

```
1 # Import necessary libraries
2 import pandas as pd
3 from sklearn.model_selection import train_test_split
4 from sklearn.tree import DecisionTreeClassifier
5 from sklearn.metrics import accuracy_score

1 df = pd.read_csv('/content/drive/MyDrive/Colab Notebooks/Dataset/application_record.csv')
2 df.head()
```

COME_TYPE	NAME_EDUCATION_TYPE	NAME_FAMILY_STATUS	DAYS_BIRTH	DAYS_EMPLOYED	FLAG_M
Working	Higher education	Civil marriage	-12005	-4542	
Working	Higher education	Civil marriage	-12005	-4542	
Working	Secondary / secondary special	Married	-21474	-1134	
Commercial associate	Secondary / secondary special	Single / not married	-19110	-3051	
Commercial associate	Secondary / secondary special	Single / not married	-19110	-3051	

```
1 # Assuming df is your DataFrame containing the dataset
2 # Step 1: Data Preprocessing
3 # Drop rows with missing values
4 df.dropna(inplace=True)

1 # Encode categorical variables
2 df_encoded = pd.get_dummies(df, drop_first=True)

1 # Split dataset into features and target variable
2 X = df_encoded.drop(columns=['CREDIT']) # Assuming 'credit_score' is the target variable
3 y = df_encoded['CREDIT']

1 # Step 2: Train-Test Split
2 X_train, X_test, y_train, y_test = train_test_split(X, y, test_size=0.2, random_state=42)

1 # Step 3: Decision Tree Model
2 # Initialize decision tree classifier
3 clf = DecisionTreeClassifier()

1 # Train the model on the training data
2 clf.fit(X_train, y_train)
```

▼ DecisionTreeClassifier

DecisionTreeClassifier()

```
1 # Step 4: Model Evaluation
2 # Predictions on the testing data
3 y_pred = clf.predict(X_test)

1 # Evaluate model performance
2 accuracy = accuracy_score(y_test, y_pred)
3 print("Accuracy:", accuracy)

Accuracy: 0.985302165585177
```

