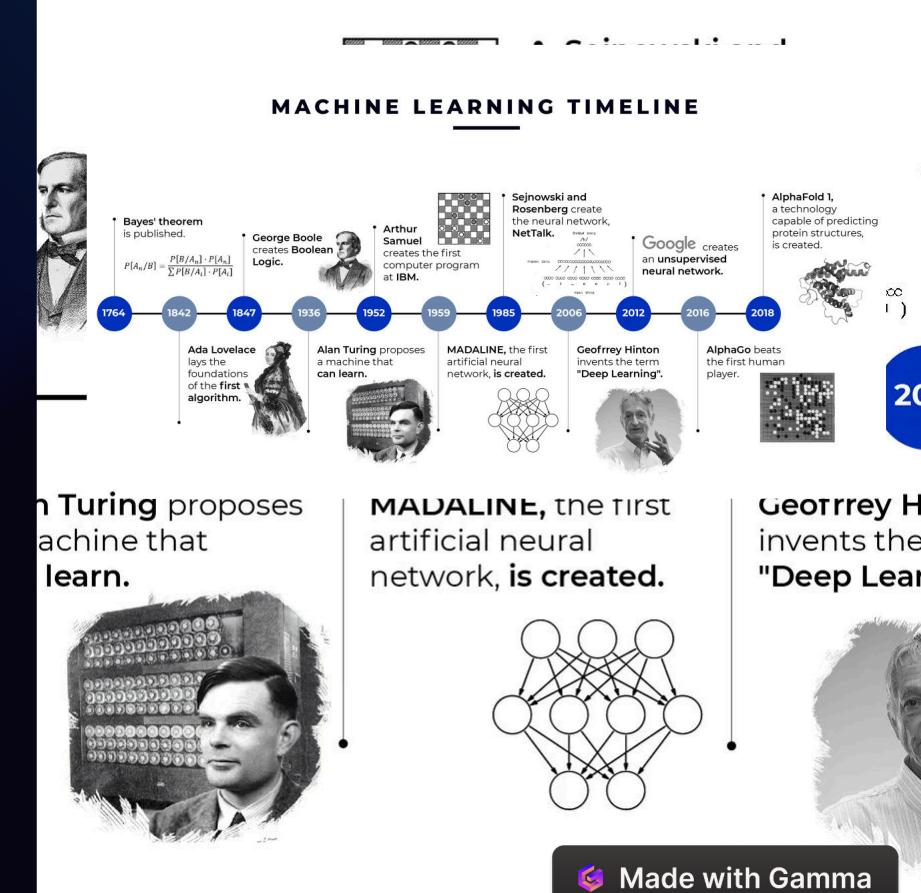


Défis et évolution du Machine Learning

Le machine learning est un domaine en constante évolution, confronté à de nombreux défis passionnants. Cette présentation explore les avantages de l'apprentissage automatique, les principes fondamentaux du MLOps et de DevOps/DataOps, offrant une perspective globale sur les tendances actuelles et futures du machine learning.

E by Etienne Koa

MACHINE LEARNING TIME



Principes fondamentaux du MLOps

Problèmes résolus par MLOps

Le MLOps vise à relever les défis liés à la mise en production et à la maintenance des modèles de machine learning, en assurant une collaboration étroite entre les équipes de développement, d'exploitation et de données.

Composants MLOps

Les principaux composants du MLOps incluent la gestion du cycle de vie des modèles, l'automatisation des pipelines, le versionnage des données et des modèles, ainsi que le monitoring et la surveillance des performances en production.

Boîte à outils MLOps

La boîte à outils MLOps comprend des outils tels que MLflow, DVC, Dagshub, Evidently AI et BentoML, qui permettent d'automatiser et de standardiser les processus liés au développement et au déploiement des modèles de machine learning.

Productivité et structuration des projets ML

1

Cookiecutter

Cookiecutter permet de gérer la structure du projet de machine learning de manière standardisée, facilitant ainsi la collaboration et la reproductibilité.

2

Gestion de projet

Des outils comme Poetry, Makefile et Hydra permettent de gérer efficacement les dépendances, l'exécution automatisée des tâches et la configuration des projets ML.

3

Qualité du code

Des outils comme Black, Flake8, Isort et Interrogate assurent la qualité du code en automatisant les vérifications et les corrections avant la validation dans Git.

4

Documentation

La génération automatique de documentation pour les projets ML facilite la compréhension et la maintenance à long terme.



Automatisation du cycle du modèle ML

1

Conception

La conception de solutions MLOps s'appuie sur des méthodologies comme Volere, qui permettent de définir les exigences et les spécifications du projet de manière rigoureuse.

2

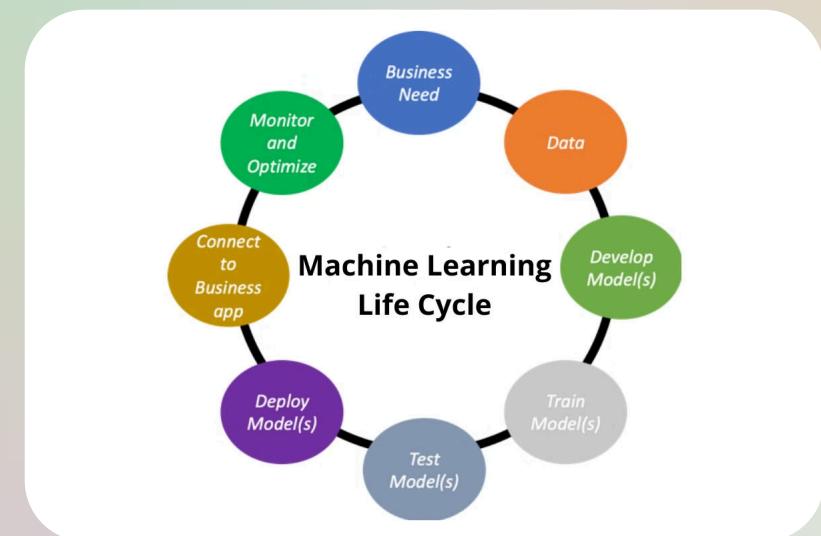
Développement

L'automatisation du cycle du modèle ML s'appuie sur des outils comme Pycaret, qui facilitent l'exploration des données, le développement de modèles avancés et le déploiement en production.

3

Déploiement

Le versionnage et l'enregistrement des modèles avec MLflow, ainsi que le versionnage des données avec DVC, assurent la traçabilité et la reproductibilité des projets ML.



Gestion de code et de données



DagsHub

DagsHub offre un dépôt de code centralisé pour les projets ML, permettant de gérer le code, les données et les modèles de manière collaborative.



DVC

DVC permet de versionner les ensembles de données et les modèles, assurant la traçabilité et la reproductibilité des expériences ML.



MLflow

MLflow gère le versionnage et l'enregistrement des modèles, facilitant le déploiement et la surveillance des performances en production.



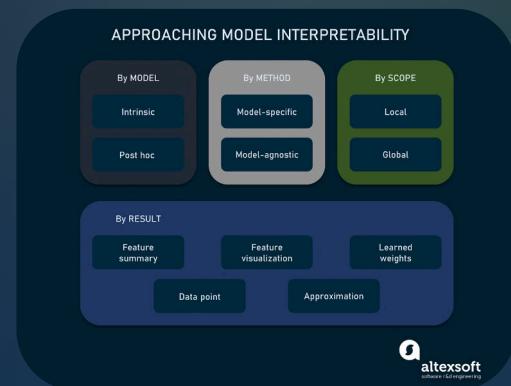
Interprétabilité et déploiement des modèles

Interprétabilité

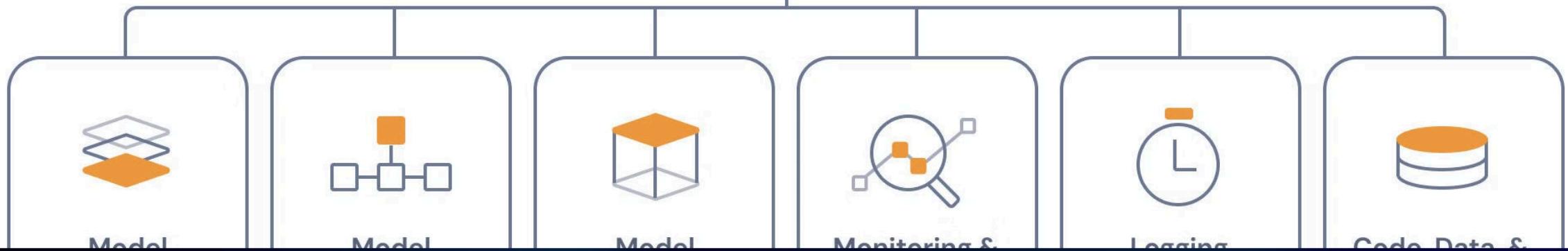
L'interprétabilité des modèles de machine learning est essentielle pour comprendre leur fonctionnement et leur prise de décision. L'utilisation d'outils comme SHAP permet d'expliquer les prédictions des modèles.

Déploiement

Le déploiement des modèles en production peut se faire via des API développées avec des outils comme FastAPI, ou des applications web construites avec Gradio ou Streamlit.



Machine Learning Model Management



Surveillance et déploiement continu



Surveillance

La surveillance des modèles et des services ML est essentielle pour détecter les dérives de données et de performances, et assurer la qualité en production. Evidently AI est un outil puissant pour cette tâche.

CI/CD

L'intégration et la livraison continues (CI/CD) avec des outils comme GitHub Actions et CML permettent d'automatiser les workflows de validation, de test et de déploiement des applications ML.

Déploiement

Le déploiement des services ML peut se faire sur des plateformes cloud comme Azure ou Heroku, en utilisant des outils comme BentoML pour générer des services ML prêts à l'emploi.

Projet MLOps de bout en bout

| Développement du modèle | Validation de la qualité | Versionnage et partage | Déploiement et surveillance |
|---------------------------------|---|-------------------------------|--|
| Exploration des données | Vérification du code, du modèle et du prétraitement | Utilisation de MLflow et DVC | Développement d'API avec BentoML |
| Entraînement de modèles avancés | Validation continue avec GitHub Actions | Partage sur DagsHub et MLflow | Applications web avec Streamlit |
| - | - | - | Déploiement automatisé avec GitHub Actions et Heroku |