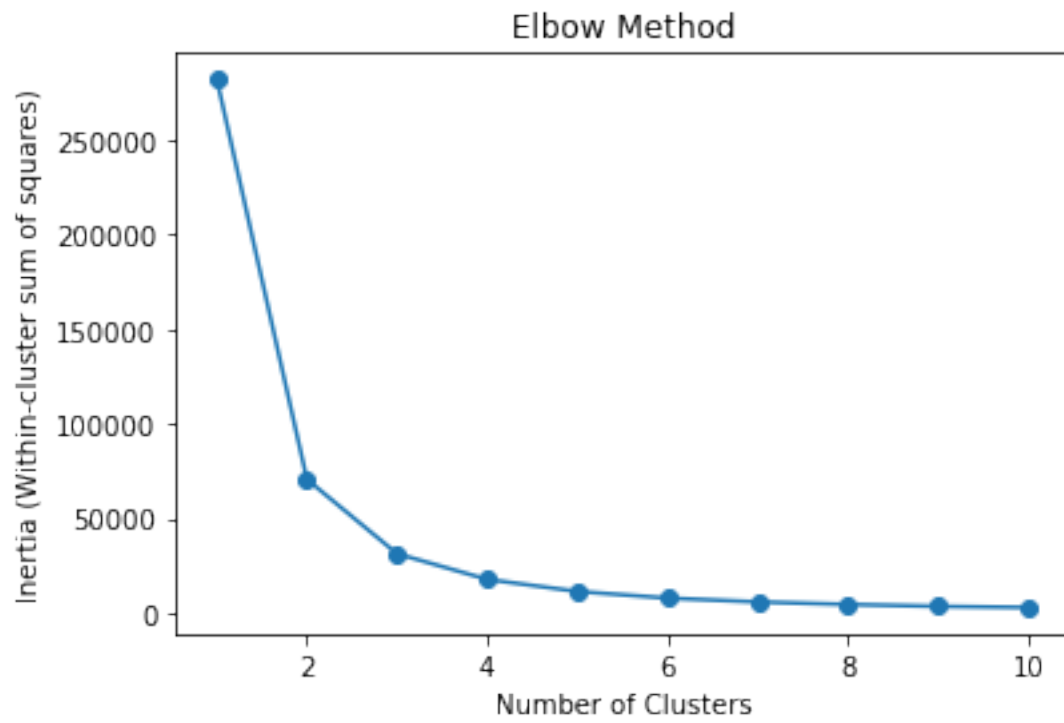
Requirement already satisfied: threadpoolctl in c:\users\hp\anaconda3\lib\site-packages (3.2.0)

Note: you may need to restart the kernel to use updated packages.

```
[3]: X = data.iloc[:, [0, 1, 2, 3]]
```

```
[4]: inertia = []
for i in range(1, 11):
    kmeans = KMeans(n_clusters=i, max_iter=300, random_state=42)
    kmeans.fit(X)
    inertia.append(kmeans.inertia_)
```

```
[5]: # Plot the Elbow Method graph
plt.plot(range(1, 11), inertia, marker='o')
plt.xlabel('Number of Clusters')
plt.ylabel('Inertia (Within-cluster sum of squares)')
plt.title('Elbow Method')
plt.show()
```



[]: