

Réaliser un dashboard et assurez
une veille technique

Plan

- Concevoir un dashboard de credit scoring
 - Cahier des charges
 - Déploiement du dashboard
 - Principaux graphiques
 - Démonstration du dashboard
- Veille technologique
 - Données utilisées
 - Méthodologie comparative
 - Résultats
 - Conclusion
 - Limites et pistes d'amélioration

Concevoir un dashboard de credit scoring

Contexte métier :

- Entreprise : Prêt à dépenser – société de crédit à la consommation ciblant des clients ayant peu ou pas d'historique de prêt.
- Objectif : Améliorer la transparence des décisions de scoring crédit auprès des clients.

Problématique :

- Comment expliquer simplement et de manière transparente les décisions d'acceptation ou de refus de crédit ?
- Comment rendre l'outil utile, lisible et accessible pour les chargés de relation client dans leurs rendez-vous ?



Cahier des charges

Objectif :

- Créer un dashboard interactif pour expliquer les décisions de crédit.
- Fonctionnalités attendues :
 - Afficher la probabilité de remboursement (score et classe prédite).
 - Donner une interprétation compréhensible du score.
 - Présenter les informations clés du client.
 - Comparer un client avec l'ensemble ou un sous-groupe de clients (filtres).
 - Assurer une accessibilité (critères WCAG).
 - Dashboard connecté à une API de scoring.
 - Déploiement sur une plateforme Cloud.



Modèle de scoring et API existante

Modèle de scoring crédit :

- Modèle LightGBM optimisé
- Prédit la probabilité de remboursement
- Classe la demande : acceptée ou refusée

API de scoring :

- Déployée sur Azure Cloud
- Intégrée dans un pipeline MLOps automatisé
- Tests automatisés et CI/CD via GitHub Actions
- Mise à jour continue après chaque push

Connexion au dashboard :

- Récupération des prédictions en temps réel
- Base pour l'affichage interactif

un modèle → une API → le dashboard → l'utilisateur

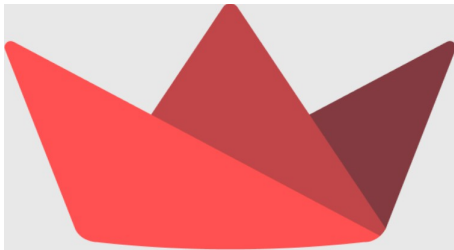
Déploiement du dashboard

Technologies utilisées :

- Streamlit pour le développement
- Docker pour le conteneur
- Azure pour l'hébergement
- Connexion à l'API FastAPI existante

Méthode :

- Création d'une image Docker
- Déploiement sur Azure
- Dashboard disponible via une URL publique



Streamlit – Dashboard



Docker – Conteneur



Azure – Hébergement

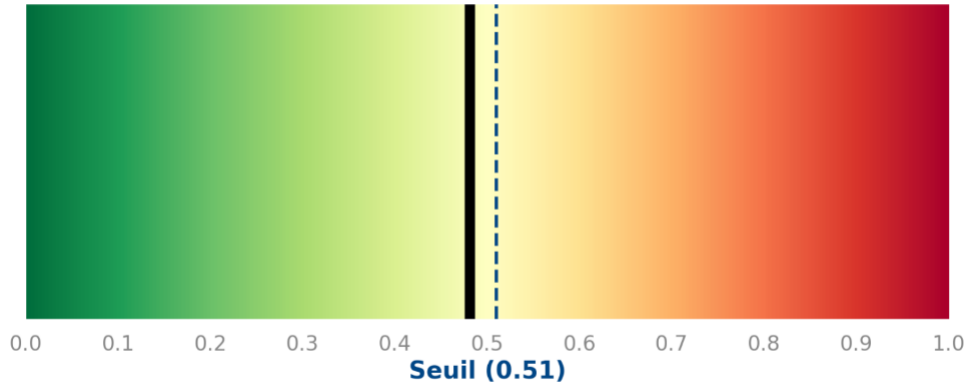


FastAPI – API

Principaux graphiques du dashboard

Probabilité de défaut du client

0.481



Le crédit est **accordé**. La probabilité de défaut (0.481) est en-dessous du seuil (0.51).

- **Jauge de score**

Visualisation de la probabilité de défaut du client et décision d'acceptation/refus selon le seuil.

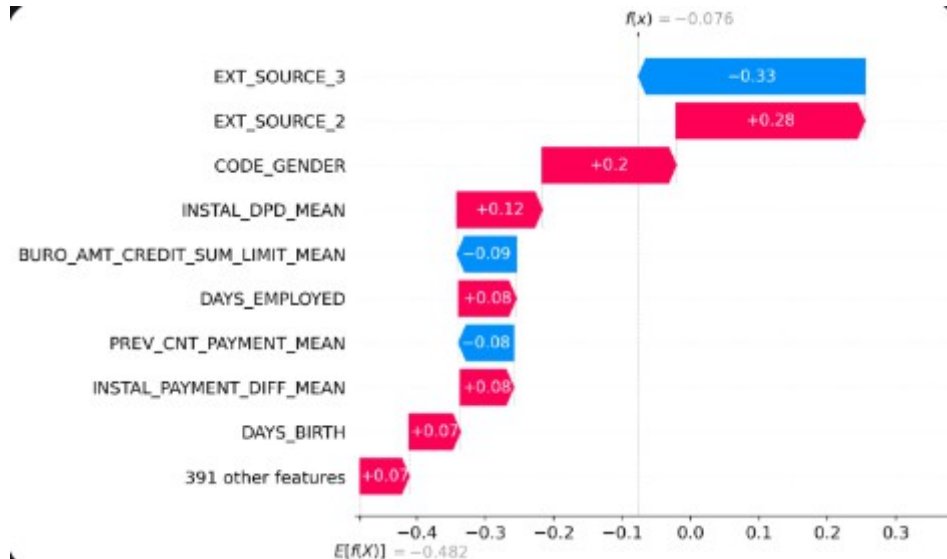
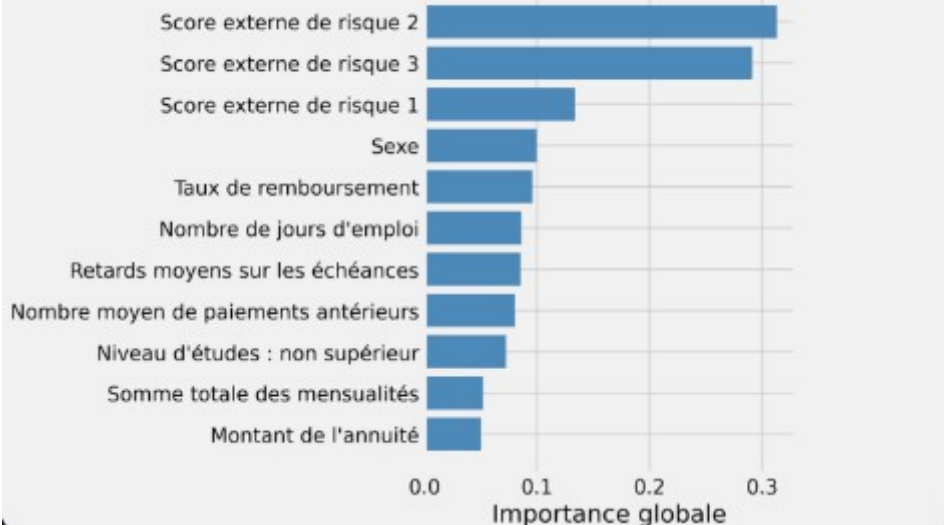
- **Variables les plus importantes (global)**

Classement des variables qui influencent le plus le modèle sur l'ensemble des clients.

- **Explication individuelle (SHAP)**

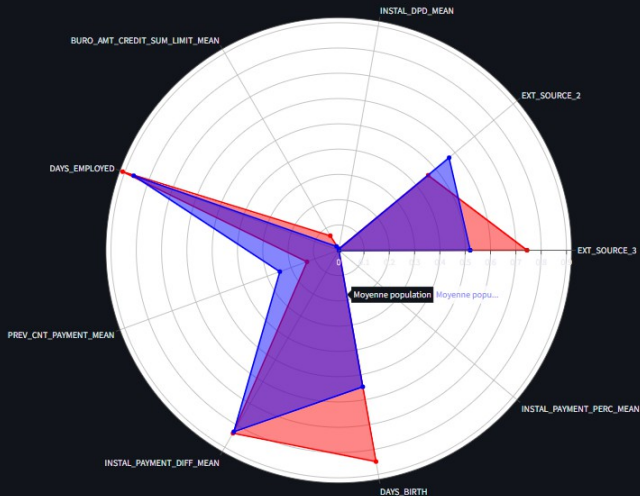
Contribution de chaque variable à la décision pour ce client précis.

Variables les plus importantes pour le modèle



Principaux graphiques du dashboard

Synthèse multi-variables du profil client



Nuage de points : EXT_SOURCE_3 vs EXT_SOURCE_2



- **Profil multi-variables du client**

Ce radar compare les principales caractéristiques du client à la moyenne de la population.

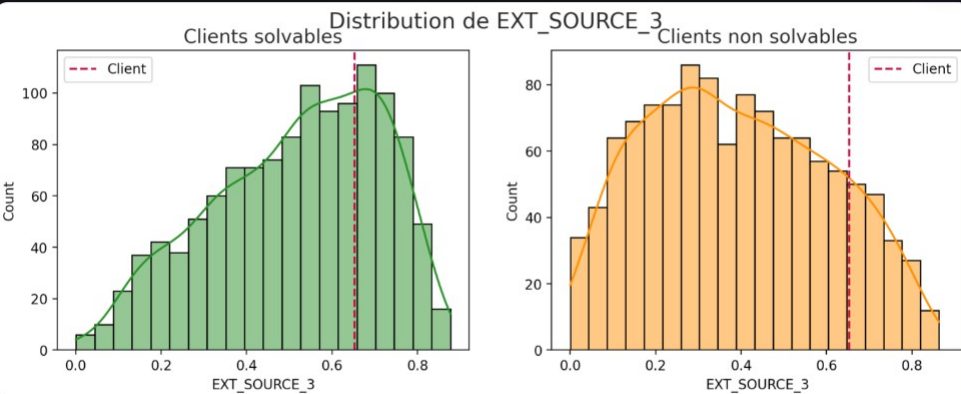
- **Distribution d'une variable clé**

Ces histogrammes montrent la répartition d'une variable clé chez les clients solvables et non solvables.

- **Positionnement du client dans la population**

Ce nuage de points permet de visualiser le client sélectionné (en rouge) parmi l'ensemble des clients selon deux variables clés.

EXT_SOURCE_3



La **ligne rouge** indique la valeur du client sélectionné.

Démonstration du dashboard interactif



Accédez au dashboard ici :

<https://projet8-dashboard-h9ancgetd4cmbjdr.westeurope-01.azurewebsites.net/>

Conclusion & Limites

Conclusion

- Un outil interactif et intuitif : Permet aux chargés de relation client d'expliquer clairement les décisions d'octroi de crédit.
- Explicabilité renforcée : Grâce aux visualisations (jauge, SHAP, spider plot, comparaisons) et aux explications textuelles, même les non-experts comprennent les décisions.
- Accessibilité prise en compte : Contraste élevé, redimensionnement du texte, légendes explicatives : le dashboard est utilisable par un large public, y compris les personnes malvoyantes.
- Connexion en temps réel à l'API de scoring : Les prédictions sont toujours à jour, ce qui garantit la fiabilité des décisions affichées.
- Déploiement cloud sur Azure : Accessibilité simple depuis n'importe quel poste de travail.

Limites

- Accessibilité : certains graphiques restent difficiles à interpréter pour les utilisateurs de lecteurs d'écran.
- Explications simplifiées : la compréhension fine du modèle peut rester complexe pour les non-experts

Introduction à la veille technologique

Contexte et objectifs

Mission : Veille technologique

Objectif

- Réaliser un état de l'art sur une technique récente (moins de 5 ans) en NLP ou vision par ordinateur, à partir de sources reconnues.

Démarche

- Tester et comparer cette technique à une approche classique sur des données déjà exploitées dans l'entreprise.
- Présenter une preuve de concept et une note méthodologique.

Enjeux

- Anticiper les évolutions technologiques et renforcer la capacité d'innovation de l'entreprise.
- Identifier des solutions performantes et adaptées au contexte métier.

Données utilisées

Objectif : classer automatiquement les produits selon leur catégorie principale, à partir de leur description textuelle uniquement.

- Jeu de données issu du site e-commerce Flipkart
- 1 050 produits, 15 colonnes (textuelles, numériques, catégorielles)
- Colonnes clés :
 - uniq_id : identifiant produit
 - product_name : nom du produit
 - description : description textuelle (utilisée pour le NLP)
 - product_category_tree : catégorie hiérarchique (variable cible)
- Classification sur le premier niveau de catégories (7 classes équilibrées)
- Seule la colonne description est utilisée pour la modélisation

Répartition des produits par niveau_categorie_1 :	
niveau_categorie_1	
Home Furnishing	150
Baby Care	150
Watches	150
Home Decor & Festive Needs	150
Kitchen & Dining	150
Beauty and Personal Care	150
Computers	150

Méthodologie comparative

Modèles comparés

- DistilBERT : version légère de BERT, plus rapide, adaptée à la production
- DeBERTa : architecture plus récente, réputée pour de meilleures performances

Approche

- Pas de fine-tuning complet
- Extraction des embeddings à partir de la description produit
- Comparaison uniquement sur la qualité des représentations textuelles

Tâches évaluées

- Classification non supervisée : K-Means pour tester la cohérence des regroupements
- Classification supervisée : Régression logistique pour prédire la catégorie

DistilBERT vs DeBERTa

Modèle	Principe & Architecture	Innovations clés	Extraction des embeddings	Apprentissage / Pré-entraînement	Points forts principaux
DistilBERT	Version distillée et allégée de BERT (architecture Transformer)	<ul style="list-style-type: none"> - Plus léger et rapide - Conserve le contexte bidirectionnel - Tokenization WordPiece 	<ul style="list-style-type: none"> - Embeddings contextuels pour chaque token - Utilisation du token [CLS] ou moyenne des tokens 	<ul style="list-style-type: none"> - Apprentissage par distillation : entraîné pour imiter BERT tout en réduisant la taille et la complexité - Utilise les mêmes données de pré-entraînement que BERT 	<ul style="list-style-type: none"> - Extraction rapide et efficace - Idéal pour des usages nécessitant rapidité et faible empreinte mémoire
DeBERTa	Évolution de BERT avec attention désentrelacée (Microsoft, 2020/ v3 en 2021).	<ul style="list-style-type: none"> - Séparation explicite contenu/position - Attention désentrelacée - Enhanced Mask Decoder - Pré-entraînement optimisé (v3) 	<ul style="list-style-type: none"> - Embeddings contextuels enrichis - Séparation contenu/position - Utilisation du token [CLS] ou moyenne des tokens 	<ul style="list-style-type: none"> - Pré-entraînement optimisé : stratégies inspirées d'ELECTRA (v3), meilleure utilisation des données - Intègre plus finement la position lors du pré-entraînement - Apprentissage plus efficace et performant, même avec moins de données 	<ul style="list-style-type: none"> - Représentations plus robustes et discriminantes - Surpasse BERT/DistilBERT sur la plupart des benchmarks - Meilleure compréhension des relations sémantiques et syntaxiques

Méthodologie de modélisation

Prétraitement des données

- DistilBERT : nettoyage simple, passage en minuscules, tokenisation automatique du modèle.
- DeBERTa : conservation de la casse, suppression des espaces multiples, tokenisation propre au modèle.
- Pas de suppression des stopwords ni de lemmatisation (on laisse les modèles gérer le contexte).

Extraction des caractéristiques

- Utilisation de DistilBERT ou DeBERTa pré-entraînés pour transformer chaque description produit en un vecteur d'embedding contextuel.
- Récupération de l'embedding du token [CLS] pour représenter chaque texte.

Réduction de dimension et visualisation

- Réduction de la dimension avec une ACP (Analyse en Composantes Principales) pour garder 99 % de la variance.
- Visualisation des données en 2D avec t-SNE.

Modélisation

- Clustering (non supervisé) : K-Means sur les embeddings réduits, pour regrouper les produits par similarité.
- Classification (supervisée) : Régression logistique sur les embeddings, prédiction de la catégorie principale.

Métriques d'évaluation

Pour le clustering (non supervisé)

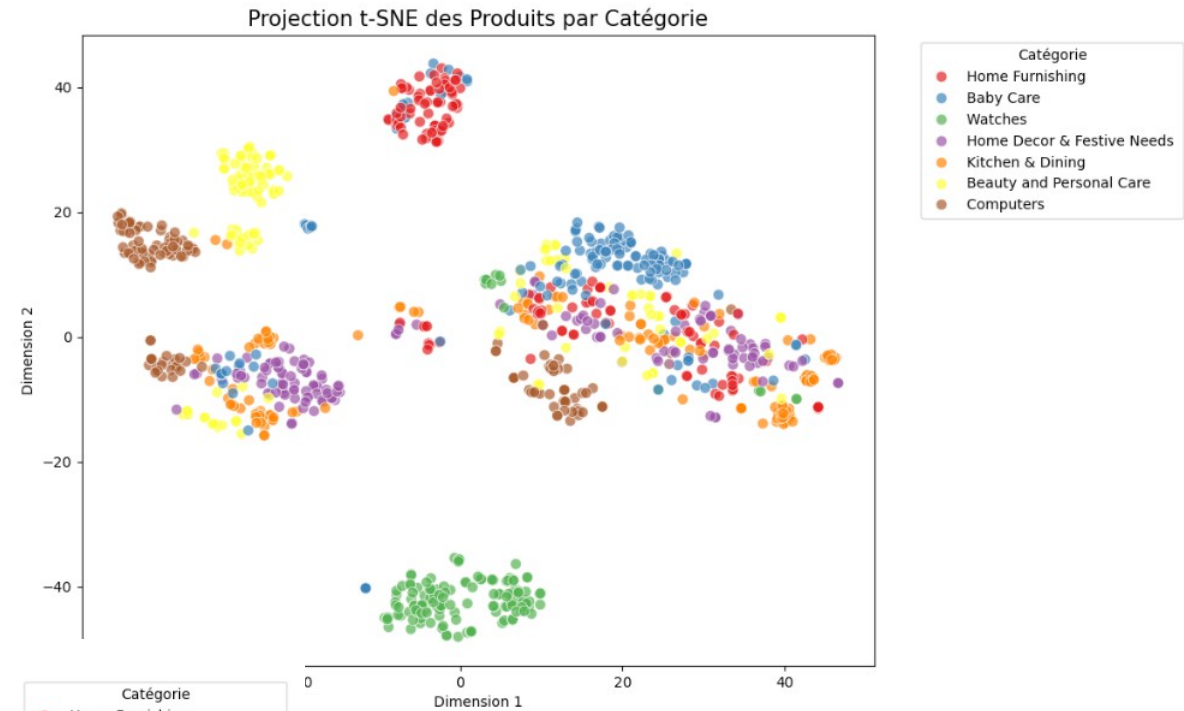
- Homogeneity Score : chaque cluster contient des éléments d'une seule classe réelle.
- Completeness Score : tous les éléments d'une même classe sont dans le même cluster.
- V-measure : synthèse des deux scores précédents.
- Adjusted Rand Index (ARI) : compare le clustering obtenu à la vraie répartition des classes.
- Silhouette Score : mesure la cohérence interne des clusters.

Pour la classification (supervisée)

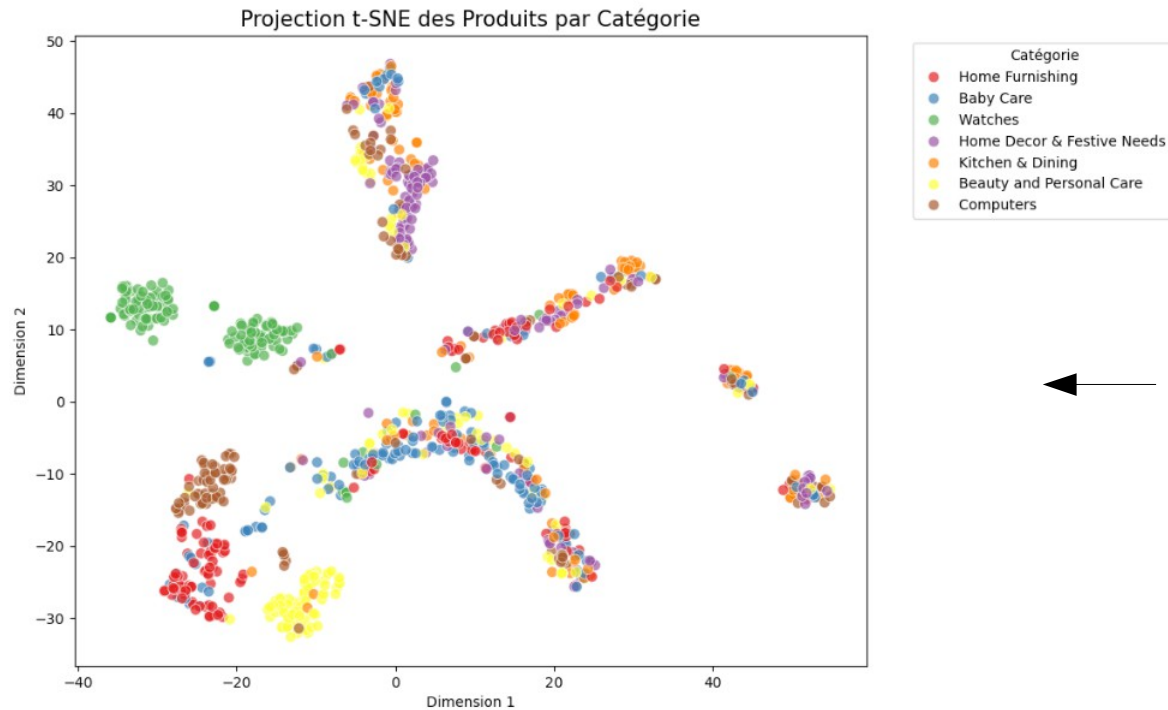
- Accuracy : proportion de bonnes prédictions.
- Précision, rappel, F1-score : évaluation détaillée par classe.
- Moyennes macro et pondérée : pour évaluer la performance globale.

Visualisation des embeddings par t-SNE

DistilBERT →



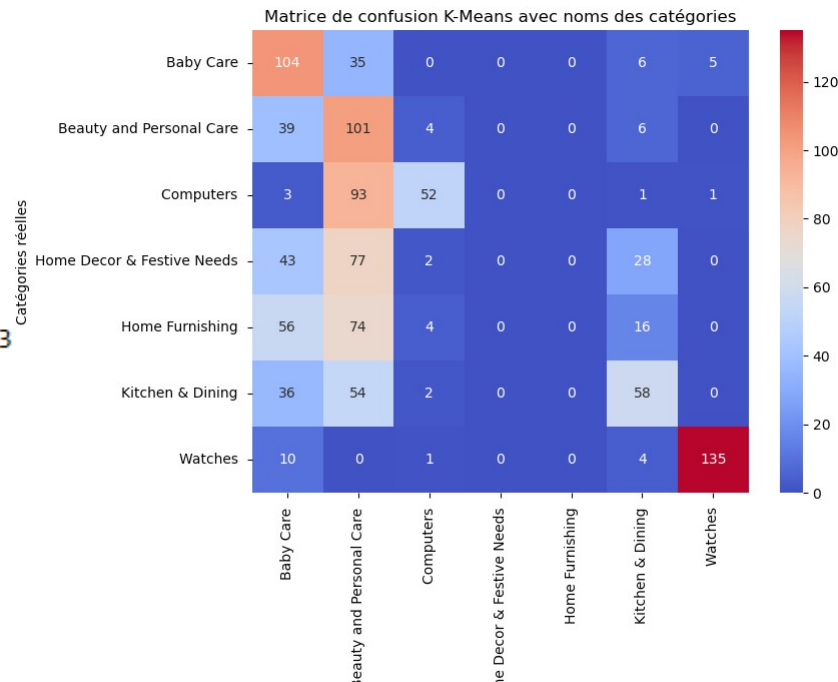
← DeBERTa



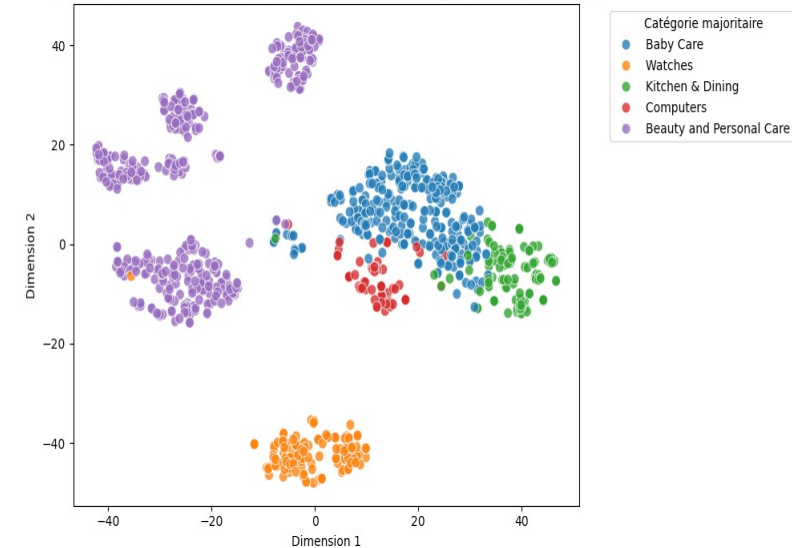
Résultats du clustering

DistilBERT

Homogeneity Score : 0.288
Completeness Score : 0.334
V-measure : 0.31
Adjusted Rand Index : 0.173
Silhouette Score : 0.161

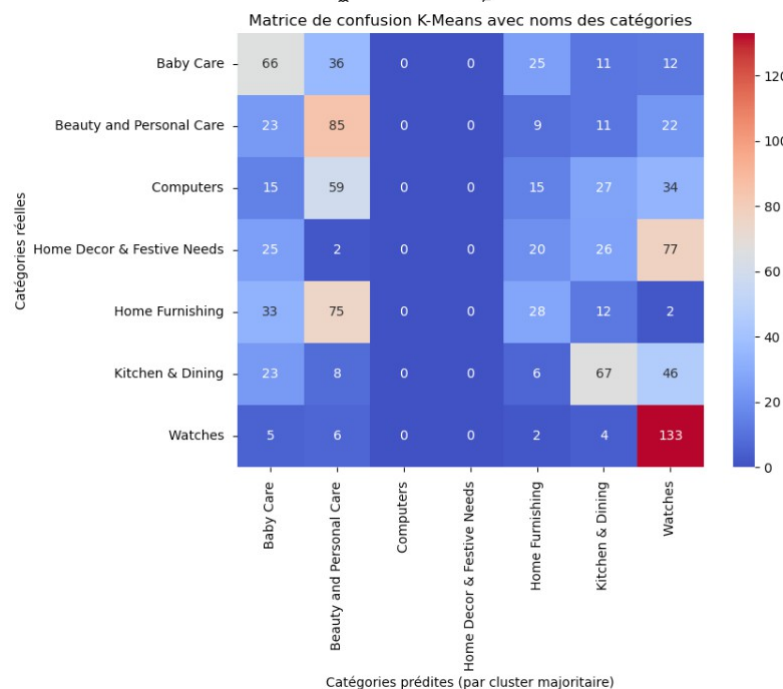


Visualisation t-SNE des Clusters K-Means (catégories majoritaires)

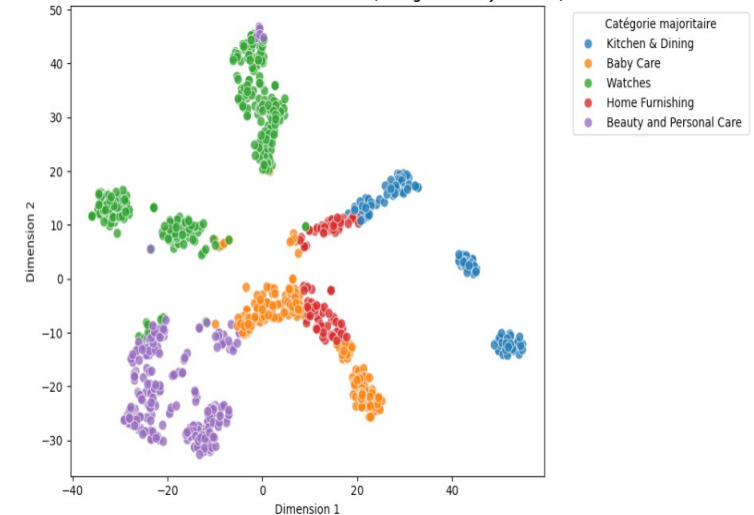


DeBERTa

Homogeneity Score : 0.182
Completeness Score : 0.201
V-measure : 0.191
Adjusted Rand Index : 0.144
Silhouette Score : 0.313



Visualisation t-SNE des Clusters K-Means (catégories majoritaires)



Résultats de la classification supervisée

DistilBERT

Accuracy sur l'ensemble d'entraînement : 0.9798

Accuracy sur l'ensemble de test : 0.9095

Rapport de classification (test) :

	precision	recall	f1-score	support
Baby Care	0.81	0.87	0.84	30
Beauty and Personal Care	0.93	0.87	0.90	30
Computers	1.00	1.00	1.00	30
Home Decor & Festive Needs	0.81	0.87	0.84	30
Home Furnishing	0.87	0.90	0.89	30
Kitchen & Dining	0.96	0.87	0.91	30
Watches	1.00	1.00	1.00	30
accuracy			0.91	210
macro avg	0.91	0.91	0.91	210
weighted avg	0.91	0.91	0.91	210



DeBERTa

Accuracy sur l'ensemble d'entraînement : 0.6798

Accuracy sur l'ensemble de test : 0.5524

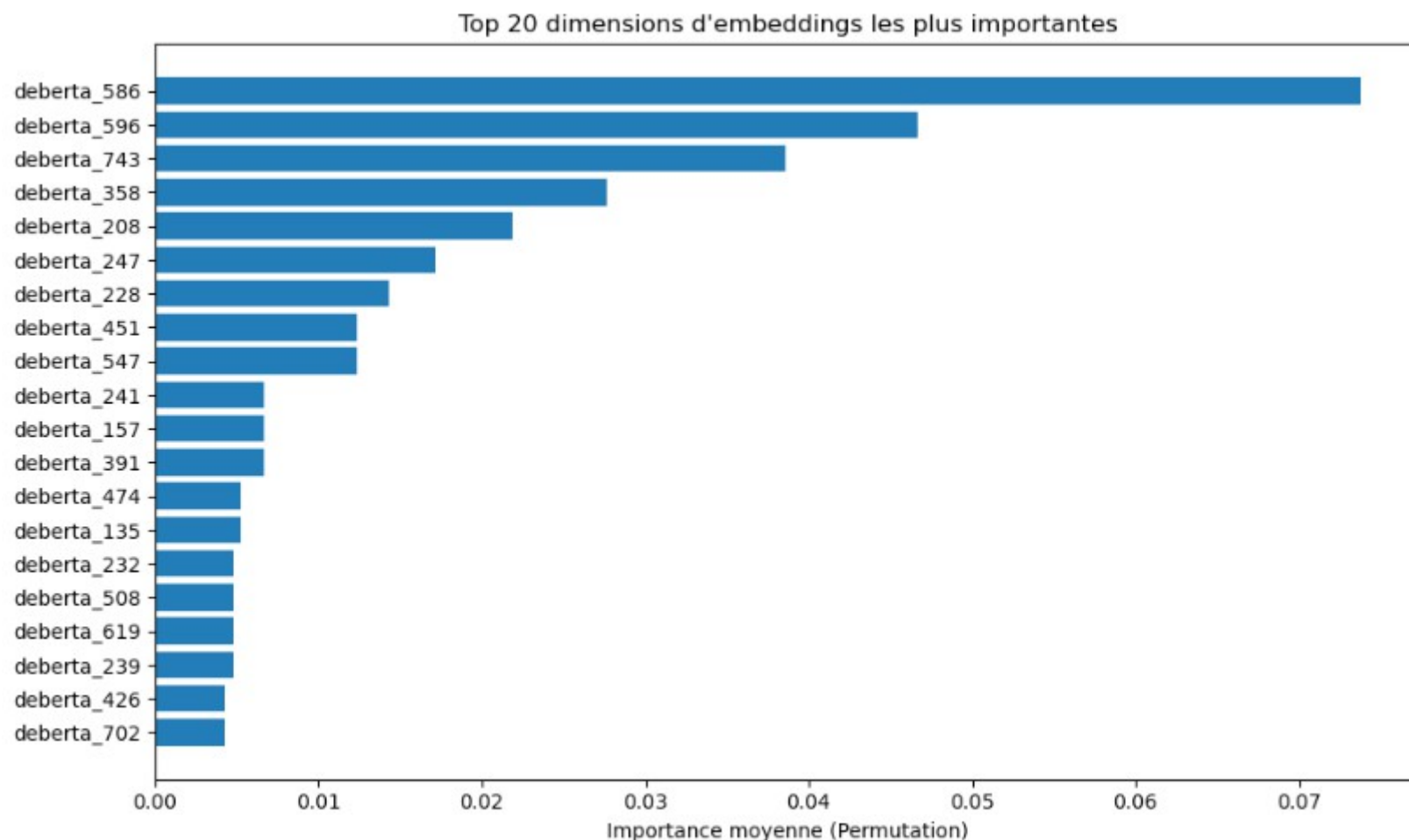
Rapport de classification (test) :

	precision	recall	f1-score	support
Baby Care	0.38	0.50	0.43	30
Beauty and Personal Care	0.94	0.50	0.65	30
Computers	0.58	0.47	0.52	30
Home Decor & Festive Needs	0.32	0.40	0.36	30
Home Furnishing	0.71	0.57	0.63	30
Kitchen & Dining	0.40	0.57	0.47	30
Watches	0.96	0.87	0.91	30
accuracy			0.55	210
macro avg	0.61	0.55	0.57	210
weighted avg	0.61	0.55	0.57	210



Importance globale des features

- Seules quelques dimensions des embeddings ont un vrai impact sur la classification.
- La majorité des dimensions sont peu utiles : l'information est répartie sur plusieurs axes.
- Permet d'identifier les composantes vraiment exploitées par le modèle.



Importance locale des features

- Analyse locale avec SHAP pour expliquer chaque prédiction individuellement.
- Certaines dimensions des embeddings influencent fortement la décision pour un exemple précis.
- Les dimensions importantes varient selon les textes, mais on retrouve souvent celles identifiées comme importantes globalement.
- Permet de mieux comprendre pourquoi le modèle classe un produit dans telle ou telle catégorie.

===== Explication locale de la prédiction =====

Texte analysé :

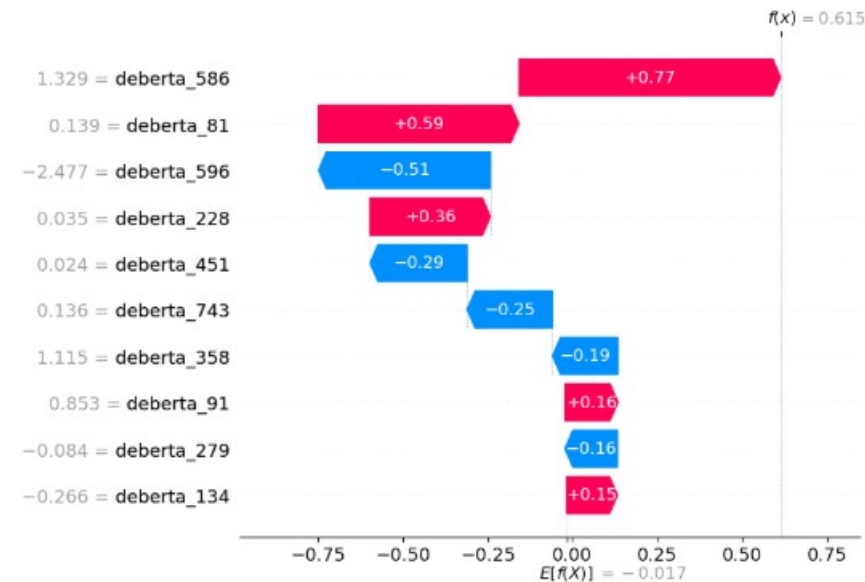
Key Features of Eurospa Cotton Terry Face Towel Set Size: small Height: 9 inch GSM: 360,Eurospa Cotton Terry Face Towel Set (20 PIECE FACE TOWEL SET, Assorted) Price: Rs. 299 Eurospa brings to you an exclusively designed, 100% soft cotton towels of export quality. All our products have soft texture that takes care of your skin and gives you that enriched feeling you deserve. Eurospa has been export...

Classe réelle : Baby Care

Classe prédite : Kitchen & Dining

Top 10 dimensions de l'embedding DeBERTa les plus influentes :

- deberta_586: 0.771
- deberta_81: 0.591
- deberta_596: -0.507
- deberta_228: 0.355
- deberta_451: -0.286
- deberta_743: -0.252
- deberta_358: -0.194
- deberta_91: 0.157
- deberta_279: -0.156
- deberta_134: 0.150



===== Explication locale de la prédiction =====

Texte analysé :

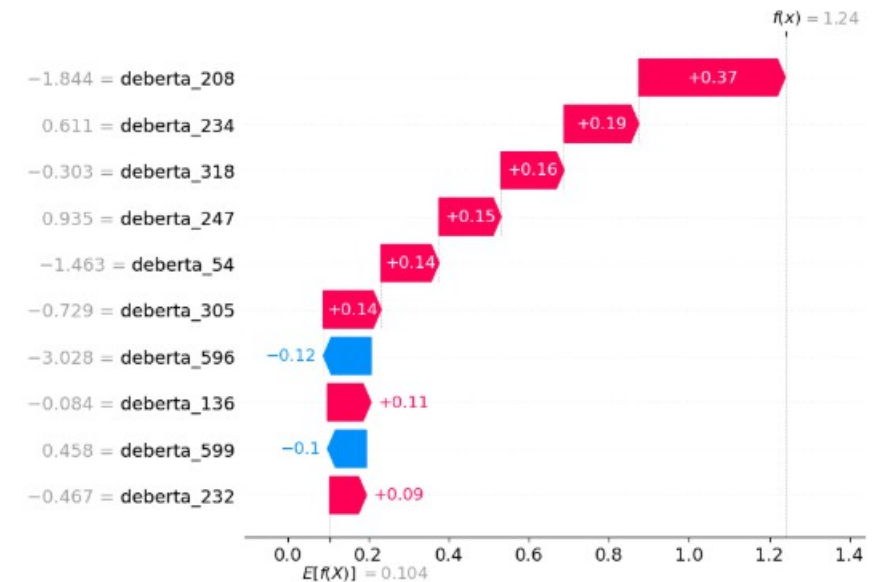
Key Features of Elegance Polyester Multicolor Abstract Eyelet Door Curtain Floral Curtain,Elegance Polyester Multicolor Abstract Eyelet Door Curtain (213 cm in Height, Pack of 2) Price: Rs. 899 This curtain enhances the look of the interiors.This curtain is made from 100% high quality polyester fabric.It features an eyelet style stitch with Metal Ring.It makes the room environment romantic and lov...

Classe réelle : Home Furnishing

Classe prédite : Baby Care

Top 10 dimensions de l'embedding DeBERTa les plus influentes :

- deberta_208: 0.365
- deberta_234: 0.186
- deberta_318: 0.157
- deberta_247: 0.155
- deberta_54: 0.144
- deberta_305: 0.144
- deberta_596: -0.119
- deberta_136: 0.108
- deberta_599: -0.097
- deberta_232: 0.092



Conclusion

- Ce POC a permis de comparer l'efficacité de DistilBERT et DeBERTa pour la classification automatique de descriptions produits.
- DistilBERT fournit de très bonnes performances et généralise bien sur les données de test.
- DeBERTa, dans la configuration actuelle, donne des résultats plus faibles.
- L'analyse globale et locale des features montre que, même en cas d'erreur, le modèle s'appuie sur des signaux pertinents pour ses décisions.
- Cette compréhension ouvre la voie à des améliorations ciblées pour de futurs tests.

Limites et pistes d'amélioration

Limites

- Les dimensions des embeddings restent difficiles à interpréter directement.
- Les performances de DeBERTa sont perfectibles, sans doute liées à la méthode d'extraction ou au manque d'adaptation au domaine.
- Certaines erreurs sont dues à des descriptions ambiguës ou proches entre catégories.
- Le classificateur linéaire (régression logistique) ne capture pas les relations complexes.
- Le contexte métier et les métadonnées ne sont pas exploités.

Améliorations possibles

- Tester d'autres stratégies d'agrégation des embeddings (ex : moyenne des tokens).
- Réduire la dimensionnalité pour simplifier l'interprétation.
- Utiliser des modèles de classification plus complexes (SVM, Random Forest, réseaux de neurones).
- Affiner les modèles par du fine-tuning sur le corpus cible.
- Enrichir et nettoyer les données (plus de métadonnées, descriptions plus précises).
- Explorer d'autres méthodes d'explicabilité (LIME, attention).
- Analyser plus finement les erreurs pour guider les prochaines améliorations.