In [23]:

```
import pandas as pd
from sklearn.model_selection import train_test_split
from sklearn.ensemble import RandomForestClassifier
from sklearn.preprocessing import OneHotEncoder
import pickle

# Assuming df is your original DataFrame with necessary features and target
df = pd.read_csv('C:/Users/asus/IMDb_Top_250_Movies.csv')
df['Rating'] = df['Rating'].str.extract('(\d+\.\d+)').astype(float)

# Only focus on 'Director', 'Genre', and 'Rating'
df = df[['Director', 'Genre', 'Rating']]

# Encode categorical data
encoder = OneHotEncoder()
X = encoder.fit_transform(df[['Director', 'Genre']])
y = df['Rating'] >= 8.5  # For example, predicting high ratings, adjust as n

# Split data
X_train, X_test, y_train, y_test = train_test_split(X, y, test_size=0.2, ran

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

# Save model and encoder
with open('rating_model.pkl', 'wb') as f:
    pickle.dump(model, f)
with open('encoder1.pkl', 'wb') as f:
    pickle.dump(encoder, f)

```

In [ ]:

```

```