

Étude sur le propagation du Covid 19 avec machine learning

Pedro ALEXANDRINE

numero d'inscription : 41758

2020-2021

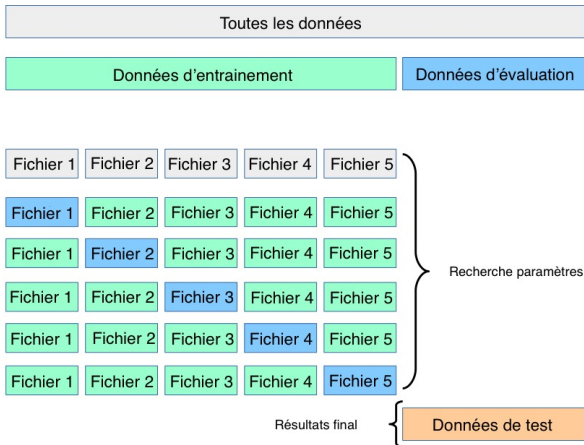
Sommaire

- 1 Première approche : simple regression
 - Validation croisée et hyperparamètres
 - Résultats avec SVR

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

Recherche du meilleur paramètre

Principe de la validation croisée:



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)

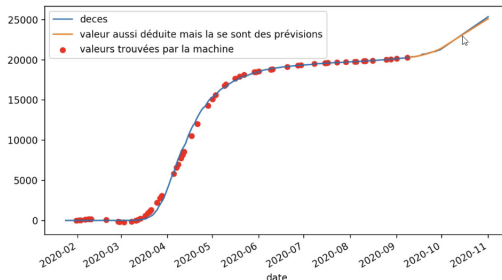


Figure: Premier résultat avec SVR et découpage inadapté, $C = 50k$

Découpage adapté pour la validation croisée

Remise en question de la méthode de découpage

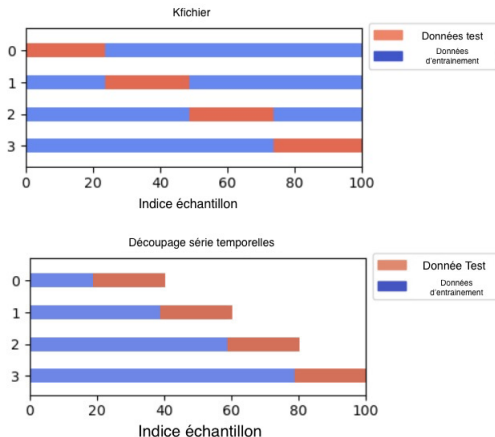


Figure: Comparaison des découpages pour la validation croisée

SVR

Avec $C = 10^5$

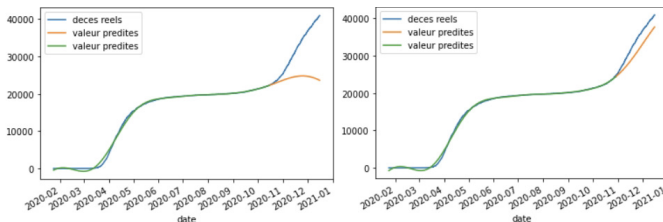


Figure: À gauche: prediction avant point d'inflexion; à droite: après.

⇒ Prédiction inefficace du point d'inflexion.

RegressorChain SVR

Multiregresseur RegressorChain

Corrélation : Cas confirmé \rightarrow réanimation \rightarrow décès

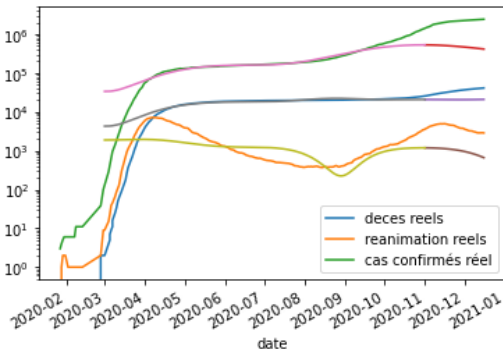


Figure: Résultat avec SVR insatisfaisant

RegressorChain TheilSenRegressor

Changement de régresseur : meilleurs résultats

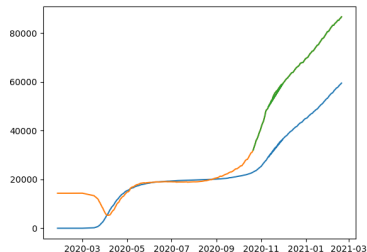
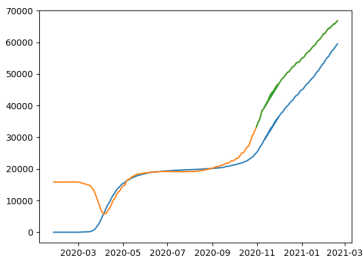
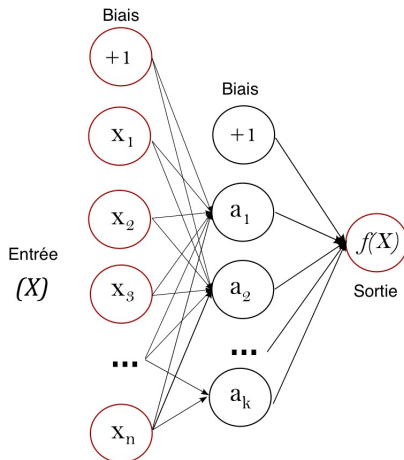


Figure: Prédiction avec et sans point d'inflexion

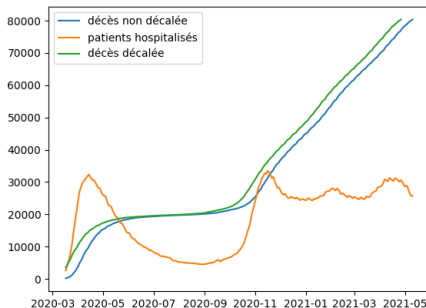
Principe du réseau neuronal



Décalage de courbes

Représentation avant et après décalage

date		date	
2020-03-17	175	2020-03-17	1100
2020-03-18	264	2020-03-18	1331
2020-03-19	372	2020-03-19	1696
2020-03-20	450	2020-03-20	1995
2020-03-21	562	2020-03-21	2314



Corrélation et Premier résultat

Corrélation cas confirmés et décès : 0.978

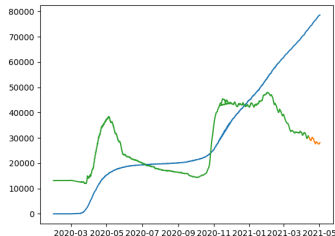
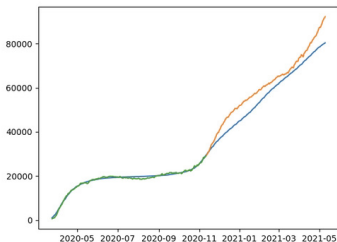
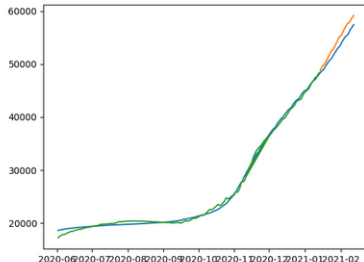
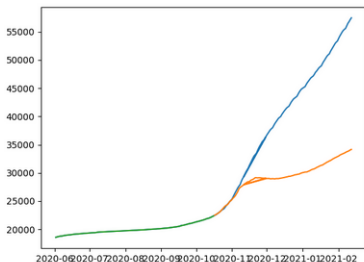


Figure: Tests avec et sans cas confirmés

Échec du modèle sans les cas confirmés → cohérent avec corrélation

Résultats

Réseaux neuronaux : Meilleurs paramètres



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

Augmentation du décalage

Résultats insatisfaisant

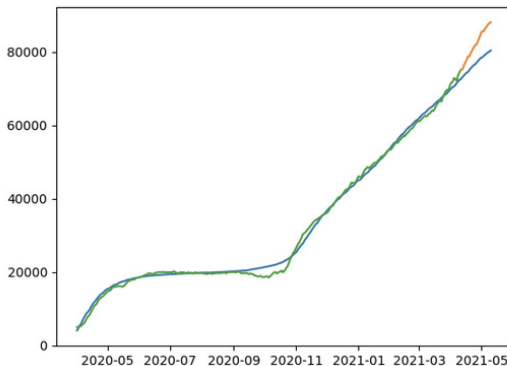


Figure: Tests avec 30 jour de décalage

Conclusion

Modèle peu fiable hors situation stabilisée ou début d'évolution