

# Veille Technologique

---

## Synthèse

### *Contexte :*

Durant notre parcours, nous avons abordé différents problèmes (clustering, régressions, classification multi-classes ou multi-labels ainsi que de l'image segmentation) sur différents types de données (images, texte, ou données mixtes). Cependant les modèles récurrents n'ont pas été abordés. De ce fait, j'ai décidé d'aborder ce sujet et de parler d'un modèle paru le 5 Novembre 2016 sur arxiv (<https://arxiv.org/abs/1611.01576>).

### *Problème :*

Afin de prédire des résultats ayant une relation dans le temps (par exemple le cours de bourse), il faut un modèle permettant de prendre en considération ce phénomène de récurrence. Pour ce faire, divers modèles existe. Nous aborderont dans un premier temps l'état de l'art de ce qui se fait avant de les évaluer sur un même dataset.

### *Approche :*

La problématique principale des RNN est le problème de Vanishing Gradient du fait que la back propagation se fait dans l'espace et le temps (qui peut faire plusieurs centaines de timesteps). Divers modèles ont été mis en place depuis 50 ans afin de palier à ce problème.

Nous aborderons dans un premier temps les Simple RNN, puis les LSTM (modèles de 1997 et 2000), les Gated Recurrent Unit avant de finir sur les Quasi Recurrent Neural Networks.

Pour finir, nous évaluerons ces différentes cellules sur un dataset préparé et fixe. La faculté d'apprentissage et leurs vitesses d'apprentissage seront comparées à topologie identique.

### *Performances des modèles :*

Concernant les Simple RNN, leur entraînement est très rapide mais leur faculté d'apprentissage très faible. Le LSTM et le GRU sont au coude à coude au long terme mais le GRU a tendance à commencer à apprendre plus tôt et est aussi plus rapide. Pour finir le QRNN, apprend très vite (de l'ordre de 2 à 3 fois plus rapidement) mais à des performances encore un peu faibles.

A l'instar du LSTM en 1997, on peut imaginer que des améliorations seront apportées dans l'avenir pour leur permettre d'atteindre des performances similaires aux LSTM.

### *Résultats :*

Lors de ce projet, nous avons donc découvert les RNN, leurs architectures et l'intérêt par rapport au problème du Vanishing Gradient. Les différentes topologies de RNN ont été abordées ainsi que l'architecture des cellules.

Quant au résultat sur le dataset, elles sont bien en deçà de ce que l'on peut faire car l'objectif était de comparer les modèles et non optimiser la classification.