Article

A Machine Learning Approach to Finding the Fastest Race Course for Professional Athletes Competing in Ironman® 70.3 Races between 2004 and 2020

Mabliny Thuany [1], David Valero [2], Elias Villiger [3], Pedro Forte [4,5,6], Katja Weiss [7], Pantelis T. Nikolaidis [8], Marília Santos Andrade [9], Ivan Cuk [10], Caio Victor Sousa [11] and Beat Knechtle [7,12,*]

1   Faculty of Sports, University of Porto, 4200-450 Porto, Portugal
2   Ultra Sports Science Foundation, 69310 Pierre-Benite, France
3   Klinik für Allgemeine Innere Medizin, Kantonsspital St. Gallen, 9000 St. Gallen, Switzerland
4   CI-ISCE, Higher Institute of Educational Sciences of the Douro, 4560-708 Penafiel, Portugal
5   Research Center in Sports, Health and Human Development, 6201-001 Covilhã, Portugal
6   Department of Sport Sciences, Instituto Politécnico de Bragança, 5300-253 Bragança, Portugal
7   Institute of Primary Care, University Hospital Zurich, 8091 Zurich, Switzerland
8   School of Health and Caring Sciences, University of West Attica, 12243 Athens, Greece
9   Department of Physiology, Federal University of São Paulo, São Paulo 04021-001, Brazil
10 Faculty of Sport and Physical Education, University of Belgrade, 11000 Belgrade, Serbia
11 Health and Human Sciences, Loyola Marymount University, Los Angeles, CA 90045, USA
12 Medbase St. Gallen Am Vadianplatz, 9001 St. Gallen, Switzerland
*   Correspondence: beat.knechtle@hispeed.ch

Abstract: Our purpose was to find the fastest race courses for elite Ironman® 70.3 athletes, using machine learning (ML) algorithms. We collected the data of all professional triathletes competing between 2004 and 2020 in Ironman 70.3 races held worldwide. A sample of 16,611 professional athletes originating from 97 different countries and competing in 163 different races was thus obtained. Four different ML regression models were built, with gender, country of origin, and event location considered as independent variables to predict the final race time. For all the models, gender was the most important variable in predicting finish times. Attending to the single decision tree model, the fastest race times in the Ironman® 70.3 World Championship of around ~4 h 03 min would be achieved by men from Austria, Australia, Belgium, Brazil, Switzerland, Germany, France, the United Kingdom, South Africa, Canada, and New Zealand. Considering the World Championship is the target event for most professional athletes, it is expected that training is planned so that they attain their best performance in this event.

Keywords: endurance; cycling; half-distance Ironman; swimming; triathlon; running

1. Introduction

Triathlon is a multi-sports discipline. Long-distance triathlons, such as the 'Ironman® Hawaii' (3.8 km swimming, 180 km cycling, and 42.195 km running), are known to many people [1] and have increased in popularity in terms of participants [2]. In a half-Ironman ('Ironman® 70.3'), the swimming, cycling, and running segments are half the distance of the corresponding segments in a full Ironman® distance triathlon. Ironman® '70.3' refers to the total distance in miles (113.0 km) of the race, consisting of a 1.2 mile swim, a 56 mile bike ride, and a 13.1 mile run. Among different research topics involving endurance athletes, the interest in understanding the environmental characteristics related to the place of competition and country of origin of the athletes has increased over the last few years [3–6].

Regarding the country of origin of long-distance triathletes, it has been reported that most finishers, as well as the fastest athletes competing in 'Ironman® Hawaii', originated

from the United States of America [7]. Although the previous literature has showed the dominance of certain nations in specific sports disciplines [4,6,8,9], little evidence is available regarding the place of competition or if these characteristics are performance-related. As the most popular long-distance triathlon in the world, athletes travel hundreds of miles to compete in Ironman® 70.3 events, with many travelling thousands of miles to compete at remote locations and achieve personal goals or to earn a spot in the World Championship.

As previously showed, travelling may have a negative effect on athletes' performance [10], due to sleep disorders, nutritional and appetite changes, travel fatigue (i.e., hours spent seated and disruption in daily rhythm and routines). Considering these negative effects on athletes' cognitive and physical levels, specific treatments to surpass these effects are necessary [10,11]. Thus, looking past the athletes' country of origin, the characteristics and place of competitions can provide important information for athletes and coaches to better prepare and select the most appropriate race course in the world. Since no information is available regarding the role of the course characteristics in runners' performance, this study intended to find the fastest race courses for elite Ironman® 70.3 athletes, using machine learning (ML) algorithms for predictive statistics. Based on previous findings, we hypothesized that the USA would be the fastest country [7].

2. Materials and Methods

2.1. Ethical Approval

This study was approved by the Institutional Review Board of Kanton St. Gallen, Switzerland, with a waiver of the requirement for informed consent of the participants as the study involved the analysis of publicly available data (EKSG 01-06-2020).

2.2. Study Design, Sample, and Data Preparation

This is a cross-sectional study, using secondary data available and downloaded from the official Ironman® website (www.ironman.com, accessed on 22 June 2022) using a Python script. We analyzed elite (PRO) finishers race data of all Ironman® 70.3 races recorded in the Ironman website between 2004 and 2020. To register as an elite triathlete, an athlete must apply in the Ironman® 70.3 community to become a "PRO" member with a copy of the athlete's license status as a professional athlete from the respective national triathlon federation, and must also be affiliated with the International Triathlon Union (www.triathlon.org, accessed on 22 June 2022). Data obtained include the athletes' sex and country of origin, the event location and year, and the race finish time. Exclusion criteria were as follows: (i) athletes who did not start or finish, (ii) disqualified athletes, (iii) records with missing split time, and (iv) inconsistent times (i.e., impossible split times or final times smaller than split times, etc.).

2.3. Statistical Analysis

Descriptive statistics are presented using mean and standard deviation, frequencies, and percentages. Four ML regression models were built and compared, using four different algorithms: random forest regressor, XG boost regressor, cat boot regressor, and decision tree regressor. All four algorithms are decision-tree-based. The simplest of all of them is the single decision tree algorithm, whilst the other three are tree-ensembles that use between a few hundred and a few thousand trees. For each algorithm, the mean absolute error (MAE) and R-squared (R^2) metrics were calculated. In order to try and understand the algorithms' logic, several model-agnostic explainability tools were used, including the algorithms' features of relative importance, partial dependence plots (PDP), and the aggregated SHAP values. In addition, a graphical illustration of the single decision tree model is also provided. During the training, the model was fitted with 75% of the available data, and the remaining 25% was used to test and compare predictions vs. real values, calculating MAE and R^2 as the accuracy metrics of the algorithm. The MAE presents the mean of the absolute values of the individual prediction errors over all instances in the test set, in which higher values

Int. J. Environ. Res. Public Health **2023**, 20, 3619

3 of 12

mean higher prediction errors. R^2 is a value between 0 and 1 and refers to the "goodness of fit", that is, the model's predictive power. The features' relative importance represents a score for each feature in a specific model. Higher values correspond to higher importance in predicting the target variable. PDP and SHAP values provide information on how each feature contributes to the model output. Since the outcome (or predicted variable) was the finish time in seconds, the fastest courses were defined as the race locations where the average finish time was the lowest. The triathletes' gender and country of origin, as well as the event location (race course) were used as predictors. As three of them were categorical variables, they were encoded into numerical values. Gender was encoded as 0 = female, 1 = male, whilst for the country and event location, a rank-encoding scheme based on the frequency was used. Data processing and analysis were performed using Python and associated ML libraries in a Jupyter notebook (Google Colab).

3. Results

A total of 16,611 athletes of both sexes from 97 different countries in 163 different race locations were sampled. Figure 1 presents the histograms of the finish times for both sexes. Based on the histograms, it is possible to see a higher frequency for men than women, and that men present a higher frequency of lower finish times.
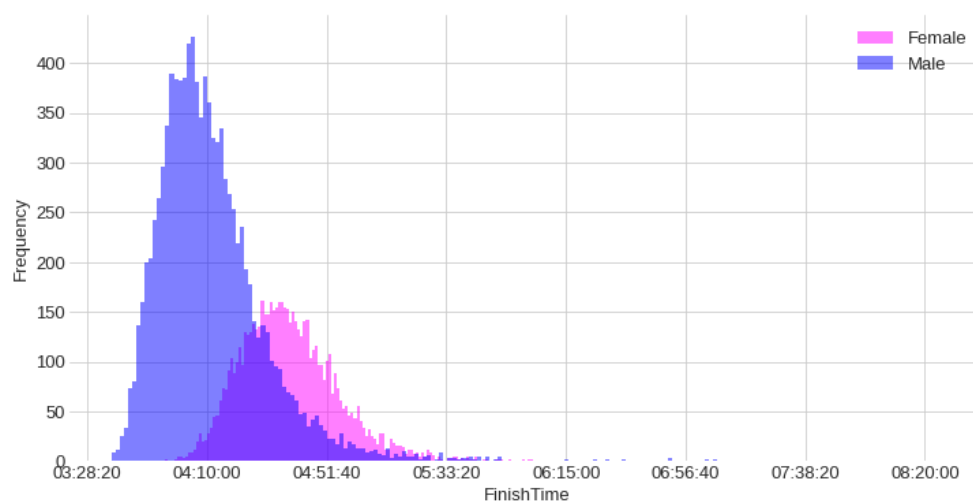

Figure 1. Histograms of the finish time of Ironman® 70.3 PRO athletes competing between 2004 and 2020.

Table 1 presents a comparative table with the performance scores (MAE and R^2) and the features' relative importance for all four predictive models considering the different algorithms and calculated using the 25% test sample. The XG boost and cat boost models were the models achieving the best predictive performance (R^2 of 0.55), while the simplest of the models (model 4, single decision tree) obtained the lowest score (R^2 of 0.43). In regards to the features' importance, gender was the most important variable to predict the finish time for all five models. However, event location and country of origin presented different importance for different algorithms. For example, event location presented a higher importance compared to country when the random forest regressor was used, whereas for the decision tree, the triathlete's country of origin was more important than the event location.

Figure 2 represents the aggregated SHAP values of each of the four models. The SHAP values show how each of the predictors contribute to the model output in respect to a reference value. For instance, the gender variable adds about 1000 s on average to the predicted finish time for females (blue points) and subtracts about 600 for males (red points). The relative low importance of the event location variable for the decision tree model can be appreciated in the respective chart, given the small size of its contributions

to the model output. However, a few blue points can be seen at the far-left of the chart (between 0 and −500 s), indicating that locations with low values for *EventLocation_ID*, which have the highest number of participants, would be among the fastest race locations. Similarly, a few red dots at the far-right end of the XG boost and cat boost charts indicate that some countries with a low number of participants would be the worst-performing ones, adding 3000 s or more to the predicted race finish time.

Table 1. Regression models results using the four different algorithms.

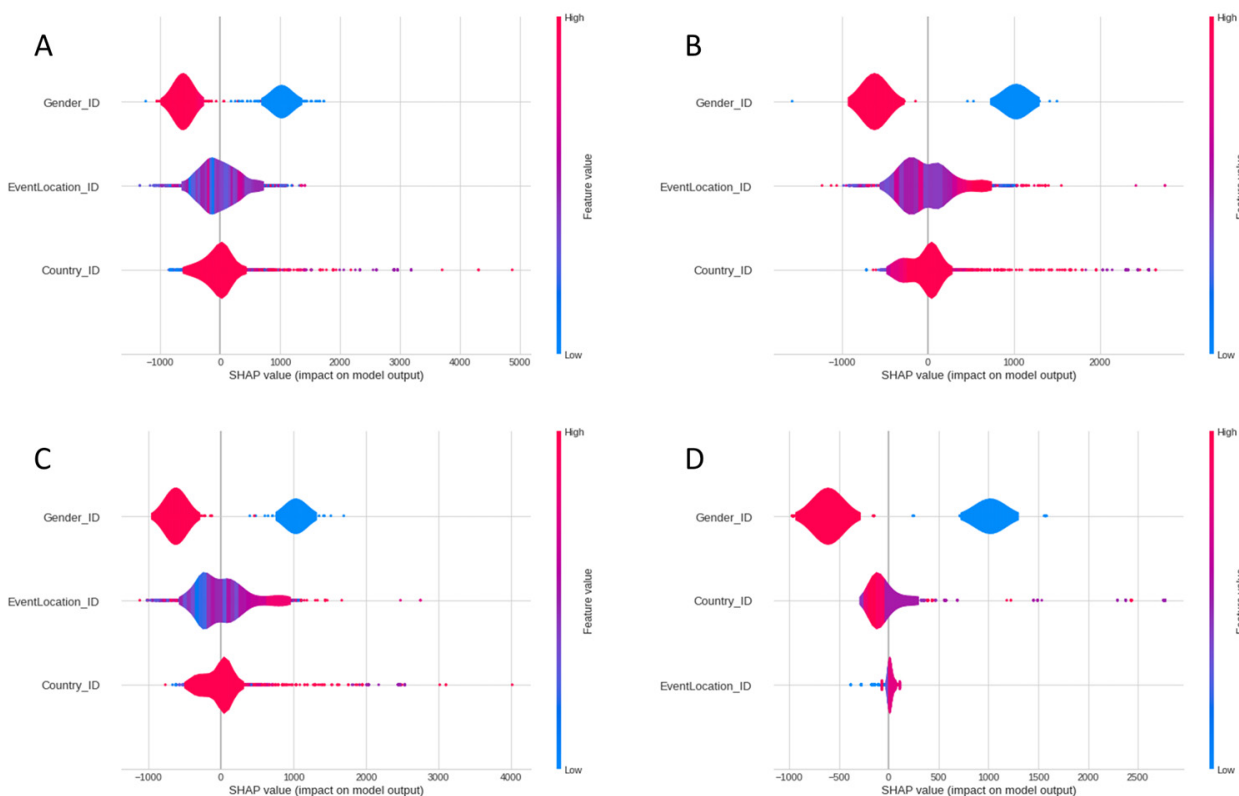|	Algorithm			
	Random Forest	**XG Boost**	**Cat Boost**	**Decision Tree**
MAE (seconds)	720.8	688.4	688.8	796.5
R-squared	0.51	0.55	0.55	0.43
Feature (variable)	Importance			
Gender	0.45	0.93	0.45	0.81
Event Location	0.32	0.04	0.30	0.01
Country	0.23	0.03	0.25	0.18

Legend: MAE (mean absolute error)

Figure 2. Aggregated SHAP values of the five predictive models: (**A**) Algorithm Random Forest Regressor; (**B**) Algorithm XG Boost Regressor; (**C**) Algorithm Cat Boot Regressor, (**D**) Algorithm Decision Tree Regressor).

Figures 3 and 4 show the partial dependence plots (PDP) of each predictive model. The PDP plots provide additional details on the specific locations and/or countries that more contribute to the models' predictions, as well as the amount of their contribution. Each of the charts show the contribution to the predictive model output for each value of each feature, with respect to a reference value of zero. For the gender predictor, 0 means females, whilst for country, a value of 0 corresponds to the USA, and for event location a value of 0 corresponds to the Ironman® 70.3 World Championship.
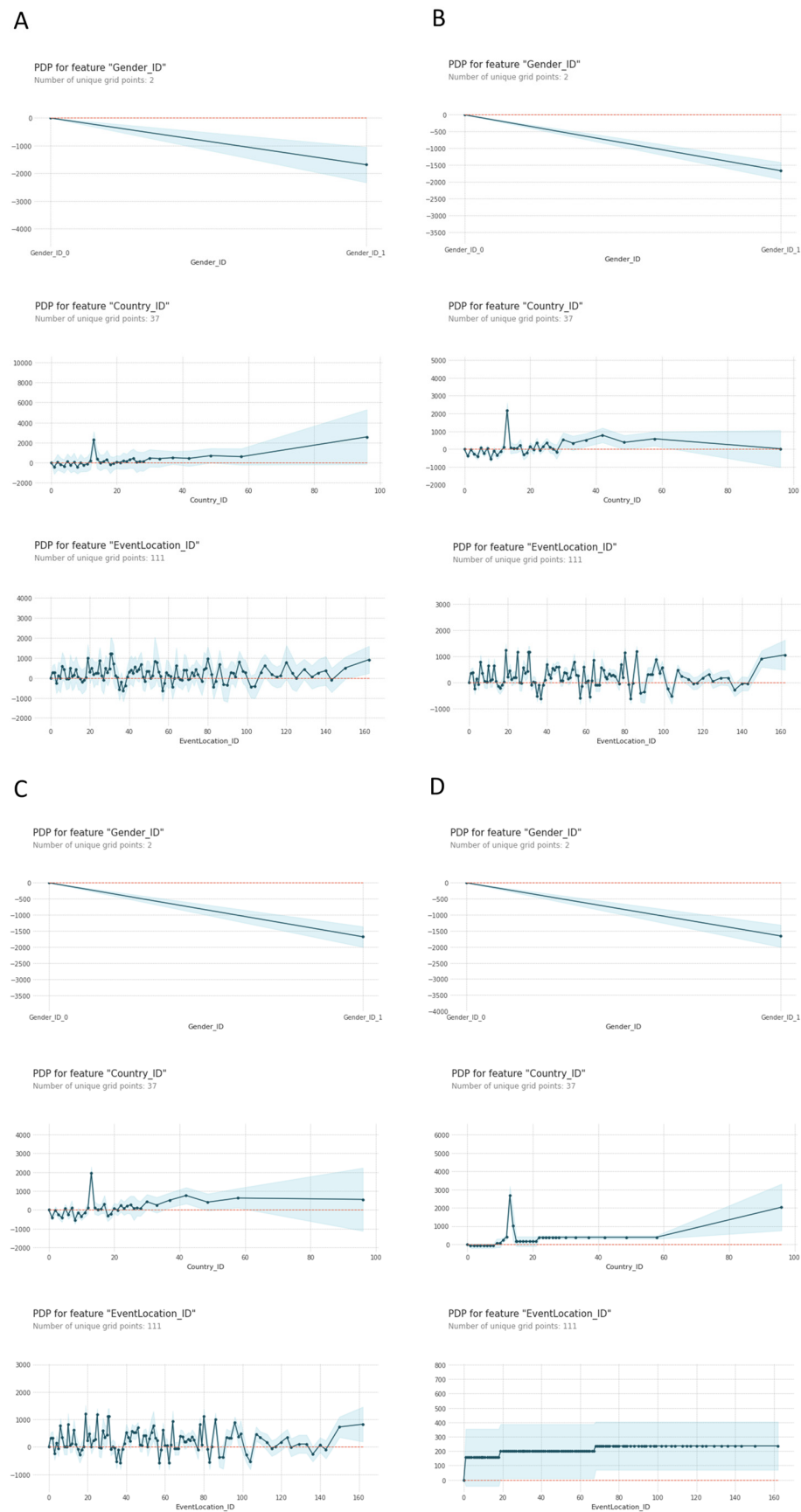
Int. J. Environ. Res. Public Health **2023**, *20*, 3619

5 of 12

Figure 3. PDP plots of PDP plots of (**A**) Algorithm Random Forest Regressor, (**B**) Algorithm XG Boost Regressor; (**C**) Algorithm Cat Boost Regressor, and (**D**) Algorithm Decision Tree.

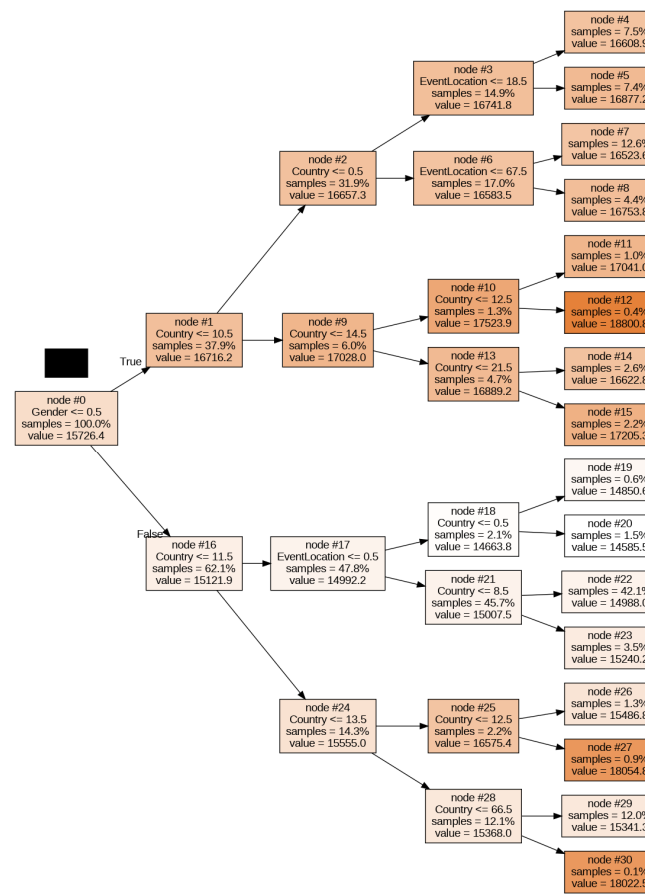*Int. J. Environ. Res. Public Health* **2023**, *20*, 3619

6 of 12

Figure 4. Decision tree (model 4) graphical representation.

All four models identify the Philippines as the worst-performing country (Country_ID = 13) adding around 2000 s on average (3000 in the case of the decision tree) to the finish time predictions with respect to the USA. Furthermore, the three ensemble algorithms identify several countries (IDs 2, 5, 9) as making negative contributions to the finish time predictions; however these contributions are relatively small (less than 1000 s) in all cases. The fourth model (single decision tree) shows a flatter country PDP chart which reflects the less complex nature of the algorithm.

In respect of the event location variable, the first three algorithms (the tree ensembles) show very similar charts with two distinctive lows at positions 35 and 37 corresponding to Ironman® 70.3 Monterrey and Dubai, subtracting around 600–800 s from the output in respect to the reference (the World Championship). The last algorithm (single decision tree) shows a simpler chart, corresponding to its simpler logic, in which all and every single race location will add either 160, 200, or 220 s (approximately) to the reference value. According to the decision tree then, the Ironman® 70.3 World Championship is the fastest race course.

Figure 4 represents the decision tree algorithm implemented as model 4. As already explained, all four predictive models are ensemble variations in the basic decision tree; hence, interpreting the latter is relevant to understanding all of them. Starting at node 0 (with gender encoded as 0 for women and 1 for men), the decision node *Gender_ID <= 0.5* leads to the women's results on the top branch (node #1) and men´s at the bottom (node #16). Note the color coding in which the darker colors represent longer finish times (represented by the "value" in seconds). As we are searching for the best times, we continue to the men's branch, in which decision node #16 *Country <= 11.5* splits the dataset into the first athletes' countries by the number of records (i.e., Austria, Australia, Belgium, Brazil, Switzerland, Germany, France, United Kingdom, United States, South Africa, Canada, and New Zealand in node #17) and the rest (node #24). The algorithm tells us at this point that 47.8% of the

Int. J. Environ. Res. Public Health **2023**, 20, 3619

7 of 12

records in the PRO sample were men from one of the first twelve countries by the number of participants, and that they achieved an average race time of 4 h 10 min (14,992.2 s). If we continue through node #18 (from node #17 when *EventLocation_ID < = 0.5*, that is, the Ironman® World Championship) and then to node #20 (when *Country_ID > 0.5*, and is less or equal than 11 as determined earlier) we reach the fastest (whitest) node in this decision tree which corresponds to male triathletes from countries with IDs from 1 to 11 (Austria, Australia, Belgium, Brazil, Switzerland, Germany, France, United Kingdom, South Africa, Canada, and New Zealand) competing in the Ironman® World Championship location and achieving an average time of 14,585.5 s (4 h 3 min) (Figures 5 and 6).
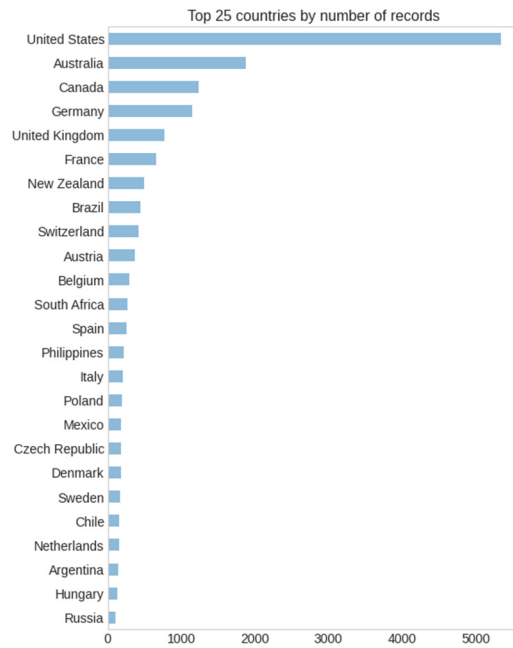

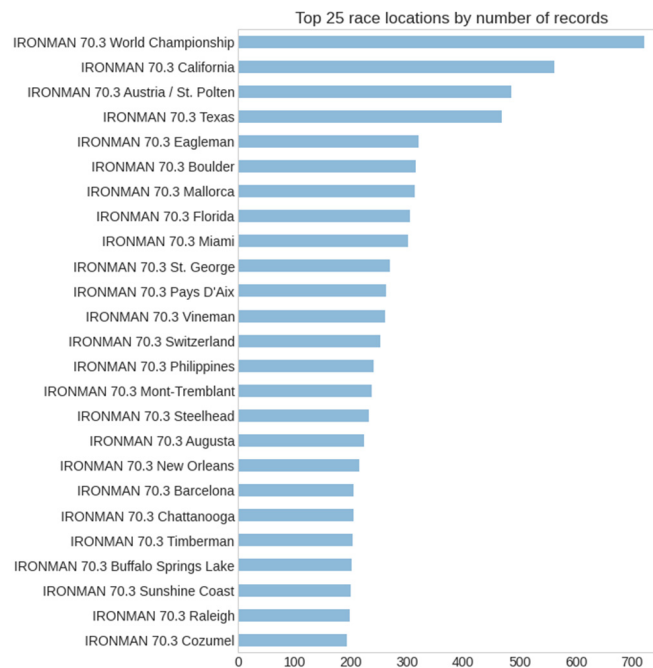**Figure 5.** Top 12 countries by number of records.

Figure 6. Top 10 locations by number of records.

4. Discussion

This study is the first to investigate the location of the fastest race course for Ironman® 70.3 using ML algorithms. The first important finding was that the best race times in professional athletes in Ironman® 70.3 competitions were achieved by men competing in the Ironman® 70.3 World Championship. The World Championship is the target event for most professional athletes, so it is expected that athletes attain their best performance in this event. In addition, only the best-placed athletes in the Ironman® 70.3 events can participate in the World Championship, which greatly raises the level of competition and competitiveness among the athletes. Since few studies have used a ML approach to predict performance or different outcomes in triathletes [12,13], comparisons are limited. Nevertheless, we expect that our data's accuracy in the prediction of results could also guide future studies. Moreover, it is important to mention that the actual race course of the Ironman® 70.3 World Championship changed between 2006 and 2020 (see Table 2). Regarding the average temperature (see Table 3), no fundamental differences exist between the various race courses; however, some differences according the month of the event as well as differences for the swimming, cycling, and running segments could be seen. These characteristics present some insights regarding the interplay between the intrapersonal (i.e., motivation), interpersonal (i.e., competitive environment), and environmental features (i.e., different race courses) for competitive athletes. Although these factors were beyond our scope in this study, future studies should investigate how these factors interact, and how performance can be conditioned by this interaction.

Table 2. Locations of Ironman® 70.3 World Championships.

Year	City, Country
2006–2010	Clearwater, FL, USA
2011–2013	Henderson, NV, USA
2014	Mont-Tremblant, QC, Canada
2015	Zell am See-Kaprun, Austria
2016	Mooloolaba, QLD, Australia
2017	Chattanooga, TN, USA
2018	Nelson Mandela Bay, South Africa
2019	Nice, Alpes-Maritimes, France
2020	Taupo, New Zealand

Table 3. Specific descriptions of Ironman® 70.3 race courses.

Race	Month	Average Temperature	Swimming Course Characteristics	Cycling Course Characteristics	Running Course Characteristics
IRONMAN® 70.3 Zell am See	August	19 °C	1 lap of 1.9 km in a lake	1 lap of 90 km, altitude difference of 1270 m per lap, overall altitude difference of 1270 m	2 laps of 21.1 km, altitude difference of 26 m per lap, at an overall altitude difference of 52 m
IRONMAN® 70.3 European Championship Elsinore	June	21 °C	1 lap of 1.9 km in a bay	1 lap of 90 km, altitude difference of 600 m per lap, overall altitude difference of 600 m	3.5 laps of 21.1 km, altitude difference of 28.6 m per lap, at an overall altitude difference 100 m
IRONMAN® 70.3 Greece Costa Navarino	October	23 °C	1 lap of 1.9 km in the ocean	2 laps of 90 km, altitude difference of 540 m per lap, overall altitude difference of 1080 m	3 laps of 21.1 km, altitude difference of 66.7 m per lap, at an overall altitude difference 200 m
IRONMAN® 70.3 Indian Wells La Quinta	December	18 °C	1 lap of 1.9 km in a lake	1 lap of 90 km, altitude difference of 161 m per lap, overall altitude difference of 161 m	2 laps of 21.1 km, altitude difference of 57.7 m per lap, at an overall altitude difference of 115 m

Table 3. *Cont.*

Race	Month	Average Temperature	Swimming Course Characteristics	Cycling Course Characteristics	Running Course Characteristics
IRONMAN® 70.3 Middle East Championship Bahrain	December	22 °C	1 lap of 1.9 km in a bay	1 lap of 90 km, altitude difference of 300 m per lap, overall altitude difference of 300 m	3 laps of 21.1 km, altitude difference of 24.3 m per lap, at an overall altitude difference 73 m
IRONMAN® 70.3 Texas	April	21 °C	1 lap of 1.9 km in a bay	1 lap of 90 km, altitude difference of 45 m per lap, overall altitude difference of 45 m	3 laps of 21.1 km, altitude difference of 10.7 m per lap, at an overall altitude difference 32 m
IRONMAN® 70.3 Tallinn	August	17 °C	1 lap of 1.9 km in a lake	1 lap of 90 km, altitude difference of 300 m per lap, overall altitude difference of 300 m	2 laps of 21.1 km, altitude difference of 57.7 m per lap, at an overall altitude difference of 115 m
IRONMAN® 70.3 Western Sydney	September	23 °C	1 lap of 1.9 km in a lake	2 laps of 90 km, altitude difference of 165 m per lap, overall altitude difference of 330 m	1 lap of 21.1 km, altitude difference of 25 m per lap, at an overall altitude difference of 25 m
IRONMAN® 70.3 Maceio	August	28 °C	1 lap of 1.9 km in the ocean	1 lap of 90 km, altitude difference of 200 m per lap, overall altitude difference of 200 m	3 laps of 21.1 km, altitude difference of 23.3 m per lap, at an overall altitude difference 70 m
IIRONMAN® 70.3 Liuzhou	September	28 °C	1 lap of 1.9 km in a river	2 laps of 90 km, altitude difference of 168.5 m per lap, overall altitude difference of 337 m	2 laps of 21.1 km, altitude difference of 171 m per lap, at an overall altitude difference of 342 m
IRONMAN® 70.3 Luxembourg	June	18 °C	1 lap of 1.9 km in a river	1 lap of 90 km, altitude difference of 580 m per lap, overall altitude difference of 580 m	2,5 laps of 21.1 km, altitude difference of 40 m per lap, at an overall altitude difference of 100 m
IRONMAN® 70.3 Marbella	May	21 °C	1 lap of 1.9 km in the ocean	1 lap of 90 km, altitude difference of 1400 m per lap, overall altitude difference of 1400 m	2 laps of 21.1 km, altitude difference of 25 m per lap, at an overall altitude difference of 50 m
IRONMAN® 70.3 Vichy	August	28 °C	1 lap of 1.9 km in a lake	1 lap of 90 km, altitude difference of 1000 m per lap, overall altitude difference of 1000 m	2 laps of 21.1 km, altitude difference of 50 m per lap, at an overall altitude difference of 100 m
IRONMAN® 70.3 Gdynia	August	25 °C	1 lap of 1.9 km in a bay	1 lap of 90 km, altitude difference of 1860 m per lap, overall altitude difference of 1860 m	2 laps of 21.1 km, altitude difference of 70 m per lap, at an overall altitude difference of 340 m
IRONMAN® 70.3 Mallorca	May	25 °C	1 lap of 1.9 km in the ocean	1 lap of 90 km, altitude difference of 850 m per lap, overall altitude difference of 850 m	3 laps of 21.1 km, altitude difference of 20 m per lap, at an overall altitude difference 60 m
IRONMAN 70.3 Panama	March	30 °C	1 lap of 1.9 km in the ocean	3 laps of 90 km	3 laps of 21.1 km
IRONMAN 70.3 Augusta	September	27 °C	1 lap of 1.9 km in the river	1 lap of 90 km	2 laps of 21.1 km
IRONMAN 70.3 North Carolina	October	18 °C	1 lap of 1.9 km in the ocean	1 lap of 90 km	1 lap of 21.1 km
IRONMAN 70.3 Dubai	March	24 °C	1 lap of 1.9 km in the ocean	1 lap of 90 km, altitude difference of 87 m per lap, overall altitude difference of 87 m	1,5 laps of 21.1 km, altitude difference of 5 m per lap, at an overall altitude difference 7.5 m
IRONMAN 70.3 Sao Paolo	September	25 °C	1 lap of 1.9 km in a reservoir	2 laps of 90 km, altitude difference of 20 m per lap, overall altitude difference of 40 m	3 laps of 21.1 km, altitude difference of 15 m per lap, at an overall altitude difference 45 m

Int. J. Environ. Res. Public Health **2023**, 20, 3619

10 of 12

Table 3. *Cont.*

Race	Month	Average Temperature	Swimming Course Characteristics	Cycling Course Characteristics	Running Course Characteristics
IRONMAN 70.3 Alagolas	August	28 °C	1 lap of 1.9 km in the ocean	1 lap of 90 km, altitude difference of 100 m per lap, overall altitude difference of 100 m	3 laps of 21.1 km, altitude difference of 20 m per lap, at an overall altitude difference 60 m
IRONMAN 70.3 Emilia Romagna	September	25 °C	1 lap of 1.9 km in the ocean	1 lap of 90 km, altitude difference of 185 m per lap, overall altitude difference of 185 m	3 laps of 21.1 km, altitude difference of 5 m per lap, at an overall altitude difference 15 m
IRONMAN 70.3 Pays D'Aix	May	20 °C	1 lap of 1.9 km in a lake	1 lap of 90 km, altitude difference of 390 m per lap, overall altitude difference of 390 m	3 laps of 21.1 km, altitude difference of 15 m per lap, at an overall altitude difference 45 m
IRONMAN 70.3 Sunshine Coast	September	26 °C	1 lap of 1.9 km in the ocean	2 laps of 90 km, altitude difference of 35 m per lap, overall altitude difference of 70 m	2 laps of 21.1 km, altitude difference of 20 m per lap, at an overall altitude difference 40 m
IRONMAN 70.3 Geelong	March	19 °C	1 lap of 1.9 km in a bay	2 laps of 90 km, altitude difference of 75 m per lap, overall altitude difference of 150 m	2,5 laps of 21.1 km, altitude difference of 25 m per lap, at an overall altitude difference 67 m
IRONMAN 70.3 Steelhead	Juna	20 °C	1 lap of 1.9 km in a lake	1 lap of 90 km	2 laps of 21.1 km
IRONMAN 70.3 Turkey	November	24 °C	2 laps of 1.9 km in the ocean	2 laps of 90 km, altitude difference of 20 m per lap, overall altitude difference of 40 m	3 laps of 21.1 km, altitude difference of 5 m per lap, at an overall altitude difference 15 m
IRONMAN 70.3 California	April	17 °C	1 lap of 1.9 km in the ocean	1 lap of 90 km, altitude difference of 220 m per lap, overall altitude difference of 220 m	2 laps of 21.1 km, altitude difference of 10 m per lap, at an overall altitude difference 20 m
IRONMAN 70.3 Astana	June	25 °C	1 lap of 1.9 km in a river	1 lap of 90 km, altitude difference of 20 m per lap, overall altitude difference of 20 m	2 laps of 21.1 km, altitude difference of 5 m per lap, at an overall altitude difference 10 m
IRONMAN 70.3 Florianapolis	April	21 °C	1 lap of 1.9 km in the ocean	1 lap of 90 km	3 laps of 21.1 km

Another interesting finding was that the locations with the highest number of participants would be among the fastest race locations. Factors that explain these findings may be related to the race course, the difficulty in accessing the race event, weather characteristics, and personal motivation, since a larger number of athletes compete in these events. Since these are speculative, future studies should better explore these findings in order to achieve a deeper understanding of the influence of each of these effects.

Our study presents some advantages for performance prediction compared to other studies which have largely focused on the role of training [14], pacing [15,16], technical adjustments [16], and nutritional and morphological characteristics [17] in endurance athletes. The importance of the event location for athletes' performance has not been sufficiently explored in the scientific literature [18,19]. Besides the influence of athletes and equipment characteristics on Ironman® 70.3 triathlon performance [20,21], as an outdoor sport, the role of environmental characteristics needs to be considered, moving beyond the weather characteristics frequently studied (i.e., temperature, wind, humidity) [22]. In this way, our results indicate that the place of competition is an important variable to be considered for athletes to achieve their best performance. In addition, altitude, terrain characteristics, water temperature, sea current, and track characteristics influence performance, and are emergent topics that also need to be considered in the future.

Finally, all our predictive models showed that the best performance would be achieved by men. Despite the reduction in gender gaps in performance over the last few years, men still perform better than women [23]. Gender differences have been previously investigated [24,25] and performance differences range between 12% to 18%, depending on race characteristics [24]. In summary, morphological characteristics (e.g., lower limbs, body fat, body mass) and physiological determinants (e.g., maximal oxygen uptake, lactate threshold, movement economy) [23] were contextualized to explain the differences in the swimming, cycling, and running segments that reduce women's performance [23]. Due to the large body of the literature available regarding sex performance differences, it is important to move beyond this topic. Future studies should investigate the role of the race course, in addition to the training background and intrapersonal motivation, in each group, as well as different performance levels and age categories. As mass sporting events have increased over the last few years and participants often compete in their leisure time, it would be desirable to obtain a good understanding of the impact of these race characteristics on their well-being.

It is acknowledged that this study is not free from limitations. The first important limitation was the lack of information regarding the race courses that could explain why the Ironman® 70.3 World Championship presented the best characteristics for athletes' performance, considering the different characteristics of the triathlon disciplines. The lack of information about the race course in each segment impaired our understanding of which discipline (swimming, cycling, running) would be the most positively affected by the race course characteristics and would have the strongest influence on the athletes' finish times. Secondly, information regarding changes in race course characteristics over the years was not taken into consideration during the data analysis. The absence of additional information about the athletes' personal characteristics, such as motivation, training background, and travel distance, also limits the generalization of these findings. However, our results bring valuable insights about the importance of the event location in performance and introduce some advances regarding the importance of statistical approaches in helping with athletes' preparation.

5. Conclusions

Between 2004 and 2020, most professional Ironman® 70.3 athletes achieved their fastest race times in the Ironman® 70.3 World Championship with an average final time of around 4 h. Future studies should consider stratification based on the performance level, as well as understand the effects of racecourses in helping athletes achieve their best results.

References

1. Lepers, R. Analysis of Hawaii ironman performances in elite triathletes from 1981 to 2007. *Med. Sci. Sports Exerc.* **2008**, *40*, 1828–1834. [CrossRef] [PubMed]

Int. J. Environ. Res. Public Health **2023**, *20*, 3619

12 of 12

2. Ironman. 170+ Triathlon Worldwide. Available online: https://www.ironman.com/im703-races (accessed on 10 November 2022).

3. Thuany, M.; Weiss, K.; Villiger, E.; Scheer, V.; Ouerghi, N.; Gomes, T.N.; Knechtle, B. A macro to micro analysis to understand performance in 100-mile ultra-marathons worldwide. *Sci. Rep.* **2023**, *13*, 1415. [CrossRef]

4. Knechtle, B.; Rosemann, T.; Nikolaidis, P. The Role of Nationality in Ultra-Endurance Sports: The Paradigm of Cross-Country Skiing and Long-Distance Running. *Int. J. Environ. Res. Public Health* **2020**, *17*, 2543. [CrossRef] [PubMed]

5. Knechtle, B.; Nikolaidis, P.T.; Valeri, F. Russians are the fastest 100-km ultra-marathoners in the world. *PLoS ONE* **2018**, *13*, e0199701. [CrossRef]

6. Rüst, C.A.; Bragazzi, N.L.; Signori, A.; Stiefel, M.; Rosemann, T.; Knechtle, B. Nation related participation and performance trends in 'Norseman Xtreme Triathlon' from 2006 to 2014. *Springer Plus* **2015**, *4*, 469. [CrossRef] [PubMed]

7. Dähler, P.; Rüst, C.A.; Rosemann, T.; Lepers, R.; Knechtle, B. Nation related participation and performance trends in 'Ironman Hawaii' from 1985 to 2012. *BMC Sports Sci. Med. Rehabil.* **2014**, *6*, 16. [CrossRef] [PubMed]

8. Irving, R.; Charlton, V.; Morrison, E.; Facey, A.; Buchanan, O. Demographic characteristics of world class Jamaican sprinters. *Sci. World J.* **2013**, *2013*, 670217. [CrossRef] [PubMed]

9. Marc, A.; Sedeaud, A.; Schipman, J.; Jacquemin, J.A.; Saulière, G.; Kryger, K.O.; Toussaint, J.F. Geographic enrolment of the top 100 in athletics running events from 1996 to 2012. *J. Sports Med. Phys. Fit.* **2017**, *57*, 418–425. [CrossRef]

10. Lee, A.; Galvez, J.C. Jet Lag in Athletes. *Sport. Health* **2012**, *4*, 211–216. [CrossRef]

11. Leatherwood, W.; Dragoo, J. Effect of airline travel on performance: A review of the literature. *Br. J. Sport. Med.* **2013**, *47*, 561. [CrossRef]

12. Martínez-Gramage, J.; Albiach, J.P.; Moltó, I.N.; Amer-Cuenca, J.J.; Huesa Moreno, V.; Segura-Ortí, E. A Random Forest Machine Learning Framework to Reduce Running Injuries in Young Triathletes. *Sensors* **2020**, *20*, 6388. [CrossRef] [PubMed]

13. Ofoghi, B.; Zeleznikow, J.; Macmahon, C.; Rehula, J.; Dwyer, D.B. Performance analysis and prediction in triathlon. *J. Sport. Sci.* **2016**, *34*, 607–612. [CrossRef] [PubMed]

14. Etxebarria, N.; Mujika, I.; Pyne, D.B. Training and Competition Readiness in Triathlon. *Sports* **2019**, *7*, 101. [CrossRef] [PubMed]

15. Wu, S.; Peiffer, J.; Brisswalter, J.; Nosaka, K.; Lau, W.; Abbiss, C. Pacing strategies during the swim, cycle and run disciplines of sprint, Olympic and half-Ironman triathlons. *Eur. J. Appl. Physiol.* **2015**, *115*, 1147–1154. [CrossRef]

16. Wu, S.; Peiffer, J.; Brisswalter, J.; Nosaka, K.; Abbiss, C. Factors influencing pacing in triathlon. *Open Access J. Sports Med.* **2014**, *5*, 223–234. [CrossRef]

17. Bentley, D.; Cox, G.; Green, D.; Laursen, P. Maximising performance in triathlon: Applied physiological and nutritional aspects of elite and non-elite competitions. *J. Sci. Med. Sport* **2008**, *11*, 407–416. [CrossRef]

18. Hollings, S.C.; Hopkins, W.G.; Hume, P.A. Environmental and venue-related factors affecting the performance of elite male track athletes. *Eur. J. Sport Sci.* **2012**, *12*, 201–206. [CrossRef]

19. Snyder, K.L.; Hoogkamer, W.; Triska, C.; Taboga, P.; Arellano, C.J.; Kram, R. Effects of course design (curves and elevation undulations) on marathon running performance: A comparison of Breaking 2 in Monza and the INEOS 1:59 Challenge in Vienna. *J. Sports Sci.* **2021**, *39*, 754–759. [CrossRef]

20. Knechtle, B.; Knechtle, R.; Stiefel, M.; Zingg, M.; Rosemann, T.; Rüst, C. Variables that influence Ironman triathlon performance— what changed in the last 35 years? *Open Access J. Sport. Med.* **2015**, *6*, 277–290. [CrossRef] [PubMed]

21. Knechtle, B.; Zingg, M.; Rosemann, T.; Stiefel, M.; Rüst, C. What predicts performance in ultra-triathlon races?—A comparison between Ironman distance triathlon and ultra-triathlon. *Open Access J. Sport. Med.* **2015**, *6*, 149–159. [CrossRef]

22. Chabert, C.; Hermand, E.; Hue, O. *Triathlon and Ultra-Endurance Events in Tropical Environments*; Springer International Publishing: Berlin/Heidelberg, Germany, 2019; pp. 283–296. [CrossRef]

23. Lepers, R. Sex Difference in Triathlon Performance. *Front. Physiol.* **2019**, *10*, 973. [CrossRef] [PubMed]

24. Rüst, C.A.; Knechtle, B.; Rosemann, T.; Lepers, R. Sex difference in race performance and age of peak performance in the Ironman Triathlon World Championship from 1983 to 2012. *Extrem. Physiol. Med.* **2012**, *1*, 15. [CrossRef] [PubMed]

25. Lepers, R.; Stapley, P.J. Differences in gender and performance in off-road triathlon. *J. Sport. Sci.* **2010**, *28*, 1555–1562. [CrossRef] [PubMed]