

CLASSIFICATION DES IMAGES DE FASHION_MNIST



UNIVERSITÉ IBN KHALDOUN

Projet intégré multidisciplinaire
Préparé en vue de valider un module universitaire

ÉLABORÉ PAR :

AMEN ALAH FELAH
&
AYOUB HANNACHI

ENCADRÉ PAR :

MME. AMAL JLASSI
&
MME. NAWRES KHLIFA

ANNÉE UNIVERSITAIRE 2024-2025

Sommaire



01

Contexte Général

02

Description
du Dataset &
Prétraitement
des données

03

Algorithmes de
Classification

04

Division du dataset
& augmentation des
données

05

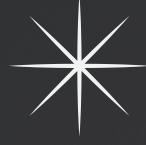
Suppression de la
Classe Chemise

06

Suppression de la Classe
Chemise avec Division du dataset

07

Conclusion &
Perspectives



1

CONTEXTE GÉNÉRAL



Dans ce projet, on explore l'application d'algorithmes de classification sur le dataset Fashion_MNIST.

L'objectif est de déterminer le modèle le plus efficace en termes de précision, robustesse et simplicité d'implémentation.



2

DÉSCRIPTION DU DATASET & PRÉTRAITEMENT DES DONNÉES

* DESCRIPTION DU DATASET

- Fashion MNIST : 70 000 images (60 000 entraînement, 10 000 test).
- Résolution : 28x28 pixels.

* CHALLENGES

- Fashion MNIST présente des images pixellisées et floues, rendant l'apprentissage plus complexe.
- Ce manque de propreté nécessite des prétraitements avancés pour des bonnes performances.



* PRÉTRAITEMENT DES DONNÉES

- **Normalisation** : Conversion des pixels en valeurs entre 0 et 1 pour une meilleure convergence.
- **Transformation** : Images 28x28 transformées en vecteurs de 784 dimensions pour les algorithmes classiques.



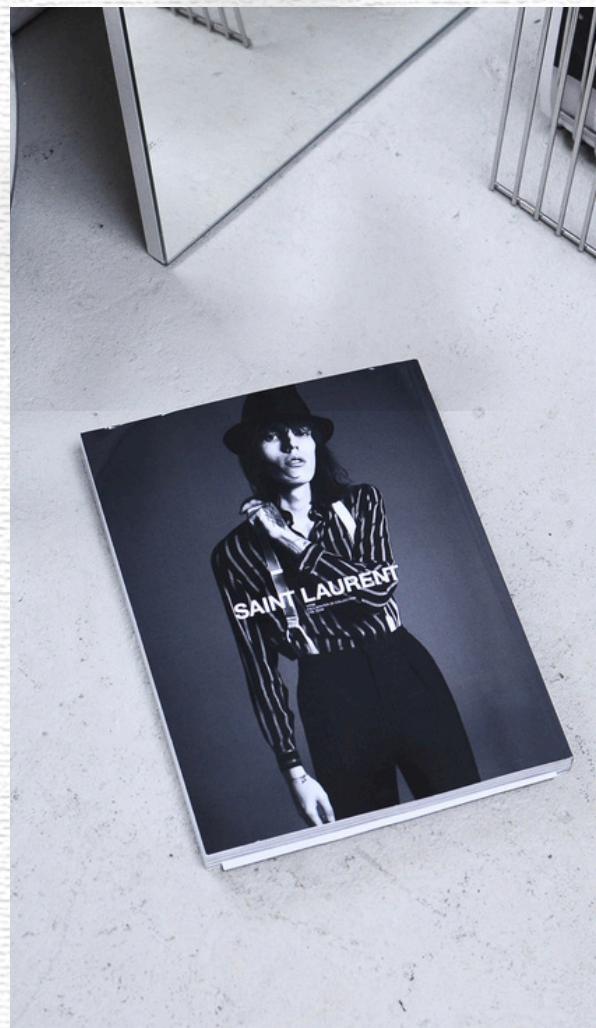


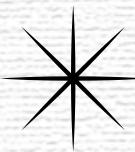
3

ALGORITHMES DE CLASSIFICATION

* ALGORITHMES DE CLASSIFICATION

- **SVM (Support Vector Machines)** : efficace pour la séparation des classes avec des marges maximales.
- **XGBoost** : un algorithme d'ensemble basé sur le boosting, connu par sa performance et sa vitesse.
- **Random Forest** : construit sur des arbres de décision pour réduire la variance et améliorer la robustesse.
- **KNN (K-Nearest Neighbors)** : un algorithme basé sur la proximité dans l'espace des caractéristiques.
- **Decision Tree** : utile pour identifier des règles décisionnelles.
- **Naïve Bayes** : Modèle probabiliste basé sur l'hypothèse d'indépendance des variables.
- **Logistic Regression** : efficace pour la classification binaire et multi-classes.





COMPARAISON ENTRE LES MODÈLES

Modèle	Précision (%)	Complexité	Temps d'entraînement	Temps de précision
Logistic Regression	84.35	Faible	Faible	Faible
KNN	85.41	Faible	Moyen	Élevé
Decision Tree	78.88	Modérée	Faible	Faible
SVM	88.28	Modérée	Élevé	Modéré
Naïve Bayes	58.56	Faible	Très faible	Très faible
Random Forest	87.64	Élevée	Élevé	Modéré
XGBoost	88.35	Élevée	Élevé	Faible



* OPTIMISATION DES HYPERPARAMÈTRES

SVM avec GridSearchCV

- GridSearchCV explore toutes les combinaisons d'hyperparamètres.
- Pour le SVM avec noyau RBF, les meilleurs résultats sont obtenus :
 - $C = 10$
 - $\text{gamma} = \text{'scale'}$

XGBoost avec RandomizedSearch

- RandomizedSearchCV échantillonne aléatoirement des combinaisons pour une optimisation plus rapide.





4

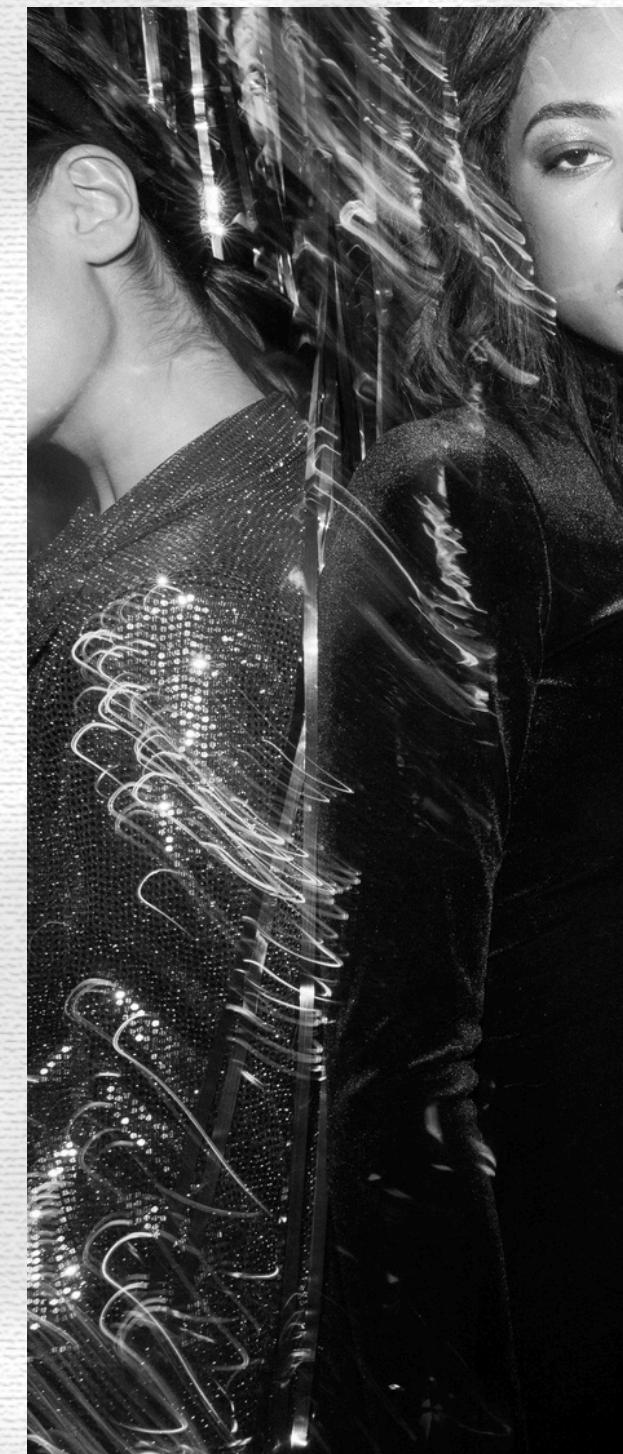
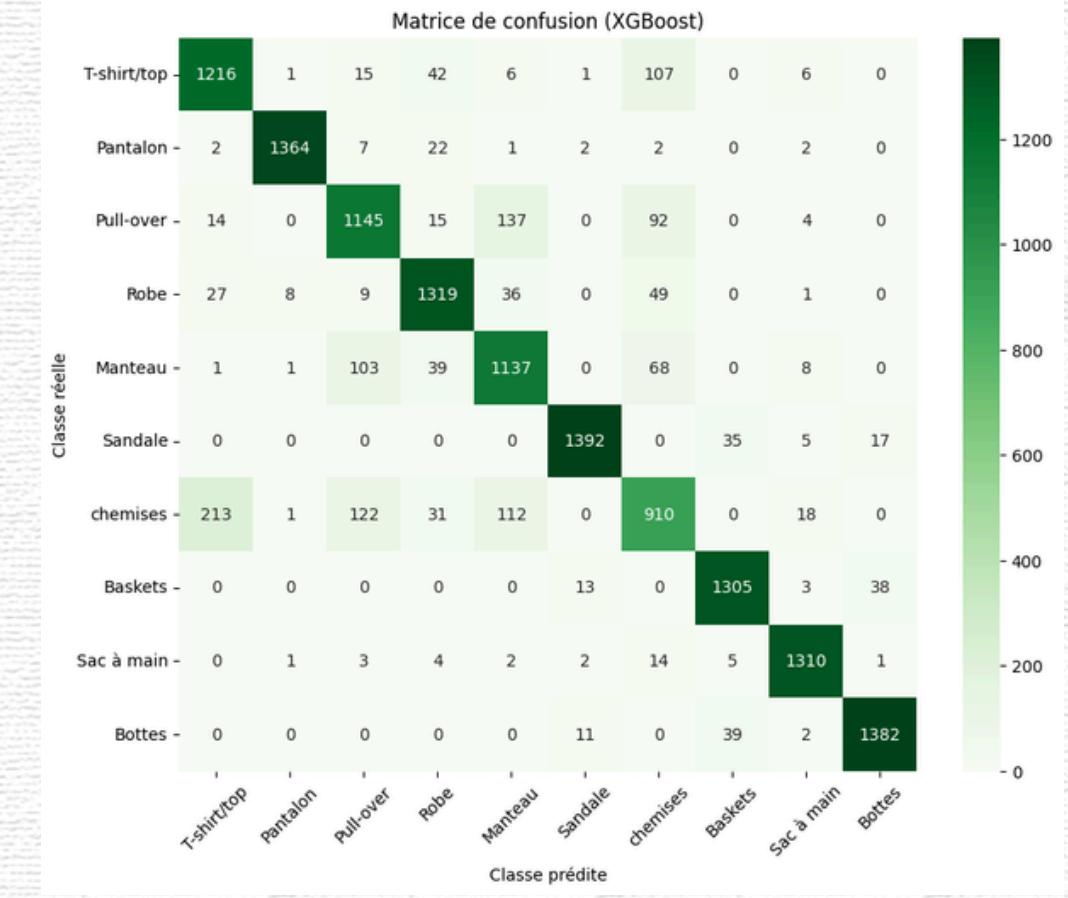
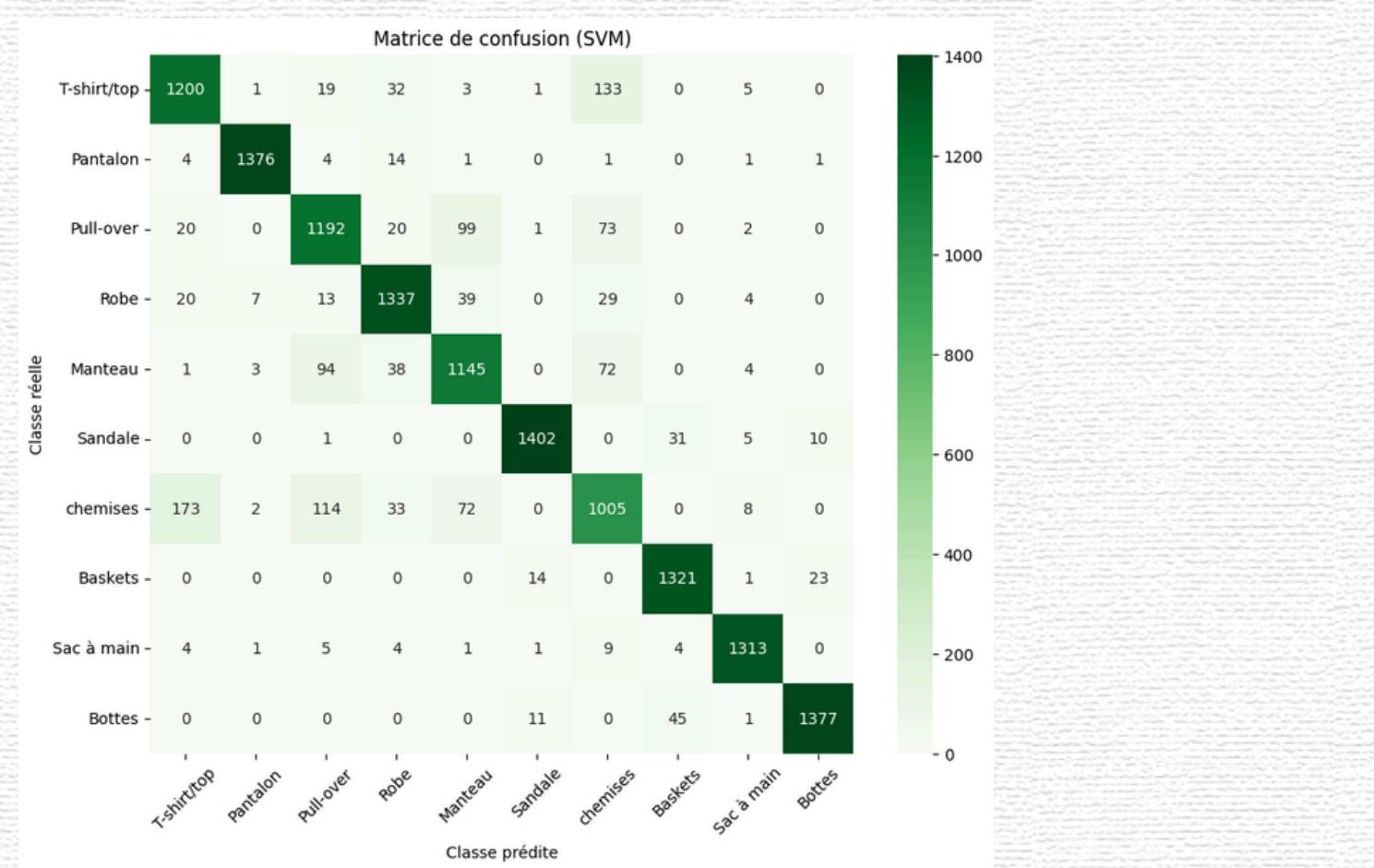
DIVISION DU DATASET & AUGMENTATION DES DONNÉES

* MOTIVATION ET RAISONS DE LA REDIVISION

La division du dataset en 80% pour l'entraînement et 20% pour le test offre un meilleur équilibre et une évaluation plus représentative que la séparation initiale de 85,71% et 14,29% (60 000 Train & 10 000 Test).

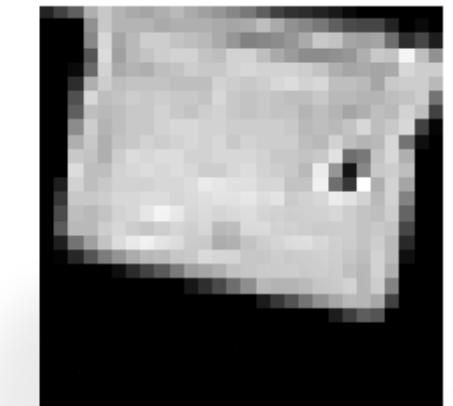
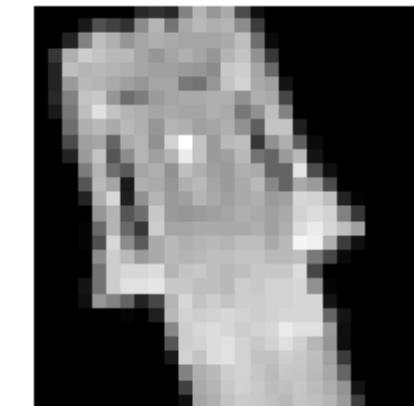
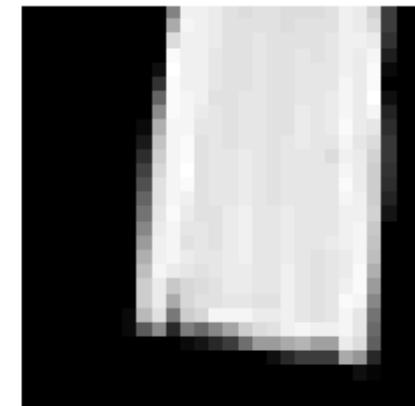
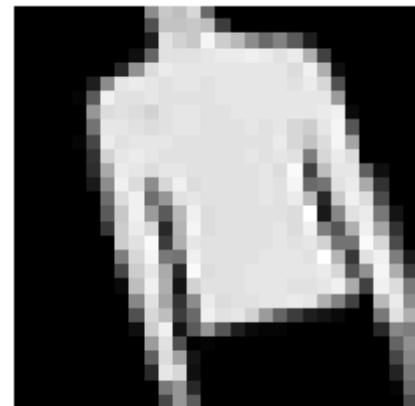
* RÉSULTATS APRÈS LA REDIVISION

- **SVM** atteint 90.94%
- **XGBoost** atteint 89.14%
- confirmant que cette répartition offre une évaluation plus équilibrée.



* AUGMENTATION DES DONNÉES

- L'augmentation des données, par rotations, translations et changements d'échelle, améliore la robustesse et la généralisation des modèles.



- Cependant, cette approche a réduit la précision des modèles, avec **SVM** à 67% et **XGBoost** à 69%. L'introduction de bruit a affecté négativement les performances des modèles.



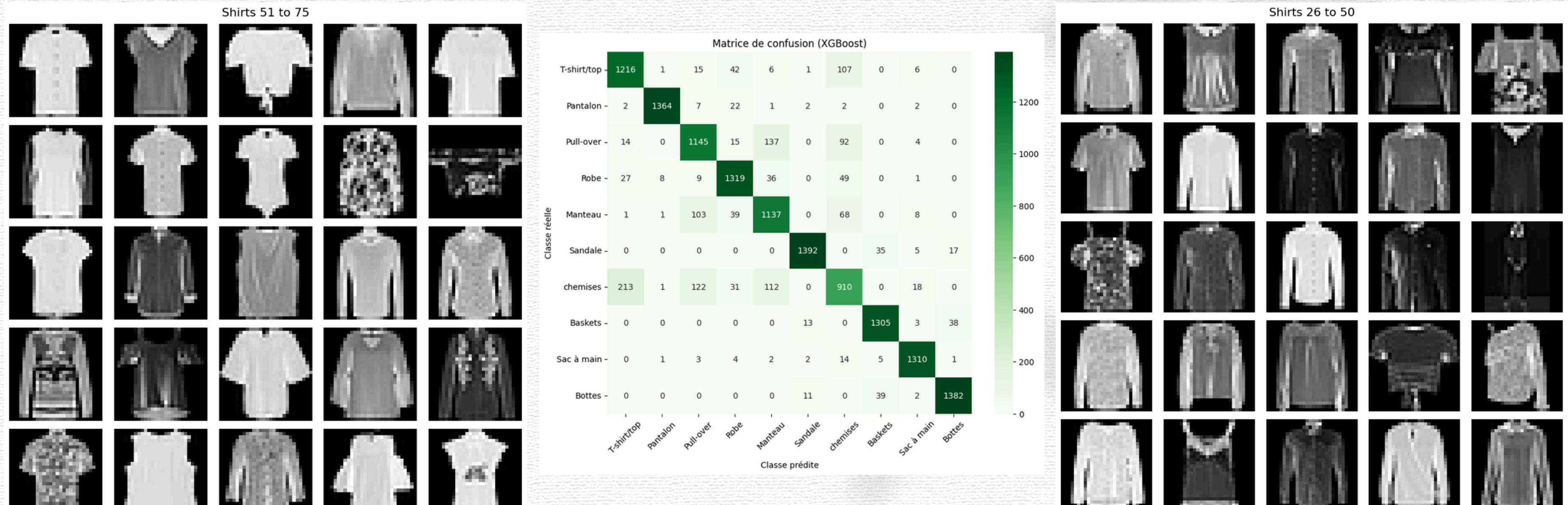


5

SUPPRESSION DE LA CLASSE CHEMISE

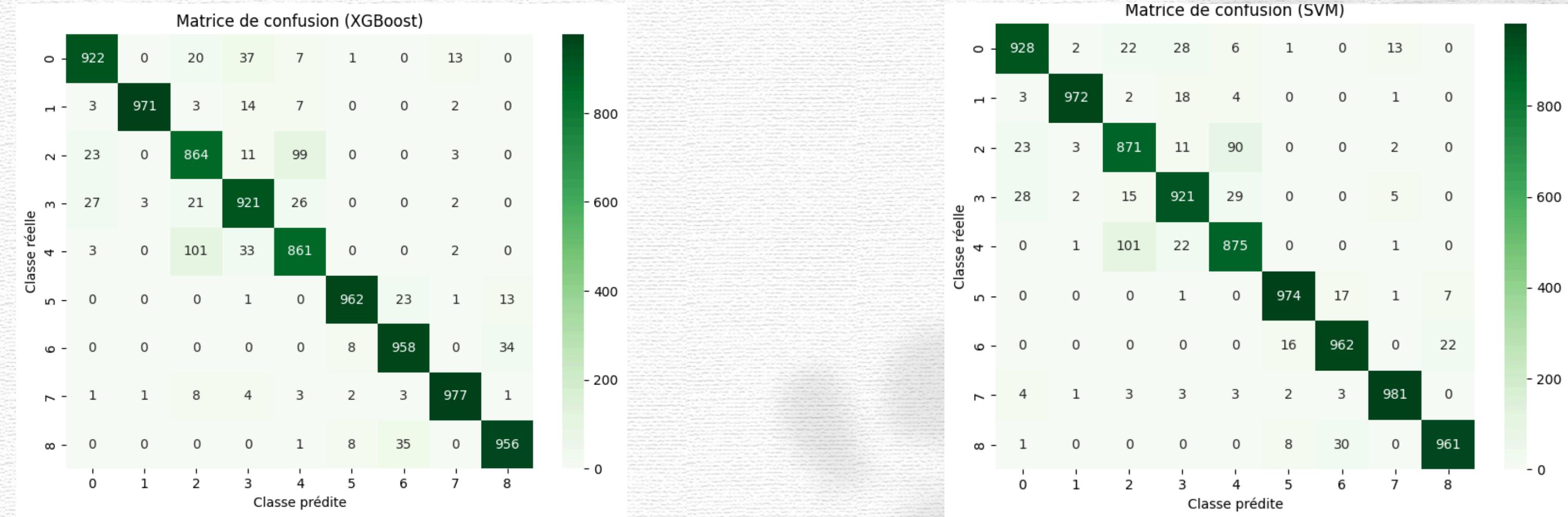
* MOTIVATION ET RAISONS DE LA SUPPRESSION

- Cette classe causait des confusions avec d'autres classes similaires, affectant la précision globale.
- Sa suppression a réduit les erreurs de classification.



RÉSULTATS APRÈS SUPPRESSION

Après suppression de la classe "chemise", **XGBoost** atteint 93.24% et **SVM** 93.83%, montrant une amélioration grâce à une meilleure distinction des autres classes.



donc il est préférable de combiner la suppression de la chemise avec la division 80/20 %.

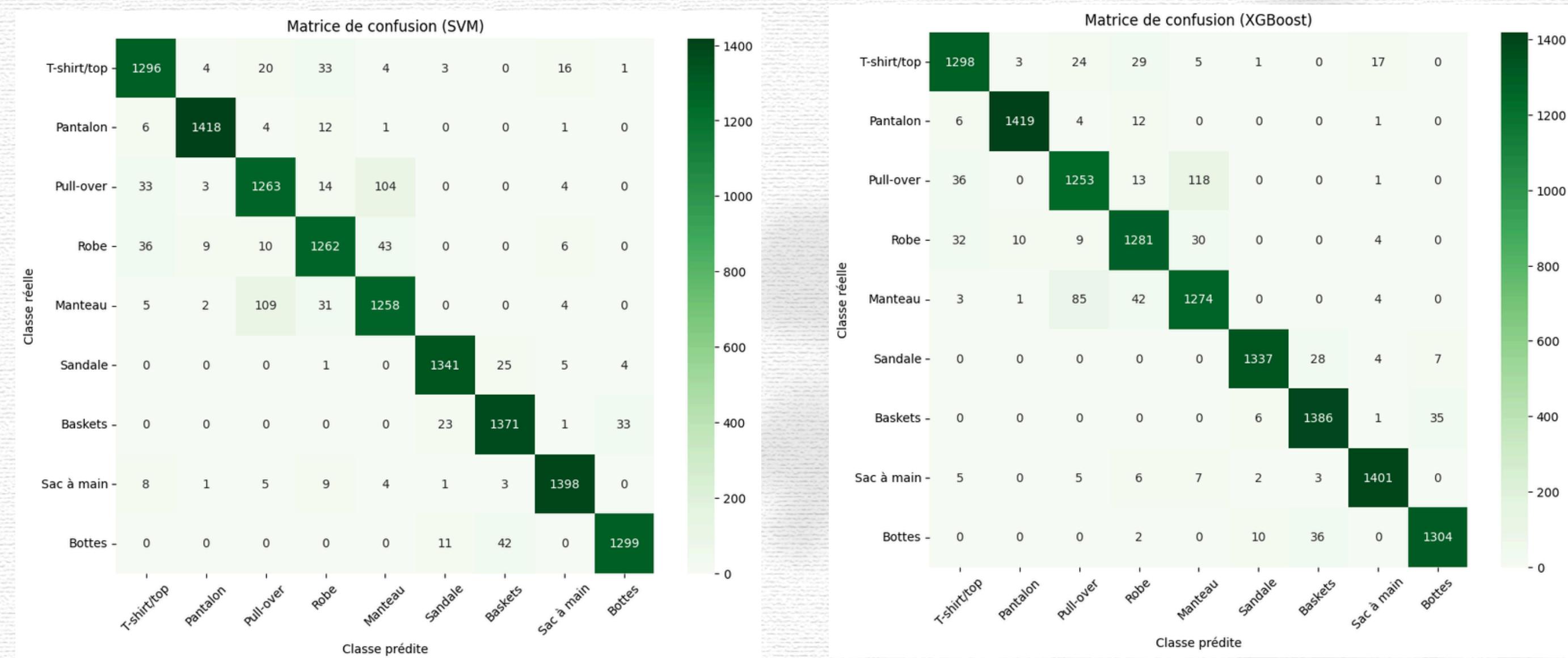


6

SUPPRESSION DE LA CLASSE CHEMISE AVEC DIVISION DU DATASET

RÉSULTATS FINAUX

Après suppression de la classe "chemise" et division 80% et 20%,
XGBoost atteint 94.86% et **SVM** 94.49% de précision, confirmant une meilleure généralisation des modèles.

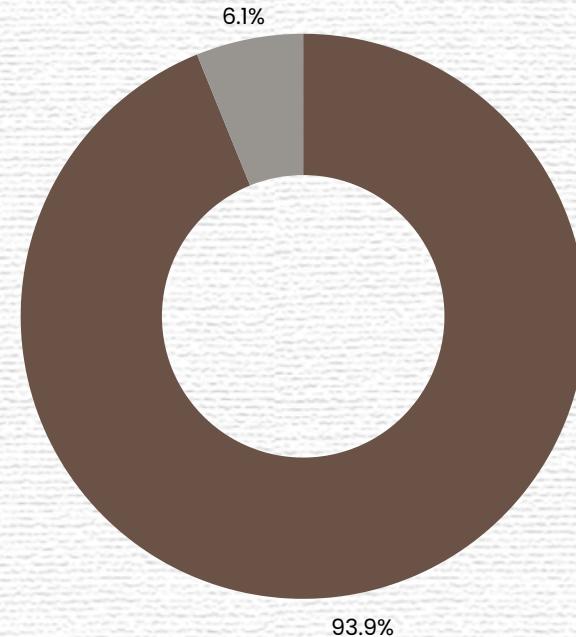




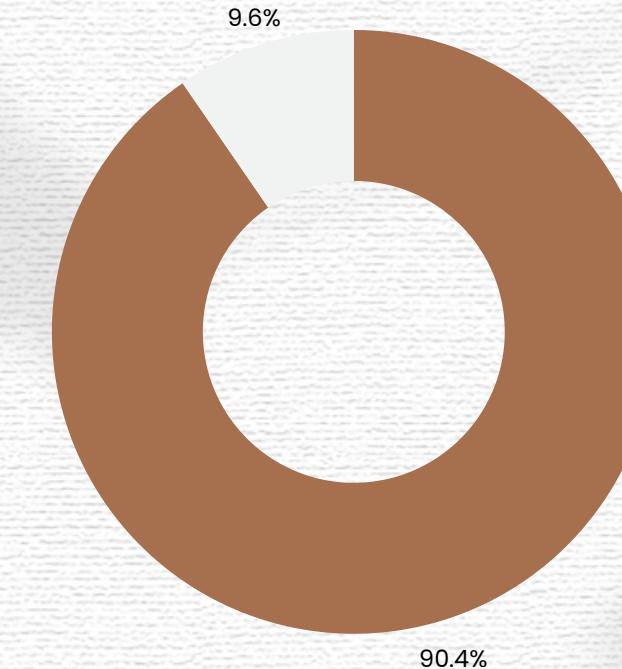
✳ VALIDATION CROISÉE

La validation croisée divise les données en plusieurs groupes. Le modèle s'entraîne sur certains groupes et est testé sur un autre, et on fait la moyenne des résultats pour voir sa performance.

SVM atteint 93.99% de score moyen



XGBoost atteint 94.21% de score moyen





7

CONCLUSION & PERSPECTIVES

★ CONCLUSION

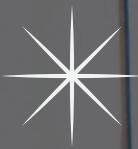
Dans ce projet, on a exploré la classification des images du dataset Fashion MNIST. Malgré l'absence de deep learning, nos modèles **SVM** et **XGBoost** ont montré une performance robuste grâce à un prétraitement soigné et des optimisations. Ces efforts ont permis de surpasser les autres travaux antérieurs sur le même dataset, démontrant l'efficacité des approches choisies.



* PERSPECTIVES

Créer une application AR qui recommande des vêtements selon la météo, les tendances mode, et les occasions spéciales. L'application propose des tenues adaptées et les affiche sur un avatar en réalité augmentée.

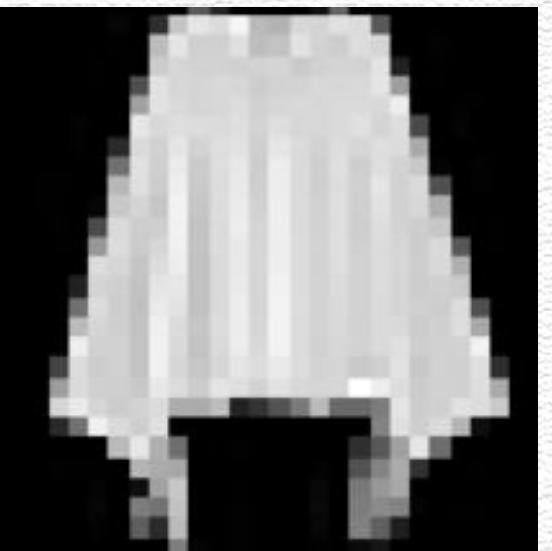




MERCI POUR
VOTRE
ATTENTION

* QUIZ ?

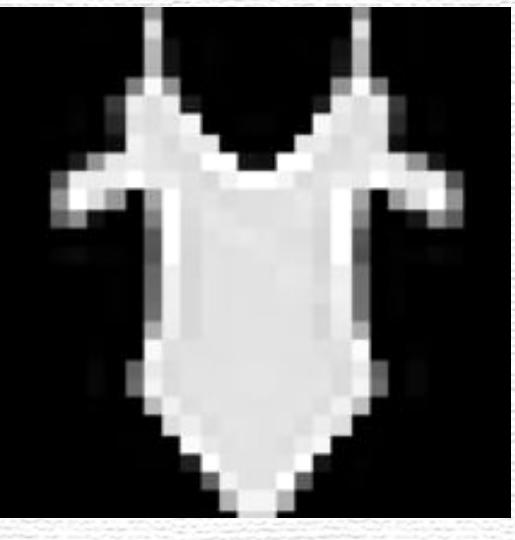
- T-shirt
- Pantalon
- Pull-over
- Robe
- Manteau
- chemise
- Sandale
- Basket
- Sac à main
- Bottes



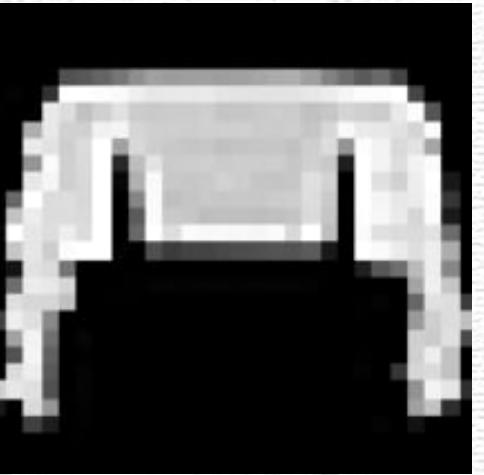
1



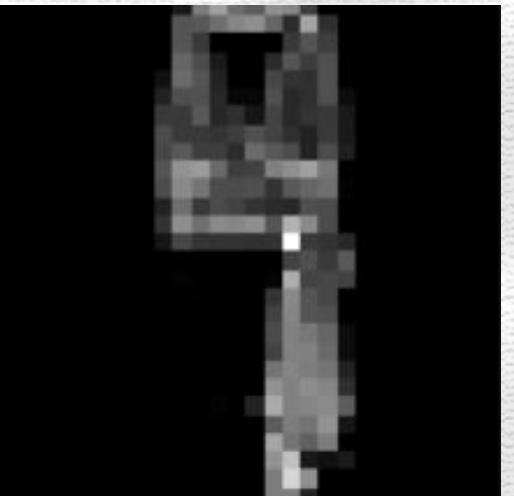
2



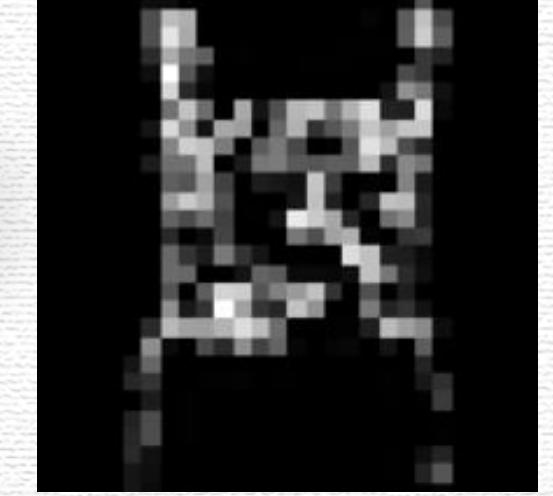
3



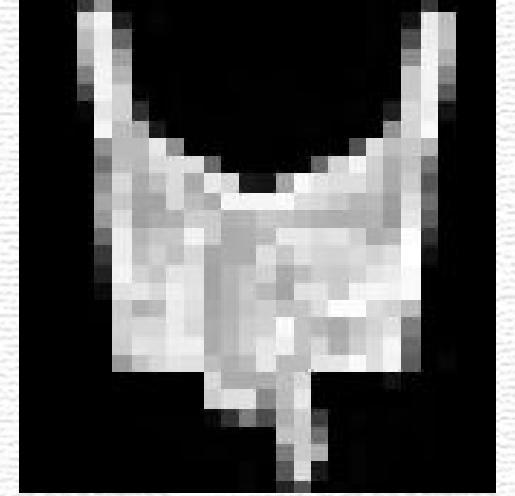
4



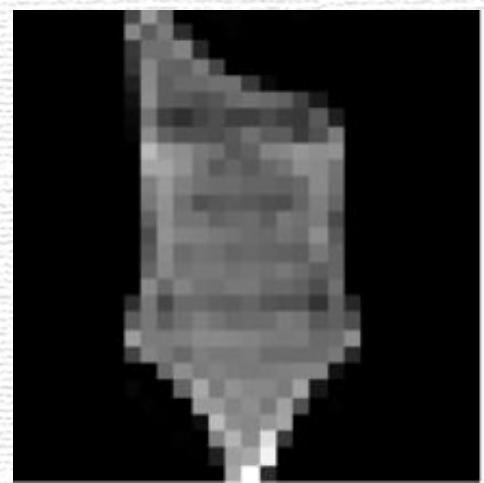
5



6



7

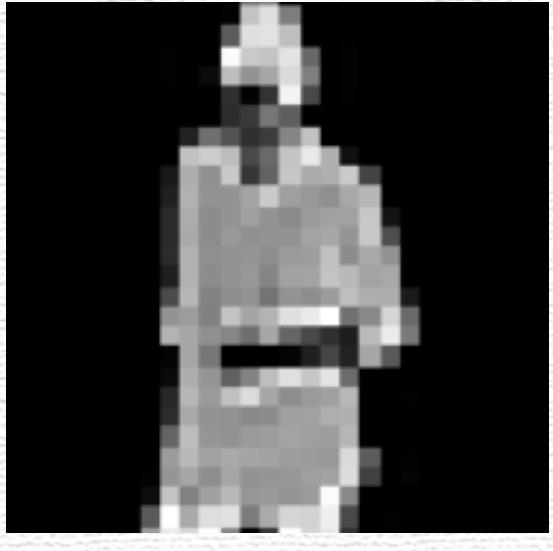


8

* QUIZ ?



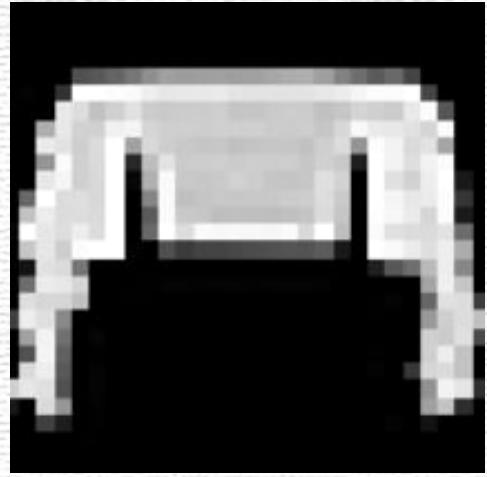
chemise



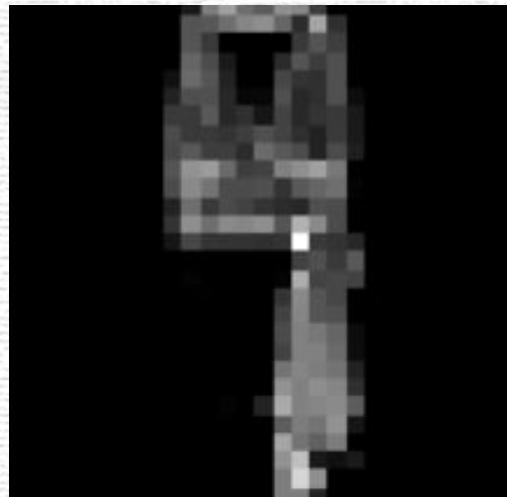
chemise



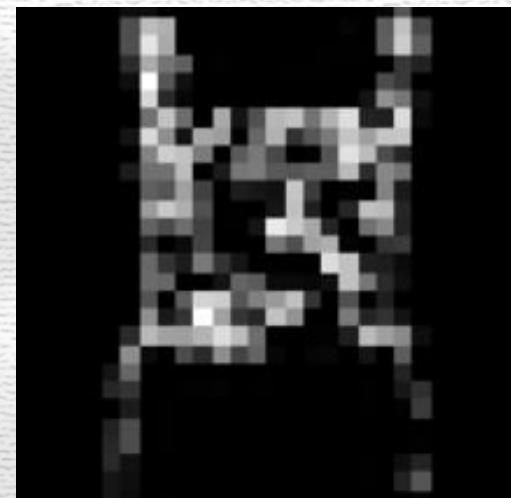
chemise



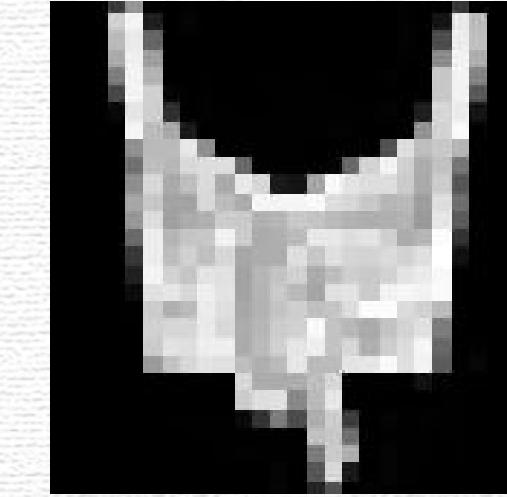
chemise



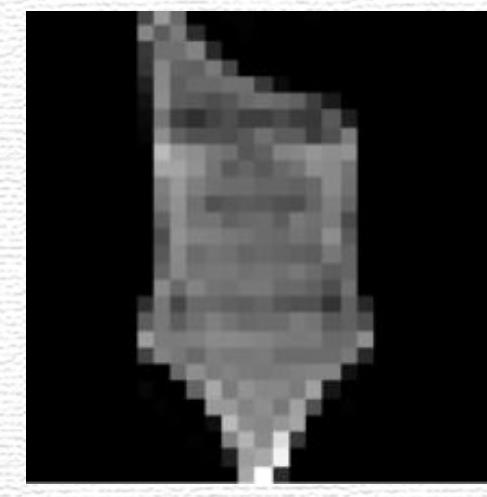
chemise



chemise



chemise



chemise