University of Haripur.

Assignment #2.

Machine Learning

Title: Linear Regression.

To: Ms. Nadia

From: Hamza Sadaat.

Depart: Information Technology.

Program: Computer Science.

Roll No: F18-0501.

Semester: 7th 'B'.

Dated: 06/12/2021.

Importing Libraries and reading dataset

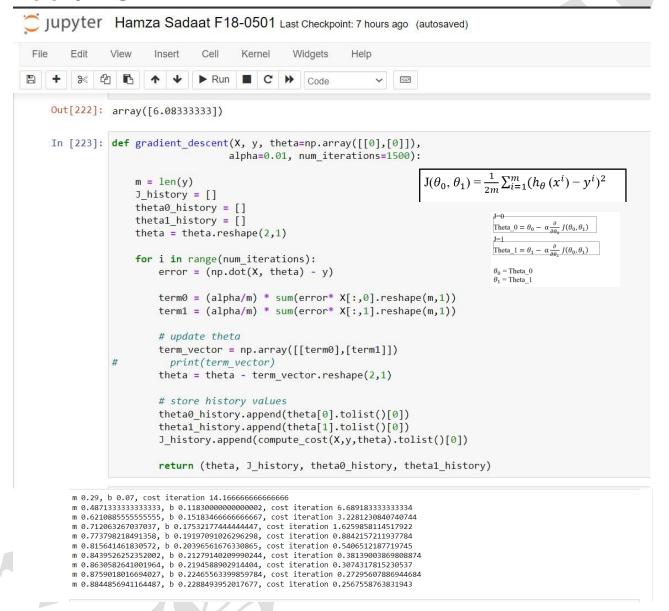


Creating function to calculate hypothesis:

Jupyter Hamza Sadaat F18-0501 Last Checkpoint: 7 hours ago (autosaved)

```
File
                                  Cell
       Edit
               View
                                          Kernel
                                                    Widgets
                                                                Help
                        Insert
                                 ► Run
                                         ■ C → Code
                                                                        F......
   In [220]: n_rows = data_df.shape[0]
   In [221]: X=data_df['X'].to_numpy().reshape(n_rows,1)
                ones = np.ones((n_rows,1))
               X = np.concatenate((ones, X), axis=1)
y=data_df['y'].to_numpy().reshape(n_rows,1)
                def compute_cost(X, y, theta=np.array([[0],[0]])):
                    m = len(y)
                    # initialize loss to zero
                    J=0
                                                         h_{\theta}(\mathbf{x}) = \theta_0 + \theta_1 \mathbf{x}
                    # reshape theta
                    theta=theta.reshape(2,1)
                    # calculate the hypothesis
                    h_x = np.dot(X,theta)
                    error_term = sum((h_x - y)**2)
                    loss = error_term/(2*m)
                    return loss
   In [222]: compute_cost(X,y)
   Out[222]: array([6.08333333])
```

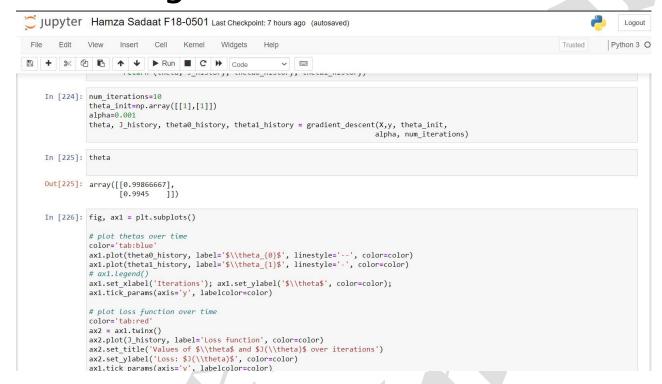
Applying Formulas



Repeat the process up to, nth iteration until you get the best fit Hypothesis.

- 4. At first kept the Learning Rate = 0.001 and check the classifier using different learning Rates.
- 5. Output clearly shows the values of theta at every iteration along with the error rate running code for 10 iterations, output is showing the updated values of theta for each 10 iterations along with the error it got at every iteration.

Visualizing the data:



Data Visualized

Jupyter Hamza Sadaat F18-0501 Last Checkpoint: 8 hours ago (unsaved changes)

```
Cell
File
       Edit
              View
                      Insert
                                       Kernel
                                                Widgets
                                                           Help
        % € 6
                     1
                               ▶ Run ■ C → Code
              ax2.plot(J_history, label='Loss function', color=color)
               ax2.set_title('Values of $\\theta$ and $J(\\theta)$ over iterations')
               ax2.set_ylabel('Loss: $J(\\theta)$', color=color)
               ax1.tick_params(axis='y', labelcolor=color)
               # ax2.legend();
               fig.legend();
                          Estimated coefficients:
                          theta_0 = 1.1428571428571428
                          theta_1 = 0.7428571428571429
                            5.0
                            4.5
                            4.0
                            3.5
                          > 3.0
                            2.5
                            2.0
                            1.5
```

Iterations