## eleven - supercase challenge

To the attention of the Data Science & Business Analytics students

February 14th, 2022







#### Your preferential interlocutors today



Morand Studer
Partner
+33 6 79 16 29 50
morand.studer
@eleven-strategy.com



+33 6 42 64 56 44 oussama.ennouri @eleven-strategy.com

Oussama Ennouri



+33 6 02 35 67 34 gregoire.lepault@eleve n-strategy.com

Grégoire Lepault



+33 6 24 09 61 73 louis.dumont@elevenstrategy.com

Consultant

1. About eleven

## 2. Case presentations

- a) The Right Price
- b) Airplane Interior Service
- c) Worksite Monitoring

#### 3. General information

- a) Expected output
- b) Practical information

eleven is Europe's first ever specialist strategy firm specifically founded to accompany clients' transformation through the AI and digital revolution, thanks to a unique combination of strategy perspective and hands-on approach



**DIGITAL & AI STRATEGY SPECIALIST** 

STRATEGY X HANDS-ON POSITIONING

supporting crievel executives and organizations from strategic ignition to project scale-up and

A UNIQUE BLEND OF SKILLS

und data science ones, thanks to its out consultants that master the continuously evolving

DISTINCTIVE ENTREPRENEURIAL MINDSET

#### CSR AT THE CORE OF OUR DNA

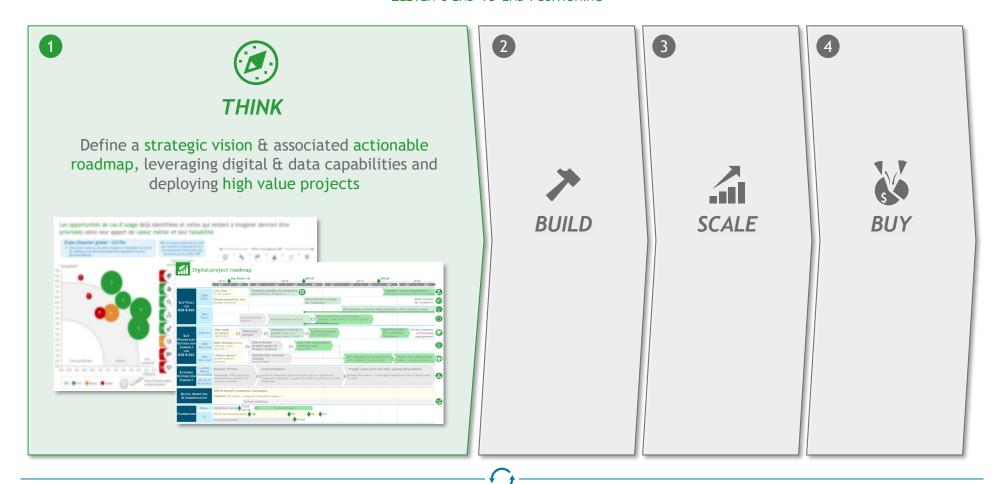
Our approach enables our clients' projects to meet CSR expectations. Both financial and CSR impacts are tracked, proven and reported thanks to our mastering of digital and AI levers

Our projects are built around ongoing major digital & data disruptions, including new modes of production, consumption and collaboration, which represent major challenges for existing players

eleven's 'plate tectonics' of disruptions Towards new internal organization Towards new customer expectations Collaborative Planning High expectations on user Forecast Replenishment experience (UX) / Mobile first Social recommendation, new Open Source, Open Data Adaptative **New Coverage** trusted third-parties Bandwidth & **Technologies Robotics** Omnichannel customer Open IT architectures based Low Latency on APIs, Open Source experiences Batteries & Agile & lean Additive Immediacy & transparency Energy project management Manufacturing Harvesting New working schemes: Staff c Sharing economy demand, collaborative development platforms, open innovation -GDand MEMS Modular IT **TECHNOLOGICAL OPERATIONAL** INNOVATIVE **EFFICIENCY OFFERS ENABLERS** Blockchain **Materials Platforms** Asset light & global market Quantum Computing AdTech / conversational bots Software is eating the world **VR** / robots Digital supply chain Mass customization Deep Learning AR New business models Zero marginal cost (Pay as you use, freemium...) Machine Unexpected convergences: Cloud Crowdsourcing Learning blurred market boundaries Towards new production models Towards a new offer



ELEVEN'S END-TO-END POSITIONING



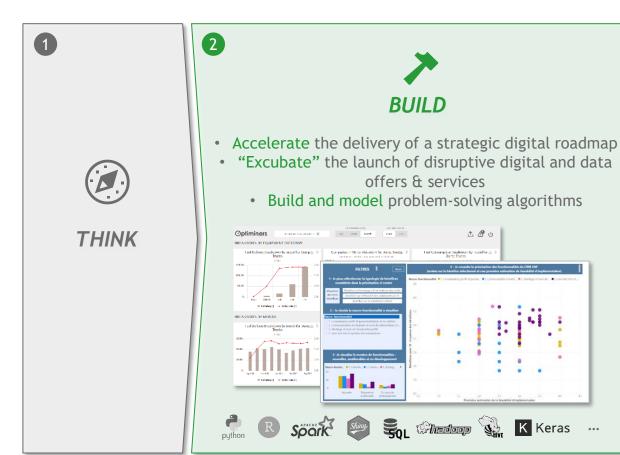
Continuous interactions with client's c-levels, business units and digital & data teams

Proven ability to lead projects and teams remotely thanks to the most advanced digital tools



eleven

ELEVEN'S END-TO-END POSITIONING







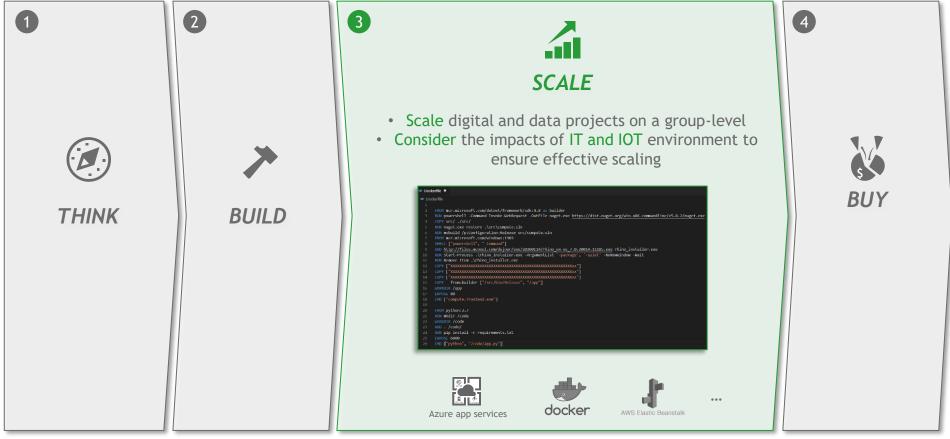
Continuous interactions with client's c-levels, business units and digital & data teams

Proven ability to lead projects and teams remotely thanks to the most advanced digital tools



Sources: eleven strategy

ELEVEN'S END-TO-END POSITIONING



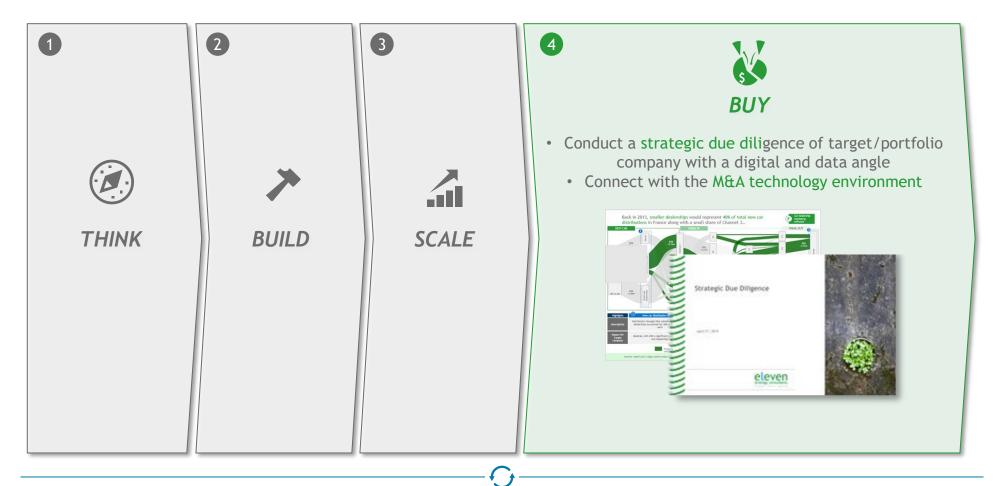


Continuous interactions with client's c-levels, business units and digital & data teams

Proven ability to lead projects and teams remotely thanks to the most advanced digital tools



ELEVEN'S END-TO-END POSITIONING



Continuous interactions with client's c-levels, business units and digital & data teams

Proven ability to lead projects and teams remotely thanks to the most advanced digital tools



eleven

eleven serves top large cap clients and leading mid cap players across several key industries as well as leading International large cap and smid cap Private Equity funds



1. About eleven

## 2. Case presentations

- a) The Right Price
- b) Airplane Interior Service
- c) Worksite Monitoring

#### 3. General information

- a) Expected output
- b) Practical information

eleven supercase challenge: work on real-life challenges coming from various industries leveraging your dual expertise in business and data science



#### **Exercise:**

- ✓ Find your group of 4/5 people
- ✓ Choose among the 3 available topics
- ✓ On your topic, put yourself in a data consultant's shoes:
  - ✓ What is the context of the company? (value chain, competition, technological trends, etc.)
  - √ How could you generate added-value for them using potential of new technological enablers? (time-series, NLP, computer vision, etc.)
  - ✓ What should be the best way to present your first results? (data visualization, 1st demo, etc.)
  - ✓ How much should we trust your first results? (accuracy of your model, etc.)
  - √ What would be the next steps if you wanted to keep improving your results? (fine tuning, new model, advanced feature engineering, etc.)
- ✓ Synthesize your work in a few slides (up to 15 slides)
- ✓ Upload your work on the Sharepoint (presentation, code, etc: see details at the end of this presentation)
- ✓ Present to the jury and prepare for questions



#### The Right Price

How to estimate a property price?



#### Airplane Interior Services

How to increase customers satisfaction?



#### Worksite Safety Monitoring

How to monitor worksites' safety?

# The Right Price

improving definition and computation of mutability score in Île-de-France



Context: the client, a player in real-estate industry, wants to build a robust purchase/sale price estimation model

Value chain of a real estate development project

For illustration purposes



Prospection



Construction



Operation



Renovation



Sale

What can be constructed?



• Data: PLU (Plan Local d'Urbanisme)

What is constructed today?



• Data: Databases of constructed buildings (height, areas, ...)

What is purchase/sale prices?



Data: Databases of historical transactions

Scope of the supercase

What is the willingness to sell/buy?



Data: NA



- Process the data from historical transactions to get geolocation information
- Build one or several price estimation model for new construction (exclusively apartments)
- Present the results to your client with the test dataset that will be provided to you

Data available: you have access to data from official mutation databases in Île-de-France, non-geo-localized except but for cadastral id, with large and various information on the properties

Note that as in any data science project, the data may require further data engineering before being fully leverageable

datemut	anneemut	moismut	coddep	libnatmut	vefa	valeur fonciere	codtypebien libtypbien
22/10/2018	2018	10	75	Vente	FALSE	147000.0	121 UN APPARTEMENT
26/10/2015	2015	10	75	Vente	FALSE	95000.0	131 UNE DEPENDANCE
13/06/2019	2019	6	75	Vente	FALSE	1900000.0	152 BATI MIXTE - LOGEME
20/04/2017	2017	4	75	Vente	FALSE	1509000.0	121 UN APPARTEMENT
06/03/2015	2015	3	75	Vente	FALSE	28250.0	131 UNE DEPENDANCE
01/08/2017	2017	8	75	Vente	FALSE	500000.0	121 UN APPARTEMENT
03/07/2019	2019	7	75	Vente	FALSE	367000.0	121 UN APPARTEMENT
27/10/2016	2016	10	75	Vente	FALSE	90000.0	121 UN APPARTEMENT
23/11/2018	2018	11	75	Vente	FALSE	135000.0	121 UN APPARTEMENT
26/08/2016	2016	8	75	Vente	FALSE	190000.0	121 UN APPARTEMENT
28/12/2017	2017	12	75	Vente	FALSE	300000.0	14 ACTIVITE
11/06/2014	2014	6	75	Vente	FALSE	730000.0	121 UN APPARTEMENT
02/05/2018	2018	5	75	Vente	FALSE	935000.0	121 UN APPARTEMENT
11/05/2016	2016	5	75	Vente	FALSE	535000.0	121 UN APPARTEMENT
07/10/2016	2016	10	75	Vente	FALSE	339050.0	121 UN APPARTEMENT
11/07/2018	2018	7	75	Vente	FALSE	413438.0	121 UN APPARTEMENT
22/10/2018	2018	10	75	Vente	FALSE	3327000.0	14 ACTIVITE
06/02/2018	2018	2	75	Vente	FALSE	650000.0	122 DEUX APPARTEMENTS
04/09/2018	2018	9	75	Vente	FALSE	375000.0	14 ACTIVITE
09/02/2015	2015	2	75	Vente	FALSE	163000.0	121 UN APPARTEMENT
18/06/2020	2020	6	75	Vente	FALSE	768490.0	121 UN APPARTEMENT
07/06/2017	2017	6	75	Vente	FALSE	15000.0	14 ACTIVITE
13/11/2014	2014	11	75	Vente	FALSE	480000.0	121 UN APPARTEMENT
30/03/2018	2018	3	75	Vente	FALSE	823000.0	121 UN APPARTEMENT
14/09/2015	2015	9	75	Vente	FALSE	250000.0	121 UN APPARTEMENT
16/02/2018	2018	2	75	Vente	FALSE	2873000.0	122 DEUX APPARTEMENTS
11/03/2019	2019	3	75	Vente	FALSE	525000.0	121 UN APPARTEMENT



√ After two days, you could be provided with a geo-localized version of this dataset

Resources: You are free to use any resources you want, here are some recommendations to help you get started





We highly recommend you use Python even though same kind of results could be achieved with similar tools

#### ---- Relevant libraries









To develop a wide range of ML models

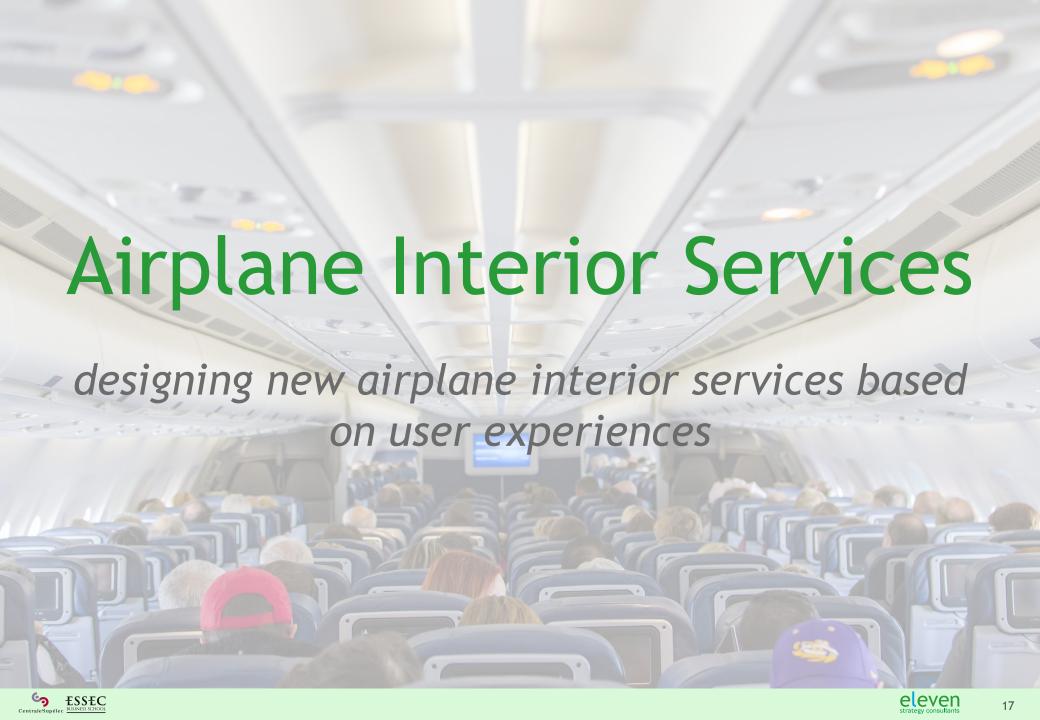
To use geo-localized visualization tools

To use geo-localized advanced ML models

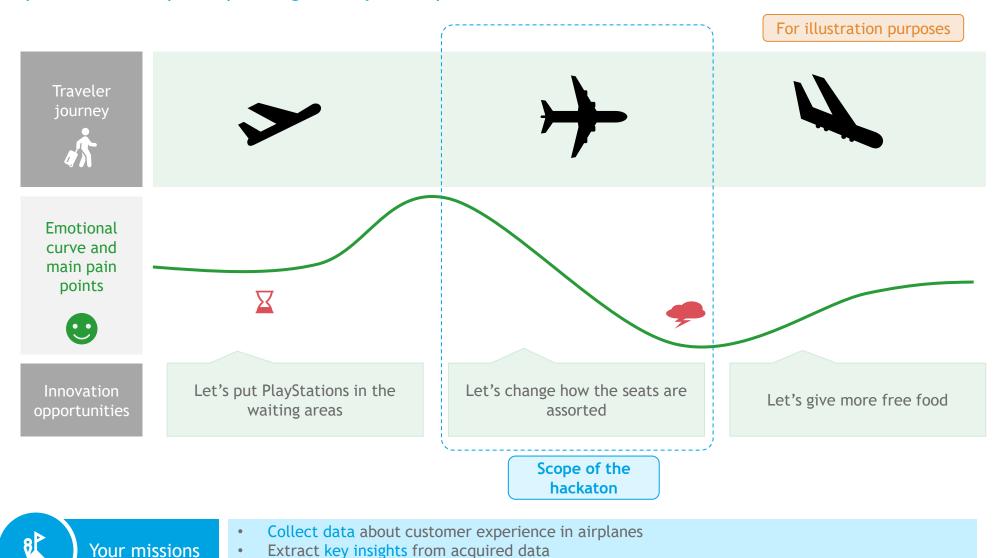
To efficiently develop a dashboard / front-end

----- For information: the process of geo-location is independent from the rest of the challenge





# Context: the client, a player in the aeronautic industry, wants to understand how the user experience of airplane passengers may be improved



Suggest to your client mitigations solutions based on extracted insights

Data available: you do not have initial access to any data, except for your computer science skills and the whole world wide web, be creative to leverage existing information

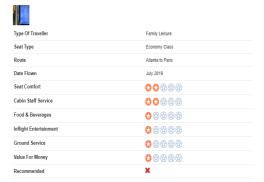
Example of the kind of data you should be looking for to complete your supercase





#### "a horrible airline to fly" Rachel Beale (United States) 28th July 2019

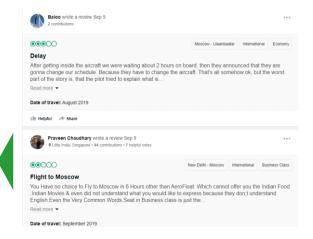
Trip Verified | Atlanta to Paris. Auf France is just a homible affiline to fly, especially compared to delta. They do not update their movies on their tiny screen they put in front of you it's the same movie names repeated over and over to make it seem like there are choices. I found the food disgusting and they have no WiFri on the plane. I would not recommend flying Air France at all unfortunately due to work and my routes? We had to use them 4 times.



country	" number_of_reviev "	date	* rating	y header	▼ body ▼ aircraft	* cabin	→ airline
Netherlands	NA	2019-11-11	9	"pretty decent airline"	✠Trip Verified   Moroni to M NA	NA	ab-aviation
United Kingdom	26 reviews	2019-06-25	1	"Not a good airline"	✠Trip Verified   Moroni to Ar E120	NA	ab-aviation
United Kingdom	26 reviews	2019-06-25	1	"flight was fortunately shor		NA	ab-aviation
Serbia	NA	2019-09-28	'n	"I will never fly again with A	Adria" Not Verified   Please do a favor • NA	NA	adria-airways
Netherlands	NA	2019-09-24	1	"it ruined our last days of he	oliday åœ Trip Verified   Do not book NA	NA	adria-airways
Austria	NA	2019-09-17	1	"Had very bad experience"	✠Trip Verified   Had very bac CR 900	NA	adria-airways
Switzerland	NA	2019-09-06	1	"worse than the budget airl	lines" Not Verified   Ljubljana to Zürii NA	NA	adria-airways
Germany	NA	2019-08-24	1	"book another company"	Not Verified   First of all, I am nc Bombardier C	RJ NA	adria-airways
Switzerland	NA	2019-08-06	1	"combined two flights"	✠Trip Verified   Worst Airlin NA	NA	adria-airways
Germany	35 reviews	2018-10-12	8	"the crew was nice"	✠Trip Verified   Ljubljana to I NA	NA	adria-airways
Germany	NA	2018-10-05	1	"Very bad experience overa	all" Not Verified   Zurich to Ljubljane NA	NA	adria-airways
United States	NA	2018-07-29	1	"bad customer service"	✠Trip Verified   Vienna to Sc NA	NA	adria-airways
France	NA	2018-07-19	2	"overall very poor"	✠Trip Verified   We were tra NA	NA	adria-airways
Slovenia	NA	2018-05-30	2	"Would not fly again"	㜠Trip Verified   Ljubljana to NA	NA	adria-airways
Czech Republic	NA	2018-05-24	3	"very unpleasant experienc	te" åæ Trip Verified   A very unple A319	NA	adria-airways
Slovenia	NA	2018-05-04	10	"Flight was very comfortabl	le" åœ Trip Verified   Frankfurt to NA	NA	adria-airways
Germany	NA	2018-03-11	1	"delayed for more than 2 ho	ours" 㜠Trip Verified   Ljubljana to NA	NA	adria-airways
United States	NA	2017-12-05	3	"should be ashamed of thei	ir operāce Trip Verified   Ljubljana to "ATR-72	NA	adria-airways
Slovenia	NA	2017-11-20	9	"Two nice short flights"	✠Trip Verified   Two nice sho CRJ700 / ATR7	2 NA	adria-airways
Finland	NA	2017-10-27	2	"extremely bad service"	âce Verified Review   We were NA	NA	adria-airways
United States	NA	2017-09-16	2	"never fly this airline again"		NA	adria-airways
Switzerland	NA	2017-04-19	9	"can't remember a flight de		NA	adria-airways
Austria	11 reviews	2017-01-27	8	"seat was quite comfortable	e " åce Verified Review   Ljubljana Canadair 700	NA	adria-airways

ate	▼ unknown	Company	▼ Flight
	From Istabul to Athens.		
29/12/2	019 Seat 3C doesnt have any extra legroom.	aegean-airlines	airbus_a319-100
	I traveled Athens-Zurich with Aegean		
	Airlines and the SX-DNB aircraft. I chose the 21A position. This		
	place is awkward. The foot space is minimal, and the window is in		
09/06/2	018 an uncomfortable position. The display is two rows ahead.	aegean-airlines	airbus a320-200
	very tiny leg room as I am quite tall.		
25/05/2	018 flight attendants friendly and meal ok.	aegean-airlines	airbus a320-200
	Exit row 12F was nice for the 2.5 hour		
	ZRH-ATH flight and back. Quite a few empty seats throughout the		
10/04/2	018 cabin both ways, and a nice meal service.	aegean-airlines	airbus a320-200
15/04/2	ota Cabin both ways, and a nice mear service.	aegean-airiines	alrous_aszo-zoo
	It is a wonderful seat at no additional		
	cost (subject to availability during online check in) - more legdom		
22/07/2	017 and reclining.	aegean-airlines	airbus a320-200







✓ After two days, you will be provided with a scrapped dataset from Skytrax and Tripadvisor

Resources: You are free to use any resources you want, here are some recommendations to help you get started

Languages



We highly recommend you use Python even though same kind of results could be achieved with similar tools

#### ---- Relevant libraries













To efficiently develop a dashboard / front-end

Streamlit

----- For information: the origin source of the data



https://www.seatguru.com/browseairlines/browseairlines.php

To develop topic-modelling algorithm as well

as sentiment-analysis tools

https://www.seatguru.com/airlines/Aegean Airlines/Aegean Airlines Airbus A320-200.php



https://www.airlinequality.com/review-pages/a-z-airline-reviews/

Illustration: Topic modelling and sentiment analysis enable to derive valuable insights from customer reviews

NLP task

#### **TOPIC MODELLING**



Aims at extracting topics from text entities.

Typical approaches:

- Latent Dirichlet Allocation
- Classification
- Vectorization and clustering
- . . .

Typical challenges

- Number of topics
- Granularity of the analysis: full document, sentences, sub-sentences...



#### SENTIMENT ANALYSIS



Aims at classifying a sentence as expressing a positive, negative or neutral opinion

Typical challenges

- Direct vs. comparative opinions
- Explicit vs. implicit opinions
- Granularity of the analysis: full document, sentences, subsentences...

Topic: crew

Sentiment: -0.8



The service was

bad to nonexistent. 2 crew members for a

rather large area.





Context: the client, a player in the construction industry is building a worksite monitoring solution to evaluate and mitigate risks on worksites

For illustration purposes

1

#### Data acquisition

- 7 camera deployed on several test construction site
- Storage in a centralized database



#### **Detection model**

- Detection automation based on deep learning
- Training on 1500 pictures with 11 000 annotations



#### Analysis model

- Clustering of images to group together images from the same camera
- Heatmap of workers' presence during the day
- Storage of results in a table



#### Reporting

- Synthesis of daily human activities contextualized in a picture
- User-friendly exposition of the results
- Identify other value propositions relying on the detection model

Scope of the supercase



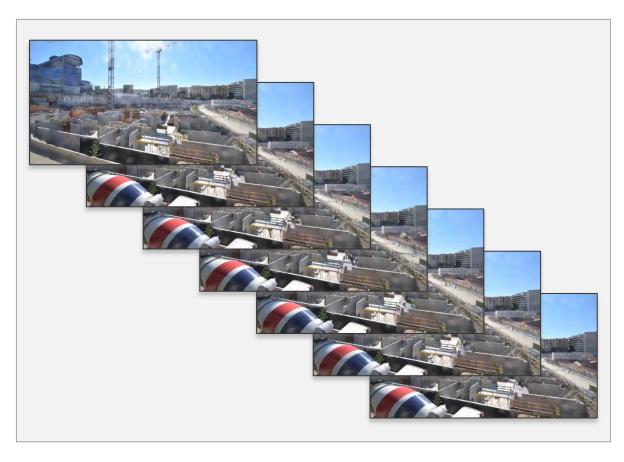
- Build a worker detection model based on annotated data
- Cluster the images to process each camera independently (prerequisite to draw heatmaps)
- Build a heatmap for each worksites with density of workers during the day
- Expose your result and performance on a dashboard with a test dataset that will be provided to you
- Identify additional value that could be brought by the detection model



Data available: you have access to pictures from different worksites, with the corresponding labels giving you information about on-site workers

Example of worksite pictures

Example of labels (json format)



```
"description": "",
"tags": [],
"size":
       {"height": 720, "width": 1280},
       "objects":
               [{"id": 346688445,
               "classId": 1285799,
               "description": "",
               "geometryType": "rectangle",
               "labelerLogin": "Raziajuthy",
               "createdAt": "2020-02-04T09:08:12.615Z",
               "updatedAt": "2020-02-28T11:38:28.963Z",
               "tags": [],
               "classTitle": "People",
               "points":
                       {"exterior": [[153, 622], [177, 647]],
                       "interior": []}},
               {"id": 346688444,
               "classId": 1285799,
               "description": "",
               "geometryType": "rectangle",
               "labelerLogin": "Raziajuthy",
               "createdAt": "2020-02-04T09:08:12.615Z",
               "updatedAt": "2020-02-28T11:38:28.963Z",
               "tags": [],
               "classTitle": "People",
               "points":
                       {"exterior": [[851, 683], [865, 720]], "interior": []}},
               {\"id": 346688443,
               "classId": 1285811,
               "description": "",
               "geometryType": "rectangle"
```

Resources: all classical Machine Learning might come handy and a good understanding of Computer Vision libraries will be helpful

----- Computer Vision libraries



Standard library for deep learning



Standard library classical computer vision



An alternative to Pytorch

..... ML + Viz libraries



To develop a wide range of ML models



To exploite model's output and aggregate them



To efficiently develop a dashboard / front-end

#### 1. About eleven

## 2. Case presentations

- a) The Right Price
- b) Airplane Interior Service
- c) Worksite Monitoring

### 3. General information

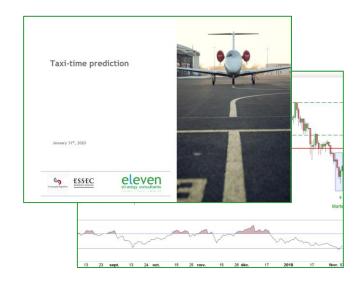
- a) Expected output
- b) Practical information

#### **Expected output**

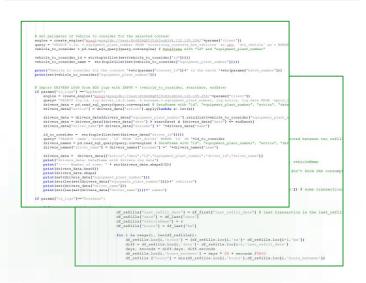
You are expected to deliver the following:

- The file with your code (for this assignment we highly recommend using Python)
- A PowerPoint presentation of your work (including your experiment process, your thoughts, the hardships you had to overcome...)

#### **PRESENTATION**



#### CODE





- 1. About eleven
- 2. Case presentations
  - a) The Right Price
  - b) Airplane Interior Service
  - c) Worksite Monitoring
- 3. General information
  - a) Expected output
  - b) Practical information

#### The Slack workplace

For this challenge, we opened a Slack workplace for you to ask your questions when eleven consultants are not on campus

On this workplace, you will find three channels:

- 1) **général**: for all questions and information related to the organization of the challenge
- **TheRightPrice**: for all questions specifically related to *TheRightPrice* case
- **3) AirplaneInteriorServices:** for all questions specifically related to *AirplaneInteriorServices* case
- 4) WorksiteSafetyMonitoring: for all questions specifically related to WorksiteMonitoring case

Additional information may also be pinned in these channels (schedules, classroom numbers, etc.)

Please use the right channel to ensure fluidity of the interactions Before asking something, also make sure that the requested information has not been given already;)



You may download the slack application on your device or access it via your usual browser



#### Download instructions & submission process

How to download datasets?

You can download datasets and potential additional information at the following links:

- Case #1: AIS - get your files here

Note: no files shared for the 2 first days

- Case #2: TRP - get your files here



Note: geo-localized data will not be shared before the 2 first days

- Case #3: WSM - get your files here



How to submit your works?

Each team will receive a link by email with a Sharepoint folder to submit their assignment (both Presentation + Code)



#### Notes:

- You can organize your folder as you wish.
- You can keep old files that should not be submitted to the jury in a *O\_Archives* folder

# Schedule for the five days: the timing may be short, do not hesitate to split the work between the members of the team



	Monday 14 <sup>th</sup>	Tuesday 15 <sup>th</sup>	Wednesday 16 <sup>th</sup>	Thursday 17 <sup>th</sup>	Friday 18 <sup>th</sup>	
AM session	9.00 am Kick-off session Amphi F3.05 Breguet - 2 <sup>nd</sup> floor	Free working session  Classrooms	Free working session  Classrooms	Free working session  Classrooms	9.00 am > 1.00 pm Pitches session  Classrooms Breguet - 2 <sup>nd</sup> floor  Pitches will take place in front of	
	Training (NLP/CV)  Classrooms Breguet - 2nd floor  Support session (TRP)  Classrooms	Breguet - 2 <sup>nd</sup> floor	Breguet - 2 <sup>nd</sup> floor	Breguet - 2 <sup>nd</sup> floor	a 3 people jury: 2 technical profiles and 1 business profile	
	<b>Q&amp;A - tech</b> Amphi F3.05  Breguet - 2 <sup>nd</sup> floor	<b>Q&amp;A - tech</b> Amphi F3.05  Breguet - 2 <sup>nd</sup> floor	<b>Q&amp;A - tech</b> Amphi F3.05  Breguet - 2 <sup>nd</sup> floor	<b>Q&amp;A - business</b> Amphi F3.05  Breguet - 2 <sup>nd</sup> floor	3.00 pm Closing session	
PM session	2.00pm - The Right Price	2.00pm - The Right Price	2.00pm - The Right Price	2.00pm - The Right Price	Amphi F3.05 Breguet - 2 <sup>nd</sup> floor	
	2:000pm - Airplane Interior Services			3.00pm - Airplane Interior Services	The best team of each topic will pitch in front of the whole class	
	2.00pm - Worksite Safety Monitoring	5.00pm - Worksite Safety Monitoring	5.00pm - Worksite Safety Monitoring	4.00pm - Worksite Safety Monitoring		

#### Final presentation details and best practices:

On Friday 18th, you will have to present your work in front of a jury during a closed-door session

The modalities of the presentation will be as follow:

- 10min group pitch based on a PowerPoint presentation
- ~10min Q&A session with the jury
- ~5min debrief from the jury

For each supercase, a winner will be announced. The three winners will then present their work to the other students (same modalities with questions from the students)

The presentation must be as professional as possible. Here are some advices and best practices that may be useful:





- Be concise and precise: focus on the most important messages, as you only have 20 minutes to present the work achieved for the entire week. You should limit the number of slides you present (you can still add appendices if needed)
- Be organized as a team: split up the speaking time between the team members beforehand to make it smoother
- Be honest: tell where you encountered issues or challenges
- C-suite level: you should convince both the CEO and the CTO/CDO of the company

#### Evaluation criteria:

\*The contribution of each criteria may not reflect the actual value in a final mission restitution

Although different in their essence, the cases will be graded based on similar criteria. NB: any provided code will be tested in order to ensure its good functioning.

Topic		Description	Contribution to the final grade*	
Engagement		The engagement of the team during the exercise (how far you've gone, how autonomous you have been, how much you have asked questions when stuck, etc.)	2 points	
aspect	Presentation quality	The quality of your final presentation: how professional it looks (slide quality), how clear and complete it is (storytelling), how pertinent your answers are, etc.	3 points	
Business aspect	Business methodology	The creativity and relevance of the methodology (i.e. scientific approach) you choose regarding the problem you try to solve, and the data provided, the business sense behind your methodology and the pragmatism of your presentation	6 points	
ct	Technical choices	The explanation of your technical choices and your ability to present them in non-technical terms	3 points	
Technical aspect	Model efficiency or Analysis relevance	The performance of your model (specific to each case), the relevance for the problem in question, the quality of the analysis led	5 points	
Tech	Code good practices	Your code must be well structured, easy to run and easy to understand with clear readme and requirements.	1 point	

Please note that all groups will be graded at the end of the week





# Enjoy the challenge!