Task 3: Decision Tree Classifier – Bank Marketing Dataset

In this task, we build a Decision Tree Classifier to predict whether a customer will subscribe to a term deposit.
We use the Bank Marketing Dataset and analyze customer demographics and behavior to train a classification model.

In [1]:
```
import pandas as pd
import matplotlib.pyplot as plt
import seaborn as sns
from sklearn.model_selection import train_test_split
from sklearn.tree import DecisionTreeClassifier
from sklearn.metrics import accuracy_score, classification_report, confusion_mat
```

In [8]:
```
import pandas as pd

df = pd.read_csv('bank-full.csv', sep=';')
df.head()
```

Out[8]:

	age	job	marital	education	default	balance	housing	loan	contact	da
0	58	management	married	tertiary	no	2143	yes	no	unknown	
1	44	technician	single	secondary	no	29	yes	no	unknown	
2	33	entrepreneur	married	secondary	no	2	yes	yes	unknown	
3	47	blue-collar	married	unknown	no	1506	yes	no	unknown	
4	33	unknown	single	unknown	no	1	no	no	unknown	

In [9]:
```
print(df.isnull().sum())
df_encoded = pd.get_dummies(df, drop_first=True)
X = df_encoded.drop('y_yes', axis=1)
y = df_encoded['y_yes']
X_train, X_test, y_train, y_test = train_test_split(X, y, test_size=0.3, random_
```

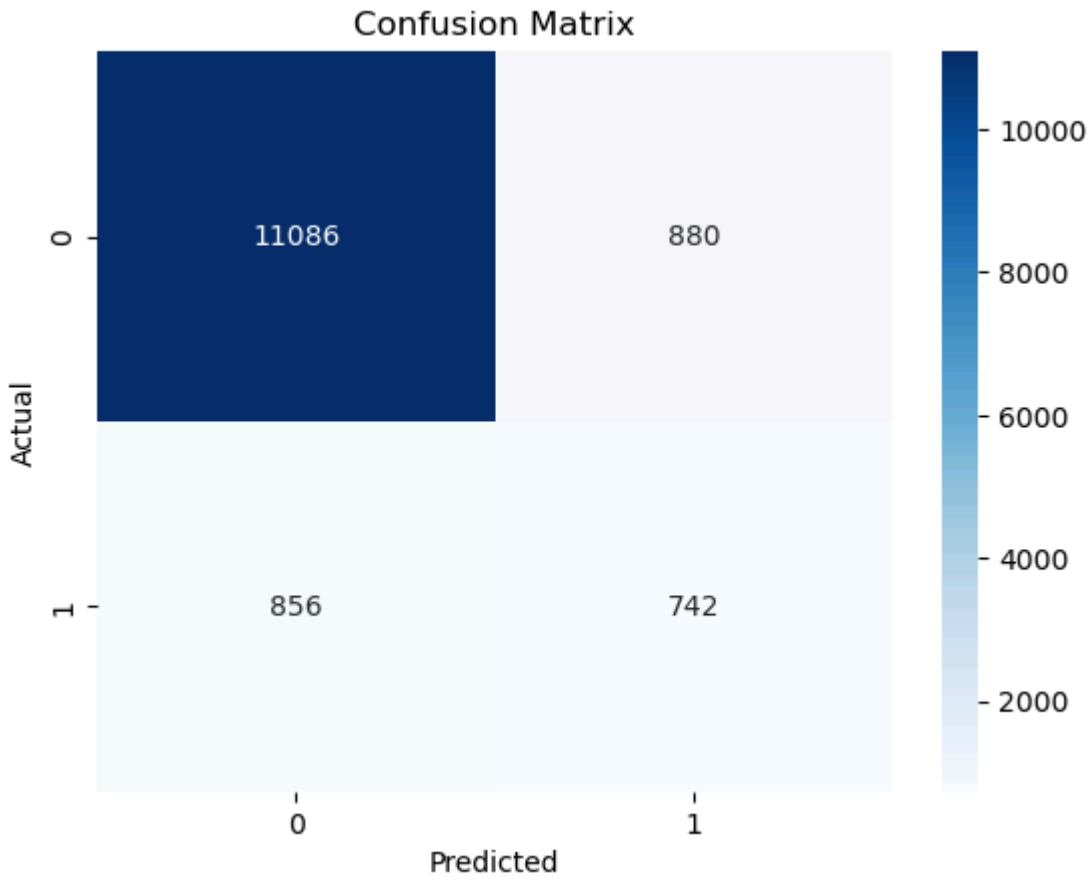
```
age          0
job          0
marital      0
education    0
default      0
balance      0
housing      0
loan         0
contact      0
day          0
month        0
duration     0
campaign     0
pdays        0
previous     0
poutcome     0
y            0
dtype: int64
```

In [10]:
```
model = DecisionTreeClassifier(random_state=42)
model.fit(X_train, y_train)
```

Out[10]:
```
DecisionTreeClassifier(random_state=42)
```

In [11]:
```
y_pred = model.predict(X_test)

# Accuracy
print("Accuracy:", accuracy_score(y_test, y_pred))

# Classification Report
print("\nClassification Report:")
print(classification_report(y_test, y_pred))

# Confusion Matrix
cm = confusion_matrix(y_test, y_pred)
sns.heatmap(cm, annot=True, fmt='d', cmap='Blues')
plt.title("Confusion Matrix")
plt.xlabel("Predicted")
plt.ylabel("Actual")
plt.show()
```

```
Accuracy: 0.8720141551164848

Classification Report:
              precision    recall  f1-score   support

       False       0.93      0.93      0.93     11966
        True       0.46      0.46      0.46      1598

    accuracy                           0.87     13564
   macro avg       0.69      0.70      0.69     13564
weighted avg       0.87      0.87      0.87     13564
```



Conclusion

- We successfully built a decision tree classifier to predict customer subscription.
- Model accuracy was reasonable, but can be improved with further tuning or cross-validation.
- Most important features were related to contact type, month, and previous campaign outcome.