```
import pandas as pd
import numpy as np
from sklearn.linear_model import LogisticRegression

dia = pd.read_excel("/content/drive/MyDrive/copy/adult.xls")
dia.head()
```


```
{"summary":"{\n  \"name\": \"dia\",\n  \"rows\": 32561,\n  \"fields\":
[\n    {\n      \"column\": \"age\",\n      \"properties\": {\n
\"dtype\": \"number\",\n        \"std\": 13,\n        \"min\": 17,\n
\"max\": 90,\n        \"num_unique_values\": 73,\n        \"samples\":
[\n          41,\n          80,\n          61\n        ],\n
\"semantic_type\": \"\",\n        \"description\": \"\"\n      }\
n    },\n    {\n      \"column\": \"workclass\",\n
\"properties\": {\n        \"dtype\": \"category\",\n
\"num_unique_values\": 9,\n        \"samples\": [\n
\"Without-pay\",\n          \"Private\",\n          \"Self-emp-inc\"\n
],\n        \"semantic_type\": \"\",\n        \"description\": \"\"\n
}\n    },\n    {\n      \"column\": \"fnlwgt\",\n      \"properties\":
{\n        \"dtype\": \"number\",\n        \"std\": 105549,\n
\"min\": 12285,\n        \"max\": 1484705,\n
\"num_unique_values\": 21648,\n        \"samples\": [\n
269687,\n          156542,\n          100295\n        ],\n
\"semantic_type\": \"\",\n        \"description\": \"\"\n      }\
n    },\n    {\n      \"column\": \"education\",\n
\"properties\": {\n        \"dtype\": \"category\",\n
\"num_unique_values\": 16,\n        \"samples\": [\n          \"HS-
grad\",\n          \"Some-college\",\n          \"Prof-school\"\n
],\n        \"semantic_type\": \"\",\n        \"description\": \"\"\n
}\n    },\n    {\n      \"column\": \"education.num\",\n
\"properties\": {\n        \"dtype\": \"number\",\n        \"std\":
2,\n        \"min\": 1,\n        \"max\": 16,\n
\"num_unique_values\": 16,\n        \"samples\": [\n          9,\n
10,\n          15\n        ],\n        \"semantic_type\": \"\",\n
\"description\": \"\"\n      }\n    },\n    {\n      \"column\":
\"marital.status\",\n      \"properties\": {\n        \"dtype\":
\"category\",\n        \"num_unique_values\": 7,\n        \"samples\":
[\n          \"Widowed\",\n          \"Divorced\",\n
\"Married-spouse-absent\"\n        ],\n        \"semantic_type\":
\"\",\n        \"description\": \"\"\n      }\n    },\n    {\n
\"column\": \"occupation\",\n      \"properties\": {\n
\"dtype\": \"category\",\n        \"num_unique_values\": 15,\n
\"samples\": [\n          \"Sales\",\n          \"Tech-support\",\n
\"?\"\n        ],\n        \"semantic_type\": \"\",\n
\"description\": \"\"\n      }\n    },\n    {\n      \"column\":
\"relationship\",\n      \"properties\": {\n        \"dtype\":
\"category\",\n        \"num_unique_values\": 6,\n        \"samples\":
[\n          \"Not-in-family\",\n          \"Unmarried\",\n
\"Wife\"\n        ],\n        \"semantic_type\": \"\",\n
\"description\": \"\"\n      }\n    },\n    {\n      \"column\":
```

\"race\",\n   \"properties\": {\n   \"dtype\": \"category\",\n
\"num_unique_values\": 5,\n   \"samples\": [\n
\"Black\",\n   \"Amer-Indian-Eskimo\",\n   \"Asian-Pac-
Islander\"\n   ],\n   \"semantic_type\": \"\",\n
\"description\": \"\"\n   }\n   },\n   {\n   \"column\":
\"sex\",\n   \"properties\": {\n   \"dtype\": \"category\",\n
\"num_unique_values\": 2,\n   \"samples\": [\n
\"Male\",\n   \"Female\"\n   ],\n
\"semantic_type\": \"\",\n   \"description\": \"\"\n   }\
n   },\n   {\n   \"column\": \"capital.gain\",\n
\"properties\": {\n   \"dtype\": \"number\",\n   \"std\":
7385,\n   \"min\": 0,\n   \"max\": 99999,\n
\"num_unique_values\": 119,\n   \"samples\": [\n   3887,\n
2329\n   ],\n   \"semantic_type\": \"\",\n
\"description\": \"\"\n   }\n   },\n   {\n   \"column\":
\"capital.loss\",\n   \"properties\": {\n   \"dtype\":
\"number\",\n   \"std\": 402,\n   \"min\": 0,\n
\"max\": 4356,\n   \"num_unique_values\": 92,\n
\"samples\": [\n   1980,\n   2258\n   ],\n
\"semantic_type\": \"\",\n   \"description\": \"\"\n   }\
n   },\n   {\n   \"column\": \"hours.per.week\",\n
\"properties\": {\n   \"dtype\": \"number\",\n   \"std\":
12,\n   \"min\": 1,\n   \"max\": 99,\n
\"num_unique_values\": 94,\n   \"samples\": [\n   34,\n
80\n   ],\n   \"semantic_type\": \"\",\n
\"description\": \"\"\n   }\n   },\n   {\n   \"column\":
\"native.country\",\n   \"properties\": {\n   \"dtype\":
\"category\",\n   \"num_unique_values\": 42,\n
\"samples\": [\n   \"Peru\",\n   \"Puerto-Rico\"\n
],\n   \"semantic_type\": \"\",\n   \"description\": \"\"\n
}\n   },\n   {\n   \"column\": \"income\",\n   \"properties\":
{\n   \"dtype\": \"category\",\n   \"num_unique_values\":
2,\n   \"samples\": [\n   \">50K\",\n   \"<=50K\"\n
],\n   \"semantic_type\": \"\",\n   \"description\": \"\"\n
}\n   }\n]\n}","type":"dataframe","variable_name":"dia"}

```
dia.isnull().sum()
```

```
age               0
workclass         0
fnlwgt            0
education         0
education.num     0
marital.status    0
occupation        0
relationship      0
race              0
sex               0
capital.gain      0
capital.loss      0
```

```
hours.per.week    0
native.country    0
income            0
dtype: int64
```

```
dia_h = dia[["age","capital.loss","hours.per.week","income"]]
```

```
ind = dia[["age","capital.loss","hours.per.week"]]
dep = dia['income']
```

```
Logr = LogisticRegression()
```

```
Logr.fit(ind, dep)
```

```
LogisticRegression()
```

```
age = int(input("Enter the age : "))
capital_loss = int(input("Enter the capital.loss : "))
hours_per_week = int(input("Enter the hours.per.week : "))
pred = Logr.predict([[age, capital_loss, hours_per_week]])
print(pred)
```

```
Enter the age : 56
Enter the capital.loss : 2357
Enter the hours.per.week : 15
['<=50K']
```

```
/usr/local/lib/python3.12/dist-packages/sklearn/utils/
validation.py:2739: UserWarning: X does not have valid feature names,
but LogisticRegression was fitted with feature names
  warnings.warn(
```