TP8 : NLP and Time-Series using RNN,LSTM,GRU

Why RNN

In some cases, we need to deal :

Sequence data eg :Text : a sentence

-time series

Sequence : an element depends on the previous elements

Fully connected NN : not take into account the dependency= Bad

Convolutional NN : capture local structure. Not sufficient : .suppose all local relations are the same across the whole text.

Does not capture anything longer than the size of windows

Not very good.

Recurrent NN

Several types of problems :

-Many to one

-Sentence-> classification

Ex : a review of a product

-guess the rate :

-one to many

-text generation

Ex : music generation

-many to many :

-language translation

.suppose all local relations are the same across the whole text.

-does not capture anything longer than the size of window

-not very good

La différence entre RNN,CNN et FCN.

In CNN/FCN, only comput gradient the update NN, don(t take the previous as the input.

The previous state of the RNN is the input of the current state.

LSTM and GRU :

Problem of RNN :

When the sentence is large enough🡪 start to forget the previous parts.

Exemple : Restaurant= good

But RNN forgets it.

-Grandient Vanishing

-Gradient Exploding

Lstm(kind of better in general) a peu près égal à GRU

Have a lne of « memory » : gates

Run throught the text to keep important information

LSTM : has 3 gates per nodes

GRU has 2 gates per nodes

Keras : LSTM(….)

GRU(….)

Should try both

-people prefer LSTM

Read

Understand

Erase the code

Rewrite

Play with RNN vs LSTM vs GRU (Tune parameters)