Data Management for the Web 2019/2020 - TEAM 1

Belcao Matteo

Gangemi Giovanni

Comolli Federico

FitYourRun

FINAL PITCH

Summary

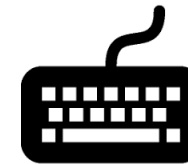Smartphone APP for IOS/ANDROID, no wearables support

Stakeholders

Runners

Sport Brands

Development Team

Summary – Customized Experience

CUSTOMIZED PLANS

👥+ **shared**

👤 **individual**

TIME PREFERENCES
by hand or/and data from 📅 Google Calendar

ENDURANCE

MOUNTAIN & TRAIL-RUNNING

BODY SKILLS IMPROVEMENT

VELOCITY

REDEMPTION & RECOVERY

CROSS-TRAINING (bike + running)

cardio
breath
muscles
.....

CUSTOMIZED PATHS

PERSONALIZED PATHS

STREET DATA

REAL-TIME CONDITIONS

OFFLINE RUNNING

AREA CLOSED

AUDIO SUGGESTIONS

FitYourRun

Summary

RANKINGS & AWARDS

best runners : users who better follow their training plans

2 kinds of rankings

- **FREE :** discounts and coupons
- **PRO** : gift cards and physical goods

LOCAL POPULAR PATHS

Sforza castle 10 km
4,8/5

Sempione park 8 km
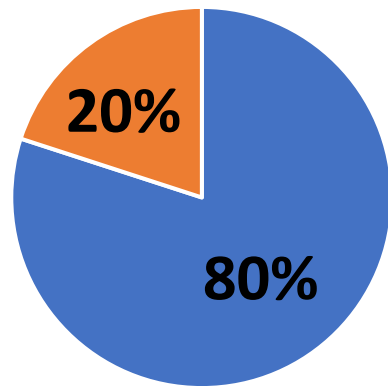4,3/5

Idroscalo 8 km
4,2/5

REVENUES

+ **PRO VERSION**

+ **SPONSORSHIPS** with sport brands

+ **SALE** of anonymized data

FitYourRun

Business plan

YEAR 1	COST €
Application server	1600
Google cloud ML engine	840
Database server	810
Application development	1000
Marketing	800
Buy data from running app	500
TOTAL	**5550**

FUNDINGS

- Development team
- Founders

20%

80%

New Compute User,

Your Estimated Bill *

Estimated Monthly Cost: EUR 270.51

Machine Learning		5.9234	EUR 69.80
Cloud Storage	Storage	2048 GB	EUR 48.42
Cloud Storage	Class A operations (millions)	300 1000	EUR 13.64
Cloud Storage	Class B operations (millions)	1500 1000	EUR 5.46
1 x FitYourRun Application Server	c2-standard-4	730 total hours per month	EUR 133.20
Total Estimated Monthly Cost			**EUR 270.51**

Business plan

YEAR 2+	COST €
Marketing	3000
In-house hardware	10000
Team salary	??
TOT	**13000+??**

REVENUES

- 🟩 Users
- 🟦 Sale data
- 🟨 Sponsorships
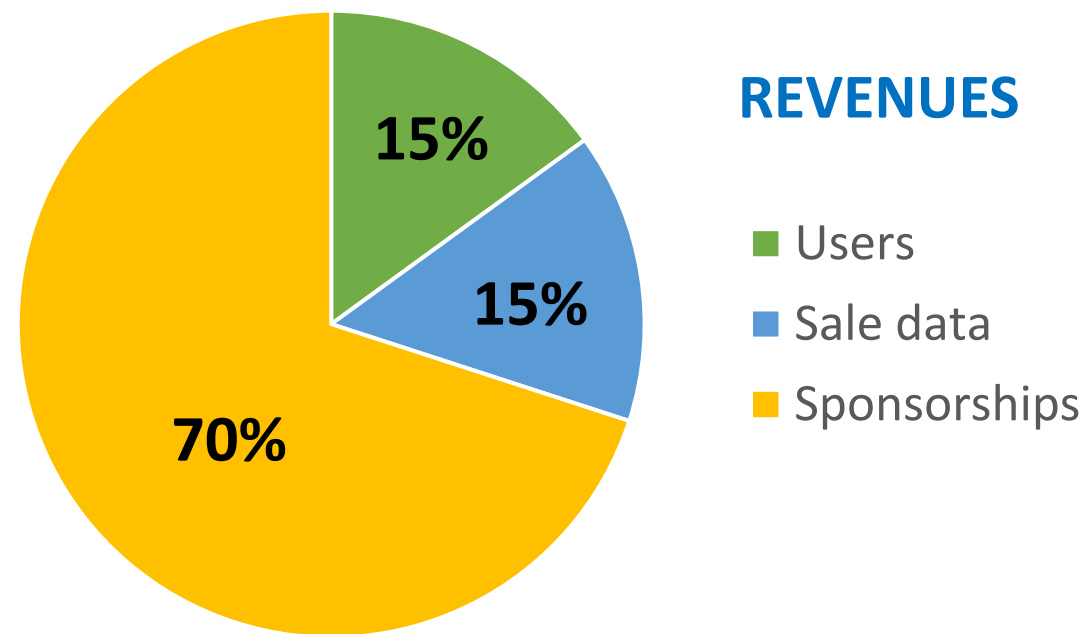
Machine Learning - POC

Clusterization Metrics

The **BATCHED CLUSTER COMPUTATION** (expensive task) is done every month and uses the following metrics

PHYSICAL METRICS

AGE

SEX

$$BMI = \frac{WEIGHT}{(HEIGHT \times HEIGHT)}$$

PERFORMANCE METRICS

OLD RUNS
added at the REGISTRATION

MANUAL INSERTION

AUTO-IMPORTATION

Google Fit **STRAVA**

adidas **RUNTASTIC**

AVERAGE PACE (10km)

$$t_{10}{}^{i} = t_{run} * (10Km / d_{run})^{k}$$

Riegel formula

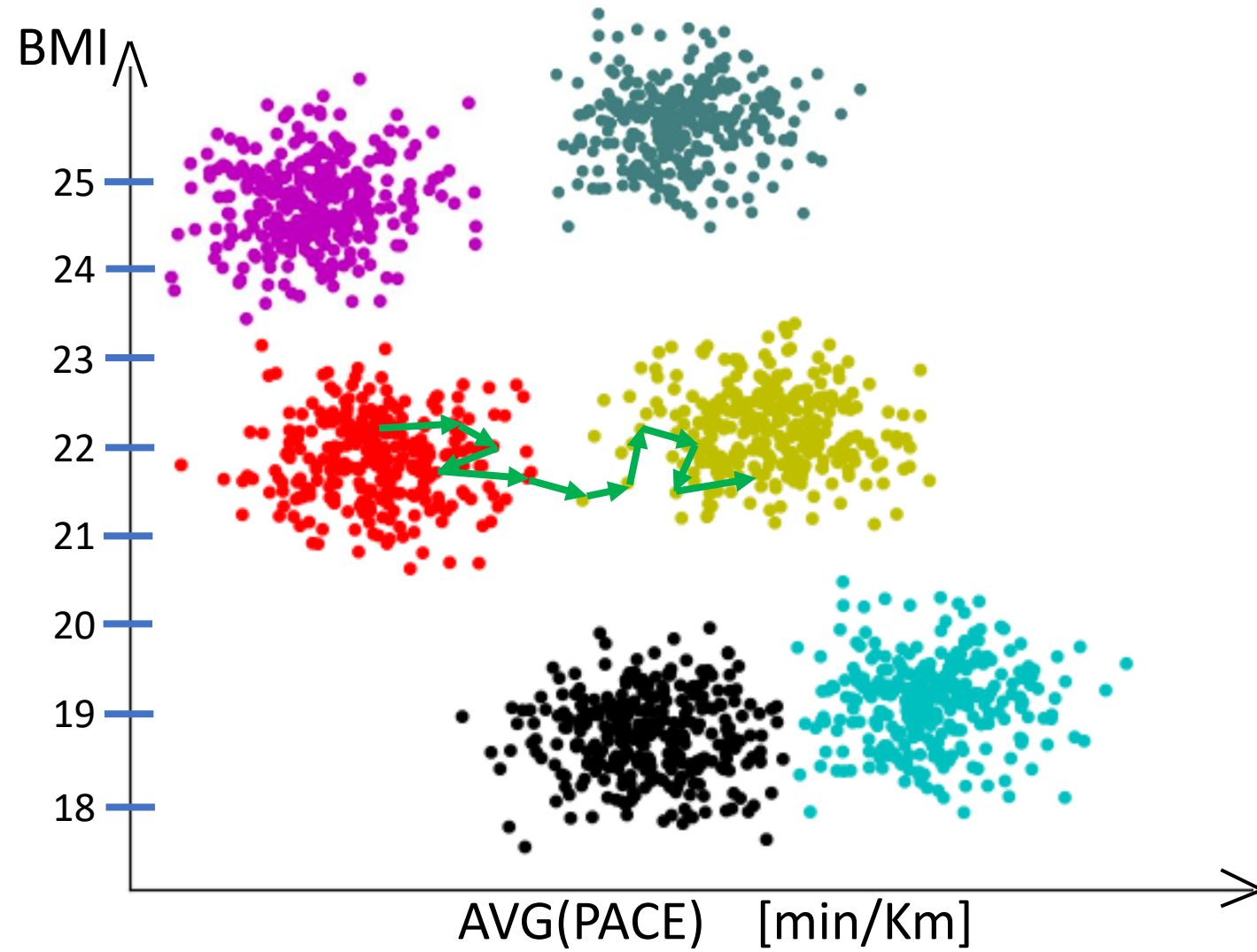$$AVG(PACE) = \frac{\sum_{i=n-10}^{n} t_{10}}{n \times 10Km}$$

TIME IMPROVEMENTS

$$\Delta T = t_{10}{}^{n-10} - t_{10}{}^{n}$$

Fit our un

Users – Cluster Membership

The specific **USER CLUSTER MEMBERSHIP** is recalculated every new user's run

This is a CLASSIFICATION TASK : we assign the user to a specific cluster

BMI

AVG(PACE)   [min/Km]

25

24

23

22

21

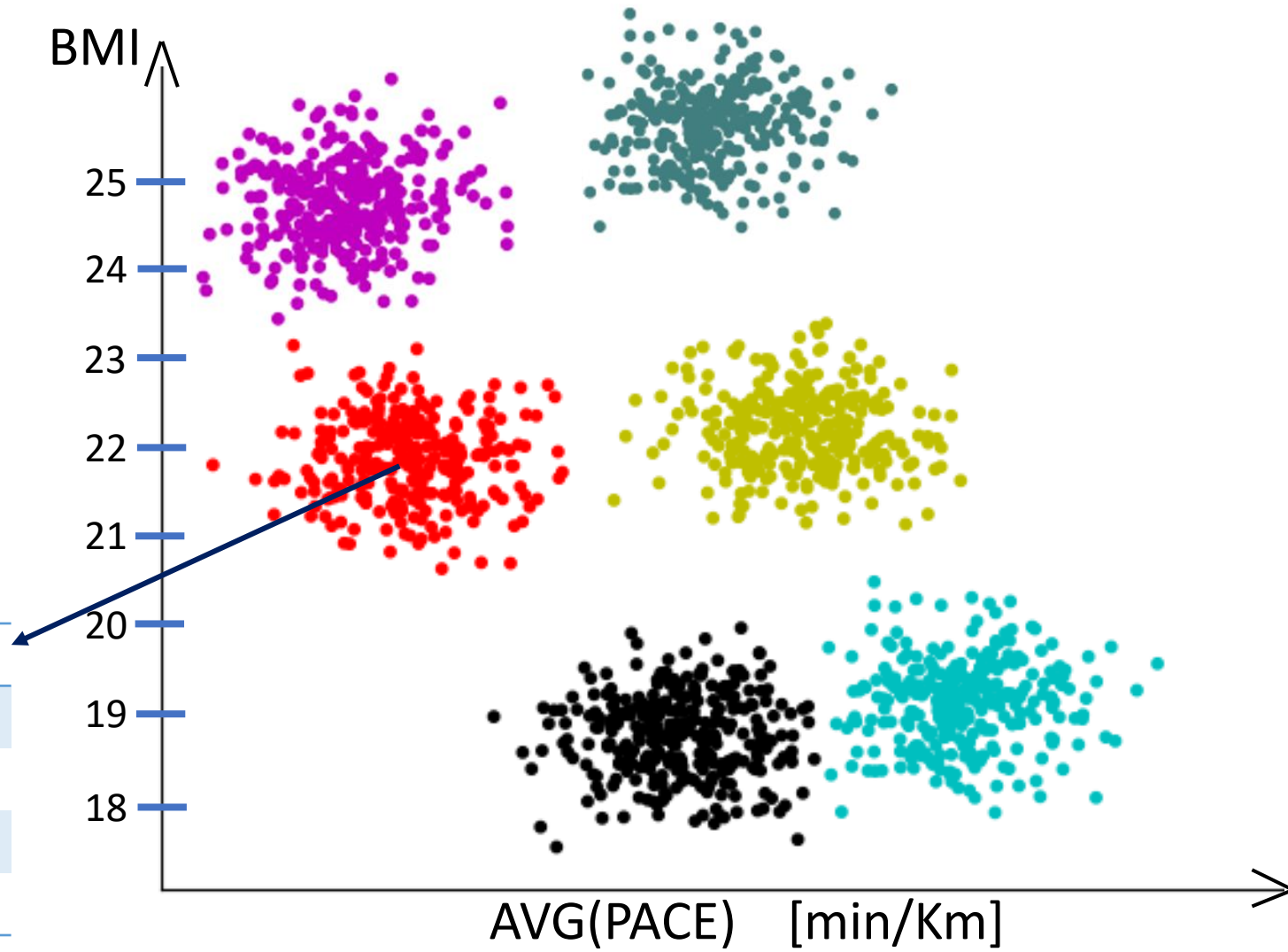20

19

18

Belcao – Gangemi - Comolli

Trainings's Exercizes

Each cluster has a specific set of **BRICKS**, that are the <u>most appropriate exercises</u>

ASSIGNED BY **EXPERTS** every new reclusterization, but in the future we want to automatize the process....

CODE	RUN [Km]	PACE [min/Km]	SCORE	REST [gg]
R1	2+2	6,50	10	3
R2	6	7,15	7	5
R3	3+2	6,25	12	4
....	....	....	....	....

BMI

25
24
23
22
21
20
19
18

AVG(PACE)   [min/Km]

Fit Your Run

Belcao – Gangemi - Comolli

Customized Paths

RUNNING PATHS

EXTRACT

COMPOSED-OF

initial DATA BOUGHT buy other running applications

OLD RUNS imported by the users at the registration

user's FREE RUNS using the application
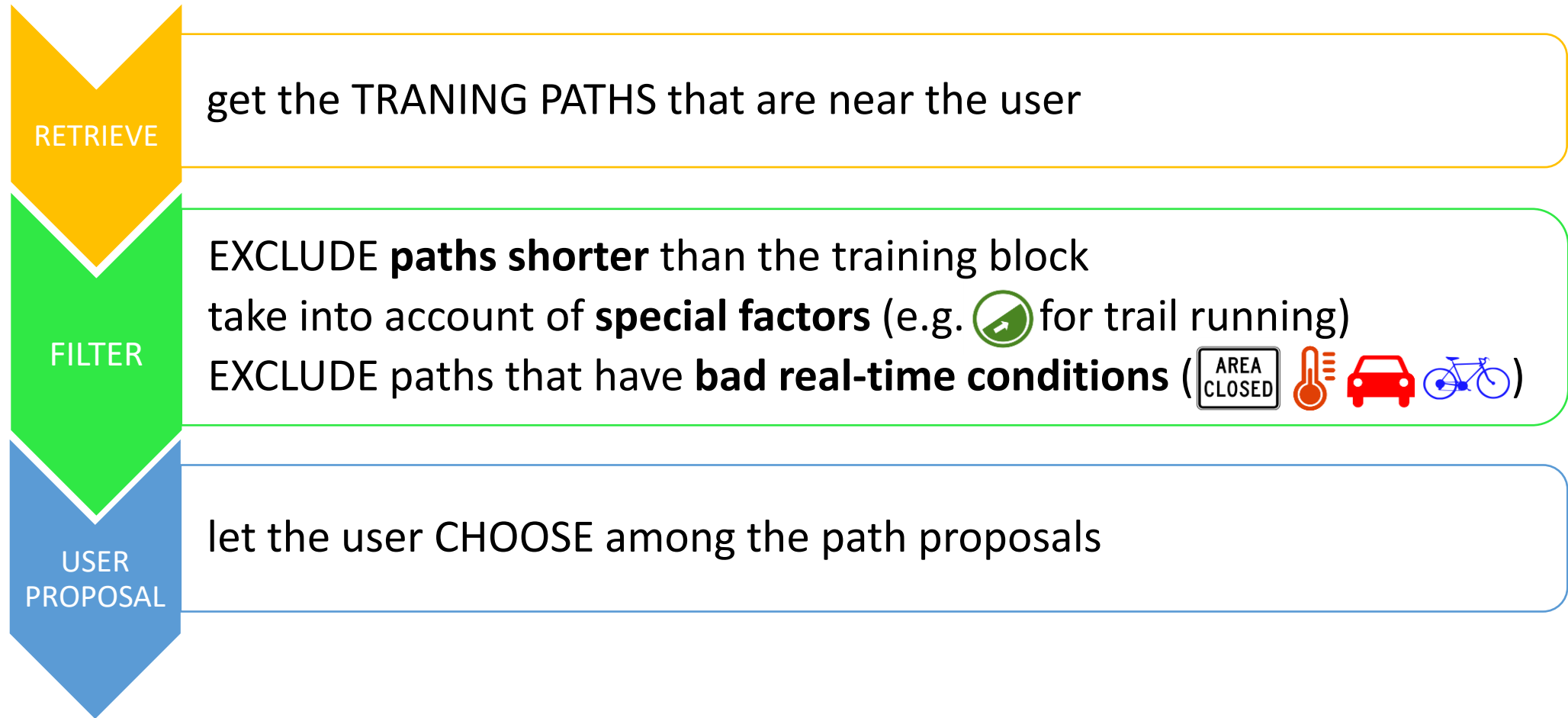
after-workout paths USER'S EVALUATIONS

STREETS

WHAT'S THE PROCESS??

1) Enrich the missing streets data retrieving the infos using the APIs

2) Use this TRAINIG SET for train the **CLASSIFICATION MODEL** and **classifying other streets**

3) Use the model for FINDING **NEW TRAING PATHS** : set of **runnable streets** that are arranged in an **adjacent manner**

Belcao – Gangemi - Comolli

Customized Paths + Training

When the user wants to **START A TRANING SESSION**…

RETRIEVE

get the TRANING PATHS that are near the user

FILTER

EXCLUDE **paths shorter** than the training block

take into account of **special factors** (e.g. for trail running)

EXCLUDE paths that have **bad real-time conditions** ()

USER PROPOSAL

let the user CHOOSE among the path proposals

Use Case

Customized Trainig pt.1

1) MARCO adds his updated **physical data** : Male, 22, 175cm, 70kg → BMI : 22,85

and imports his **old runs** from Google Fit

Run1	5 km	5,58 min/km	20/10/2019
Run2	10 km	6,12 min/km	25/10/2019
Run3	4 km	5,55 min/km	07/11/2019

→ **CLUSTER X**

2) MARCO chooses his final **goal**

MARATHON (42,2 Km)

average pace : 6,30 min/Km — Riegel (10Km) → **5,58 min/Km**

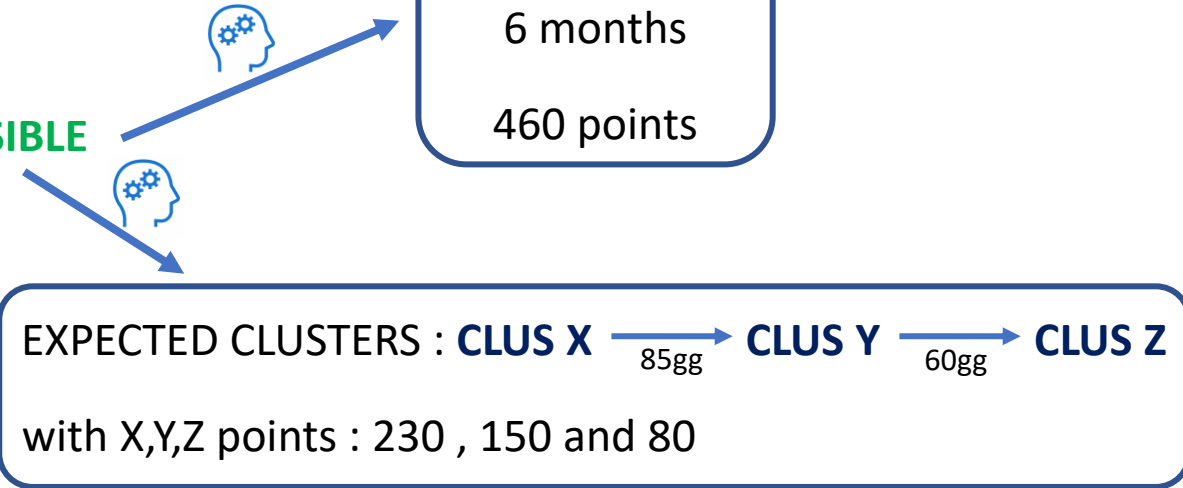3) MARCO specifies his **timing preferences**

SCHEDULE from 31 Google Calendar

maximum of **3 trainings/week** for a maximum amount of **8 months**

Customized Trainig pt.2

2 trainings/week

6 months

460 points

4) The **Fit Your Run** algorithm judges goals & preferences **FEASIBLE**

EXPECTED CLUSTERS : **CLUS X** →85gg→ **CLUS Y** →60gg→ **CLUS Z**

with X,Y,Z points : 230 , 150 and 80

5) MARCO accepts the **PROPOSED SCHEDULE**

BLOCK	RUN	PACE (10Km)	SCHEDULE	OBTAINABLE SCORE	OBTAINED SCORE	REST TIME	CLUSTER
R3	2 + 2 Km	7,50 min/km	10/12/2019	+10 pt	+8 pt	3gg	X
R16	6 Km	8,00 min/km	13/12/2019	+5 pt	+5 pt	4gg	X
R4	3 + 2 Km	7,45 min/km	17/12/2019	+10 pt	+10 pt	4gg	X
R23	3+3+3 Km	6,45 min/Km	03/03/2020	+8 pt		3gg	Y
R34	12 Km	6,40 min/Km	06/03/2020	+12 pt		4gg	Y
R84	8+8 Km	5,35 min/Km	17/05/2020	+15 pt		4gg	Z
R35	15Km	5,30 min/Km	21/05/2020	+18 pt		5gg	Z

33/230 to CLUS Y

0/150 to CLUS Z

0/80 to **COMPLETE GOAL**

Customized Trainig pt.3

6) <u>Before each training</u>, **FitYourun** extracts the **SUGGESTED PATHS** and proposes them to MARCO

7) MARCO **CHOOSE** one among the paths and **STARTS TRAINING SESSION**

IDROSCALO

PARCO SEMPIONE

NAVIGLIO MARTESANA

SCHEDULE EVOLUTION

At the end of every running session the schedule is recomputed (in most of the cases doesn't change)
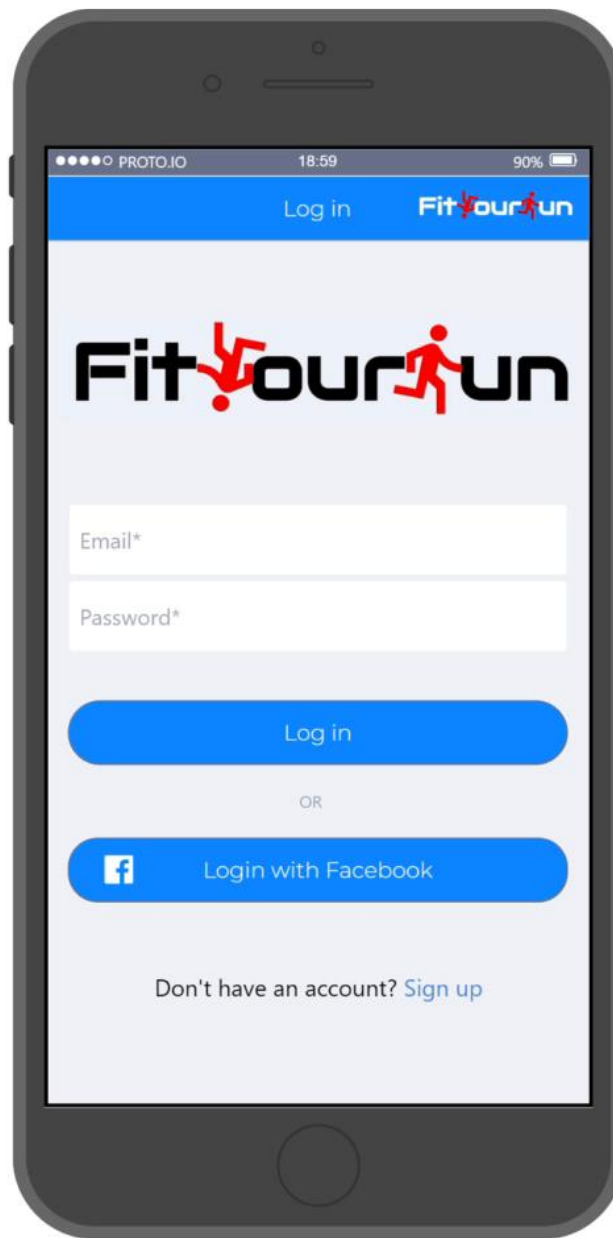
ADD NEW RUNNING SESSIONS if

MARCO doesn't do the training --> +0 pt

Some MARCO's workouts didn't get the full score

ADD ADDITIONAL POINTS if MARCO performs workouts in the "free runner" mode

User Interface - POC

UI Demo

Thank You