
Faculté des Mathématiques et d'Informatique


**République Algérienne Démocratique et Populaire**

**Ministère de l'Enseignement Supérieur et de la Recherche Scientifique**

**Université de Bordj Bou Arreridj Mohammed El Bachir El Ibrahimi**

**Faculté des Mathématiques et d'Informatique**

**Département d'Informatique**

---

# Mémoire

**En vue de l'obtention du diplôme de Master**

**Domaine : Mathématiques et Informatique**
**Filière : Informatique**
**Option : Ingénierie de l'Informatique Décisionnelle (IID)**

## Intitulé

# Détection de fausses informations sur le web et les réseaux sociaux :

# Cas de l'infodemie sur coronavirus

**Présenté par** :
- Cheyma TAGUIA
- Zahra SID

**Proposé et dirigé par** :
- Djamila MOHDEB

**Devant le jury composé de** :
- Mouhoub BELAZZOUG Président
- Meriem LAIFA Examinatrice

**Année Universitaire** 2019/2020

# Dédicace

*Au nom de Dieu Clément et Miséricordieux*

*Je dédie ce modeste travail de fin de formation à mes très chers parents qui m'ont été un grand soutien moral dans les moments difficiles durant la formation*

*À tous mes enseignants pour leur contribution à ma formation*

*Ainsi qu'à mes chères sœurs et mon frère et à toute ma famille*

*Sans oublier mon binôme et mes amies proches Cheyma et Ibtissam*

*A tous qui ont contribué de près ou de loin à la réalisation de ce travail, qu'ils trouvent ici la traduction de ma gratitude et ma reconnaissance.*

*Zahra Sid*

# Dédicace

*Au nom de Dieu Clément et Miséricordieux*

*Je dédie ce modeste travail de fin d'études à mes grands-parents et mes très chers parents, particulièrement à mon père qui est mon idole dans la vie et ma mère qui m'a été un grand soutien moral dans les moments difficiles durant mes études*

*À tous mes enseignants pour leur contribution à ma formation*

*Ainsi qu'à ma sœur et mes frères et à toute ma famille*

*Sans oublier mes collègues étudiants et mes amies Zahra et Ibtissam*

*A tous qui ont contribué de près ou de loin à la réalisation de ce travail, qu'ils trouvent ici la traduction de ma gratitude et ma reconnaissance.*

*Cheyma Taguia*

# Remerciements

C’est avec un immense plaisir que nous tenons à remercier très sincèrement toutes les personnes qui nous ont aidé et qui ont ainsi contribué à la réalisation de ce mémoire.

Nous tenons à remercier notre encadrante Dr. Mohdeb Djamila d’avoir dirigé ce travail.

Nos vifs remerciements vont également aux membres du jury pour l’intérêt qu’ils ont porté à notre projet de fin d’études en acceptant d’examiner ce travail et de l’enrichir par leurs propositions.

Nous souhaitons exprimer notre gratitude envers nos familles et tous nos amis pour leur soutien et encouragements tout au long de ce travail.

Enfin, nous voudrions également remercier toutes les personnes qui ont participé de près ou de loin à la réalisation de ce travail.

# ملخص

مع اتساع رقعة استعمال الانترنت وبروز تكنولوجيات الإعلام والإتصال ونظرا للطلب المتزايد والدائم على المعلومات، عرفت ظاهرة الأخبار الكاذبة و المعلومات الخاطئة انتشارا كبيرا وقد أثرت سلبا على جميع مجالات الحياة العصرية. نتيجة لذلك، أصبح وجود أنظمة أو تطبيقات للتعرف على الأخبار الكاذبة ضرورة ملحة. هاته العملية يصعب إنجازها يدويا، لذلك يتم الإستعانة بعدد كثير من التطبيقات و الحلول الرقمية التي وضعت لأجل هذا الغرض.

في بحثنا هذا نقترح نموذجا معلوماتيا يعمل على التعرف على المعلومات الكاذبة التي تم نشرها عبر مواقع الانترنت ومواقع التواصل الاجتماعي خلال الأزمة الصحية التي رافقت وباء كورونا الجديد. وهذا باستخدام تقنيات التعلم الآلي و التعلم العميق مع احترام منهجية تصنيف النصوص التي تعود لمجال المعالجة الآلية للغة الطبيعية.

**كلمات مفتاحية:** المعلومات الخاطئة، الأخبار الكاذبة، فيروس كورونا، كوفيد 19، التعلم الآلي، التعلم العميق، تصنيف النص، المعالجة الآلية للغات الطبيعية.

# Abstract

With the emergence of the use of internet and ICT techniques, and the growing demand for information, the phenomenon of fake news has spread widely by negatively influencing different sectors of modern life. Thus, it has become a necessity to develop solutions for detecting and suspending the dissemination of this type of information that circulate on the web and social platforms. As this problem is difficult to manage manually, many applications and systems have already been developed to solve it automatically.

In this project, we propose an automatic system to accomplish the task of detecting fake news published on the web and social media during the health crisis of the Coronavirus pandemic. This is performed using machine learning and deep learning techniques while respecting the methodology of text classification related to the field of natural language processing.

**Keywords**: false information, fake news, coronavirus, covid-19, machine learning, deep learning, text classification, NLP.

# Résumé

Avec l'émergence de l'utilisation de l'internet et des TIC, et suite à la demande croissante et continuée d'information, le phénomène de fausses informations a connu une large propagation en influençant négativement des différents secteurs de vie moderne. Ainsi, il est devenu une nécessité de développer des solutions pour la détection et la suspension de la dissémination de ce type d'informations circulant sur le web et les plateformes sociales. Comme ce problème est difficile à gérer manuellement, de nombreuses applications et systèmes ont déjà été développés pour le résoudre automatiquement.

Dans ce projet, nous proposons un système informatique pour accomplir la tâche de détection de fausses nouvelles publiées sur le web et les médias sociaux durant la crise sanitaire de la pandémie du Coronavirus. Ceci est réalisé grâce aux techniques de l'apprentissage automatique et l'apprentissage approfondi tout en respectant la méthodologie de la classification du texte liée au domaine du traitement automatique du langage naturel.

**Mots-clés** : fausses informations, fausses nouvelles, coronavirus, covid-19, apprentissage automatique, apprentissage approfondi, classification de texte, TALN.

# Table des matières

# Liste des figures

# Liste des tableaux

# Liste des algorithmes

# Introduction générale

## 1. Contexte

De nos jours, Internet, le plus vaste réseau mondial de communication, est devenu universel et accessible grâce aux sites web et aux média sociaux qui ont réussi à rassembler les individus, les entreprises et les organisations côte à côte. Ce nouvel environnement a offert un moyen facile et efficace de partage de publications, d'articles, de nouvelles et des informations par rapport aux médias traditionnels en raison de ses services gratuits qui offrent une large diversité de fonctionnalités. La popularité des médias sociaux est liée à la demande croissante et continuée d'information, devenue plus importante dans nos sociétés modernes et se multipliant à un niveau sans précédent.

Durant cette décennie, le phénomène de fausses nouvelles, *fake news* en anglais et *infox* en français, a attiré autant d'inquiétudes que d'attention. Les fausses informations portent sur tous les niveaux allant des individus ordinaires jusqu'aux gouvernements et états. Elles ont connu une large propagation en influençant négativement des différents secteurs de vie. De ce fait, l'information est devenue plus précieuse que jamais, et l'existence des solutions pour détecter et bloquer la dissémination des fausses nouvelles qui circulent sur le web est devenue une nécessité.

De nombreuses applications et systèmes ont déjà été développés pour cette tâche. Le processus de détection de la désinformation et la mésinformation vise, automatiquement ou manuellement, à une meilleure détermination d'une fausse information en vue de bloquer sa diffusion ou de minimiser son ampleur.

## 2. Problématique

L'an 2020 a connu une crise qui a affecté la vie de la population du monde entier sur plusieurs domaines, plus précisément le secteur commercial et social. Des villes et provinces ou même des zones géographiques sont entièrement fermées. Cette crise fait suite au déclenchement de la pandémie de Coronavirus (Covid-19) causée par le SARS-CoV-2.

Un très grand nombre d'infox et cas de désinformation ont été relevés à travers les divers types des médias, amenant l'Organisation Mondiale de la Santé à parler d'une infodémie[1] tout aussi dangereuse que la maladie elle-même. Tout autour du monde, les compagnes de

[1] https://www.who.int/fr/dg/speeches/detail/munich-security-conference

mésinformation ont emmené à des situations de panique et de désespération d'un côté, et des situations de négligence d'autre part. D'autres ont incité les gens à sous-estimer la sériosité de cette maladie, ignorer les mesures de prévention nécessaires, et croire à la fausse médecine et aux faux remèdes.

Devant cette situation inquiétante, il est devenu indispensable d'étudier et comprendre ce phénomène pour en trouver des solutions techniques performantes et efficaces.

## 3. Objectif et contribution

Dans ce projet, notre objectif est d'explorer le domaine des fausses informations, en particulier la tâche de détection de fausses nouvelles publiées sur le web et les médias sociaux durant la crise sanitaire de la pandémie du Coronavirus. Nous avons également exploité les techniques de l'apprentissage automatique et l'apprentissage approfondi pour développer un modèle de détection de fausses informations tout en respectant la méthodologie de la classification du texte liée au domaine du traitement automatique du langage naturel.

## 4. Plan de mémoire

La suite de ce mémoire est organisée en quatre chapitres :

Le premier chapitre consistera en la présentation du problème des fausses informations, ses types, ses acteurs et les motivations derrière la diffusion de ce type de contenu. Le deuxième portera sur les fondements théoriques et techniques de la tâche de détection des fausses informations. Le troisième, fera l'objet de l'étude conceptuelle et technique du modèle proposé de détection en se basant sur une description, claire et précise, des techniques utilisées de l'apprentissage automatique et l'apprentissage approfondi. Par la suite, nous présenterons brièvement la chronologie de la crise sanitaire de la Covid-19 et la désinformation liée à elle. Enfin, dans le quatrième et dernier chapitre, nous présenterons les résultats obtenus par notre modèle après avoir mené une analyse exploratoire de données. Puis, nous dresserons une conclusion générale du projet.

# Chapitre I
# Les fausses informations

# Introduction

A cette ère des médias sociaux, la consommation de l'information a été multiplié à un niveau sans précédent. Le phénomène de fausses nouvelles, a attiré beaucoup d'attention ces dernières années vu son influence majeure sur les individus et les sociétés modernes.

Dans ce qui suit, nous essayons de donner une vision claire sur la notion de fausses nouvelles et ses dérivés. Nous avons utilisé un terme plus général qui est « les fausses informations » afin de bien cerner les différentes facettes dont une fausse nouvelle exhibe. De plus amples informations ont été présentées aussi sur les principaux diffuseurs des fausses informations et leurs objectifs et motivations.

## 1. Définition

Les fausses informations sont des renseignements mensongères ou inventés dont l'authenticité est ambiguë ou jamais confirmée. Elles portent sur une situation, un évènement, une personnalité publique ou inconnue, une organisation, un gouvernement voire un état.

Les informations erronées sont délivrées via les médias traditionnels ou les médias sociaux non institutionnels tels les blogs et les réseaux sociaux en ligne. L'objectif se varie entre manipuler et tromper l'opinion publique ou un auditoire bien déterminé sur des sujets spécifiques ; modifier leurs décisions et actions ; faire du mal aux opposants (organisations, partis, politiciens, intellectuels …etc.) ou bien pour le but d'augmenter le profit et la popularité.

## 2. Les types des fausses informations

Les fausses informations apparaissent sous différentes formes. Il est important de noter qu'une fausse information peut être le résultat de chevauchement de plusieurs types à la fois, parmi les huit types suivants [1] :

- **Information fabriquée** : une information complètement imaginaire qui n'a aucune relation avec des faits réels. Ce type de fausse information n'est pas nouveau. Il existe depuis la naissance du journalisme. Des exemples très populaires incluent les histoires fabriquées liées aux extraterrestres et aux objets volants non identifiés (UFO).
- **Propagande** : il s'agit d'un sous type particulier des informations fabriquées. Ce genre de fausses nouvelles était largement utilisé pendant la Seconde Guerre Mondiale et la Guerre Froide. La propagande est fortement utilisée dans des contextes politiques pour influencer

un public cible, dans le but de propager des idées, des plans politiques, des idéologies ou de discréditer un parti politique ou un état-nation particulier. La propagande peut même changer le cours de l'histoire, prenant par exemple l'invasion de l'Irak en 2003.

- **Théories du complot** : se réfèrent à des récits théoriques qui expliquent des situations ou des évènements par l'invocation des complots sans aucune preuve. Ces récits parlent des actes illégaux des gouvernements ou des personnes puissantes agissant dans l'ombre. Les informations conspirationnistes expliquent les évènements par recours à la conspiration politique en réfutant toutes les autres explications officielles ou scientifiques fournies par les systèmes et les grands médias.
- **Canulars** : des nouvelles qui contiennent des faits faux ou inexacts qui ont pour but d'abuser de la crédulité de quelqu'un. Ils s'appellent aussi les demi-vérités ou les *factoïdes*. Prenant par exemple les nouvelles qui parlent de la mort des célébrités.
- **Information biaisée ou unilatérale** : réfère à un récit extrêmement partiel ou subjectif. Dans le contexte politique, ce type est appelé les nouvelles hyper-partisanes ; des faits qui sont biaisées au profit d'une personne, un parti, une situation ou un évènement.
- **Rumeur** : un récit dont la source est inconnue et la véracité est ambiguë.
- **Titres accrocheurs** : désigne l'utilisation délibérée des titres trompeurs sur le web dans le but de captiver plus de lectorat. C'est un ancien type de fausses informations connu sous le nom de journalisme jaune. Il s'agit de l'un des catégories les moins graves de fausses informations parce qu'il est facile de vérifier son authenticité en consultant le contenu lié au titre.
- **Nouvelles satiriques** : des écrits ou du contenu visuel moqueurs critiquant des situations, des personnes, des organisations, ou des gouvernements. Ce genre de nouvelles attire plus d'attention sur le web. Cependant, comme les écrits satiriques sont principalement diffusés via les réseaux sociaux, leur caractère ironique est souvent négligé et ignoré par les lecteurs qui les prennent au sérieux, sans vérification supplémentaire.

## 3. Les acteurs de fausses informations

Divers acteurs peuvent obtenir un gain personnel en diffusant de fausses informations. Il s'agit d'une catégorie très large allant des utilisateurs simples jusqu'à des entités plus organisés et plus puissantes. On peut citer entre autres [1] :

- **Bots** : des programmes qui font partie d'un réseau de bot (Botnet). Ils sont généralement reliés à un grand nombre de faux comptes qui collaborent pour contrôler l'activité en ligne et propager de fausses informations dans le web.
  Il existe plusieurs types de Bots informatiques ayant des capacités variables. Certains bots republient du contenu ou le promeuvent ; d'autres affichent du contenu original.
- **Organisations criminelles et terroristes** : les gangs criminels et les organisations terroristes exploitent les réseaux sociaux en ligne pour diffuser de fausses informations dans le but d'atteindre leurs objectifs logistiques ou politiques. Un exemple récent est l'organisation terroriste EI (Etat Islamique ou Daech) qui diffuse de fausses informations dans les réseaux sociaux à des fins de propagande, en particulier pour le recrutement de nouveaux membres terroristes.
- **Activistes et organisations politiques** : différentes organisations et activistes politiques partagent des fausses informations, soit pour promouvoir leurs organisations, discréditer leurs opposants ou bien imposer certaines narrations spécifiques, notamment avant les grandes évènements politiques.
- **Gouvernements** : historiquement, les gouvernements participent régulièrement à la diffusion de fausses informations pour diverses raisons. Plus récemment, avec la prolifération du web, les gouvernements utilisent les médias sociaux pour manipuler l'opinion publique local ou étranger sur des sujets précis. Exemple : l'intervention du gouvernement russe aux élections américaines de 2016.
- **Trolls** : le terme *Troll* se réfère aux utilisateurs qui publient des informations controversées ou hors-sujet afin de perturber le flux normal de discussion sur un site web, de provoquer d'autres utilisateurs ou bien de pratiquer une pression émotionnelle sur eux. L'objectif est généralement l'amusement personnel.
- **Trolls financés** : il s'agit d'un groupe spécial des internautes qui sont payés pour publier du contenu ciblant une population bien déterminée. Habituellement, ils sont employés pour faire avancer un agenda politique et influencer les gens à adopter certaines tendances sociales ou économique. Ce type d'acteurs travaille à son profit. Il est beaucoup plus difficile de les distinguer par rapport aux bots parce qu'ils présentent des caractéristiques semblables à celles des utilisateurs ordinaires.
- **Journalistes** : les journalistes sont les principales entités responsables de la diffusion de l'information à la fois en ligne et hors ligne. Habituellement, Ils peuvent publier de fausses

informations pour diverses raisons. Par exemple rendre une nouvelle plus attirante, pour augmenter la popularité de leur plateforme, blog, site, ou journal.

- **Idiots utiles** : les idiots utiles sont des utilisateurs ordinaires qui partagent de fausses informations principalement parce qu'ils sont naïfs ou manipulés par d'autres personnes ou organisations. Généralement, ils sont inconscients des objectifs de leurs manipulateurs.
- **Vrais croyants (les conspirationnistes)** : il s'agit de personnes qui sentent la responsabilité de partager et diffuser des informations, généralement conspirationnistes, parce qu'elles croient entièrement à son authenticité et son importance pour les autres.

## 4. Les motifs de la propagation de la fausse information

Les acteurs qui participent à la propagation de l'information fausse se diffèrent selon leurs intentions, leurs objectifs et leurs motivations. Les motifs les plus courants sont [1] :

- **Intention malveillante** : par exemple, porter atteinte à l'image publique d'une personne ou d'une entité. Les intentions malveillantes poussent à nuire ou faire du mal aux autres personnes par des moyens et manières différents.
- **Influence** : induire en erreur d'autres personnes afin d'influencer leurs décisions, manipuler l'opinion publique sur des sujets précis, ou pousser un changement de normes au profit d'un programme donné.
- **Semer la discorde** : au cours des évènements ou situations particuliers (généralement politiques), des personnes ou des organisations partagent de fausses informations pour semer la confusion et la discorde au public. Ces pratiques peuvent aider à faire avancer un agenda d'une entité particulière.
- **Profit** : de nombreux acteurs de l'information fausse cherchent la popularité et le profit monétaire pour leurs organisations ou leurs plateformes média.
- **Passion** : un nombre considérable d'utilisateurs sont passionnés par une idée ou une organisation. Cela influence leur jugement et peut contribuer à la diffusion de fausses informations. Plus précisément, les utilisateurs passionnés sont aveuglés par leurs idéologies et aperçoivent les fausses informations comme correctes, et contribuent à leur propagation.
- **Amusement** : les trolls en ligne diffusent généralement de fausses informations pour leur propre amusement. Leurs actions peuvent parfois causer des dommages considérables aux autres personnes.


**Figure 1.1** : Les fausses informations partagées sur les réseaux sociaux : différents types de contenu, différentes motivations, différents acteurs[2]

[2] Ces exemples sont extrait de la page Facebook de *fact-checking* https://www.facebook.com/Toubal.TT/

## Conclusion

Une fausse information prend une multitude de formes allant d'une simple blague jusqu'à des compagnes bien organisés de propagande et de rumeurs. Les acteurs qui contribuent à la propagation de ce type d'informations ont des intentions et motivations diverses, et des finalités variées. Bien comprendre les différentes dimensions de ce phénomène peut aider à concevoir des solutions plus efficaces et plus convaincantes.

# Chapitre II

# Détection de fausses informations : revue de littérature

## Introduction

La croissance explosive des fausses informations et des fausses nouvelles en particulier, et son effet négatif sur tous les domaines ont augmenté la demande sur les études qui détectent la désinformation sur le Web et les médias sociaux.

Dans ce chapitre, on va identifier les principaux fondements théoriques sur lesquelles sont fondées ces études. Nous caractérisons ensuite l'ensemble des méthodes, techniques et stratégies qui sont habituellement appliquées dans ce vaste domaine de recherche.

## 1. Problématique

Laissons $a$ se référer à un article qui contient de fausses informations (publié sur les médias sociaux ou les média traditionnels). L'article $a$ comporte deux éléments essentiels, l' « **éditeur »** et le « **contenu »**. L'**éditeur** $\overrightarrow{p_a}$ inclut un ensemble de caractéristiques de profile définissant l'auteur de l'article (nom, domaine, âge …etc.). Le **contenu** $\overrightarrow{c_a}$ inclut un ensemble des attributs qui représentent le contenu de l'article (gros titre, texte, image, liens, …etc.).

La détection des fausses informations est généralement définie en tant qu'un problème de classification binaire. La tâche de détection consiste à prédire si l'article $a$ est une fausse nouvelle (information) ou non, i.e., $F : \varepsilon \rightarrow \{0 , 1\}$ tel que :

$$\begin{cases} F(a) = 1 \; si \; a \; contient \; de \; fausses \; informations \\ \qquad\quad 0 \quad sinon \end{cases}$$

Où $F$ est la fonction de prédiction que nous voulons entrainer [2].

## 2. Les méthodes et les techniques de détection de fausses informations

Les techniques de détection de fausses informations se distribuent sur quatre catégories principales [3] :

### 2.1. L détection basée sur l'expertise (FactChecking)

Cette technique est initialement développée en journalisme. Elle vise à évaluer l'authenticité des informations en comparant les connaissances extraites du contenu des nouvelles que l'on veut vérifier avec les faits réels (connaissances valides). La vérification du contenu peut être manuelle ou automatique.

### 2.1.1. La vérification manuelle

La vérification manuelle des informations s'effectue par des personnes qualifiées ou moins qualifiées :

- **La vérification manuelle par des experts :** elle s'appuie sur des experts du domaine pour vérifier le contenu des informations. Elle mène à des résultats très précis mais coûteuses.
- **La vérification manuelle auprès de la foule :** elle repose sur un grand nombre des utilisateurs réguliers agissant comme vérificateurs des faits. Ce type de vérification est difficile à gérer et moins crédible. Par conséquent, lors de la vérification des faits par la foule, il faut souvent filtrer les utilisateurs non crédibles et résoudre les résultats de vérification contradictoires.

### 2.1.2. La vérification automatique

Elle s'appuie sur les techniques d'extraction d'information (Information Retrieval), de traitement automatique du langage naturel (NLP) et de la théorie des graphes. Le processus global de cette vérification peut être divisé en deux étapes :

- **L'extraction des faits :** les connaissances sont souvent extraites du Web ouvert en tant que « faits bruts ». Ces données brutes et non structurées sont ensuite traitées et nettoyées pour construire une base de connaissances ou un graphe de connaissances qui incluent un ensemble de faits réels (connaissance correctes).
- **La vérification des faits :** l'authenticité de l'information à vérifier est évaluée en comparant les connaissances extraites à partir du contenu de cette information avec les faits stockés dans la base de connaissances construite ou le graphe de connaissances.

## 2.2. La détection basée sur le style de l'information

Tandis que les méthodes de détection basées sur les connaissances visent à évaluer l'authenticité de l'information, les méthodes basées sur le style cherchent à évaluer l'intention de l'information, i.e. existe-t-elle une intention d'induire le public en erreur ou non ? Ces méthodes étudient un ensemble de caractéristiques quantifiables qui peuvent bien représenter la déception dans les informations et la différencier de la vérité.

### 2.2.1. Les théories de style de déception

Intuitivement, le style de l'information trompeuse est différent de celui de l'information réelle. Des études psychologiques ont montré que les récits fondés sur la réalité diffèrent par leur contenu et leur qualité de ceux fondés sur l'imagination. Exemple des théories de déception les plus connues :

- **La théorie des 4 facteurs** : les mensonges sont exprimés d'une manière différente en termes de : l'excitation (niveau de provocation), le contrôle de comportement, l'émotion et le raisonnement.
- **Le réflexe de Semmelweis** : les individus tendent à rejeter des nouvelles preuves parce qu'elles contredisent leurs normes et convictions établies.
- **L'illusion de la vision asymétrique** : les gens perçoivent leurs connaissances comme supérieures à celles des autres.
- **L'effet de l'écho-chambre** : les croyances et les convictions sont amplifiées ou renforcées par la communication et la répétition dans un système fermé.

### 2.2.2. Les caractéristiques et modèles basés style

Le style de contenu est généralement représenté par un ensemble de caractéristiques quantifiables exploitées pour un apprentissage automatique. Ces caractéristiques peuvent être regroupées en fonction de :

- **Les attributs linguistiques :** les attributs linguistiques d'une fausse information peuvent être regroupés selon différentes dimensions : quantité (ex. nombre de mots, verbes, noms…etc.), complexité (ex. nombre moyen de mots par phrase), subjectivité (ex. pourcentage des verbes subjectifs), incertitude (ex. nombres des points d'interrogation), non-immédiateté (ex. pourcentage de l'utilisation de la voix passive), émotion (ex. pourcentage des mots négatifs), diversité (ex. répétition), simplicité (ex. fautes d'orthographe), spécificité (ex. termes exclusifs) et lisibilité (ex. indexe de Flesch-Kincaid et Gunning-Fog).
- **La structure linguistique :** les caractéristiques issues de la structure linguistique décrivent le style de contenu selon quatre niveaux linguistiques : lexique, syntaxe, sémantique et discours. La quantification de ce type de caractéristiques est basée principalement sur les techniques de traitement automatique de la langue (NLP).

### 2.2.3. Les stratégies de détection de la déception

Une stratégie commune de détection consiste à utiliser un vecteur de caractéristiques représentant le style de contenu de l'information dans un cadre d'apprentissage automatique pour prédire si l'information est trompeuse (problème de classification) ou comment est-elle trompeuse (problème de régression). La plupart des études utilisent des techniques d'apprentissage supervisé où l'étiquetage initial des données (c.-à-d. information fausse, information vraie) est nécessaire.

## 2.3. La détection basée sur la propagation de l'information

Les méthodes basées sur la propagation de l'information étudient les données reliées à la diffusion de l'information fausse, c.-à-d. comment propagent-elles et comment sont-elles diffusées par les utilisateurs.

### 2.3.1. Les techniques basées sur la propagation en cascade

La plupart des études ont adopté le modèle en cascade comme une représentation formelle pour la propagation de l'information fausse. Un modèle en cascade est une structure arbre ou pseudo-arbre qui représente la propagation de la fausse information dans les réseaux sociaux. Le nœud racine représente l'utilisateur qui a publié la fausse information pour la première fois (le créateur) ; les autres nœuds représentent les utilisateurs connectés au nœud racine qui ont partagés et republié la fausse information par la suite.

Détecter l'information fausse consiste à calculer la similarité de son modèle en cascade avec le modèle en cascade d'une autre information fausse (ou réelle).

### 2.3.2. Les techniques basées sur le réseau de la diffusion

Les méthodes basées sur le réseau de diffusion construisent des réseaux flexibles pour capturer indirectement la propagation de la fausse information. Ces réseaux contiennent différents types de nœuds (ex. éditeurs de l'information, diffuseurs de l'information, information) reliés entre eux par des relations de différents types. Ces relations sont analysées dans le but d'attraper la fausse information en utilisant des méthodes multiples telles que les modèles probabilistes, l'apprentissage multitâche et les algorithmes PageRank.

### 2.4. La détection basée sur l'évaluation de la crédibilité

Les méthodes de cette catégorie supposent que l'authenticité de l'information dépend la crédibilité des données reliées à elle, en particulier :

- **L'évaluation de la crédibilité de titre de l'information :** consiste à évaluer la crédibilité des titres de nouvelles par la détection des titres accrochants (en anglais, *clickbaits*). Les études actuelles de détection utilisent des caractéristiques linguistiques et non linguistiques dans un cadre d'apprentissage supervisé. La détection basée sur l'apprentissage profond a également émergé récemment.
- **L'évaluation de la crédibilité de la source de l'information :** consiste à exploiter les techniques classiques de classement Web (Web Ranking) tels que les algorithmes PageRank et HITS pour évaluer la crédibilité d'une source d'information.
- **L'évaluation de la crédibilité des commentaires sur l'information :** les commentaires des utilisateurs sur une nouvelle sur les sites Web et les médias sociaux contiennent des informations précieuses sur la valeur de cette nouvelle. Evaluer les commentaires pour vérifier la fiabilité d'une information a été longtemps utilisé pour la détection de fausses évaluations des produits sur les sites e-commerce.
- **L'évaluation de la crédibilité des diffuseurs de l'information :** la crédibilité des diffuseurs d'une information souvent reflète sa crédibilité elle-même. Les utilisateurs peuvent être regroupés en utilisateurs malveillants, avec une faible crédibilité et utilisateurs normaux, avec une crédibilité relativement élevée.

## Conclusion

Les mécanismes pour la détection de fausses informations peuvent être regroupés selon quatre perspectives : l'expertise, le style d'écriture, la propagation et la crédibilité. Chacune de ces techniques possède des avantages et des points faibles. On peut utiliser également des approches hybrides fusionnant ces techniques pour améliorer la tâche de la détection. D'autres perspectives peuvent être exploitées prochainement vue la nouveauté et l'importance de ce domaine de recherche.

# Chapitre III

# Conception d'un modèle de détection de fausses informations

## Introduction

Dans le but de mettre techniquement en évidence la tâche de la détection de fausses informations, nous présentons dans ce chapitre la conception d'un cas d'étude concret de classification du texte basé sur l'apprentissage supervisé, et appliqué sur les fausses informations au sujet de la nouvelle pandémie de Coronavirus.

Cette étude est une modeste contribution aux efforts de sensibilisation contre les compagnes acharnées de désinformation, qui ont accompagné la crise sanitaire de la Covid-19 en mettant la vie de nombreuses personnes dans le monde en danger.

## I. Description et objectif de projet

Notre projet est une application directe de la détection de fausses informations sur le web et les plateformes en ligne en utilisant l'apprentissage automatique (Machine Learning) et l'apprentissage approfondi (Deep Learning) (voir Algorithme 3.1 et Algorithme 3.2).

Parmi les différentes approches de détection mentionnées précédemment dans la littérature (voir le Chapitre.2), nous avons suivi une approche basée sur le style de l'information, ce qui se croise aussi avec les techniques de traitement automatique du langage naturel (NLP).

Les informations à examiner et classifier concernent le premier sujet de l'an 2020 ; la nouvelle pandémie du Coronavirus qui a apparu en décembre 2019 en touchant une partie majoritaire de la planète. La désinformation à ce sujet a suscité des potentiels dommages indirects sur la santé publique et des mouvements de panique inappropriés. Les causes sont toujours les rumeurs, les théories du complot, les fausses mesures de prévention et les faux remèdes.

La figure 3.1 illustre la démarche de conception que nous avons suivi pour la réalisation de la tâche de détection de fausses informations.

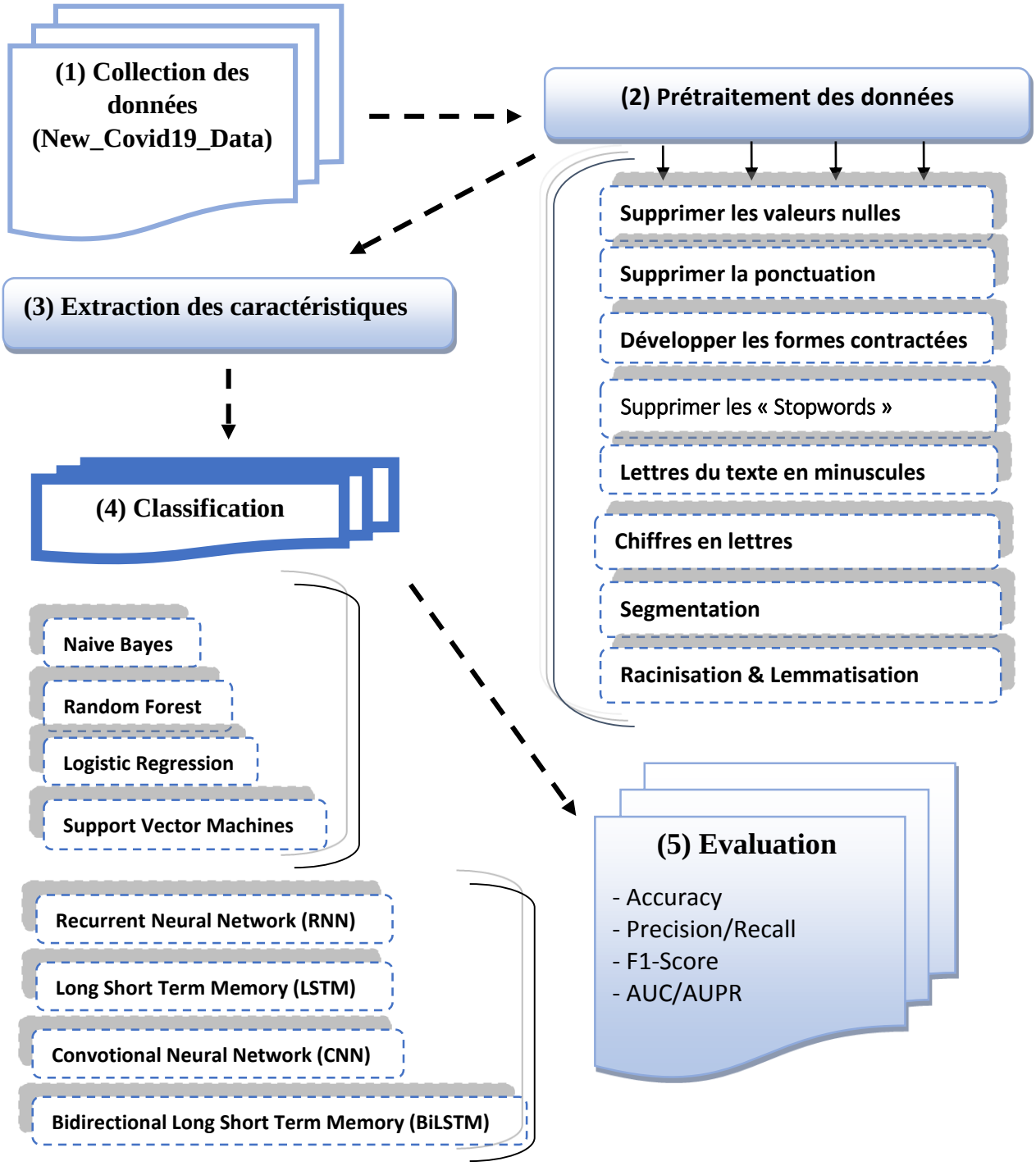



**Figure 3.1** : Processus de réalisation de système de détection de fausses informations

# II. L'apprentissage automatique et l'apprentissage approfondi pour la classification du texte

## 1. L'apprentissage automatique (Machine Learning)

L'apprentissage automatique (ML : Machine Learning) est un sous-domaine de l'intelligence artificielle, qui donne à un système une capacité de compréhension grâce à ses algorithmes. Il est basé sur l'idée de faire apprendre des algorithmes à partir de données et de faire des prédictions avec ces données et par cela les ordinateurs apprennent à résoudre des tâches spécifiques, sans avoir besoin de les programmer. Il existe trois types ou catégories de l'apprentissage machine [4] :

i. L'apprentissage supervisé où, après avoir présenté les données et les résultats souhaités aux ordinateurs, ils auront la capacité de faire des prédictions pour de nouvelles données d'entrée.
ii. L'apprentissage non supervisé où l'on ne donne à l'ordinateur que les données et où il doit trouver une structure avec un sens par lui-même sans l'intervention d'une supervision extérieure. Elle dépend principalement du clustering.
iii. L'apprentissage par renforcement où la machine se comporte comme un agent qui apprend de son environnement d'une manière interactive jusqu'à ce qu'il découvert les comportements qui produisent des récompenses.

Nous avons déjà défini la détection de fausses informations comme une tâche de classification binaire (voir Chapitre 2). Par conséquent, notre conception automatiquement suit le premier type de l'apprentissage : l'apprentissage supervisé.

## 2. La classification du texte (Text Classification)

La classification des textes est une tâche générique de traitement automatique de la langue naturelle qui consiste à assigner une ou plusieurs catégories, parmi une liste prédéfinie ou non à un document en trouvant une liaison fonctionnelle entre un ensemble de textes et un ensemble de catégories (étiquettes, classes) selon des critères.

Classifier des textes consiste souvent en cinq étapes : collection de données, prétraitement de données, extraction des caractéristiques, classification et évaluation du modèle de classification [5].

**Algorithme 3.1** : Algorithme d'entrainement de classification de fausses informations

**input** : New_Covid19_Data

**output** : ModèleClassificationEntrainé

tf-idf = []

bow = []

**pour** $article \in New\_Covid19\_Data$ **faire**

    prétraitement (article)

    segmentation (article)

    stemming (articles) // ou lemmetisation (article)

    bow += BoW (article)

    tf-idf += TF-IDF(article)

**fin**

ModèleClassificationEntrainé = AlgorithmeClassification (bow, tf-idf)

valider (ModèleClassificationEntrainé)

**retourner** ModèleClassificationEntrainé

**Algorithme 3.2** : Algorithme de test de détection (prédiction) de fausses informations

**input** : texte_à_classifier, ModèleClassificationEntrainé

**output** : classe_du_texte (*Fake* ou *True*)

prétraitement (texte_à_classifier)

segmentation (texte_à_classifier)

stemming (texte_à_classifier) // ou lemmetisation (texte_à_classifier)

Prédiction = Charger (ModèleClassificationEntrainé)

classe_du_texte = Prédiction (texte_à_classifier, ModèleClassificationEntrainé)

évaluer (ModèleClassificationEntrainé)

**retourner** classe_du_texte

## 2.1. La collection des données

Cette étape consiste à collecter les données brutes destinées à la classification depuis des sources crédibles de données.

Pour notre projet, nous avons collecté un ensemble de données contenant des informations au sujet de la nouvelle pandémie de Covid-19. Les fausses informations sont étiquetées « Fake » et les vraies informations sont étiquetées « True ». L'ensemble de données est essentiellement une concaténation de deux autres ensembles disponibles publiquement en ligne :

- L'ensemble de données **CoAID** (**C**ovid-19 **H**ealthcare **M**isinformation **D**ataset) [32] : un nouvel ensemble de données incluant des faits et de mésinformation sur Covid-19 collectés entre 01 décembre 2019 et 01 mai 2020 depuis des journaux et plateformes sociaux en ligne. Nous avons choisi une partie assemblant les fausses et les vraies prétentions en plus des fausses et des vraies nouvelles[3].
- Un ensemble de plus de 1,100 données publiques fausses et vraies, collectés à partir des articles des journaux et des publications sur les réseaux sociaux[4].

De plus amples informations et statistiques sur l'ensemble de données résultant (New_Covid19_Data) seront détaillées dans le chapitre 4.

## 2.2. Le prétraitement de donnés

Le pré-traitement linguistique est un outil puissant pour préparer les données textuelles au traitement automatique. Généralement, les prétraitements comprennent la normalisation textuelle des mots, la normalisation linguistique et la segmentation.

La segmentation (en anglais *tokenisation*) consiste à séparer une suite de caractères en éléments sémantiques, ou mots (en anglais *tokens*). La normalisation est une opération qui consiste à fournir une forme canonique pour chaque mot. Deux types de normalisation à distinguer :

- **La normalisation textuelle** ou ***surfacique*** qui consiste à effectuer quelques transformations superficielles sur les séquences de caractères de ces mots. Prenant par exemple : supprimer les valeurs nulles et les mots vides (en anglais *stopwords*), éliminer la ponctuation, mettre les lettres du texte en minuscules et transformer les chiffres en lettres, développer les formes contractées (exemple : *can't* devient *can not*) …etc.
- **La normalisation linguistique** avec ses deux types. **La racinisation** (en anglais *stemming*) se rapporte au procédé qui cherche à supprimer les flexions et les suffixes des mots. Il est fortement dépendant de la langue utilisée. **La lemmatisation** (en anglais *lemmatization*) fait une analyse linguistique poussée destinée à enlever les variantes

[3] https://github.com/cuilimeng/CoAID
[4] https://raw.githubusercontent.com/susanli2016/NLP-with-Python/master/data/corona_fake.csv

flexionnelles des mots afin de les ramener sous leur forme lemmatisée ou encyclopédique.

## 2.3. L'extraction des caractéristiques

La vectorisation du texte (en anglais *word embedding*) est un ensemble de techniques de traitement du langage naturel où les mots ou les phrases sont transformées à des vecteurs numériques. Les deux techniques de vectorisation les plus utilisées sont :

- **Bag Of Word (BoW)** : le modèle de sac-à-mots est une représentation simplificatrice. Il repose sur le principe de nombrer les occurrences des mots composant le texte à traiter.
- **Term Frequency-Inverse Document Frequency (TF-IDF) :** est une méthode d'évaluation de la pertinence d'un document par rapport à un terme, en tenant compte de deux facteurs : la fréquence de ce mot dans le document (TF) et le nombre de documents contenant ce mot (IDF) dans le corpus étudié.

$$tf - idf_{t,d} = tf_{t,d} \times idf_t = tf_{t,d} \times \log\left(\frac{N}{df_{t,d}}\right)$$

Tel que : $tf_{t,d}$ désigne la fréquence d'un terme ou mot $t$ dans le document $d$, $df_{t,d}$ le nombre de documents $d$ contenant le terme $t$, $N$ est le nombre total de documents dans le corpus.

## 2.4. La classification

Il existe différents classificateurs d'apprentissage supervisé qui sont utilisés pour la classification du texte. Les modèles de base sont les suivants :

- **Naïve de Bayes (Naive Bayes)**

Cette méthode se base sur le théorème de Bayes permettant de calculer les probabilités conditionnelles. Dans le cas de la classification du texte, la méthode Naïve Bayes est utilisée comme suit : on cherche la classification qui maximise la probabilité d'observer les mots du document. Lors de la phase d'entraînement, le classificateur calcule les probabilités qu'un nouveau document appartient à telle catégorie à partir de la proportion des documents d'entraînement appartenant à cette catégorie. Il calcule aussi la probabilité qu'un mot donné soit présent dans un texte, sachant que ce texte appartient à telle catégorie. Quand un nouveau

document doit être classé, on calcule les probabilités qu'il appartienne à chacune des catégories à l'aide de la règle de Bayes La formule [6] :

$$P(A/B) = \frac{P(A \cap B)}{P(B)}$$

- **La régression logistique (Logistic Regression)**

La régression logistique est essentiellement un algorithme de classification supervisé. Dans un problème de classification, la variable cible (ou sortie), $y$, ne peut prendre que des valeurs discrètes pour un ensemble donné de caractéristiques (ou d'entrées), $X$. Le modèle construit un modèle de régression pour prédire la probabilité qu'une entrée de données donnée appartienne à la catégorie numérotée « 1 ». La régression logistique modélise les données à l'aide de la fonction sigmoïde [7].

- **Les machines à support de vecteurs (SVM, Support Vector Machines)**

Le but de SVM est de trouver un classificateur qui sépare au mieux les données et maximise la distance entre ces deux classes. Ce dernier est un classificateur linéaire appelé hyperplan. Comme montré dans la figure ci-dessous (figure 3.2), cet hyperplan sépare les deux ensembles de points. Les points les plus proches, qui seuls sont utilisés pour la détermination d'hyperplan, sont appelés vecteurs de support [6].

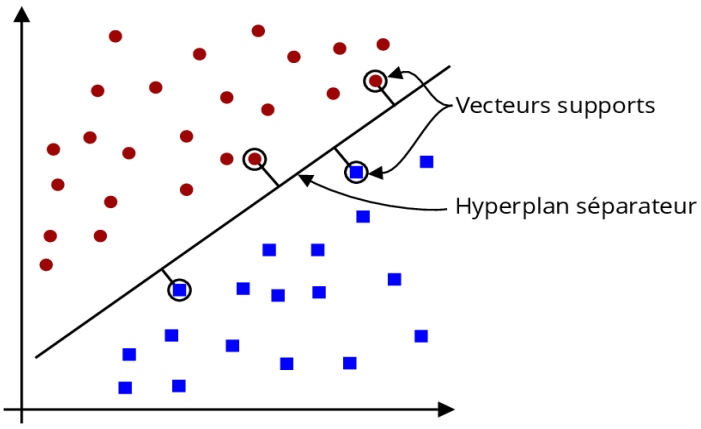



**Figure 3.2** : Représentation graphique de l'algorithme SVM

- **Les forêts aléatoires (Random Forest)**

C'est une application de graphe en arbres de décision permettant ainsi la modélisation de chaque résultat sur une branche en fonction des choix précédents. On prend ensuite la

meilleure décision en fonction des résultats qui suivront. On peut considérer ceci comme une forme d'anticipation [8].

## 2.5. L'évaluation

La performance des modèles de classification est généralement basée sur la façon dont ils prédisent les résultats pour les nouveaux ensembles de donnés. Cette performance est mesurée par rapport à un ensemble de test. Plusieurs métriques déterminent les performances de prédiction d'un modèle, mais nous allons principalement se concentrer sur les métriques suivantes [6] :

- **La matrice de confusion :** une matrice de confusion est un résumé des résultats de prédiction sur un problème de classification. Le nombre de prédictions correctes et incorrectes est résumé avec des valeurs de comptage et ventilé par classe sous forme tabulaire comme la montre le tableau suivant :

| ***Ensemble des catégories*** | **Documents appartenant à la catégorie** | **Documents n'appartenant pas à la catégorie** |
|---|---|---|
| *Documents assignés à la catégorie par le classifier* | True Positif (TP) | False Negative (FN) |
| *Documents rejetés à la catégorie par le classifier* | False Positif (FP) | True Negative (TN) |

**Tableau 3.1** : Matrice de confusion

On définit à partir des statistiques de cette table les mesures suivantes :

- **Le taux de succès ou d'erreur** (en anglais *accuracy*) : désigne le taux des prédictions réussites obtenu par le modèle de classification. C-à-d :

$$Accuracy = \frac{TP + TN}{\mathrm{TP} + \mathrm{TN} + \mathrm{FP} + \mathrm{FN}}$$

- **La précision** (en anglais *precision*) : est définie comme le nombre de prédictions faites qui sont réellement correctes ou pertinent parmi toutes les prédictions basées sur la classe positive. Ceci est également connu comme **valeur prédictive** positive et peut être représentée par la formule :

$$Precision = \frac{TP}{\mathrm{TP} + \mathrm{FP}}$$

- **Le rappel** (en anglais *recall*) : est défini comme le nombre d'instances de la classe positive qui étaient correctement prédit. Ceci est également connu sous le nom de couverture ou de sensibilité et peut être représenté par la formule :

$$Recall = \frac{TP}{\mathrm{TP} + \mathrm{FN}}$$

- **F1-Score** : est une autre mesure de précision qui est calculée en prenant la moyenne harmonique de la précision et du rappel et peut être représentée comme suit :

$$\mathrm{F1-Score} = \frac{2 * Precesion * Recall}{\mathrm{Precision} + \mathrm{Recall}}$$

## 3. L'apprentissage approfondi (Deep Learning)

Un réseau de neurones artificiels (Artificial Neural Network) est un système informatique s'inspirant du fonctionnement du cerveau humain pour apprendre**.** Les réseaux de neurones artificiels sont conçus pour reproduire certaines caractéristiques des mémoires biologiques par le fait qui sont massivement parallèles, capables d'apprentissage, capables de mémoriser l'information dans les connexions inter-neurones et capables de traiter des informations incomplètes [8].

### 3.1. Le modèle mathématique des réseaux de neurones

Le neurone artificiel (Perceptron) est l'élément de base dans un réseau de neurone artificiel. Chaque neurone artificiel reçoit un nombre variable d'entrées $X$. A chacune de ces entrées est associée un poids synaptique $W$. Chaque neurone est doté d'une sortie unique, qui se ramifie ensuite pour alimenter un nombre variable de neurones avals [9].

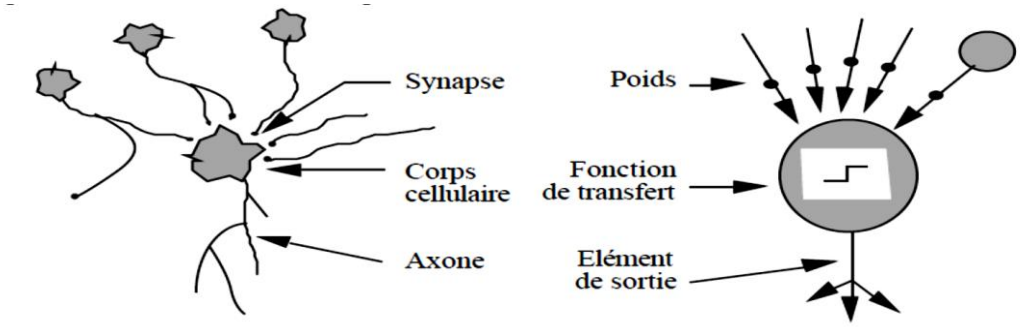



**Figure 3.3 :** Le neurone artificiel

Dans l'exemple de la figure 3.4, on peut interpréter la décision du perceptron comme classe 1 si la valeur de $x$ est $+1$ et classe 2 si la valeur de $x$ est $-1$. Les connexions des deux entrées $e1$ et $e2$ au neurone sont pondérées par les poids $w1$ et $w2$. La valeur de sortie du neurone est notée $x$. Elle est obtenue après somme pondérée des entrées $(a)$ et comparaison à une valeur de seuil $S$ [9].

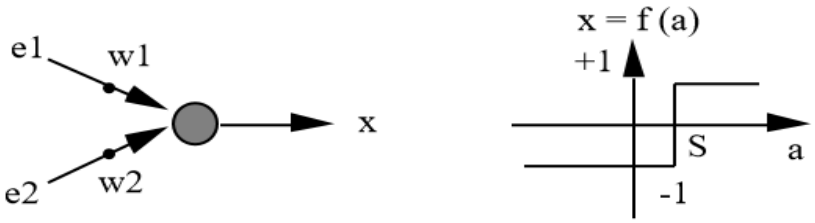



**Figure 3.4 :** Structure et comportement d'un perceptron

## 3.2. Les réseaux de neurones multicouches

- **Les réseaux de neurones à convolution (Convotional Neural Network CNN)**

Le CNN est une classe du Deep Neural Network qui est largement utilisé pour le NLP. Durant le processus d'apprentissage, les éléments constitutifs du réseau sont alternés pour atteindre la performance optimale pour classifier les textes aussi précisément que possible. Les réseaux de neurones à convolution excellent dans l'apprentissage de la structure spatiale des données d'entrée. Un réseau de neurone à convolution se distingue par les couches suivantes :

- **La couche de vectorisation** où les textes à classifier seront vectorisées.
- **La couche à convolution** qui sert à extraire les caractéristiques essentielles des donnés en entrée.
- **La couche « Pooling » :** *pooling* c'est réduire les dimensions de chacune des caractéristiques et conserver les caractéristiques les plus importants des donnés. Le *pooling* peut être de différents types : Max (maximum), Average (moyenne), Sum (somme)…etc. Dans notre conception, nous avons utilisé le *Max Pooling*.
- **La couche « Flatten » :** dans cette couche, on a besoin de convertir la sortie de la partie convolution et *pooling* en un vecteur de caractéristiques unidimensionnel (01 dimension).
- **La couche entièrement connectée** : désigne le réseau de neurones multicouches liée aux couches précédentes. Il prend comme entrée le résultat final des trois couches précédentes afin de réaliser la tâche de classification demandée.

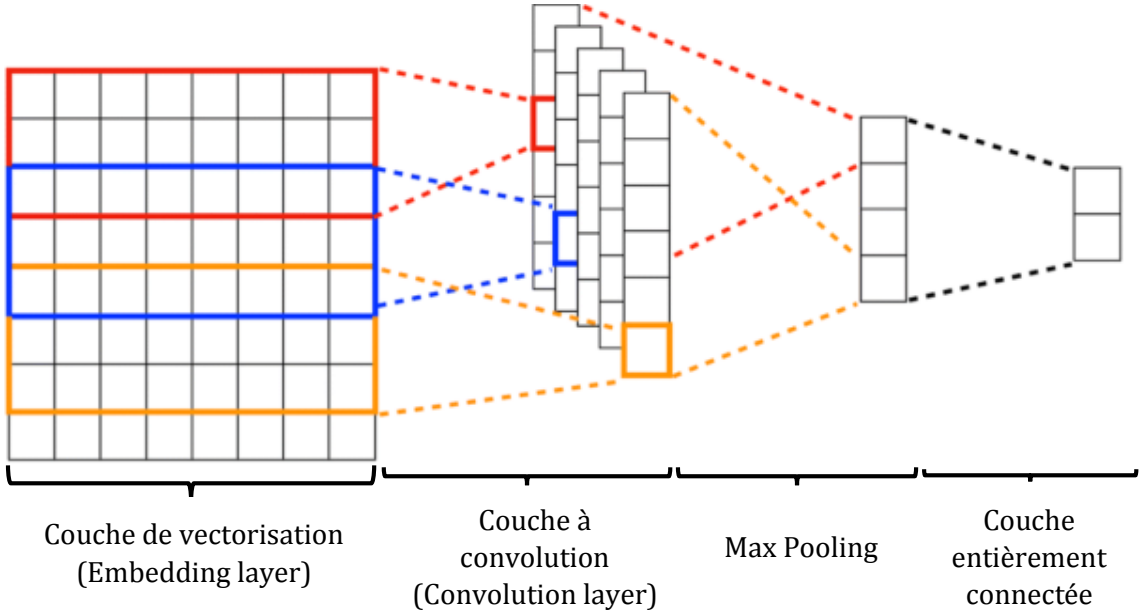



**Figure 3.5** : Structure d'un réseau de neurones à convolutions

- **Les réseaux de neurones récurrents (Recurrent Neural Network RNN)**

Un réseau neuronal récurrent (RNN) est une classe de réseau neuronal artificiel où les connexions entre les nœuds forment un graphe orienté le long d'une séquence. Cela lui permet de présenter un comportement temporel dynamique pour une séquence temporelle. Les RNNs ont des boucles de rétroaction dans lesquelles la sortie du déclenchement précédent ou de l'indice de temps $T$ est fournie comme l'une des entrées à l'indice de temps $T + 1$. Il peut y avoir des cas dans lesquels la sortie du neurone est alimentée à lui-même en tant qu'entrée [10].

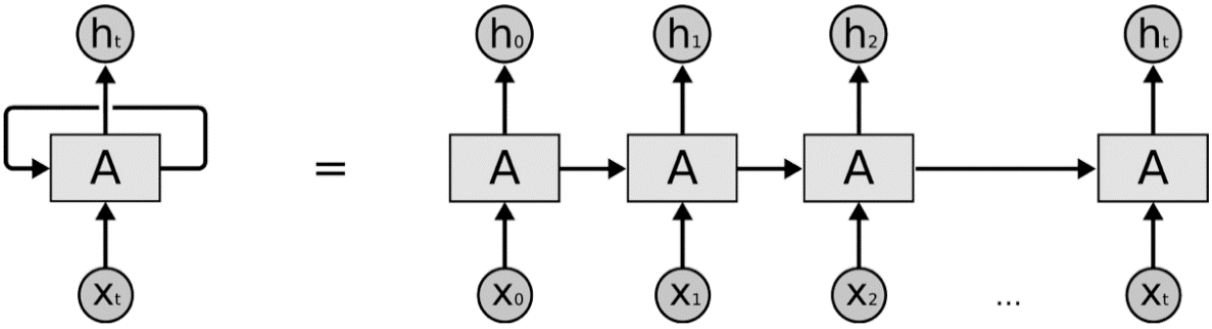



**Figure 3.6 :** Structure d'un réseau de neurones récurrent

- **Les réseaux de neurones LSTM (Long Short Term Memory)**

Un réseau LSTM c'est le type d'architecture RNN qui aide à former le modèle sur de longues séquences et dans la rétention de la mémoire des étapes de temps précédentes de l'entrée fournie au modèle. Idéalement, il résout le problème de la disparition du gradient ou l'explosion

du gradient, en introduisant des portes supplémentaires, qui permettent un meilleur contrôle du gradient, en permettant quelles informations à conserver et ce qu'il faut oublier. Ils contrôlent ainsi l'accès aux informations sur l'état actuel de la cellule, ce qui permet une meilleure conservation des « dépendances à longue portée ». Un LSTM se compose de :

- **Porte d'entrée (Input Gate)** : pour contrôler la contribution d'une nouvelle entrée dans la mémoire.
- **Porte d'oubli (Forget Gate)** : pour contrôler la limite jusqu'à laquelle une valeur appartenait à la mémoire.
- **Porte de sortie (Output Gate)** : pour contrôler la contribution d'une valeur au bloc d'activation de la sortie [10].

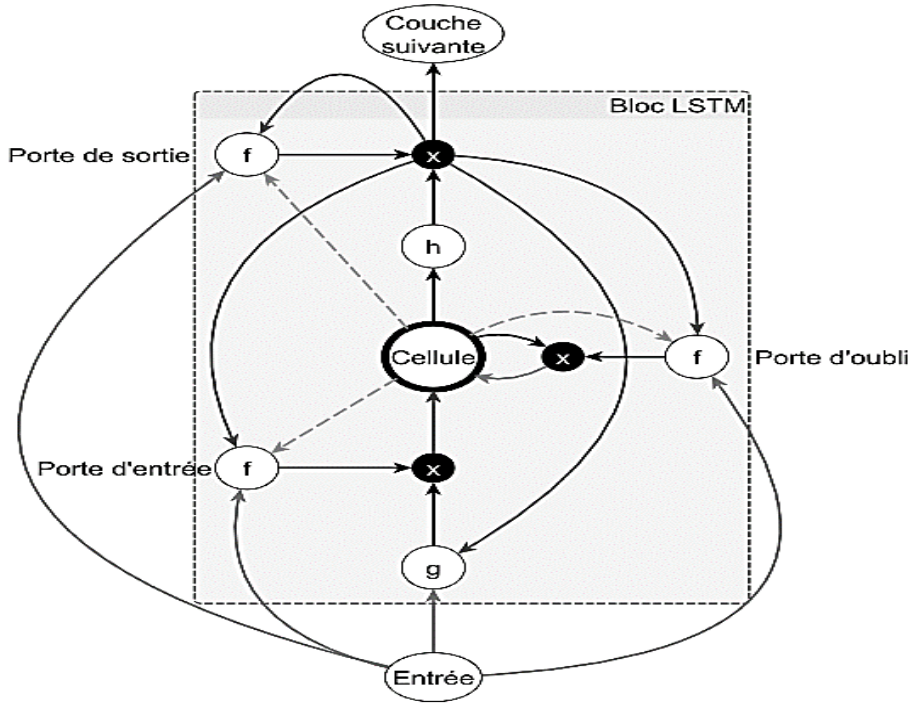



**Figure 3.7 :** Structure d'un LSTM

- **Les réseaux de neurones LSTM bidirectionnels (Bidirectional Long Short Term Memory) :**

Le réseau BLSTM est constitué de deux couches cachées « *Forward* » et « *Backward* », composées de blocs mémoires LSTM remplaçant les neurones récurrents de départ. Ces deux couches sont indépendantes sur la phase d'apprentissage, l'une donnant accès au passé et l'autre, au futur. Les blocs LSTM gèrent l'influence des informations qui sont diffusées dans le réseau. Puis la somme pondérée des activations des deux couches cachées est fournie en entrée de la couche de sortie pour chaque instant [11].

# III. La crise sanitaire du Coronavirus

## 1. Que ce que Covid-19 ?

D'après l 'OMS (Organisation Mondiale de la Santé), la Covid-19 (SRAS-CoV-2) est la maladie infectieuse causée par le dernier coronavirus qui a été découvert. Ce nouveau virus et cette maladie étaient inconnus avant l'apparition de la flambée à Wuhan (Chine) en décembre 2019. La Covid-19 est maintenant pandémique et touche de nombreux pays dans le monde [12].

## 2. La chronologie de la pandémie Covid-19

Ce qui suit sont les dates clés de la pandémie mondialement et localement dans l'Algérie [13][14] :

- **08 décembre 2019** : l'apparition du virus à Wuhan, en Chine.
- **13 janvier 2020** : un cas de Covid-19 en Thaïlande, premier cas signalé hors de Chine.
- **24 janvier 2020** : premiers cas européens, en France.
- **28 janvier 2020** : le premier cas de contamination au Proche-Orient est détecté à Dubaï dans les Emirates Arabes Unies.
- **30 janvier 2020** : l'OMS déclare le Coronavirus comme une urgence mondiale.
- **11 février 2020** : l'OMS annonce que la nouvelle maladie infectieuse sera baptisée Covid-19 (**Co**rona**vi**rus **d**isease 20**19**).
- **25 février 2020** : le premier cas de Covid-19 est confirmé en Algérie, chez un ressortissant italien arrivé dans le pays le 17 février. C'est le second pays africain touché par l'épidémie après l'Égypte.
- **11 mars 2020** : l'OMS déclare que la flambée de Coronavirus est une pandémie.
- **22 mars 2020** : le ministre de la Santé annonce que l'Algérie est passé au niveau trois de l'épidémie. De nouvelles mesures de confinement après une réunion du Haut Conseil de Sécurité sont annoncées le lendemain[5].
- **14 juin 2020** : les autorités algériennes annoncent l'assouplissement des horaires de confinement et la reprise de certaines activités commerciales.[6]

## 3. La chronologie de l'infodémie sur Covid-19

Parmi les évènements les plus influenceurs dans la chronologie de l'infodémie Covid-19, nous citons :

[5] https://itriinsights.com/coronavirus-algerie/
[6] https://itriinsights.com/coronavirus-algerie/

- **02 février 2020**: l'Organisation mondiale de la santé (OMS) a parlé d'une « infodémie massive » à propos du virus[7].
- **24 janvier 2020**: la BBC a publié deux articles du *Washington Times* qui affirmaient que le virus faisait partie d'un programme d'armes biologiques chinois. De nombreux messages sur les réseaux sociaux ont soutenu une théorie du complot selon laquelle le virus était connu et qu'un vaccin était déjà disponible [15].
- **Février 2020** : la BBC a rapporté que des groupes de théoriciens du complot sur les réseaux sociaux ont prétendu qu'il y avait un lien entre le coronavirus et les réseaux de 5G mobile [16].
- **06 février 2020** : selon l'Institut de recherche des médias du Moyen-Orient (MEMRI) nombreux écrivains dans la presse arabe ont soutenu la théorie selon laquelle la Covid-19 a été délibérément créé et diffusé par les États-Unis afin de vendre des vaccins contre ces maladies, et cela fait partie d'une guerre économique et psychologique menée par les États-Unis [17].
- **22 février 2020** : la Russie fait circuler des informations que « Coronavirus : une guerre biologique américaine contre la Russie et la Chine » [18][19][20].
- **09 mars 2020 :** l'ancien président iranien Mahmoud Ahmadinejad a envoyé une lettre aux Nations Unies affirmant que coronavirus a été produit en laboratoire pour maintenir une suprématie politique et économique mondiale [21].
- **Début mars 2020** : une fausse théorie stipule que le VIH (virus du Sida) aurait été combiné au SARS-CoV-2 [22].
- **17 mars 2020**, l'Institut Pasteur (situé en France) est la cible d'une vidéo conspirationniste l'accusant à tort, sur la base d'un brevet daté de 2004, d'avoir inventé le nouveau Coronavirus responsable de la pandémie de Covid-19[8].
- **Mars 2020** : des fausses informations ont laissé croire que certains animaux transmettent la Covid-19. Plusieurs informations ont circulé sur les réseaux sociaux concernant certains médicaments qui pourraient être dangereux (les antibiotiques sont inefficaces contre le virus) [16][23][24].

[7] https://www.who.int/fr/dg/speeches/detail/munich-security-conference
[8] https://www.sciencesetavenir.fr/fondamental/biologie-cellulaire/covid-19-non-le-coronavirus-sars-cov-2-n-est-pas-une-invention-de-l-institut-pasteur_142628

## 4. Les fausses informations les plus répandues sur Covid-19

La demande d'une information rapide et fiable a conduit à l'aggravation de la désinformation disséminée via les sites web et les médias sociaux. Parmi les fausses informations les plus répandues sur la Covid-19 :

- **Covid-19 est une arme biologique chinoise**

En janvier 2020, de nombreux quotidiens célèbres (Washington Times , Washington Post) ont publié des articles qui affirment que cette pandémie est une arme biologique chinoise basé à l'Institut de virologie de Wuhan. Cette théorie est ensuite réfutée par des experts y compris des américains qui disent que la plupart des pays ont abandonné les armes biologiques [15][25].

- **Covid-19 est une arme biologique américaine**

En février 2020, la Russie est accusée par des responsables américains de mener une campagne de désinformation, utilisant des milliers de comptes sur Twitter, Facebook et Instagram, pour faire circuler des théories du complot sans fondements affirmant que le virus serait une arme biologique fabriquée par la CIA. De nombreux media russes déclarant que le virus est destiné à nuire à l'économie chinoise, et de l'affaiblir lors du prochain cycle de négociations commerciales [18][19] [20].

- **Préexistence du vaccin**

De nombreux messages sur les réseaux sociaux ont soutenu une théorie du complot selon laquelle le vaccin était déjà disponible. L'OMS a rapporté à partir du 5 février 2020 que parmi les nouvelles de « percées » dans les médicaments qui ont été découverts pour traiter les personnes infectées par le virus, il n'y avait pas de traitement efficace connu et les brevets cités par les réseaux sociaux faisaient référence à des brevets déjà existants concernant des vaccins pour d'autres souches de coronavirus comme le coronavirus du SRAS [26][27].

- **Coronavirus propage via 5G**

la BBC a rapporté que des groupes de théoriciens du complot sur les réseaux sociaux ont prétendu qu'il y avait un lien entre le Coronavirus et les réseaux de 5G mobile, affirmant que la propagation à Wuhan était directement causée par des champs électro-magnétiques et par l'introduction de la 5G et des technologies sans fil [16].

## Conclusion

Les compagnes de désinformation sur la Covid-19 propagent rapidement et largement en parallèle avec la propagation de la pandémie elle-même. Vu la sensibilité de cette double crise sanitaire et informationnelle, nous avons proposé une étude suivant l'apprentissage automatique pour la classification et la détection de fausses informations dans ce sujet. Nous avons présenté les fonctions les plus importantes du prétraitement des textes et mentionné les différentes techniques d'extraction des caractéristiques. Ensuite, nous avons détaillé la tâche de classification en décrivant les algorithmes de base et les modèles de *Deep Learning*. Nous avons également introduit les différentes mesures d'évaluation de performance d'un classificateur. Dans le chapitre suivant les résultats de ce cas d'étude.

# Chapitre IV
# Implémentation et résultats

# Introduction

Après avoir décrit notre approche de conception dans le précèdent chapitre, le présent chapitre rapporte la mise en œuvre de la tâche de détection de fausses informations liées à la pandémie de la Covid-19 en se basant sur les deux approches d'apprentissage automatique et apprentissage profond. Nous déterminerons l'environnement et les outils de travail d'abord. Puis, après une analyse exploratoire nécessaire, nous listerons les étapes de la détection de fausses informations en les supportant par des illustrations statistiques ou graphiques. Nous clôturons ensuite par une discussion sur les résultats finaux et leurs implications.

## 1. L'environnement et les outils de travail

### 1.1. Le matériel

Nous avons réalisé notre projet à l'aide de deux postes de travail dont les caractéristiques sont décrites dans le tableau ci-dessous :

| | **Poste de travail N°01** | **Poste de travail N°02** |
|---|---|---|
| ***PC*** | DELL | TOSHIBA |
| ***Système d'exploitation*** | Windows 10 Professionnel | Windows 7 Edition Intégrale |
| ***Processeur*** | Intel(R) Core(tm) i5-4310M CPU @ 2.70ghz | Intel(R) Core(tm) i3-3217M CPU @ 1.80ghz |
| ***RAM*** | 8,00 Go | 4,00 Go |
| ***Type de système*** | SE 64 bits | SE 64 bits |

**Tableau 4.1** : Caractéristiques du matériel utilisé

### 1.2. Le langage de programmation

Nous avons choisi Python, version 3.7.4 pour implémenter notre système de détection de fausses informations. Python est un langage de programmation interprété, multi-paradigme et multiplateformes. Il est conçu pour optimiser la productivité des programmeurs en offrant des outils de haut niveau et une syntaxe simple à utiliser. Il est doté d'un typage dynamique fort, d'une gestion automatique de la mémoire par ramasse-miettes et d'un système de gestion d'exceptions. Le langage Python est placé sous une licence libre et fonctionne sur la plupart des plates-formes informatiques [28].

### 1.3. L'éditeur de code

Pour éditer le code de notre système, nous avons utilisé Jupyter Notebook, version 6.0.3, qui est une application web *open source* qui permet de créer et de partager des documents contenant du code en direct, des équations, des visualisations et du texte narratif. Les utilisations comprennent : le nettoyage et la transformation des données, la modélisation statistique, la visualisation des données, l'apprentissage automatique et bien plus encore [29].

### 1.4. Les librairies et les bibliothèques Python


**Figure 4.1 :** Editeur de code et bibliothèques Python utilisés

- **Pandas** : Pandas est une bibliothèque écrite pour le langage de programmation Python permettant la manipulation et l'analyse des données. Elle propose en particulier des structures de données et des opérations de manipulation de tableaux numériques et de séries temporelles.
- **Nltk :** est une plateforme Python leader pour travailler sur les données du langage naturel. Elle offre des interfaces faciles à manipuler et plus de 50 corpus et ressources lexiques. Nltk offre également des librairies de prétraitement de données, de classification, segmentation, racinisation, et plus d'autres.
- **Scikit-learn** : est une bibliothèque libre Python destinée à l'apprentissage automatique. Elle comprend de nombreuses méthodes pour la classification, la régression, et le clustering. Elle est conçue pour s'harmoniser avec d'autres bibliothèques libres Python, notamment NumPy et SciPy [30].

- **Keras** : Keras est un framework *open source* d'apprentissage profond pour le Python, capable de s'exécuter sur TensorFlow. Keras a été développé pour permettre des expérimentations rapides avec les réseaux de neurones profonds. Il se distingue par son extensibilité, compréhensibilité, rapidité et notamment sa simplicité [31].
- **Matplotlib :** est une bibliothèque de visualisation 2D de données, conçue pour Python. Elle offre des possibilités variées de visualisations statiques, personnalisées et interactifs dans des différents formats.

## 2. L'analyse exploratoire de données

### 2.1. Informations générales

Nous avons nommé notre ensemble de données « New_Covid_Data ». Comme nous l'avons déjà décrit dans le chapitre 03, cet ensemble est extrait à partir de deux autres ensembles de données disponibles publiquement. Pour notre objectif, nous avons gardé un corpus avec seulement trois attributs : Label (*Fake* ou *True*), Type (une publication sur les média sociaux « *Post* » ou un article de presse « *Article* »), et Statement (texte). Les caractéristiques de l'ensemble de données final sont décrites ci-après (tableau 4.2).

| ***Nom de l'ensemble de données*** | **New_Covid_Data** |
|---|---|
| ***Nombre de lignes*** | 3321 |
| ***Nombre de colonnes*** | 3 |
| ***Type de données*** | Textuelles |
| ***Noms des colonnes*** | (Label, Type, Statement) |
| ***Utilisation de la mémoire*** | 78 KB |
| ***Intervalle d'index*** | 3321 entrés de 0 à 3320 |
| ***Nombre de valeurs nulles*** | 0 |

**Tableau 4.2** : Caractéristiques de l'ensemble de données « New_Covid_Data)

### 2.2. La distribution de données

L'ensemble de données « New_Covid_Data » contient 1703 articles publiés sur le Web et 435 publications postées sur les médias sociaux, 2174 sont étiquetées « True », et 1147 sont étiquetées « Fake ». La figure 4.2 illustre la distribution de données selon les attributs « Type » et « Label ».

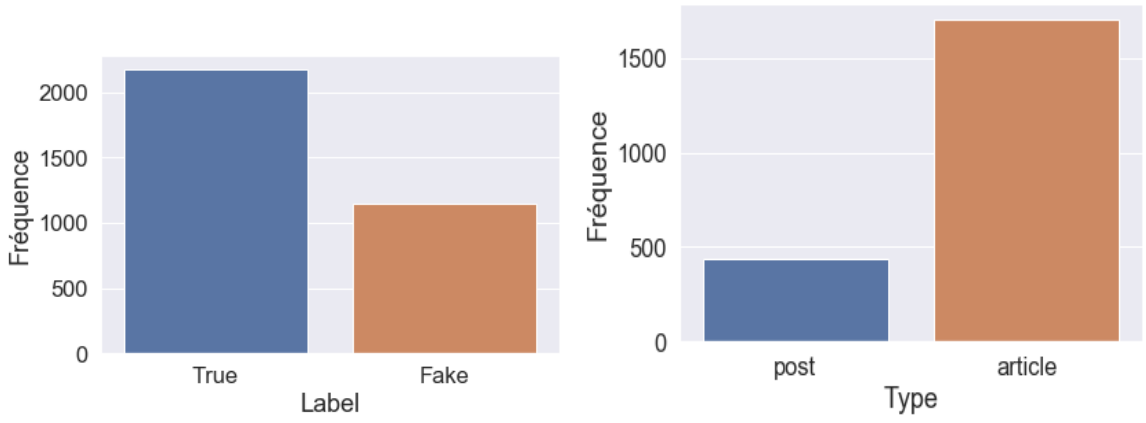



**Figure 4.2** : Distribution des classes et des types des données « New_Covid_Data »

## 2.3. Unigrammes, Bigrammes, Trigrammes

Dans le domaine du traitement automatique du langage naturel, analyser la fréquence des mots ou des phrases dans un corpus donné, constitue une étape indispensable pour comprendre la nature du problème de traitement automatique qu'on veut l'étudier. Dans notre étude, nous sommes intéressés particulièrement par les *unigrammes* (mots simples), les *bigrammes* (structure de deux mots consécutifs), et les *trigrammes* (structure de trois mots consécutifs). La figure 4.3 montre les mots les plus fréquents dans notre corpus. Les figures 4.4, 4.5, 4.6 et 4.7 montrent les bigrammes et les trigrammes fréquents dans les deux classes des données « *True* » et « *Fake* ».

En général, nous notons la ressemblance des termes utilisés dans les deux classes des données. Néanmoins, nous distinguons la fréquence remarquable des structures « *bill gates* », « *gates foundation* », « *bill melinda gates* », « *melinda gates foundation* » dans les données de classe « Fake ». En effet, depuis le début de l'épidémie de Coronavirus, le milliardaire américain Bill Gates et sa fondation « Bill & Melinda Gates » sont régulièrement accusés par les complotistes d'être à l'origine de la Covid-19. Chose que le co-fondateur de « *Microsoft* » a dénoncé plus tard en rappelant les investissements et les efforts de sa fondation contre les épidémies dans l'Afrique et le monde[9]. La fréquence des structures « *wuhan coronavirus* » et « *wuhun institute virology* » peut être interprétée par les doutes qui ont accompagné cet institut

[9] https://www.lefigaro.fr/flash-eco/bill-gates-balaie-les-theories-du-complot-l-accusant-de-la-pandemie-20200724

de virologie chinois accusé lui aussi par les désinformateurs d'avoir inventé le Coronavirus dans un cadre d'un armement biologique.

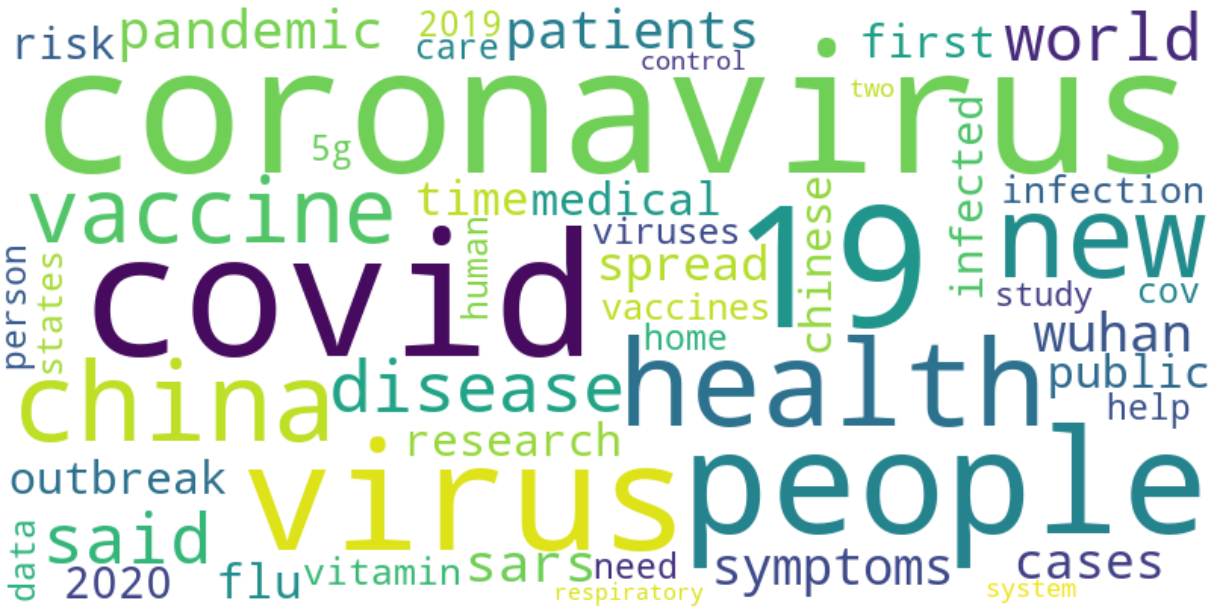



**Figure 4.3** : Les mots les plus fréquents dans l'ensemble de données « New_Covid_Data)

## 3. Le prétraitement

La phase de prétraitement de données textuelle est une étape clé pour un modèle de détection performant. Nous avons appliqué sur le corpus « New_Covid_Data » les fonctions de normalisation et prétraitement mentionnées précédemment (Chapitre 3). L'exemple suivant montre un texte brut avant et après la normalisation superficielle et la normalisation linguistique :

**Texte brut**

| |
|---|
| CORONAVIRUS OUTBREAK IN EUROPE: CRIMINAL NEGLIGENCE OR PREPLANNED ACTION. There is little doubt that the coronavirus threat is overestimated by mainstream media and governemnts. Covid-19 is in fact an ordinary viral disease with a slightly higher mortality from complications for people of old age or peopel with weakened immunity. Another open secret is that the current hysteria over the outbreak is being successfully used by some players to achieve their own economic and geopolitical goals. Looking at the current situation in Europe, one could suppose that some forces have seized an opportunity and are now fueling the coronavirus crisis intentionally. |

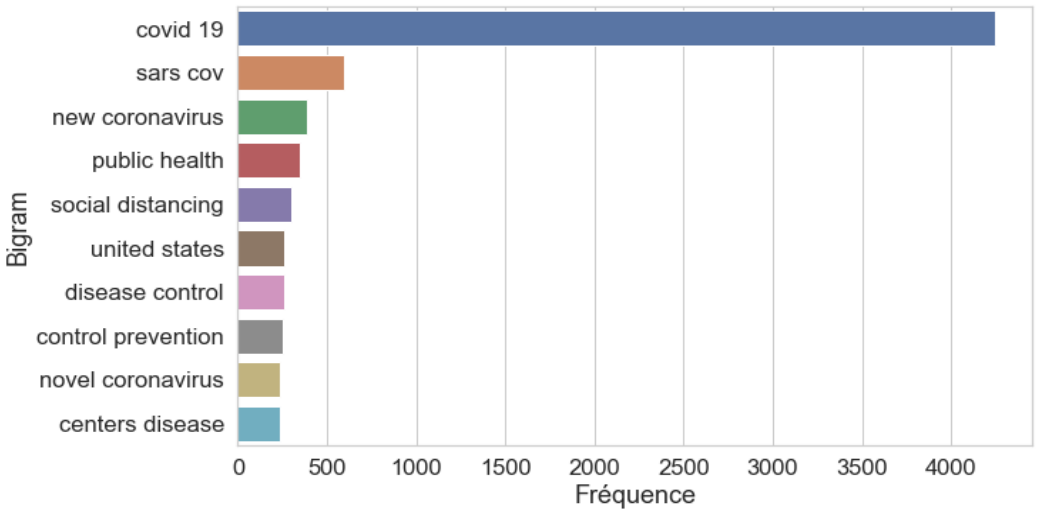



**Figure 4.4:** Top 10 *"True"* bigrams

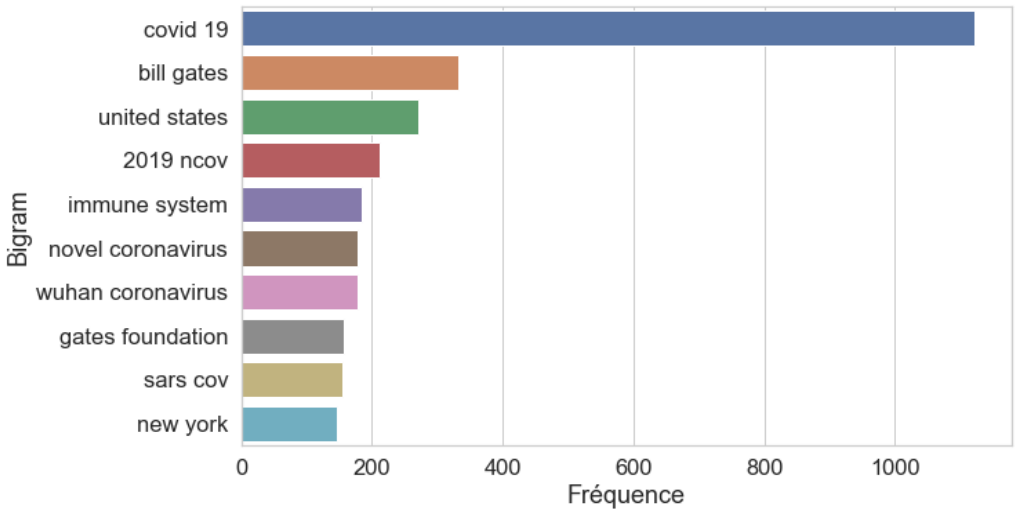



**Figure 4.5:** Top 10 *"Fake"* bigrams

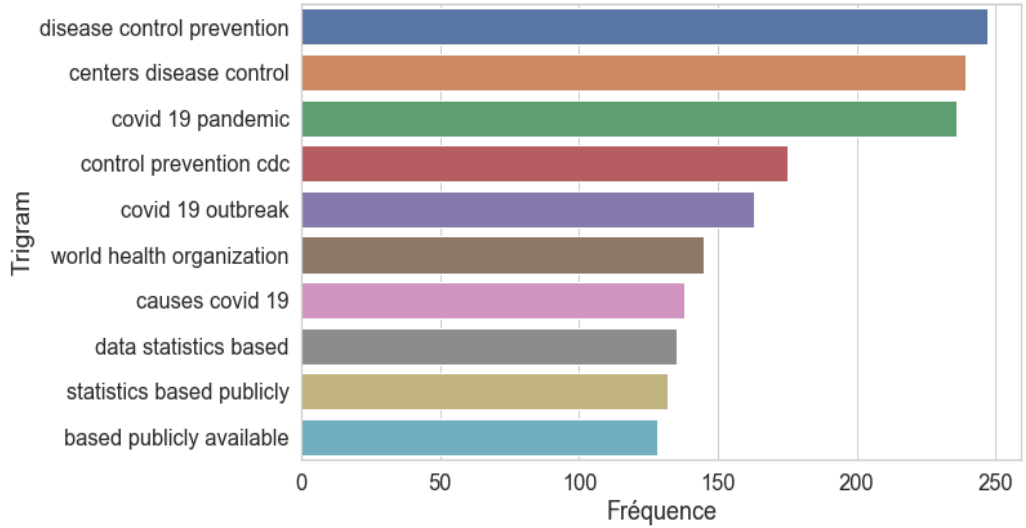



**Figure 4.6:** Top 10 *"True"* trigrams

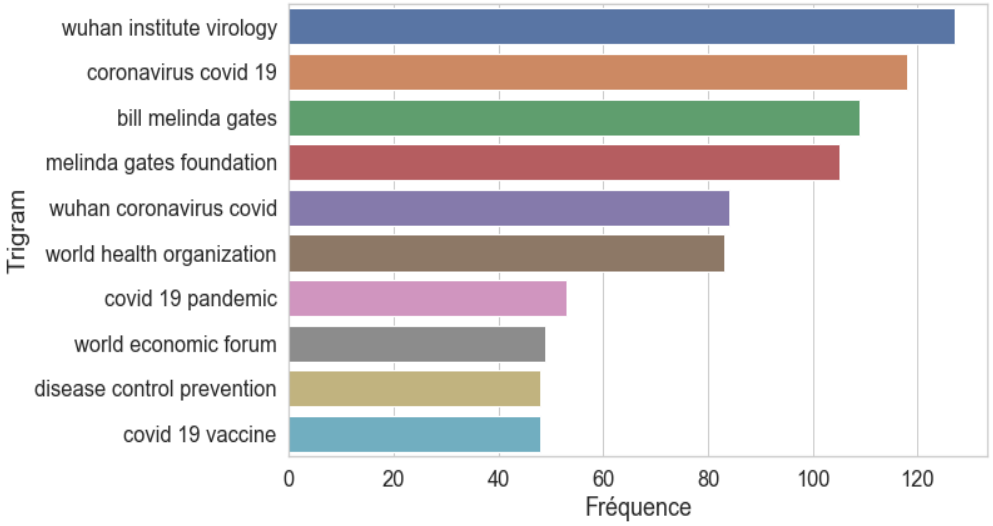



**Figure 4.7:** Top 10 *"Fake"* trigrams

**Texte normalisé**

['coronavirus', 'outbreak', 'europe', 'criminal', 'negligence', 'preplanned', 'action', 'there', 'little', 'doubt', 'coronavirus', 'threat', 'overestimated', 'mainstream', 'media', 'governemnts', 'covid19', 'fact', 'ordinary', 'viral', 'disease', 'slightly', 'higher', 'mortality', 'complications', 'people', 'old', 'age', 'peopel', 'weakened', 'immunity', 'another', 'open', 'secret', 'current', 'hysteria', 'outbreak', 'successfully', 'used', 'players', 'achieve', 'economic', 'geopolitical', 'goals', 'looking', 'current', 'situation', 'europe', 'one', 'could', 'suppose', 'forces', 'seized', 'opportunity', 'fueling', 'coronavirus', 'crisis', 'intentionally']

**Texte final prétraité**

['coronaviru', 'outbreak', 'europ', 'crimin', 'neglig', 'preplan', 'action', 'there', 'littl', 'doubt', 'coronaviru', 'threat', 'overestim', 'mainstream', 'media', 'governemnt', 'covid19', 'fact', 'ordinari', 'viral', 'diseas', 'slightli', 'higher', 'mortal', 'complic', 'peopl', 'old', 'age', 'peopel', 'weaken', 'immun', 'anoth', 'open', 'secret', 'current', 'hysteria', 'outbreak', 'success', 'use', 'player', 'achiev', 'econom', 'geopolit', 'goal', 'look', 'current', 'situat', 'europ', 'one', 'could', 'suppos', 'forc', 'seiz', 'opportun', 'fuel', 'coronaviru', 'crisi', 'intent']

## 4. Les résultats de classification

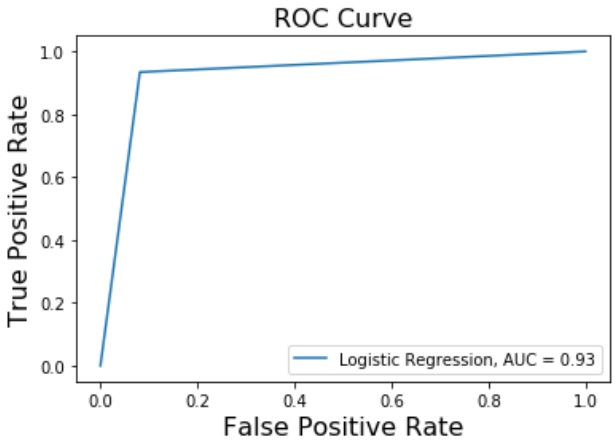
Pour classifier les textes de l'ensemble de données « New_Covid_Data », nous avons appliqué 08 modèles de classification. Quatre modèles de base : Naïve de Bayes, Régression logistique, Forêts aléatoires, et Machines à vecteurs de supports SVM. Quatre modèles d'apprentissage profond : Réseaux de neurones à convolution CNN, : Réseaux de neurones récurrents RNN, Réseaux de neurones à mémoire long-proche terme LSTM, et les LSTM bidirectionnels.

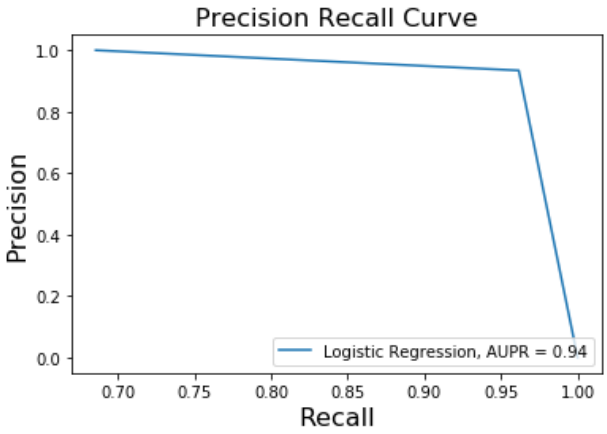
Les résultats de classification (tableau 4.3 et 4.4) sont évalués à l'aide des métriques : Précision (*Precision*), Rappel (*Recall*), F1 Score, et Taux de succès (*Accuracy*). Plus d'illustration en courbes ROC (Receiver Operating Characteristic) et PRC (Precision-Recall Curve), avec respectivement leurs sous-aires AUC (Area Under ROC Curve) et AUPR (Area Under Precision-Recall Curve) sont montrées aussi dans les figures 4.8 et 4.9.

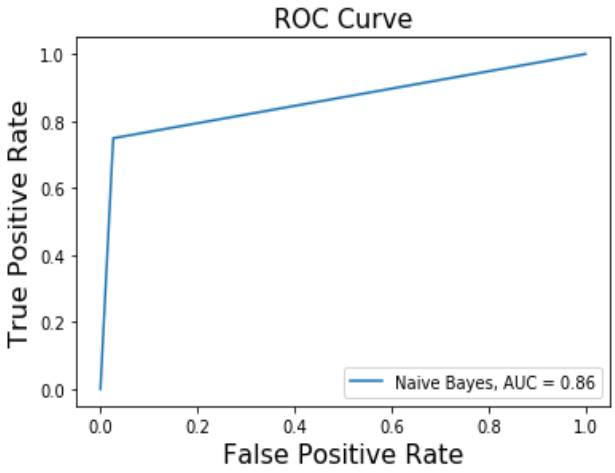
### 4.1. La classification par modèles de base

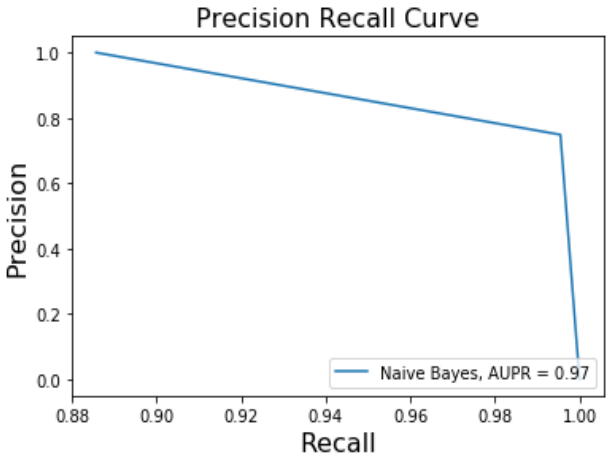
| | Precision | Recall | F1_Score | Accuracy | AUC | AUPR |
|---|---|---|---|---|---|---|
| **Logistic Regression** | 0.96 | 0.93 | 0.95 | 0.93 | 0.93 | 0.94 |
| **Naïve Bayes** | 1.00 | 0.75 | 0.85 | 0.77 | 0.86 | 0.97 |
| **SVM** | 0.96 | 0.94 | 0.95 | 0.94 | 0.93 | 0.94 |
| **Random Forests** | 0.93 | 0.93 | 0.93 | 0.91 | 0.90 | 0.92 |

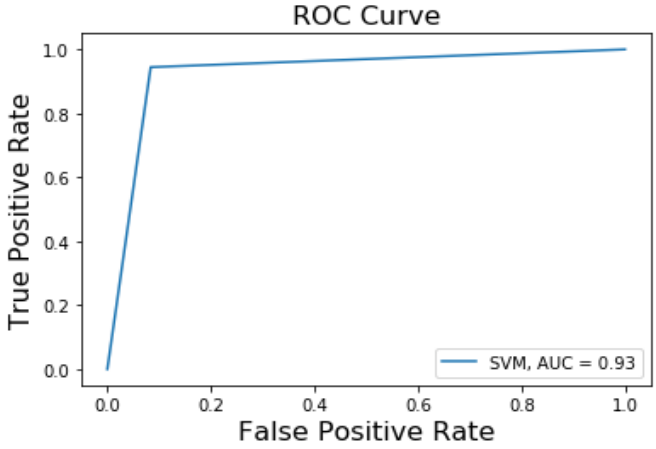
**Tableau 4.3** : Résultats de classification par modèles de base

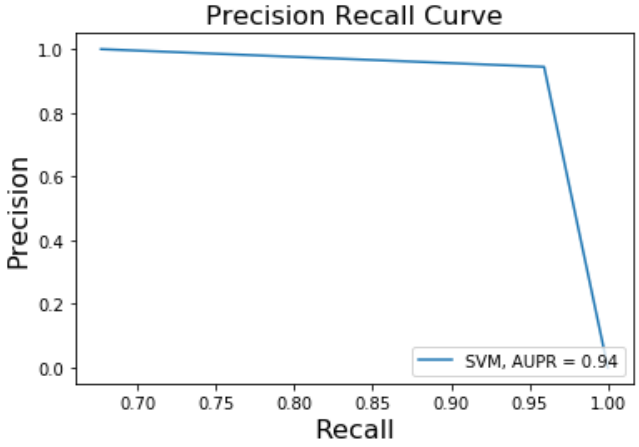

ROC Curve
True Positive Rate
False Positive Rate
Logistic Regression, AUC = 0.93


Precision Recall Curve
Precision
Recall
Logistic Regression, AUPR = 0.94


ROC Curve
True Positive Rate
False Positive Rate
Naive Bayes, AUC = 0.86


Precision Recall Curve
Precision
Recall
Naive Bayes, AUPR = 0.97


ROC Curve
True Positive Rate
False Positive Rate
SVM, AUC = 0.93


Precision Recall Curve
Precision
Recall
SVM, AUPR = 0.94

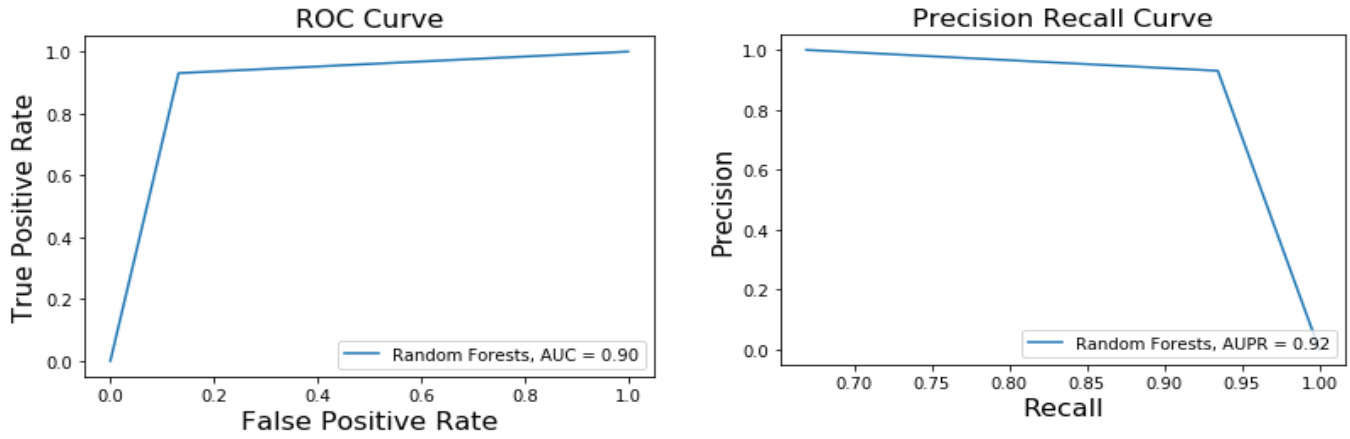
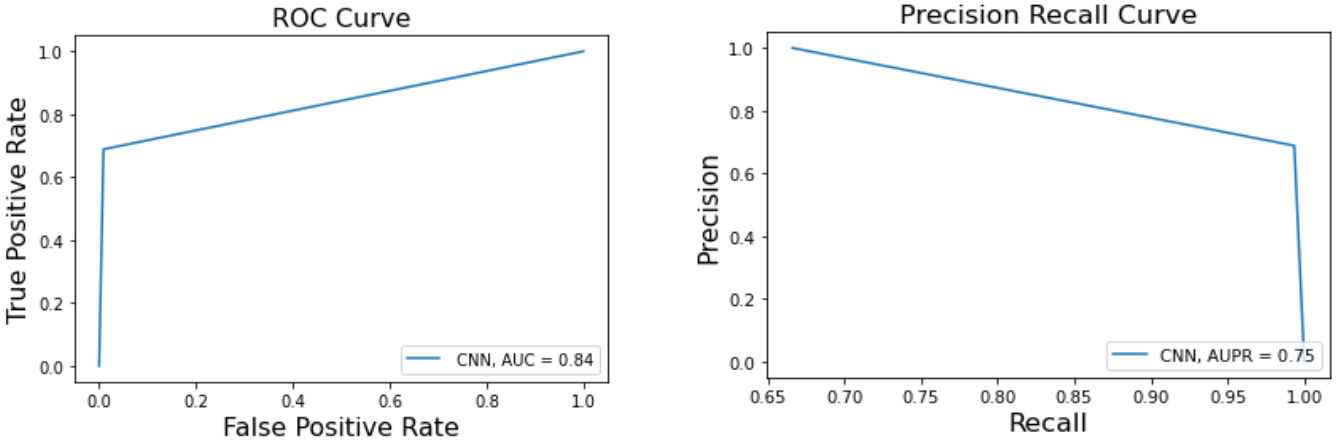



**Figure 4.8:** Courbes ROC et PR des modèles de classification de base

## 4.2. La classification par modèles profonds

| | Precision | Recall | F1_Score | Accuracy | AUC | AUPR |
|---|---|---|---|---|---|---|
| **CNN** | 0.80 | 0.84 | 0.79 | 0,78 | 0.84 | 0.75 |
| **RNN** | 0.71 | 0.72 | 0.71 | 0,73 | 0.72 | 0.64 |
| **LSTM** | 0.89 | 0.88 | 0.89 | 0,90 | 0.88 | 0.84 |
| **BLSTM** | 0.90 | 0.92 | 0.91 | 0,91 | 0.92 | 0.86 |

**Tableau 4.4** : Résultats de classification par modèles profonds



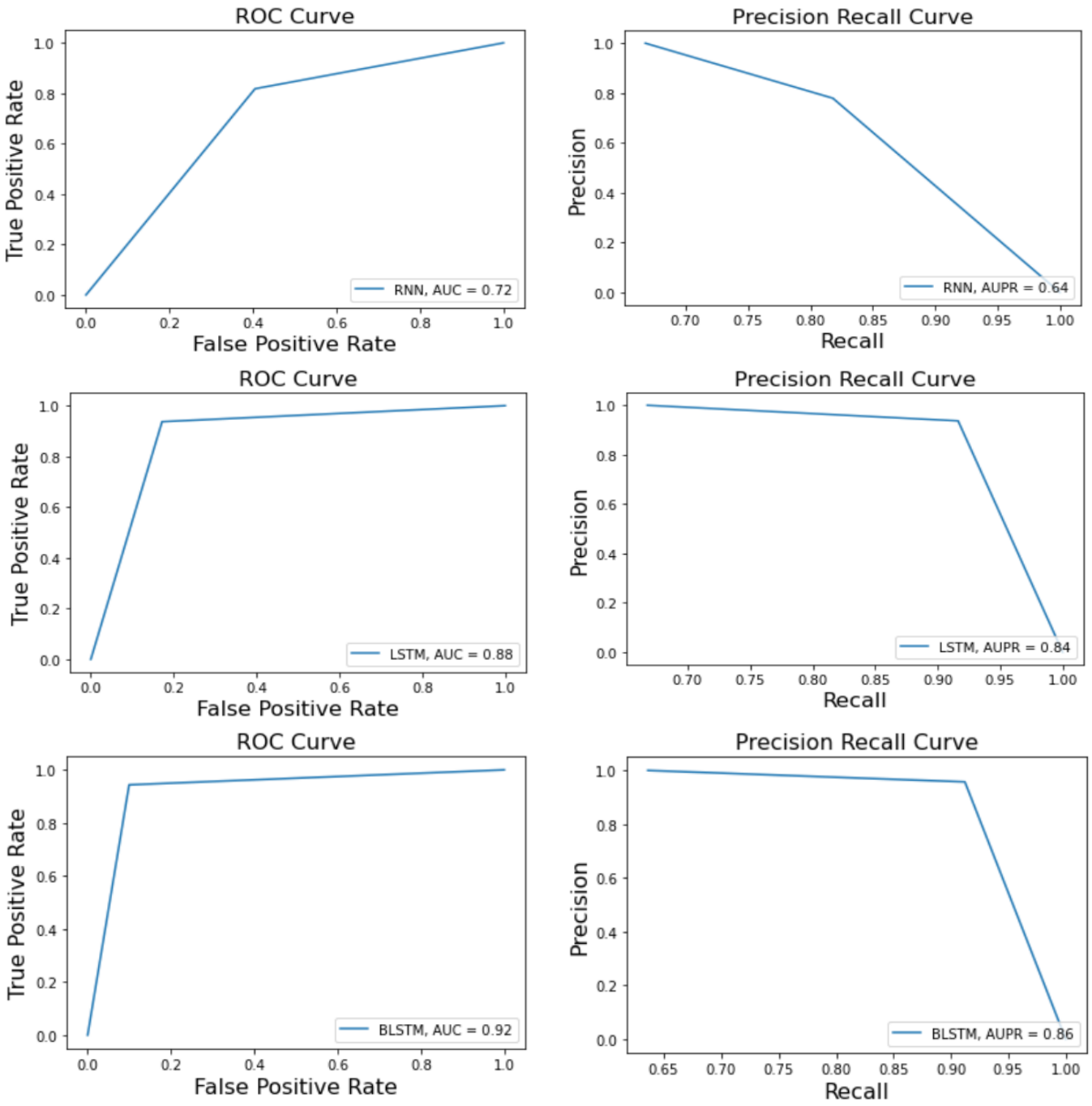



**Figure 4.9:** Courbes ROC et PR des modèles de classification profonds

## 5. La discussion des résultats de classification

A partir de ce que nous avons comme résultats, nous remarquons que les modèles de classification de base Naïve de Bayes, Régression logistique, Forêts aléatoires, et Machines à vecteurs de supports SVM, donnent de bonnes performances en comparaison avec les modèles d'apprentissage profond. D'autre part, nous constatons que le modèle SVM et le modèle de la régression logistique surpassent les autres classificateurs de base et fonctionnent mieux en

termes de toutes les mesures. Ceci n'est pas surprenant parce que SVM est connu de sa capacité optimale à supporter le type catégorique de données (comme les données textuelles) par rapport aux autres classificateurs classiques. En plus, sa performance s'augmente avec les ensembles de données de taille petite et moyenne. De même pour la régression logistique qui donne des bons résultats dans le cas où les classes des données peuvent être linéairement séparables. La performance d'un SVM avec un noyau linéaire est presque semblable à la performance de la régression logistique dont sa linéarité est sa caractéristique spécifique.

Les résultats des modèles de DL le LSTM, le CNN, le RNN et le BLSTM ont été affectés par la quantité de données dans l'ensemble « New_Covid_Data ». En effet, la taille de l'ensemble de données qui était bénéfique pour les classificateurs de base, a été de l'autre côté négative pour les classificateurs profonds. Les modèles d'apprentissage profond exigent des très grandes quantités de données pour aboutir à des meilleurs scores de classification.

## 6. La prédiction de fausses informations

La dernière étape dans notre projet est la prédiction de fausses informations. Cette étape consiste à prédire la véracité d'un texte en entrée en se basant sur le modèle de classification le plus performant parmi les modèles présentés ci-dessus (dans notre cas, les SVM ou la régression logistique). Le résultat est la classe « Fake » ou « True » avec sa probabilité d'authenticité. Comme exemple, nous avons testé le modèle de détection final avec un texte fabriqué aléatoirement. Le résultat est illustré dans la figure 4.10.

```
In [4]: var = input("Veuillez entrer le texte à vérifier ----------: ")

Veuillez entrer le texte à vérifier ----------: China is to blame because the
culture where people eat bats and snakes and dogs and things like that  these
viruses are transmitted from the animal to the people and that s why China
has been the source of a lot of these viruses like SARS  like MERS  the Swine
Flu and now the coronavirus

Le résultat :
Le texte est: Fake
La probabilité de véracité est= 0.32
```

**Figure 4.10 :** Exemple de détection d'une fausse information

## Conclusion

Nous avons conclu notre travail dans ce mémoire par un rapport décrivant la réalisation de la tâche de détection de fausses informations appliquée sur des données liées à l'infodémie de Coronavirus. Nous avons présenté les différentes étapes de la mise en œuvre ainsi que les résultats obtenus par les différents modèles présentés préalablement dans le chapitre de conception précédent. Une interprétation des résultats sur la base des scores des métriques adéquates est offerte à la fin du chapitre afin de bien comprendre les limitations et les performances des modèles de classification appliqués.

# Conclusion générale

Le concept de détection de fausses informations sur le web est particulièrement nouveau et important. La vitesse de diffusion de ce type de contenu a récemment suscité une véritable inquiétude autour du monde notamment avec la crise sanitaire de Covid-19 prévue d'avoir d'énormes impacts politiques et sociaux, aussi que des graves conséquences sur le long terme.

A l'heure actuelle, après presque neuf mois de la crise sanitaire, il y a encore de désinformation et de mésinformation au sujet de Coronavirus sur le web et les réseaux sociaux malgré les efforts sérieux des géants du web Google, Facebook, Twitter, et WhatsApp qui ont essayé depuis le déclenchement de cette situation à limiter la propagation de ce genre d'informations. Les compagnes de désinformation actuellement par exemple mettent en doute l'utilité de la politique de confinement, du port de masque et la crédibilité d'un vaccin prochain possible.

Dans ce contexte, notre travail a visé plus d'un objectif. En premier lieu, il nous a permis d'explorer le domaine d'échange informationnel humain sur le web et les plateformes sociaux et comprendre toutes ses complexités en termes de traitement, de contrôle et d'orientation vers la bonne voie. En outre, nous avons eu l'opportunité de mettre en pratique toutes nos connaissances dans le domaine d'apprentissage automatique et d'apprentissage approfondi sur un thème sanitaire aussi important. Nous avons eu également la chance de travailler sur le domaine de traitement automatique de langage naturel, qui est un domaine très actif et très prometteur et dont la nécessité pour la société moderne s'accroit jour après jour.

Comme extension de notre travail, nous proposons d'appliquer les idées de ce mémoire sur des données textuelles en langue arabe. En ce qui nous concerne, vu l'universalité de la crise sanitaire, nous avons choisi de travailler sur un corpus en anglais parce qu'il y a de la ressemblance des fausses idées diffusées dans la région MENA (Middle East and North Africa) et les autres fausses informations diffusées autour du monde. En plus, autre que l'approche de détection basée sur le style d'information que nous avons suivi, d'autres approches de détection peuvent être combinées pour plus de performance et d'efficacité.

On espère aujourd'hui que nos objectifs ont été atteint à un niveau raisonnable et que notre application contribue à aider le lecteur de notre mémoire à comprendre le phénomène des fausses informations, à sensibiliser aux effets négatifs de la désinformation, et à se familiariser avec les solutions techniques de détection de la désinformation diffusée sur le web.

# Bibliographie

## Articles & livres

[1] ZANNETTOU, Savvas, SIRIVIANOS, Michael, BLACKBURN, Jeremy, *et al.* The web of false information: Rumors, fake news, hoaxes, clickbait, and various other shenanigans. *Journal of Data and Information Quality (JDIQ)*, 2019, vol. 11, no 3, p. 1-37.

[2] SHU, Kai, SLIVA, Amy, WANG, Suhang, *et al.* Fake news detection on social media: A data mining perspective. *ACM SIGKDD explorations newsletter*, 2017, vol. 19, no 1, p. 22-36.

[3] ZHOU, Xinyi et ZAFARANI, Reza. Fake news: A survey of research, detection methods, and opportunities. *arXiv preprint arXiv:1812.00315*, 2018.p 4-30.

[4] PRADHAN, Vidisha M., VALA, Jay, et BALANI, Prem. A survey on sentiment analysis algorithms for opinion mining. *International Journal of Computer Applications*, 2016, vol. 133, no 9, p. 7-11.

[5] BOUARARA H, D. HAMOU, M. Apprentissages et métaheuristique basée sur les algorithmes génétiques pour la détection de plagiat. Mémoire de Master. *Université de Saida*, 2013.

[6] SARKAR, Dipanjan. Text analytics with Python: a practical real-world approach to gaining actionable insights from your data. New York: *Apress*; 2016.

[10] GOYAL, Palash, PANDEY, SUMIT, et JAIN, Karan. Deep learning for natural language processing. *Apress*, 2018, p. 138-143.

[9] TOUZET, Claude. Les réseaux de neurones artificiels, introduction au connexionnisme : cours, exercices et travaux pratiques. *EC2, Collection de l'EERIE, N. Giambiasi*, 1992.

[11] GRANET, A., MORIN, E., MOUCHERE, H., QUINIOU, S., & VIARD-GAUDIN, C. Étude préliminaire de reconnaissance d'écriture sur des documents historiques. *CORIA*, 2017.

[32] CUI, Limeng, and DONGWON Lee. "CoAID: COVID-19 Healthcare Misinformation Dataset." *arXiv preprint arXiv:2006.00885*, 2020.

## Sites web

[7] Geeksforgeeks: Understanding Logistic Regression. https://www.geeksforgeeks.org/understanding-logistic-regression/ (Site consulté en juillet 2020).

[8] OpenClassrooms. Initiez-vous au deep learning / découvrez le neurone formel. https://openclassrooms.com/fr/courses/5801891-initiez-vous-au-deep-learning/5801898-decouvrez-le-neurone-formel. (Site consulté en juin 2020).

[12] WHO. Qu'est-ce que la COVID-19 ? https://www.who.int/fr/emergencies/diseases/novel-coronavirus-2019/advice-for-public/q-a-coronaviruses. (Site consulté en juillet 2020).

[13] FREDIRIQUE ScEneider. « Coronavirus : les grandes dates de la pandémie ». La croix, le 25 avril 2020. https://www.la-croix.com/Monde/Coronavirus-grandes-dates-pandemie-2020-04-25-1201091132. (Site consulté en juillet 2020).

[14] « Les étapes de la propagation du coronavirus dans le monde », LE TEMPS, publié le 21 avril 2020. https://www.letemps.ch/monde/etapes-propagation-coronavirus-monde. (Site consulté en juillet 2020).

[15] BBC Monitoring and UGC Newsgathering. «China coronavirus: Misinformation spreads online about origin and scale», BBC Nwes, publié en 30 Janvier 2020. https://www.bbc.com/news/blogs-trending-51271037. (Site consulté en juillet 2020).

[16] Rory Cellan-Jones. « Coronavirus: Fake news is spreading fast». BBC News, publié le 26 février 2020. https://www.bbc.com/news/technology-51646309. (Site consulté en juillet 2020).

[17] « Arab Writers: The Coronavirus Is Part Of Biological Warfare Waged By The U.S. Against China». MEMRI, publié le 6 février 2020. https://www.memri.org/reports/arab-writers-coronavirus-part-biological-warfare-waged-us-against-china. (Site consulté en juillet 2020).

[18] Agence France-Presse. «Coronavirus: Russia pushing fake news about US using outbreak to 'wage economic war' on China, officials say », South China Morning Post, publié le 23 février 2020. https://www.scmp.com/news/world/russia-central-asia/article/3051939/coronavirus-russia-pushing-fake-news-about-us-using. (Site consulté en juillet 2020).

[19] KATE Ng. «US accuses Russia of huge coronavirus disinformation campaign». Yahoo News publié le 22 février 2020. https://news.yahoo.com/us-accuses-russia-huge-coronavirus-171730161.html. (Site consulté en juillet 2020).

[20] JESSICA Glenza. «Coronavirus: US says Russia behind disinformation campaign». THE GUARDIAN, publié le 22 février 2020.

https://www.theguardian.com/world/2020/feb/22/coronavirus-russia-disinformation-campaign-us-officials. (Site consulté en juillet 2020).

[21] ZACHARY Halaschak. « 'Biologic war': Former Iranian president says coronavirus was 'produced in laboratories' ». Washington Examiner, publié le 9 mars 2020. https://www.washingtonexaminer.com/news/biologic-war-former-iranian-president-says-coronavirus-was-produced-in-laboratories. (Site consulté en juillet 2020).

[22] WILLIAM Audureau. « Non, le Covid-19 n'est pas une combinaison du SRAS et du sida ». LE MONDE, Publié le 09 mars 2020. https://www.lemonde.fr/les-decodeurs/article/2020/03/09/non-le-covid-19-n-est-pas-une-combinaison-du-sras-et-du-sida_6032401_4355770.html. (Site consulté en juillet 2020).

[23] CHARLIE Campbell. « 'They Are Overwhelmed.' China's Animal Shelters Can't Cope with the Number of Pets Abandoned Due to COVID-19 ». TIME publié le 2 mars 2020. https://time.com/5793363/china-coronavirus-covid19-abandoned-pets-wuhan/. (Site consulté en juillet 2020).

[24] Allen Kim. «Cats and dogs abandoned at the start of the coronavirus outbreak are now starving or being killed». CNN, publié le 15 mars 2020. https://edition.cnn.com/2020/03/15/asia/coronavirus-animals-pets-trnd/index.html. (Site consulté en juillet 2020).

[25] Adam Tylor. « Experts debunk fringe theory linking China's coronavirus to weapons research». The Washington Post, publié le 29 janvier 2020. (Site consulté en juillet 2020).

[26] WHO. « Nouveau coronavirus (2019-nCoV) : conseils au grand public - En finir avec les idées reçues ». https://www.who.int/fr/emergencies/diseases/novel-coronavirus-2019/advice-for-public/myth-busters. (Site consulté en juillet 2020).

[27] Tom Kertscher. «Says a coronavirus patent expired just as there is a "sudden outbreak" and, despite "media fear-mongering," there is already a vaccine available. ». PLITIFACT, publié le 27 juin 2020. https://www.politifact.com/factchecks/2020/jan/23/facebook-posts/there-outbreak-china-wuhan-coronavirus-there-not-v/. (Site consulté en juillet 2020).

[28] https://fr.wikipedia.org/wiki/Python_(langage). (Site consulté en aout 2020).

[29] https://jupyter.org/. (Site consulté en aout 2020).

[30] https://fr.wikipedia.org/wiki/Scikit-learn. (Site consulté en aout 2020).

[31] https://keras.io/. (Site consulté en aout 2020).