PREDICTING CALORIE EXPENDITURE

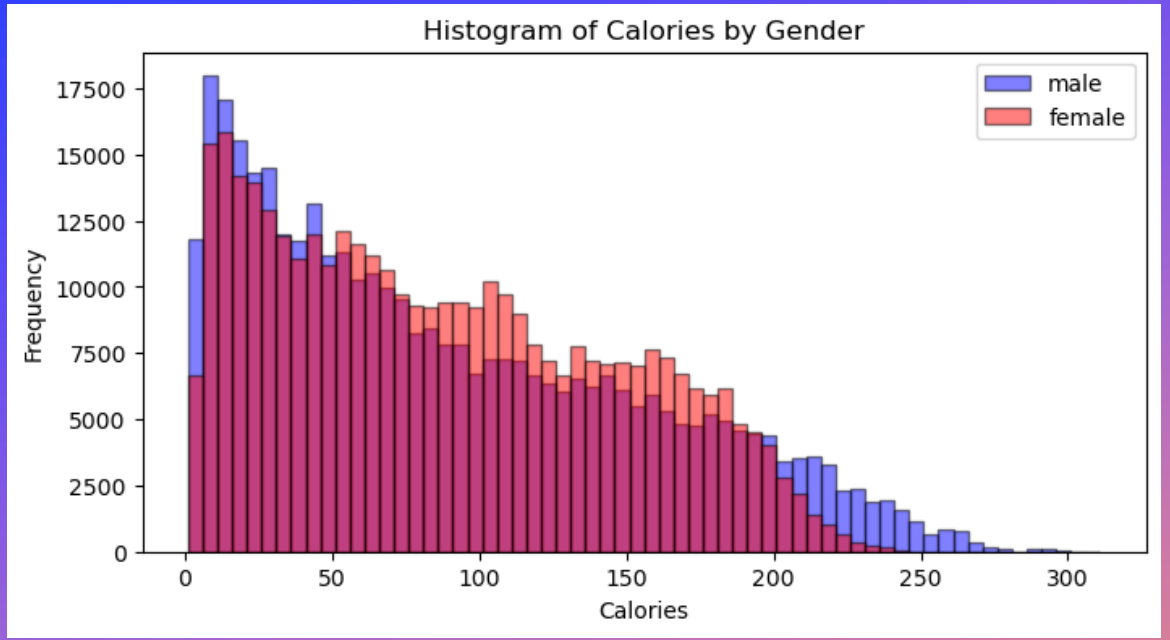
Elliot Blackstone
eblackst@umich.com
https://github.com/ElliotBlackstone

DATASET

Training set has 750,000 rows and 9 columns:

- Duration
- Heart Rate
- Gender
- Body Temperature
- Height
- Weight
- Age
- id
- Calories (target)

Testing set has 250,000 entries. The objective is to minimize root mean squared log error (RMSLE).



PREPROCESSING, FEATURE ENGINEERING

Preprocessing is comprised of the following steps:

- Log transform calories
- One-hot encoding for gender
- Delete id column
- Change columns height, weight, duration, heart rate from float to int

This reduces file size by 10.5%!

For feature engineering, we add:

- Body mass index (BMI)
- Interaction terms for numerical features
- Body temperature squared

MODEL PERFORMANCE

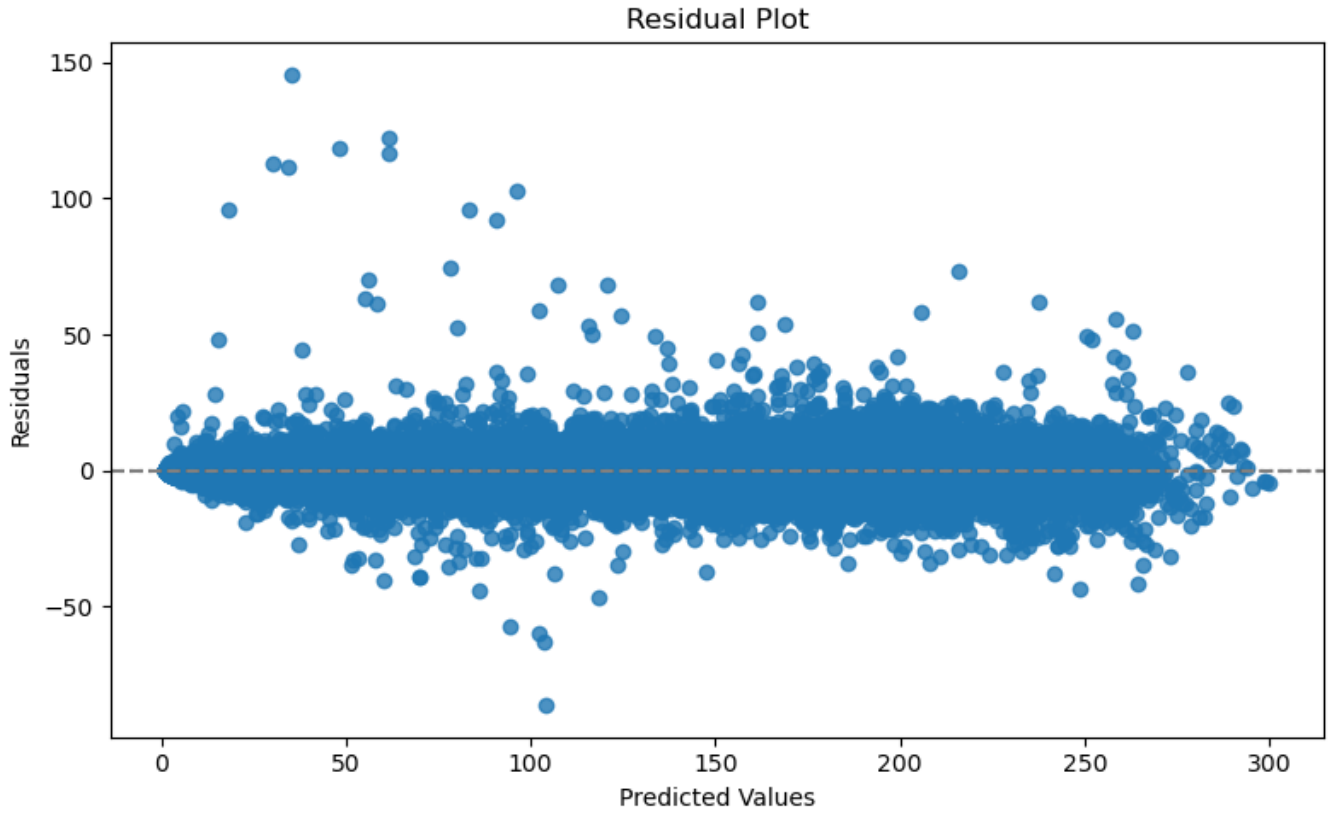
Model	Feature Engineering	RMSLE	LB (out of 4318)
Linear Regression	No	0.17926	4025
Linear Regression	Yes	0.09431	3826
GAM	No	0.08945	3811
GAM	Yes	0.06949	3605
XGBoost	Yes	0.05922	1324
LightGBM	Yes	0.05921	1314
CatBoost	Yes	0.05903	1050
Ensemble	N/A	0.05879	678
AutoGluon	No	0.05846	4

AUTOGLUON MODEL

AutoGluon trained 91 different models. The top performing model is a stacked ensemble of

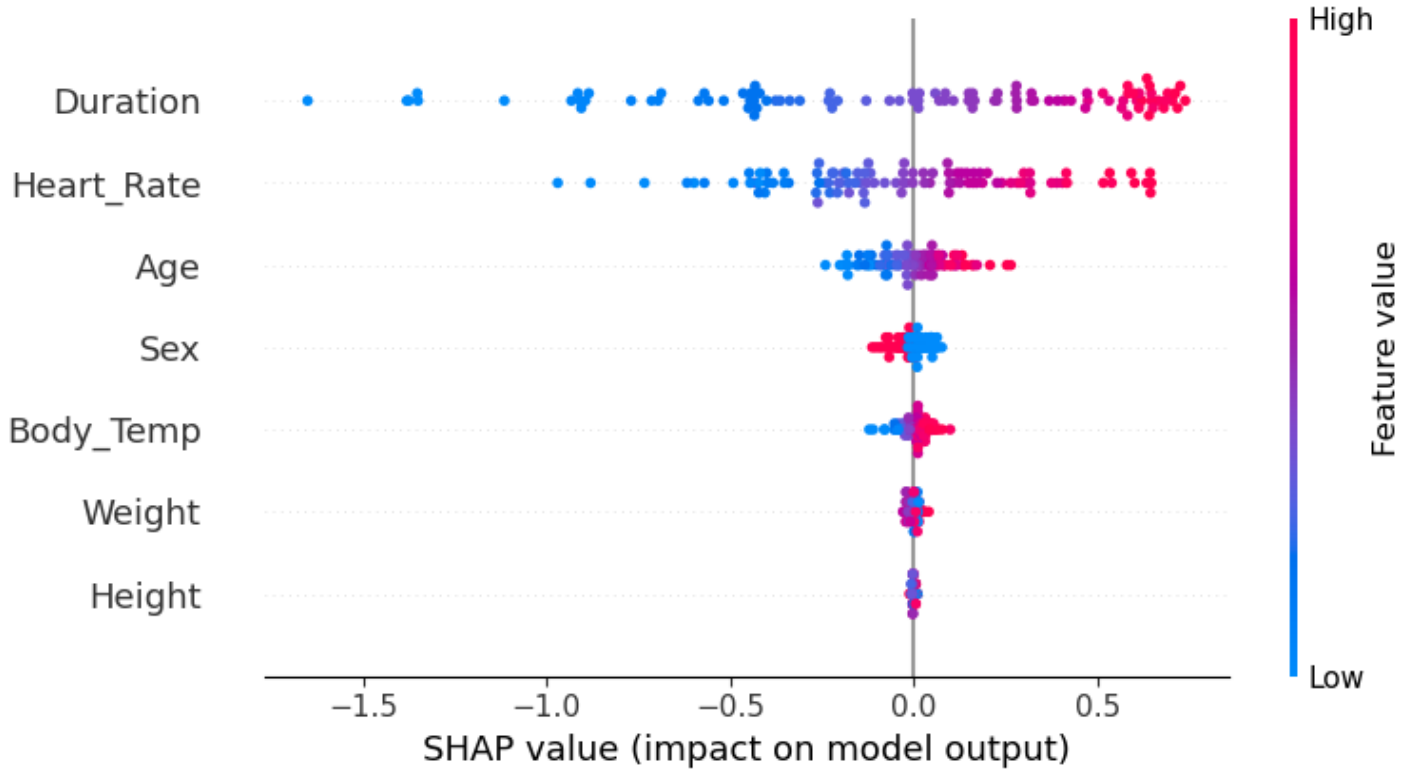
- 3 CatBoost,
- 4 XGBoost,
- 2 ExtraTrees,
- 2 RandomForest,
- 3 NeuralNet,
- 2 LightGBM.

This model is the simplest of all the top performing models on Kaggle!



EXPLAINABILITY

How to understand the contributions of each feature towards a prediction made by a machine learning model? We can use SHAP values!



CONCLUSIONS

The features that are most predictive of calorie expenditure are:

- Duration,
- Heart Rate.

The features that are NOT predictive of calorie expenditure are:

- Height,
- Weight.

According to data, if you want to burn the most calories, you should workout for a long time with a high heart rate!

Future directions for improvement:

- Expanded feature engineering
- Experiment more with ensembles and stacking
- Adjust many of the features within AutoGluon

THANK YOU

Elliot Blackstone

eblackst@umich.com

https://github.com/ElliotBlackstone