

Multiple Linear Regression Algorithm

```
In [2]: import warnings
warnings.filterwarnings("ignore")
```

```
In [34]: import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
import seaborn as sns
from sklearn.model_selection import train_test_split
from sklearn.linear_model import LinearRegression
from sklearn.metrics import mean_squared_error
from sklearn.metrics import r2_score
from sklearn.datasets import fetch_openml
from sklearn.model_selection import cross_val_score
from sklearn.preprocessing import OneHotEncoder
from sklearn.compose import ColumnTransformer
import pickle
```

```
In [25]: data = pd.read_csv(r"C:\Users\JANHAVI\Downloads\50_Startups.csv")
```

```
In [39]: x = data.iloc[:, :-1]
y = data.iloc[:, 4]
```

```
In [40]: ##convert the column into categorical columns
states = pd.get_dummies(x['State'], drop_first=True)
```

```
In [41]: # Drop the states column
X = x.drop('State', axis=1)
```

```
In [42]: # Concat the dummy variable
X = pd.concat([X, states], axis=1)
```

```
In [43]: # Split the dataset into training and testing sets
x_train, x_test, y_train, y_test = train_test_split(X, y, test_size=0.20, random_st
```

```
In [44]: # Fit Multiple Linear Regression
regressor = LinearRegression()
regressor.fit(x_train, y_train)
```

```
Out[44]: 

▼ LinearRegression ⓘ ?
  ► Parameters


```

```
In [45]: # Predicting the test results
y_pred = regressor.predict(x_test)
```

```
In [46]: # Evaluate model
score = r2_score(y_test, y_pred)
print("R2 Score:", score)
```

R² Score: 0.9347068473282423