

## Analyse automatique de commentaire Amazon/film



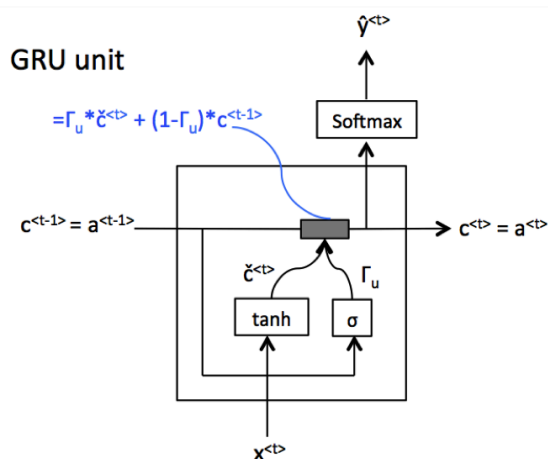
A l'inverse d'un algorithme basique de NLP (topic modelling, word embeddings...), un modèle de réseaux de neurones récurrents à l'avantage de garder une trace du sens des mots qu'il a analysés précédemment (La phrase est-elle au singulier? Le sujet au masculin?). Ce modèle est très performant lorsqu'il s'agit d'identifier le sens global d'une phrase ainsi que sur des problématiques de traduction.

J'ai pu l'utiliser sur un exercice de classification concernant les commentaires qu'a pu laisser un client d'amazon après un achat:

```
In [53]: X[0]
```

```
Out[53]: 'Great CD: My lovely Pat has one of the GREAT voices of her generation. I have listened to this CD for YEARS and I still LOVE IT. When I\'m in a good mood it makes me feel better. A bad mood just evaporates like sugar in the rain. This CD just oozes LIFE. Vocals are just STUUNNING and lyrics just kill. One of life\'s hidden gems. This is a desert isle CD in my book. Why she never made it big is just beyond me. Everytime I play this, no matter black, white, young, old, male, female EVERYBODY says one thing "Who was that singing ?"'
```

Afin d'identifier si le commentaire laissé était positif, j'ai mis en place un réseau GRU (Gated recurrent unit) afin qu'il puisse garder en mémoire l'information analysée.



```
from keras.models import Sequential
from keras.layers import GRU, LSTM, Dense, Embedding
def GRU_RNN():
    model = Sequential()
    model.add(Embedding(input_dim=10000,
                        output_dim=32, input_length=100))
    model.add(GRU(units=32, return_sequences=True))
    model.add(GRU(units=32, return_sequences=False))
    model.add(Dense(units=1, activation='sigmoid'))
    return model
```