```python
import pandas as pd
import numpy as np
from sklearn.model_selection import train_test_split
from sklearn.preprocessing import StandardScaler, LabelEncoder
from sklearn.linear_model import LogisticRegression
from sklearn.metrics import classification_report, accuracy_score

# 1. قراءة البيانات
df = pd.read_csv("df_final_features.csv")

# 2. إزالة التكرارات
df = df.drop_duplicates()

# 3. معالجة القيم المفقودة
df = df.dropna(subset=['Height', 'Weight'])  # حذف صفوف تحتوي على NaN في هذه الأعمدة
df['Avg_Temp'] = pd.to_numeric(df['Avg_Temp'], errors='coerce')  # تحويل
Avg_Temp إلى float
df = df.dropna(subset=['Avg_Temp'])  # حذف صفوف فيها Avg_Temp غير قابلة للتحويل

# 4. تحويل الأعمدة النصية إلى رقمية
df['Sex'] = df['Sex'].map({'M': 0, 'F': 1})  # تحويل الجنس
label_enc_cols = ['Team', 'Season', 'Host_City', 'Host_Country', 'Sport',
'Event']
for col in label_enc_cols:
    df[col] = LabelEncoder().fit_transform(df[col])

# 5. تحديد الميزات والهدف (Features and Target)
X = df.drop(columns=['Name', 'Medal', 'Medal_Binary'])  # الميزات
y = df['Medal_Binary']  # الهدف

# 6. تقسيم البيانات إلى تدريب واختبار
X_train, X_test, y_train, y_test = train_test_split(X, y, test_size=0.2,
random_state=42)

# 7. تطبيع البيانات
scaler = StandardScaler()
X_train = scaler.fit_transform(X_train)
X_test = scaler.transform(X_test)

# 8. تدريب نموذج (لوجستك ريجريشن)
model = LogisticRegression()
model.fit(X_train, y_train)
```