```
import pandas as pd
df = pd.read_csv('/content/train_delay_sample_10.csv')
print(df.head())
print(df.info())
from sklearn.preprocessing import LabelEncoder

le = LabelEncoder()
df['Train_Name'] = le.fit_transform(df['Train_Name'])
df['Departure_Station'] = le.fit_transform(df['Departure_Station'])
df['Arrival_Station'] = le.fit_transform(df['Arrival_Station'])
df['Departure_Time'] = le.fit_transform(df['Departure_Time'])
df['Scheduled_Arrival'] = le.fit_transform(df['Scheduled_Arrival'])
df['Actual_Arrival'] = le.fit_transform(df['Actual_Arrival'])
df['Weather_Condition'] = le.fit_transform(df['Weather_Condition'])
df['Is_Delayed']=le.fit_transform(df['Is_Delayed'])
from sklearn.model_selection import train_test_split

X = df.drop('Is_Delayed', axis=1)
y = df['Is_Delayed']
X_train, X_test, y_train, y_test = train_test_split(X, y, test_size=0.2, random_state=42)
from sklearn.ensemble import RandomForestClassifier
from sklearn.metrics import classification_report

model = RandomForestClassifier()
model.fit(X_train, y_train)

y_pred = model.predict(X_test)
print(classification_report(y_test, y_pred))
from sklearn.metrics import accuracy_score

accuracy = accuracy_score(y_test, y_pred)
print(f"Accuracy: {accuracy}")
from sklearn.model_selection import cross_val_score

scores = cross_val_score(model, X, y, cv=5)
print(f"Average CV Accuracy: {scores.mean()}")
import pandas as pd
df['Departure_Time'] = pd.to_datetime(df['Departure_Time'])
df['Scheduled_Arrival'] = pd.to_datetime(df['Scheduled_Arrival'])
df['Actual_Arrival'] = pd.to_datetime(df['Actual_Arrival'])

df['Delay_Minutes'] = (df['Actual_Arrival'] - df['Scheduled_Arrival']).dt.total_seconds() / 60
df['Departure_Hour'] = df['Departure_Time'].dt.hour
df['Day_of_Week'] = df['Departure_Time'].dt.dayofweek
```

```
  Train_Name Departure_Station Arrival_Station       Departure_Time  \
0    Express            Mumbai       Bangalore  2024-01-28 11:45:12
1      Local           Chennai            Pune  2024-03-31 23:42:55
2       Mail           Chennai         Lucknow  2024-07-28 13:48:12
3  Superfast              Pune          Jaipur  2024-06-04 14:10:53
4  Superfast           Kolkata          Jaipur  2024-09-19 05:08:03

     Scheduled_Arrival       Actual_Arrival Weather_Condition  Is_Delayed
0  2024-01-29 10:45:12  2024-01-29 11:49:12             Storm           1
1  2024-04-01 09:42:55  2024-04-01 09:42:55             Storm           0
2  2024-07-29 06:48:12  2024-07-29 08:14:12             Clear           1
3  2024-06-05 04:10:53  2024-06-05 05:49:53              Rain           1
4  2024-09-20 09:08:03  2024-09-20 09:08:03             Clear           0
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 10 entries, 0 to 9
Data columns (total 8 columns):
 #   Column             Non-Null Count  Dtype
---  ------             --------------  -----
 0   Train_Name         10 non-null     object
 1   Departure_Station  10 non-null     object
 2   Arrival_Station    10 non-null     object
 3   Departure_Time     10 non-null     object
 4   Scheduled_Arrival  10 non-null     object
 5   Actual_Arrival     10 non-null     object
 6   Weather_Condition  10 non-null     object
 7   Is_Delayed         10 non-null     int64
dtypes: int64(1), object(7)
memory usage: 772.0+ bytes
None
              precision    recall  f1-score   support

           0       0.00      0.00      0.00       2.0
           1       0.00      0.00      0.00       0.0

    accuracy                           0.00       2.0
   macro avg       0.00      0.00      0.00       2.0
weighted avg       0.00      0.00      0.00       2.0
```

```
Accuracy: 0.0
/usr/local/lib/python3.11/dist-packages/sklearn/metrics/_classification.py:1565: UndefinedMetricWarning: Precision is il
  _warn_prf(average, modifier, f"{metric.capitalize()} is", len(result))
/usr/local/lib/python3.11/dist-packages/sklearn/metrics/_classification.py:1565: UndefinedMetricWarning: Recall is ill-d
  _warn_prf(average, modifier, f"{metric.capitalize()} is", len(result))
/usr/local/lib/python3.11/dist-packages/sklearn/metrics/_classification.py:1565: UndefinedMetricWarning: Precision is il
  _warn_prf(average, modifier, f"{metric.capitalize()} is", len(result))
/usr/local/lib/python3.11/dist-packages/sklearn/metrics/_classification.py:1565: UndefinedMetricWarning: Recall is ill-d
  _warn_prf(average, modifier, f"{metric.capitalize()} is", len(result))
/usr/local/lib/python3.11/dist-packages/sklearn/metrics/_classification.py:1565: UndefinedMetricWarning: Precision is il
  _warn_prf(average, modifier, f"{metric.capitalize()} is", len(result))
/usr/local/lib/python3.11/dist-packages/sklearn/metrics/_classification.py:1565: UndefinedMetricWarning: Recall is ill-d
  _warn_prf(average, modifier, f"{metric.capitalize()} is", len(result))
/usr/local/lib/python3.11/dist-packages/sklearn/model_selection/_split.py:805: UserWarning: The least populated class in
  warnings.warn(
Average CV Accuracy: 0.3
```