INPT
المعهد الوطني للبريد والمواصلات
ⵓⵔⵣⵣⵓ ⵏⵓⵎⵙⵙⵓ ⵏ ⵜⵓⵏⵓⵖⵜ ⴰ ⵙⵎⵖⵓⵏⵓⵎ
Institut National des Postes et Télécommunications

orange™ Business Services

PHISHING DETECTION

USING AI AND NLP

Réalisé par :

ER-RAFI Chaimae ;
KARROUM Salim ;
SENHAJI Boutayna ;
SBIAA Ayoub ;
DOUBALI Salma ;
CHARIFI Zakaria.

Encadré par :

Mr EL KABIR Taoufik ;
Mr LAHSEN CHERIF Iyad.

Plan

1. **Contexte général ;**
2. **Approche de travail ;**
3. **Architecture de l’application ;**
4. **Feature Engineering et NLP ;**
5. **Récupération des e-mails via l’API Gmail ;**
6. **Modélisation et classification des e-mails de phishing à l’aide de modèles de ML ;**
7. **Visualisation des résultats ;**
8. **Structure de l’interface ;**
9. **Démonstaration ;**
10. **Points restants à traiter.**

Contexte général

- Le phishing est aujourd’hui la cybermenace la plus répandue et la plus coûteuse, tant par son volume d’attaques que par son efficacité.
- Phishing : menace globale et croissance
- Le phishing est devenu la forme de cybercriminalité la plus courante, avec 3,4 milliards d’e-mails malveillants envoyés chaque jour en 2021 ;
- En 2023, les attaques ont bondi de 58,2 % par rapport à l’année précédente et le nombre d’URL frauduleuses a atteint 1,76 milliard.

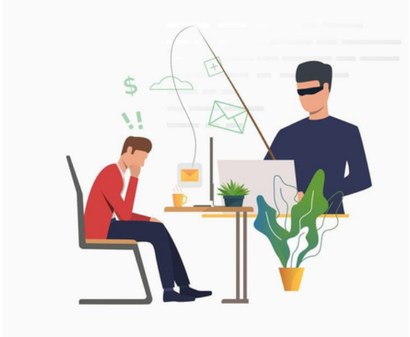
Contexte général

=> **Volume et évolution selon l'APWG :**

- En 2022, l'APWG (Anti-Phishing Working Group) a recensé plus de 4,7 millions d'attaques, en croissance de 150 % par an depuis 2019 ;
- Au 4e trimestre 2022, l'APWG a observé 1 350 037 attaques, en légère hausse par rapport aux 1 270 883 du trimestre précédent.

=> **Impacts financiers :**

- Le coût mondial du phishing est estimé à 3,5 milliards de dollars en 2024, incluant pertes directes et frais de remédiation.

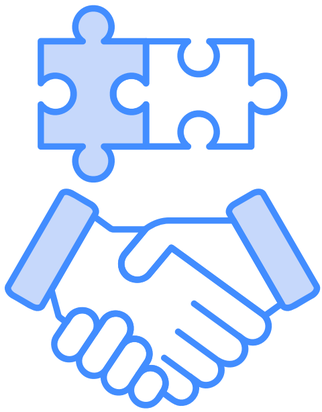


Objectifs

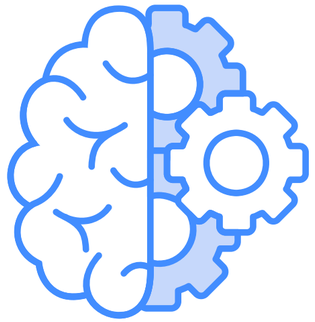
Face à l’ampleur et à la croissance rapide des attaques de phishing, ce projet vise à :

- *Développer une interface web capable de détecter automatiquement les e-mails de type phishing en exploitant le traitement automatique du langage naturel (NLP) ;*
- *Récupérer automatiquement les e-mails via l’API Gmail, afin d’analyser directement leur contenu sans que l’utilisateur n’ait besoin de copier ou soumettre manuellement un message ;*
- *Appliquer un modèle d’apprentissage automatique sur le contenu des e-mails pour prédire s’ils sont "safe" ou "phishing" ;*
- *Afficher le résultat directement sur l’interface web, en indiquant clairement à l’utilisateur si l’e-mail est considéré comme sûr ou frauduleux.*

Approche de travail

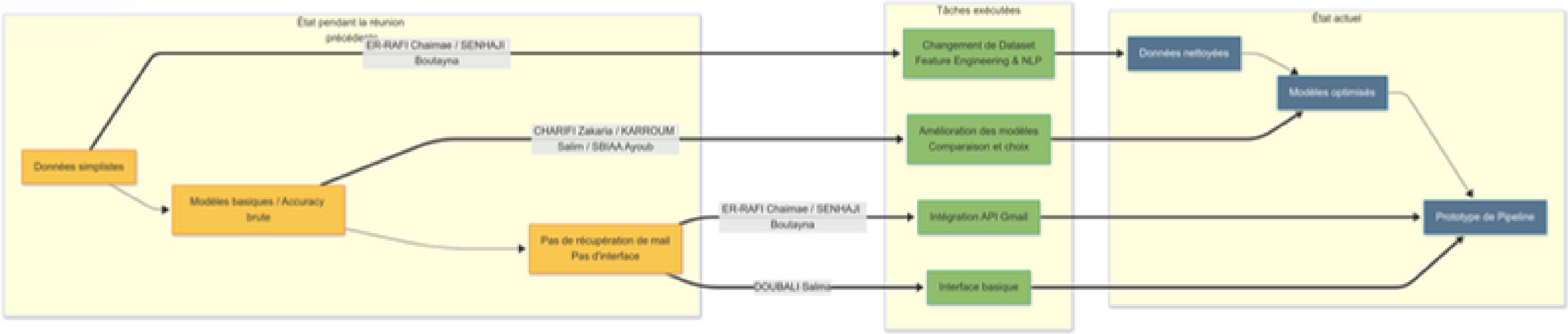


- **Définition des objectifs et documentation ;**
- **Élaboration du planning ;**

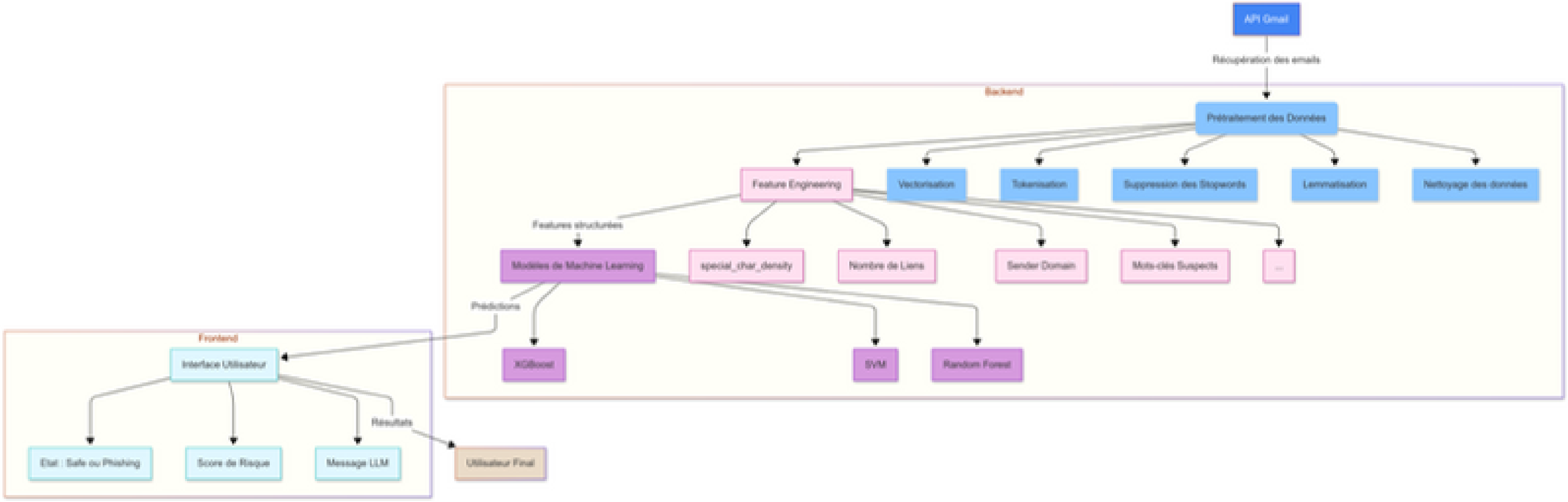


- **Répartition des taches ;**
- **Suivi de l’avancement et évaluation ;**
- **Organisation des réunions de suivi (en présentielle et en ligne).**

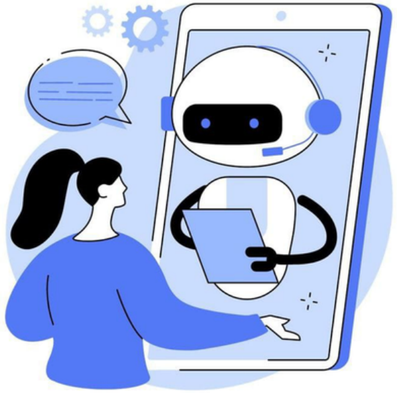
Approche de travail et avancement



Architecture de l'application



Feature Engineering & NLP

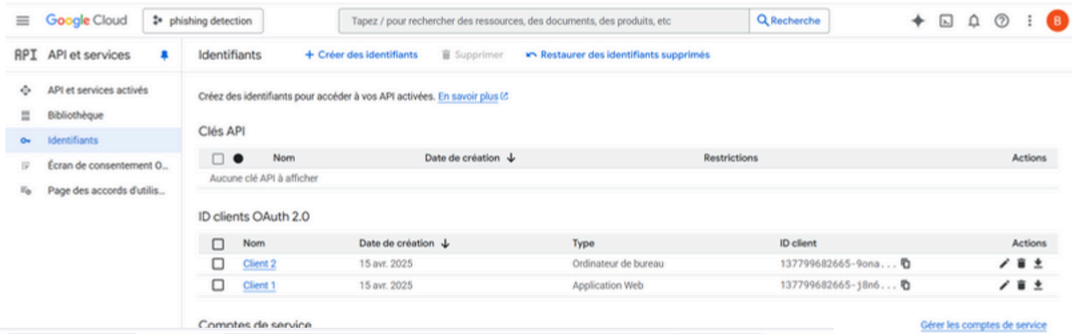


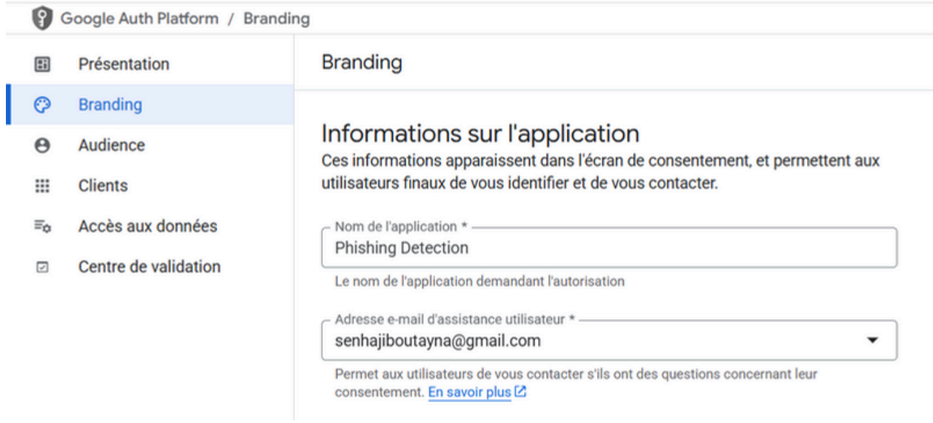
- Suppression : doublons, NaN ;
- Extraction des domaines d’expéditeurs et de destinataires
- Détection d’expéditeurs suspect (*suspicious_sender)* ;
- Extraction des URLs, comptage (*num_urls*) + détection de domaines suspects (*suspicious_urls*) ;
- Densité des caractères spéciaux dans body ;
- Analyse des mots-clés de phishing dans subject et body ;
- NLP → tokenisation, suppression des stopwords, lemmatisation, TF-IDF.

Récupération des Emails-API

Configuration Google Cloud Console :

- Création d’un projet Google Cloud dédié à l’accès Gmail.
- Configuration d’un écran d'autorisation OAuth (type externe, accès testeur).
- Création d’un identifiant (Desktop App).
- Téléchargement du fichier credentials.json → utilisé pour l’authentification locale via google_auth_oauthlib et google.oauth2.credentials





Récupération des Emails-API

Connexion et extraction automatisée des emails :

- Lancement du processus d'authentification : ouverture d'un lien dans le navigateur pour autoriser l'accès Gmail.
- Récupération automatique d'emails récents avec leur métadonnées (expéditeur, destinataire, sujet...).
- Détection et extraction automatique des URLs .
- Génération de sorties intermédiaires normalisées pour le prétraitement.
- Prévision du niveau de risque à partir du modèle (phishing/safe.) intégré dans la fonction.

📩 Email 2
From: Unity <info@unity3d.com>
Subject: Unite 2025 Call for Proposals is now open!
Body:
This is your chance to join us on stage at this year’s event.

...
URLs found: ['https://app.response.unity3d.com/e/er?utm_campaign=Other_global_Announcement_MESD-10505-2025-04-GLBL-Unite',
3d.com/e/es?s=795651218&e=26812880&elqTrackId=a7fc0838320d4ab99c43a716a87aa41c&target=_blank&elq=abf947fe0b0c43eb919ded9dc5
&elqak=8AF525EEE0356EE6C5323E1A2FF11A67B6863761D2D3C04D991D684A52E7C49A19F9>', 'https://app.response.unity3d.com/e/er?utm_ca
cement_MESD-10505-2025-04-GLBL-Unite', 'https://app.response.unity3d.com/e/er?utm_campaign=Other_global_Announcement_MESD-10
https://app.response.unity3d.com/e/er?utm_campaign=Other_global_Announcement_MESD-10505-2025-04-GLBL-Unite', 'https://app.re
m_campaign=Other_global_Announcement_MESD-10505-2025-04-GLBL-Unite', 'https://app.response.unity3d.com/e/er?utm_campaign=Oth
D-10505-2025-04-GLBL-Unite', 'https://app.response.unity3d.com/e/er?utm_campaign=Other_global_Announcement_MESD-10505-2025-0
eate.unity3d.com/Unity_MyPage_W_AccountID?elqTrackId=3b50880375c94617b58fc46ef01c3b63&elq=abf947fe0b0c43eb919ded9dc55aac87&e
paignId=17695', 'https://create.unity3d.com/unsubscribe-email-group?elqTrackId=dd1e9540012c409ba682e0ff99d467f5&elqc=11&elq=
5aac87&elqaid=37223&elqat=1&elqCampaignId=17695', 'http://unity3d.com/legal/privacy-policy?elqTrackId=f18f96fcb4bb47ef97f3a5
c43eb919ded9dc55aac87&elqaid=37223&elqat=1&elqCampaignId=17695', 'http://unity3d.com/legal/terms-of-service?elqTrackId=e74f
4&elq=abf947fe0b0c43eb919ded9dc55aac87&elqaid=37223&elqat=1&elqCampaignId=17695']
Prediction: Safe

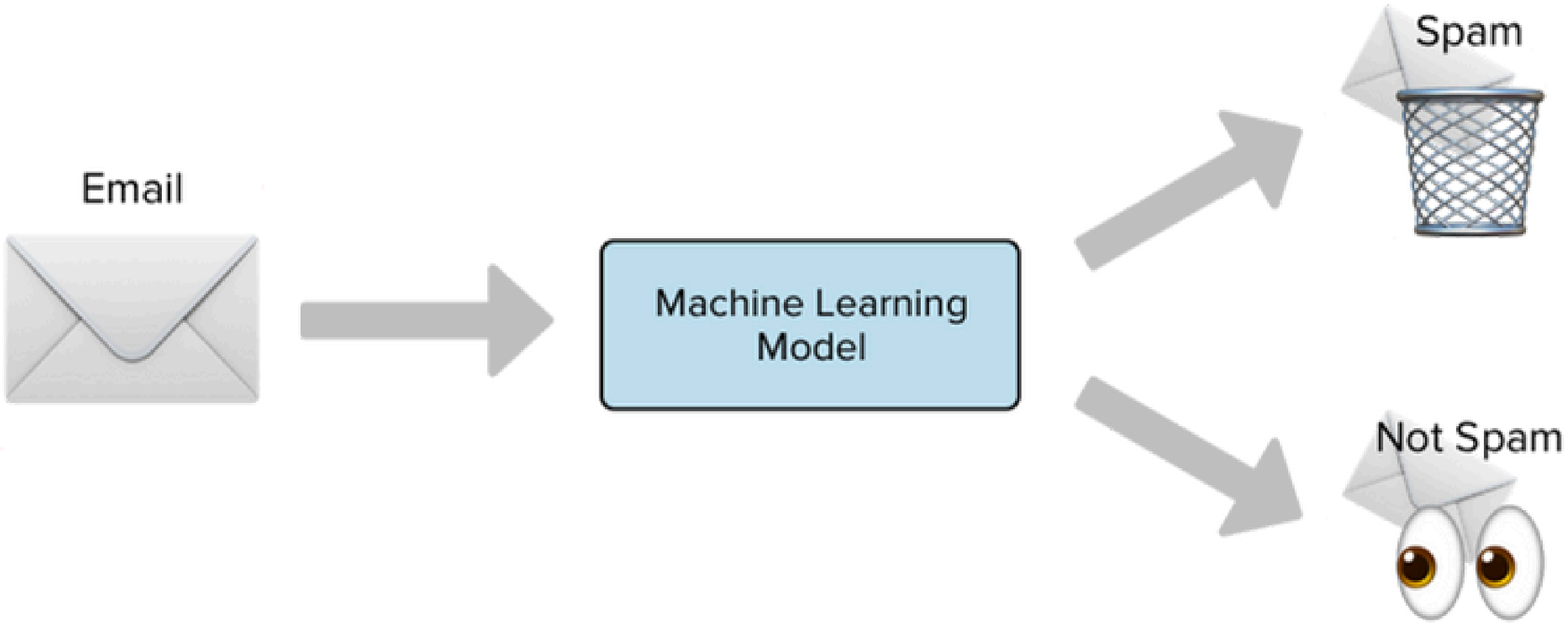
📩 Email 4
From: J Ia <iaj82706@gmail.com>
Subject: We Work You WIN
Body:
Relax and have fun with poker, blackjack, roulette, progressive video slots
at your own leisure from your couch. Our safe, secure games will get you
smiling when you start seeing dollars pouring in. We're serious about fun.
When YOU WIN, we win!

We Work You WIN

We Pay! You Play!

htt...
URLs found: ['http://casinojackpotsdeals.com']
Prediction: Phishing

Modélisation et Classification des E-mails Phishing a l’aide des Modèles de ML



XGBoost

Machine Learning

Les e-mails de phishing constituent une menace de plus en plus préoccupante pour la cybersécurité.

L'objectif de cette partie est de développer un système de classification permettant de prédire si un e-mail est frauduleux ou légitime, en s'appuyant sur des caractéristiques extraites de son contenu.

Pour cela, plusieurs algorithmes de machine learning ont été mis en œuvre, notamment **XGBoost**, **SVM**, **Random Forest**, ainsi qu'un modèle de **Stacking** combinant leurs performances.

Bibliothèques utilisées

manipulation de données

visualisation

Machine Learning

Modéles ML

XGBoost

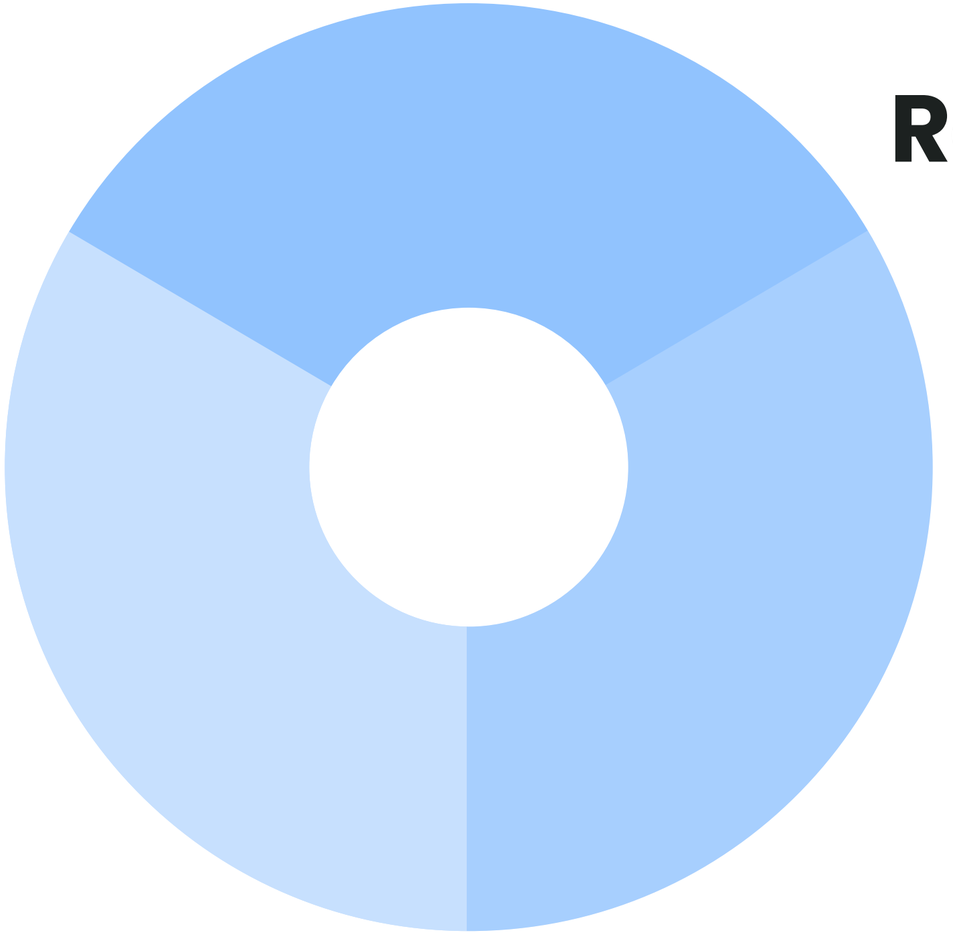
Accuracy : 0.995731

ROC AUC : 0.999794

SVM

Accuracy : 0.993791

ROC AUC : 0.998523



Random Forest

Accuracy : 0.974130

ROC AUC : 0.997868

Modéle de Stacking

Accuracy : 0.9960

ROC AUC : 0.9996

Visualisation des résultats

✓ Fast

✓ Simple

✓ well-suited for ML projects

Structure de l’interface

1

Chargement des modèles ML

2

Authentification Gmail via API Google

3

Récupération des emails

4

Prétraitement du contenu

5

XGBoost

Machine Learning

Prédiction par le modèle XGBoost

Bibliothèques utilisées

```
import streamlit as st
```

Pour l’interface

```
import base64
import os
import re
import pandas as pd
import joblib
from bs4 import BeautifulSoup
```

pour le traitement

```
from googleapiclient.discovery import build
from google.auth.transport.requests import Request
from google.oauth2.credentials import Credentials
from google_auth_oauthlib.flow import InstalledAppFlow
```

pour l'accès aux mails

Démo de l’interface

app

localhost:8501

Deploy

Détecteur de Phishing Gmail avec IA

1

Scanner mes 10 derniers emails

app

Connexion : comptes Google

accounts.google.com/signin/oauth/v2/consentsummary?authuser=0&part=AJi8hAMtuUHrZSISesVyXXJNCZbzogmSqrOy3E5ehLQgbkADY0rxdj7l21tN8H1uwwDC_hNuvuLChQQlLpp_pdVA...

Se connecter avec Google

Phishing Detection veut accéder à votre compte Google

salmasalmadoubli@gmail.com

Lorsque vous accorderez cet accès, Phishing Detection pourra

Afficher vos e-mails et vos paramètres de messagerie. En savoir plus

Vérifiez que l'application Phishing Detection est digne de confiance

Découvrez pourquoi vous ne voyez pas de liens vers les règles de confidentialité ou les conditions d'utilisation de Phishing Detection.

Consultez les règles de confidentialité et les conditions d'utilisation de Phishing Detection pour comprendre comment Phishing Detection traitera et protégera vos données.

Pour apporter des modifications à tout moment, accédez à votre compte Google.

Découvrez comment Google vous aide à partager vos données de manière sécurisée.

Annuler

Continuer

2

Taper ici pour rechercher

19°C

ENG

12:58

22/04/2025

app

localhost:8501

Deploy

Détecteur de Phishing Gmail avec IA

Scanner mes 10 derniers emails

10 emails récupérés.

Email 1

De : ne-pas-repondre.etudesenfrance@diffusion.diplomatie.gouv.fr

Sujet : MA23-04605-P02 / Etudes en France : Paiement enregistré

Prévision : *Safe*

Contenu de l'email

Email 2

De : ne-pas-repondre.etudesenfrance@diffusion.diplomatie.gouv.fr

Sujet : Etudes en France : paiement des frais de dossier / votre identifiant à conserver et à présenter au Crédit du Maroc : MA23-04605-P02

Prévision : *Safe*

Taper ici pour rechercher

22°C

ENG

10:07

23/04/2025

Les points restants à traiter.

- **Amélioration et finalisation de l'interface utilisateur** : c’est l’un des aspects prioritaires afin de garantir une meilleure expérience pour l'utilisateur final.
- **Intégration d'une réponse personnalisée générée par un LLM**, conformément à la proposition de M. Taoufik, et dans le but d’apporter une valeur ajoutée intelligente au système.
- **Les autres points à aborder, y compris les éventuelles modifications ou ajouts, seront discutés avec M. Taoufik** pour aligner les prochaines étapes sur ses orientations et nos propositions.

Merci pour votre attention