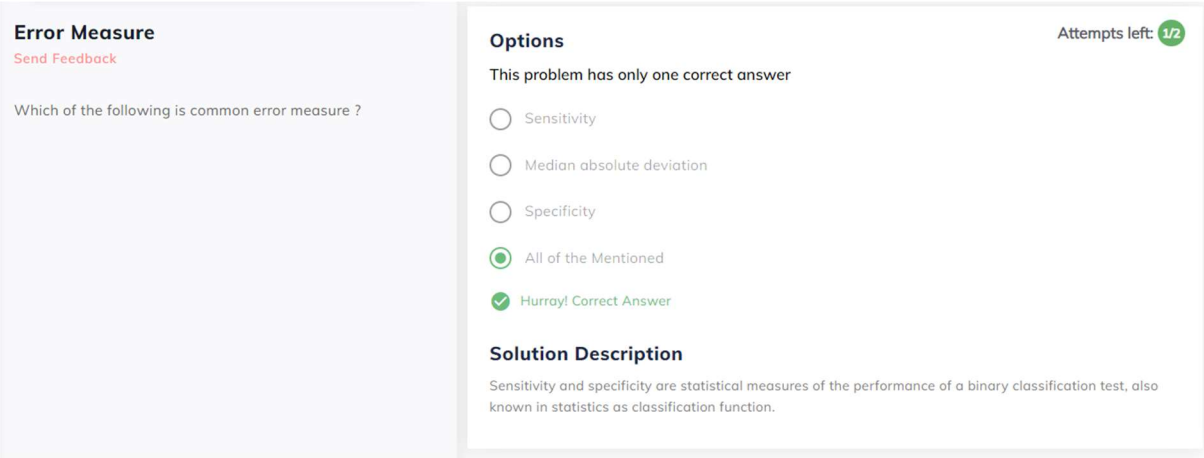
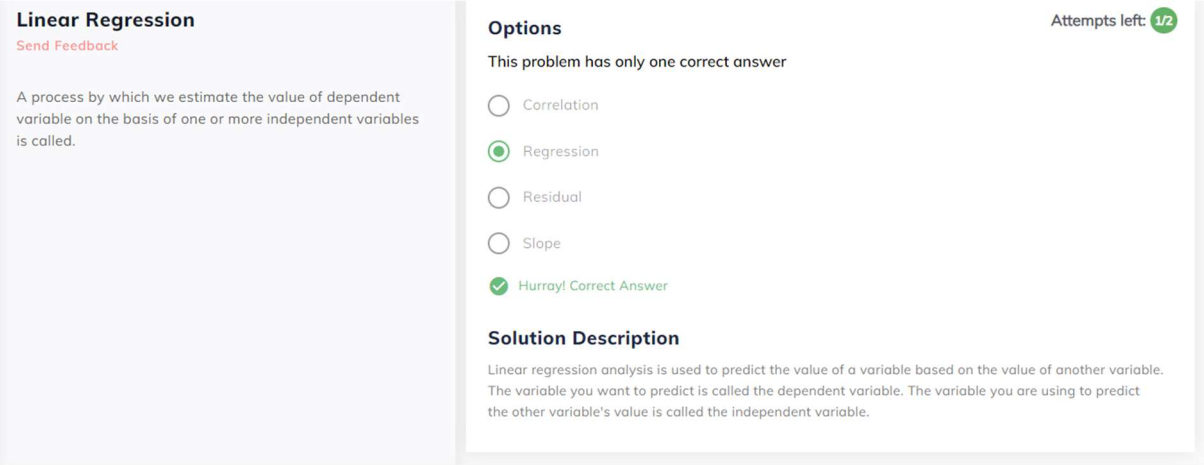
Linear Regression Intuition

MCQ Questions:

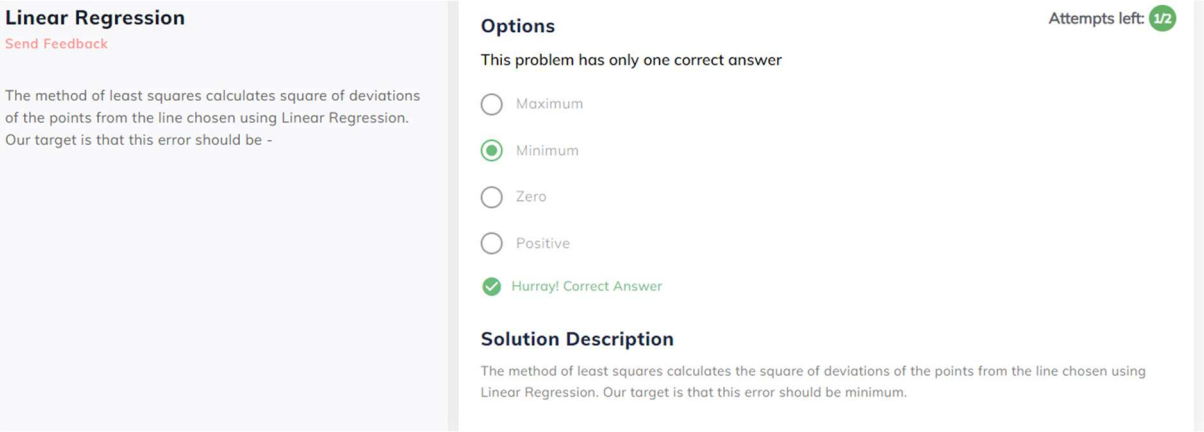
1.



2.



3.



4.

Linear Regression

Send Feedback

All data points falling along a straight line is called:

Options Attempts left: 1/2

This problem has only one correct answer

- (•) Linear relationship
- ( ) Non linear relationship
- ( ) Residual
- ( ) Scatter diagram

Hurray! Correct Answer

Solution Description

All data points falling along a straight line is called Linear relationship.

5.

Linear Regression

Send Feedback

In simple linear regression, the numbers of unknown constants are:

Options Attempts left: 0/2

This problem has only one correct answer

- ( ) One
- (•) Two
- ( ) Three
- ( ) Four

Hurray! Correct Answer

Solution Description

In simple linear regression, the numbers of unknown constants are two.

The equation for simple linear regression is :

$y = mx + c$

where m and c are unknown constants.

6.

Linear Regression

Send Feedback

In simple regression equation, the numbers of variables involved are:

Options Attempts left: 1/2

This problem has only one correct answer

- ( ) 0
- ( ) 1
- (•) 2
- ( ) 3

Hurray! Correct Answer

Solution Description

In simple linear regression, the numbers of variables involved are two.

The equation for simple linear regression is :

$y = mx + c$

where y and x are variables.

7.

Linear Regression

Send Feedback

If the value of any regression coefficient is zero, then two variables are:

Options Attempts left: 1/2

This problem has only one correct answer

- Qualitative
- Correlation
- Dependent
- Independent

Hurray! Correct Answer

Solution Description

If the value of any regression coefficient is zero, then two variables are independent.

8.

Linear Regression

Send Feedback

In Which of the following case , A Straight line slope will be upward given linear equation is Y = a+bX.

Options Attempts left: 1/2

This problem has only one correct answer

- b = 0
- b < 0
- b > 0
- b ≠ 0

Hurray! Correct Answer

Solution Description

A Straight line slope will be upward when the value of b > 0.

9.

Linear Regression

Send Feedback

The straight line graph of the linear equation Y = a + bX, slope will be downward If:

Options Attempts left: 1/2

This problem has only one correct answer

- b > 0
- b < 0
- b = 0
- b ≠ 0

Hurray! Correct Answer

Solution Description

The slope will be downward when b < 0.

Analysis of LR using dummy data

MCQ Questions:

1.

Numpy loadtxt Function

Send Feedback

What all statements are correct about numpy loadtxt() function:

Options Attempts left: 1/2

This problem may have one or more correct answers

- [x] file path is input to this function
- [x] It loads data from text file
- [x] By-default delimiter is space
- [ ] By-default delimiter is comma

Hurray! Correct Answer

Solution Description

1. file path is input to this function.
2. It loads data from the text file
3. By-default delimiter is space

2.

Is Statement Correct

Send Feedback

Consider below code:

```
import numpy as np
data = np.loadtxt("data.csv",delimeter=",")
X = data[:,0]
Y = data[:,1]
```

Statement: X has all the rows of 0th Col and Y has all the rows of 1st Col

Options Attempts left: 1/2

This problem has only one correct answer

- (•) Statement is Correct
- ( ) Statement is Incorrect
- ( ) It will produce error

Hurray! Correct Answer

Solution Description

With this line, X = data[:,0]
X will have all the rows of 0th Col

With this line, Y = data[:,1]
Y will have all the rows of 1st Col

That's why the statement is true

3.

Numpy Shape Function

Send Feedback

Consider below code:

```
import numpy as np
X = np.array([1,2,3,4])
X.shape
```

What will be the output ?

Options Attempts left: 1/2

This problem has only one correct answer

- (•) (4,)
- ( ) (,4)
- ( ) (4,4)

Hurray! Correct Answer

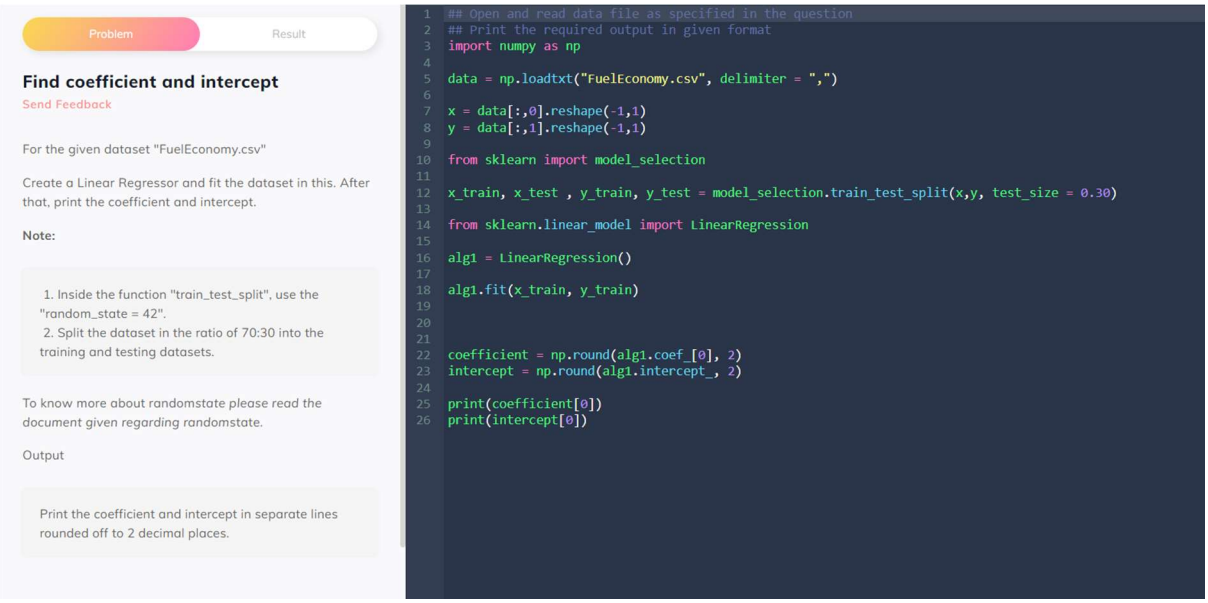
Solution Description

np.array([1,2,3,4]) is a one dimensional array. The shape of this is (4,).

Read : Using Random State Notes

Question: 1

Find coefficient and intercept



Coefficient of Determination

Question: 2

Find the testing and training score

Cost Function

MCQ Questions:

1.

Objective Way to Define Best Fit Line

Send Feedback

What is the best objective way to define the best fit line ?

Options Attempts left: 1/2

This problem has only one correct answer

- (•) Minimization of Error function
- ( ) Minimization of Sum function
- ( ) Minimization of Multiplication Function

Hurray! Correct Answer

Solution Description

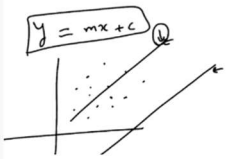
The best objective way to define the best fit line is to the Minimization of the Error function.

2.

Error in Line

Send Feedback

Which line has more error as taught in video:



Options Attempts left: 1/2

This problem has only one correct answer

- ( ) Upper Line
- (•) Lower Line

Hurray! Correct Answer

Solution Description

The lower line has more error, as most of the data points are far from the lower line than from the upper line.

3.

Mean Squared Error

Send Feedback

For simple linear equation y=mx+b we can calculate Mean Square Error as:

$$MSE = \frac{1}{N}\sum_{i=1}^{n}(y_i - (mx_i + b))^2$$

Where

- N is the total number of observations (data points)
- $\frac{1}{N}\sum_{i=1}^{n}$ is the mean
- y_i is the actual value of an observation and $mx_i + b$ is our prediction

Options Attempts left: 1/2

This problem has only one correct answer

- (•) True
- ( ) False

Hurray! Correct Answer

Solution Description

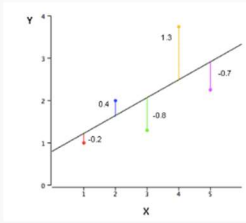
The formula described here is for Mean Square Error.

4.

Sum of Squared Error

Send Feedback

The graph below represents a regression line predicting Y from X. The values on the graph shows the residuals for each predictions value. Use this information to compute the Sum of Squared Error.



Options Attempts left: 0/2

This problem has only one correct answer

- (•) 3.02
- ( ) 0.75
- ( ) 1.01
- ( ) None of these

Hurray! Correct Answer

Solution Description

SSE is the sum of the squared errors of prediction, so SSE = (-.2)^2 + (.4)^2 + (-.8)^2 + (1.3)^2 + (-.7)^2 = 3.02

5.

M and C of Linear Regression

Send Feedback

Choose the correct equation for m and c once you derive m and c of linear regression:

Equation-1:

```
y.mean() - m* x.mean()
```

Equation-2:

```
(x*y).mean() - x.mean()*y.mean()
---------------------------------
(x*x).mean() - x.mean()*x.mean()
```

Note: mean() function in above equation is used to calculate the mean of value, y.mean() will give you the mean of y(y being 1D array) and similarly x.mean() will give you mean of x(x being 1D array)

Options Attempts left: 1/2

This problem has only one correct answer

- (•) Equation 1 represent c and Equation 2 represent m
- ( ) Equation 1 represent m and Equation 2 represent c

Hurray! Correct Answer

Solution Description

Equation 1 represent c and Equation 2 represent m

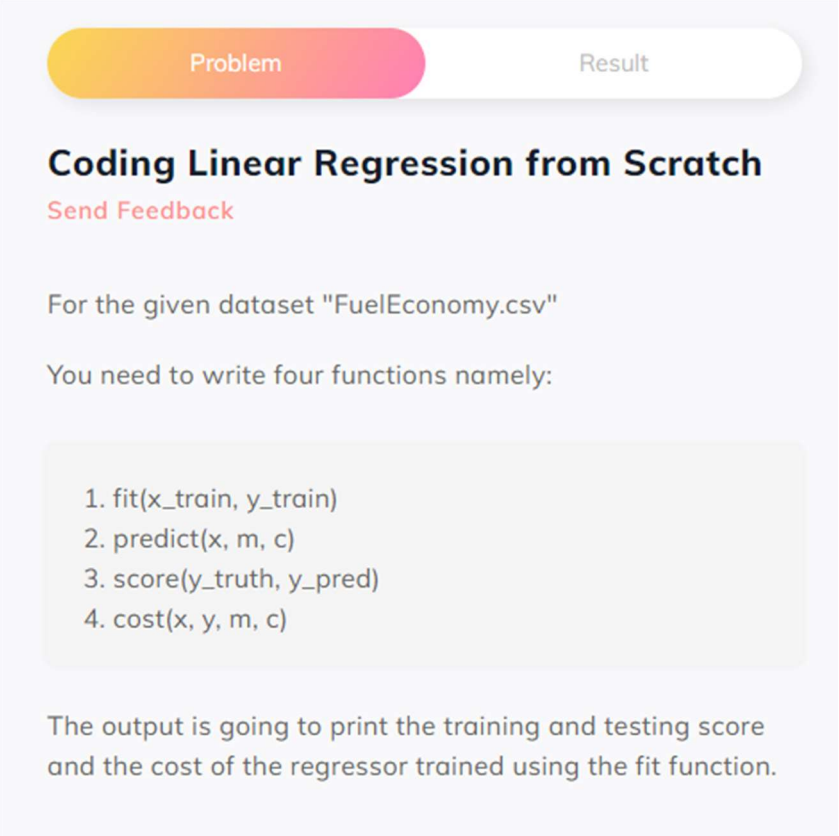
Optimal Coefficients:

--> "please refer to notes"

Linear Regression from Scratch

Question: 3

Coding Linear Regression from Scratch



Source Code:

```
## Open and read data file as specified in the question
## Print the required output in given format
import numpy as np
from sklearn.model_selection import train_test_split

data = np.loadtxt('FuelEconomy.csv', delimiter=',')
X = data[:, 0]
y = data[:, 1]

X_train, X_test, Y_train, Y_test = train_test_split(X, y, test_size = 0.3, random_state=42)
```

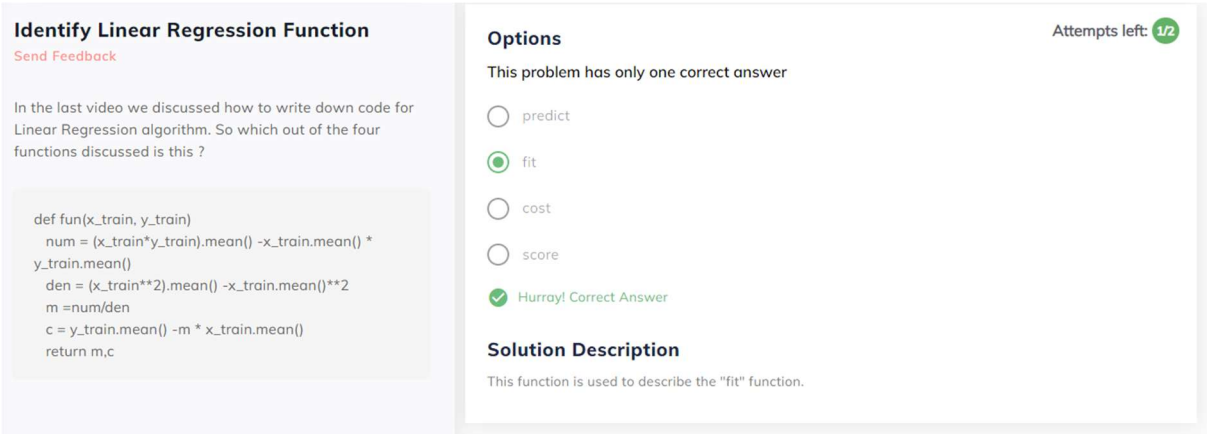
```
############################
##### Start from here #####
############################

def fit(x_train, y_train):
    ## Write code for fit function
    n = len(x_train)

    # Calculate the mean of x and y
    mean_x = np.mean(x_train)
    mean_y = np.mean(y_train)

    # Calculate the slope (m) and intercept (c) of the regression line
    numerator = np.sum((x_train - mean_x) * (y_train - mean_y))
    denominator = np.sum((x_train - mean_x) ** 2)

    m = numerator / denominator
    c = mean_y - m * mean_x

    return m, c



def predict(x, m, c):
    ## Write code for predict function
    # Predict the values using the regression line equation: y = mx + c
    return m * x + c



def score(y_truth, y_pred):
    ## Write code for score function
    # Calculate the R-squared score
    mean_y = np.mean(y_truth)
    ss_total = np.sum((y_truth - mean_y) ** 2)
```

```
    ss_residual = np.sum((y_truth - y_pred) ** 2)

    r_squared = 1 - (ss_residual / ss_total)
    return r_squared




def cost(x, y, m, c):
    ## Write code for cost function
    # Calculate the mean squared error
    y_pred = predict(x, m, c)
    mse = np.mean((y - y_pred) ** 2)
    return mse



############################
############################

m, c = fit(X_train, Y_train)

y_test_pred = predict(X_test, m, c)
y_train_pred = predict(X_train, m, c)

print(round(score(Y_test, y_test_pred), 2))
print(round(score(Y_train, y_train_pred), 2))
print(round(cost(X_train,Y_train, m, c ), 2))
```

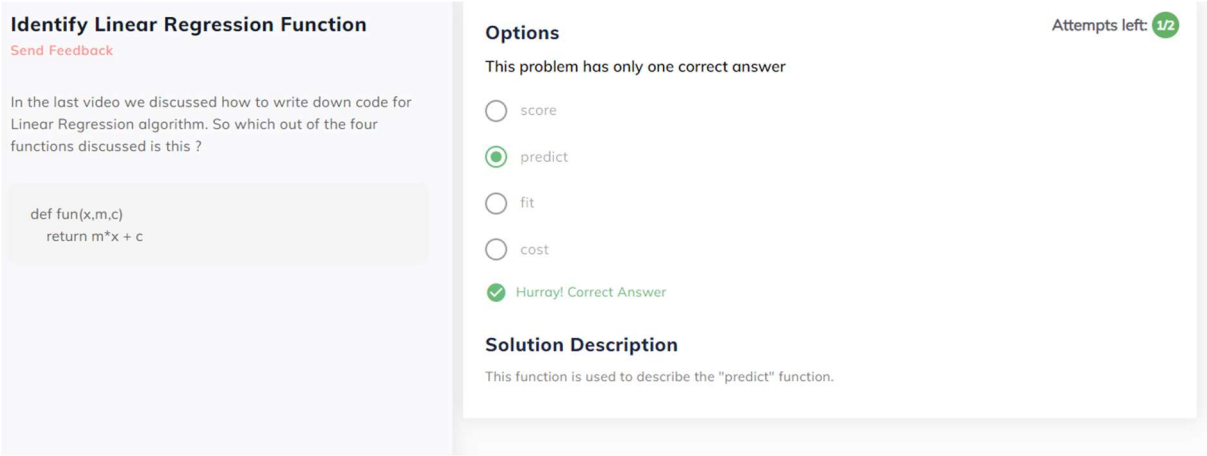
Linear Regression from Scratch:

MCQ Questions:

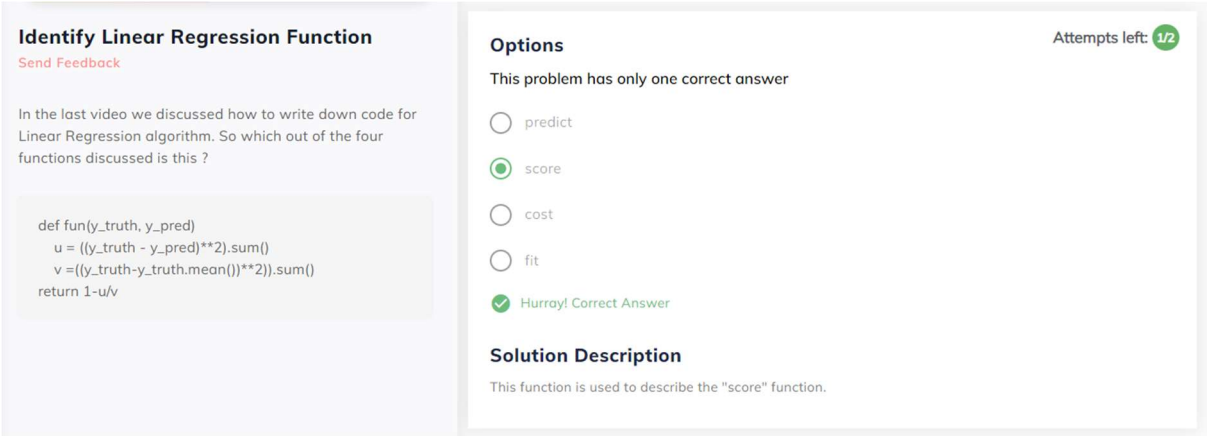
1.



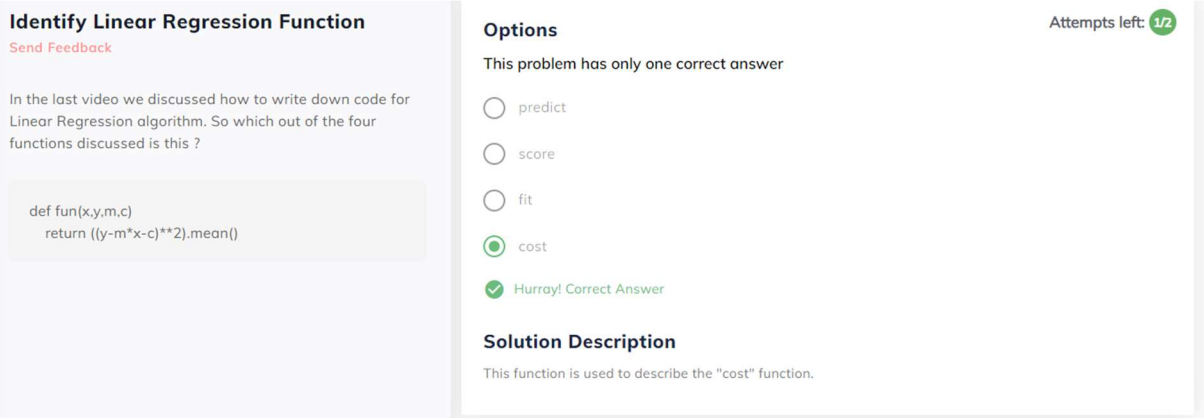
2.



3.



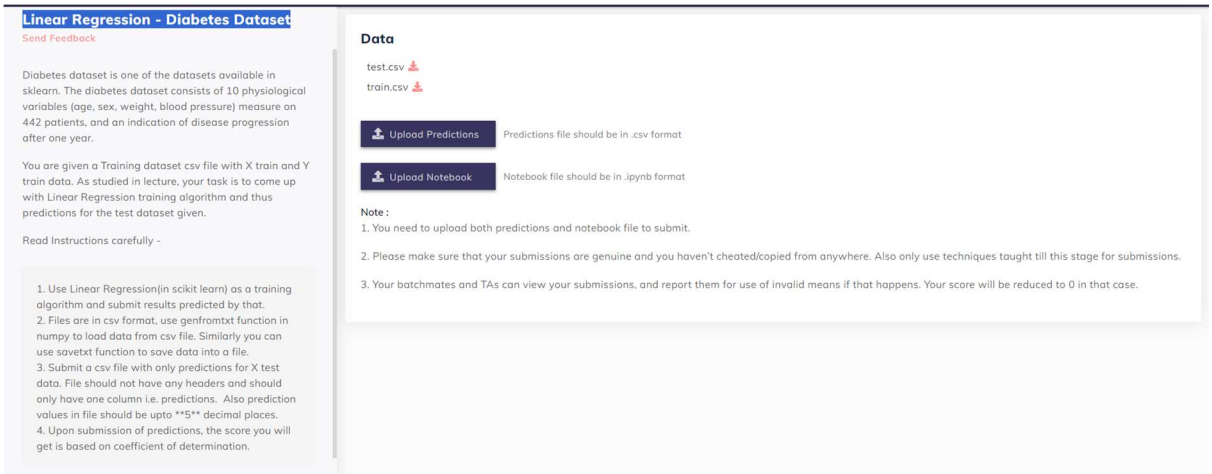
4.



Read: Linear Regression Notes

Assignment

Linear Regression - Diabetes Dataset



Source code: → check the repo..