

```
In [24]: import pandas as pd
import matplotlib.pyplot as plt
```

```
In [25]: pure = pd.read_csv("../reports/knn_0.2.csv")
pca = pd.read_csv("../reports/pca_then_knn_0.2.csv")
```

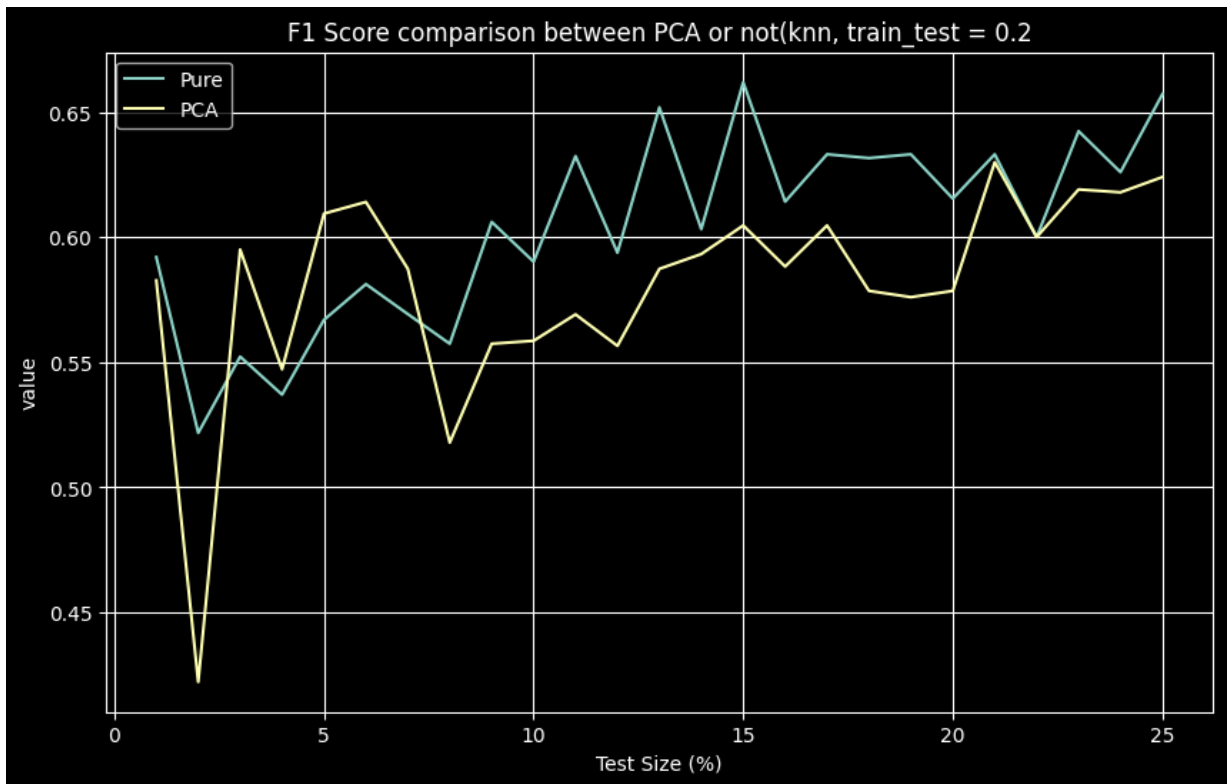
```
In [26]: pure.head()
pca.head()
```

```
Out[26]:
```

	k	F1 Score	Accuracy	Precision	Recall
0	1	0.582677	0.592308	0.552239	0.616667
1	2	0.422222	0.600000	0.633333	0.316667
2	3	0.595041	0.623077	0.590164	0.600000
3	4	0.547170	0.630769	0.630435	0.483333
4	5	0.609375	0.615385	0.573529	0.650000

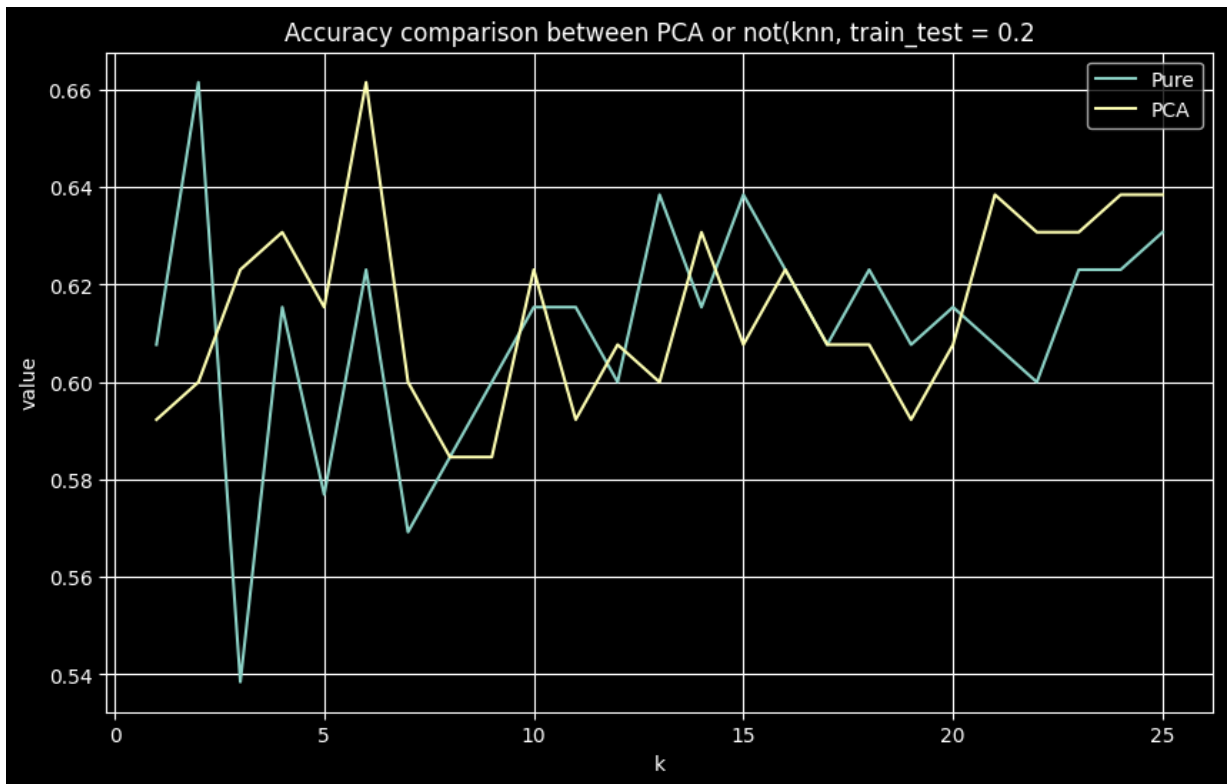
```
In [27]: plt.style.use('dark_background')
plt.figure(figsize=(10, 6))
plt.plot(pure['k'], pure['F1 Score'], label='Pure')
plt.plot(pca['k'], pca['F1 Score'], label='PCA')

plt.xlabel('Test Size (%)')
plt.ylabel('value')
plt.title('F1 Score comparison between PCA or not(knn, train_test = 0.2)')
plt.legend()
plt.grid(True)
plt.savefig('../plots/F1 Score comparison between PCA or not(knn, train_test = 0.2)')
plt.show()
```



```
In [28]: plt.style.use('dark_background')
plt.figure(figsize=(10, 6))
plt.plot(pure['k'], pure["Accuracy"], label='Pure', )
plt.plot(pca['k'], pca['Accuracy'], label='PCA')

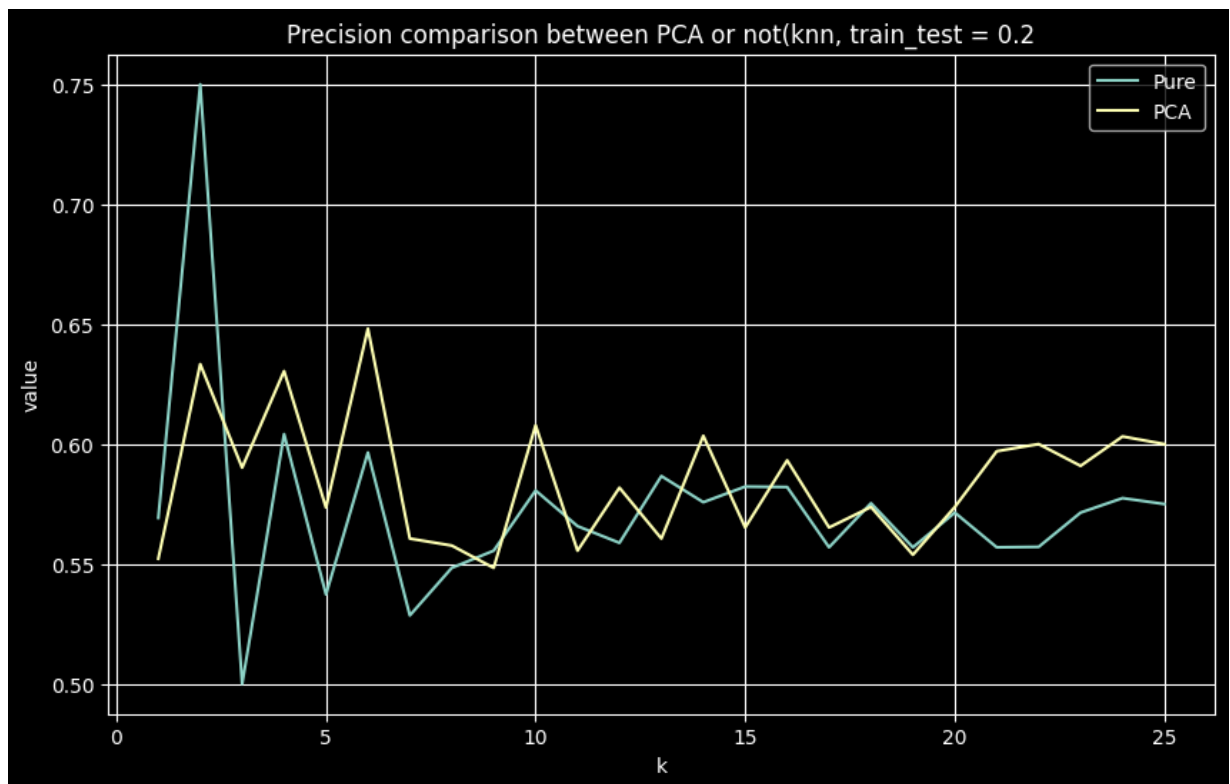
plt.xlabel('k')
plt.ylabel('value')
plt.title('Accuracy comparison between PCA or not(knn, train_test = 0.2)')
plt.legend()
plt.grid(True)
plt.savefig('../plots/Accuracy comparison between PCA or not(knn, train_test = 0.2.
plt.show())
```



```
In [29]: plt.style.use('dark_background')
plt.figure(figsize=(10, 6))
plt.plot(pure['k'], pure["Precision"], label='Pure', )
plt.plot(pca['k'], pca['Precision'], label='PCA')

plt.xlabel('k')
plt.ylabel('value')

plt.title('Precision comparison between PCA or not(knn, train_test = 0.2')
plt.legend()
plt.grid(True)
plt.savefig('../plots/Precision comparison between PCA or not(knn, train_test = 0.2
plt.show()
```



```
In [30]: plt.style.use('dark_background')
plt.figure(figsize=(10, 6))
plt.plot(pure['k'], pure["Recall"], label='Pure', )
plt.plot(pca['k'], pca['Recall'], label='PCA')

plt.xlabel('k')
plt.ylabel('value')
plt.title('Recall comparison between PCA or not(knn, train_test = 0.2)')
plt.legend()
plt.grid(True)
plt.savefig('../plots/Recall comparison between PCA or not(knn, train_test = 0.2.png')
plt.show()
```

