

```

In [23]: import zipfile
import os
import pandas as pd
import matplotlib.pyplot as plt
from sklearn.model_selection import train_test_split
from sklearn.tree import DecisionTreeClassifier, plot_tree
from sklearn.metrics import accuracy_score, confusion_matrix, classification_report
from IPython.display import display, HTML

# Step 1: Extract Outer ZIP
outer_zip_path = r"C:\Users\ahire\Downloads\bank+marketing.zip"
outer_extract_path = "outer_extracted"

if not os.path.exists(outer_extract_path):
    with zipfile.ZipFile(outer_zip_path, 'r') as outer_zip:
        outer_zip.extractall(outer_extract_path)

# Step 2: Extract Inner ZIP
inner_zip_path = os.path.join(outer_extract_path, "bank-additional.zip")
inner_extract_path = "inner_extracted"

if not os.path.exists(inner_extract_path):
    with zipfile.ZipFile(inner_zip_path, 'r') as inner_zip:
        inner_zip.extractall(inner_extract_path)

# Step 3: Load CSV
csv_path = os.path.join(inner_extract_path, "bank-additional", "bank-additional.csv")
df = pd.read_csv(csv_path, sep=';')

# Step 4: Preprocessing
df['y'] = df['y'].map({'yes': 1, 'no': 0})
df_encoded = pd.get_dummies(df, drop_first=True)
X = df_encoded.drop('y', axis=1)
y = df_encoded['y']
X_train, X_test, y_train, y_test = train_test_split(X, y, test_size=0.2, random_state=42)

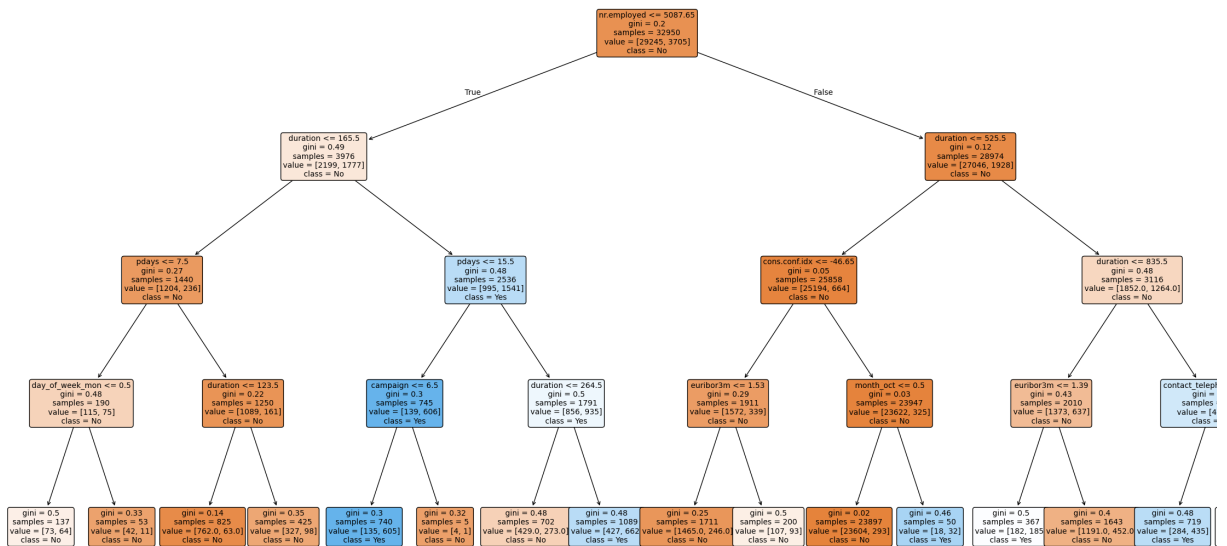
# Step 5: Train Decision Tree
clf = DecisionTreeClassifier(max_depth=4, random_state=42)
clf.fit(X_train, y_train)

# Step 6: Visualize Decision Tree
fig = plt.figure(figsize=(24, 12), facecolor='white')
ax = fig.add_subplot(1, 1, 1)
plot_tree(
    clf,
    filled=True,
    feature_names=X.columns,
    class_names=["No", "Yes"],
    rounded=True,
    precision=2,
    fontsize=10,
    ax=ax
)
ax.set_title("Decision Tree: Predicting Customer Subscription", fontsize=14)
plt.tight_layout()
plt.show()

# Step 7: Evaluate Model
y_pred = clf.predict(X_test)

```

Decision Tree: Predicting Customer Subscription



Accuracy: 0.9124

Confusion Matrix

	Predicted: No	Predicted: Yes
Actual: No	7052	251
Actual: Yes	471	464

Classification Report:

	precision	recall	f1-score	support
No	0.94	0.97	0.95	7303
Yes	0.65	0.50	0.56	935
accuracy			0.91	8238
macro avg	0.79	0.73	0.76	8238
weighted avg	0.90	0.91	0.91	8238

In []: