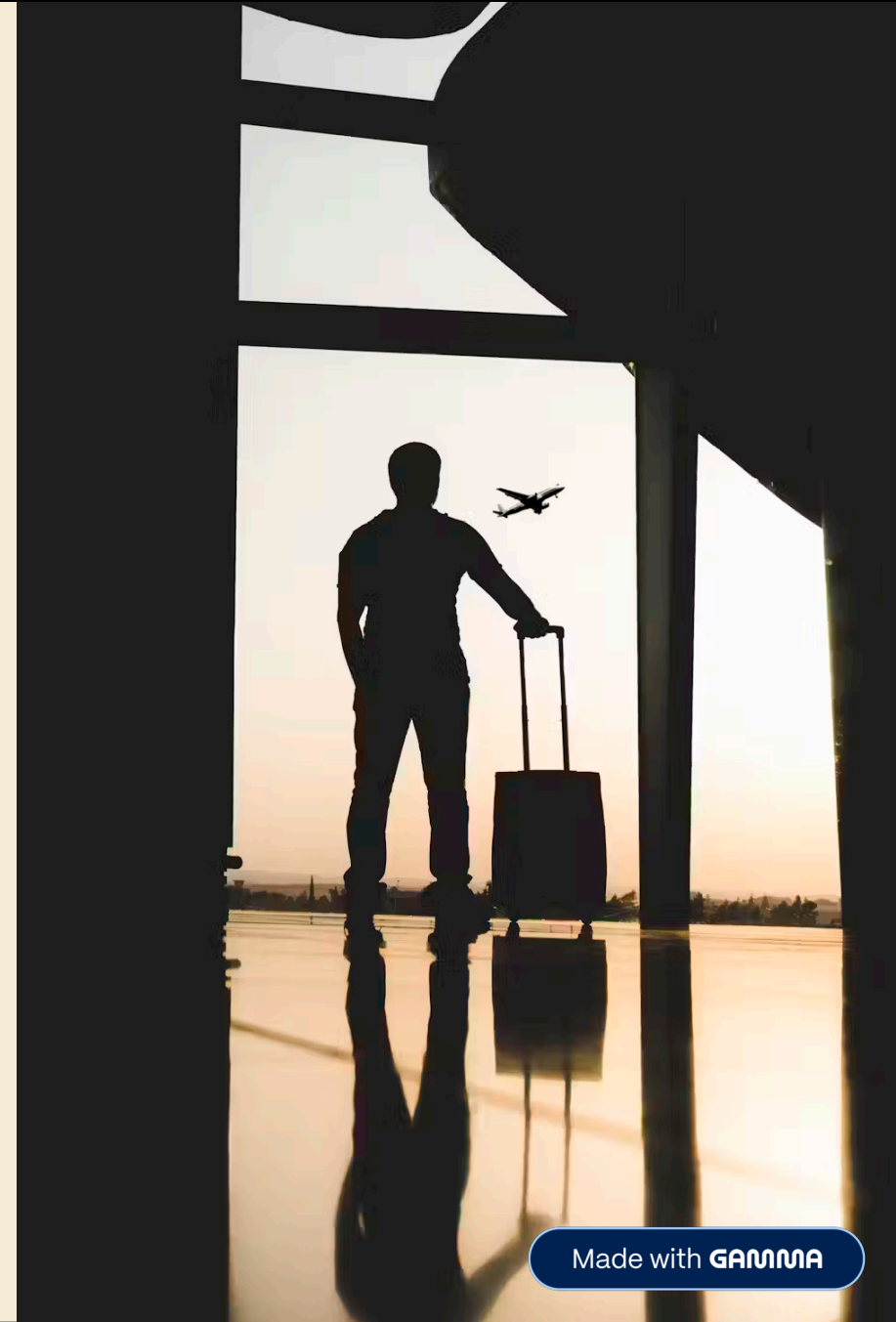


Réalisez une analyse de sentiments grâce au Deep Learning

 par Ethan VUILLEMIN



Made with GAMMA



Contexte & Objectifs

Contexte

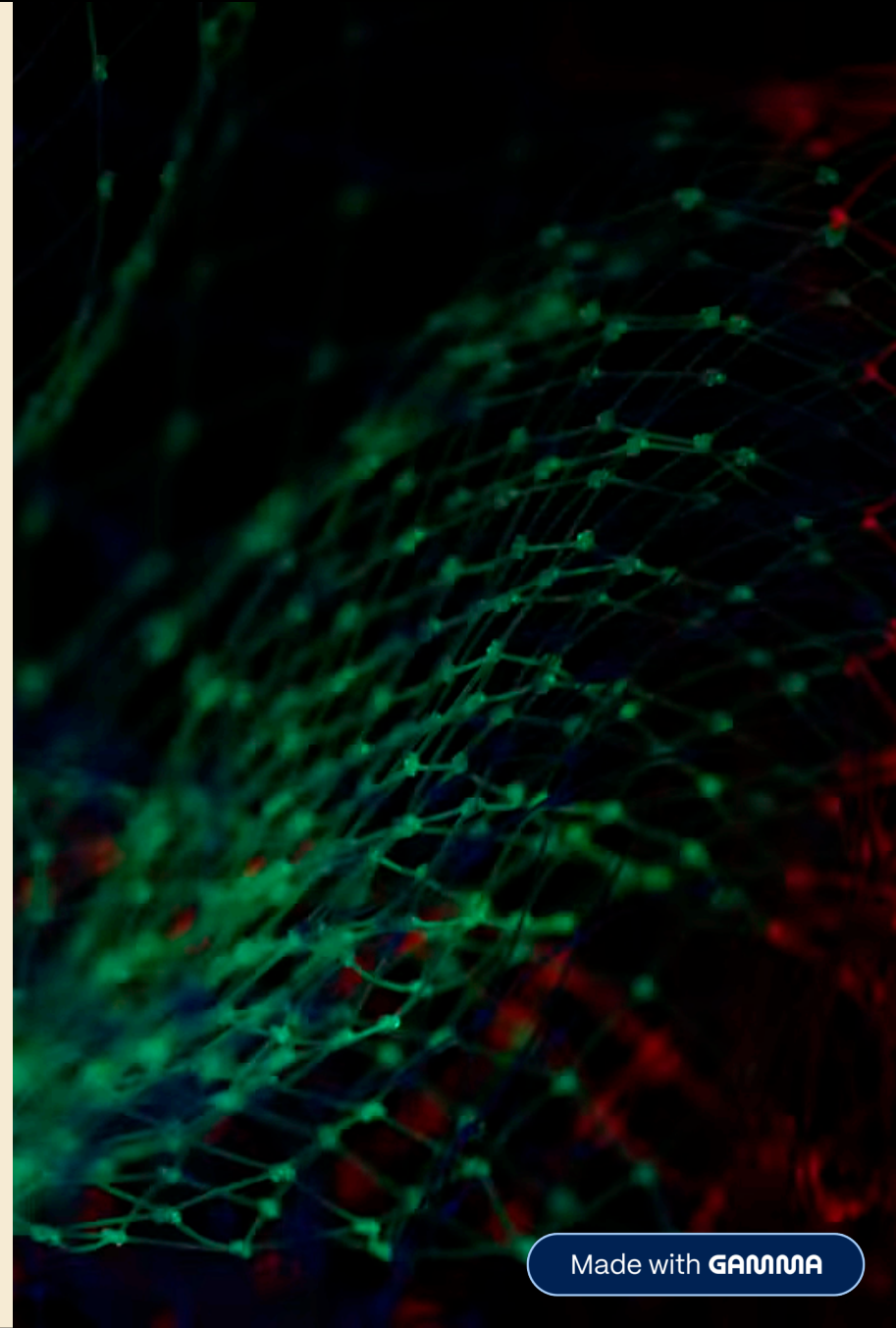
Consultant pour la **société MIC**

Air Paradis, notre client, souhaite **anticiper les avis négatifs** sur les réseaux sociaux grâce à un modèle d'analyse de sentiments afin de **predire les badbuzz**

Objectif

Crée le **prototype** d'une application d'**analyse de sentiment** d'un ou plusieurs tweet grace a un **modèle de classification de sentiment** tout en respectant une **demarche MLOps**

Lecture & Preprocessing des données



Règlement général sur la protection des données (RGPD)

Licéité

Fondé sur une base légale

Finalité

Collectées pour un but précis et légitime

Minimisation

Seules les données strictement nécessaires

Exactitude

Les données doivent être exactes et à jour

Limitation de la conservation

Ne doivent pas être conservées plus longtemps que nécessaire

Intégrité et confidentialité

Protégées contre les accès non autorisés ou les fuites

Transparence

Les personnes concernées doivent être informées

Droits des personnes concernées

Les individus ont des droits sur leurs données (accès, rectification, effacement, opposition)

Creation des modèles

Sur-mesure simple

Logistic Regression

1

2

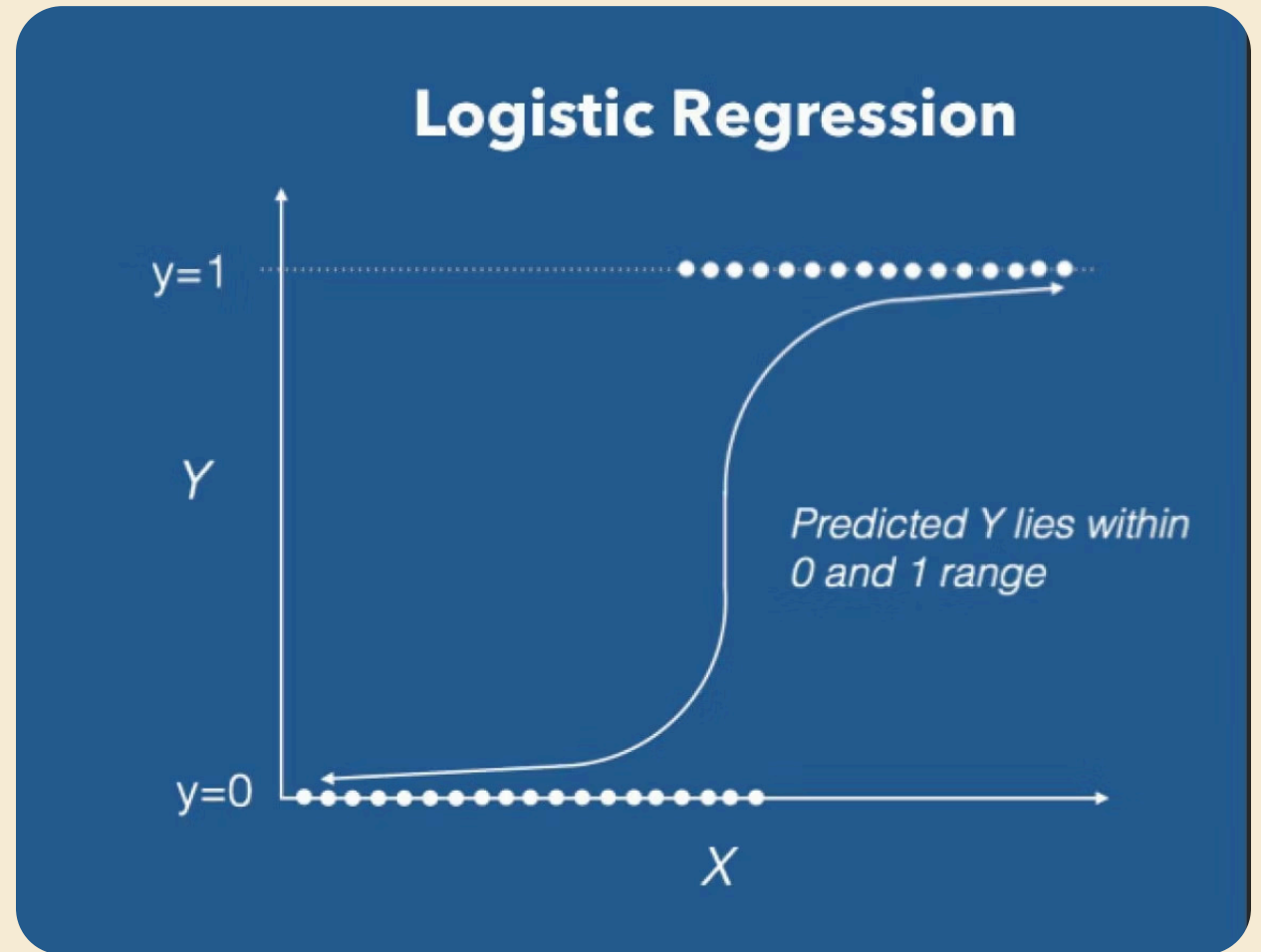
3

BERT

Sur-mesure avancé

RNN, LSTM

Modèle Sur Mesure Simple

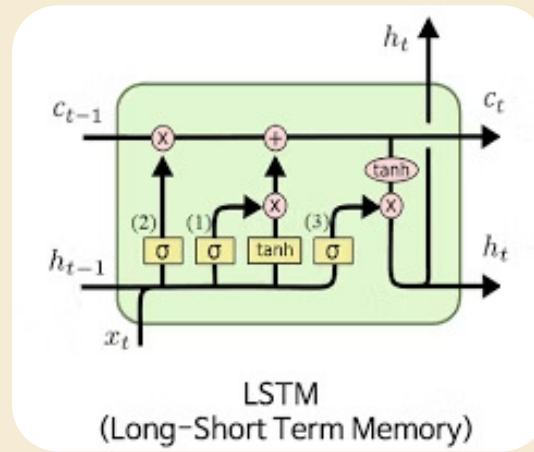
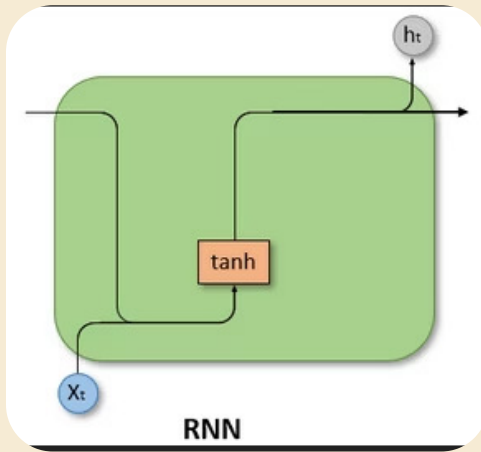


Comparaison des approches

❗ Erreur lors de l'envoi de l'image.



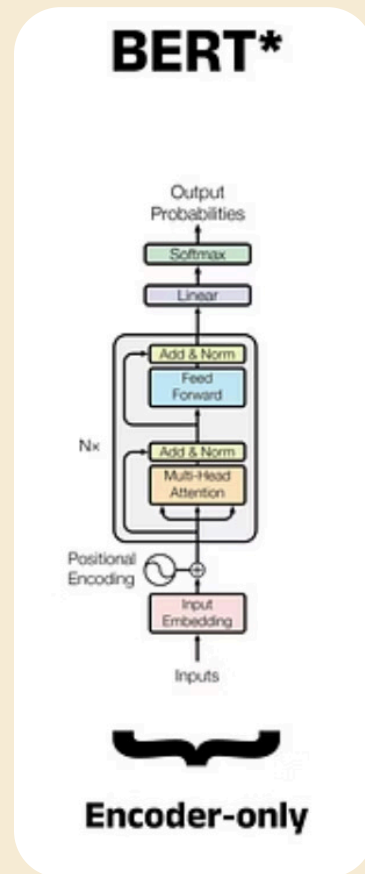
Modèle Sur Mesure Avancé



Comparaison des approches



BERT



Comparaison des approches





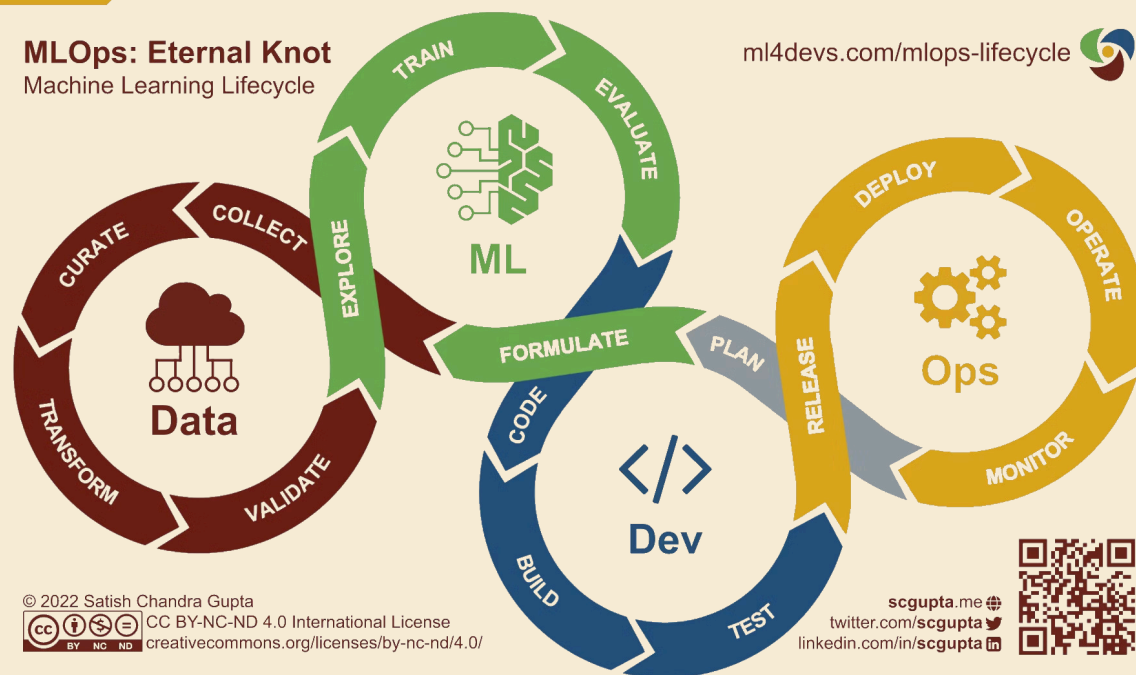
Comparaison des modèles



Introduction au MLOps

MLOps: Eternal Knot
Machine Learning Lifecycle

ml4devs.com/mlops-lifecycle



© 2022 Satish Chandra Gupta
CC BY-NC-ND 4.0 International License
creativecommons.org/licenses/by-nc-nd/4.0/

scgupta.me
twitter.com/scgupta
linkedin.com/in/scgupta



Article de Blog



Gestion des Expérimentations avec MLFlow

mlflow

Tracking

Record and query
experiments: code,
data, config, results

Projects

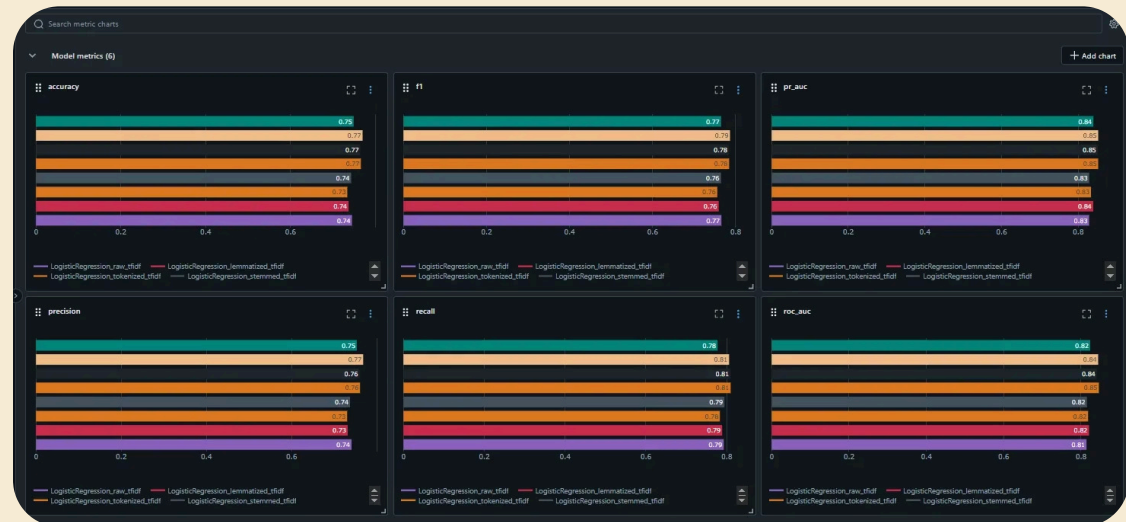
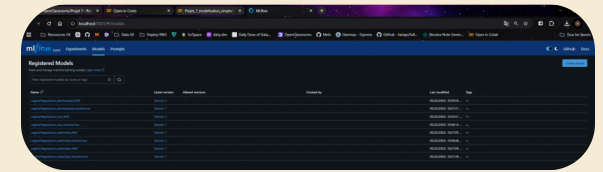
Packaging format for
reproducible runs
on any platform

Models

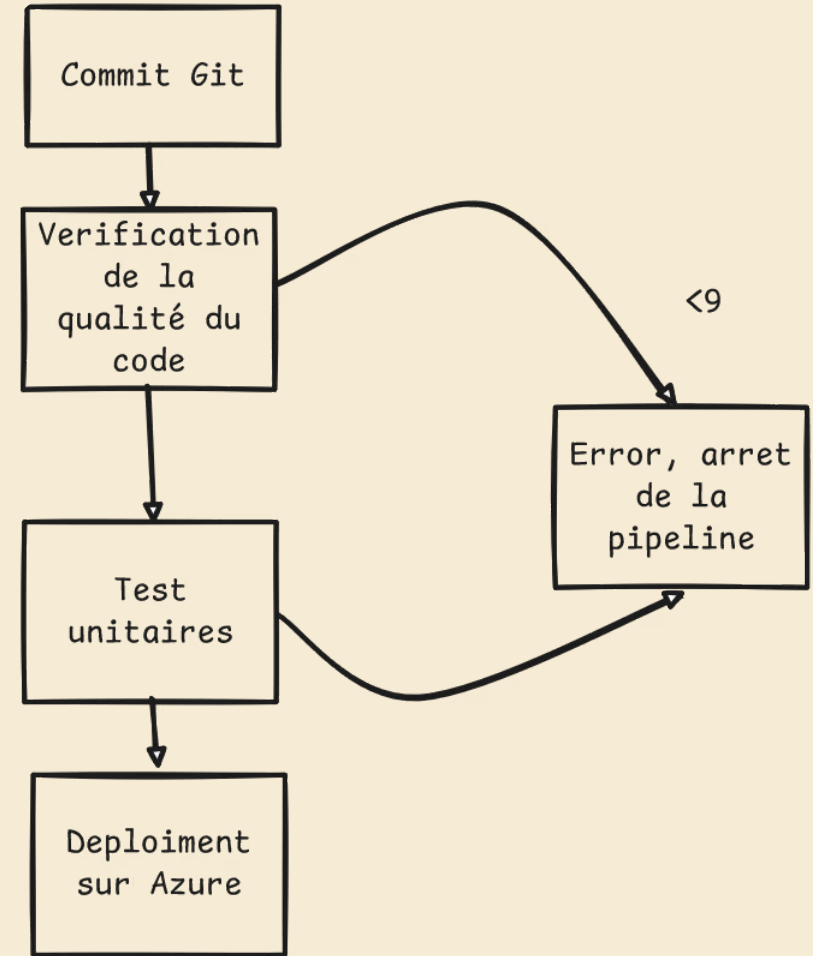
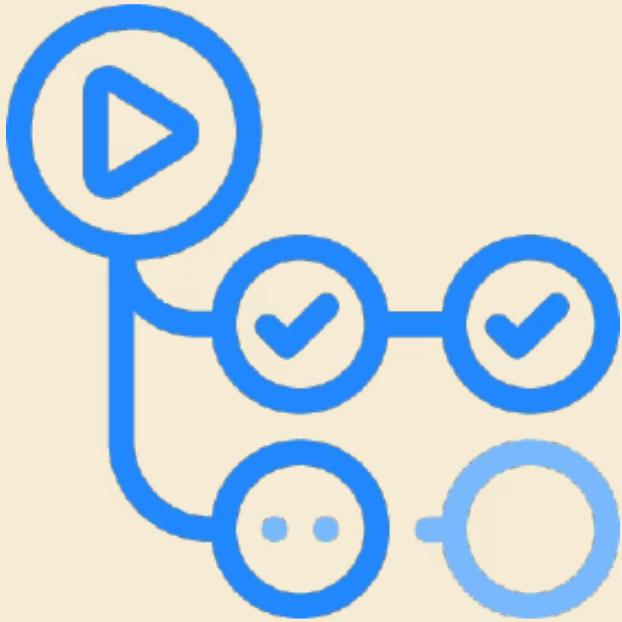
General format for
sending models to
diverse deploy tools

Tracking, stockage model, gestion version

Run Name	Created	Duration	Metrics		Tags			
			accuracy	F ₁	dataset_used	embedding_met	preprocessing	sample_size
LogisticRegression_tokent...	28 minutes ago	4.4s	0.77		sentiment140	transformer	tokenized	5000
LogisticRegression_raw_t...	29 minutes ago	3.6s	0.766		sentiment140	transformer	raw	5000
LogisticRegression_lexm...	29 minutes ago	3.6s	0.765		sentiment140	transformer	lemmatized	5000
LogisticRegression_stem...	29 minutes ago	3.8s	0.748		sentiment140	transformer	stemmed	5000
LogisticRegression_raw_t...	29 minutes ago	8.3s	0.765		sentiment140	tfidf	raw	5000
LogisticRegression_stem...	29 minutes ago	4.7s	0.763		sentiment140	tfidf	stemmed	5000
LogisticRegression_lexm...	29 minutes ago	3.7s	0.737		sentiment140	tfidf	lemmatized	5000
LogisticRegression_tokent...	29 minutes ago	4.0s	0.734		sentiment140	tfidf	tokenized	5000



Pipeline de Déploiement Continu



Prototype (demo)

Application



API





Suivi de la Performance en Production

Conclusion

Prototype valide

- Possibilité de prédire sur une phrase ou un dataset entier

Modèle sur mesure avancé
moins performant que
BERT

Cycle MlOps respecté

- Mise en place de pipeline de mise en production
- Log & Monitoring des modèles

Amelioration Proposé

Centralisation du monitoring

Dashboard Graphana +
Opentelemetry

Monitoring des données

Prevention du datadrift
Cout du stockage des données

Modèle plus performant

BERT model

Test plus poussés

Performance, Integration, Sécurité

Solutions de deployment

Edge, Cloud, Local ...