# SOUTENANCE DU PROJET 6:
# CLASSIFIEZ AUTOMATIQUEMENT DES BIENS DE CONSOMMATION

Check KOUTAME

# PLAN

I/ Projet et donnés

- Mission & Description du projet
- Observations des données

II/ Traitement des données textuelles

- Pre-processing
- Classification et clustering

III/ Traitement des données visuelles

- Pre-processing
- Classification et clustering
- Mixte avec les données textuelles

IV/ Comparaison de tous les modèles: bilan

# I. PROJET ET DONNÉES
# MISSION & DESCRIPTION DU PROJET:


- L'objectif pour **Place de marché**: Moteur de classification
- **Problématique :**
  - Site e-commerce Flickpart a mis à disposition une base de données avec plusieurs articles
    - Vendeurs sur le Marketplace
    - Articles (description + nom+ images…)
      - **Attribution manuelle de la catégorie de l'article: sources d'erreurs**
  - **Comment automatiser la tâche de classification ?**
- **Mission :**
  - Réaliser une première étude de faisabilité d'un moteur de classification en se basant sur une image et une description pour l'automatisation de l'attribution de la catégorie de l'article
    - Améliorer l'interaction utilisateurs
    - Rendre plus fiable la catégorisation des articles

# I. PROJET ET DONNÉES: OBSERVATION DES DONNÉES

| Dataset |
| --- |
| uniq_id |
| crawl_timestamp |
| product_url |
| product_name |
| product_category_tree |
| pid |
| retail_price |
| discounted_price |
| image |
| is_FK_Advantage_product |
| description |
| product_rating |
| overall_rating |
| brand |
| product_specifications |

- **15 variables: 1050 observations**

## Textes

### Brand

- Sathiyas

### Product-specification

- [{"key"=>"Bra nd", "value"=>"Elegance"}, {"key"=>"Designed For", "value"=>"Door"},../..

### description

- Specifications of Sathiyas Cotton Bath Towel (3 Bath Towel, Red, Yellow, Blue) Bath Towel Features Machine

### Product_name

- Sathiyas Cotton Bath Towel

## Images

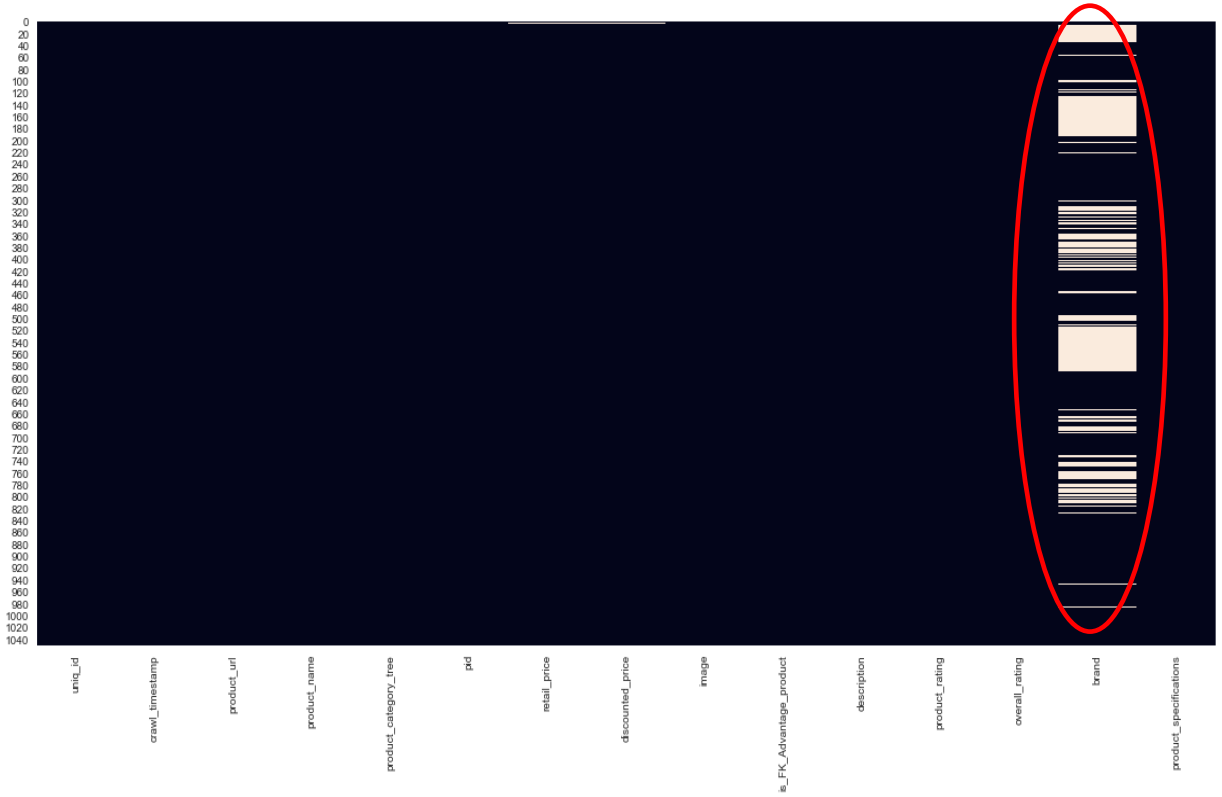
### Images


## Cibles

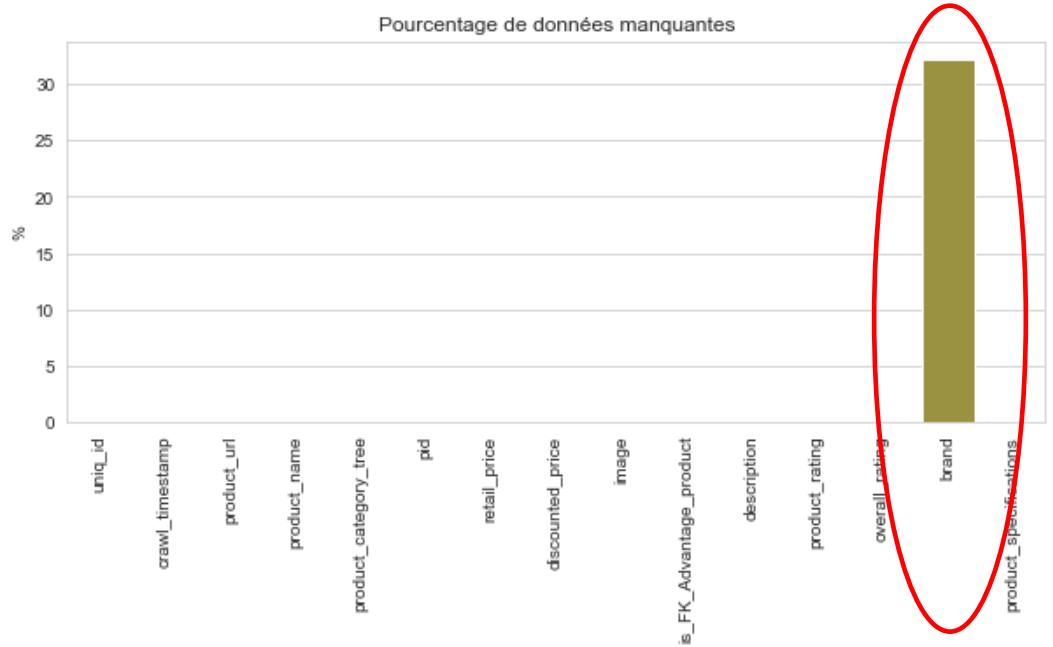
### product_category_tree

- "Baby Care >> Baby Bath & Skin >> Baby Bath Towels >> Sathiyas Baby Bath Towels >> Sathiyas Cotton Bath Towel (3 Bath Towel, Red, Y..."

# I. PROJET ET DONNÉES: OBSERVATION DES DONNÉES




Valeurs manquantes :341 NaN pour 15750 données (2.17 %)




La feature 'brand' contient plus de 30 % de données manquantes

| Jeu de données | Nb lignes | Nb variables | Nb nan | Nb var avec nan | Var avec nan |
|---|---|---|---|---|---|
| Data | 1050 | 15 | 341 | 4 | retail_price, discounted_price, brand, product_specifications |

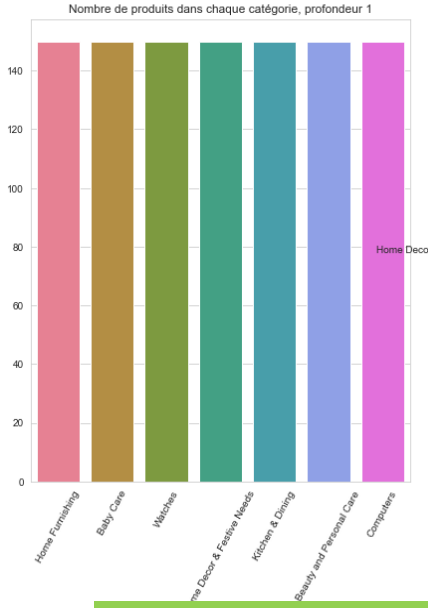
# I. PROJET ET DONNÉES: OBSERVATION DES DONNÉES: ANALYSE DES NIVEAUX

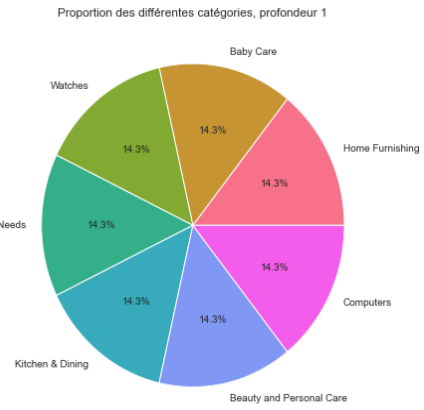
La feature « product_category_tree » représente l'arborescence complète d'un article.

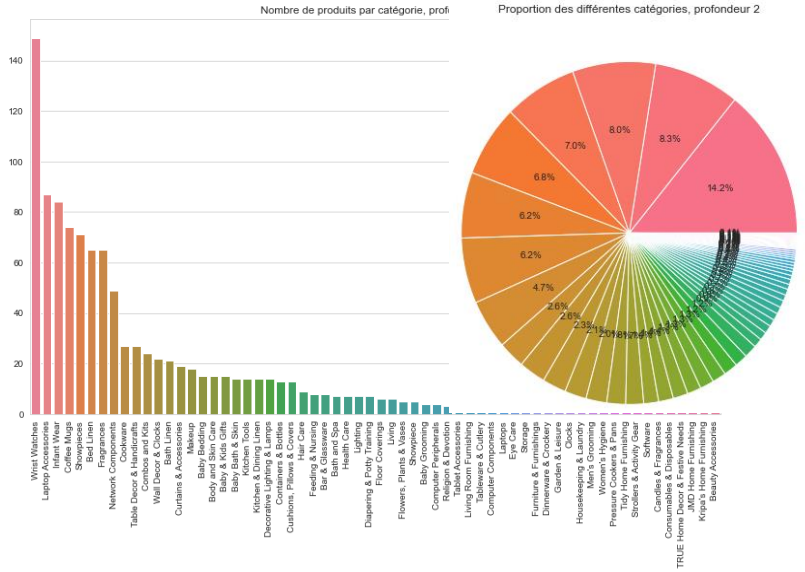
["Home Furnishing >> Curtains & Accessories >> Curtains >> Elegance Polyester Multicolor Abstract Eyelet Do..."] .... 1

.... 1

Niveau 1 | Niveau 2 | Niveau 3 | Niveau 4



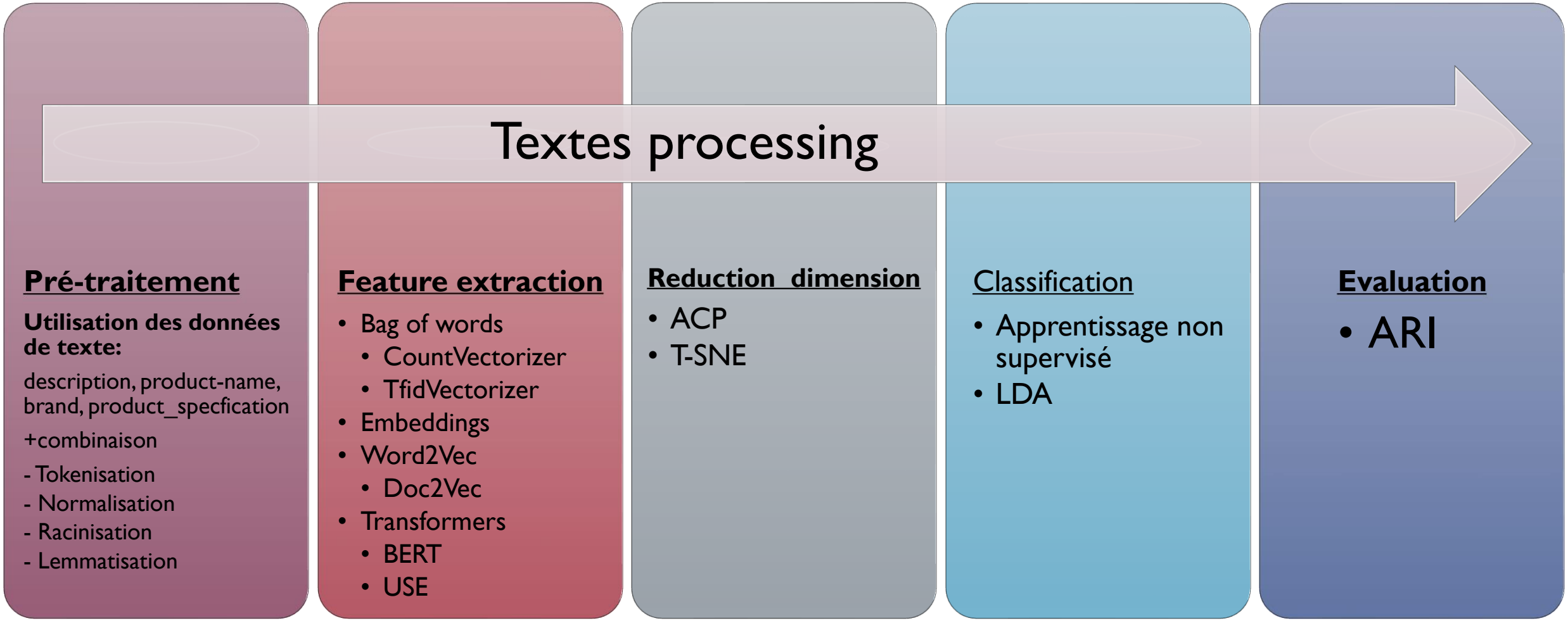

| Niveau | Nb catégories |
|---|---|
| 2 | 62 |
| 3 | 241 |
| 4 | 349 |
| 5 | 297 |
| 6 | 117 |
| 7 | 57 |

Niveau1: 7 catégories

Niveau 2: 62 catégories

Niveaux 3 à 7

# II. TRAITEMENT DES DONNÉES TEXTUELLES: PROCESSUS



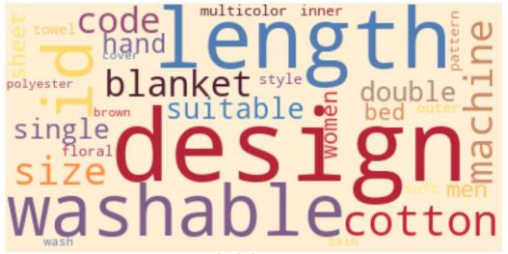
# II. TRAITEMENT DES DONNÉES TEXTUELLES: PROCESSUS: PRÉ-TRAITEMENT – EXEMPLE D'ÉTUDE

| Action | Contenu |
|---|---|
| **Texte original** | Wallmantra Large Vinyl Sticker Sticker (Pack of 1) Price: Rs. 1,896 Bring home this exclusive Piece of Wall Art to give your home a refreshing look it deserves ! Wall Decals are the latest trend, sweeping the world of interior design, as a quick and easy way to personalise and transform your home. We at Wallmantra use only the highest quality premium self-adhesive vinyl for our products to ensure you receive the best quality product. The sizes are in inches (Width x Height), rounded to the nearest inch. The size of the product is the final size that will be up on the walls . These are very easy to apply and remove. You can apply these yourself. We send ready to apply Decal/Sticker with 2 practice items with instruction manual. You can also see the how to apply video. These are not re-usable. Contact Wallmantra for more options Bring home this exclusive Piece of Wall Art to give your home a refreshing look it deserves ! Wall Decals are the latest trend, sweeping the world of interior design, as a quick and easy way to personalise and transform your home. We at Wallmantra use only the highest quality premium self-adhesive vinyl for our products to ensure you receive the best quality product. The sizes are in inches (Width x Height), rounded to the nearest inch. The size of the product is the final size that will be up on the walls . These are very easy to apply and remove. You can apply these yourself. We send ready to apply Decal/Sticker with 2 practice items with instruction manual. You can also see the how to apply video. These are not re-usable. Contact Wallmantra for more options |
| **Tokenisation** | ['Wallmantra', 'Large', 'Vinyl', 'Sticker', 'Sticker', '(', 'Pack', 'of', '1', ')', 'Price', ':', 'Rs', '.', '1,896', 'Bring', 'home', 'this', 'exclusive', 'Piece', 'of', 'Wall', 'Art', 'to', 'give', 'your', 'home', 'a', 'refreshing', 'look', 'it', 'deserves', '!', 'Wall', 'Decals', 'are', 'the', 'latest', 'trend', ',', 'sweeping', 'the', 'world', 'of', 'interior', 'design', ',', 'as', 'a', 'quick', 'and', 'easy', 'way', 'to', 'personalise', 'and', 'transform', 'your', 'home', '.', 'We', 'at', 'Wallmantra', 'use', 'only', 'the', 'highest', 'quality', 'premium', 'self-adhesive', 'vinyl', 'for', 'our', 'products', 'to', 'ensure', 'you', 'receive', 'the', 'best', 'quality', 'product', '.', 'The', 'sizes', 'are', 'in', 'inches', '(', 'Width', 'x', 'Height', ')', ',', 'rounded', 'to', 'the', 'nearest', 'inch', '.', 'The', 'size', 'of', 'the', 'product', 'is', 'the', 'final', 'size', 'that', 'will', 'be', 'up', 'on', 'the', 'walls', '.', 'These', 'are', 'very', 'easy', 'to', 'apply', 'and', 'remove', '.', 'You', 'can', 'apply', 'these', 'yourself', '.', 'We', 'send', 'ready', 'to', 'apply', 'Decal/Sticker', 'with', '2', 'practice', 'items', 'with', 'instruction', 'manual', '.', 'You', 'can', 'also', 'see', 'the', 'how', 'to', 'apply', 'video', '.', 'These', 'are', 'not', 're-usable', '.', 'Contact', 'Wallmantra', 'for', 'more', 'options', 'Bring', 'home', 'this', 'exclusive', 'Piece', 'of', 'Wall', 'Art', 'to', 'give', 'your', 'home', 'a', 'refreshing', 'look', 'it', 'deserves', '!', 'Wall', 'Decals', 'are', 'the', 'latest', 'trend', ',', 'sweeping', 'the', 'world', 'of', 'interior', 'design', ',', 'as', 'a', 'quick', 'and', 'easy', 'way', 'to', 'personalise', 'and', 'transform', 'your', 'home', '.', 'We', 'at', 'Wallmantra', 'use', 'only', 'the', 'highest', 'quality', 'premium', 'self-adhesive', 'vinyl', 'for', 'our', 'products', 'to', 'ensure', 'you', 'receive', 'the', 'best', 'quality', 'product', '.', 'The', 'sizes', 'are', 'in', 'inches', '(', 'Width', 'x', 'Height', ')', ',', 'rounded', 'to', 'the', 'nearest', 'inch', '.', 'The', 'size', 'of', 'the', 'product', 'is', 'the', 'final', 'size', 'that', 'will', 'be', 'up', 'on', 'the', 'walls', '.', 'These', 'are', 'very', 'easy', 'to', 'apply', 'and', 'remove', '.', 'You', 'can', 'apply', 'these', 'yourself', '.', 'We', 'send', 'ready', 'to', 'apply', 'Decal/Sticker', 'with', '2', 'practice', 'items', 'with', 'instruction', 'manual', '.', 'You', 'can', 'also', 'see', 'the', 'how', 'to', 'apply', 'video', '.', 'These', 'are', 'not', 're-usable', '.', 'Contact', 'Wallmantra', 'for', 'more', 'options'] |
| **Normalisation TextHero:**<br>- **Minuscule**<br>- **Suppr ponctuation**<br>- **Supp stop words**<br>- **Supp fré-rares** | wallmantra large vinyl sticker sticker pack bring home exclusive piece wall art give home refreshing look wall decals latest trend world interior design quick easy way transform home wallmantra use highest quality premium self adhesive vinyl ensure best quality product sizes inches width height rounded nearest inch size product size walls easy apply remove apply send ready apply decal sticker items manual also see apply video usable contact wallmantra options bring home exclusive piece wall art give home refreshing look wall decals latest trend world interior design quick easy way transform home wallmantra use highest quality premium self adhesive vinyl ensure best quality product sizes inches width height rounded nearest inch size product size walls easy apply remove apply send ready apply decal sticker items manual also see apply video usable contact wallmantra options |
| **Racinisation (radical)** | wallmantra larg vinyl sticker sticker pack bring home exclus piec wall art give home refresh look wall decal latest trend world interior design quick easi way transform home wallmantra use highest qualiti premium self adhes vinyl ensur best qualiti product size inch width height round nearest inch size product size wall easi appli remov appli send readi appli decal sticker item manual also see appli video usabl contact wallmantra option bring home exclus piec wall art give home refresh look wall decal latest trend world interior design quick easi way transform home wallmantra use highest qualiti premium self adhes vinyl ensur best qualiti product size inch width height round nearest inch size product size wall easi appli remov appli send readi appli decal sticker item manual also see appli video usabl contact wallmantra option |
| **Lemmatisation (canonique)** | wallmantra large vinyl sticker sticker pack bring home exclusive piece wall art give home refreshing look wall decal latest trend world interior design quick easy way transform home wallmantra use highest quality premium self adhesive vinyl ensure best quality product size inch width height rounded nearest inch size product size wall easy apply remove apply send ready apply decal sticker item manual also see apply video usable contact wallmantra option bring home exclusive piece wall art give home refreshing look wall decal latest trend world interior design quick easy way transform home wallmantra use highest quality premium self adhesive vinyl ensure best quality product size inch width height rounded nearest inch size product size wall easy apply remove apply send ready apply decal sticker item manual also see apply video usable contact wallmantra option |

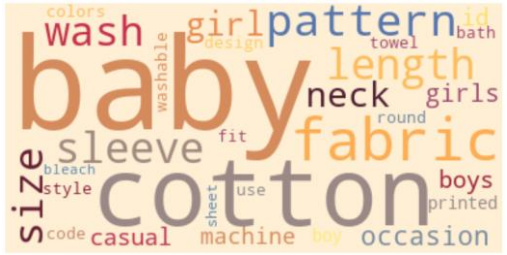
# II. TRAITEMENT DES DONNÉES TEXTUELLES: PROCESSUS: PRÉ-TRAITEMENT – FRÉQUENCE DES MOTS

product_specification

Visualisation des mots les plus fréquents pour chacune des 7 catégories après le pré-traitement de la variable

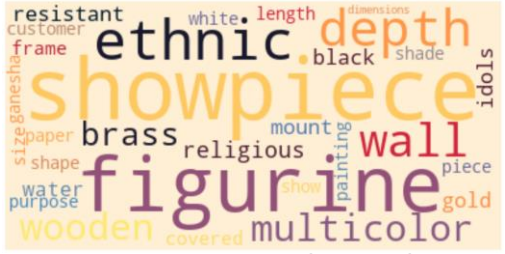



Home Furnishing




Baby Care


Watches




Home Decor & Festive Needs

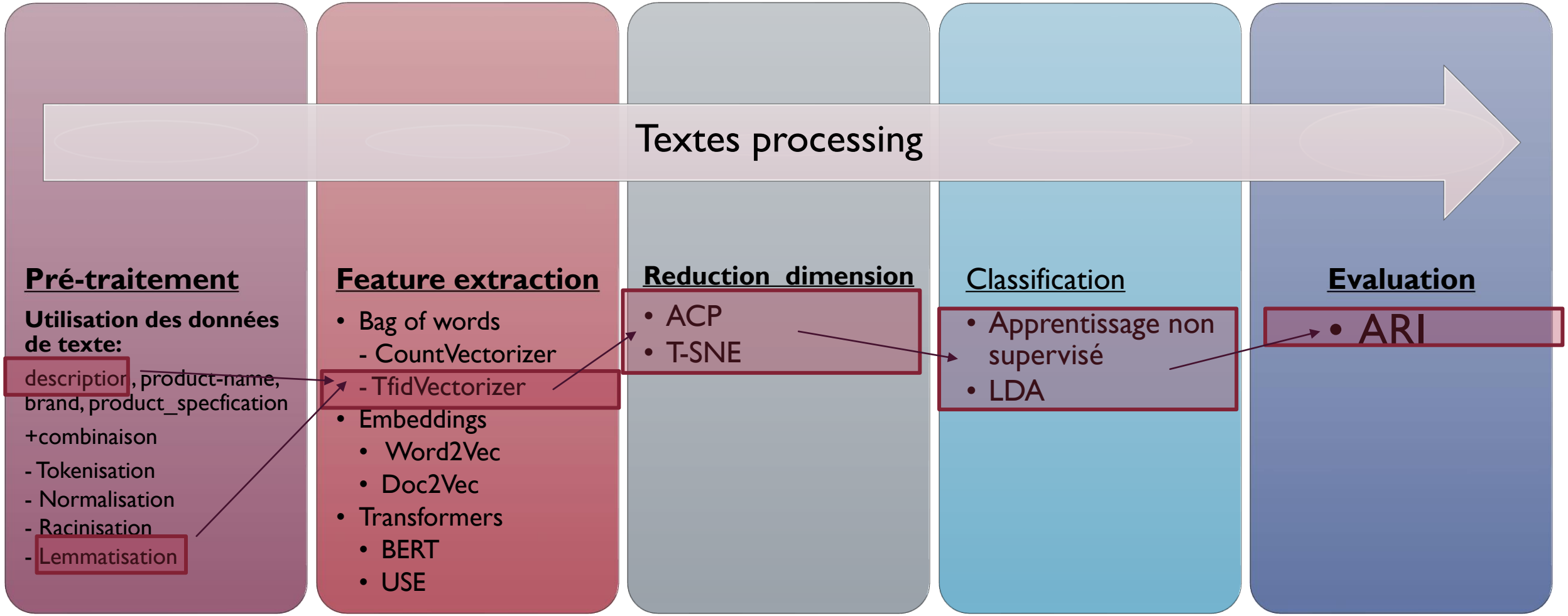

Kitchen & Dining


Beauty and Personal Care


Computers

| | Mot | Frequence | %_frequence |
|---|---|---|---|
| 0 | product | 1244 | 33.56 |
| 6 | type | 1160 | 31.29 |
| 1 | specification | 1049 | 28.30 |
| 16 | color | 864 | 23.31 |
| 8 | model | 837 | 22.58 |
| 23 | package | 810 | 21.85 |
| 22 | sales | 808 | 21.80 |
| 20 | number | 806 | 21.74 |
| 26 | material | 800 | 21.58 |
| 19 | cm | 780 | 21.04 |
| 113 | warranty | 744 | 20.07 |

# II. TRAITEMENT DES DONNÉES TEXTUELLES: PROCESSUS: EXEMPLE DE MODÈLE –TF-IDF



# II. TRAITEMENT DES DONNÉES TEXTUELLES: PROCESSUS: WORD EMBEDDING – EX DE MODÈLE: TF-IDF

- Une transformation tf-idf (term frequency-inverse document frequency), permet de pondérer les fréquences d'apparition des termes par leur nombre d'occurrences dans l'ensemble des documents.
- La pondération tf-idf permet de contrebalancer l'importance d'un mot utilisé très fréquemment dans tous les documents du corpus par rapport aux termes plus spécifiques à certains documents.
- - TF-IDF est un produit de deux parties :
  - TF (Term Frequency) - Elle est définie comme le nombre de fois qu'un mot apparaît dans une phrase donnée.
  - IDF (Inverse Document Frequency) - Il est défini comme le logarithme à la base e du nombre total de documents divisé par les documents dans lesquels le mot apparaît.
- Une mesure possible est la suivante (Jones 1973) :

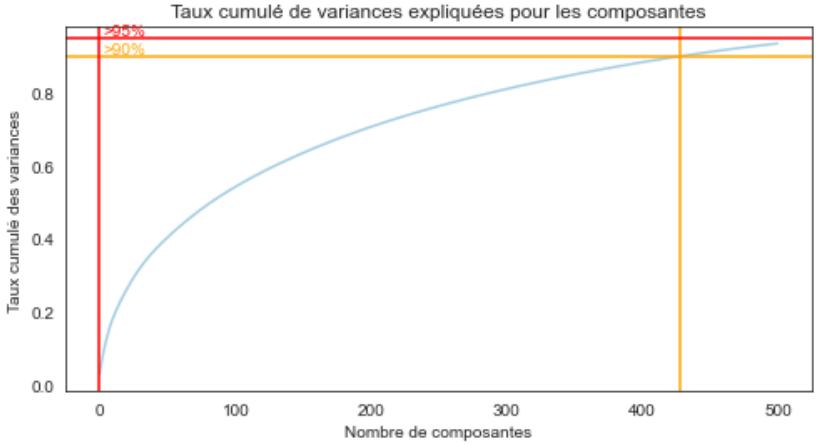
$$w_{i,j} = tf_{i,j} \times \log\left(\frac{N}{df_i}\right)$$

$tf_{ij}$ = number of occurrences of $i$ in $j$
$df_i$ = number of documents containing $i$
$N$ = total number of documents

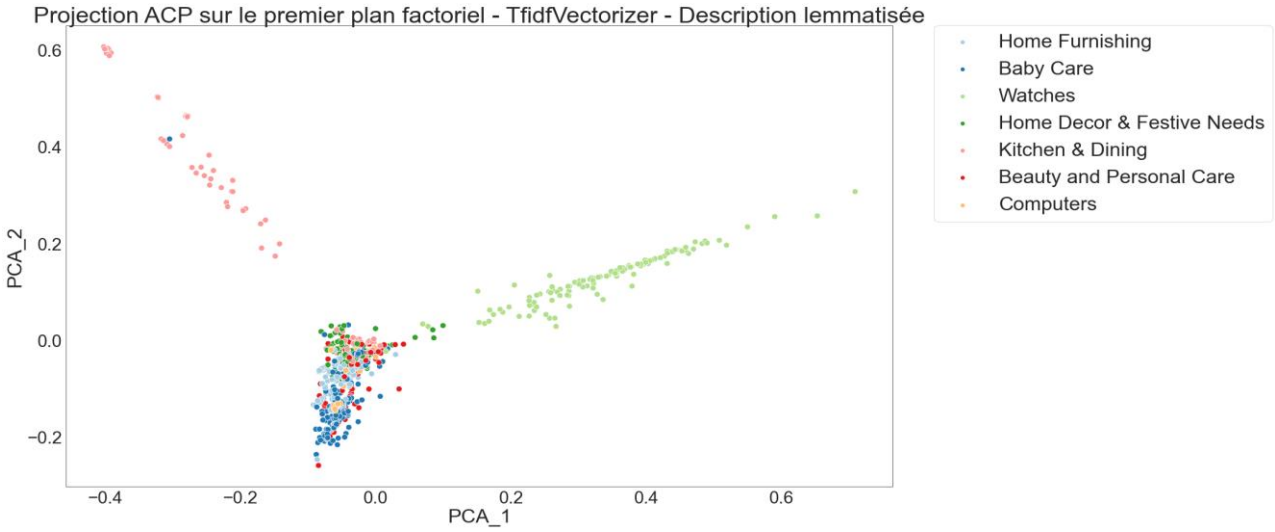
- - Cette transformation permet de produire des premiers éléments d'analyses, au premier titre duquel apparaissent les nuages de mots (wordcloud) qui représentent la fréquence relative des termes et qui donnent une première idée des significations contenues dans le corpus.

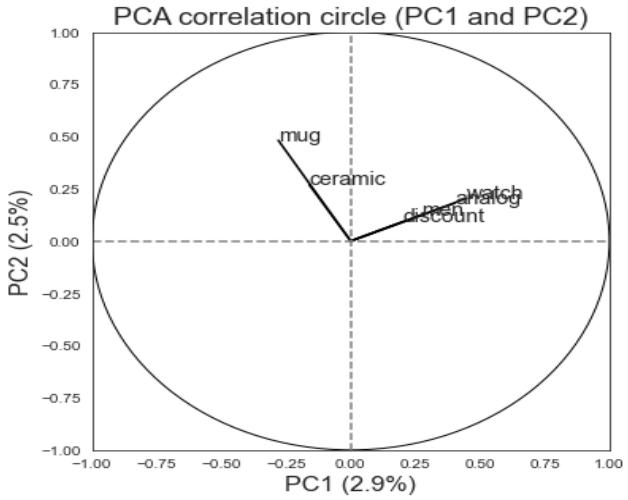
# II. TRAITEMENT DES DONNÉES TEXTUELLES: PROCESSUS: WORD EMBEDDING – EX DE MODÈLE: WORD2VEC

## Réduction de dimension: PCA




Nombre de composantes expliquant 90% de la variance : 427

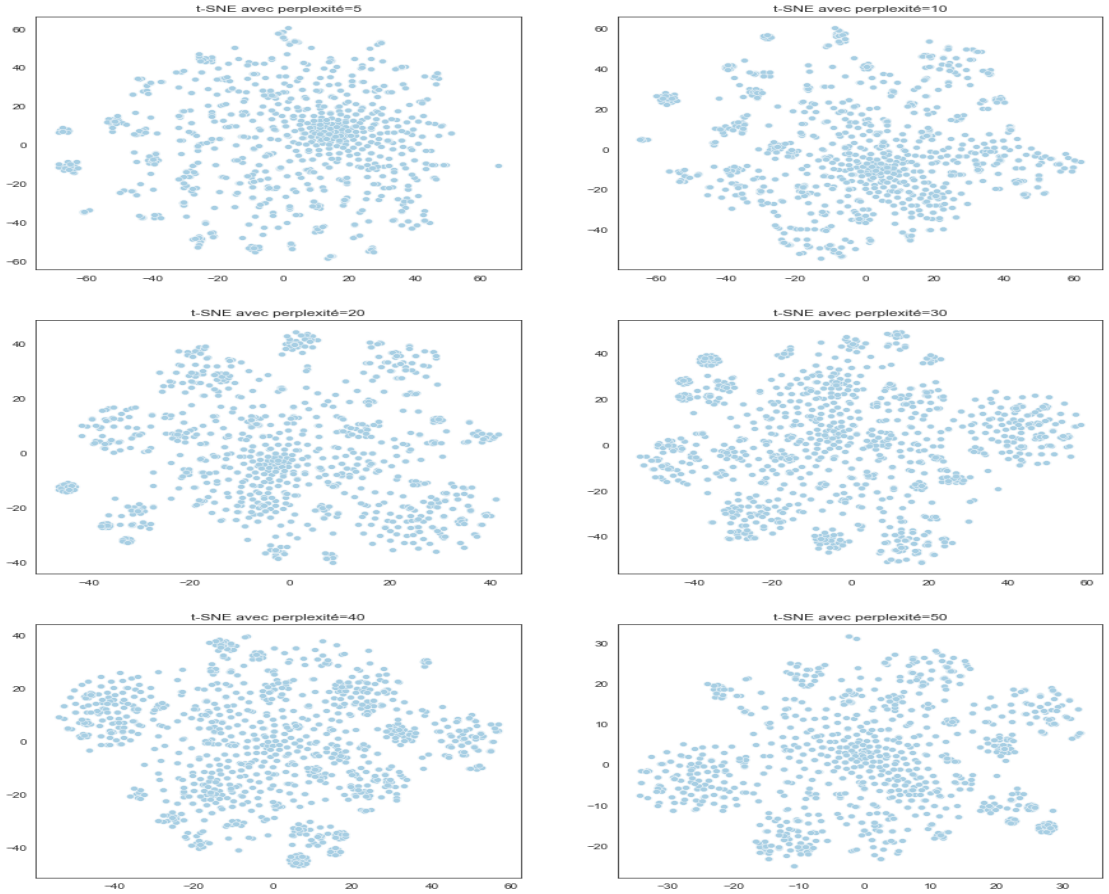


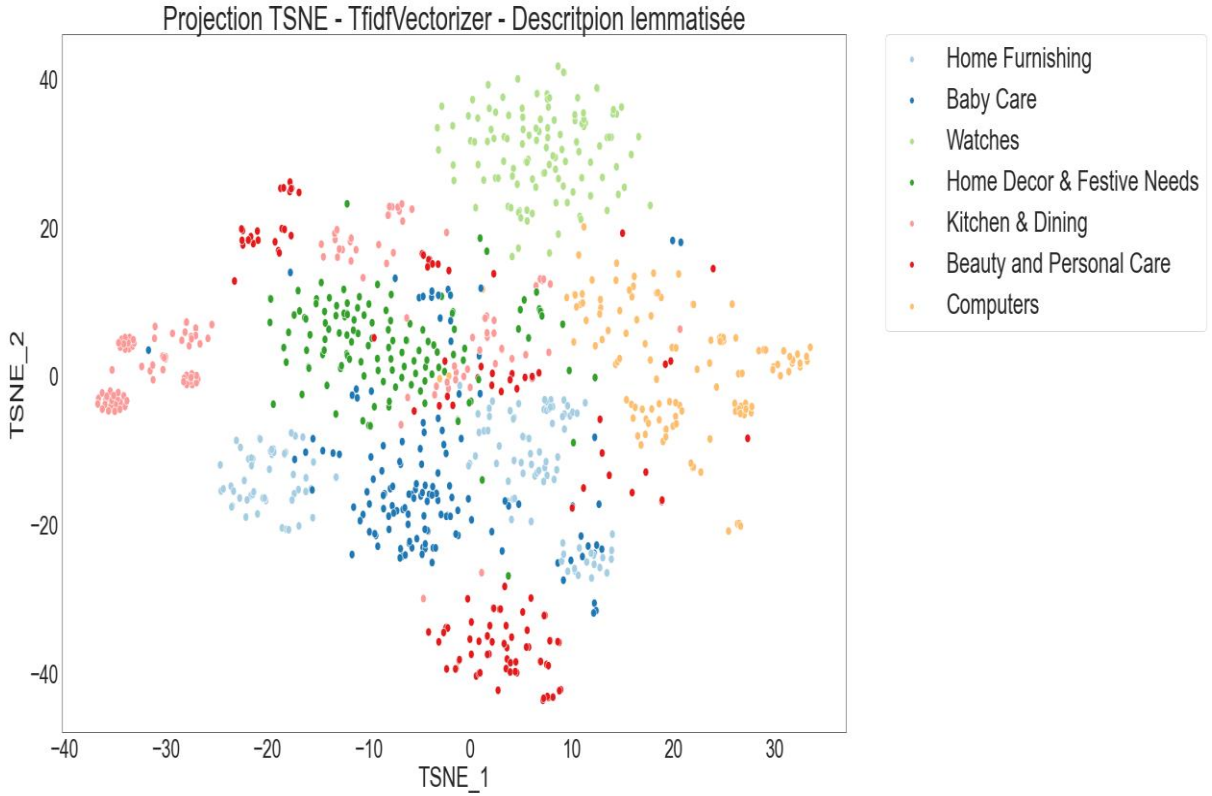




•La **première composante** (watch, men, discount) représente la catégorie **Watches** :
•La **seconde composante** (ceramic, mug) représente la catégorie : **Kitchen & Dining**

# II. TRAITEMENT DES DONNÉES TEXTUELLES: PROCESSUS: WORD EMBEDDING – EX DE MODÈLE: TF-IDF

Réduction de dimension: t-SNE

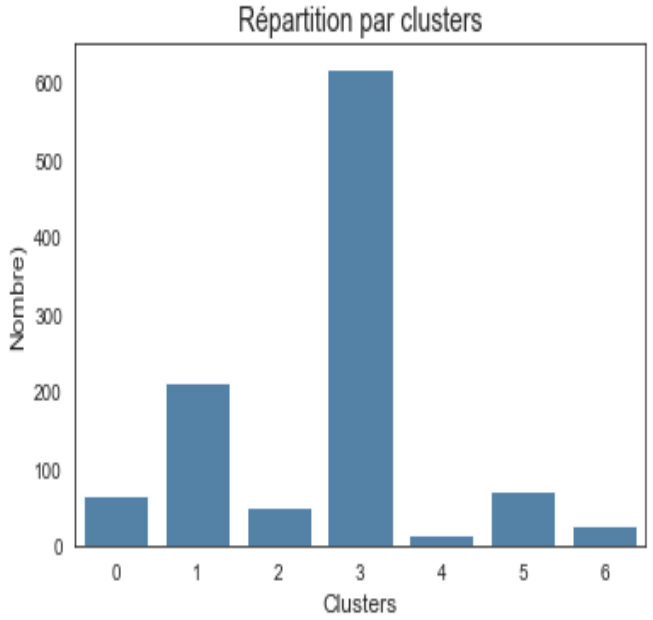


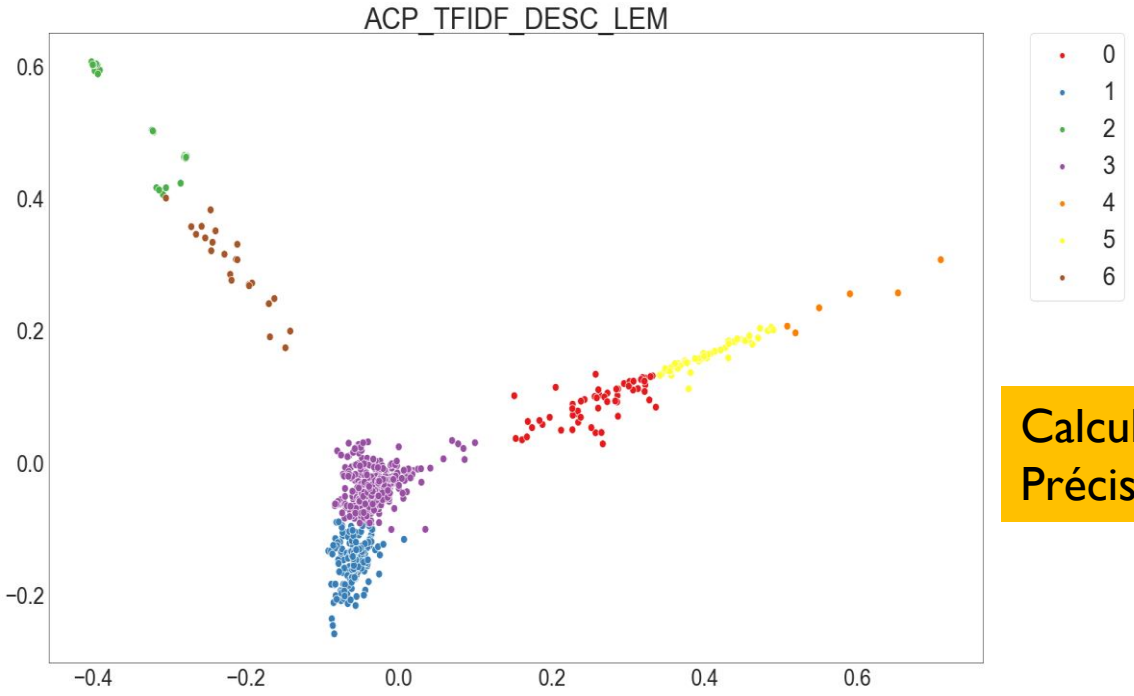




•groupes plus homogènes pour perplexité=40
•on laissera les paramètres learning_rate et n_iter par défaut (1000) et le nombre de composants par défaut (2)

# II. TRAITEMENT DES DONNÉES TEXTUELLES: PROCESSUS: WORD EMBEDDING – EX DE MODÈLE: TF-IDF

Clusterisation



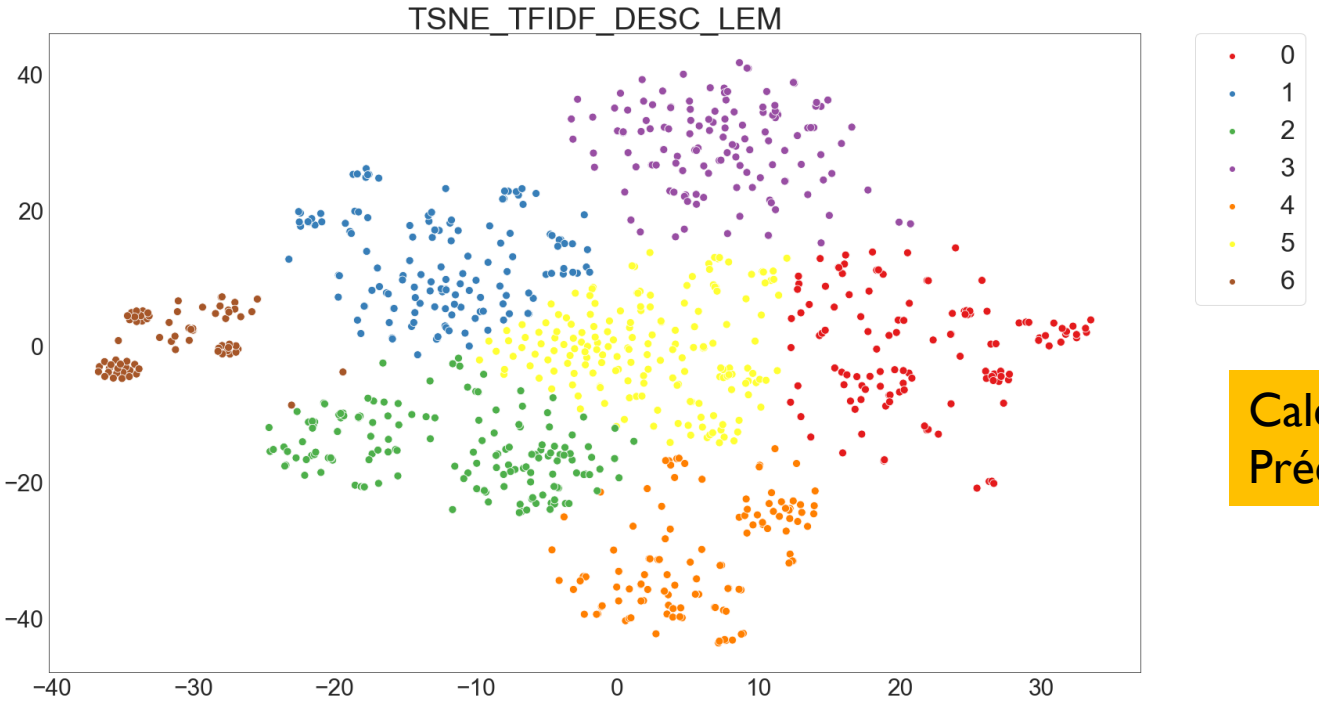




Calcul du ARI: 0.149083
Précision: 64%

•Répartition inégale (2 gros clusters), les vraies catégories sont mélangées dans les clusters, certains clusters sont presque vides, le score ARI est faible (mais meilleur résultat qu'avec ACP CountVectorizer).
•Cette combinaison ACP TfidfVectorizer+ description lemmatisée parvient à retrouver les 7 catégories comme le souhaite notre client.

# II. TRAITEMENT DES DONNÉES TEXTUELLES: PROCESSUS: WORD EMBEDDING – EX DE MODÈLE: TF-IDF

Clusterisation


Calcul du ARI: 0.42
Précision: 64%

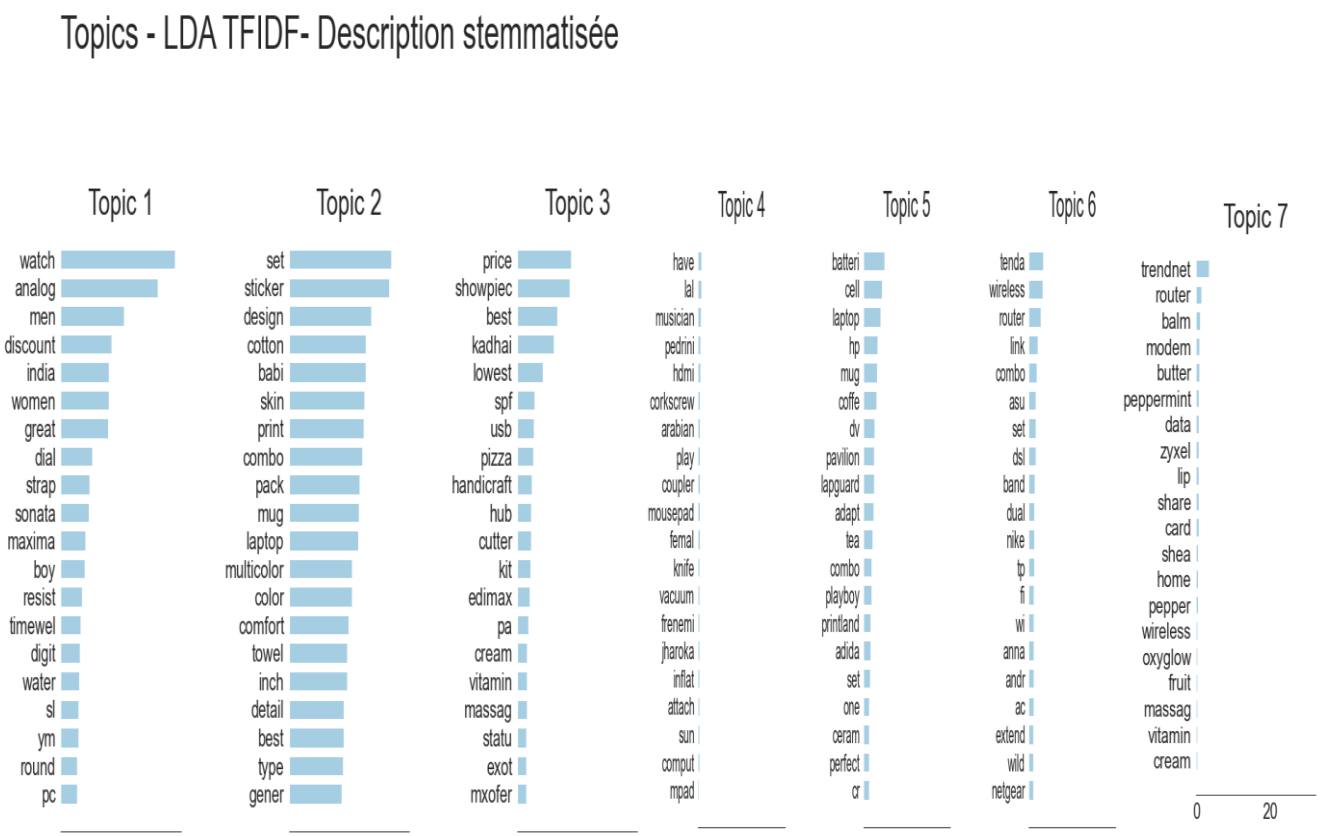
•Cette combinaison t-SNE TfidfVectorizer+ description lemmatisée parvient à retrouver les 7 catégories comme le souhaite notre client.

# II. TRAITEMENT DES DONNÉES TEXTUELLES: PROCESSUS: MODÉLISATION DE TOPIC - LDA

Explication Modèle

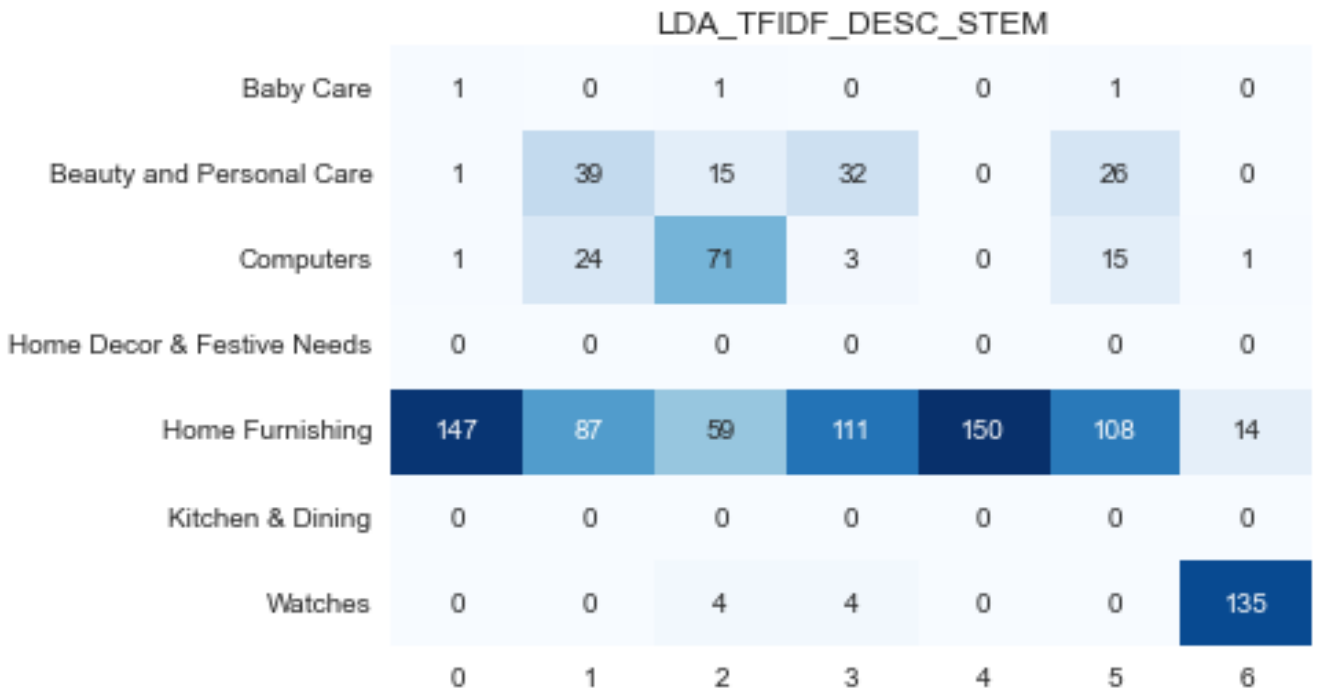
- LDA - Latent Dirichlet Allocation : est un algorithme d'apprentissage non supervisé utilisé pour découvrir les thèmes présents dans un corpus.
- LDA est basé sur la modélisation graphique probabiliste.
- L'algorithme prend en entrée une matrice de sac de mots (c'est-à-dire que chaque document est représenté par une ligne, chaque colonne contenant le nombre de mots dans le corpus). L'objectif est ensuite de produire deux matrices plus petites, une matrice document/sujet et une matrice mot/sujet qui, une fois multipliées ensemble, reproduisent la matrice du sac de mots avec l'erreur la plus faible.

# II. TRAITEMENT DES DONNÉES TEXTUELLES: PROCESSUS: MODÉLISATION DE TOPIC - LDA




- On retrouve uniquement 4 des 7 catégories :
  - Computers
  - Home Furnishing
  - Baby Care
  - Home Decor & Festive Needs
- Il manque les catégories :
  - Kitchen & Dining
  - Beauty and Personal Care
  - et Watches.
- LDA TfidfVectorizer description stemmatisée ne parvient pas à catégoriser les produits en 7 catégories comme le souhaite le client.

# II. TRAITEMENT DES DONNÉES TEXTUELLES: PROCESSUS: MODÉLISATION DE TOPIC - LDA

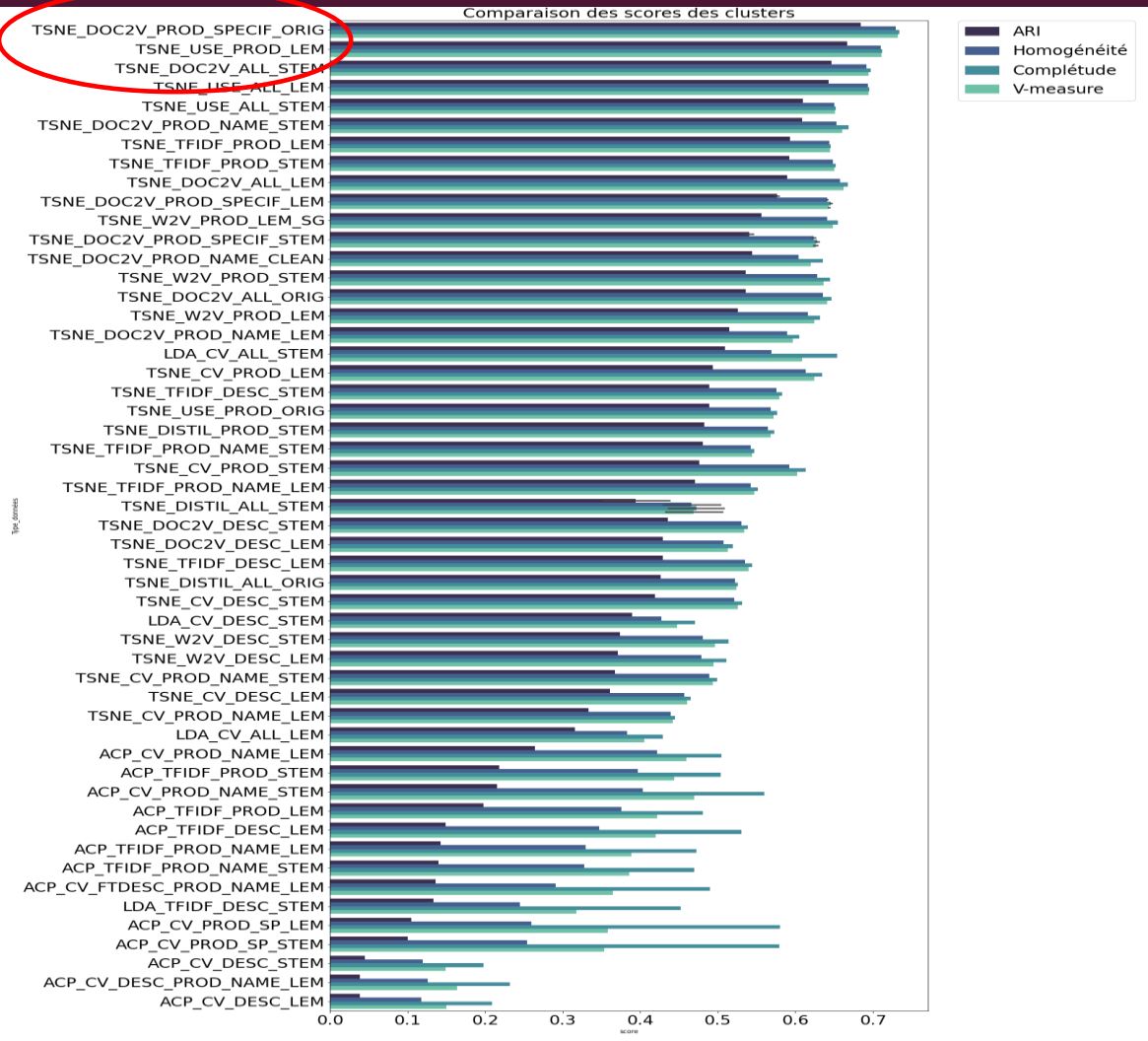


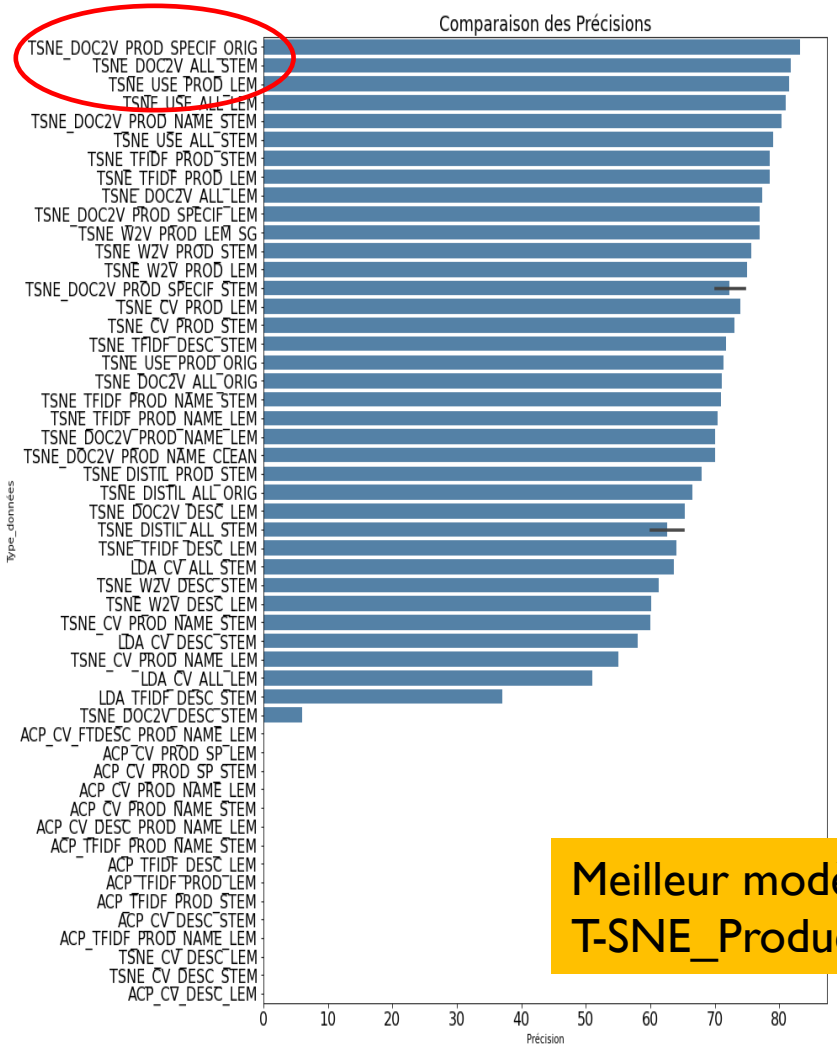

'Précision: 37.71%

ARI: 0.133662

- On retrouve uniquement 4 des 7 catégories :
  - Computers
  - Home Furnishing
  - Baby Care
  - Home Decor & Festive Needs
- Il manque les catégories :
  - Kitchen & Dining
  - Beauty and Personal Care
  - et Watches.
- LDA TfidfVectorizer description stemmatisée ne parvient pas à catégoriser les produits en 7 catégories comme le souhaite le client.

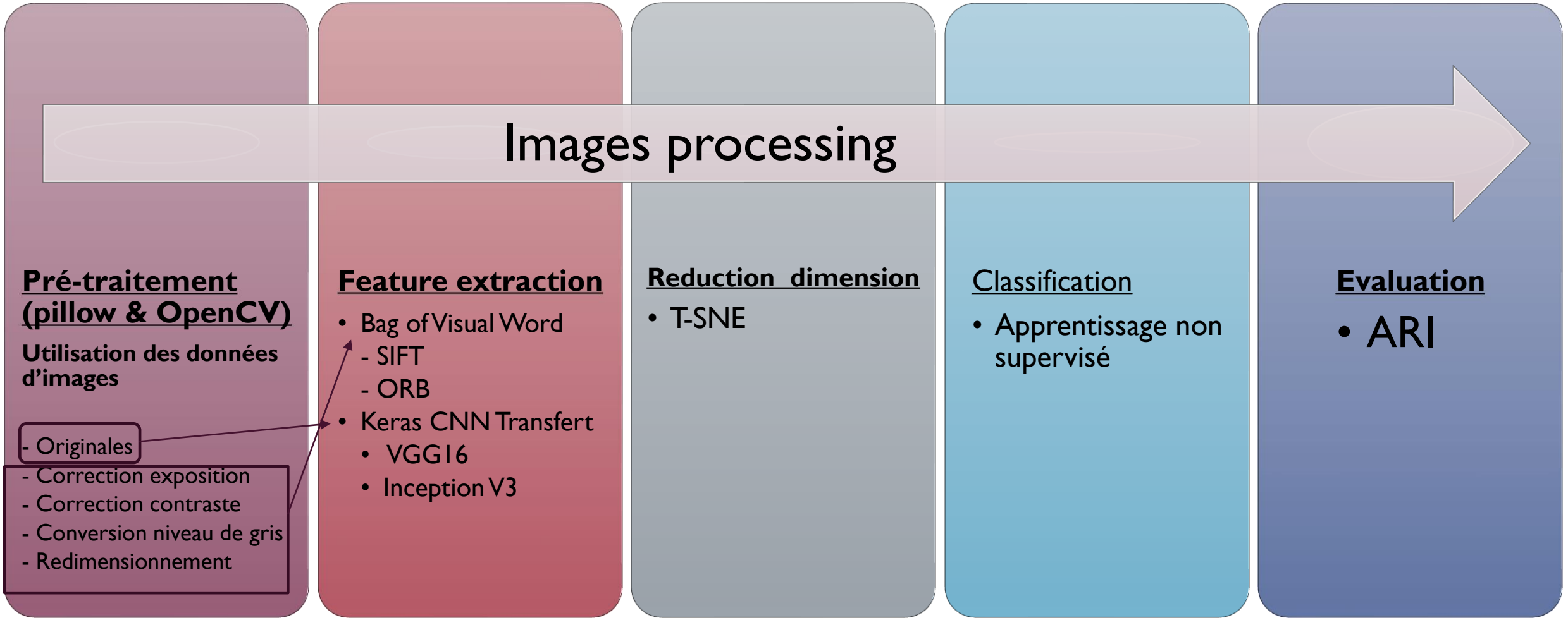
# II. TRAITEMENT DES DONNÉES TEXTUELLES: COMPARAISON DE TOUS LES MODÈLES



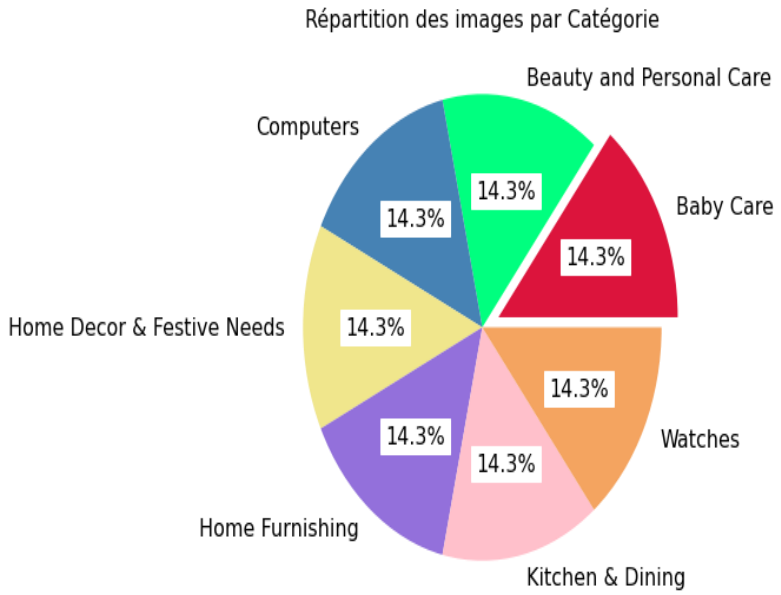




Meilleur modèle:
T-SNE_Produc_specif_origine

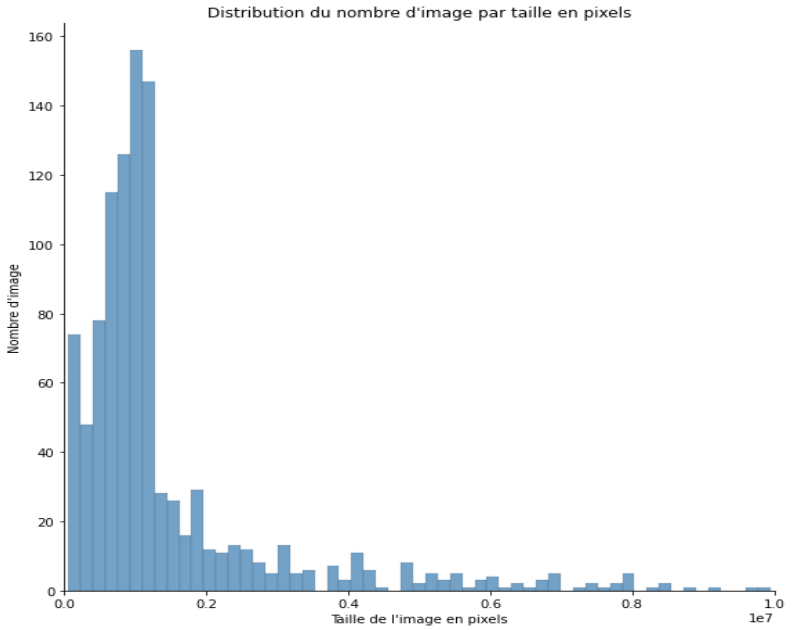
# III. TRAITEMENT DES DONNÉES VISUELLES: PROCESSUS



# III. TRAITEMENT DES DONNÉES VISUELLES: ANALYSE EXPLORATOIRE

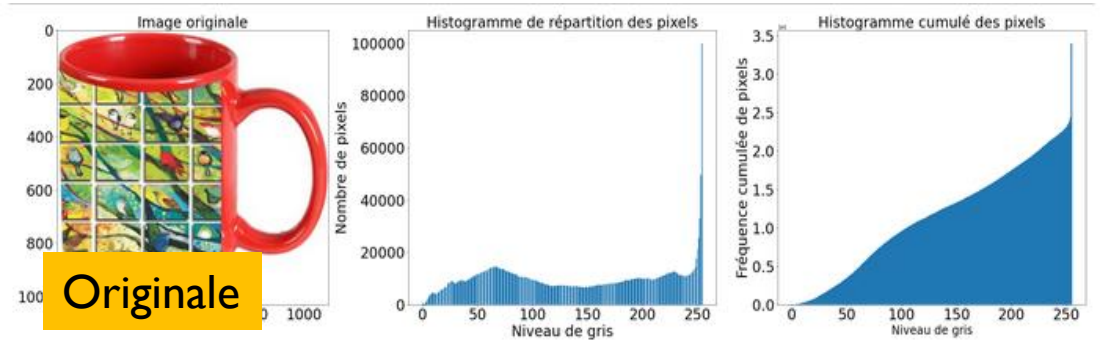



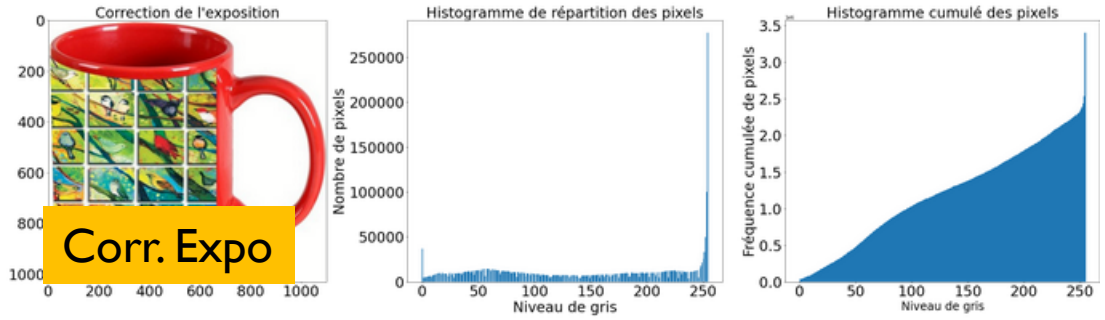


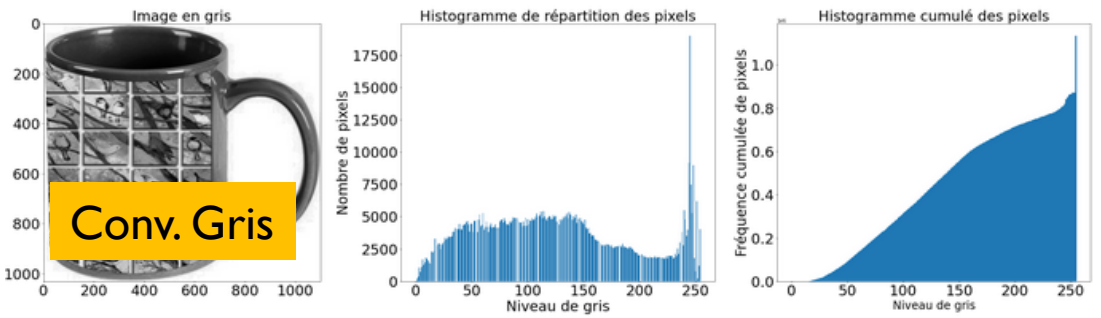


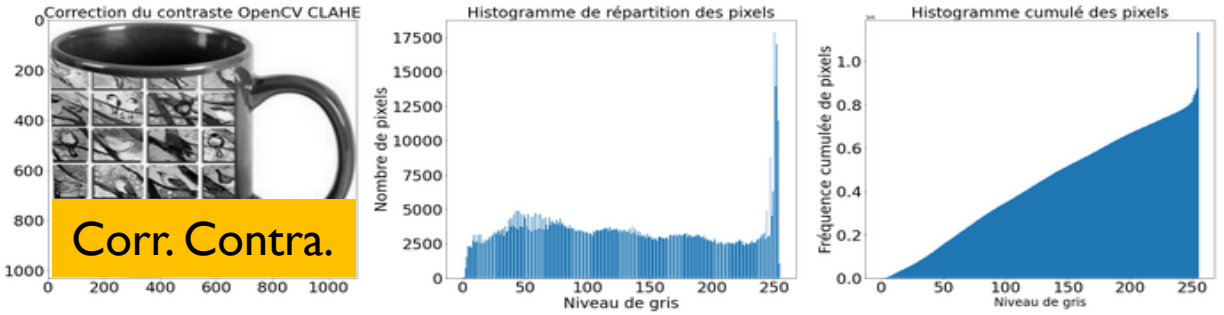

- Les catégories sont facilement identifiables ou non.
- La classification nous confirmera, ou non, ce point de vue.
- Les catégories sont bien équilibrées
- images sont de tailles très inégales, il convient donc de bien les uniformiser en les redimensionnant pour pouvoir utiliser les réseaux de neurones et les CNN de transfert learning.

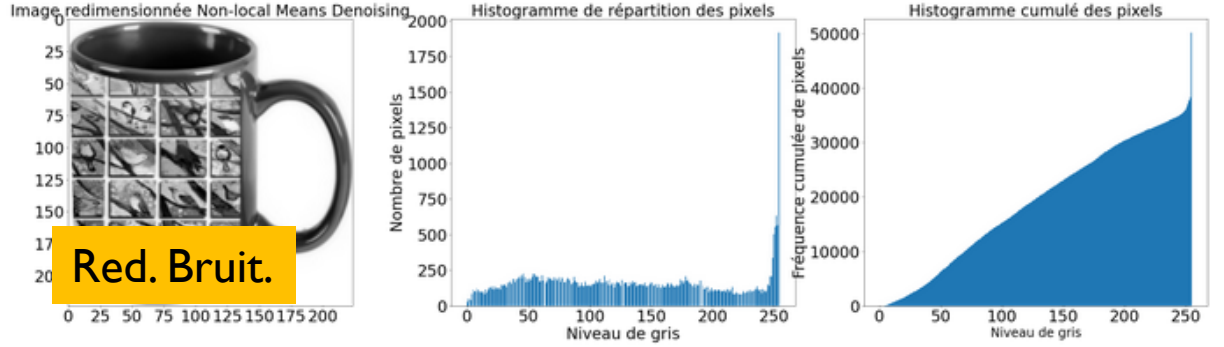
# III. TRAITEMENT DES DONNÉES VISUELLES: PRÉ-TRAITEMENT AVEC OPENCV & PILLOW

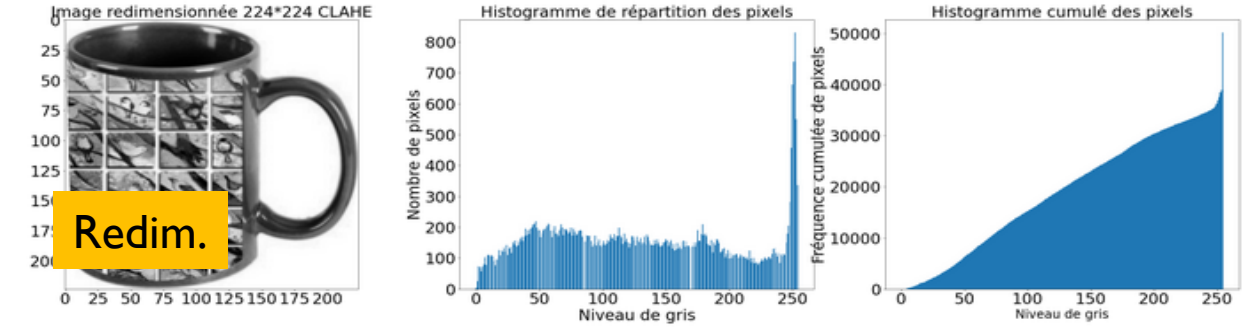


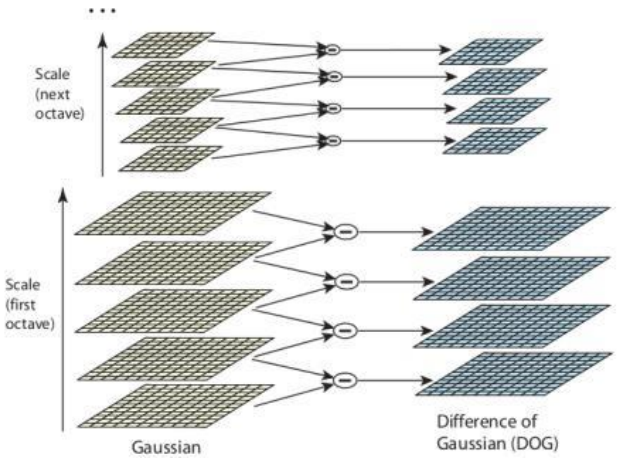



# III. TRAITEMENT DES DONNÉES VISUELLES: BVW – SIFT & ORB

SIFT

- SIFT (Scale Invariant Feature Transform), cette méthode, développée en 1999 et très populaire dans le domaine de la vision par ordinateur, permet d'extraire des features (ou points d'intérêt) de l'image et de calculer leurs descripteurs.
- L'algorithme SIFT se divise en plusieurs étapes :
  - Détection : création de l'espace des échelles, calcul des "DoG" (Différence of Gaussian), localisation des points d'intérêt.
  - Description : assignation d'orientation, création des descripteurs

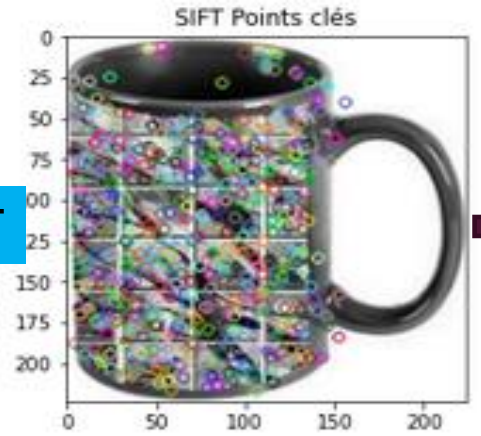



ORB

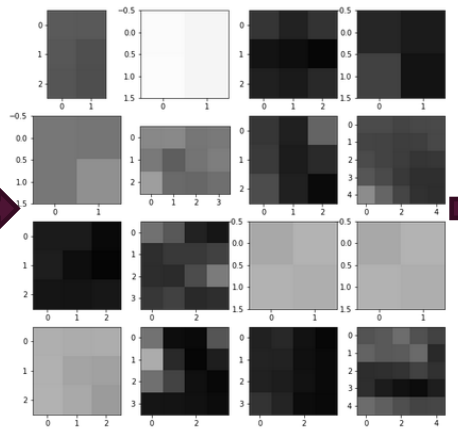
•**ORB** (Oriented FAST and Rotated BRIEF) a été développé dans les laboratoires d'OpenCV par Ethan Rublee, Vincent Rabaud, Kurt Konolige et Gary R. Bradski en 2011, comme une alternative efficace et viable à SIFT et SURF.
•ORB a été conçu principalement parce que SIFT et SURF sont des algorithmes brevetés. ORB, cependant, est libre d'utilisation.
•ORB est aussi performant que SIFT pour la détection des caractéristiques (et est meilleur que SURF) tout en étant presque deux ordres de grandeur plus rapide.
•ORB s'appuie sur le célèbre détecteur de points clés FAST et le descripteur BRIEF. Ces deux techniques sont intéressantes en raison de leurs bonnes performances et de leur faible coût.
•Les principales contributions d'ORB sont les suivantes :
•L'ajout d'une composante d'orientation rapide et précise à FAST.
•Le calcul efficace des caractéristiques BRIEF orientées
•L'analyse de la variance et de la corrélation des caractéristiques BRIEF orientées
•Une méthode d'apprentissage pour la décorrélation des caractéristiques BRIEF sous invariance rotationnelle, conduisant à de meilleures performances dans les applications de type "nearest-neighbor"

# III. TRAITEMENT DES DONNÉES VISUELLES: BVW – SIFT & ORB

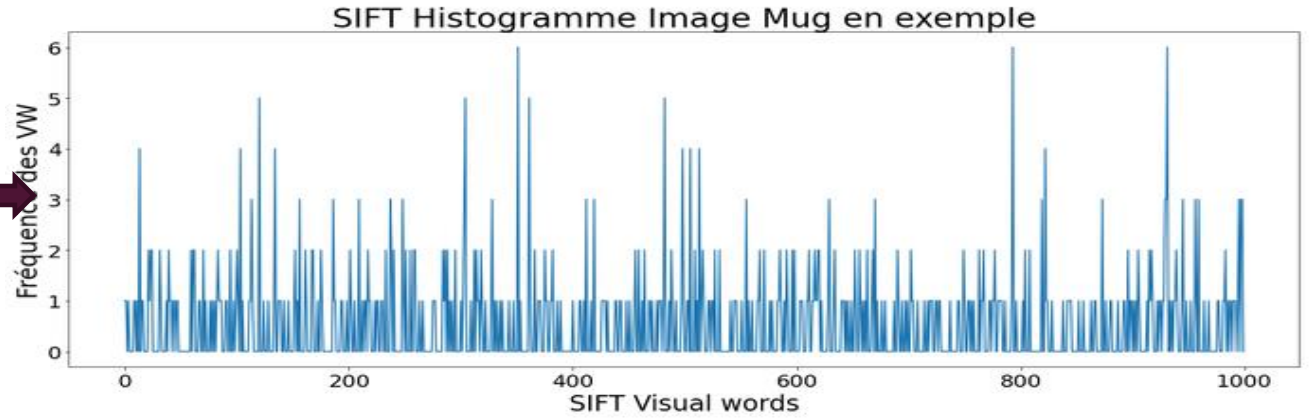
SIFT



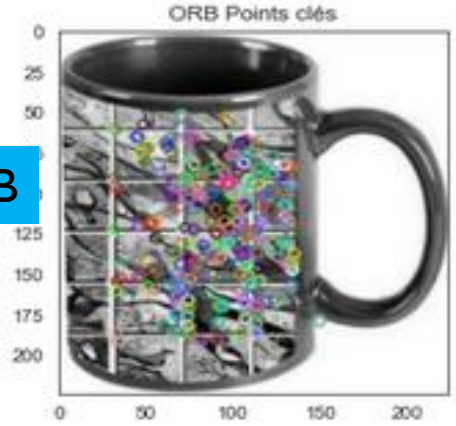

334110



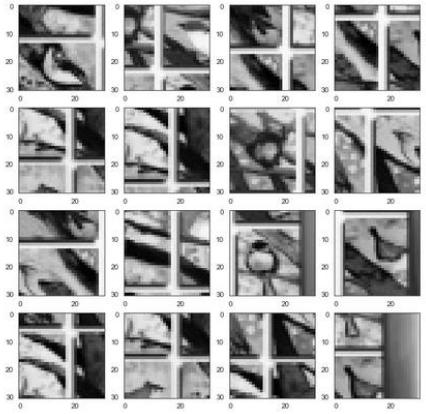
1000 clusters

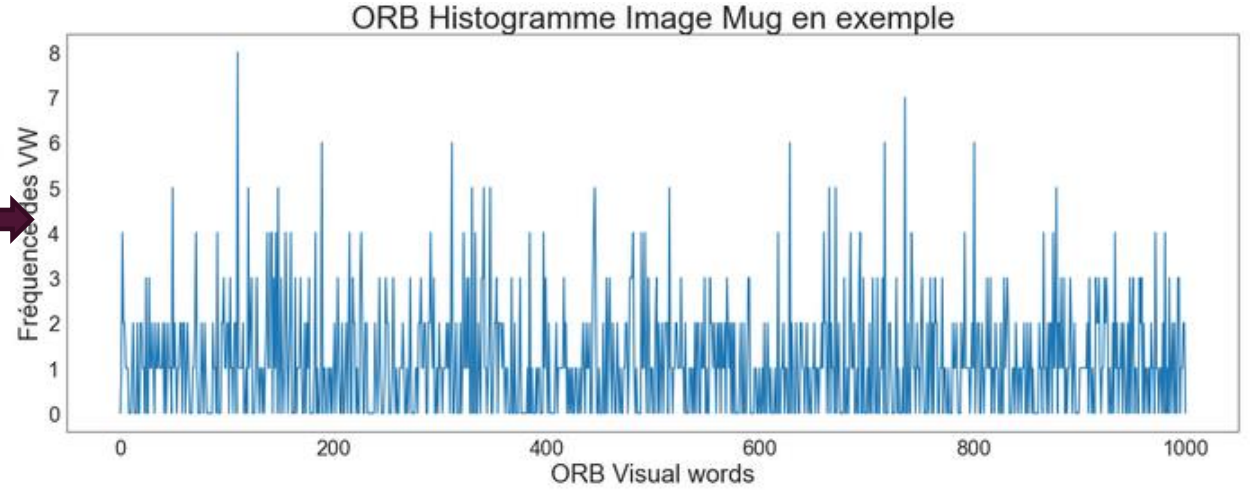



ORB



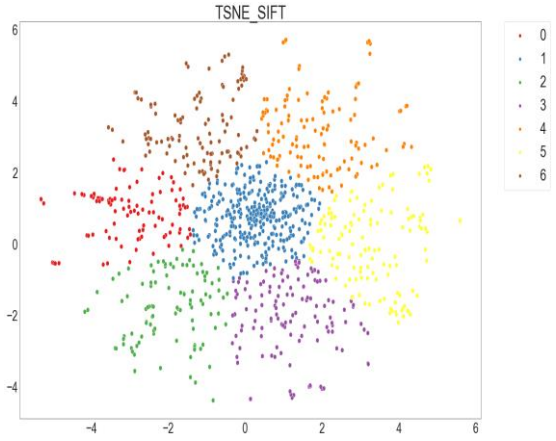

735579

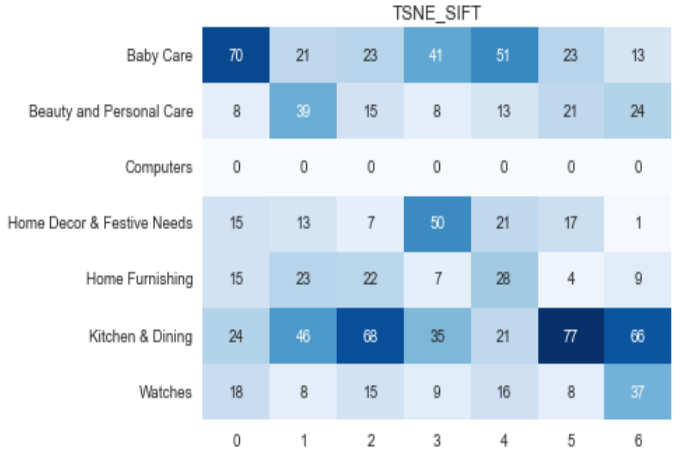




# III. TRAITEMENT DES DONNÉES VISUELLES: BVW – SIFT & ORB

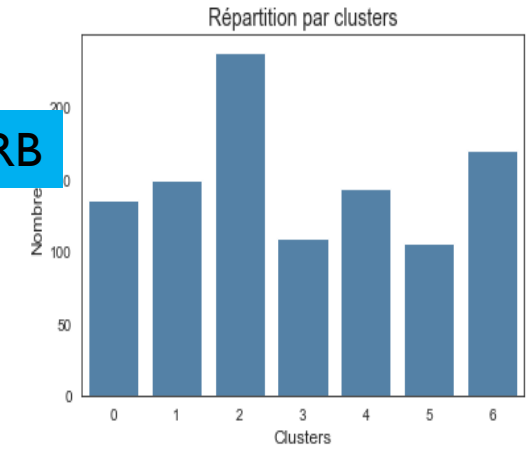
**SIFT**

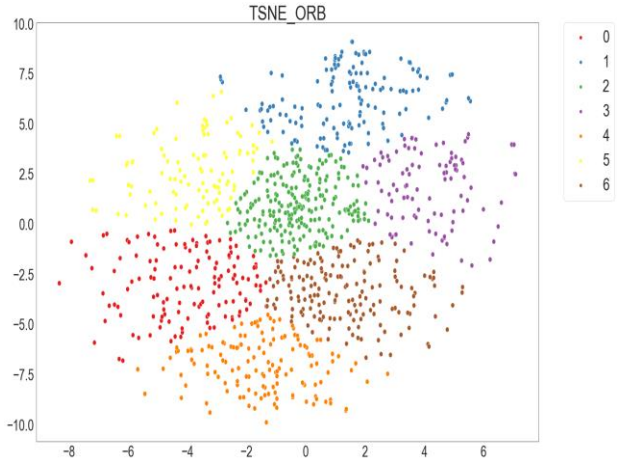



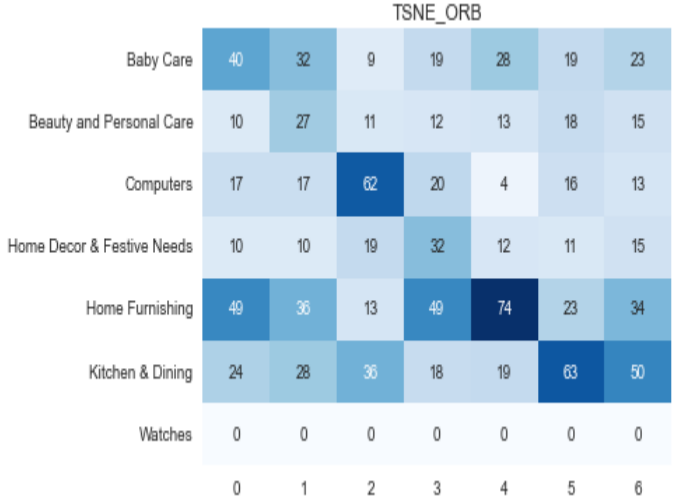




- Le score ARI est faible (0.04).
- Cette combinaison TSNE SIFT sur les images parvient à retrouver les 7 catégories comme le souhaite notre client
- Précision: 28,7%

**ORB**



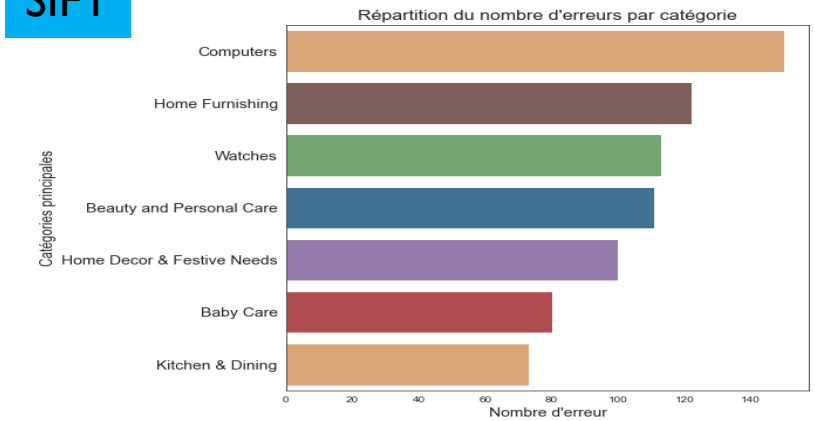



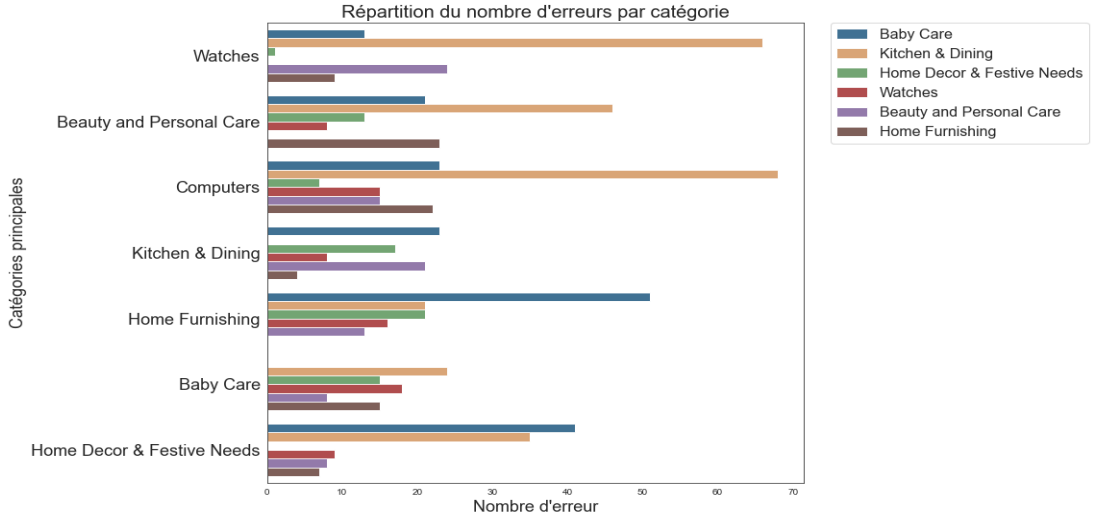




- Le score ARI est faible (0.029).
- Cette combinaison TSNE ORB sur les images parvient à retrouver les 7 catégories comme le souhaite notre client
- Précision: 28,3%

# III. TRAITEMENT DES DONNÉES VISUELLES: BVW – SIFT & ORB - ERREURS

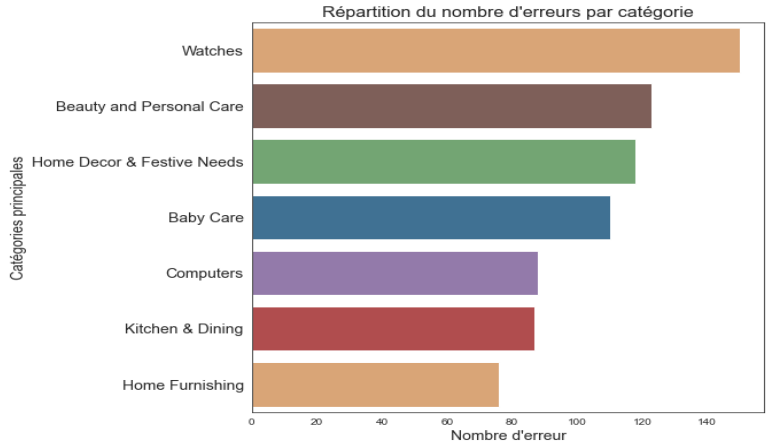
SIFT

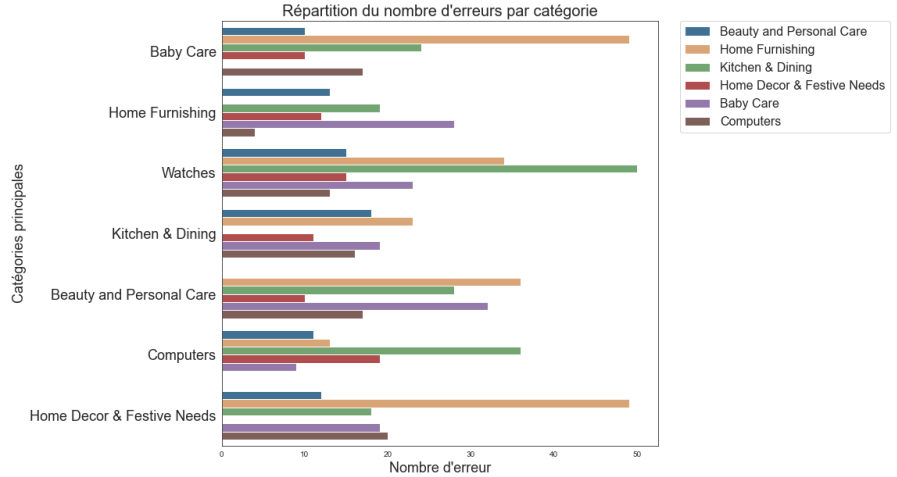



Nombre d'erreurs : 749

ORB




Nombre d'erreurs : 752

# III. TRAITEMENT DES DONNÉES VISUELLES: CNN TRANSFERT LEARNING

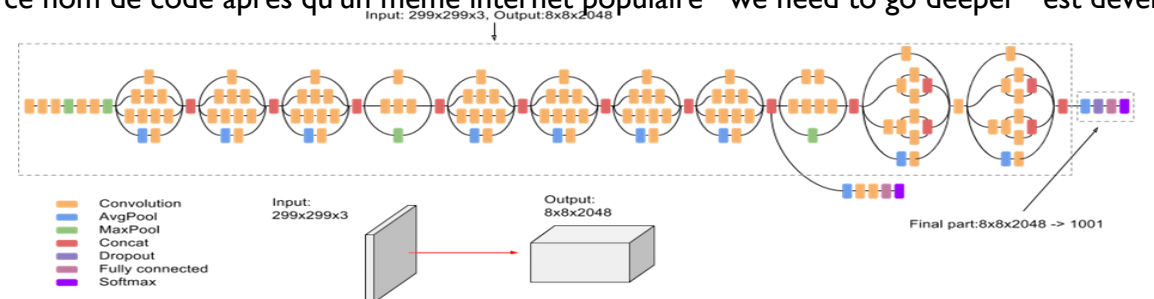
## VGG16

VGG-16 est une version du réseau de neurones convolutif VGG-Net.

- VGG-16 est constitué de plusieurs couches, dont 13 couches de convolution et 3 fully-connected. Il doit donc apprendre les poids de 16 couches.
- Il prend en entrée une image en couleurs de taille 224 × 224 px et la classifie dans une des 1000 classes. Il renvoie donc un vecteur de taille 1000, qui contient les probabilités d'appartenance à chacune des classes.

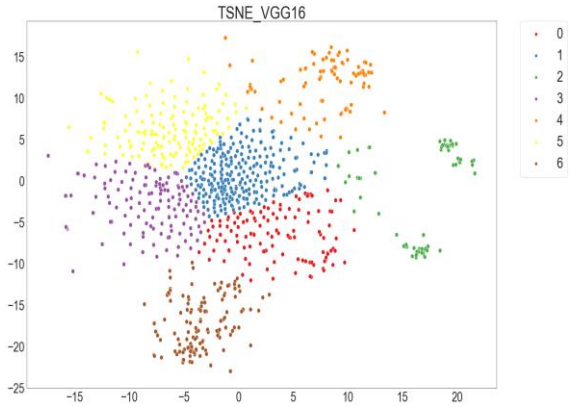

## InceptionV3

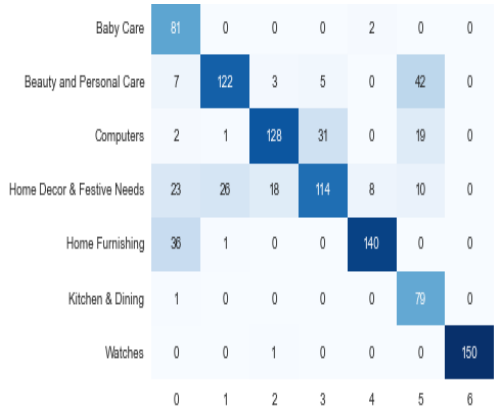
- Inception v3 est un réseau de neurones convolutif destiné à aider à l'analyse d'images et à la détection d'objets, et a débuté comme module pour Googlenet.
- Il s'agit de la troisième édition du réseau de neurones convolutif Inception de Google, initialement présenté lors du défi de reconnaissance ImageNet.
- Tout comme ImageNet peut être considéré comme une base de données d'objets visuels classés, Inception aide à la classification d'objets dans le monde de la vision par ordinateur.
- Le nom original (Inception) a reçu ce nom de code après qu'un mème internet populaire "'we need to go deeper'" est devenu viral, citant une phrase du film Inception de Christopher Nolan.



# III. TRAITEMENT DES DONNÉES VISUELLES: CNN TRANSFERT LEARNING

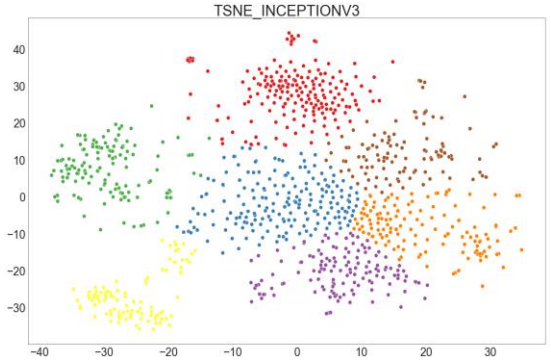
## VGG16

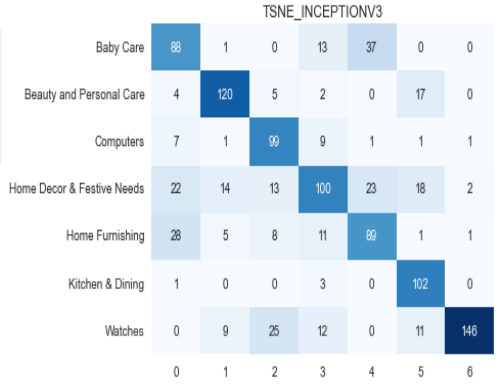




- Le score ARI est bon (0.31).
- Cette combinaison sur les images parvient à retrouver les 7 catégories comme le souhaite notre client
- Précision: 58%

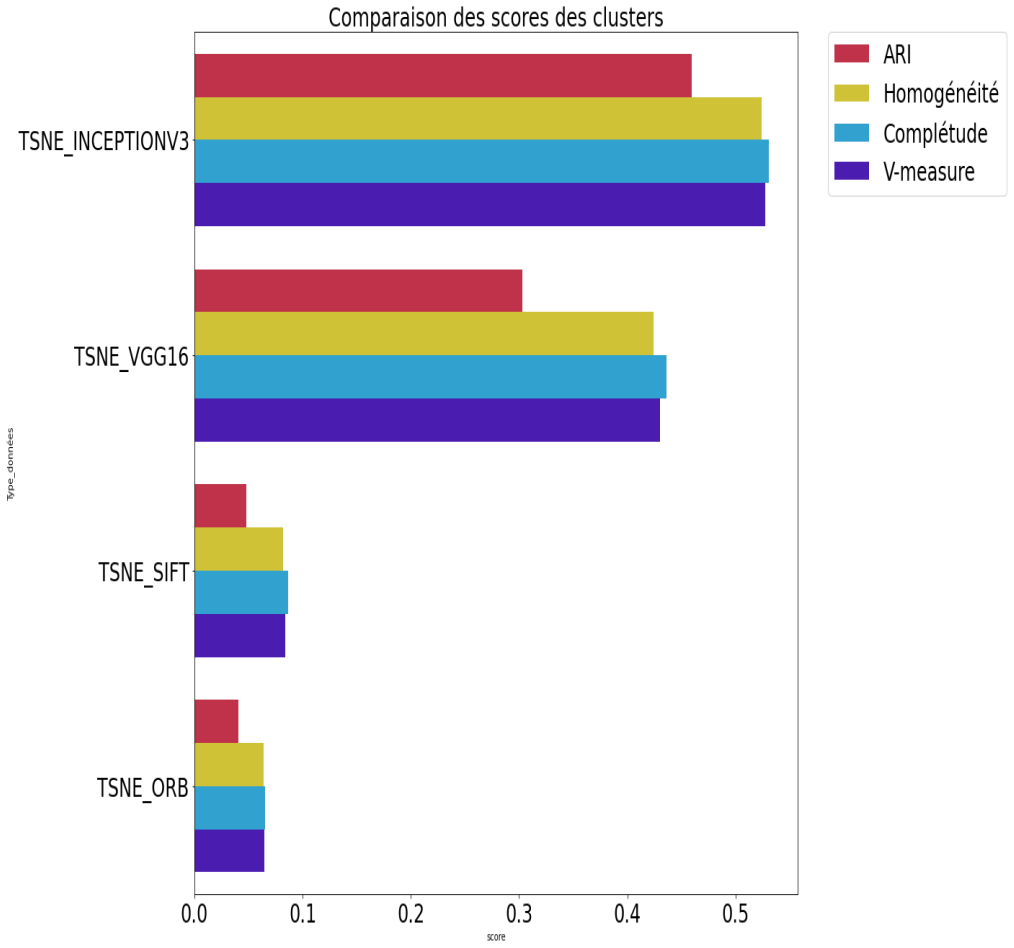
## InceptionV3

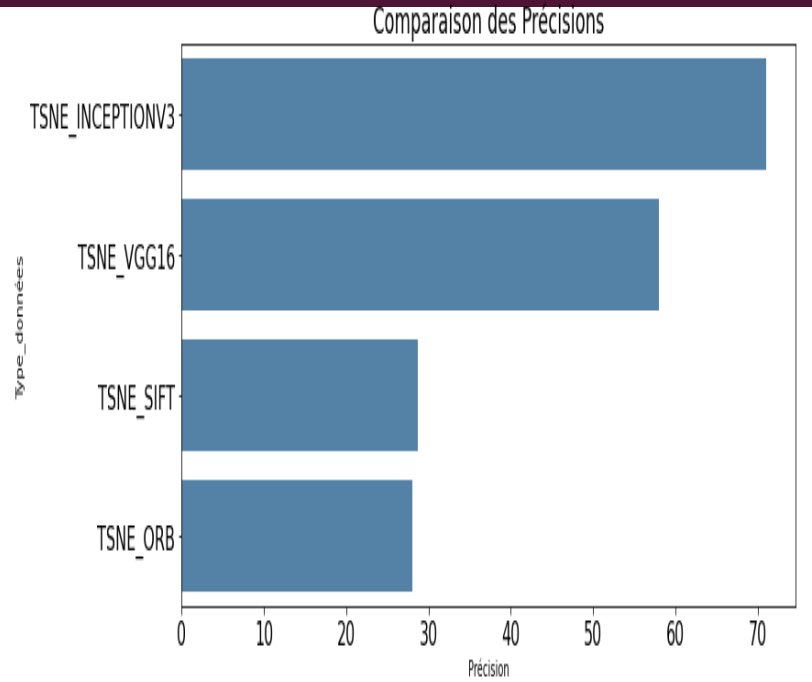




- Le score ARI est bon (0.45).
- Cette combinaison sur les images parvient à retrouver les 7 catégories comme le souhaite notre client
- Précision: 70,1%

# III. TRAITEMENT DES DONNÉES VISUELLES: COMPARAISON DE TOUS LES SCORES ET PRÉCISION

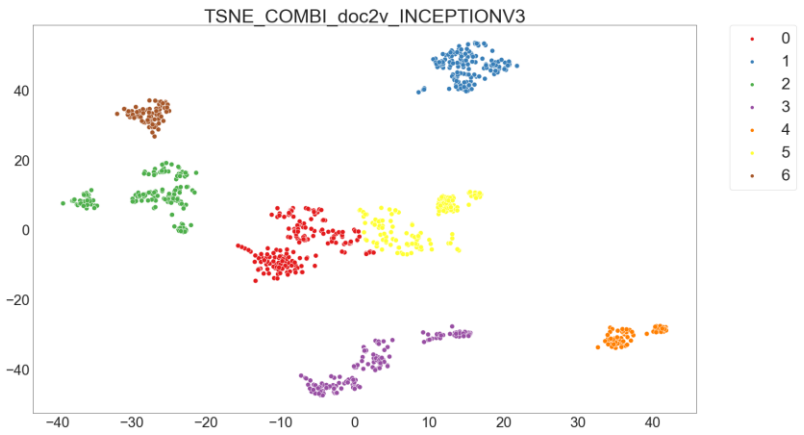


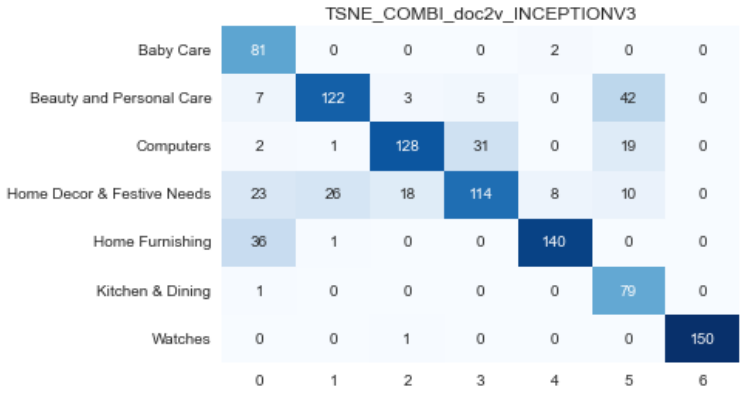




- Les algorithmes InceptionV3 (ARI=0.45 et 71% de précision) et VGG16 (ARI=0.30 et 58% de précision), sont les 2 algorithmes les plus performants sur les images et parviennent à classer les images en 7 catégories comme souhaité par le client

# III. TRAITEMENT DES DONNÉES VISUELLES: MIXTE AVEC LES DONNÉES TEXTUELLES

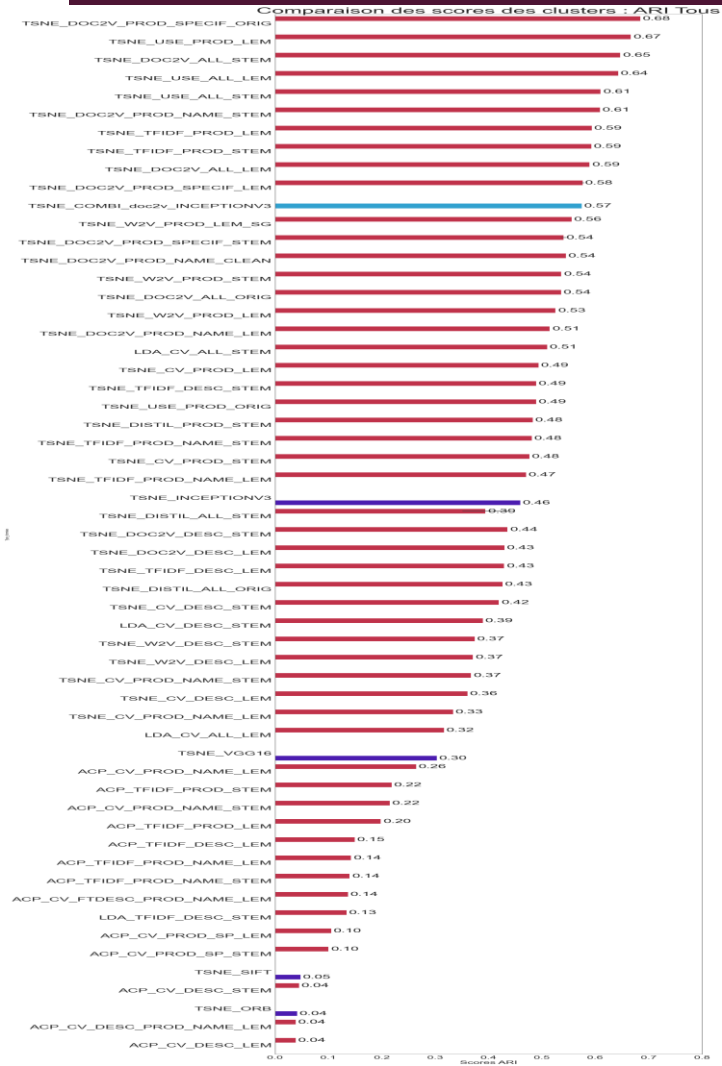
- Combinaison des meilleurs modèles des données textuelles et visuelles
  - Word2vec + Inception V3

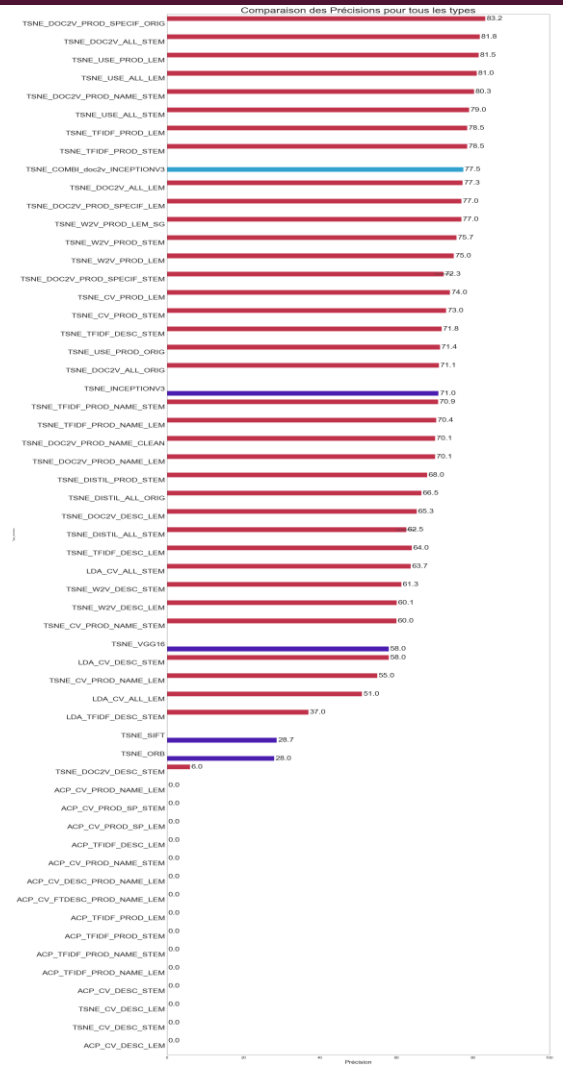



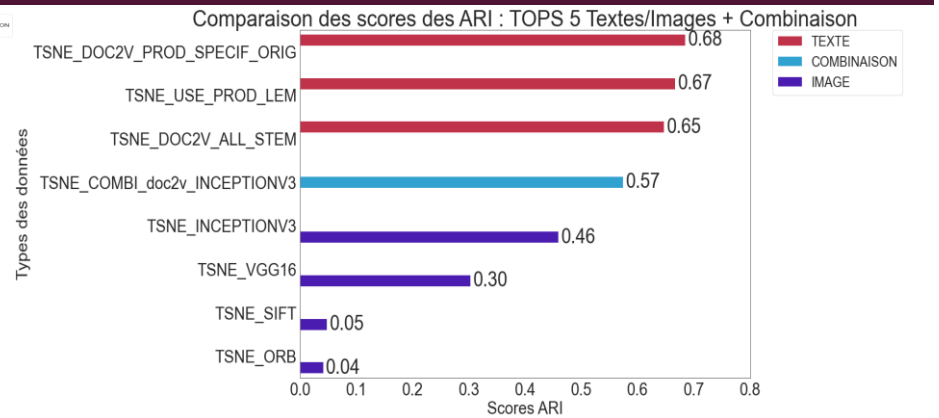



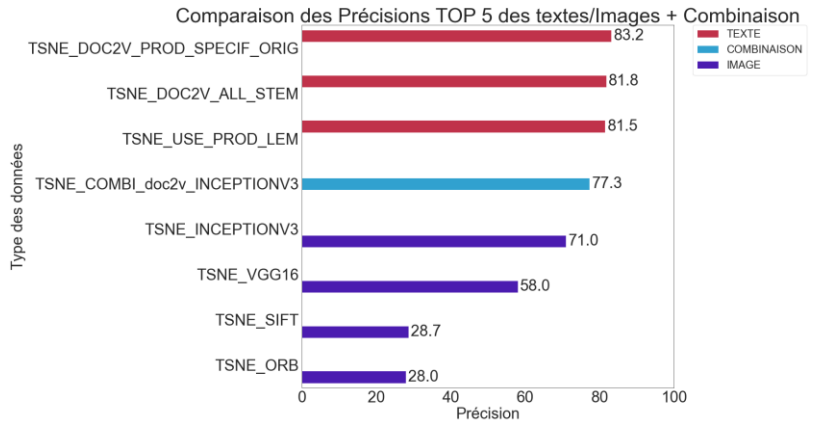

•Le score ARI est assez haut (0.60).
•Cette combinaison TSNE Combinaison données textuelles produits spécifications lemmatisés avec word2vec et données image avec InceptionV3 sur les images parvient à retrouver les 7 catégories comme le souhaite notre client avec une précision de 79%.

# IV/ COMPARAISON DE TOUS LES MODÈLES: BILAN CONCLUSION




- On remarque la combinaison entre Texte et Image donne des résultats intermédiaires entre Texte seul et Image seule
- Utilisation d'autre modèles cnn
- Il faut densifier la base de données pour améliorer les performances