

ANALYSE DE SENTIMENT D'INWI SUR TWITTER

TEST TECHNIQUE PRÉSENTÉ POUR INWI

25/11/2024

Réalisé par:

OUMAIMA DGUIGOU

Plan de travail

- 01** INTRODUCTION
- 02** PROBLÉMATIQUE
- 03** OBJECTIF DU PROJET
- 04** MÉTHODOLOGIE ET OUTILS
- 05** RÉALISATION
- 06** CONCLUSION & PERSPECTIVES

Quels sentiments cachent les commentaires sur Inwi et le secteur télécom marocain, et comment ces perceptions peuvent-elles impacter leur réputation ?



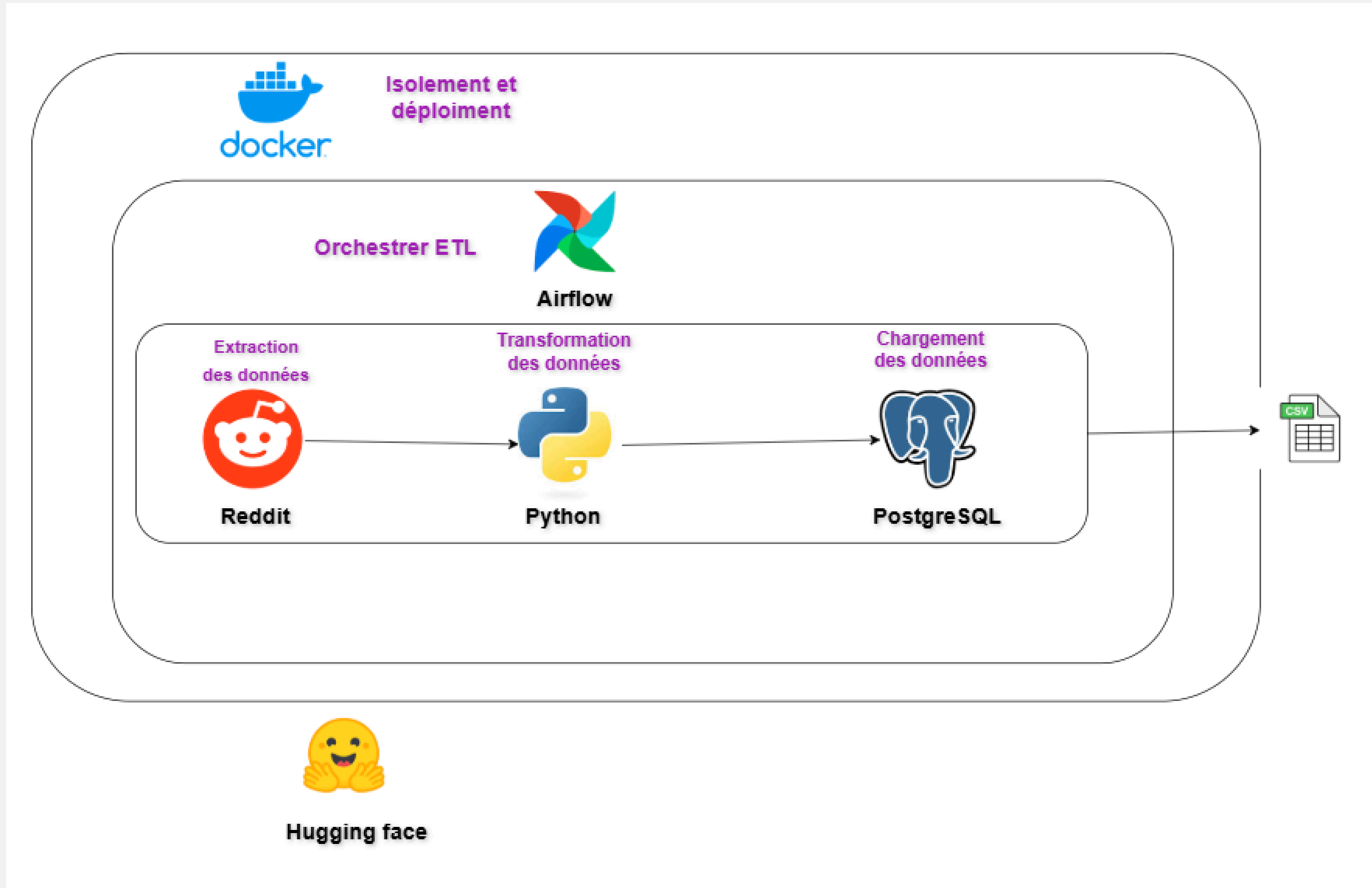
Analyser les discussions publiques sur Twitter concernant Inwi et le secteur télécom marocain afin de déterminer le sentiment global et d'identifier des thèmes récurrents dans les tweets négatifs.

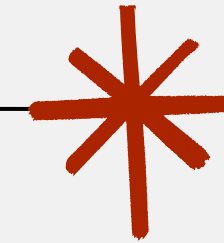
Collecte des données

- 50 tweets sur Inwi
- 50 sur le secteur télécom marocain

Modélisation des données

Exploration des données





EXTRACTION DES DONNÉES

- Connexion à l'API Reddit
- Recherche des publications pertinentes
- Récupération des données des publications.
- Extraction des commentaires

TRANSFORMATION DES DONNÉES

- Supprimer les emojis
- Supprimer les nombres flottants
- Supprimer les doublons dans le texte
- Convertir en minuscules
- Supprimer la ponctuation
- Supprimer les stopwords

MODÉLISATION

- Les modèles appropriés.

Collecte des données

01

Choix de l'API

- Accès limité à Twitter
- Problème avec YouTube
- Choix de Reddit

.

Stockage des données

02

Pourquoi Postgresql

- **Initialement** : Pas pris en compte l'exigence clé-valeur.
- **Setup** : Docker avec Airflow et PostgreSQL, deux tables créées.
- **Solution** : PostgreSQL supporte JSONB, idéal pour les données clé-valeur.

Les modèles de machine learning utilisés

Twitter-roBERTa-base for
Sentiment Analysis

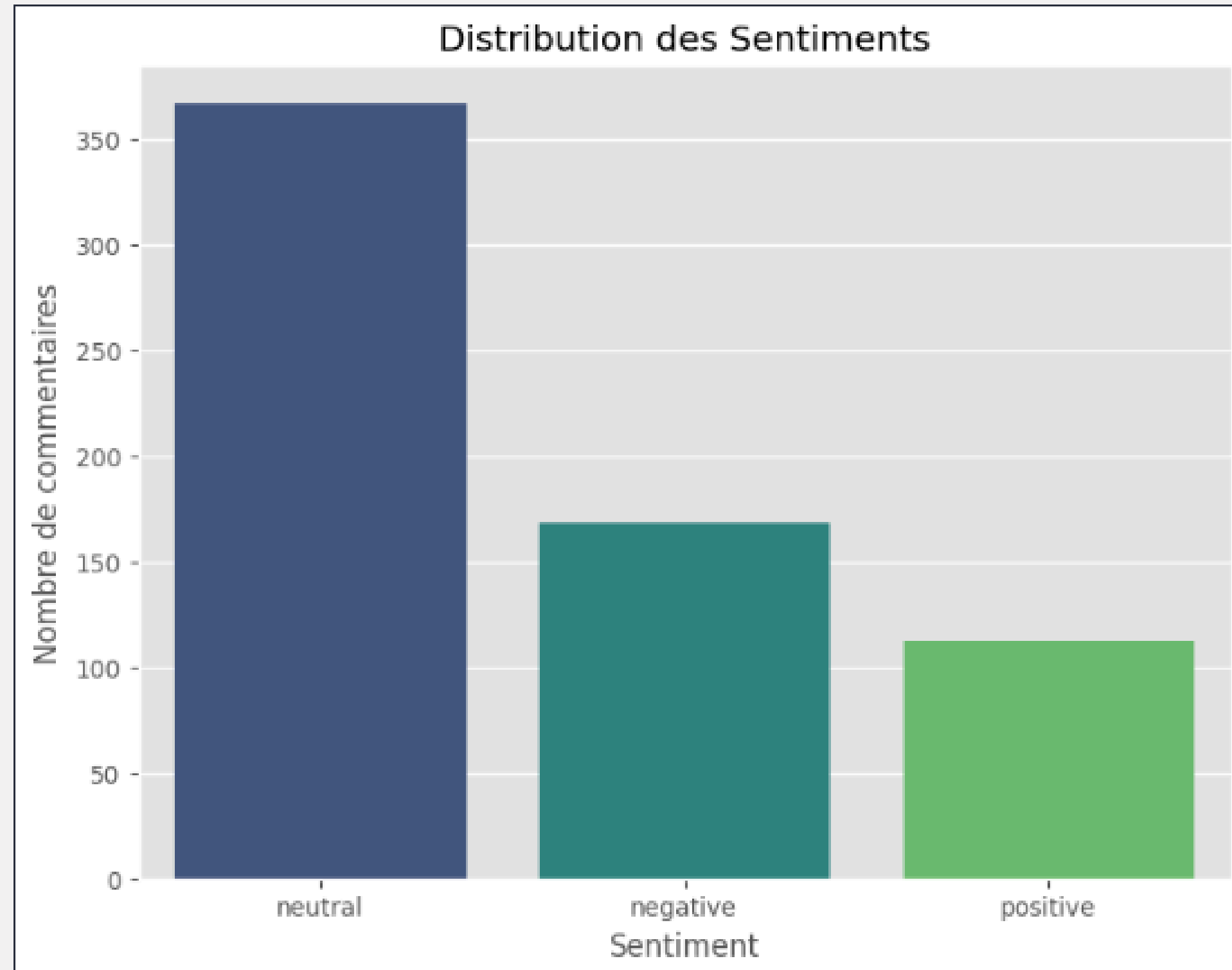
VS

VADER (Valence Aware Dictionary
and sEntiment Reasoner)

- **un modèle RoBERTa-base entraîné sur environ 124 millions de tweets**
- **collectés entre janvier 2018 et décembre 2021**

- **VADER utilise un lexique de mots courants associés à des scores de sentiment prédéfinis. Chaque mot a une valeur de sentiment :**
 - **Positif (+ve)**
 - **Négatif (-ve)**
 - **Neutre (0)**

DISTRIBUTION DES SENTIMENTS SELON “ROBERTA”



- Majorité de commentaires neutres : Indique un manque d'émotion forte ou un faible engagement.
- Environ 150 des commentaires neutres sont négatifs : Suggère des préoccupations sous-jacentes non exprimées clairement.
- Peu de commentaires positifs : Montre un sentiment général de mécontentement ou d'indifférence.

les sujets des commentaires négatifs

Expérience client
décevante

Difficultés avec la
connexion
mobile/fibre et le
service client

Prix élevés et
faible qualité de
service

Services et lieux
de faible qualité

Mécontentement
général avec
l'internet et les
tarifs

Conclusion

- Extraction réussie des posts et commentaires sur Inwi et le secteur télécom au Maroc.
- Stockage des données et analyse de sentiment (0 à 1).
- Identification des sujets négatifs (promotions, pannes).

Perspectives

- **Améliorer l'analyse de sentiment avec du machine learning.**
- **Étendre la collecte en temps réel à d'autres acteurs.**
- **Créer des dashboards interactifs.**
- **Utiliser les analyses pour la gestion de réputation et la veille concurrentielle.**

**MERCI POUR
VOTRE ATTENTION**