

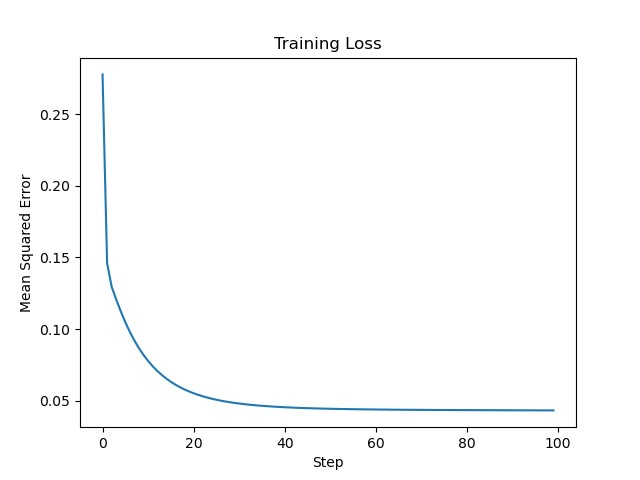
|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | | | | | | | | |
|  | | CSE 6363: MACHINE LEARNING ASSIGNMENT - 1 | | | | |  | |
|  |  | | | | | | |  |
|  | | | |  |  | | | |
|  | | | | Sanjana ReddyThigulla 1002142811 |  | | | |
|  | | | | 09/15/2023—Fall 2023—Alex J Dillhoff |  | | | |
|  | | |  | | |  | | |

1. Implementation of the LinearRegression class
   1. The fit method: the data has been divided into train set and validation set (1:10 ratio). Weights and biases have been initialized (set to zero). After gradient calculation based on the validation predictions validation loss is calculated and the minimal loss parameters are constantly updated based on the given patience = 3.
   2. The predict method: the predict function is implemented.
   3. The score method: the mean square error is found using “mse = m \* np.mean((predictions - y) \*\* 2)”.
   4. Saving and Loading Weights: save and load methods are updated to save and load parameters to and from a file.
   5. Regression with a single output:

Model 1:

Using Petal Length and Sepal Width as input features and

Using Petal Width as the output feature

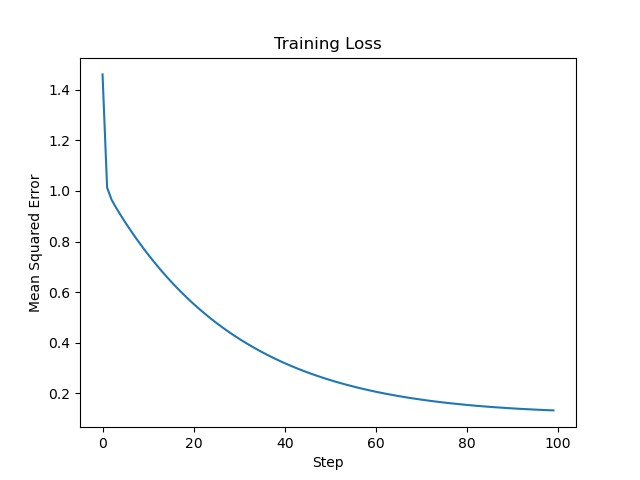


Mean Squared Error on Test Data (Model 1): 0.0571

Model 2:

Using Petal Width and Sepal Length as input features and

Using Petal Length as the output feature

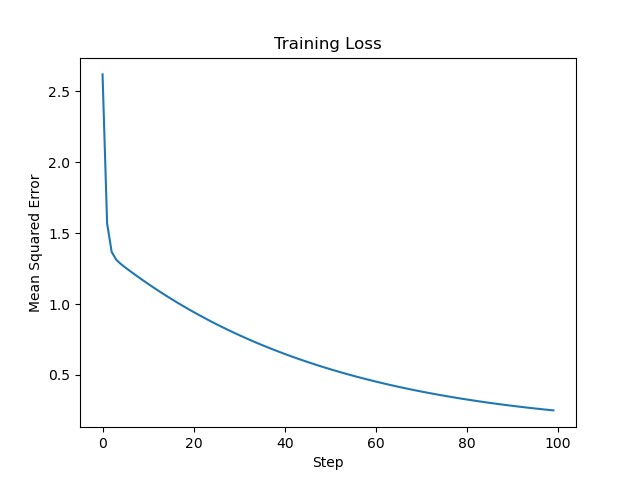


Mean Squared Error on Test Data (Model 2): 0.2055

Model 3:

Using Petal Length and Petal Width as input features and

Using Sepal Width as the output feature

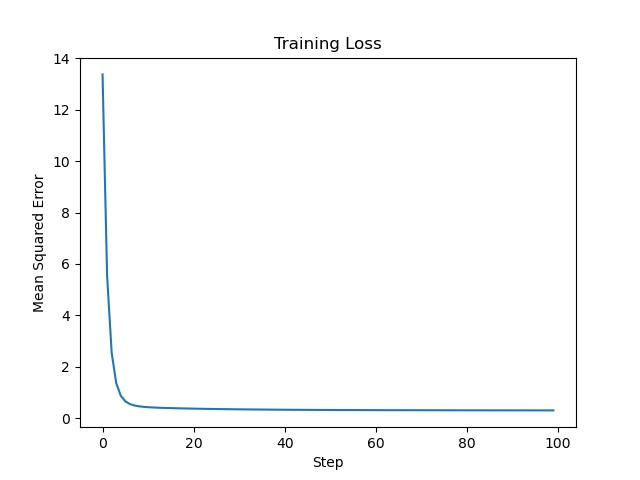


Mean Squared Error on Test Data (Model 3): 0.5251

Model 4:

Using Sepal Width and Petal Width as input feature

Using Sepal Length as the output feature



Mean Squared Error on Test Data (Model 4): 0.3633

Better prediction:

As depicted above, the first model shows least mean square error of 0.0571. It uses Petal Length and Sepal Width as input features and uses Petal Width as the target feature.

Though MSE isn’t the only factor to determine whether an input feature is the most predictive of an output feature, it can be considered as one of the important factors. Hence from the results it can be concluded that Petal length and Sepal Width are the most predictive input features for the target feature petal width.

Regularization:

After this l2 regularization is applied on model 1 and the difference in parameters with and without regularization has been recorded as follows

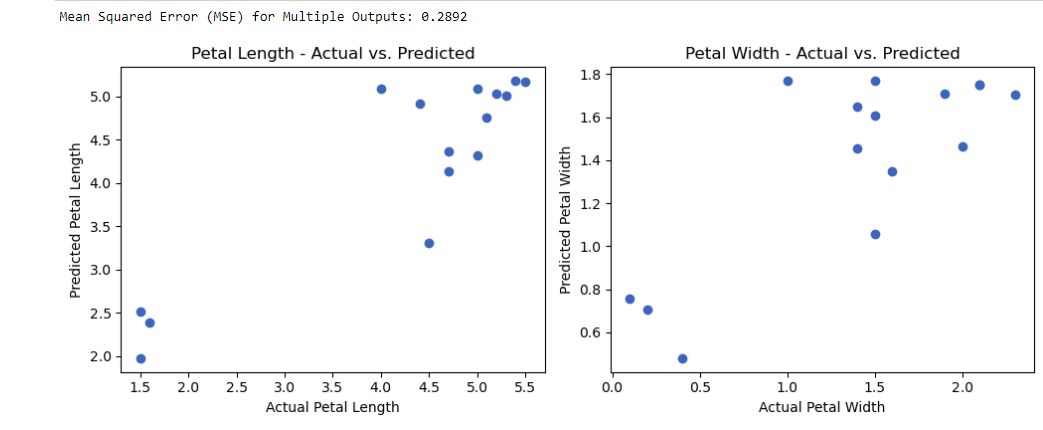
Weight Differences : [0.08811735 0.08430523]

Bias Difference : 0.059734542473980946

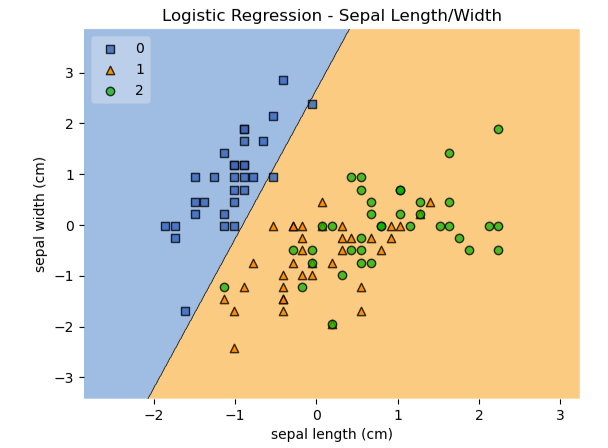
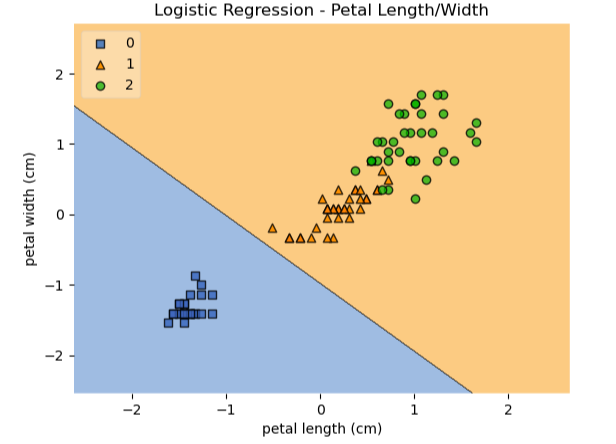
* 1. Regression with Multiple Outputs:

A single model that predicts the petal length and petal width given the sepal length and width is created similar to the previous single output linear regression models. MSE has been calculated using

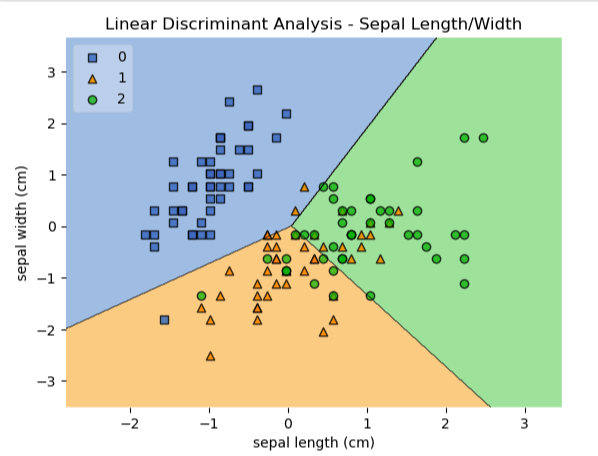
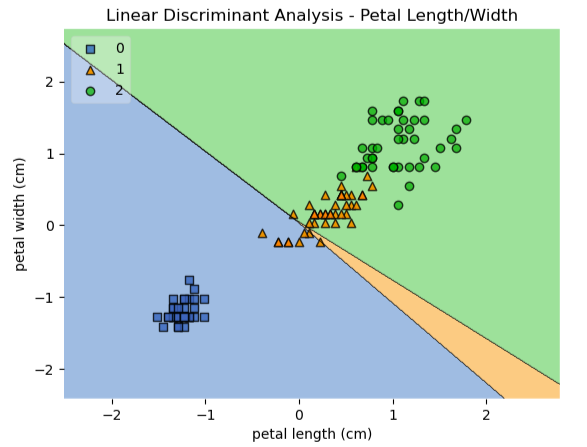
“mse = (1/(n\*m))\*np.sum((predictions - y) \*\* 2)” and 2 diff plots have been plotted for the actual vs. predicted values for both petal length and petal width.



1. Classification
   1. Logistic Regression: logistic regression has been implemented on 3 models and the 2 plots have been displayed.

* 1. Linear Discriminant Analysis: LDA has been implemented on 3 models and the 2 plots have been displayed.

* 1. Testing: 3 evaluation scripts for each model have been created where the parameters are loaded and the accuracy score is found.