

Réseaux de Neurones Récurrents (RNN)

- Architecture d'un réseau de neurones récurrents
- Les types de réseaux de neurones récurrents : Vanilla RNN, LSTM, GRU
- Application (Reproduire une étude depuis Kaggle en assimilant les étapes et les interprétations)

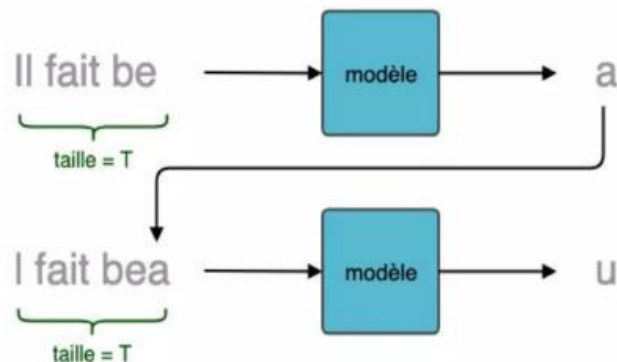
1. Architecture d'un réseau de neurones récurrents

Définition

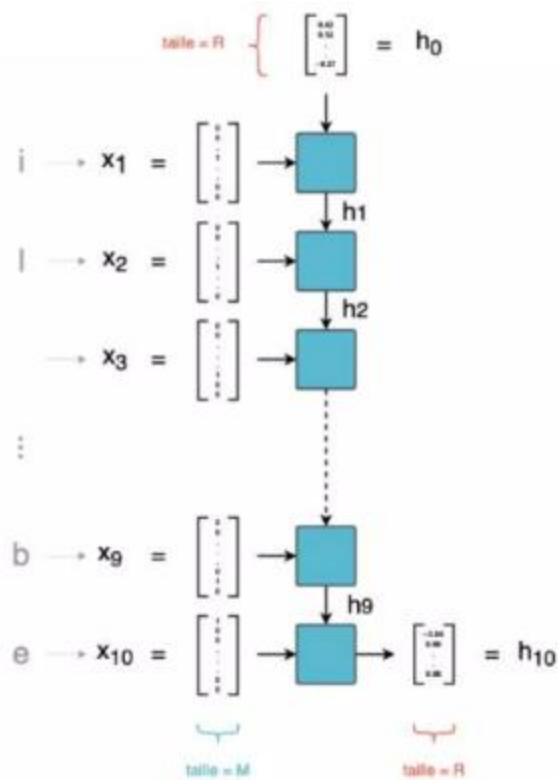
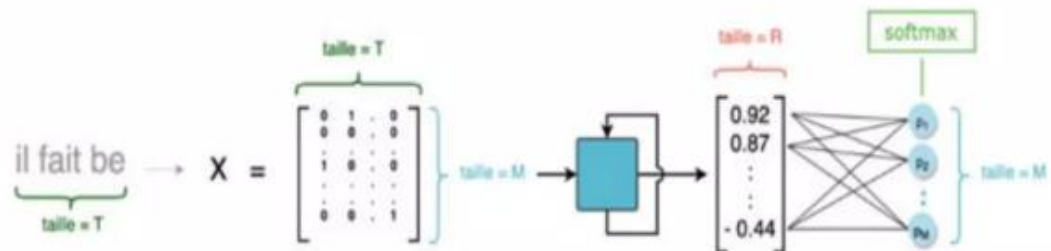
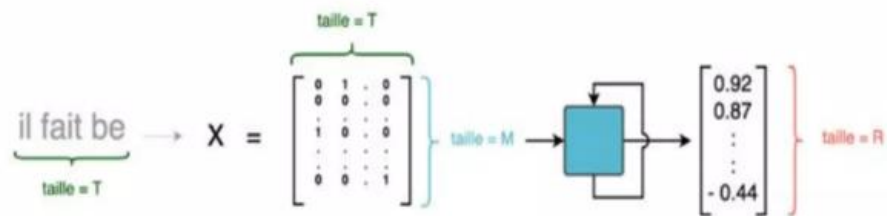
Un réseau de neurones récurrent (RNN, recurrent neural network) est un type de réseau de neurones artificiels principalement utilisé dans la reconnaissance vocale et le traitement automatique du langage naturel (TAL, NLP, TNL). Les RNN sont conçus de manière à reconnaître les caractéristiques séquentielles et les modèles d'utilisation des données requis pour prédire le scénario suivant le plus probable.



Architecture globale



$\text{nb_échantillons (N)} * \text{taille_séquence (T)} * \text{nb_variables (M)}$



Dans une couche RNN, on parcourt successivement les entrées x_1 à x_T . À l'instant t , la $t^{\text{ème}}$ cellule combine l'entrée courante x_t avec la prédiction au pas précédent h_{t-1} pour calculer une sortie h_t de taille R .

Le dernier vecteur calculé h_T (qui est de taille R) est la sortie finale de la couche RNN. Une couche RNN définit donc une relation de récurrence de la forme :

$$h_t = f(x_t, h_{t-1})$$

2. Les types de réseaux de neurones récurrents

- **Vanilla RNN**

Les Vanilla RNN sont les RNN les plus simples. Ils sont constitués d'une couche d'entrée, d'une couche cachée et d'une couche de sortie. La couche cachée est connectée de manière récurrente à elle-même, ce qui permet au RNN de mémoriser les données du passé.

- **LSTM (Long Short-Term Memory)**

Les LSTM sont une variante des Vanilla RNN qui sont conçues pour mieux gérer les problèmes de long terme. Les LSTM utilisent des portes pour contrôler le flux d'informations dans la couche cachée. Cela permet au RNN de retenir les informations importantes du passé, même si elles sont masquées par des informations moins importantes.

- **GRU (Gated Recurrent Unit)**

Les GRU sont une autre variante des Vanilla RNN qui sont conçues pour mieux gérer les problèmes de long terme. Les GRU utilisent des portes pour contrôler le flux d'informations dans la couche cachée. Cela permet au RNN de retenir les informations importantes du passé, tout en évitant les problèmes de vanishing gradient.

- **Comparaison des trois types de RNN**

Les trois types de RNN ont leurs propres avantages et inconvénients. Les Vanilla RNN sont les plus simples à comprendre et à implémenter, mais ils sont moins performants que les LSTM et les GRU pour traiter des séquences temporelles de longue durée. Les LSTM sont plus performants que les Vanilla RNN pour traiter des séquences temporelles de longue durée, mais ils sont plus complexes à implémenter. Les GRU sont une alternative aux LSTM qui est plus simple à implémenter, tout en offrant des performances similaires.