



République Algérienne Démocratique et Populaire  
Ministère de l'enseignement supérieur et de la  
recherche scientifique

Université Larbi Tébessi - Tébessa

Faculté des Sciences Exactes et des Sciences de la Nature et de la Vie  
Département : Mathématiques et Informatique



Mémoire de fin d'étude  
Pour l'obtention du diplôme de **MASTER**  
Domaine : Mathématiques et Informatique  
Filière : Informatique  
Option : Systèmes d'information

Thème

## **Détection de fausses informations dans les réseaux sociaux**

Présenté Par :

**Ktir Abassia**

Devant le jury :

Mr Djeddi Chawki	MCA	Université Larbi Tébessi	Président
Mr Bendib Issam	MCB	Université Larbi Tébessi	Examineur
Mr Benjenna Hakim	Pr	Université Larbi Tébessi	Encadreur

Date de soutenance : 14 septembre 2020



# Résumé

Actuellement, les médias sociaux sont une source d'informations pour de nombreux utilisateurs. Les informations sont publiées sur ces plateformes très rapidement sans vérifier leur authenticité ou leur source. La plupart de ces informations n'ont pas de source et sont classées comme des rumeurs. Notre travail vise à classer automatiquement les informations erronées ou correctes en appliquant l'approche d'apprentissage automatique.

À cette fin, nous avons introduit un modèle qui repose sur le traitement du langage naturel « NLP » en utilisant la technologie de « Indexation des mots » avec un algorithme d'apprentissage en profondeur CNN-1D afin d'obtenir la meilleure classification possible de l'information.

**Les mots clés :** réseaux sociaux, fausses rumeurs, classification des informations, apprentissage automatique, apprentissage profond, traitement du langage naturel, Indexation des mots.

# Abstract

Currently, social media is a source of information for many users. Information is published on these platforms very quickly without verifying its authenticity or its source. Most of this information has no source and is classified as rumors. Our work aims to automatically classify erroneous or correct information by applying the machine learning approach.

To this end, we have introduced a method based on Natural language processing "NLP" using "word embedding" technology with a CNN-1D deep learning algorithm in order to obtain the best possible classification of information.

**Keywords:** social networks, false rumors, classification of information, machine learning, deep learning, NLP, word embedding.

# ملخص

حاليا تعتبر وسائل التواصل الاجتماعي مصدر معلومات للكثير من المستخدمين، تنتشر المعلومات في هذه المنصات بسرعة كبيرة من دون التحقق من صحتها او مصدرها، معظم هذه المعلومات ليس لها اي مصدر وتصنف ضمن الشائعات. يهدف عملنا الى تصنيف المعلومة الى خاطئة او صحيحة تلقائيا من خلال تطبيق نهج التعلم الالي . تحقيقا لهذه الغاية، قدمنا طريقة تعتمد على معالجة اللغة الطبيعية باستخدام تقنية تضمين المعلومات مع خوارزمية التعلم العميق CNN-1D من أجل الحصول على أفضل تصنيف ممكن للمعلومات.

**الكلمات المفتاحية:** شبكات اجتماعية، الشائعات الكاذبة، تصنيف المعلومات، التعلم الالي، التعلم العميق، معالجة اللغة الطبيعية، تضمين الكلمات.

# **Remerciements**

*Tout d'abord, je remercie Dieu*

*Je voudrais remercier dans un premier temps, mon encadreur **Pr. Benjenna Hakím**, pour sa patience, sa disponibilité et ses bons conseils qui ont alimenté ma réflexion.*

*Je voudrais remercier les membres de jury qui ont accepté de juger ce modeste travail **Dr. Djeddi Chawki** et **Dr. Issam Bendib** et tous les enseignants de département math&info de l'université de Tébessa.*

*Je voudrais remercier, ma famille pour leur soutien, ainsi que mes amis d'être ma deuxième famille.*

*Enfin Je tiens à remercier toutes les personnes qui ont contribué au succès de la rédaction de ce mémoire.*

# ***Dédicace***

*Je dédie ce modeste travail*

*A mes parents et à toute la famille sans exception*

*A tous mes enseignants*

*A tous mes collègues d'étude et de travail*

*A tous mes amis*

*Ktir Abassia*

# Sommaire

## Table des matières

Introduction Générale.....	1
1. Introduction.....	2
2. Objectif.....	2
3. La description des chapitres.....	2
Chapitre 1 : Les réseaux sociaux.....	3
1.1. Introduction.....	4
1.2. Le réseau.....	4
1.3. Naissance de réseaux sociaux.....	5
1.4. Définition.....	6
1.5. Concepts à ne pas confondre avec les réseaux sociaux.....	7
1.5.1. Les réseaux sociaux et les médias sociaux.....	7
1.5.2. Les réseaux sociaux et les communautés.....	7
1.5.3. Comparaison avec les médias traditionnels.....	7
1.6. L'objectif des réseaux sociaux.....	8
1.7. Les indicateurs des réseaux sociaux.....	8
1.8. Types de réseaux sociaux.....	8
1.8.1. Selon l'évolution et l'apparition.....	9
1.8.2. Selon la fonctionnalité.....	10
1.8.3. Selon le point de vue des chercheurs.....	11
1.9. Intérêts des réseaux sociaux.....	11
1.10. Principaux réseaux sociaux.....	15
1.11. Cartographie des réseaux sociaux numériques.....	18
1.12. Fonctions et fonctionnalités.....	18
1.13. Enjeux des réseaux sociaux numériques.....	19
1.14. Problèmes ouverts des réseaux sociaux.....	21
1.15. Conclusion.....	23
Chapitre 2 : Les rumeurs dans les réseaux sociaux.....	24
2.1. Introduction.....	25
2.2. Définitions d'une rumeur.....	25

2.3.	Les caractéristiques de la rumeur .....	28
2.4.	Rumeur via les réseaux sociaux .....	28
2.5.	Représentation de réseau social .....	29
2.6.	Méthodes de détection des rumeurs .....	31
2.6.1.	Les approches textuelles .....	32
2.6.2.	Approches basées sur les sources .....	34
2.6.3.	Approches multimédia .....	34
2.7.	Critiques sur les méthodes de détection des fausses informations .....	37
2.8.	Synthèse et discussion .....	40
2.9.	Conclusion.....	41
Chapitre 3 : L'apprentissage profond & Le traitement de langage naturelle .....		42
3.1.	Introduction.....	43
3.2.	L'intelligence artificielle .....	43
3.3.	Apprentissage automatique.....	43
3.4.	L'apprentissage profond .....	44
3.5.	Le traitement du langage naturel.....	49
3.6.	L'apprentissage profond et le traitement de langage naturel .....	49
3.7.	Conclusion .....	52
Chapitre 4 : L'approche proposée .....		54
4.1.	Introduction .....	55
4.2	. L'architecture globale .....	55
4.3.	Le prétraitement de données .....	57
4.4.	Le modèle proposé .....	60
4.5.	Implémentation et évaluation.....	62
1.	Présentation des outils utilisés .....	62
2.	Présentation de la base de données usuelle .....	64
3.	Word embedding (Glove) .....	66
4.6.	Résultats obtenus et discussions .....	66
4.7.	Conclusion .....	69
Conclusion Générale.....		70

## *Liste des figures*

<b>Figure</b>	<b>Page</b>
Figure 1 : représentation d'un réseau [1]	<b>4</b>
Figure 2 : Évolution du Web de l'ère des ordinateurs personnels (PC) au Web 4.0.[2]	<b>5</b>
Figure 3 : Graphe non orienté représente les relations sociométriques entre les internautes[5]	<b>6</b>
Figure 4 : Panorama des médias sociaux 2019 [12]	<b>12</b>
Figure 5 : Carte mondiale des réseaux sociaux [16]	<b>18</b>
Figure 6 : Pyramide des besoins d'information	<b>25</b>
Figure 7 : représentation nœud-lien	<b>30</b>
Figure 8 : représentation matricielle	<b>30</b>
Figure 9 : Exemple d'une rumeur répandue sur les réseaux sociaux [76]	<b>31</b>
Figure 10 : Les approches de prédiction de la rumeur [75]	<b>33</b>
Figure 11 : Les approches de prédiction du rumeur base sur les images [75]	<b>35</b>
Figure 12: Topologie de réseau de neurones [64]	<b>45</b>
Figure 13 : Topologie de réseau de neurones profond.[65]	<b>46</b>
Figure 14: Architecture de réseau de neurone convolutionnel [66]	<b>48</b>
Figure 15 : application de NLP Classique vs NLP avec le deep learning [69]	<b>50</b>
Figure 16 : L'architecture global du système	<b>55</b>
Figure 17 : Code de tokenisation en python	<b>56</b>
Figure 18 : Les quinze premières mots vides de la langue anglais	<b>57</b>
Figure 19 : Code d'élimination des mots vides en python	<b>57</b>
Figure 20 : Division de données en python	<b>58</b>
Figure 21 : Configuration du notre modèle	<b>61</b>
Figure 22 : Représentation de deux classes de jeux de données	<b>63</b>
Figure 23 : Quelques lignes de jeux de donnes	<b>64</b>
Figure 24: Lien de téléchargement de word embedding	<b>65</b>



Figure 25 : Précision et Erreur pour notre modèle	66
Figure 26 : Matrice de confusion du notre modèle	66
Figure 27 : Précision, Rappelle et F-score du notre modèle	67
Figure 28: modèle 2 avec 100 époques	68

## ***Liste des tableaux***

<b>Tableau</b>	<b>Page</b>
Tableau 1 : Types des réseaux sociaux selon l'évolution et l'apparition	<b>10</b>
Tableau 2 : Tableau comparatif entre les réseaux sociaux selon la fonctionnalité	<b>10</b>
Tableau 3 : Tableau comparatif entre les réseaux sociaux selon le point de vue des chercheurs	<b>11</b>
Tableau 4 : Tableau comparatif entre les réseaux sociaux	<b>17</b>
Tableau 5 : Tableau de concept associé à la fausse information	<b>27</b>
Tableau 6 : Synthèse des méthodes de détection de fausses informations dans les réseaux sociaux	<b>40</b>
Tableau 7 : Description de dataset	<b>63</b>
Tableau 8 : Description des colonnes de dataset	<b>64</b>
Tableau 9 : Comparaison des résultats	<b>69</b>

---

# Introduction Générale

---

## 1. Introduction

Ces dernières années, Internet est devenu un élément indispensable de toute vie humaine, en fonction de leur niveau d'éducation et de leurs intérêts, en raison du développement des médias sociaux auxquels tout le monde et partout peuvent accéder afin de localiser des lieux sur la carte, rechercher des informations et acheter différents produits et exprimer des sentiments et des opinions, ainsi que lire des articles et des journaux.

À l'instar des médias traditionnels, les médias sociaux sont rapides et faciles à afficher, à partager et à consulter en raison de la diversité de l'information (sous forme de message, vidéo, photo, événement, tweet, etc.). De plus, l'échange d'informations à travers des interactions et des organisations.

Avec l'augmentation de la diffusion des informations sur les moyens de communication, il est devenu difficile pour le lecteur et les entreprises de distinguer les informations fausses des informations correctes. Par conséquent, il est nécessaire de fournir des programmes et des sites pour analyser les informations et réduire la propagation des rumeurs.

## 2. Objectif

Dans notre projet de fin d'étude, l'objectif est d'explorer le domaine de l'analyse de rumeurs, de donner un aperçu de ce qu'est ce problème et de représenter les différentes techniques et approches proposées et leurs résultats. Nous avons également tenter de présenter une approche Deep Learning pour le traitement de langage nature en tant qu'une contribution à la problématique de détection de rumeurs

## 3. La description des chapitres

- ❖ Chapitre 1 : Le premier chapitre présente le contexte de réseaux sociaux et notre problématique.
- ❖ Chapitre 2 : Le second chapitre sera un état de l'art sur les rumeurs dans les réseaux sociaux
- ❖ Chapitre 3 : Le troisième sera une description de la technique utilise qui est l'apprentissage approfondi et le traitement de langage naturelle (NLP).
- ❖ Chapitre 4 : Le quatrième chapitre sera une description de nos modèles proposés, les outils utilisés et les résultats obtenus.

# Chapitre 1

---

## Les réseaux sociaux

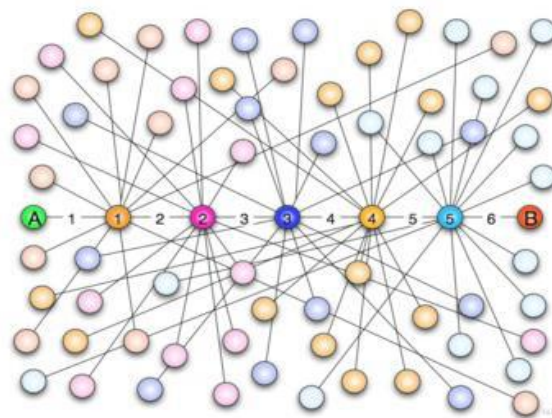
## 1.1.Introduction

Avec l'avènement du Web en 2000, les réseaux sociaux sont devenus un support de communication basé sur le Web, qui permet aux utilisateurs d'avoir des conversations, de partager des informations et de créer du contenu. Cela permet aux internautes de diffuser les informations, de donner leur avis ou leurs critiques sur un sujet donné, D'où la nécessité d'analyser les postes, des tweets ou des informations pour vérifier l'exactitude des informations et faire une évaluation de ces informations.

Dans ce premier chapitre nous détaillons la notion de réseaux sociaux online (OSNs<sup>1</sup>), leur intérêt et leur enjeux leur problèmes dont notre problématique de recherche.

## 1.2.Le réseau

Historiquement et étymologiquement, le réseau est un filet, un tissu ou un entrelacement de fils et les figures de sa généalogie montrent que cette référence originelle est persistante jusqu'à nos jours [1].



**Figure 1** : représentation d'un réseau [1]

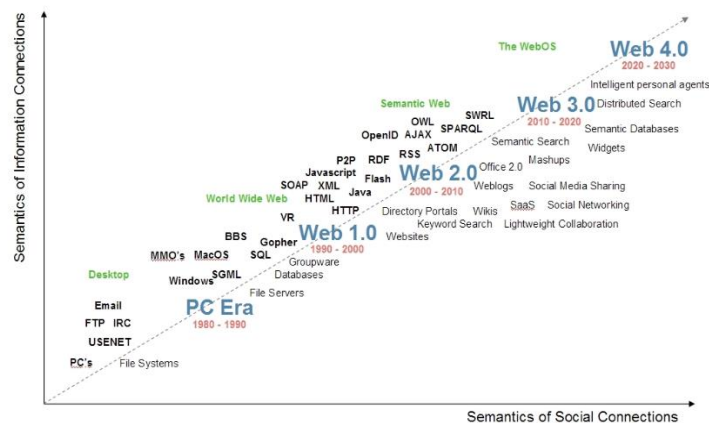
Mais c'est « l'objet enserré dans les mailles du réseau-filet qui a changé au fil du temps : corpscosmos, nature et planète, société et organisations ».

Au XVII<sup>e</sup> siècle et jusqu'à la fin du XVII<sup>e</sup> siècle, le terme « réseau » ne sort pas du langage des médecins : le mot n'entretient donc aucun support avec la communication.

<sup>1</sup> Online Social Newtork

### 1.3.Naissance de réseaux sociaux

Internet a connu une évolution avec l'arrivée du web 2.0 (figure 1) , qui permet grâce aux divers appareils comme les ordinateurs, Smartphone, tablette tactile, une interaction permanente entre les utilisateurs d'internet à travers les blogs et les réseaux sociaux, qui peuvent aussi diffuser des informations ou créer du contenu par eux-mêmes grâce au développement des navigateurs [2].



**Figure 2 :** Évolution du Web de l'ère des ordinateurs personnels (PC) au Web 4.0.[2]

Cependant une nouvelle notion est apparue sur internet et qui est celle du web social, qui considère internet comme un endroit de socialisation entre les personnes.

2003 est l'année officielle du lancement de nombreux réseaux sociaux, depuis ce moment ils ont connu une popularité grandissante que ce soit pour une utilisation personnelle ou professionnelle, avec le développement de la technologie, les réseaux sociaux deviennent plus riches et présentent diverses utilisations.

Mais c'est en 2004 que les réseaux sociaux ont connu une grande évolution, avec l'arrivée de Facebook qui a connu rapidement une grande extension du nombre d'utilisateurs. Facebook a été créé à Harvard et trois ans après il a atteint la barre des huit millions de membres, puis de nombreux réseaux sociaux ont vu le jour tel que Twitter, YouTube ... Ces réseaux sociaux ne consacrent pas de budgets de publicité pour se faire connaître mais sont de plus en plus populaires.

Ces dernières années les réseaux sociaux ont connu une montée fulgurante, ils sont actuellement présents pratiquement dans le monde entier, et prennent une importance et une réelle utilité que ce soit de type personnelle ou professionnelle.

## 1.4.Définition

Classiquement, un réseau social est défini comme une entité constituée d'un ensemble d'individus et des relations qu'ils entretiennent les uns avec les autres, directement ou indirectement par le biais de chaînes de relations. Il existe des dizaines de définitions de réseaux sociaux, nous présentons quelques-unes [2]:

### 1) Selon Tichy :

« Un réseau social décrit une société comme un système de participants " individus, groupe, organisations" reliés par une variété de relations. Chaque paire de participant n'est pas jointe forcément et certaines sont jointes par plusieurs relations ». [3]

### 2) Selon Giles Hogben et ENISA :

« Un réseau social consiste en un ensemble fini d'acteurs et les relations définies entre eux. Un acteur peut être une seule personne, ou un groupe de personnes. Les acteurs dans un réseau social sont liés par des relations. Le type et le niveau de confiance de ces relations peut varier dépendamment des acteurs impliqués. Les amis, la famille, ou encore les collègues sont autant d'exemples de types de relations ». [4]

### 3) Selon Dajana Kapusova Leconte :

« Un réseau social est une structure formée par des relations entre des personnes. Cette structure sociale est composée de noeuds, généralement représentés par des individus ou des organisations. Les noeuds sont reliés entre eux par diverses connaissances sociales qui peuvent aller d'une connaissance simple jusqu'à un lien familial très fort » (figure 3). [5]

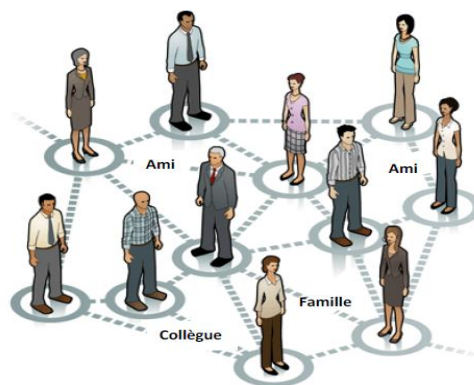


Figure 3 : Graphe non orienté représente les relations sociométriques entre les internautes [5]



## 1.5. Concepts à ne pas confondre avec les réseaux sociaux

### 1.5.1. Les réseaux sociaux et les médias sociaux

On confond souvent Web 2.0 et réseaux sociaux. Le Web 2.0 regroupe les différentes innovations d'ordre technique et sociologique qui permettent aujourd'hui aux internautes d'échanger, de prendre la parole et, d'une certaine manière, le pouvoir. Au même titre que les blogs et les forums de discussions, les réseaux sociaux ne représentent qu'une partie du Web 2.0. Cependant, ils en constituent un des éléments essentiels et des plus populaires auprès du grand public.

### 1.5.2. Les réseaux sociaux et les communautés

Il faut distinguer les termes « les réseaux sociaux » et « les communautés ». En fait, il y a une relation d'ordre entre ces deux concepts : pour qu'il y ait communauté, il faut qu'il y ait réseau social (mais l'inverse n'est pas vrai).

On peut donc considérer qu'une communauté est un réseau social auquel se greffent des caractéristiques propres. Dans une thèse consacrée à « la nouvelle place des communautés virtuelles dans la construction et la gestion de la relation à la marque », il faut ajouter à un réseau social quatre éléments fondamentaux pour constituer une communauté :

- Un objectif commun partagé.
- Des rites, un programme.
- Une structure sociale.
- Un sentiment d'appartenance.

### 1.5.3. Comparaison avec les médias traditionnels

Les médias sociaux diffèrent des médias traditionnels en de nombreux points [6] :

- **Expressivité.** En contraste avec les articles publiés par les médias traditionnels, les messages publiés par les utilisateurs des médias sociaux sont courts, voire parfois très courts. La longueur des messages est parfois limitée par le service.
- **Volume.** Là où les médias traditionnels reposent sur un petit nombre de contributeurs, chaque média social compte un grand nombre d'utilisateurs, chacun d'entre eux publiant plus ou moins régulièrement des messages.

- **Hétérogénéité.** Les messages publiés abordent des sujets divers et variés, allant d'évènements banals de la vie quotidienne à des évènements importants et/ou globaux. Qui plus est, contrairement aux médias traditionnels, les messages ne sont pas catégorisés, ni structurés.
- **Rapidité.** Enfin, la grande force des médias sociaux par rapport aux médias traditionnels est l'immédiateté de la publication. Il n'y a en effet (en principe) aucun filtrage sur le contenu publié. Cette rapidité est amplifiée par le fait que pour la plupart des médias sociaux, une part importante des utilisateurs accèdent au service depuis un terminal mobile

### 1.6.L'objectif des réseaux sociaux

Il s'agit tout d'abord de se donner les bons objectifs : les réseaux sociaux font aussi un travail sur d'autres items que les ventes, comme la présence à l'esprit, la notoriété, l'implication dans la marque.

La capacité des réseaux sociaux à créer un bruit positif sur une marque, en un temps record et à l'échelle mondiale [6].

### 1.7. Les indicateurs des réseaux sociaux

Parmi les indicateurs des réseaux sociaux on trouve :

- **Reach** : traduit « portée » en français, il permet d'avoir une idée du nombre d'internautes touchés par une publication sur un réseau.
- **Taux d'engagement** : il reflète le niveau d'interaction entre votre communauté et votre marque ou enseigne.

### 1.8.Types de réseaux sociaux

Dans le monde des réseaux sociaux, on trouve plusieurs classifications suivant plusieurs critères, parmi ces typologies [7][8][9]:

### 1.8.1. Selon l'évolution et l'apparition

Nous retrouvons également une tentative de typologisation des réseaux sociaux numériques, selon laquelle les auteurs dressent l'évolution et l'apparition des réseaux. Pour eux, il existe huit types de réseaux :

<b>Réseaux généralistes</b>	<b>Centrés autour de l'individu, et non plus de l'activité professionnelle. Ces sites permettent indirectement de nouer des affinités personnelles, sans pour autant avoir vocation unique d'être des sites de rencontre [7].</b>
<b>Réseaux politiques</b>	Du fait que le monde politique s'interroge beaucoup sur les réseaux sociaux, ils entrent de plein pied dans la réflexion des politiques avec deux expériences : la campagne d'Obama <sup>2</sup> et celle de Ségolène Royale <sup>3</sup> .
<b>Réseaux hyperlocal</b>	Dans le but de renforcer les liens au niveau local, ce type de réseau a vu le jour. L'idée sous-jacente est de mieux connaître ses voisins de promouvoir et d'encourager les solidarités [8].
<b>Réseaux d'universités</b>	Par exemple en trouve : Réseau campus <sup>4</sup> , etnoka <sup>5</sup> .
<b>Réseaux d'entreprises</b>	Via les OSNs, ces entreprises peuvent évaluer leur présence en termes d'opportunité et de risque.
<b>Réseaux associatifs</b>	Les auteurs donnent deux exemples : l'Association Française de Sociologie et les réseaux des créatifs culturels.
<b>Réseaux sociaux de communautés d'intérêts</b>	Ce sont des réseaux sociaux qui proposent des types de relations beaucoup plus spécifiques et qui pour certains s'apparentent à des communautés d'intérêts. D'un côté, il existe des réseaux liés aux âges de la vie comme Beboomer <sup>6</sup> qui s'adresse aux seniors actifs de la génération. De l'autre

<sup>2</sup> <http://bababillgates.free.fr/index.php/comment-obama-a-utilise-le-webmarketing-pour-remporter-lelection- americaine/>

<sup>3</sup> Voir la Ségosphère, une cartographie réalisée par Linkfluence sur le site : <http://www.observatoire-presidentielle.fr/?pageid=12>

<sup>4</sup> <http://reseau-campus.com/>

<sup>5</sup> <http://etnoka.fr/>

<sup>6</sup> <http://fr.beboomer.com/>

	côté, il y a des réseaux sociaux réservés aux enfants comme Globe2child <sup>7</sup> .
<b>Réseaux de passionnés</b>	Les auteurs terminent leur typologie en notant que des réseaux comme Culture visuelle <sup>8</sup> qui traite de l'image sous toutes ses formes est déjà plus qu'une communauté d'intérêts.

**Tableau 1 :** Types des réseaux sociaux selon l'évolution et l'apparition

### 1.8.2. Selon la fonctionnalité

D'autres chercheurs ont tenté de classer et de catégoriser les réseaux sociaux. Pascal Faucompré [9] a essayé dans un ticket intitulé : Ras le bol des réseaux sociaux ? d'élaborer un classement suivant la fonctionnalité :

<b>Networkings</b>	<b>Qui permettent les échanges entre les professionnels.</b>
<b>Bloglikes</b>	Ils ressemblent vaguement aux blogs et sont souvent le refuge d'ados en mal de reconnaissance.
<b>Spécialisés</b>	Regroupent des communautés autour d'un thème bien précis.
<b>Micro-blogging</b>	Chat publique instantané.
<b>Fourre-tout</b>	Ce sont les inclassables qui se servent du collaboratif ou du participatif pour alimenter leur service.
<b>Open-sources</b>	Plateformes qui permettent aux utilisateurs de créer leurs propres réseaux.

**Tableau 2 :** Tableau comparatif entre les réseaux sociaux selon la fonctionnalité

<sup>7</sup> <http://www.globe2child.org/>

<sup>8</sup> <http://culturevisuelle.org/>

### 1.8.3. Selon le point de vue des chercheurs

Thelwall [10] catégorise les OSNs selon leurs trois objectifs qu'il nomme respectivement :

<b>Réseaux sociaux de socialisation</b>	<b>Caractérisé par son aspect récréatif et conçu pour les loisirs de communication sociale entre les membres. Les connexions sont souvent utilisées pour trouver d'amis existants hors ligne, comme par exemple : Facebook .</b>
<b>Réseaux sociaux de réseautage</b>	Utilisés davantage pour trouver de nouveaux contacts et peuvent servir pour trouver de nouveaux contacts et entrer en connexions avec des personnes inconnues auparavant comme c'est le cas de LinkedIn ou Viadeo.
<b>Réseaux sociaux de navigation</b>	Comme Digg <sup>9</sup> , qui est un site de partage de liens . un moyen pour aider les utilisateurs à trouver une information talque les listes de contacts, listes permettant l'accès à l'information et aux ressources associés à ceux-ci. Les membres peuvent soit lire les propositions mises en avant en page d'accueil, soit utiliser la navigation sociale en lisant les informations postées ou recommandées par leur amis [11].

**Tableau 3** : Tableau comparatif entre les réseaux sociaux selon le point de vue des chercheurs

## 1.9. Intérêts des réseaux sociaux

L'écosystème des médias sociaux illustré dans la figure 2 s'organise autour de quatre grands usages : la publication, le partage, la discussion et le réseautage. À chacun de ces usages correspondent des services, certains étant dédiés à une fonction bien particulier (ex. Instagram qui ne sert qu'à publier des photos depuis son Smartphone), tandis que d'autres sont plus versatiles (ex. Tumblr est une plateforme de blogs créée en 2007 à New York qui est difficile à caser).

<sup>9</sup> <http://digg.com>

## Social Media Landscape 2019



Figure 4 : Panorama des médias sociaux 2019 [12]

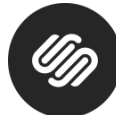
Les quatre grands usages sont complémentaires : les utilisateurs publient des contenus, en partageant d'autres, cela génère des conversations qui leur permettent de développer leur réseau de contacts :

### 1.9.1. La publication

Avec les plateformes d'hébergement de blog tel que :



Wordpress



SquareSpace



Bloogger

### 1.9.2. Les services de partage

#### 1) De photos :



Printerset



500px



Flickr

2) De vidéos :



*Youtube*



*Dailymotion*



*Vimeo*

3) De musique :



*SoundCloud*



*MySpace*



*Spotify*

1.9.3. La discussion

1) Avec les plateformes conversationnelles :



*Quora*



*Github*



*Reddit*

2) Les outils de communication grand public



*Viber*



*Skype*



*Whatsapp*

3) Les applications mobiles de communication :



*Messenger*



*MessageMe*



*Telegram*

1.9.4. Le réseautage

1) Les équivalents asiatiques et russes :



*Mixi*



*RenRen*



*Qzone*

2) Les réseaux sociaux B2B



*LinkedIn*



*Viadeo*



*Xing*

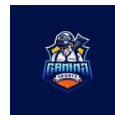
3) Services de jeux en ligne



*ClubPenguin*



*Zynga*



*pubg*



## 1.10. Principaux réseaux sociaux

### 1) Exemple de réseaux sociaux grands publics



Poche d'un milliard de membres d'inscrits. Le principe est d'échanger avec sa communauté d'amis sur tout et n'importe quoi. L'inscription est obligatoire. Permet également de réagir sur les commentaires et news postés par ses amis via le « Like ». Il est devenu fréquent d'entendre le verbe « Liker » dans une conversation. Permet beaucoup d'autres chose : discussion instantanée, envoi de message, identifier des amis sur une photo ... [13]



C'est le « petit » qui monte en flèche. Une plateforme de microblogging. Le fonctionnement est toutefois différent de Facebook : une limitation à **140** caractères par message, la possibilité de suivre d'autres comptes, pas de demande d'invitation, le partage de photo, de vidéo ou d'article se fait par l'utilisation de lien. Sur Twitter, nous n'avons pas un mur mais une timeline ou fil. La notion d'influence y est très présente de par sa logique de suivi d'informations données par d'autres. Avec Twitter, nous lisons les tweets de personnes que nous suivons les Retweeter pour les partager avec vos followers.[13]



Il peut aussi être classé dans les réseaux sociaux puisqu'il permet de partager ses vidéos et de commenter les vidéos postées. Youtube appartient à Google. Il n'est pas nécessaire d'être inscrit pour regarder les vidéos postées. Mais, pour y déposer vidéo et commentaires, une inscription à Youtube est obligatoire. La logique d'amis avec les autres internautes inscrits existe et Youtube permet également de suivre des thèmes pour être informé de leurs actualités. La grande majorité du contenu présent provient de particuliers [1].

## 2) Exemple de réseaux sociaux professionnels



Un réseau professionnel international permet la mise en relation entre des professionnels. Il offre un espace de présentation de ses compétences et expériences qui peuvent être consultable par le public. En Mars 2011, le site revendique plus de 130 millions de membres issus de 170 secteurs d'activités.[13]



C'est une plateforme allemande qui permet de construire et d'agréger son réseau professionnel. Il possède 3,5 millions d'utilisateurs répartis sur plus de 190 pays [13][14].



Il est le pendant français du réseau social LinkedIn. Il permet lui aussi de construire et de gérer son réseau professionnel. Viadeo est plus populaire et plus connu en France que LinkedIn. Il offre à peu près les mêmes possibilités que LinkedIn.[13]

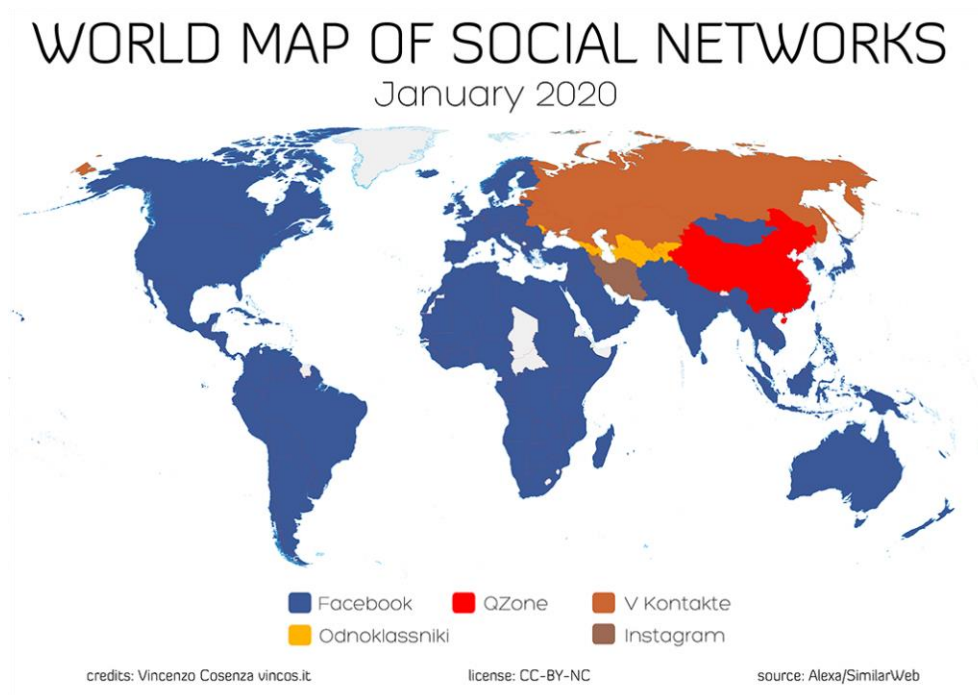
Le tableau suivant décrit les avantages et les inconvénients des réseaux sociaux :

	<b>Facebook</b>	<b>Twitter</b>	<b>LinkedIn</b>
<b>Avantages</b>	Créer une page Facebook simple et la plateforme offre de plus en plus d'outils qui permettent notamment aux entreprises et marques de suivre la progression de leur nombre d'adaptes ainsi que leurs données démographiques.	On peut s'informer en temps réel et faire la même chose pour diffuser une information importante et pertinente. Il n'y a aucun problème à publier fréquemment puisque les tweets se suivent rapidement.	<b>C'est un réseau social professionnel par excellence, qui permet de bâtir un réseau à partir de ses expériences de travail mais également autour de ses intérêts et compétences.</b>
<b>Inconvénients</b>	<b>Facebook demande une interaction avec les adeptes. On ne peut pas diffuser une information sur Facebook et ignorer la réaction des adeptes.</b>	<b>Sur Twitter, on ne peut pas envoyer un message qui contient plus de 140 caractères. Twitter oblige les entreprises à être concises et claires.</b>	<b>L'interaction sur LinkedIn, à part dans certains groupes, est vraiment limitée. Une interface un peu moins accessible ne rend pas les choses faciles.</b>

**Tableau 4** : Tableau comparatif entre les réseaux sociaux

### 1.11. Cartographie des réseaux sociaux numériques

Une cartographie des réseaux sociaux édité en Janvier 2020, montrant les sites de réseautage social les plus populaires par pays, selon [15].



**Figure 5 :** Carte mondiale des réseaux sociaux [16]

Nous remarquons que la très forte implantation de Facebook en Europe Occidentale, sur le continent américain, en Afrique du Nord, ainsi qu'en Australie.

### 1.12. Fonctions et fonctionnalités

Le fonctionnement de ces sites est toujours le même. Pour créer sa page d'accueil, on demande au nouvel arrivant de définir un profil qui constituera la représentation qu'il a de lui-même sous la forme de textes, de photos, de vidéos, de musiques et de liens. Des questionnaires lui imposent parfois de préciser ses goûts, ses affinités, ses opinions, etc.

Facebook permet aux utilisateurs d'enrichir leur profil avec toutes sortes d'applications. Les listes d'amis font le lien entre les profils. Chaque participant au site peut demander à d'autres participants au même site de faire partie de ses amis ; si sa proposition est

acceptée, le nom et une photo (ou un avatar) de l'amis sont affichés sur le site de l'autre. Par « contagion » les participants du site étendent ainsi continuellement leur liste « amis ». Mais qui dit « amis » sur Facebook ne dit pas forcément camarade dans la vraie vie [16].

La liste d'amis ne se résume pas à l'exposition des liens forts d'amitié qui ont une réalité dans la « vraie vie ». Elle permet aussi de mesurer l'audience que chaque participant a auprès d'autres participants, souvent inconnus dans un premier temps, qui partagent les mêmes activités, goûts, opinions, etc. Ces liens sont dits faibles.

Deux pratiques différentes tournent autour de cette distinction : l'accès d'un profil peut être restreint aux « amis », liens forts ou faibles, ou au contraire ouvert à tous, pour faciliter la croissance de la liste.

La troisième par ordre d'importance est la fonction de revue publique ; « Témoignages », « Commentaires », « Panneau d'affichages [17]. Cette fonctionnalité essentielle des réseaux sociaux permet à tout visiteur de déposer des commentaires sur une page d'accueil. Ces commentaires sont lisibles par toutes les personnes autorisées à accéder au profil (Mur sur Facebook). Toutes ces possibilités permettent une médiatisation de soi via un profil détaillé et des contenus diversifiés (statuts, photos, liens), ainsi qu'une socialisation active avec les membres de son entourage au moyen des commentaires.

La plupart des sites sociaux propose l'échange de messages privés entre participants. Ils peuvent aussi offrir des services de partage de photos, de vidéos, et offrir des interactions avec les téléphones mobiles. Ces trois fonctionnalités « profils, listes d'amis, commentaires » constituent la structure de base des réseaux sociaux numériques, même si certains d'entre eux proposent des fonctions complémentaires. Les réseaux sociaux numériques permettent à leurs visiteurs de naviguer d'un ami à l'autre et de correspondre avec quiconque possède un profil visible [17].

Les réseaux sociaux numériques permettent la navigation d'un profil à l'autre et rendent accessible la communication avec tout participant. Le mode d'accès le plus répandu est cependant basé sur l'existence de groupes d'amis préexistants, ayant rejoint le site ensemble, pour conserver ensuite le contact entre eux (élèves d'une classe, d'une école, etc.) et pouvoir communiquer deux à deux tout en faisant partie d'un groupe [16].

### **1.13. Enjeux des réseaux sociaux numériques**

Les enjeux dans le domaine des réseaux sociaux sont multiples tant pour les utilisateurs que pour les acteurs. Du côté des utilisateurs, les réseaux sociaux regroupent toutes sortes

d'individus provenant de tous secteurs. Ainsi, sur le réseau LinkedIn, les utilisateurs du service viennent de tous les horizons : plus de 120 secteurs d'activités sont représentés et aucun secteur ne représente plus de 11% de la base des inscrits. D'après Scott Allen<sup>32</sup> que 11,8% des utilisateurs de LinkedIn sont PDG, 10,2% sont vice-présidents ou directeurs généraux et 1,3% sont membres d'un conseil d'administration. Cela ne vaut pas que pour les réseaux sociaux professionnels. Sur MySpace, par exemple, il est possible de trouver la plupart des grands groupes de musiques qui ont créé leur page MySpace officielle.

De ce fait, il est possible de retrouver quasiment n'importe quelle personne quel que soit sa situation géographique ou son poste au sein de la société [18]. En plus de pouvoir visualiser des informations sur les personnes, les réseaux sociaux offrent l'opportunité d'entrer en contact avec toutes ces personnes. Les liens entre tous les membres d'un réseau sont les profils personnalisés, ce que l'on peut appeler la carte d'identité numérique. Ainsi, les réseaux sociaux permettent de gérer son identité numérique ainsi que sa réputation en ligne.

Lorsque l'utilisateur remplit sa fiche, il a le choix d'y intégrer les informations qu'il souhaite et de cacher celles qu'il estime privées. Les réseaux sociaux lui offrent donc de la visibilité et lui permettent de contrôler son « extimité ». Plus l'internaute arrive à se mettre en avant et se rendre visibles seront les informations choisies. De même selon les informations entrées sur la fiche d'identité, l'utilisateur va pouvoir se mettre en avant devant telles ou telles personnes et dans un cadre bien précis (recherche d'emploi, contact pour développer une entreprise...).

Et en raison de la popularité croissante des réseaux sociaux en ligne (OSN) et énorme quantité de données partagées sensibles, la préservation de la vie privée devient un enjeu majeur pour les utilisateurs OSN. Les services de réseaux sociaux existants sont centralisés et les entreprises qui fournissent les services ont l'autorité exclusive de contrôler toutes les données des utilisateurs. Les utilisateurs ont également peu de contrôle sur comment et ce que les informations les concernant sont présentées à leurs amis en ligne.

La présentation de leur information dépend en grande partie sur la conception du service de réseau social que les utilisateurs utilisent plusieurs architectures décentralisées ont récemment été proposées pour des OSN décentralisée basée sur les réseaux P2P comme une alternative des architectures OSN centralisées.

Il s'agit essentiellement là d'enjeux économiques mais aussi de visibilité. Ainsi les milliers d'utilisateurs inscrits offrent, indirectement, une source de revenu importante. MySpace compte entre 80 et 100 millions de profils créés dont un million rempli en détail.

Il est ainsi possible de les cibler très précisément pour leur proposer du contenu publicitaire en adéquation avec leurs passions et leurs centres d'intérêts. Outre la publicité directe, les réseaux sociaux, grâce à les nombres importants d'utilisateurs, offrent aux grands groupes (audiovisuels, musicaux, informationnels ...) un beau support de diffusion avec un large public qu'ils peuvent toucher de manières très pertinentes.

Ainsi, des chaînes comme celles du groupe Fox ou encore des maisons de disques profitent des réseaux sociaux pour diffuser des contenus adaptés au profil des membres de communautés. De plus, il est très facile d'infiltrer la communauté (création de page personnelle pour un utilisateur fictif, mise en ligne de vidéo marketing...) afin de mettre en place une opération de marketing viral.

Ainsi, il est essentiel pour les acteurs du monde des réseaux sociaux d'accroître le nombre d'utilisateurs qui est au final leur vrai fonds de commerce.

#### **1.14. Problèmes ouverts des réseaux sociaux**

Parmi les problèmes connus et ouverts des réseaux sociaux on peut citer :

##### **1) Problème de détection des communautés**

Le problème de la détection de communauté dans les réseaux est un sujet relativement récent, mais qui a très rapidement conduit à une grande quantité de travaux.

Lorsque l'on étudie des réseaux de terrain, de grande taille et/ou représentent des données complexes tel que les réseaux sociaux, le nombre de groupes que l'on cherche à obtenir ne peut être connu à l'avance. Ce qui a à nous intéresser à un autre problème plus complexe, celui de la détection de communautés.

On peut définir le problème de la manière suivante : pour un réseau donné, comment le décomposer en un nombre inconnu de groupes de nœuds de manière à ce que ces groupes de nœuds satisfasse efficacement le problème de la minimisation des liens inter-communautés, et la maximisation des liens intra-communautés.[19]

##### **2) Problème de sécurité**

En plus d'être l'origine de fuites d'informations nuisibles, les sites de réseaux sociaux peuvent être utilisés comme plateforme d'attaque contre votre système ou pour mener des escroqueries. Voici quelques étapes à suivre afin de vous protéger.

- **Ouverture de session** : Protégez le compte de réseau social avec un mot de passe fort.
- **Chiffrement** : De nombreux sites comme Facebook, Google+ ou Twitter vous permettent de forcer le chiffrement de toutes les communications (appelé HTTPS). Si cela est possible, activez cette option.
- **Applications** : Certains sites de réseaux sociaux vous permettent d'ajouter ou d'installer des applications d'une tierce partie comme des jeux. Des applications malveillantes peuvent utiliser cet accès dans le but d'interagir avec vos amis en votre nom et ainsi voler ou utiliser des données personnelles. Soyez attentifs et installez uniquement les applications provenant de sites connus, assurez-vous également qu'après installation elles soient régulièrement mises à jour [20].

### 3) Problème de détection des rumeurs

Les réseaux sociaux permettent une diffusion massive et rapide des informations. Un des problèmes principaux de ces canaux de communication est l'absence de vérification associée à la viralité de l'information partagée.

Puisque le problème d'Internet et des media sociaux n'est pas la quantité, mais la qualité des informations : quel sens donner à une information ? Mais également quelle valeur donner à une information reçue ou lue sur les media sociaux ?



### 1.15. Conclusion

L'un des outils de communication les plus importants est les réseaux sociaux de toutes sortes. Dans ce chapitre, nous avons présenté les réseaux sociaux numériques, regardé par définition, puis cité certains types d'OSN, leurs intérêts, leurs enjeux.

Le développement rapide des réseaux sociaux a encouragé l'échange d'une grande quantité de données, ce qui a conduit à la propagation de fausse information. Ce qui ouvre la porte à un intérêt pour le classement des informations diffusées sur les sites de réseaux sociaux.

Dans le chapitre suivant, nous allons présenter une synthèse sur les différentes approches de détection des fausses informations ou de la prévention des rumeurs.

# Chapitre 2

---

## Les rumeurs dans les réseaux sociaux

## 2.1. Introduction

« Le rumeur » est un terme utilisé pour représenter des informations fabriquées ou de la propagande comprenant des informations erronées communiquées par le biais des canaux médiatiques traditionnels comme la presse écrite et la télévision ainsi que des canaux médiatiques non traditionnels comme les médias sociaux. Le motif général pour diffuser de telles nouvelles est d'induire les lecteurs en erreur, de nuire à la réputation de toute entité ou de tirer profit du sensationnalisme. Elle est considérée comme l'une des plus grandes menaces pour la démocratie, le libre débat et l'ordre occidental. Dans ce chapitre, nous allons présenter les différentes méthodes de détection des rumeurs et quelques algorithmes et approches.

## 2.2. Définitions d'une rumeur

Dans le besoin d'information, il y a une hiérarchie à l'image de la pyramide des besoins de MASLOW<sup>11</sup>. Cette hiérarchie est définie dans la Figure 6 : Pyramide des besoins d'information. Dans ce besoin, nous abordons donc la question de la fiabilité de l'information. Un besoin stratégique pour la prise de décision. Alors que le besoin de communiquer est déjà satisfait par les moyens technologiques ou non, récents ou non.

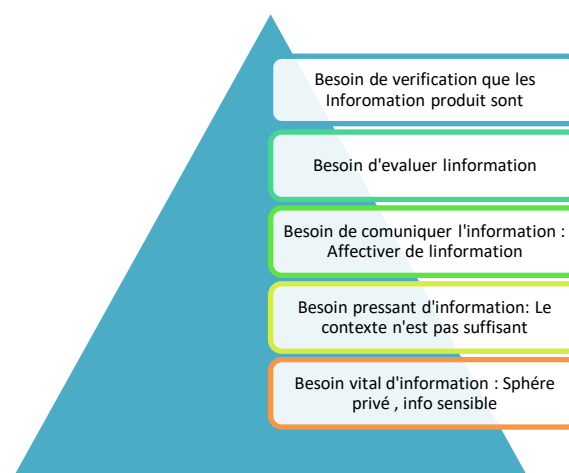


Figure 6 : Pyramide des besoins d'information

<sup>11</sup> Abraham Maslow est un célèbre psychologue américain du XXème siècle. Beaucoup le considère d'a i l leurs comme le père del'approche humaniste. Ils'est notamment fait connaître à partir des années 1940 avec *A Theory of Human Motivation*, parue pour la première fois en 1943.

Les médias sociaux sont donc des supports d'échanges de messages et d'informations. Ces messages et informations ont un contenu. Ce contenu a de la valeur pour l'expéditeur et le destinataire. Cette valeur est plus ou moins importante en fonction de la problématique de chacun.

Mais la valeur de l'information envoyée par un autre membre du même media peut être perçue comme plus importante uniquement sur la base de la confiance faite. Sur les media sociaux, il n'y a pas entre les membres de hiérarchie ni d'autorité supérieur. Tous les membres ont la même quantité de connaissances et d'intérêts à participer à l'intelligence collective, donc à répondre à un besoin d'information exprimé.

Les échanges d'informations avant prise de décision ont donc aussi lieu, en partie, sur les media sociaux. Pour chacune de ces informations, l'origine de l'information, sa source, est évaluée différemment que s'il s'agissait d'une communication de masse.

Actuellement, il n'y a pas eu de définition universelle des fausses informations, même dans le journalisme. Une définition claire et précise permet de jeter des bases solides pour l'analyse de fausses informations et l'évaluation des études connexes.

### 2.2.1. Concepts associés

Les études existantes relient souvent les fausses informations à des termes et concepts tels que *les fausses informations malveillantes* [21] [22], *fausses nouvelles* [23], *informations satiriques* [24], *désinformation* (c.-à-d. Tromperie) [25], *informations satire* [26] et *rumeur* [27]. Sur la base de ces termes et concepts définis, nous pouvons distinguer l'un des autres en fonction de trois caractéristiques :

2. L'authenticité (fausse ou non),
3. L'intention (mauvaise ou non),
4. Si l'information est nouvelles ou non. Le tableau 5 contient les détails

	<b>Authenticité</b>	<b>Intention</b>	<b>Nouvelles ?</b>
<b>Fausse information Malveillantes</b>	Fausse	Mauvaise	Oui
<b>Fausse nouvelles</b>	Fausse	Inconnu	Oui
<b>Information satiriques</b>	Inconnu	Non	Oui
<b>Désinformation</b>	Fausse	Mauvaise	Inconnu
<b>Information satire</b>	Fausse	Inconnu	Inconnu
<b>Rumeur</b>	Inconnu	Inconnu	Inconnu

**Tableau 5** : Tableau de concept associé à la fausse information

### 2.2.2. Définitions

#### **Information**

Pour la différencier entre autres de la « rumeur », une information est définie par Yves- François LE COADIC comme : « une connaissance inscrite (enregistrée) sous forme écrite (imprimé ou numérisée), orale ou audiovisuelle sur un support spatiotemporel [28].

#### **Opinion**

Une opinion est un jugement de valeur individuel sur un objet par une personne dans un instant limité et identifiable. Une opinion n'est donc pas une information car elle est particulièrement applicable et valable pour la personne qui l'exprime. Elle est subjective [28].

#### **Rumeur**

D'après Laurent Gaildraud, une rumeur est avant tout un bruit informel. Il existe, il persiste, il s'évapore. Il repart aussi vite qu'il est arrivé. Il s'agit d'une ou de plusieurs informations qui circulent et qui sont transmises par un vecteur (individus, internet...). Il n'y a pas forcément de source déterminée [29].

## 2.3. Les caractéristiques de la rumeur

Tout d'abord, on peut décrire la rumeur selon 7 caractéristiques réparties en 3 classes [30] :

### 2.3.1. La situation

- La rumeur apparaît dans une situation de crise, mais elle n'est pas toujours le signe d'un dysfonctionnement social,
- Les canaux formels de communication ne véhiculent qu'une information réduite sur certains événements ou aspects de cette situation, c'est à dire que devant la privation d'information les individus vont créer des rumeurs.

### 2.3.2. Le processus de transmission

- La rumeur ne se transmet oralement de personne à personne, par le bouche-à-oreille, mais aussi par les médias (presse, radios, internet). Les canaux sont donc formels (médias) et informels (bouche-à-oreille)
- Cette communication a lieu entre des individus également impliqués dans la situation.

### 2.3.3. Le contenu :

- Le contenu de la rumeur connaît différentes distorsions au cours de son processus de transmission.
- Ce contenu traduit la pensée de désir de la population, elle témoigne de l'exercice d'une pensée sociale, la rumeur devient une sorte d'écran projectif où se déchiffre une dynamique socioaffective.
- Il entretient un rapport avec l'actualité.

## 2.4. Rumeur via les réseaux sociaux

Une rumeur via les réseaux sociaux est une nouvelle qui est diffusée via un site de réseautage.

Dans un laps de temps très court, elle va prendre une ampleur considérable de manière virale. Sur internet, et dans les réseaux sociaux en particulier, quatre caractéristiques sont ajoutées à celles reconnues habituellement à la rumeur [31]:

1. **Instantanéité** : connectés en permanence, grâce à des outils technologiques de plus en plus abordables, les internautes ont très facilement accès à leur messagerie instantanée et aux réseaux sociaux. Cela favorise évidemment la diffusion rapide d'une information.
2. **Décentralisation** : sur internet, la diffusion de l'information n'est pas réglementée par les principes déontologiques du journalisme professionnel comme ce l'est dans les médias d'information traditionnels. Tout un chacun a accès aux outils qui lui permettent d'éditer du contenu en ligne. Nous sommes donc tous des informateurs potentiels. L'information surgit sur nos profils Facebook, nos chronologies Twitter, dans le flux de nos mails ou sur le blog d'une personnalité influente. Nous sommes tous susceptibles d'être les « relayeurs » d'une rumeur.
3. **Visibilité** : les réseaux sociaux sont basés sur le principe d'une liste de contacts qui s'échangent des informations. Relayer les propos « d'amis » dans sa propre sphère est donc le premier pas vers une visibilité accrue par les relais successifs.
4. **Internationalité** : les réseaux « d'amis » ne connaissant pas de frontières, les nouvelles, et donc les rumeurs, font rapidement le tour du globe de mail en mail ou de profil en profil. La confirmation ou l'infirmité d'une rumeur ne parviendra donc pas forcément aux oreilles « électroniques » de chaque internaute. Internet et les réseaux sociaux laissent donc libre cours aux croyances et offrent ainsi une plus longue vie aux rumeurs que lorsque celles-ci se propagent par la bouche à oreille humain.

## 2.5.Représentation de réseau social

Les représentations les plus connues de réseau social sont la représentation nœuds-liens et la représentation matricielle.

### 2.5.1. La représentation nœuds-liens

La première personne à avoir représenté un réseau social est Jacob Levy Moreno au début des années 1930 [32]. Cette représentation a donc l'avantage d'être familière à la majorité des scientifiques. Cependant, elle souffre de problèmes de lisibilité lorsque les réseaux représentés sont soit grands (beaucoup de nœuds), soit denses (beaucoup de liens). Ces problèmes de passage à l'échelle sont d'autant plus importants que les données à analyser sont de plus en plus nombreuses ; ils sont devenus rédhibitoires avec l'apparition des réseaux sociaux en ligne qui sont à la fois grands et denses.

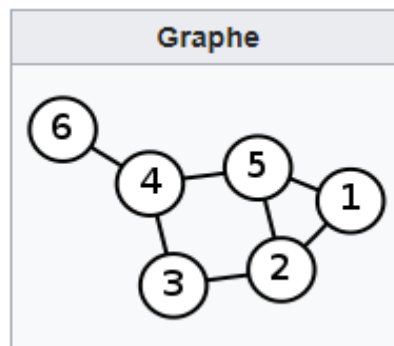


Figure 7 : représentation nœud-lien

### 2.5.2. La représentation matricielle :

Une matrice d'adjacence représente chaque sommet d'un réseau à la fois comme une ligne et comme une colonne. Si deux sommets sont connectés, la case correspondant à l'intersection de la ligne et de la colonne est marquée. Traditionnellement, on utilise une valeur numérique (0 marquant l'absence de connexion, 1 marquant la présence),

Représentation par une matrice d'adjacence					
0	1	0	0	1	0
1	0	1	0	1	0
0	1	0	1	0	0
0	0	1	0	1	1
1	1	0	1	0	0
0	0	0	1	0	0

Figure 8 : représentation matricielle



## 2.6. Méthodes de détection des rumeurs

L'analyse de la véracité des informations est un axe de recherche qui est étudié dans le cadre de plusieurs projets.

La vérification des rumeurs signifie généralement vérifier l'authenticité d'une rumeur. Une rumeur est une information qui peut être vraie ou non, il est donc très important de déterminer si elle est fausse ou authentique avant d'y croire. La figure 9 montre comment la rumeur s'est répandue sur les réseaux sociaux.

Par conséquent, un système de résolution de rumeurs doit subir un ensemble d'étapes, de la détection qu'une nouvelle allégation en circulation est une rumeur, à l'étape ultime de détermination de sa valeur de véracité.

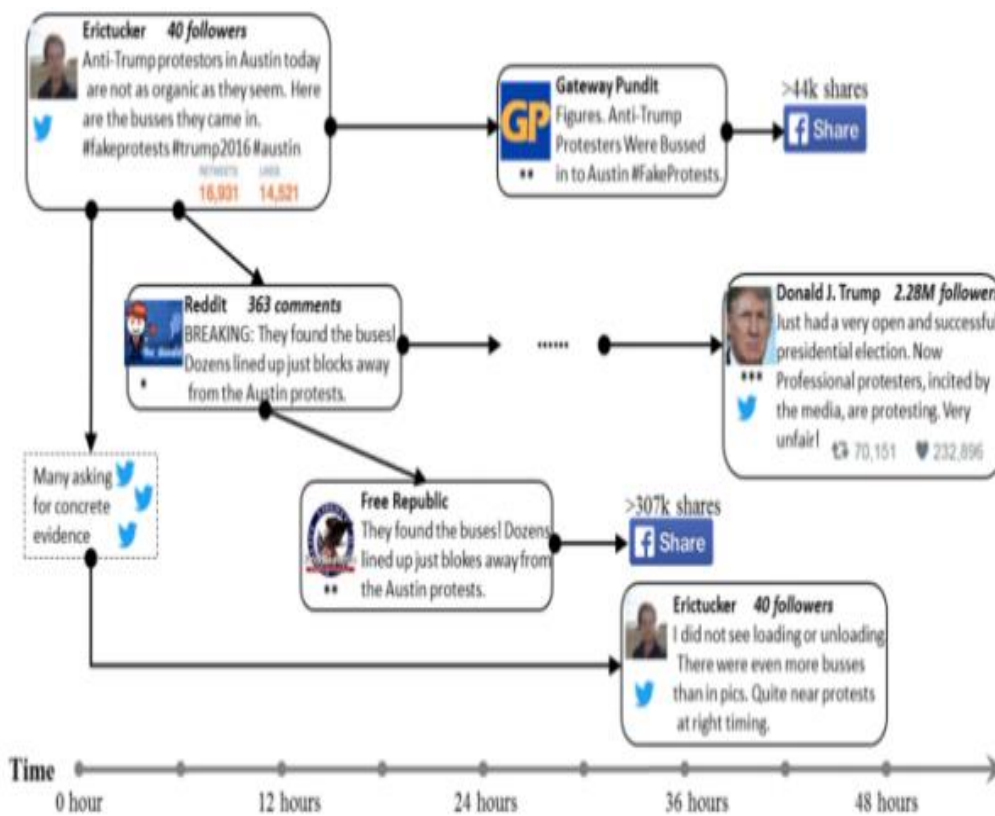


Figure 9 : Exemple d'une rumeur répandue sur les réseaux sociaux [76]

La méthode de résolution des rumeurs comprenant quatre sous-tâches [33]:

- **Détection des rumeurs** : déterminante si une réclamation est une vérification de valeur plutôt que l'expression d'une opinion.
- **Suivi des rumeurs** : rassembler les sources et les opinions sur une rumeur au fur et à mesure qu'elle se déroule.
- **Classification des positions** : déterminer l'attitude des sources ou des utilisateurs envers la véracité de la rumeur.
- **Vérification de la rumeur** : comme l'étape ultime où la valeur de véracité de la rumeur est prédite.

### 2.6.1. Les approches textuelles

Les techniques adoptées pour réaliser les tâches précédents sont classées en trois familles :

#### 1) Basées sur le contenu textuel.

Un ensemble de caractéristiques est tout d'abord extrait du contenu des messages, des propriétés de leur diffusion et des profils des utilisateurs. Ensuite, des algorithmes de classification supervisée permettent de prédire la véracité des messages [34][35][36][37]. Dans cette approche, la qualité des caractéristiques extraites des rumeurs est une étape cruciale pour obtenir des résultats de classification fiables.

#### 2) Basées sur la structure de la propagation ou l'optimisation des graphes

Contrairement aux méthodes de la première famille, ces méthodes évaluent la crédibilité des messages et des événements dans leur ensemble. Ils commencent par la création d'un graphe de crédibilité où les entités impliquées dans la détection de la rumeur, comme les messages et les utilisateurs, constituent les nœuds, et les relations entre ces entités les arêtes. Les arêtes

sont pondérées par l'intensité de la relation. Chaque entité a une valeur initiale de crédibilité, puis ces valeurs de crédibilité sont propagées dans le graphe jusqu'à convergence et évaluation de la crédibilité finale de chaque entité [38] [39][40]. L'inconvénient majeur de cette technique est qu'elle ignore le contenu textuel des messages.

### 3) Basées sur l'apprentissage profond.

Ces méthodes automatiques utilisent essentiellement deux structures de réseaux de neurones : les réseaux de neurones récurrents (RNN), qui modélisent les données textuelles des messages comme des données séquentielles [41][42][43], et les CNN, qui peuvent apprendre la représentation textuelle latente des données de la rumeur et améliorer la précision de la classification [44][45]. Grâce à leur capacité d'apprendre la représentation profonde des données de la rumeur, ces approches améliorent considérablement les performances de prédiction par rapport aux deux familles d'approches précédentes.

En guise de synthèse, nous proposons dans la Figure 10 une cartographie des tâches et des familles de méthodes utilisées pour la détection des fausses informations.

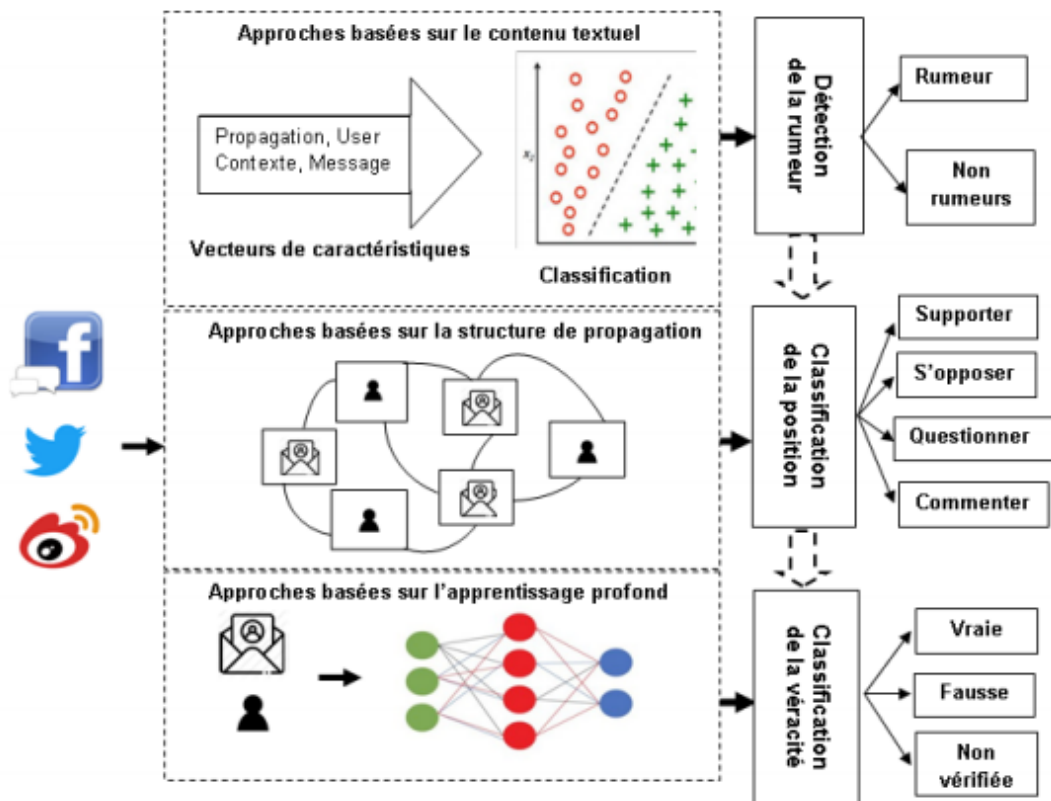


Figure 10 : Les approches de prédiction de la rumeur [75]

### 2.6.2. Approches basées sur les sources

Deux prédictions sont identifiées comme utilisant des sources : *Link-media* (LK) et *Verifying Multimedia Use* (VMU). Alors que les deux approches se basent sur une liste de sources de confiance, la prédiction VMU considère en plus une source de non-confiance. On peut noter que les deux listes de source de confiance n'étant pas identiques, ces dernières peuvent donc se compléter. Une seconde différence se fait quant au choix de la classe à attribuer en cas d'absence de source. Alors que VMU choisit la classe faux, qui est la classe majoritaire de l'ensemble d'apprentissage, la prédiction LK fait le choix de la classe inconnu qui donnera obligatoirement un message mal classé (puisque aucun message ne possède réellement cette classe) mais qui permet une haute précision des messages classés comme vrai ou faux (respectivement 100 % et 92; 97 %) aux dépens du rappel (respectivement 41; 22 % et 87; 47 %).

### 2.6.3. Approches multimédia

Dans cette section, nous étudions les méthodes utilisées pour analyser une composante importante des messages, en l'occurrence leur contenu visuel et plus particulièrement les images, du point de vue de leur véracité.

- **Fausse images dans les réseaux sociaux**

Il existe dans la littérature plusieurs définitions d'une fausse image. Pour [46], il s'agit de toute image attachée à un message qui ne représente pas d'une manière fidèle l'événement auquel elle fait référence. Pour [47], c'est une image attachée à un faux message.

Identifient trois types de fausses images sur les réseaux sociaux : des images anciennes utilisées pour décrire des événements récents, des images délibérément manipulées et des images qui sont utilisées d'une manière imprécise pour décrire un faux événement.

L'analyse de la véracité des images est une tâche difficile qui nécessite de relever de multiples défis. Nous classons les approches qui abordent cette problématique en trois catégories (Figure 11) : celles qui analysent le contenu de l'image pour en détecter les altérations [47] ; celles qui analysent les caractéristiques textuelles des messages pour la classification de la

véracité des images associées ; et celles qui utilisent des informations externes à l'image pour déterminer sa véracité.

Il s'agit dans ce dernier cas de rechercher dans une base de données ou sur le Web des images similaires ou identiques afin de déterminer si l'image étudiée a été modifiée ou détournée.

- **Analyse du contenu de l'image**

Les méthodes décrites dans cette section exploitent uniquement le contenu des images. Il existe principalement deux types d'algorithmes de détection des images contrefaites : les algorithmes actifs et passifs

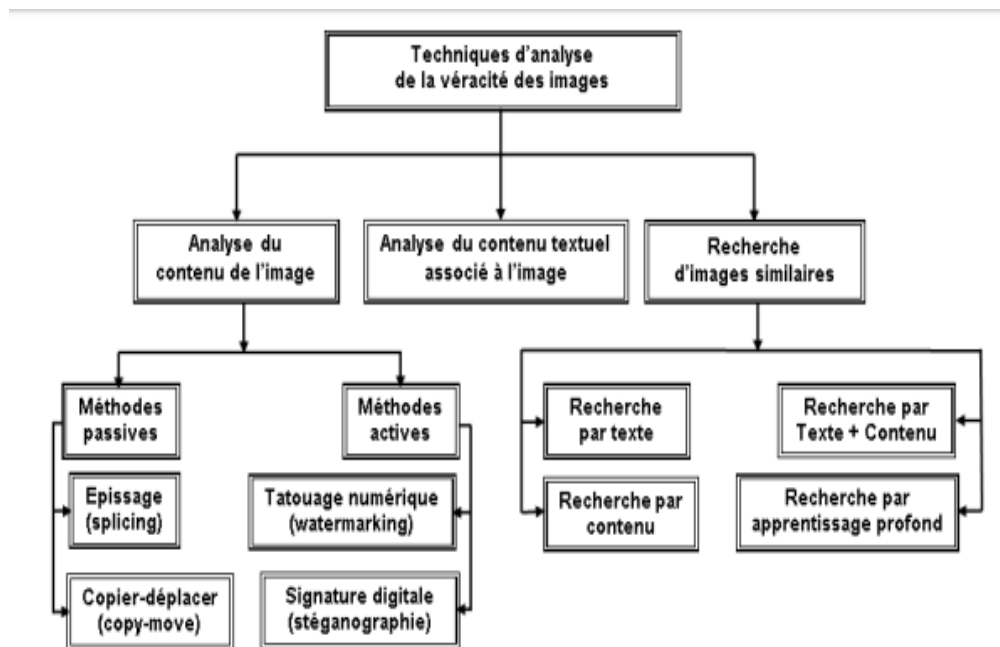


Figure 11 : Les approches de prédiction du rumeur base sur les images [75]

Les algorithmes actifs exploitent une signature, comme un filigrane, la double compression des images JPEG [49] ou la signature laissée par les appareils de capture [49]. Cette signature est mise en correspondance avec la signature de l'image originale pour détecter toute altération. La famille des algorithmes actifs se décompose en deux sous-familles : la signature digitale (stéganographie) qui permet de cacher un message dans une image numérique, et les tatouages numériques (watermarking), qui incluent des informations de copyright ou d'identification dans une image. Les algorithmes passifs exploitent le contenu même de

l'image. Il existe deux familles principales d'algorithmes passifs : les algorithmes de détection d'épissage (splicing) et de copier-déplacer (copymove). Dans une falsification par épissage, des parties de deux images ou plus sont assemblées pour former une nouvelle image. Dans une falsification par copier-déplacer, une partie d'une image est copiée et collée dans une autre partie de la même image.

- **Analyse du contenu textuel associé à une image**

Gupta en [50] propose une méthode d'apprentissage supervisé pour la classification des images dans les messages (tweets). Pour cela, ils développent deux ensembles de caractéristiques extraites du profil des twittos et du contenu des tweets, qui sont exploités par classifieurs, des arbres de décision J48 et un classifieur Bayésien naïf. Ils obtiennent les meilleurs résultats de la classification avec les arbres de décision et les caractéristiques basées sur le contenu des tweets. Par ailleurs, lors d'une investigation relative à la perception des utilisateurs concernant la crédibilité du contenu sur Twitter, [51] ont découvert que les indicateurs importants sur lesquels les utilisateurs jugent la crédibilité sont les informations visibles au premier regard, notamment celles relatives au profil de l'utilisateur (son nom et son image), qui ont un grand impact sur la crédibilité des messages publiés par cet utilisateur.

- **Recherche d'images similaires**

Cette approche utilise des données externes à l'image étudiée par la recherche d'images similaires dans une collection d'images de référence répertoriées comme vraies ou fausses. Une image requête dont on cherche la classe reçoit la classe de l'image la plus similaire de la base de données, si elle existe. Sinon, l'image requête reçoit la classe « inconnue ». Plusieurs techniques sont utilisées pour rechercher les images similaires.

- **Recherche textuelle.** Cette technique utilise les données décrivant l'image, comme des métadonnées (nom, taille, format, titre de la page web, . . . ), des mots clés, des tags ou même un message associé à l'image dans le cas des réseaux sociaux. Les images sont indexées dans une base de données selon ces données textuelles et l'utilisateur formule des requêtes textuelles pour rechercher des [52][53].

- **Recherche par le contenu.** À partir du contenu de l'image étudiée, on calcule une signature et on la stocke dans une base de données indexée offline. La recherche d'images similaires implique le calcul online de la signature de l'image étudiée et sa comparaison avec toutes les signatures stockées dans la base de données. La signature est calculée à partir des caractéristiques visuelles de l'image : les pixels ou toute information dérivée de l'image elle-même comme les couleurs, les textures, les formes ou une combinaison de ces caractéristiques. Une fois les caractéristiques extraites, la comparaison consiste à définir diverses distances entre ces caractéristiques et à définir une mesure de similarité globale entre deux images. On peut alors calculer la similarité entre une image requête et celles de la base de données. Les images résultats sont classées selon leur score de similarité [54].
- **Recherche mixte.** L'exploitation simultanée des contenus textuel et visuel d'une image s'avère une bonne alternative. [55] ont en effet testé trois combinaisons pour concevoir un système de recherche d'images similaires sur le Web : utiliser la recherche textuelle en premier, puis la recherche par le contenu parmi les images résultant de la première recherche ; l'inverse ; et les deux types de recherche simultanément. Les meilleurs résultats, évalués en fonction de la précision et du rappel, sont obtenus avec la première approche. [56] proposent un système qui prend en charge trois modes de recherche : par image requête, par mot-clé et les deux ensembles. Ils fusionnent les caractéristiques textuelles et visuelles de l'image dans un seul vecteur en utilisant des techniques de sac de mots textuels et visuels, respectivement.
- **Recherche par apprentissage profond.** Afin de prédire automatiquement la véracité d'images à partir de leur contenu seulement, [57] calculent les descripteurs des images en utilisant les CNN. La similarité entre l'image requête et les images de la base de référence est calculée par une mesure de similarité cosinus.

## 2.7. Critiques sur les méthodes de détection des fausses informations

L'analyse des différentes approches de détection des fausses informations a permis de tirer la synthèse et les conclusions suivantes :

- Les solutions proposées ont fixé comme seul objectif, la minimisation de l'influence négative en négligeant le temps d'exécution, un paramètre très important dans les réseaux sociaux. La plupart des approches proposées sont des algorithmes qui donnent des résultats qui approchent la valeur optimale mais qui sont longs en termes de temps d'exécution.

### 1) Approche basée sur le contenu textuel

- Cette approche exploite les caractéristiques traditionnelles du contenu des messages, des propriétés de leur diffusion ou du profil de l'utilisateur.
- Elle ignore le contenu visuel lorsqu'il faut statuer sur la véracité des images.

### 2) Approche basée sur l'analyse du contenu des images

- Les algorithmes actifs de détection de falsification ne sont pas utiles, car les techniques de tatouages numériques ou de stéganographie modifient les valeurs de quelques pixels permettant de cacher un message dans l'image sans modifier son aspect visuel.
- Si une telle image est diffusée sur les réseaux sociaux, elle ne suscite aucun intérêt car les modifications sont invisibles à l'œil nu.
- Les techniques basées sur un format particulier d'image (par exemple JPEG) ou sur le type de l'appareil de capture ne sont pas applicables dans le contexte des réseaux sociaux.
- Les formats d'image peuvent subir des altérations lors de la publication du message. L'information relative au type de l'appareil de capture de la photo n'est pas non plus toujours disponible.
- Les techniques de détection des images épissées sont difficilement applicables dans le contexte des réseaux sociaux, en raison des opérations de redimensionnement et de recompressions d'images automatiquement appliquées par les plateformes de réseaux sociaux à tous les contenus téléchargés.
- Les techniques de détection de copier-déplacer, bien qu'elles soient fiables sur des petits jeux de données, ne sont pas susceptibles de passer à. Elles n'ont de fait jamais été



utilisées dans le contexte des médias sociaux.

- Un autre défi pour l'analyse de la véracité des images est que les plateformes de médias sociaux ont tendance à effacer les métadonnées, en particulier les données Exif des images, qui sont des informations utiles pour la détection de l'altération

### 3) Approche basée sur la recherche d'images similaires

- La recherche d'images similaires via des éléments textuels nécessite la description ou l'annotation manuelle des images, qui est une tâche difficile, notamment quand les données sont volumineuses. Elle est également fortement dépendante de la langue utilisée.
- Les images étant riches en contenu et présentant différents niveaux de détail, un même annotateur peut donner, de par sa subjectivité, une description différente de deux images avec le même contenu visuel.
- Les résultats pertinents de recherche pourraient être accompagnés par un grand nombre de résultats non pertinents à cause du niveau de description faible de l'image requête ou/et de l'annotation des images de la base d'images.
- Dans la recherche d'images similaires par le contenu, les caractéristiques visuelles calculées sont dites de bas niveau, car elles sont très proches du signal et ne comportent aucune information sémantique.
- Les difficultés liées au calcul des caractéristiques des images et leur stockage peuvent survenir, particulièrement si la recherche des images similaires est effectuée sur le Web qui contient un nombre illimité d'images.
- L'utilisation de techniques d'apprentissage profond pour la recherche d'images similaires pourrait être une solution au problème de l'écart sémantique qui existe entre les caractéristiques visuelles de bas niveau capturées par des machines et les concepts sémantiques de haut niveau perçus par l'être humain.
- Des expériences poussées menées par [58] ont donné des résultats encourageants par rapport aux approches précédentes. Cette section a permis de présenter les travaux relatifs à l'analyse de la véracité des images.

## 2.8. Synthèse et discussion

Le tableau 6, résume les travaux cités précédemment. Ou les signe ‘+’ signifie que le résultat est acceptable , ‘++’ une résultat moyenne et le ‘+++’ signifie la meilleur résultat.

Approche	Contribution	Techniques	Sémantique préservé	Résultat	
				Performance de prédiction	Temps d'exécution
Basé texte	Base contenue du texte	Extraction des caractéristiques, Algorithme de classification supervisé	Oui	+	++
	Base sur la structure de propagation	Graphe de caractéristique	Non	+	+
	Base apprentissage profond	Apprentissage profond (RNN, CNN)	Oui	+++	+++
Basé source	LK	Une liste de sources de confiance	Non	+++	+
	VMU	Une liste de sources de confiance	Non	++	+
Basé multimédia	Analyse de contenu d'image	Méthodes passives (Splicing, Cryptographie)	Non	++	+
		Méthodes actives (Watermarking, Stéganographie)	Non	++	+
	Analyse de contenu textuelle associée à une image	Recherche texte et contenu	Non	++	+
	Recherche d'image similaire	Recherche par Apprentissage profond (CNN)	Oui	+++	+++

**Tableau 6 :** Synthèse des méthodes de détection de fausses informations dans les réseaux sociaux

## 2.9. Conclusion

L'influence dans les réseaux sociaux est un sujet en plein essor. La problématique de détection de fausse information, par ses applications et ses avantages apportés aux réseaux sociaux, suscite l'intérêt de la communauté de recherche. Dans ce chapitre nous avons présenté la rumeur et les différentes approches de détection et de minimisation, ensuite nous avons cités les quelques approches de détection de rumeur, en particulier leurs principes de fonctionnement.

# Chapitre 3

---

## L'apprentissage profond & Le traitement de langage naturelle

### **3.1.Introduction**

Le « Deep Learning » est une solution pour chaque sujet de classification, de régression ou même d'exploration, et cela est dû à ses résultats satisfaisants pour de nombreuses recherches, telles que le traitement d'images et le traitement du langage naturel.

Dans ce chapitre, nous présentons les groupes de l'intelligence artificielle à savoir l'apprentissage automatique et l'apprentissage profond. Nous nous soumettons aux concepts du traitement du langage naturel comme deuxième point du chapitre, et comme point final, ce chapitre présente le rôle de l'apprentissage profond dans le traitement du langage naturel.

### **3.2.L'intelligence artificielle**

L'intelligence artificielle (IA, ou AI en anglais pour *Artificial Intelligence*) vise à mimer le fonctionnement du cerveau humain, ou du moins sa logique lorsqu'il s'agit de prendre des décisions. L'IA consiste à mettre en œuvre un certain nombre de techniques visant à permettre aux machines d'imiter une forme d'intelligence réelle.

L'IA englobe des différents sous-domaines tels que les règles métiers, Machine Learning, Deep Learning, Traitement du langage Naturel ...etc

### **3.3.Apprentissage automatique**

Selon [59] L'apprentissage automatique (machine Learning) est le champ de l'IA qui permet à une machine d'apprendre. C'est-à-dire, d'améliorer progressivement ses performances sur une tâche spécifique en se basant sur des données sans être explicitement programmé.

L'apprentissage automatique passe par les réseaux de neurones conçus pour imiter les capacités de prise de décision de l'homme. Nous devons l'appliquer pour résoudre tout problème qui nécessite une réflexion, qu'il soit humain ou artificiel. L'apprentissage automatique est généralement divisé en :

### a) Apprentissage supervisé

#### Définition

La forme la plus commune d'apprentissage automatique est l'apprentissage supervisé. L'apprentissage supervisé est une méthode permettant de transformer un jeu de données en un autre, le programme est formé sur un ensemble prédéfini d'exemples de formation, ce qui facilite ensuite sa capacité à parvenir à une conclusion précise lorsque de nouvelles données sont fournies [60] [61]. Les algorithmes de classification supervisée de ML sont : Forêt aléatoire, Arbres de décision, Régression logistique, et le plus connu c'est Machines à vecteurs de support SVM.

### b) Apprentissage non supervisé

#### Définition

L'apprentissage non supervisé, encore appelé apprentissage à partir d'observations, partage une propriété commune avec l'apprentissage supervisé : il transforme un jeu de données en un autre. Mais l'ensemble de données dans lequel il se transforme n'est pas connu ou compris auparavant. Contrairement à l'apprentissage supervisé sera quant à lui alimenté uniquement par des exemples, et créera lui-même les classes qui lui semblent les plus judicieuses (clustering) ou des règles d'associations (algorithmes Apriori). L'algorithme K-moyen (K-means) permet de comprendre facilement le concept de classification non supervisée [62].

## 3.4.L'apprentissage profond

### 3.4.1 Définition

D'après les travaux de Deng L et al [63], le « Deep Learning » est une classe de techniques d'apprentissage automatique appartient au domaine de « Machine Learning » dans lesquelles de multiples couches de traitement de calcul itératif dans des architectures hiérarchisées supervisées sont exploitées pour les algorithmes d'apprentissage non supervisé pour des tâches d'analyse et de classification. L'apprentissage profond consiste essentiellement à calculer des caractéristiques hiérarchiques des paramètres des réseaux de neurones artificiels pour les représentations vectorielles des données d'observation ou d'entrées. La famille des méthodes d'apprentissage en profondeur s'enrichit de plus en plus, englobant celles des réseaux de neurones, des modèles probabilistes hiérarchiques, ainsi que de nombreux algorithmes d'apprentissage des fonctionnalités supervisés et non supervisés.

### 3.4.2 L'apprentissage profond VS l'apprentissage automatique

Deux caractéristiques principales qui distinguent "l'apprentissage profonde" de "l'apprentissage automatique" qui sont :

#### A. La phase d'extraction des caractéristiques

La plupart des fonctionnalités de l'application sont requises par un expert, puis codées manuellement par domaine et type de données dans la Machine Learning, contrairement au Deep Learning dont les algorithmes tentent de connaître les fonctionnalités de haut niveau des données. C'est une partie très distincte de l'apprentissage en profondeur et une étape majeure avant l'apprentissage automatique traditionnel. Par conséquent, l'apprentissage en profondeur réduit la tâche de développer un nouvel extracteur de fonctionnalités pour chaque problème.

#### B. La performance

Les performances sont associées à un volume de données accru. Lorsque les données sont petites, les performances des algorithmes d'apprentissage approfondi donnent de mauvais résultats car elles nécessitent une grande quantité de données pour bien les comprendre. L'inverse se produit avec les algorithmes d'apprentissage automatique traditionnels où les résultats prévalent à mesure que la quantité de données augmente. La figure ci-dessous résume ce fait.

### 3.4.3 Principe de fonctionnement

Le réseau de neurone se compose de trois couches importantes comme indiqué sur la figure :

- ✓ La couche d'entrée ;
- ✓ La / Les couches cachées ;
- ✓ La couche de sortie.

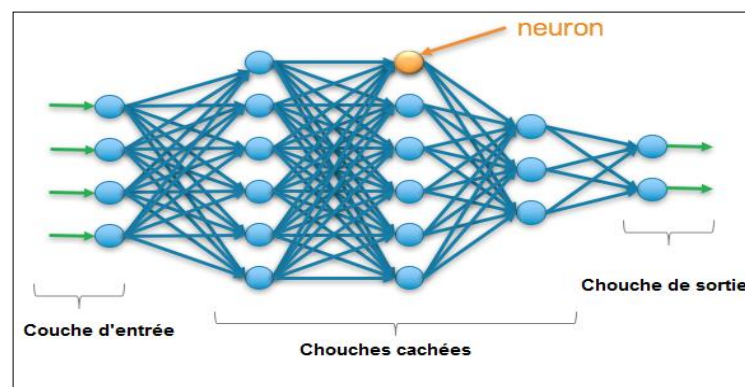
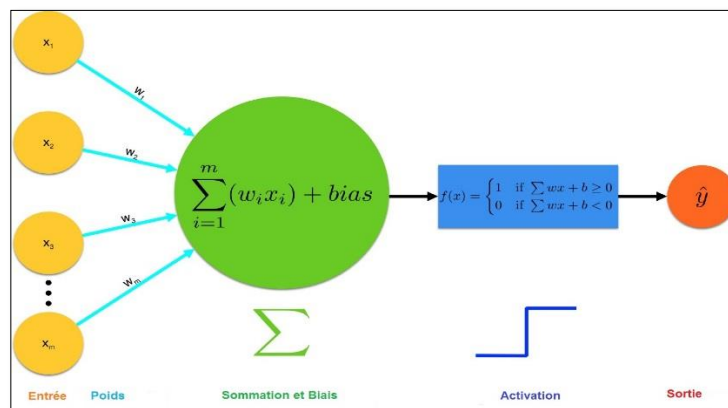


Figure 12: Topologie de réseau de neurones [64]

Le sens du mot "profond" fait référence au nombre de couches cachées dans le réseau neuronal où les réseaux neuronaux traditionnels ne contiennent que un petit nombre de couches cachés ( 2 ou 3 ) , tandis que les réseaux profonds peuvent compter jusqu'à 150 couches cachées [64].

Le réseau de neurones est composé d'un ensemble de nœuds (neurones) connectés via des liaisons dirigées (flèche), chaque flèche représente une liaison entre la sortie d'un neurone et l'entrée d'un autre, Chaque flèche porte un poids (W) , reflétant son importance, chaque nœud étant une unité de traitement qui exécute une fonction de nœud statique sur son signal entrant pour générer une sortie de nœud unique[65].



**Figure 13 :** Topologie de réseau de neurones profond.[65]

Les valeurs d'entrée, ou en d'autres termes, nos données sous-jacentes, sont transmises via ce « réseau » de couches masquées jusqu'à ce qu'elles convergent vers la couche de sortie. La couche en sortie correspond à notre prédiction : il peut s'agir d'un nœud (on dit une classification binaire) ou de quelques nœuds s'il s'agit d'un problème de classification multi-classe.

La forme à l'intérieur des neurones dans les couches centrales représente une fonction d'activation qui peut être un Cube ,Elu , Hardsigmoid , Hardtanh , Identity , Leakyrelu ,Rationaltanh ,Relu ,RRelu ,Sigmoid , Softmax , Softplus ,Softsign , Tanh.

On va détaillés quelques fonctions les plus utilisées :

- **Fonction Sigmoid**
- **Fonction Tanh**



- **Fonction ReLu**
- **Fonction Softmax**

### **Back propagation**

Une partie importante des réseaux de neurones, y compris des architectures profondes modernes, est la propagation en arrière des erreurs à travers un réseau. Chaque exemple d'entrées faire une propagation et back propagation afin de mettre à jour les poids utilisés par les neurones plus proches de l'entrée.

#### **3.4.4 Modèles d'apprentissage profond**

Dans cette section on va présenter quelques types de modèle ou bien des algorithmes d'apprentissage profond.

Dans les années 1980, la plupart des réseaux de neurones ne formaient qu'une seule couche en raison du coût de calcul et de la disponibilité des données. De nos jours, nous pouvons nous permettre d'avoir plus de couches cachées dans nos réseaux de neurones, d'où le surnom d'apprentissage profond. Les différents types de réseaux de neurones disponibles à l'utilisation ont également proliféré, des modèles tels que les réseaux de neurones Convolutionnels (CNN), les réseaux de neurones récurrents (RNN) et (LSTM)

#### **Les réseaux de neurones Convolutionnels (CNN)**

D'après les travaux de *Yin W et al* [66], les réseaux de neurones convolutif « CNN » sont des types spécifiques de réseaux de neurones artificiels qui utilise des perceptrons et un algorithme d'unité d'apprentissage automatique pour l'apprentissage supervisé, permettant d'analyser des données. Les CNN s'appliquent beaucoup plus au traitement d'images, aussi au traitement du langage naturel « NLP » et à d'autres types de tâches cognitives. Comme d'autres types de réseaux de neurones artificiels, un réseau de neurones convolutionnel comporte une couche d'entrée, une couche de sortie et une ou plusieurs couches cachées (voir figure 11). Certaines de ces couches sont convolutionnelles et utilisent un modèle mathématique et algébrique pour transmettre les résultats aux couches successives.

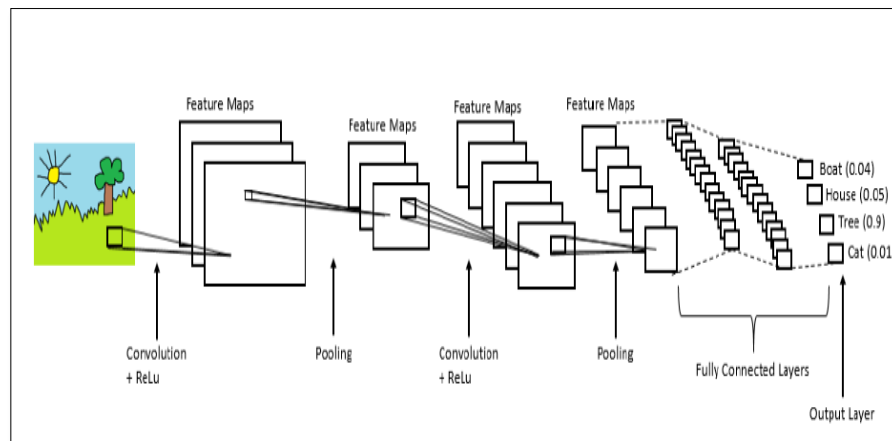


Figure 14: Architecture de réseau de neurone convolucional [66]

Les réseaux « CNN » sont des exemples typiques de la technique d'apprentissage profond, où un modèle plus sophistiqué accélère l'évolution de l'intelligence artificielle en proposant des systèmes simulant différents types d'activités biologiques du cerveau humain.

Généralement les réseaux « CNN » sont composés de quatre principales couches :

- **La couche de convolution « CONV »** qui traite les données d'entrées ou les résultats des couches intermédiaires ou cachées.
- **La couche de Pooling « POOL »**, qui permet de compresser l'information en réduisant la taille de l'image intermédiaire.
- **La couche d'activation**, c'est des fonctions de rectification linéaire souvent appelée par « ReLU » qui permet la mise à jour des paramètres de la couche actuelle ainsi que les couches précédentes
- **La couche entièrement connectée « Fully Connected »**, qui est une couche de type perceptron

### 3.4.5 Domaines d'application

L'éducation en profondeur a été développée dans une large mesure ; aujourd'hui, les entreprises utilisent de nombreuses applications pratiques du domaine de l'apprentissage en profondeur, notamment :

- **Reconnaissance de la parole**
- **Reconnaissance d'Image**
- **Systèmes de recommandation**
- **Traitement du langage naturel**

### 3.5. Le traitement du langage naturel

#### 3.5.1 Définition

Selon [68], le traitement automatique du langage naturel (NLP) est l'élément qui fait le lien entre la conversation humaine et la compréhension programmée par ordinateur (quels que soient les humains qui parlent, quelle langue est parlée et la façon dont ils parlent, peut-être grammaticalement). Le composant PNL permet à l'ordinateur d'interpréter le langage humain vaste et compliqué, de comprendre ce qui est dit, de tout traiter, de refléter ce qui est demandé et de réagir efficacement, comme le font les humains

#### 3.5.2 Objectif

L'objectif du traitement du langage naturel (PNL) est de rendre le langage humain compréhensible pour la machine en convertissant le langage naturel en une représentation formelle facilement manipulable par les ordinateurs pour étudier les problèmes de base, qui sont bien adaptés à la modélisation de données textuelles afin d'en extraire des informations. .

### 3.6. L'apprentissage profond et le traitement de langage naturel

Dans les techniques classiques de NLP , il faut prétraité les données dans les premiers stades avant de générer des fonctionnalités à partir des données. Après, nous utilisons des fonctionnalités artisanales. Nous alimentons ces fonctionnalités en entrée de l'algorithme d'apprentissage automatique (ML) et formons le modèle. Nous vérifierons la précision, et si la précision n'est pas bonne, nous optimiserons certains des paramètres de l'algorithme et essayerons de générer un résultat plus précis. Selon l'application NLP, vous pouvez inclure le module qui détecte la langue et génère ensuite des fonctionnalités. Dans les techniques de Deep Learning pour la PNL, nous effectuons un prétraitement de base sur les données dont nous disposons. Ensuite, nous convertissons nos données d'entrée de texte en une forme de vecteurs denses. Pour générer les vecteurs denses, nous utiliserons des techniques d'intégration de mots tels que word2vec, GloVe, etc., et alimenterons ces vecteurs denses en intégration au DNN. Ici, nous n'utilisons pas de fonctionnalités artisanales mais différents types de DNN selon l'application NLP.. [69]

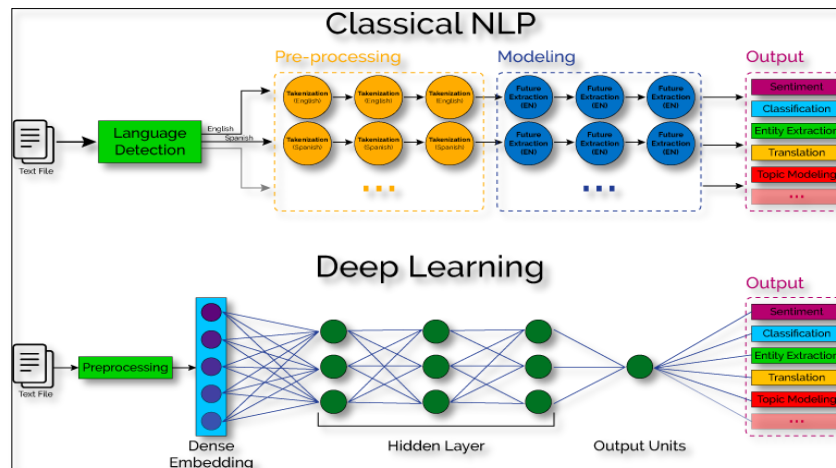


Figure 15 : application de NLP Classique vs NLP avec le deep learning [69]

### 3.6.1 Les techniques de vectorisation

Comme cela est bien connu, la machine ne comprend pas les langages naturels. Le texte doit être convertie en nombres, ce processus appelé « Vectorisation ». Il existe plusieurs techniques de vectorisation chaque une a son utilisation :

#### 1. Bag Of Words

Le modèle du sac de mots est une représentation simple utilisée en PNL. Dans ce modèle, un texte est représenté comme le sac de ses mots, sans tenir compte de la grammaire ni même de l'ordre des mots mais en gardant la multiplicité.

#### 2. TF-IDF

TF-IDF est l'abréviation de terme fréquence – fréquence du document inverse. Il est conçu pour refléter l'importance d'un mot pour un document dans une collection ou un corpus. La valeur TF-IDF augmente proportionnellement au nombre de fois qu'un mot apparaît dans le document et est compensée par le nombre de documents dans le corpus qui contiennent le mot, ce qui permet d'ajuster le fait que certains mots apparaissent plus fréquemment en général.

Pour un terme « i » dans un document « j » :

$$W_{i,j} = tf_{i,j} * \log\left(\frac{N}{df_i}\right)$$

Où :

$tf_{i,j}$  est nombre d'occurrence de i dans j

$df_i$  est le nombre de document contenant i ;

$N$  est le nombre total des documents.

### 3. Word Embedding

La technique de « Word Embedding » qui désigne la représentation distributionnelle vectorielle des mots dans les documents. Ces techniques reposent sur des méthodes pour représenter la distribution des mots avec une représentation vectorielle de chaque mot avec un vecteur de taille fixe selon leurs contextes.

Les techniques de « Word Embedding » offrent la possibilité de création des modèles qui représentent les documents et ces termes avec un modèle vectoriel de taille fixe qui prend en charge les relations sémantiques et syntaxiques entre ces termes et ressource. En effet, nous trouvons plusieurs techniques automatiques de capture de contexte, parmi ces techniques nous trouvons : word2vec, Glove.

- **Word2Vec**

Word2vec est un groupe de modèles associés qui sont utilisés pour produire des incorporations de mots. Word2vec prend comme entrée un grand corpus de texte et produit un espace vectoriel, généralement de plusieurs centaines de dimensions, chaque mot unique du corpus se voyant attribuer un vecteur correspondant dans l'espace.

- **Glove**

Comme word2vec, Glove est une autre méthode d'intégration de mots couramment utilisée. Glove est l'abréviation de factorisation matricielle globale, c'est le processus d'utilisation des méthodes de factorisation matricielle de l'algèbre linéaire pour effectuer une réduction de rang sur une grande matrice terme-fréquence.

Et c'est ce que fait une partie d'extraction de fonctionnalités du pipeline NLP.

### 3.7. Conclusion

L'apprentissage en profondeur a surmonté les méthodes traditionnelles de manipulation d'images et de texte en rendant les systèmes plus rapides et plus faciles. Dans des recherches récentes, l'apprentissage en profondeur a été utilisé dans le traitement du langage naturel dans des défient problèmes, et il a mis en évidence plusieurs résultats satisfaisants. Dans ce chapitre, nous avons présenté les quatre techniques à savoir l'intelligence artificiel, l'apprentissage automatique, la technique d'apprentissage en profondeur et ses différents algorithmes, ainsi le traitement du langage naturel, et nous concluons le chapitre avec les techniques d'apprentissage en profondeur utilisées dans le traitement du langage naturel.

# Chapitre 4

---

## L'approche proposée

---

## **4.1. Introduction**

Nous rappelons que le but principal de notre travail est la détection des fausses informations, pour cette fin nous avons choisi la détection à partir des textes ce qui est connu par le traitement de langage pour analyser et classifier les informations.

Les fausses nouvelles sont de la désinformation ou des nouvelles manipulées qui se propagent sur les réseaux sociaux dans le but d'induire en erreur l'opinion publique, de nuire à une personne ou de discréditer une institution.

Avec la propagation de fausses nouvelles ces derniers temps et le manque de systèmes pour détecter les rumeurs, des méthodes intelligentes sont nécessaires pour révéler les dernières nouvelles. La détection de fausses nouvelles vise à aider les utilisateurs à découvrir différents types de nouvelles fabriquées.

## **4.2. L'architecture globale**

La figure suivante illustre notre architecture globale :



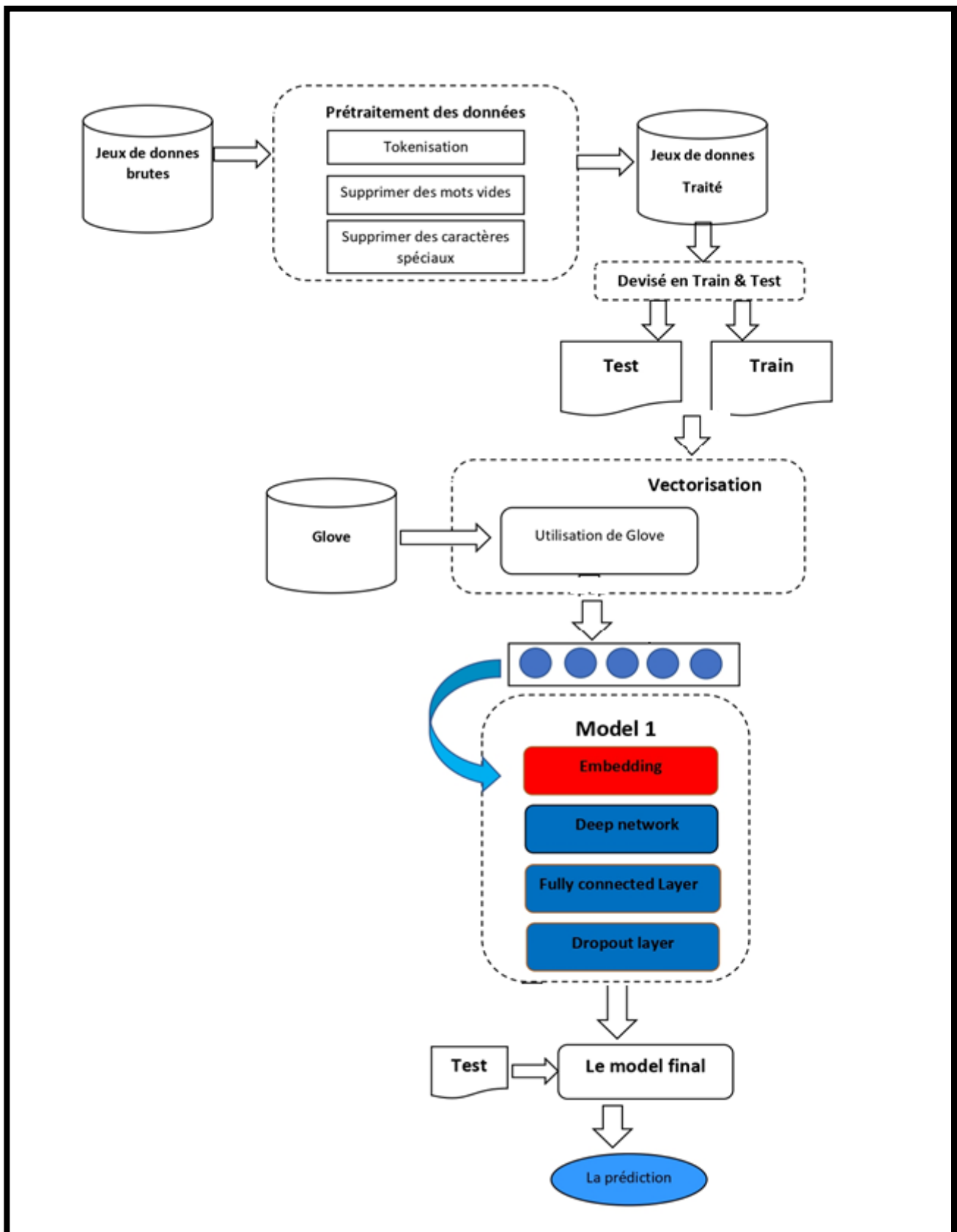


Figure 16 : L'architecture global du système

Les étapes du système proposé sont considérées comme des prétraitements du texte des (Tokenisation, Les mots vides ...), d'autre sont considérés comme étant étapes de division de données et de représentation distribuée (vectorisation) pour le modèle de traitement. Ci-dessous nous décrivons en détail cette architecture.

### 4.3. Le prétraitement de données

L'étape de prétraitement englobe différentes techniques de traitement de texte tel que la normalisation, la tokenisation, la lemmatisation... Dans notre cas on s'intéresse seulement par l'application de tokenisation et l'élimination des mots vides.

#### 1. Tokenisation

La tokenisation consiste essentiellement à diviser une phrase, un paragraphe ou un document texte entier en unités significatives plus petites, telles que des mots ou des termes individuels. Chacune de ces petites unités est appelée jeton. Dans la figure ci-dessous nous présentons

```
# tokenization --
from nltk.tokenize import RegexpTokenizer
tokenizer = RegexpTokenizer(r'\w+')
df_clean['text']=df_clean['text'].str.lower()
df_clean['tokenisation']=df_clean['text'].str.split()
df_clean[['text','label','filtered']]
```

`/usr/local/lib/python3.6/dist-packages/ipykernel_launcher.py:4: SettingWithCopyWarning: A value is trying to be set on a copy of a slice from a DataFrame. Try using .loc[row_indexer,col_indexer] = value instead`

See the caveats in the documentation: [https://pandas.pydata.org/pandas-docs/stable/user\\_guide/indexing.html](https://pandas.pydata.org/pandas-docs/stable/user_guide/indexing.html) after removing the cwd from sys.path.

	text	label	filtered
0	house dem aide: we didn't even see comey's let...	1	[house, dem, aide:, we, didn't, even, see, com...
1	ever get the feeling your life circles the rou...	0	[ever, get, the, feeling, your, life, circles,...
2	why the truth might get you fired october 29, ...	1	[why, the, truth, might, get, you, fired, octo...
3	videos 15 civilians killed in single us airstr...	1	[videos, 15, civilians, killed, in, single, us...
4	print \nan iranian woman has been sentenced to...	1	[print, an, iranian, woman, has, been, sentenc...
...	...	...	...

Figure 17 : Code de tokenisation en python

## 2. Élimination des mots vides

Les mots vides sont les mots dans n'importe quelle langue qui n'ajoutent pas beaucoup de sens à une phrase. Ils peuvent être ignorés en toute sécurité sans sacrifier le sens de la phrase. Pour certains moteurs de recherche, ce sont quelques-uns des mots de fonction courts les plus courants, par exemple pour la langue anglaise on trouve :

```
import nltk
nltk.download('stopwords')
load_stop_words
stop_words = stopwords.words('english')
print(stop_words[:15])
```

↳ ['i', 'me', 'my', 'myself', 'we', 'our', 'ours', 'ourselves', 'you', "you're", "you've", "you'll", "you'd", 'your', 'yours']

Figure 18 : Les quinze premiers mots vides de la langue anglais

Dans ce cas, les mots vides peuvent poser des problèmes lors de la recherche de phrases qui les incluent, le code suivant exprime la phase d'élimination des mots vides :

```
#####suppression des mots vides #####
df_clean['clean'] = df_clean['text'].astype('str')
df_clean['clean'] = df_clean['clean'].apply(lambda x: ' '.join([word for word in x.split() if word not in (stop)]))
df_clean['tokenisation']=df_clean['text'].str.split()
df_clean["tokens"] = df_clean["text"].apply(tokenizer.tokenize)
# delete Stop Words
#print('Time to tokenize everything: {} mins'.format(round((time() - t) / 60, 2)))
df_clean[['text', 'label', 'clean']]
```

↳ /usr/local/lib/python3.6/dist-packages/ipykernel\_launcher.py:3: SettingWithCopyWarning:  
A value is trying to be set on a copy of a slice from a DataFrame.  
Try using .loc[row\_indexer,col\_indexer] = value instead

See the caveats in the documentation: [https://pandas.pydata.org/pandas-docs/stable/user\\_guide/indexing.html#returning](https://pandas.pydata.org/pandas-docs/stable/user_guide/indexing.html#returning)  
This is separate from the ipykernel package so we can avoid doing imports until

/usr/local/lib/python3.6/dist-packages/ipykernel\_launcher.py:4: SettingWithCopyWarning:  
A value is trying to be set on a copy of a slice from a DataFrame.  
Try using .loc[row\_indexer,col\_indexer] = value instead

See the caveats in the documentation: [https://pandas.pydata.org/pandas-docs/stable/user\\_guide/indexing.html#returning](https://pandas.pydata.org/pandas-docs/stable/user_guide/indexing.html#returning)  
after removing the cwd from sys.path.

/usr/local/lib/python3.6/dist-packages/ipykernel\_launcher.py:5: SettingWithCopyWarning:  
A value is trying to be set on a copy of a slice from a DataFrame.  
Try using .loc[row\_indexer,col\_indexer] = value instead

See the caveats in the documentation: [https://pandas.pydata.org/pandas-docs/stable/user\\_guide/indexing.html#returning](https://pandas.pydata.org/pandas-docs/stable/user_guide/indexing.html#returning)  
"""

	text	label	clean
0	house dem aide: we didn't even see comey's let...	1	house dem aide: didn't even see comey's letter...
1	ever get the feeling your life circles the rou...	0	ever get feeling life circles roundabout rathe...
2	why the truth might get you fired october 29, ...	1	truth might get fired october 29, 2016 tension...
3	videos 15 civilians killed in single us airst...	1	videos 15 civilians killed single us airstrike...
4	print \nan iranian woman has been sentenced to...	1	print iranian woman sentenced six years prison...
...	...	...	...

Figure 19 : Code d'élimination des mots vides en python

### 3. Division des données

Dans la phase de division des données nous avons choisi de souffler les données et les diviser en 3 parties :

1. **Train** : Partie de jeux de données (échantillon) pour la phase l'apprentissage dans les réseaux de neurones.
2. **Test** : Partie de jeux de données (échantillon) utilisées uniquement pour évaluer les performances d'un modèle final.
3. **Validation** : Partie de jeux de données (échantillon) utilisé pour ajuster les paramètres d'un classificateur et fournir une évaluation impartiale d'un modèle.

```
#####Division : Train test validation Split#####  
from sklearn.model_selection import train_test_split  
indices = np.arange(data.shape[0])  
np.random.shuffle(indices)  
data = data[indices]  
#print(data)  
labels = labels[indices]  
x_train, x_test, y_train, y_test = train_test_split(data, labels, test_size=0.25, random_state=42) # 25/75  
x_test, x_val, y_test, y_val = train_test_split( data, labels, test_size=0.50, random_state=42)
```

Figure 20 : Division de données en python

### 4.Représentation distribuée

Pour utiliser le texte dans les algorithmes d'apprentissage automatique, nous devons les convertir en une représentation numérique. Est comme il est mentionné dans le chapitre précédent, pour extraire les caractéristiques numériques du contenu textuel en a choisi d'utiliser la technique d'indexation « Word Embedding ».

#### ❖ Utilisation de Glove

Il existe une autre approche pour construire le modèle de clarification de la rumeur consiste à apprendre séparément les « word embedding », puis passer à la couche embedding. Cette approche permet également d'utiliser n'importe quel word embedding prétraité et permet également de gagner du temps dans la formation du modèle de classification.

A cette raison nous utiliserons l'implémentation Gensim de Glove. La première étape consiste à préparer le corpus de texte pour l'apprentissage de l'intégration en créant des jetons de mots, en supprimant la ponctuation, en supprimant les mots vides, etc. L'algorithme Glove traite les documents phrase par phrase.

#### 4.4. Le modèle proposé

Dans notre proposition nous avons créé un modèle avec « la couche embedding », où on a utilisé une Word embedding prétraité 'Glove' avec le même algorithme CNN-1D.

Dans ce qui suit on présente l'architecture de notre modèle :

Le modèle que nous présentons dans la figure 33 est composé d'une couche de embedding, deux couches de convolution et deux couches de maxpooling et deux couches de sortie.

Les informations en entrée sont de nombre maximum des mots 1000, et la dimension de la représentation vectoriel 100.

$$\begin{aligned} \text{Le nombre de paramètre} &= (\text{vocabulaire}) * (\text{dimension}) + (\text{biais}). \\ \text{Param} &= 253853 * 100 + 100 = 25385400 \end{aligned}$$

L'information passe d'abord à la première couche de convolution. Cette couche est composée de 5 filtres de taille 128

$$\begin{aligned} \text{Le nombre de paramètre} &= \text{output} * \text{filtre} * \text{input dimension} + \text{biais}. \\ \text{Param} &= 128 * 5 * 100 + 128 = 64128. \end{aligned}$$

Ensuite on applique Maxpooling pour réduire la taille de l'information ainsi la quantité de paramètres et de calcul. À la sortie de cette couche, nous aurons 199 features maps de taille 128.

Puis dans la deuxième couche de convolution. Cette couche est aussi composée de 5 filtres de taille 128

$$\text{Param} = 128 * 5 * 128 + 128 = 82048$$

Ensuite on applique Maxpooling pour réduire la taille de l'information ainsi la quantité de paramètres et de calcul. À la sortie de cette couche, nous aurons 39 feature maps de taille 128.

Le vecteur de caractéristiques issu de convolutions à une dimension de 4992.

La dernière couche utilisée la fonction d'activation un softmax qui permet de calculer la distribution de probabilité des 2 classes (positive et négative).

$$\begin{aligned} \text{Le nombre de paramètre} &= \text{input} * \text{output} + \text{biais.} \\ \text{Param} &= 128 * 2 + 2 = 258. \end{aligned}$$

Totale paramètre =  $\Sigma$  (*Param*)

$$\text{Total param} = \Sigma (25385400 + 64128 + 512 + 82048 + 639104 + 258) = 26171450.$$

```

Time to tokenize everything: 0.0 mins
Model: "model_10"

```

Layer (type)	Output Shape	Param #
input_10 (InputLayer)	(None, 1000)	0
embedding_21 (Embedding)	(None, 1000, 100)	25385400
dropout_73 (Dropout)	(None, 1000, 100)	0
conv1d_49 (Conv1D)	(None, 996, 128)	64128
max_pooling1d_49 (MaxPooling)	(None, 199, 128)	0
dropout_74 (Dropout)	(None, 199, 128)	0
batch_normalization_49 (Batch Normalization)	(None, 199, 128)	512
conv1d_50 (Conv1D)	(None, 195, 128)	82048
max_pooling1d_50 (MaxPooling)	(None, 39, 128)	0
flatten_25 (Flatten)	(None, 4992)	0
dense_49 (Dense)	(None, 128)	639104
dense_50 (Dense)	(None, 2)	258

```

Total params: 26,171,450
Trainable params: 26,171,194
Non-trainable params: 256

```

Figure 21 : Configuration du notre modèle

## 4.5. Implémentation et évaluation

### 1. Présentation des outils utilisés

#### Le software



« Python », le langage de programmation orienté objet, classé comme un langage de haut niveau interprété. Il est très sollicité par une large communauté de développeurs et peut s'utiliser dans de nombreux contextes et s'adapter à tout type d'utilisation grâce à des bibliothèques spécialisées. Python est un langage simple, facile à apprendre et permet une bonne réduction du coût de la maintenance des codes.



« TensorFlow<sup>12</sup> », est une bibliothèque unitaire et complète créée par l'équipe de Google Brain orienté vers les travaux de recherches du domaine « Machine Learning ». Nous avons choisi d'utiliser « Tensorflow » comme un backend pour la bibliothèque « Keras ».



« Keras<sup>13</sup> », représente la bibliothèque la plus sollicitée pour la création des classifieurs. Elle est souvent utilisée en collaboration avec les bibliothèques « Tensorflow » et

« Theano ». Dans notre système, nous avons utilisé cette bibliothèque pour la création du modèle de classification « CNN »

### Le hardware

Toutes les expériences ont été réalisées sur une machine qui offre des performances acceptables. Voici la configuration du matériel utilisé dans notre implémentation :

Système d'exploitation	Windows 7 64 bits
RAM	4 GO
CPU	Intel® Core™ i3 M380 2.53 GHz
GPU	Pas de GPU

Nous avons utilisé une plate-forme de calcul affiliée à Google pour fournir un espace et des exigences arithmétiques élevés.



« Google Colab » C'est une plateforme gratuite pour des sessions de temps limitées sur le web « envirent 12 heure / jours » offre la possibilité de calcul GPU et TPU. Elle utilise l'environnement « Jupyter Notebook » pour les différentes routines de programmation en langage Python.

<sup>12</sup> <https://www.tensorflow.org/>

<sup>13</sup> <https://keras.io/>



## 2.Présentation de la base de données usuelle

Notre but est de développer un programme d'apprentissage automatique pour identifier quand une source d'informations peut produire de fausses informations. Nous visons à utiliser un corpus de nouveaux articles étiquetés réels et faux pour construire un classificateur capable de prendre des décisions sur les informations en fonction du contenu du corpus [73]. Le modèle se concentrera sur l'identification de fausses sources d'informations, sur la base de plusieurs articles provenant d'une source. Une fois qu'une source est étiquetée comme producteur de fausses nouvelles, nous pouvons prédire avec une grande confiance que tout futur article de cette source sera également de fausses nouvelles. Le fait de se concentrer sur les sources élargit notre tolérance aux erreurs de classification des articles, car nous aurons plusieurs points de données provenant de chaque source.

Cet ensemble de données pour la classification des informations binaires contient un ensemble de 20 800 lignes d'information. Les informations sont étiquetées respectivement 1 ou 0 pour les informations positifs et négatifs.

Les données	Etiquette	Nombre	Type
Données fausse	0	10413	Texte anglais
Données Vraie	1	10387	Texte anglais
Total	/	20800	Texte anglais

Tableau 7 : Description de dataset

La figure suivante montre le nombre de fausse information en rouge et de vraie information en bleu :

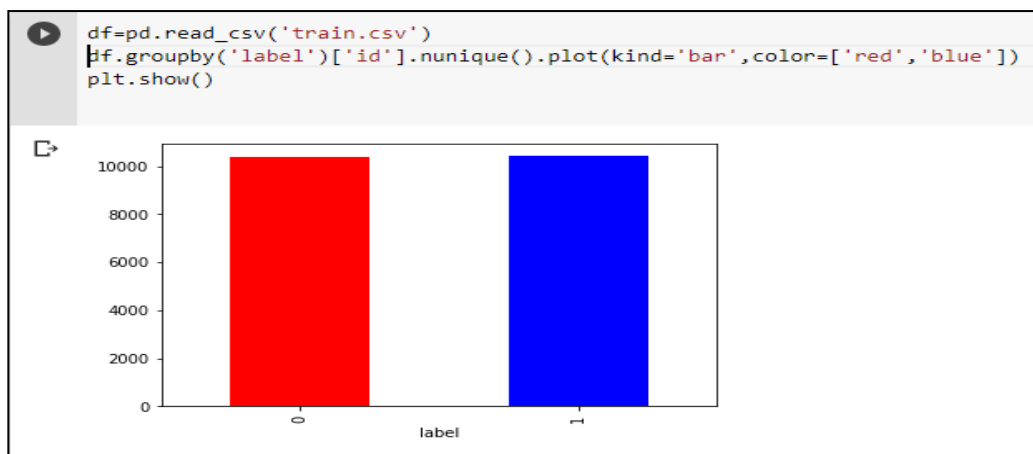


Figure 22 : Représentation de deux classes de jeux de données

Le fichier de jeux de données est un ensemble complet de données de formation avec les attributs suivants :

<b>Id</b>	<b>Identificateur unique pour chaque texte (article)</b>
<b>Titre</b>	Le titre de l'article
<b>Auteur</b>	Auteur de l'article
<b>Texte</b>	Le texte peut être incomplet
<b>Etiquette (label)</b>	Considérer comme '1' pour l'information vrai et '0' pour l'information fausse

**Tableau 8 :** Description des colonnes de dataset

Dans La figure 23, nous représentons le jeu de données en data frame :

	id	title	author	text	label
0	0	House Dem Aide: We Didn't Even See Comey's Let...	Darrell Lucus	House Dem Aide: We Didn't Even See Comey's Let...	1
1	1	FLYNN: Hillary Clinton, Big Woman on Campus - ...	Daniel J. Flynn	Ever get the feeling your life circles the rou...	0
2	2	Why the Truth Might Get You Fired	Consortiumnews.com	Why the Truth Might Get You Fired October 29, ...	1
3	3	15 Civilians Killed In Single US Airstrike Hav...	Jessica Purkiss	Videos 15 Civilians Killed In Single US Aistr...	1
4	4	Iranian woman jailed for fictional unpublished...	Howard Portnoy	Print InAn Iranian woman has been sentenced to...	1
...	...	...	...	...	...
20795	20795	Rapper T.I.: Trump a 'Poster Child For White S...	Jerome Hudson	Rapper T. I. unloaded on black celebrities who...	0
20796	20796	N.F.L. Playoffs: Schedule, Matchups and Odds -...	Benjamin Hoffman	When the Green Bay Packers lost to the Washing...	0
20797	20797	Macy's Is Said to Receive Takeover Approach by...	Michael J. de la Merced and Rachel Abrams	The Macy's of today grew from the union of sev...	0
20798	20798	NATO, Russia To Hold Parallel Exercises In Bal...	Alex Ansary	NATO, Russia To Hold Parallel Exercises In Bal...	1
20799	20799	What Keeps the F-35 Alive	David Swanson	David Swanson is an author, activist, journa...	1

20800 rows x 5 columns

Figure 23 : Quelques lignes de jeux de données

### 3. Word embedding (Glove)

Comme nous avons mentionné dans le chapitre précédent que l'entrée de l'algorithme de CNN est une couche de embedding c-à-d chaque mot est représenté avec un vecteur de taille fixe, pour cela nous avons choisi d'utiliser le modèle de Glove à dimension 100, en peut le télécharger en python par le code en figure suivant.

```
#####telecharger word embedding pretraité --glove-- #####
!wget https://storage.googleapis.com/laurencemoroney-blog.appspot.com/glove.6B.100d.txt
```

Figure 24: Lien de téléchargement de word embedding

## 4.6. Résultats obtenus et discussions

Pour notre modèle on a utilisé le modèle CNN-1d à l'aide de modèle de word embedding « Glove » de dimension 100 et le nombre d'époque de 5.

Dans la Figure (25), La précision de l'apprentissage augmente avec le nombre d'époque, jusqu'à elle est stable, ceci reflète qu'à chaque époque le modèle apprend plus d'informations. Aussi la validation augmente de la même façon.

De même, l'erreur d'apprentissage et diminue, d'autre part la validation diminue aussi avec le nombre d'époque.

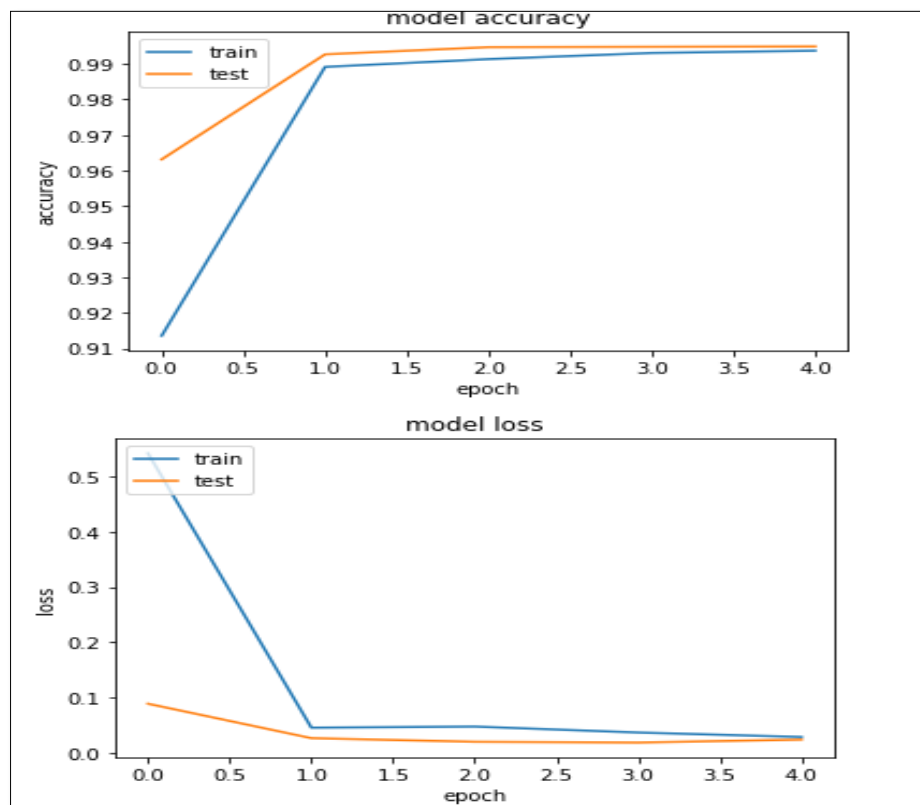


Figure 25 : Précision et Erreur pour notre modèle

La figure 26 montre la matrice de confusion de la prédiction du notre model, le taux de précision est de 99%. Le modèle a mal classé 16 informations. Ce résultat est très satisfaisable.

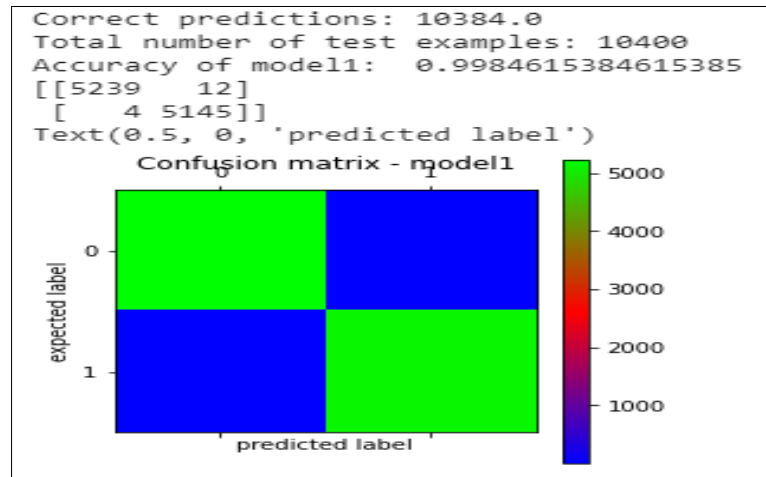


Figure 26 : Matrice de confusion du notre modèle

La figure 27, montre la Précision, le Rappel et le F-score de la prédiction du notre modèle, le rappel pour les données étiqueté en 0 est de 99% et pour les donnes étiquetées en 1 est de presque le même taux (100%) .

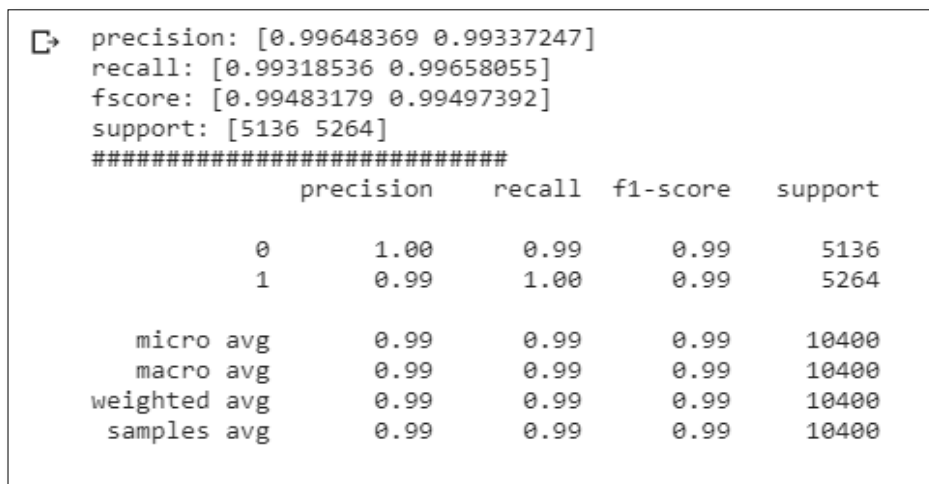


Figure 27 : Précision, Rappel et F-score du notre modèle

Dans ce modèle en va prendre le modèle précédent et en va faire une augmentation de 100 dans le nombre d'itération.

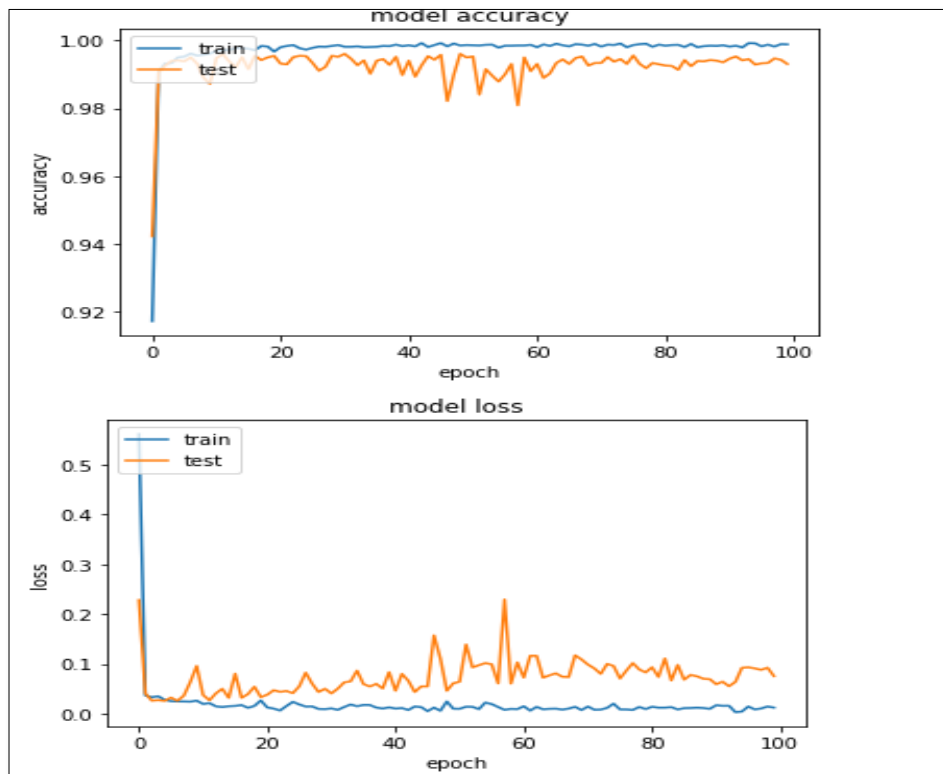


Figure 28: modèle 2 avec 100 époques

On remarque qu'avec l'augmentation de nombre des époques, le modèle ne tombe pas dans le problème de overffing.

Ainsi, nous pouvons constaté comme résultat final le meilleur modèle qui donne des bons résultats c'est le model CNN-1D avec une Word embedding prétraité.

## Discussion

Dans le tableau 9 on compare notre résultat avec le travail de Shu [74] qui a presque applique tous les algorithmes d'apprentissage automatique sur le même corpus de données que nous avons utilisée.

---

Le modèle	La précision
Naive Bayes	72.94%
SVM	88.42%
Neural Network with TF	81.42%
Neural Network with Keras	92.62%
LSTM	94.53%
<b>Notre modèle CNN-1D</b>	<b>99,8%</b>

**Tableau 9 :** Comparaison des résultats

## 4.7. Conclusion

Dans ce dernier chapitre nous avons expliqué notre proposition et montrer son implémentation pour la détection de fausse information. L'idée est basé sur la technique d'indexation des mots « Word Embedding » et la sous technique « Glove» comme un support de représentation de connaissances nous avons expliqué notre modèle ainsi que la validation et les divers résultats obtenus. Et nous avons conclure que le model «CNN-1D » avec le Word Embedding prétraités nous donne le meilleur résultat.

# Conclusion Générale

---



## 1. Conclusion

L'objectif de ce mémoire est la détection des rumeurs dans les réseaux sociaux selon deux voies, une rumeur positive, une rumeur négative. Pour cela nous avons choisi de réaliser une application sous Python qui utilise une source de données (Corpus en format .csv) contient des textes annotés par des valeurs 1 et 0. Nous avons commencé par la définition de quelques concepts de contexte de travail. Ensuite, nous avons mis l'accent sur les approches connexes, puis nous avons étudié le traitement de langage naturel. L'étape qui a pris plus de temps est le choix de méthode de vectorisation des mots, et finalement nous avons choisi la méthode d'indexation des mots « Glove » du « Word Embedding » et faire une comparaison avec les travaux déjà réalisées. Le classificateur que nous avons appliqué c'est le CNN-1D. Les résultats obtenus sont très encourageant, où nous avons atteint un taux de 99% dans le cas de l'utilisation de Word embedding prétraité.

## 2. Les perspectives

Pour les travaux futurs, nous avons fixé les buts suivants :

- ❖ Tester le modèle pour d'autres langues par exemple : arabe.
- ❖ L'utilisation des autres techniques de NLP comme le word2vec et doc2vec.
- ❖ L'application d'autres classificateurs et l'utilisation d'autres fonctionnalités.
- ❖ Tester le modèle avec des classes Mixte c'est-à-dire autres que les classes positive et négative.
- ❖ Développer une interface graphique pour que le modèle soit mieux utilisable..
- ❖ L'enrichissement de Corpus par d'autres informations afin d'obtenir des résultats bien précis.



## ***Les références***

[1] : Zammar, N. (2012). Réseaux Sociaux numériques: essai de catégorisation et cartographie des controverses (Doctoral dissertation).

[2] : I. Nishanbaev, E. Champion, and D. A. McMeekin, "A survey of geospatial semantic web for cultural heritage," *Heritage*, vol. 2, no. 2, 2019, pp. 1471–1498

[3] : Tichy, N. M "Networks in organizations", in Nystrom and Starbuck (Eds) *Handbook of Organizational Design*, Vol. 2, Oxford University Press, New York, 1981.

[4] : Giles Hogben & ENISA, "security issues and recommendations for online social networks", ENISA Report, Greece, [http://www.enisa.europa.eu/doc/pdf/deliverables/enisa\\_pp\\_social\\_networks.pdf](http://www.enisa.europa.eu/doc/pdf/deliverables/enisa_pp_social_networks.pdf), 2007.

[5] : Dajana Kapusova Leconte " développement d'un logiciel de réseau social comme soutien à une communauté de pratique ", Mémoire présenté pour l'obtention du DESS STAF Sciences et Technologies de l'Apprentissage et de la Formation TECFA, Juin 2008.

[6]: RIOU, (Nicolas): « le consommateur digital » : nouvelles approches pour le séduire, Édition Eyrolles, Paris, 2017, p .76.

[7] : <http://www.lien-optionnel.com/reseaux-sociaux.html> .( Le 25/04/2020)

[8] : <https://fredcavazza.net/2015/05/29/panorama-desm%C3%A9dias-sociaux-2015/>. (Le 26/04/2020)

[9] : [http://www.pointsdactu.org/article.php3? Id article=1293](http://www.pointsdactu.org/article.php3?Id%20article=1293). (Le 26/04/2020)

[10] : THELWALL Mike, "Social network sites: Users and uses, "M.Zelkowitz (Ed.) *Advances Incomputers Elsevier*, vol. 76, pp. 19-73, 2009.

[11] : ZEMMAR Nisrine, "Réseaux Sociaux numériques : essai de catégorisation et

cartographie des controverses, "Université Rennes 2, Thèse de doctorat 2012.

[12]: <http://www.fredcavazza.net/2015/05/29/panorama-desmedias-sociaux-2015/>

[13]: CREFF Marie, "Réseaux sociaux : quelles opportunités pour les services ? Le cas de l'assistance en ligne d'Orange, "Institut national des techniques de la documentation, Mémoire de fin d'étude INTD 2010.

[14]: <http://reseauxlapie.canalblog.com/> .( Le 05/05/2020)

[15] : <https://vincos.it/world-map-of-social-networks/> (le 05/05/2020)

[16] :FILLIETTAZ Francois, GREGORI Macro, "Comprendre les réseaux sociaux numériques,"Direction des systèmes d'information et service écoles- médias, 2011.

[17] :BOYD danah, "Social Network Sites: Public, Private, or What? " Knowledge Tree May 2007.

[18] :TORLOTING Philippe, "Enjeux et perspectives des réseaux sociaux, "Institut Supérieur du Commerce, Paris, mémoire de fin d'étude 2006.

[29] : NEDIOUI MOHAMED ABDELHAMID, Fouille et apprentissage automatique dans les réseaux sociaux dynamiques, 2014/2015.

[20] : OUCH ! Sécurité des réseaux sociaux, Septembre 2011.

[21] Hunt Allcott and Matthew Gentzkow. 2017. Social media and fake news in the 2016 election. *Journal of Economic Perspectives* 31, 2 (2017), 211–36

[22] Kai Shu, Amy Sliva, Suhang Wang, Jiliang Tang, and Huan Liu. 2017a. Fake news detection on social media: A data mining perspective. *ACM SIGKDD Explorations Newsletter* 19, 1 (2017), 22–36

[23] Soroush Vosoughi, Deb Roy, and Sinan Aral. 2018. The spread of true and false news online. *Science* 359, 6380 (2018), 1146–1151

[24] Dan Berkowitz and David Asa Schwartz. 2016. Miley, CNN and The Onion: When fake news becomes realer than real. *Journalism Practice* 10, 1 (2016), 1–17

[25] Nir Kshetri and Jeffrey Voas. 2017. The Economics of "Fake News". *IT Professional* 6 (2017), 8–12

- [26] Adam Kucharski. 2016. Post-truth: Study epidemiology of fake news. *Nature* 540, 7634 (2016), 525
- [27] Cody Buntain and Jennifer Golbeck. 2017. Automatically Identifying Fake News in Popular Twitter Threads. In *Smart Cloud (SmartCloud)*, 2017 IEEE International Conference on. IEEE, 208–215
- [28] : Christine DU, M. Olivier POMMERET, thèse professionnelle en intelligence économique et management des connaissances, promotion 2012.
- [29]: FRANÇOIS-BERNARD HUYGHE, Faux, rumeurs et désinformation dans le cyberspace, Edition L'IRIS, 2013.
- [30]: Adeline Michel, A.C.Sordet et E. Moraillon, les rumeurs en tant que phénomène d'influence sociale, Ecole de psychologues praticiens de Lyon, 2004.
- [31] Jean-Noël Kapferer : Rumeurs. Le plus vieux média du monde, Le Seuil, Coll. Points actuels, 1990.
- [32] J. MORENO, Emotions mapped by new geography, <http://diana-jones.com/wpcontent/uploads/Emotions-Mapped-by-New-Geography.pdf>, 1933.
- [33] Gustavo Aguilar, Suraj Maharjan, Adrian Pastor Lopez Monroy, and Tamar Solorio, A multi-task approach for named entity recognition in social media data, In *Proceedings of the 3rd Workshop on Noisy User-generated Text*, 2017
- [34] Qazvinian, V., E. Rosengren, D. R. Radev, et Q. Mei (2011). Rumor has it : Identifying misinformation in microblogs. In *Proceedings of the CEMNLP*, pp. 1589–1599. ACL
- [35] Hamidian, S. et M. Diab (2015). Rumor detection and classification for twitter data. In *SOTICS 2015*, pp. 71–77.
- [36] Gupta, A., P. Kumaraguru, C. Castillo, et P. Meier (2014). Tweetcred : Real-time credibility assessment of content on twitter. In *ICSI*, pp. 228–243. Springer.
- [37] Castillo, C., M. Mendoza, et B. Poblete (2011). Information credibility on

twitter. In WWW 2011, pp. 675–684. ACM.

[38] Gupta, M., P. Zhao, et J. Han (2012). Evaluating event credibility on twitter. In Proceedings of the 2012 SIAM International Conference on Data Mining, pp. 153–164. SIAM.

[39] Jin, Z., J. Cao, Y.-G. Jiang, et Y. Zhang (2014). News credibility evaluation on microblog with a hierarchical propagation model. In ICDM 2014, pp. 230–239. IEEE

[40] Zhao, Z., P. Resnick, et Q. Mei (2015). Enquiring minds : Early detection of rumors in social media from enquiry posts. In WWW 2015, pp. 1395–1405.

[41] Ma, J., W. Gao, P. Mitra, S. Kwon, B. J. Jansen, K.-F. Wong, et M. Cha (2016). Detecting rumors from microblogs with recurrent neural networks. In IJCAI, pp. 3818–3824.

[42] Chen, T., X. Li, H. Yin, et J. Zhang (2018). Call attention to rumors : Deep attention based recurrent neural networks for early rumor detection. In PAKDD 2018, pp. 40–52. Springer.

[43] Jin, Z., J. Cao, Y. Zhang, J. Zhou, et Q. Tian (2017). Novel visual and statistical image features for microblogs news verification. IEEE transactions on multimedia 19(3), 598–608.

[44] Yu, F., Q. Liu, S. Wu, L. Wang, T. Tan, et al. (2017). A convolutional approach for misinformation identification. In IJCAI, pp. 3901–3907.

[45] Nguyen, T. N., C. Li, et C. Niederée (2017). On early-stage debunking rumors on twitter : Leveraging the wisdom of weak learners. In ICSI, pp. 141–158. Springer.

[46] Boididou, C., S. Papadopoulos, L. Apostolidis, et Y. Kompatsiaris (2017). Learning to detect misleading content on twitter. In ICMR 2017, pp. 278–286. ACM.

[47] Marra, F., D. Gragnaniello, D. Cozzolino, et L. Verdoliva (2018). Detection of gan-generated fake images over social networks. In 2018 IEEE Conference on MIPR, pp. 384–389. IEEE.

- [48] Bianchi, T. et A. Piva (2012). Image forgery localization via block-grained analysis of jpeg artifacts. *IEEE Transactions on IFS* 7(3), 1003–1017.
- [49] Goljan, M., J. J. Fridrich, et M. Chen (2011). Defending against fingerprint-copy attack in sensor-based camera identification. *IEEE Transactions on IFS* 6(1), 227–236.
- [50] Gupta, A., H. Lamba, P. Kumaraguru, et A. Joshi (2013). Faking sandy : characterizing and identifying fake images on twitter during hurricane sandy. In *WWW 2013*, pp. 729–736. ACM.
- [51] Morris, M. R., S. Counts, A. Roseway, A. Hoff, et J. Schwarz (2012). Tweeting is believing ? : understanding microblog credibility perceptions. In *ACM conference on CSC*, pp. 441–450
- [52] Karthikram, G. M. P. et G. Parthiban (2014). Tag based image retrieval (tbir) using automatic image annotation. *IJRET* 3(03)
- [53] Alkhawlan, M., M. Elmogy, et H. El Bakry (2015). Text-based, content-based, and semantic-based image retrievals : A survey. *IJCIT* 4(01).
- [54] Rashno, A. et E. Rashno (2019). Content-based image retrieval system with most relevant features among wavelet and color features. *arXiv preprint arXiv :1902.02059*.
- [55] Luo, B., X. Wang, et X. Tang (2003). World wide web based image search engine using text and image content features. In *Electronic Imaging 2003*, Volume 5018, pp. 123–130.
- [56] Unar, S., X. Wang, C. Wang, et Y. Wang (2019). A decisive content based image retrieval approach for feature fusion in visual and textual images. *Knowledge-Based Systems* 179, 8–20
- [57] Maigrot, C., E. Kijak, et V. Claveau (2017). Détection de fausses informations dans les réseaux sociaux : l'utilité des fusions de connaissances. In *CONFérence Recherche d'Information et Applications*, pp. 107–122.
- [58] Wan, J., D. Wang, S. C. H. Hoi, P. Wu, J. Zhu, Y. Zhang, et J. Li (2014). Deep learning for contentbased image retrieval : A comprehensive study. In *ACM*

international conference on Multimedia, pp. 157–166.

[59] A. L. Samuel, "Some Studies in Machine Learning Using the Game of Checkers," in IBM Journal of Research and Development, vol. 3, no. 3, pp. 210-229, July 1959.

[60] Annina S., Mahima S, S. Venkatesan<sup>3</sup>, D.R. Ramesh Babu, An Overview of Machine Learning and its Applications. International Journal of Electrical Sciences & Engineering (IJESE).

[61] Andrew W. Trask, 2019. grokking Deep Learning.

[62] Yann LeCun, Yoshua Bengio & Geoffrey Hinton., may 2015 REVIEW Deep learning

[63] : L. Deng, et D. Yu, Deep learning. Boston, p.217.2014.

[64] : <https://fr.mathworks.com/discovery/deep-learning.html> . (le 17/05/2020)

[65] : <https://www.saagie.com/fr/blog/qu-est-ce-que-le-deep-learning/> (le 28/05/2020)

[66] W. Yin, K. Kann, M. Yu, et H. Schütze. Comparative Study of CNN and RNN for Natural Language Processing, 2017.

[67] (2015, 8 27). Récupéré sur le blog de colah: <https://colah.github.io/posts/2015-08-Understanding-LSTMs/>

[68] liu, B. (2012). Sentiment analysis and opinion mining. Synthesis lectures on human.

[69] : <https://www.slideshare.net/EIHabibNFAOUI/conceptual-foundations-of-text-mining-and-preprocessing-steps-nfaoui-elhabib> (Le 28/05/2020)

[70] T. Mikolov ,I. Sutskever, K. Chen, G. Corrago, J. Dean, Distributed representations of words and phrases and their compositionality. In: 27th Conference on Neural Information Processing Systems; 5-10 December 2013; Lake Tahoe, Nevada, USA. pp. 1-9.2013.

[71] : <https://medium.com/voice-tech-podcast/nlp-pipeline-101-with-basic-code->



[example-feature-extraction-ea9894ed8daf](#) (Le 25/06/2020)

[72] : <https://stackabuse.com/text-classification-with-python-and-scikit-learn/> (le 26/06/2020)

[73] : [https://www.kaggle.com/c/fake-news/overview\\_1](https://www.kaggle.com/c/fake-news/overview_1) (le 12/07/2020)

[74] : Shu , 2017 <https://github.com/rockash/Fake-news-Detection> (le 12/07/2020)

[75] Azri, A., Favre, C., Harbi, N., & Darmont, J. (2019, October). Vers une analyse des rumeurs dans les réseaux sociaux basée sur la véracité des images: état de l'art.