Data Science Project

Aug 19, 2020

Our Team

Sarah Kurmulis

Engineering Physics & Business Development

Olena Levchun

Entrepreneurship & Project Management

Matthieu Bornet

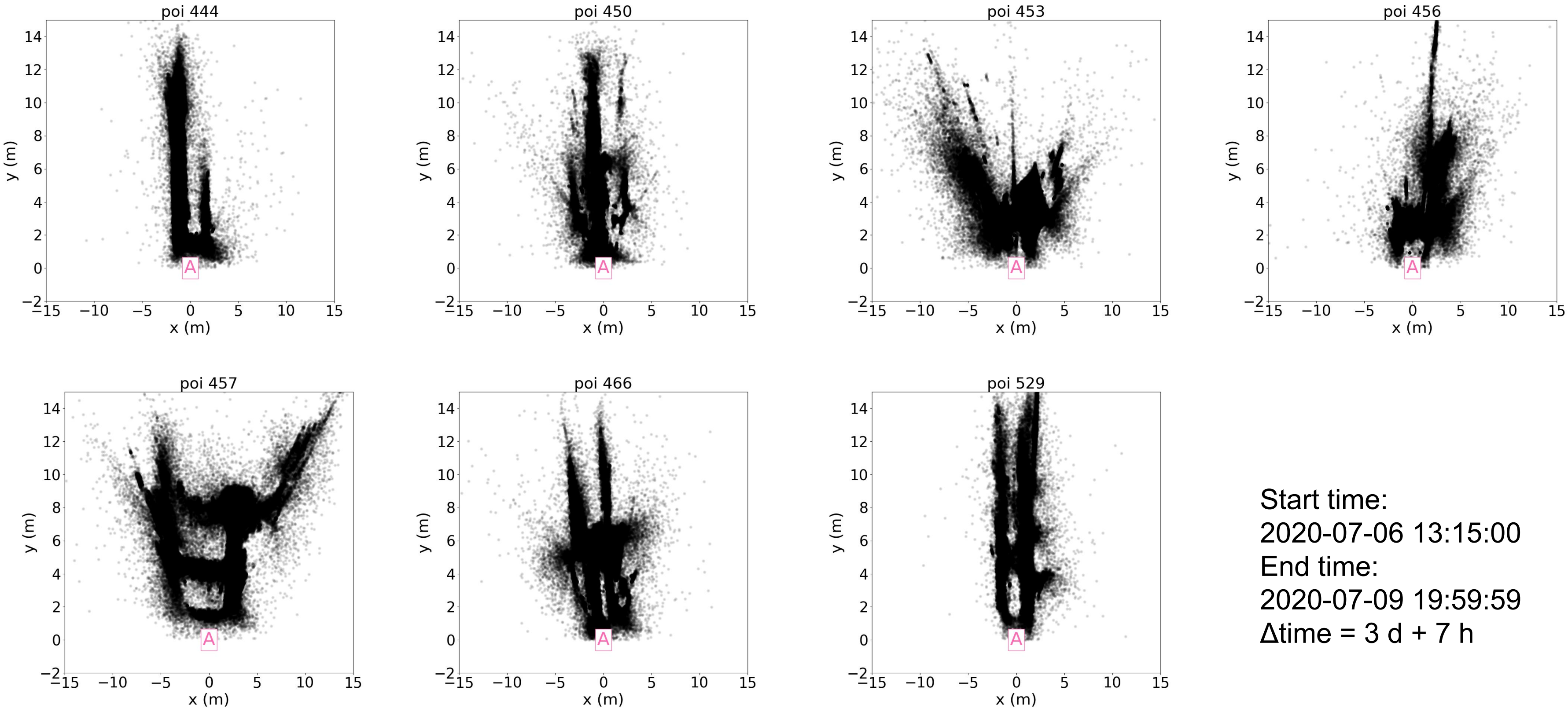
Finance & Digital Strategy

Business Problem

Improve **Smart Signage** & **Targeting**

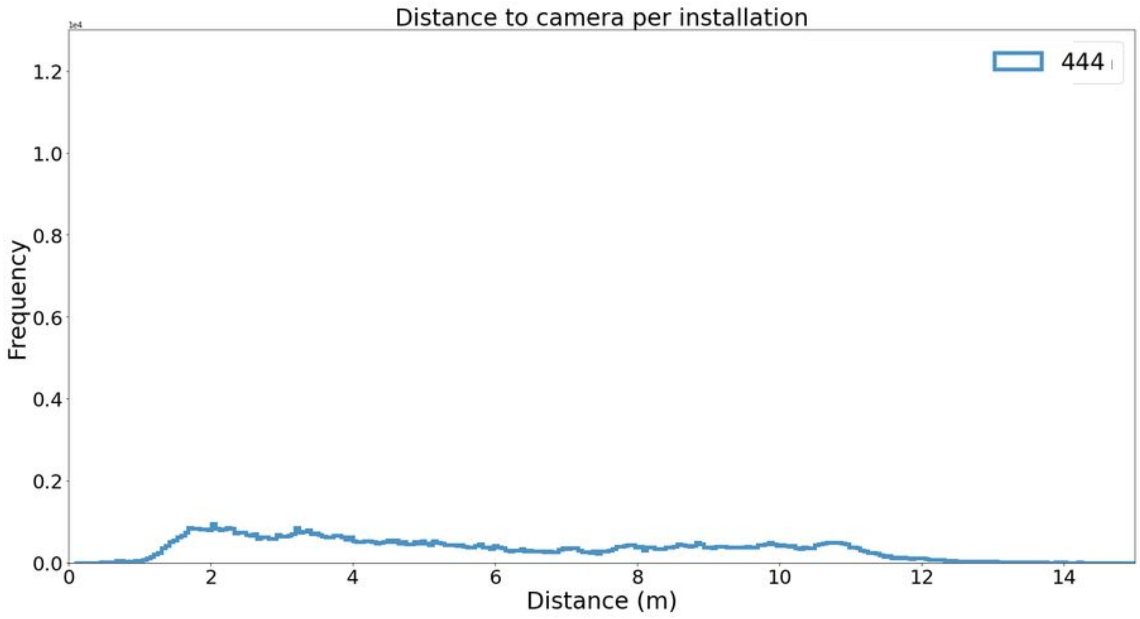
Presence → **Position** → **Attention**

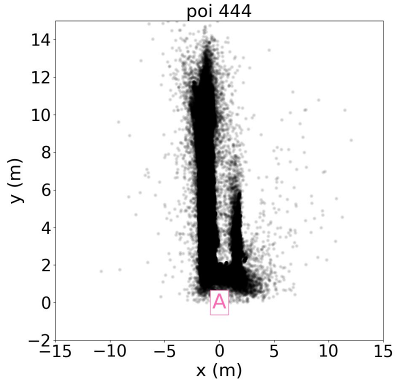
Data Exploration - Store Layouts



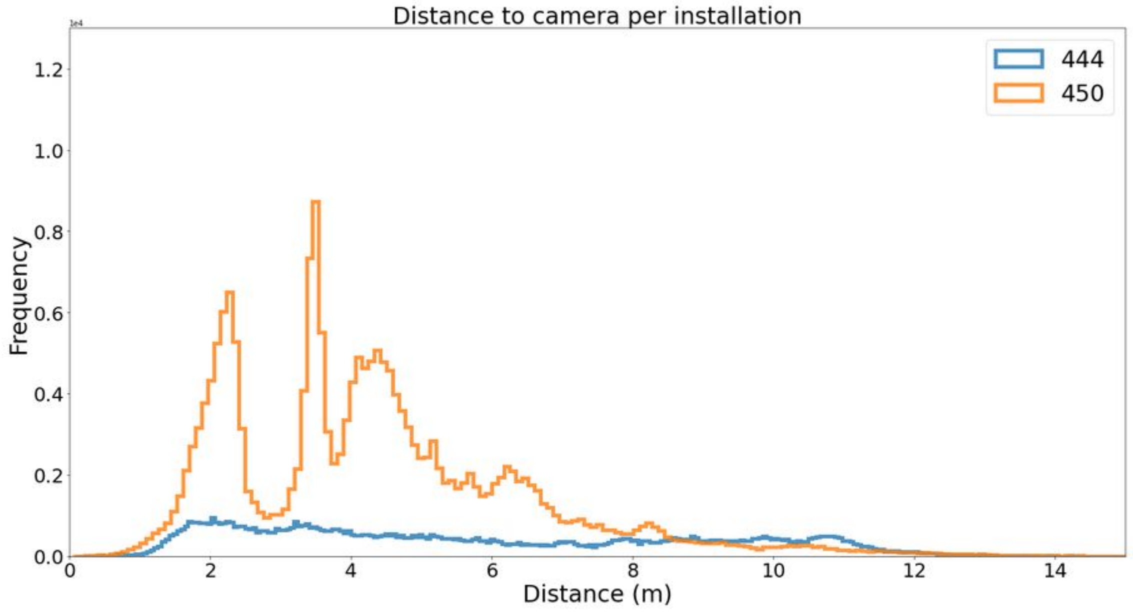
Start time:
2020-07-06 13:15:00
End time:
2020-07-09 19:59:59
Δtime = 3 d + 7 h

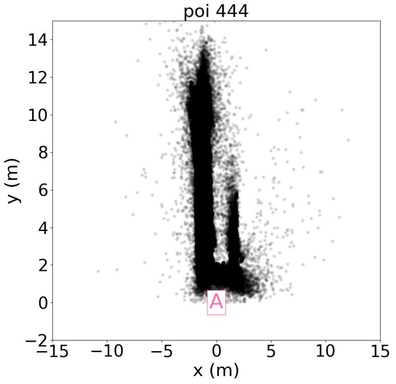
Data Exploration

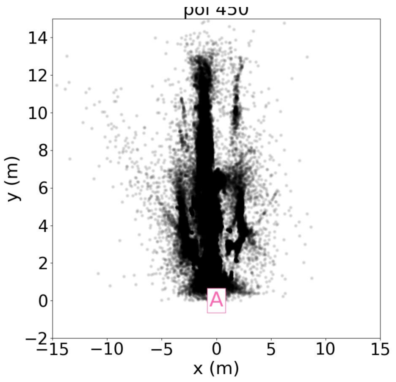




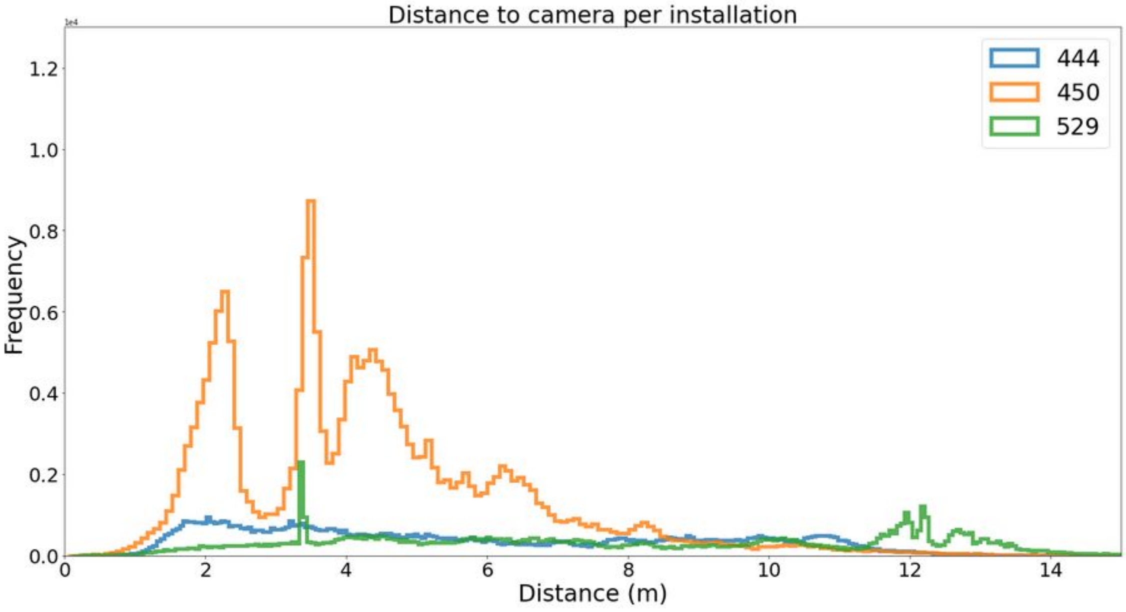
Data Exploration

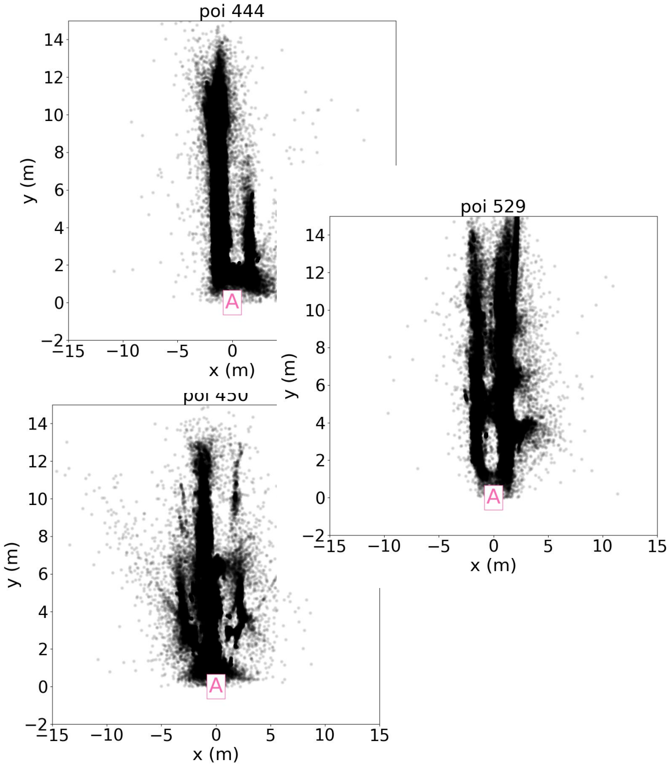




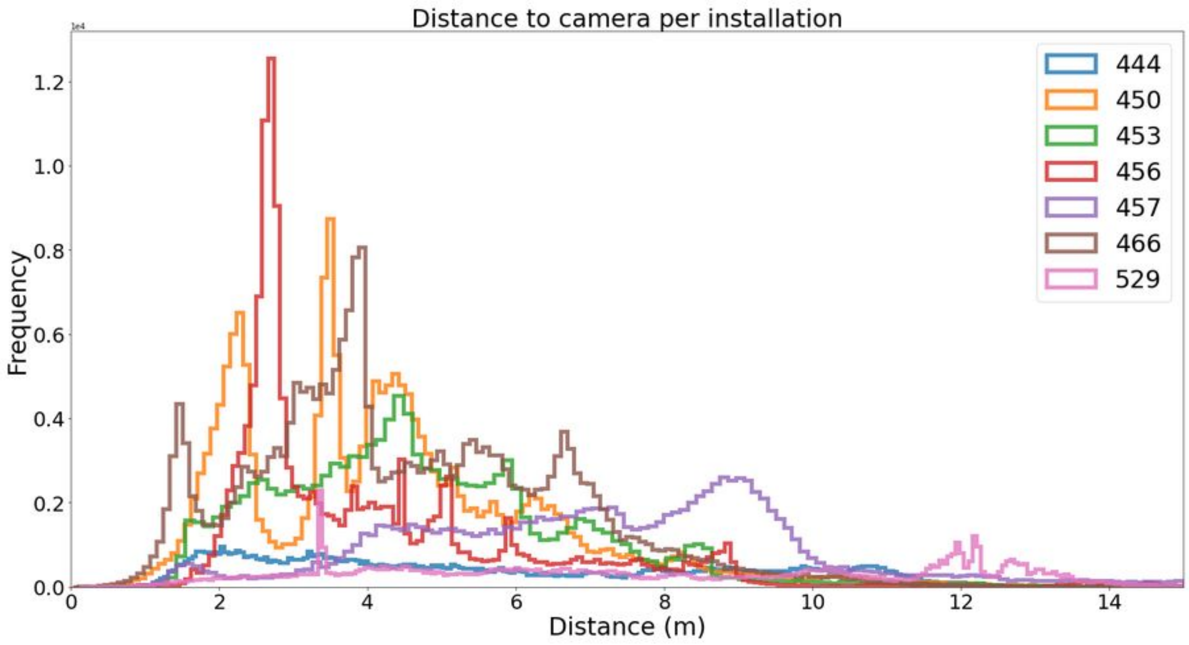


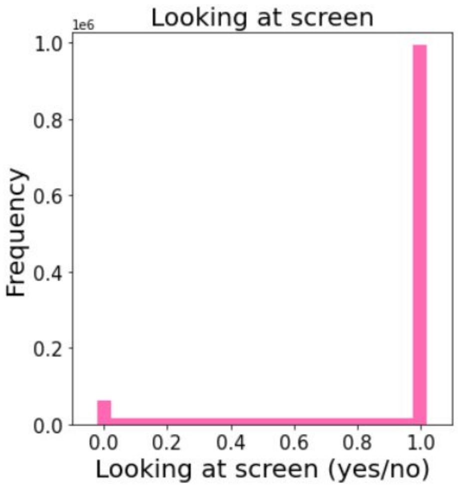
Data Exploration





Data Exploration





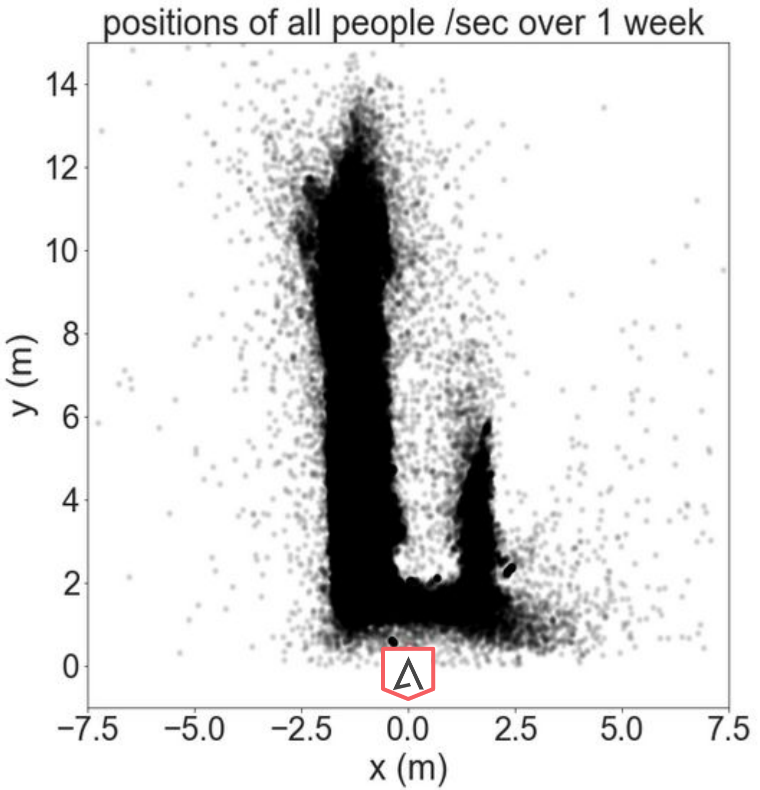
- N/A head position - 25% of data
- Looking at screen - 4% of 75%

Data Engineering

- **Data imbalance**
  - Predicting attention: only 4% of person_looking
  - Predicting presence: 1, 2, 5 seconds prediction windows

- **Engineering**
  - Outlier and error removal
  - Added features: exponential moving averages, velocity, cumulative time, distance to camera
  - Data format: ms -> sec created some duplicate timestamps needing adjusting
  - Missing data: linearly extrapolated
  - Age/Gender rolling estimate

Customer Trajectory & Store Layout



Predict Presence

For 85 people out of 100 we can accurately predict if person will leave within the 5 sec

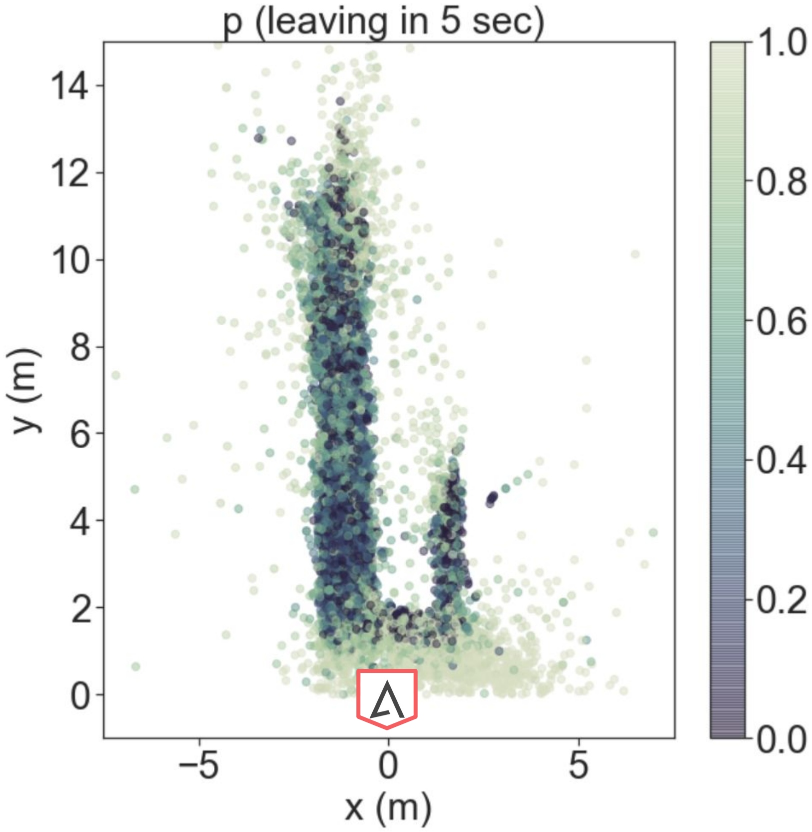


Illustration of one installation with 5-second predictive time window

Predict Presence - Modeling Results

Predict presence (1/0) in n seconds							
Predictive window		**5 sec**		**2 sec**		**1sec**	
Approach	**KPI**	**Precision**	**Recall**	**Precision**	**Recall**	**Precision**	**Recall**
keep every 2nd row	0 - absent	83	64	75	52	63	32
	1 - present	84	87	92	96	94	98
keep every 5th row	0 - absent	87	78	75	54	63	34
	1 - present	75	71	87	93	91	97
data balance 50/50	0 - absent	85	84	85	83	83	86
	1 - present	84	85	84	85	85	83

Precision: For 100 customers that we predict will be present in 5 sec, 16 have already left.

Recall: For 100 customers that will be present in 5 sec, we only predicted 85, i.e. we missed 15.

Predict Presence - Modeling Results + Data Imbalance

Predict presence (1/0) in n seconds							
Predictive window		**5 sec**		**2 sec**		**1sec**	
Approach	**KPI**	**Precision**	**Recall**	**Precision**	**Recall**	**Precision**	**Recall**
keep every 2nd row	0 - absent	83	64	75	52	63	32
	1 - present	84	87	92	96	94	98
keep every 5th row	0 - absent	87	78	75	54	63	34
	1 - present	75	71	87	93	91	97
data balance 50/50	0 - absent	85	84	85	83	83	86
	1 - present	84	85	84	85	85	83

Data Imbalance per Approach				
Predictive window		**5 sec**	**2 sec**	**1sec**
0s in the dataset	No rebalacing	22%	9%	5%
	Keep every 5th row	36%	15%	7%
	Keep every 2nd row	58%	24%	12%
	50/50 split	50%	50%	50%

Predict Position

Mean position prediction accuracy

In 1 second: **17cm**

In 2 seconds: **68cm**

In 5 seconds: **134cm**

Up to 40% better accuracy
(depending on layout)

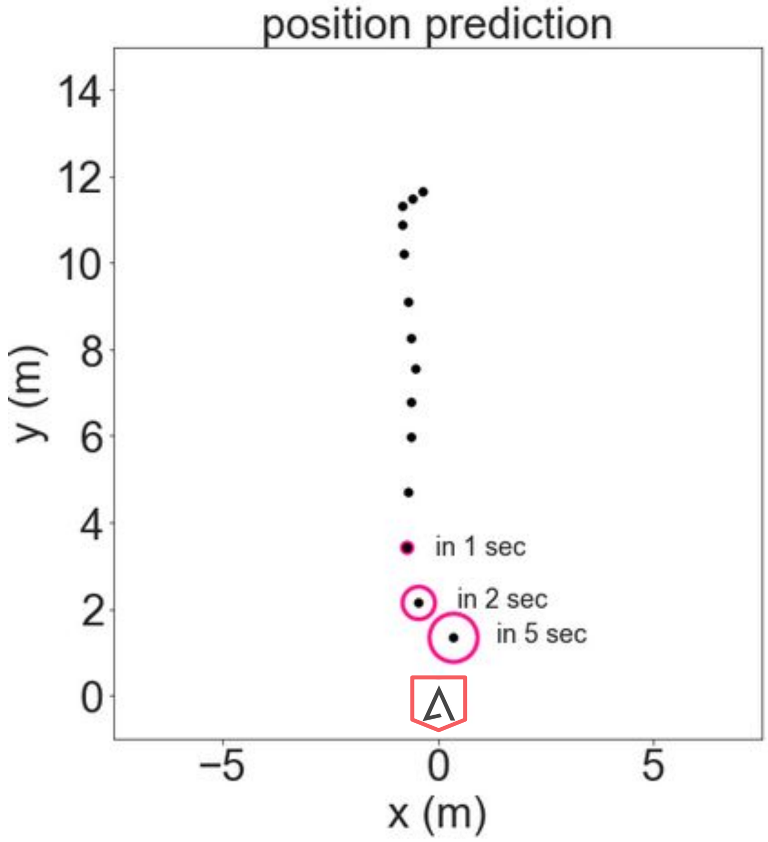


Illustration of one installation

Predict Position - Modeling Results

(x,y) position prediction			
Timeframe	**1 second**	**2 seconds**	**5 seconds**
POI \ KPI	**RMSE**	**RMSE**	**RMSE**
450	0.12	0.44	0.85
466	0.14	0.51	0.92
444	0.15	0.56	1.10
456	0.16	0.54	1.15
453	0.16	0.78	1.63
457	0.20	0.81	1.67
529	0.23	1.14	2.07
Mean	**0.17**	**0.68**	**1.34**

Physical store layout impacts the prediction error

Predict Position - Modeling Results

(x,y) position prediction						
Timeframe	1 second		2 seconds		5 seconds	
POI \ KPI	RMSE	d(ŷ - y)	RMSE	d(ŷ - y)	RMSE	d(ŷ - y)
529	0.23	0.12	1.14	0.62	2.07	1.61
466	0.14	0.05	0.51	0.28	0.92	0.74
457	0.20	0.11	0.81	0.58	1.67	1.57
456	0.16	0.06	0.54	0.35	1.15	0.96
453	0.16	0.08	0.78	0.50	1.63	1.34
450	0.12	0.05	0.44	0.26	0.85	0.67
444	0.15	0.08	0.56	0.37	1.10	0.98
Mean	**0.17**	**0.08**	**0.68**	**0.42**	**1.34**	**1.12**

Delta (ŷ - y): distance between actual position and prediction, in meter

Predict Attention

For 7 people out of 10 we can accurately predict if person will look at the screen within the next few seconds

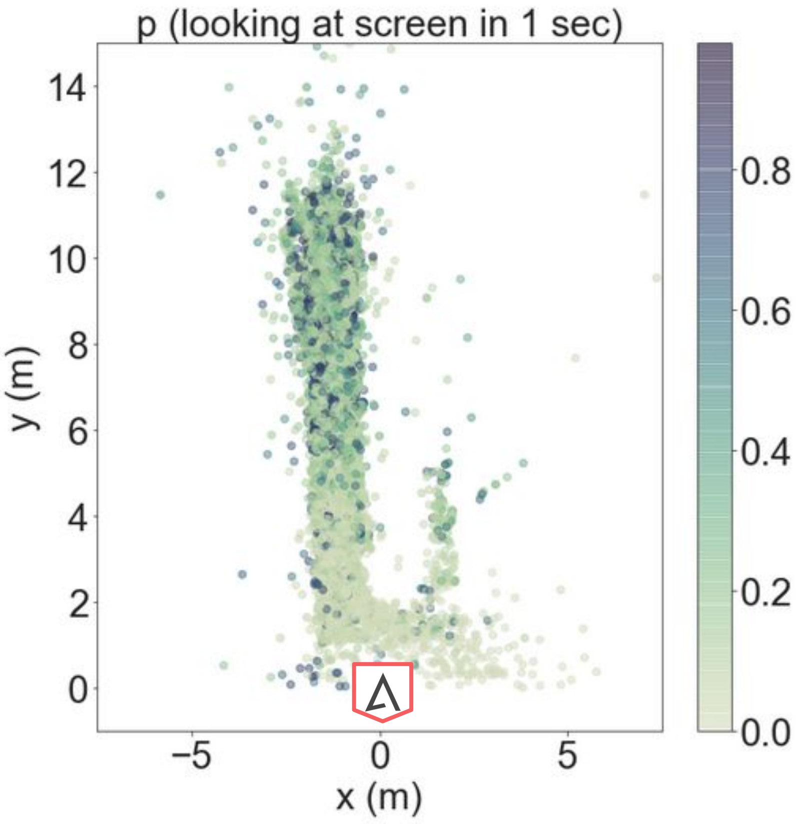


Illustration of one installation with 1-second predictive time window

Predict Attention - Modeling Results

Predict Attention to Screen in n seconds							
Predictive window		5 sec		2 sec		1 sec	
KPI		Precision	Recall	Precision	Recall	Precision	Recall
Not Looking	0	0.9	0.94	0.84	0.93	0.72	0.83
Looking	1	0.68	0.42	0.68	0.39	0.66	0.46

Precision: For 100 customers that we predict will be looking at the screen in 5 sec, 32 actually will not.

Recall: For 100 customers that will be looking at the screen in 5 sec, we predicted only 42, i.e. we missed 58 by tagging them as not looking.

Potential Next Steps

Implementation & fine-tuning of **Machine Learning** models

Further exploration of **Deep Learning** models

Add **product position information** for improved targeting

Thank you!

Sarah Kurmulis

linkedin.com/in/sarahkurmulis

Olena Levchun

linkedin.com/in/olenalevchun

Matthieu Bornet

linkedin.com/in/matthieubornet