Assignment #1 **Possible Points: 100**
CS-770 ML **Due date: 12th October 2022**

Name: Hemil Mehul Shah

Email id:
hxshah4@shockers.wichita.edu

Assignment should be done on individual basis.

Q1 Fit a predictive linear regression model to estimate weight of the fish from its length, height and width? (the data source fish.csv can be found here: https://www.kaggle.com/aungpyaeap/fish-market) (50 points)

-Report the coefficients values by using the standard Least Square Estimates
Length1: 234.9989
Length2: -123.6765
Length3: -85.2295
Height: 66.6378
Width: -35.0027
-What is the standard error of the estimated coefficients, R-squared term, and the 95% confidence interval?
Length1: 74.062
Length2: 78.440
Length3: 30.355
Height: 41.9060
Width: -10.2263
-Is there any dependence between the length and weight of the fish?
The weight impacts the vertical length, but does not have much of an impact on the diagonal and cross length.

Summary of Results:

Dep. Variable:	Weight	**R-squared (uncentered):**	0.871
Model:	OLS	**Adj. R-squared (uncentered):**	0.865
Method:	Least Squares	**F-statistic:**	143.1
Date:	Wed, 12 Oct 2022	**Prob (F-statistic):**	1.82e-45
Time:	22:02:45	**Log-Likelihood:**	-734.65
No. Observations:	111	**AIC:**	1479.
Df Residuals:	106	**BIC:**	1493.
Df Model:	5		
Covariance Type:	nonrobust		

	coef	std err	t	P>|t|	[0.025	0.975]
Length1	234.9989	76.799	3.060	0.003	82.737	387.261
Length2	-123.6765	78.054	-1.585	0.116	-278.425	31.072
Length3	-85.2295	29.452	-2.894	0.005	-143.622	-26.837
Height	66.6378	15.254	4.369	0.000	36.395	96.880
Width	-35.0027	36.160	-0.968	0.335	-106.694	36.688

Omnibus:	49.876	**Durbin-Watson:**	2.077
Prob(Omnibus):	0.000	**Jarque-Bera (JB):**	147.805
Skew:	1.644	**Prob(JB):**	8.03e-33
Kurtosis:	7.599	**Cond. No.**	324.

Q2 Using the data source in Q1 fit the Ridge and Lasso Regression Models. (25 points)

- Report the coefficients for both the models

Ridge Regression Coefficients:

```
[ 56.04223714  -1.09488211 -28.48041569  27.73300025  22.55304673]
```

Lasso Regression Coefficients:

```
[ 58.67015974  -6.03941838 -26.23168025  26.90224938  24.58239574]
```

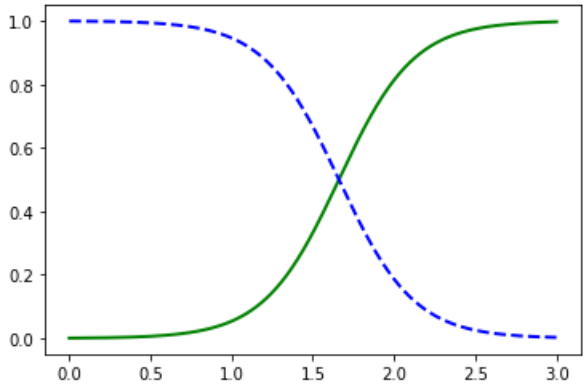
- Report the attribute(s) least impacting the weight of the fish.

Cross and Diagonal length least impact the weight of the fish

Q3 Modify the example code for Logistic Regression to include all the four attributes in iris dataset for two class and multi-class classification. (25 points)

In the training data set: use X = iris[“data”]
In the testing: use:
x0, x1, x2, x3 = np.meshgrid(
np.linspace(4,8,9).reshape(-1,1),
np.linspace(2,5,7).reshape(-1,1),
np.linspace(2.9,7,10).reshapre(-1,1))
x_new = numpy.c[x0.ravel(), x1.ravel(), x2.ravel(), x3.ravel()]
y_proba = log_reg.predict_proba(X_new)

The original Iris data graph is as follows:



The changed Iris data graph is as follows: