```python
import pandas as pd

df = pd.read_csv("/content/drive/MyDrive/Colab Notebooks/utility_data_final.csv")

df.head()
```

	Electricity (kWh)	Water (litres/day)	Gas Flow (m^3/min)	Consumption Group
0	0.001298	107.0	0.029881	Normal
1	0.001597	97.0	0.032672	Normal
2	0.002266	64.0	0.038864	Normal
3	0.002279	87.0	0.040210	Normal
4	0.001024	86.0	0.029695	Normal

```python
from sklearn.model_selection import train_test_split

from sklearn.neighbors import KNeighborsClassifier


y = df['Consumption Group'].values
X = df.drop('Consumption Group', axis=1).values

X_train, X_test, y_train, y_test = train_test_split(X, y, random_state=42, test_size=0.3)


knn = KNeighborsClassifier(n_neighbors=6)
knn.fit(X_train, y_train)
y_pred = knn.predict(X_test)

from sklearn.metrics import classification_report
from sklearn.metrics import accuracy_score

print(classification_report(y_test, y_pred))
```

```
              precision    recall  f1-score   support

        High       0.54      0.59      0.56       146
         Low       0.60      0.74      0.66        87
      Normal       0.82      0.80      0.81       435
   Very High       0.88      0.60      0.71        82

    accuracy                           0.73       750
   macro avg       0.71      0.68      0.69       750
weighted avg       0.75      0.73      0.73       750
```

```python
print(accuracy_score(y_test, y_pred))
```

```
print(accuracy_score(y_test, y_pred))
```

```
0.732
```

```python
import seaborn as sns
import matplotlib.pyplot as plt
sns.heatmap(confusion_matrix2, annot=True, fmt='d', cmap='Blues', cbar=False)

# Add axis labels and title
plt.xlabel('Predicted label')
plt.ylabel('True label')
plt.title('Confusion Matrix')

# Add a legend
class_names = ['high', 'low', 'normal', 'v_high']
plt.xticks(range(len(class_names)), class_names, rotation=45)
plt.yticks(range(len(class_names)), class_names)
plt.show()
```
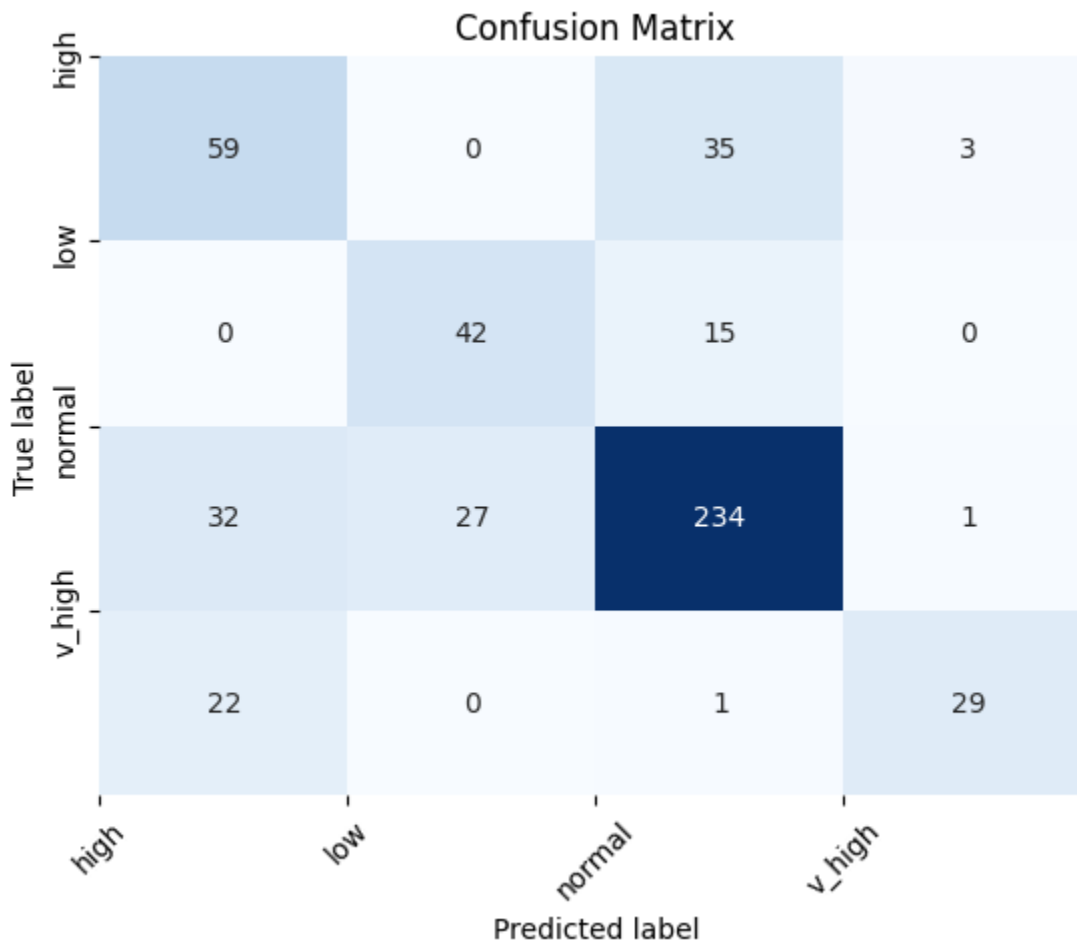

Start coding or generate with AI.