



OPENCLASSROOMS

## Projet 8

-

Réalisez un dashboard et  
assurez une veille technique



CentraleSupélec

Pierrick BERTHE

Formation Expert en Data Science  
*Openclassrooms – CentraleSupélec*

*août 2023 → juin 2024*



# Problématique

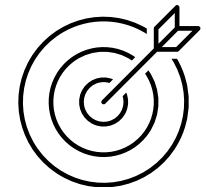
L'entreprise « Prêt à dépenser » propose des crédits à la consommation. Elle souhaite mettre en œuvre un «*scoring credit*» pour accorder ses crédits selon la probabilité qu'un client rembourse son crédit.



**=> Automatiser la prise de décision d'accord de prêt grâce à un algorithme de classification**

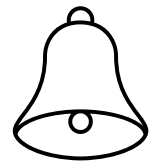
## Missions :

1. Construire un dashboard interactif basé sur le modèle de scoring déployé sur l'API.
2. Réaliser une veille des outils de Data Science et démontrer un POC.

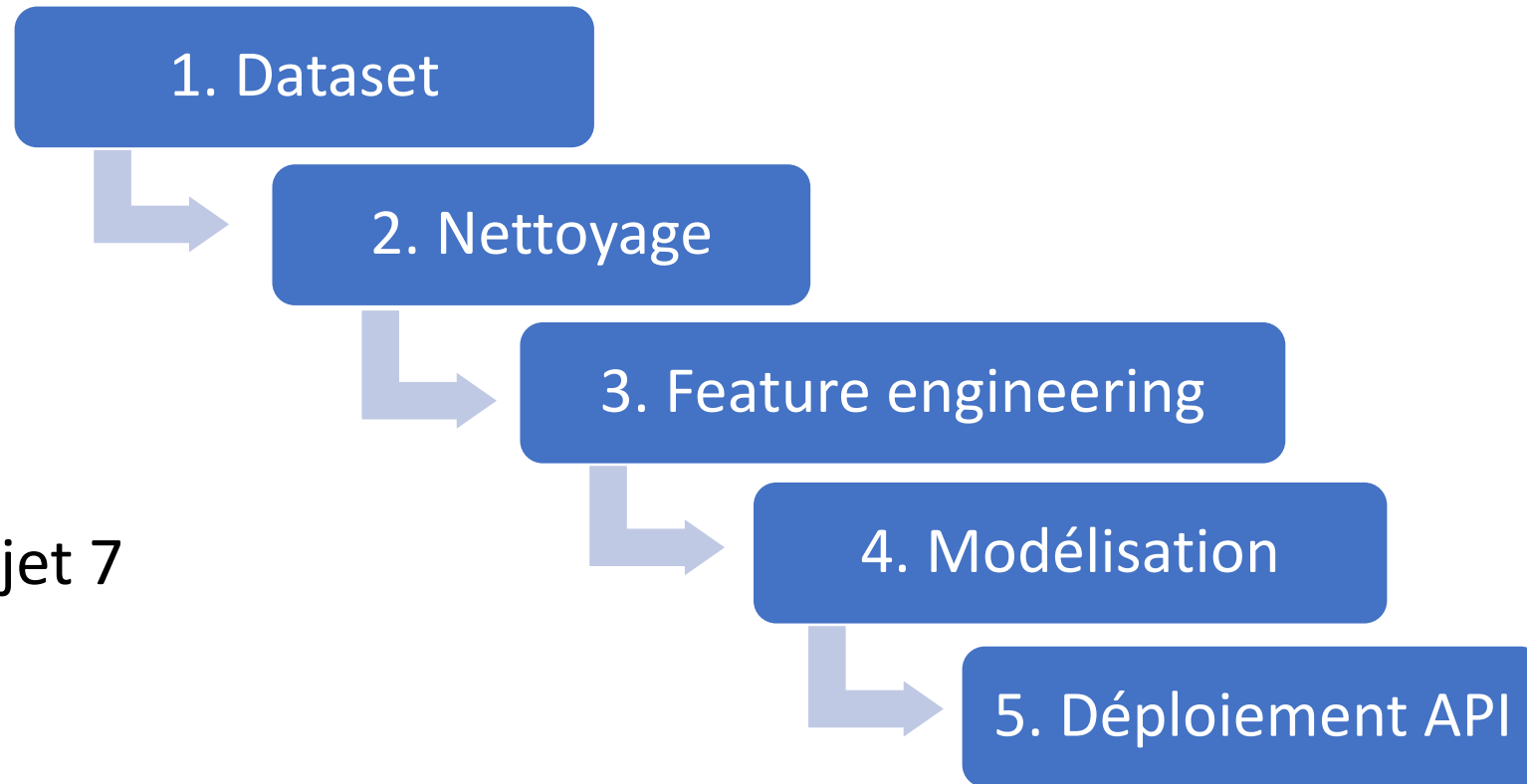




# Dashboard



Projet 7





# Dashboard



## Projet 8

### Réalisez un Dashboard

Formation Data Scientist de Pierrick BERTHE

Prédiction client DataViz features Analyse bivariée Feature importance Globale

### Prédiction client

Choisir un id de client :

454704

### Client sélectionné : 454704

Choisir les features à afficher :

Choose an option

	NAME_CONTRACT_TYPE	CODE_GENDER	FLAG_OWN_CAR	FLAG_OWN_REALTY	CNT_CHILDREN	AMT.
0	0	1	0	1	0	

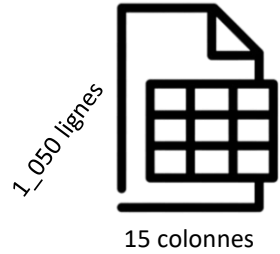
Calculer la prédiction

Le prêt n'est pas accordé.

<https://projet8ocrdatascientist-pierrick-berthe.streamlit.app>



# Veille scientifique / POC



**Flipkart**



Computer vision

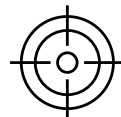


**VS**

**ultralytics**  
YOLOv8

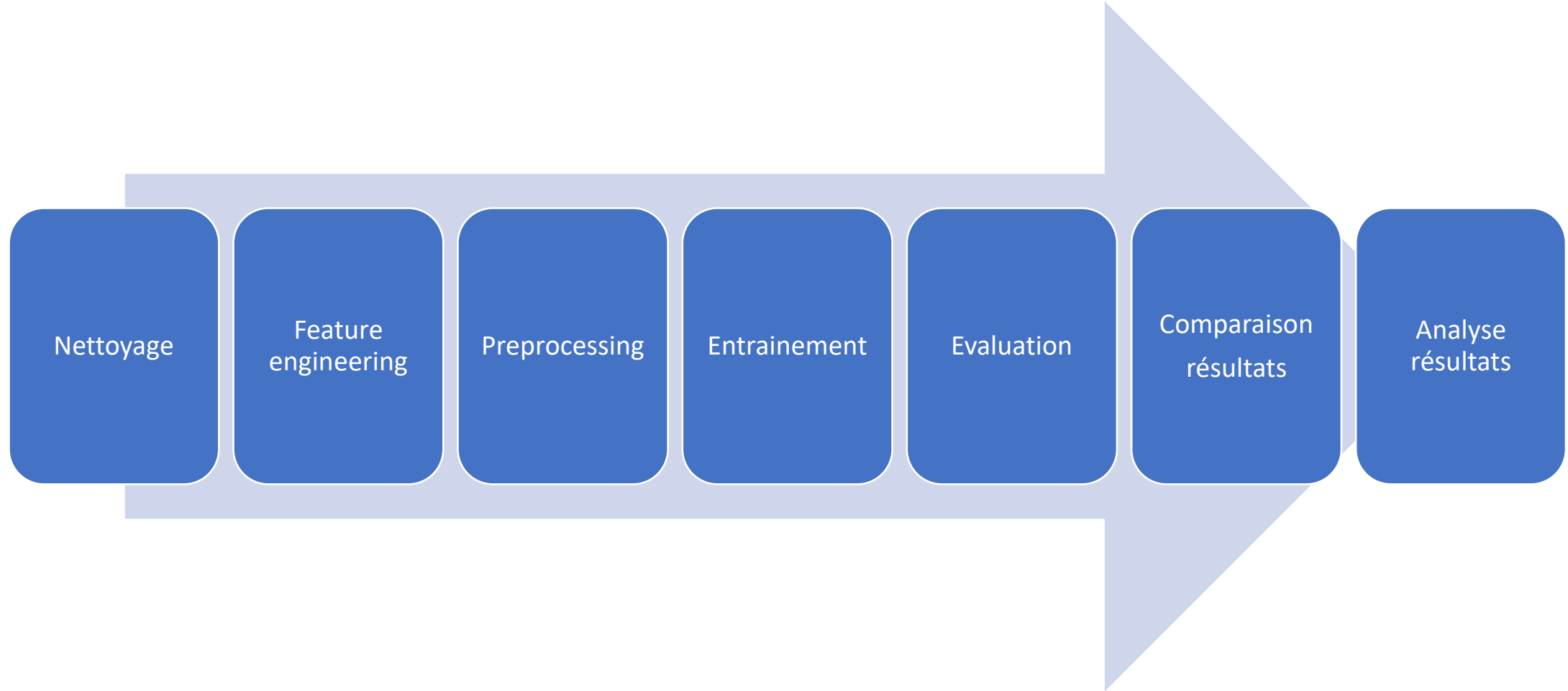
**2014**

JANUARY  
**2023**





# Veille scientifique / POC





# Veille scientifique / POC



Nettoyage

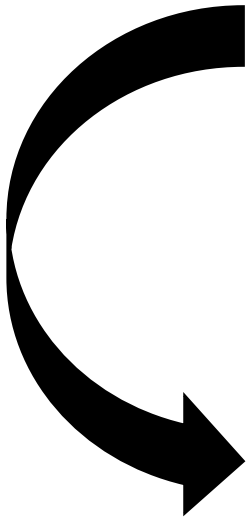
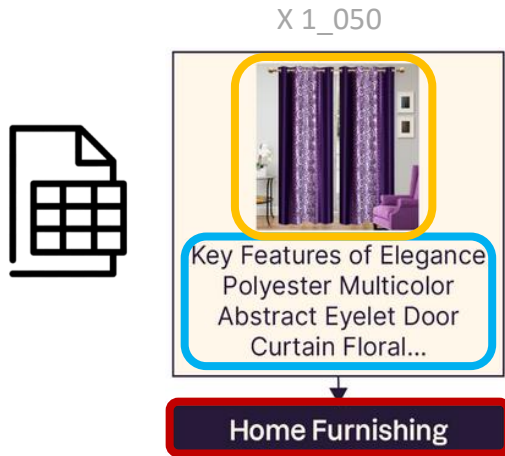
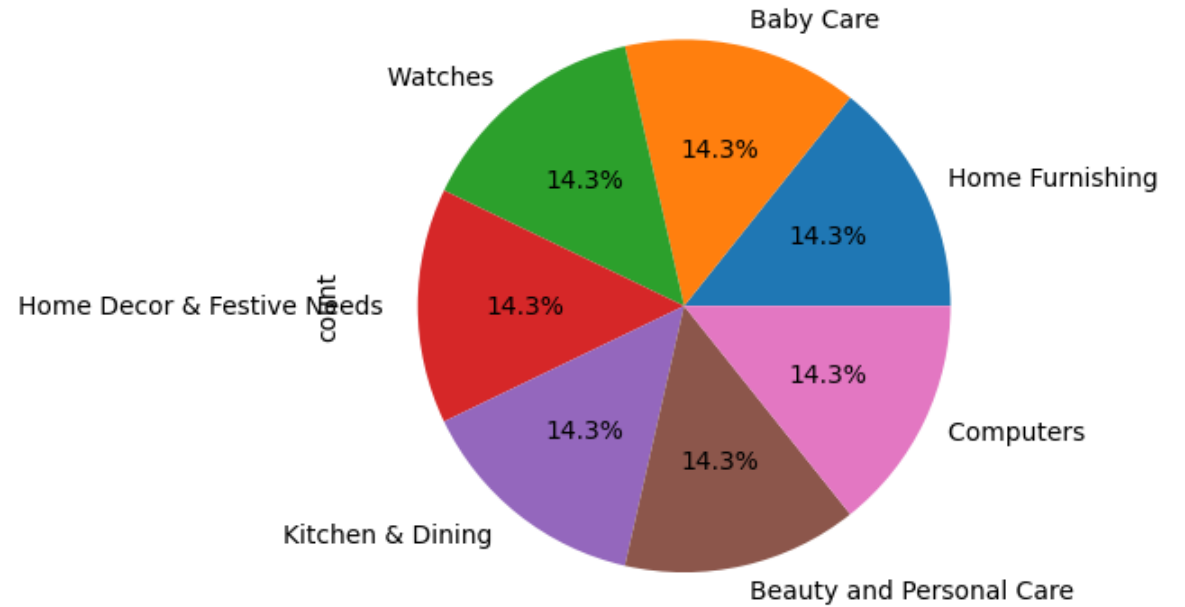


	image	description	product_name	label
0	55b85ea15a1536d46b7190ad6fff8ce7.jpg	Key Features of Elegance Polyester Multicolor ...	Home Furnishing	4
1	7b72c92c2f6c40268628ec5f14c6d590.jpg	Specifications ... on Bath Towel (...)	Baby Care	0
2	64d5d4a258243731dc7bbb1eef49ad74.jpg	Key Features ... Face Towe...	Baby Care	0



7 catégories



# Veille scientifique / POC

Feature  
engineering



X 1\_050

	image	product_name	label
0	55b85ea15a1536d46b7190ad6fff8ce7.jpg	Home Furnishing	4
1	7b72c92c2f6c40268628ec5f14c6d590.jpg	Baby Care	0
2	64d5d4a258243731dc7bbb1eef49ad74.jpg	Baby Care	0

1. Séparation

Features / Target

2. Train test split

70% train  
15% validation  
15% test





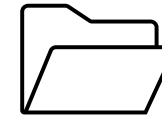
# Veille scientifique / POC



Target



VS



Features

- Redimensionnement 224 px
- Conversion tableau Numpy
- Certification dimension (h x l x c)
- Normalisation pixels
- RGB to BGR

- Redimensionnement
- Rognage en carré
- Normalisation pixels
- Conversion image en tenseur

```
-- train/  
| |-- Baby Care/  
| | |-- 10008.png  
| | |-- ...
```

```
-- Home Furnishing/  
| |-- 1000.png  
| |-- ...
```

```
-- test/  
| |-- ...
```

```
-- val/  
| |-- ...
```



# Veille scientifique / POC



Entrainement

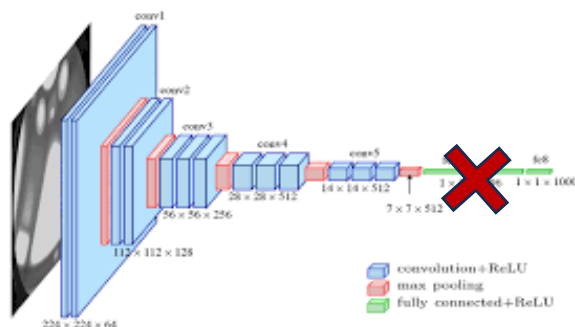


VS



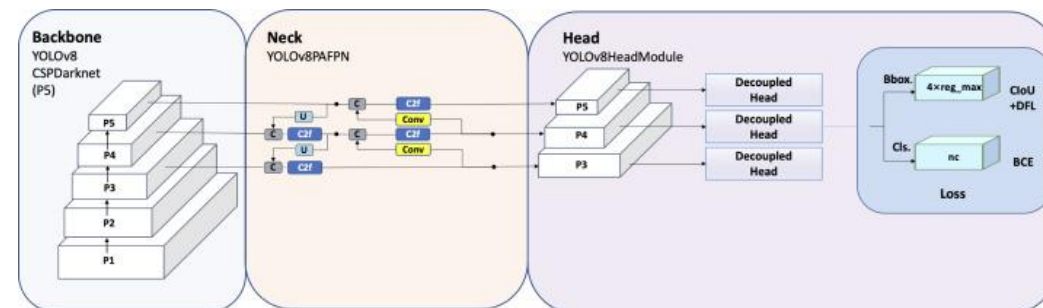
nano

Transfer learning



**22 couches :**

- Modèle VGG16 : 5 blocs de convolution/Max pooling (18 couches)
- GlobalAveragePooling 2D
- Dense (256) avec fonction ReLU
- Dropout (0,5)
- Dense (7) avec fonction Softmax



**225 couches dans 3 parties :**

- Backbone : feature extraction
- Neck : combine les features extraites
- Head : prédiction objet et classe

*Terven, Córdoba-Esparza, et  
Romero-González 2023*



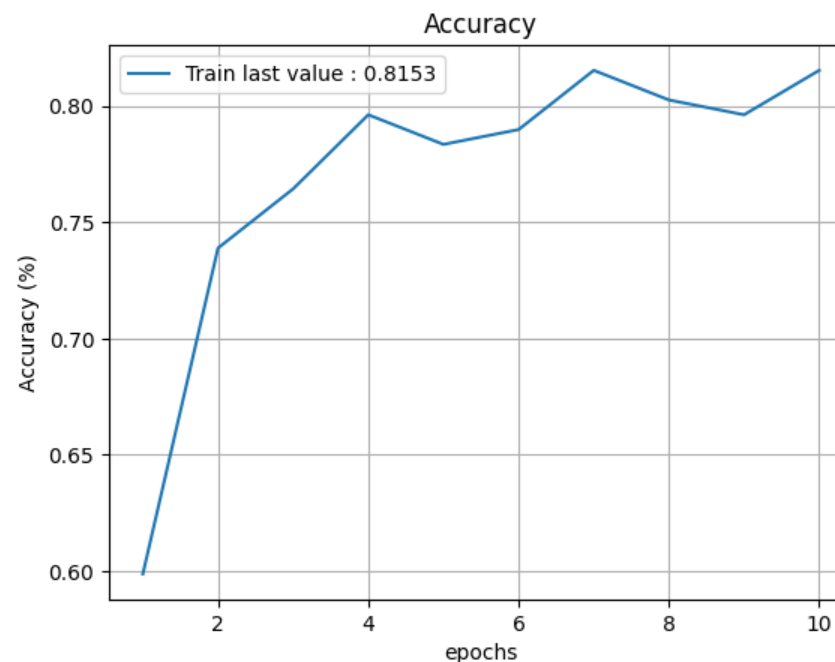
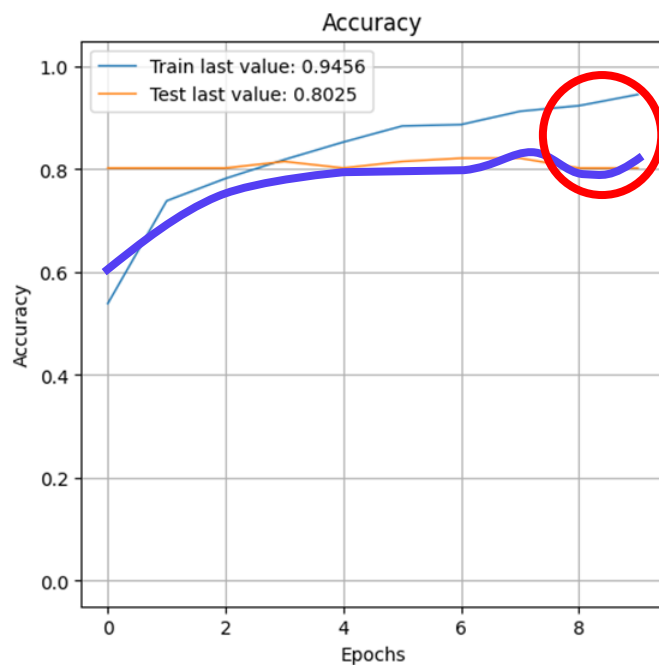
# Veille scientifique / POC



Evaluation



VS



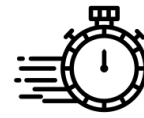
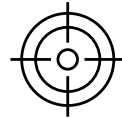


# Veille scientifique / POC

Comparaison  
résultats



vs



	Test Accuracy	Temps (sec)	Epochs
vgg16	0.81	582.69	10
yolo_v8	0.83	436.46	10

YOLOv8 est **plus précis** (+2%) et **plus rapide** à entrainer (-25%)



# Veille scientifique / POC

Analyse  
résultats

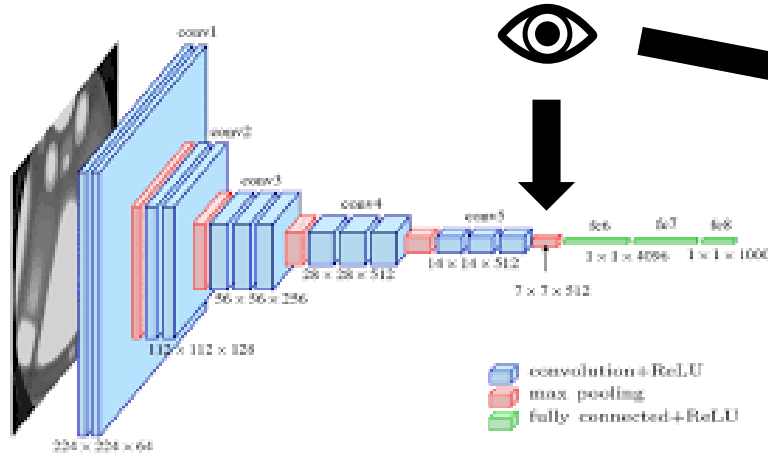


Vgg16

Grad-CAM

*Selvaraju et al. 2020*

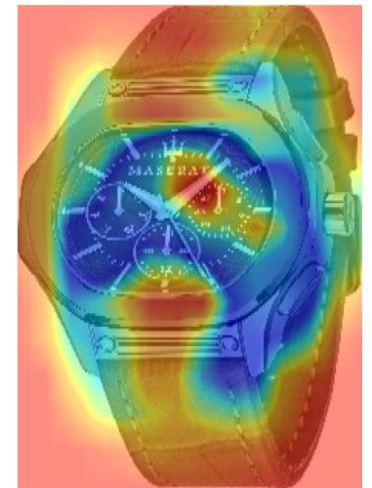
Interprétabilité  
Locale



Originale



Grad-CAM





# Veille scientifique / POC



Analyse  
résultats

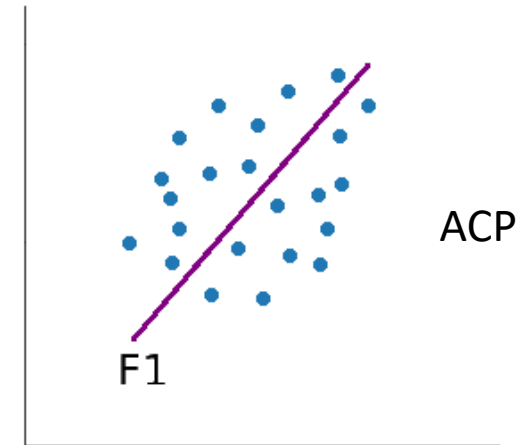
Eigen-CAM

*Muhammad et  
Yeasin 2020*

ultralytics  
YOLOv8

PyTorch

Interprétabilité  
Locale



ACP

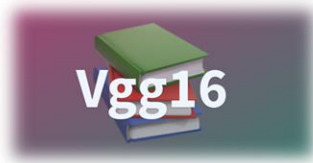
F1



# Veille scientifique / POC



Analyse  
résultats



VS

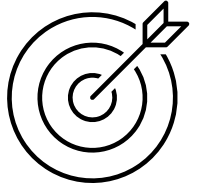


Interprétabilité  
Globale





# Conclusion



## Missions :

1. Construire un dashboard interactif basé sur le modèle de scoring déployé sur l'API. ✓



2. Réaliser une veille des outils de Data Science et démontrer un POC. ✓





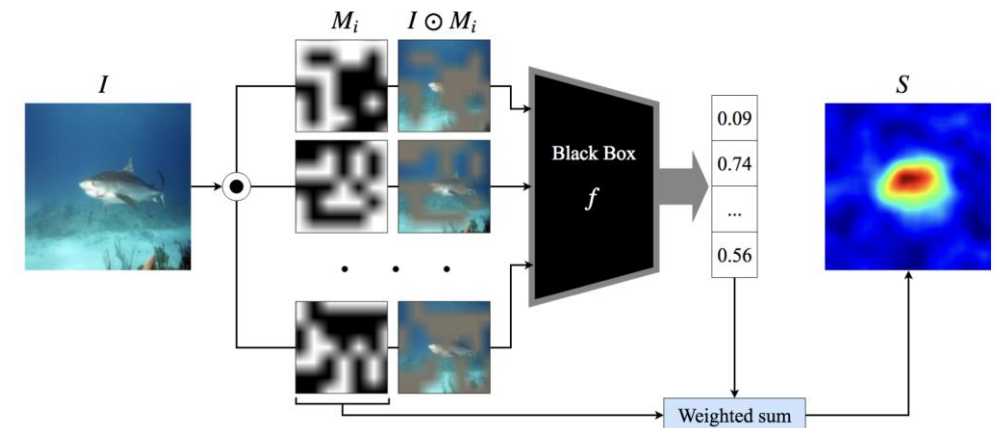


# Conclusion

## Limites et améliorations :

- Dashboard avec possibilité de modification des données
- Utilisation de format plus performant de YOLOv8 (s, m, l et x)
- Optimisation hyperparamètres via GridSearchCV
- Randomized Input Sampling for Explanation (RISE) →

*Petsiuk, Das, et Saenko 2018*



Model	size (pixels)	mAP <sup>val</sup> <sub>50-95</sub>
<a href="#">YOLOv8n</a>	640	37.3
<a href="#">YOLOv8s</a>	640	44.9
<a href="#">YOLOv8m</a>	640	50.2
<a href="#">YOLOv8l</a>	640	52.9
<a href="#">YOLOv8x</a>	640	53.9



OPENCLASSROOMS

Merci pour votre attention



CentraleSupélec

Pierrick BERTHE

Formation Expert en Data Science  
*Openclassrooms – CentraleSupélec*

*août 2023 → juin 2024*