

A photograph of Donald Trump speaking at a podium. He is wearing a dark blue suit, a white shirt, and a red tie. His right hand is raised in a fist, and his mouth is open as if he is shouting or speaking passionately. The background is blurred, showing a crowd of people. The text 'Présentation TER' and 'Génération de tweets de Donald Trump' is overlaid on the image in white.

# Présentation TER

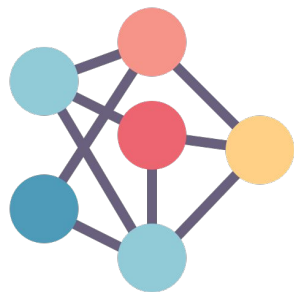
## Génération de tweets de Donald Trump

Gabin Marc Telly Mberi Kongo, Florent Jakubowski

# Sommaire

1. Introduction
2. Modèles utilisés
3. Résultats
4. Conclusion et perspectives

# 1. Introduction



Récents progrès en  
génération de texte  
SeqtoSeq (2014),  
transformeur (2017)



Ancien président  
américain très actif sur  
tweeter



Langage simple et peu  
structuré

# 1. Données

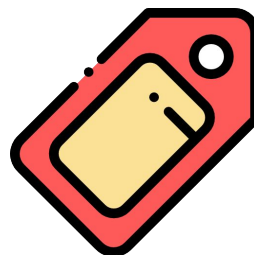
52 183

Nombre de tweets  
utilisés



Nettoyage de la données

- Suppression des retweets pour les deux premiers modèles que nous avons utilisé
- Texte minuscule



- Tokenisation des mots du datasets
- Utilisation de la librairie de représentations vectorielles pré-entraînées du sens des mots, *Fasttext*

## 2. Modèles utilisés

### LSTM

- Architecture de réseau de neurones récurrents
- Représentation vectorielles des mots

### GRU

avec attention

- Architecture de réseau de neurones récurrents
- Représentation vectorielles des mots
- Mécanisme d'attention

### Markov

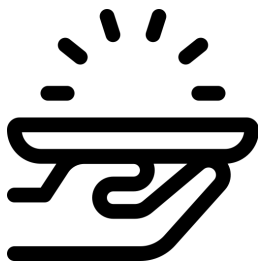
Chain

- Modèle statistique
- Entraîné sur les caractères

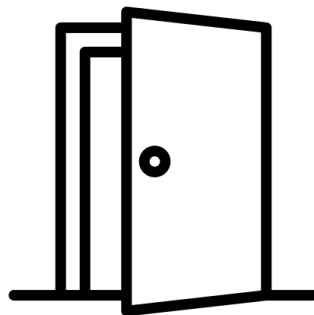
### GPT2

- Modèle de langue pré-entraîné
- Architecture basée sur le transformeur

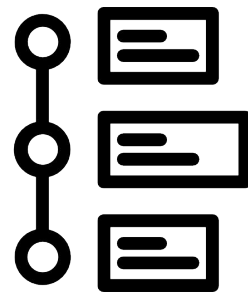
## 2. Recurrent Neural Networks de type LSTM et GRU



LSTM et GRU pallie au problème d'évanescence du gradient

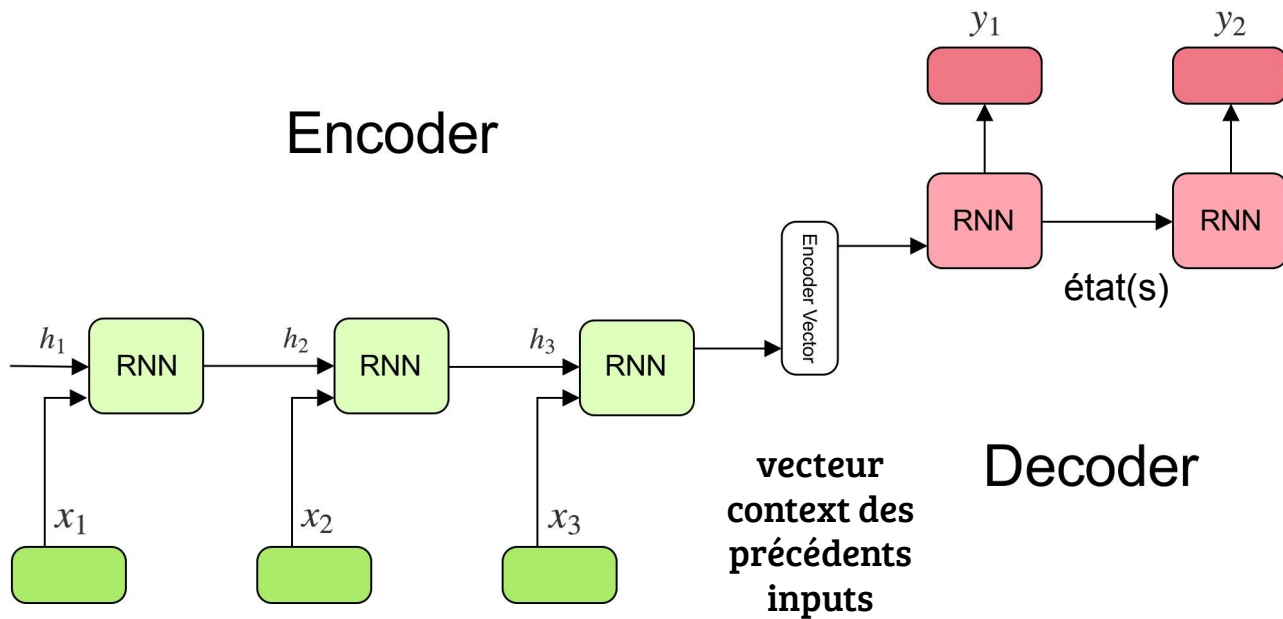


Utilise un système de portes et d'état(s) pour stocker et transférer le "contexte"

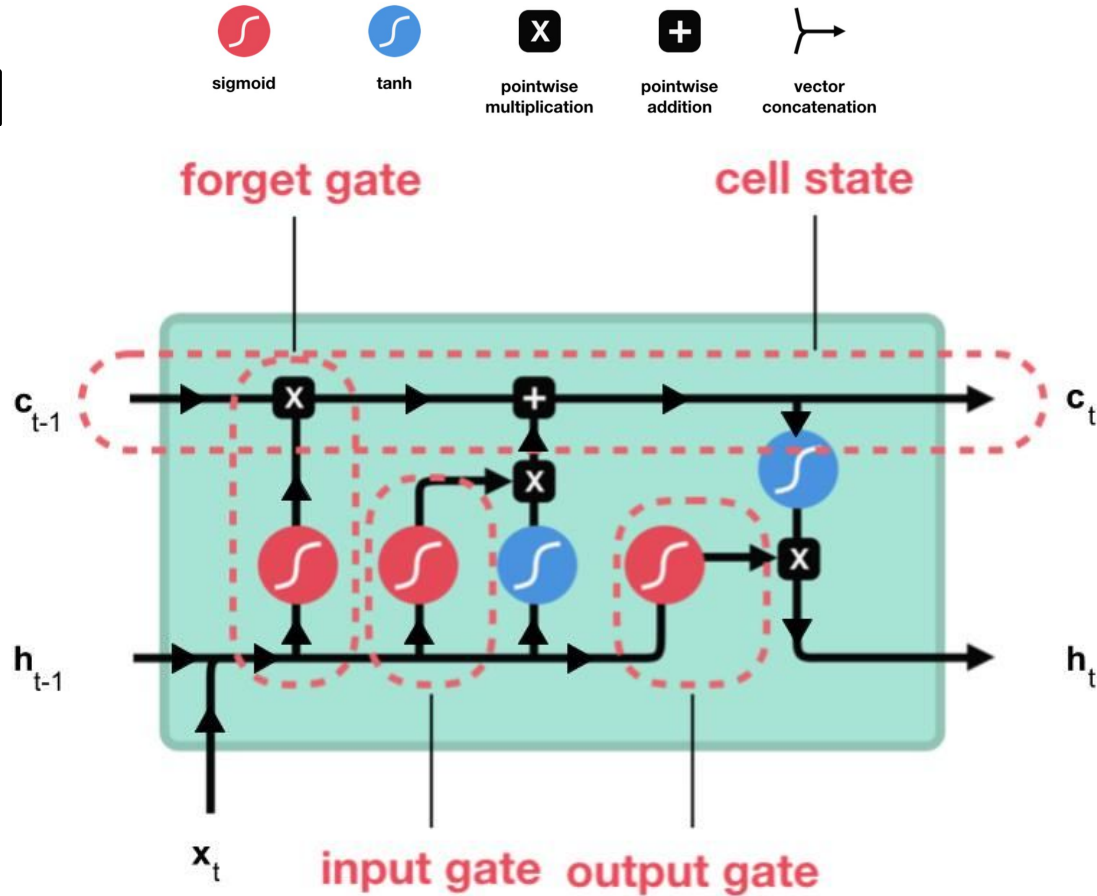


Utile quand il existe de l'information dans l'ordonnancement des données et le sens de lecture (time series, texte)

## 2. Architecture encodeur-décodeur



## 2. LSTM

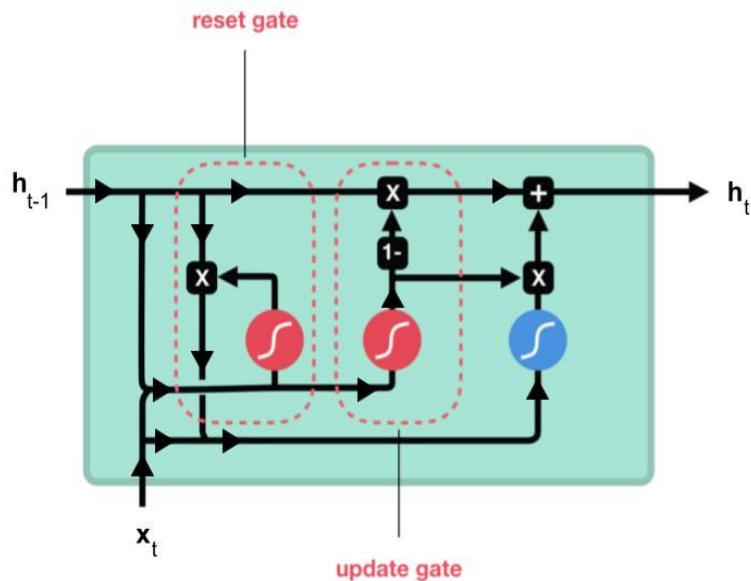




## 2. LSTM résultats

Entrée	Sortie
“barack obama is the”	“worst president in the world. all of the world is laughing at the same time ?”
“71 000 000 votes incredible”	“- we are going to win !”
“thanks to all of you”	“to the great state of texas. i will be a great governor of montana !”

## 2. GRU avec mécanisme d'attention



sigmoid



tanh



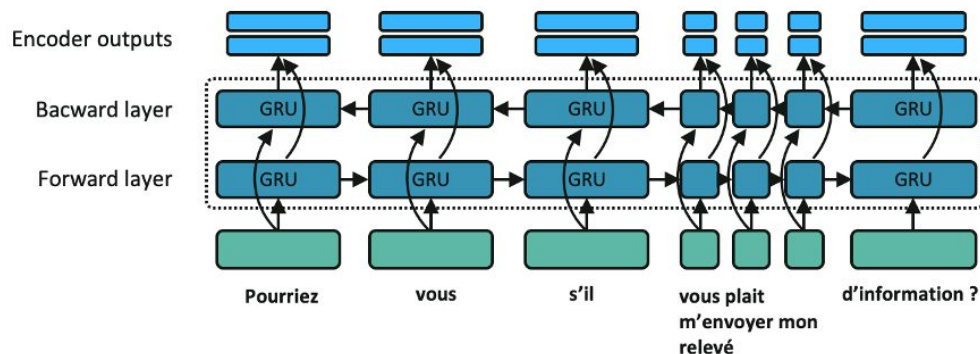
pointwise  
multiplication



pointwise  
addition



vector  
concatenation

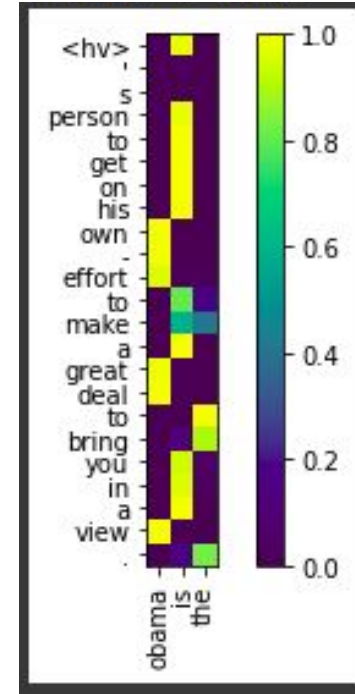


**— université  
— lumière  
— LYON 2**

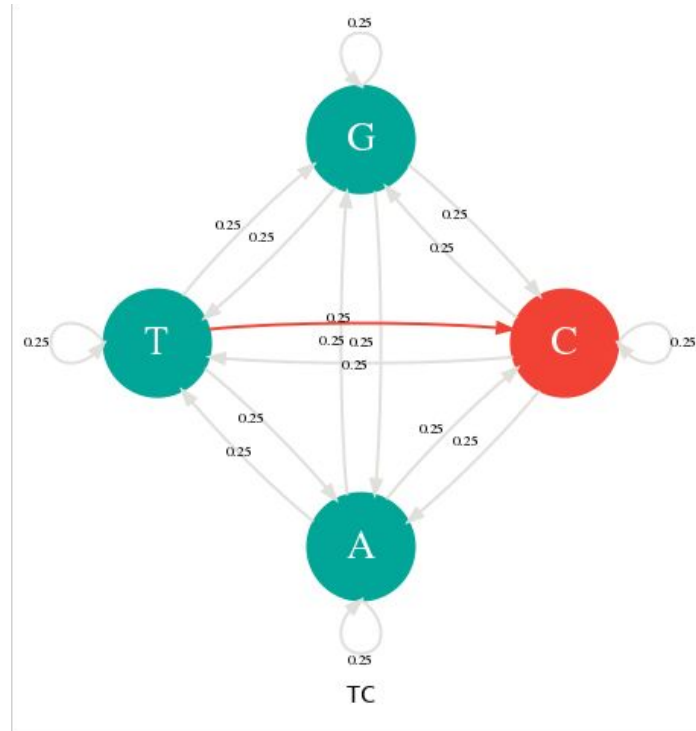


## 2. GRU, résultats

Entrée	Sortie
"obama is"	"<hv> 's person to get on is own - effort to make a great deal to bring you in a view."
"we need that for the"	" @ <hv> speech on your passing the # <hv> . we need it ! "
"my people are"	"for all the @ god party 's card - a great commodity is a terrific person."



## 2. Chaîne de Markov



-N-grammes

-Fréquence  
d'apparition par  
rapport aux n-gramme

## 2. Chaîne de Markov, résultats

n-gramme initial	longueur du n-gramme	longueur du tweet	Sortie
"Hill"	4	70	"Hillars they way use othere of in the phone who do sold at @Trump. Nice.If"
"Hillary"	7	65	"Hillary Clinton like him! I will never given? RT @newtgingrich on an int"
"Barack Obama"	12	50	"Barack Obama was president! RT @tashlutsa: @CLewandowski_: Wat"

- Sens sémantique limité
- Présence des RT
- Présences de compte tweeter réels

## 2. Modèles de langues pré-entraînés, GPT2

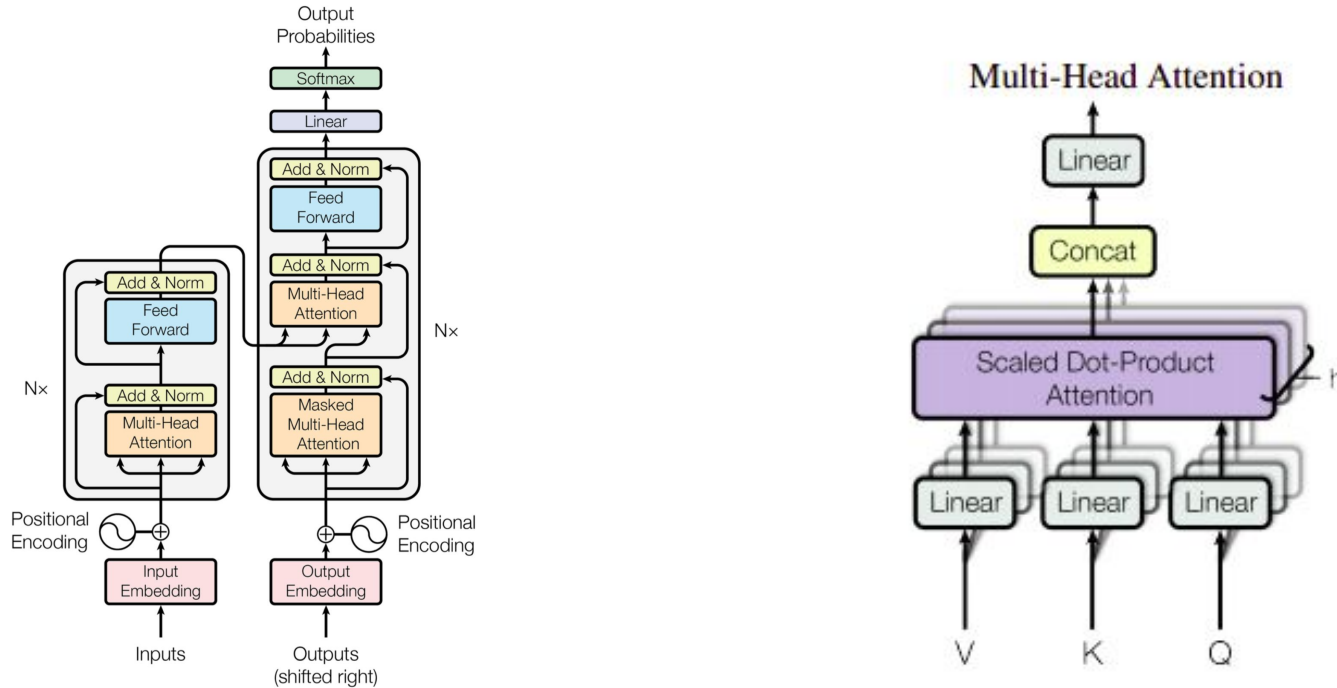
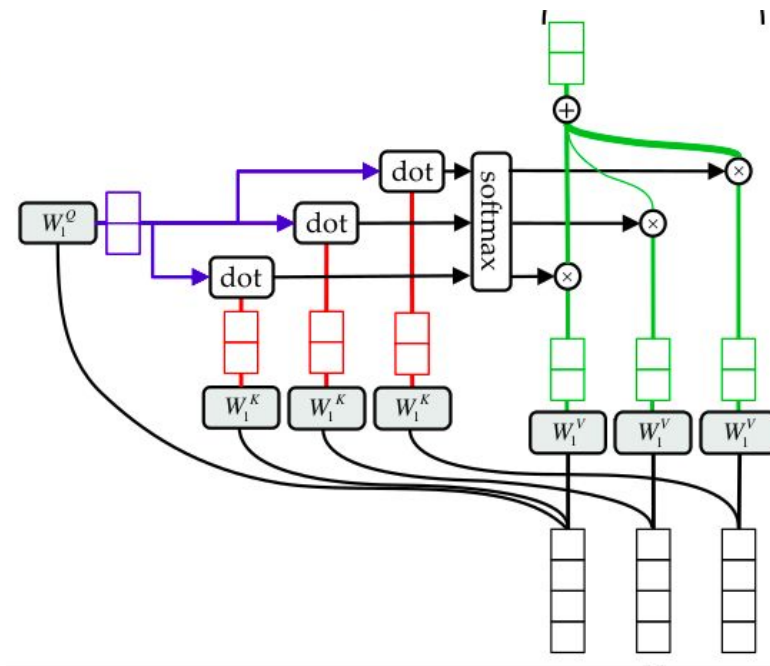
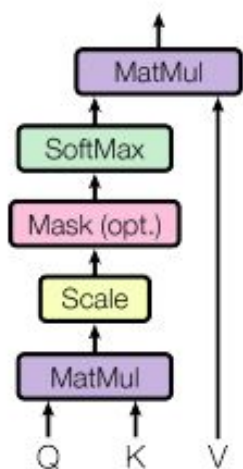


Figure 1: The Transformer - model architecture.

## 2. Modèles de langues pré-entraînés, GPT2

Scaled Dot-Product Attention



self	attention	mechanism
$x_1 = (0.7, 1.1)$	$x_2 = (-0.4, 0.3)$	$x_3 = (0.3, -0.5)$
$q_1 = x_1 W^q$	$q_2 = x_2 W^q$	$q_3 = x_3 W^q$
$k_1 = x_1 W^k$	$k_2 = x_2 W^k$	$k_3 = x_3 W^k$
$v_1 = x_1 W^v$	$v_2 = x_2 W^v$	$v_3 = x_3 W^v$
$q_1 \cdot k_1 = 29.61$	$q_1 \cdot k_2 = 31.02$	$q_1 \cdot k_3 = 26.79$
20.94	21.93	18.94
$\alpha_{12} = 0.26$	$\alpha_{12} = 0.70$	$\alpha_{13} = 0.04$
$0.26 \times v_1$	$0.70 \times v_2$	$0.04 \times v_3$
$y_1 = \sum_{i=1}^3 \alpha_{1i} v_i$	$y_2 = \sum_{i=1}^3 \alpha_{2i} v_i$	$y_3 = \sum_{i=1}^3 \alpha_{3i} v_i$

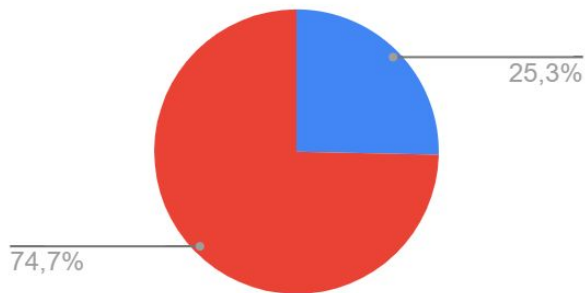


## 2. Résultats de GPT2

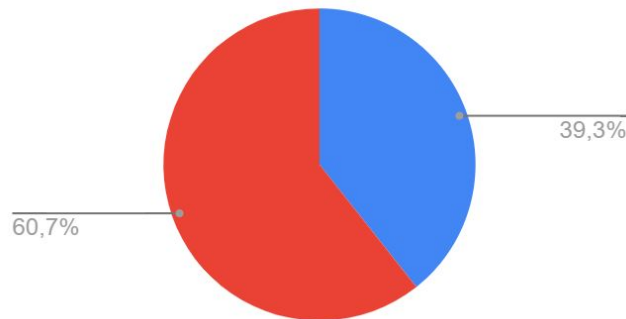
Entrée	Sortie
Barack Obama and democrats	"Barack Obama and democrats can't get their way...we're at a standstill", 'Congressman Mike Crapo (@CumpoGOP) has been an incredible advocate for America and has my Complete and Total End"
"Joe"	'Joe. My time is going to come when I'm 100% there and you will WIN!'
"Thank to our"	'Thank to our beautiful @GOPladers...', 'RT @TomFitton: "Senator Mark Warner did an amazing job on Fox & Friends last night. Great job on your new book. You are the man for it but you still"
sleepy Joe	'sleepy Joe Biden? He is also a totally stupid person!'

### 3. Résultats google form

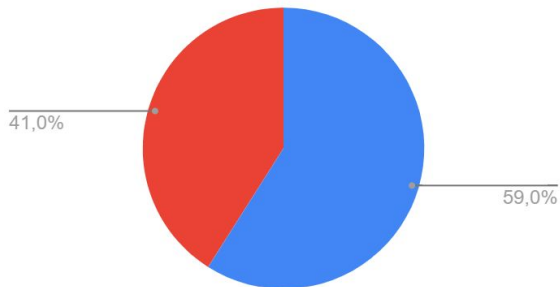
Markov



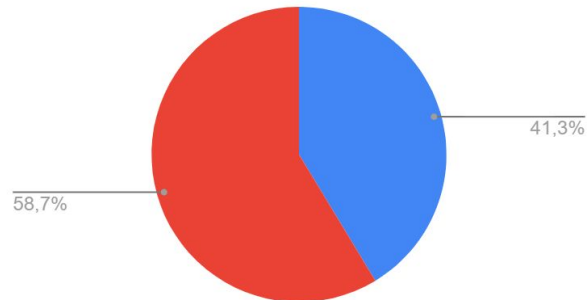
LSTM



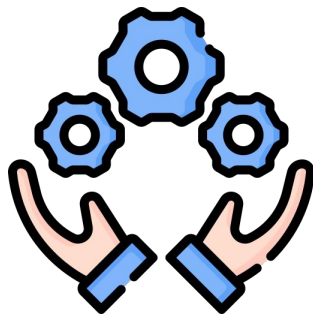
GRU avec attention



Trump



# Conclusion et perspectives



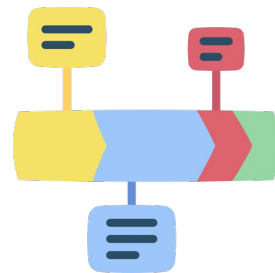
Expérience  
pratique



Réelle avantage  
de l'attention



Nous aurions voulu  
explorer plus les  
modèles langues  
pré-entraînés et  
inclure les résultats au  
formulaire soumis aux  
étudiants



Découper le jeu de  
données en fonction  
des périodes de la  
présidence de  
Donald Trump.



Avez-vous des questions ?