

Super titre de TIPE

sous-titre

Alexandrine, Pedro
De Carvalho, Enzo

2020-2021

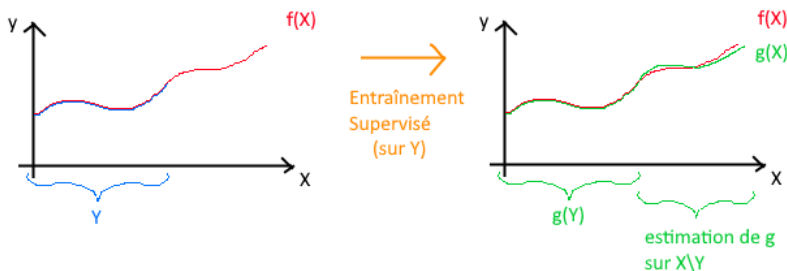
Sommaire

- 1 Première approche : simple regression
 - Principe
 - Optimisation d'hyperparamètres et stratégies
 - Cross-validation
 - Résultats avec SVR
 - Découpages adaptés

- 2 Approche multivariées
 - Multiregresseur : 'RegressorChain'
 - Réseau neuronal

Principe de l'apprentissage supervisé

Donné une fonction $f : X \rightarrow y$ (X et y des données corrélées), un algorithme d'apprentissage supervisé cherche alors à obtenir un modèle $g : X \rightarrow y$ qui estime f à partir d'une partie $Y \subset f(X)$ déjà connue de l'algorithme.



Principe de l'apprentissage supervisé

La création du modèle passe par l'optimisation d'une fonction d'objectif, où les paramètres à optimiser sont ceux du modèle. Exemple avec une regression linéaire dont le modèle est de la forme :

$$\hat{y}(w, x) = w_0 + w_1x_1 + \dots + w_px_p$$

Figure: x les données, w les paramètres

N.B : refaire la figure (celle-ci est volée du manuel scikit)

ElasticNet

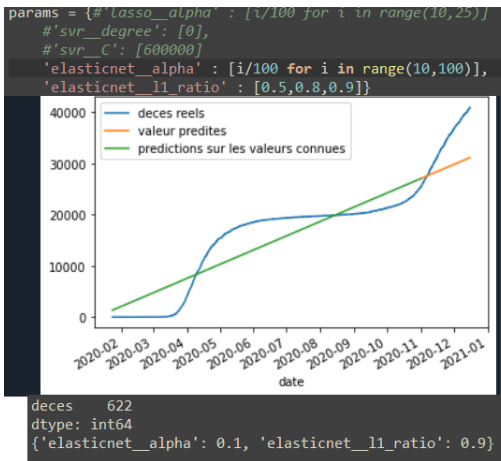
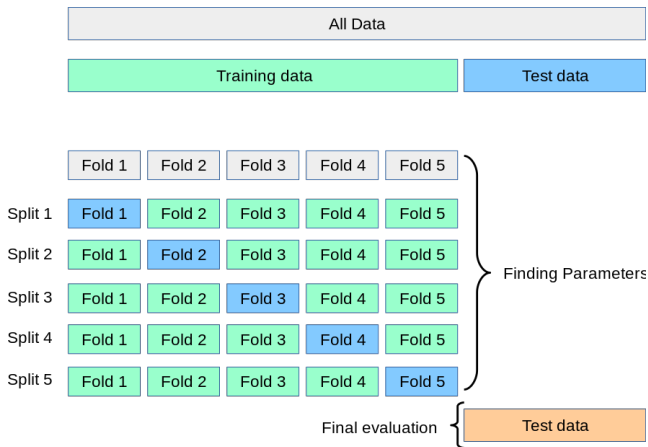


Figure: résultats peu intéressants dans notre cas..

Recherche du meilleur paramètre

Principe de la "cross-validation" :

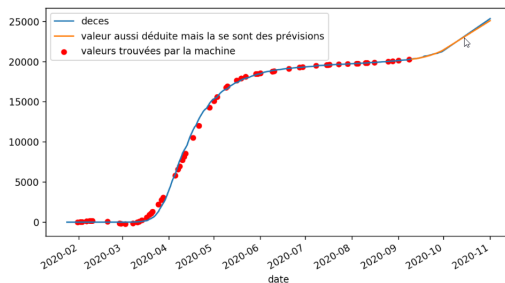


SVR ; premier résultat

Approche à l'aide du modèle SVR.

Noyau 'rbf' → ajustement du paramètre C

— Modèle SVR (prédit les données entre le 02/12/20 et 16/12/20)



```
{'svr_C': 500000, 'svr_cache_size': 200, 'svr_degree': 0}  
0.9980841978694834
```

Figure: Premier résultat avec SVR et découpage inadapté

Découpage adapté pour la CV

Remise en question de la méthode de découpage

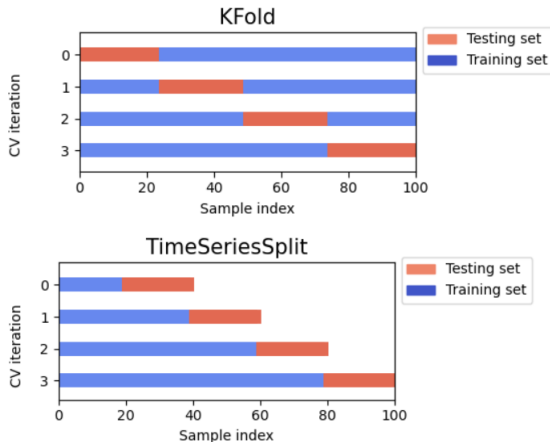


Figure: Comparaison des découpages pour la CV

SVR

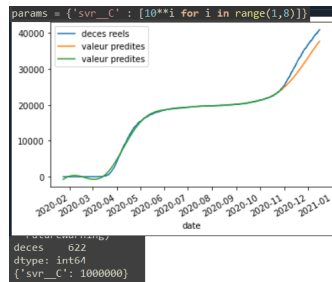
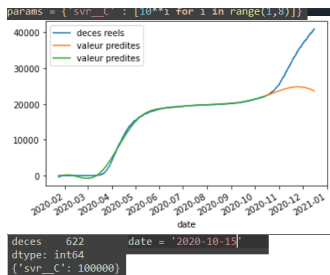


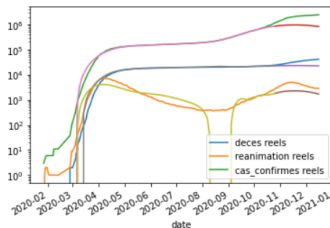
Figure: à gauche la prediction avant le pt d'inflexion, à droite, après.

⇒ Le modèle n'arrive pas à "suivre" sans le point d'inflexion.

RegressorChain SVR

Multiregresseur RegressorChain

Corrélation: Cas confirmé → hospitalisé → décès



Multiregresseur

```
In [52]: runfile("C:/Users/Anzo/Documents/TIPE/repo/covid_deces_multiregresseur.py", wdir='C:/Users/Anzo/Documents/TIPE/repo')
C:\Users\Anzo\Documents\TIPE\repo\covid_deces_multiregresseur.py:1: DtypeWarning: Columns (17,18)
have mixed types.Specify dtype option on import or set low_memory=False.
### Multi regresseur avec le modele SVR ###
{'regressorchain_base_estimator__C': 100000, 'regressorchain_base_estimator__epsilon': 0.001}
```

Figures now render in the Plots pane by default. To make them also appear inline in the Console, uncheck "Mute Inline Plotting" under the Plots pane options menu.

Figure: résultat avec SVR (insatisfaisant)

Explications

Approche avec des réseau neuronaux

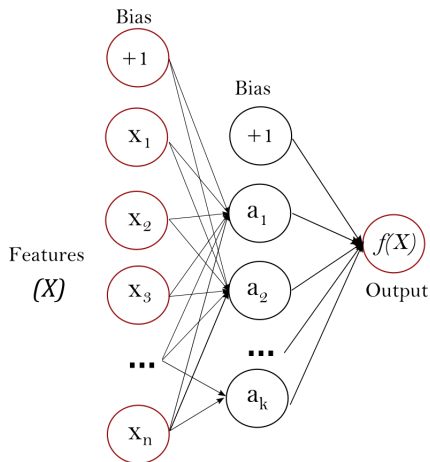
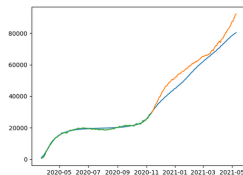
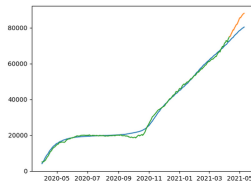


Figure: échec du modèle sans la courbe des cas confirmés

Résultats



Neural network avec 7 jour de décalage; max_iter=100k; prédit à partir du 11 novembre

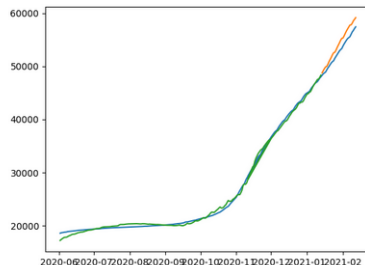
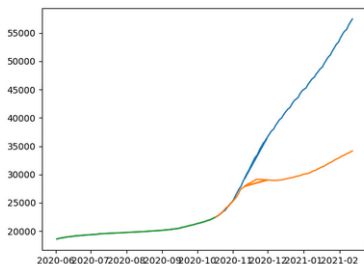


{'mlpregressor__max_iter': 90000, 'mlpregressor__n_iter_no_change': 3, 'mlpregressor__tol': 0.0001}

NN: décès prédits 30 jours avec les données de 15 jours avant

Figure: échec du modèle sans la courbe des cas confirmés

Résultats



Neural network avec 7 jour de décalage; max_iter=90k

Figure: échec du modèle sans la courbe des cas confirmés

Notes

N.B les diapos sont trop verbeux; il faut encore un effort de concision, ou alors elles doivent être expliquées à l'oral
N.B sourcer les images