

Présentation IAR

Yoann Taillé
Gaspard Ducamp

Androide 2016

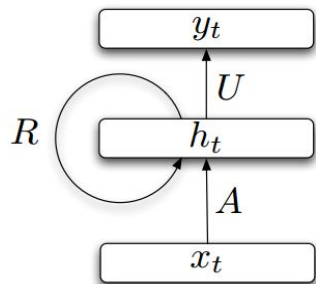
Problématique

Comment résoudre le problème de “disparition du gradient” lors de l’utilisation d’une rétropropagation de gradient dans un réseau de neurones récurrent ?

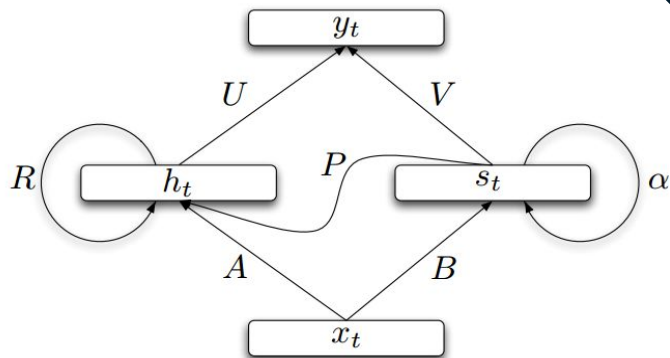
SCRN [Mikolov et al, 2015]

- › Structurally Constrained Recurrent Nets
- › Se base sur un SRN
- › Ajout d'une couche contextuelle

SRN(a) vs SCRNN (b)



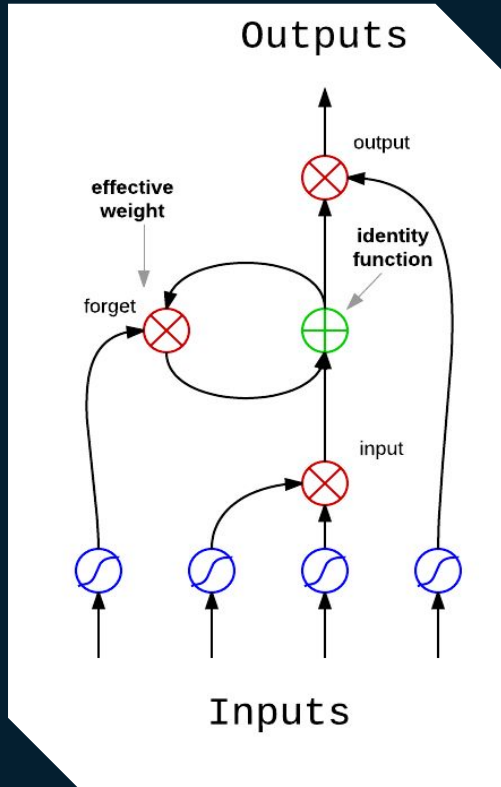
(a)



(b)

LSTM [Hochreiter & Schmidhuber,97]

- › Long Short-Term Memory
- › Méthode "State of the art"
- › Très utilisée dans les tâches de reconnaissance d'écriture, de prédiction linguistique
- › Fournit de bons résultats et permet une mémoire à long terme



Utilisation de différentes portes permettant de bloquer l'apprentissage pendant un temps

Protocole expérimental

- › Travail sur le corpus Penn Treebank
- › Code pour TensorFlow 0.7.1
- › Utilisation de CUDA 7.5 et CUDNN 5.1 pour accélération GPU
- › I5 6600K + GTX 1070 + 16go de RAM

Métriques étudiées

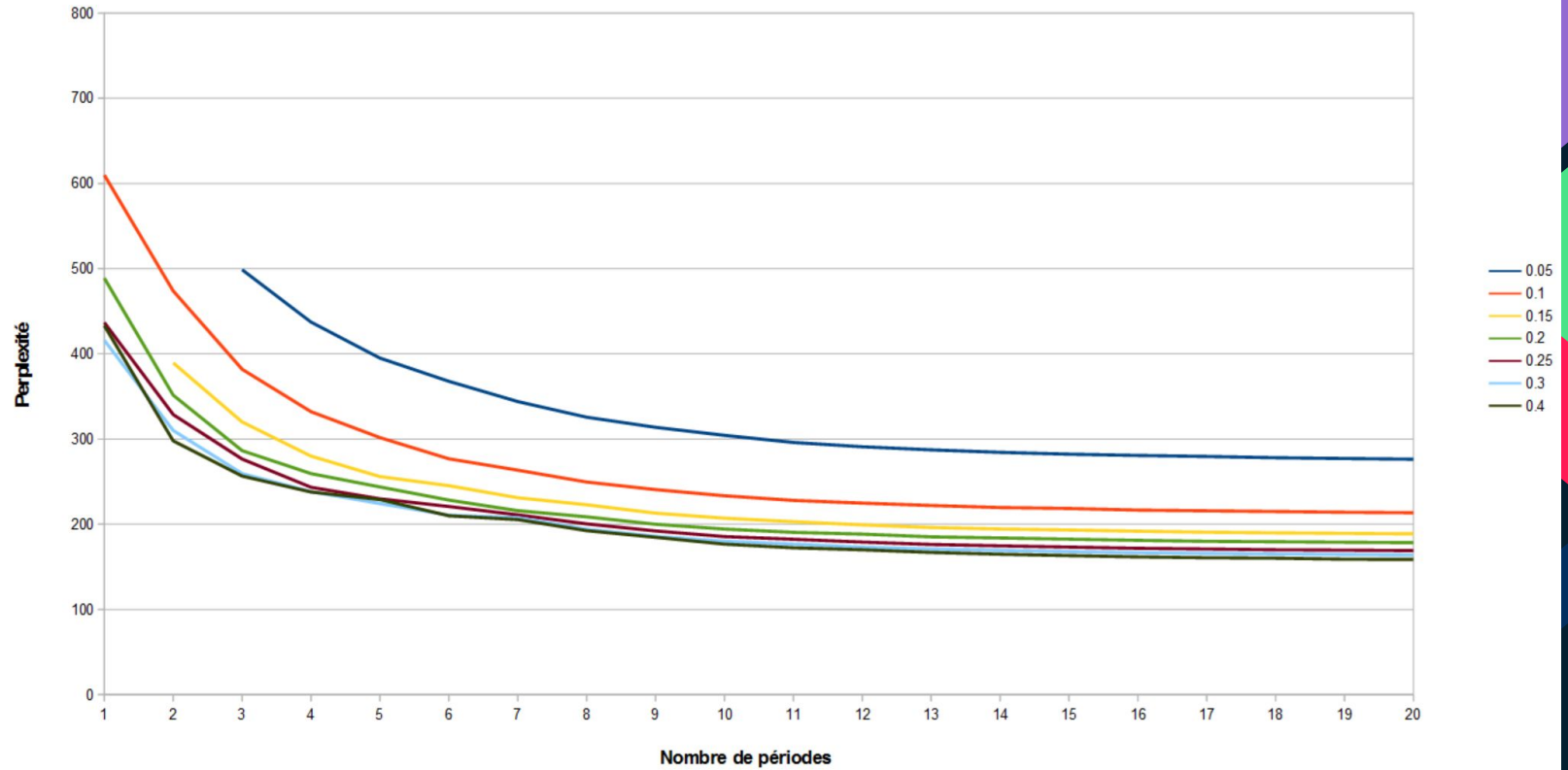
- › Perplexité
- › Temps de calcul
- › Taux d'apprentissage

Tableau comparatif sur PTB

	#Cachées	#Contextuelles	Valid perplexity	Test perplexity	Temps d'exécution
SRN	50	-	212.781	202.488	220.364
SRN	100	-	183.919	175.402	453.054
SRN	300	-	140.508	134.858	1054.153
LSTM	50	-	214.493	205.328	362.046
LSTM	100	-	174.298	166.922	734.399
LSTM	300	-	<u>112.713</u>	<u>109.344</u>	1541.348
SCRN	40	10	<u>206.755</u>	<u>197.445</u>	335.252
SCRN	90	10	<u>169.057</u>	<u>161.948</u>	329.768
SCRN	100	40	<u>161.222</u>	<u>155.274</u>	626.139
SCRN	300	40	122.367	117.377	1328.508

*

Evolution de la perplexité en fonction du taux d'apprentissage initial



GRU [Cho & al,14]

- Gated Recurrent Unit
- Logique similaire à LSTM

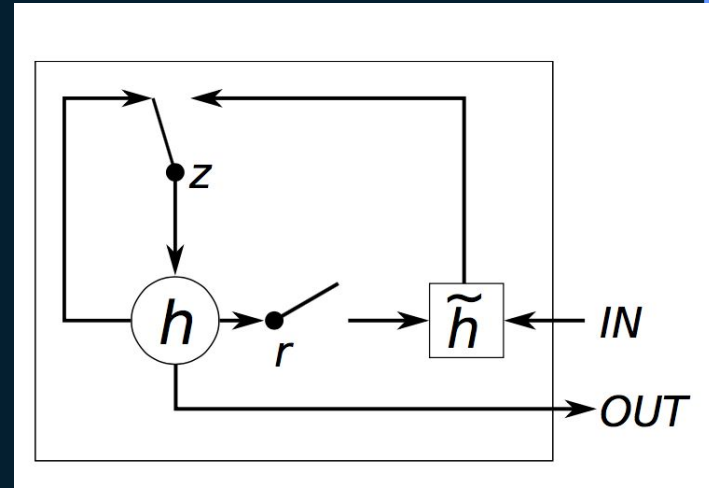


Tableau comparatif sur PTB

	#Cachées	#Contextuelles	Valid perplexity	Test perplexity	Temps d'exécution
SRN	50	-	212.781	202.488	220.364
SRN	100	-	183.919	175.402	453.054
SRN	300	-	140.508	134.858	1054.153
LSTM	50	-	214.493	205.328	362.046
LSTM	100	-	174.298	166.922	734.399
LSTM	300	-	112.713	109.344	1541.348
SCRN	40	10	206.755	197.445	335.252
SCRN	90	10	169.057	161.948	329.768
SCRN	100	40	161.222	155.274	626.139
SCRN	300	40	122.367	117.377	1328.508
GRU	50	-	185.366	177.085	351.190
GRU	100	-	152.255	146.488	656.664
GRU	300	-	107.124	103.048	1506.527



**Merci de
votre
attention!**



Références :

- › [Empirical Evaluation of Gated Recurrent Neural Networks on Sequence Modeling](#) [Cho & al, 14]
- › [Long Short-Term Memory](#) [Hochreiter & Schmidhuber, 97]
- › [Learning Longer Memory in Recurrent Neural Networks](#) [Mikolov & al, 15]
- › [WildML](#)