

YNOV Tweespy

API TWITTER

ABID Sabri
RAMDANI Rayan
HAROUNE Rayane

Master 1 Data engineer & Data science
Ynov Paris 2021/2022

Sommaire

- ▶ Introduction API twitter
- ▶ Objectif du projet
- ▶ Notre architecture
- ▶ La démonstration
- ▶ Difficultés rencontrées
- ▶ Pistes explorées
- ▶ Conclusion

Introduction Twitter API

- ❖ L'API Twitter permet un accès programmatique à Twitter de manière unique et avancée. Exploitez les éléments essentiels de Twitter tels que : les tweets, les messages directs, les espaces, les listes, les utilisateurs, etc.

Niveaux d'accès V2

Essentiel

Avec l'accès Essentiel, vous pouvez désormais accéder à l'API Twitter v2 rapidement et gratuitement !

- Récupérez 500 000 Tweets par mois
- 1 projet par compte
- 1 environnement d'application par projet
- Pas d'accès à la v1.1 standard, à la v1.1 premium ou à l'entreprise

Élevé

Avec l'accès élevé, vous pouvez obtenir un accès supplémentaire gratuit aux points de terminaison et aux données, ainsi qu'à des environnements d'application supplémentaires.

- Récupérez 2 millions de Tweets par mois
- 1 projet par compte
- 3 environnements d'application par projet
- Accès à la v1.1 standard, à la v1.1 premium et à l'entreprise

Recherche académique

Si vous êtes admissible à notre niveau d'accès à la recherche universitaire, vous pouvez accéder à encore plus de données et à des points de terminaison de recherche avancée.

- Récupérez 10 millions de Tweets par mois
- Accès à la recherche d'archives complètes et au nombre de Tweets d'archives complètes
- Accès aux opérateurs de recherche avancée

Objectif du projet

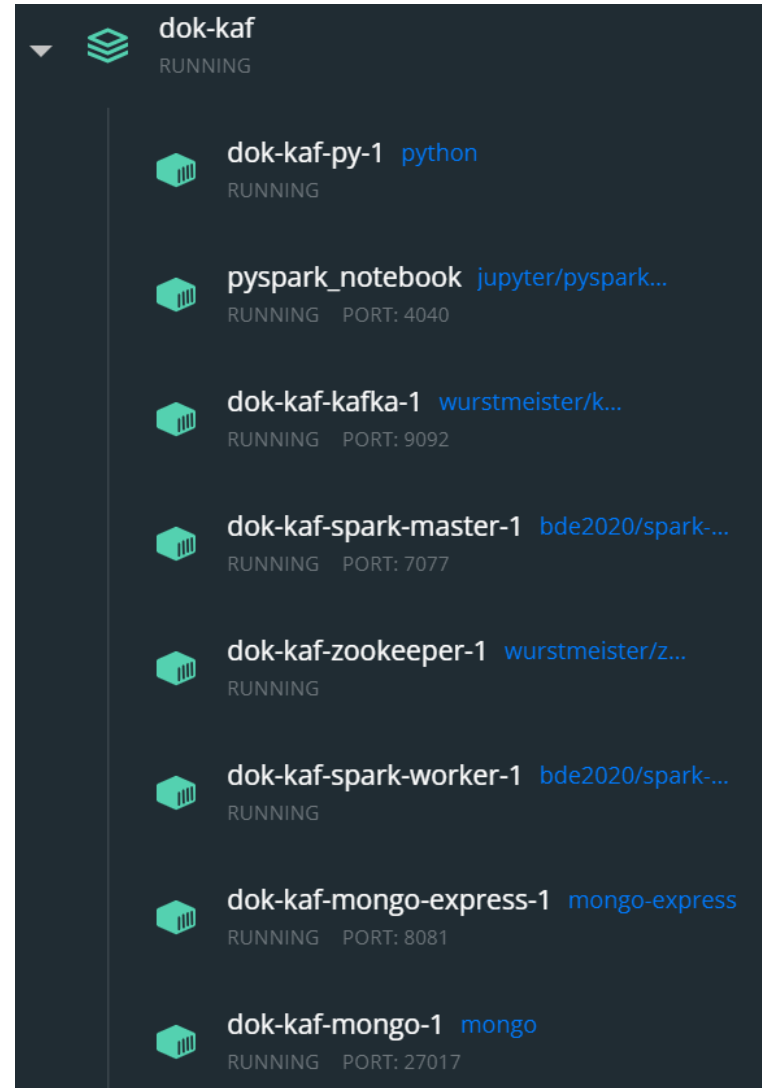
Ce projet a pour objectif de récupérer en temps réel les tweet (d'un compte en particulier) afin d'appliquer notre modèle sur ces derniers pour faire une classification tweet positif/négatif dessus ce qui nous permettra de connaître le sentiment sur certains sujets, situations...

Technologies utilisées

- ❖ Pour ce projet nous avons opté pour les technologies suivantes :
 - 1) Docker.
 - 2) Kafka.
 - 3) Spark.
 - 4) Streamlit.
 - 5) Mongo DB.

Technologies utilisées

❖ DOCKER :



E-reputation 🔥

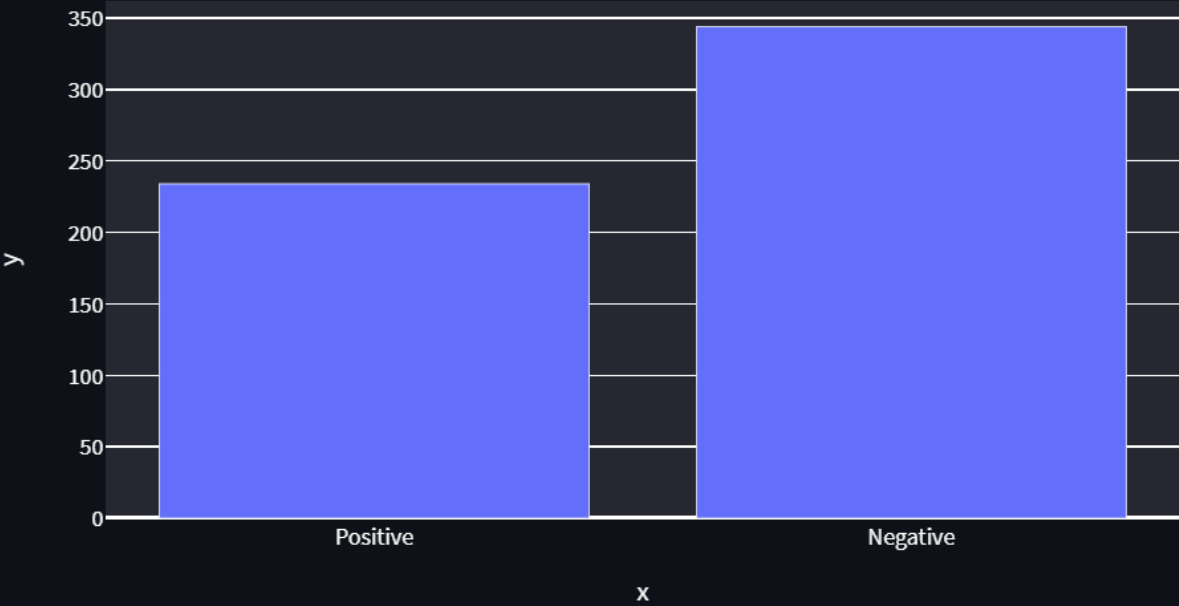
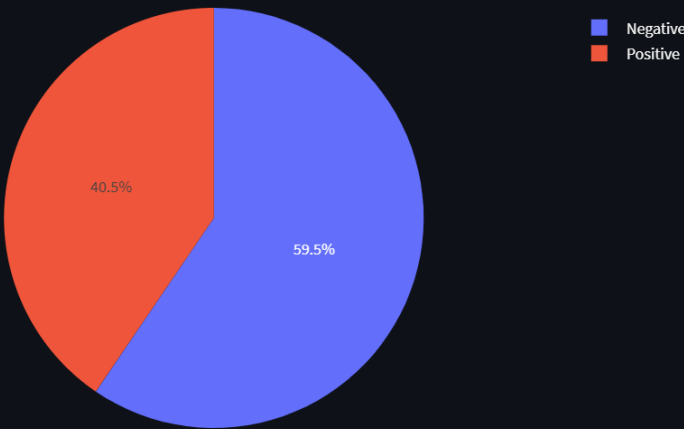
Enter Your name

CVEannounce

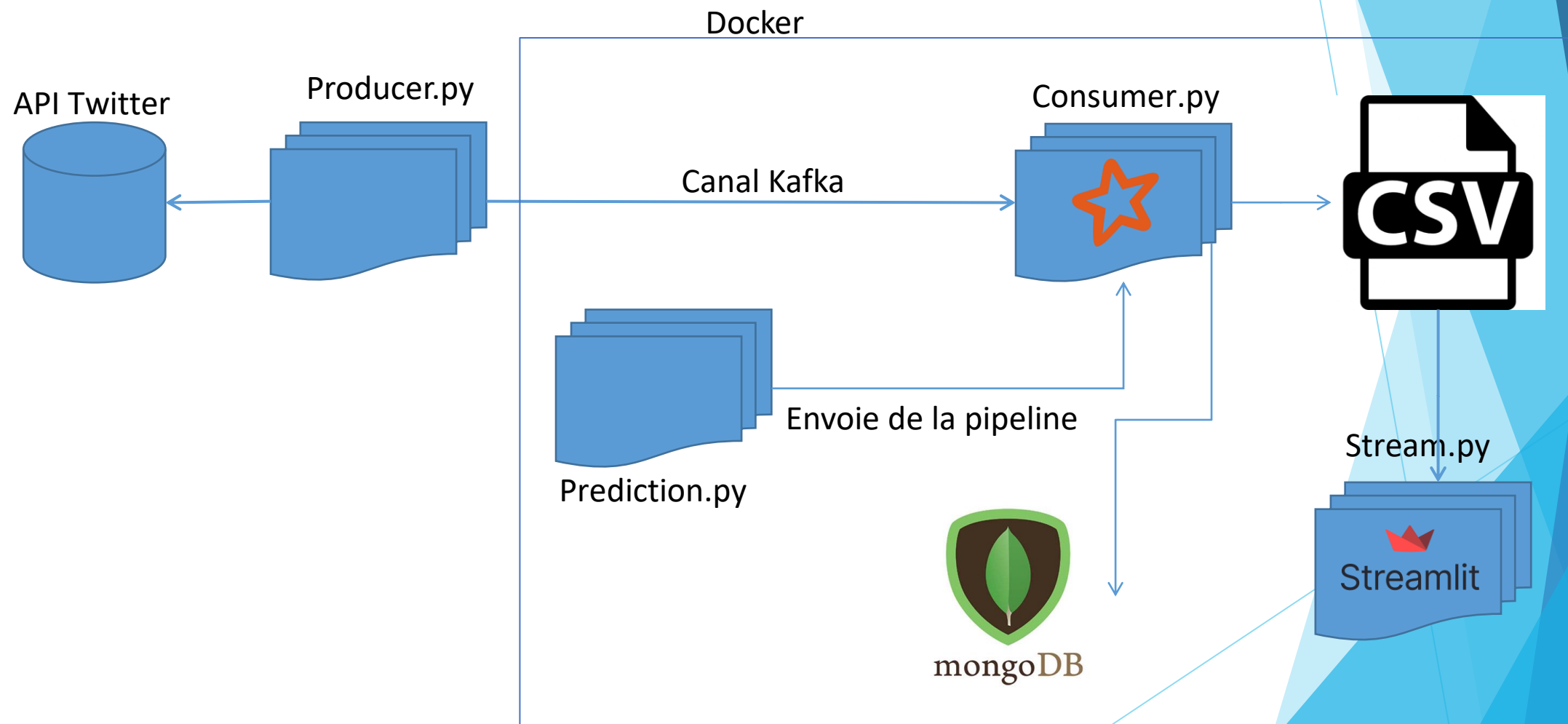
Submit

	Tweet content	date	prediction	Sentiment
400	As Ohio senate hopefuls vie fo...	07/04/2022	0.0000	Negative
401	BREAKING: Tiger Woods finish...	07/04/2022	1.0000	Positive
402	SURPRISE!: Republican gover...	07/04/2022	1.0000	Positive
403	US Navy intends to decommis...	07/04/2022	0.0000	Negative
404	GOP senator blasts Biden for ...	07/04/2022	0.0000	Negative
405	Leo Terrell rips California's 'lu...	07/04/2022	0.0000	Negative
406	'VERY DANGEROUS': @SenJoh...	07/04/2022	1.0000	Positive
407	iPhones, Androids, tablets an...	07/04/2022	0.0000	Negative
408	OPINION: @greggutfeld calls ...	07/04/2022	1.0000	Positive
409	Biden administration slamme...	07/04/2022	1.0000	Positive

Pourcentage de tweet positifs et negatifs



Notre architecture :



Démonstration

Difficultés rencontrés

- ❖ Récupération de l'API Twitter
- ❖ intégration de streamlit dans le consumer
- ❖ Architecture distribuée
- ❖ Connexion entre le consumer et mongo DB

Les pistes potentiels

- ▶ Nous pouvons apportés d'autres calculs ou d'autres statistique
- ▶ Faires des corrélations
- ▶ Faire une détection de langue sur les tweet
- ▶ Améliorer notre visualisation
- ▶ Faire une connexion entre streamlit et mongo DB

Conclusion

- ▶ Ce projet nous a permis de découvrir beaucoup de nouveaux outils pour nous aider à valoriser des données de manière rapide et continue, nous avons aussi appris à visualiser ses données et à effectuer des calculs dessus.

Des questions ?