Optimizing Fitness: Integrating Fitness Metrics into Tailored Workout Recommendations

Jeffrey Jacob and **Amar Jilani** and **Spencer Uyematsu**
(**YouTube Presentation Video**)

Abstract

This study delves into optimizing workout strategies using the Endomondo social fitness application dataset, addressing the gap in personalized exercise regimens. Focusing on three key questions, we developed five distinct models: two (an Ordinary Least Squares model and an artificial neural network) for predicting average heart rate, another OLS model for estimating calorie expenditure, a Random Forest classifier for sport type identification, and a collaborative filter based model to recommend new sports and connect users. Our findings reveal the OLS models' effectiveness in predicting calories burned, while highlighting limitations in heart rate prediction due to the complex, non-linear nature of the data. The Random Forest classifier successfully categorized sports types, indicating the potential for enhanced personalized workout recommendations. The collaborative filter successfully produced sport recommendations and potential user-user connections. To fully measure the effectiveness of these recommendations, however, user input is still required. This research contributes to the evolving field of fitness technology, emphasizing the importance of tailored workout plans for improved health outcomes.

Keywords: Ordinary Least Squares, Random Forest, Fitness Recommendation, Collaborative Filtering

1. Introduction

Exercise has long been recognized as a fundamental aspect of good health. Several studies over the past decades have consistently demonstrated the various benefits of regular physical activity, ranging from improved cardiovascular health, reduced risk of disease, and improved mental health [1][2]. Despite the clear benefits, adherence to exercise regimens has decreased over the past decade, coinciding with increased time spent in sedentary behavior [3]. Additionally, creating and adhering to an effective workout plan can be difficult for many individuals. This growing gap between the importance of exercise and the practical application of exercise routines presents a unique challenge.

Bridging this gap, fitness trackers have emerged as a pivotal tool in promoting healthier and more active lifestyles[4]. Fitness trackers have given us the ability to accurately measure statistics about our workouts. Not only does this help a user track their progress, but it can also influence their future workout performance by providing valuable insights. Furthermore, they allow us to analyze workout patterns and statistics across groups of people at a large scale. In this project, we examined a rich dataset from Endomondo, a social fitness application, that helps users track their fitness metrics. We intend to utilize this data to gain a more nuanced understanding of the various workout metrics and how they can be used to inform personalized and optimized workout regimens.

There has been considerable research delving into the benefits and impacts of exercise on health and fitness, however there remains a notable gap in the exploration of optimized workout strategies. Optimized workout strategies focus on tailoring exercise regimens to a specific individual to maximize effectiveness while ensuring consistent engagement.

Previous studies have looked into recommending exercises based on a user's personal workout history using recurrent neural networks. For example, Mahyari and Pirolli focused on predicting the probability of exercise completion based on a user's personal exercise history [5]. In their study they trained an RNN on a sequence of data that included user information on the successful completion of a workout, allowing the RNN model to predict the probability of successfully completing a recommended workout. Other tools, such as FitRec, which also utilized the Endomondo dataset and variation of the recurrent neural network, a long short-term memory (LSTM) model, have been developed to analyze and predict individual exercise performance in the short-term, focusing on route recommendations and pacing to maintain target heart rates [6]. This study utilized a user's time-series heart rate data, along with a user's workout history, to train the LSTM model to accurately predict a user's heart rate during a workout. The motivation for this was to provide a user personalized insights during a workout based on their predicted upcoming heart rate, i.e. whether the increase or decrease pace. Despite these important developments, there is a notable scarcity of research dedicated to the development of individualized workout routines that are tailored to specific heart rate or caloric goals, which are two common and important metrics when optimizing personal health outcomes.

In this paper we address three separate questions:

1. What specific workout metrics are most predictive of key performance indicators like heart rate and calorie expenditure, and what models can we use to best capture this relationship?
2. Can the workout features be used to accurately classify the sport being performed?
3. How can we facilitate user-user connections and create personalized recommendations for users to explore new sports?

To answer these questions, we created five different models.

The first model is an ordinary least squares (OLS) model used to predict the average heart rate of a user throughout a workout. We then modeled this same issue using an artificial neural network in an attempt to achieve more accurate results. These two models differ from the approach outlined in the FitRec dataset as they approached the problem with an LSTM, whereas we utilized a standard deep neural network. Additionally, they focused on instantaneous heart rate throughout a workout, while we focus on average heart rate. Next, we created another OLS model to predict a user's total calorie expenditure throughout a workout. Then, we utilized a random forest to accurately predict the sport based on the different features collected from the workout features. Finally, we built a model based on collaborative filtering to connect users with others who have similar workout patterns. Each user is then recommended new sports to try based on the sports their connections enjoy.

2. Description of Data

In this project, we used a dataset from Endomondo, a social fitness application, that helps users track their fitness metrics. The data is broken into two separate datasets. The first includes continuous measurements for heart rate, speed, altitude, latitude, longitude, and a unique workout ID. The second includes additional information on each workout such as the sport, weather, altitude, duration, calories, time, and distance. There is also information about the user such as a unique identification number and gender. The dataset contains 253,020 workout records across 1,104 different users. Additional information about each feature in the dataset can be found in table 1.

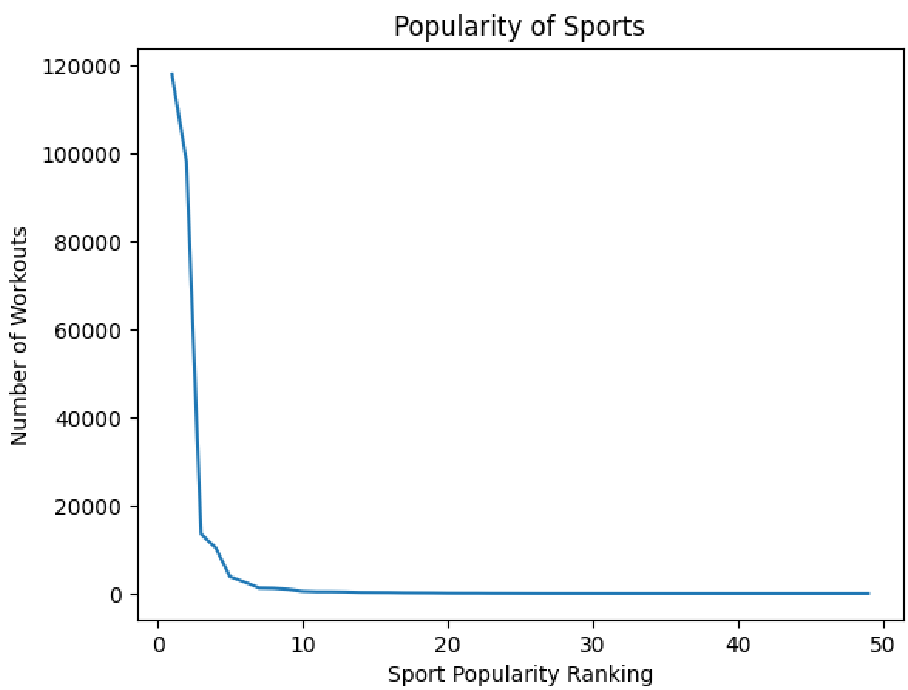

Figure 1: Number of recorded Workouts for each sport by popularity ranking

There are 61 different sports present in the dataset, however, the three most popular, running, biking, and walking makeup 76% of the total entries. Measurements such as distance and heart rate vary greatly across different sports. Furthermore, most of the entries center around Europe, however, users have entered workouts all over the world.

Feature		**Description**
UserID		Unique identification for each Endomondo user.
ID		Unique identification number for each individual workout.
Sport		Activity conducted in a given workout.
Weather		Weather during workout.
Gender		Endomondo user's gender
Feature	**Unit**	**Description**
Ascent	M	The total negative vertical distance traveled during a workout.
Descent	M	The total negative vertical distance traveled during a workout.
Gender		Endomondo user's gender
Duration	sec	Length of workout.
Timestamp	Unix Timestamp	Continuous time measurements throughout the workout.
Distance	Mile	Total distance traveled during workout.
Latitude & Longitude	Degrees	Continuous location measurement throughout the workout.
Heart Rate	BPM	Continuous heart rate measurement throughout the workout.
Calories	kCal	Total calories burned during workout.

Table 1: User provided information and Workout Measurements

While there is a large amount of exercise data in the Endomondo dataset, it is important to recognize its limitations as our goal is to make models that can extend beyond the dataset. To begin, the data was recorded through the Endomondo mobile application by individual users. The measurements found in our data set were taken by a wide range of devices (e.g. smartphones, smart watches, etc.), so both the accuracy and the number of measurements for every data point are not consistent. This could cause a potential selection bias as users with more advanced fitness tracking devices have more data. Furthermore, we understand that the individual who chooses to use Endomondo might not be representative of the general population.

Additionally, there is a strong bias toward male users as over 86.5% of the users within the dataset are male. These users account for around 89.5% of the total workouts in the dataset. The dataset contains a high number of missing values. For example, we only have heart rate information (the first dataset outlined above) on 250,000 workouts of the around 900,000 total workouts contained within the second dataset.

3. Methodology

In this project, we address three key research questions, each requiring distinct models.

1. *What specific workout metrics are most predictive of key performance indicators like heart rate and calorie expenditure?*

Our first approach for predicting average heart rate was using an Ordinary Least Squares (OLS) regression model. We believed that OLS was an appropriate model to start with as it is simple and interpretable, while also being able to use the various features of the dataset. For this model, we chose to focus on the features: ascent, descent, duration, distance, type of sport, and gender. For feature engineering, since sport and gender are categorical features, we one-hot encoded these values, whereas the quantitative values were normalized along each column. Lastly, for the training dataset, we also chose to remove outliers, as there were various abnormal or irregular workout entries. The data was split into a training and testing set, where the training set was further split into training and validation during 3-fold cross-validation. Cross-validation allows us to evaluate the model's performance without having to utilize the test set.

For our second model, we chose to implement a deep neural network for predicting average heart rate, as we believed that this approach would do a better job of capturing the more complex and non-linear relationships between the features. We designed a straightforward model with 3 hidden layers that used the ReLU activation function. The input size was the number of features of our data, and the output layer was a single value that we are predicting, i.e. average heart rate. We used similar feature engineering and normalization techniques as was done for the OLS model. The model utilized MSE as the loss function. Furthermore, we also added L2 regularization to the Adam optimizer in the form of weight decay to avoid overfitting. This model was also trained and validated using 3-fold cross-validation. We manually tuned the model with various learning rates, levels of regularization, and different activation functions. The models were trained for 300 epochs.

We also used an OLS model to predict the calories burned during the workout based on various workout metrics, including type of sport, ascent, descent, duration, distance, and gender. After observing the relationship between the different features and calories, we found that there was a strong linear relationship between log duration and log calories, therefore, we log-transformed the duration feature and the output (calories). We also performed one-hot encoding for the categorical values and normalized the quantitative features along each column. Again, we removed some of the outliers and irregularities from the data. The model was fitted and evaluated using 3-fold cross-validation.

2. *How can we classify different sports types from workout data, such that we can make recommendations?*

We utilized the Random Forest Classifier for our classification model using the following features for prediction gathered from the dataset and accompanying metadata: 'avg_latitude', 'avg_longitude', 'avg_heart_rate', 'avg_altitude', 'avg_speed', 'gender', 'ascent', 'descent', 'calories', 'duration', 'distance', and 'weather'. Random forests are a good choice for our model as it can capture complex non-linear relationships in data similar to those present in our data. For feature engineering, we took the averages of the heart rate and latitude/longitude data, as they were in a time series. Additionally, we one-hot encoded the gender column as well as the weather column. In total, we used around 250,000 entries in our model. There was a great deal of missing values which required extensive processing.

In the data processing routine, the training and test datasets are treated differently to prevent data leakage and ensure the model's evaluation is fair and unbiased. For the training set, One-Hot Encoding (OHE) is applied to the target 'sport' column, transforming categorical sports data into a numerical format, and this encoded data is then used along with the features in an Iterative Imputation process to fill in any missing values (there were many). Iterative imputation from sklearn.experimental is more advanced than traditional imputation strategies like filling with the mean as it uses a regression model. This process helps the model learn complex patterns specific to each sport. After imputation, the features are standardized to normalize their scale. For the test set imputation and scaling, however, the approach is slightly different to maintain the integrity of unseen data: the OHE is applied as per the encoding learned from the training set, but imputation does not leverage the 'sport' information. Dummy columns are temporarily added to align the test set's features with the training set, ensuring that the model evaluates its performance on data structured identically to what it was trained on, but without using any specific information from the train set or being able to peek at the sports column to influence imputation. This careful separation in processing steps, of scaling and imputation is crucial for avoiding data leakage, where information from the test set could inadvertently influence the model during training, leading to overly optimistic and unrealistic performance evaluations.

The classification problem on this dataset is inherently challenging due to the disparity in occurrences of each sport, certain sports appear thousands of times more than others skewing predictions towards them. To combat this we employed SMOTE (Synthetic Minority Oversampling Technique) from the imblearn package to address this in the training data. Additionally we also employed under sampling of the most common classes to further fix the class imbalance. We included the top 25 most common sports in the classifier. We then used Grid Search CV along with a 3 fold Cross validation to find the best hyperparameters which are 'max_depth': 30, 'n_estimators': 300.

3. *How can we facilitate user-user connections and create personalized recommendations for user's to explore new sports?*

As shown in the previous analysis, a substantial portion of our dataset is spread across the most popular sports activities such as running, biking, and walking. The exploration of new sports can help maintain an individual's interest and engagement in their fitness regimen. We devised a model, akin to a collaborative filtering approach, aimed at linking similar users and crafting a tailored list of new sports to pursue.

Similar to the approach used in the OLS and neural network models above, our initial steps involved the elimination of workouts that contained outlier measurements, normalization of continuous features (using sklearn.standardscaler), and one-hot encoding of categorical features. Additionally, we introduced a normalized feature indicating the count of workouts recorded by each user within the dataset. After this, we aggregated all workouts for each user, computing the average of each feature to create an overall summary of individual profiles. Then, we reduced the dimensionality of the data set from 62 features to 8 principal components using Principal Component Analysis (PCA).

Subsequently, we implemented a nearest-neighbor algorithm to identify similarities among users, facilitating potential recommendations for connections or workout partners with similar fitness profiles. In the final step, we compared the sports each user had recorded a workout for to those engaged in by users with similar profiles. Through this model, we propose suggesting sports, in order of popularity, undertaken by similar users that the original user has not yet participated in.

The model described above is an ideal fit for curating targeted sport suggestions as it does not suggest only the most popular sports to every user, but the sports that are underrepresented in the dataset as well. In the process of creating sport recommendations, the model also identifies similar users, providing a route to recommend possible user-user connections.

4. Results

Average Heart Rate Prediction

As discussed in the previous section, we created two models to predict the average heart rate of a user given various aspects of a workout. Our first model was an OLS model, thus we evaluated the model's performance using RMSE. In this model, we were able to predict a user's heart rate with a root mean squared error of 14.84 bpm. Figure 2 shows a comparison between the predicted and observed average heart rates of the test set. Furthermore, we have also included a residual plot to analyze the distribution of the residuals along the predicted values. These plots provide significant insight into the nature of our model's predictions and the relationship between the various features of a workout and heart rate.

As shown by the plot predicted vs. observed values plot, it is clear that the model tends to predict in a tight range between 130-150 bpm. This finding is also reinforced through the residual plot. Though there is generally even positive and negative residual distribution, the residuals are largest in the center of the plot, where the model is predicting within that range of 130-150. This likely shows that the model is simply predicting the average range for average heart rates over all the data points, rather than being able to infer from the features. These results indicate one of two possibilities: either the relationship between the selected workout metrics and heart rate cannot be represented linearly, or heart rate cannot be predicted effectively using the features selected.

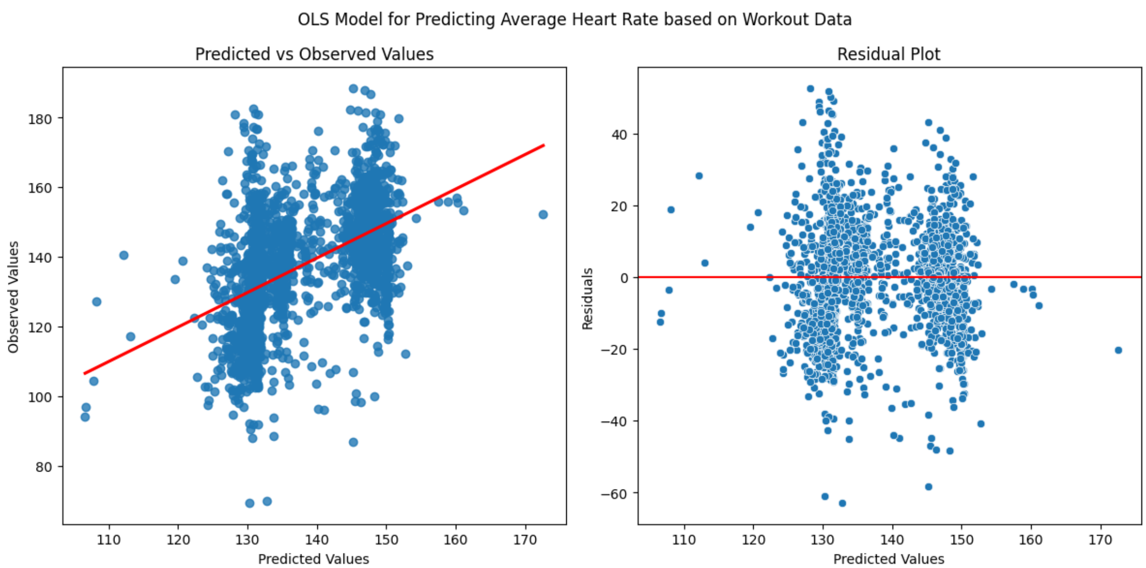

Figure 2: OLS heart rate model prediction accuracy and residual distribution

To address the problem we faced with the OLS model, we chose to design and train an artificial neural network with the hopes that it may capture the more complex relationships between the features, if present. The model was trained with 3-fold cross-validation and both the training losses and validation losses were plotted, as shown in Figure 3. The plots show that the models were able to converge fairly quickly, with no notable signs of overfitting, as the validation error does not show any signs of increasing. The model converges quickly, but at a relatively high loss, with the lowest validation loss at around 17 bpm, and the testing error at around 18 bpm. This is slightly worse than what was achieved with the

OLS model.

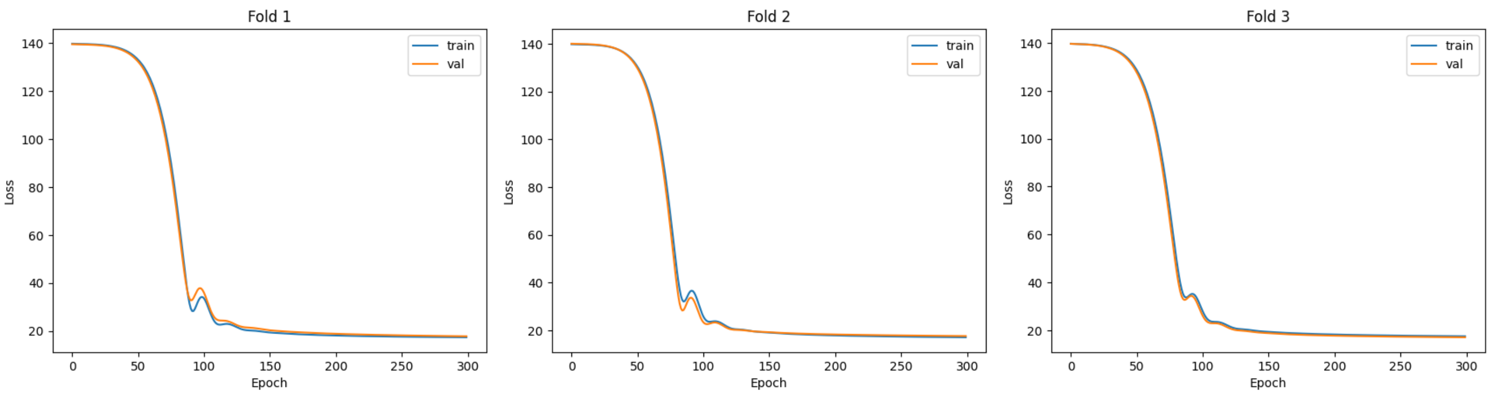

Figure 3: ANN heart rate prediction model learning curves

The plots in Figure 4 show a very similar pattern in terms of the model tending to predict within a specific range, however, the range is much tighter compared to the OLS model. This is shown in both the accuracy and residual plots shown below. Unlike the OLS model, there is a slight negative correlation between the predicted values and the observed values. The residual plot shows that the ANN model tends to overestimate average heart rate compared to the OLS model, as there are a greater number of negative residuals, as well as of larger magnitude.

Our research question was to investigate the relationship between workout metrics and performance indicators, such as heart rate, and whether those metrics could then be used to make meaningful predictions. For predicting heart rate, the results show that there is a fundamental limitation in the relationship between workout data and the heart rate outcome. Both models were unable to make meaningful predictions, thus it is likely that heart rate during a workout is less related to the specific features of the workout, such as duration, ascent, descent and the type of sport, and likely more related to the specific user's individual characteristics and general fitness. This explains why both models generally performed poorly and tended to predict close to the mean average heart rate, rather than capturing any meaningful relationship. Furthermore, both models performed similarly, despite the ANN being much more complex than the OLS model, which shows that both models faced high bias and were generally under-fitting. Surprisingly, the artificial neural network performed slightly worse than the linear regression model when comparing RMSE on the test set, which may be due to the relatively small size of the dataset, as well as the general lack of feature relevance, making it difficult for the ANN to learn.

Calorie Expenditure Prediction

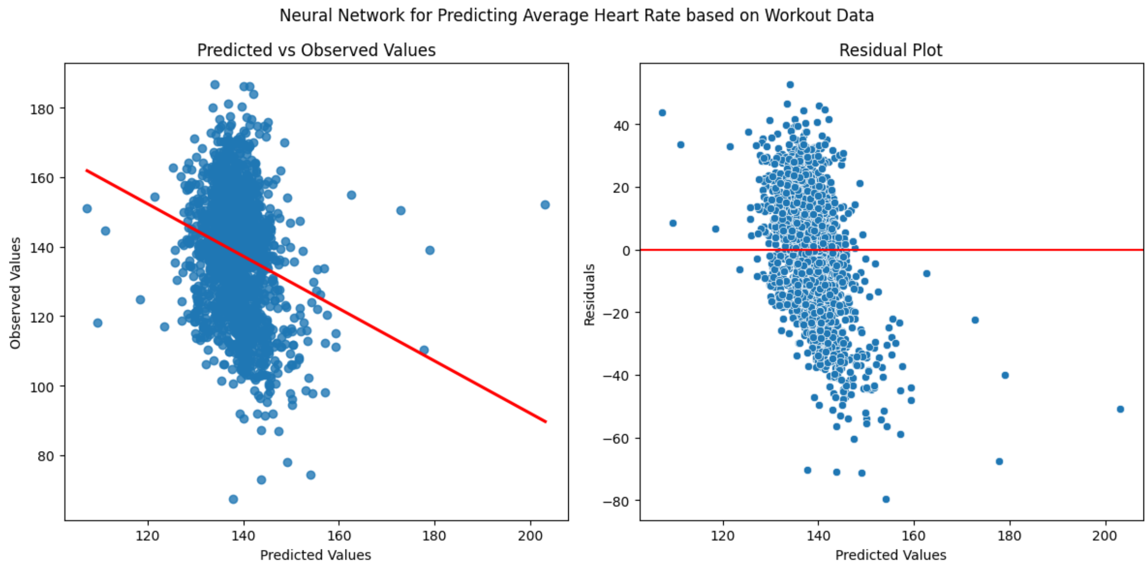

Figure 4: ANN heart rate prediction model accuracy and residual distribution

Our next analysis focused on another important performance indicator of a workout, the calorie expenditure during exercise. We used an OLS model using the same features that were used for the heart rate model. In contrast to our heart rate predictor, our model was able to accurately predict the log calories burned during a workout with an RMSE of 0.32 log calories. These results are shown in the accuracy and residual distribution plots in Figure 5. This accuracy plot shows that there is a strong positive correlation between the predicted log calories and the observed log calories. Furthermore, the positive and negative residuals are generally evenly distributed, however, there is some level of heteroskedasticity. This means that the variance of the residuals is not constant at all values of values of log calories. This violates the assumption that each data point's error term should have the same variance. This does not increase the bias of the model, however, the predictions may be less reliable. This approach may require further transformations or alternative models.

Returning to our initial research question, we found that workout metrics are generally good predictors of calorie expenditure and a linear regression model is capable of modeling this relationship. As mentioned, there are some limitations with the model due to heteroskedasticity, which may be due to limitations of the data. There are likely many more factors that play a role at higher levels of calorie expenditure, such as fatigue, intensity, and efficiency.

Figure 5: OLS calorie prediction versus measured calories burned values

Sports Classification

Classification Report:

	precision	recall	f1-score	support
aerobics	0.71	0.71	0.71	7
bike	0.98	0.96	0.97	19395
bike (transport)	0.86	0.87	0.86	2085
circuit training	0.83	0.76	0.79	38
core stability training	0.63	0.71	0.67	248
cross-country skiing	0.80	0.87	0.83	191
downhill skiing	0.91	0.91	0.91	23
elliptical	0.75	0.75	0.75	16
fitness walking	0.87	0.76	0.81	79
gymnastics	0.93	0.87	0.90	30
hiking	0.58	0.74	0.65	68
indoor cycling	0.70	0.85	0.77	552
kayaking	0.79	0.83	0.81	53
mountain bike	0.90	0.95	0.93	2667
orienteering	0.68	0.78	0.73	247
roller skiing	0.92	0.74	0.82	88
rowing	0.54	0.39	0.45	18
run	0.99	0.98	0.99	23773
skate	0.88	0.72	0.79	114
snowshoeing	0.50	0.33	0.40	6
soccer	0.85	0.81	0.83	21
swimming	0.43	0.38	0.40	8
treadmill running	0.67	0.14	0.24	14
walk	0.85	0.92	0.88	775
weight training	0.94	0.78	0.85	58
accuracy			0.96	50574
macro avg	0.78	0.74	0.75	50574
weighted avg	0.96	0.96	0.96	50574

Accuracy: 0.9602364851504726

Figure 6: Classification report for random forest sports classification

Figure 5 shows the results of the random forest classifier on the test set of 50,000 entries. The Classifier has a high Weighted Precision, Recall, and Accuracy all at 96 percent. It is clear that class imbalance in the test set is severe with the most common sport running appearing 23,000 times while the least common snowshoeing only appearing 6 times. High precision indicates that the model is able to correctly identify the sport based on the features, whereas a high recall shows that the model is generally able to capture instances of a specific sport.

Despite these imbalances, the classifier was able to achieve a macro average precision, recall, and accuracy at around 75%, which indicates that the model was generally accurate at predicting the sport. Our efforts in treating class imbalance using the imblearn package to create synthetic data for the minority class and under-sampling the majority were successful in increasing our macro scores as seen in the following 2 graphs

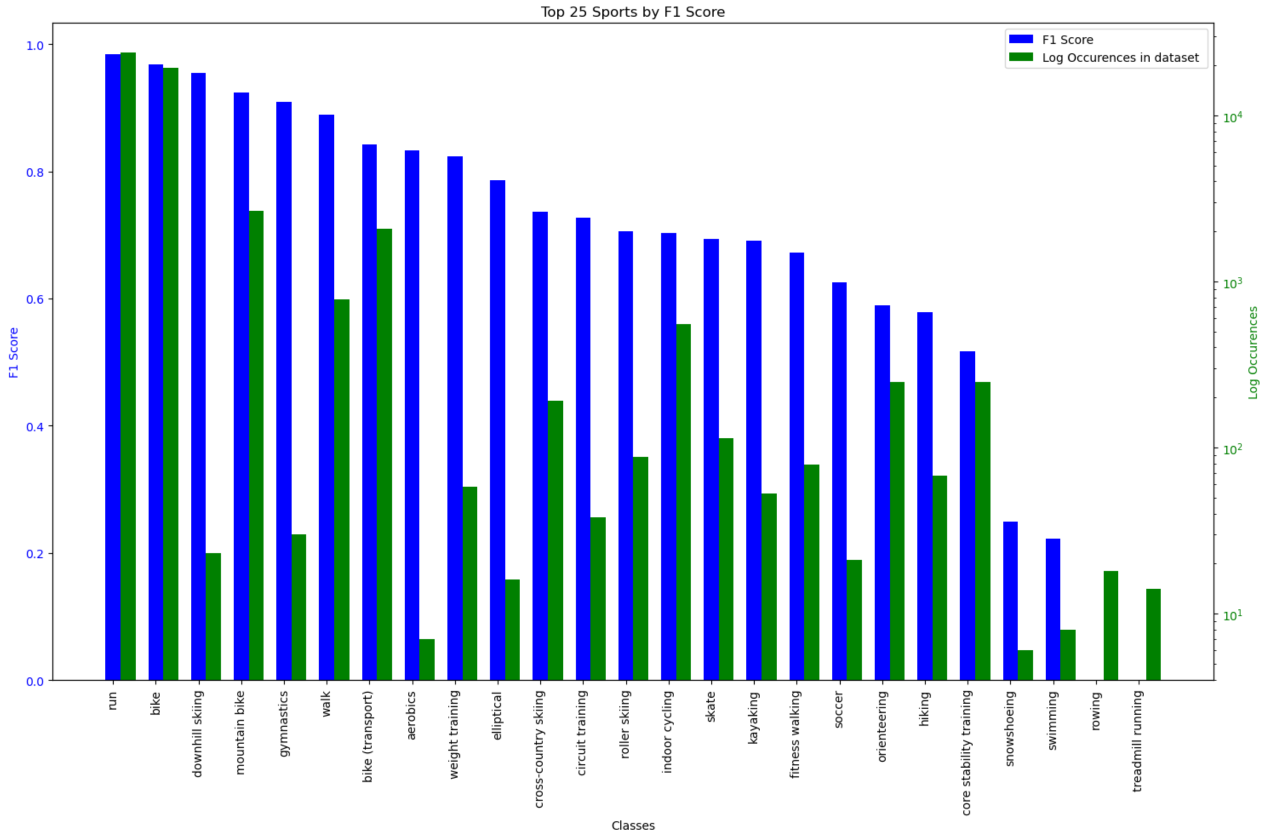

Figure 7: F1 score and occurrence trained on non-re-sampled data

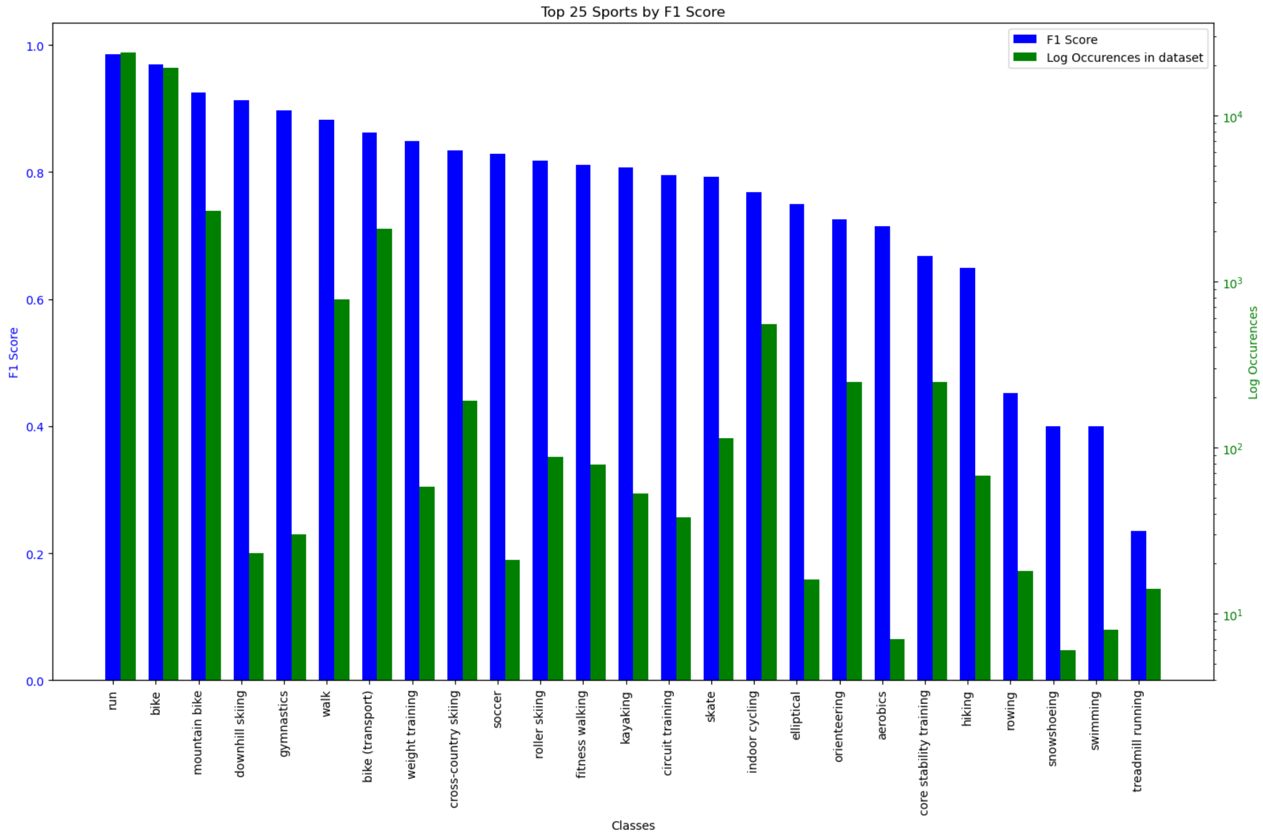

Figure 8: F1 score and occurrence trained on re-sampled data

We can see that in general (there are a few exceptions) as the number of occurrences increases the classifier tends to be better at classifying said class. without re-sampling 2 out of the 25 sports had f1 scores of 0, after re-sampling while the imbalance still exists it is lessened creating a more accurate model on the macro level.

Sports Recommendations and User-User connections

The recommendation model was able to successfully identify similar users as well as suggest an array of new sports to pursue for each user. The average number of recommendations per user was 4.192 new sports (Figure 9). It failed to produce a new sport recommendation for only 3.86% of users. Additionally, the model was able to produce four possible user-user connections for each user.

This model is a theoretical example of how personalized recommendations could be implemented in fitness data. Additional user input is required to fully evaluate the effectiveness of the recommendations produced by this model.

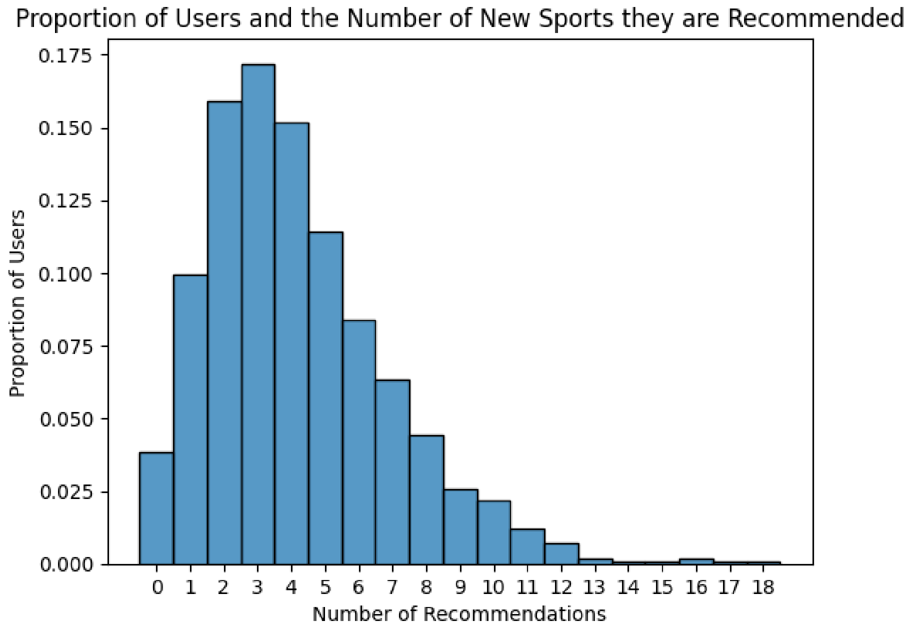

Figure 9: Number of Recommendations per user produced by the collaborative filtering model.

5. Discussion

In this paper, we present five different models with the goal of answering three different questions. First, we created both an OLS model and a deep artificial neural network to predict the average heart rate of a user given a variety of features. Heart rate is an important indicator of workout intensity, therefore accurate prediction is vital for personalized exercise regimens. Contrary to initial expectations, we found that the OLS model was able to predict heart rate more accurately than the neural network, however, both model's predictions converged around the mean average heart rate across all users. This suggests that additional features beyond ascent, descent, duration, distance, type of sport, and gender are needed to predict a user's average heart rate. Heart rate is more closely tied to more physiological features, such as a user's age, weight and height, which were not present in this dataset. Thus it is difficult to predict heart rate based solely on the workout metrics, and thus personalized workout regimens based on heart rate would require additional, more personal features. In future studies, incorporating information such as age, weight, and experience could lead to more accurate models. Furthermore, though fitness trackers are able to provide many metrics, they can be inconsistent due to different hardware/software limitations across the various devices.

In addition to heart rate, our research also focused on calorie expenditure. Calorie expenditure is another factor that is important to many users trying to lose or maintain weight. We used a similar OLS model to predict the log of calories burned throughout a workout. This model performed much better than the heart rate prediction model. This is likely because there is a strong relationship between calorie expenditure and workout duration. Furthermore, duration is an easily measurable metric across most fitness trackers, in contrast to highly variable and less accurate measurements such has heart rate. This likely contributed to consistency in the dataset and accuracy in calorie prediction. However, as mentioned in the results, there is evidence of heteroskedasticity, which affects the reliability of the model. Further studies could look into additional transformations or features, as well as other models such as weighted least squares, which does not assume uniform variance. This model could be very useful to users trying to count the total calories they burn. Users could potentially select or input a predetermined workout, and the model could output the predicted total calorie expenditure. Alternatively, users could input their desired calorie expenditure and be recommended appropriate workouts.

Additionally, using a random forest, we were able to predict the sport being performed given a user's exercise data at an accuracy of around 96%. This model could potentially be used to help recommend activities to users based on their desired workout features. Our model was generally more accurate for the sports with a large number of data points such as biking and running. In future studies, we could create a more accurate model curating the data in our train sets to have equal proportions of classes, the disparity in classes in the dataset is very severe so this might mean scaling back from 25 sports to only classifying 10. Additional techniques other than SMOTE (Synthetic Minority Over-sampling Technique) can be used as well like ADASYN (Adaptive Synthetic Sampling) which focuses specifically on classes that are hard to learn by adapting to the intrinsic data distribution using a density distribution as a criterion to automatically decide the number of synthetic samples to generate for each minority sample, focusing on the samples that are difficult to learn. Ensemble methods combining gradient boosting and k-nearest-neighbor (KNN) with the Random forest model could leverage the strengths of each to make a more well-rounded model as well. In the end, some sports may just be impossible to accurately classify due to differences in the people who do them, Hiking was the worst classified sports despite having relatively high occurrence which makes sense given the diversity of hiking trails and skill levels in hikers.

Finally, our collaborative filter model has successfully laid the foundation for recommending new sports and potential connections among users. The model first recognizes users with similar profiles, proposing them as potential workout partners or challengers to keep a given user consistent and motivated. these identified similar profiles are leveraged to recommend new sports for each user. This strategic setup ensures that recommendations encompass a variety of sports, including both less popular activities.

It is important to note that user input is required to validate and refine the model. Currently, it is impossible to determine if the recommendations the model makes are successful, but implementing the recommendations into the fitness tracker system could allow us to receive feedback on the model by measuring user responses. Incorporating a rating system where users have the option to rate the sports they have participated in or rate the workouts they completed could also help to produce a more accurate model.

Additionally, the integration of more data points and the incorporation of more extensive location information could further elevate the user-user connection aspect of our model. While virtual workout partnerships are beneficial, establishing connections among users in the same geographic area holds significant potential.

Overall, our research contributes to the evolving field of fitness technology, highlighting three important aspects of creating tailored workout plans. Predicting average heart rate and total calorie expenditure during a workout are essential to predicting/measuring a workout's intensity and overall impact on a person's health. Classifying a user's sport based on their workout is an essential step in being able to recommend sports for people with specific fitness goals. Finally, recommending new sports and connecting users based on their workout profiles adds a social and exploratory dimension to fitness which could keep users more engaged and consistent. Personalized fitness plans, informed by data-driven models, are a great well to help people reach their fitness goals by keeping them informed, consistent, and engaged.

References

[1] DE Warburton, CW Nicol, and SS Bredin. Health benefits of physical activity: The evidence. 174(6), pages 801–9, CMAJ, March 2006.

[2] SR Chekroud, R Gueorguieva, AB Zheutlin, M Paulus, HM Krumholz, JH Krystal, and AM Chekroud. Association between physical exercise and mental health in 1·2 million individuals in the usa between 2011 and 2015: A cross-sectional study. 5(9), pages 739–746, Lancet Psychiatry, September 2018.

[3] Y Du, B Liu, Y Sun, LG Snetselaar, RB Wallace, and W Bao. Trends in adherence to the physical activity guidelines for americans for aerobic activity and time spent on sedentary behavior among us adults, 2007 to 2016. 2(7), page e197597, JAMA Netw Open, July 2019.

[4] MSS Tang, K Moore, A McGavigan, RA Clark, and AN Ganesan. Effectiveness of wearable trackers on physical activity in healthy adults: Systematic review and meta-analysis of randomized controlled trials. 8(7), page e15576, JMIR Mhealth Uhealth, July 2020.

[5] Arash Mahyari and Peter Pirolli. Physical Exercise Recommendation and Success Prediction Using Interconnected Recurrent Neural Networks. In *2021 IEEE International Conference on Digital Health (ICDH)*, pages 148–153, September 2021.

[6] Jianmo Ni, Larry Muhlstein, and Julian McAuley. Modeling Heart Rate and Activity Data for Personalized Fitness Recommendation. In *The World Wide Web Conference*, WWW '19, pages 1343–1353, New York, NY, USA, May 2019. Association for Computing Machinery.