```
import pandas as pd
from  sklearn.linear_model import LinearRegression
from sklearn.preprocessing import LabelEncoder

house=pd.read_csv("/content/house_prices.csv")

house.head()
```

```
{"summary":"{\n  \"name\": \"house\",\n  \"rows\": 1000,\n
\"fields\": [\n    {\n      \"column\": \"area\",\n
\"properties\": {\n        \"dtype\": \"number\",\n        \"std\":
60,\n        \"min\": 40,\n        \"max\": 249,\n
\"num_unique_values\": 208,\n        \"samples\": [\n          158,\n
191,\n          121\n        ],\n        \"semantic_type\": \"\",\n
\"description\": \"\"\n      }\n    },\n    {\n      \"column\":
\"rooms\",\n      \"properties\": {\n        \"dtype\": \"number\",\n
\"std\": 1,\n        \"min\": 1,\n        \"max\": 5,\n
\"num_unique_values\": 5,\n        \"samples\": [\n          4,\n
1,\n          5\n        ],\n        \"semantic_type\": \"\",\n
\"description\": \"\"\n      }\n    },\n    {\n      \"column\":
\"age\",\n      \"properties\": {\n        \"dtype\": \"number\",\n
\"std\": 14,\n        \"min\": 0,\n        \"max\": 49,\n
\"num_unique_values\": 50,\n        \"samples\": [\n          33,\n
47,\n          9\n        ],\n        \"semantic_type\": \"\",\n
\"description\": \"\"\n      }\n    },\n    {\n      \"column\":
\"location\",\n      \"properties\": {\n        \"dtype\":
\"category\",\n        \"num_unique_values\": 4,\n        \"samples\":
[\n          \"Downtown\",\n          \"Countryside\",\n
\"Suburb\"\n        ],\n        \"semantic_type\": \"\",\n
\"description\": \"\"\n      }\n    },\n    {\n      \"column\":
\"price\",\n      \"properties\": {\n        \"dtype\": \"number\",\n
\"std\": 91698,\n        \"min\": 230,\n        \"max\": 438845,\n
\"num_unique_values\": 998,\n        \"samples\": [\n
288086,\n          64267,\n          256737\n        ],\n
\"semantic_type\": \"\",\n        \"description\": \"\"\n      }\
n    }\n  ]\n}","type":"dataframe","variable_name":"house"}
```

```
house.isnull().sum()
```

```
area        0
rooms       0
age         0
location    0
price       0
dtype: int64
```

```
le=LabelEncoder()

house['location']=le.fit_transform(house['location'])

int_data=house[['area','rooms','age','location']]
```

```
dep_data=house['price']

LR=LinearRegression()

LR.fit(int_data,dep_data)
```

```
LinearRegression()
```

```
from google.colab import drive
drive.mount('/content/drive')
```

```
Mounted at /content/drive
```

```
import pickle
with open('house_prices.pkl','wb') as f:
  pickle.dump(LR,f)

LR.predict([[32,1,28,880]])
```

```
/usr/local/lib/python3.11/dist-packages/sklearn/utils/
validation.py:2739: UserWarning: X does not have valid feature names,
but LinearRegression was fitted with feature names
  warnings.warn(

array([14556470.30097486])
```