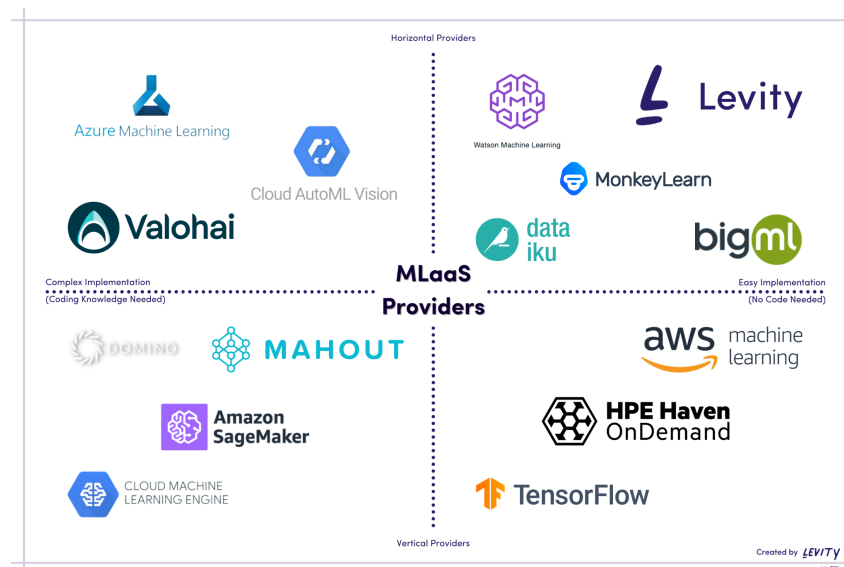


ML in the clouds

"Aide toi, le cloud t'aidera." – Jean de La Fontaine

Le Machine Learning, un service



Nous vivons actuellement une véritable **révolution de l'intelligence artificielle**, marquée par des avancées technologiques rapides et des applications de plus en plus diversifiées. Dans ce contexte, **le Machine Learning as a Service (MLaaS)** émerge comme **une solution stratégique**, permettant aux entreprises d'utiliser des services cloud pour **développer, déployer et gérer des modèles de machine learning sans avoir à gérer l'infrastructure sous-jacente**. Avec le MLaaS, des outils et des ressources dédiées à l'intelligence artificielle sont accessibles via **des API ou des interfaces intuitives**, ce qui simplifie le processus de développement et réduit les barrières à l'adoption de l'IA.



En utilisant des services MLaaS, les entreprises peuvent exploiter les avantages du machine learning **sans avoir à investir massivement dans l'infrastructure et l'expertise requises pour construire et maintenir des systèmes de machine learning complexes.**



Durant votre parcours plein **d'aventures palpitantes et de péripéties**, vous avez amassé un savoir considérable et précieux dont vous êtes bien fiers ! Plus rien ne vous arrête. Sur cet élan, **vous embarquez dans une nouvelle quête, l'AutoML et le MLaaS.**

Vous réalisez une **veille concise** sur les notions et les différentes plateformes de MLaaS ci-dessous :

- **L'AutoML et ses avantages,**
- **Le MLaaS et ses avantages,**
- **Google Cloud Platform et Vertex AI,**
- **AWS et Sagemaker,**
- **Azure et Azure Machine Learning,**
- **DataRobot.**



Analyse de sentiments

Votre travail ne s'arrête pas là, vous testez ce que vous avez appris sur une problématique qui vous intéresse, le **traitement naturel du langage** pour **l'analyse de sentiments**.



A l'aide de l'**AutoML**, vous devrez proposer plusieurs modèles de classification multi-classes permettant de distinguer au moins 6 émotions humaines différentes afin de répondre à la question : **est-il possible d'identifier des émotions dans des phrases narratives issues de communications écrites ?**

Pour ce faire, vous utilisez le jeu de données d'entraînement fourni [ici](#) pour réaliser **vos apprentissage et l'évaluation de vos modèles**. Dans l'objectif d'enrichir votre prédiction, **vous augmentez votre jeu de données** avec des données de test fournies [ici](#).

Vous testez l'outil **Pycaret**, une bibliothèque d'apprentissage automatique en Python, open-source et Lowcode, qui automatise le workflow de



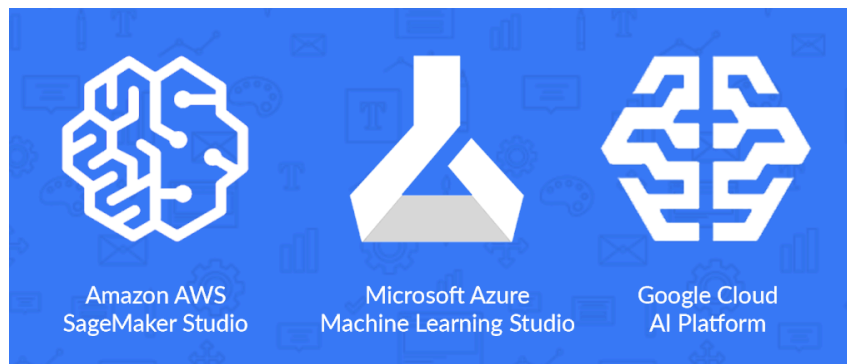
l'apprentissage automatique. Vous passez par les étapes habituelles pour la construction d'**un modèle de prédiction** (à l'aide de **Pycaret** bien sûr) :

- Vous pré-traitez vos données et vous les analysez. Vous réalisez une visualisation de données bien propre au NLP.
- Vous entraînez plusieurs modèles et vous les comparez selon différents métriques d'évaluation.



Vers le cloud et au delà

Fascinés par les services de cloud proposés par Google, Amazon, Microsoft et **DataRobot**, **vous décidez de tester l'un d'eux**. Vous utilisez les versions d'essais offertes afin de manipuler les services dédiés au MLaaS et **vous déployez un modèle d'analyse de sentiment**. Pensez à garder des captures d'écran **des différentes étapes de la construction au déploiement et au monitoring** (ou des scripts s'il y en a) afin de ne pas perdre votre travail à la fin de la période d'essai.





Compétences visées

→ MLops

Rendu

L'évaluation de ce projet se fera sur deux aspects :

1. Une **présentation explicative de votre travail** sous forme de diapositives. Celle-ci doit raconter la réflexion derrière votre travail et inclure votre **Trello d'organisation**. **Pensez à inclure les captures d'écran (voir une démonstration) de la section 3.**
2. Un repository github public nommé **ml-in-the-clouds**, contenant les éléments suivants :
 - a. Un notebook Jupyter **propre et commenté** (introduction, analyses, conclusion, etc) contenant le procédé de développement de votre outil, du **prétraitement** à la **modélisation des données**, en passant par l'**analyse exploratoire**. **Pensez à répondre à la problématique.** Vous pouvez avoir au **maximum** deux notebooks, un pour l'exploration et l'autre pour la modélisation de données.
 - b. Un fichier **README.md** présentant le contexte du projet, les données et leur analyse, les différents outils utilisés et une conclusion sur votre travail. **Pensez à inclure la veille réalisée.**



Base de connaissances

- [Experience the DataRobot AI Platform](#)
- [Google Cloud Platform](#)
- [Amazon AWS](#)
- [Microsoft Azure](#)
- [PyCaret 3.0 : An open-source, low-code machine learning library in Python](#)
- [PyCaret Multiclass Classification Tutorial - Notebook](#)
- [Nltk.org : Natural Language Toolkit Documentation](#)
- [Text Classification with NLTK and Scikit-Learn by Benjamin Bengfort](#)