



Sommaire

1. Objectifs: étude de faisabilité d'un moteur de classification
2. Données
 - 2.1 Présentation
 - 2.2 Nettoyage
3. Classification via les Descriptions
 - 3.1 Bag of Words
 - 3.2 Tf-Idf
 - 3.3 LDA
4. Classification via les Images
 - 4.1 SIFT
 - 4.2 CNN
5. Conclusion

Données : Présentation

product	category	description	image
Elegance Polyester Multicolor Abstract Eyelet Door Curtain	[" Home Furnishing >> Curtains & Accessories >> Curtains >> Elegance Polyester Multicolor Abstract Eyelet Do..."]	Key Features of Elegance Polyester Multicolor Abstract Eyelet Door Curtain Floral Curtain, Elegance Polyester Multicolor Abstract Eyelet Door Curtain (213 cm in Height, Pack of 2) Price: Rs. 899 This curtain enhances the look of the interiors.This curtain is made from 100% high quality polyester fabric. It features an eyelet style stitch with Metal Ring.It makes the room environment romantic and loving. This curtain is ant- wrinkle and anti shrinkage and have elegant apparance.Give your home a bright and modernistic appeal with these designs. The surreal attention is sure to steal hearts. These contemporary eyelet and valance curtains slide smoothly so when you draw	17.jpg

1050 produits
(150 / catégorie)

7 Catégories niveau_1

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

109 à 3490 caractères

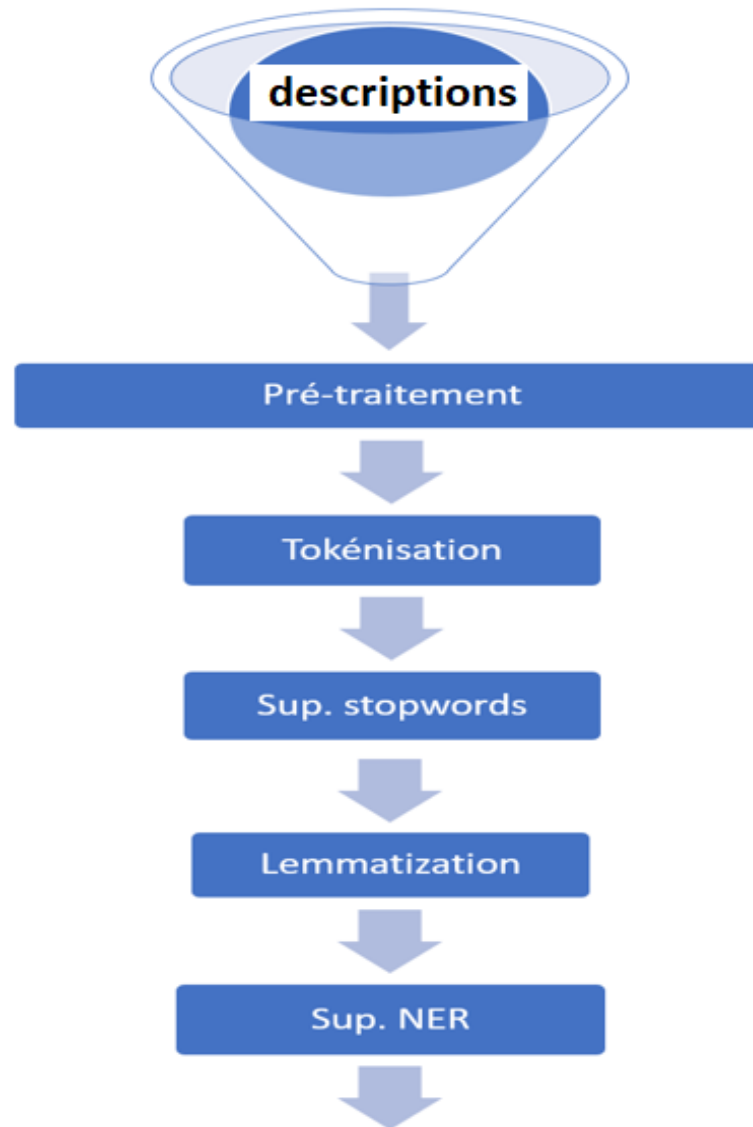
image (couleur)

Largeur (px) : 145 ... 8484

Hauteur (px) : 160 ... 11042

Ex : "**Baby Care** > **Baby Bath & Skin** > **Baby Bath Towels** > **Sathiyas Baby Bath Towels** > **Sathiyas Cotton Bath Towel Y...**"
niveau_1 niveau_2 ...

Traitement des descriptions



Pré-traitement

Calibro SW-**125** Analog-Digital Watch - For Men, Boys - Price: Rs. **699** **#CALIBRO** presents MTG Black Dial Round Watch. This is a fashionable watch from MTG that adds to your style statement. This brand is known for its usage of quality material in making appealing watches. The variety in design, style and youthful approach is very well executed by this brand. Get one and see how brilliantly it leads other brands in style, finish and everything else that matters. CALIBRO presents MTG Black Dial Round Watch. This is a fashionable watch from MTG that adds to your style statement. It understands the needs of youth so well that it is able to present wide range of watches to its customers. Get one at **@Calibro** or **<https://calibro.com/>**.



Calibro SW- Analog-Digital Watch - For Men, Boys - Price: Rs. presents MTG Black Dial Round Watch. This is a fashionable watch from MTG that adds to your style statement. This brand is known for its usage of quality material in making appealing watches. The variety in design, style and youthful approach is very well executed by this brand. Get one and see how brilliantly it leads other brands in style, finish and everything else that matters. CALIBRO presents MTG Black Dial Round Watch. This is a fashionable watch from MTG that adds to your style statement. It understands the needs of youth so well that it is able to present wide range of watches to its customers. Get one at or

- liens hypertexte <https://fr.trustpilot.com/review/www.amazon.fr>
- Tags @Amazon, #SAV
- Emoji 😊😏😐😞
- Numérique 0123456789

Tokénisation

Calibro **SW- Analog-Digital Watch** - For Men, Boys - Price: Rs. presents MTG Black Dial Round Watch. This is a fashionable watch from MTG that adds to your style statement. This brand is known for its usage of quality material in making appealing watches. The variety in design, style and youthful approach is very well executed by this brand. Get one and see how brilliantly it leads other brands in style, finish and everything else that matters. CALIBRO presents MTG Black Dial Round Watch. This is a fashionable watch from MTG that adds to your style statement. It understands the needs of youth so well that it is able to present wide range of watches to its customers. Get one at or



'calibro', 'sw', 'analog', 'digital', 'watch', 'for', 'men', 'boys', 'price', 'rs', 'presents', 'mtg', 'black', 'dial', 'round', 'watch', 'this', 'is', 'a', 'fashionable', 'watch', 'from', 'mtg', 'that', 'adds', 'to', 'your', 'style', 'statement', 'this', 'brand', 'is', 'known', 'for', 'its', 'usage', 'of', 'quality', 'material', 'in', 'making', 'appealing', 'watches', 'the', 'variety', 'in', 'design', 'style', 'and', 'youthful', 'approach', 'is', 'very', 'well', 'executed', 'by', 'this', 'brand', 'get', 'one', 'and', 'see', 'how', 'brilliantly', 'it', 'leads', 'other', 'brands', 'in', 'style', 'finish', 'and', 'everything', 'else', 'that', 'matters', 'calibro', 'presents', 'mtg', 'black', 'dial', 'round', 'watch', 'this', 'is', 'a', 'fashionable', 'watch', 'from', 'mtg', 'that', 'adds', 'to', 'your', 'style', 'statement', 'it', 'understands', 'the', 'needs', 'of', 'youth', 'so', 'well', 'that', 'it', 'is', 'able', 'to', 'present', 'wide', 'range', 'of', 'watches', 'to', 'its', 'customers', 'get', 'one', 'at', 'or'

Suppression des Stop Words

'calibro', 'sw', 'analog', 'digital', 'watch', 'for', 'men', 'boys', 'price', 'rs', 'presents', 'mtg', 'black', 'dial', 'round', 'watch', 'this', 'is', 'a',
'fashionable', 'watch', 'from', 'mtg', 'that', 'adds', 'to', 'your', 'style', 'statement', 'this', 'brand', 'is', 'known', 'for', 'its', 'usage', 'of',
'quality', 'material', 'in', 'making', 'appealing', 'watches', 'the', 'variety', 'in', 'design', 'style', 'and', 'youthful', 'approach', 'is', 'very',
'well', 'executed', 'by', 'this', 'brand', 'get', 'one', 'and', 'see', 'how', 'brilliantly', 'it', 'leads', 'other', 'brands', 'in', 'style', 'finish', 'and',
'everything', 'else', 'that', 'matters', 'calibro', 'presents', 'mtg', 'black', 'dial', 'round', 'watch', 'this', 'is', 'a', 'fashionable', 'watch',
'from', 'mtg', 'that', 'adds', 'to', 'your', 'style', 'statement', 'it', 'understands', 'the', 'needs', 'of', 'youth', 'so', 'well', 'that', 'it', 'is',
'able', 'to', 'present', 'wide', 'range', 'of', 'watches', 'to', 'its', 'customers', 'get', 'one', 'at', 'or'



'calibro', 'sw', 'analog', 'digital', 'watch', 'men', 'boys', 'price', 'rs', 'presents', 'mtg', 'black', 'dial', 'round', 'watch', 'fashionable',
'watch', 'mtg', 'adds', 'style', 'statement', 'brand', 'known', 'usage', 'quality', 'material', 'making', 'appealing', 'watches', 'variety',
'design', 'style', 'youthful', 'approach', 'well', 'executed', 'brand', 'get', 'one', 'see', 'brilliantly', 'leads', 'brands', 'style', 'finish',
'everything', 'else', 'matters', 'calibro', 'presents', 'mtg', 'black', 'dial', 'round', 'watch', 'fashionable', 'watch', 'mtg', 'adds', 'style',
'statement', 'understands', 'needs', 'youth', 'well', 'able', 'present', 'wide', 'range', 'watches', 'customers', 'get', 'one'

Lemmatisation

'calibro', 'sw', 'analog', 'digital', 'watch', 'men', 'boys', 'price', 'rs', 'presents', 'mtg', 'black', 'dial', 'round', 'watch', 'fashionable', 'watch', 'mtg', 'adds', 'style', 'statement', 'brand', 'known', 'usage', 'quality', 'material', 'making', 'appealing', 'watches', 'variety', 'design', 'style', 'youthful', 'approach', 'well', 'executed', 'brand', 'get', 'one', 'see', 'brilliantly', 'leads', 'brands', 'style', 'finish', 'everything', 'else', 'matters', 'calibro', 'presents', 'mtg', 'black', 'dial', 'round', 'watch', 'fashionable', 'watch', 'mtg', 'adds', 'style', 'statement', 'understands', 'needs', 'youth', 'well', 'able', 'present', 'wide', 'range', 'watches', 'customers', 'get', 'one'



'calibro', 'sw', 'analog', 'digital', 'watch', 'man', 'boy', 'price', 'rs', 'present', 'mtg', 'black', 'dial', 'round', 'watch', 'fashionable', 'watch', 'mtg', 'add', 'style', 'statement', 'brand', 'know', 'usage', 'quality', 'material', 'make', 'appeal', 'watch', 'variety', 'design', 'style', 'youthful', 'approach', 'well', 'execute', 'brand', 'get', 'one', 'see', 'brilliantly', 'lead', 'brand', 'style', 'finish', 'everything', 'else', 'matter', 'calibro', 'present', 'mtg', 'black', 'dial', 'round', 'watch', 'fashionable', 'watch', 'mtg', 'add', 'style', 'statement', 'understand', 'need', 'youth', 'well', 'able', 'present', 'wide', 'range', 'watch', 'customer', 'get', 'one'

- ✓ Un verbe deviendra son infinitif
- ✓ Un nom deviendra son masculin singulier, idem pour les adjectifs

Suppression des entités nommées (NER)

Organisations : ['gsm', 'terry design shuvam', 'eurospa gsm', 'santosh royal fashion cotton print king', 'santosh royal fashion machine washable', 'goldbedi', 'jaipur print cotton floral king', 'jaipur', 'maserati', 'maserati time r analog watch', 'alrams great', 'calibro sw', 'mtg black dial round', 'mtg black', 'lcs analog watch', 'titan', 'srushti art jewelry digital', 'digital watch', 'fbc', 'skmei', 'rs india', 'skmei dg red sports digital watch', 'fastrack pl', 'com stainless steel', 'qibla compass', 'disney analog watch', 'disney', 'vizion purple sports series', 'elantra', 'lenco bdblue', 'analog watch man', 'santosh royal fashion cotton', 'diwan sheet cushion', 'riva carpet cotton free bath', 'myesquire ceramic burner', 'southampton university u k']

Personnes : ['camerii wm', 'gregorian calnedars qur', 'cau ba', 'lois caron', 'srushti girlswatch', 'jack klein', 'franck bella fbc', 'ridas apple', 'ducjug ducgl', 'akup', 'mug akup', 'roq', 'giorgio armani', 'chinmay', 'aapno rajasthan', 'vaishnav', 'radha krishna', 'lollipop lane', 'recrafto buddha', 'tatvaarts url', 'purpledip kattle', 'agromech yiboo', 'rajo raja magan', 'ganesh ganpati', 'moshai baby', 'vincent valentine paris', 'perfume anna andre', 'neo', 'hamilton combo', 'anna andre', 'tenda te wr', 'yardley jasmine', 'vincent valentine' ...]

Lieux : ['mediam', 'shuvampcftsetassorte', 'north moon', 'north moon', 'inch cm', 'diameter mm', 'diameter mm', 'lbs g', 'mars', 'mars', 'radha krishna', 'cm brass', 'asia', 'asia', 'cm crystal', 'rajasthan', 'asia', 'asia', 'mayursflora bonsai aloe', 'ft x ft stiche', 'laxmi brass', 'laxmi brass', 'ac', 'jaycoknit mediterranean sea', 'jaycoknit mediterranean sea', 'mediterranean', 'mediterranean', 'jaycoknit mediterranean sea', 'mediterranean', 'west bengal', 'west bengal', 'mhz dimm']

Produits : ['sathiyas', 'camerii wm', 'titan sl', 'ridas apple', 'alfajr wq', 'print xcm', 'ios', 'android', 'cookart casserole', 'casserole', 'casserole handi', 'android smart', 'usb', 'android smart', 'android smart', 'jrb type usb', 'intenso creme g', 'katori bowl', 'vareesha', 'treasure', 'cvpc', 'vii', 'nvidia geforce', 'polyresin', 'model', 'razai quilt', 'dohar', 'omic poto blue ml', 'xp vista shopizone', 'bmw', 'shopizone bmw wheel gb pen', 'sumeet hard', ...]

Norp : ['luman', 'polish', 'rajasthan', 'indian', 'rajasthan', 'indian', 'prithish', 'chinese', 'indian', 'beautiful indian rajasthani', 'beautiful indian rajasthani', 'rajasthani', 'rajasthani', 'metallino', 'rajasthani', 'italian', 'italian', 'adino', 'pooja', 'ganesha', 'hindu', 'jains', 'buddhists', 'ganesha', 'zingalalaa', 'polish', 'polish', 'polish', 'theseuberstylish', 'masala', 'chinese', 'vatika', 'titrit', 'hindu', 'puja upasana hindu', 'hindu', 'puja upasana hindu', 'persian', 'persian', 'furnishingland', 'persian', 'persian', 'arabic', 'arabian', ...]

2677 mots ont été retirés du corpus

Classification des descriptions – Bag of Words

	mot1	mot2	mot3	mot4	mot5	mot6	885
description1	0	0	0	1	1	0	0	0	0
description2	1	1	0	0	0	2	0	0	1
description3	0	0	0	0	0	0	0	0	1
description4	1	1	0	0	0	0	1	1	1
description5	3	0	0	0	1	0	0	0	0
...									
description(n)									

mot5 apparaît une fois dans description5

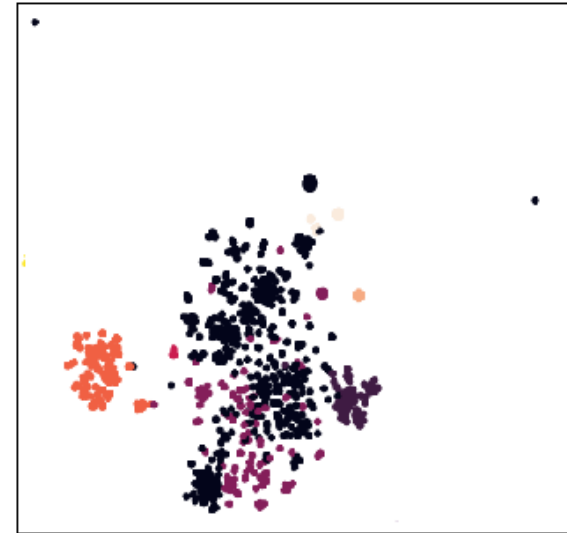
* **min_df=5** : Un mot est ajouté au vocabulaire s'il apparaît au moins 5 fois.

Résultat - Bag of Words (2)

Matrice de confusion

Baby Care	59	78	13	0	0	0	0
Beauty and Personal Care	140	0	10	0	0	0	0
Computers	116	0	27	7	0	0	0
Home Decor & Festive Needs	113	0	37	0	0	0	0
Home Furnishing	105	0	45	0	0	0	0
Kitchen & Dining	101	0	18	0	0	11	20
Watches	1	0	1	0	148	0	0
	0	1	2	3	4	5	6
	Cluster						

Représentation t-SNE des produits



Nb produits
/ cluster

cluster	nb_produits
0	635
1	78
2	151
3	7
4	148
5	11
6	20

Rand Score ajusté : 0.1593

Matching Clustering Kmeans -
Classification humaine

Classification des descriptions – Tf-Idf

	additional	also	analog	baby	black	body	box	boy	brand	buy	care	cash	ceramic	...
description1	0	0	0	0	0	0,125344	0,082055	0	0,087738	0	0	0	0	
description2	0	0	0	0	0	0	0,120724	0	0,129084	0	0	0	0	
description3	0	0	0,073561	0	0,174814	0	0	0	0,372679	0	0	0	0	
description4	0	0	0,409667	0	0	0	0	0,469889	0	0,106337	0	0,107968	0	
description5	0	0	0,44848	0	0	0	0	0,257204	0	0,116412	0	0,118197	0	
description	0	0	0,44848	0	0	0	0	0,257204	0	0,116412	0	0,118197	0	
	0	0	0	0	0	0	0	0,300113	0	0,067916	0	0,068958	0	
	0	0	0,438489	0	0,347349	0	0	0,167649	0	0	0	0	0	

85 mots

$$\text{TF-IDF} = \text{TF} * \text{IDF}$$

- **TF** : Nombre de fois que le mot apparaît dans une description
- **IDF** : $\text{LOG}(\text{nb descriptions} / \text{nb descriptions contenant le terme})$

* Plus le TF-IDF est élevé, plus le mot est représentatif de la description.

* **max_df = 0.99** : ignore les mots qui apparaissent dans plus de 99% des descriptions.

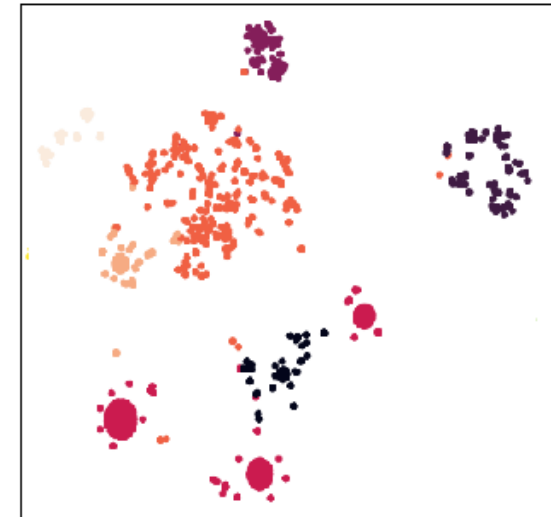
* **min_df = 0.06** : ignore les mots qui apparaissent dans moins de 6% des descriptions.

Résultat - Tf-Idf (2)

Matrice de confusion

Baby Care	20	0	85	14	30	0	1
Beauty and Personal Care	0	0	0	100	49	1	0
Computers	0	0	0	94	56	0	0
Home Decor & Festive Needs	0	0	0	12	54	84	0
Home Furnishing	74	0	0	0	76	0	0
Kitchen & Dining	1	0	0	31	44	0	74
Watches	0	147	0	0	3	0	0
	0	1	2	3	4	5	6
		Cluster					

Représentation t-SNE des produits



Nb produits
/ cluster

cluster	nb_produits
0	95
1	147
2	85
3	251
4	312
5	85
6	75

Rand Score ajusté : 0.3492

Matching Clustering Kmeans -
Classification humaine

Classification des descriptions – LDA

20 mots les plus représentatifs sur 885

- TOPIC 1 :** cm baby cotton pack color general type design fabric box feature specification number girl package sale detail print brand polyester
- TOPIC 2 :** product eye beauty lip price quality cream face craft oxyglow make range bring massage beautifully art vitamin get colour canvas
- TOPIC 3 :** free delivery product buy cash genuine shipping day replacement guarantee online watch analog price good man india great discount showpiece
- TOPIC 4 :** skin warranty shape print pad inch mouse set multicolor size quality product type feature adapter power brand color high specification
- TOPIC 5 :** glass bottle water pizza steel price design cutter pack nutcase soap copper ml wrap stainless use durable frame body come
- TOPIC 6 :** mug sticker wall ceramic make cm home product gift material price design coffee perfect feature one apply size use give

* **Topic** = **Sujet** = **Catégorie**

LDA (2)

Assigination du topic dominant à chaque description

description	topic
['key', 'feature', 'elegance', 'polyester', 'multicolor', 'abstract', 'eyelet', 'door', 'curtain', 'floral', 'curtain', 'elegance', 'polyester', 'multicolor', 'abstract', 'eyelet', 'door', 'curtain', 'cm', 'height', 'pack', 'price', 'curtain', 'enhance', 'look', 'interior', 'curtain', 'make', 'high', 'quality', 'polyester', 'fabric', 'feature', 'eyelet', 'style', 'stitch', 'metal', 'ring', 'make', 'room', 'environment', 'romantic', 'love', 'curtain', 'ant', 'wrinkle', 'anti', 'shrinkage', 'elegant', 'apparance', 'give', 'home', 'bright', 'modernistic', 'appeal', 'design', 'surreal', 'attention', 'sure', 'steal', 'heart', 'contemporary', 'eyelet', 'valance', 'curtain', 'slide', 'smoothly', 'draw', 'apart', 'first', 'thing', 'morning', 'welcome', 'bright', 'sun', 'ray', 'want', 'wish', 'good', 'morning', 'whole', 'world', 'draw', 'close', 'evening', 'create', 'special', 'moment', 'joyous', 'beauty', 'give', 'soothe', 'print', 'bring', 'home', 'elegant', 'curtain', 'softly', 'filter', 'light', 'room', 'get', 'right', 'amount', 'sunlight', 'specification', 'elegance', 'polyester', 'multicolor', 'abstract', 'eyelet', 'door', 'curtain', 'cm', 'height', 'pack', 'general', 'brand', 'elegance', 'design', 'door', 'type', 'eyelet', 'name', 'abstract', 'polyester', 'door', 'curtain', 'set', '-PRON-', 'would', 'duster', 'color', 'multicolor', 'dimension', 'length', 'cm', 'box', 'number', 'content', 'sale', 'package', 'pack', 'sale', 'package', 'curtain', 'body', 'design', 'material', 'polyester']	3
['specifications', 'cotton', 'bath', 'towel', 'bath', 'towel', 'red', 'yellow', 'blue', 'bath', 'towel', 'feature', 'machine', 'washable', 'yes', 'material', 'cotton', 'd	3
['key', 'feature', 'eurospa', 'cotton', 'terry', 'face', 'towel', 'set', 'size', 'small', 'height', 'inch', 'eurospa', 'cotton', 'terry', 'face', 'towel', 'set', 'piece', 'face',	3
['key', 'feature', 'santosh', 'royal', 'fashion', 'cotton', 'print', 'king', 'size', 'double', 'bedsheet', 'royal', 'bedsheet', 'perfact', 'wedding', 'gifting', 'specifica	3
['key', 'feature', 'print', 'cotton', 'floral', 'king', 'size', 'double', 'bedsheet', 'cotton', 'print', 'cotton', 'floral', 'king', 'size', 'double', 'bedsheet', 'bed', 'shee	3
['time', 'r', 'analog', 'watch', 'boy', 'buy', 'time', 'r', 'analog', 'watch', 'boy', 'r', 'online', 'india', 'great', 'discount', 'genuine', 'product', 'day', 'replacement',	6
['camerii', 'elegance', 'analog', 'watch', 'man', 'boy', 'buy', 'camerii', 'elegance', 'analog', 'watch', 'man', 'boy', 'online', 'india', 'great', 'discount', 'genuine	6
['star', 'uft', 'tsw', 'bk', 'br', 'analog', 'watch', 'boy', 'price', 'r', 'whether', 'way', 'work', 'travel', 'abroad', 'family', 'lifestyle', 'accessory', 'like', 'watch', 'wal	2



Sujet dominant d'une description

Résultat - LDA (3)

Matrice de confusion

Baby Care	4	0	0	106	4	2	34
Beauty and Personal Care	17	0	3	16	4	17	93
Computers	0	0	43	4	0	11	92
Home Decor & Festive Needs	0	0	3	13	67	0	67
Home Furnishing	6	0	2	63	3	3	73
Kitchen & Dining	23	35	1	28	6	6	51
Watches	0	0	5	5	1	0	139
	0	1	2	3	4	5	6
	Cluster						

Nb produits
/ cluster

topic	nb_produits
0	50
1	35
2	57
3	235
4	85
5	39
6	549

Rand Score ajusté : 0.1027

Matching Clustering Kmeans -
Classification humaine

Classification des produits via leur image

7 catégories d'images



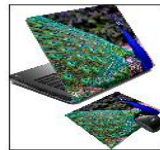
Soins bébé



Ameublement



Beauté et soins personnels



ordinateurs



Cuisine et salle à manger



Décor à la maison et besoins festifs



Montres

catégorie la plus
homogène
visuellement

Classification des images - SIFT



1 image -> X Key Point

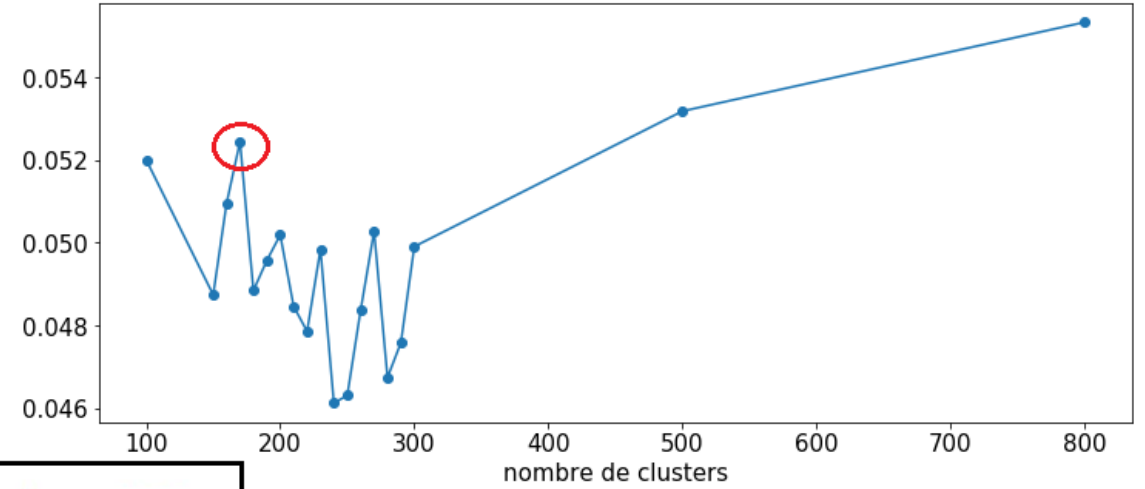
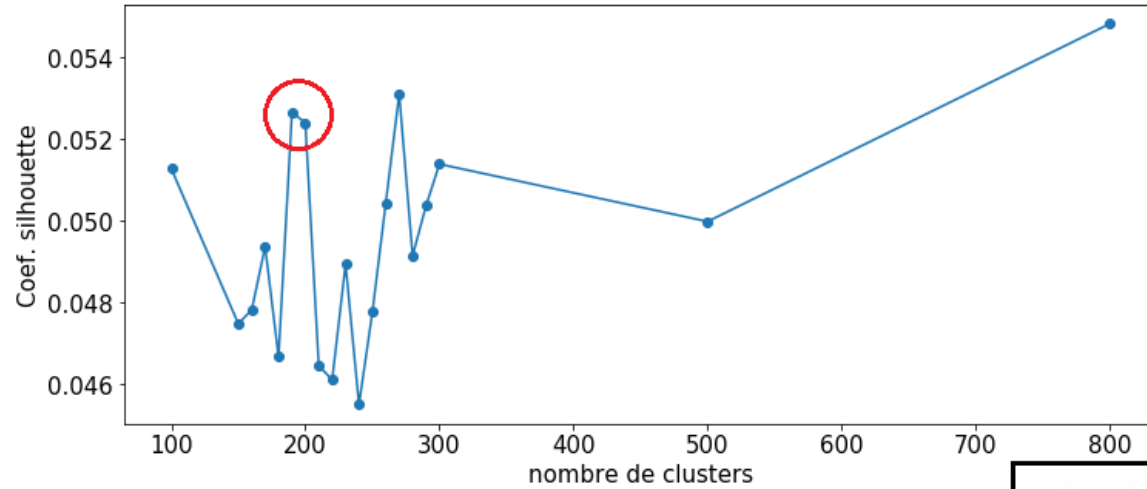
(1 Key Point => 1 descripteur)

image	nombre descripteurs
0	181
1	115
2	229
3	208
4	176
5	148
6	145
7	145
....

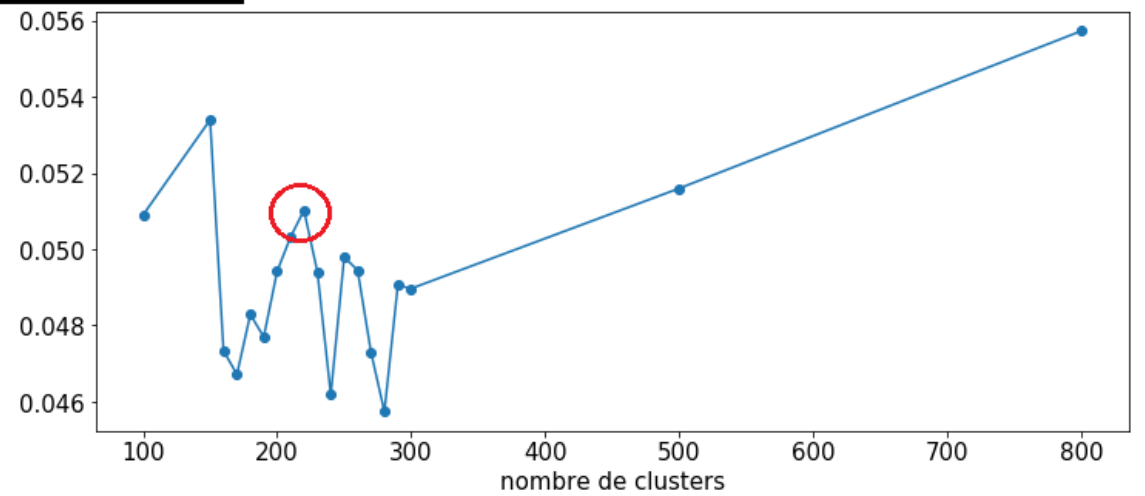
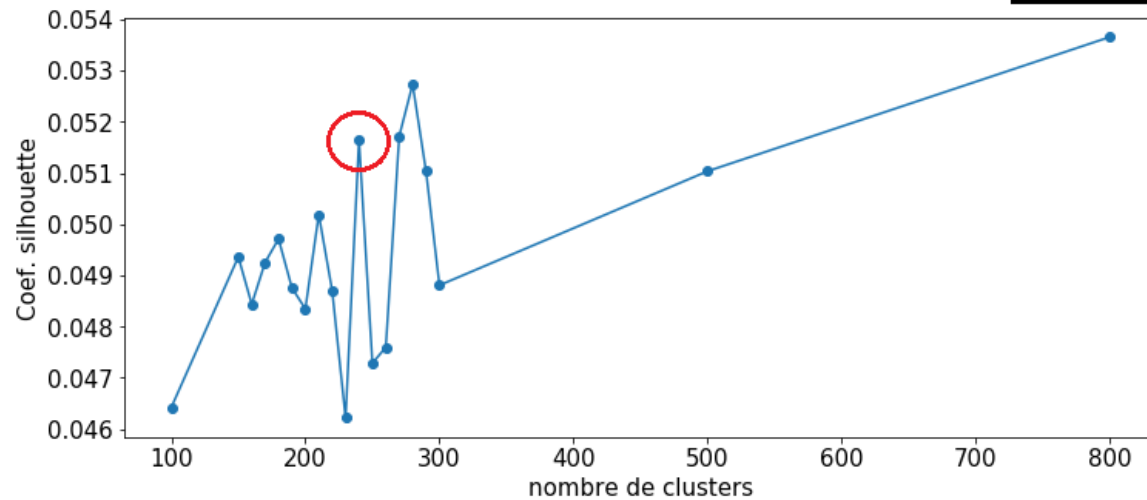
SIFT (2)

0	1	2	3	4	5	6	7	126	127	<- 128 features SIFT
5.0	3.0	0.0	25.0	57.0	21.0	17.0	3.0		0.0	1.0	image 1
2.0	10.0	33.0	139.0	124.0	19.0	17.0	1.0		16.0	14.0	
120.0	53.0	8.0	3.0	6.0	8.0	14.0	120.0		2.0	2.0	
2.0	1.0	1.0	1.0	1.0	0.0	2.0	3.0		0.0	3.0	
41.0	26.0	1.0	0.0	2.0	17.0	26.0	7.0		1.0	2.0	image 2
75.0	123.0	51.0	6.0	9.0	1.0	0.0	1.0		42.0	52.0	
1.0	8.0	21.0	2.0	1.0	20.0	90.0	5.0		12.0	0.0	image 3
140.0	1.0	0.0	0.0	0.0	0.0	0.0	17.0		8.0	10.0	
64.0	60.0	18.0	0.0	0.0	1.0	6.0	13.0		3.0	4.0	
5.0	0.0	0.0	2.0	18.0	9.0	1.0	5.0		0.0	3.0	
0.0	0.0	0.0	3.0	3.0	3.0	0.0	0.0		0.0	1.0	
39.0	0.0	0.0	2.0	169.0	10.0	0.0	1.0		0.0	0.0	
30.0	1.0	0.0	1.0	127.0	29.0	0.0	0.0		2.0	0.0	<- 1 descripteur
77.0	5.0	0.0	0.0	8.0	1.0	0.0	0.0		4.0	2.0	image 4
69.0	2.0	0.0	1.0	9.0	0.0	0.0	1.0		20.0	20.0	
23.0	60.0	118.0	10.0	1.0	0.0	0.0	0.0		1.0	5.0	
1.0	37.0	134.0	11.0	0.0	0.0	0.0	0.0		5.0	7.0	
⋮											⋮
											image 1050

SIFT(3)



K optimal \simeq 230



Histogramme des descripteurs – SIFT(4)

230 classes de descripteurs												
image	0	1	2	3	4	5	6	7	8	9	...	229
1	3.0	4.0	0.0	1.0	0.0	4.0	3.0	1.0	0.0	0.0		0.0
2	0.0	0.0	0.0	0.0	8.0	1.0	0.0	0.0	0.0	1.0		0.0
3	0.0	1.0	5.0	6.0	0.0	0.0	1.0	0.0	10.0	0.0		0.0
4	0.0	0.0	0.0	1.0	2.0	0.0	1.0	0.0	0.0	17.0		0.0
5	0.0	0.0	1.0	0.0	1.0	0.0	2.0	0.0	1.0	3.0		0.0
6	0.0	0.0	2.0	0.0	2.0	0.0	0.0	0.0	0.0	5.0		0.0
7	1.0	3.0	0.0	0.0	2.0	1.0	0.0	0.0	1.0	0.0		0.0
8	1.0	0.0	0.0	0.0	5.0	0.0	0.0	0.0	1.0	0.0		0.0
9	2.0	0.0	0.0	0.0	4.0	0.0	0.0	0.0	0.0	0.0		0.0
10	0.0	4.0	0.0	1.0	2.0	0.0	1.0	0.0	1.0	0.0		0.0
.....											
.....												
1047	0.0	0.0	0.0	0.0	3.0	0.0	0.0	0.0	0.0	0.0		0.0
1048	0.0	2.0	1.0	1.0	0.0	3.0	1.0	2.0	0.0	0.0	...	0.0

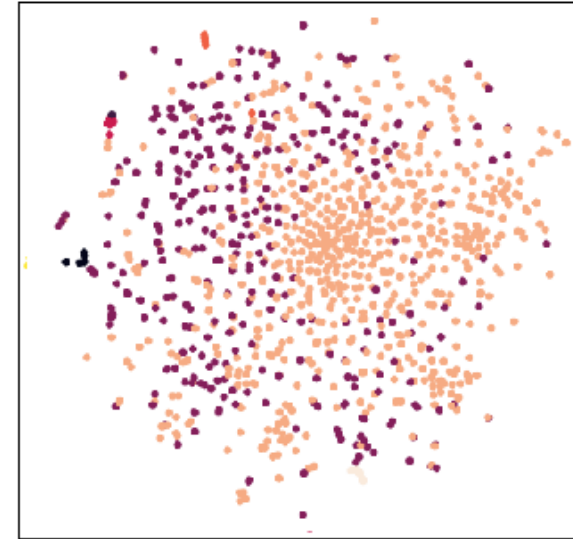
L'image **10** contient **4** descripteurs de classe **1**

Résultat – SIFT (5)

Matrice de confusion

Baby Care	0	0	41	0	6	103	0
Beauty and Personal Care	0	0	49	0	0	101	0
Computers	0	0	12	1	0	133	2
Home Decor & Festive Needs	1	0	78	3	0	68	0
Home Furnishing	5	1	79	1	0	60	4
Kitchen & Dining	1	0	33	1	0	115	0
Watches	0	0	42	0	0	108	0
	0	1	2	3	4	5	6
	Cluster						

Représentation t-SNE des produits



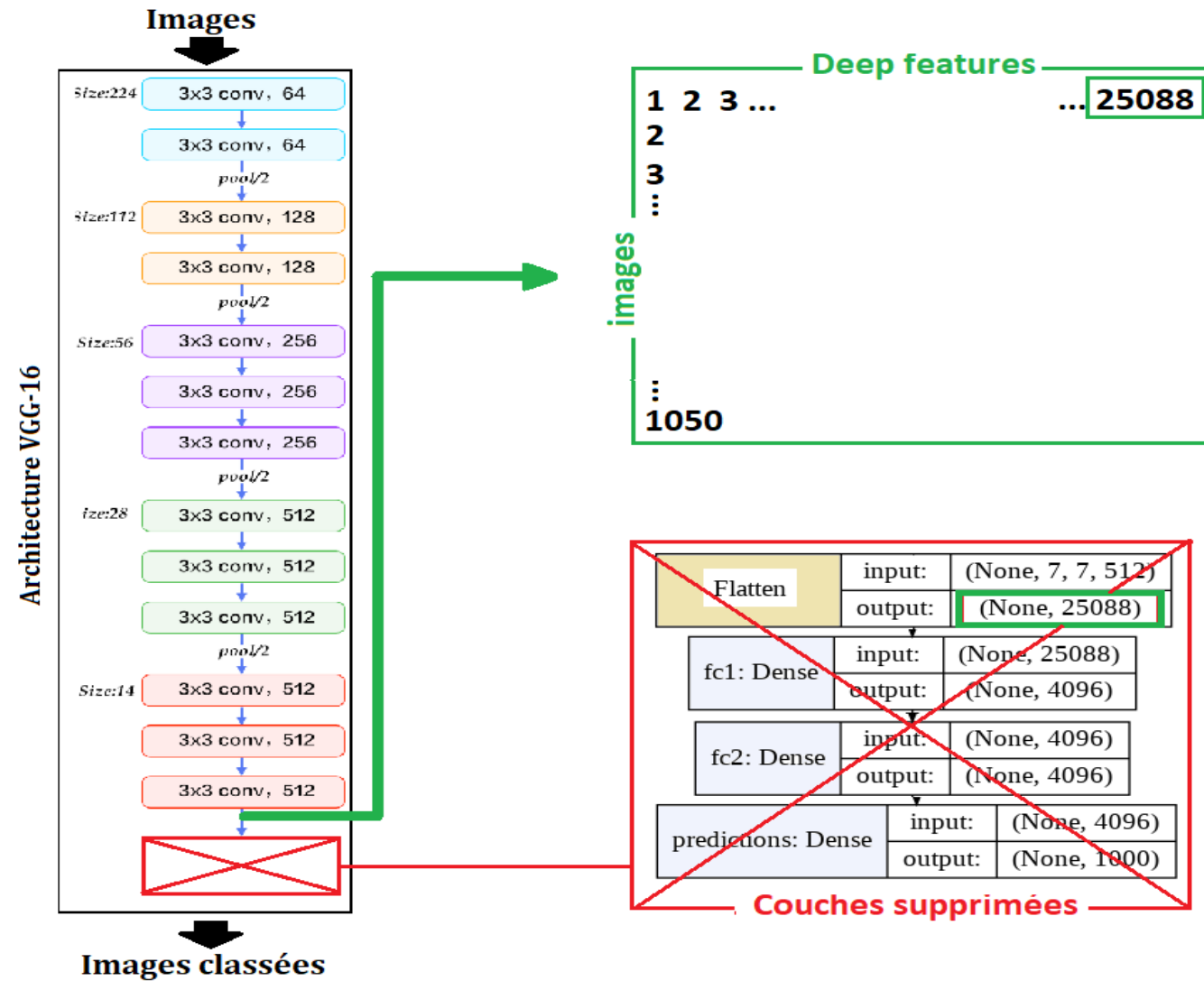
Nb produits
/ cluster

cluster	nb_produits
0	7
1	1
2	334
3	6
4	6
5	688
6	6

Rand Score ajusté : 0.0247

Matching Clustering Kmeans -
Classification humaine

Classification - CNN

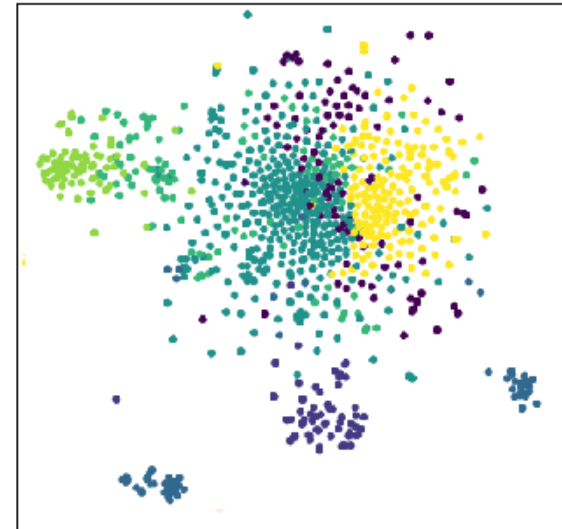


Résultat - CNN(2)

Matrice de confusion

Baby Care	13	0	1	33	4	0	99
Beauty and Personal Care	1	77	1	62	4	0	5
Computers	7	1	0	139	1	0	2
Home Decor & Festive Needs	30	0	0	80	30	1	9
Home Furnishing	58	0	0	25	0	0	67
Kitchen & Dining	0	1	79	54	16	0	0
Watches	0	0	0	13	61	76	0
	0	1	2	3	4	5	6
	Cluster						

Représentation t-SNE des produits



Nb produits
/ cluster

cluster	nb_produits
0	109
1	79
2	81
3	406
4	116
5	77
6	182

Rand Score ajusté : **0.2482**

Matching Clustering Kmeans -
Classification humaine

Résultats comparatifs - Conclusion

	Méthode de classification	Rand score ajusté
texte	Bag of Words	0,1593
	TF-Idf	0,3492 *
	LDA	0,1027
image	SIFT	0,0247
	CNN	0,2482 *

Variabilité du nombre de caractères selon la catégorie

