In [43]:
```
import pandas as pd
import matplotlib.pyplot as plt
```

In [44]:
```
pure = pd.read_csv("../reports/knn_0.3.csv")
pca = pd.read_csv("../reports/pca_then_knn_0.3.csv")
```

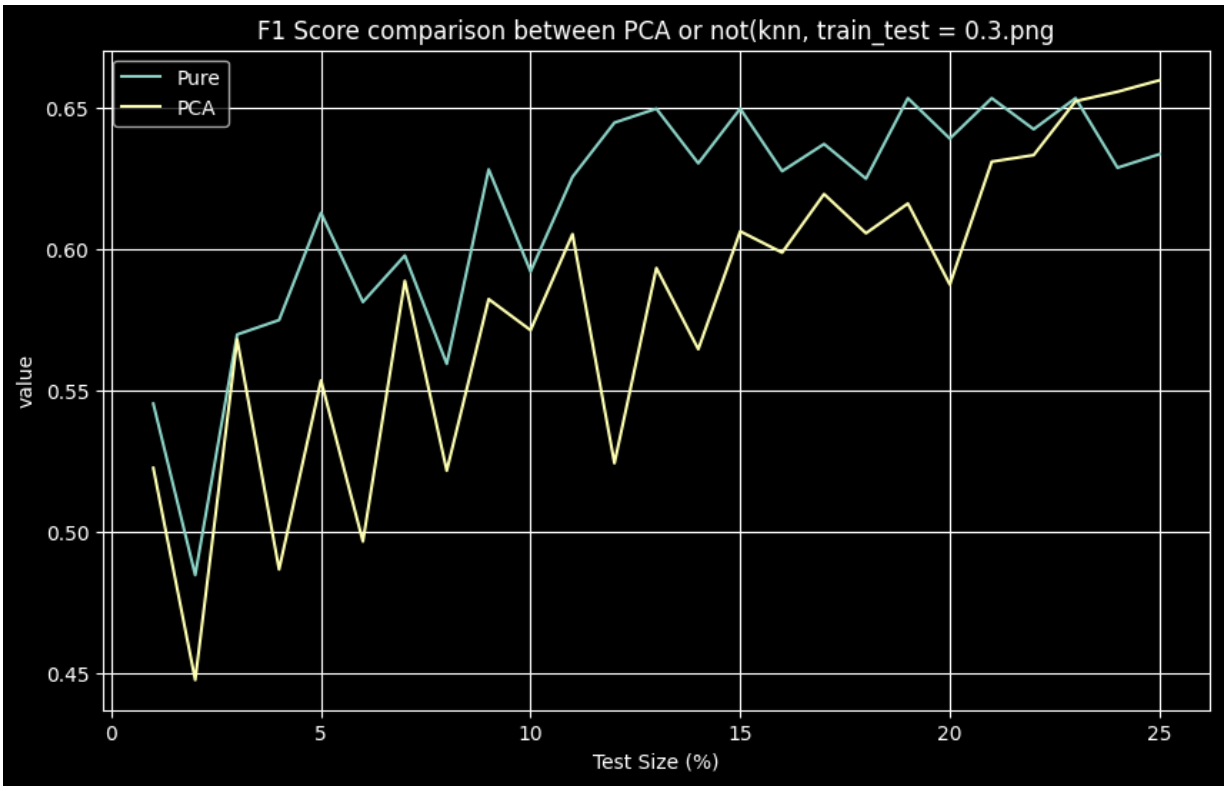
In [45]:
```
pure.head()
pca.head()
```

Out[45]:

	k	F1 Score	Accuracy	Precision	Recall
0	1	0.522727	0.569231	0.505495	0.541176
1	2	0.447761	0.620513	0.612245	0.352941
2	3	0.568182	0.610256	0.549451	0.588235
3	4	0.486842	0.600000	0.552239	0.435294
4	5	0.553672	0.594872	0.532609	0.576471

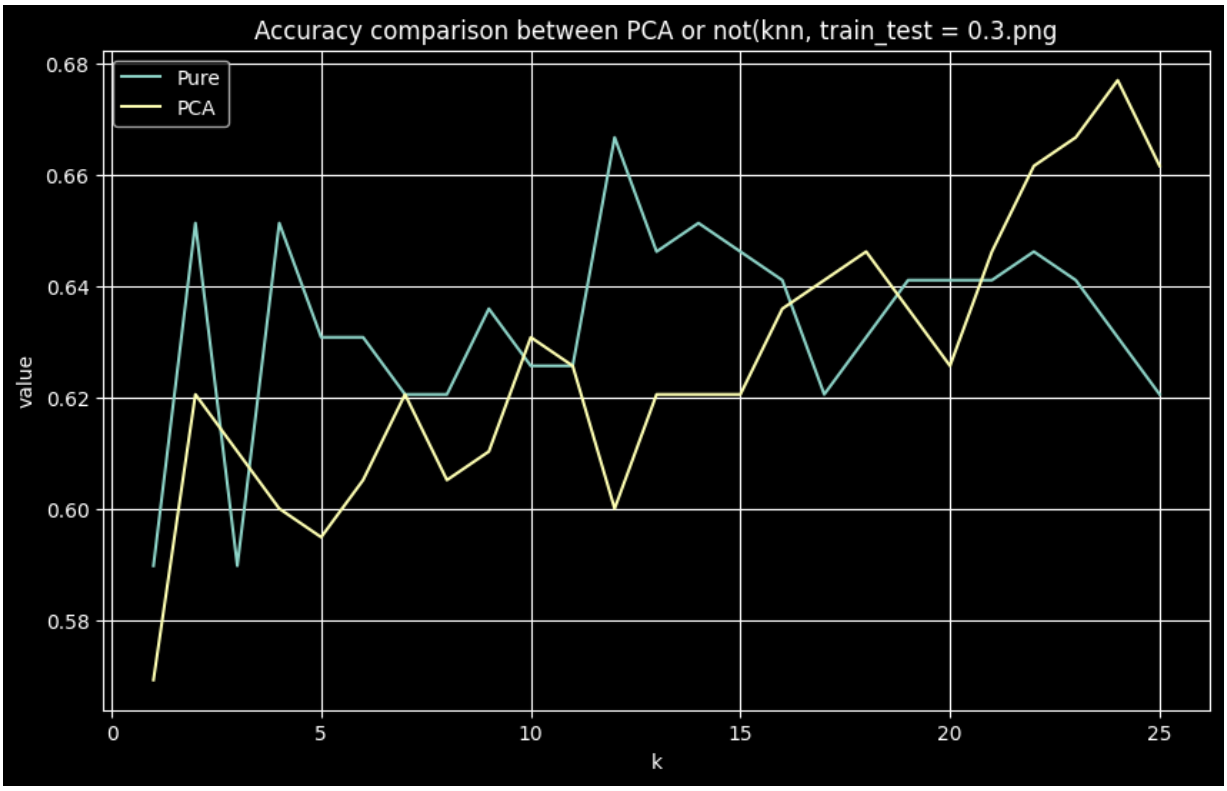
In [46]:
```
plt.style.use('dark_background')
plt.figure(figsize=(10, 6))
plt.plot(pure['k'], pure['F1 Score'], label='Pure')
plt.plot(pca['k'], pca['F1 Score'], label='PCA')

plt.xlabel('Test Size (%)')
plt.ylabel('value')
plt.title('F1 Score comparison between PCA or not(knn, train_test = 0.3.png')
plt.legend()
plt.grid(True)
plt.savefig('../plots/F1 Score comparison between PCA or not(knn, train_test = 0.3.
plt.show()
```



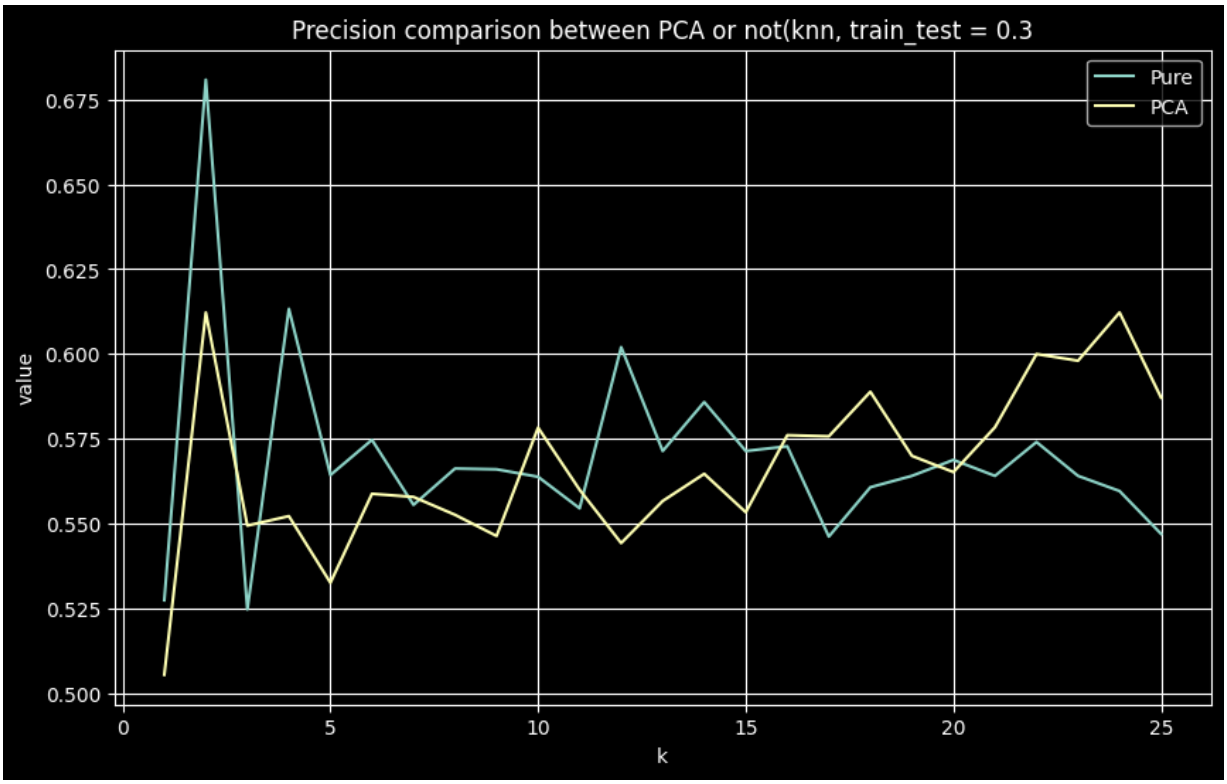
In [47]:
```
plt.style.use('dark_background')
plt.figure(figsize=(10, 6))
plt.plot(pure['k'], pure["Accuracy"], label='Pure', )
plt.plot(pca['k'], pca['Accuracy'], label='PCA')

plt.xlabel('k')
plt.ylabel('value')
plt.title('Accuracy comparison between PCA or not(knn, train_test = 0.3.png')
plt.legend()
plt.grid(True)
plt.savefig('../plots/Accuracy comparison between PCA or not(knn, train_test = 0.3.
plt.show()
```



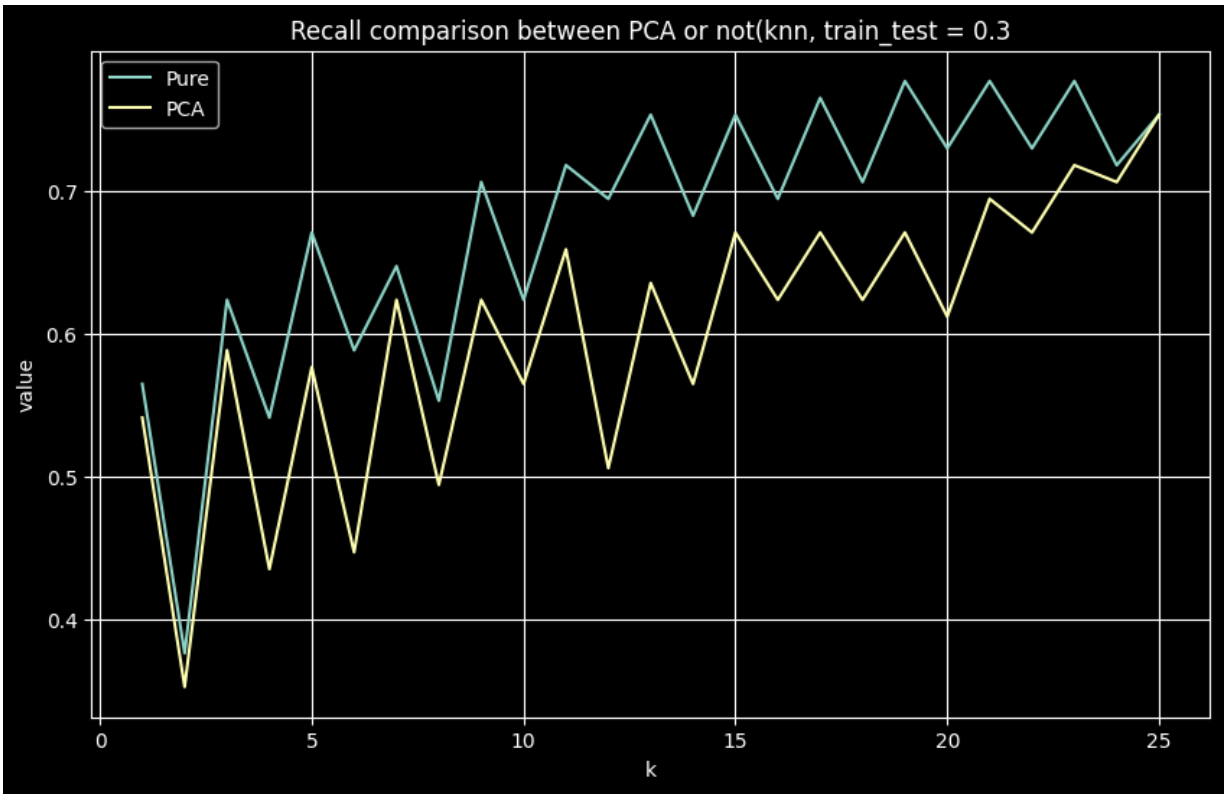
In [48]:
```
plt.style.use('dark_background')
plt.figure(figsize=(10, 6))
plt.plot(pure['k'], pure["Precision"], label='Pure', )
plt.plot(pca['k'], pca['Precision'], label='PCA')

plt.xlabel('k')
plt.ylabel('value')
plt.title('Precision comparison between PCA or not(knn, train_test = 0.3')
plt.legend()
plt.grid(True)
plt.savefig('../plots/Precision comparison between PCA or not(knn, train_test = 0.3
plt.show()
```



In [49]:
```
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 = 0.3')
plt.legend()
plt.grid(True)
plt.savefig('../plots/Recall comparison between PCA or not(knn, train_test = 0.3.pn
plt.show()
```


Recall comparison between PCA or not(knn, train_test = 0.3
Pure
PCA
value
k
0.4
0.5
0.6
0.7
0
5
10
15
20
25