Bachelor's thesis

Degree programme in Information and Communications Technology

2019

Sakari Ekqvist

DEVELOPING AND TESTING A DISTANCE PREDICTION ALGORITHM

– For a mobile running app

TURKU AMK
TURKU UNIVERSITY OF APPLIED SCIENCES

BACHELOR'S THESIS | ABSTRACT

TURKU UNIVERSITY OF APPLIED SCIENCES

Degree programme in infofmation and communications technology

2019 | 29 pages

Sakari Ekqvist

DEVELOPING AND TESTING A DISTANCE PREDICTION ALGORITHM

– For a mobile running app

This thesis introduces a few already existing mobile applications designed for runners and their benefits as well as their shortcomings, mostly in their personalization for the runner.

The objective of the thesis was to develop a distance prediction algorithm for ExerGo. ExerGo is a mobile application developed by Turku Game Lab designed for runners, that provides the user with a running program depending on their physical condition, helps the user to train more effectively and helps the user by creating a running loop on the map for their run so the user does not have to plan the route themselves.

The algorithm would estimate the distance of the run depending on the runner's results. Three different versions of the algorithm were designed and tested. The tests were planned to include eight runs for each algorithm and then the results of each estimation were compared to the actual distances of the runs after they were completed.

The experiment did not give a conclusive result for the correct algorithm version, but the results showed that any of the algorithms can provide usable estimations for use in ExerGo, but using more than a single run was in the estimation was beneficial. The experiment also revealed how other runners might be more prone to inaccurate estimations.

KEYWORDS:

Mobile application, algorithm, exercise, running, data analysis

OPINNÄYTETYÖ (AMK) | TIIVISTELMÄ

TURUN AMMATTIKORKEAKOULU

Tieto- ja viestintätekniikka

2019 | 29 sivua

Sakari Ekqvist

MATKAA ENNUSTAVAN ALGORITMIN KEHITYS JA TESTAUS

– Juoksijoille kehitetyssä mobiiliaplikaatiossa

Tässä opinnäytetyössä esitellään muutama jo olemassa-oleva juoksijoille tehty mobiiliapplikaatio ja niiden hyötyjä käyttäjälle, ja niiden puutteita, pääasiassa applikaatioiden yksityistämisen kannalta.

Tavoitteena oli kehittää algoritmi Turku Game Labin kehittämään ExerGo-mobiiliapplikaatioon. ExerGo on juoksijoille kehitetty applikaatio, joka tarjoaa juoksijalle juoksuohjelman hänen juoksukuntonsa mukaisesti, auttaa juoksemaan tehokkaammin ja luo juoksulenkin kartalle poistaen juoksijalta sen vastuun.

Algoritmin tehtävänä oli arvioida juoksujen pituus juoksijan tulosten mukaisesti. Algoritmistä kehitettiin ja testattiin kolmea eri versiota. Testit suunnteltiin sisältämään kahdeksan juoksua jokaiselle algoritmille. Kun juoksut olivat suoritettu, tuloksia verrattiin arvioihin.

Tutkimus ei tuottanut ratkaisevaa lopputulosta siitä, mikä olisi paras algoritmiversio lopulliseen käyttöön, mutta sen tuloksista selvisi, että mikä tahansa versio pystyy tuottamaan käytettäviä arvioita, mutta useamman kuin yhden juoksun käyttäminen arvioissa osoittautui hyödylliseksi. Tulokset myös näyttivät juoksijoiden eroja arvioiden tarkkuuksissa.

ASIASANAT:

Mobiiliapplikaatio, algoritmi, liikunta, jokseminen, data-analyysi

CONTENTS

FIGURES

TABLES

LIST OF ABBREVIATIONS

BEE	Business Ecosystems in Effective Exergaming
HRZ	Heart rate zone
PC	Personal computer

1 INTRODUCTION

Exercising can be anything that makes the body move and requires the body to burn calories. Whatever the exercising method is, exercising has multiple health benefits that can affect the person either physically or mentally. (WEIL and STÖPPLER, 2019)

There have been multiple studies showing that regular exercise will help the person's overall mood and reduce possible anxiety. There are several reasons for this. Physical activity can increase hormone production and can act as a distraction. There are multiple connections between different biological and psychological mechanisms and exercise. (SEMECO, 2017) Regular exercise has also been shown to increase the quality and quantity of sleep and help with relaxation. (KREDLOW, CAPOZZOLI, HEARON, CALKINS and OTTO, 2015)

Inactivity and the lack of exercise has been shown to be a major cause for weight gain and obesity. Exercise also can help with skin health, chronic pains and increase energy levels. It is also important for building and maintaining muscles and bones. Maintaining bones and muscles becomes vital as people age. (WEIL and STÖPPLER, 2019)

1.1 Running as a form of exercise

Running is a convenient and popular way of exercising. Running does not require any special equipment and can be done almost anywhere. Running can cause some problems, mainly in joints. If a person has joint pain or is overweight, they should not practice excessively. Otherwise running is a safe way to exercise. (LEE, BRELLENTHIN, THOMPSON, SUI, LEE, and LAVIE, 2017)

1.2 Training programs

To gain the maximum benefit from running, a runner should run regularly. It is useful to create a training program for goal-oriented exercise to make the running more motivating and challenging, but still keeping the training and the goal realistic. The plan should be set up so that it will help the runner to improve but the runner can recover between different exercises. Training plans are constructed in a way that they increase in intensity as the program progresses. To achieve this type of training program, it needs to be somewhat personalized. (DOUGLAS and KEFLEZIGHI, 2015)

For example, Nike+ Run Club offers training plans that consist of different pieces the runner can choose themselves. The plans are split into weeks that include speed, endurance and recovery exercise that the runner can complete as they see fit. The runs mostly consist of different distances that should be run at a certain pace. These different paces need to be setup by the runner, or the runner can complete few calibration runs to determine their pace. (NIKE.COM, 2019)

Boston Athletic Association, the group behind the Boston Marathon, also offers runners training plans for marathons and all the other different distances they have races for. These training programs also include the pace chart which determines the pace of each run depending on the runners goal time for the race. (BAA.ORG, 2019)

There are also time-based training programs, which just give the runner a time to run or have exercises that include running and walking. (RUNNERSWOLD.COM, 2019) Additionally in running programs it is possible to use the runner’s heart rate to improve the program’s effectiveness. In these programs the program gives a specific heart rate zone (HRZ) for the runner for each part of the run so that the runner can try to stay at an optimal heart rate during the run. These training programs require a way to monitor the runners heart rate, which is commonly a heart rate belt and a sport watch (HOTTENROTT, 2007). The zones all have their place in improving the runner’s overall running condition, which makes the HRZ training a suitable and effective way to train for running. The running program that was used for this thesis was a heart rate based running program and was made by a doctor of human movement science, based on years of coaching experience. (SACHS, 2011; IRVINE, 2019)

1.3 Mobile applications for running

Mobile applications are software that run on smartphones or tablet computers. They are usually small and designed to fulfil an individual task or service. Many mobile applications are similar to applications that are usable on PCs, but since the mobile device market has grown huge, different applications that are more suitable for mobile environment have become commonplace. (TECHNOPEDIA.COM, 2019)

There are countless mobile applications that are designed to help people exercise more effectively. Since smartphones are compact while also being able to do tasks required to keep track of the running activities, such as recording the location data required to visualize a route on a map to from the run and keep track of the times and speeds of the user, smartphones are a useful tool that can help the user exercise more efficiently.

There are mobile applications that give the user different running exercises to perform and they often try to motivate the user to perform the exercises regularly. Most often applications that are designed for running are tracking the users progress, letting the user see their statistics and possibly a route of their past runs. Amer Sports Digital's Suunto application is an application made for tracking the users progress. It allows the user to set a weekly goal for exercising and track different types of exercises, including running and cycling. The Suunto application records distance, pace, and heart rate during the exercise among other statistics. The application also gives the user the option to plan routes beforehand and then use them as guides when running. There is a social element with the exercise tracking and the application allows the user to share their progress and past exercises or planned routes with their friends. These applications do not give personalized exercises to follow and give the user the responsibility to decide where to run and at what pace. (AMER SPORTS DIGITAL, 2019)

There are applications that give the user a generic exercise to follow. For example, Samsung's health application allows the user to use Samsung's wireless earphones to track the user's heart rate and the application can then give the user audio feedback to guide the user to stay in a optimal heart rate zone for the exercise. In the application, there are a few preset exercises with different lengths and intensities that the user can use to optimize their exercising. The health application additionally records statistics around the clock and is designed to be a more of a general health guide. (SAMSUNG ELECTRONICS CO., LTD., 2019)

Zen Labs Fitness has made a series of applications that are designed for goal-oriented running exercises. C25K – 5K Running Trainer is an application that gives a training program to follow for a complete beginner targeting to be able to run 5 km. They also have applications for users wanting to train up to a marathon, but the applications all function similiarly. The application gives the user a set of exercises that are separated in different parts depending on the intensity. A typical beginner exercise includes warm up, and then few repeats of walk and jog cycles one after the other. The intensity and the length then gradually increases when the user continues on the training program. Zen

Labs Fitness' applications are less about the statistics and aim to provide a general training program for the user. (ZEN LABS FITNESS, 2019)

From the applications that are available, it is clear that most are able to track a runner's performance for post-run analysis, while others may also suggest generic training programs to follow toward some running goal. Unfortunately, these generic programs do not address personal fitness levels and running ability and are not as effective as they could be. The objective of personalized training programs would be to mimic a personal trainer as closely as possible. There are some limitations with trying to achieve that with a mobile application that is targeted for general use. Having a professional personal trainer would be beneficial to many, since that would direct the user to do the correct exercises at the correct times and also monitor that the exercises are done correctly without harming the user's body.

2 PROBLEM

Companies refrain from supplying truly personalized training programs. The reason lies in that it is difficult to estimate the running condition of a runner and estimate how they will perform during different types of runs and how their condition will progress. The main shortcoming with many of the existing running applications is that they do not offer the user a distance estimation for their next exercise. When an application gives the user a timed run to complete, but does not give the user an estimation of the distance that would be ran during the exercise, the user needs to plan their run if they do not want to end the run in an undesirable location leaving them to walk the rest of the way or having to run unnecessary distance, which compromises their optimal training, after they have reached the end of their training session.

Depending on the run, the planning might get complicated. For example, a run where there is a warm up part, a jogging part, a walking part and a cooldown part, would require the user to have an average speed for each of those different parts and calculate separately the distance that would be covered in each part of the next run. Using a generic average speed would result in inaccurate results, since the next run might not have the same parameters as the previous one. Also, as the person’s fitness level improves, the speed per zone also increases—meaning they run farther as they get fitter.

In a familiar area, it might be sufficient to provide the runner with a distance for their run, since they might be able to plan the route without much of effort. When running in an unfamiliar place, it can be difficult to plan a running route, since in addition to the location being unfamiliar the exercise time might be unfamiliar leaving the user with a lot of responsibilities to plan their exercise.

3 AIM & OBJECTIVES

The aim of this thesis was to provide important features to a mobile exercising application ExerGo, developed by Turku Game Lab. Turku Game Lab is a learning and working environment of the University of Turku and Turku University of Applied Sciences. It provides education for students and allows them to take part in various projects using the latest game technologies, so the students are well equipped to work in the field after graduation. (TURKUGAMELAB.FI, 2019)

The aim of this thesis is to develop a mobile app that serves a personalized running program, based on HRZ, with the ability to estimate and plan the running route according to the runner's current ability.

The primary objective of this thesis was to develop and test an algorithm to estimate the distance of the user's next running session in the training program by utilizing the data gathered during their previous sessions. The data gathered and used was the time spent in each part of the session and the average speed of the user at each different part of the session. The distance estimation is then fed into a path-finding algorithm to plot a suitable route on the app's mapping function. This allows the user the freedom to run in unfamiliar areas and not have to plan their run loop to not end too far from their starting point. The scope of this thesis does not inspect the path finding algorithm of ExerGo, but rather attempts to find the most effective algorithm for predicting the next running session distance.

4 BACKGROUND AND METHOD

ExerGo is a mobile application that aims to encourages users to exercise more and lowers the bar to start exercising with an exercising program. It is designed to give users everything they need to start exercising using exercise programs tailored for their needs and their physical abilities.

ExerGo is part of the contribution of Turku Game Lab to the BEE project. The BEE project aims to improve the quality of exercise games and prove their effectiveness. The BEE project also aims to increase the level of attractiveness and motivation effect of exercise games.

ExerGo is being developed with the Unity game engine. ExerGo has augmented reality features making Unity a suitable engine for the project. The augmented reality features are made using Google's ARCore plugin for Unity. Unity is a game engine suitable for 2D and 3D games and it is a flexible engine to develop all types of applications.

To develop the best possible algorithm for use in ExerGo, three different versions of the same base algorithm were designed and compared in a test where they were used for eight sessions and then the data was analyzed to determine which one was the most accurate. The first version used all the data gathered during the training program, the second version used only a small part of the data gathered, more specifically from the three latest run exercises, and the third version used only the latest finished exercise.

The main benefit for the user would be to determine the actual distance of their next run exercise and to create a running route for the user according to their own running condition and training program.

5 DESIGN PROCESS OF THE ALGORITHMS

Algorithms are a set of step by step instructions that are given to the computer to execute. In other words, an algorithm is a computational tool that transforms given inputs to the correct output for those input values. For the algorithms developed during this thesis, the outputs are the distance the next exercise would cover. That could then be used as an input in ExerGo for another algorithm that would plan a route and output it for the user.

Since the exercises for ExerGo were implemented in a way that would encourage the user to stay in a certain HRZ rather than running at a specific speed the speed of each user would differ and the speed would even change during the exercise, it was important to in each case use data equally from the start, the middle and the end of the exercise to output an estimation as accurate as possible. That means it would be convenient to use complete exercises as the minimum sets of data to be used.

The challenge in calculating an accurate distance was determining what information to use for the algorithm as the input. Since the algorithm would be used to calculate distances according to data gathered during past exercises, the input for the algorithm would change for every time the user would go or the next exercise.

As the environment for the algorithm is an application that aims to improve the physical condition of the user, the expectation is that the speed of which the user runs would increase during the exercise program. To determine accurately how the speed would change for a specific user, there would need to be more personal information that is conceivable to gather from the person that is starting an exercise program aimed at beginners. The aim is just to be as accurate as possible for the general user but to increase that personalization as the user completes more exercises.

ExerGo tries to optimize the training by guiding the user to exercise in a specific heart rate zone. Because there are different heart rate zones used in the training programs that determine the intensity of each part of an exercise, the average speed of the user is recorded according to the zone the training program determines so that each of the zones would get a separate average speed. There are five zones and they are calculated from the user’s theoretical maximum heart rate. The theoretical maximum can be calculated using the user’s age and gender. The user is in the first zone when their heart rate is 50 – 60% of their maximum heart rate. The first zone is very light exercise and is mostly used in the training programs in the warmup phase or in the cooldown phase. The second zone up to 70% of the maximum heart rate and is also used in the warmup and cooldown phases but is also used in the main work phase in the beginner training program. The third zone which goes up to 80% and the fourth zone which goes up to 90% are used in the work phases. The fifth zone that goes from 90% up to the theoretical maximum is not in use in the current programs.

Since it is established that it would be beneficial to use complete exercises as the basis to use the gathered data. Depending on how fast the algorithm needs to react to changes in the user's physical condition or speed, the number of days should be decided. Using fewer days to calculate the distance of the user’s next exercise would allow the algorithm to react faster to changes in the user’s conditions, but that has the danger of overreacting to the changes in the data and if the algorithm takes into account too many exercises, it may not react fast enough to the changes in the user’s physical condition.

The hypothesis for the correct inputs for this case is to use a somewhat dynamic number of most recent exercises and the speed data collected during them. That would mean that the program would react fast enough to the changes in the user’s progress but would not be affected as much by a single day with a big deviation from the average. Such deviations could for example, stem from the runner becoming ill or environmental conditions (harsh or favourable).

5.1 Algorithm 1: Using all the past exercises

The first algorithm that was tested was an algorithm that accounts for all the exercises the user has completed during the exercise program they were currently using. The idea behind using all the data collected to calculate the average speed of the user would be to allow the consideration of the continually improving physical condition of the user, but not to overreact to the change. This method would minimize the effect from an exercise where the speeds were far from the average speeds since it would consider all the highest and the lowest results and calculate an average from them. The only phase when this algorithm couldn’t be used is the very first exercise of the exercise program when there is no data gathered yet. This is however, common with all the algorithms.

The structure of the algorithm is rather simple. The first part of the program is to create a list of all the completed runs in the current training program. There is a possibility to stop an exercise without fully completing it or skip it. Exercises that are not fully completed, are skipped by the algorithm because the data collected in them might be inaccurate. When all the eligible exercises are listed, those exercises are then searched for the specific training heart rate zones that are used in the next exercise. If there are values for that zone, then the values are gathered in a list. This process is done for all the heart rate zones used in the next exercise. Then from those speeds, averages are calculated. To calculate an accurate distance estimation, for each zone the average speed is then multiplied by the time each zone will be on in the next exercise.

5.2 Algorithm 2: Using last three exercises

The second algorithm that was tested was a modified version of the first algorithm. This algorithm would take only a few most recent exercises into account during the calculations. The idea behind this version of the algorithm was that it would allow the application to react faster to the improvements in the user's physical condition without overreacting to changes. If the user for example had a single day where their heart rate was higher than usual, forcing them to run slower to train optimally, the algorithm would give a lower estimation to the next exercise than if all the exercises would be taken into account.

The structure of the algorithm is very similar to the first algorithm but when the eligible exercises are listed, only the three most recent ones are added to the list. Again, only fully completed exercises are considered and the ones that are not fully completed are skipped and the next most recent exercise is used instead. After the exercises that are used in the calculations are determined the algorithm is the same as for the first algorithm

The part of this algorithm that would need more testing is finding the optimal amount of exercises that are taken into account during the calculations. For testing purposes, to create a larger difference between the algorithm during the tests, only three most recent exercises were taken into consideration.

5.3 Algorithm 3: Using only the most recent exercise

The third algorithm that was considered, was the algorithm that only used the data from the most recent exercise. The idea behind this algorithm was to create an option that would react as fast as possible to any improvements on the physical condition of the user, or any setbacks in the process. Only using the most recent data gathered, while still maintaining the principle of having the speed data gathered from a full run, could possibly make the algorithm be as accurate as possible if the progress of the user stays predictable. Another way of looking at the benefit of this option is that it forgets all the data from the start of the exercise program while the user is continuing to train with the exercise program and would not use the data that could be misleading for the distance calculations.

6 EXPERIMENT DESIGN

To determine the best possible algorithm a test was conducted. The test was conducted by having a person go through a shortened exercise program and have the estimations of the exercise distances and the actual distance covered during an exercise recorded. Then the estimated distances and the actual distances could be compared, and the result would show how accurately that algorithm can predict the distance that the user would cover during their future exercises. This test used two different test subjects. They completed the runs first and then the estimations they would have got from each of the algorithms were calculated. This way all of the algorithms were able to be tested simultaneously. Using more test subjects would have increased the amount of data that could have been gathered from the tests and made them more reliable, but for this case, it was decided that testing with only two test subjects would lead to a result that could be analysed for the purpose of this thesis.

To make this test useful in the environment the algorithm is created for, the test was done using an example exercise program that would follow the same format as the real exercise programs that would be used with ExerGo. The exercise programs in ExerGo have runs with different durations, so it’s essential that the tests include the different types of runs and use them in the data gathering and calculations. The length of the test was decided to be eight sessions so that there would be at least two program changes (shown in tables 1 through 3). That is, eight runs per algorithm were planned to be recorded and compared with their estimation counterparts.

Table 1. Week 1 running program.

Run state	Duration (min)	HRZ
Warm up	15	2
Work	10	3
Cooldown	10	1-2

Table 2. Week 2 running program.

Run state	Duration (min)	HRZ
Warm up	10	2
Work	12	3
Cooldown	10	1-2

Table 3. Week 3 running program.

Run state	Duration (min)	HRZ
Warm up	10	2
Work	15	3
Cooldown	10	1-2

Possible problems with this type of test are that the speed of which the user runs is controlled by their heart rate. Since each person's physical capabilities and condition differ, the way their heart rate will go up and down will be different. There are multiple variables that can affect the users heart rate.

Since ExerGo will encourage the user to run having their heart rate at a specific zone, some users might be more inclined to raise their heart rate faster than other people. In the same way, the speed of which the user's heart rate goes up at the start of the exercise is possibly affected by their motivation for the exercise that day. Not only that, but since the heart rate zone changes multiple times during an exercise the reaction speed to that change might differ even for a single user. The heart rate data of the user was collected during these tests, but the data was only used to determine possible problems during the exercise.

7 FINDINGS

As a result of the experiment, eight runs were completed by the first test subject and six runs were completed by the second test subject. Also, for all of these runs an estimation was calculated using all three algorithms. For the purposes of this thesis we named the algorithms: (a) all-averages algorithm; (b) last-three-averages algorithm; and (c) last-run algorithm.

In each run, the average speed of the user was tracked separately for each different HRZ From these average speeds it is possibly to see the improvement of the runner's running condition and find any outliers that would affect the future estimations negatively.

For all the runs completed after the first run for each test subject, a difference was calculated (Tables 5 through 7 and 9 through 11). This difference represents the error in the estimation compared to the actual distance covered during the run. If the difference was positive, that means the estimation went over the actual distance and if the difference was negative the estimation was under the actual distance. The results of the estimations are identical at the start of the program, since there is not enough runs completed for the algorithms to work differently. All of the algorithms function the same way for the second day of the program, meaning the first estimation and the all-averages algorithm and the three-last-averages algorithm start to differ after the third estimation.

Table 4. Average running speeds in each HRZ for the first test subject.

Day	Z1-Z2 (km/h)	Z2 (km/h)	Z3 (km/h)
1	9,58	10,13	10,87
2	9,16	10,27	10,94
3	9,28	9,95	12,00
4	9,78	11,11	11,95
5	9,13	11,86	11,72
6	9,22	11,04	11,24
7	9,82	12,36	12,25
8	10,00	11,38	12,05

Table 5. Estimations with the all-averages algorithm for the first test subject.

Day	Actual distance (m)	Estimation (m)	Difference (m)
1	5940		
2	5915	5940	25
3	6034	5927	-107
4	5872	5496	-376
5	5844	5590	-254
6	5624	5640	16
7	6758	6210	-584
8	6575	6288	-287

Table 6. Estimations with the last-three-averages algorithm for the first test subject.

Day	Actual distance (m)	Estimation (m)	Difference (m)
1	5940		
2	5915	5940	25
3	6034	5927	-107
4	5872	5496	-376
5	5844	5633	-211
6	5624	5773	149
7	6758	6361	-397
8	6575	6458	-117

Table 7. Estimations with the last-run algorithm for the first test subject.

Day	Actual distance (m)	Estimation (m)	Difference (m)
1	5940		
2	5915	5940	25
3	6034	5915	-119
4	5872	5605	-267
5	5844	5872	28
6	5624	5843	129
7	6758	6186	-572
8	6575	6758	183

Table 8. Average running speeds in each HRZ for the second test subject.

Day	Z1-Z2 (km/h)	Z2 (km/h)	Z3 (km/h)
1	7,47	7,93	8,40
2	6,89	8,26	8,54
3	6,57	7,68	7,53
4	6,39	8,69	8,73
5	6,32	9,66	9,06
6	6,58	8,56	8,25

Table 9. Estimations with the all-averages algorithm for the second test subject.

Day	Actual distance (m)	Estimation (m)	Difference (m)
1	4629		
2	4637	4628	-9
3	4269	4632	363
4	4261	4120	-141
5	4475	4155	-320
6	4174	4210	45

Table 10. Estimations with the last-three-averages algorithm for the second test subject.

Day	Actual distance (m)	Estimation (m)	Difference (m)
1	4629		
2	4637	4628	-9
3	4269	4632	363
4	4261	4120	-141
5	4475	4124	-351
6	4174	4205	31

Table 11. Estimations with the last-run algorithm for the second test subject.

Day	Actual distance (m)	Estimation (m)	Difference (m)
1	4629		
2	4637	4628	-9
3	4269	4637	368
4	4261	3880	-381
5	4475	4260	-215
6	4174	4474	300

8 RESULT ANALYSIS

When comparing the results of the algorithms to each other (figures 1 through 4), it is apparent that using more than a single run to determine the runners speed is beneficial. When the training program progresses and more data is gathered, the differences between algorithm prediction and average distances seem to smooth out for the all-averages algorithm and the last-three-averages algorithm (Figure 1 & Figure 3). Even though the last-run algorithm still does estimate the run distance better for some runs individually, the differences between two consecutive estimations were higher than for the other algorithms. For example the jump from estimating the distance 572m too low in day 7 for the first test subject to estimating the distance 183m too high in day 8 is substantial.

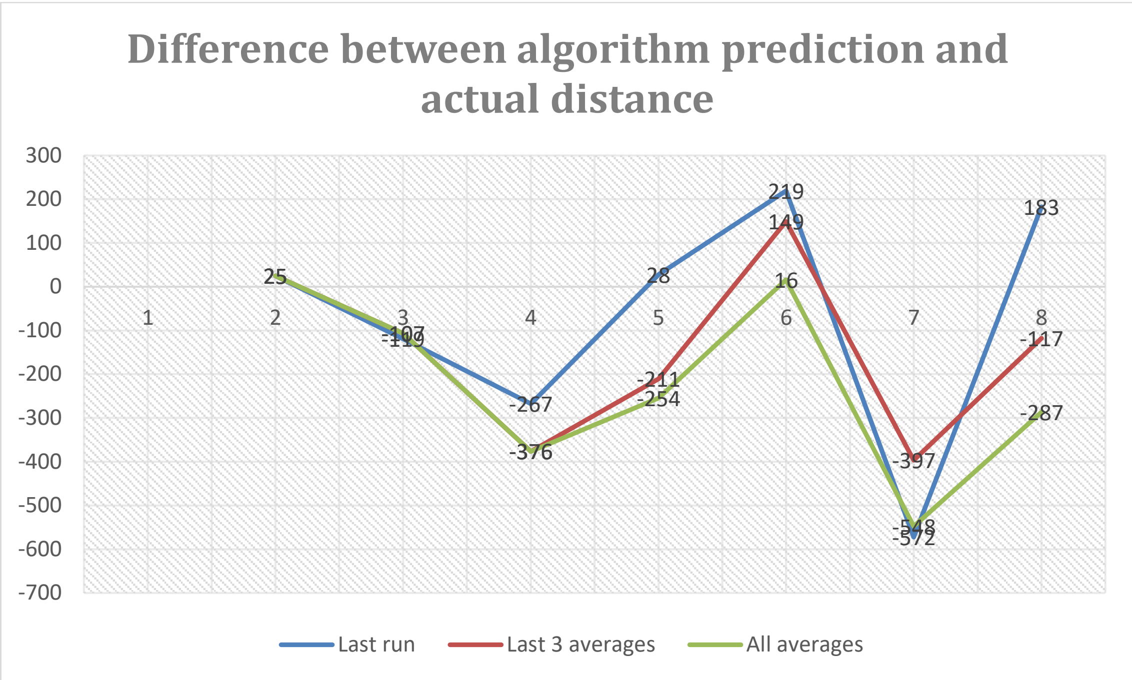

Figure 1. Difference between algorithm prediction and actual distance for first test subject

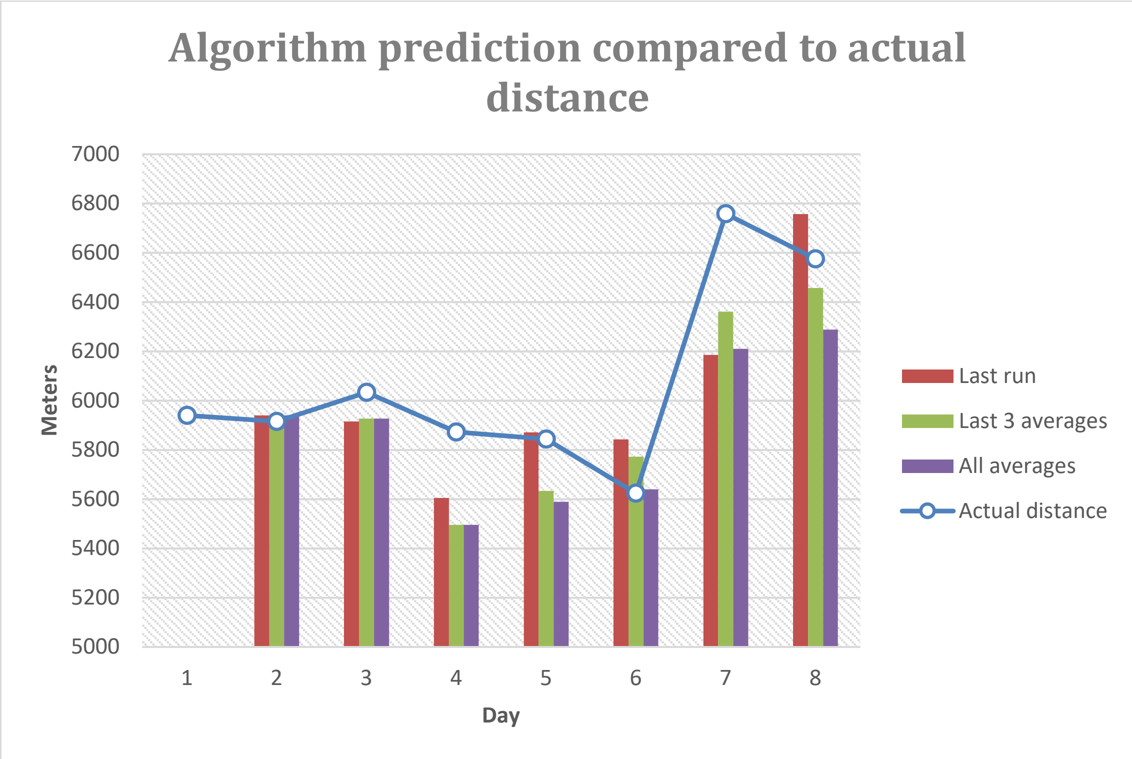

Figure 2. Algorithm prediction compared to actual distance for the first test subject

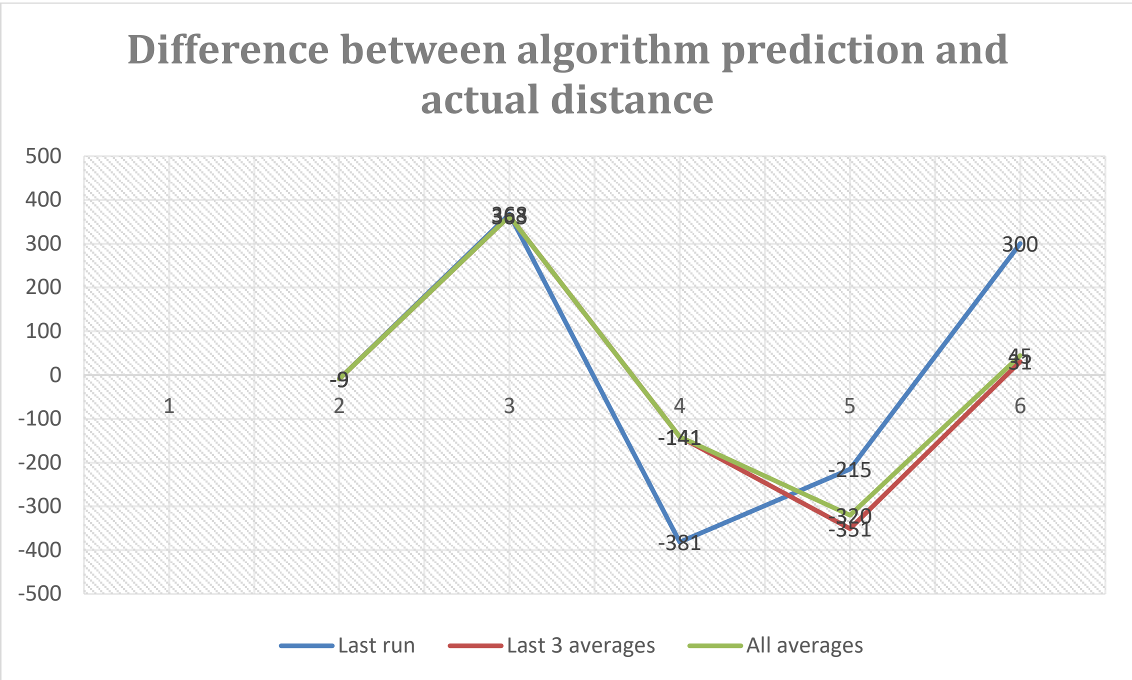

Figure 3. Difference between algorithm prediction and actual distance for second test subject

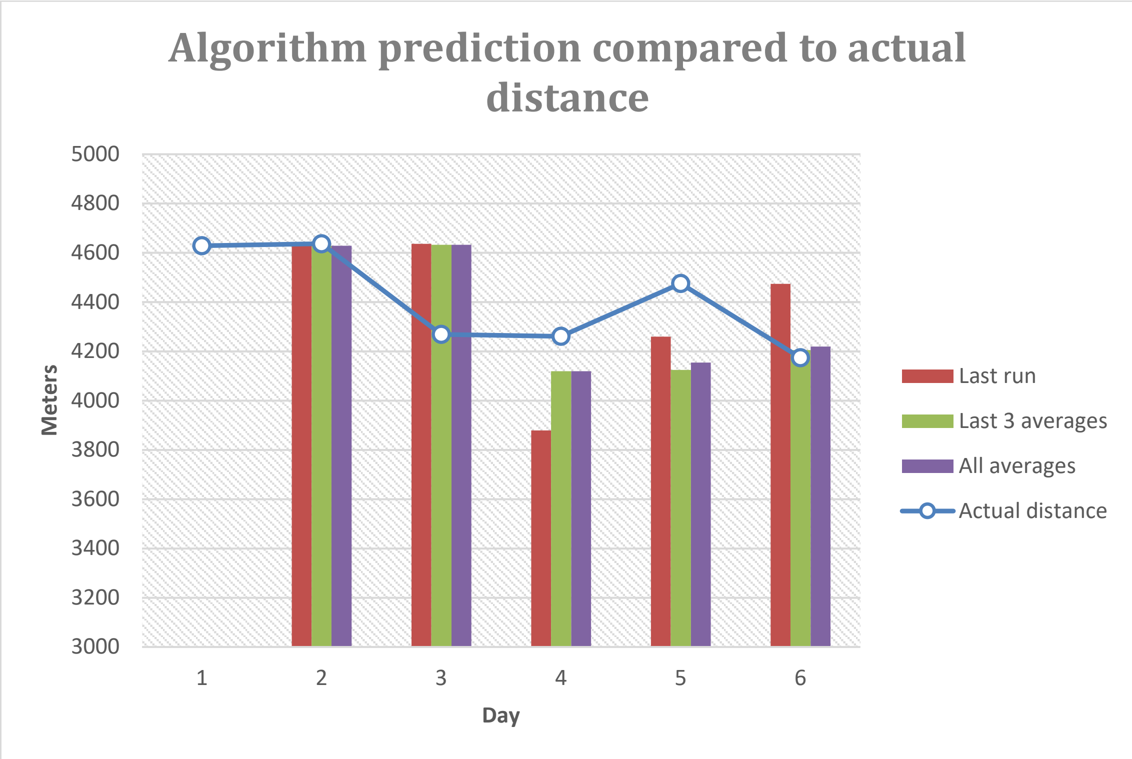

Figure 4. Algorithm prediction compared to actual distance for the second test subject

The last-run algorithm option has clear flaws. The biggest flaw in this option is that it overreacts to the runner's immediate results. If the user has one day when they are in very good shape, having slept well and eaten well, they might outperform their standard level. Then again in a few days, the user might be coming up with a cold, or just has not eaten well or slept well, their results may suffer. That might lead to the distance estimations varying greatly in just a few days, even though the actual condition of the runner has not changed. More factors that could affect the results are weather conditions and the amount of height differences in the route ran. Since the algorithm is planned to be used alongside a premade running program, the day in which the run is completed can be a factor in making the results unreliable and using only a single exercise as the data removes the ability to average the results out by balancing the good days with the bad days. These problems with the last-run algorithm might be more noticeable with some runners than others, since the inconsistency of the runner would worsen the estimation by a greater amount than for the last-run algorithm than other algorithms.

The results for the last-run algorithm show that some runners will have more fluctuation in their run distances than others. In this experiment the first test subject has consistently lower difference between the estimation and the actual distance if the run distance was taken into account.

The main problem with the all-averages algorithm is that it’s estimations would always be too low as the user’s physical condition improves and the training program would progress because it would keep using the data from the runs completed early in the program. There is a possibility that this algorithm option would need a multiplier that would increase the estimation as the runner progresses in the running program. However, this would just add another factor that would contribute to the variance of the results.

The three-last-averages algorithm does show the most potential as the changes in the running program do not affect it as much. This is apparent in the day 7 (Figure 1 & Figure 2) when the overall distance of the run gets longer. This algorithm also has the least variation. With continued testing and longer experiment the distance estimation would probably start to get converge with the actual distance because of the increased consistency of the runner that comes from the continuous training. For a longer program, the amount of days taken into account could also be adjusted to be higher, since that would reduce the effect a single run has to the estimation and the assumption is that the runner’s improvement would slow down as the training program progresses. Still, not taking into account all the past runs but only a few does give the benefit of the being more sensitive to the continuous conditioning of the runner. The key benefit of the three-last-averages algorithm is also visible by the end of the experiment for the first test subject where the estimations for the all-averages algorithm are lowered by the results from the runs at the start of the program but the three-last-averages algorithm is ignoring them (Figure 1 & Figure 2).

With these results it does seem that the experiment was not extensive enough to grant a conclusive best algorithm option, but the last-three-averages algorithm does show the most potential in this experiment as the training program progresses and it would probably benefit the most out of the runner’s fitness level raising and the results becoming more consistent.

9 CONCLUSION

There are countless mobile applications to help runners run in more effectively. Some applications have personalization in the form of training programs or exercise guidance, while others also provide analytics of the user's performance. The personalization aspect is an important part of the application that enhances its effectiveness. Accurate analytics give the mobile app user possibility to personalize the user experience further. The ability to estimate the distance for the next run is an example of such personalisaton as it very much depends on the current ability and fitness levels of individual runners.

From the experiment results of this thesis, it was evident that more testing was needed to find the most accurate way to estimate a runners next session performance. The experiment showed that the algorithms used were able to provide estimations to the run distance with a deviation of under ten percent of the actual distance. The three-last-averages algorithm showed the most potential as it shows the best results as the training program progresses and changes.

The errors in the estimations are already low enough that it is possible to utilize them in a routing algorithm in ExerGo which would create a running loop for the user on the map before their run to help the user run in new locations where they possibly cannot predict the route effectively without manually planning it beforehand. With more in-depth and comprehensive experiments, the error in the estimation could possibly be lowered further and the algorithm made more reliable. It is also expected that the results would improve as the training program progresses. Having more test subjects complete entire training programs would be beneficial to determine how the algorithms handle the changing physical condition of the runner and the increasing intensity of the program. This experiment also showed that some runners have more fluctuation in their results than others and to determine the algorithm that would be used in a general use, more test subjects would need to complete the extensive tests.

10 REFERENCES

ANDERSON, E. and SHIVAKUMAR, G., 2013. Effects of Exercise and Physical Activity on Anxiety.

AMER SPORTS DIGITAL, 2019. Suunto. Mobile app. Version 4.2.4. Available from: https://play.google.com/store/apps/details?id=com.stt.android.suunto [September 12, 2019]

BAA.ORG, Boston Marathon Training Plan. Available: https://www.baa.org/races/boston-marathon/enter/training/levelone [September 12, 2019].

DOUGLAS, S. and KEFLEZIGHI, M., February 26, 2015, Set a Goal for Any Running Distance. Available: https://www.runnersworld.com/runners-stories/a20853240/how-to-set-good-running-goals/.

HOTTENROTT, K. 2007, Training with the Heart Rate Monitor, Meyer & Meyer.

IRVINE, H.M. 2019, February 1,-last update, Everything you need to know about heart-rate training. Available: https://www.runnersworld.com/uk/training/beginners/a760176/heart-rate-training-the-basics/ [September 18, 2019].

KREDLOW, M.A., CAPOZZOLI, M.C., HEARON, B.A., CALKINS, A.W. and OTTO, M.W., 2015. The effects of physical activity on sleep: a meta-analytic review. *Journal of Behavioral Medicine,* **38**(3), pp. 427-449.

LEE, D., BRELLENTHIN, A.G., THOMPSON, P.D., SUI, X., LEE, I. and LAVIE, C.J., 2017. Running as a Key Lifestyle Medicine for Longevity. Progress in Cardiovascular Diseases, 60(1), pp. 45-55.

NIKE.COM, 5 km – 8-week Training Programme. Available: https://content.nike.com/content/dam/one-nike/en_lu/Evergreen/PDFs/nrc-training-plans/nike-run-club-5k-en_LU.pdf [September 12, 2019].

RUNNERSWORLD.COM, How to start running today: a beginners guideApril 30, 2018-last update. Available: https://www.runnersworld.com/uk/training/beginners/a772727/how-to-start-running-today/ [September 12, 2019].

SACHS, L. 2011, May 18,-last update, Heart Rate Training. Available: https://www.ideafit.com/fitness-library/heart-rate-training [September 18, 2019].

SAMSUNG ELECTRONICS CO., LTD., 2019. Samsung Health. Mobile app. Version 6.6.0.029. Available from: https://play.google.com/store/apps/details?id=com.sec.android.app.shealth [September 12, 2019]

SEMECO, A., February 10, 2017, The Top 10 Benefits of Regular Exercise. Available: https://www.healthline.com/nutrition/10-benefits-of-exercise [September 2, 2019].

TECHNOPEDIA.COM, Mobile Application (Mobile App). Available: https://www.techopedia.com/definition/2953/mobile-application-mobile-app [September 3, 2019].

TURKUGAMELAB.FI, About Game Lab. Available: https://www.turkugamelab.fi/?page_id=72 [September 27, 2019]

WEIL, R. and STÖPPLER, M.C., February 1, 2019, Running. Available: https://www.emedicinehealth.com/running/article_em.htm#more_famous_runners [September 2, 2019].

ZEN LABS FITNESS, 2019. C25K – 5K Running Trainer. Mobile app. Version 143.30. Available: https://play.google.com/store/apps/details?id=com.c25k [September 12, 2019]