



ÉCOLE CENTRALE NANTES

INFO IA - PIIA MARÉES
RAPPORT

Rapport 5 - PIIA1 Marées

Élèves :

Daniel MACEDO GALEMBECK
Nicolas CONTRERAS
Alexandre THOMASSIN
Liushuangfei XIE

Enseignant :
Mr. ROUX

17 février 2023

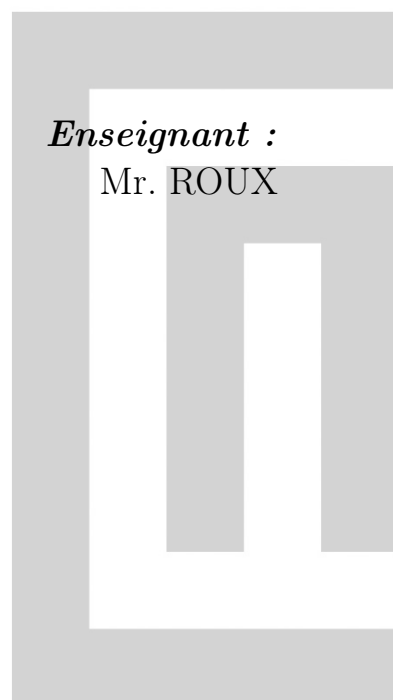


Table des matières

1	Rapport des deux semaines passées	2
1.1	Utilisation de PyTorch et Tensorflow sur GPU :	2
1.2	Comparaison PyTorch / Tensorflow :	2
1.3	Modèles testés	2
1.3.1	Prophet	2
1.3.2	LSTM - PyTorch	3
2	Pour les deux semaines à venir	5
2.1	Création d'un script de prédictions	5
2.2	Tester d'autres modèles	5

1 Rapport des deux semaines passées

1.1 Utilisation de PyTorch et Tensorflow sur GPU :

Nous avons eu l'occasion de pouvoir tester les entraînements des modèles précédents sur GPU. Les entraînements sont bien plus rapides que lors de l'utilisation du CPU (MSE) d'environ 0.026.

On a donc pu entraîner les modèles sur des jeux de données un peu plus grands (1.4M en prenant toutes les données de sources 3), cela nous permet donc d'avoir des résultats un peu plus précis. On peut quasiment atteindre une perte de 0.003 et une Erreur moyenne carrée de 0.026.

1.2 Comparaison PyTorch / Tensorflow :

Après ces quelques semaines d'apprentissage des deux bibliothèques et de différents modèles, on peut dégager quelques différences.

En effet, PyTorch Forecasting à l'avantage de pouvoir prendre en compte la date sans trop changer la précision de l'entraînement, cela augmente cependant le nombre de paramètres donc rend les entraînements significativement plus longs.

Cependant Tensorflow ou PyTorch classique permettent de plus facilement personnaliser les modèles utilisés en changeant les différentes couches cachées et le nombre de neurones. Cela permet de comprendre la façon dont est fait un modèle, mais cela reste un peu plus compliqué car il faut gérer la taille et le shape de chaque couche de neurones.

1.3 Modèles testés

1.3.1 Prophet

Prophet est une méthode de prédiction du type time series forecasting basée sur un modèle additif dans lequel les tendances non linéaires sont ajustées à la saisonnalité annuelle, hebdomadaire et/ou quotidienne, ainsi qu'aux effets des week-ends. Cette procédure fonctionne mieux avec les séries temporelles qui ont de forts effets saisonniers et plusieurs saisons de données historiques.

Nous avons modélisé avec la bibliothèque de prophètes en utilisant les données de 2020 et 2021. Après avoir entraîné les données, nous avons fait une prédiction pour les prochaines 24 heures, que l'on peut voir dans l'image ci-dessous 1.

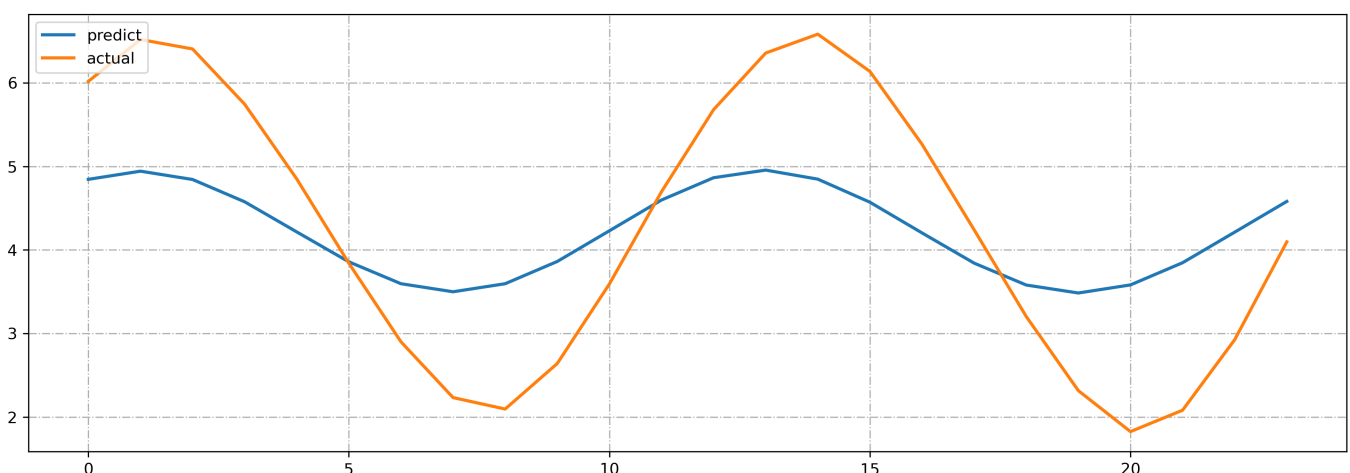


FIGURE 1 – Les prédictions pour les prochaines 24 heures.

Cette prédiction a une précision de 75,2%, elle peut donc être considérablement améliorée.

1.3.2 LSTM - PyTorch

Enfin, nous avons créé un autre modèle LSTM, mais cette fois en utilisant la bibliothèque PyTorch. Dans l'image 2, nous pouvons voir la distribution des données réelles pour les années 2020 et 2021. Nous n'avons pas ajouté les années 2022 en raison de données manquantes.

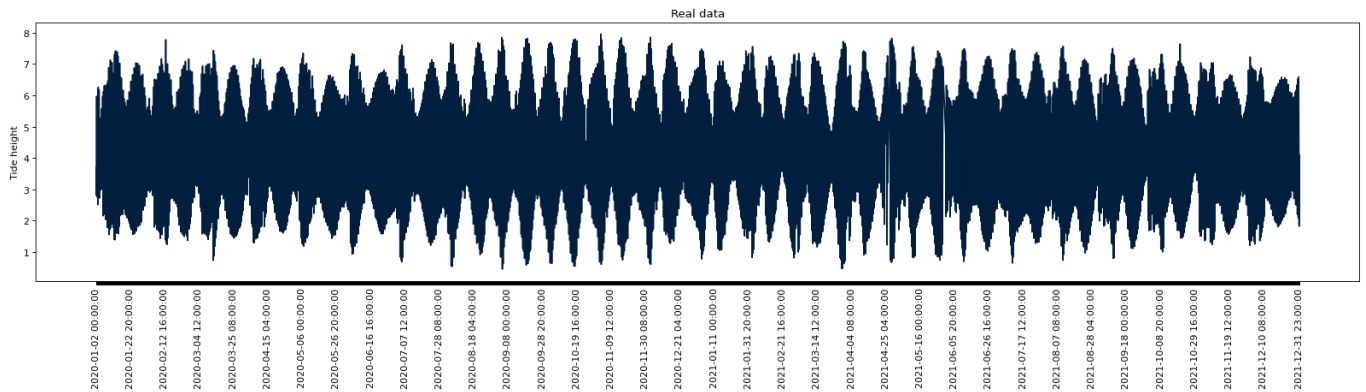


FIGURE 2 – Données 2020 et 2021

Pour tester correctement le modèle, nous séparons les données en deux ensembles : un ensemble d'entraînement et un ensemble de validation. Nous pouvons voir la séparation des données dans l'image 3.



FIGURE 3 – Séparation *Training* et *Testing*

Après la formation, nous avons testé le modèle avec les ensembles de formation et de validation. Dans l'image 4 nous pouvons voir les résultats des deux ensembles et dans l'image 5 nous pouvons voir la prédiction de l'ensemble de validation. En général, les données réelles et les données prédites sont assez similaires dans les deux tests.



FIGURE 4 – Comparaison entre la hauteur de marée prévue et la hauteur de marée réelle

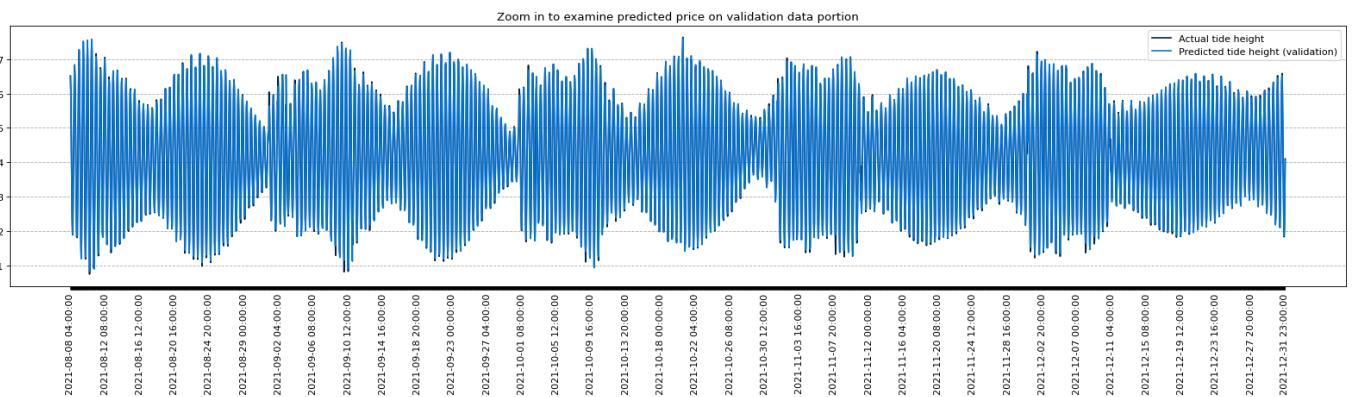


FIGURE 5 – Prédiction de la hauteur des marées sur l'ensemble de validation

Lorsque nous avons vu que le modèle s'adaptait correctement aux données, nous avons fait une prédiction pour l'heure suivante, que l'on peut voir dans l'image 6.

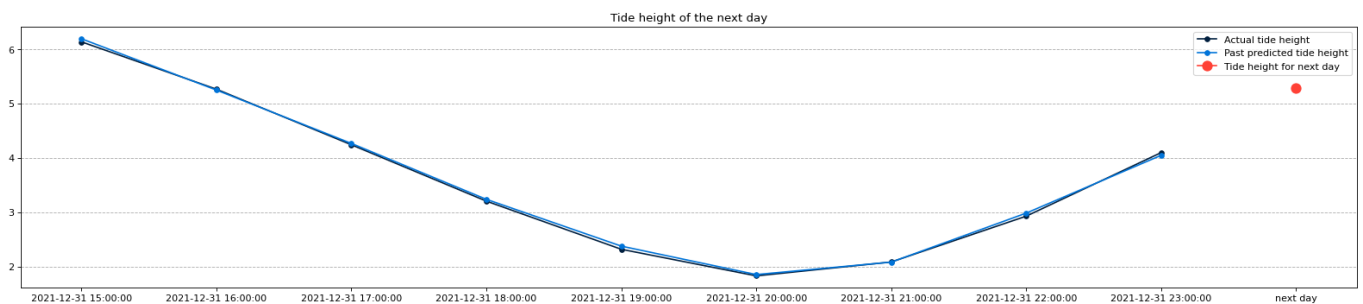


FIGURE 6 – La prédiction du jour suivant

La valeur prédite est de 5,29 mètres pour le 01/01/2022 à 00 :00 :00hrs, alors que la valeur réelle est d'environ 5,32 mètres.

De cette façon, nous pourrions dire que ce modèle pourrait prédire correctement les données sur les marées dans un avenir proche.

2 Pour les deux semaines à venir

2.1 Création d'un script de prédictions

Maintenant que nous avons une bonne connaissance des modèles et des précisions assez suffisantes on peut essayer de prédire des nouvelles valeurs.

L'idée principale est d'écrire un script ou même un module permettant de fournir les hauteurs prédites par un modèle choisis, cependant il est possible que les prédictions soit assez longues, le but étant d'optimiser celles-ci afin que le script soit facilement utilisable.

2.2 Tester d'autres modèles

Au cours de ces deux semaines, nous avons également testé d'autres modèles de prédiction, mais nous n'avons pas obtenu de bons résultats. Cela peut être dû au fait que nous n'avons pas choisi les bons paramètres ou que nous n'avons pas traité les données d'entrée correctement.

De cette manière, nous continuerons à tester de nouveaux modèles de prévision jusqu'à ce que nous obtenions de bons résultats. En particulier, nous testerons un modèle de prévision ARIMA.