الجمهورية الجزائرية الديمقراطية الشعبية
وزارة التعليم العالي والبحث العلمي
جامعة سعيدة د. مولاي الطاهر
كلية التكنولوجيا
قسم: الإعلام الآلي

Mémoire de Master

Spécialité : Réseaux Informatique Systèmes Repartis

Thème

Système de détection de Fake news utilisant machine learning

Présenté par :

MEBARKI Ahmed Abdelmoumne

AGGOUN brahim

Dirigé par :

DR.Siham KOUIDRI

Année universitaire 2022-2023

DÉDICACE

Au nom de Dieu Clément et Miséricordieux,
Je dédie ce modeste travail de fin d'études à mes très chers parents, qui ont été un grand soutien moral dans les moments difficiles tout au long de ce parcours.
À tous mes enseignants, pour leur contribution à ma formation intellectuelle.
Ainsi qu'à mes chères sœurs, mon frère et à toute ma famille, pour leur soutien constant et leur amour inconditionnel.
Sans oublier mon binôme et mes amies proches, qui ont été présentes à mes côtés et qui ont apporté leur soutien précieux.
À tous ceux qui ont contribué de près ou de loin à la réalisation de ce travail, qu'ils trouvent ici l'expression de ma profonde gratitude et reconnaissance.
Que Dieu vous récompense pour votre soutien et votre présence tout au long de ce cheminement.

AGGOUN BRAHIM

DÉDICACE

Au nom de Dieu Clément et Miséricordieux,
Je souhaite dédier ce modeste travail de fin d'études à mes chers parents, qui ont été un soutien inestimable durant les moments difficiles tout au long de ce parcours. Je tiens également à exprimer ma reconnaissance à tous mes enseignants pour leur contribution à ma formation intellectuelle. Je suis reconnaissant envers mes chères sœurs, mon frère et toute ma famille pour leur soutien constant et leur amour inconditionnel. Je n'oublie pas de mentionner mon binôme et mes amies proches qui ont été présents à mes côtés et qui m'ont apporté un soutien précieux. Je suis profondément reconnaissant envers tous ceux qui ont contribué, de près ou de loin, à la réalisation de ce travail. Que vous trouviez ici l'expression de ma gratitude sincère et de ma reconnaissance. Que Dieu vous récompense pour votre soutien et votre présence tout au long de ce parcours.

MEBARKI Ahmed Abdelmoumne

Remerciement

Nous tenons à exprimer notre profonde gratitude à toutes les personnes qui ont apporté leur aide et contribué à la réalisation de ce mémoire. Nos remerciements les plus sincères vont à notre encadrante, **Mme KOUIDRI Siham**, pour sa direction et ses précieux conseils tout au long de ce travail.

Nous souhaitons également remercier chaleureusement les membres du jury pour l'intérêt qu'ils ont porté à notre projet de fin d'études, en acceptant de l'examiner et en nous faisant part de leurs propositions constructives pour son amélioration.

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

Enfin, nous tenons à remercier toutes les personnes qui ont contribué de près ou de loin à la réalisation de ce mémoire. Leur implication et leur collaboration ont été essentielles, et nous leur sommes très reconnaissants.

الملخص

تواجه منصات التواصل الاجتماعي تحولًا جذريًا في طريقة نشر المعلومات، مما يعزز بشكل كبير سرعة وحجم وتنوع نقل المعلومات. ومع ذلك، تُيسّر شبكات التواصل الاجتماعي أيضًا انتشار المعلومات الصحيحة والمعلومات الزائفة بسرعة.

تعتبر جائحة مرض فيروس كورونا (كوفيد 19) ربما أكبر تحد صحي عالمي في القرن الماضي. وترافق هذه الجائحة ءنفوديمياْ موازية، تشمل التسويق والبيع عبر الإنترنت لمنتجات غير معتمدة وغير قانونية ومقلدة. ويشمل ذلك المنتجات الأساسية والمنتجات الحساسة مثل الغذاء والمنتجات الصيدلانية والمستلزمات الطبية، والتي يمكن أن تسبب آثارًا خطيرة على الصحة الفردية والصحة العامة.

الهدف هو تحليل مقارن بين نماذج التصنيف لتحقيق مهمة اكتشاف الأخبار الزائفة المنشورة على الويب ووسائل التواصل الاجتماعي خلال أزمة جائحة فيروس كورونا. يتم تحقيق ذلك من خلال استخدام تقنيات التعلم الآلي.

الكلمات المفتاحية : التعلم الآلي، المعلومات الزائفة، الأخبار الكاذبة، فيروس كورونا، كوفيد 19 ، تصنيف النصوص.

Abstract

Social media platforms have revolutionized the way information is disseminated, greatly improving the speed, volume, and variety of information transmission. However, social networks also facilitate the rapid spread of both factual and false information.

The COVID-19 pandemic is perhaps the greatest global health challenge of the last century. This pandemic has been accompanied by a parallel "infodemic," which includes the online marketing and sale of unapproved, illegal, and counterfeit products. This is particularly concerning for essential goods and sensitive products such as food, pharmaceuticals, and related items, as they can have serious consequences on individual and public health.

The objective is comparative analysis between models of classification to accomplish the task of detecting fake news published on the web and social media during the Coronavirus pandemic crisis. This is achieved through the use of machine learning techniques.

Keywords : machine learning, false information, fake news, coronavirus, COVID-19, text classification.

Résumé

Les plateformes de réseaux sociaux ont révolutionné le mode de diffusion de l'information, ce qui améliore considérablement la vitesse, le volume et la variété de la transmission de l'information. Cependant, les réseaux sociaux facilitent la diffusion rapide d'informations factuelles et fausses.
La pandémie de maladie à coronavirus (COVID-19) est peut-être le plus grand défi sanitaire mondial du siècle dernier. Cette pandémie s'accompagne d'une "infodémie" parallèle, comprenant le marketing et la vente en ligne des produits non approuvés, illégaux et contrefaits, Notamment les produits de première nécessité et les produits à caractères sensibles comme la nourriture, les produits pharmaceutiques et parapharmaceutiques, qui peuvent causer des conséquences très grave sur la santé individuelle et publique.
L'objectif est de étude comparative entre les modèles des classification 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

Mots clés :machine learning, fausse information, fausse nouvelle, coronas virus, covid'19, classification du texte.

Table des matières

Figure List

Table List

Introduction Générale

Contexte

Ces dernières années, les Fake News ont connu une propagation massive, largement favorisée par l'essor des réseaux sociaux. Ces fausses informations sont diffusées dans le but d'atteindre différents objectifs. Certaines sont créées simplement pour augmenter le nombre de clics et de visiteurs sur un site web, tandis que d'autres visent à influencer l'opinion publique sur des décisions politiques ou les marchés financiers, impactant ainsi la réputation des entreprises et des institutions en ligne. Une étude menée par l'économiste Roberto Cavazos de l'université de Baltimore met en évidence l'impact direct de ces Fake News sur l'économie. Selon cette étude, au moins 400 millions de dollars ont été dépensés pour diffuser des Fake News lors des récentes élections en Inde (140 millions de dollars en 2019), au Brésil (34 millions de dollars en 2018), au Royaume-Uni (1 million de dollars), en France (586 000 euros), en Australie (828 000 dollars), au Kenya (20 millions de dollars), en Afrique du Sud (2,7 millions de dollars) et au Mexique (642 000 dollars), par exemple. Une étude menée par l'université de Princeton sur la consommation de Fake News pendant la campagne présidentielle américaine de 2016 a révélé que les fausses informations représentaient 2,6 Le contexte de la détection des fausses informations basée sur l'apprentissage automatique sur Twitter est marqué par la propagation rapide et généralisée des informations sur cette plateforme de médias sociaux. Twitter est un réseau social très populaire où les utilisateurs peuvent partager des messages courts appelés "tweets". Cependant, cette facilité de partage peut également permettre la diffusion de fausses informations, ce qui peut avoir un impact significatif sur les opinions publiques, les débats en ligne et même les événements du monde réel. Les fausses informations sur Twitter peuvent prendre différentes formes, allant des informations trompeuses aux rumeurs infondées et aux manipulations intentionnelles. En raison de la nature en temps réel de Twitter, les fausses informations peuvent se propager rapidement et atteindre un large public en un laps de temps très court. Par conséquent, la détection rapide et précise des fausses informations sur cette plateforme est d'une importance cruciale. Les algorithmes d'apprentissage automatique offrent une approche prometteuse pour la détection des fausses informations sur Twitter. Ces algorithmes peuvent être entraînés à reconnaître des modèles et des caractéristiques spécifiques dans les tweets, tels que le choix des mots, les relations entre les utilisateurs, les comportements de partage, les informations de profil et les métadonnées associées aux tweets. En analysant ces caractéristiques, les modèles d'apprentissage automatique peuvent tenter de distinguer les tweets véridiques des tweets potentiellement faux. Cependant, la détection des fausses informations sur Twitter présente des défis uniques. Le caractère rapide et évolutif des fausses informations, le volume élevé de données, la présence de sarcasme et d'ironie, la crédibilité variable des sources et les déséquilibres de données sont quelques-uns des défis auxquels les chercheurs et les praticiens sont confrontés dans ce domaine. Dans ce contexte, l'utilisation de l'apprentissage automatique pour la détection des fausses in-

formations sur Twitter vise à développer des modèles et des techniques avancés pour identifier et contrer efficacement les informations trompeuses.

Problématique

La propagation des fausses informations constitue un défi majeur dans notre ère numérique, avec des conséquences significatives sur la société, la politique, et la confiance du public dans les médias. Les algorithmes d'apprentissage automatique offrent des solutions prometteuses pour détecter et contrer ces fausses informations. Comment améliorer la détection des fausses informations sur Twitter en utilisant des algorithmes d'apprentissage automatique ? La résolution de ce probléme est cruciale pour développer des systèmes de détection des fausses informations robustes, précis et adaptatifs, contribuant ainsi à promouvoir une information de qualité et à lutter contre la désinformation dans notre société.

Objectif

La détection des fausses informations basée sur des algorithmes d'apprentissage automatique est devenue une méthode prometteuse pour lutter contre la propagation de la désinformation dans notre société numérique. Avec la prolifération rapide des médias sociaux et la facilité de partage d'informations en ligne, il est devenu essentiel de pouvoir distinguer les informations réelles des fausses. L'approche de détection des fausses informations basée sur des algorithmes d'apprentissage automatique repose sur la capacité des modèles d'apprentissage à apprendre des modèles à partir de grandes quantités de données et à généraliser ces modèles pour de nouvelles informations. Le processus de détection des fausses informations commence par la collecte d'un ensemble de données étiquetées, comprenant à la fois des exemples tweets réels et des exemples de fausses informations. Cette étape est cruciale car elle fournit aux algorithmes d'apprentissage automatique les informations nécessaires pour apprendre à discriminer entre les deux types de tweet. Ensuite, des caractéristiques pertinentes sont extraites des tweets, telles que la fréquence des mots, les n-grammes, les caractéristiques linguistiques, les métadonnées (telles que la source et la date de publication), et les caractéristiques de twitter (telles que le nombre de partages, de likes et de commentaires). Ces caractéristiques servent de base à l'apprentissage des modèles. Avant d'entraîner les modèles, les données textuelles sont prétraitées en les nettoyant, en les normalisant et en les transformant en une représentation appropriée pour l'apprentissage automatique, par exemple en les tokenisant, en supprimant les mots vides, en effectuant une lemmatisation ou une racinisation, et en gérant les caractères spéciaux. En utilisant des algorithmes d'apprentissage automatique tels que Naive Bayes, SVM, Random Forest, Gradient Boosting ou Régression Logistique, les modèles sont entraînés sur les données d'apprentissage et évalués sur les données de test. Différentes métriques d'évaluation, telles que l'entropie, la précision, le rappel et le F1-score, sont utilisées pour mesurer les performances du modèle. Une fois que le modèle atteint des performances satisfaisantes, il peut être déployé en tant qu'application autonome ou intégré à des systèmes existants pour aider à la classification des tweets en tant que réels ou faux. Il convient de noter que l'efficacité de cette approche repose sur la qualité et la diversité de l'ensemble de données d'apprentissage, la sélection appropriée des caractéristiques et l'algorithme choisi.

L'objectif final est de promouvoir un environnement d'information plus fiable et de

garantir que les utilisateurs de Twitter sont mieux informés et protégés contre les effets néfastes des fausses informations.

Organisation du manuscript

1. **Les Fausses Information :** Le premier chapitre présente les différentes définitions concernant les fausse nouvelles et le domaine du Web Mining.
2. **Machine Learning :**Le deuxième chapitre présente un état de l'art sur les méthodes et les principe utilisées dans la détection des fausses nouvelles
3. **Conception du système,Implémentation et résultats**Le troisième chapitre sera réservé à la description détaillée de l'architecture générale et l'architecture détaillée de notre système et présentant la mise en oeuvre de notre travail et les résultats obtenus.
4. **Conclusion generale et perspectives**Enfin, Une synthèse et un ensemble de perspectives feront l'objet de notre conclusion.

CHAPITRE 1. FAUSSES INFORMATIONS

1

1.1 Introduction

L'objet de ce chapitre est d'apporter quelques rappels élémentaires et nécessaires pour clarifier la notion de fake news, ses types et son concept, d'une part. D'autre part, nous définissons les médias sociaux et la relation entre les deux. Nous présenterons d'abord, les fausses nouvelles puis leur relation avec les médias sociaux.

1.2 Définition :

De fausses nouvelles ou de fausses informations peuvent être diffusées à diverses fins. Certains sont destinés à distraire le lecteur ou à influencer ses opinions sur un sujet particulier. L'autre a été fabriqué de toutes pièces avec un titre attractif pour attirer du trafic et augmenter le nombre de visiteurs sur le site. Ces dernières années, le phénomène des fake news s'est répondu sur Internet au détriment des internautes. Le phénomène des fake news a pris une ampleur médiatique au cours de l'élection présidentielle de 2016 aux États-Unis. Cette tendance a fait son apparition en France quelques mois plus tard lorsque le peuple a également été appelé aux urnes pour élire son nouveau président. [20]. Ainsi au cour de la pandemie Coronas Virus voici quelques exemples des fausses informations :

- Théories du complot : Certaines fausses informations affirment que le virus a été créé intentionnellement dans un laboratoire ou qu'il s'agit d'une arme biologique. Ces théories n'ont aucune base scientifique solide et ont été réfutées par les experts de la santé.
- Remèdes miracles : Plusieurs prétendus remèdes miracles ont été promus, affirmant guérir ou prévenir le coronavirus. Certains exemples incluent la consommation d'eau de javel, l'injection de désinfectant ou la prise de certains médicaments sans preuve scientifique valable. Il est important de faire confiance aux recommandations des autorités sanitaires et des professionnels de la santé.
- Désinformation sur les symptômes et la prévention : Des informations erronées ont circulé concernant les symptômes du COVID-19 et les mesures préventives. Par exemple, certaines fausses informations affirmaient que porter un masque était inutile ou dangereux, alors que les experts recommandent maintenant largement le port du masque pour réduire la propagation du virus.
- Statistiques trompeuses : Des données et des statistiques ont été manipulées ou déformées pour minimiser ou exagérer l'ampleur de la pandémie. Il est important de consulter des sources fiables et d'obtenir des informations auprès d'organismes de santé officiels.

1.3 Concepts sur Fake News

Fausses information diffusées sur Internet ou d'autres médias. Conçues pour influencer les opinions politiques, social, santé .....etc. Elle sont caractérisées par ce qui suit [27] :

- Volume : le volume des fausses nouvelles est très grand tel que tout le monde peut facilement écrire de fausses nouvelles sur Internet et sans aucune procédure de vérification.
- Variété : il existe plusieurs sources des fausses nouvelles : les rumeurs, les nouvelles satiriques, les fausses critiques, les fausses informations, les fausses publicités, les théories du complot, les fausses déclarations des politiciens,...etc.

— Rapidité : l'apparition des fausses nouvelles est très rapide, Ceci complexifie leur détection par les systèmes de contre-mesures.

1.4 Acteurs de fausses informations

Divers acteurs peuvent rechercher un gain personnel en diffusant de la désinformation. Il s'agit d'une catégorie très large, des simples utilisateurs aux entités plus organisées et plus fort. On peut citer [26] :

— 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.
— Organisations criminelles et terroristes : Les gangs criminels et les organisations terroristes profitent des réseaux sociaux en ligne pour diffuser de fausses informations. Afin d'atteindre leurs objectifs logistiques ou politiques. Un exemple récent est l'organisation terroriste (Daech) qui diffuse de fausses informations sur les réseaux sociaux à des fins de propagande et le recrutement de nouveaux membres terroristes.
— Activistes et organisations politiques : Diverses organisations politiques et militants partagent de fausses informations, soit pour promouvoir leurs organisations, discréditer leurs opposants, soit imposer certains récits, notamment avant les grands événements politiques.
— Trolls : Troll fait référence aux utilisateurs qui publient un contenu controversé ou hors sujet afin de perturber le flux normal de discussion sur un site Web, ou de provoquer ou d'exercer une pression émotionnelle sur d'autres utilisateurs pour rien de plus que leur propre amusement personnel.
— Bots : Un programme qui appartient au Bot Network (Botnet). Ils sont souvent associés à un grand nombre de faux comptes qui surveillent collectivement l'activité en ligne et propagent la désinformation sur le Web. Il existe plusieurs types de robots informatiques avec des fonctions différentes. Certains bots republient ou promeuvent du contenu ; d'autres affichent du contenu original.
— 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.
— Trolls finances : Il s'agit d'un groupe privé d'internautes qui sont payés pour publier du contenu destiné à un groupe démographique spécifique. Ils sont employés pour faire avancer un programme politique et influencer les gens à prendre certaines orientations sociales ou économiques. Ces types de représentants travaillent pour leur propre bien. Il est beaucoup plus difficile de les distinguer des bots car ils ont des caractéristiques similaires aux utilisateurs réguliers.
— Gouvernements : Historiquement, les gouvernements se livrent régulièrement à la diffusion de fausses informations pour diverses raisons. Récemment, et surtout avec la prolifération du Web, les gouvernements s'appuient sur les médias sociaux pour manipuler l'opinion publique nationale et étrangère sur des sujets spécifiques. Par exemple, l'ingérence du gouvernement russe dans les élections américaines de 2016.

— 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.

1.5 Aspects liés aux fausses informations

Tout d'abord, nous pouvons décrire de fausses informations en fonction de trois catégories. Ce classement est officiellement déterminé par M.L. Rouquette1 [17] :

1-processus d'envoi : Les fausses informations sont transmises verbalement de personne à personne par le bouche à oreille ou par les médias. Par conséquent, les canaux sont des canaux médiatiques et des canaux oraux. Cet échange se produit entre des individus qui sont également impliqués dans la situation.

2-Situation : La désinformation apparaît dans les situations de crise, mais elle n'est pas toujours le signe d'un dysfonctionnement social. Les canaux de communication formels transmettent des informations qui se limitent uniquement à des événements ou des situation. En d'autres termes, les individus créent de fausses informations face à un manque d'informations.

3-Contenu : Le contenu des fausses nouvelles subit diverses perturbations au cours du processus de transmission. Ce contenu traduit l'idée de désir de la population, témoigne de la pratique de la pensée sociale, et la désinformation devient une sorte d'écran projectif où le dynamisme est décodé.

1.6 Types de fausses nouvelles

La désinformation apparaît sous différentes formes. Il est important de noter que la désinformation peut être le résultat du chevauchement de plusieurs espèces en même temps, Parmi les huit types suivants [26] :

1-Information fabriquée : information fictive qui n'a rien à voir avec les faits réels. Ce genre de désinformation n'est pas nouveau. Il existe depuis les débuts du journalisme. Des exemples très courants incluent des histoires apocryphes sur les extraterrestres et les objets volants non identifiés (OVNI).

2-Canulars : nouvelles qui contiennent de faux faits avec l'intention d'offenser la crédulité de quelqu'un. On les appelle aussi des demi-vérités. Par exemple, les nouvelles sur la mort de personnes célèbres.

3-Information biaisée ou unilatérale : indique un récit partiel ou très personnel. Dans le contexte politique, ce type est appelé nouvelles hyper-partisanes. Faits biaisés en faveur d'une personne, d'un parti, d'une situation ou d'un événement.

4-Propagande : Il s'agit d'un type d'informations fabriquées. Ce type de fausses nouvelles a été largement utilisé pendant la Seconde Guerre mondiale et la guerre froide. La propagande est couramment utilisée dans des contextes politiques pour influencer un public cible, dans le but de diffuser des idées, des plans ou des idéologies politiques, ou de discréditer un parti politique ou un État-nation particulier. La propagande peut changer le cours de l'histoire, par exemple l'invasion de l'Irak en 2003.

5-Rumeur : Une histoire anonyme, la vérité est obscure.

6-Théories du complot : faites référence à des récits théoriques qui expliquent des situations ou des événements en invoquant des complots sans aucune preuve. Ces histoires parlent des actions illégales des gouvernements ou de personnes puissantes agissant dans

les coulisses. Les informations complotistes expliquent les événements en recourant au complot politique en réfutant toutes les autres explications officielles ou scientifiques proposées par les régimes et les médias grand public.

7-Titres attractifs : fait référence à l'utilisation délibérée de titres trompeurs sur le Web dans le but d'attirer plus de lecteurs. Il s'agit d'un ancien type de fausses nouvelles connu sous le nom de journalisme jaune. Il s'agit de l'une des catégories de fake news les moins dangereuses car il est facile de vérifier son authenticité en visualisant le contenu associé au titre.

8-Satire : texte satirique ou contenu visuel qui critique des situations, des personnes, des organisations ou des gouvernements. Ce type de nouvelles attire de plus en plus l'attention sur le Web. Cependant, comme les écrits satiriques sont principalement diffusés via les médias sociaux, leur nature satirique est souvent ignorée et ignorée par les lecteurs qui les prennent au sérieux sans vérifier.

1.7 Composants des fausses nouvelles

Le terme Fake News se base sur quatre composants principaux : ,victime ciblées , contenu de l'actualité, créateur/ diuseur, contexte social [20] :

- Victimes ciblées : les victimes des fausses nouvelles peuvent être des utilisateurs de médias sociaux ou d'autres récentes plates-formes. Selon les objectifs de la nouvelle, les cibles peuvent être des étudiants, des électeurs, des parents, des personnes âgées...etc.
- Contenu de l'actualité : le contenu de l'actualité fait référence au corps de la nouvelle. Il contient à la fois un contenu physique (par exemple, titre, corps de texte, multimédia) et contenu non physique (par exemple, but, sentiment, sujets).
- Créateur / Diuseur : les créateurs de fausses nouvelles en ligne peuvent être des humains ou non.
- Contexte social : le contexte social indique comment les nouvelles sont diusées sur Internet. L'analyse du contexte social inclut l'utilisateur, l'analyse du réseau (comment les utilisateurs en ligne sont impliqués dans les actualités) et l'analyse du modèle de diusion.

1.7.1 Contenu d'actualité

Chaque nouvelle est constituée d'un contenu d'actualité physique et d'un contenu d'actualité non-physique.[20]

- Contenu d'actualité physique : le contenu physique des fausses nouvelles contient le titre de la nouvelle, le corps de la nouvelle, des images, des vidéos, des hashtags, des signaux de mention, des emojis,...etc. En raison de certaines significations et fonctionnalités, ces composants sont des fonctionnalités importantes pour la détection de fausses nouvelles.
- Contenu d'actualité non-physique : le contenu non physique concerne les opinions, les émotions, attitudes et sentiments que les créateurs de nouvelles veulent exprimer. Exemple, chaque jour, des millions de commentaires sont publiés sur des plateformes de vente en ligne, ces avis biaisés sont de gros problèmes pour les marques et les clients en ligne, non seulement ils vont aecter la décision du processus de la fabrication, mais aussi ils peuvent facilement détruire la réputation d'une marque.

1.7.2 Créateur/Diuseur de fausses nouvelles

Il est important de montrer qui est derrière la fausse nouvelle et pourquoi elle est écrite et partagée tout au long de la vie sociale. Le créateur/diuseur des fausses nouvelles peut être soit un humain, soit non.[20]

— Humain : les robots sociaux ou les Cyborgs ne sont que les porteurs de fausses nouvelles sur les médias sociaux. Ces comptes automatiques sont programmés pour diuser de faux messages par des humains, qui cherchent à perturber la crédibilité de la communauté sociale en ligne.

— Non-humain : les robots sociaux ou les Cyborgs sont les créateurs non-humains les plus courants, de fausses nouvelles. Ce sont des programmes informatiques conçus pour acher des comportements similaires à ceux des humains, et pour produire automatiquement du contenu et interagir avec les humains via les médias sociaux, diuser des rumeurs, des spams, des logiciels malveillants, des informations erronées.

1.7.3 Contexte social

Le contexte social fait référence à l'ensemble du système d'activité et de l'environnement social dans lequel la diusion de la nouvelle a eu lieu. Aujourd'hui, les modes de partage de la nouvelle sont de plus en plus dominés par les technologies interactives sur les médias sociaux. Les utilisateurs en ligne peuvent non seulement apprendre d'avantage sur les tendances, mais aussi partager leurs histoires et défendre leurs intérêts. S'ils partagent ces expériences et interactions au sein de certains groupes sociaux, et les membres de ces groupes partageant les mêmes idées, l'influence de ces dernières peut être amplifiée. Ceci facilite la diusion de fausses nouvelles.[20]
Voici quelques pages et journaux qui diffusent de fausses nouvelles

FIGURE 1.1 – Les fausses informations partagées sur les réseaux sociaux : différents types de contenu, différentes motivations, différents acteurs

1.8 Médias Sociaux : définitions, spécificités et fonctionnement

Appartenant aux outils d'expression et d'échanges sur Internet, les réseaux sociaux se distinguent cependant de leurs homologues, les blogs, les fils RSS, les forums de discussion et les mondes virtuels. A la fois dispositif informationnel et communicationnel, un réseau social numérique s'affiche alors de manière très singulière par rapport aux autres dispositifs TIC. Et se distingue clairement de la télévision par leurs multiples interfaces, de la radio par sa capacité interactionnelle. Les réseaux sociaux numériques possèdent des points de similitudes avec les blogs par l'initiation des dialogues interpersonnelles, avec les forums avec l'analyse possible de pratiques communicationnelles en ligne et avec les mondes virtuels par la capacité de création et du maintien de relations personnelles en ligne.[25]

1.8.1 Définition de réseau social

Dans le domaine des technologies, un réseau social consiste en un service permettant de regrouper diverses personnes afin de créer un échange sur un sujet particulier ou non. En quelque sorte, le réseau social trouve ses origines dans les forums, groupes de discussion et salons de chat introduits dès les premières heures d'Internet. Depuis le début des années 2000, la présence des réseaux sociaux, également appelés réseaux communautaires, devient de plus en plus importante et tend à se multiplier selon diverses caractéristiques. Les premiers réseaux sociaux de grande envergure (MySpace et Facebook) se sont positionnés en tant que services généralistes sur lesquels chacun peut partager le contenu de son choix, quel qu'en soit le sujet, avec ses contacts.[21]

1.8.2 Les spécificités des réseaux sociaux

Il existe toutefois une multitude de réseaux communautaires plus spécifiques ciblant soit :

- un âge de population, par exemple Snapchat regroupe principalement des adolescents.
- une passion commune,par exemple Bëhance convoite les créateurs graphiques.
- un type de publication, par exemple le réseau social Instagram cible les photographes, YouTube et Vimeo s'adressent aux vidéastes .
- un objectif précis, par exemple LinkedIn et Viadeo ont pour but d'élargir ses opportunités professionnelles.

Notons que l'avènement des smartphones a transformé les réseaux sociaux pour en revoir leurs usages ou les confiner uniquement au téléphone. À l'heure où nous écrivons ces lignes, c'est le cas de WhatsApp, Facebook Messenger ou Instagram. Selon une étude de Statista publiée en janvier 2018, Facebook serait le réseau social le plus important avec 2.167 milliards d'utilisateurs devant YouTube (1,5 milliard), WhatsApp (1,3 milliard) et Facebook Messenger (1,3 milliard). WeChat, QQ, Instagram et Tumblr seraient respectivement en 5e, 6e, 7e et 8e position.[21]

1.8.3 fonctionnement

Les réseaux sociaux s'alimentent des émotions des personnes, de leur besoin de communiquer et d'interagir. Les fonctionnalités existantes comme les célèbres likes de Facebook ou hashtags de Twitter sont destinées à faire réagir les utilisateurs, à les encourager

à partager des informations et leurs avis. La nécessité de s'accomplir et le bénéfice émotionnel, qui sont bel et bien réels, s'obtiennent par un moyen virtuel. Le réseau de chaque utilisateur est souvent comparé à une étoile. Chaque branche correspond à une caractéristique qui le définit : contacts, centres d'intérêt, groupes, données personnelles... Une fois votre toile tissée, la plateforme utilise vos liens forts (nos contacts proches) pour créer des liens faibles (les amis de vos amis). Ainsi, la communauté s'agrandit et le réseau social prospère.[21]

1.9 Fausses informations via les réseaux sociaux

Les fausses informations fournies par les médias sociaux sont des informations diffusées sur les sites suivants : Réseaux. Dans peu de temps, il occupera une grande échelle. méthode virale. Sur les réseaux sociaux, les fonctionnalités sont généralement ajoutées par de fausses informations. La décentralisation transitoire, locale et mondiale est la plus importante. [14] :

— Décentralisation : Sur Internet, l'éthique professionnelle des médias traditionnels n'est pas invoquée. Tout le monde peut utiliser des outils permettant de modifier du contenu en ligne. Par conséquent, nous sommes tous des informateurs potentiels. Des informations apparaissent dans nos profils Facebook, nos fils d'actualité Twitter, nos flux d'e-mails ou sur le blog d'un influenceur personnel. Il est probable que nous soyons tous sur la même page d'envoi de désinformation.
— 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.
— Internationalité : les réseaux « d'amis » ne connaissant pas de frontières, les nouvelles, et donc les fausses informations, font rapidement le tour du globe de mail en mail ou de profil en profil. La confirmation ou l'infirmation d'une fausse information 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 fausses informations que lorsque celles-ci se propagent par la bouche à oreille humain.
— Instantanéité : Une connexion permanente grâce à des outils de plus en plus techniques, les internautes accèdent rapidement aux messageries instantanées et aux réseaux sociaux, ce qui favorise la diffusion rapide de l'information.

1.10 Conclusion

Dans ce chapitre, nous avons exploré les différents concepts des fausses nouvelles. Nous avons examiné leur signification et leurs caractéristiques. L'application de la fouille de données (Data Mining) sur le Web, et plus particulièrement l'analyse d'opinions (Opinion Mining), a permis de proposer des solutions intéressantes à ce problème.

Dans le prochain chapitre, notre attention se portera spécifiquement sur les méthodes d'apprentissage automatique. Les algorithmes d'apprentissage automatique peuvent etre utilis pour analyser et classifier les information afin d'identifier les contenus potentiellement trompeurs ou errons.

CHAPITRE 2. MACHINE LEARNING ET DETECTION DE FAUSSES NOUVELLE

2

2.1 Introduction

La propagation des fausses nouvelles en ligne et des rumeurs peut avoir des conséquences considérables. Les approches traditionnelles basées sur les techniques linguistiques et l'encodage, telles que Arbres de decision , Gradient Boosting Classifier ,Logistic Regression ou Naïve bayes :, offrent une bonne approche pour la détection et l'analyse de ces informations trompeuses. Cependant, ces méthodes ne sont pas suffisantes pour extraire pleinement les fausses nouvelles, en raison de l'évolution constante du langage utilisé et des limites qu'elles présentent pour prendre en compte différents aspects tels que les relations entre les mots et les caractéristiques spécifiques des réseaux sociaux, comme les URL, les auteurs, les horaires, etc.

Pour remédier à cela, les techniques d'apprentissage automatique, qui sont des algorithmes puissants d'intégration du langage naturel, sont utilisées dans la détection des fausses nouvelles, des rumeurs et même des satires. Dans ce chapitre, nous examinerons d'abord la description, les types et la modélisation de l'apprentissage automatique, ainsi que les méthodes les plus couramment utilisées dans les systèmes de détection des fausses nouvelles, en mettant en évidence les principales différences entre elles.

2.2 Définition :

L'apprentissage automatique, également connu sous le nom de machine learning en anglais, est un sous-domaine de l'intelligence artificielle qui se concentre sur le développement d'algorithmes et de modèles permettant aux machines d'apprendre à partir des données et d'améliorer leurs performances sans être explicitement programmées [10]

2.3 Types de Machine Learning

Les algorithmes d'apprentissage automatique sont une vieille chose, mais ce n'est que récemment qu'il est devenu possible d'appliquer des calculs mathématiques complexes aux mégadonnées ; (Un très grand ensemble de données qui nécessiterait une outil de gestion de base de données).
L'apprentissage automatique est utilisé aujourd'hui dans divers domaines, tels que le développement de véhicules autonomes, les systèmes de recommandation en ligne (tels que Netflix et Amazon), l'analyse du sentiment client ou la détection de fraude [15]. Les modèles d'apprentissage automatique peuvent être résumés comme suit : modèles non supervisés, semi-supervisés, supervisés, et de renforcement.

2.3.1 Apprentissage non supervise

L'apprentissage non supervisé est une branche de l'apprentissage automatique où les algorithmes sont utilisés pour explorer et découvrir des structures et des modèles intrinsèques dans un ensemble de données non étiquetées, c'est-à-dire des données pour lesquelles les sorties désirées ne sont pas fournies. Contrairement à l'apprentissage supervisé, il n'y a pas de guidage externe ou de supervision pour l'algorithme. L'objectif principal de l'apprentissage non supervisé est de trouver des structures inhérentes dans les données, tels que des regroupements (clustering) ou des relations de similarité entre les données. Cela permet de regrouper les données similaires ensemble, d'identifier des sous-groupes ou des motifs récurrents, ou de réduire la dimensionnalité des données

en les représentant de manière plus compacte. Voici quelques techniques couramment utilisées en apprentissage non supervisé[13] :

1. **Clustering (regroupement) :** Les algorithmes de clustering sont utilisés pour regrouper les données similaires dans des clusters ou des groupes. Les exemples populaires incluent l'algorithme k-means et la classification hiérarchique.
2. **Réduction de dimensionnalité :** Ces techniques permettent de réduire la dimensionnalité des données en extrayant les caractéristiques les plus significatives. L'analyse en composantes principales (PCA) et le t-SNE sont des exemples courants de techniques de réduction de dimensionnalité.
3. **Règles d'association :** Ces techniques identifient des relations fréquentes ou des modèles d'association entre des éléments dans les données. L'algorithme Apriori est largement utilisé pour extraire des règles d'association à partir de données transactionnelles.
4. **Détection d'anomalies :** Les algorithmes de détection d'anomalies identifient les points de données qui sont considérés comme inhabituels ou aberrants par rapport au reste des données. Ils sont utilisés pour détecter des comportements suspects ou des valeurs aberrantes dans les données.
5. **Apprentissage de représentation :** Ces techniques apprennent une représentation plus informative des données, généralement sous la forme d'un espace vectoriel de plus basse dimension. Des exemples incluent les réseaux de neurones auto-encodeurs et les réseaux de neurones génératifs.

L'apprentissage non supervisé trouve de nombreuses applications pratiques, comme la segmentation des clients, l'analyse des réseaux sociaux, la détection de fraude, l'exploration de données, la recommandation de contenu, etc. Il permet d'extraire des informations utiles à partir de données non étiquetées et peut servir de point de départ pour des tâches d'apprentissage plus avancées.

2.3.2 Apprentissage semi-supervisé

L'apprentissage semi-supervisé est une approche de l'apprentissage automatique qui combine des données étiquetées et non étiquetées pour entraîner un modèle. Contrairement à l'apprentissage supervisé pur où toutes les données sont étiquetées, et à l'apprentissage non supervisé où aucune donnée n'est étiquetée, l'apprentissage semi-supervisé utilise à la fois des exemples étiquetés et des exemples non étiquetés pour améliorer les performances du modèle.
L'idée principale de l'apprentissage semi-supervisé est que les exemples non étiquetés contiennent également des informations utiles pour l'apprentissage. En exploitant ces données non étiquetées, le modèle peut découvrir des structures ou des motifs latents qui améliorent sa capacité de généralisation. L'apprentissage semi-supervisé est particulièrement utile lorsque l'étiquetage manuel des données est coûteux, difficile ou limité.
Une approche courante en apprentissage semi-supervisé est d'utiliser des algorithmes d'apprentissage supervisé traditionnels en y intégrant des exemples non étiquetés. Voici une procédure générale pour l'apprentissage semi-supervisé avec référence [7] :

1. **Collecte des données :** Rassemblez un ensemble de données qui contient à la fois des exemples étiquetés (avec leurs sorties désirées) et des exemples non étiquetés.
2. **Entraînement initial supervisé :** Utilisez les exemples étiquetés pour entraîner un modèle supervisé de base à l'aide d'un algorithme d'apprentissage supervisé tel

que les arbres de décision, les réseaux de neurones, les SVM, etc. Ce modèle initial servira de point de départ.

3. **Sélection d'exemples non étiquetés :** À partir de l'ensemble d'exemples non étiquetés, sélectionnez un sous-ensemble d'exemples non étiquetés pour lesquels vous avez une certaine confiance dans leurs étiquettes potentielles. Cela peut être fait en utilisant des heuristiques, des techniques de clustering, ou en exploitant les caractéristiques des données.
4. **Étiquetage semi-supervisé :** Étiquetez les exemples non étiquetés sélectionnés à l'étape précédente à l'aide d'un expert ou d'un processus automatisé. Ces étiquettes servent de référence.
5. **Entraînement semi-supervisé :** Combiner les exemples étiquetés et les exemples non étiquetés avec leurs étiquettes référencées pour former un nouvel ensemble d'entraînement. Réentraînez le modèle initial en incluant ces exemples non étiquetés pour améliorer sa performance. Cela peut être réalisé en utilisant des techniques telles que l'apprentissage par renforcement, l'auto-encodage, ou des algorithmes spécifiques à l'apprentissage semi-supervisé.
6. **Évaluation et ajustement :** Évaluez les performances du modèle entraîné sur un ensemble de données de test. Si nécessaire, ajustez le modèle en itérant les étapes précédentes jusqu'à ce que vous obteniez des résultats satisfaisants.

2.3.3 Apprentissage supervisé

L'apprentissage supervisé est une approche de l'apprentissage automatique où un modèle est entraîné sur un ensemble de données étiquetées, c'est-à-dire des données pour lesquelles les sorties désirées sont fournies. Le modèle apprend à partir de ces exemples étiquetés pour faire des prédictions précises sur de nouvelles données non étiquetées[1]. Voici les étapes principales de l'apprentissage supervisé :

1. **Collecte des données :** Rassemblez un ensemble de données qui comprend à la fois les entrées (caractéristiques) et les sorties désirées (étiquettes) correspondantes. Les exemples doivent être représentatifs du problème que vous souhaitez résoudre.
2. **Préparation des données :** Avant de commencer l'apprentissage, il est souvent nécessaire de prétraiter les données. Cela peut inclure des étapes telles que le nettoyage des données en éliminant les valeurs manquantes ou aberrantes, la normalisation des caractéristiques pour les mettre à une échelle commune, la conversion des données catégorielles en format numérique, etc.
3. **Séparation des ensembles d'entraînement et de test :** Divisez l'ensemble de données en ensembles d'entraînement et de test. L'ensemble d'entraînement est utilisé pour entraîner le modèle, tandis que l'ensemble de test est utilisé pour évaluer les performances du modèle sur de nouvelles données non vues. Il est important de garder l'ensemble de test séparé pour évaluer objectivement les performances du modèle.
4. **Choix d'un algorithme d'apprentissage :** Sélectionnez un algorithme d'apprentissage approprié en fonction de la nature du problème et des caractéristiques des données. Certains algorithmes couramment utilisés comprennent les arbres de décision, les réseaux de neurones, les machines à vecteurs de support (SVM), les forêts aléatoires, les méthodes ensemblistes, etc. Chaque algorithme a ses propres forces, limitations et hypothèses.

5. **Entraînement du modèle :** Appliquez l'algorithme d'apprentissage sur l'ensemble d'entraînement pour ajuster les paramètres du modèle. Le modèle apprend à partir des exemples étiquetés en minimisant une fonction de coût ou en maximisant une fonction de performance, selon l'algorithme choisi. Les itérations d'optimisation sont effectuées jusqu'à ce que le modèle atteigne un état satisfaisant.
6. **Évaluation du modèle :** Évaluez les performances du modèle sur l'ensemble de test pour mesurer sa capacité à généraliser sur de nouvelles données. Des mesures d'évaluation courantes incluent la précision, le rappel, la F-mesure, l'aire sous la courbe ROC (AUC-ROC), etc. Ces mesures permettent de comprendre la qualité des prédictions du modèle.
7. **Réglage des hyperparamètres :** Certains algorithmes d'apprentissage ont des hyperparamètres qui doivent être réglés pour optimiser les performances du modèle. Il est courant d'utiliser des techniques telles que la recherche par grille, la recherche aléatoire ou l'optimisation bayésienne pour trouver les valeurs optimales des hyperparamètres.
8. **Prédiction sur de nouvelles données :** Une fois que le modèle est entraîné et évalué, il peut être utilisé pour faire des prédictions sur de nouvelles données non étiquetées. Le modèle utilise les caractéristiques de ces nouvelles données pour générer des prédictions sur les sorties correspondantes.

L'apprentissage supervisé est largement utilisé dans de nombreux domaines tels que la classification d'images, la prédiction de prix, la détection de fraudes, le diagnostic médical, la reconnaissance vocale, etc. L'efficacité de l'apprentissage supervisé dépend de la qualité et de la représentativité des données d'entraînement, du choix approprié de l'algorithme d'apprentissage, ainsi que de la capacité à régler les hyperparamètres de manière adéquate

2.3.4 Apprentissage par renforcement

Ce type d'apprentissage multiplie les tentatives pour découvrir quelles actions apportent le plus de récompenses. Ce type d'apprentissage comprend trois composants principaux : l'agent (qui apprend ou prend des décisions), l'environnement (tout ce avec quoi l'agent interagit) et les actions (ce que l'agent peut faire). L'objectif était que l'agent choisisse des actions qui maximisent les récompenses attendues sur une période donnée. Il atteindra cet objectif plus rapidement en suivant les règles établies, apprenant ainsi les meilleures règles[15].

2.4 Avantages et incovenients de quelques algoritmes d'apprentissage automatique

2.4.1 Arbres de decision

Un arbre de décision est un modèle prédictif qui, comme son nom l'indique, peut être considéré comme un arbre. Plus précisément, chaque branche de l'arbre est une question de classification et les feuilles de l'arbre sont des partitions de l'ensemble de données avec leur classification .Un arbre de décision est une représentation simple des connaissances et ils classent les exemples dans un nombre fini de classes, les nœuds sont étiquetés avec des noms d'attributs, les bords sont étiquetés avec des valeurs possibles pour cet attribut et les feuilles étiquetées avec différentes classes. Les objets sont classés en

suivant un chemin dans l'arborescence, en prenant les bords, correspondant aux valeurs des attributs d'un objet. Il y a des points intéressants à propos de l'arbre [18] :

- Il divise les données sur chaque point de branchement sans perdre aucune des données (le nombre total d'enregistrements dans un nœud parent donné est égal à la somme des enregistrements contenus dans ses deux enfants).
- Il est assez facile de comprendre comment le modèle est construit (contrai- rement aux modèles des réseaux de neurones ou des statistiques standard).
- Il serait également assez facile d'utiliser ce modèle si vous deviez réellement cibler les clients susceptibles de proposer une offre marketing ciblée.

1-Fonctionnement :[18]

- La décision est tirée de l'apprentissage de l'ensemble de données (DataSet).
- Chaque nœud non-feuille dans une arborescence représente une fonc- tionnalité et chaque branche représente une valeur que la fonctionna- lité peut prendre.
- Les instances sont classées en suivant un chemin qui commence au nœud racine et se termine à une feuille en suivant les branches en fonction des valeurs de fonctionnalité d'instance.
- La construction des arbres de décision est basée sur l'heuristique.

2-Les forces : [18]

- Plusieurs arbres de décision peuvent être informés à partir du même DataSet.
- Effectuez implicitement un filtrage de variables ou une sélection de fonctionnalités.
- Nécessite relativement peu d'efforts de la part des utilisateurs pour la préparation des données.
- Les relations non linéaires entre les paramètres n'affectent pas les performances de l'arborescence.
- Facile à interpréter et à expliquer.
- Permettre d'ajouter de nouveaux scénarios possibles.
- Interprétabilité : Les arbres de décision sont des modèles facilement interprétables et compréhensibles. Les décisions prises par l'algorithme peuvent être expliquées de manière simple et compréhensible, même pour les non-experts en apprentissage automatique.
- Adaptabilité : Les arbres de décision sont des modèles adaptables qui peuvent être utilisés pour la classification, la régression et la classification multi-étiquettes. Ils peuvent également être utilisés avec des données numériques et catégoriques, et sont utiles pour des problèmes de modélisation non linéaires.
- Rapidité : Les arbres de décision sont des modèles rapides et efficaces. Une fois que l'arbre est construit, la prédiction est rapide car elle implique simplement de traverser l'arbre.
- Traitement des données manquantes : Les arbres de décision peuvent gérer les données manquantes. Les données manquantes sont automatiquement traitées lors de la construction de l'arbre, ce qui permet de les inclure dans le modèle.
- Robustesse : Les arbres de décision sont robustes aux valeurs aberrantes dans les données. Ils peuvent également être utilisés pour réduire le bruit dans les données.
- Pas de supposition sur la distribution des données : Contrairement à de nombreux autres algorithmes d'apprentissage automatique, les arbres de décision ne font pas

d'hypothèses sur la distribution des données. Ils peuvent donc être utilisés avec des données de différentes distributions.

Ces avantages font des arbres de décision un choix populaire pour de nombreuses applications d'apprentissage automatique, notamment dans le domaine de la classification et de la prédiction.

Bien que les arbres de décision aient de nombreux avantages, ils présentent également quelques limites :

- Les arbres de décision sont construits de manière récursive à par- tir de données d'apprentissage en utilisant une approche gloutonne descendante dans laquelle les caractéristiques sont sélectionnées séquentiellement.
- Faible performance car vous devez « redessiner l'arborescence » chaque fois que vous souhaitez mettre à jour votre modèle CART et mauvaise résolution des données avec des relations complexes entre les variables.
- Ce ne sont que deux possibilités (gauche-droite) à chaque nœud, il y a donc des relations variables que les arbres de décision ne peuvent tout simplement pas apprendre.
- Pratiquement limité à la classification.
- Surapprentissage : Les arbres de décision peuvent facilement surapprendre les données d'entraînement. Cela se produit lorsque l'arbre est trop complexe et capture le bruit dans les données d'entraînement, plutôt que le véritable signal. Cela peut entraîner des performances médiocres sur de nouvelles données.
- Instabilité : Les arbres de décision peuvent être instables, c'est-à-dire que de petits changements dans les données d'entraînement peuvent entraîner de grands changements dans l'arbre final. Cela peut rendre l'interprétation de l'arbre plus difficile.
- Biais de sélection : Les arbres de décision peuvent être biaisés en faveur des classes majoritaires ou des variables fortement corrélées avec la variable cible. Cela peut entraîner une prédiction inexacte pour les classes minoritaires ou pour les nouvelles données qui ne sont pas corrélées avec les variables utilisées pour construire l'arbre.
- Sensibilité aux paramètres : Les performances des arbres de décision dépendent des paramètres utilisés pour construire l'arbre, tels que la profondeur maximale de l'arbre et le nombre minimal d'exemples dans une feuille. La sélection des bons paramètres peut être difficile et peut nécessiter une recherche intensive.
- Manque de généralisation : Les arbres de décision peuvent être spécifiques à un ensemble de données particulier. Ils peuvent ne pas généraliser bien à de nouveaux ensembles de données, en particulier si les distributions des données sont différentes.

Ces inconvénients doivent être pris en compte lors de l'utilisation des arbres de décision pour résoudre des problèmes d'apprentissage automatique, et des stratégies doivent être mises en place pour atténuer ces problèmes, tels que la sélection de paramètres appropriés et la validation croisée.

2.4.2 Gradient Boosting Classifier

Le gradient boosting classifier (GBC) est un algorithme d'apprentissage automatique supervisé qui utilise une méthode d'ensemble pour construire un modèle prédictif à partir de plusieurs modèles plus simples. Plus précisément, le GBC construit des arbres de décision à partir de plusieurs itérations en utilisant un gradient de perte pour minimiser

l'erreur de prévision[23].

1-Fonctionnement :[23]

— Initialisation : Le GBC commence par initialiser les prédictions en utilisant une constante qui représente la valeur moyenne de la variable cible dans les données d'entraînement.
— Calcul du gradient : Le GBC calcule le gradient de la fonction de perte par rapport aux prédictions actuelles pour chaque exemple d'entraînement. Le gradient mesure la pente de la fonction de perte par rapport à la prédiction actuelle, indiquant ainsi la direction et la magnitude de la correction requise pour améliorer la précision du modèle.
— Mise à jour des prédictions : Les prédictions sont mises à jour en ajoutant les prédictions de chaque nouvel arbre construit, pondérées par un facteur de taux d'apprentissage qui contrôle la contribution relative de chaque arbre à la prédiction globale.
— Construction d'un arbre de décision : Le GBC construit un arbre de décision pour capturer les relations entre les variables d'entrée et la variable cible, en utilisant les gradients calculés à l'étape précédente pour guider la construction de l'arbre. Plus précisément, l'arbre est construit pour minimiser la perte de manière itérative en choisissant la variable d'entrée et le seuil de division qui réduisent la perte maximale pour chaque sous-ensemble des données d'entraînement.
— Répétition des étapes 2 à 3 : Les étapes 2 à 3 sont répétées plusieurs fois pour construire plusieurs arbres de décision et améliorer la précision de la prédiction.
— Prédiction finale : La prédiction finale est obtenue en agrégeant les prédictions de tous les arbres de décision construits.

2-Les forces : [23]

— Précision : Le GBC est un algorithme très précis qui peut atteindre des performances proches de l'état de l'art pour la plupart des tâches de classification et de régression.
— Gestion de données hétérogènes : Le GBC est capable de gérer des données hétérogènes, notamment des variables catégoriques et numériques.
— Interprétabilité : Les modèles GBC sont relativement faciles à interpréter car ils peuvent fournir des informations sur l'importance relative des différentes variables d'entrée.
— Capacité à gérer de grandes quantités de données : Le GBC peut gérer efficacement de grandes quantités de données, grâce à sa capacité à construire des arbres de décision en parallèle.
— Possibilité d'ajustement des paramètres : Le GBC offre de nombreuses options de paramétrage, permettant ainsi d'ajuster les paramètres pour optimiser les performances du modèle.
— Robustesse : Le GBC est robuste aux données manquantes et aux valeurs aberrantes, grâce à sa capacité à utiliser des arbres de décision résilients.

— Haute précision : Le GBC est connu pour sa haute précision dans les tâches de classification et de régression.
— Flexibilité : Le GBC est flexible car il peut être utilisé pour des tâches de classification et de régression, ainsi que pour des données catégoriques et numériques.
— Interprétabilité : Les modèles GBC sont relativement faciles à interpréter car ils fournissent des informations sur l'importance relative des différentes variables d'entrée.
— Possibilité de gérer de grandes quantités de données : Le GBC peut gérer de grandes quantités de données grâce à sa capacité à construire des arbres de décision en parallèle.
— Adaptabilité : Le GBC peut être adapté pour de nombreuses tâches différentes en ajustant les paramètres tels que la profondeur de l'arbre, le taux d'apprentissage et le nombre d'estimateurs.
— Robustesse : Le GBC est robuste aux valeurs aberrantes et aux données manquantes, car il utilise des arbres de décision résilients pour construire le modèle.
— Résistance à l'overfitting : Le GBC est moins susceptible de surapprentissage que d'autres algorithmes de boosting tels que l'AdaBoost, car il utilise une technique de régularisation appelée "shrinkage" pour réduire la contribution de chaque arbre.

Malgré ses nombreux avantages, le Gradient Boosting Classifier (GBC) peut également présenter certains inconvénients, notamment :

— Coût computationnel élevé : Le GBC peut être relativement lent pour les grandes quantités de données et nécessiter des ressources informatiques importantes pour être entraîné.
— Sensibilité aux données de mauvaise qualité : Comme tout algorithme d'apprentissage automatique, le GBC est sensible aux données de mauvaise qualité, telles que les données manquantes ou les valeurs aberrantes.
— Risque de surapprentissage : Comme mentionné précédemment, le GBC est sensible au surapprentissage, en particulier si les paramètres ne sont pas correctement ajustés.
— Limitations de la représentation de l'information : Les arbres de décision utilisés par le GBC sont limités dans leur capacité à représenter des relations complexes entre les variables, ce qui peut affecter la qualité du modèle.
— Difficulté à interpréter les résultats : Bien que les modèles GBC soient relativement faciles à interpréter, la complexité de l'algorithme peut rendre difficile la compréhension des relations entre les variables.
— Nécessité de connaissances spécialisées : La mise en place et l'utilisation du GBC peuvent nécessiter une connaissance spécialisée des techniques d'apprentissage automatique et des outils informatiques.
— Sensible au surapprentissage : Comme tout algorithme d'apprentissage automatique, le GBC peut être sensible au surapprentissage si les paramètres ne sont pas correctement ajustés.
— Temps de traitement : Le GBC peut être relativement lent pour les grandes quantités de données, en particulier si la profondeur de l'arbre est importante.
— Sensibilité aux valeurs aberrantes : Bien que le GBC soit généralement robuste aux valeurs aberrantes, il peut être sensible aux valeurs aberrantes extrêmes qui peuvent influencer le modèle de manière disproportionnée.
— Difficulté à choisir les paramètres optimaux : Le GBC offre de nombreuses options

de paramétrage, ce qui peut rendre difficile la sélection des paramètres optimaux pour maximiser les performances du modèle.
bien que le GBC soit un algorithme d'apprentissage automatique puissant et flexible, il peut présenter des limitations qui nécessitent une attention particulière lors de son utilisation pour l'analyse de données. Il est important de bien comprendre les forces et les faiblesses de l'algorithme pour l'utiliser de manière efficace et éviter les erreurs courantes.

2.4.3 Logistic Regression

La régression logistique est un algorithme de classification supervisée qui est utilisé pour prédire une variable binaire (0 ou 1). Elle utilise une fonction logistique pour calculer la probabilité qu'une observation appartienne à une classe particulière.
La régression logistique est largement utilisée dans les domaines de la recherche médicale, de la biologie, de la psychologie, de l'économie et de la finance, entre autres, pour prédire des résultats binaires tels que la survie ou la mort, la guérison ou la non-guérison, le succès ou l'échec, etc.
Le modèle de régression logistique calcule la probabilité que la variable dépendante prenne l'une ou l'autre de ses deux valeurs possibles en fonction des valeurs des variables indépendantes. Le modèle est généralement estimé en utilisant la méthode du maximum de vraisemblance, qui consiste à trouver les valeurs des coefficients de régression qui maximisent la probabilité d'observer les données réelles[8].

1-Fonctionnement :

Le fonctionnement de la régression logistique peut être décrit en plusieurs étapes [8] :
— Préparation des données : Les données sont collectées et préparées en vue de l'analyse. Les variables d'entrée doivent être indépendantes les unes des autres et il doit y avoir suffisamment de données pour permettre une analyse statistique.
— Définition de la variable cible : La variable cible doit être une variable binaire qui prend les valeurs 0 ou 1.
— Sélection des variables d'entrée : Les variables qui ont une relation avec la variable cible sont sélectionnées pour être utilisées dans le modèle de régression logistique.
— Estimation des coefficients : Les coefficients de régression sont estimés à partir des données d'entrée et de la variable cible. Les coefficients indiquent l'impact de chaque variable d'entrée sur la variable cible.
— Calcul de la probabilité : La fonction logistique est utilisée pour calculer la probabilité qu'une observation appartienne à une classe particulière.
— Détermination de la classe prédite : La classe prédite est déterminée en fonction de la probabilité calculée. Si la probabilité est supérieure à un seuil défini (0,5 par défaut), l'observation est classée dans la classe positive, sinon elle est classée dans la classe négative.
— Évaluation du modèle : Le modèle de régression logistique est évalué en utilisant des métriques de performance telles que la précision, le rappel et la courbe ROC.

2-Les forces :

La régression logistique a plusieurs forces qui en font une méthode statistique populaire et efficace dans de nombreux domaines [8] :

— Facilité d'interprétation : Les coefficients de régression dans la régression logistique ont une interprétation intuitive qui facilite la compréhension de la relation entre la variable dépendante et les variables indépendantes. Les coefficients indiquent l'impact de chaque variable indépendante sur la variable dépendante.
— Bonne performance pour des ensembles de données de taille moyenne : La régression logistique est souvent performante pour des ensembles de données de taille moyenne, ce qui en fait un outil utile pour les applications courantes.
— Robustesse : La régression logistique est relativement robuste aux violations mineures des hypothèses sous-jacentes.
— Peut être régularisée : La régression logistique peut être régularisée pour éviter le surapprentissage et améliorer la généralisation du modèle.
— La régression logistique est l'un des algorithmes d'apprentissage automatique les plus simples et est facile à mettre en uvre tout en offrant une grande efficacité de formation dans certains cas. Aussi, pour ces raisons, la formation d'un modèle avec cet algorithme ne nécessite pas une puissance de calcul élevée.
— Les paramètres prédits (poids entraînés) donnent une inférence sur l'importance de chaque caractéristique. La direction de l'association, c'est-à-dire positive ou négative, est également indiquée. Nous pouvons donc utiliser la régression logistique pour découvrir la relation entre les caractéristiques.
— Cet algorithme permet aux modèles d'être facilement mis à jour pour refléter de nouvelles données, contrairement aux arbres de décision ou aux machines vectorielles de support. La mise à jour peut être effectuée à l'aide de la descente de gradient stochastique.
— La régression logistique produit des probabilités bien calibrées ainsi que des résultats de classification. C'est un avantage par rapport aux modèles qui ne donnent que la classification finale comme résultats. Si un exemple de formation a une probabilité de 95% pour une classe et un autre a une probabilité de 55% pour la même classe, nous obtenons une inférence sur les exemples de formation les plus précis pour le problème formulé.
— Dans un Dataset de faible dimension comportant un nombre suffisant d'exemples d'entraînement, la régression logistique est moins sujette au sur-ajustement.
— Plutôt que de commencer tout de suite avec un modèle complexe, la régression logistique est parfois utilisée comme modèle de référence pour mesurer les performances, car elle est relativement rapide et facile à mettre en œuvre.
— La régression logistique s'avère très efficace lorsque le jeu de données comporte des entités linéairement séparables.

3-Les limites :[8]

— La classification de texte est un problème classique.
— Il est instable quand un prédicteur pourrait presque expliquer la variable de réponse, car le coefficient de cette variable sera augmenté au plus haut possible, voici un cas où les gens se tournent vers l'analyse discriminante.
— Requiert plus d'hypothèses et est sensible aux valeurs aberrantes.
— La régression logistique suppose que la relation entre les variables indépendantes et la variable dépendante est linéaire. Si la relation n'est pas linéaire, la régression logistique ne sera pas adaptée.
— La régression logistique peut être sensible aux valeurs aberrantes, car elle est basée

sur une fonction de probabilité. Les valeurs aberrantes peuvent affecter la distribution des données et donc la précision des prédictions.
— La régression logistique peut être sujette au surajustement si le modèle est trop complexe ou si le nombre de variables indépendantes est supérieur au nombre d'observations dans l'ensemble de données.
— La régression logistique ne peut pas gérer les valeurs manquantes de manière automatique. Si les données contiennent des valeurs manquantes, il est nécessaire de traiter ces valeurs avant d'appliquer la régression logistique.
— La régression logistique peut être sensible à la multicollinéarité, qui est une situation où deux ou plusieurs variables indépendantes sont hautement corrélées entre elles. Cela peut rendre difficile l'interprétation des coefficients de régression et peut également réduire la précision des prédictions.

la régression logistique est un outil statistique puissant pour la prédiction de variables binaires, mais elle présente également certaines limitations qu'il convient de prendre en compte lors de son utilisation.

2.4.4 Naïve bayes :

Naïve Bayes est un algorithme d'apprentissage automatique qui est couramment utilisé pour les tâches de classification. Il est basé sur le théorème de Bayes, qui permet de calculer la probabilité qu'un événement se produise en fonction de la probabilité de certaines hypothèses.
Dans le cas de Naïve Bayes, chaque observation est représentée par un ensemble de caractéristiques (variables) qui sont indépendantes les unes des autres (d'où le terme "naïf"). L'algorithme utilise ces caractéristiques pour estimer la probabilité que l'observation appartienne à chaque classe possible, puis attribue l'observation à la classe avec la probabilité la plus élevée.
Naïve Bayes est souvent utilisé pour des tâches de classification de texte, telles que la catégorisation de courriels en spam et non-spam, ou la classification de commentaires en positifs et négatifs. Il est également connu pour être rapide et efficace, même avec de grandes quantités de données[11].

1-Fonctionnement :

— La technique Naïve Bayes est généralement utilisée pour la détection d'intrusions en combinaison avec des schémas statistiques, une procé- dure qui présente plusieurs avantages, notamment la capacité de coder les interdépendances entre les variables et de prédire les événements, ainsi que la capacité d'incorporer à la fois des connaissances et des données antérieures[11].
— Le réseau bayésien est un modèle qui code les relations probabilistes entre les variables d'intérêt
— Dans l'apprentissage automatique, les classificateurs de Bayes naïfs sont une famille de classificateurs probabilistes simples basés sur l'ap- plication du théorème de Bayes avec des hypothèses d'indépendances fortes (naïves) entre les caractéristiques.

2-Les forces :[11]

— Facile à utiliser.
— Efficace si l'ensemble d'entraînement est suffisamment grand.

— Capacité d'apprendre, au fur et à mesure que l'ensemble d'entraînement s'agrandit, les résultats deviennent de plus en plus précis (intelligence).
— Malgré sa simplicité, Naïve Bayes peut souvent surpasser les méthodes de classification plus sophistiquées.
— Si l'hypothèse d'indépendance conditionnelle NB se vérifie, un classificateur Naïf Bayes convergera plus rapidement que les modèles discriminants comme la régression logistique, vous aurez donc besoin de moins de données d'entraînement .
— Les classificateurs Naïve Bayes simplifient les calculs et présentent une précision et une rapidité élevées lorsqu'ils sont appliqués à de grandes bases de données.
— Les classificateurs bayésiens donnent des résultats satisfaisants car l'accent est mis sur l'identification des lasses pour les instances, pas sur les probabilités exactes).
— Il est rapide et facile à utiliser et à prédire la classe de l'ensemble de données de test.
— Lorsque l'hypothèse d'indépendance est valable, cela fonctionne mieux que les autres algorithmes de classification.
— Cela fonctionne bien dans le cas de variables d'entrée catégorielles à la place de variables numériques.

3-Les limites :[11]

— Ne tient pas compte de la séquence de mots (fonction de mot non pertinente).
— Il ne peut pas apprendre les interactions entre les fonctionnalités.
— Est basé sur le théorème donc bayésien et est particulièrement adapté lorsque la dimensionnalité des entrées est élevée.
— Dans l'approche Naïve Bayes, on suppose que les attributs des don- nées sont conditionnellement indépendants.
— Un effort de calcul considérablement plus élevé est requis.
— Elle se limite à l'indépendance de l'hypothèse du prédicteur.
— Si la variable catégorielle a une catégorie (dans la base de test), qui n'a pas été observée dans l'ensemble de données d'apprentissage, alors le modèle attribuera une probabilité de 0 (zéro) et ne pourra pas faire de prédiction.

2.4.5 Random Forest Classification :

Modèle de classificateur aléatoire : est un modèle de classification qui assigne des étiquettes de classe de manière aléatoire, sans prendre en compte les caractéristiques des données d'entrée. En d'autres termes, le modèle attribue une classe à chaque instance d'entrée de manière totalement arbitraire, sans aucune relation avec les caractéristiques de cette instance [22].

1-Fonctionnement :

Le fonctionnement d'un modèle de classificateur aléatoire est très simple : il assigne une classe de manière aléatoire à chaque instance d'entrée, sans tenir compte des caractéristiques de cette instance. Par exemple, si nous avons un ensemble de données contenant des images de chats et de chiens, un modèle de classificateur aléatoire attribuera une étiquette de classe "chat" ou "chien" à chaque image de manière aléatoire, sans considérer les caractéristiques de l'image, telles que la couleur, la forme ou la texture[22].

Bien que cela puisse sembler inutile, un modèle de classificateur aléatoire peut être utilisé comme référence de base pour évaluer les performances de modèles plus avancés. En effet, tout modèle de classification que nous construisons doit être capable de surpasser la performance d'un modèle de classificateur aléatoire.
Cependant, il convient de noter que le modèle de classificateur aléatoire ne fournit aucune information sur les données d'entrée et ne peut donc pas être utilisé pour résoudre des problèmes de classification réels.

2-Les forces :

— Le modèle de classificateur aléatoire est rapide à exécuter, car il ne nécessite aucun calcul ou traitement complexe. Cela en fait un choix idéal pour les ensembles de données de grande taille qui peuvent prendre beaucoup de temps à traiter avec des modèles de classification plus avancés[22].
— Le modèle de classificateur aléatoire est très utile pour comprendre l'importance des caractéristiques d'entrée pour un problème de classification donné. Si un modèle de classification plus avancé donne des résultats similaires à ceux du modèle de classificateur aléatoire, cela peut indiquer que les caractéristiques d'entrée n'ont pas beaucoup d'importance pour résoudre le problème de classification.
— Le modèle de classificateur aléatoire est également utile pour évaluer les performances d'un modèle de classification lorsqu'il y a un déséquilibre entre les classes d'entrée. Dans un tel cas, le modèle de classificateur aléatoire fournit une mesure de référence pour évaluer les performances du modèle de classification sur les classes minoritaires.
— Enfin, le modèle de classificateur aléatoire est souvent utilisé comme une comparaison de base pour d'autres modèles de classification aléatoires, tels que les forêts aléatoires et les réseaux de neurones aléatoires. Cela permet de déterminer si ces modèles plus avancés ont réellement un impact sur les performances de classification par rapport au modèle de classificateur aléatoire.
— Le modèle de classificateur aléatoire est très simple à mettre en œuvre et à utiliser. Il ne nécessite pas de connaissances techniques avancées ou de préparation de données complexes.
— Le modèle de classificateur aléatoire fournit une référence de base pour évaluer les performances de modèles de classification plus avancés. En comparant les performances des modèles de classification avancés à celles du modèle de classificateur aléatoire, on peut déterminer s'il y a une amélioration significative dans les performances de classification.
— Le modèle de classificateur aléatoire est rapide à exécuter, car il ne nécessite aucun calcul ou traitement complexe. Cela en fait un choix idéal pour les ensembles de données de grande taille qui peuvent prendre beaucoup de temps à traiter avec des modèles de classification plus avancés.
— Le modèle de classificateur aléatoire est très utile pour comprendre l'importance des caractéristiques d'entrée pour un problème de classification donné. Si un modèle de classification plus avancé donne des résultats similaires à ceux du modèle de classificateur aléatoire, cela peut indiquer que les caractéristiques d'entrée n'ont pas beaucoup d'importance pour résoudre le problème de classification.
— Le modèle de classificateur aléatoire est également utile pour évaluer les performances d'un modèle de classification lorsqu'il y a un déséquilibre entre les classes d'entrée. Dans un tel cas, le modèle de classificateur aléatoire fournit une mesure

de référence pour évaluer les performances du modèle de classification sur les classes minoritaires.

3-Les limites :

— Le modèle de classificateur aléatoire ne fournit aucune information sur les données d'entrée et ne peut donc pas être utilisé pour résoudre des problèmes de classification réels[22].

— Le modèle de classificateur aléatoire attribue des classes de manière totalement aléatoire, sans tenir compte des caractéristiques des données d'entrée. Cela signifie que le modèle ne peut pas apprendre des relations complexes entre les caractéristiques d'entrée et les classes de sortie, ce qui le rend inutile pour résoudre des problèmes de classification réels.

— Bien que le modèle de classificateur aléatoire puisse être utilisé comme une référence pour évaluer les performances de modèles plus avancés, il ne fournit pas d'informations sur les raisons pour lesquelles ces modèles sont meilleurs ou pires que le hasard. Par conséquent, il est important de considérer d'autres mesures d'évaluation, telles que la précision, le rappel, la F-mesure, etc.

— Enfin, le modèle de classificateur aléatoire peut être trompeur dans certains cas, en particulier lorsque les classes d'entrée sont déséquilibrées. Par exemple, si une classe minoritaire a une probabilité de 0,1 d'être prédite correctement par le modèle de classificateur aléatoire, cela peut sembler acceptable, mais cela signifie que le modèle est très mauvais pour cette classe et ne fournit aucune information utile pour la résoudre.

— Le modèle de classificateur aléatoire est totalement incapable de fournir des résultats précis ou utiles pour la classification des données. Il attribue les classes au hasard, sans aucune compréhension des relations entre les caractéristiques d'entrée et les classes de sortie, ce qui le rend inutile pour résoudre des problèmes de classification réels.

— Bien que le modèle de classificateur aléatoire puisse être utilisé pour évaluer les performances de modèles de classification plus avancés, il ne fournit pas de mesures d'évaluation précises sur la performance du modèle. Les résultats obtenus avec le modèle de classificateur aléatoire ne sont pas suffisants pour fournir des informations précises sur la qualité du modèle.

— Le modèle de classificateur aléatoire ne prend pas en compte les différences entre les caractéristiques d'entrée et ne peut pas être utilisé pour effectuer une sélection de caractéristiques efficace. Les caractéristiques d'entrée qui sont inutiles pour la classification sont traitées de la même manière que les caractéristiques importantes, ce qui peut fausser les résultats de l'évaluation des performances.

— Enfin, le modèle de classificateur aléatoire peut être trompeur dans certains cas, en particulier lorsque les classes d'entrée sont déséquilibrées. Le modèle peut fournir des résultats acceptables pour les classes majoritaires, mais peut être totalement inutile pour les classes minoritaires. Cela peut conduire à des résultats biaisés qui ne reflètent pas la performance réelle du modèle.

2.5 Modèles de détection des fausses nouvelles

Nous allons maintenant discuter des modèles de détection des fausses nouvelles[6]

2.5.1 Modèles du contenu d'actualités :

les modèles de contenu des actualités peuvent être classifiés en ce qui suit :

— **Modèle basé sur les connaissances :** Approche basée sur les connaissances, il vise à utiliser des sources pour vérifier l'authenticité du contenu des nouvelles.

— **Modèle basé sur le style :** Les rédacteurs de fausses nouvelles utilisent certaines des techniques d'écriture spécifiques nécessaires pour réaliser un véritable reportage, comme les mots neutres, qui décrivent les événements avec des faits. Meilleure qualité d'écriture (prise en compte des mots accentués, de la ponctuation et de la longueur des phrases).

2.5.2 Modèles du contexte social :

les réseaux sociaux fournissent des ressources supplémentaires aux chercheurs pour compléter et améliorer les modèles de contexte d'actualité, les modèles du contexte social sont aussi utilisés pour la détection des rumeurs et l'identification des faux contenus sur Facebook. Ces modèles sont en fonction de la position et de la propagation.

— **Basé sur la position :** C'est un processus qui peut déterminer les sentiments exprimés par l'utilisateur, pour, contre ou neutre. Cela signifie que la position de l'utilisateur est explicite ou implicite. Les emplacements explicites sont ceux où les lecteurs donnent des expressions directes, comme un pouce levé ou l'utilisation d'emojis. Les situations implicites sont des situations où les émotions sont extraites des publications sur les réseaux sociaux.

— **Détection des rumeurs :** Le but de la détection des rumeurs est de classer les nouvelles en tant que rumeurs ou faits. C'est un modèle en quatre étapes : tracer, détecter, positionner et valider.

— **Basé sur la propagation :** C'est un processus qui peut définir la relation entre les messages. Il définit donc les événements liés entre eux. Il existe deux catégories, la propagation homogène et hétérogène. La première contient une seule entité, comme un message ou un événement, et la seconde contient plusieurs entités à la fois.

2.6 Travaux connexes

— Les auteurs de [9] proposent une typologie de plusieurs méthodes d'évaluation de la véracité émergeant de deux catégories principales : les approches de repères linguistiques (avec apprentissage automatique) et les approches d'analyse de réseau, pour la détection des fausses nouvelles.

— Les auteurs de [12] présentent une approche simple pour la détection de fausses nouvelles en utilisant un classificateur NaiveBayse. Cette approche est testée par rapport à un ensemble de données extraites des messages d'actualités Facebook. Ils proclament pouvoir atteindre une précision de 74%.

— Les auteurs de [5] proposent un modèle détection de fausses nouvelles qui utilise l'analyse n-gram et les techniques d'apprentissage automatique en comparant deux techniques d'extraction de caractéristiques diérentes et six techniques de classification . Les expérimentations menées montrent que les meilleures performances sont obtenues en utilisant la méthode d'extraction des caractéristiques dite (TF-IDF). Le classieur utilisé est la machine à vecteurs de support linéaire (LSVM), qui a donné une précision de 92%.

— Les auteurs de [19] décrivent Comment les utilisateurs des réseaux sociaux peuvent s'assurer de la vérité des information, comment les critères du vrai et du faux se définissent. Ils décrivent aussi les mécanismes qui permettent leur validation et le rôle des journalistes ou ce qu'il faut attendre des chercheurs et des institutions ocielles.

— Les auteurs de [16] proposent plusieurs stratégies et types d'indices relevant de diérentes modalités (texte, image, informations sociales). Ils explorent également l'intérêt de combiner et fusionner ces approches pour évaluer et vérifier l'information partagée

2.7 Conclusion

Dans ce chapitre, nous avons examiné la description et les principes de fonctionnement des différents types d'apprentissage automatique, ainsi que les définitions, les avantages et les inconvénients des algorithmes d'apprentissage supervisé les plus couramment utilisés dans la détection des fausses nouvelles.

Dans le prochain chapitre, nous aborderons l'implementation de notre solution pour résoudre ce problème.

CHAPITRE 3. CONCEPTION DU SYSTÈME, IMPLÉMENTATION ET RÉSULTATS

3

3.1 Introduction

Ce chapitre vise à présenter les étapes de mise en œuvre d'un système informatique, allant de sa réalisation à son déploiement pour permettre l'interaction avec les utilisateurs. L'objectif principal est de présenter les différents outils utilisés dans notre système, tels que les logiciels, les langages de programmation, les bibliothèques et les données. Ensuite, nous aborderons l'application concrète de ces outils, suivie d'une discussion des résultats obtenus.

3.2 Environnement et les outils de travail

3.2.1 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 :

Table 3.1 – Caractéristiques du matériel utilisé

	Poste de travail N01	Poste de travail N01
PC	HP	ACER
System exploitation	Windows 10 Professional	Windows 8
Processor	Intel(R) Core(tm) i5-4310M CPU @ 2.70ghz	Intel(R) Core(tm) i3-3217M
RAM	8,00 Go	4,00 Go
Type de system	SE 64 bits	SE 64 bits

3.2.2 Langage de programmation :

Nous avons choisi Python, version 3.11.3 pour implémenter notre système de détection de fausses informations. Python est un langage de programmation interprété, multiparadigme 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 .

Pour éditer le code de notre système, nous avons utilisé Google Colaboratory, une plateforme cloud gratuite fournie par Google, pour éditer le code de notre système. Colab est basé sur Jupyter Notebook, version 6.0.3, et 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 de Colab incluent le nettoyage et la transformation des données, la modélisation statistique, la visualisation des données, l'apprentissage automatique et bien plus encore. Avec Colab, nous avons pu travailler sur notre système de manière pratique et efficace dans un environnement de développement en ligne sans avoir besoin d'installer localement Jupyter sur nos propres machines.

FIGURE 3.1 – Editeur de code et bibliothèques Python utilisés

3.2.3 Librairies et bibliothèques Python

— **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[4].

— **Matplotlib :** Matplotlib est une bibliothèque Python utilisée pour la visualisation de données. Elle offre une grande variété de fonctionnalités pour créer des graphiques, des diagrammes et des visualisations de qualité. Matplotlib permet de personnaliser de nombreux aspects des visualisations, tels que les couleurs, les étiquettes et les légendes. Elle est largement utilisée dans la recherche, l'analyse de données et d'autres domaines pour représenter les données de manière claire et informative[24].

— **NumPy :** est une bibliothèque libre pour le langage de programmation Python qui fournit des fonctionnalités pour la manipulation efficace de tableaux multidimensionnels (tels que des matrices et des vecteurs) ainsi que des fonctions mathématiques pour effectuer des opérations sur ces tableaux. NumPy est largement utilisé dans le domaine de la science des données, de l'apprentissage automatique et de la recherche scientifique en raison de sa performance optimisée et de sa syntaxe concise. En comparaison[3].

— **Scikit-learn** : est une bibliothèque Python dédiée à l'apprentissage automatique. Elle propose des fonctionnalités pour la classification, la régression, le clustering et la réduction de dimensionnalité. Scikit-learn facilite la création, l'entraînement et l'évaluation de modèles d'apprentissage automatique. Elle inclut également des outils pour la préparation des données, la validation croisée, la sélection de modèles et les métriques de performance. Scikit-learn est largement utilisée dans les domaines de l'analyse de données et de la science des données pour résoudre divers problèmes d'apprentissage automatique[2].

3.3 Analyse du Dataset

3.3.1 Informations generales sur Dataset

Nous avons créé un nouvel ensemble de données appelé "Constraint_train" à partir de deux autres ensembles de données disponibles publiquement. Notre ensemble de données final contient trois colonnes, à savoir "ID", "Tweet" et "Label" (avec les valeurs "Fake" ou "True"). Les données ont été extraites de publications sur twitter(étiquetées comme "Post") . Les détails de notre ensemble de données sont résumés dans le tableau 3.2 ci-dessous.

Table 3.2 – Caractéristiques de Dataset

Nom de l'ensemble de données	Constraint-train
Nombre de lignes	6420
Nombre de colonnes	3
Type de données	Textuelles
Noms des colonnes	(Id, Tweet, Label)
Utilisation de la mémoire	1.19 MB
Nombre de valeurs nulles	0

3.3.2 Distribution de Dataset

l'ensemble de données "Constraint-train" contient 6420 articles publiés sur Twitter, 3360 sont étiquetées "True", et 3060 sont étiquetées "Fake".

Figure 3.2 – Distribution de données selon attribut « Label ».

3.4 Implementation

3.4.1 Les Etapes de pretraitement de donnees

Dans le domaine de la détection des fausses informations liées au COVID-19 sur les tweets, le prétraitement des données joue un rôle essentiel. Il s'agit d'une série d'étapes nécessaires pour préparer les données textuelles avant d'appliquer des techniques d'apprentissage automatique. Dans cette section, nous présenterons en détail les techniques de prétraitement utilisées sur les tweets afin de mieux comprendre et identifier automatiquement les cas de fausses informations.

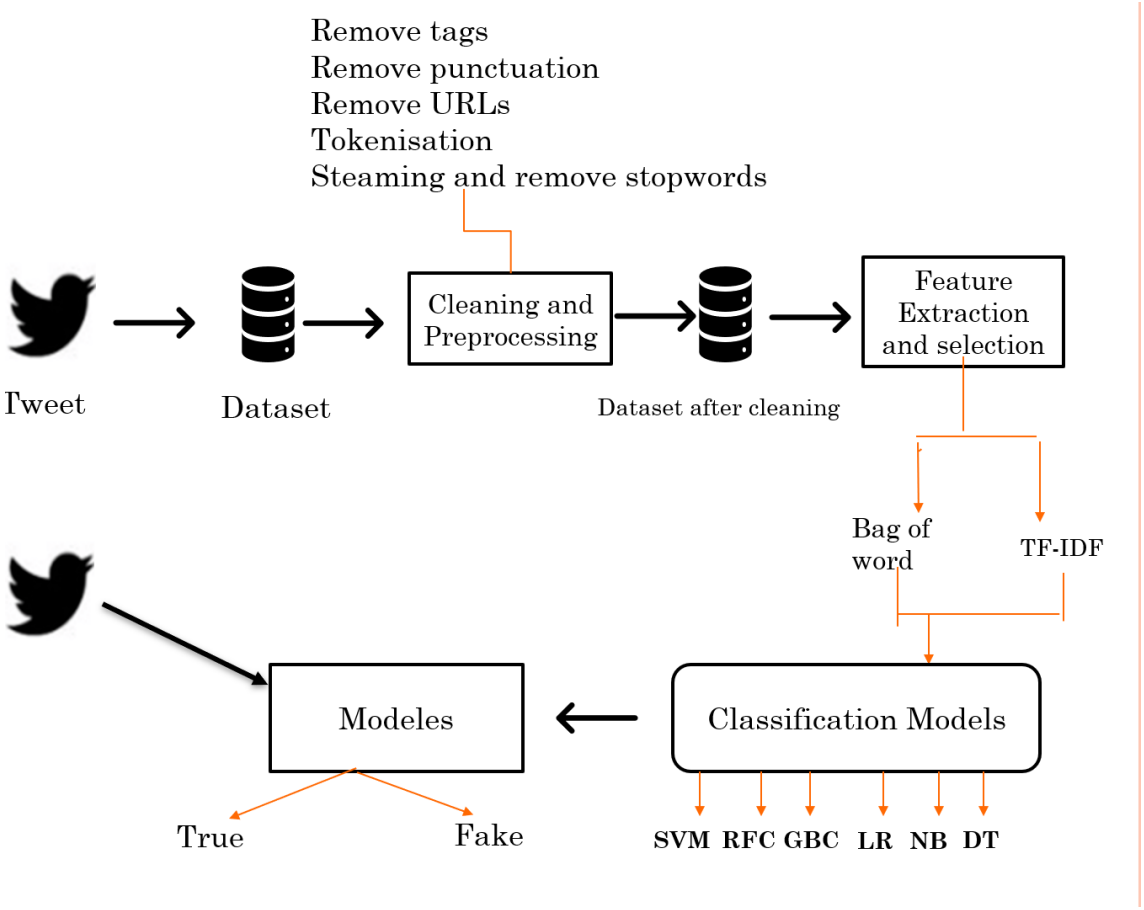

Figure 3.3 – Architecture generale de notre proposition

La première étape du prétraitement des données a consisté à collecter les tweets. Après avoir constitué notre ensemble de données, nous avons appliqué diverses méthodes de prétraitement du texte afin de préparer les tweets à la classification et d'éliminer le bruit qu'ils contiennent. Ceci est cohérent avec les recherches passées, qui suggèrent que l'application de techniques de prétraitement de texte peut améliorer la qualité des résultats de classification obtenus à partir d'algorithmes d'apprentissage automatique.

• **Nettoyer le tweet**

Cette étape comprend la suppression des URL (qui commencent par http ://" ou "https ://"), les mentions d'utilisateurs (par exemple, "@utilisateur") et les caractères spéciaux tels que les signes de ponctuation, les hashtags et les émoticônes.

• **Repartition des hashtags :**

Les hashtags sont utilisés pour indiquer des sujets ou des thèmes spécifiques dans les tweets. Dans notre prétraitement, nous avons décomposé les hashtags en mots individuels afin d'améliorer la précision de l'analyse des sentiments. Par exemple, le hashtag HealthAndSafety a été décomposé en trois tokens distincts : "Health" et "and" et" Safety". Dans l'ensemble, les différentes techniques de prétraitement appliquées dans notre étude visaient à réduire la dimension des données et le bruit tout en améliorant la qualité des résultats de la classification.

• **Capturez les mots allongés :**

Les utilisateurs de Twitter utilisent fréquemment des formes allongées de mots pour mettre l'accent ou intensifier leur expression. Par exemple, le mot "coviiiiiiiiiiiiid" est une version allongée de "covid" pour amplifier le sentiment. Cette pratique linguistique permet aux utilisateurs d'éviter les mots répétitifs (par exemple, "covid", "covidd", "coviidddd", etc.) tout en transmettant le même sens de grande admiration ou d'étonnement.

• **Éliminez la ponctuation et mettez le texte en minuscules :**

Le maintien d'une casse cohérente pour les mots est crucial lors de la classification des textes, car il garantit que tous les termes sont correctement appariés, qu'ils soient en majuscules ou en minuscules. Cette cohérence est particulièrement importante pour notre étude, car les microblogs contiennent souvent des modèles de capitalisation irréguliers (par exemple, " VacCiNeS WoRk"). Pour ce faire, nous convertissons tous les mots du texte en minuscules et supprimons tout signe de ponctuation ou caractère spécial. Ce processus nous permet de normaliser le texte et d'éliminer les divergences potentielles causées par des majuscules différentes ou des symboles superflus.

• **Supprimer les mots vides (stopwords :) :**

Les mots vides sont des mots qui ont peu de sens et qui contribuent moins à l'analyse d'un tweet. Il s'agit généralement d'articles, de conjonctions et de prépositions qui ne fournissent pas d'informations contextuelles ou significatives. La suppression des mots vides n'a pas d'impact négatif sur notre modèle final. Parmi les exemples de mots vides couramment utilisés en anglais, citons "is", "and" et "are". Dans notre approche, nous avons utilisé le corpus de mots vides de la bibliothèque Spacy pour éliminer les mots

vides. Ce corpus se compose de 305 mots vides spécifiquement adaptés à la langue anglaise.

• **Tokenization du texte :**

Pour traiter efficacement des données textuelles, il est nécessaire d'identifier les frontières des mots. La tokenisation est le processus de décomposition du texte en jetons, qui peuvent être des mots, des phrases ou des symboles individuels. Ces jetons jouent un rôle déterminant dans l'analyse et d'autres tâches d'exploration. La tokenisation est généralement effectuée à l'aide de bibliothèques telles que Spacy.

• **Stemming :**

Le stemming consiste à générer différentes formes d'un mot de base ou racine. Son but est de réduire les mots à leur forme fondamentale. La recherche de radicaux est utilisée pour simplifier la recherche de mots et normaliser les phrases pour une meilleure compréhension.

Steaming Example :

— **Original Sentence :** I was running in the park and saw many runners.

— **Stemmed Sentence :** I was run in the park and see mani runner.

Dans cet exemple, la phrase d'origine est segmentée en mots individuels. Le processus de radicalisation convertit ensuite chaque mot en sa forme de base ou racine. Par exemple, le mot "running" est stemmed de "run" et le mot " runners " est stemmed de " runner ". Le but du stemming est de réduire les mots à leurs formes essentielles, en simplifiant l'analyse de texte et en améliorant l'efficacité des calculs.

3.5 Extraction des caractéristiques

L'extraction de caractéristiques est est le processus de transformation des données brutes en caractéristiques ou attributs significatifs qui peuvent être utilisés comme entrée pour un modèle d'apprentissage automatique. Cette étape est essentielle pour la plupart des tâches d'analyse de données, car les modèles d'apprentissage automatique ont besoin de données structurées et informatives pour obtenir de bonnes performances.

3.5.1 Representation of the tweet :

Nous extrayons les caractéristiques Unigram, Bigram et Trigram des tweets prétraités et les pondérer en fonction de leurs valeurs TF-IDF et BOW. Le but de l'utilisation de TF-IDF et BOW est de réduire l'effet des termes moins informatifs qui apparaissent très fréquemment dans les données corpus. Cette fonctionnalité peut être une fonctionnalité standard.

3.5.2 Bag of word

La méthode la plus simple pour convertir du texte en caractéristiques structurées consiste à utiliser l'approche du sac de mots (bag of words). Le sac de mots consiste simplement à diviser les mots du texte de l'examen en statistiques de décompte de mots individuels. La Figure 1.2 présente comment transformer du texte brut en une représentation sac de mots.

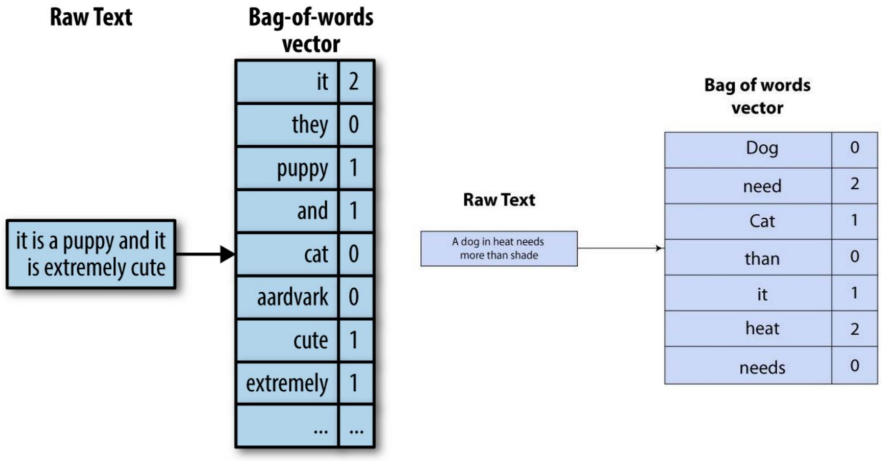

FIGURE 3.4 – Transformation d'un texte brut en représentation d'un sac de mots

3.5.3 TF-IDF (Term Frequency-Inverse Document Frequency) :

est une mesure statistique utilisée pour évaluer l'importance d'un terme dans un document ou une collection de documents. Il est couramment utilisé dans des tâches de recherche d'information et d'analyse de texte, telles que le classement de documents, la classification de texte et l'analyse de contenu.
Le score TF-IDF d'un terme est calculé en combinant deux facteurs :

— **Fréquence du terme (TF) :** Il mesure la fréquence à laquelle un terme apparaît dans un document. Il est calculé en divisant le nombre de fois où un terme apparaît dans un document par le nombre total de termes dans le document. Une valeur de TF élevée indique qu'un terme est plus important dans le document.

— **Fréquence inverse du document (IDF) :** Elle mesure la rareté ou l'unicité d'un terme dans une collection de documents. Elle est calculée comme le logarithme du nombre total de documents divisé par le nombre de documents qui contiennent le terme. Une valeur IDF élevée indique qu'un terme est plus unique ou informatif dans l'ensemble de la collection.

Le score TF-IDF est obtenu en multipliant les valeurs de TF et IDF pour un terme dans un document. Ce score permet de mesurer l'importance d'un terme dans un document par rapport à l'ensemble de la collection.
La technique TF-IDF permet d'identifier les termes importants dans un document et de les distinguer des termes courants qui apparaissent fréquemment dans plusieurs documents. Elle permet la représentation des documents sous forme de vecteurs numériques, ce qui facilite diverses tâches d'analyse de texte telles que le calcul de similarité entre documents et l'extraction de mots-clés.
TF - IDF(d,t) = tf(t) * idf(d,t)

$$TF(t) = \frac{\text{Number of times 't' appears in 'd'}}{\text{Total number of terms in 'd'}}$$

$$idf(t) = \frac{\text{Total number of tweets}}{\text{Number of tweets containing the term 't'}}$$

3.5.4 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 3.6 montre les mots les plus fréquents dans notre corpus. Les figures 3.8, 3.9, 3.10 et 3.11 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. La fréquence des structures « wuhan coronavirus » et « wuhun institute virology » peut être interprétée par les doutes qui ont accompagné cet institut de virologie chinois accusé lui aussi par les désinformateurs d'avoir inventé le Coronavirus dans un cadre d'un armement biologique.

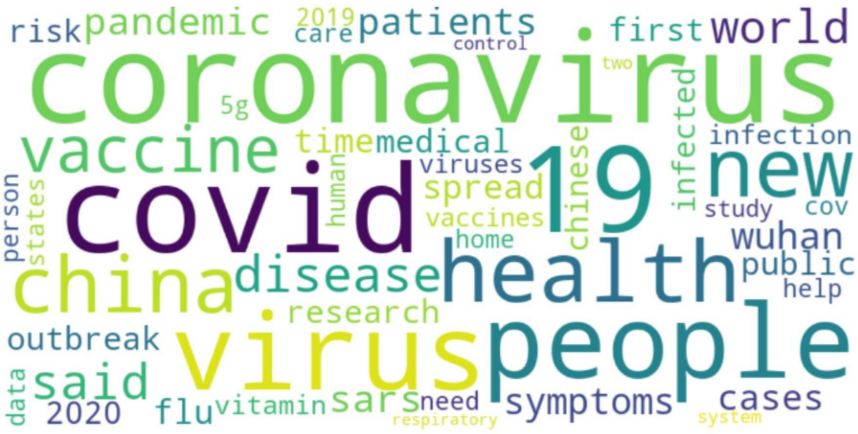

FIGURE 3.5 – Les mots les plus fréquents dans l'ensemble de données

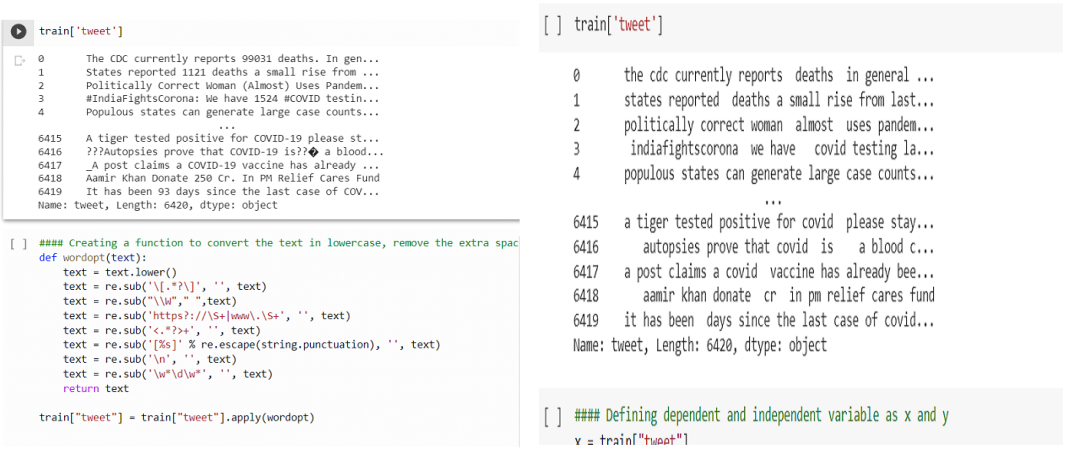

FIGURE 3.6 – prétraitement de texte pour la normalisation des données.

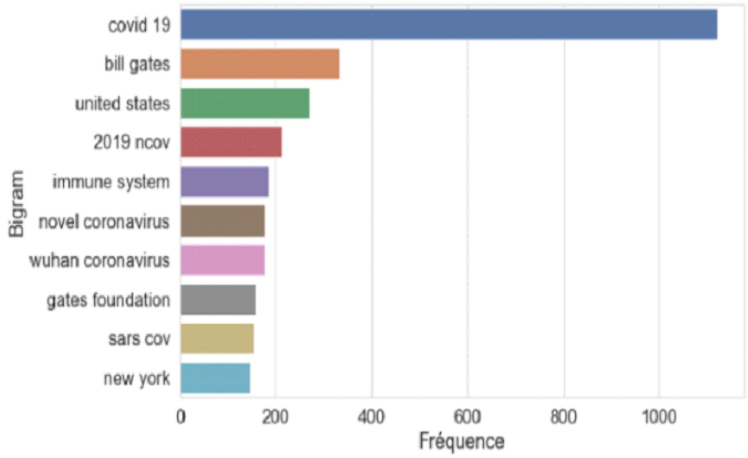

FIGURE 3.7 – Top 10 "True" bigrams

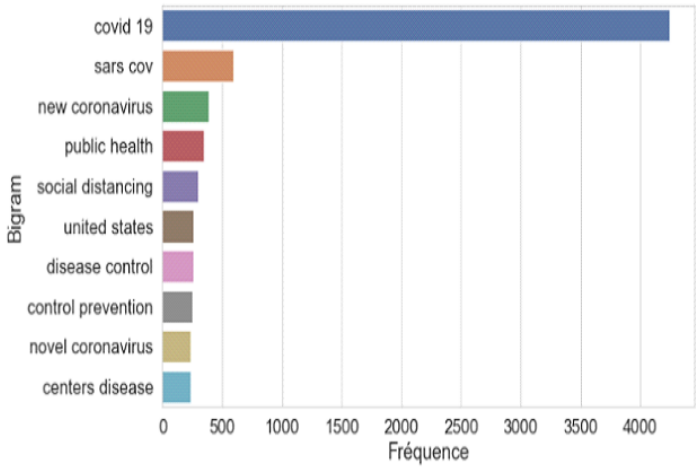

FIGURE 3.8 – Top 10 "Fake" bigrams

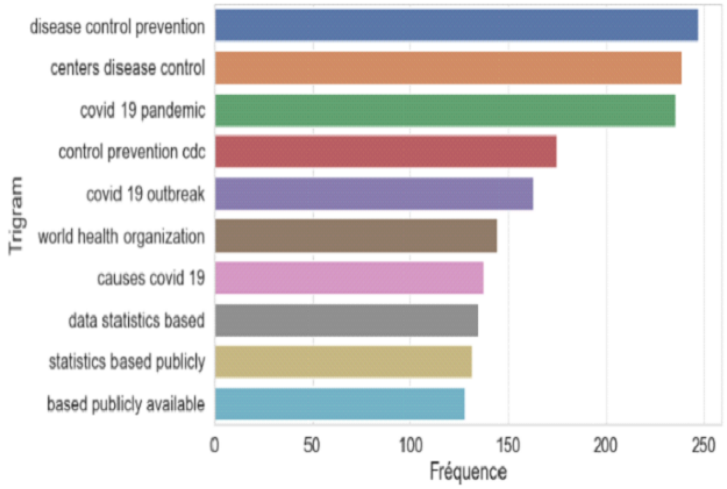

FIGURE 3.9 – Top 10 "True" trigrams

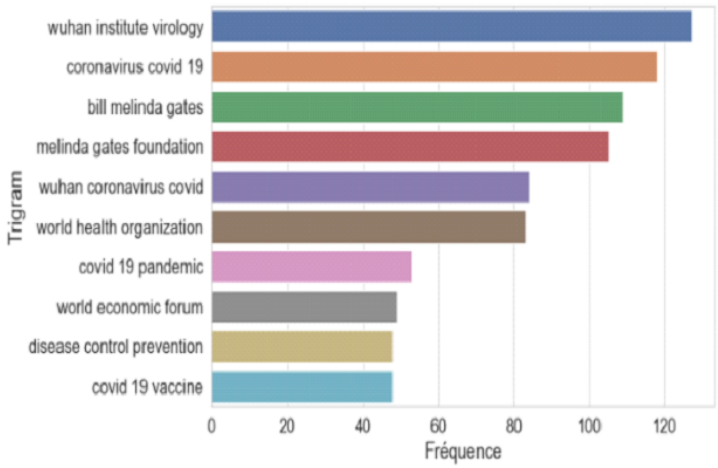

FIGURE 3.10 – Top 10 "Fake" trigrams

3.6 CLASSIFICATION

La classification supervisée est la prédiction de la valeur d'une variable de réponse qualitative. Ce variable de réponse est communément appelée catégorie ou classe. De nombreux algorithmes de classification (également appelés classificateurs) et leurs performances dépendent du problème et de sa nature. Depuis là n'y a pas de classificateur qui soit le meilleur pour toutes sortes de problèmes (pas de théorème de repas gratuit), pour notre travaux, un ensemble de modèles a été évalué :

— **Support Vector Machine SVM :** sont des classificateurs dont le résultat est basé sur une frontière de décision générés par des vecteurs de support, les points les plus proches de la frontière de décision. La forme de la frontière est déterminée par une fonction du noyau. De cette façon, il est possible de résoudre problèmes qui ne peuvent pas être résolus par une frontière linéaire. Intuitivement, une bonne séparation est obtenu par l'hyperplan le plus éloigné des vecteurs supports. En général, plus la marge est grande, plus l'erreur de généralisation du classifieur est faible[2] .

— **Régression logistique :** L'algorithme de classification par régression logistique utilise une fonction sigmoïde, exprimée par : P(x)= $\frac{1}{1+e-x}$ comme une fonction de prédiction qui renvoie une valeur de probabilité qui peuvent ensuite être mappés à des classes discrètes. Les résultats de la fonction sigmoïde sont la probabilité estimations qui vont de 0 à 1. Pour prédire l'étiquette d'un point de données, la classe avec le le score ou la probabilité le plus élevé est choisi .

— **Arbre de décision (DT) :** L'arbre de décision est l'une des techniques de classification dans sification se fait par les critères de découpage. L'arbre de décision est un organigramme comme un arbre

structure qui classe les instances en les triant en fonction des valeurs d'attribut (fonctionnalité).
Chaque nœud d'un arbre de décision représente un attribut dans une instance à classer. fié. Toutes les branches représentent un résultat du test, chaque nœud feuille contient l'étiquette de classe.
L'instance est classée en fonction de sa valeur de fonctionnalité. Il existe de nombreuses méthodes pour trouver la fonctionnalité qui divise le mieux les données d'entraînement telles que le gain d'informations, le rapport de gain,Indice de Gini, etc. Le moyen le plus courant de créer des arbres de décision en utilisant des partitionnement de méthode, en commençant par l'ensemble d'apprentissage et en trouvant de manière récursive une fonction de fractionnement qui maximise un critère local. L'arbre de décision génère la règle pour la classification de l'ensemble de données (DATA SET).

— **Naive Bayes NB :** est un classificateur probabiliste simple basé sur l'application du théorème de Baye avec de fortes hypothèses d'indépendance. Pour la classification de texte, cet algorithme calcule la probabilité a posteriori du document appartient à différentes classes, et elle attribue la document à la classe avec la probabilité a posteriori la plus élevée. Ce modèle de probabilité serait un modèle d'entité indépendant afin que la présence d'une entité n'affecte pas autres entités dans les tâches de classification .

— **Random Forest Classification :** Modèle de classificateur aléatoire : est un modèle de classification qui assigne des étiquettes de classe de manière aléatoire, sans prendre en compte les caractéristiques des données d'entrée. En d'autres termes, le modèle attribue une classe à chaque instance d'entrée de manière totalement arbitraire, sans aucune relation avec les caractéristiques de cette instance. Le modèle de classificateur aléatoire est rapide à exécuter, car il ne nécessite aucun calcul ou traitement complexe. Cela en fait un choix idéal pour les ensembles de données de grande taille qui peuvent prendre beaucoup de temps à traiter avec des modèles de classification plus avancés. utilisé comme une comparaison de base pour d'autres modèles de classification aléatoires, tels que les forêts aléatoires et les réseaux de neurones aléatoires. Cela permet de déterminer si ces modèles plus avancés ont réellement un impact sur les performances de classification par rapport au modèle de classificateur aléatoire

— **Gradient Boosting Classifier :** Le gradient boosting classifier (GBC) est un algorithme d'apprentissage automatique supervisé qui utilise une méthode d'ensemble pour construire un modèle prédictif à partir de plusieurs modèles plus simples. Plus précisément, le GBC construit des arbres de décision à partir de plusieurs itérations en utilisant un gradient de perte pour minimiser l'erreur de prévision.
  — Le GBC est capable de gérer des données hétérogènes, notamment des variables catégoriques et numériques.
  — Le GBC est un algorithme très précis qui peut atteindre des performances proches de l'état de l'art pour la plupart des tâches de classification et de régression.

— Le GBC peut gérer efficacement de grandes quantités de données, grâce à sa capacité à construire des arbres de décision en parallèle.
— Possibilité d'ajustement des paramètres : Le GBC offre de nombreuses options de paramétrage, permettant ainsi d'ajuster les paramètres pour optimiser les performances du modèle.

3.7 L'évaluation

En règle générale, l'évaluation de la performance des modèles de classification se fait en fonction de leur capacité à prédire les résultats pour de nouveaux ensembles de données. Cette évaluation est réalisée à l'aide d'un ensemble de test, et plusieurs métriques sont utilisées pour déterminer la performance de prédiction d'un modèle. Dans ce contexte, nous allons nous concentrer principalement sur les métriques 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 :
$$\text{Accuracy} = \frac{TP + TN}{TP + TN + FP + FN}$$
— **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 :
$$\textbf{Precision} = \frac{TP}{TP + FP}$$
— **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 :
$$\textbf{Recall} = \frac{TP}{TP + FN}$$
— **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 :
$$\textbf{F1 Score} = \frac{2 * Precision * Recall}{Precision + Recall}$$

3.8 Résultats de classification

Pour classifier les textes de l'ensemble de données «**Constraint-train**», nous avons appliqué 06 modèles de classification. modèles de base : Naïve de Bayes, Régression logistique, Forêts aléatoires, et Machines à vecteurs de supports SVM. GBC ,DT Les résultats de classification sont évalués à l'aide des métriques : Précision (Precision), Rappel (Recall), F1 Score, et Taux de succès (Accuracy).

3.9 Évaluation des modèles de classification :

Nous avons pris en compte six modèles différents : la machine à vecteurs de support (SVM), la régression logistique (LR), l'arbre de décision (DT), le classifieur de forêt aléatoire (RFC), le classifieur de renforcement par gradient (GBC) et le classifieur naïf de Bayes (NB). Chaque modèle a été entraîné et testé sur des données appropriées afin de déterminer son exactitude et precision dans la détection de fausses informations.

3.9.1 Comparaison en term de Accuracy

Voici les résultats de cette comparaison :
— Le modèle SVM a une accuracy de 93%.
— Le modèle LR a une accuracy) de 92%.
— Le modèle DT a une accuracy de 84%.
— Le modèle RFC a une accuracy de 90%.
— Le modèle GBC a une accuracy de 87%.
— Le modèle NB a une accuracy de 91%.
Pour représenter ces résultats visuellement, nous pouvons utiliser un graphique à barres :

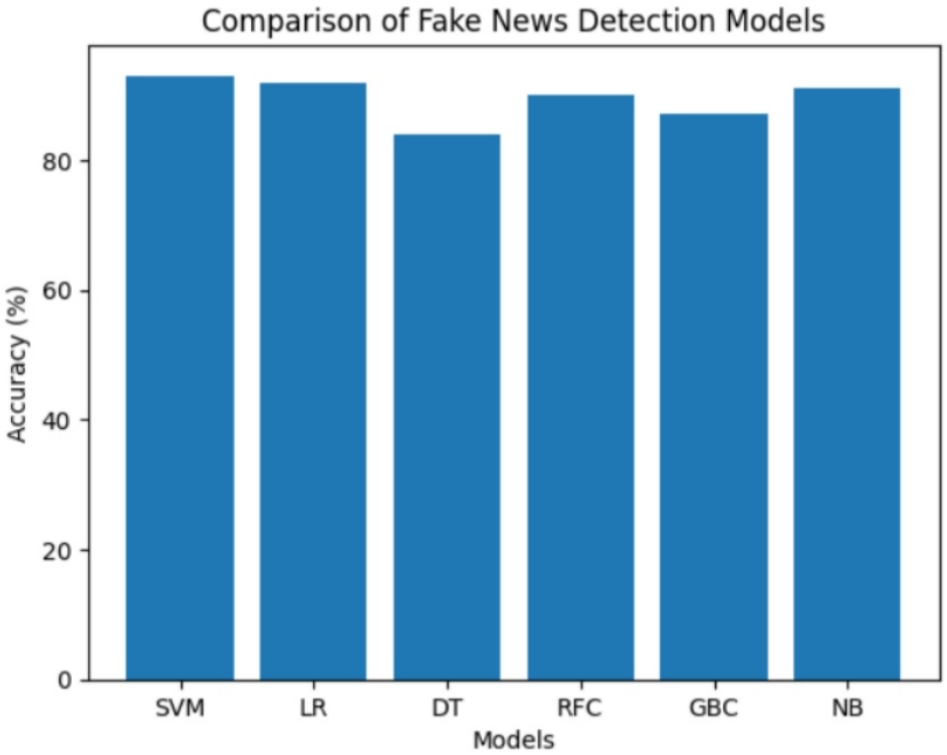

FIGURE 3.11 – comparaison entre les modèles de détection des fakes news

Dans ce graphique, chaque barre représente un modèle différent, et sa hauteur correspond à l'exactitude de ce modèle. Cela permet de visualiser facilement les performances relatives de chaque modèle dans la détection de fausses informations.
En utilisant l'exactitude comme métrique de performance, nous pouvons conclure que le modèle SVM présente la meilleure exactitude avec 93%, suivi de près par le modèle LR avec 92%

3.9.2 Comparaison en term de precision

Cette comparaison vise à évaluer les performances de différents modèles en termes de précision, à la fois pour les faux positifs (fausses informations classées comme vraies) et pour les vrais positifs (vraies informations classées comme vraies).

Voici les résultats de cette comparaison :
Modèle SVM
— Précision (fausses informations) : 0,93

— Précision (vraies informations) : 0,93

Modèle LR :

— Précision (fausses informations) : 0,90
Précision (vraies informations) : 0,94

Modèle DT :

— Précision (fausses informations) : 0,84
— Précision (vraies informations) : 0,83

Modèle GBC :

— Précision (fausses informations) : 0,83
— Précision (vraies informations) : 0,91

Modèle RFC :

— Précision (fausses informations) : 0,89
— Précision (vraies informations) : 0,92

Modèle NB :

— Précision (fausses informations) : 0,94
— Précision (vraies informations) : 0,89

Pour représenter visuellement ces résultats, nous pouvons utiliser un graphique à barres :

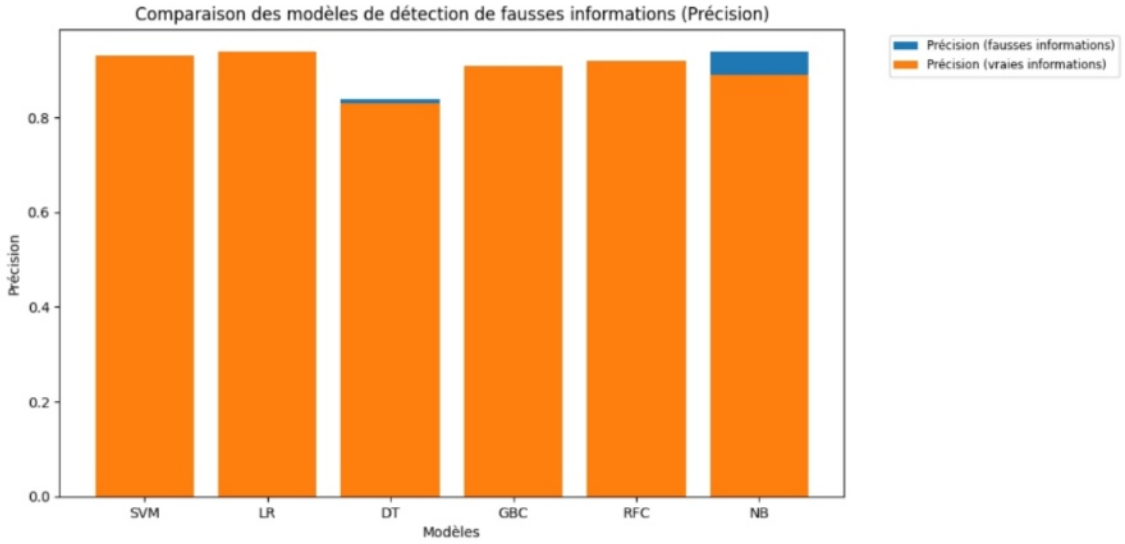

FIGURE 3.12 – comparaison des modèles de détection de fausses informations(précision

Dans ce graphique, chaque modèle est représenté par deux barres : une pour la précision des fausses informations et une pour la précision des vraies informations. Cela permet de visualiser facilement les performances relatives de chaque modèle dans la détection des fausses informations selon ces deux aspects.

En utilisant la précision comme métrique de performance, nous pouvons voir que le modèle NB présente la meilleure précision dans la détection des fausses informations (0,94), tandis que le modèle LR présente la meilleure précision dans la détection des vraies informations (0,94).

En analysant ces résultats, nous pouvons conclure que :

— Le modèle SVM présente la meilleure exactitude (93%) et une précision élevée pour les fausses informations et les vraies informations (93% pour les deux).
— Le modèle LR a une exactitude légèrement inférieure (92%), mais une précision plus élevée pour les vraies informations (94%).

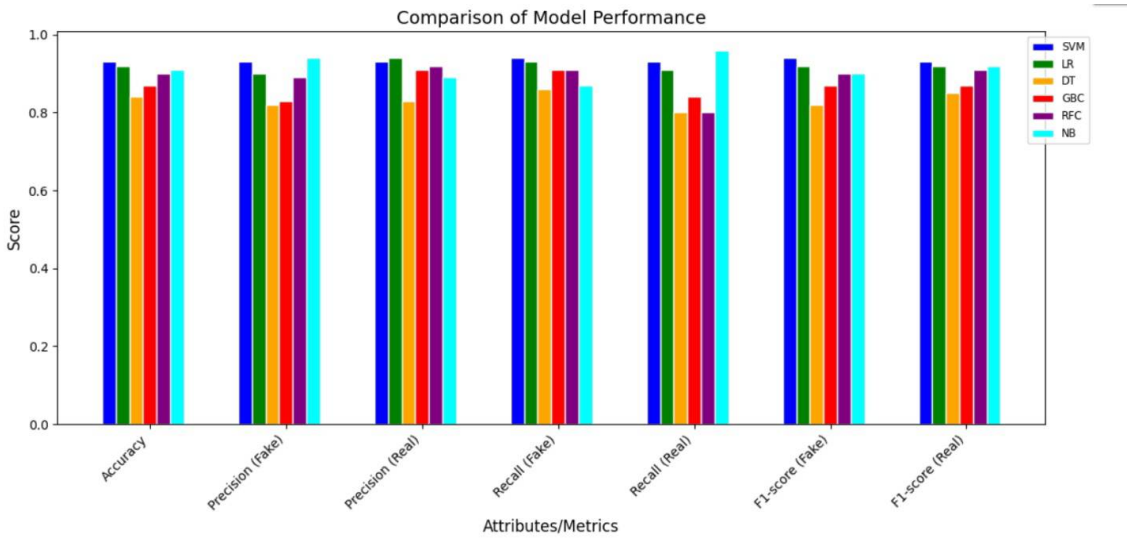

FIGURE 3.13 – comparison of models classification

— Le modèle DT a la plus faible exactitude (84%) et la précision la plus basse pour les fausses informations et les vraies informations (82% et 83% respectivement).
— Le modèle GBC et le modèle RFC ont des performances similaires en termes d'exactitude (87% et 90% respectivement) et de précision.
— Le modèle NB présente une exactitude élevée (91%) et la meilleure précision pour les fausses informations (94%).

Il est important de prendre en compte d'autres métriques et facteurs lors de la sélection du modèle le plus approprié pour un problème spécifique de détection de fausses informations.

3.10 La classification par modèles de base

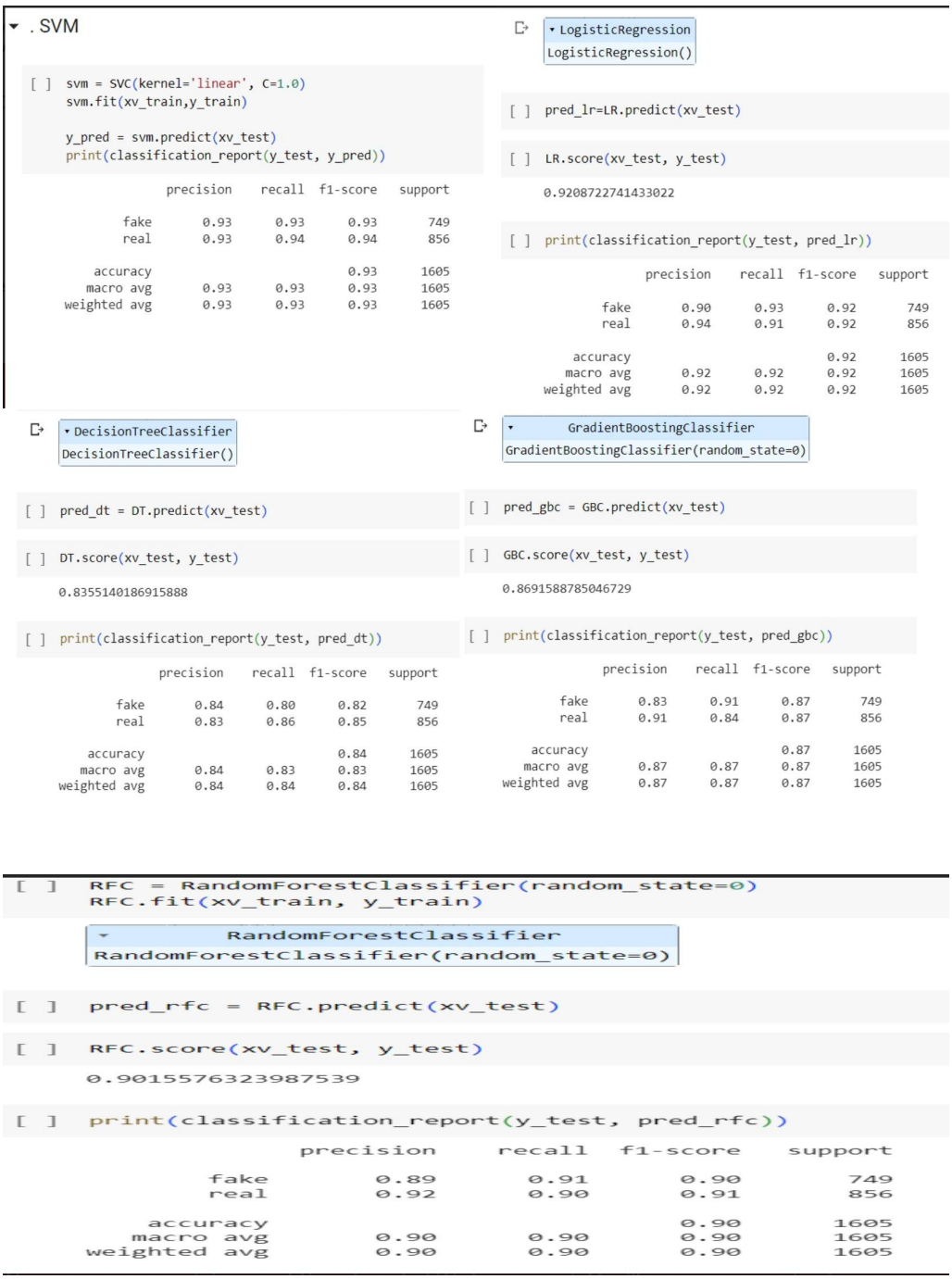

Figure 3.14 – Résultats de classification par modèles de base

3.11 Utilisation :

3.11.1 Implementation(interface)

Le programme Python a été intégré au framework Flask, ainsi qu'aux langages HTML, CSS pour créer une version web de l'application. Avec le framework Web Flask, nous fournissons une interface conviviale où les utilis ateurs peuvent facilement saisir des articles d'actualité à des fins d'analyse. Notre programme traite ensuite l'entrée à l'aide du modèle de détection de fausses nouvelles et génère des résultats en temps réel, indiquant l'authenticité de l'article soumis. En intégrant de manière transparente notre programme Python à Flask, nous visons à donner aux utilisateurs les moyens de prendre des décisions éclairées et de contribuer à un écosystème d'informations plus fiable et digne de confiance. La figure 3.2 contient le code de notre application flask de basse sert d'interface Web .

C: > Users > aggou > Downloads > app.py

```python
from flask import Flask, render_template, request
import detection

app = Flask(__name__, static_folder='static')


@app.route('/')
def home():
    return render_template('index.html', result=None)

@app.route('/submit', methods=['POST'])
def submit():
    user_input = request.form['input_field']
    result = detection.manual_testing(user_input)
    return render_template('index.html', result=result)

if __name__ == '__main__':
    app.run(debug=True)

```

Figure 3.15 – flask application

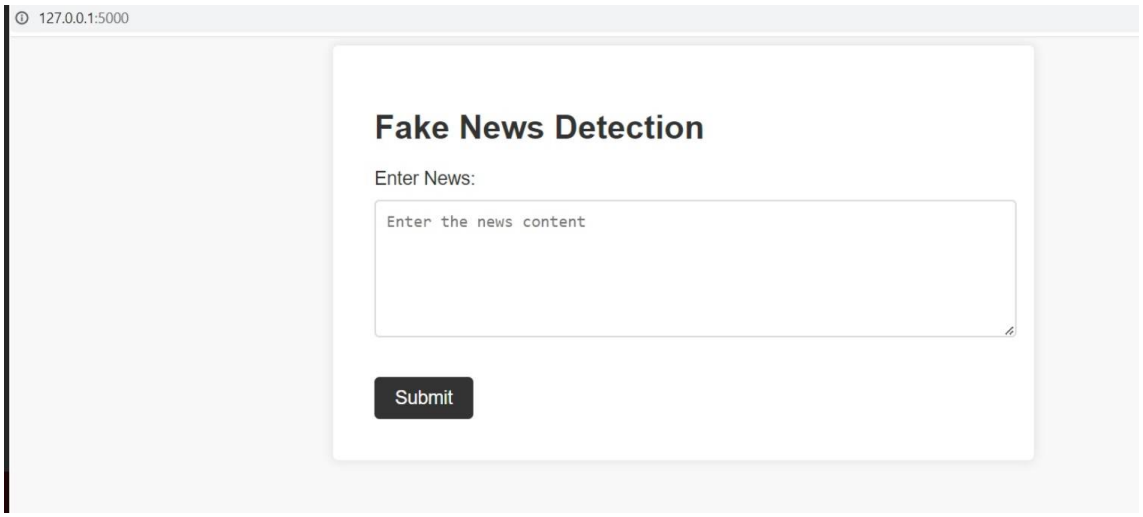

Figure 3.16 – L'interface

3.11.2 Le Test

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 3.16.

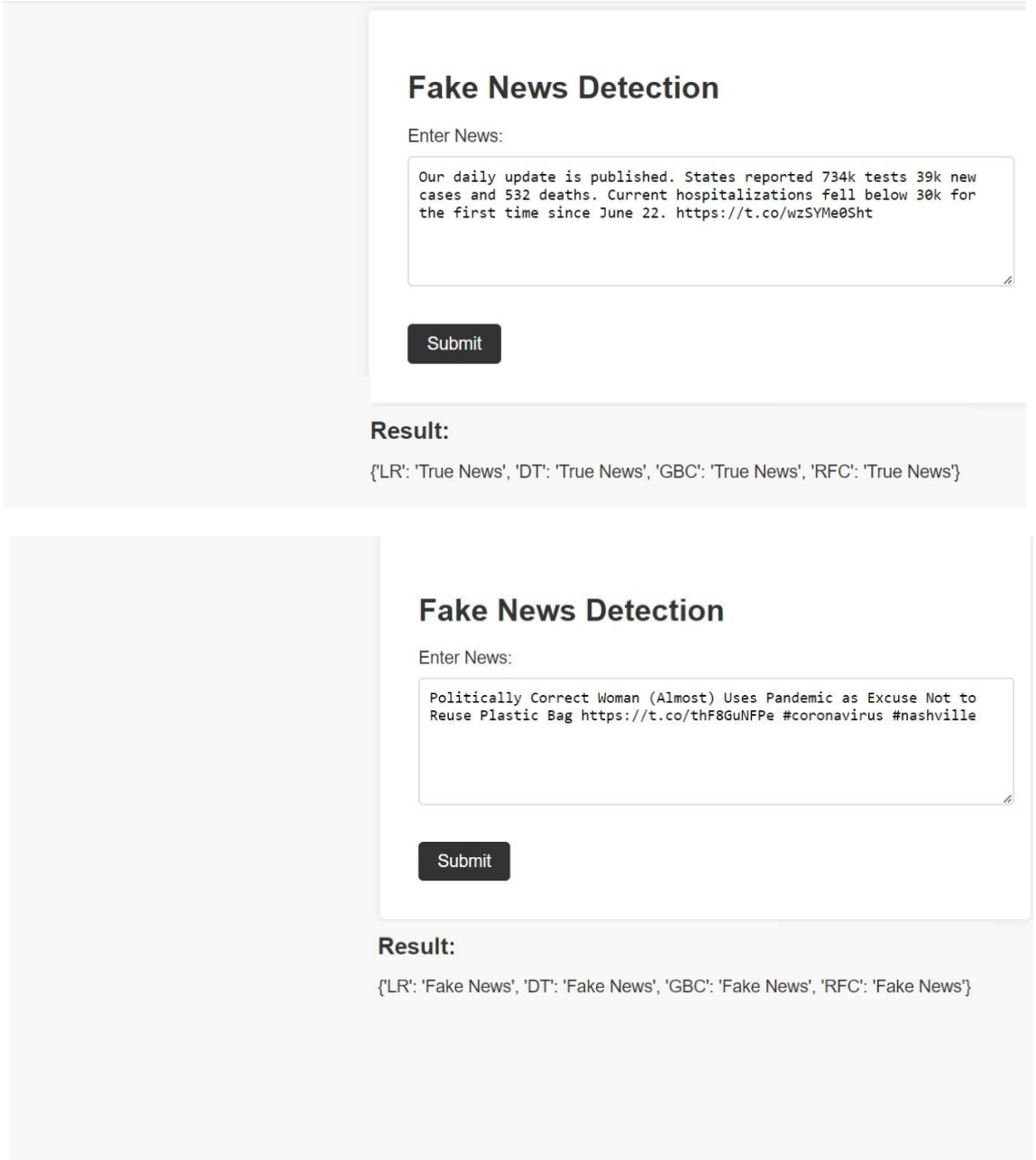

FIGURE 3.17 – Exemple de détection d'une fausse information

3.12 Conclusion :

Alors que la détection du contenu lié au coronavirus et de la désinformation dans les médias sociaux continue d'être une préoccupation urgente, le besoin de systèmes de détection automatique du coronavirus devient de plus en plus évident. Tout au long de ce chapitre, nous présentons nos différentes expérimentations menées dans le cadre de l'analyse des tweets pour la détection du cdu fake news liee au coronas virus à l'aide d'algorithmes d'apprentissage automatique supervisé.

Dans nos expériences, nous explorons l'efficacité de divers algorithmes d'apprentissage automatique supervisé pour détecter les fakes news lies au coronavirus. Ces algorithmes sont formés à l'aide d'ensembles de données étiquetés qui consistent en des publications sur les réseaux sociaux, des articles de presse ou d'autres sources d'informations liées au coronavirus.
Grâce à nos expériences et évaluations, nous évaluons les performances et l'efficacité de différents algorithmes d'apprentissage automatique supervisé pour identifier les sentiments liés au coronavirus. Nous prenons en compte des mesures telles que l'exactitude, la précision, le rappel et le score F1 pour évaluer les performances des modèles dans la classification du contenu lié au coronavirus comme positif, négatif ou neutre.

Conclusion générale

Notre système de détection de fausses informations liées au coronavirus présente des résultats prometteurs pour atténuer la propagation de la désinformation. Nous avons évalué les performances des différents modèles en utilisant plusieurs métriques telles que l'exactitude (Accuracy), la précision (Precision), le rappel (Recall) et le score F1 (F1-score).

Parmi les modèles, SVM a démontré la plus haute précision de 93%, suivi de près par LR avec une précision de 92%. GBC a obtenu la plus haute précision pour les fausses informations, atteignant 83%, et la plus haute précision pour les informations réelles, atteignant 91%. NB a affiché la plus haute précision pour les fausses informations, avec un taux de 94%, tandis que LR a atteint la plus haute précision pour les informations réelles, à 94%.

En ce qui concerne le rappel, NB a obtenu le plus haut rappel pour les fausses informations, à 96 %, et SVM a eu le plus haut rappel pour les informations réelles, à 94 %. GBC a affiché des performances équilibrées avec des valeurs de rappel élevées pour les fausses informations (91 %) et les informations réelles (84 %).

En ce qui concerne les scores F1, NB a atteint le plus haut score F1 pour les fausses informations, à 90 %, tandis que SVM a obtenu le plus haut score F1 pour les informations réelles, à 94 %. GBC a démontré des scores F1 compétitifs pour les fausses informations et les informations réelles, à 87 %.

Sur la base de ces résultats, on peut conclure que chaque modèle présente des points forts dans différentes mesures de performance. SVM et LR ont montré une précision élevée, tandis que GBC a démontré des taux de rappel équilibrés. NB a affiché des valeurs de précision solides, ce qui indique sa capacité à identifier avec précision les fausses informations.

Notre système de détection de fausses informations liées au coronavirus offre une contribution précieuse pour lutter contre la propagation de la désinformation pendant la pandémie. Les plateformes de médias sociaux, les agences de presse et autres parties prenantes peuvent utiliser notre système comme un outil automatisé pour détecter et signaler les informations potentiellement trompeuses, ce qui contribue à promouvoir des informations précises et à protéger la santé publique et le bien-être des individus.

Naturellement de nombre perspectives envisageables pour améliorer ce travail. Tout d'abord, il se base principalement sur des tweets comme source d'informations, ce qui limite la diversité des sources d'informations considérées. De plus, la généralisation de notre système à d'autres scénarios et à l'évolution des tendances des fausses informations nécessite une adaptation continue et une surveillance régulière.

Bibliographie

[1] Guide de l'intelligence artificielle : Apprentissage supervisé.

[2] Classification model - svm classifier python example. https://vitalflux.com/classification-model-svm-classifier-python-example/?fbclid=IwAR21au1hvb_gl9vZ9s4ngoeBJKRt7mNC0AKd2lPFtTMryRlAP9jemmEdnag, consulté le 14 juin 2023.

[3] Numpy. https://fr.wikipedia.org/wiki/NumPy, consulté le 14 juin 2023. Page Wikipedia sur NumPy.

[4] Pandas. https://fr.wikipedia.org/wiki/Pandas#cite_note-2, consulté le 14 juin 2023. Page Wikipedia sur Pandas.

[5] Hadeer Ahmed, Issa Traore, and Sherif Saad. Detection of online fake news using n-gram analysis and machine learning techniques. In *International Conference on Intelligent, Secure, and Dependable Systems in Distributed and Cloud Environments*, pages 127–138. Springer, 2017. Pages 12, 33.

[6] Sajjad Ahmed, Knut Hinkelmann, and Flavio Corradini. Combining machine learning with knowledge engineering to detect fake news in social networks - a survey. In *Proceedings of the AAAI 2019 Spring Symposium*, volume 12, page 8, 2019.

[7] AI. chatgbt.

[8] Evangelos Christodoulou, Jiefei Ma, Gary S. Collins, Ewout W. Steyerberg, Jan Y. Verbakel, and Ben Van Calster. A systematic review shows no performance benefit of machine learning over logistic regression for clinical prediction models. *Journal of Clinical Epidemiology*, 110 :12–22, 2019.

[9] Niall J Conroy, Victoria L Rubin, and Yimin Chen. Automatic deception detection : Methods for finding fake news. *Proceedings of the Association for Information Science and Technology*, 52(1) :1–4, 2015. Page 12.

[10] DataScientest. Machine Learning . https ://datascientest.com/machine-learning-tout-savoir.

[11] Maksym Granik and Volodymyr Mesyura. Fake news detection using naive bayes classifier. In *2017 IEEE First Ukraine Conference on Electrical and Computer Engineering (UKRCON)*, pages 687–692. IEEE, 2017.

[12] Mykhailo Granik and Volodymyr Mesyura. Fake news detection using naive bayes classifier. In *2017 IEEE First Ukraine Conference on Electrical and Computer Engineering (UKRCON)*, pages 900–903. IEEE, 2017. Page 12.

[13] Journal du Net. Guide de l'intelligence artificielle : Apprentissage non supervisé.

[14] Jean-Noël Rumeurs Kapferer. Le plus vieux media du monde. *Paris : Éditions du Seuil, coll."Points Actuel*, 1990.

[15] Bastien L. Machine learning définition fonctionnement et secteurs d'application. http://www.artificiel.net/machine-learning-definition, 2017. Pages 15, 16, 25, 26.

[16] Cédric Maigrot, Ewa Kijak, and Vincent Claveau. Fusion par apprentissage pour la détection de fausses informations dans les réseaux sociaux. *Document numérique*, 21(3) :55–80, 2018. Page 12.

[17] Adeline Michel, AC Sordet, and E Moraillon. Les rumeurs en tant que phénomène d'influence sociale. *Ecole de psychologues praticiens de Lyon*, 2004.

[18] S. Moro, P. Cortez, and P. Rita. A data-driven approach to predict the success of bank telemarketing. *Decision Support Systems*, 62 :22–31, June 2014.

[19] Florian Sauvageau. *Les fausses nouvelles, nouveaux visages, nouveaux défis. Comment déterminer la valeur de l'information dans les sociétés démocratiques ?* Presses de l'Université Laval, 2018. Page 12.

[20] Futura Sciences. Fake news - définition et explications. https://www.futura-sciences.com/tech/definitions/informatique-fake-news-17092/.

[21] Futura Sciences. Réseau social - définition et explications. https://www.futura-sciences.com/tech/definitions/informatique-reseau-social-10255/.

[22] Scikit-learn. RandomForestClassifier - Scikit-learn Documentation.

[23] Stack Abuse. Gradient Boosting Classifiers in Python with Scikit-learn.

[24] Sandro Tosi. *Matplotlib for Python Developers : Build Remarkable Publication Quality Plots the Easy Way*. Packt Publishing, 2009.

[25] Nisrine Zammar. *Réseaux Sociaux numériques : essai de catégorisation et cartographie des controverses*. PhD thesis, Université Rennes 2, 2012.

[26] Savvas Zannettou, Michael Sirivianos, Jeremy Blackburn, and Nicolas Kourtellis. The web of false information : Rumors, fake news, hoaxes, clickbait, and various other shenanigans. *Journal of Data and Information Quality (JDIQ)*, 11(3) :1–37, 2019.

[27] Xichen Zhang and Ali A Ghorbani. An overview of online fake news : Characterization, detection, and discussion. *Information Processing & Management*, 57(2) :5,6,7, 2020.