University of Asia Pacific

Department of Computer Science & Engineering

Project Report

Course Code: CSE 404

Course Title: Artificial Intelligence and Expert Systems Lab

Project Name**: Implementation of Multivariable Linear Regression.**

Submitted By:

Aiatul Al-Amin Ador

ID: 18201019, Section: A

4th Year 1st Semester

April 20, 2022

1. ***Problem Statement***

Implement the Multivariable Linear Regression Using Open-Source Dataset with and without SK-Learn.

Dataset: Wine Dataset

2. ***Multiple Linear Regression***

Multiple Linear Regression is one of the important regression algorithms which models the linear relationship between a single dependent continuous variable and more than one independent variable. The goal of multiple linear regression is to model the linear relationship between the explanatory (independent) variables and response (dependent) variables.

By looking at a linear regression model, analysts can make predictions about the dependent variable based on data gleaned from multiple independent variables

Formula and Calculation of Multiple Linear Regression.

$y_i = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + ... + \beta_p x_{ip} + \epsilon$

where, for i=n observations:

y_i = dependent variable

x_i = explanatory variables

β_0 = y-intercept (constant term)

β_p = slope coefficients for each explanatory variable

ϵ=the model's error term (also known as the residuals)

3. ***Tools And Languages***

- ☐ **Programing Language: Python**
- ☐ **IDE: Jupyter Notebook**

4. ***Linear Regression Without SK-Learn: Output***

Cost Function & Gradient Decent:

```
Iteration number:  120
Hypothesis function value is: h0(x)=theta_0+theta_1 * x
Cost function is: j(theta)=1/(2*m) * i=1_samtionSign_m (h_theta_(x)-y)**2

Cost function is: 0.34341058105000594

gradient decent:
New parameter value is:  [5.153505116837199, 1.599216592292546, 8.002076981208337, 9.1458173945597]
result coefficient is [5.153505116837199, 1.599216592292546, 8.002076981208337, 9.1458173945597]
```

Graph Plotting:

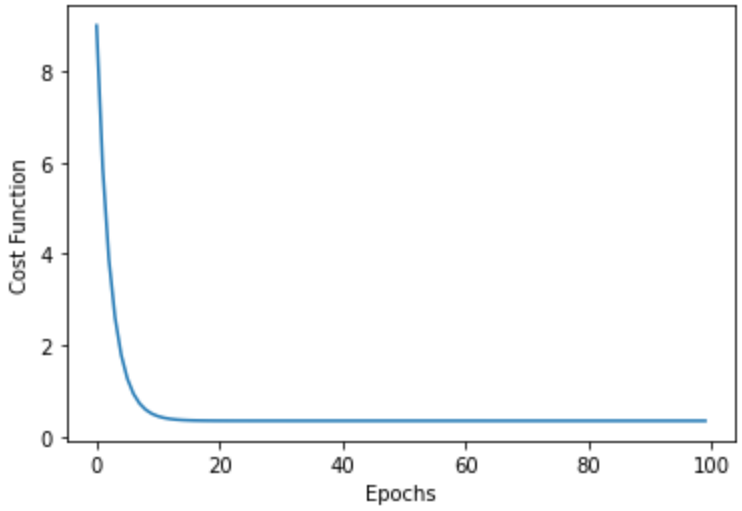
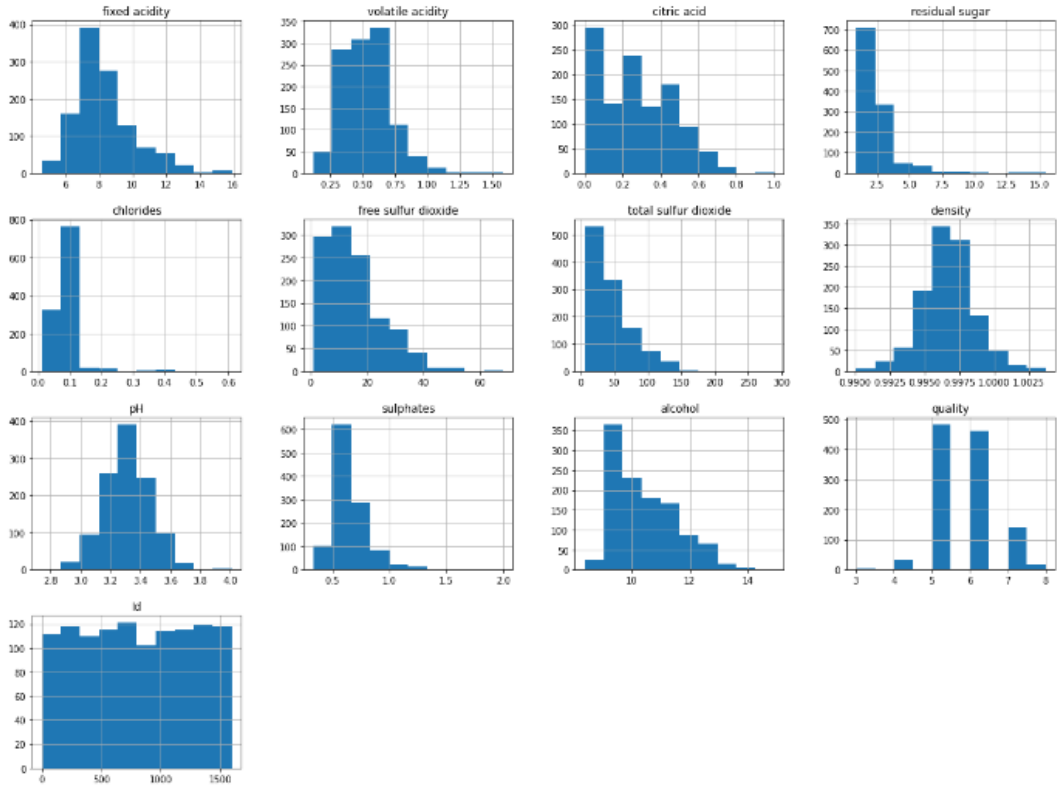
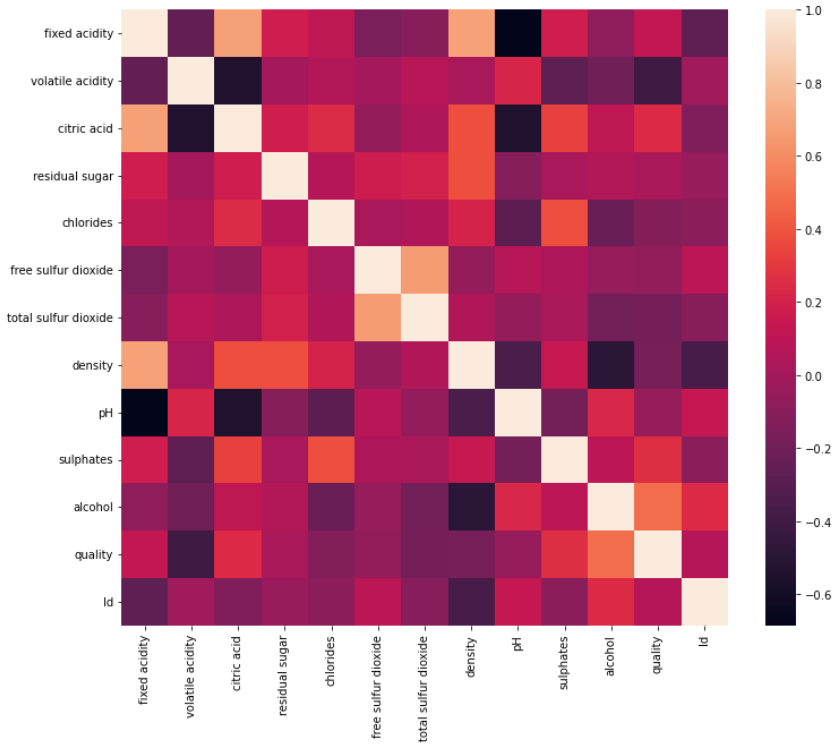


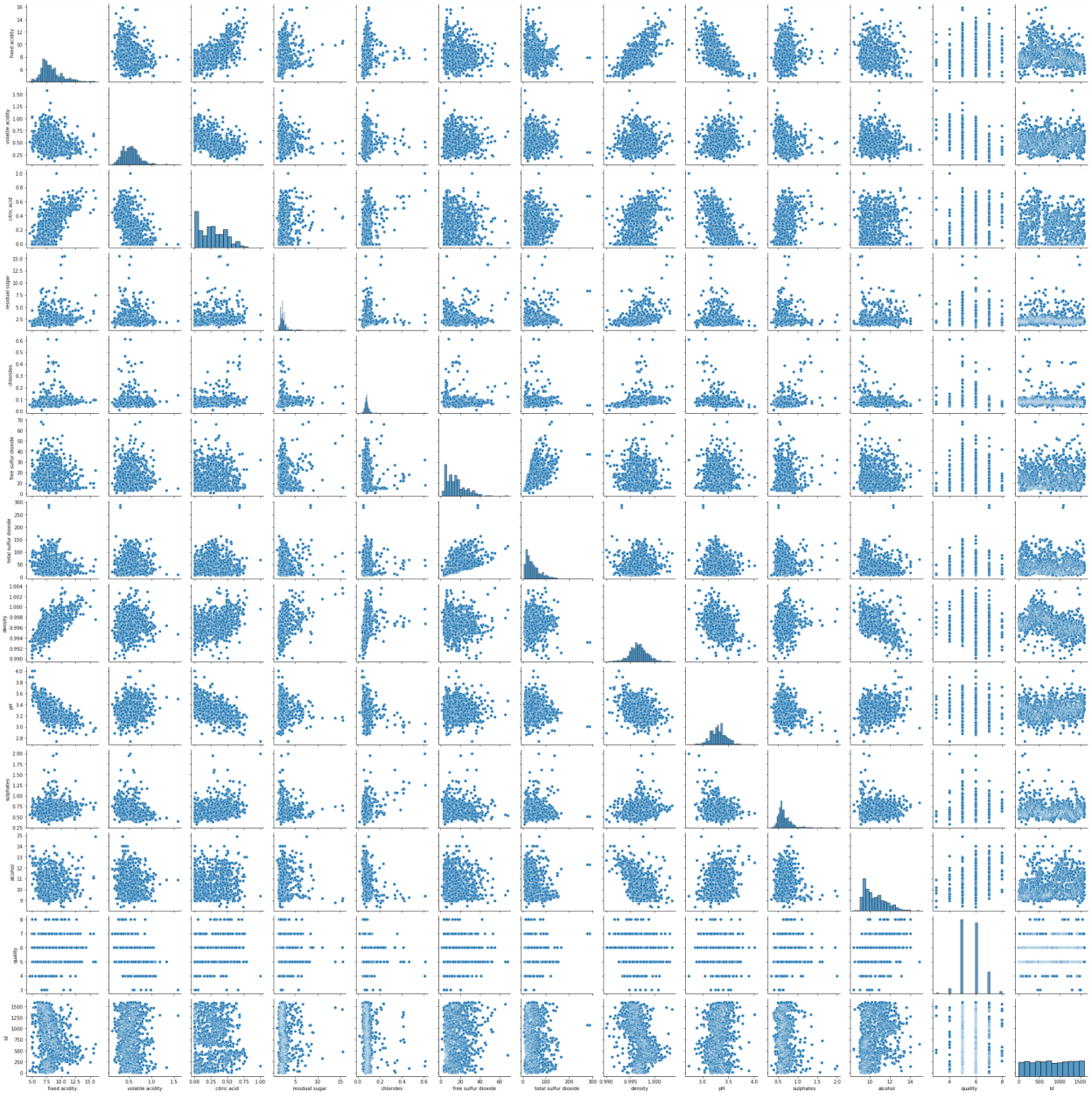
Figure Plotting:



HeatMap:



Pairwise Figure:



5. Linear Regression With SK-Learn: Output

```
print('Coefficients: \n', regr.coef_)
# The Mean Squared Error
print("Mean squared error: %.2f" % np.mean((regr.predict(X_test) - y_test) ** 2))
# Explained Variance Score: 1 is perfect prediction
print('Variance score: %.2f' % regr.score(X_test, y_test))
```

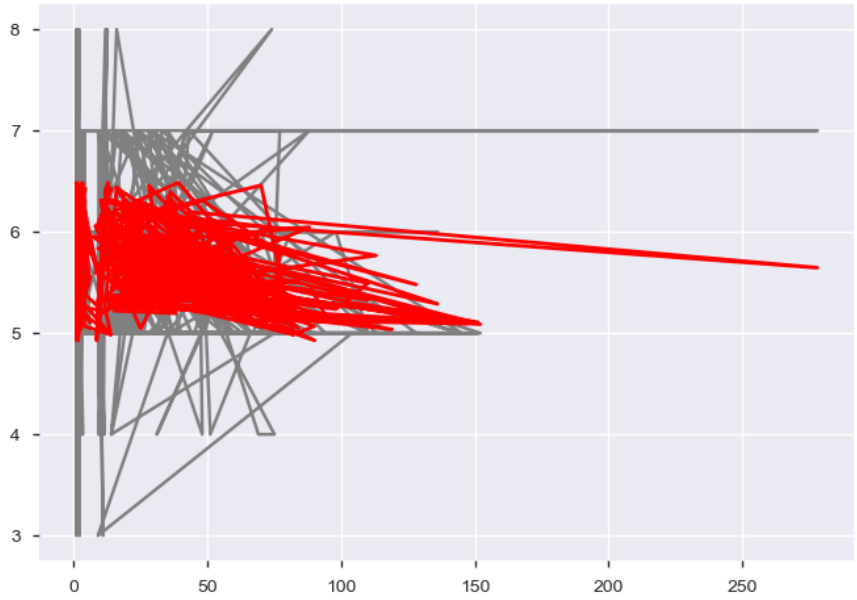
```
Coefficients:
 [-1.70025277e-02 -2.09473413e-03  4.03017510e+01  3.85813074e-01]
Mean squared error: 0.42
Variance score: 0.31
```

```
# Predicting the Test set results
y_pred = regr.predict(X_test)

from sklearn.metrics import r2_score
score=r2_score(y_test,y_pred)
print(score)
```

```
0.31225169543763764
```

6. Graph Plotting:



7. Source Code:

8. Conclusion

Multiple regression is an extension of simple linear regression. In this project, after successful implementation, I have a brief knowledge about linear regression. Multiple linear regression models are useful in helping an enterprise to consider the impact of multiple independent predictors and variables on a dependent variable and can be beneficial for forecasting and predicting results. So, the outcome knowledge of this project will help me a lot in real-life problem predicting and solving in various fields.

-The End-