

```
In [1]: import pandas as pd
import matplotlib.pyplot as plt
import seaborn as sns
from sklearn.cluster import AgglomerativeClustering
from sklearn.metrics import silhouette_score
```

```
In [2]: df=pd.read_csv("Iris.csv")
df.head()
```

```
Out[2]:
```

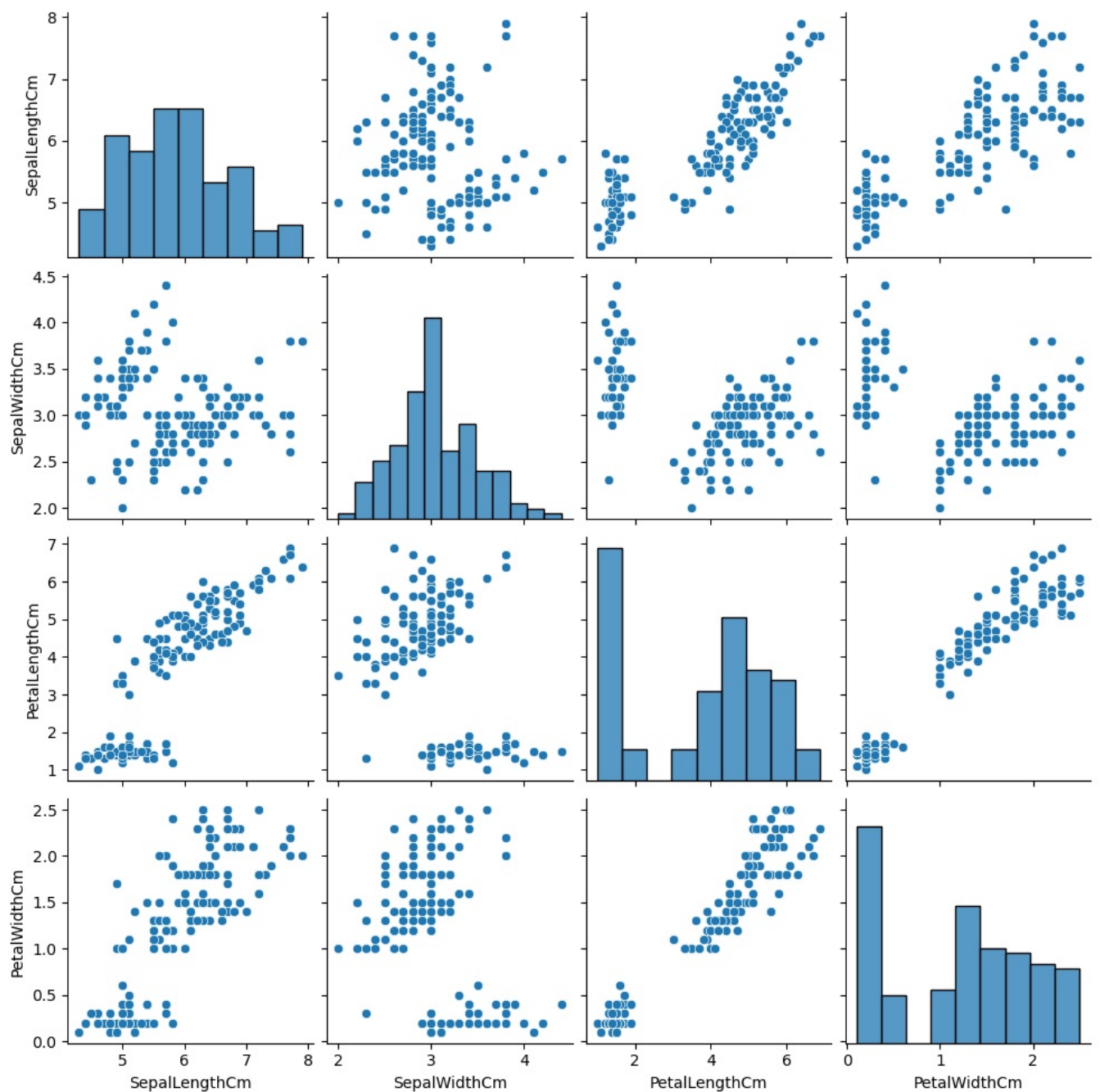
| | SepalLengthCm | SepalWidthCm | PetalLengthCm | PetalWidthCm |
|---|---------------|--------------|---------------|--------------|
| 0 | 5.1 | 3.5 | 1.4 | 0.2 |
| 1 | 4.9 | 3.0 | 1.4 | 0.2 |
| 2 | 4.7 | 3.2 | 1.3 | 0.2 |
| 3 | 4.6 | 3.1 | 1.5 | 0.2 |
| 4 | 5.0 | 3.6 | 1.4 | 0.2 |

```
In [3]: df.isnull().sum()
```

```
Out[3]: SepalLengthCm    0
SepalWidthCm          0
PetalLengthCm         0
PetalWidthCm          0
dtype: int64
```

```
In [4]: sns.pairplot(data=df)
plt.show()
```

```
C:\ProgramData\anaconda3\Lib\site-packages\seaborn\axisgrid.py:118: UserWarning: The figure layout has changed to tight
self._figure.tight_layout(*args, **kwargs)
```

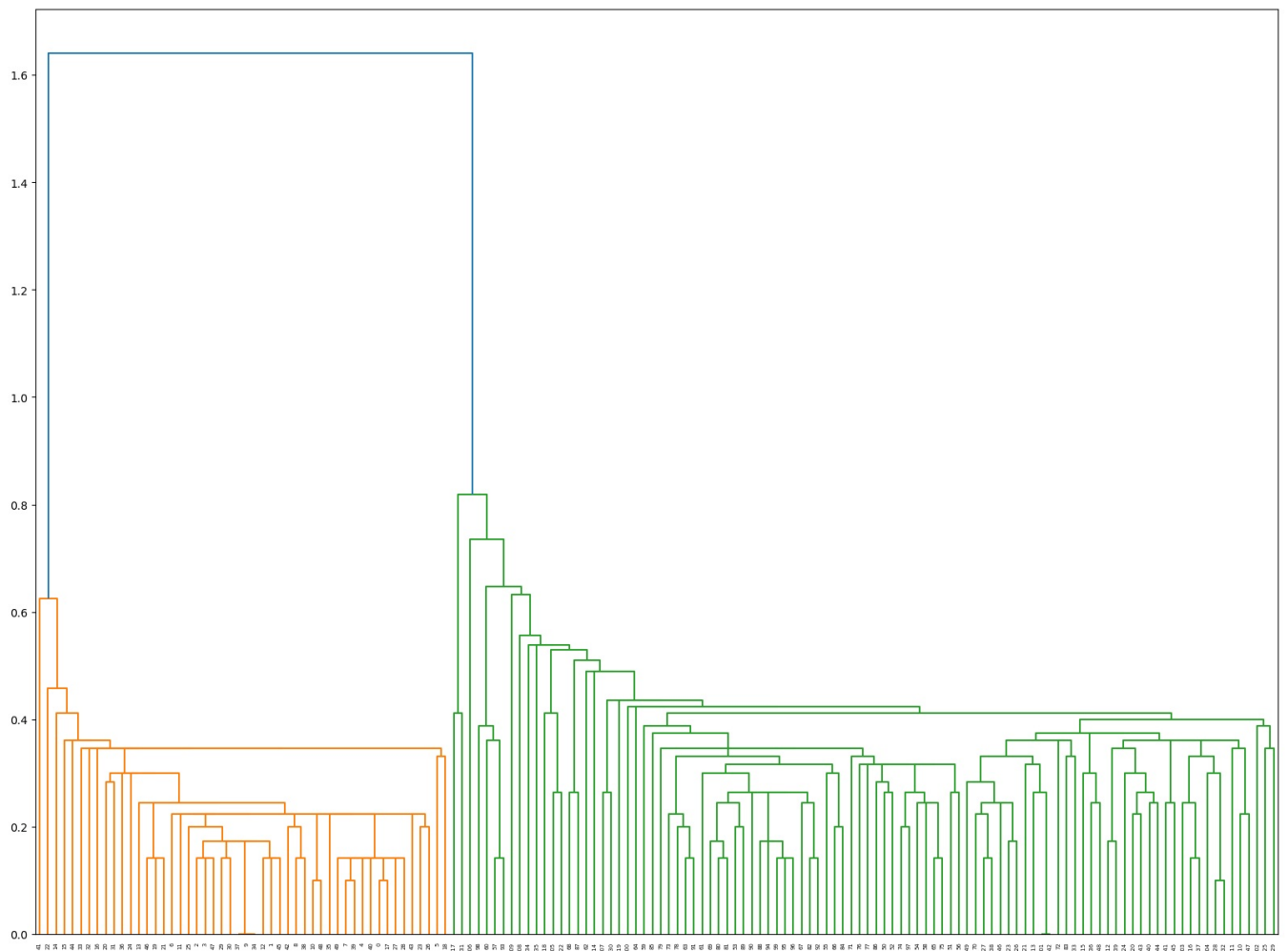


```
In [5]: ! pip install scipy
```

Requirement already satisfied: scipy in c:\programdata\anaconda3\lib\site-packages (1.11.1)
Requirement already satisfied: numpy<1.28.0,>=1.21.6 in c:\programdata\anaconda3\lib\site-packages (from scipy) (1.24.3)

```
In [6]: import scipy.cluster.hierarchy as sp
```

```
In [7]: plt.figure(figsize=(20,15))
sp.dendrogram(sp.linkage(df, method='single', metric='euclidean'))
plt.show()
```



```
In [8]: agc=AgglomerativeClustering(n_clusters=2,linkage="single")
df["predict"]=agc.fit_predict(df)
```

```
In [9]: df
```

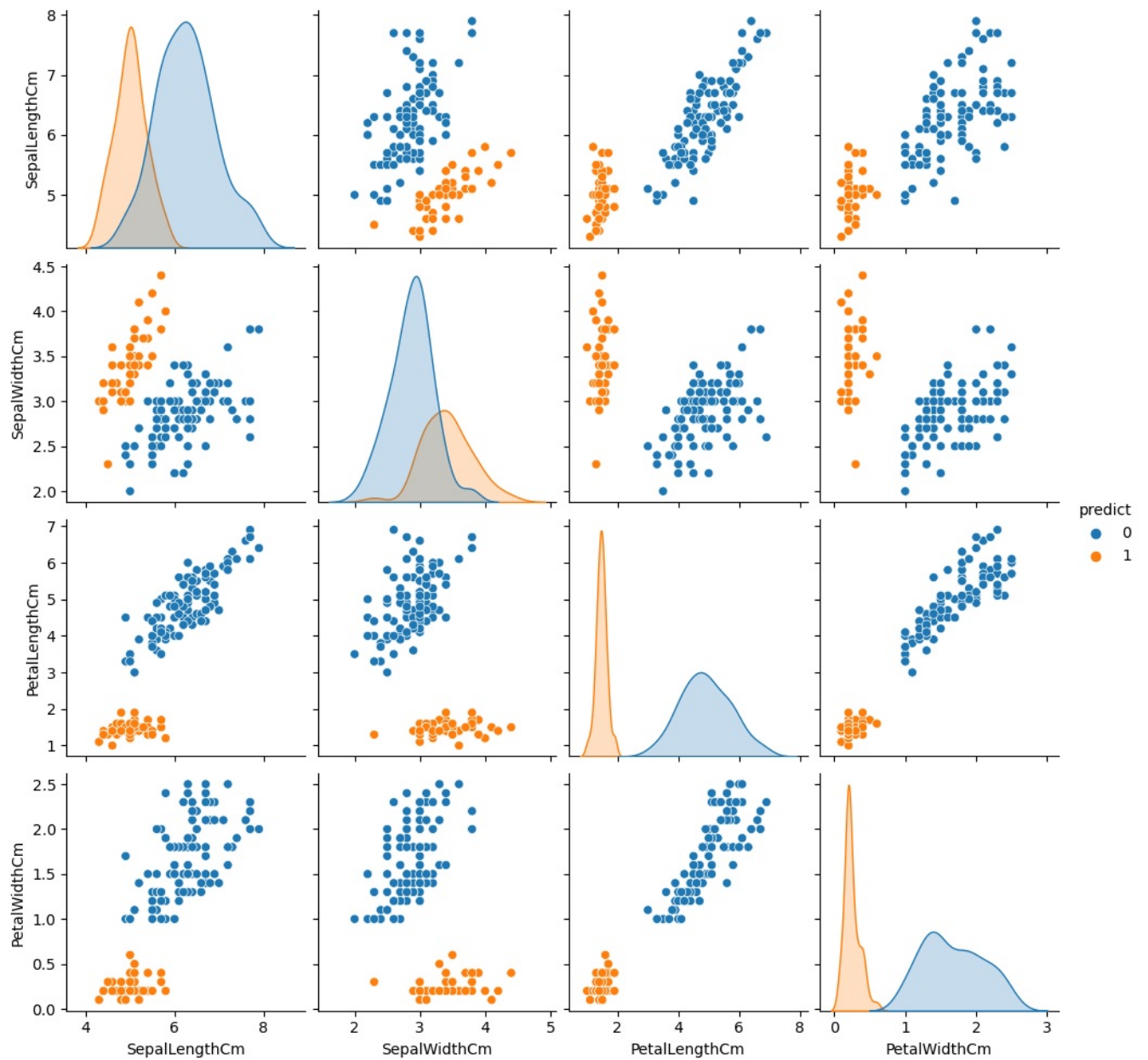
Out[9]:

| | SepalLengthCm | SepalWidthCm | PetalLengthCm | PetalWidthCm | predict |
|-----|---------------|--------------|---------------|--------------|---------|
| 0 | 5.1 | 3.5 | 1.4 | 0.2 | 1 |
| 1 | 4.9 | 3.0 | 1.4 | 0.2 | 1 |
| 2 | 4.7 | 3.2 | 1.3 | 0.2 | 1 |
| 3 | 4.6 | 3.1 | 1.5 | 0.2 | 1 |
| 4 | 5.0 | 3.6 | 1.4 | 0.2 | 1 |
| ... | ... | ... | ... | ... | ... |
| 145 | 6.7 | 3.0 | 5.2 | 2.3 | 0 |
| 146 | 6.3 | 2.5 | 5.0 | 1.9 | 0 |
| 147 | 6.5 | 3.0 | 5.2 | 2.0 | 0 |
| 148 | 6.2 | 3.4 | 5.4 | 2.3 | 0 |
| 149 | 5.9 | 3.0 | 5.1 | 1.8 | 0 |

150 rows × 5 columns

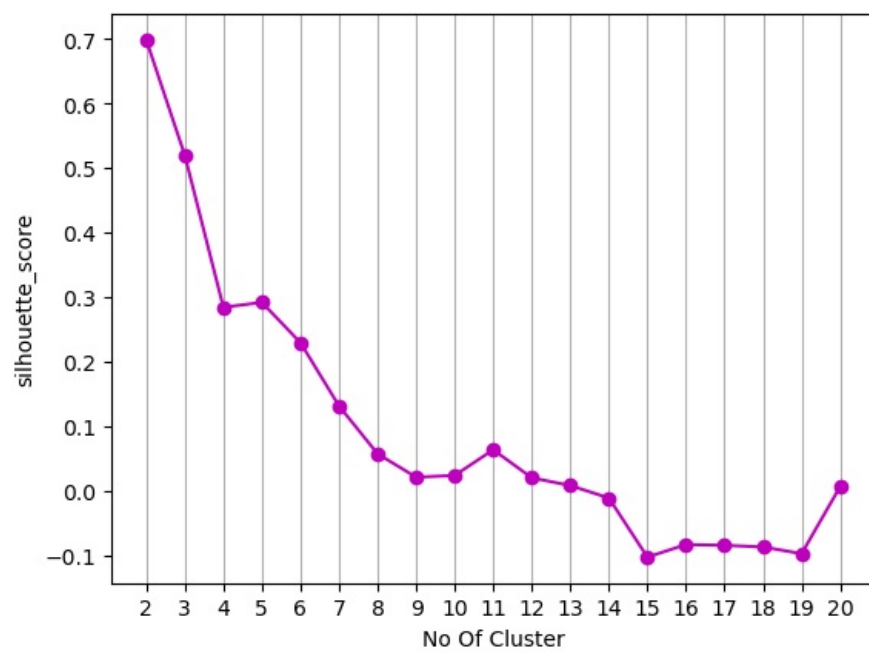
```
In [10]: sns.pairplot(df,hue="predict")
plt.show()
```

C:\ProgramData\anaconda3\Lib\site-packages\seaborn\axisgrid.py:118: UserWarning: The figure layout has changed to tight
self._figure.tight_layout(*args, **kwargs)



```
In [12]: ss=[]
for i in range(2,21):
    agc=AgglomerativeClustering(n_clusters=i,linkage="single")
    agc.fit(df)
    ss.append(silhouette_score(df,agc.labels_))
```

```
In [23]: plt.plot([i for i in range(2,21)],ss,marker="o",color="m")
plt.xlabel("No Of Cluster")
plt.xticks([i for i in range(2,21)])
plt.grid(axis="x")
plt.ylabel("silhouette_score")
plt.show()
```



In []:

Loading [MathJax]/jax/output/CommonHTML/fonts/TeX/fontdata.js