

Projet 6

Classifiez automatiquement des biens de consommation

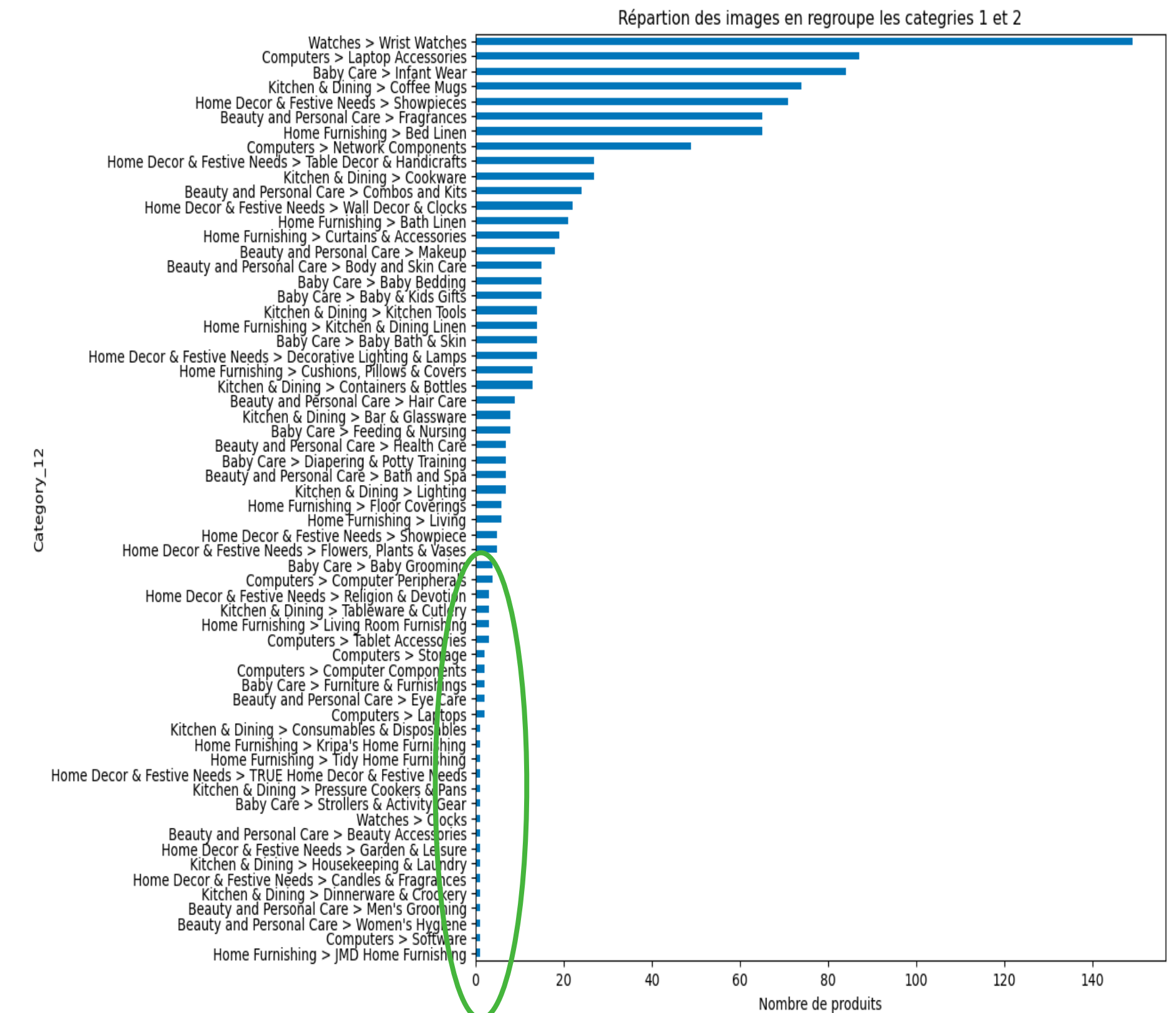
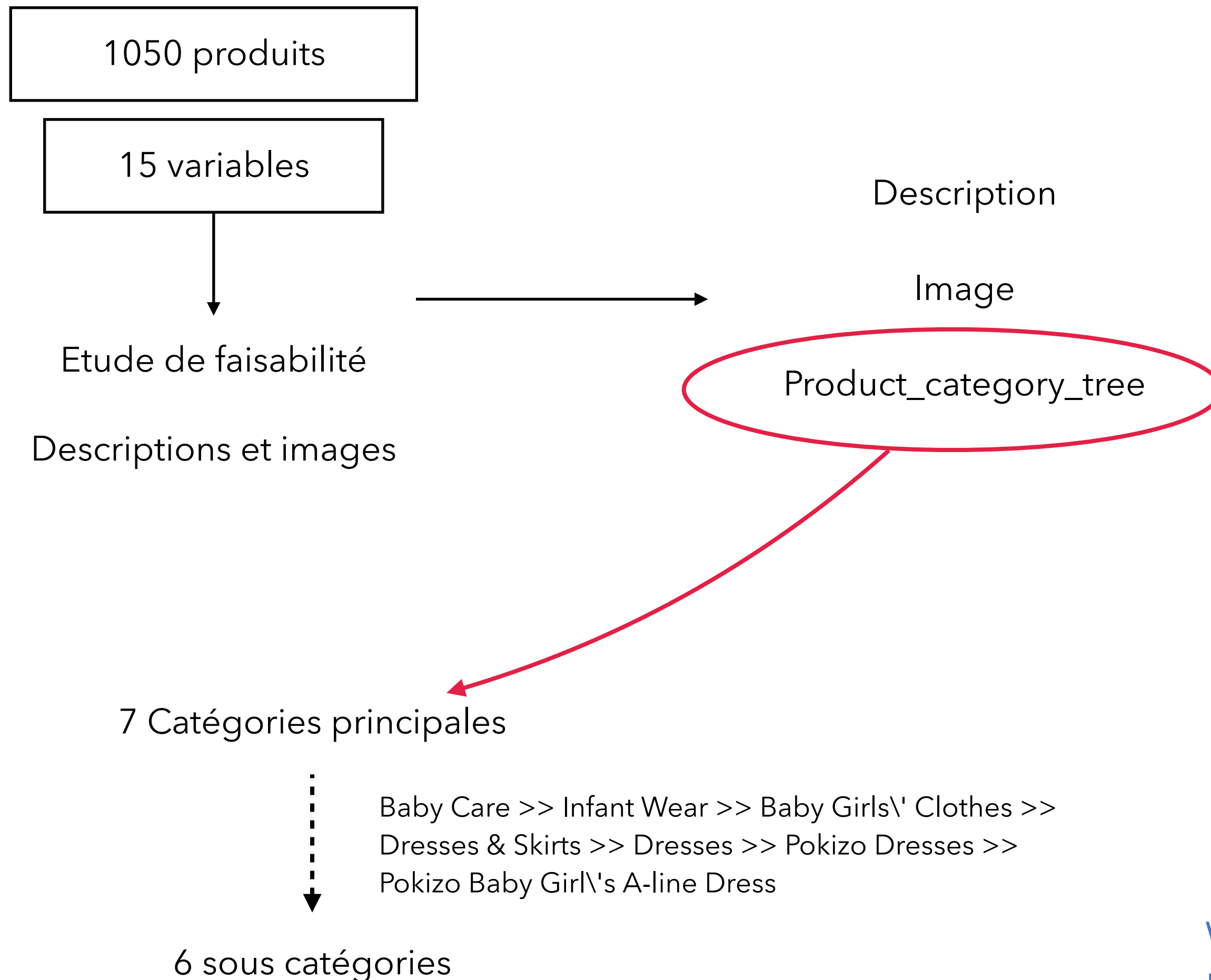
Lancement d'une marketplace e-commerce de « Place de marché »

Etude de l'automatisation de la classification d'articles

- 1 - Présentation du jeu de données
- 2 - Etude de faisabilité
 - 2.1 - Descriptions produits
 - 2.2 - Images produits
 - 2.3 - Résultats de l'étude
- 3 - Classification supervisée à partir des images
- 4 - Collecte de nouveaux produits
- 5 - Conclusion

1 - Présentation du jeu de données

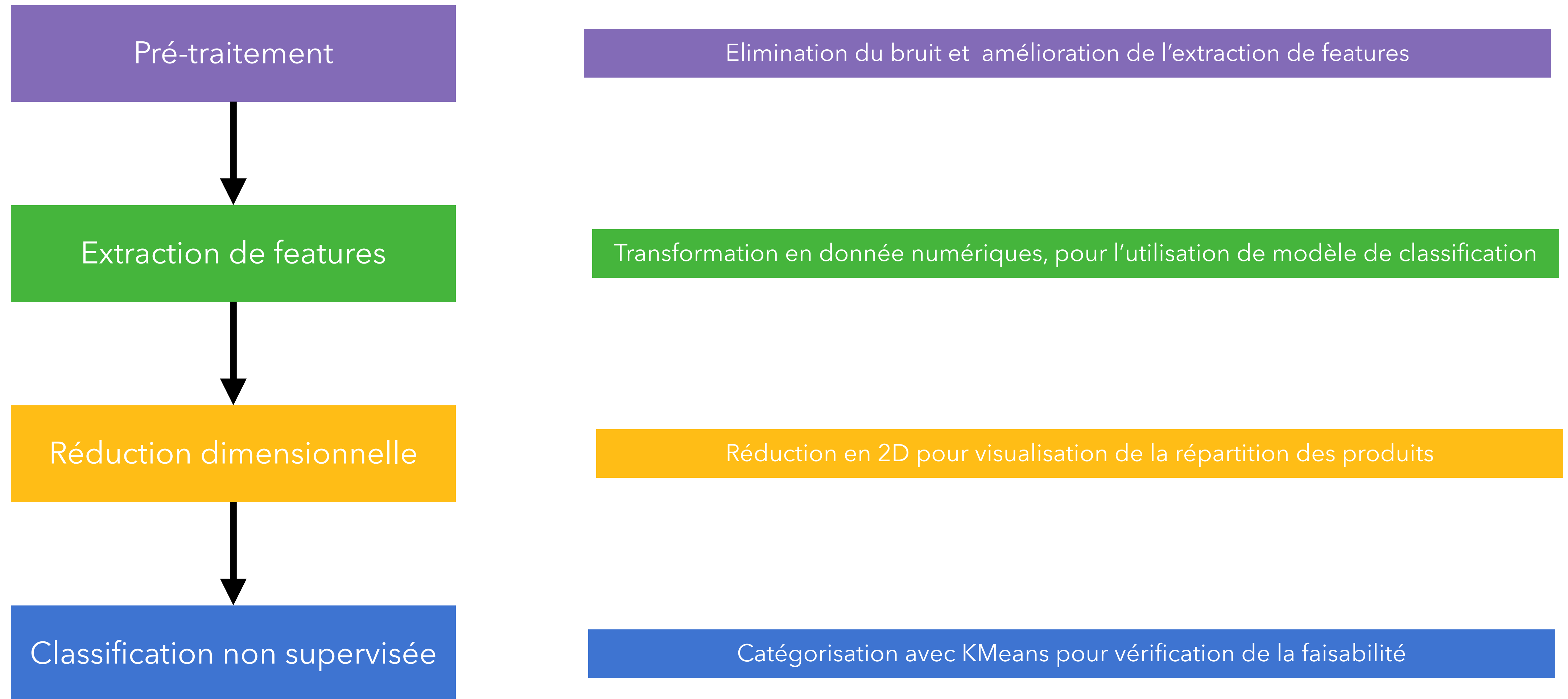
Jeu de donnée fournie par Place du marché



Watches / Kitchen & Dinning / Home Furnishing /
Home Decor & Festive Needs / Computers / Beauty and Personal Care /
Baby Care

2- Etude de faisabilité

Descriptions de produits



2- Etude de faisabilité

Descriptions de produits

Pré-traitement

Descriptions brutes

Tokenization

Normalisation

Passage en minuscule

Suppression des stop_words

Suppression des petits tokens

Suppression des tokens numériques

Stemmatisation ou Lemmatisation

Extraction de features

Bags of Words :
CounterVectorizer
TfidfVecotrizer

Méthodes classiques
Comptage d'apparition de chaque mots

Words Embeddings :
Word2Vec
Doc2Vec

Représentations vectorielles des mots
Permits de capturer
la similarité sémantique

Réseaux de Neurones :
DistilBERT-base-uncased
DistilBERT-base-uncased-finetuned-
SST-2
RoBERTa-Large
USE-v4
USE-Large-v5

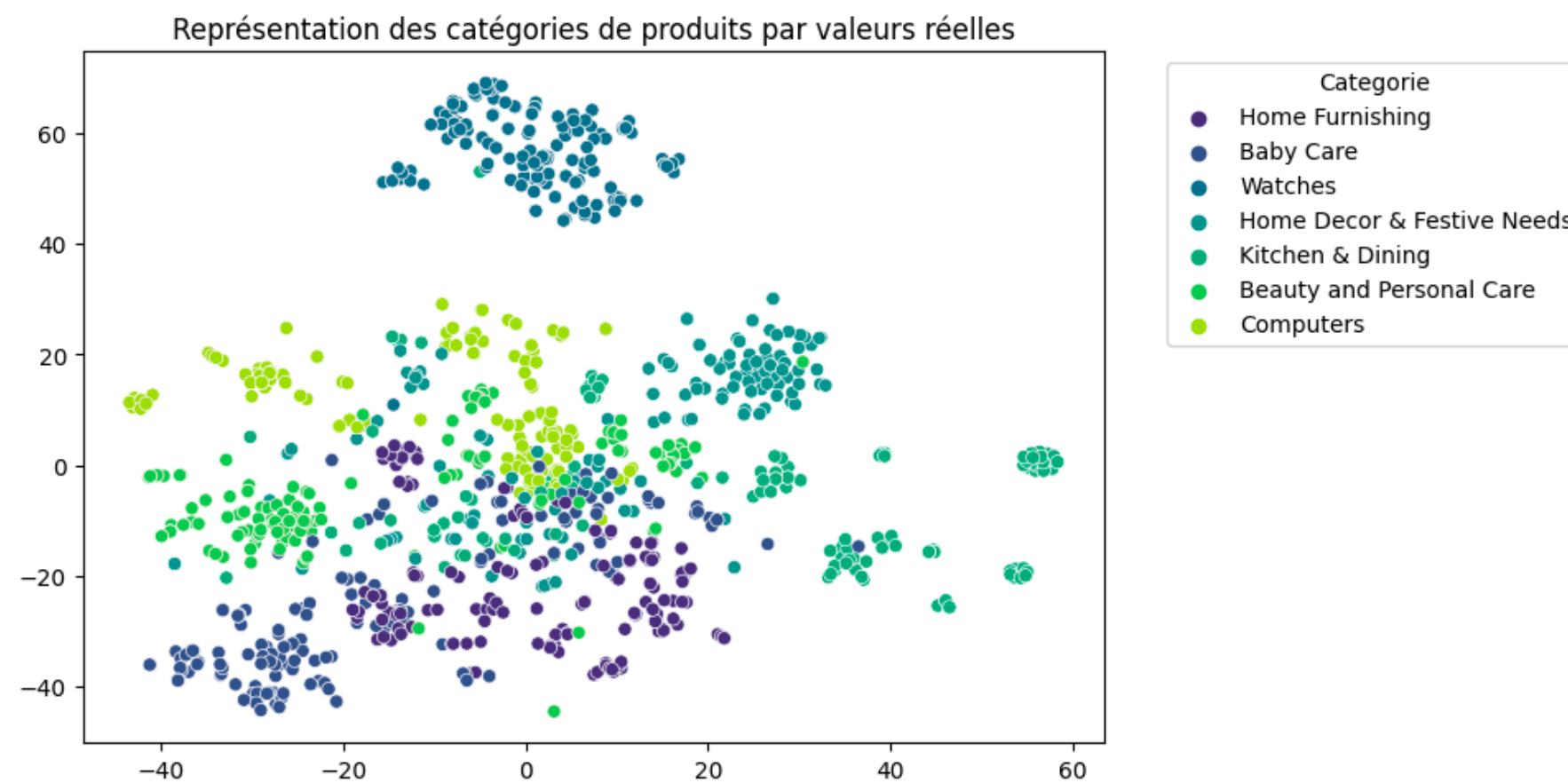
Réseaux de neurones interconnectés
Utilisation de modèle pré-entraînés
sur de très grands corpus

2- Etude de faisabilité

Descriptions de produits

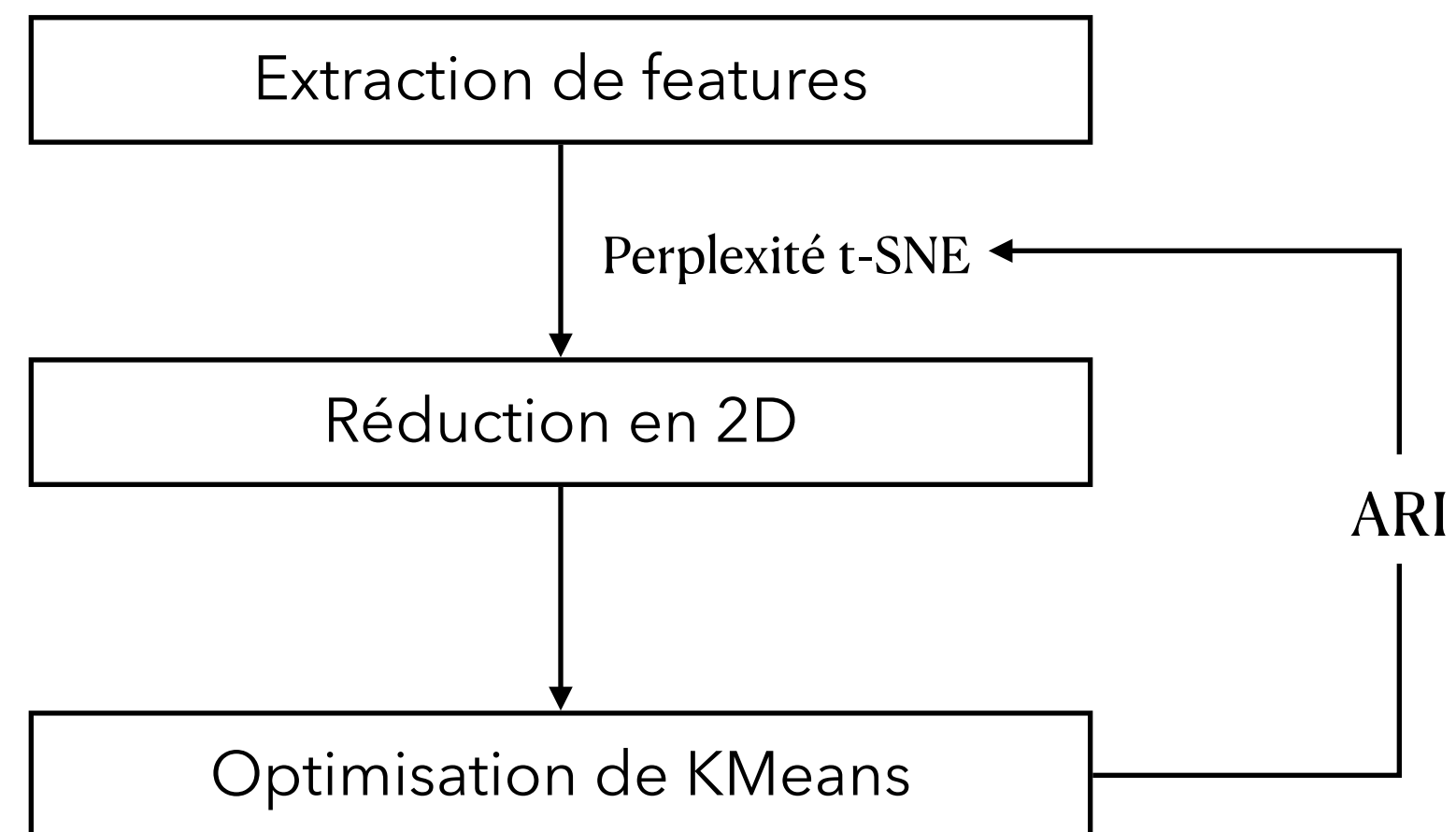
Réduction dimensionnelle

t-SNE → projection en 2D



Classification non supervisée

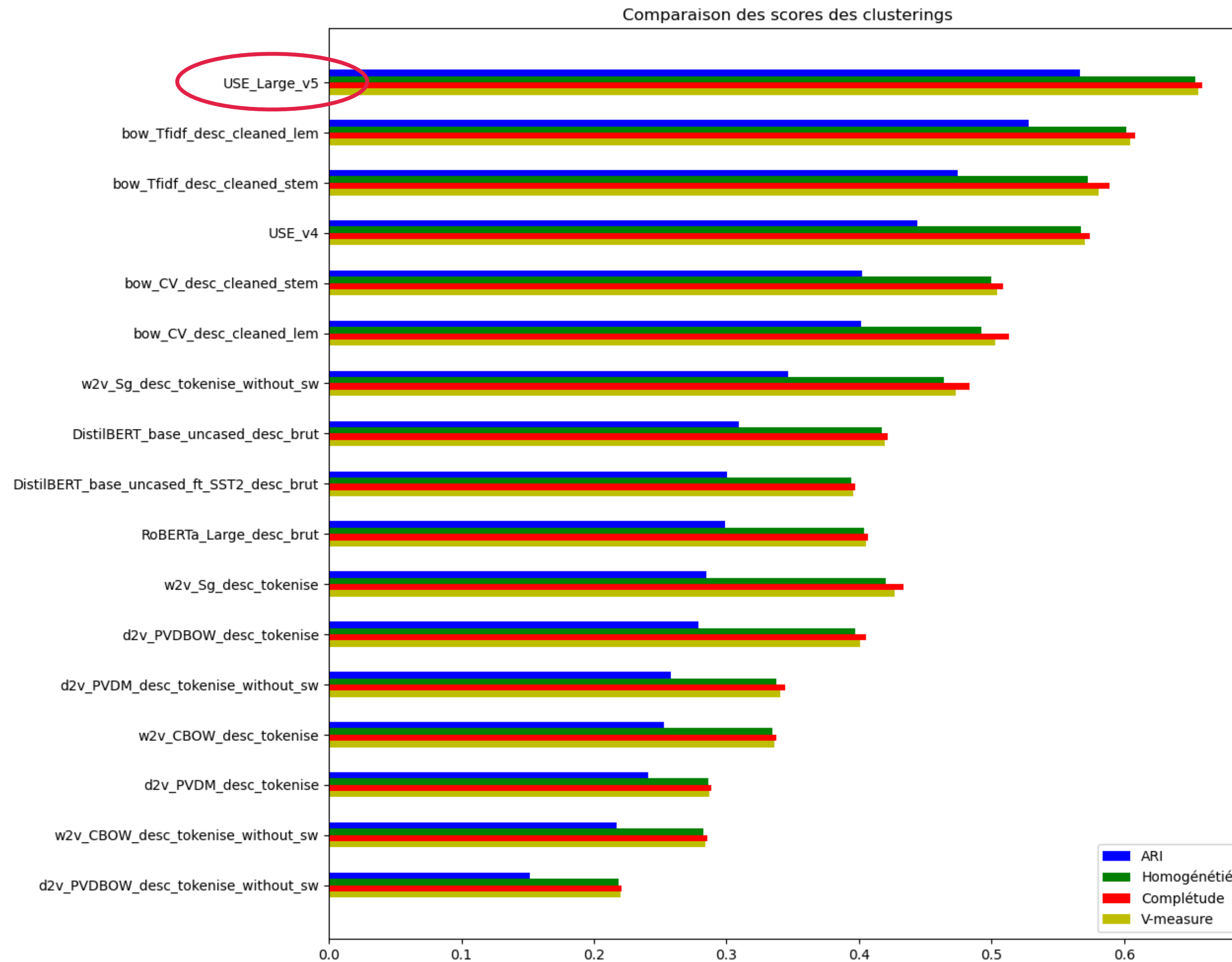
KMeans → 7 clusters



Extraction de features → Meilleures combinaison : perplexité t-SNE / hyperparamètres KMeans → Comparaison des résultats

2- Etude de faisabilité

Descriptions de produits

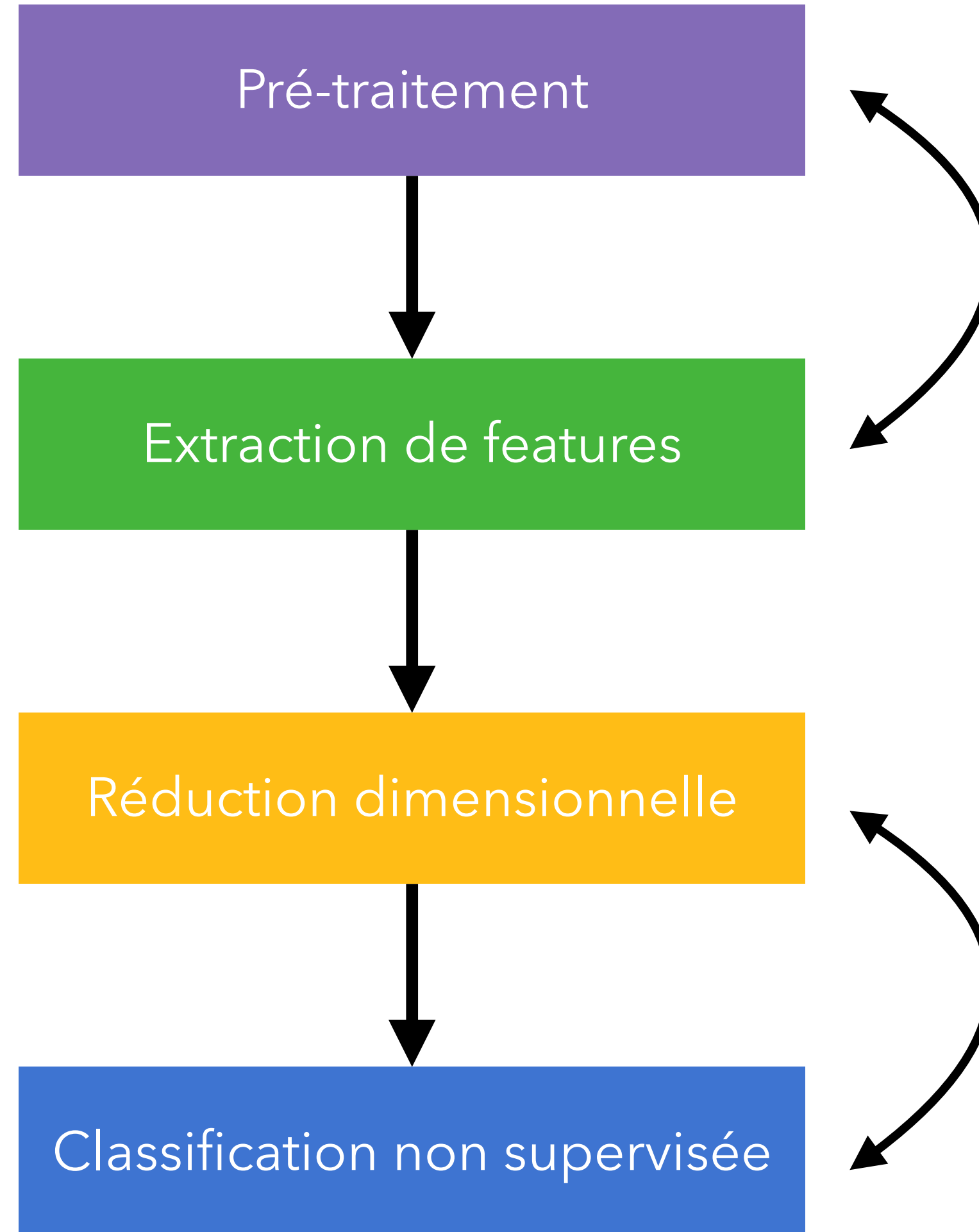


		USE_Large_v5						
Vraies Catégories	Kitchen & Dining	102	2	0	14	20	12	0
	Watches	3	123	3	6	0	15	0
	Home Furnishing	0	0	149	0	0	1	0
	Home Decor & Festive Needs	0	0	14	88	0	48	0
	Baby Care	37	0	0	0	86	27	0
	Computers	1	3	7	51	1	87	0
	Beauty and Personal Care	0	0	0	0	0	1	149
		0	1	2	3	4	5	6
		Prédictions						

Précision du clustering de 75 %

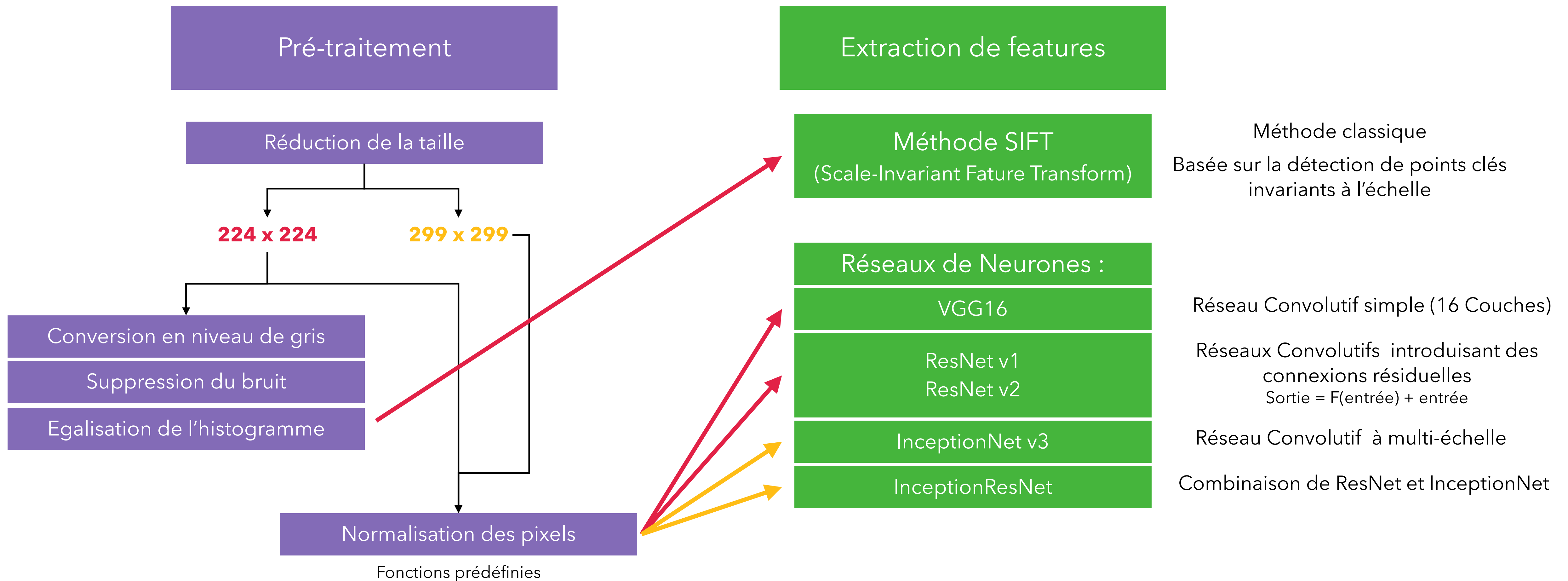
2- Etude de faisabilité

Images produits



2- Etude de faisabilité

Images produits



2- Etude de faisabilité

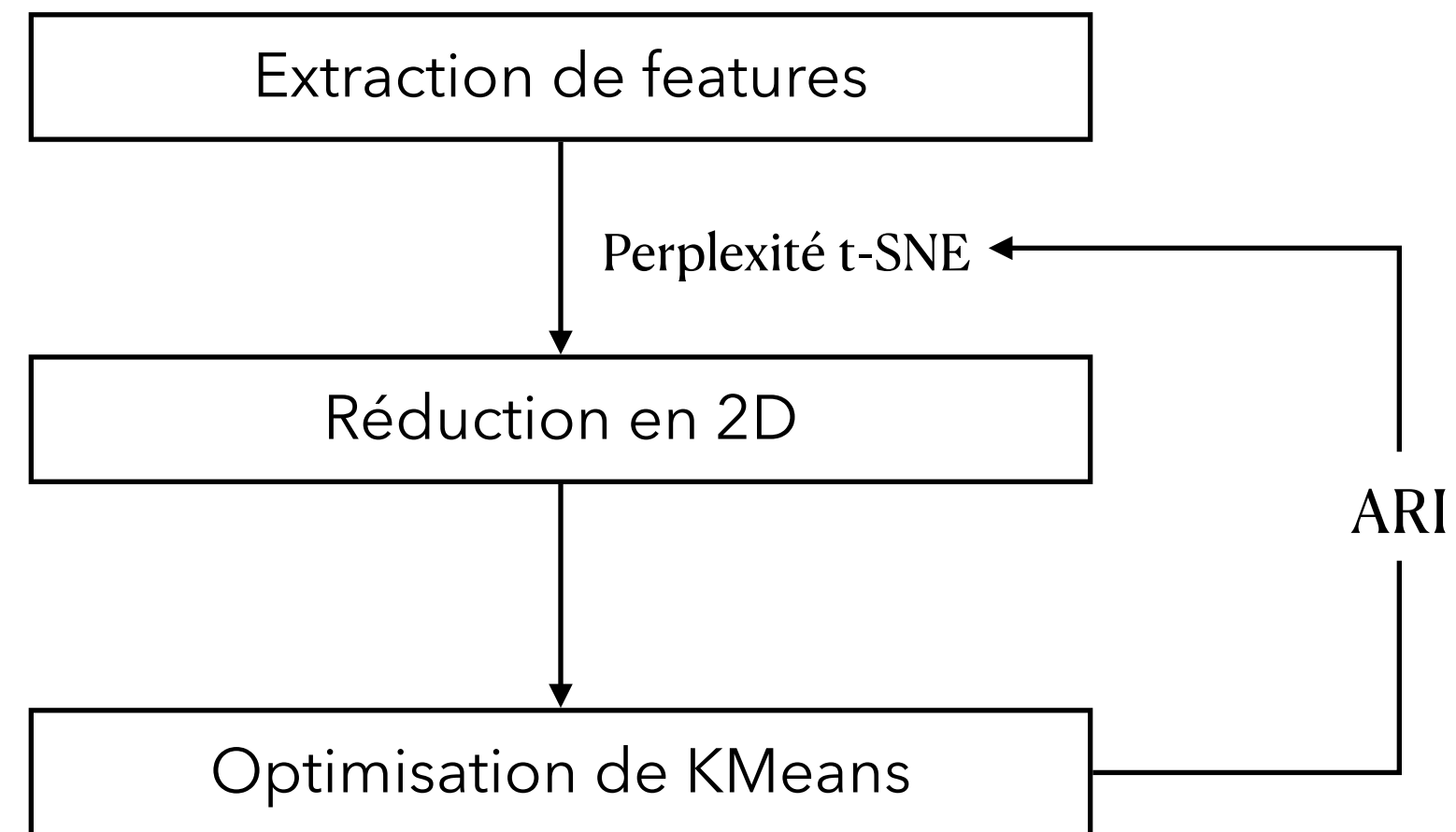
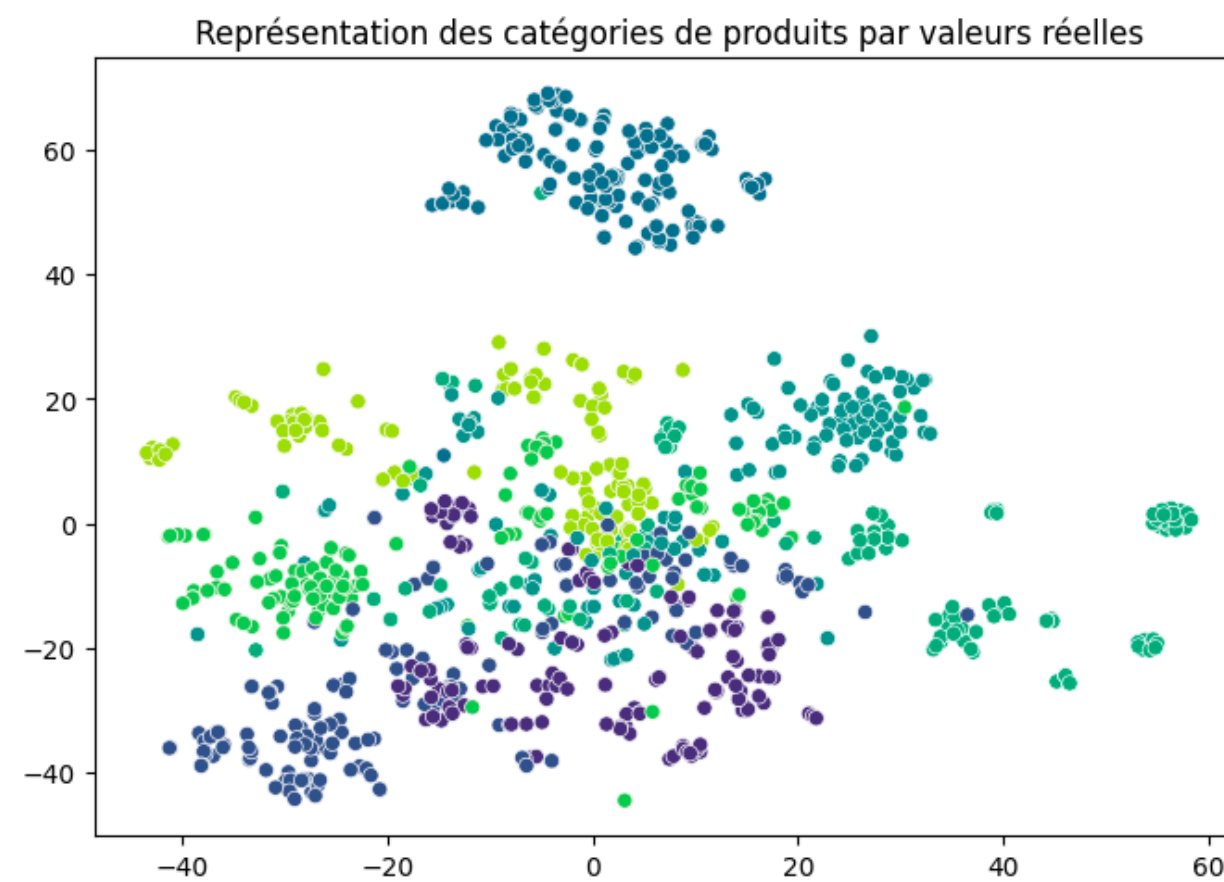
Images produits

Réduction dimensionnelle

Classification non supervisée

t-SNE → projection en 2D

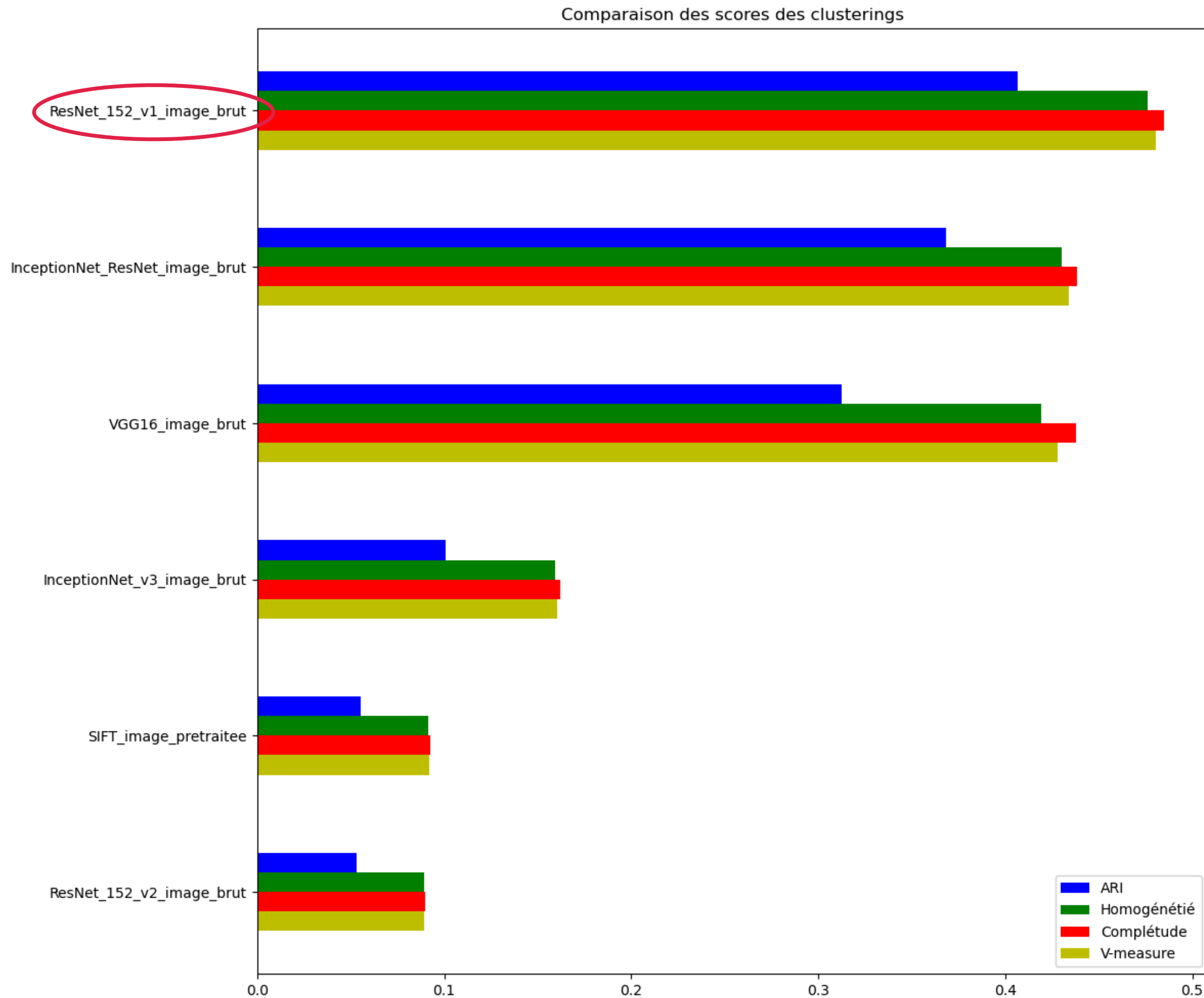
KMeans → 7 clusters



Meilleures combinaison :
perplexité t-SNE / hyperparamètres KMeans

2- Etude de faisabilité

Images produits

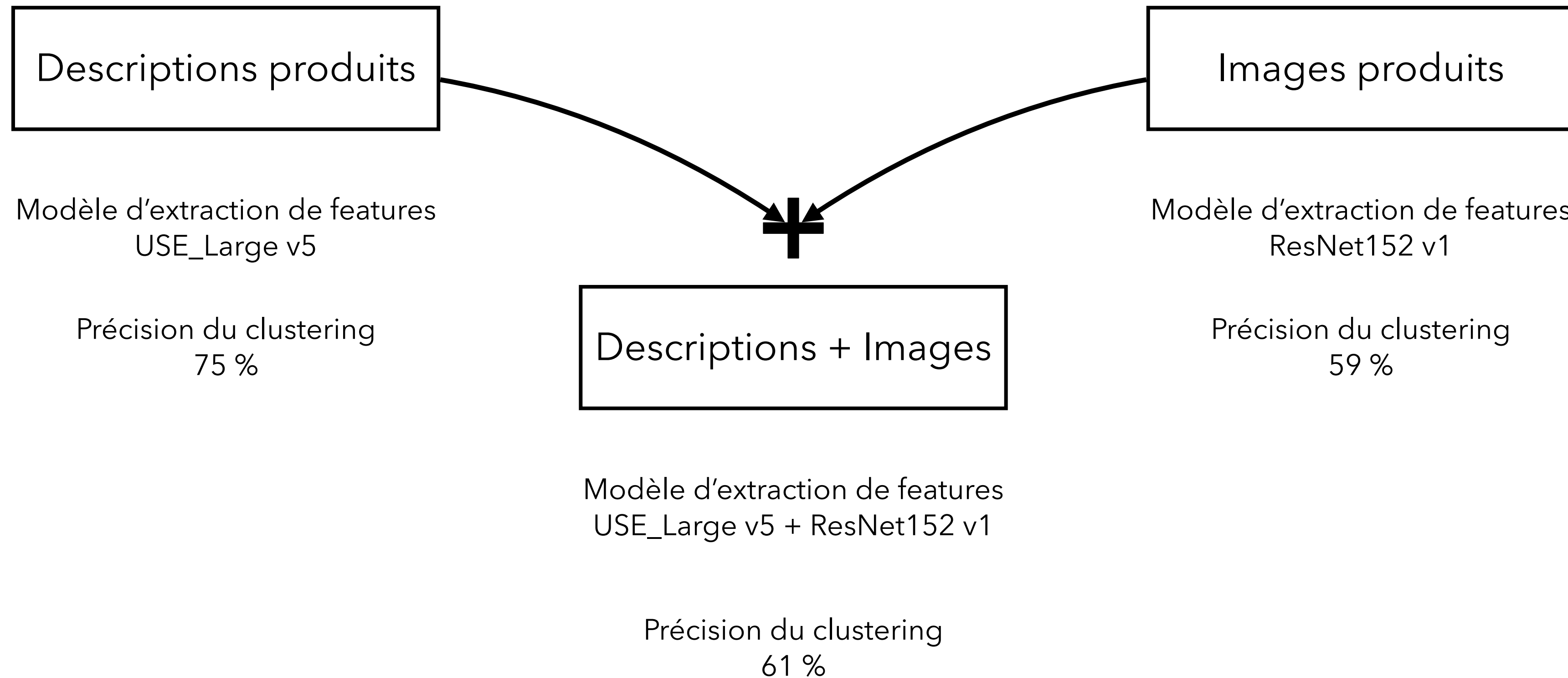


		ResNet_152_v1						
Vraies Catégories	Computers	3	4	24	50	63	3	3
	Beauty and Personal Care	4	119	16	5	3	0	3
	Kitchen & Dining	40	5	76	4	2	2	21
	Baby Care	8	2	48	71	7	2	12
	Watches	29	0	10	10	99	2	0
	Home Decor & Festive Needs	6	4	28	1	0	104	7
	Home Furnishing	0	1	3	0	0	0	146
		0	1	2	3	4	5	6
		Prédictions						

Précision du clustering de 59 %

2- Etude de faisabilité

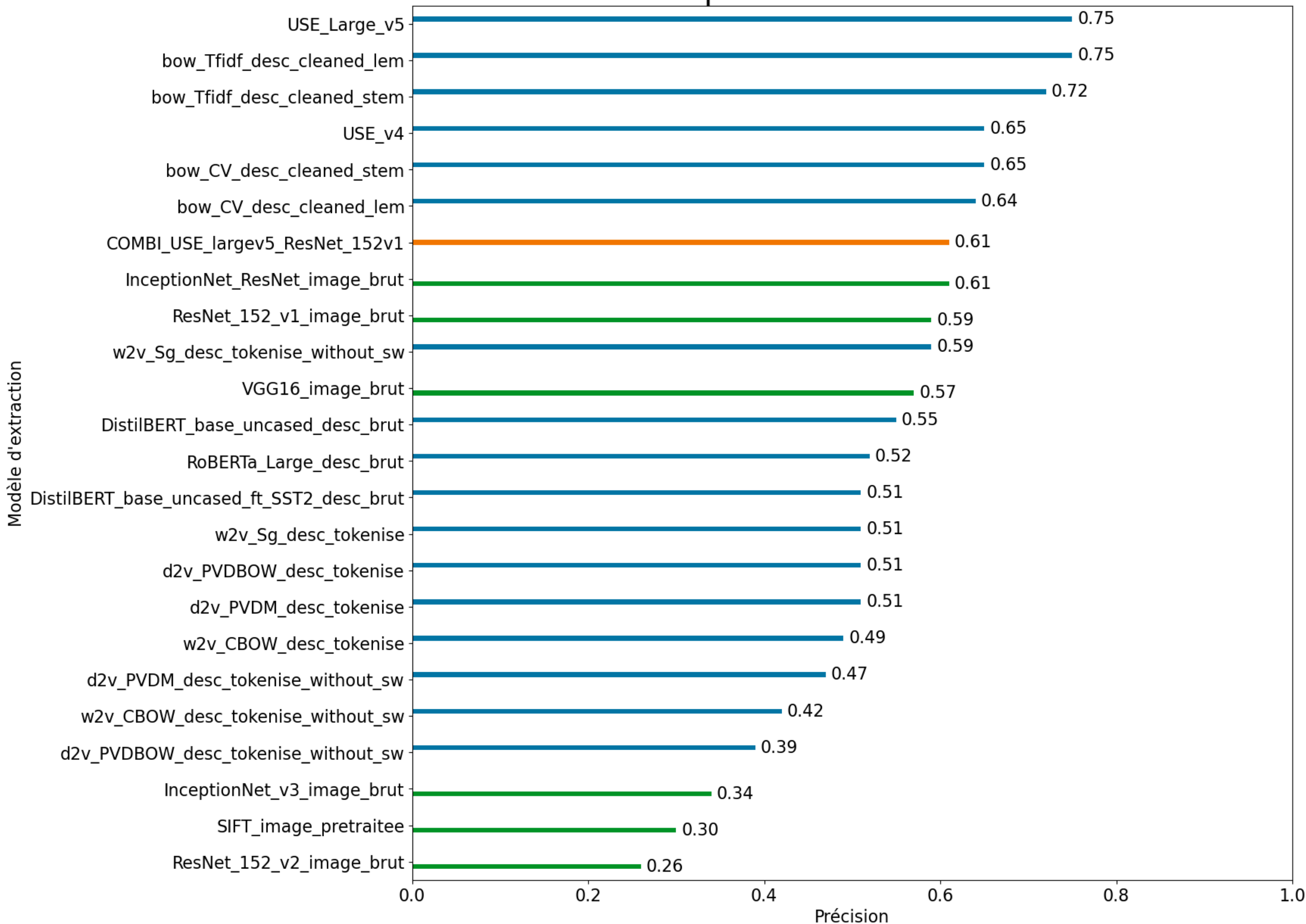
Résultats



2- Etude de faisabilité

Résultats

Comparaison des Précisions

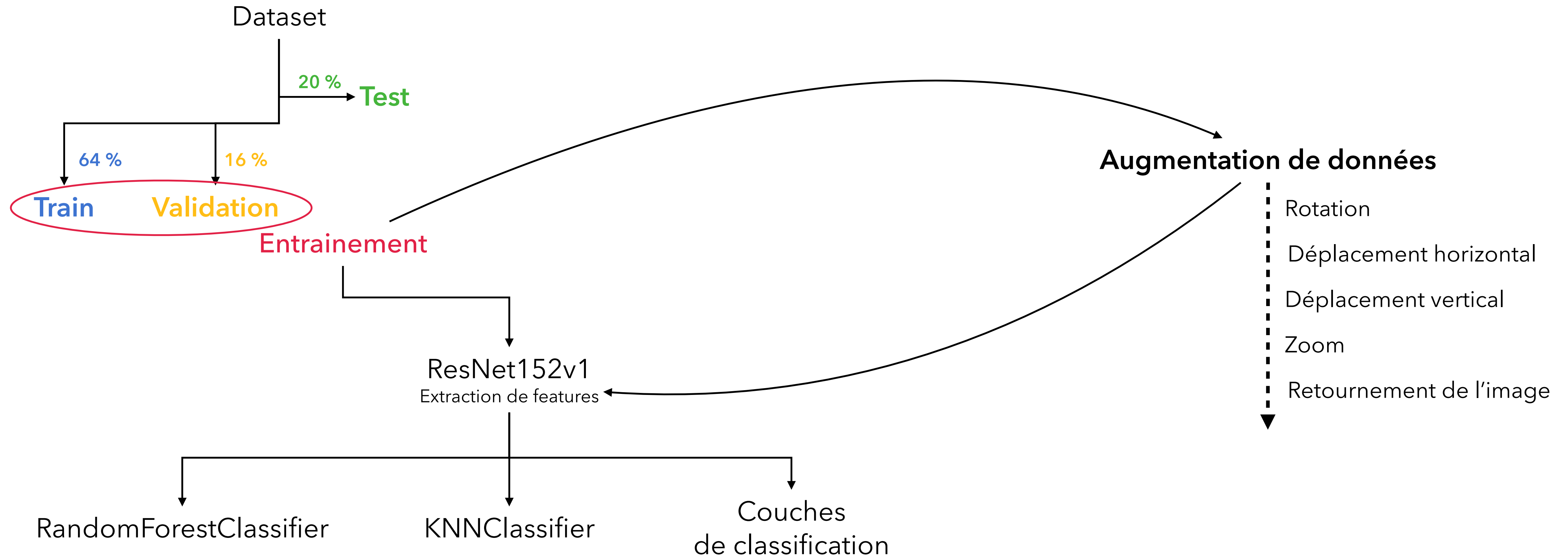


Classification supervisée
basée sur les descriptions semble prometteuse

Utilisation unique des images ou d'une combinaison
avec les descriptions semble plus difficile
à mettre en place

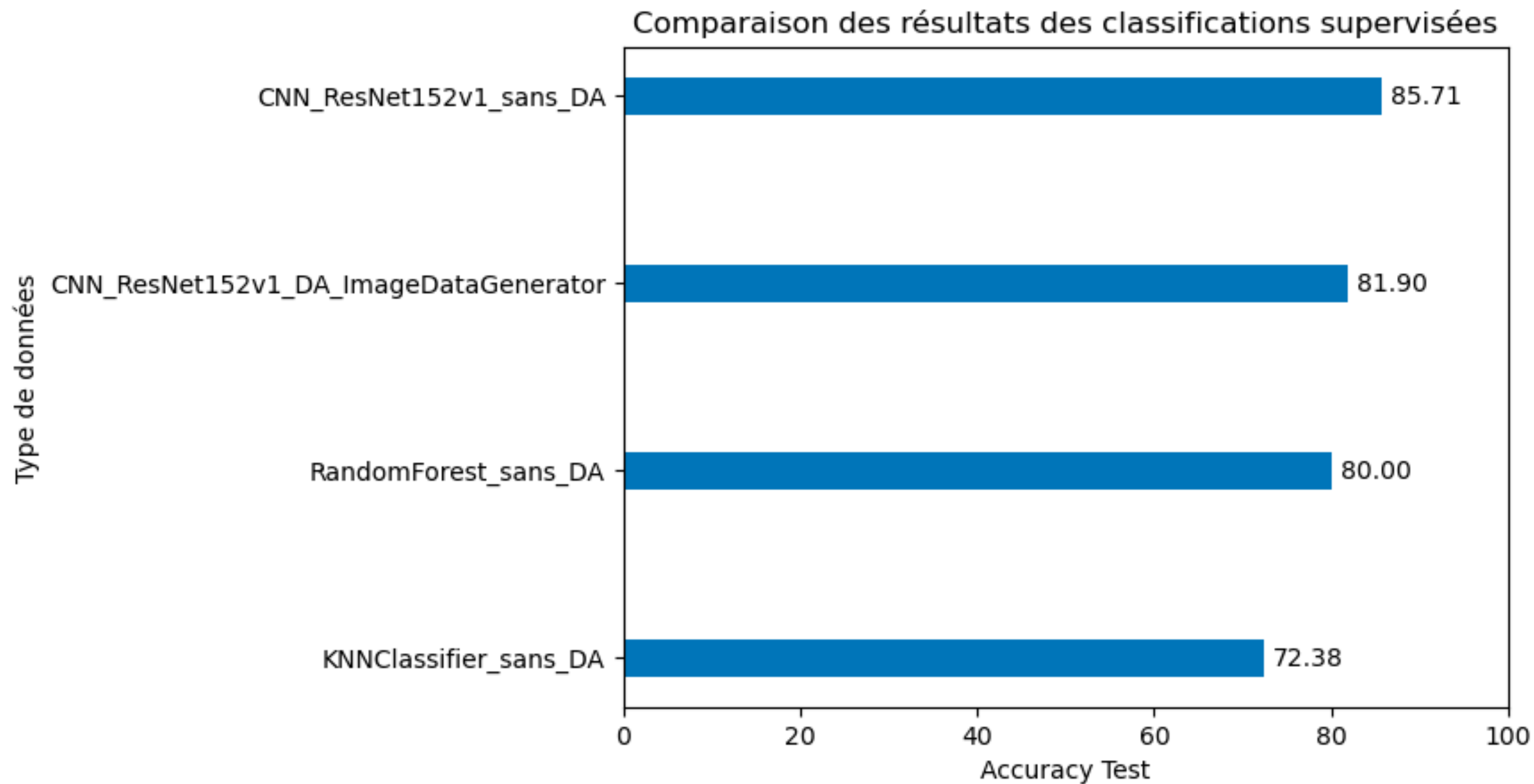
3- Classification supervisée

Images produits



3- Classification supervisée

Images produits



Meilleur Résultat :
Réseau de neurones convolutif

Supériorité du réseau de neurones
sur les modèles de classification plus classique

Augmentation de donnée :
- Pas d'amélioration de performances
- Modèle plus robuste

4- Collecte de nouveaux produits

Elargissement de la gamme de produits



Extraction_10_prduits_champagne

	foodid	label	category	foodContentsLabel	image
0	food_a656mk2a5dmqb2adiamu6beihduu	Champagne	Generic foods	None	a718cf3c52add522128929f1f324d2ab.jpg
1	food_b753ithamdb8psbt0w2k9aquo06c	Champagne Vinaigrette, Champagne	Packaged foods	OLIVE OIL; BALSAMIC VINEGAR; CHAMPAGNE VINEGAR; GARLIC; DIJON MUSTARD; SEA SALT.	None
2	food_b3dyababjo54xobm6r8jzbghjgqe	Champagne Vinaigrette, Champagne	Packaged foods	INGREDIENTS: WATER; CANOLA OIL; CHAMPAGNE VINEGAR; SUGAR; OLIVE OIL; SALT; DRIED GARLIC; DRED SHALLOTS; f	d88b64d97349ed062368972113124e35.jpg
3	food_a9e0ghsamvoc45bwa2ybsa3gken9	Champagne Vinaigrette, Champagne	Packaged foods	CANOLA AND SOYBEAN OIL; WHITE WINE (CONTAINS SULFITES); WATER; VINEGARS (CHAMPAGNE AND WHITE WINE); SU	None
4	food_an4jjueaucpus2a3u1ni8auhe7q9	Champagne Vinaigrette, Champagne	Packaged foods	WATER; CANOLA AND SOYBEAN OIL; WHITE WINE (CONTAINS SULFITES); VINEGARS (CHAMPAGNE AND WHITE WINE); SU	None
5	food_bmu5dmkazwuvpaa5prh1daa8jxs0	Champagne Dressing, Champagne	Packaged foods	SOYBEAN OIL; WHITE WINE (PRESERVED WITH SULFITES); WATER; WHITE WINE VINEGAR; SUGAR; SALT; SPICES (INCLUD	ab2459fc2a98cd35f68b848be2337ecb.jpg
6	food_alpl44taoyv11ra0lic1qa8xculi	Champagne Buttercream	Generic meals	sugar; butter; shortening; vanilla; champagne; milk	None
7	food_byap67hab6evc3a0f9w1oag3s0qf	Champagne Sorbet	Generic meals	Sugar; Lemon juice; brandy; Champagne; Peach	None
8	food_am5egz6aq3fpjlaf8xpkdbc2asis	Champagne Truffles	Generic meals	butter; cocoa; sweetened condensed milk; vanilla extract; champagne; powdered sugar	None
9	food_bcz8rhiajk1fuva0vkfmeakbouc0	Champagne Vinaigrette	Generic meals	champagne vinegar; olive oil; Dijon mustard; shallot; honey; Salt; pepper	None

5- Conclusion

Etude de faisabilité d'une classification
automatique d'articles

Descriptions
produits

Précision de clustering de 75 %

Images produits

Précision de clustering de 59 %

Classification basée sur les descriptions semble plus prometteuse

Classification supervisée à partir des images montre de bonne performances

Utilisation de réseau de neurones convolutif

Réseau de neurone hybride, réalisation d'une classification sur les descriptions et les images simultanément

Classification
supervisée

Précision de 85,71 % sans DA

Précision de 81,90 % avec DA

Question ?