

# Classifiez automatiquement des biens de consommation

## Projet 6 du parcours Data Scientist

Dernière Màj  
22 mai 2023

Raquel Sanchez Pellicer



## Contexte



## Étude faisabilité

NLP  
Image Processing



## Classification supervisée



## Test collecte images, API



## Conclusions

# Contexte - objectifs

**place de marché** souhaite lancer une marketplace e-commerce.

Actuellement sur la **place de marché**, des vendeurs proposent des articles à des acheteurs en postant une photo et une description.

Pour l'instant, l'attribution de la catégorie d'un article est effectuée manuellement par les vendeurs.



**Mission :** Réaliser un **étude de faisabilité** d'un moteur de classification des articles en différentes catégories, avec un niveau de précision suffisant.

## Objectifs :

- Fiabiliser l'attribution de la catégorie d'un article.
- rendre l'**expérience** utilisateur des vendeurs (faciliter la mise en ligne de nouveaux articles) et des acheteurs (faciliter la recherche de produits) la **plus fluide** possible.
- Automatiser la tâche dans l'optique d'un passage à l'échelle.

# Contexte - données disponibles

place de marché



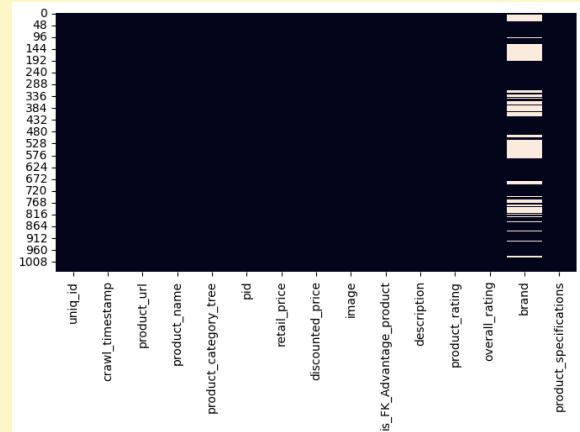
L'absence de propriété intellectuelle sur le texte ou les images a été vérifiée.

## Données textuelles : descriptions des produits

- 1050 lignes
- 15 colonnes
- Peu de données manquantes
- Arbre de classification à 7 niveaux



	Nombre catégories	Taux remplissage
Categ_N0	7	0.000000
Categ_N1	62	0.000000
Categ_N2	241	0.002857
Categ_N3	349	0.353333
Categ_N4	297	0.614286
Categ_N5	117	0.879048
Categ_N6	57	0.945714



description	product_name	product_specifications
Key Features of Elegance Polyester Multicolor ...	Elegance Polyester Multicolor Abstract Eyelet ...	{"product_specification":>[{"key":>"Brand", "v...
Specifications of Sathiya's Cotton Bath Towel (...	Sathiya's Cotton Bath Towel	{"product_specification":>[{"key":>"Machine Wa...
Key Features of Eurospa Cotton Terry Face Towe...	Eurospa Cotton Terry Face Towel Set	{"product_specification":>[{"key":>"Material", ...
Key Features of SANTOSH ROYAL FASHION Cotton P...	SANTOSH ROYAL FASHION Cotton Printed King size...	{"product_specification":>[{"key":>"Brand", "v...
Key Features of Jaipur Print Cotton Floral Kin...	Jaipur Print Cotton Floral King sized Double B...	{"product_specification":>[{"key":>"Machine Wa...

## Images

- 1050 images





Contexte



**Étude faisabilité**  
NLP  
Image Processing



Classification supervisée



Test collecte images, API



Conclusions

# Approche – Étude de faisabilité



## Pré-traitement

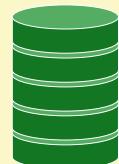
### Données texte

(description, nom du produit et spécifications du produit)



**9 opérations** (mise en minuscule, tokenization, suppression des mots de liaison...)

### Données Image



**5 opérations** (correction exposition, conversion en niveau de gris...)

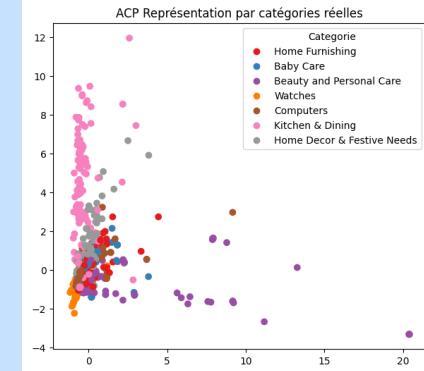
## Extraction des features

Comptage simple des mots  
Words Embeddings  
Sentence Embeddings

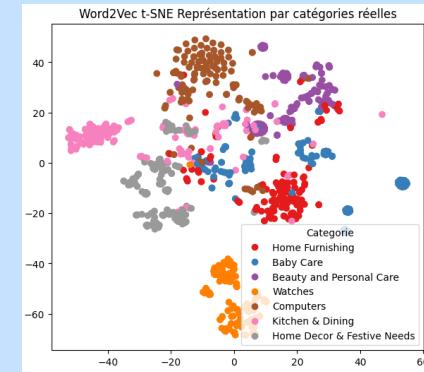
Bag-of-images  
CNN

## Réduction dimension

### PCA

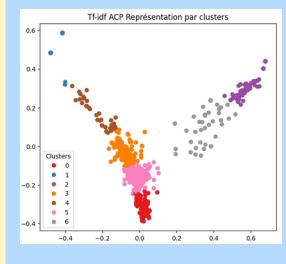


### t-SNE



## Classification

- Classification non supervisée
- K-Means, k=7
- Optimisation hyperparamètres d'initiation



## Évaluation

- Score ARI
- Temps de calcul

# Étude faisabilité : NLP - Pré-traitement



Action	code	Explication
Mise en minuscule	.lower().strip()	convertit tout le texte en minuscules et supprime les espaces vides
Tokenization	word_tokenize() wordpunct_tokenize()	La tokenisation est le processus qui consiste à délimiter et éventuellement à classer les sections d'une chaîne de caractères d'entrée. Les tokens ainsi obtenus sont ensuite transmis à une autre forme de traitement.
Suppression des mots de liaison	[w for w in list_words if not w in all_stop_words]	(stop_words en anglais) ce sont des mots tellement communs qu'il est inutile de les indexer ou de les utiliser dans une recherche. Une liste de 1780 mots de liaison a été créée par combinaison de 3 listes. Le code ne garde dans la liste de mots que ceux n'apparaissant pas dans la liste des mots de liaison.
Suppression de la ponctuation et les caractères non alphabétiques	[w for w in list_words if w.isalpha()]	Ne garde un token que si tous les caractères sont des lettres de l'alphabet (a-z)
Filtre en fonction de la fréquence des mots sur le corpus	[w for w in list_words if w not in unique_words] [w for w in list_words if w not in mostfrequent_words]	Après le comptage des tokens deux listes ont été créées : une liste avec les tokens uniques et l'autre avec les tokens les plus fréquents jugés non pertinentes. Le code ne garde dans la liste de mots que ceux n'apparaissant pas dans ces listes.
Suppression des mots avec moins de N caractères	[w for w in list_words if len(w) >= min_len_word]	Une longueur minimale a été sélectionnée pour les tokens à conserver (dans ce cas, 3). Le code ne garde dans la liste de mots que ceux avec une longeur égale ou supérieure à la longueur minimale.
Stem et Lem	[PorterStemmer().stem(w) for w in list_words] [WordNetLemmatizer().lemmatize(w) for w in list_words]	Le <b>Stemming</b> (ou Racinisation) est un procédé de transformation des flexions en leur radical ou racine. La racine d'un mot correspond à la partie du mot restante une fois que l'on a supprimé son (ses) préfixe(s) et suffixe(s), à savoir son radical. La <b>lemmatisation</b> désigne un traitement lexical apporté à un texte en vue de son classement dans un index ou de son analyse. Ce traitement consiste à appliquer aux occurrences des lexèmes sujets à flexion un codage renvoyant à leur entrée lexicale commune, que l'on désigne sous le terme de lemme.
Mots communes à toutes les catégories	[w for w in text if w not in list_doublons]	Une liste a été créée avec les tokens présents dans toutes les catégories. Le code ne garde dans la liste de mots que ceux n'apparaissant pas dans cette liste.
Mots présents sur plusieurs catégories	[w for w in text if w not in list_mots_plusieurs_categ]	Après un premier traitement du texte, une liste a été créée avec les tokens présents dans plusieurs catégories et jugés non pertinents.

# Étude faisabilité : NLP – Extraction features



## Comptage simple des mots

### CountVectorizer

Convertit une collection de documents textuels en une matrice de comptages de tokens. Cette implémentation produit une représentation éparsée des décomptes.

### TfidfVectorizer

Équivalent à CountVectorizer suivi de TfidfTransformer . Tf signifie fréquence des termes, tandis que tf-idf signifie fréquence des termes multipliée par l'inverse de la fréquence dans les documents. Il s'agit d'un schéma de pondération des termes.

## Word2Vec

Les mots individuels sont convertis en une représentation numérique (vecteur). Les vecteurs tentent de capturer les différentes caractéristiques de ce mot par rapport à l'ensemble du texte.

Le vecteur qui caractérise un mot dépend de l'ensemble du texte considéré

## Word embeddings

### BERT

(Bidirectional Encoder Representations from Transformers)

BERT crée des représentations numériques en conditionnant conjointement le contexte gauche et droit. BERT pré-entraîné peut être affiné avec une seule couche de sortie supplémentaire. La représentation numérique ne dépend pas de l'ensemble du texte considéré

### USE

(Universal Sentence Encoder)

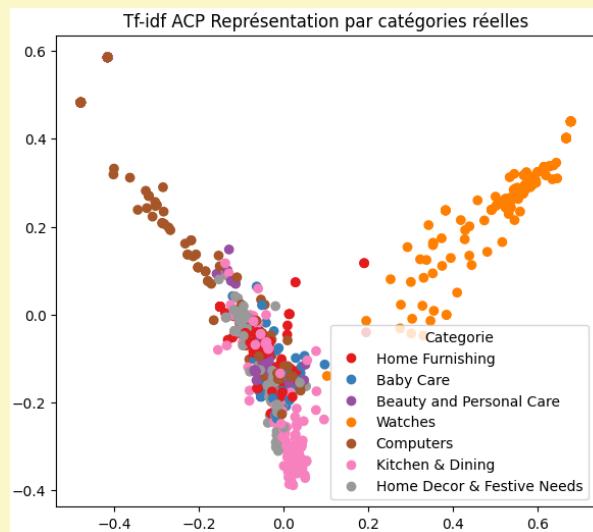
Le modèle calcule d'abord les word embeddings, puis détermine le vecteur de la phrase en calculant la somme par élément des vecteurs de mots. L'entrée de l'encodeur est une séquence de tokens et la sortie est un vecteur de 512 de taille.

# Étude faisabilité : NLP – Réduction des dimensions

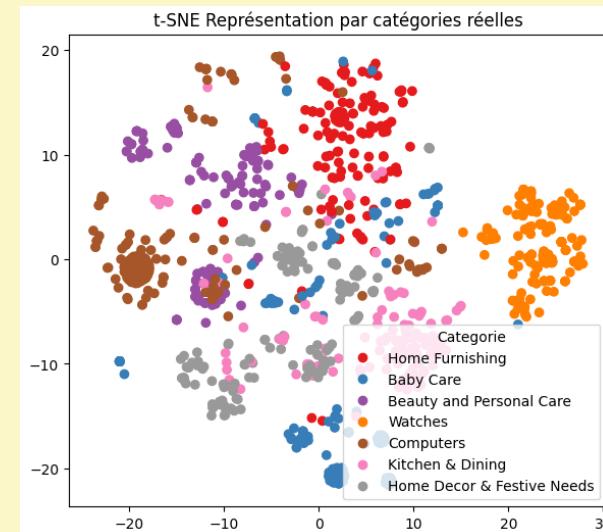


Tous les modèles génèrent des vecteurs à haute dimensionnalité, pour faciliter la comparaison des vecteurs et identifier les potentielles relations entre eux il est nécessaire de réduire la dimensionnalité.

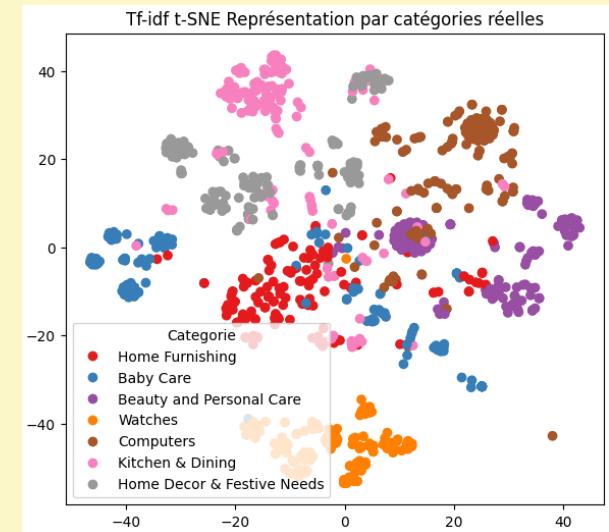
ACP, composantes expliquant 99 % de la variance



t-SNE



T-SNE avec initialisation ACP

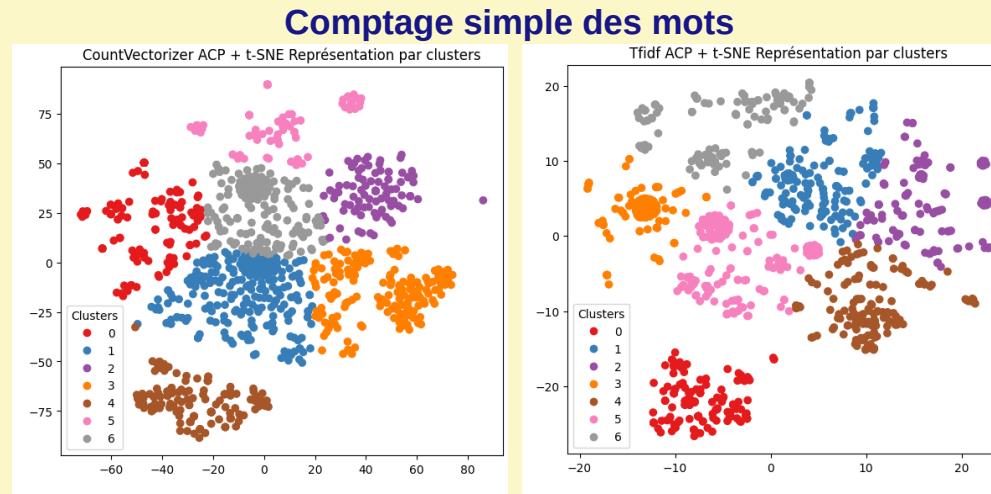


# Étude faisabilité : NLP – Classification

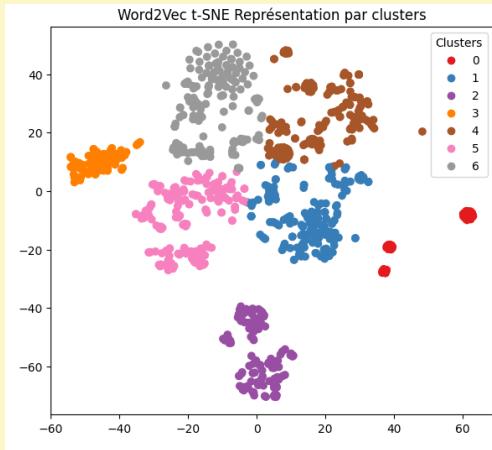


Analyse clusters, k-Means, k = 7

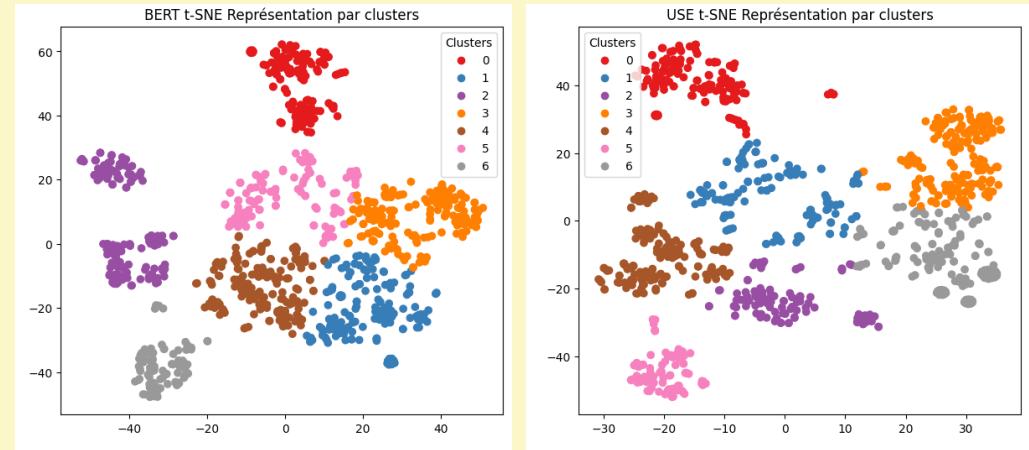
Optimisation des hyperparamètres d'initialisation



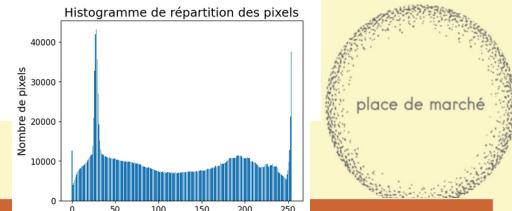
### Word embeddings



### Sentence embeddings

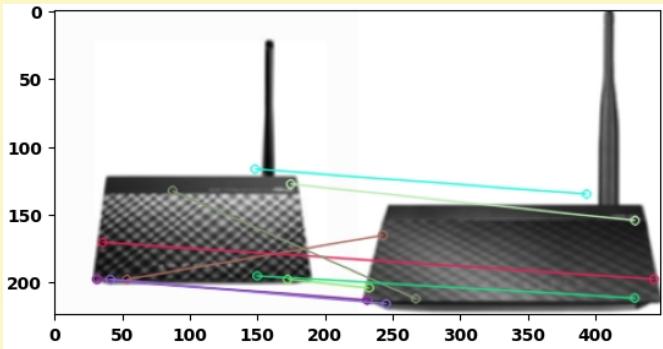


# Étude faisabilité : Images - Pré-traitement



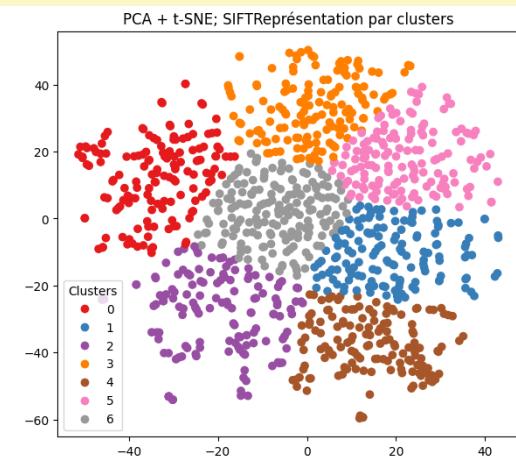
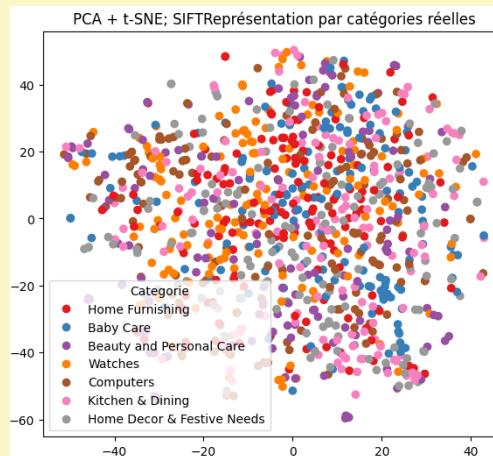
Action	Explication	Exemple
Correction de l'exposition	Étirement de l'histogramme	 <p>Histogramme de répartition des pixels pour l'image du mug. L'axe horizontal est 'Niveau de gris' (0 à 255) et l'axe vertical est 'Nombre de pixels' (0 à 35000). L'histogramme est étiré vers les extrêmes, montrant des pics à 0 et 255.</p>
Conversion en niveau de gris de l'image		 <p>Histogramme de répartition des pixels pour l'image convertie en niveaux de gris. L'axe horizontal est 'Niveau de gris' (0 à 1000) et l'axe vertical est 'Nombre de pixels' (0 à 16000). La distribution est plus large et centrée.</p>
Correction du contraste	Égalisation de l'histogramme	 <p>Histogramme de répartition des pixels pour l'image égalisée. L'axe horizontal est 'Niveau de gris' (0 à 255) et l'axe vertical est 'Nombre de pixels' (0 à 8000). L'histogramme est plus uniformément distribué.</p>
Réduction du bruit avec l'algorithme Non-local Means Denoising	Réalise une moyenne de la totalité des valeurs des pixels contenus dans l'image, pondérées en fonction de leur similarité avec le pixel cible.	 <p>Histogramme de répartition des pixels pour l'image traitée par Non-local Means Denoising. L'axe horizontal est 'Niveau de gris' (0 à 1000) et l'axe vertical est 'Nombre de pixels' (0 à 20000). La distribution est plus large et régulière.</p>
Redimensionnement en 224 * 224	Réduit et égale la taille de toutes les images disponibles	 <p>Histogramme de répartition des pixels pour l'image redimensionnée à 224x224. L'axe horizontal est 'Niveau de gris' (0 à 1000) et l'axe vertical est 'Nombre de pixels' (0 à 800). La distribution est très étroite et régulière.</p>

# Étude faisabilité : Bag-of-images, SIFT

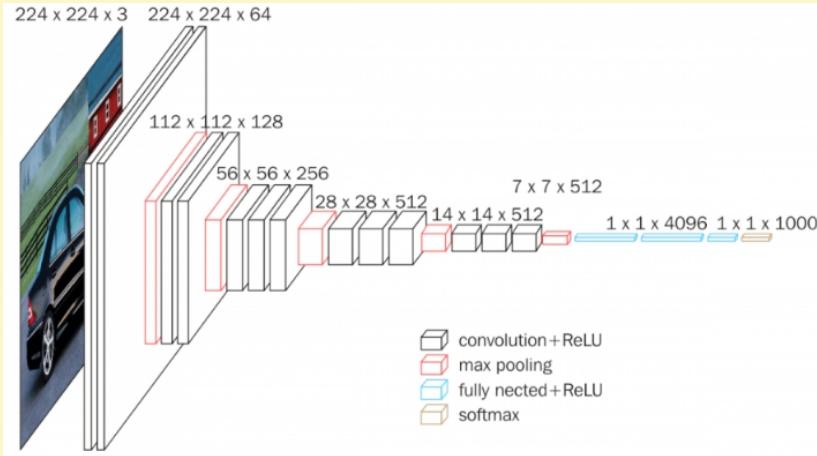


- création d'une liste de descripteurs par image
- création d'une liste de descripteurs pour l'ensemble des images qui sera utilisé pour créer les clusters de descripteurs

- Pour chaque image :
  - prédiction du numéro de cluster pour chaque descripteur
  - création d'un histogramme = combien des descripteurs de l'image sont associées à chaque numéro de cluster (bag-of-images)
- Réduction de dimensions
- Classification non supervisée

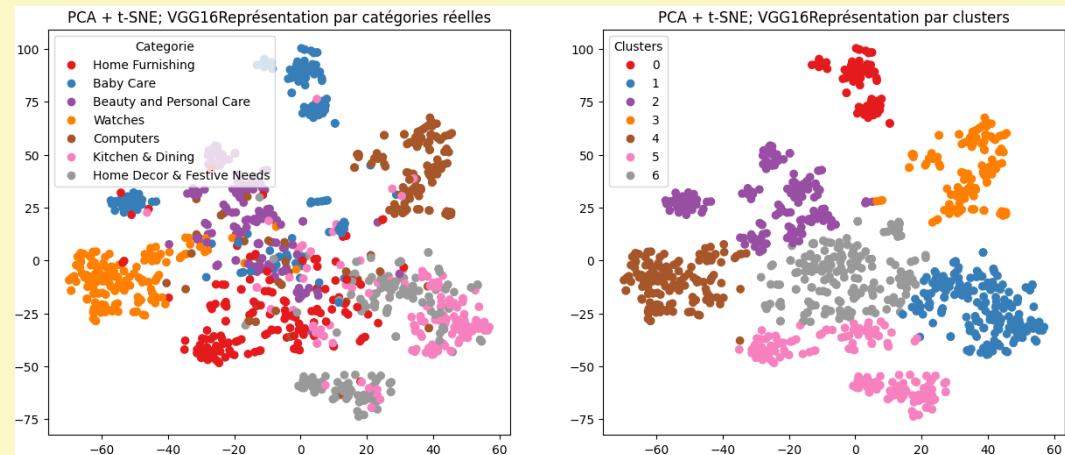


# Étude faisabilité : Images, CNN, VGG-16



- › Réduction de dimensions
- › Classification non supervisée, k-means

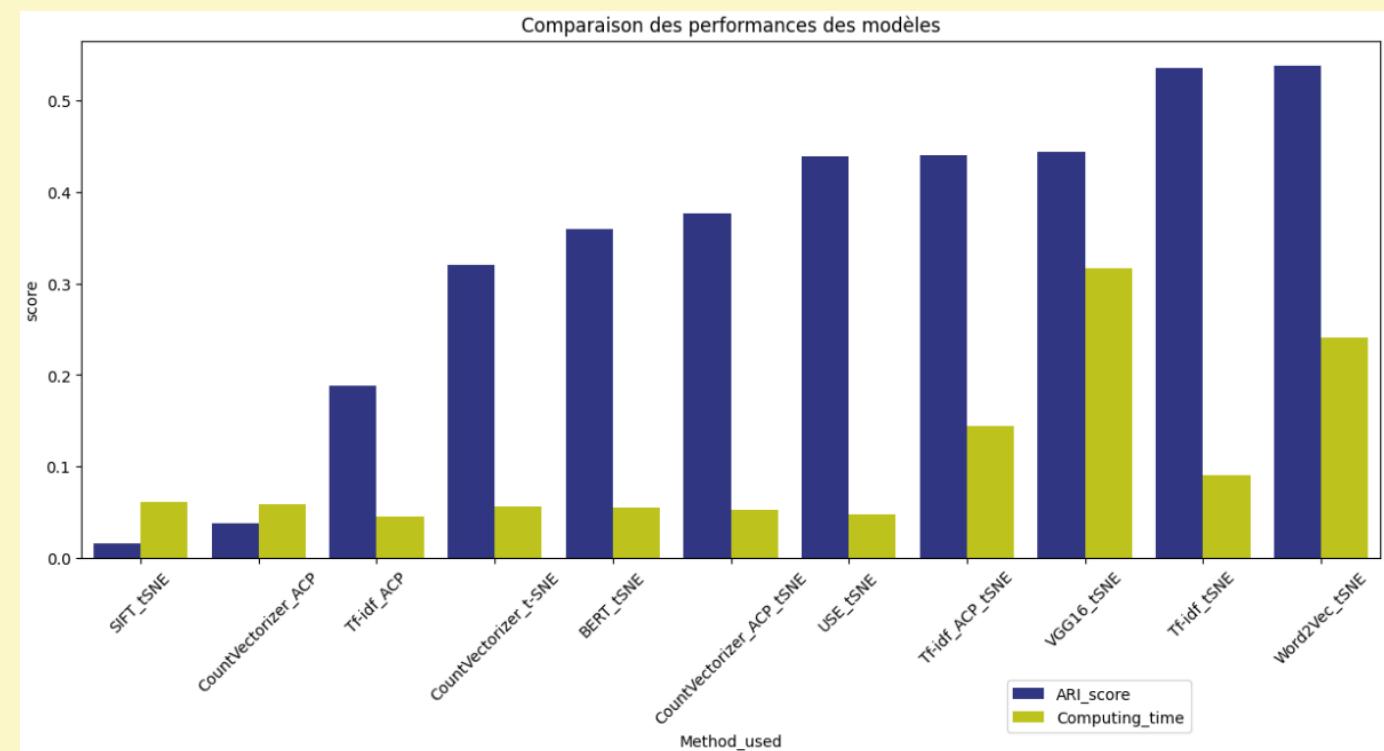
- › Réseau de neurones convolutional pré-entraîné, VGG-16
- › Préparation des images
- › Création d'un modèle pré-entraîné (suppression dernières couches, prédiction)



# Étude faisabilité - Comparaison résultats



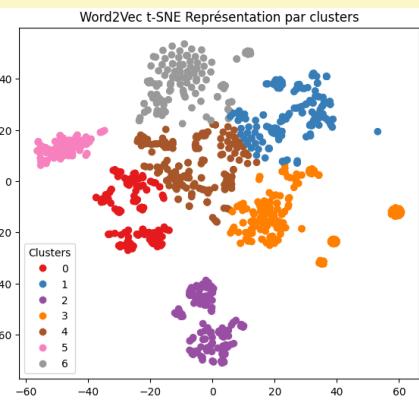
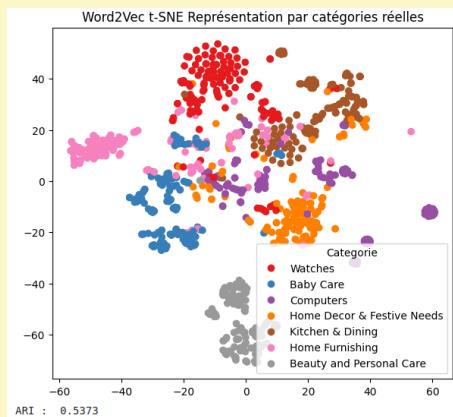
Method_used	ARI_score	Computing_time
SIFT_tSNE	0.0163	0.0613
CountVectorizer_ACP	0.0387	0.0588
Tf-idf_ACP	0.1882	0.0449
CountVectorizer_t-SNE	0.3199	0.0560
BERT_tSNE	0.3589	0.0558
CountVectorizer_ACP_tSNE	0.3764	0.0531
USE_tSNE	0.4390	0.0483
Tf-idf_ACP_tSNE	0.4397	0.1439
VGG16_tSNE	0.4430	0.3170
Tf-idf_tSNE	0.5352	0.0905
Word2Vec_tSNE	0.5373	0.2414



# Étude faisabilité - Meilleurs résultats



## Word2Vec

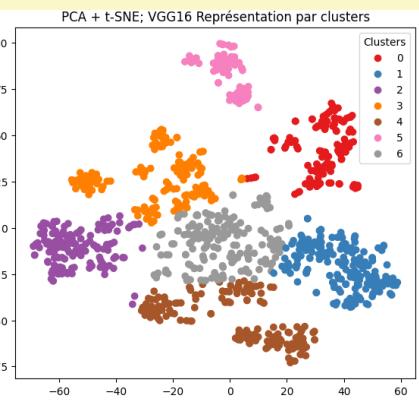
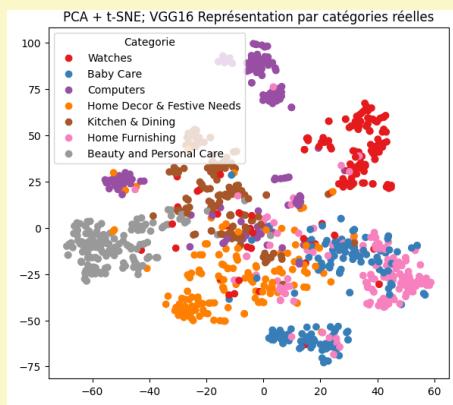


	Method_used	ARI_score	Computing_time
	Word2Vec_tSNE	0.5373	0.2414

	Computers	5	5	4	9	43	0
Baby Care -	0	1.1e+02	19	8	1	9	0
Watches -	0	11	1.1e+02	0	0	29	0
Kitchen & Dining -	0	1	14	1.1e+02	8	19	0
Home Decor & Festive Needs -	0	0	0	1	1.2e+02	24	0
Home Furnishing -	0	0	13	74	2	61	0
Beauty and Personal Care -	0	0	0	0	0	1	1.5e+02
	0	1	2	3	4	5	6

	precision	recall	f1-score	support
0	1.00	0.56	0.72	150
1	0.87	0.75	0.81	150
2	0.68	0.73	0.71	150
3	0.55	0.72	0.63	150
4	0.86	0.83	0.85	150
5	0.33	0.41	0.36	150
6	1.00	0.99	1.00	150
accuracy			0.71	1050
macro avg	0.76	0.71	0.72	1050
weighted avg	0.76	0.71	0.72	1050

## VGG-16



	Method_used	ARI_score	Computing_time
	VGG16_tSNE	0.4430	0.3170

	Computers	3	2	36	0	1	0
Baby Care -	3	1.2e+02	6	20	0	0	3
Watches -	1	1	1e+02	44	0	0	1
Kitchen & Dining -	6	1	5	1.3e+02	0	0	4
Home Decor & Festive Needs -	72	0	1	77	0	0	0
Home Furnishing -	0	5	36	31	0	78	0
Beauty and Personal Care -	0	0	10	2	0	0	1.4e+02
	0	1	2	3	4	5	6

	precision	recall	f1-score	support
0	0.57	0.72	0.64	150
1	0.92	0.79	0.85	150
2	0.63	0.69	0.66	150
3	0.39	0.89	0.54	150
4	0.00	0.00	0.00	150
5	0.99	0.52	0.68	150
6	0.95	0.92	0.93	150
accuracy			0.65	1050
macro avg	0.63	0.65	0.61	1050
weighted avg	0.63	0.65	0.61	1050



Contexte



Étude faisabilité  
NLP  
Image Processing



**Classification supervisée**



Test collecte images, API



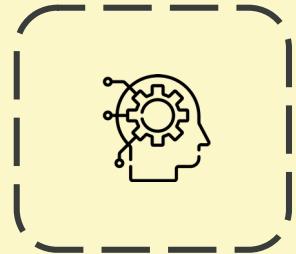
Conclusions

# Approche – Classification supervisée des images



## Création du modèle

VGG-16  
Extraction des features

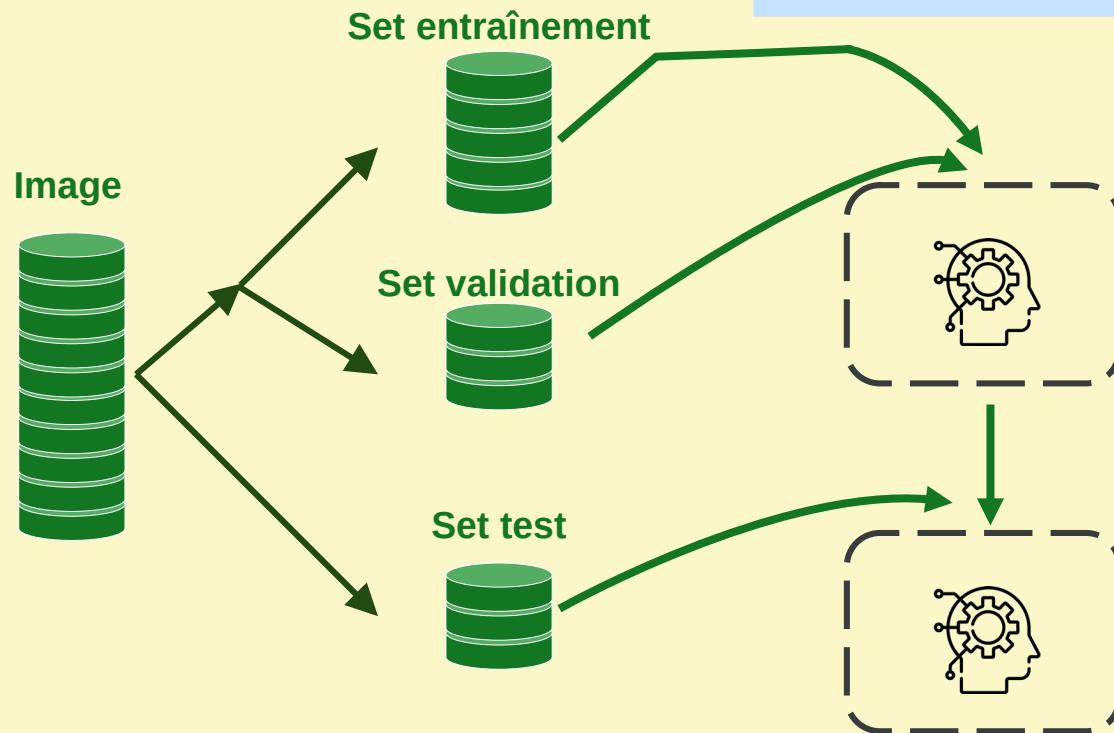


## Données

## Pre-processing split

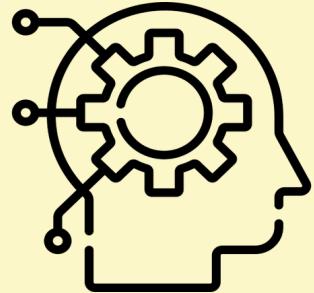
## Entraînement et validation du modèle

## Évaluation



- Loss validation
- Accuracy validation

# Classification supervisée des images



Aspects communs :

- VGG-16
- Extraction des features
- Modèle complété avec :
  - 1 couche Global Average Pooling 2D
  - Couche Dense (relu)
  - Couche dropout 0,5
  - Coche dense, 7 classes, activation softmax

## 2 tests différents

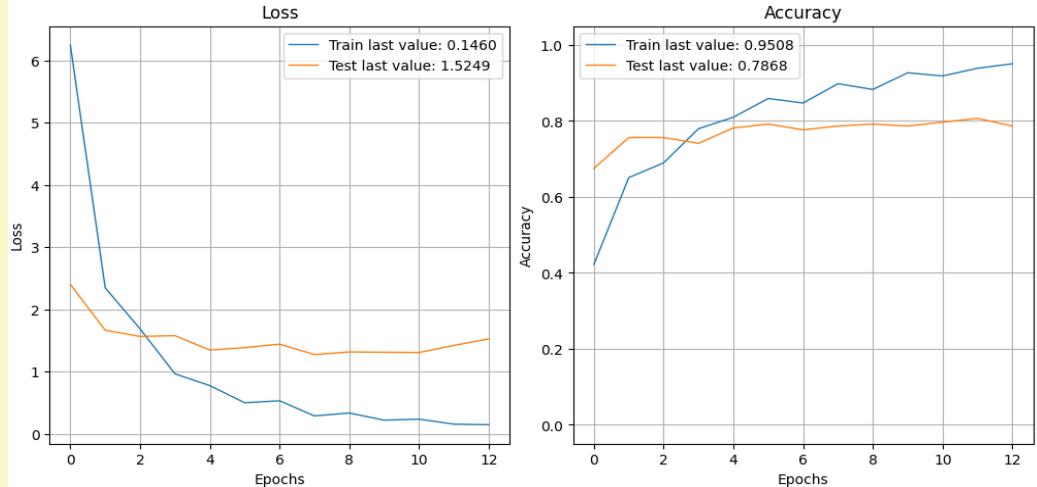
1. Préparation initiale des images
2. Data augmentation avant d'entraîner le modèle

# Classification supervisée des images



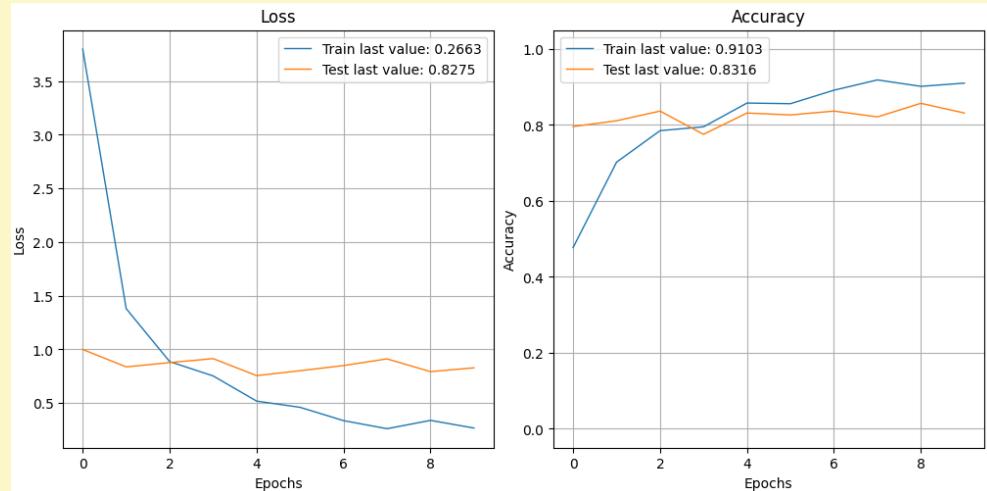
## Test 1. Préparation initiale des images

Préparer les images de la même manière que celles utilisées pour l'entraînement de VGG-16.

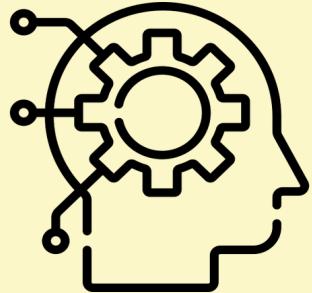


## Test 2. Data augmentation avant d'entraîner le modèle

Création d'un générateur (data augmentation) personnalisé.



# Classification supervisée des images



Aspects communs :

- VGG-16
- Extraction des features
- Modèle complété avec :
  - 1 couche Global Average Pooling 2D
  - Couche Dense (relu)
  - Couche dropout 0,5
  - Coche dense, 7 classes, activation softmax

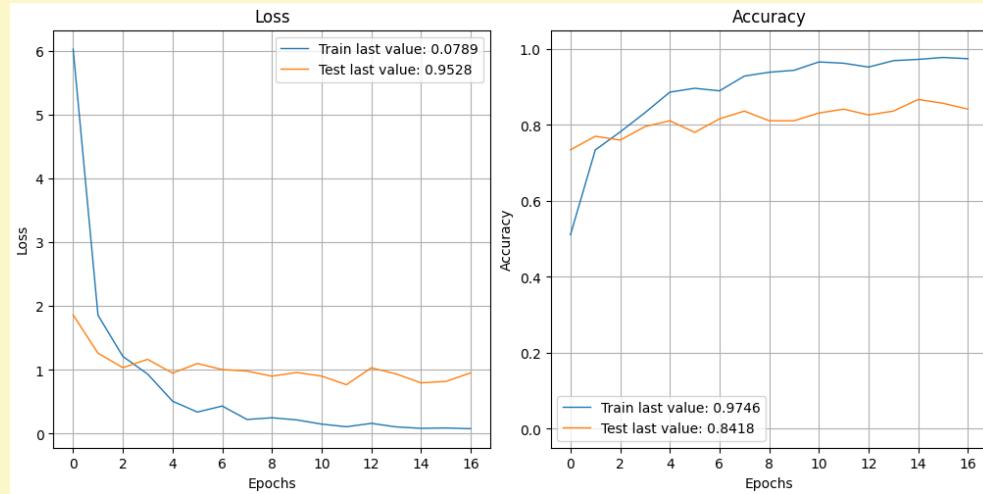
## 2 tests avec approches récentes

3. Entraînement (et test) du modèle directement sur les Datasets
4. Intégration de la data augmentation dans le modèle et entraînement (et test) du modèle directement sur les Datasets

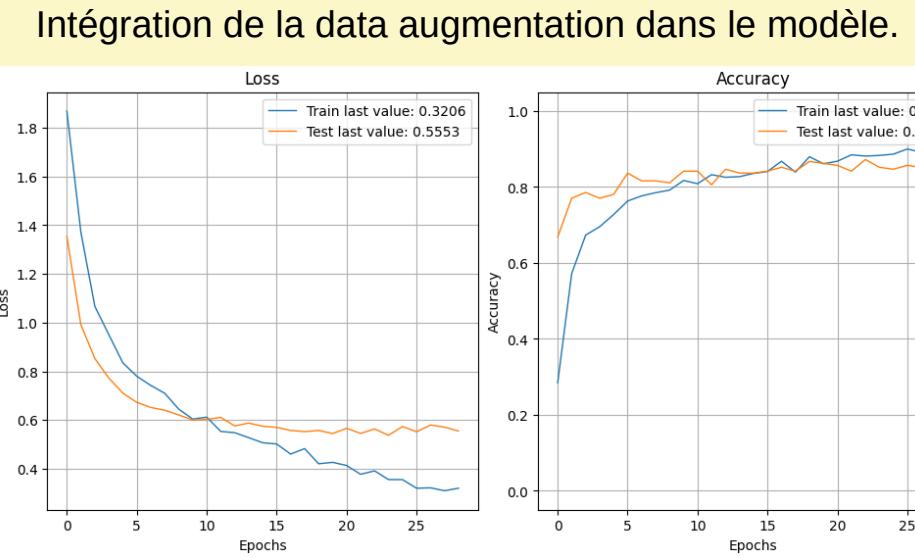
# Classification supervisée des images



## Test 3. Entraînement (et test) du modèle directement sur les Datasets



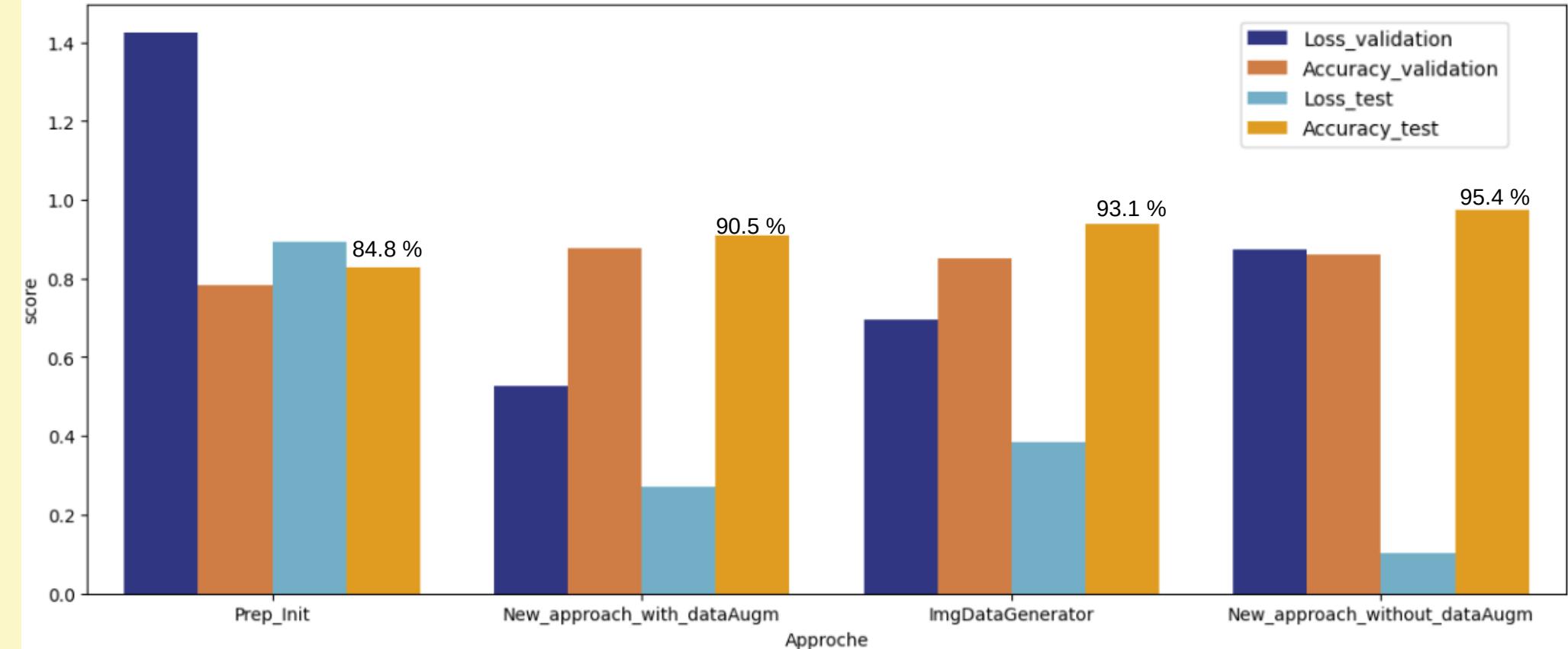
## Test 4. Intégration de la data augmentation dans le modèle et entraînement (et test) du modèle directement sur les Datasets



# Classification supervisée des images



Comparaison des performances des approches





Contexte



Étude faisabilité

NLP

Image Processing



Classification supervisée



**Test collecte images, API**



Conclusions

# Test collecte images, API



**Objectif :** tester la collecte des images pour enrichir la base de données en images « d'épicerie fine »

EDAMAM

Products Nutrition Wizard Recipes Partners

Signup API Login

## Leading provider of nutrition data and analytics

**Nutrition Analysis API**  
Copy/paste any food recipe and learn its nutrition details in under a second

**Food Database API**  
Get free access to a database with close to 900,000 foods and over 680,000 unique UPC codes

**Recipe Search API**  
Search over 2.3 million recipes by diets, calories and nutrient ranges

**Produit à tester:** « champagne »

**Catégories target :** 'foodId', 'label', 'category',  
'categoryLabel', 'image', 'foodContentsLabel'



La conformité au GDPR a été assurée  
à toutes les étapes de la collecte et  
du stockage des données.

Edamam Food and Grocery Database FREEMIUM

Subscribed

Endpoints About Tutorials Discussions Pricing

GET /api/food-database/v2/parser

Test Endpoint

The parser access point handles text search for foods as well as filters for the foods like presence specific nutrient content or exclusion of allergens.

- Search for a phrase or keyword using NLP to get food entities from it.
- Get basic nutrition facts and...

Personal Account raquel.sanchez-pellicer

RapidAPI App default-application\_7364205 REQUIRED

Request URL rapidapi.com REQUIRED

```
(Node.js Axios) Copy Code
const axios = require('axios');

const options = {
  method: 'GET',
  url: 'https://edamam-food-and-grocery-database.p.rapidapi.com/api/food-database/v2/parser',
  params: {
    'nutrition-type': 'cooking',
    'category[0]': 'generic-foods',
    'health[0]': 'alcohol-free'
  },
  headers: {
    'X-RapidAPI-Key': '704dbaf6dfnshd5cb3516523976fp1dfcfejsn52770b5e2a98',
    'X-RapidAPI-Host': 'edamam-food-and-grocery-database.p.rapidapi.com'
  }
};

try {
  const response = await axios.request(options);
  console.log(response.data);
} catch (error) {
  console.error(error);
}
```

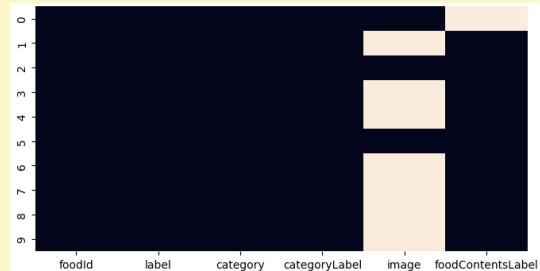
# Test collecte images, API



## Enregistrement des 10 premiers produits dans un fichier en format .csv

- Nombre de « NAN » important

(10, 6)	foodId	label	category	categoryLabel	image	foodContentsLabel
0	food_a656mk2a5dmqb2adiamu6beihduu	Champagne	Generic foods	food	https://www.edamam.com/food-img/a71/a718cf3c52...	NaN
1	food_b753ithamdb8psbt0w2k9aqu006c	Champagne Vinaigrette, Champagne	Packaged foods	food	NaN	OLIVE OIL; BALSAMIC VINEGAR; CHAMPAGNE VINEGAR...
2	food_b3dyababjo54xobm6r8jzbghjgqe	Champagne Vinaigrette, Champagne	Packaged foods	food	https://www.edamam.com/food-img/d88/d88b64d973...	INGREDIENTS: WATER; CANOLA OIL; CHAMPAGNE VINE...
3	food_a9e0ghsamvoc45bwa2ybsa3gken9	Champagne Vinaigrette, Champagne	Packaged foods	food	NaN	CANOLA AND SOYBEAN OIL; WHITE WINE (CONTAINS S...
4	food_an4jjeauacpus2a3u1ni8auhe7q9	Champagne Vinaigrette, Champagne	Packaged foods	food	NaN	WATER; CANOLA AND SOYBEAN OIL; WHITE WINE (CON...
5	food_bmu5dmkazwuvpaa5prh1daa8jxs0	Champagne Dressing, Champagne	Packaged foods	food	https://www.edamam.com/food-img/ab2/ab2459fc2a...	SOYBEAN OIL; WHITE WINE (PRESERVED WITH SULFIT...
6	food_apl44taoyv11ra0lic1qa8xculi	Champagne Buttercream	Generic meals	meal	NaN	sugar; butter; shortening; vanilla; champagne;...
7	food_byap67hab6evc3a0f9w1oag3s0qf	Champagne Sorbet	Generic meals	meal	NaN	Sugar; Lemon juice; brandy; Champagne; Peach
8	food_am5egz6aq3fpjlaf8xpkdbc2asis	Champagne Truffles	Generic meals	meal	NaN	butter; cocoa; sweetened condensed milk; vanil...
9	food_bcz8rhiajk1fuva0vkfmeakbouco	Champagne Vinaigrette	Generic meals	meal	NaN	champagne vinegar; olive oil; Dijon mustard; s...



## Suppression des cellules vides et enregistrement de lignes restantes dans un fichier en format .csv

- Seulement 3 lignes complètes

(3, 6)	foodId	label	category	categoryLabel	image	foodContentsLabel
0	food_b3dyababjo54xobm6r8jzbghjgqe	Champagne Vinaigrette, Champagne	Packaged foods	food	https://www.edamam.com/food-img/d88/d88b64d973...	INGREDIENTS: WATER; CANOLA OIL; CHAMPAGNE VINE...
1	food_bmu5dmkazwuvpaa5prh1daa8jxs0	Champagne Dressing, Champagne	Packaged foods	food	https://www.edamam.com/food-img/ab2/ab2459fc2a...	SOYBEAN OIL; WHITE WINE (PRESERVED WITH SULFIT...
2	food_bu12urpb tuo9v6b4jpdk2a1fh4hh	Champagne Simply Dressed Vinaigrette, Champagne	Packaged foods	food	https://www.edamam.com/food-img/736/736a3e27a6...	FILTERED WATER; CANOLA OIL; CHAMPAGNE AND WHIT...



## Contexte



Étude faisabilité  
NLP  
Image Processing



## Classification supervisée



## Test collecte images, API



## Conclusions

# Conclusions



## ➤ Analyse de faisabilité

- À partir de texte, meilleur résultat avec **Word2Vec, ARI 0.5245**
- À partir d'images, meilleur résultat avec **VGG-16, ARI 0.4435**



## ➤ Test classification supervisée d'images

- Meilleur résultat **approche récente DataSet sans data augmentation 95.4 %**



- **Test collecte données** pour enrichir la base de données en images « d'épicerie fine », ex. produits avec champagne à travers **API Edamam**
  - **Données disponibles**
  - **Insuffisants**

# Conclusions

place de marché

## Possibles pistes d'amélioration :

### ➤ Jeu de données

Ajouter des données, seulement 150 par catégorie

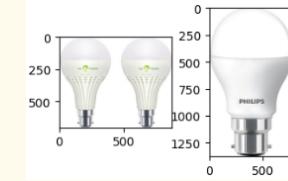
### ➤ Étiquettes

NOMBREUSES erreurs dans les catégories Home Decor & Festive Needs / Kitchen & Dinning / Baby care

#### Home Decor & Festive Needings



#### Kitchen and dinning



#### Baby care



### ➤ Classification non supervisée

Tester des valeurs de  $k > 7$  (subdivision des catégories ou se cumulent les erreurs?)

### ➤ Tester d'autres méthodes (BERT TensorFlow, RoBERTa, ResNet)

### ➤ Élargir la recherche d'API pour la collecte d'images



# Annexes

<https://www.economie.gouv.fr/entreprises/reglement-general-protection-donnees-rgpd>

## Le RGPD, qu'est-ce que c'est ?

Le règlement général de protection des données (RGPD) est un texte réglementaire européen qui encadre le traitement des données de manière égalitaire sur tout le territoire de l'Union européenne (UE). Il est entré en application le 25 mai 2018.

Le RGPD s'inscrit dans la continuité de la loi française « Informatique et Libertés » de 1978, modifiée par la loi du 20 juin 2018 relative à la protection des données personnelles, établissant des règles sur la collecte et l'utilisation des données sur le territoire français. Il a été conçu autour de trois objectifs :

renforcer les droits des personnes

responsabiliser les acteurs traitant des données

crédibiliser la régulation grâce à une coopération renforcée entre les autorités de protection des données.

<https://www.cnil.fr/fr/cnil-direct/question/quels-sont-les-grands-principes-des-regles-de-protection-des-donnees>

[La Commission Nationale de l'Informatique et des Libertés (CNIL)]

## **Quels sont les grands principes des règles de protection des données personnelles ?**

Les 5 grands principes des règles de protection des données personnelles sont les suivants :

- **Le principe de finalité** : le responsable d'un fichier ne peut enregistrer et utiliser des informations sur des personnes physiques que dans un but bien précis, légal et légitime ;
- **Le principe de proportionnalité et de pertinence** : les informations enregistrées doivent être pertinentes et strictement nécessaires au regard de la finalité du fichier ;
- **Le principe d'une durée de conservation limitée** : il n'est pas possible de conserver des informations sur des personnes physiques dans un fichier pour une durée indéfinie. Une durée de conservation précise doit être fixée, en fonction du type d'information enregistrée et de la finalité du fichier ;
- **Le principe de sécurité et de confidentialité** : le responsable du fichier doit garantir la sécurité des informations qu'il détient. Il doit en particulier veiller à ce que seules les personnes autorisées aient accès à ces informations ;
- **Les droits des personnes.**