# 逻辑回归分类模板（自动处理无用特征）

import pandas as pd

from sklearn.linear\_model import LogisticRegression

from sklearn.preprocessing import StandardScaler

# 1. 加载数据

train = pd.read\_csv("train.csv") # 训练集

test = pd.read\_csv("test.csv") # 测试集

# 2. 指定列（根据实际数据修改）

target\_col = "target" # 目标列名（必须是二分类0/1）

useless\_cols = ["id"] # 要删除的无用列名

# 3. 自动选择有效特征

used\_features = [col for col in train.columns

if col not in useless\_cols + [target\_col]]

# 4. 数据标准化（重要！）

scaler = StandardScaler()

X\_train = scaler.fit\_transform(train[used\_features])

X\_test = scaler.transform(test[used\_features])

y\_train = train[target\_col]

# 5. 训练模型

model = LogisticRegression(max\_iter=1000) # 增加迭代次数确保收敛

model.fit(X\_train, y\_train)

# 6. 预测并保存结果

pd.DataFrame({

"id": test["id"], # 保留ID列

"prediction": model.predict(X\_test)

}).to\_csv("predictions.csv", index=False)

print("预测完成！结果已保存至 predictions.csv")

print(f"使用的特征：{used\_features}")

print(f"模型准确率（训练集）：{model.score(X\_train, y\_train):.2%}")