

Sébastien ZHOU  
Félix MARCOCCIA  
Jean-Philippe VICHAIKIT  
Annie LIM

## Time Serie Forecasting with Seq2Seq RNN

### Compte rendu de la réunion

#### 7ème réunion (Mercredi 17 juin) :

Pour se débarrasser des nombres négatifs, on pourrait essayer d'utiliser une autre loss :  $\hat{y}\log(\hat{y})$ .

Pour chercher pourquoi la loss ne diminue pas, on pourrait également essayer de changer le learning rate, enlever les dropouts.

Essayer de faire du topic modeling pour déduire le thème de chaque page car celui-ci doit très probablement influencer le nombre de visites. Chercher s'il y a des catégories pour chaque article : historique, scientifique, ... Essayer de récupérer d'autres informations (nombre de pages, nombre de mots, nombre d'images...)

A la place du MSE loss, on pourrait utiliser le SMAPE pour la compétition Kaggle, ou la variante avec epsilon comme l'a fait le gagnant de la compétition pour gérer les valeurs  $\sim 0$ .

Pour le rapport, on devrait intégrer le domaine applicatif avec l'article sur la calcination et aussi les times series du Kaggle.

Pour nous aider nous pouvons lire ces papiers :

- Jason Brownlee - Deep Learning for Time Series Forecasting - Predict the Future with MLPs, CNNs and LSTMs in Python (2018)
- Jason Brownlee - Introduction to Time Series Forecasting with Python - How to Prepare Data and Develop Models to Predict the Future-v1.9 (2020)  
[dl101.zlibcdn.com/dtoken/8386c9e0afe71bbcd74799654f521dbd](https://dl101.zlibcdn.com/dtoken/8386c9e0afe71bbcd74799654f521dbd)
- N D Lewis - Deep Time Series Forecasting with Python\_ An Intuitive Introduction to Deep Learning for Applied Time Series Modeling-CreateSpace Independent Publishing Platform (2016)

Prochaine réunion : 8h45-9h30