sCrabAgePrediction

CLASS: CPSC–483 Machine Learning Section–02

LAST UPDATE: May 5, 2022

PROJECT NAME: Crab Age Prediction

PROJECT GROUP:

Name	Email	Student
Brian Lucero	13rianlucero@csu.fullerton.edu	Undergraduate
Justin Heng	justinheng@csu.fullerton.edu	Graduate

PROJECT PAPER: Here

Overview

1. Abstract

Paper Summary ✓

Machine learning can be used to predict the age of crabs. It can be more accurate than simply weighing a crab to estimate its age. Several different models can be used, though support vector regression was found to be the most accurate in this experiment.

2. Introduction

The Problem ✓	Why it's important? ✓	Our Solution Strategy ✓
It is quite difficult to determine a crab's age due to their molting cycles which happen throughout their whole life. Essentially, the failure to harvest at an ideal age, increases cost and crab lives go to waste.	Beyond a certain age, there is negligible growth in crab's physical characteristics and hence, it is important to time the harvesting to reduce cost and increase profit.	Prepare crab data and use it to train several machine learning models. Thus, given certain physcial chraracteristics and the corresponding values, the ML models will accurately determine the age of the crabs.

3. Background

Process Activities ✓

- Feature Selection & Representation
- Evaluation on variety of methods
- Method Selection
- Parameter Tuning
- Classifier Evaluation
- Train-Test Split
- Cross Validation
- Eliminating Data
- Handle Categorical Data
- One-hot encoding
- Data Partitioning
- Feature Scaling
- Feature Selection
- Choose ML Models

Models ✓

- K-Nearest Neighbours (KNN)
- Multiple Linear Regression (MLR)
- Support Vector Machine (SVM)

Analysis ✓

- Evaluate Results
- Performance Metrics
- Compare ML Models using Metrics

4. Methods

Approach ✓

```
graph TD
A[Crab Dataset] --> B[Data Preprocessing]
B[Data Preprocessing] --> [KNN Model]
B[Data Preprocessing] --> [MLR Model]
B[Data Preprocessing] --> [SVM Model]
```

```
graph TD
```

```
A[Problem Domain: Age Prediction of Crabs]
B[Data Representation: Physical Attribute Values]
C[Objective Function: Average Error Rate on Training
Crab Data]
D[Evalutation: Average Error Rate on Test Crab Data]
E[Learning Algorithm: KNN, MLR, SVM]

F[Predictive Model: KNN Model]
G[Predictive Model: MLR Model]
H[Predictive Model: SVM Model]

I[Prediction System: Aggregate Model Results]
J[Useful Predictions: Low Bias, Low Variance]
K[Domain Insights: Types of Crabs]

A -- Discrete and Continuous --> B
A -- Training dataset --> C
A -- Test dataset --> D

B --> E
C --> E
D --> E

E --> F
E --> G
E --> H

F --> I
G --> I
H --> I

I --> J
I --> K
```
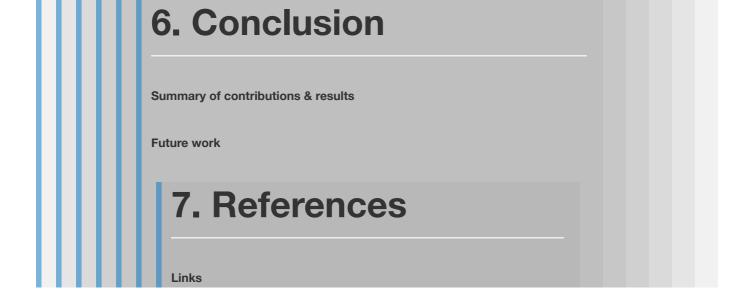
Key Contributions ✔

- **Justin**
    - KNN
    - SVM
- **Brian**
    - MLR

5. Experiments

Description of ML process workflow

Code

6. Conclusion

Summary of contributions & results

Future work

7. References

Links

SexValue	0.0337
Length	0.555
Diameter	0.574
Height	0.552
Weight	0.539
Shucked Weight	0.419
Viscera Weight	0.501
Shell Weight	0.625

_{Table 1. Pearson correlation coefficients}

[1] --- https://www.kaggle.com/datasets/sidhus/crab-age-prediction

kaggle **KAGGLE**

For a commercial crab farmer knowing the right age of the crab helps them decide if and when to harvest the crabs. Beyond a certain age, there is negligible growth in crab's physical characteristics and hence, it is important to time the harvesting to reduce cost and increase profit.

Technologies used:

- K-Nearest Neighbours (KNN) - Machine Learning Model
- Multiple Linear Regression - Machine Learning Model
- Support Vector Machine (SVM) - Machine Learning Model
- Feature Selection & Representation

- Evaluation on variety of methods
- Method Selection
- Parameter Tuning
- Classifier Evaluation
- Train-Test Split
- Cross Validation