Predicting Average Heart Rate during Cardiovascular Exercise

Miles Mena

Lewis UniversityColorado Spring, United States  
milestmena@lewisu.edu

*Abstract*— Wearable GPS devices with heart rate monitoring capabilities enable athletes to collect insightful data on their workouts and recovery. Heart rate can be used to estimate exertion level and to determine the training load that an athlete has experienced, so having an accurate understanding of an athlete's personal heart rate response to exercise stimulus provides athletes and coaches information to optimize training, keep athletes healthy, and compete more effectively. We propose and test a multiple regression machine learning algorithm that takes in data from an athletes run and predicts what their estimated average heart rate was during exercise.

# Introduction (*Heading 1*)

In a paper by [Matthias Füller](https://link.springer.com/chapter/10.1007/978-3-319-27695-3_7#auth-Matthias-F_ller) et al. [3] , they predict real time heart rates during indoor running exercises and outdoor running exercises. They used several machine models with varying parameters, such as Linear Regression, Support Vector Regression, Multilayer Perceptron, and some analytical models, such as LTI, ODE Cheng, ODE Paradiso, Takagi-Sugeno to predict heart rates of a varying amount of time in the future. They trained these models to make predictions 10 seconds into the future, 20 seconds into the future, and so on.

Similarly, [Xiaoli Liu](https://www.researchgate.net/profile/Xiaoli-Liu-20) et al., in [2] adopted a model based on an Ordinary Differential Equation to predict an entire session's heart rate, and they utilized machine learning techniques, Levenberg-Marquardt algorithm, to parameterize their equations. Their model produced a RMSE of 9.35 bpm, in context of heart rate-based training 10 bpm can be the difference between two target heart rate zones, but the authors would like to improve the RMSE to 5 bpm.

Whereas the other two papers predicted heart rate during real time stimulus and for an entire session, Dur-e-Zehra Baig et al. used Linear Time Varying Models to predict heart rate from specific exercise intensities in [1].

In this paper, we expand upon previous heart rate prediction techniques by predicting what an athlete’s average heart rate was during a run from a collection of summary statistics regarding the run.

# methods

## Data Collection

Garmin GPS watches are popular GPS devices for collecting and analyzing data from an activity that uses GPS. Garming offers a free API, called Garmin Connect to export every activity that has been record by a device. In doing so, we have access to summary statistics of every activity tracked by the user.

For this research, we are using data collected from three collegiate cross-country athletes. Each athlete provided around a years’ worth of activities, accounting for a total of 2,116 runs. These athletes are between the ages of 21-23, and two are male and one is female.

## Data Cleaning

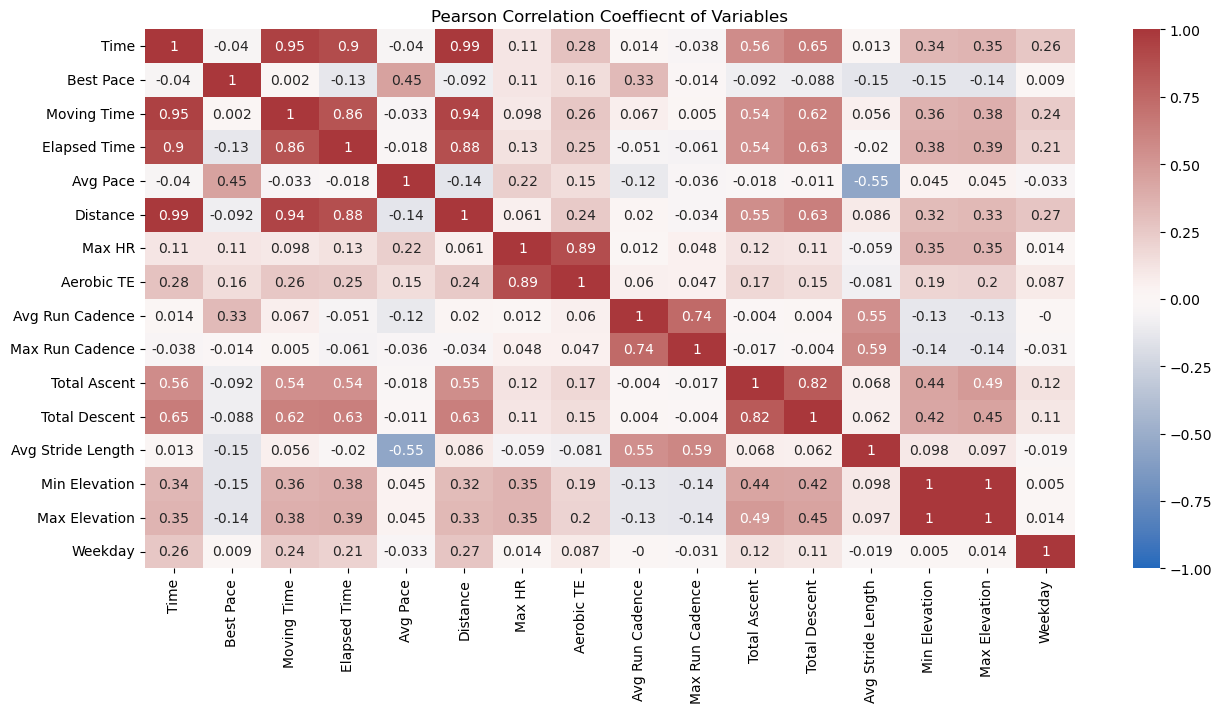
Since Garmin’s exportation of data standardizes the features, we only have to clean the features themselves and don’t have to worry about formatting columns before concatenating the datasets. The two main issues with this dataset unmeasured features on activities and converting features that track time into seconds. We change data that wasn’t collect for that feature to ‘0’, and convert time focused features like pace, elapsed time, moving time into seconds.

## Data Preproccessing

Before we can train and test our model, we have to preprocess the data so that it exists in a format conducive to an accurate model.

First, we add a feature by extracting the weekday from the date. Collegiate training follows a strict, cyclic training schedule, so workouts and easy days fall on the same day of the week. This would imply that the day that the activity takes place on might predict what the average heart rate was for a run.

Secondly, to decrease the complexity of our model we find features that are highly correlated with each other, higher than .9, and use only one of those features. Running our predictive features through Pearson correlation tests yields the results in the following heatmap.



Before spitting the data into an training and testing data, we normalize the data using max absolute normalization using the formula in (1).

*X’ = X / max (| X|)* 

Finally, we would like to test the performance of the model on each individual athlete. To ensure that our testing data has a representative sample, we split each athlete’s data individually into X training, X testing, y training, and y testing using an 80:20 split. After each athlete has been split, we concatenate each athletes X training and y training so that we have one X training and one y training that holds every athlete. When we test our model, we can now determine how well the model training on all three athletes performs on each individual athlete.

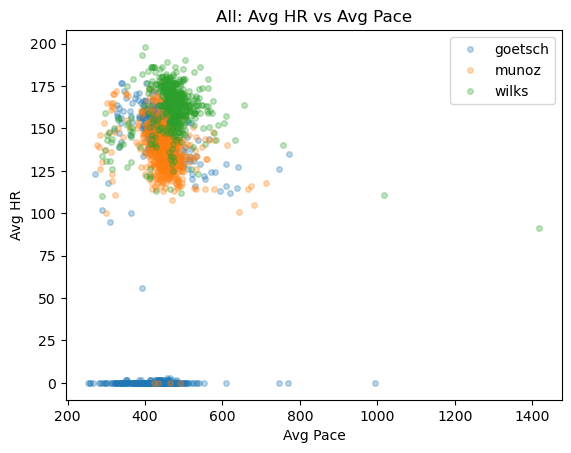
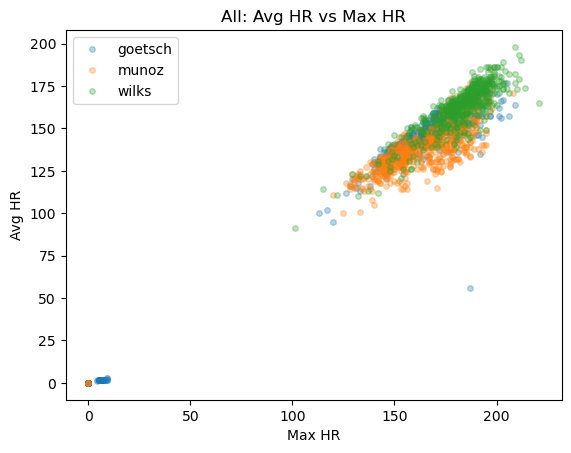
# THE MODEL

The model we use is a multiple linear regression model from the python library sklearn. Multiple linear regression is a variant of a single regression model, whereas instead of using a single feature to predict a target feature, Multiple linear regression uses multiple features to predict a target feature. In other words, multiple linear regression is a statistical algorithm used to predict a continuous response variable with a collection of explanatory variables.

*Yi = b0 + b1 x1 + b2x2 + … + bp xip + error* 2

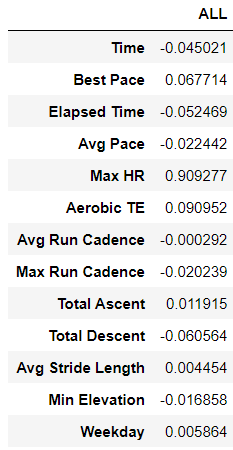
Where Yi is the response variable, average heart rate in this case, xi is the feature, and bi is the coefficient corresponding to its feature.

In this scenario it is appropriate to use a multiple linear regression because some features exhibit a linear relationship. For example, an athlete’s max heart rate seems to be related linearly with average hear rate.



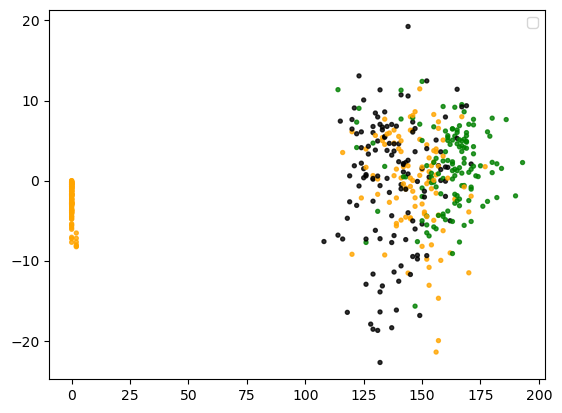
However, it is not the case that all variables are linear. For example, average pace, which is measured in average seconds per mile, does not seem hold a linear relationship with average heart rate. This is a bit contradictory to what one might assume as it indicates that an activity where a runner is running faster doesn’t mean that their hear rate will be higher.

After training the multiple linear regression model, we have a model that predicts an average heart rate using the following coefficients,



Notice that this model’s highest coeefient is max heart rate, and that weekday isn’t as high of a predictor as suspected.

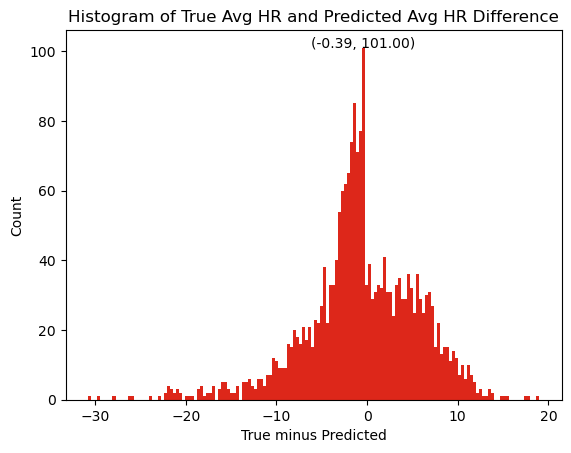
In addition, after testing this model on the testing data set containg all althetes we get the following perfomance scores: Mean Squared Error of 0.00099, and an R^2 value of 0.9922. While the model appears to be producing extremely accurate predictions there seems to something inaccurate with the performance scores we are receiving. For example, examining the residual plot shows that the predictions are further off than the performance score indicate.



On this plot shows the true heart rate on the x axis and on the y axis is the difference between true heart rate and predicted heart rate of 500 randomly sampled activities. This shows that the performance scores aren’t telling the full story since there is a random distribution of points that aren’t cluster around the y = 0 line. It is possible that the training data is producing a strong performance score because it is scaled.

Regardless, the difference between the true heart rate and predicted heart rate follows a pattern and there is a group of data that is predict quietly accurately. The following graph demonstrates the histogram of this pattern. There are several bins that are extremely close to 0, and a lot of the activities are predicted within about 5 beats per minute of the average heart rate.

On the other hand, this graph demonstrates that there are just as many predictions that don’t fall close enough to be considered accurate. Predictions off by over 5 beats per minute wouldn’t help an athlete or coach understand the response to a workout, so for these predictions we can consider them inaccurate.



Despite the strange accuracy results, we still tested how the model performs on individual athletes testing data. The results are displayed in the table below.

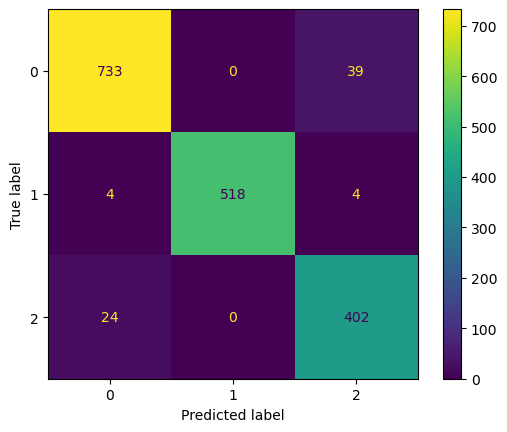
|  |  |
| --- | --- |
| Athlete | MSE |
| Munoz | 0.00166 |
| Wilks | 0.00063 |
| Goetsch | 0.00073 |

The MSE of Munoz is the highest, while the MSE of Wilks is the lowest in this test. Regardless, the MSE of all three athletes is quite close, relative to each other. It is interesting that predicting Munoz preforms the worst with this model considering that Goetsch’s data is the most different that the other athletes.

# Classifying Athletes and Weekday

Since we have all the athletes in one dataset, we build a logistic regression model, using a one versus rest approach, that predicts what athlete completed an activity based on their performance statistics. For such a model, we split and train the data differently than how we test the multiple linear regression model. Instead of breaking the testing data into three different testing sets, we keep them as one large testing sets.

After training, our model has a training accuracy of 95.8817. In addition, the confusion matrix looks like the following image, where 0 is Goetsch, 1 is Munoz and 2 is Wilks.



This indicates that there is some quality in an activity that makes an athlete distinguishable. The most missed runners is confusing wilks with goetsch, which makes sense since they often run together. This means that their performance stats are essentially the same besides a few features that are unique to each runner. The logistic regression for classifying runner is therefore more effective than this confusion matrix would indicate.

# Conclusion

A multiple linear regression using performance variables to predict the average heart rate from run can be quite effective given the right circumstances. Using data from multiple athletes is as effective if not more effective than using a single runners’ data. Thus, if a model were to be scaled up to predict other college athletes heart rate it should be possible to use only a handful of college athletes’ data in training. This would save the task of constantly retraining.

Also, classifying which runner ran a particular activity is possible using a one versus rest logistic regression. Such a model would work for runners who don’t often run together, as was the case for wilks and goetsch.

# Future Work

Despite the multiple linear regression model having a high R^2 and a low mean squared error, the predictions are not working as they should. Fixing this would give an insight into what is going on in the model. To fix this we would unscale the mean square error so that we can have a context for the MSE. We can also produce better visualization of the model and it’s residuals.

To increase the performance of the model we can add more data from other collegiate athletes. This will give the model a wider view. We could also add data from non-collegiate athletes.

##### Acknowledgment *(Heading 5)*

The authors of this paper would like to thank Dr. Al-Khassaweneh for is support and guidance in creating our model and in learning various machine learning techniques.

##### References

1. B. Dur-E-Zehra, S. Hao, C. Teddy and S. Andrey, S. Steven, C. B.G., “Modeling of human heart rate response during walking, cycling and rowing,” Conference proceedings :Annual International Conference of the IEEE Engineering in Medicine and Biology Society. IEEE Engineering in Medicine and Biology Society, Conference. 2010
2. L. Xiaoli, S. Xiang, S. Tamminen, and T. Korhonen, and J. Röning, “Predicting the heart rate response to outdoor running exercise,” 2019
3. M. Füller, A Sundaram, M. Ludwig, A. Asteroth, and E Prassler, “Modeling and predicting the human heart rate during running Exercise,” Springer International Publishing, pp. 106-125, December 2015

As for the logistic regression model, we can add other runners

Data and see if the model works just as well.