

Ce document est destiné principalement aux personnes citées dans la section [Direction Projet]. Il présente les informations nécessaires à fournir à l'équipe projet pour planifier et développer efficacement votre application web orientée IA. L'équipe de développement utilisera ces réponses de façon à ce que votre application réponde précisément à vos exigences.

Nous utiliserons ce document pour faciliter la discussion qui permettra de préciser le cahier des charges. Ce document deviendra notre contrat concernant ce à quoi ressemblera l'application et sur sa date de livraison. Nous utilisons un processus réactif grâce auquel vous serez étroitement impliqué dans toutes les étapes du développement et qui vous permettra de voir le site se construire au fur et à mesure.

Les données relatives au projet sont organisées dans Jira, et des réunions régulières sont organisées pour évaluer les avancements et ajuster la direction du projet en fonction des besoins.

## DIRECTION DE PROJET

Le projet s'aligne bien avec nos objectifs et nos valeurs. Il est caractérisé par:

- **Nom et type du projet** : film prédiction, application web / mobile orientée IA
- **Date** : 01/04/2025
- **Nom du client** : Jérémy Vangansberg
- **Coordonnées du client**:
  - [linkedin.com/in/jvangansberg](https://www.linkedin.com/in/jvangansberg)
  - jeremy.vangansberg@gmail.com
- **Nom et coordonnées de notre agence** : XXXXXX

Le projet est géré par une équipe de 2 développeurs data/IA, sous la responsabilité d'un Scrum Master et d'un Product Owner. La gestion est orientée sur la méthodologie Agile avec des sprints définis de manière à livrer des versions progressives du produit, en étroite collaboration avec le client. Les étapes sont bien définies dans ce document et sont suivies dans Jira, avec un synchronisme entre les tâches de développement et la gestion du projet.

## OÙ EN SOMMES-NOUS MAINTENANT?

Actuellement, le client utilise une méthode manuelle pour sélectionner les films à projeter, basée sur la veille des nouveautés et son expérience dans les festivals. Cependant, cette approche est chronophage et manque d'automatisation, ce qui le pousse à envisager une solution plus efficace. Le client souhaite intégrer l'intelligence artificielle pour prédire la fréquentation des films dès leur première semaine de sortie, ce qui serait une aide précieuse pour optimiser la programmation

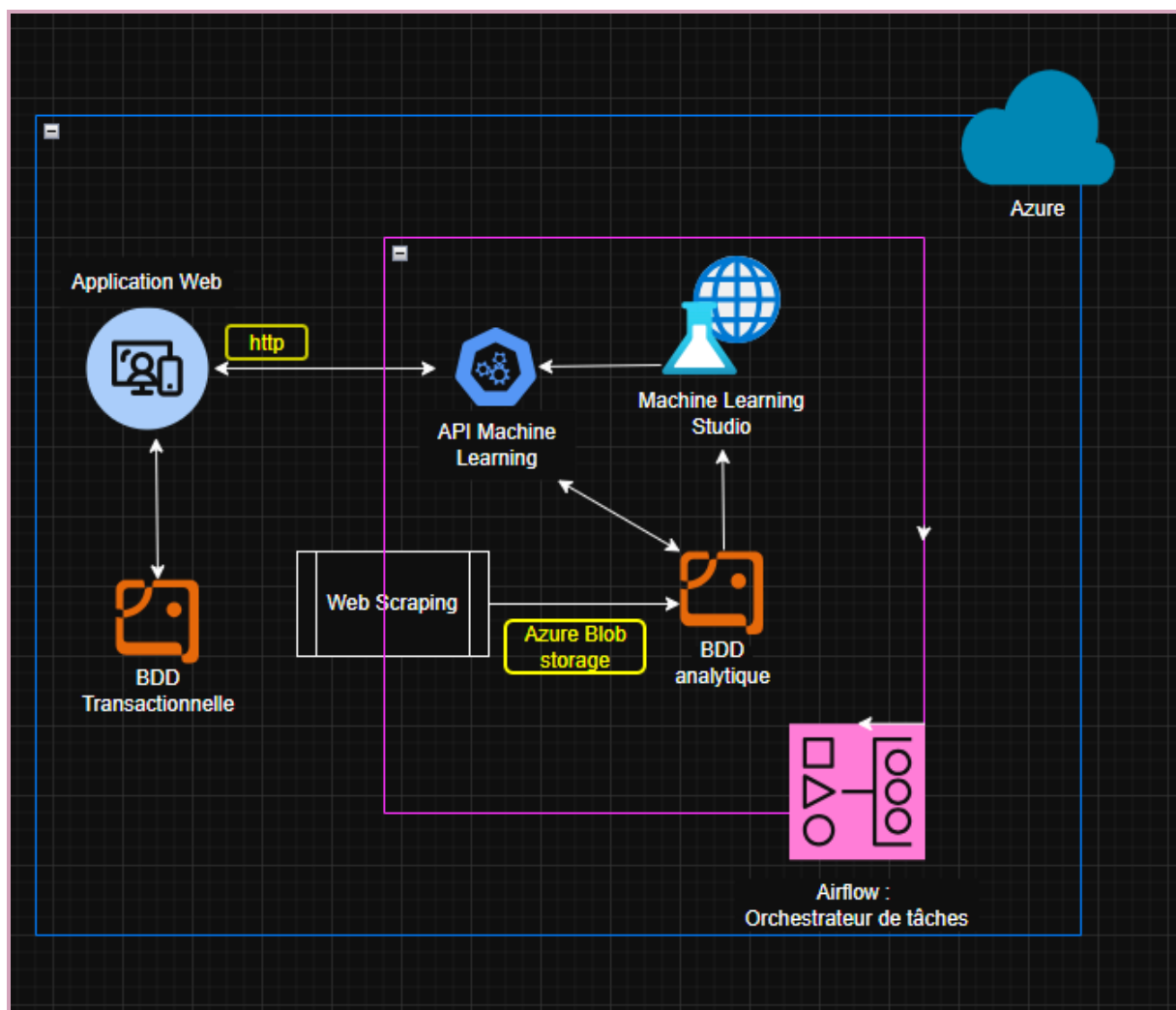
du cinéma et maximiser les revenus. Le projet en est à la phase de définition du besoin, avec un premier livrable à venir, le cahier des charges fonctionnel.

Les principaux problèmes rencontrés sont :

- La gestion manuelle du choix des films, ce qui ralentit la prise de décision.
- L'impossibilité d'avoir une estimation fiable et rapide de l'audience pour chaque film à projeter.
- La difficulté d'adaptation aux différents types de films (internationaux et nationaux).

## OÙ VOULONS-NOUS ARRIVER?

On présente le schéma d'architecture applicative, voir ci-dessous:



**Fig 1.** Schéma d'architecture applicative

Les **objectifs** du projet sont les suivants :

- **Objectifs commerciaux** : Maximiser les revenus en optimisant la programmation des films selon leur potentiel d'audience.

- **Raison de l'investissement** : Automatiser le processus de sélection des films pour gagner du temps et améliorer la rentabilité du cinéma.

- **Indicateurs de performance (KPIs)** :

- Estimation de l'audience pour chaque film.
- Prédiction du chiffre d'affaires hebdomadaire.
- Suivi du taux d'occupation et de la croissance par rapport aux semaines et mois précédents.

- **Retour sur investissement (ROI) optimal** : réduction des coûts et augmentation des recettes par une programmation plus précise.

La **logique** métier se base principalement sur la prédiction du nombre d'entrées pour chaque film. La logique suivante sera appliquée :

1. **Estimation du potentiel des films** : estimation du nombre d'entrées nationales, puis calcul du potentiel pour le cinéma (1/2000ème du trafic national).

2. **Sélection et programmation des films** : le film avec le plus grand potentiel d'audience est affecté à la salle 1 (120 places), le suivant à la salle 2 (80 places).

3. **Calcul de la recette et de l'audience quotidienne** : la recette hebdomadaire est calculée en fonction du nombre d'entrées et du prix moyen d'une place (10 €).

4. **Gestion des coûts** : soustraction des frais fixes de 4900 € pour déterminer la rentabilité de la semaine.

On a deux types de besoins **fonctionnels** (voir Fig 2.) et **non fonctionnels**:

### 1) Besoins fonctionnels

- **Authentification et gestion des paramètres** (ex. utilisateurs, fréquence de scraping): système d'authentification simple pour le gérant du cinéma.

- **Estimation de fréquentation des films** : prédiction des entrées pour chaque film à partir de données pertinentes.

- **Gestion de la programmation** : sélection des films et allocation aux salles.

- **Reporting financier** : calcul du chiffre d'affaires, des recettes et du taux d'occupation.

- **Historique des prédictions** : comparaison entre prédictions et résultats réels, avec des métriques de performance.

- **Interface utilisateur simple** : tableau de bord visuel pour le gérant, incluant les informations essentielles (films, recettes, historique).

## **2) Besoins non fonctionnels**

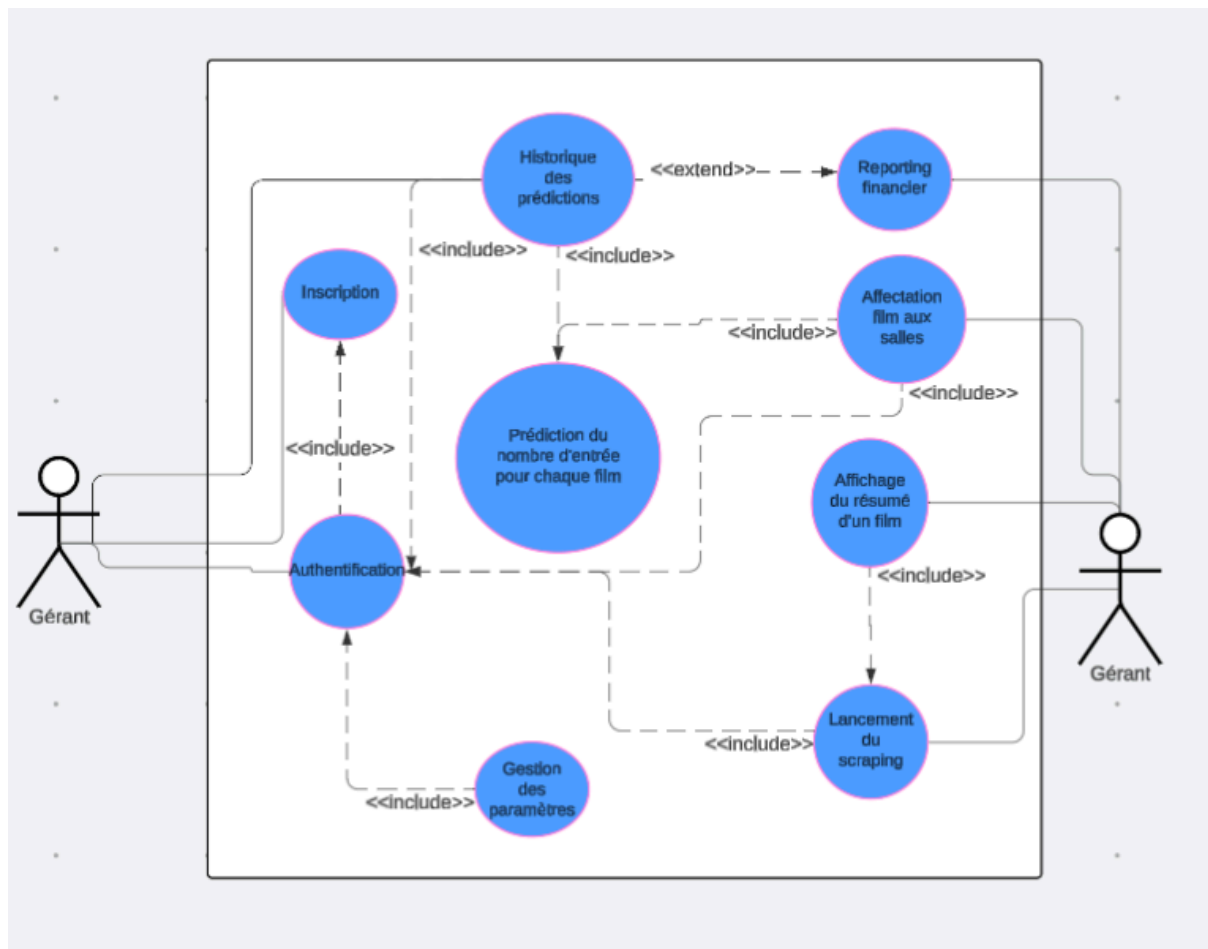
- **Haute disponibilité et scalabilité** : le système doit pouvoir gérer des volumes de données croissants avec un accès rapide et fiable.

- **Sécurité et confidentialité** : respect de la confidentialité des données et des normes de sécurité.

- **Temps de réponse optimisé** : prédictions effectuées rapidement pour permettre une programmation efficace.

- **Interface utilisateur fluide** : design simple et accessible.

- **Maintenance et évolutivité** : Solution facile à maintenir et à faire évoluer.



**Fig 2.** Diagramme de cas d'utilisation

## QU'ALLONS-NOUS FAIRE POUR Y PARVENIR ?

Le plan d'action au niveau organisationnel et technologies:

- **Équipes impliquées** : 2 développeurs IA, Scrum Master, Product Owner.
- **Technologies** : Python, Django, Scrapy, Azure Machine Learning, FastAPI, Docker, Jira, MLflow.
- **Documentation et réunions** : planification de sprints, révisions et feedback réguliers avec le client.
- **Pair programming** : collaborer sur des parties spécifiques du modèle ou de l'intégration.

Le plan d'action au niveau développement des modules:

1. **Intégration des processus de gestion de projet Agile** : Suivi via Jira, avec des sprints réguliers et une revue du projet par le client à chaque sprint.
2. **Automatisation du scraping** : Utiliser Scrapy et des APIs pour collecter les données des films à venir, en prenant soin de ne pas entraîner des biais dans

les données (target leaking). Ces données seront utilisées pour enrichir les prédictions.

3. **Développement de l'application** : Développer une application Django avec un tableau de bord pour le gérant. Ce tableau de bord affichera les prédictions, les films sélectionnés pour chaque semaine, le chiffre d'affaires et l'historique des prédictions.
4. **Création d'un modèle de prédiction** : Développer un modèle de machine learning basé sur des données historiques et des caractéristiques des films pour prédire le nombre d'entrées. Ce modèle sera alimenté par des données de scraping sur les films à venir.
5. **API de machine learning** : Exposer le modèle de prédiction sous forme d'API pour intégrer facilement les prédictions dans l'application web.
6. **Microservices et architecture cloud** : Utiliser des microservices pour déployer les composants du projet, avec Docker pour la production et Azure pour l'hébergement des services.

On présente un diagramme de flux de données (voir fig 3.)

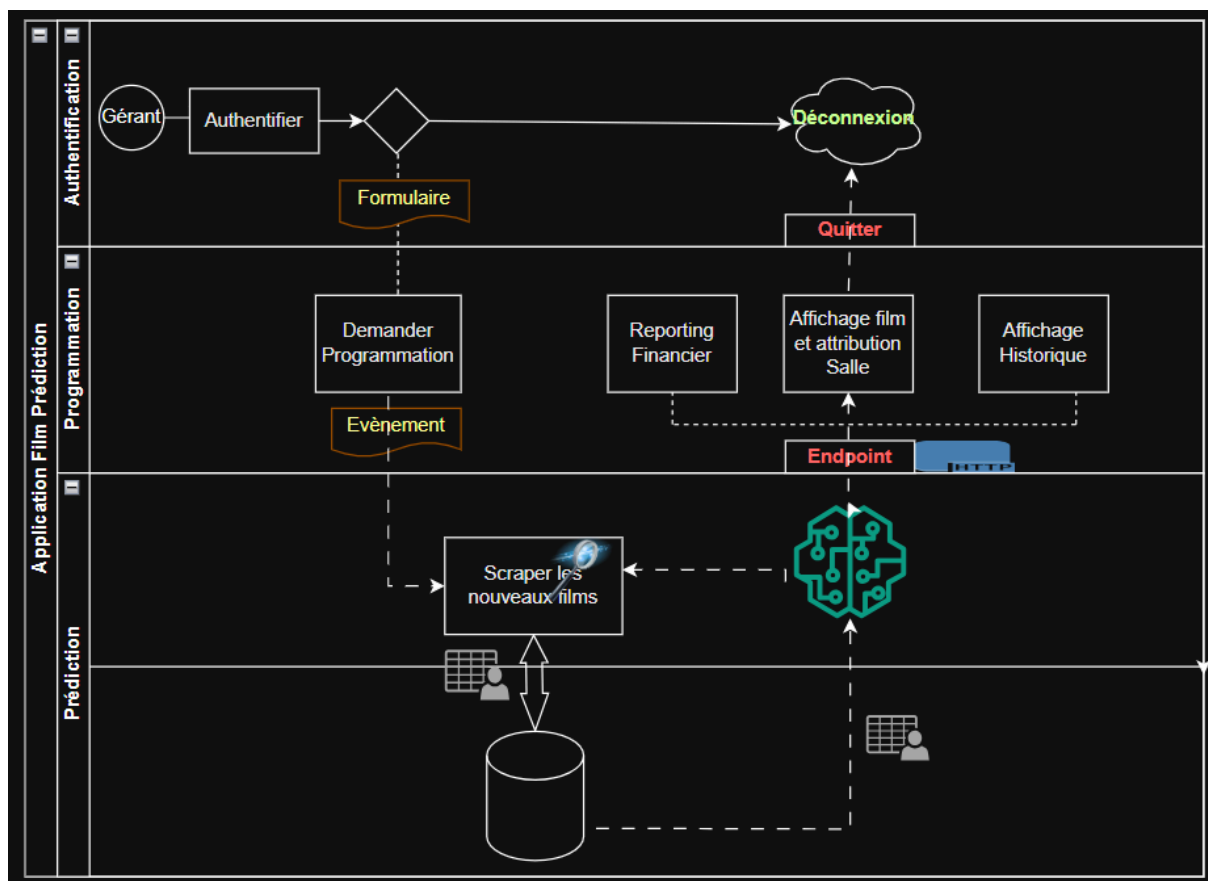


Fig 3. Diagramme de flux de données

## Sprint Planning

Sprint	Période	Livrables
Sprint 1	31/03 → 09/04	Cahier des charges fonctionnel
Sprint 2	10/04 → 16/04	Première version de l'outil
Sprint 3	17/04 → 23/04	Modèle de machine learning
Sprint 4	24/04 → 30/04	Présentation finale de l'outil

Les principales contraintes sont :

- **Qualité des données** : L'outil dépendra fortement de la qualité des données collectées via le scraping. Il faut s'assurer que les informations sont fiables et ne contiennent pas de biais.
- **Délais** : Le projet doit être livré dans un mois, avec une organisation en 4 sprints. Cela impose des délais serrés pour chaque phase du projet.
- **Accessibilité pour le client** : L'outil doit être simple à utiliser, sans que le client ait besoin de connaissances techniques en IA.
- **Précision des prédictions** : Le modèle de machine learning doit être suffisamment précis pour générer des prédictions utiles pour le client.
- **Coût des ressources Azure** : Le client devra être informé du coût des ressources utilisées sur Azure pour l'hébergement du projet:
  - Les modèles de prédiction sauvegardés coûtent de **l'argent**, il faut garder que les trois derniers modèles.

S'il existe des **dépendances** et des **contraintes**, il faut revoir les besoins en tenant compte des contraintes. **Dans ce cas, il faut prévoir un plan de gestion des intervenants (PGI)** : dans notre cas, ce sera plus le changement de rôle au sein de l'équipe interne (PO, SM, 2 DEV). Pour cela, il faut:

- prévoir un pair programming dans les parties non maîtrisées
- capitaliser:
  - documenter les développements
  - assurer la montée en compétence par exemple à travers des tutos [\[https://zipppy-twig-11a.notion.site/Azure-Machine-learning-studio-1c91f9041c968025a134d722ba5be3b\]](https://zipppy-twig-11a.notion.site/Azure-Machine-learning-studio-1c91f9041c968025a134d722ba5be3b)
  - faire des réunion pour les parties en avance
- une organisation de sprint est planifiée, voir tableau ci-dessous:

	PO	SM	Dev	ML
Sprint 1	W.B	M.A	L.G / M.A / N.C	N.C
Sprint 2				
Sprint 3				
Sprint 4				

**Tab 1.** Organisation Sprint

**Le PGI est un document distinct qui n'est pas nécessairement lié au cahier des charges fonctionnel, mais qui est essentiel au début de tout projet.**

## À QUI DEVONS-NOUS NOUS ADRESSER ?

Dans ce projet, voici la liste des différents publics classés par ordre de priorité en fonction de leurs besoins en termes de contacts.

**Client** : Le gérant du cinéma "New is always better" sera la personne principale à consulter pour les retours sur l'application, les besoins et les attentes.

**Product Owner** : Responsable de l'alignement des exigences business avec le produit.

**Scrum Master** : Responsable de la gestion des sprints, de la méthodologie Agile et du bon déroulement du projet.

**Équipe de développement** : Les 2 développeurs IA, chacun ayant une spécialité (scraping, machine learning, backend Django, etc.).

## COMMENT SAURONS-NOUS SI LE PROJET EST RÉUSSI ?

Suite à l'envoi du livrable du troisième sprint contenant une première version de l'outil, le gérant peut commencer les tests des différentes fonctionnalités. Le projet sera considéré comme réussi si :

- L'application est fonctionnelle et permet de sélectionner les films à projeter de manière optimisée.
- Les prédictions d'audience sont fiables et permettent de maximiser les recettes du cinéma.



- L'outil est simple à utiliser et répond aux besoins du gérant sans nécessiter de compétences techniques.
- Les performances du modèle de machine learning sont mesurées (via des métriques comme l'erreur quadratique moyenne) et satisfont les attentes.
- Le projet est livré dans les délais impartis (1 mois) avec des résultats concrets lors des présentations aux sprints.

## QUELS SONT LES ASPECTS PRATIQUES ?

Les aspects pratiques incluent :

- Utilisation d'une infrastructure cloud sur Azure avec Docker pour la production.
- Automatisation du scraping via des cronjobs ou un pipeline Airflow.
- Utilisation de Jira pour le suivi du projet et de GitHub pour le versionnement.
- Préparation des prévisions des films sur une base hebdomadaire, avec des ajustements possibles en fonction des données collectées en temps réel.

## APPROBATIONS

Dans notre cas, la validation est nécessaire et opérée au niveau:

- du cahier des charges par le gérant du cinéma.
- des livrables par le gérant du cinéma et le product Owner.

**Approbation du cahier des charges fonctionnel** lors du Sprint 1 pour valider les exigences du client.

**Approbation de la première version de l'outil** lors du Sprint 2 pour vérifier les différents composants de l'application.

**Approbation du modèle de machine learning** lors du Sprint 3 pour confirmer que les prédictions sont suffisamment précises pour répondre aux besoins du client et sont bien intégrées dans l'application.

**Approbation finale** lors du Sprint 4 pour valider le produit final et sa mise en production.

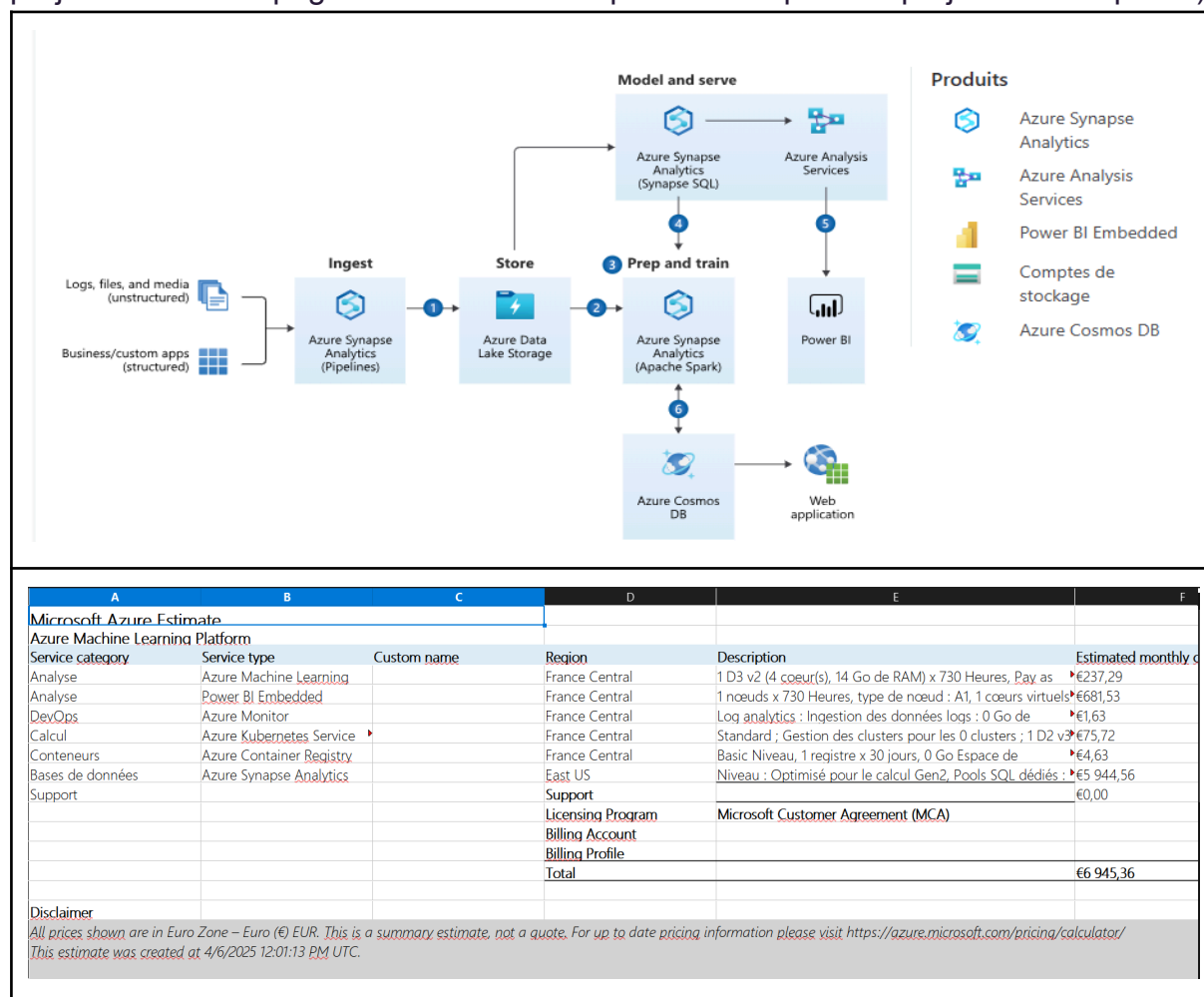
## Demandes spécifiques

### Sprint 2:

- Carte géographique des cinémas : utilisation pour affiner les prédictions en fonction de la localisation.
- Un modèle performant doit avoir une performance minimale de quoi?
- Utiliser ML Studio, en utilisant Calculator Pricing pour Europe, on aura 6945 euros/mois contre 6866 euros/mois en USA, (voir Tab 2.)

### Sprint 3:

- Mécanisme de feedback utilisateur : affiner les recommandations basées sur les interactions.
- Intégration avec d'autres activités de développement ou de marketing (par exemple, le client planifie-t-il une campagne de marketing pour le déploiement de ce projet et cette campagne a-t-elle une date précise à laquelle le projet doit être prêt ?)



Azure Machine Learning Platform

<div> <div>▼</div> <div>Azure Machine Learning</div> </div>	<div> <div>①</div> <div>1 D3 v2 (4 coeur(s), 14 Go de RAM) x 730 Heures, P...</div> </div>	<div> <div>🔗</div> <div>📄</div> </div>	<div>Frais initiaux : 0,00 €</div> <div>Frais mensuels : 154,8...</div>
<div> <div>▼</div> <div>Power BI Embedded</div> </div>	<div> <div>①</div> <div>1 nœuds x 730 Heures, type de nœud : A1, 1 cœurs ...</div> </div>	<div> <div>🔗</div> <div>📄</div> </div>	<div>Frais initiaux : 0,00 €</div> <div>Frais mensuels : 681,5...</div>
<div> <div>▼</div> <div>Azure Monitor</div> </div>	<div> <div>①</div> <div>Log analytics : Ingestion des données logs : 0 Go de...</div> </div>	<div> <div>🔗</div> <div>📄</div> </div>	<div>Frais initiaux : 0,00 €</div> <div>Frais mensuels : 1,63 €</div>
<div> <div>▼</div> <div>Service Azure Kubernetes (AKS)</div> </div>	<div> <div>①</div> <div>Standard ; Gestion des clusters pour les 0 clusters ; ...</div> </div>	<div> <div>🔗</div> <div>📄</div> </div>	<div>Frais initiaux : 0,00 €</div> <div>Frais mensuels : 79,10 €</div>
<div> <div>▼</div> <div>Azure Container Registry</div> </div>	<div> <div>①</div> <div>Basic Niveau, 1 registre x 30 jours, 0 Go Espace de s...</div> </div>	<div> <div>🔗</div> <div>📄</div> </div>	<div>Frais initiaux : 0,00 €</div> <div>Frais mensuels : 4,63 €</div>
<div> <div>▼</div> <div>Azure Synapse Analytics</div> </div>	<div> <div>①</div> <div>Niveau : Optimisé pour le calcul Gen2, Pools SQL dé...</div> </div>	<div> <div>🔗</div> <div>📄</div> </div>	<div>Frais initiaux : 0,00 €</div> <div>Frais mensuels : 5 944...</div>

Assistance

ASSISTANCE:

De base (inclus) ▼

①

0,00 €

Sélectionner votre programme/offre

PROGRAMME DE LICENCES:

Contrat client Microsoft (CCM) ▼

①

[Connectez-vous](#) pour voir la tarification de votre contrat Azure.

•

Afficher La Tarification Développement/Test

①

Frais initiaux estimés

0,00 €

Coût mensuel estimé

6 866,26 €

**Tab 2.** Exemple de Calculator Pricing