



Prévision de séries temporelles

YOUR 5-DAY FORECAST	38°F 	41°F 	36°F 	40°F 	44°F
YOUR 5-MONTH FORECAST	38°F 	29°F 	21°F 	24°F 	35°F
YOUR 5-YEAR FORECAST	38°F 	25°F 	36°F 	37°F 	41°F
YOUR 5-MILLION-YEAR FORECAST	38°F 	52°F 	40°F 	275°F 	40°F
YOUR 5-BILLION-YEAR FORECAST	38°F 	105°F 	371°F 	71,488,106°F 	-452°F
YOUR 5-TRILLION-YEAR FORECAST	38°F 	-452°F 	-452°F 	-452°F 	-453°F

Your 5 day forecast

Source: [xkcd](#)

Objectifs du module

Une introduction de base à la prévision des séries temporelles

- Comprendre en quoi les étapes de prévision des séries temporelles sont différentes des autres prédictions d'apprentissage automatique
- Utiliser des modèles classiques de séries chronologiques linéaires pour faire des prévisions univariées
- Utiliser une librairie de prévision de séries chronologiques pour des prévisions univariées

Modalités

- Durée du projet : 3 jours
- Travailler en équipes de deux
- Produire vos propres scripts et mémos individuels pour terminer le projet

Contexte

Dans le monde numérique actuel, nous collectons des quantités incroyables de données. Tout ce que nous créons dans le monde numérique est horodaté: de nos interactions, clics, achats en ligne, à nos moniteurs de santé comme les «fitbits», à notre position GPS à partir de nos téléphones, à notre «selfie» le plus récent... Plus les capteurs deviennent bon marché et plus nous les intégrons dans nos vies. Nous collectons ainsi de nombreux points de données pour construire des séries chronologiques de notre vie et de notre société. Si certaines données que vous avez actuellement ne sont PAS une série chronologique, elles le seront à l'avenir.

Par extension, les tentatives de prédictions sur l'avenir se multiplient, aussi qualifiées de prévisions. Ce podcast [«Les prévisions sont toujours fausses \(mais nous en avons besoin quand même\)»](#) explique à quoi ressemblent les prévisions modernes. Dans la prévision, la terminologie est particulièrement fournie et un temps certain peut être requis afin de comprendre l'ensemble des méthodes utilisées, leurs spécificités utiles pour différents types de données. Le but de ce module est de vous faire utiliser python pour faire des prévisions robustes et de commencer à appréhender la terminologie utilisée dans la prévision des séries temporelles.

Etape 1 (2 jours)

Modélisation avec ARIMA

Objectifs de l'activité

- Comprendre en quoi les étapes de prévision des séries temporelles sont différentes des autres prédictions d'apprentissage automatique
- Utiliser des modèles classiques de séries temporelles linéaires pour faire des prévisions univariées
- Comprendre le fonctionnement du modèle ARIMA

Compétences

- Expliquer une méthodologie typique utilisée lors de la réalisation de prévisions de séries temporelles
- Utiliser des modèles classiques de séries temporelles statistiques pour faire des prévisions

Consignes

- Effectuer les tâches définies dans le jupyter notebook 01ARIMA_modelling.ipynb. Ce NB vous guidera à travers le processus d'utilisation des modèles ARIMA sur les données de séries chronologiques pour faire des prévisions.
- Il existe de nombreuses descriptions des modèles ARIMA et de leur fonctionnement en ligne (et dans le livre suggéré ci-dessous).

Ressources

- Former les fractionnements des tests et la validation croisée dans les séries chronologiques :
<https://stats.stackexchange.com/questions/14099/using-k-fold-cross-validation-for-time-series-model-selection>
<https://robjhyndman.com/hyndsight/tscv/>
- Forecasting: Principles and Practice livre (Rob Hyndman) :
<https://otexts.com/fpp2/>
<https://otexts.com/fpp2/non-seasonal-arima.html>
<https://otexts.com/fpp2/transformations.html>
- Documentation de Pmdarima :
<http://alkaline-ml.com/pmdarima/>

Livrables

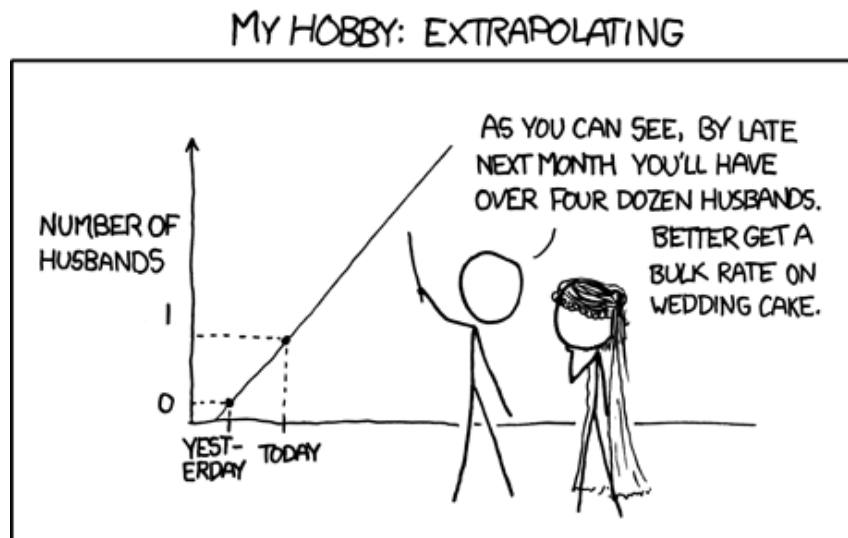
- Un mémoire qui explique (dans vos propres mots) ce qui suit:
 - Comment le fractionnement des trains / tests et la validation croisée des données de séries chronologiques doivent être effectués
 - Une brève description des modèles de séries temporelles ARIMA
- Script / cahier Python (ou Memo) qui contient:
 - Exercice 1 terminé dans 01ARIMA_modelling.ipynb

Pour aller plus loin

- Essayer de créer un modèle ARIMA sans utiliser la fonction `auto_arima`. Voir section 'ACF and PACF plots' here :
<https://otexts.com/fpp2/non-seasonal-arima.html>
- Trouver un ensemble de données de séries chronologiques qui vous intéresse et essayer d'en construire une prévision

Etape 2 (1 jour)

Utiliser un Prophet pour faire des prophéties ...*hum*...prévisions



L'extrapolation est une forme de prévision ...

Source : [XKCD](#)

Objectifs de l'activité

- Utiliser un logiciel de prévision de séries chronologiques automatisé pour faire des prévisions univariées
- Faire des évaluations robustes de l'exactitude des prévisions du modèle
- Comprendre quels sont les avantages et les inconvénients de l'utilisation d'une librairie de prévisions automatisées comme Prophet
- Réaliser un projet de prévision sur un ensemble de données réelles, en prenant des données brutes, en définissant votre objectif, en évaluant votre modèle et en prenant des mesures pour l'améliorer

Compétences

- Savoir utiliser Facebook Prophet pour faire des prévisions

Consignes

- Parcourir le cahier jupyter 02Prophet_forecasting pour comprendre le fonctionnement du prophète
- Lire les ressources ci-dessous pour comprendre ce qu'est un prophète
- Faire l'exercice 1 du NB

Ressources

- Documentation du prophète
<https://facebook.github.io/prophet/>
- Document académique décrivant le prophète
<https://peerj.com/preprints/3190/>
- Pourquoi le prophète existe - "Time series are not sexy"
<https://www.youtube.com/watch?v=pOYAXv15r3A>
- Blog sur le réglage des hyperparamètres dans Prophet
<https://towardsdatascience.com/implementing-facebook-prophet-efficiently-c241305405a3>
- Prophète appliqué à la prévision des données énergétiques
<https://www.kaggle.com/robikscube/time-series-forecasting-with-prophet>

Livrables

- Script / NB Python (ou Memo) qui contient:
 - Exercice 1 terminé dans 02Prophet_forecasting.ipynb

Pour aller plus loin

- Résumer les différents types de modèles de séries chronologiques et leurs différences
https://www.youtube.com/watch?v=d4Sn6ny_5LI
- Essayer d'ajouter des variables exogènes (telles que les prévisions météorologiques) à l'un de vos modèles. Améliorent-ils vos prévisions de modèle? Ceci est possible dans les modèles ARIMA.
https://alkaline-ml.com/pmdarima/modules/generated/pmdarima.arima.auto_arima.html
- Essayer de construire un LSTM et comparer ses prédictions à vos autres modèles
<https://machinelearningmastery.com/time-series-forecasting-long-short-term-memory-network-python/>
<https://machinelearningmastery.com/multi-step-time-series-forecasting-long-short-term-memory-networks-python/>
<https://medium.com/@cdabakoglu/time-series-forecasting-arima-lstm-prophet-with-python-e73a750a9887>

Step 3 (Optionnel)

Ecrire votre propre librairie python de forecasting

Objectifs de l'activité

- Créer une librairie avec une classe (simple mais puissante) permettant de faire du forecasting.

Consignes

Voir: [03_ts_forecast.md](https://github.com/03-ts-forecast).

Livrables

Une librairie fonctionnelle avec la classe décrite.