<u>**ML Project Proposal:**</u> **Crab Age Prediction**

1 **Introduction**
Crab is very tasty and many countries of the world import huge amounts of crabs for consumption every year. The main benefits of crab farming are, labor cost is very low, production cost is comparatively lower and they grow very fast. Commercial crab farming business is developing the lifestyle of the people of coastal areas. By proper care and management we can earn more from crab farming business than shrimp farming. You can raise mud crabs in two systems. Grow out farming and fattening systems. 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.

2 **Problem**
The following traits of a crab are given: sex, length, diameter, height, weight, shucked weight, viscera weight, and shell weight. The measurements are given in feet, ounces, and months. Using these features, the age of the crab is to be determined. Note: not all of these features may be relevant, and some features may be discarded as needed.

3 **Input**
The input is a vector of the crab's sex, length, diameter, height, weight, shucked weight, viscera weight, and shell weight in feet or ounces.

4 **Output**
The output is the age of the crab in months.

5 **ML Technique**
We will try out several different machine learning algorithms in order to find the best results. The simplest algorithm we can implement is multiple linear regression where we try to fit a line to the data. This could be problematic if the age of the crab does not linearly correlate with its other features. Next we will perform K-nearest neighbors using python libraries such as sklearn. Finally, we will implement the Support Vector Regression (SVM) algorithm which uses the same principle as SVM, but for this regression problem. The analysis would involve comparing the results of each regression model's age predictions using various precision metrics. Thus, we can conclude on which model performs the best predictions and determine the reasons why it outperformed the other models in this particular setting.

6 **Dataset**
The detailed description and the dataset itself can be found under the following
URL: https://www.kaggle.com/sidhus/crab-age-prediction