Predicted Abalone's age by Machine Learnings models
And evaluating best model with Hyperparameter Tuning

The dataset was obtained from UCI Machine Learning Repository. This data consist of physical measurements of abalone from which we have to predict the age.

Brief Summary on the Topic:

Source: UCI Repository

https://archive.ics.uci.edu/ml/datasets/Website+Phishing#

Data Set Information:

- Predicting the age of abalone from physical measurements.
- The age of abalone is determined by cutting the shell through the cone, staining it, and counting the number of rings through a microscope -- a boring and time-consuming task.
- Other measurements, which are easier to obtain, are used to predict the age. Further information, such as weather patterns and location (hence food availability) may be required to solve the problem.
- Given is the attribute name, attribute type, the measurement unit and a brief description. The number of rings is the value to predict: either as a continuous value or as a classification problem
- Dataset consists of 4177 instances with 8 attributes.

Data: Name / Data Type / Measurement Unit / Description

1. Sex / nominal / -- / M, F, and I (infant)
2. Length / continuous / mm / Longest shell measurement
3. Diameter / continuous / mm / perpendicular to length
4. Height / continuous / mm / with meat in shell
5. Whole weight / continuous / grams / whole abalone
6. Shucked weight / continuous / grams / weight of meat
7. Viscera weight / continuous / grams / gut weight (after bleeding)
8. Shell weight / continuous / grams / after being dried
9. Rings / integer / -- / +1.5 gives the age in years

Head of the data frame

	0	1	2	3	4
Sex	M	M	F	M	I
Length	0.455	0.35	0.53	0.44	0.33
Diameter	0.365	0.265	0.42	0.365	0.255
Height	0.095	0.09	0.135	0.125	0.08
Wholeweight	0.514	0.2255	0.677	0.516	0.205
Shuckedweight	0.2245	0.0995	0.2565	0.2155	0.0895
Visceraweight	0.101	0.0485	0.1415	0.114	0.0395
Shellweight	0.15	0.07	0.21	0.155	0.055
Rings	15	7	9	10	

Description of Data

	Length	Diameter	Height	Wholeweight	Shuckedweight	Visceraweight	Shellweight	Rings
count	4177.000000	4177.000000	4177.000000	4177.000000	4177.000000	4177.000000	4177.000000	4177.000000
mean	0.523992	0.407881	0.139516	0.828742	0.359367	0.180594	0.238831	9.933684

std	0.120093	0.099240	0.041827	0.490389	0.221963	0.109614	0.139203	3.224169
min	0.075000	0.055000	0.000000	0.002000	0.001000	0.000500	0.001500	1.000000
25%	0.450000	0.350000	0.115000	0.441500	0.186000	0.093500	0.130000	8.000000
50%	0.545000	0.425000	0.140000	0.799500	0.336000	0.171000	0.234000	9.000000
75%	0.615000	0.480000	0.165000	1.153000	0.502000	0.253000	0.329000	11.000000
max	0.815000	0.650000	1.130000	2.825500	1.488000	0.760000	1.005000	29.000000

Description of analysis procedure and objectives:

Objective- To predict the value of age in years based on number of rings predicted based on attribute name, attribute type, the measurement unit and a brief description.
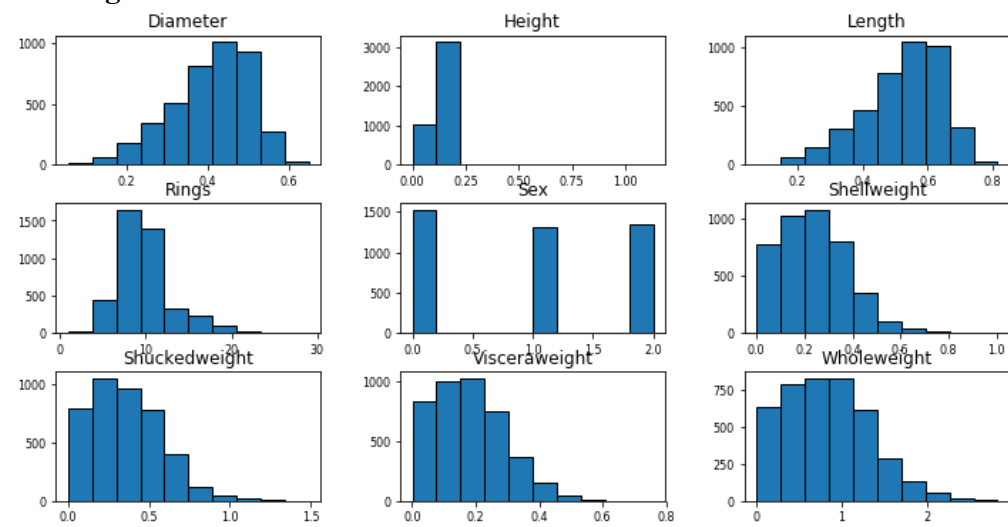
HEATMAP:
Checking : 'Missing values in the dataset'

There are no null values in data

Looking over how different features are distributed in data:

Grid-search for hyper-parameter tuning:

S NO	SUPERVISED LEARNING MODEL	PERFORMANCE OF ALGORITHMS	
		SCORE BEFORE TUNING	BEST SCORE AFTER TUNING (GRID SEARCH)
1.	Linear Regression	52.18%	50.47%
2.	Polynomial Regression	51.95%	52.18%
3.	Random Forest Regression	50.5%	52.13% 'criterion': 'mae', 'max_depth': 8, 'n_estim ators': 15, 'oob_score': 'True' (BEST PARAMETERS)
4.	Support Vector Regression	54.28%	53.93% Best Parameters: {'C': 50, 'degree': 1, 'gamma': 0.01, 'kernel': 'rbf', 'tol': 0.01}

Conclusion:
1. Maximum age of abalone is 30.5 years.

Among all the models we can compare and see that SVR model without hyperparameter tuning has highest score i.e.54.28 although overall scoring of all the models are not good so, its need further feature addition, such as weather patterns and location (hence food availability) features may be required to improve the accuracy of the model. So, that better prediction of model can be done, But for now we can consider SVR model for age prediction.