DATATHON
AUGUST 2019
One Day Challenge

WHAT IS A DATATHON?

A datathon is a one day challenge event where you will use your analytical and technical skills to tackle a given problem.

WHY ARE WE DOING THIS?

GOALS

» 2 projects per module → Datathon = Mini ML project.

» Challenge yourself.

» Great learning opportunity.

» Learn from teammates.

» Have a great time!

CHALLENGE

Cookie Flea is in trouble

Cookie Flea is a newcomer in the cookie industry.

They want to be the best cookie makers so their goal is to win the Best Cookie Rookie Award this year!

How can they win?

In order to get first place, they want to know which are the characteristics of the best quality cookies.

Then, they can create the ultimate cookie recipe!

What do they need?

Cookie Flea wants a tool to be able to predict the quality of a cookie before starting its production.

Then, they can save money and time during their search of the perfect cookie!

Solution: predict the cookie quality using ML

They bought a cookies dataset!

» **16 columns with cookies measurements and their quality**

» **The rows are the values of different cookie types**

YOUR TASK

» Predict the cookies quality using the Cookie Flea dataset.

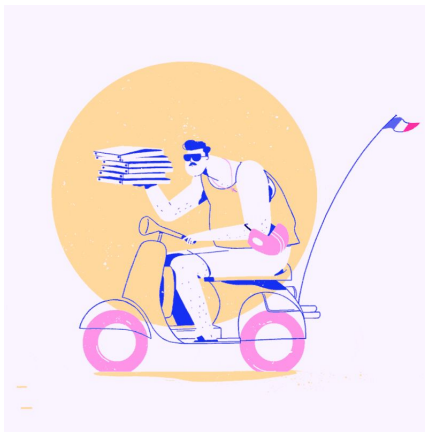» Present your results at the end of the day!

CONSIDERATIONS

DATABASE

You can find the available data here:

driver = 'mysql+pymysql'

ip = '35.239.232.23'

username = 'ironhacker_read'

password = 'ir0nhack3r'

db = cookies

table = cookies_quality

WHAT DO YOU NEED TO DELIVER?

» Jupyter notebook with your analysis.

» Presentation slides.

NOTEBOOK REQUIREMENTS

» You **must include a cleaning method**.

» Include the metric you used to evaluate and your result.

» Add comments and descriptions of the steps you take.

» Justify each decision you make.

» Give the notebook a clear structure.

» Clean code!

WHAT ABOUT THE PRESENTATIONS?

» Keep presentations **under 10 minutes**.

» Use slides.com or Google Slides.

» Present yourselves.

» Remember to tell a story!

HOW WILL YOU BE EVALUATED?

» Quality of your thinking and analysis.

» Technical checkpoints.

» Independent test set.

» Organization & Teamwork.

» Presentation.

INDEPENDENT TEST SET

» At the end of the day, we will provide you with an independent test set.

» This set has a column called quality. It is filled with 0s!

» You will need to fill the quality column with your predicted quality value and send the test set back to us.

REMEMBER: The test set we receive needs to be exactly the same as the one we sent you!

Same columns, same rows!!

* Be careful if you drop columns/rows in your code!

You must only change the values in column quality!

INDEPENDENT TEST SET

You can find a small example of an independent test set in the database:

db = cookies

table = cookies_evaluate_example

Use it to make sure that later in the evening you will be able to provide
the predictions we ask for in the correct format.

CHECKPOINTS

MORNING CHECKPOINT - BY 11:00

» Planning

» Repository

» README file

» Data cleaning

» Analysis plan

◊ Hypotheses

◊ Questions to investigate

AFTERNOON CHECKPOINT - BY 14:30

» Appropriate model selection

» Appropriate metric selection

» Clean code

» Good practices

SCHEDULE

MORNING STAGE

» **8:30h**   **KICK-OFF & TEAM BUILDING ACTIVITY**
» **11:00h**   **MORNING CHECKPOINT**
» **13:00h**   **LUNCH**

MORNING STAGE

- » **8:30h** **KICK-OFF & TEAM BUILDING ACTIVITY**
- » **11:00h** **MORNING CHECKPOINT**
- » **13:00h** **LUNCH**

AFTERNOON STAGE

- » **14:30h** **AFTERNOON CHECKPOINT**
- » **16:00h** **CODE DELIVERY**
- » **16.20h** **SLIDES DELIVERY**
- » **16:30h** **PRESENTATIONS !**
- » **17:30h** **BREAK & SURVEY**
- » **17:55h** **WINNER ANNOUNCEMENT**

TIPS AND TRICKS

LAST TIPS & TRICKS!

» Use a Kanban Board.

» Remember: pull and push.

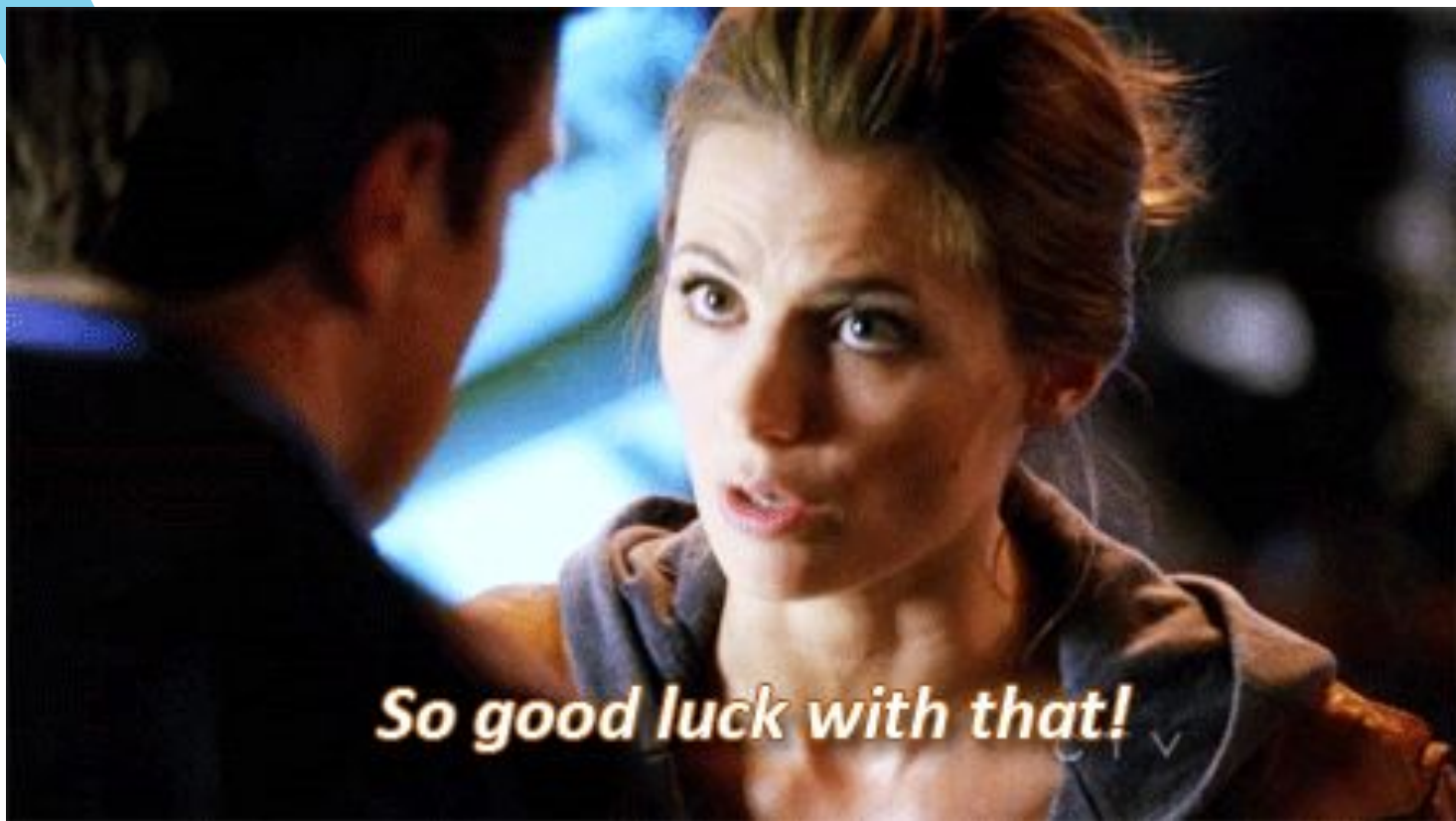» Deliver on time!

» Learn from each other.

I'm not even mad.

LAST TIPS & TRICKS!

» Splitting the work is OK.

» You're stronger together.

WORK FAST, HAVE FUN!

So good luck with that!

WINNER, WINNER CHICKEN DINNER

WINNING CATEGORIES

Best Overall

Honorable Mention

WINNING CATEGORIES

Honorable

Mention

5 POINTS !!!

WINNING CATEGORIES

Best Overall

15 POINTS !!!

Now....

DO YOU WANT TO MEET YOUR TEAM FOR THE DAY?

TEAM 1

Chao-Ting Chang
Miquel Petit
Pau Ferri

TEAM 2

Josep Foradada

Kerim Basbug

Salva Rocher

TEAM 3

Fabia Höhne
Ingrid Munné
Laurent Guerguy

TEAM 4

Javier Pardo

Emanuele De Sanctis

Alberto Abreu

Onur Taskin

TEAM BUILDING ACTIVITY

Rock, Paper & Scissors

Attack Defense Edition

LET THE TAs SHOW YOU HOW TO PLAY!

1. Play the regular Rock, Paper & Scissors game.

2. Did you win? Attack!!

3. Did you lose? Defense!!

4. Eliminate the other teams doing relay.

TEAM BUILDING ACTIVITY

Let's Play!

GOOD LUCK!

If you have any questions please address them to teachers, TA's & mentors directly.

Also, you can use the cohort's Slack channel for this event.