Body Fat Measurement Analysis

Kimberly Humphrey - October 2020

The Dataset

- 252 Adult Males
- Hydrostatic Method and the Brozek body fat algorithm
- 13 additional separate body measurements

The problem

DoD Measurement

Cons:

Not Accurate

Mean Error:

7.65%

Error Range:

0-30%

Hydrostatic Weighing

Cons:

Expensive/Long

Mean Error:

0.8%

Error Range:

0-2%

Solution

Accurate +

Cheap +

Efficient =

Solution

Data Analysis Insights

The skeletal frame accounts for a large percentage of total body weight

- Can we use a standard formula for estimating frame size as a feature?

Feature Engineering

DoD Algorithm

86.010 x log10(abdomen - neck) - 70.041 x log10(height) + 36.76

&

Margin of Error from the Dependent Variable

BMI & Hip/Waist Ratios

(Commonly used)

BMI (Body Mass Index)
Ratio of Weight to Height

H/W Ratio of Hip to Waist measurements

Additional Body Ratios

Abs/Waist

Abs/Height

Hip/Weight

Thigh/Hip

Thigh/Abs

Chest/Abs

Abs/ (Chest/Hips)

Feature Engineering

Frame Size

Using Wrist measurements

Frame Size of either -1,0,1

(small/medium/large)

Data Analysis Insights

The BMI correlated with Error rate of the DoD measurement, at a polynomial curve against the dependent variable ‘Body Fat’

- We incorporated a second independent variable to model, based on this observation.

Data Analysis Insights

Methods Tested

5 in Total

- Linear Regression
- Lasso, Ridge Regression
- Random Forest Regression
- XgB Regressor

Solution

Lasso Regression of the Error Rate

Predicting the Error of the Current Measurement

BEST - Lasso Regression

MEAN ERROR -

ERROR MARGIN -

Comparison to Current Method

(Visual)

Future Methods

- Polynomial Regression
- Stacking Regression on top of clustering
- JackKnife Regression (Robustness)

I believe there is more potential in this data, given more research

Thank You