

RESEAUX RNN

Objectif : Construire et tester différentes architectures pour des réseaux RNN, afin de déterminer l'architecture la plus pertinente et de la comparer avec les LSTM et GRU.

Architectures RNN :

- RNN « simple »

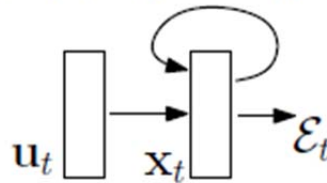
$$x_t = W_{rec}\sigma(x_{t-1}) + W_{in}u_t + b$$

Réseau à l'étape t

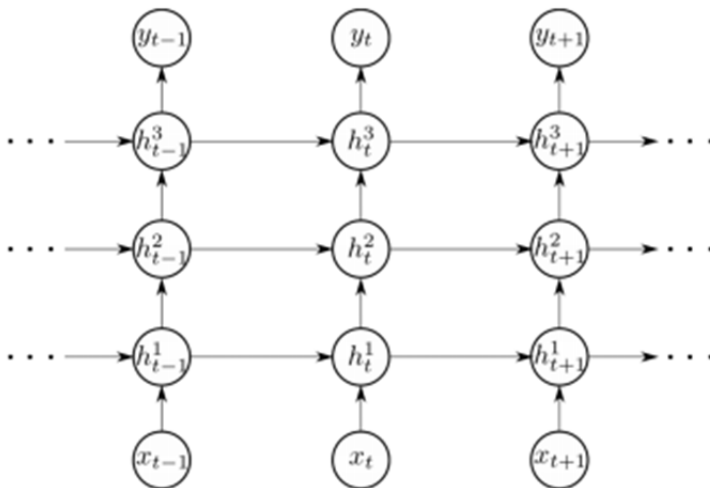
u : entrée

x : état de la couche cachée

ε : erreur en sortie



- RNN multicouches (plusieurs couches cachées)



Réseau déplié à 3 couches cachées

x : entrée

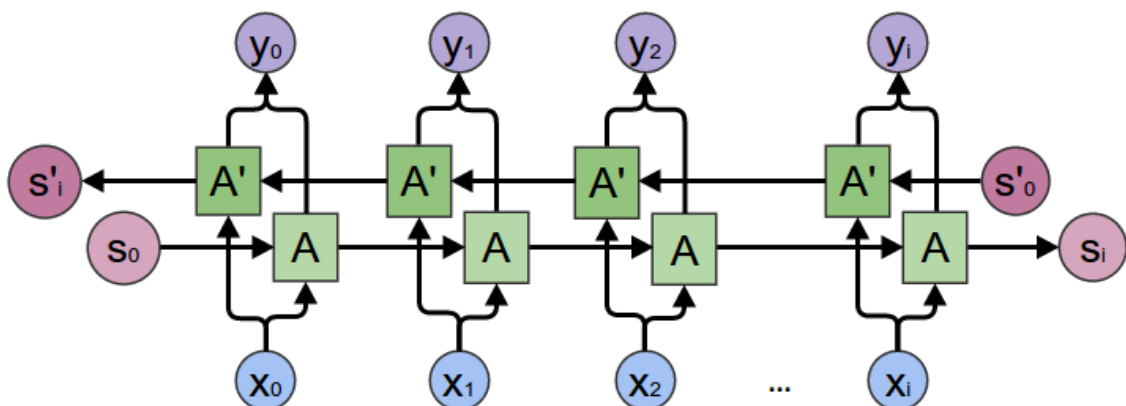
h : état de la couche cachée

l : numéro de la couche cachée

y : sortie

$$h_t^{(l)} = f_h^{(l)}(h_t^{(l-1)}, h_{t-1}^{(l)})$$

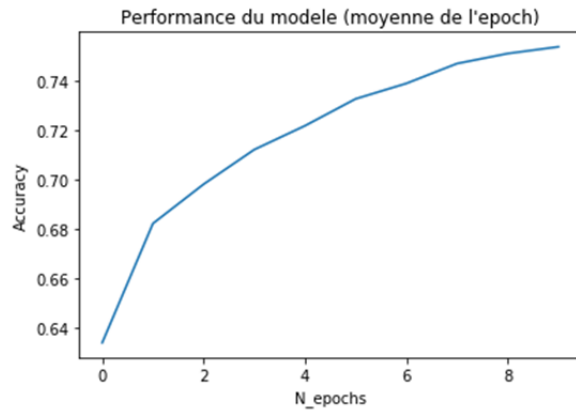
- RNN bidirectionnel



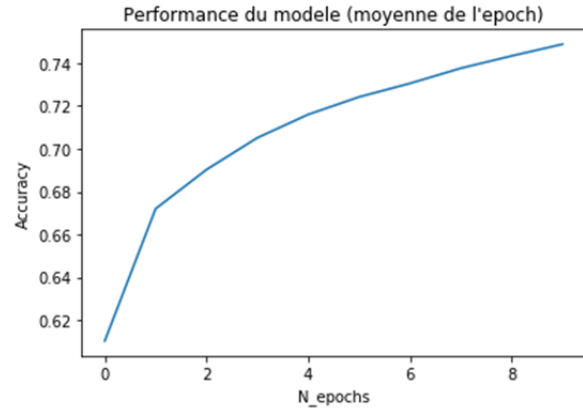
Pour chacune des architectures ci-dessus, on fait plusieurs essais en modifiant la fonction d'activation ou encore en ajoutant un Dropout entre les couches cachées.

Evolution du score f1 sur le jeu de validation en fonction du nombre d'epoch :

RNN Simple



RNN 2 couches cachées



Résultats sur le jeu test :

RNN simple	RNN 2 couches cachées	RNN bidirectionnel
72.6%	73.3%	