Pace my race: recommendations for marathon running

Conference Paper · September 2019

DOI: 10.1145/3298689.3346991

Jakim Berndsen
University College Dublin

Pace My Race: Recommendations for Marathon Running

Jakim Berndsen
jakim.berndsen@insight-centre.org
Insight Centre for Data Analytics
Belfield, Dublin

Barry Smyth
barry.smyth@insight-centre.org
Insight Centre for Data Analytics
Belfield, Dublin

Aonghus Lawlor
aonghus.lawlor@insight-centre.org
Insight Centre for Data Analytics
Belfield, Dublin

ABSTRACT

We propose marathon running as a novel domain for recommender systems and machine learning. Using high-resolution marathon performance data from multiple marathon races ($n = 7931$), we build in-race recommendations for runners. We show that we can outperform the existing techniques which are currently employed for in-race finish-time prediction, and we demonstrate how such predictions may be used to make real time recommendations to runners. The recommendations are made at critical points in the race to provide personalised guidance so the runner can adjust their race strategy. Through the association of model features and the expert domain knowledge of marathon runners we generate explainable, adaptable pacing recommendations which can guide runners to their best possible finish time and help them avoid the potentially catastrophic effects of *hitting the wall*.

CCS CONCEPTS

• **Information systems → Recommender systems**; • **Human-centered computing → Social recommendation**.

KEYWORDS

recommender systems, running, marathon pacing, sports analytics

ACM Reference Format:
Jakim Berndsen, Barry Smyth, and Aonghus Lawlor. 2019. Pace My Race: Recommendations for Marathon Running. In *Thirteenth ACM Conference on Recommender Systems (RecSys '19),September 16–20, 2019, Copenhagen, Denmark.* ACM, New York, NY, USA, 5 pages. https://doi.org/10.1145/3298689.3346991

1 INTRODUCTION

Over the past decades marathons have become mass participation events with the most popular races attracting upwards of 40,000 runners annually. The proliferation of smart, GPS-enabled devices in recent years has lead to a large increase in data surrounding marathon training and performance. This data, coupled with the growing interest amongst runners to improve all facets of their marathon performance opens up a number of new avenues for recommendation, along with a large cohort of potential target users.

The use of running data in the marathon is nothing new. Work by Riegel [17] from as early as 1981 is still in use today that allows a runner to predict their marathon finish time based on the finish time they have achieved in a previous, shorter race. These finish time predictors are still in common use by first time marathon runners to determine a goal time and they play an important role in informing the race strategy. Training data has also been used to make such a prediction [18], offering a different avenue from which a runner can build a race strategy. Other work has shown how to make explicit pacing recommendations for the marathon. Given a runner that has previously finished one or more marathons, a realistic personal best can be recommended to that runner, alongside an exact pacing plan that has previously worked for similar runners attempting to achieve their personal best times [19, 20].

Beyond pacing, recommender systems have shown promise in providing training programmes in the form of an e-coach. Training sessions can be recommended by mimicking the actions of a coach through estimation of a runner's abilities which drive the expectation of a runner's performance in a given session [15]. Social recommendations have also been utilised to schedule training sessions, by finding athletes with similar abilities for a runner to train with [8]. Human-in-the-loop systems have also shown promise in the e-coaching domain. Models have been built that predict whether a runner is likely to lose motivation [16]; or determine whether a training session was of substandard quality [4]; allowing an expert coach to intervene before any adverse consequences occur that may affect a runner's outcome in a race.

However, while various avenues of prediction and recommendation have been explored for the marathon the use of these systems have been largely limited to the period before the start of the race. Once the marathon race begins, training and pace recommendation can no longer be altered with current systems, and even in the situation where session quality can be evaluated - a system that could be extended to determine race quality - this evaluation only occurs after a session has finished. No known work has been done to use recommender systems to guide and direct runners *during* the running of the race, despite the finding that the pacing decisions an athlete makes during the race can account for as much as 15% of their marathon finish time [1]. The only known use of in-race prediction is currently a crudely predicted estimated finish time, whereby the current pace is scaled up to the full marathon distance. Runners tend to have no access to this prediction during the race and it is mainly used by organisers as a guide for spectators when certain athletes are likely to finish the race.

While a preordained strategy is certainly helpful to a runner there is little to help a runner should they need to deviate from this strategy. Starting too fast, for example, can lead to the depletion of energy reserves before the end of the race which causes a premature slowdown and can even lead to a runner *hitting the wall* [21]. In such situations an adaptable in-race recommendation is required to successfully steer the athlete to the finish line, by

RecSys '19, September 16–20, 2019, Copenhagen, Denmark

ACM ISBN 978-1-4503-6243-6/19/09…$15.00
https://doi.org/10.1145/3298689.3346991

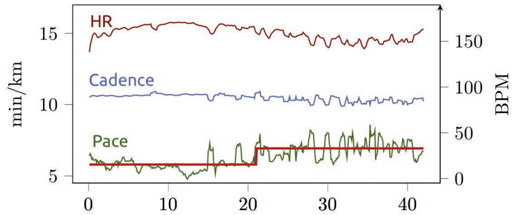

Figure 1: Race profile of heart rate, cadence and pace for a sample athlete. A positive split is highlighted for pace (the second half of the race is slower than the first half).

avoiding such a slowdown. The aim of this paper is to determine how a recommender system can be used to help runners in such a situation during the race. We use anonymised data from $n = 7931$ marathon finishers. This data has been collected from users of the Strava app[1] and contains information on a runner's pace, heart rate (HR) and cadence during the marathon. This data will be used to determine the physical state of a user during the marathon with the aim to: (a) improve the current in-race finish time predictions and allow runners to determine if they are on track for their desired finish time; and (b) make a recommendation, if deemed necessary, of a personalised race pacing strategy for the remainder of the race to guide the runner to the finish line without slowing down or hitting the wall.

2 FINISH TIME PREDICTION

2.1 Data Generation

Our initial data set comprises $n = 13000$ runners who have completed one of the New York, London or Dublin marathons, and for each of these we have pacing, HR and cadence data. Pace, the inverse of speed, is measured in minutes per kilometer with higher values indicating slower speeds, and gives an indication of how fast the runner is moving [7, 11]. HR, measured in beats per minute, can be used to determine how hard a runner is working to achieve their current level of performance [13]. Finally, cadence, the number of steps taken per minute, can be used to gauge an athlete's running form and efficiency [10]. Each race is sampled at 100 meter intervals and we have the average value of each feature (pace, HR, cadence) over that interval at each point in the race.

We remove faulty traces (eg. GPS errors) from the dataset, injured runners or unusual race strategies, such as extended periods of walking and truncate the profiles at 42.2km. After cleaning we are left with $n = 7931$ runners with a mean finish time of 230($\pm$40) minutes.

In order to generate data from which we can build a predictive model we extract features from these time series. We take the mean of the pace, HR, and cadence over multiple window sizes, namely 1km, 5km and the full race to that point. These windows correspond to short, medium and long term efforts of a runner which can not only be used to track changes in pacing, physical effort or running form, but also can be related to the types and lengths of intervals runners may currently use to evaluate their own performance during a run.

[1]Data, in part, provided by Strava under limited research license.

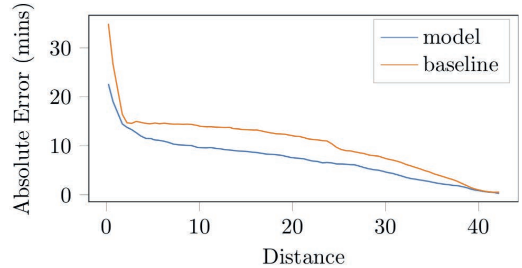

Figure 2: Error in minutes for model and baseline at each interval through the race.

2.2 Model

In order to build a model to predict the finish time we generate the above features at every 500m interval in the race. We then use these features to build a separate model at each interval using XGBoost [5] to predict a runner's finish time from that point. This allows us to make accurate finish time predictions for any runner as they progress through the race.

2.3 Model Performance

To evaluate our model we compare it against the performance of a baseline prediction. The baseline we are using is that often used by marathon race organisers to inform spectators of a specific runner's finish time, and assumes a runner will run the race with an *ideal* even pacing strategy; that is, runners will aim to run at the same pace for the duration of the race. The baseline prediction for a runner can be calculated by

$$\text{Predicted Finish Time(s)} = MRP * 42.2 * 60 \qquad (1)$$

where *MRP* is the mean pace of the runner to that point. This baseline offers a useful comparison because it also corresponds to the suggested pacing strategy for recreational runners and thus allows us to see the scale of the errors these non-adaptable pacing plans may cause. Figure 2 shows the mean absolute error of finish time prediction for both the model and the baseline at each interval in the race. The model substantially improves on the the prediction made by the baseline, outperforming it by over 4 minutes between 7.5 and 23.5km. After the half way point of the marathon the errors begin to converge as runners begin to slow down (and for some hit the wall) and better reflect the average pace of their full race.

On this basis we conclude that our model is capable of predicting whether a runner will slow down. The model uses the race features to establish the degree to which the current performance is sustainable. If a feature is deemed too high for a given pace the model will predict that the runner will slow down before the end of the race and will reflect this in the predicted finish time.

Our model allows us to fulfil our first objective of making a better finish time prediction. As most runners already carry some form of smart device during the race, this information could be easily relayed to ensure a runner is on target for their goal time. However, if this performance improvement is due to the ability to predict slowdown such a model will also be useful for our second aim; identifying runners that are likely to hit the wall and making recommendations to ensure they can mitigate the effect on their finish time.

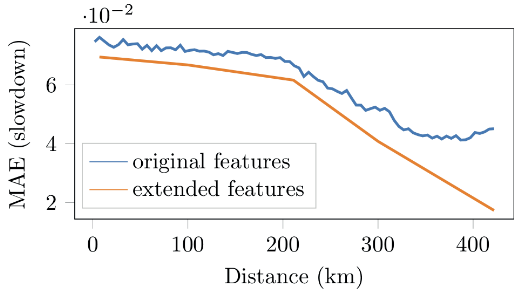

Figure 3: Mean Absolute Error of split magnitude prediction.

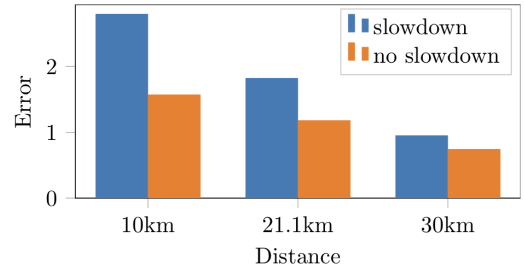

Figure 4: Distances between suggested profile and real profile for runners that receive a recommendation.

3 PACE RECOMMENDATION

3.1 Identifying Slowdown

An important aspect of marathon pacing is the idea of a pacing split [2]. An even split is a race run at the same pace for the first and second half of the marathon, and is often the recommended strategy for recreational runners. Elites may employ a negative split, a second half run faster than the first half. Finally, a positive split is where the second half is slower than the first. This is the most prevalent pacing split amongst recreational runners and is caused by a tendency to tire over the long distance [9]. The magnitude of the positive split can indicate problems pacing the race, with very large positive splits considered evidence of *hitting the wall*.

Our first goal is to identify those runners that have a larger than average positive split. We start with the same features as in Section 2.1 and build a further XGBoost regression model to predict the race split (slowdown) at every point in the race. However, Figure 3 shows that these features perform poorly at predicting the magnitude of the slowdown until the second half of the race (after 21km). It is clear that our model is able to predict whether a slowdown will occur from early in the race, but has difficulties predicting the magnitude of the slowdown. We now look for features in the race profile that are able to improve on this prediction.

We employ an automated time series feature extraction method, tsfresh [6], to generate a new, highly detailed feature set from our time series data (pace, HR, cadence). We use feature selection to identify the most relevant of these features and this allows us to substantially improve on the split predictions made by our model. Due to computational cost we decide to only extract features for some of the most critical points of the marathon, namely the 10km, half-marathon, 30km and finish line points of the race. After making a prediction as to the pacing split of the runner, we then classify whether that slowdown was greater than the average. Figure 3 demonstrates that the extended feature set is better capable of predicting the split of a runner, and we are satisfied that the model is capable of recognising a large quantity of runners that will slow down at these points.

3.2 Making Recommendations

We look to make recommendations at crucial points of the race to runners that we believe are likely experience a significant slow down. During the race, runners are continually making decisions about how to adjust their pace, either to to reach a certain target finish time, to avoid a slow down or to accommodate conditions on the course. To make useful recommendations, we leverage the fact that some runners will make decisions about their pacing that will mitigate these effects and avoid a significant slowdown. The runners who successfully avoid a significant slowdown can become exemplars to those at risk of a slow down. If our model predicts a runner is at risk of slowing down significantly we will suggest a pacing plan based on the strategies of similar, at-risk runners that managed to finish the race without a detrimental slowdown. We will make recommendations to all runners that we predict will slow down.

We recommend an updated pacing plan to a runner at risk of slowing down by utilising user-based recommendation. The race features (pace, HR, cadence) are used to find similar runners and these are used to recommend the pacing profile for the remainder of the race. The recommended pacing plan for runner X at risk of slow down is generated with the following steps:

(1) Find the most similar runners to X that do not slow down.
(2) Calculate the average pacing profile these similar runners.
(3) Normalise this average pacing profile. $\frac{P_i}{Mean(P)}$
(4) Make a further finish time prediction based model trained on runners that do not slow down.
(5) Calculate the required average pace to finish in that time.
(6) Multiply the average pace over the normalised pacing profile to return a personalised pacing profile to the runner.

3.3 Model Validation

In order to validate our recommendations we examine more closely the group of runners that our model predicts will experience a slow down. Of this group, some will naturally adjust their pace and finish safely, whereas others will experience a major slow down and hit the wall.

We wish to show that runners that are at risk of slowdown (predicted to slowdown at each critical point) but finish safely without intervention, follow a pacing strategy more similar to our recommendations than those runners that do experience a slowdown.

To do this we calculate the error between the recommended pace and the actual pace. Figure 4 demonstrates that for recommendations made at each critical point, the recommendations are close to the actual performance of the runner, showing that our recommendations are similar to those strategies followed by at risk runners to correct their pacing. Meanwhile, the recommendations are less similar to the strategies run by runners that do slow down. This means that our recommendations, if followed, allow a runner to adapt their pace to achieve a minimal slowdown.

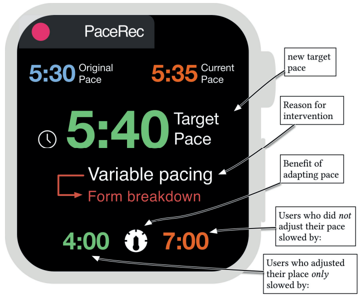

Figure 5: Example of recommendation on a smartwatch app.

3.4 Explainability

Recommendations made during a race may seem counter intuitive to the athlete, particularly when the athlete is advised to slow down in order to achieve a better finish time (because there is a risk of hitting the wall). The need for explanation of these recommendations has been recognised [3]. In the case of in-race recommendation a user is unlikely to feel tiredness at the point a recommendation is being made, and thus may ignore suggested pace adaptations. Indeed, during a race it may be difficult to envisage that a slight reduction in pace may lead to a faster finish time. For this reason explanations are vital to convince an athlete that a pace adaptation is the correct course of action; both why a recommendation is being made and the benefit that will follow from accepting that recommendation. We seek to make these explanations simple for users. Runners tend to only carry small devices and the space to make explanations is limited. Additionally runners are likely to find it challenging to focus on complex explanations while running a marathon, making simple explanations crucial to any recommendation.

The timing of these explanations and recommendations is also of importance. In Section 3.2 we generate recommendations only at the 10km, half marathon and 30km stages of the race. This is partially due to computational cost of the feature generation, but also because these are considered major distance milestones of the marathon. Common advice on marathon running suggests runners break the race into segments and to focus on hitting intermediate time goals [12]. At the end of these segments runners begin to assess their race and make decisions about future pacing. This makes these points a natural stage for presenting recommendations. Runners are already assessing their performance and are more susceptible to recommendation than they would be while focused solely on running. We chose 10km intervals for convenience; however 5km intervals, race aid stations, course specific landmarks or user defined intervals are all possible. Our foremost goal is to make recommendations when they are most likely to be observed.

It is noted in section 2.1 that the features used in prediction are correspond closely to those used by runners to evaluate their performance. This, alongside our tree based XGBoost model allows us to generate an explanation as to why a recommendation is being made for the user. The decisions that cause the model to predict a slowdown can be relayed to the user in terms they understand. For example a variability in cadence may suggest a forthcoming breakdown in form or a change in HR without an associated change in pace may suggest they are running a pace that will be unsustainable for the remaining distance. These reasons can be prompted to the user as explanations that trained runners will immediately understand. This association between the model and easily generated explanations are why we use an XGBoost model and forgo the use of deep learning methods for making predictions.

We also wish to explain why a recommendation should be followed. While it is of benefit to know why a recommendation was made we must also demonstrate the benefit of the pace adaptation. To do this we leverage our use of user-based recommendation and the neighbourhood approach used to generate a pacing plan. To demonstrate the effectiveness of the recommendation to a runner we show the future performance of runners in their neighbourhood that have been used as exemplars against the performance of those that did not adapt their pace. This gives a simple explanation - *"Users that have run similarly to you and adapted their pace slowed down by X compared to users that did not adapt who slowed by Y"*. Generating hybrid explanations from multiple sources has proved useful in different domains [14] and we believe our use of item and user-based explanations will lead to a user trusting and following our recommendations. An example of how these recommendations may be presented to the user can be seen in Figure 5.

4 CONCLUSIONS AND FUTURE WORK

In this paper we built a model to predict the finish time of a runner during the race, and we demonstrate a clear improvement over the current baselines. We extended this model to accurately predict whether a runner will have a future slowdown in the race to identify runners that will experience premature fatigue. This identification provides us with the set of runners who require an in-race recommendation to mitigate the effect of the potential slowdown. A recommendation method using runner similarity to provide suggested pace profiles to those runners at risk of fatigue can successfully help runners avoid a detrimental slowdown. By using features that are related to a runner's understanding of marathon effort we are also able to provide meaningful explanations to runners in terms they associate with fatigue and slowdown.

We note that it may also be useful to provide recommendations for runners that are unable to finish the marathon due to injury. However, there is currently no way of easily identifying runners that have not completed the full marathon distance within our dataset. Furthermore we believe that our recommendations may allow runners experiencing excessive fatigue to at least finish the race. Nonetheless, we plan on including such runners in the future.

One limitation of this work is that we do not have access to a runner's previous performance data. Incorporating training data in our model would allow us to better identify whether or not a particular effort is sustainable. Such information would lead to an improvement in model performance in both finish time and slowdown prediction. This in turn would allow us to make recommendations earlier in the race to a group of runners who are more likely to slow down. Therefore we look to expand on this work by utilising detailed training data from these athletes to improve recommendations, particularly during early stages of the race.

ACKNOWLEDGMENTS

This work is supported by the Insight Centre for Data Analytics under Grant Number SFI/12/RC/2289_P2

REFERENCES

[1] 2019. Dublin Marathon Results 2018 - Part 2. https://www.pace-man.com/dublin-marathon-results-2018-part-2

[2] Chris R Abbiss and Paul B Laursen. 2008. Describing and Understanding PacingÂăStrategiesÂăduring AthleticÂăCompetition. *Sports Medicine* 38, 3 (2008), 239–252. https://doi.org/10.2165/00007256-200838030-00004

[3] Jakim Berndsen, Aonghus Lawlor, and Barry Smyth. 2017. Running with Recommendation.. In *HealthRecSys@ RecSys*. 18–21.

[4] Ludovico Boratto, Salvatore Carta, Walid Iguider, Fabrizio Mulas, and Paolo Pilloni. 2018. Predicting workout quality to help coaches support sportspeople. *CEUR Workshop Proceedings* 2216 (2018), 8–12.

[5] Tianqi Chen and Carlos Guestrin. 2016. XGBoost: A Scalable Tree Boosting System. In *Proceedings of the 22Nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD '16)*. ACM, New York, NY, USA, 785–794. https://doi.org/10.1145/2939672.2939785

[6] Maximilian Christ, Nils Braun, Julius Neuffer, and Andreas W. Kempa-Liehr. 2018. Time Series FeatuRe Extraction on basis of Scalable Hypothesis tests (tsfresh âĂŞ A Python package). *Neurocomputing* 307 (2018), 72–77. https://doi.org/10.1016/j.neucom.2018.03.067

[7] Robert O Deaner. 2006. More males run fast: A stable sex difference in competitiveness in US distance runners. *Evolution and Human Behavior* 27, 1 (2006), 63–84.

[8] Saumil Dharia, Magdalini Eirinaki, Vijesh Jain, Jvalant Patel, Iraklis Varlamis, Jainikkumar Vora, and Rizen Yamauchi. 2018. Social recommendations for personalized fitness assistance. *Personal and Ubiquitous Computing* 22, 2 (2018), 245–257. https://doi.org/10.1007/s00779-017-1039-8

[9] C FOSTER, M SCHREAGER, A C SNYDER, and N N THOMPSON. 1994. Pacing strategy and athletic performance (Choix strategique d'une allure et performance sportive). *Sports Medicine* 17, 2 (1994), 77–85. https://doi.org/10.2165/00007256-199417020-00001

[10] Jocelyn F. Hafer, Allison M. Brown, Polly DeMille, Howard J. Hillstrom, and Carol Ewing Garber. 2015. The effect of a cadence retraining protocol on running biomechanics and efficiency: a pilot study. *Journal of Sports Sciences* 33, 7 (2015), 724–731. https://doi.org/10.1080/02640414.2014.962573

[11] T A Haney and J A Mercer. 2010. Variability of pacing in marathon distance running. *Department of Kinesiology and Nutrition Sciences* Master of (2010), 73. http://digitalscholarship.unlv.edu/cgi/viewcontent.cgi?article=1780{&}context=thesesdissertations

[12] Hal Higdon. 2005. *Marathon: The ultimate training guide.* Rodale.

[13] Juha Karvonen and Timo Vuorimaa. 1988. Heart Rate and Exercise Intensity During Sports Activities: Practical Application. *Sports Medicine: An International Journal of Applied Medicine and Science in Sport and Exercise* 5, 5 (1988), 303–311. https://doi.org/10.2165/00007256-198805050-00002

[14] Pigi Kouki, James Schaffer, Jay Pujara, John O'Donovan, and Lise Getoor. 2019. Personalized explanations for hybrid recommender systems. In *Proceedings of the 24th International Conference on Intelligent User Interfaces*. ACM, 379–390.

[15] Shiwali Mohan, Anusha Venkatakrishnan, Michael Silva, and Peter Pirolli. 2017. On Designing a Social Coach to Promote Regular Aerobic Exercise. *Proceedings of the 29th Conference on Innovative Applications of Artificial Intelligence (to appear)* February (2017), 4721–4727.

[16] Paolo Pilloni, Luca Piras, Ludovico Boratto, Salvatore Carta, Gianni Fenu, and Fabrizio Mulas. 2017. Recommendation in persuasive eHealth systems: An effective strategy to spot users' losing motivation to exercise. *CEUR Workshop Proceedings* 1953 (2017), 6–9.

[17] Peter S Riegel. 1981. Athletic Records and Human Endurance. (1981).

[18] Daniel Ruiz-mayo, Estrella Pulido, and Gonzalo Mart. 2013. Marathon Performance Prediction of Amateur Runners based on Training Session Data. (2013).

[19] Barry Smyth and Padraig Cunningham. 2017. 'Running with Cases: A CBR Approach to Running Your Best Marathon'. *Proceedings of ICCBR 2017, Trondheim, Norway, June 2017* (2017).

[20] Barry Smyth and Pádraig Cunningham. 2018. Marathon race planning: A case-based reasoning approach. *IJCAI International Joint Conference on Artificial Intelligence* 2018-July (2018), 5364–5368. https://doi.org/10.24963/ijcai.2018/754

[21] C. D. Stevinson and S. J. Biddle. 1998. Cognitive orientations in marathon running and "hitting the wall". *British Journal of Sports Medicine* 32, 3 (1998), 229–234. https://doi.org/10.1136/bjsm.32.3.229