

# Design document

## Time Series Prediction

### Objectif

La prévision de série chronologique une approche pour objectif d'utiliser des méthodes statistiques pour prévoir le comportement futur en fonction de données historiques.

En effet elle permet d'analyser, de décrire et d'expliquer un phénomène au cours du temps et d'en tirer des conséquences pour des prises de décision (marketing...)

### Description

La théorie des séries chronologiques (ou temporelles) est appliquée de nos jours dans des domaines aussi variés que l'économetrie, la médecine ou la démographie, pour n'en citer qu'une petite partie. On s'intéresse à l'évolution au cours du temps d'un phénomène, dans le but de décrire, expliquer puis prévoir ce phénomène dans le futur. On dispose ainsi d'observations à des dates différentes, c'est à dire d'une suite de valeurs numériques indicées par le temps.

Time Series Prediction, en particulier, « Extrapolation-TS »- Extrapolation of Time Series - vise à déduire ce qui est arrivé à une série de points de données dans le passé et à tenter de prédire ce qu'il adviendra de l'avenir.

Dans ce contexte, nous allons adopter une approche statistique quantitative des séries chronologiques du « Production mensuelle de lait: pounds par vache dans une période du Jan 62 jusqu'à Déc 75" »

En effet, « Extrapolation-TS » permet d'analyser, de décrire et d'expliquer un phénomène au cours du temps et d'en tirer des conséquences pour des prises de décision (marketing...).

## Les mérites d'Extrapolation-TS

- ❖ Faire un contrôle, par exemple pour la gestion des stocks, Un autre problème intéressant est la détection de ruptures résultantes,
  - ❖ Obtenir la meilleure description d'une série ou de pouvoir faire de la prédiction d'évènements.
  - ❖ Déterminer des tendances au sein des séries ainsi que la stabilité des valeurs (et de leur variation) au cours de temps.
  - ❖ Déterminer les composantes de la série en termes de modèles saisonniers, les tendances, par rapport à des facteurs externes, et autres.
- ⇒ Les prévisions constituent une aide importante pour une planification efficace dans le futur.

## Caractéristiques

La prévision de série chronologique devrait utiliser des techniques d'apprentissage automatique et des modèles de prédiction pour traiter des données historiques afin de prévoir les valeurs futures.

Pour réduire les incertitudes à l'avenir, l'utilisation de méthodes multivariées peut intégrer des informations pertinentes à la prévision de la série chronologique. Ces informations peuvent être incorporées dans des méthodes statistiques (régression linéaire multiple, régression dynamique) ou dans des méthodes de calcul informatisé (telles que les réseaux de neurones).

En outre, les prévisions de séries chronologiques ont une importance commerciale énorme car ils impliquent la compréhension de divers aspects de la nature inhérente de la série afin que vous soyez mieux informé pour créer des prévisions significatives et précises.

## Features

| Feature                 | Description  |
|-------------------------|--|
| Évaluation de besoin    | la validation de la pertinence de l'interrogation et son adaptabilité au modèle prédictif. |
| Importation des données | Sur la base d'interrogation, nous  |

|   |   |
|---|---|
|   | sélectionnons les données nécessaires pour y apporter une réponse adaptée.  |
| Correction des données                          | Avant de se lancer dans l'étude d'une série chronologique, il est souvent nécessaire de traiter, modifier les données brutes.   |
| Entraînement du modèle prédictif                | On va retourner le modèle afin de l'habituer à recevoir une requête / question  |
| Analyse de la série à partir de ses composantes | Une fois l'étape de modélisation effectuée, on étudie les composantes du modèle les unes après les autres.  |
| Diagnostic du modèle                            | Une fois le modèle construit et ses paramètres estimés, on vérifie que le modèle proposé est bon c'est-à-dire l'ajustement au modèle  |
| Envoyer une requête de prédiction               | Le modèle de prévision renvoie une réponse qui est : soit une valeur numérique estimée, soit une catégorisation de votre objet de requête, en fonction de vos données d'entraînement. |

## High level design of the system

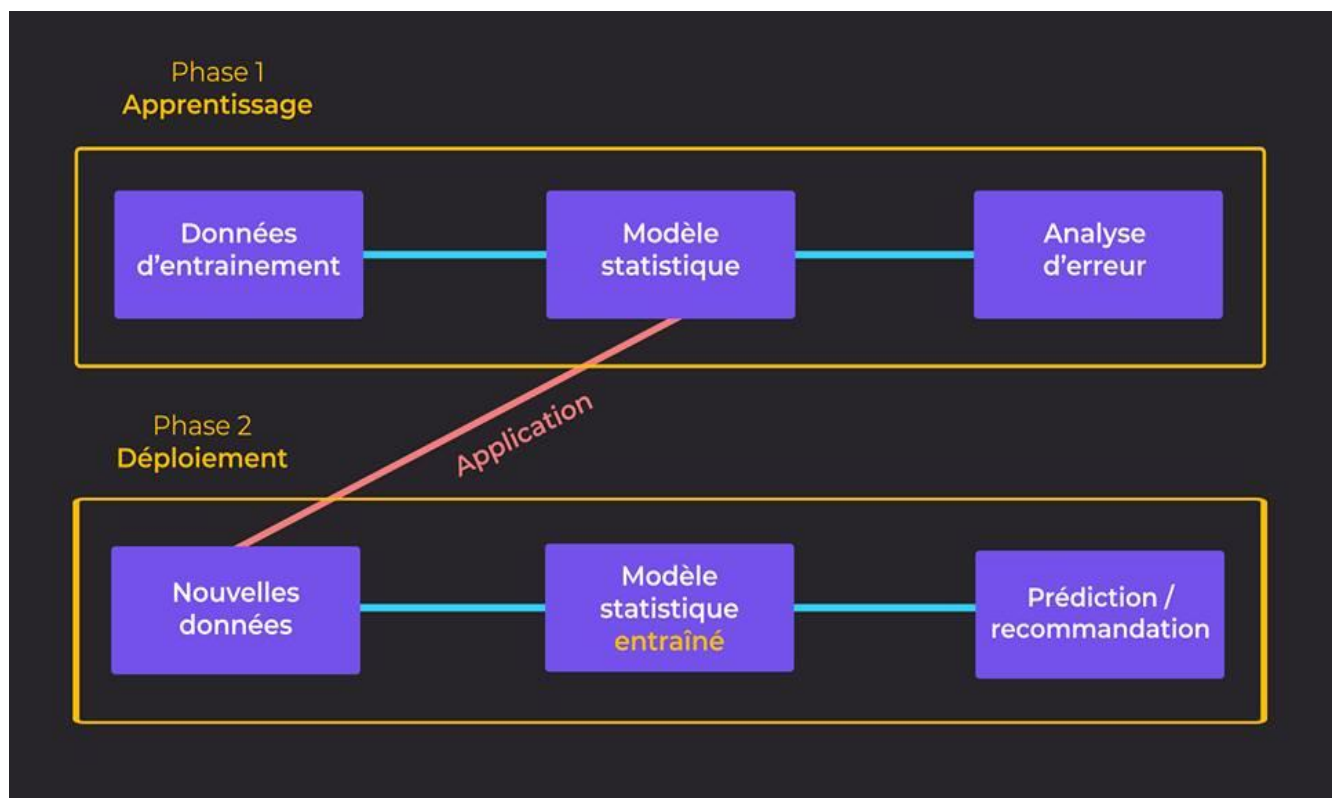


Figure 1 : Workflow of Time Series Prediction

Un processus de prévision de série chronologique est un ensemble d'étapes ou une recette qui vous amène de la définition du problème au résultat de la création d'un modèle de prévision de série chronologique ou d'un ensemble de prévisions.

| Fonctionnalités                           | Descriptions  | Outils  |
|---|---|---|
| Importation des données et Pré-traitement | Récupérer les données et les stocker dans un DataFrame afin de les transformer.   | La bibliothèque Pandas et SQLAlchemy de python  |
| Visualiser les données                    | Une fois les données nettoyées, stockées dans un DataFrame et prêtes à être analysées, vous pouvez maintenant appliquer des outils de machine learning (apprentissage automatique) et représenter les résultats par des visualisations. | Les deux bibliothèques de python <a href="#"><u>scikit-learn</u></a> et <a href="#"><u>matplotlib</u></a> . |
| Entrainement de modèle prédictif          | Prédire les séries chronologiques pour garantir une meilleure performance à notre modèle  | XGBoost<br>RandomForest   |
| Vérifier le modèle prédictif              | On vérifie que le modèle est le meilleur à travers son accuracy   |   |
| Tester le modèle prédictif                | Le modèle choisi renvoie une réponse en fonction des données d'entraînement.  |   |
| Visualiser les prédictions                | Pour nous aider à comprendre l'exactitude de nos prévisions, on établit les prévisions pour une période de temps donnée avec des courbes explicatifs.   | Dash<br>Kibana  |