

# Demande de proposition: Solution de plateforme MLOps

## Table des matières

1. Introduction et contexte
2. Objectifs du projet
3. Exigences techniques
4. Exigences fonctionnelles
5. Soutien et maintenance
6. Critères d'évaluation
7. Lignes directrices pour la soumission
8. Chronologie

### 1. Introduction et contexte

[Nom de l'entreprise] lance un appel d'offres pour une plateforme MLOps (Machine Learning Operations) complète afin de rationaliser nos opérations d'apprentissage automatique. Cet appel d'offres décrit nos besoins pour une solution de bout en bout qui nous permettra de gérer efficacement l'ensemble du cycle de vie de nos projets d'apprentissage automatique.

#### 1.1 Historique de l'organisation

- Secteur d'activité et objectif principal de l'entreprise
- Initiatives actuelles en matière de ML/AI
- Ampleur des opérations
- Environnement réglementaire
- Motivations spécifiques des entreprises pour la mise en œuvre de MLOps

#### 1.2 Environnement actuel

- Outils et plateformes existants
- Structure et taille de l'équipe

- Points faibles actuels
- Exigences d'intégration
- Processus de déploiement du modèle actuel

## 2. Objectifs du projet

### 2.1 Objectifs principaux

- Mettre en œuvre une plateforme MLOps évolutive pour gérer et surveiller les modèles d'apprentissage automatique.
- Rationaliser le processus de développement, de déploiement et de maintenance des modèles de ML
- Améliorer la collaboration entre les scientifiques des données, les ingénieurs et les parties prenantes de l'entreprise
- Veiller au respect des exigences réglementaires et des normes industrielles
- Permettre des itérations rapides dans les cycles de développement des modèles
- Réduire le temps de déploiement des modèles de ML
- Standardiser les pratiques de développement du ML au sein des équipes
- Améliorer la reproductibilité et la traçabilité des modèles
- Optimiser l'utilisation des ressources et la gestion des coûts
- Mettre en place des processus d'assurance qualité cohérents

## 3. Exigences techniques

### 3.1 Architecture de la plate-forme

- Options de déploiement en nuage (public, privé, hybride)
- Possibilités de déploiement sur site
- Prise en charge multirégionale
- Architecture à haute disponibilité
- Capacités de reprise après sinistre

- Prise en charge de la conteneurisation
- Compatibilité avec l'architecture des microservices

### 3.2 Capacités d'intégration

- Prise en charge de l'API REST pour les intégrations personnalisées
- Intégration avec les technologies existantes
- Prise en charge des frameworks ML courants (TensorFlow, PyTorch, scikit-learn)
- Intégration d'un système de contrôle de version (Git)
- Compatibilité avec le pipeline CI/CD
- Connecteurs de sources de données
- Intégration du système d'authentification

### 3.3 Performance et évolutivité

- Spécifications de la taille maximale du modèle
- Capacité d'utilisateurs simultanés
- Exigences en matière de temps de réponse
- Limites d'utilisation des ressources
- Capacités de mise à l'échelle horizontale et verticale
- Spécifications de l'équilibrage de charge
- Capacités de traitement par lots

### 3.4 Exigences en matière de sécurité

- Cryptage des données (au repos et en transit)
- Contrôle d'accès basé sur les rôles (RBAC)
- Intégration de l'authentification unique (SSO)
- Journalisation des audits
- Certifications de conformité (SOC 2, ISO 27001, etc.)

- Exigences en matière de sécurité des réseaux
- Normes de sécurité de l'API

### 3.5 Gestion des ressources

- Allocation et gestion des GPU/CPU
- Optimisation de la mémoire
- Gestion du stockage
- Orchestration de conteneurs
- Surveillance des ressources et alertes
- Fonctions d'optimisation des coûts

## 4. Exigences fonctionnelles

### 4.1 Gestion des données

**Conseil : une gestion efficace des données constitue la base du MLOps. Concentrez-vous sur les fonctionnalités garantissant la qualité, la version et l'accessibilité des données tout en maintenant la conformité. Prenez en compte les besoins de traitement par lots et en temps réel, et assurez-vous que la solution peut gérer votre volume de données.**

Exigence	Sous-exigence	O/N	Notes
Version des données	Contrôle de version pour les ensembles de données		
	Suivi des données		
	Documentation sur l'historique des modifications		
Ingénierie des fonctionnalités	Capacités de la boutique en ligne		
	Pipelines de calcul des caractéristiques		
	Version des fonctionnalités		

Qualité des données	Outils de contrôle de la qualité		
	Cadres de validation		
	Capacités de profilage des données		
Intégration des données	Prise en charge des données structurées		
	Prise en charge des données non structurées		
	Connectivité de sources multiples		
Traitement en temps réel	Capacité de traitement des flux		
	Validation des données en temps réel		
	Traitement à faible latence		
Conservation des données	Gestion des politiques		
	Archivage automatisé		
	Contrôle de la conformité		

#### 4.2 Développement du modèle

**Conseil : prenez en charge l'ensemble de votre flux de travail en science des données, de l'expérimentation à la production, grâce à des fonctions robustes de contrôle des versions et de collaboration. Assurez-vous de la compatibilité de la plateforme avec les outils et frameworks préférés de votre équipe.**

Exigence	Sous-exigence	O/N	Notes
Suivi des expériences	Versionnement de l'expérience		
	Suivi des paramètres		
	Comparaison des résultats		
Soutien linguistique	Intégration de Python		

	R intégration		
	Prise en charge d'autres langues		
Sélection des caractéristiques	Sélection automatisée des caractéristiques		
	Analyse de l'importance des caractéristiques		
	Analyse de corrélation des caractéristiques		
Intégration du cadre	Prise en charge de TensorFlow		
	Support de PyTorch		
	Prise en charge de Scikit-learn		
Environnement de développement	Intégration d'un carnet de notes Jupyter		
	Support IDE		
	Version du code		

#### 4.3 Formation au modèle

***Conseil : garantir un support de formation évolutif et efficace pour les différents paradigmes. Équilibrez les ressources informatiques et les capacités d'orchestration tout en maintenant la reproductibilité et une validation appropriée.***

Exigence	Sous-exigence	O/N	Notes
Infrastructure de formation	Prise en charge du GPU		
	Formation distribuée		
	Capacités multi-nœuds		

Méthodes d'apprentissage	Apprentissage supervisé		
	Apprentissage non supervisé		
	Apprentissage par renforcement		
	Apprentissage par transfert		
Gestion des ressources	Mise à l'échelle dynamique		
	Allocation des ressources		
	Optimisation des coûts		
Gestion des ensembles de données	Traitement des données de validation		
	Versionnement du jeu de données de test		
	Capacités de fractionnement des ensembles de données		
Visualisation de la formation	Affichage des données en temps réel		
	Suivi métrique personnalisé		
	Visualisation des performances		

#### 4.4 Déploiement du modèle

**Conseil : permettre un déploiement automatisé et fiable avec la prise en charge de plusieurs modèles. Concentrez-vous sur les capacités de déploiement continu tout en conservant les fonctionnalités de contrôle de version et de retour en arrière.**

Exigence	Sous-exigence	O/N	Notes
Options de déploiement	Déploiement de l'API REST		
	Inférence par lots		

	Déploiement en périphérie		
Essais	Capacité de test A/B		
	Déploiement de Canary		
	Tests d'intégration		
Gestion de l'environnement	Environnement de développement		
	Environnement de transit		
	Environnement de production		
Santé du déploiement	Contrôle de l'état des services		
	Suivi de l'utilisation des ressources		
	Mesures de performance		
	Contrôles de santé automatisés		

#### 4.5 Surveillance du modèle

**Conseil : Un suivi complet est essentiel pour maintenir la performance et la fiabilité des modèles en production. La plateforme doit offrir des capacités de surveillance en temps réel avec des alertes automatisées et la détection des dérives, afin de garantir que les modèles restent précis et efficaces au fil du temps.**

Exigence	Sous-exigence	O/N	Notes
Suivi des performances	Mesures en temps réel		
	Analyse historique		
	Mesures personnalisées		
Détection de la dérive	Contrôle de la dérive des données		
	Détection de la dérive des concepts		

	Alertes sur la dérive des performances		
Modèle de notation de la santé	Définition des indicateurs de santé		
	Algorithmes de notation		
	Analyse des tendances en matière de santé		
Alerte	Configuration des alertes		
	Canaux de notification		
	Priorité aux alertes		
Rapports	Rapports automatisés		
	Tableaux de bord personnalisés		
	Rapports de conformité		

#### 4.6 Gestion des modèles

**Conseil : une gestion efficace des modèles nécessite un suivi et une organisation complets de tous les actifs ML. La plate-forme doit offrir de solides capacités de catalogage, de versionnement et de documentation afin de maintenir une lignée de modèles claire et une bonne gouvernance au sein de l'organisation.**

Exigence	Sous-exigence	O/N	Notes
Registre des modèles	Catalogage des modèles		
	Suivi des versions		
	Gestion des métadonnées		
Comparaison des modèles	Comparaison des performances		
	Comparaison de l'utilisation des ressources		

	Comparaison de l'importance des caractéristiques		
Suivi des dépendances	Dépendances de la bibliothèque		
	Dépendances des données		
	Dépendances de l'environnement		
Documentation	Documentation automatisée		
	Modèles de cartes		
	Guide d'utilisation		
Flux d'approbation	Processus d'examen des modèles		
	Gestion de la chaîne d'approbation		
	Suivi des signatures		
Gestion du cycle de vie	Suivi de l'état d'avancement		
	Processus de retraite		
	Gestion des archives		

#### 4.7 Outils de collaboration

**Conseil : Permettre une collaboration transparente entre les scientifiques des données, les ingénieurs et les parties prenantes grâce à des outils et des flux de travail intégrés. La plateforme doit permettre le partage de code, le transfert de connaissances et une communication efficace tout en respectant les normes de sécurité.**

Exigence	Sous-exigence	O/N	Notes
Espaces de travail partagés	Gestion de l'espace de travail de l'équipe		
	Partage des ressources		
	Contrôle d'accès		

Contrôle des versions	Version du code		
	Gestion de l'agence		
	Fusionner les capacités		
Modèles de projets	Création de modèles		
	Gestion des modèles		
	Partage de modèles		
Partage des connaissances	Partage de la documentation		
	Bibliothèque de bonnes pratiques		
	Modèles de code		
Analyse de la collaboration	Mesures de l'activité de l'équipe		
	Suivi des contributions		
	Modèles de collaboration		
Communication	Notifications d'équipe		
	Systèmes de commentaires		
	Examiner les flux de travail		

#### 4.8 Gouvernance et conformité

**Conseil : mettez en place des mécanismes de gouvernance solides pour garantir la conformité réglementaire et des pratiques d'IA responsables. La plateforme doit offrir des capacités d'audit complètes, des contrôles d'accès et une application des politiques tout en maintenant l'efficacité opérationnelle.**

Exigence	Sous-exigence	O/N	Notes
Contrôle d'accès	Approvisionnement des utilisateurs		
	Accès basé sur les rôles		

	Gestion des autorisations		
Pistes d'audit	Enregistrement des activités		
	Suivi des changements		
	Enregistrement des accès		
Application de la politique	Politiques de conformité		
	Application automatisée		
	Alertes en cas de violation de la politique		
Flux de travail de gouvernance	Flux de travail pour la création de politiques		
	Procédures d'approbation		
	Contrôle de conformité		
	Gestion des exceptions		
Protection des données	Traitement des IPI		
	Masquage des données		
	Restrictions d'accès		

#### 4.9 Explicabilité et transparence

**Conseil : les capacités d'explication des modèles sont essentielles pour instaurer la confiance et répondre aux exigences réglementaires. Veillez à disposer d'outils complets pour comprendre les décisions des modèles et identifier les biais potentiels dans tous les modèles déployés.**

Exigence	Sous-exigence	O/N	Notes
Interprétation du modèle	Importance de la caractéristique		
	Valeurs du SHAP		

	Analyse LIME		
Analyse des décisions	Visualisation du chemin de décision		
	Explications de la prédiction		
	Analyse contrefactuelle		
Explications personnalisées	Intégration de méthodes personnalisées		
	Modèles d'explication		
	Explications spécifiques au domaine		
Détection des biais	Mesures des biais		
	Analyse de l'équité		
	Évaluation démographique		
Rapports	Rapports d'explication		
	Documentation de conformité		
	Communication avec les parties prenantes		

#### 4.10 Capacités AutoML

***Conseil : Accélérer le développement de modèles tout en maintenant la qualité grâce à des fonctions d'apprentissage automatique. La plateforme doit automatiser les tâches répétitives tout en permettant à un expert de superviser et de personnaliser le pipeline de développement.***

Exigence	Sous-exigence	O/N	Notes
Sélection des caractéristiques	Sélection automatisée des caractéristiques		
	Classement des caractéristiques		

	Ingénierie des fonctionnalités		
Sélection du modèle	Sélection d'un algorithme		
	Comparaison des modèles		
	Optimisation des performances		
Personnalisation du pipeline	Définition d'un pipeline personnalisé		
	Modèles de pipelines		
	Configuration des composants		
Réglage des hyperparamètres	Mise au point automatisée		
	Définition de l'espace de recherche		
	Stratégies d'optimisation		
Documentation du modèle	Documentation automatisée		
	Rapports de performance		
	Enregistrement de la configuration		

#### 4.11 Intégration du pipeline CI/CD

**Conseil : permettre une intégration transparente avec les pratiques DevOps existantes tout en ajoutant des capacités spécifiques au ML. La plateforme doit prendre en charge les tests automatisés, le déploiement et la validation des modèles dans le cadre des flux de travail CI/CD établis.**

Exigence	Sous-exigence	O/N	Notes
Cadre de test	Tests unitaires		
	Tests d'intégration		
	Tests de performance		

Automatisation des pipelines	Constructions automatisées		
	Déploiement automatisé		
	Contrôles de validation		
Surveillance des pipelines	Contrôle des performances		
	Analyse des pipelines		
	Suivi des erreurs		
Intégration des outils	Intégration de Git		
	Intégration de Jenkins		
	Soutien aux conteneurs		
Automatisation du retour en arrière	Déclencheurs de retour en arrière automatisés		
	Intégration du contrôle de version		
	Gestion de l'État		
Portes de qualité	Contrôles de la qualité du code		
	Contrôles de la qualité des modèles		
	Analyse de sécurité		

#### 4.12 Gestion et optimisation des coûts

**Conseil : maintenir la visibilité et le contrôle sur l'utilisation des ressources et les coûts associés. La plateforme devrait fournir un suivi détaillé, des recommandations d'optimisation et des capacités de prévision pour toutes les opérations de ML.**

Exigence	Sous-exigence	O/N	Notes
Suivi des ressources	Contrôle de l'utilisation		

	Répartition des coûts		
	Utilisation des ressources		
Gestion du budget	Fixation du budget		
	Seuils d'alerte		
	Rapport sur les coûts		
Coût Détection des anomalies	Règles de détection des anomalies		
	Seuils d'alerte		
	Comparaison historique		
Optimisation	Optimisation des ressources		
	Recommandations en matière de coûts		
	Mise à l'échelle automatisée		
Prévisions	Prévisions d'utilisation		
	Prévision des coûts		
	Analyse des tendances		

## 5. Soutien et maintenance

### 5.1 Accords de niveau de service

- Engagements en matière de temps de réponse
- Engagements en matière de temps de résolution
- Garanties de disponibilité du système
- Mesures de performance
- Clauses pénales

- Structure des crédits de service
- Méthodes de mesure et de rapport

#### 5.2 Services d'appui

- Procédures d'assistance en cas d'urgence (assistance 24/7 pour les problèmes critiques)
- Équipe d'assistance sur appel
- Procédure d'escalade en cas d'urgence
- Définition du soutien de niveau 1/2/3
- Temps de réponse par niveau
- Critères d'escalade
- Processus d'escalade de la gestion

#### 5.3 Accès à la base de connaissances

- Documentation en ligne
- Guides de bonnes pratiques
- Guides de dépannage
- Forums communautaires
- Tutoriels vidéo
- Documentation de l'API
- Fenêtres d'entretien régulier
- Procédures de gestion des correctifs
- Prise en charge de la mise à jour des versions
- Soutien au développement personnalisé

#### 5.4 Formation et habilitation

- Programme de formation initiale
- Formation avancée des utilisateurs

- Formation des administrateurs
- Mise à jour régulière des connaissances
- Options de formation personnalisées
- Programmes de certification
- Matériels et ressources de formation

## 6. Critères d'évaluation

### 6.1 Complétude de la solution (20 %)

- L'exhaustivité de la solution MLOps
- Couverture de toutes les exigences fonctionnelles et techniques requises
- Intégralité de la méthodologie de mise en œuvre
- Qualité de l'interface utilisateur et de l'expérience
- Capacités d'intégration
- Maturité de la plate-forme

### 6.2 Architecture technique (20%)

- Capacités d'évolutivité et de performance
- Fiabilité et disponibilité de la plate-forme
- Caractéristiques de sécurité et mesures de conformité
- Flexibilité d'intégration
- Innovation technique
- Qualité de la conception architecturale

### 6.3 Capacités d'intégration (15%)

- Facilité d'intégration avec les systèmes existants
- Complétude de l'API et documentation
- Prise en charge des protocoles et formats standard

- Options d'extension
- Capacités d'intégration personnalisées
- Support d'outils tiers

#### 6.4 Expérience des fournisseurs (15%)

- Expérience dans la mise en œuvre de MLOps
- Expertise sectorielle et présence sur le marché
- Stabilité financière
- Références clients
- Feuille de route pour le développement
- L'histoire de l'innovation

#### 6.5 Services d'appui (15%)

- Qualité de l'assistance technique
- Formation et documentation
- Services de mise en œuvre
- Maintenance et mises à jour continues
- Disponibilité des ressources
- Temps de réponse

#### 6.6 Coût et retour sur investissement (15%)

- Coût total de possession
- Clarté de la structure des prix
- Valeur de l'investissement
- Rendement attendu de l'investissement
- Prévisibilité des coûts
- Coûts de mise à l'échelle

## 7. Lignes directrices pour la soumission

### 7.1 Contenu obligatoire de la proposition

#### 1. Résumé

- Présentation de l'entreprise
- Points forts de la solution
- Résumé de l'approche de mise en œuvre
- Calendrier et coûts estimés

#### 2. Description de la solution technique

- Architecture détaillée
- Capacités de la plate-forme
- Spécifications techniques
- Mesures de sécurité

#### 3. Approche de la mise en œuvre

- Méthodologie
- Phases du projet
- Besoins en ressources
- Gestion des risques

#### 4. Modèle de soutien

- Niveaux de soutien
- Temps de réponse
- Procédures d'escalade
- Calendrier d'entretien

#### 5. Structure des prix

- Frais de licence

- Coûts de mise en œuvre
- Coûts de formation
- Coûts de l'assistance permanente
- Frais de service supplémentaires

#### 6. Historique de l'entreprise

- Historique de l'entreprise
- Informations financières
- Qualifications de l'équipe
- Expérience en matière de MLOps

#### 7. Références clients

- Au moins trois références
- Mises en œuvre similaires dans l'industrie
- Portée et résultats du projet
- Informations sur les contacts

#### 8. Exemple de documentation

- Documentation de la plate-forme
- Matériel de formation
- Spécifications techniques
- Guides d'utilisation

#### 9. Calendrier du projet

- Calendrier détaillé de la mise en œuvre
- Définitions des jalons
- Allocation des ressources

- Plan de communication

## 10. Plan de gestion des risques

- Identification des risques
- Stratégies d'atténuation
- Plans d'urgence
- Processus de résolution des problèmes

### 7.2 Format de soumission

- Format du fichier : PDF
- Longueur maximale : [X] pages
- Méthode de soumission : [Préciser la livraison électronique/physique]
- Exemplaires requis : [Préciser le nombre]

## 8. Calendrier

### 8.1 Calendrier de l'appel d'offres

- Date de publication de l'appel d'offres : [Date]
- Questions à remettre : [Date]
- Réponse aux questions : [Date]
- Date limite de dépôt des propositions : [Date]
- Évaluation initiale : [Date]
- Présentations des fournisseurs : [Fourchette de dates]
- Sélection finale : [Date]
- Négociation du contrat : [Période]
- Lancement du projet : [Date]

### 8.2 Informations de contact

Pour toute question concernant cet appel d'offres, veuillez contacter

[Nom] [Titre] [Courriel] [Téléphone]

### 8.3 Informations complémentaires

- Contraintes budgétaires (le cas échéant)
- Processus de décision
- Exigences en matière de présentation du fournisseur
- Exigences en matière de preuve de concept (le cas échéant)
- Conditions contractuelles
- Toute exigence ou préférence spécifique de l'entreprise