Capstone Project

Abalone Age Classification

By Ramalee Rodrigo

Highlights

- Introduction
- Objectives
- Descriptive Analysis
- Final Findings
- Advanced Analysis
- Final Model
- Data Product
- Problems Faced
- References

What is an Abalone?

A sea snail

Can be found along the cold coastal waters of every continent which has a worldwide distribution. There are 7 species on the West Coast, and about 60-100 species recognized globally.

Why is it popular?

"The Best and Rarest Seafood In The World"

About the dataset

Source: Study: named "The Population Biology of Abalone (Haliotis species) in Tasmania. I. Blacklip Abalone (H. rubra) from the North Coast and Islands of Bass Strait". By the authors, Warwick J Nash, Tracy L Sellers, Simon R Talbot, Andrew J Cawthorn and Wes B Ford

About: lab is measurements of Blacklip abalone collected from the Bass Strait, which separates Tasmania from mainland Australia.

Variables: Categorical -1 , Numerical - 8

Observations: 4177

Objective: Abalone age Prediction by using Physical measurements.

Pre-processing:

Missing values: Height 2 records Removed (Anomaly)

Duplicates: None

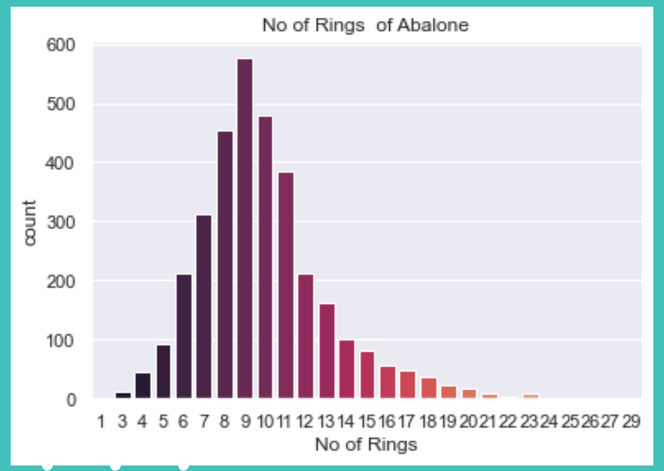
Recategorization and New Variable:

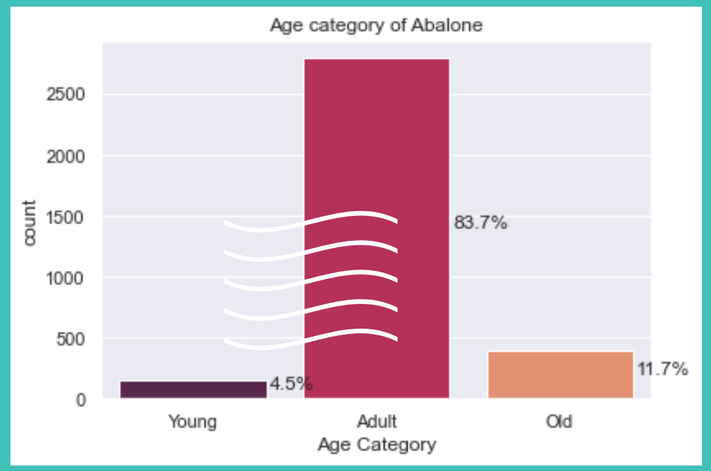
Ring variable → Age, Age_Category

Shucked Weight, Viscera weight,

Shell weight, Whole weight → weight_diff

Rings and Age





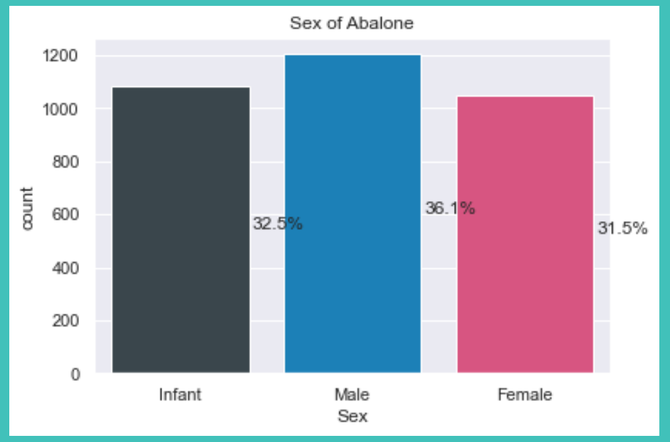
"the researchers believed adding 1.5 to the ring count is a reasonable approximation of the abalones age."

- Econometric Ways to Estimate the Age and Price of Abalone | MPRA Paper No. 91210

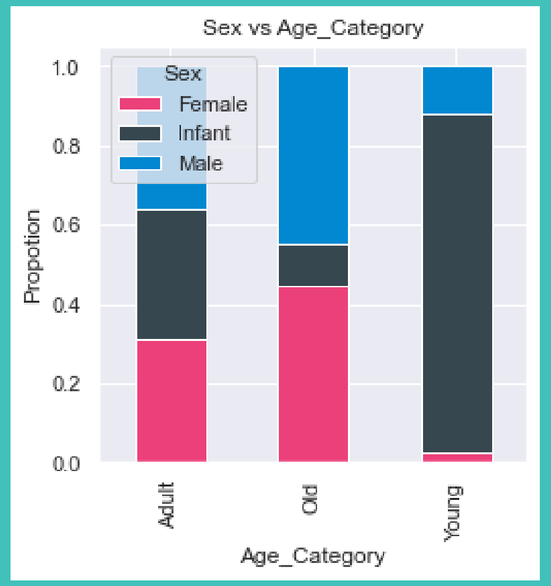
Age = Rings + 1.5yrs

1-5=Young | 6-13=Adult | 14-30=Old

Sex – Gender of Abalone



For every age there are abalones from each gender.

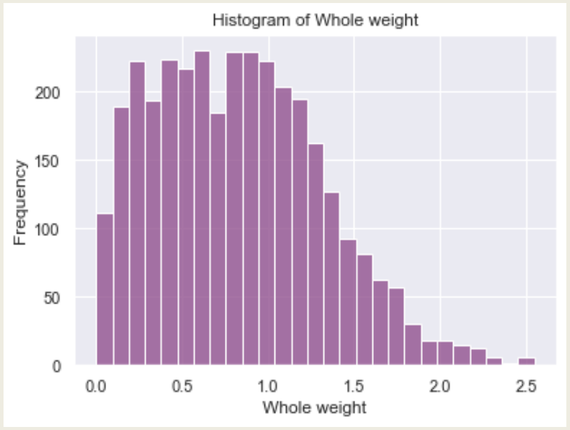


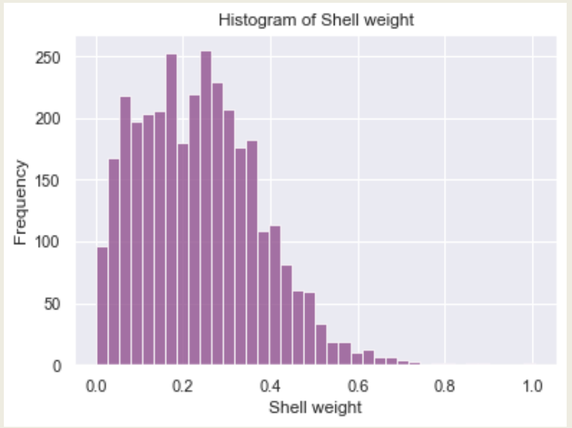
Infant ≠ young

But majority of the Infants are young

"Abalone were considered to be sexually differentiated if gonad (either testis or ovary) could be discerned." -original Study

Weight Measures (dg)



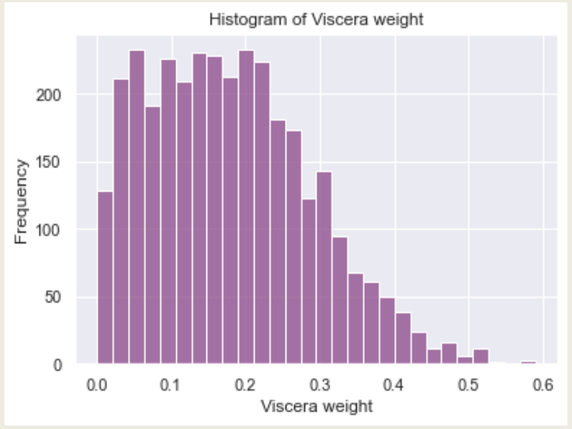


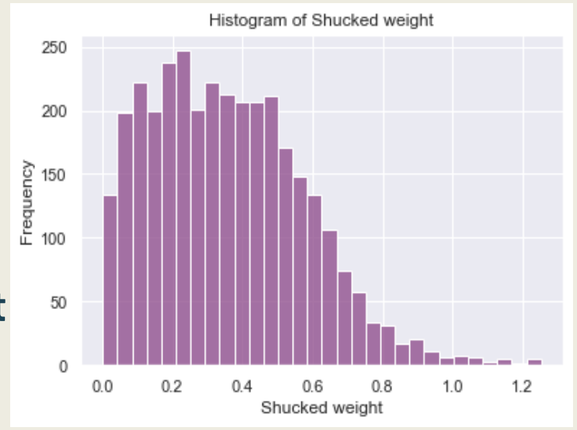
Whole weight

Viscera weight

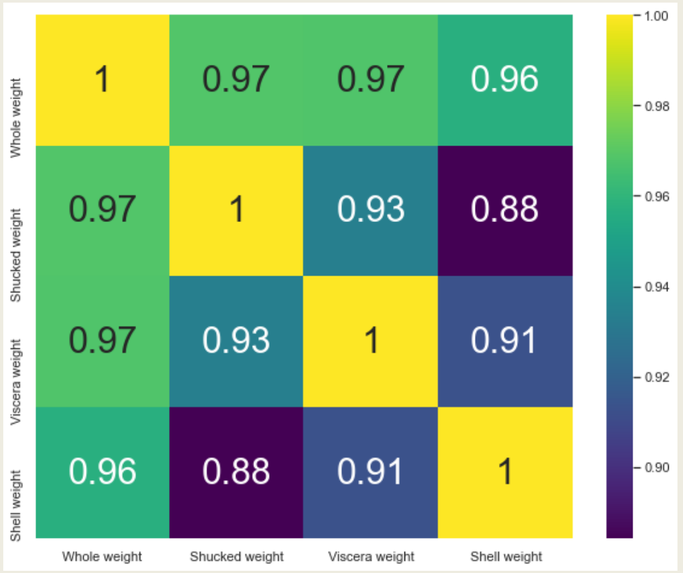
Shell weight

Shuck weight





Positively Correlated!



$$Whole\ weight = shuck\ weight + shell\ weight + viscera\ weight + weight\ loss\ due\ to\ \ bleeding/water$$

Size measures (dm)

After a certain age, the size increment rate reduces.

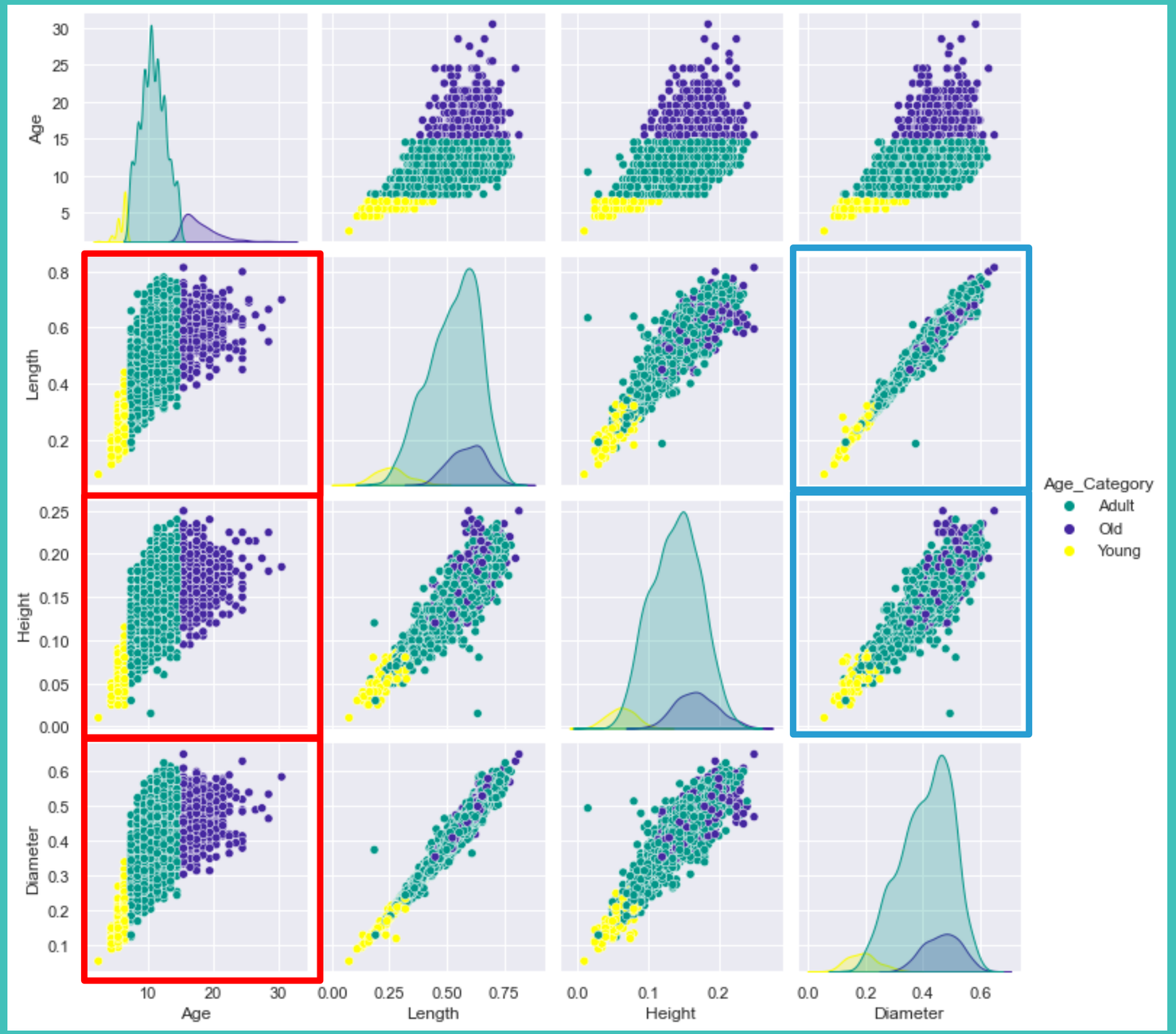
> *"growth rate reduces when maturity is reached."*
>
> *-Red Abalone Morphology*

> *"Abalones grow at about 2–3 cm a year; it takes 4 years to grow to the marketable size of more than 9 cm. In a few cases, however, it grows only about 1 cm a year, so that it takes more than 8 years to reach the commercial size."*
>
> *-Biology and culture of abalone(Chapter 1)*

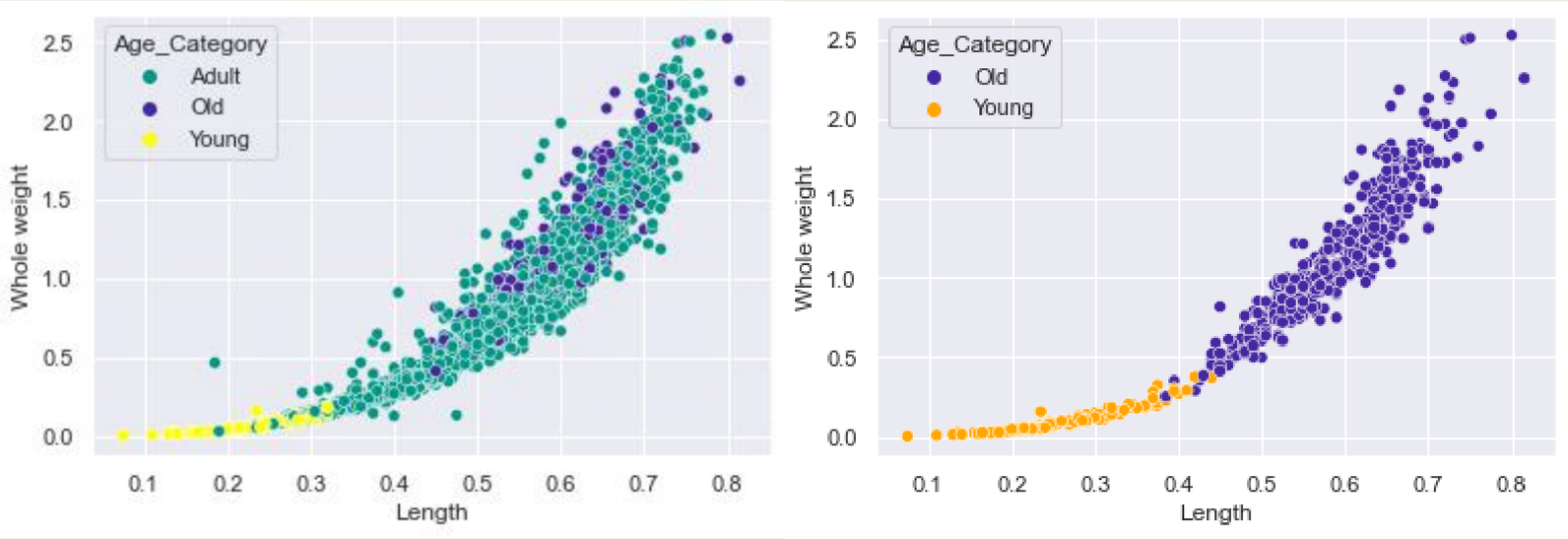
Diameter and length is highly correlated.
Height and Diameter is positively correlated.

Isometric growth

> growth that occurs at the same rate for all parts of an organism so that its shape is consistent throughout development.
>
> -genscript.com



Size and weight relationship with age



Allometric growth?

"The weight increases faster than the growth of shell. For example, if the length of the shell doubles, the weight increases eight-fold."

-Biology and Culture of Abalone Chapter 1

Allometric growth: phenomenon whereby parts of the same organism grow at different rates.

- These results agree with the findings of the study: *Biometric relationships of the Indian abalone Haliotis varia Linnaeus 1758 from Mandapam waters of Gulf of Mannar, south-east coast of India by T. M. Najmudeen*
- The **L-W relationship in different size groups** indicates that the growth in length and weight was **allometric**

"Living organisms show both types of growth during their development."

–differencebetween.com

Findings: Descriptive

- Non-Infants growth is more significant than the Infants. Because Infants are under-developed.
- Majority of the Infants are young because sex determination is unclear during early stages
- All weight measures are correlated. Weight loss due to bleeding/water. increases with age.
- Abalone growth is both allometric and isometric.

Old Abalones Are Larger and Heavier!

"The economic value of abalone is positively correlated with its age"

-Econometric Ways to Estimate the Age and Price of Abalone | MPRA Paper 91210

Old abalones properties:

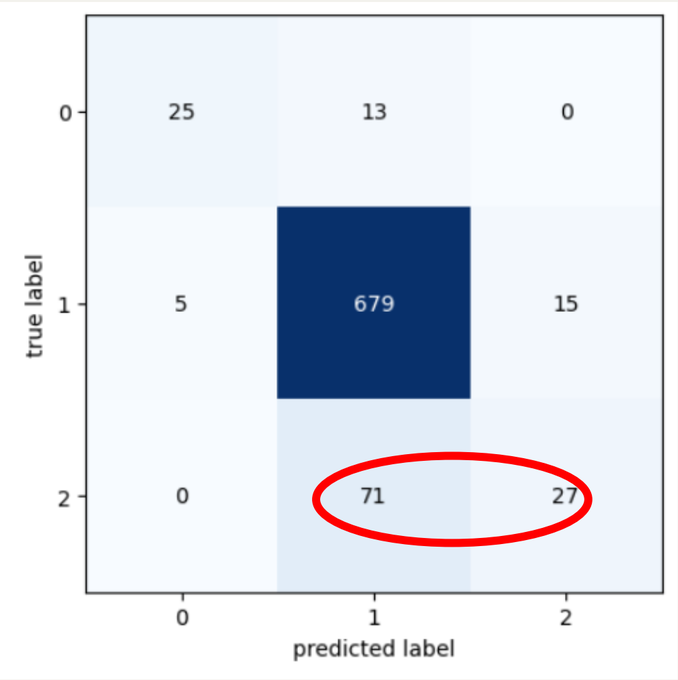
- Whole weight- (0.8-1.48) hg
- Length- (0.55-0.64) dm
- Diameter- (0.42-0.51) dm
- Height- (0.15-0.18) dm

Advanced Analysis: Model Selection

Considering All Variables

Classifier	Accuracy	F1-Score
Logistic Ridge	0.8754	0.86
Logistic Lasso	0.8742	0.85
Logistic Elastic Net	0.8742	0.85
KNN	0.8671	0.84
Random Forest	0.8682	0.85
Gradient Boost	0.8671	0.85

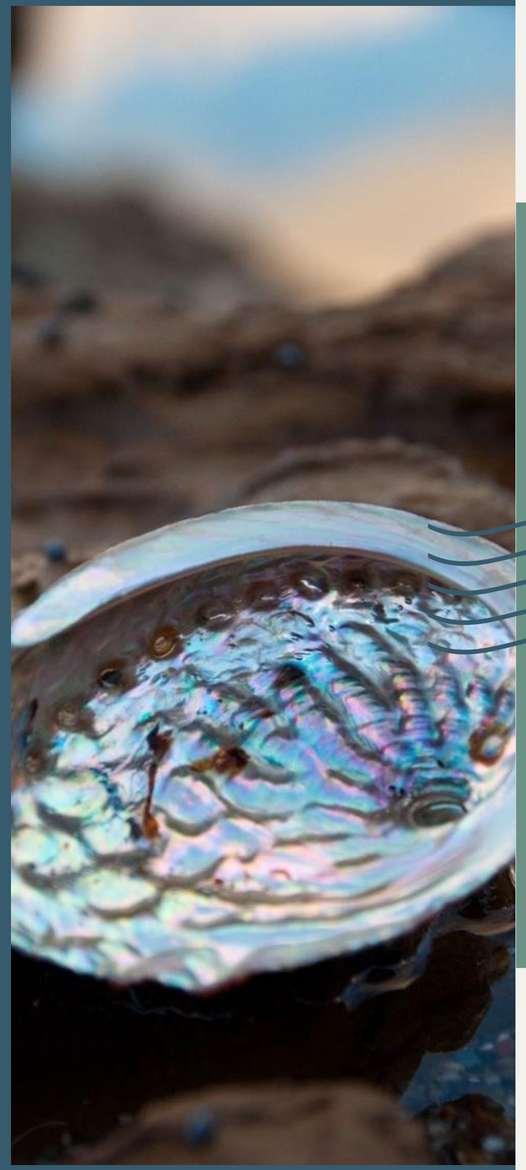
CONFIUSION MATRIX OF LOGISTIC Ridge



Best model : Logistic Ridge

Accuracy an be increased by

- Sampling Techniques
- Gender Wise Modelling



Regression Model 2

For the convenience of the user: predict the value of age

Regressor	RMSE	R square
Linear Regression	2.222	0.521
Ridge Regression	2.250	0.509
Lasso Regression	3.214	-0.001
E-net Regression	3.084	0.078
KNN Regression	2.349	0.465
CART Regression	2.868	0.202
Random Forest Regression	2.186	0.536
Gradient Boost Regression	2.196	0.532

Selected Models for Data Product

Age Prediction :

- Categorical Output Model

  Logistic Ridge Regression

- Regression Model

  Random Forest Model

Data product Provides the output of the Age category as well as the age of the abalone.


THANK YOU