# CSE 4020 - MACHINE LEARNING

# Lab 29+30

# Lab Task4(Logistic Regression)

# Submitted by: Alokam Nikhitha(19BCE2555)

**Question:**

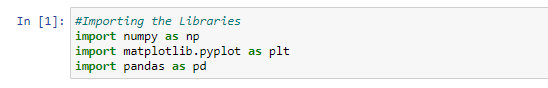
Train a Logistic regression model to predict the charges is there or not for insurance company from given features.

**Dataset Used:** “insurance.csv” as provided.

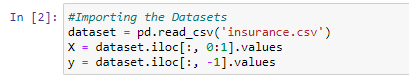
**Procedure:**

* We first import the dataset into our workspace the use of pandas.
* We then need to determine at the impartial and based attributed for use in our regression version.
* We then need to initialize our Linear regression version and logistic regression and match it to the X and y attributes.
* Next, we need to create any other variable to save the outcomes of X set as anticipated with the aid of using our regression version.
* We then can discover the scatter plot of our units and the exceptional match Regression line.
* Finally, we calculate our assessment metrics to test the accuracy of our version.

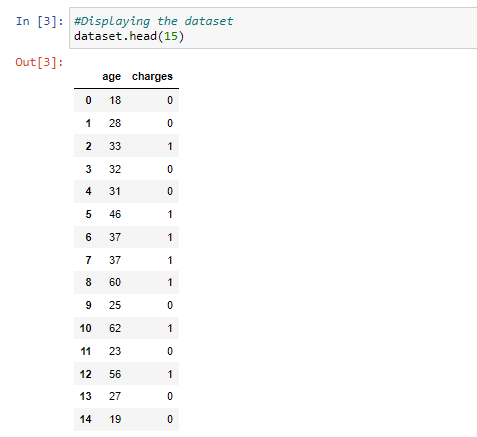
**Code Snippets and Explanation:**



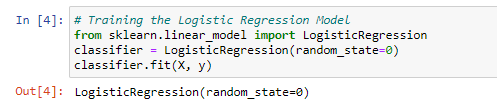
Here we are importing the libraries.



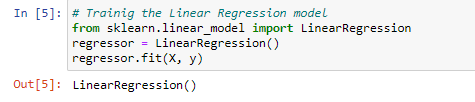
We're importing the dataset into our workspace and naming the set of independent attributes X and the set of dependent attributes y.



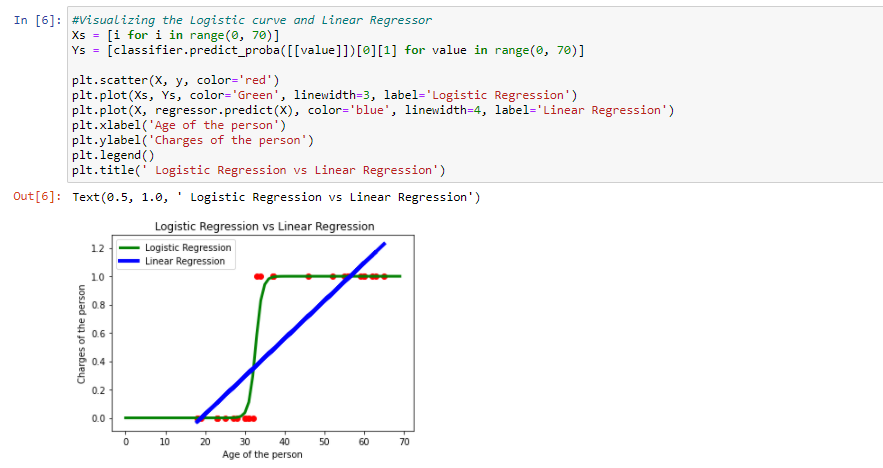
The first 15 rows of our dataset are seen here. The charges attribute is a categorical attribute, as we can see, with ‘No' labelled as 0 and ‘Yes' tagged as 1. With values ranging from 18 to 62, the age property is both continuous and discrete.



Here we have trained our Logistic regression model with X and y set.

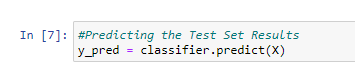


Here we have trained our linear regression model with X and y set.

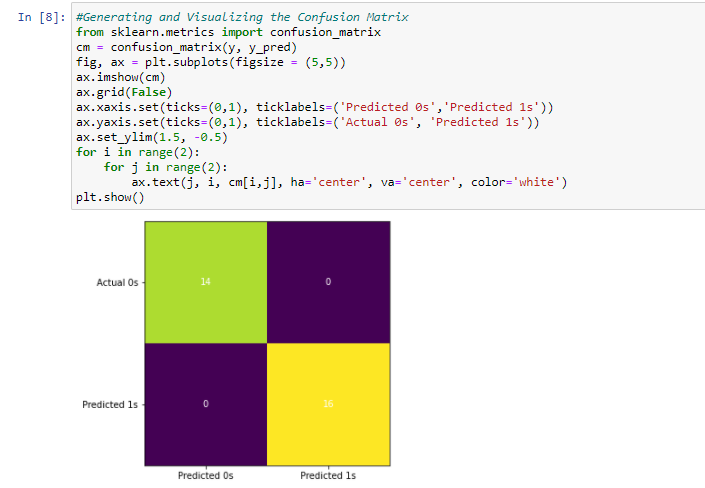


The probability line of our logistic regression is presented here. Any value more than 0.5 is defined as "charged," while any value less than 0.5 is classified as "no charge." We have also plotted the values as predicted by linear regressor.

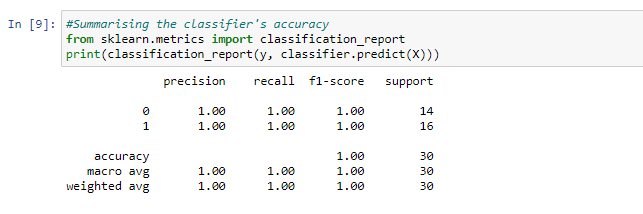
The Logistic Regression is plotted in Green color and Linear Regression in Blue color.



We create another vector to store the result of our X set as predicted by the trained classifier.



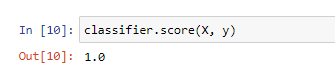
Here, we have visualised our predicted results and actual results. We can clearly see that the true positives and true negatives account for complete results and thus our logistic regression classifier is 100% accurate, that is, actual set is same as the predicted set.



Here we have analysed the performance of our logistic regression classifier.

As we can see, the precision and recall for both 1 and 0 accounts for 1.00, that is all the values of 0 and 1 were identified correctly.

Also, the accuracy of out model is 1.00 (100%) as macro average and weighted average are both 1.00.



Finally, the classifier score is also 1 as all the values were correctly identified.

**Results and Conclusion:**

1. Identified ‘no charge’ = 14
2. True ‘no charge’ = 14
3. Identified ‘charged’ = 16
4. True ‘charged’ = 16
5. Precision of ‘no charge’ = 1.00
6. Precision of ‘charged’ = 1.00
7. Recall of ‘no charge’ = 1.00
8. Recall of ‘charged’ = 1.00
9. Model Accuracy = 100%

* Logistic Regression vs Linear Regression

