

# **Cours 2 : Deep Learning pour les données séquentielles**

**François HU - 13/10/2020**

Data Scientist au DataLab de la Société Générale Assurances

Doctorant à l'ENSAE-CREST

Les cours se trouvent ici : <https://curiousml.github.io/>

# Sommaire

## 1. Introduction

## 2. Recurrent Neural Network (RNN)

- Modèles RNN « classique »
- D'autres architectures RNN
- Modèle Bidirectional-RNN (BRNN)
- Modèle Deep RNN

## 3. Problème de la disparition du gradient

- [option] LSTM / GRU
- [option] BERT

## Programme

### Introduction

### Représentations vectorielles

### Deep Learning pour NLP

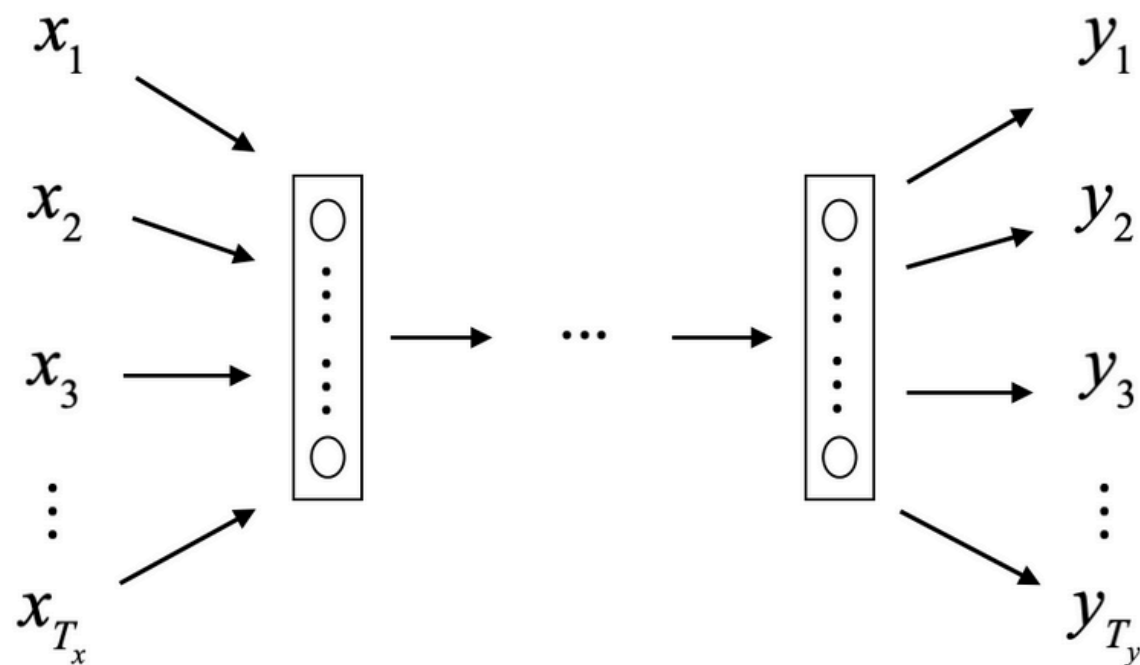
### Active Learning

# Introduction

## Pourquoi les modèles séquentiels ?

- classification de textes / analyse de sentiment
- Named Entity Recognition (NER)
- génération de textes / de musiques
- traducteur de langue automatique
- ...

## Pourquoi pas les réseaux de neurones « standard » ?

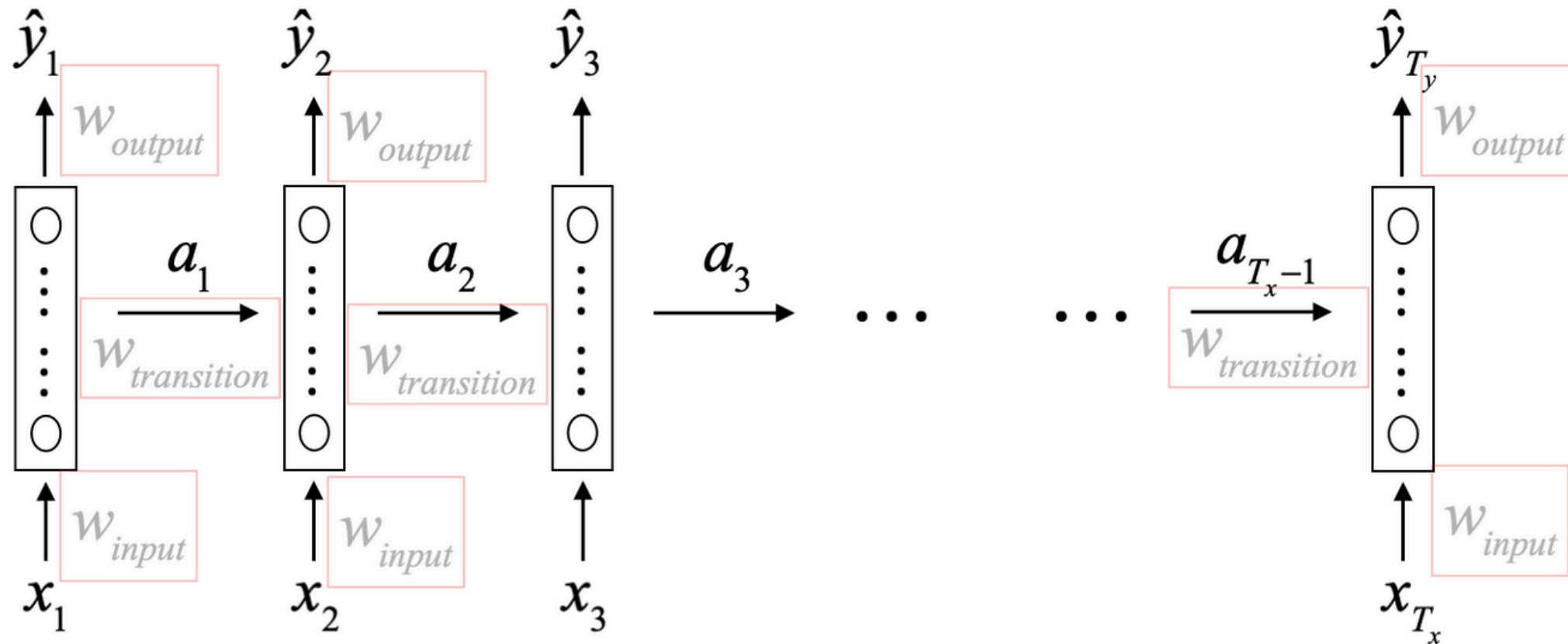


- inputs / outputs peuvent être de **tailles différentes**
- ne tient pas en compte des **différentes positions** des mots

## **2. Recurrent Neural Networks (RNN)**

# Modèle RNN « classique »

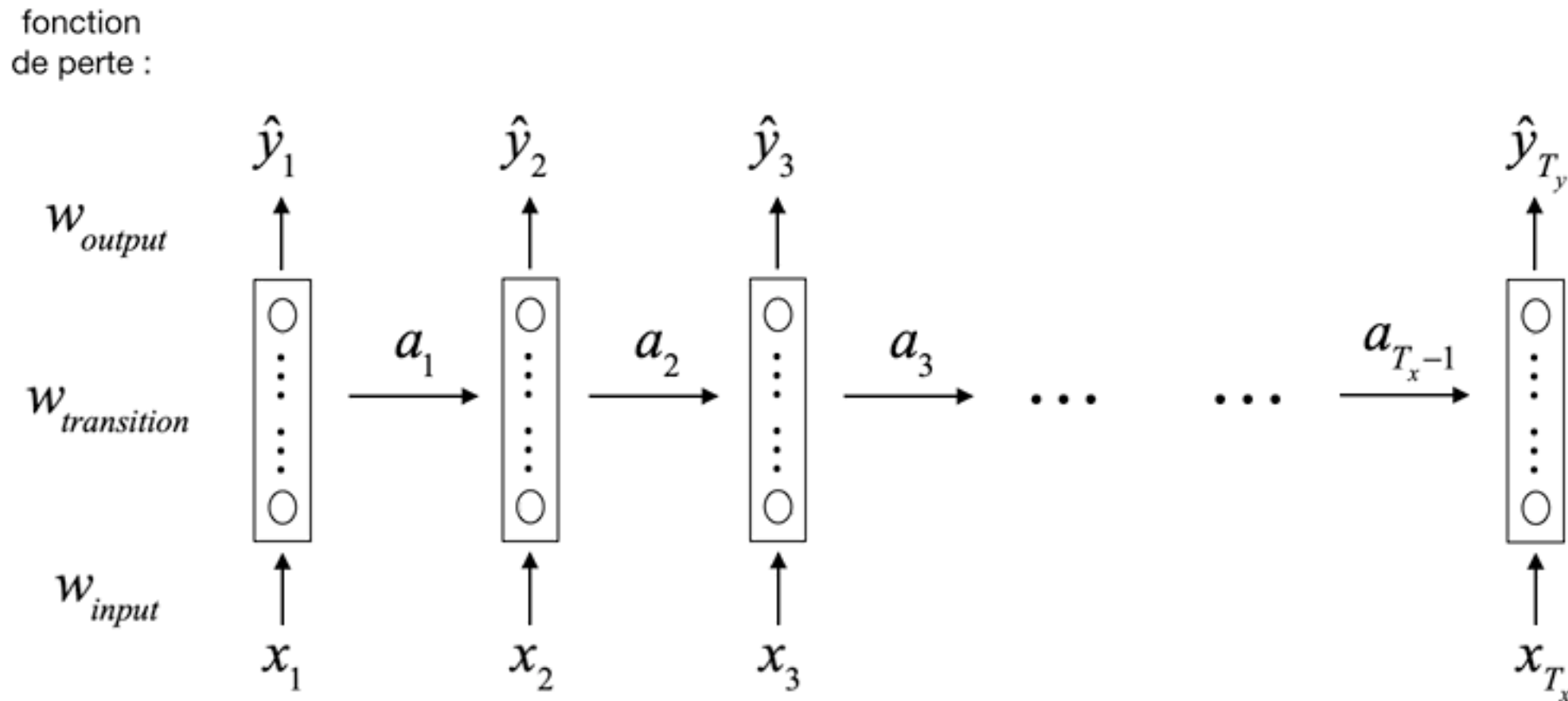
**Word embedding** (plongement de mot en français) : vectorisation des mots de sorte que les mots apparaissant dans des contextes similaires ont des significations apparentées



- possibilité de manipuler des **séquences de taille variable**
- tout calcul futur **tient compte des calculs passés**
- les poids / **paramètres sont partagés** dans le temps

# Modèle RNN « classique »

**Propagation avant** (forward propagation) d'un RNN :



- possibilité de manipuler des **séquences de taille variable**
- tout calcul futur **tient compte des calculs passés**
- les poids / **paramètres sont partagés** dans le temps