

```
import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
import seaborn as sns

from sklearn.model_selection import train_test_split
from sklearn.preprocessing import StandardScaler
from sklearn.ensemble import RandomForestClassifier
from sklearn.metrics import accuracy_score

df = pd.read_csv("Crop_recommendation.csv")
df.drop_duplicates(inplace=True)

if df.isnull().sum().sum() > 0:
    raise ValueError("Dataset contains null values.")

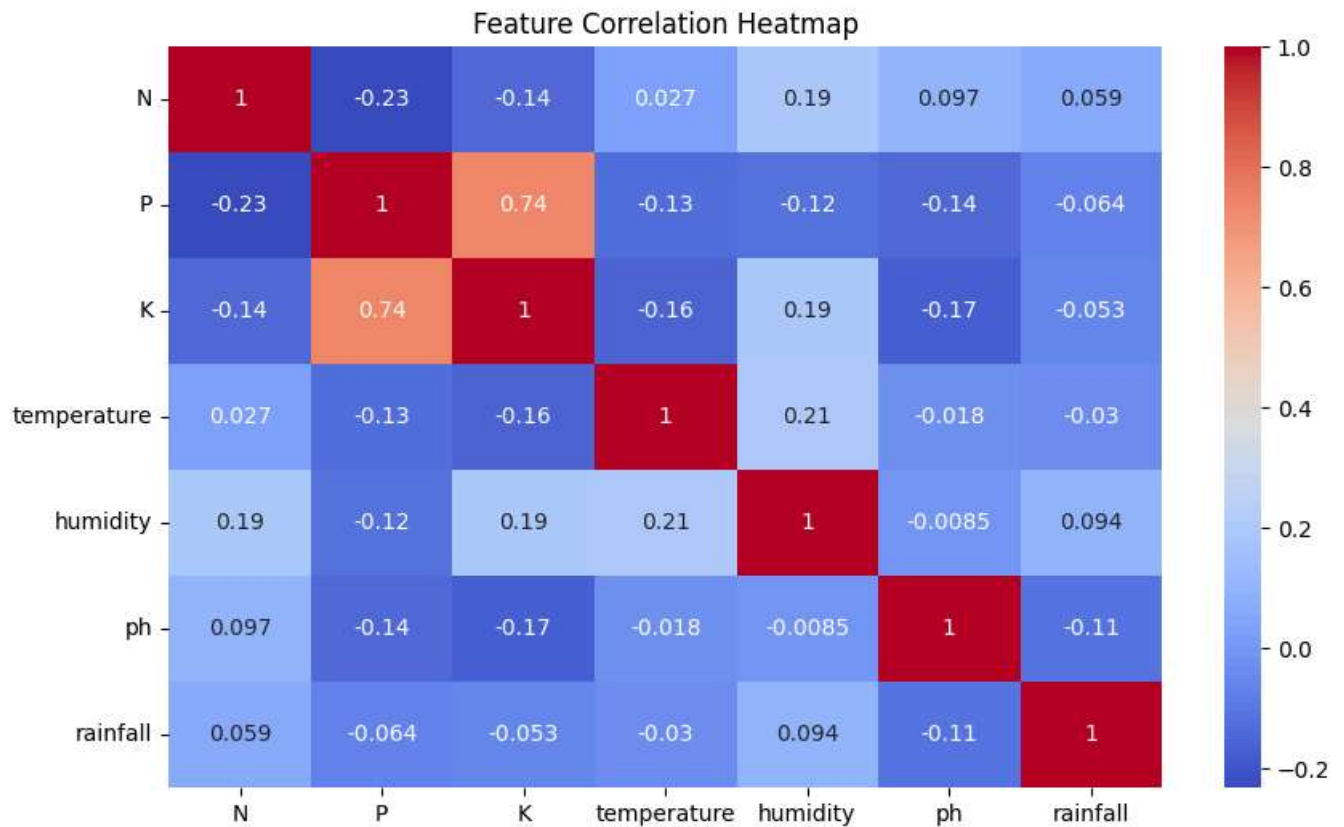
plt.figure(figsize=(10, 6))
sns.heatmap(df.corr(numeric_only=True), annot=True, cmap='coolwarm')
plt.title("Feature Correlation Heatmap")
plt.show()
```

Command palette	Ctrl+Shift+P
-----------------	--------------

Settings	
----------	--

Keyboard shortcuts	Ctrl+M H
--------------------	----------

Diff notebooks	
----------------	--



```
label_to_num = {label: idx for idx, label in enumerate(df['label'].unique(), 1)}
num_to_label = {v: k for k, v in label_to_num.items()}
df['crop_num'] = df['label'].map(label_to_num)
df.drop(columns=['label'], inplace=True)
```

```
X = df.drop(columns=['crop_num'])
y = df['crop_num']
```

```
X_train, X_test, y_train, y_test = train_test_split(X, y, test_size=0.2, random_state=42)
```

```
scaler = StandardScaler()
X_train = scaler.fit_transform(X_train)
X_test = scaler.transform(X_test)
```

```
model = RandomForestClassifier(random_state=42)
model.fit(X_train, y_train)
```



▼ RandomForestClassifier ⓘ ?

```
RandomForestClassifier(random_state=42)
```

```
y_pred = model.predict(X_test)
print(f"Random Forest Accuracy: {accuracy_score(y_test, y_pred):.4f}")
```



Random Forest Accuracy: 0.9932

```
feature_names = X.columns.tolist() # ['N','P','K','temperature','humidity','ph','rainfall']
```

```
def recommend_crop(N, P, K, temperature, humidity, ph, rainfall):
    """Return the best crop name for the supplied soil & climate parameters."""
    data = pd.DataFrame(
        [[N, P, K, temperature, humidity, ph, rainfall]],
        columns=feature_names # <-- gives scaler the names it expects
    )
    scaled = scaler.transform(data) # no warning now
    pred_num = model.predict(scaled)[0]
    return num_to_label[pred_num]
```

```
result = recommend_crop(90, 50, 50, 21, 90, 9.5, 200)
print(f"Recommended Crop: {result}")
```



Recommended Crop: papaya