



Sujet



Déterminer si le mail est un spam



Maïmouna BACAR



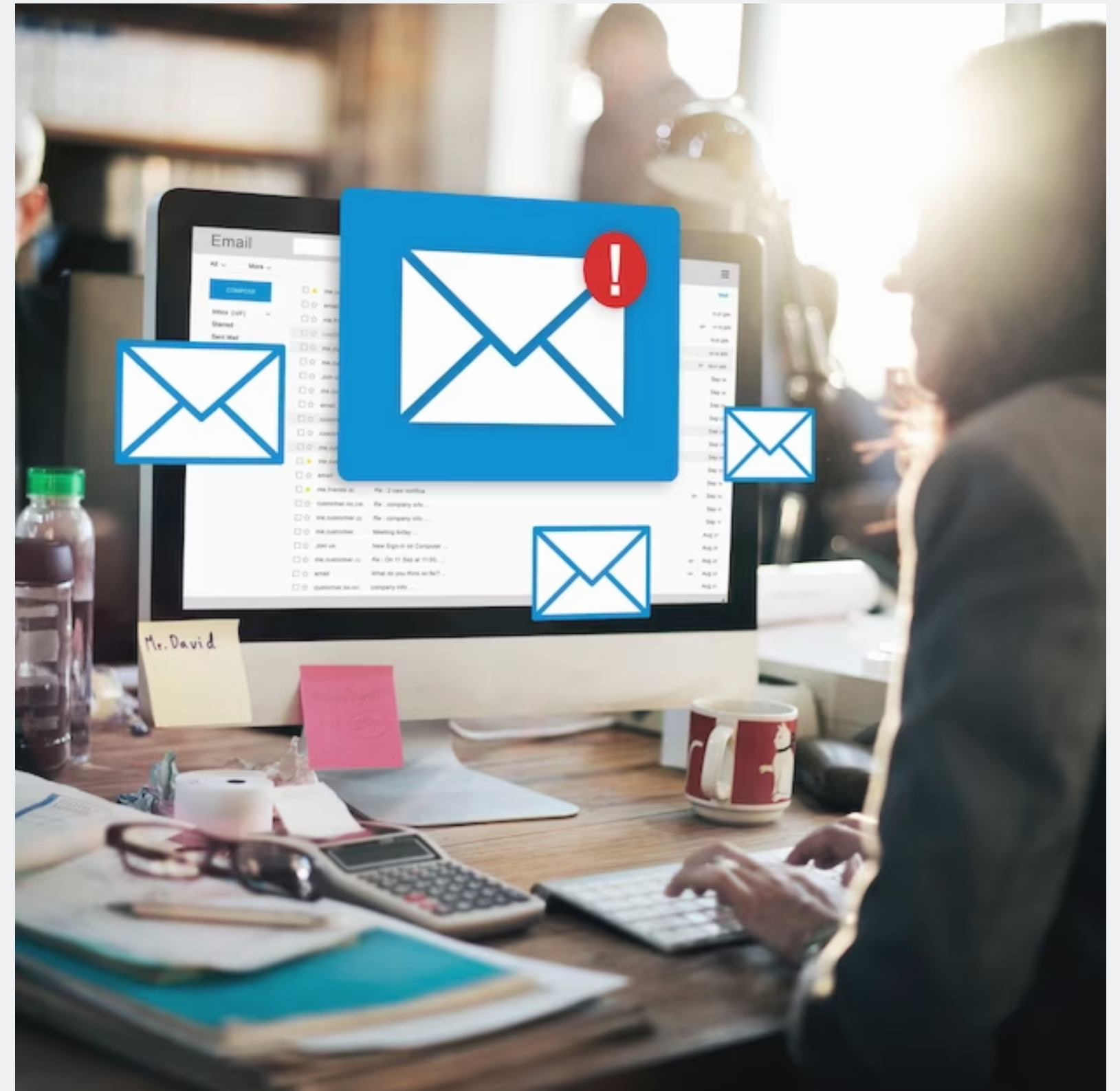
Sujet

Introduction



Objectif

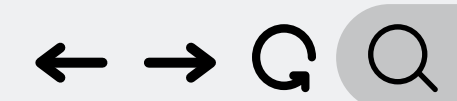
- Classification des emails indésirables
- Développer un Réseau de Neurones Artificiels





Sujet

Introduction



Plan

- Présentation des données
- ANN
- Amélioration ANN
- Conclusion



Sujet

Introduction

Données



← → 🔍 <https://www.kaggle.com/datasets/balaka18/email-spam-classification-dataset-csv>



Présentation des données

Email anonyme

Fréquence de mots

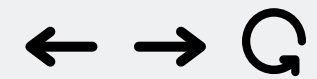
Spam ou non



Sujet

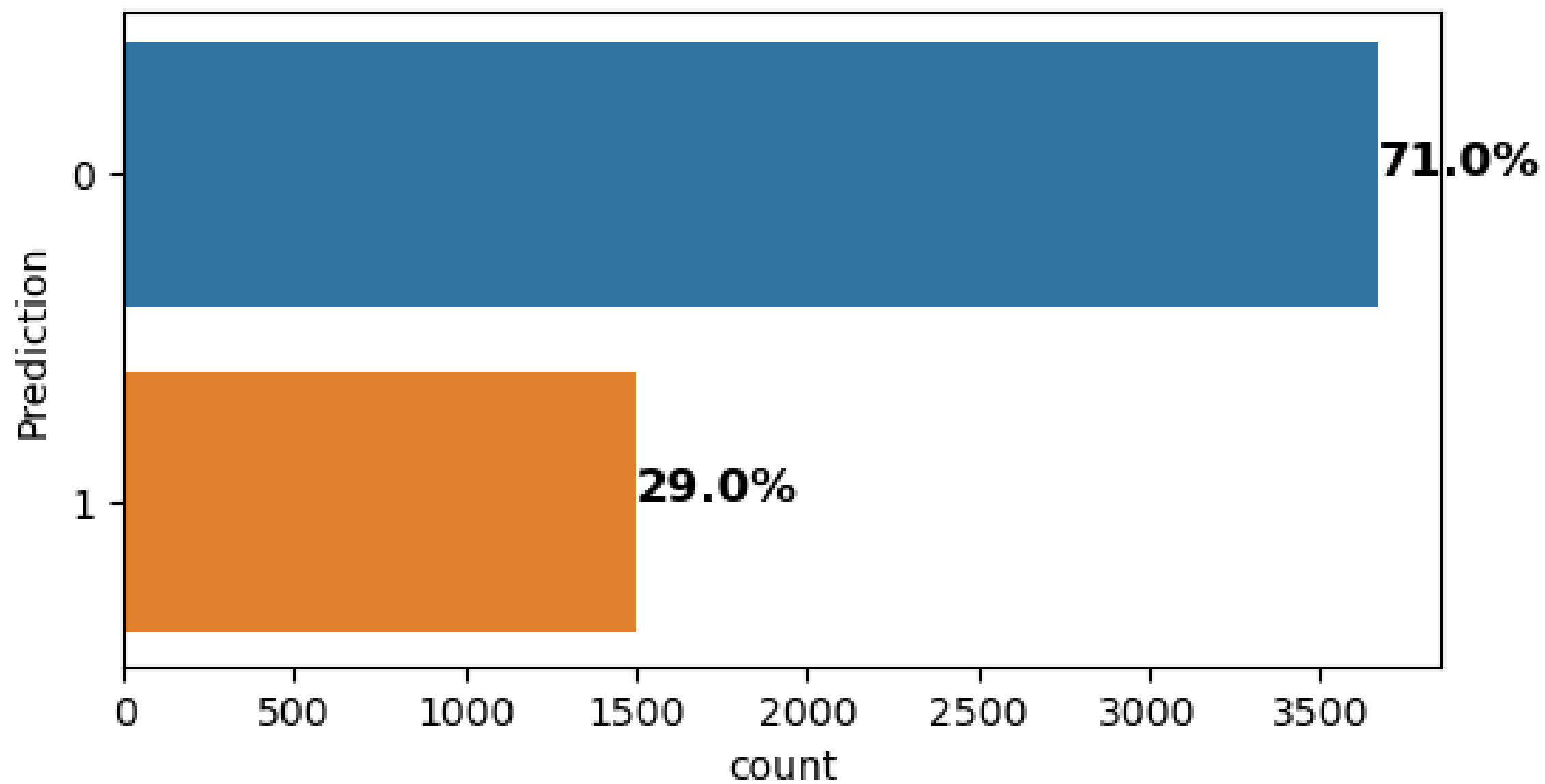
Introduction

Données



Nombre de spams

La distribution des spams





Sujet

Introduction

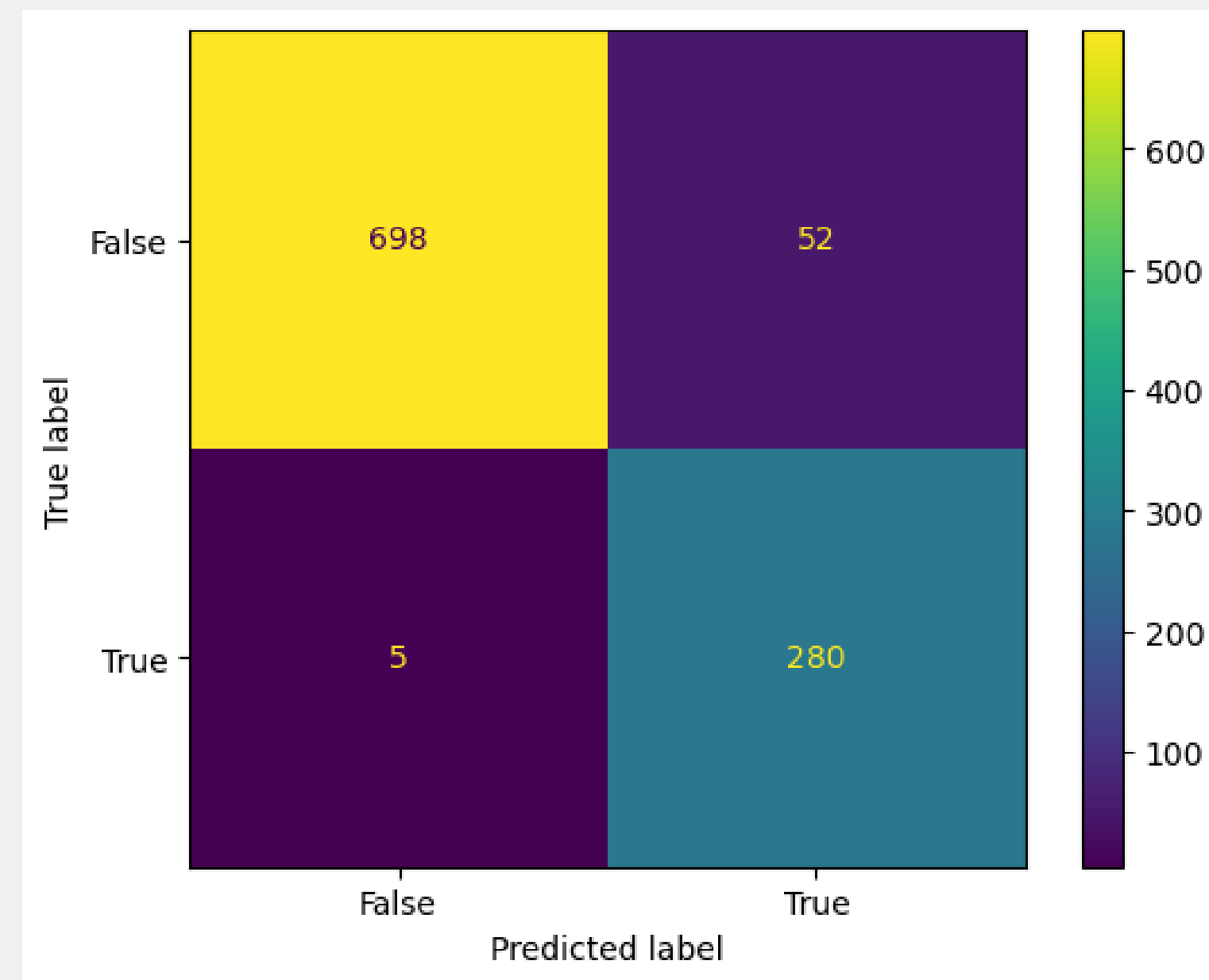
Données

ANN



← → ↺ 🔍 Réseau de neurones artificiels

Matrice de confusion



accuracy_score : 0.9449



Sujet

Introduction

Données

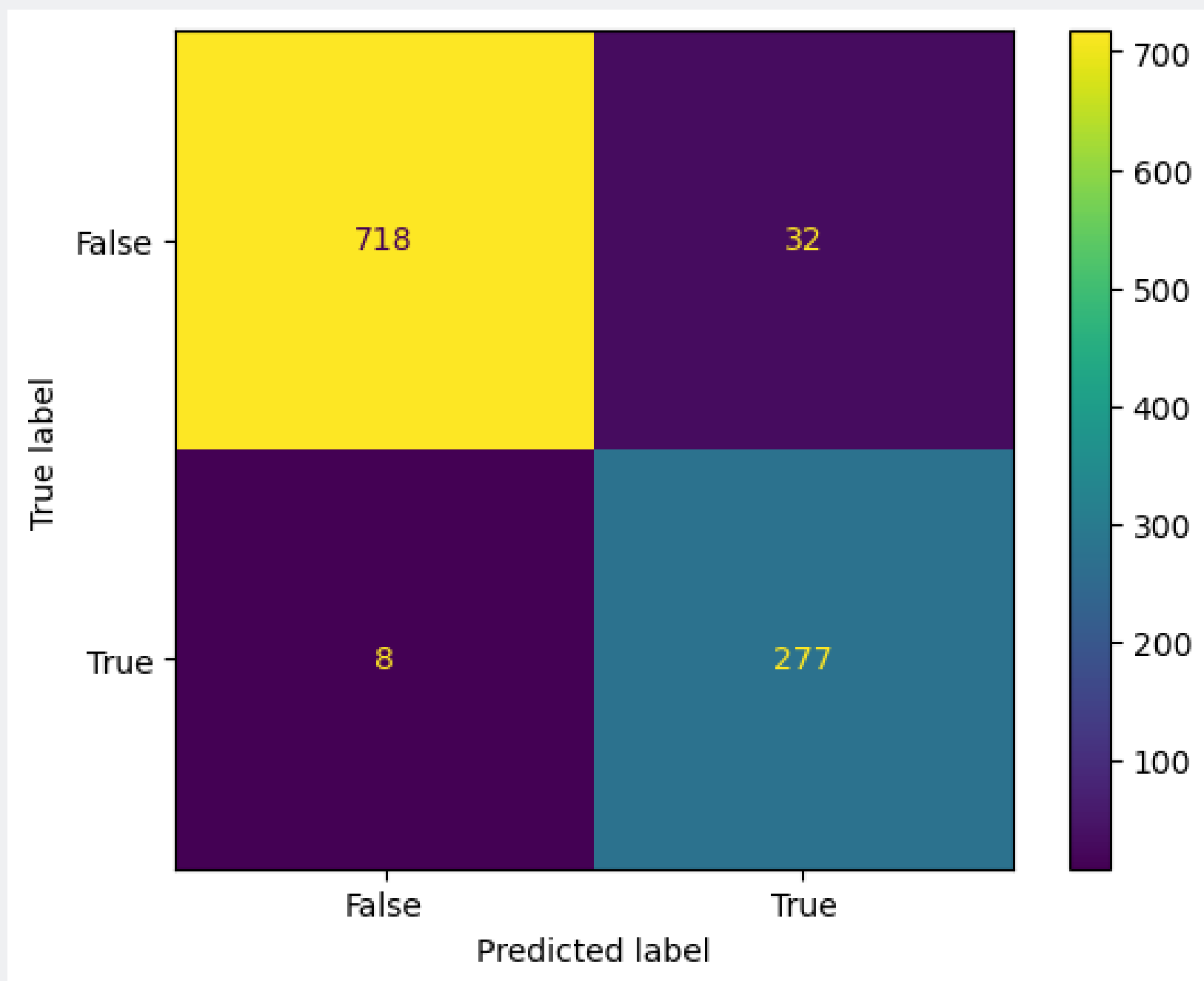
ANN

Améliorer ANN



← → 🔍 GridSearchCV

Matrice de confusion



optimizer : rmsprop

epochs : 150

batch_size : 30

accuracy_score : 0.9614



Sujet

Introduction

Données

ANN

Améliorer

Conclusion



← → 🔍 🔍 Comment conclure ?

Conclusion



**ANN est une méthode efficace
afin de trier les spams**

**4% d'erreur : principalement
des mails classés par erreur
dans les spams**



Sujet

Introduction

Données

ANN

Améliorer ANN

Conclusion

Merci



← → ↺ 🔍 remerciement

Merci pour votre attention

