الجمهورية الجزائرية الديمقراطية الشعبية

République algérienne démocratique et populaire

وزارة التعليم العالي والبحث العلمي

Ministère de l'enseignement supérieur et de la recherche scientifique

جامعة عين تموشنت بلحاج بوشعيب

Université –Ain Temouchent- Belhadj Bouchaib

Faculté des Sciences et de Technologie

Département Mathématiques et Informatique

Projet de Fin d'Etudes

Pour l'obtention du diplôme de Master en Informatique

Domaine : Mathématique et Informatique

Filière : Informatique

Spécialité : CySIA

Thème

Détection des rumeurs dans les réseaux publics

Présenté par: Mr.Dahou Sidi Okba

Devant le jury composé de :

Dr. Saidi Samira	M.A.A	UAT.B.B (Ain Temouchent)	Présidente
Dr. Messaoudi Amine	M.C.B	UAT.B.B (Ain Temouchent)	Examinateur
Dr. Berrakem Fatima Zahra	M.A.A	UAT.B.B (Ain Temouchent)	Encadrante

Année Universitaire 2023/2024


بسم الله الرحمن الرحيم
١٤٢٠

Dédicaces

Je dédie ce modeste travail

A mes parents et à toute la famille sans exception

A tous mes enseignants

A tous mes collègues d'étude et de travail

A tous mes amis

Remerciement

Tout d'abord, je remercier Dieu

Je voudrais remercier dans un premier temps mon encadreur Dr.Berrakem .f , pour sa patience, sa disponibilité et ses bons conseils qui ont alimenté ma réflexion.

Je voudrais remercier les membres de jury qui ont accepté de juger ce modeste travail Dr. Saidi .S et Dr. Messaoudi .A .

Résumé

Le développement rapide des médias sociaux a facilité l'échange de grandes quantités de données, mais a également accéléré la propagation de fausses informations. Plusieurs études se sont concentrées sur la détection des rumeurs en analysant principalement le contenu textuel des messages. Toutefois, le contenu visuel, notamment les images, reste largement sous-utilisé. Pourtant, les images sont omniprésentes sur les réseaux sociaux, et leur utilisation est essentielle pour une analyse exhaustive des rumeurs. Dans cette étude, nous présentons une synthèse des travaux actuels sur la classification des rumeurs, en récapitulant les étapes clés de ce processus et les approches utilisées pour les étudier.

L'objectif de notre travail est de mettre en place un système de détection automatiques des rumeurs en utilisant l'apprentissage profond basé sur le modèle de réseau neuronal récurrent (RNN) afin de reconnaître les spams de la base de données KAGGLE.

Nous avons évalué les capacités et les performances de notre système en le testant sur un ensemble de données de test après avoir intégré et validé le modèle RNN sur l'ensemble des données prétraitées. Les résultats montrent une précision d'environ 99,89 % avec un taux d'erreur négligeable.

Mots clés : Intelligence Artificielle (IA), détection des rumeurs, Réseaux de Neurones récurrents (RNNs), Apprentissage en profondeur, apprentissage automatique, rembourrage (Padding), Tensorflow, spam.

Abstract

The rapid development of social media has facilitated the exchange of large amounts of data but has also accelerated the spread of false information. Several studies have focused on rumor detection by primarily analyzing the textual content of messages. However, visual content, particularly images, remains largely underutilized. Yet, images are ubiquitous on social media, and their use is essential for a comprehensive analysis of rumors. In this study, we present a synthesis of current research on rumor classification, summarizing the key steps of this process and the approaches used to study it.

The objective of our work is to develop an automatic rumor detection system using deep learning based on the recurrent neural network (RNN) model to recognize spam in the KAGGLE database.

We evaluated the capabilities and performance of our system by testing it on a test dataset after integrating and validating the RNN model on the preprocessed data. The results show an accuracy of approximately 99.89% with a negligible error rate.

Keywords: Artificial Intelligence (AI), rumor detection, Recurrent Neural Networks (RNNs), Deep Learning, Machine Learning, Padding, TensorFlow, spam.

ملخص

لقد سهل التطور السريع لوسائل التواصل الاجتماعي تبادل كميات كبيرة من البيانات، ولكنه أيضًا أدى إلى تسريع انتشار المعلومات الخاطئة. ركزت العديد من الدراسات على اكتشاف الشائعات من خلال تحليل المحتوى النصي للرسائل بشكل أساسي. ومع ذلك، فإن المحتوى المرئي، وخاصة الصور، لا يزال غير مستغل بشكل كبير. ومع أن الصور منتشرة بشكل واسع على وسائل التواصل الاجتماعي، إلا أن استخدامها ضروري لتحليل شامل للشائعات. في هذه الدراسة، نقدم ملخصًا للأبحاث الحالية حول تصنيف الشائعات، مع تلخيص الخطوات الرئيسية لهذا العملية والأساليب المستخدمة لدراستها.

هدف عملنا هو تطوير نظام للكشف التلقائي عن الشائعات باستخدام التعلم العميق، بناءً على نموذج الشبكة العصبية المتكررة (RNN) للتعرف على الرسائل غير المرغوب فيها (Spam) في قاعدة بيانات كاجل.(KAGGLE)

قمنا بتقييم قدرات وأداء نظامنا من خلال اختباره على مجموعة بيانات اختبار بعد دمج وتأكيد نموذج RNN على البيانات المعالجة مسبقًا. أظهرت النتائج دقة تصل إلى حوالي 99.89% مع معدل خطأ لا يكاد يُذكر.

الكلمات المفتاحية: الذكاء الاصطناعي(AI) ، اكتشاف الشائعات، الشبكات العصبية المتكررة (RNNs)، التعلم العميق، التعلم الآلي، الحشو(Padding) ، TensorFlow، الرسائل غير المرغوب فيها.(Spam)

Table des matières

Contenu

Liste des figures

Liste des tableaux

Introduction Générale

La détection et l'analyse des rumeurs sont devenues des enjeux cruciaux dans l'ère numérique moderne, où la propagation rapide de l'information peut avoir des impacts significatifs sur les individus et les sociétés. Les rumeurs, souvent diffusées de manière virale à travers les réseaux sociaux, peuvent influencer l'opinion publique, créer des paniques, et même affecter les marchés financiers. La détection automatique des rumeurs est donc essentielle pour contenir leur propagation et minimiser leurs effets néfastes.

La technique de détection des rumeurs repose sur l'utilisation d'algorithmes d'apprentissage automatique capables d'analyser des volumes massifs de données textuelles et de détecter des patterns indicatifs de désinformation. Les réseaux de neurones récurrents (RNN) sont particulièrement adaptés pour cette tâche en raison de leur capacité à traiter des séquences de données textuelles et à capturer les dépendances temporelles dans les conversations. Les RNN, et plus spécifiquement leurs variantes avancées comme les Long Short-Term Memory (LSTM) et Gated Recurrent Units (GRU), offrent des capacités améliorées pour comprendre le contexte et la dynamique des rumeurs.

Dans ce projet, nous nous proposons d'explorer l'application des RNN pour la détection des rumeurs sur les réseaux sociaux, en utilisant une base de données Kaggle réputée pour sa richesse en données pertinentes. Les RNN seront entraînés pour identifier des caractéristiques textuelles associées aux rumeurs et pour distinguer les messages fiables des messages trompeurs. La mise en œuvre et l'évaluation du modèle seront réalisées en utilisant des techniques d'apprentissage supervisé sur des ensembles de données annotés, permettant ainsi une approche rigoureuse pour le développement et la validation de notre système de détection.

Nous discuterons des spécificités des RNN et de leur adaptation à la tâche de détection des rumeurs, ainsi que des défis associés à l'analyse des données textuelles provenant de diverses sources de réseaux sociaux. Ce projet vise non seulement à développer un modèle efficace, mais aussi à contribuer à une meilleure compréhension des mécanismes de diffusion des rumeurs et des stratégies pour les contrer.

L'objet de ce mémoire est de développer un système de détection des rumeurs plus efficace et

fiable. Le travail est structuré en trois chapitres, organisés comme suit :

Dans le premier chapitre, nous présenterons une vue d'ensemble des différents aspects de la diffusion des rumeurs sur les réseaux sociaux, ainsi qu'un cadre général pour la mise en place d'un système de détection basé sur l'analyse textuelle.

Le deuxième chapitre sera consacré à l'apprentissage profond et aux différents types de réseaux de neurones artificiels utilisés pour la détection des rumeurs. Nous y détaillerons les réseaux de neurones récurrents (RNN), y compris les architectures telles que LSTM et GRU.

Le dernier chapitre présentera la conception de notre modèle. Nous y exposerons les résultats obtenus pour la validation et l'évaluation du système. Enfin, une comparaison entre les hyper paramètres, et discuterons les résultats obtenus en termes de performance et de fiabilité.

Chapitre 01

Généralités sur la détection des rumeurs dans les réseaux publics

1. Chapitre01 : Généralités sur la détection des rumeurs dans les réseaux publics

1. *Introduction*

En raison de l'essor des réseaux sociaux, la détection des rumeurs est devenue un défi majeur. Les rumeurs peuvent être rapidement diffusées, ce qui peut avoir des conséquences potentiellement graves.

Nous présentons dans ce chapitre une description générale des différents aspects de la détection des rumeurs dans les réseaux publics. En fait, dans un premier temps, nous présentons un ensemble de concepts préliminaires qui introduisent le domaine. Ensuite, nous décrivons le système de détection des rumeurs à travers un schéma qui explique les phases du système. Ainsi que, on parlera sur les différentes techniques et les méthodes pour la détection. Nous donnons ensuite les Enjeux et Défis de la détection des rumeurs. A la fin de ce chapitre on présentant quelques travaux connexes.

2. *Definition d'une rumeur*

Les rumeurs sont le sujet de recherches dans plusieurs disciplines, comme la philosophie, la psychologie sociale, les sciences politiques, les sciences de gestion et, plus récemment, l'informatique, notamment avec l'expansion des réseaux sociaux.[2]

Les rumeurs sont des informations non vérifiées, souvent transmises de manière informelle. Elles peuvent être basées sur des faits réels ou être complètement inventées.

De nombreuses définitions différentes de la rumeur ont été proposées dans la littérature. Zhao et al. (2015) définissent une rumeur comme étant une déclaration controversée et vérifiable, Zubiaga et al. (2018) comme une information en circulation dont le statut de véracité n'a pas encore été vérifié au moment de la publication. Pour Hamidian et Diab (2015), une rumeur peut être à la fois vraie et fausse. C'est une affirmation dont la véracité est mise en doute et n'a pas de source claire, même si ses origines et ses intentions idéologiques ou partisanes sont claires. Cao et al. (2018) vont plus loin en proposant des familles de rumeurs. Les rumeurs générales ont valeur de vérité non vérifiée. Les rumeurs générales englobent deux sous familles de rumeurs : les rumeurs objectives, dont la valeur de vérité est confirmée par une source fiable ou crédible ; et les rumeurs subjectives, dont la valeur de vérité est déterminée par les jugements subjectifs des utilisateurs des réseaux sociaux. Soit incertaine, une rumeur n'implique pas forcément de fausses informations.

La définition la plus utilisée par la communauté scientifique est celle évoquée par Di Fonzo et Bordia (2006) et Qazvinian et al. (2011), où une rumeur est définie comme une information qui émerge et se propage, et dont la valeur de vérité est non vérifiée ou délibérément fausse.

3. *La typologie des rumeurs*

La typologie des rumeurs peut varier selon plusieurs critères :

1. Rumeurs basées sur des faits réels : Ces rumeurs s'appuient sur des événements réels, mais sont souvent déformées ou exagérées par la transmission orale ou écrite.[3]
2. Rumeurs spéculatives : Elles résultent de la mauvaise interprétation de données ou d'informations, souvent en l'absence de preuves solides.[4]
3. Rumeurs intentionnelles : Celles-ci sont délibérément propagées pour tromper ou manipuler l'opinion publique. Elles sont souvent utilisées dans des contextes politiques ou commerciaux.[5]
4. Rumeurs de complot : Ces rumeurs impliquent des théories de conspiration, souvent sans fondement, mais peuvent avoir un impact significatif sur les croyances et les comportements des gens.[6]
5. Rumeurs de peur : Propagées en période de crise, ces rumeurs jouent sur les émotions et les peurs des gens pour se répandre rapidement.[7]

Les rumeurs ont une influence importante sur la société et peuvent affecter les comportements individuels et collectifs.

4. *Evolution et propagation des rumeurs*

Les rumeurs se propagent sur les réseaux sociaux selon des dynamiques spécifiques qui ont évolué avec l'avènement de ces plateformes. Voici les principaux aspects :

1. Propagation rapide des fausses informations : Les réseaux sociaux facilitent la diffusion rapide des rumeurs, particulièrement des fausses informations, qui voyagent souvent plus vite et atteignent un public plus large que les informations vérifiées. Ce phénomène est amplifié par les algorithmes des plateformes qui privilégient les contenus engageants, souvent sensationnalistes.[8]
2. Influence des écosystèmes médiatiques : La structure des réseaux sociaux et l'accès facilité à des plateformes variées favorisent la diffusion de rumeurs. Les utilisateurs partagent des informations sans les vérifier, souvent influencés par les croyances ou les biais cognitifs.[9]
3. Impact des événements récents : Des événements d'actualité, comme la pandémie de COVID-19, ont vu une augmentation des rumeurs et des théories du complot. Ces périodes de crise sont propices à la diffusion rapide de fausses informations, accentuant la désinformation.[10]

4. Outils de détection et de correction : La lutte contre la propagation des rumeurs inclut le développement d'outils de détection automatique sur les plateformes sociales, visant à identifier et corriger les fausses informations rapidement.[2]

5. Historique des rumeurs dans les réseaux sociaux

Les rumeurs ont une longue histoire et se sont adaptées aux technologies de communication de chaque époque. À l'ère des réseaux sociaux, leur propagation a pris de nouvelles dimensions :

a) **Origines et nature des rumeurs :** Les rumeurs existent depuis des siècles, se répandant à travers divers moyens de communication, comme le bouche-à-oreille. Elles peuvent être vraies ou fausses, mais leur véracité est souvent mise en doute.[11] Toute rumeur est condamnée à s'éteindre, puisqu'elle-même crée les mécanismes de sa disparition. Cependant, comme l'indique Kapferer, certaines rumeurs semblent résister au temps en réapparaissant plus tard et en étant absorbées par un nouveau public. En ce sens, un des cas emblématiques est la rumeur d'Orléans, étudiée par Edgar Morin dans un ouvrage de référence à ce sujet.

La rumeur d'Orléans En mai 1969 se répand dans la ville française ou la rumeur Qu'un magasin de vêtements pour femmes pratiquerait ce que l'on a appelé la « traite de Blanches ». La rumeur a atteint très rapidement six établissements semblables. Elle était la suivante :

« Des jeunes filles sont droguées grâce à des piqûres dans les cabines d'essayage et sont déposées dans le sous-sol, d'où elles seront envoyées le soir même vers des points exotiques de prostitution. Les magasins accusés sont tenus par des commerçants juifs ».

La rumeur s'est répandue malgré un récit invraisemblable, mais localisé dans le temps, au sein d'une société qui disposait de moins d'informations qu'aujourd'hui pour confirmer la nouvelle. [11]

Figure 1. 1: Des rumeurs naturelles.[1]

b) **Évolution avec les technologies** : De la tradition orale aux plateformes numériques, les rumeurs se sont toujours appuyées sur les technologies disponibles. Aujourd'hui, les réseaux sociaux amplifient leur portée, rendant les fausses informations plus visibles et influentes.[1]
À l'ère de Facebook et de Twitter, le processus d'enquête journalistique est de plus en plus remplacé par l'instantanéité des réseaux sociaux. Sur ce point, tant le citoyen que la presse présentent le même intérêt à véhiculer et à diffuser de l'information sur Internet, ce qui, compte tenu des éléments que nous soulevons, peut en compromettre la qualité. La création de sites spécialisés visant à transmettre au lecteur les dernières rumeurs sur des sujets tels que le sport ou la technologie démontre la magie que ce type d'information suscite même à l'ère de l'Internet.[11]

c) **Propagation accélérée** : Avec l'avènement des réseaux sociaux, la diffusion des rumeurs est devenue instantanée et mondiale. Ces plateformes permettent un partage rapide d'informations, souvent sans vérification préalable, ce qui peut mener à des conséquences sérieuses.[2]

L'un des changements introduits par les nouvelles technologies, et qui s'ajoute au récit traditionnel de la rumeur, est la possibilité de joindre des images à une narration. Un faux portrait-robot par exemple ou la photographie d'un individu peuvent être manipulés ou retirés de leur contexte afin de créer une situation gênante pour la victime, comme l'illustrent des cas récents sur Facebook et WhatsApp. Le recours visuel apparaît de la sorte comme un moyen de rapprocher la fiction de la réalité, ce qui confère de la vraisemblance au récit.[11]

Une conséquence extrême observée récemment au Brésil est liée à l'épisode connu du cas Guarujá 44. Le 3 mai 2014, la femme au foyer Fabiane Maria de Jesus est décédée deux jours après avoir été lynchée, après avoir été confondue avec une supposée kidnappeuse d'enfants, à la suite de la publication d'un portrait robot sur une page Facebook. Un groupe de résidents du quartier de Morrinhos, à Guarujá (SP 45), a alors essayé de rendre lui-même justice après que la rumeur ait provoqué des répercussions dans cette communauté-là bien qu'aucun enfant n'ait été séquestré dans la région. La rumeur du cas Guarujá a provoqué un phénomène de panique à partir d'une histoire irréelle et a eu un dénouement tragique.[11]

La simplicité du processus de partage a conduit à de grands volumes de contenu d'actualités se propageant sur les réseaux sociaux et atteignant un grand nombre de lecteurs en très peu de temps. En particulier, le contenu multimédia (images, vidéos) peut rapidement atteindre des audiences massives et devenir viral en raison de sa consommation facile et de sa grande valeur de divertissement.[12]

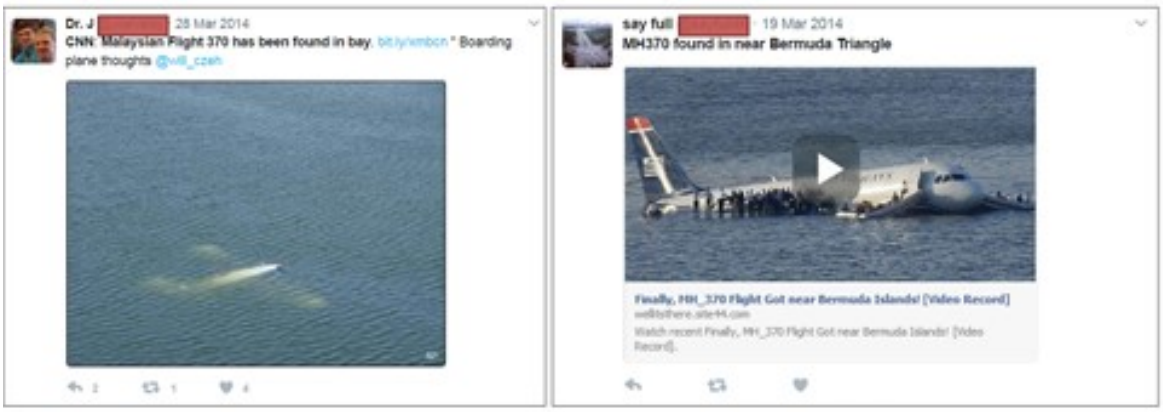

Figure 1.2:Exemples d'images fausses qui se sont propagées sur Twitter lors de l'incident des lignes aériennes Malaisiennes en mars 2014.[12]

d) **Rôle des réseaux sociaux :** Les réseaux sociaux ont un impact significatif sur la propagation des rumeurs, facilitant le partage de contenus controversés ou sensationnalistes. Les algorithmes de ces plateformes favorisent l'engagement, ce qui peut accentuer la diffusion des rumeurs.[13]

Silverman souligne en outre que l'apparition de sites spécialisés dans la diffusion de fausses nouvelles est en train de « polluer » l'information présente sur les réseaux en engendrant des hoax qui attirent un nombre significatif de partages et de vues. Au Brésil, ce genre de sites a trouvé une grande répercussion sur les réseaux sociaux numériques ces dernières années. Selon l'auteur, il n'est pas rare que les grands groupes de médias diffusent ces informations comme si elles étaient vraies, ce qui contribue à augmenter la désinformation sur l'Internet.[11]

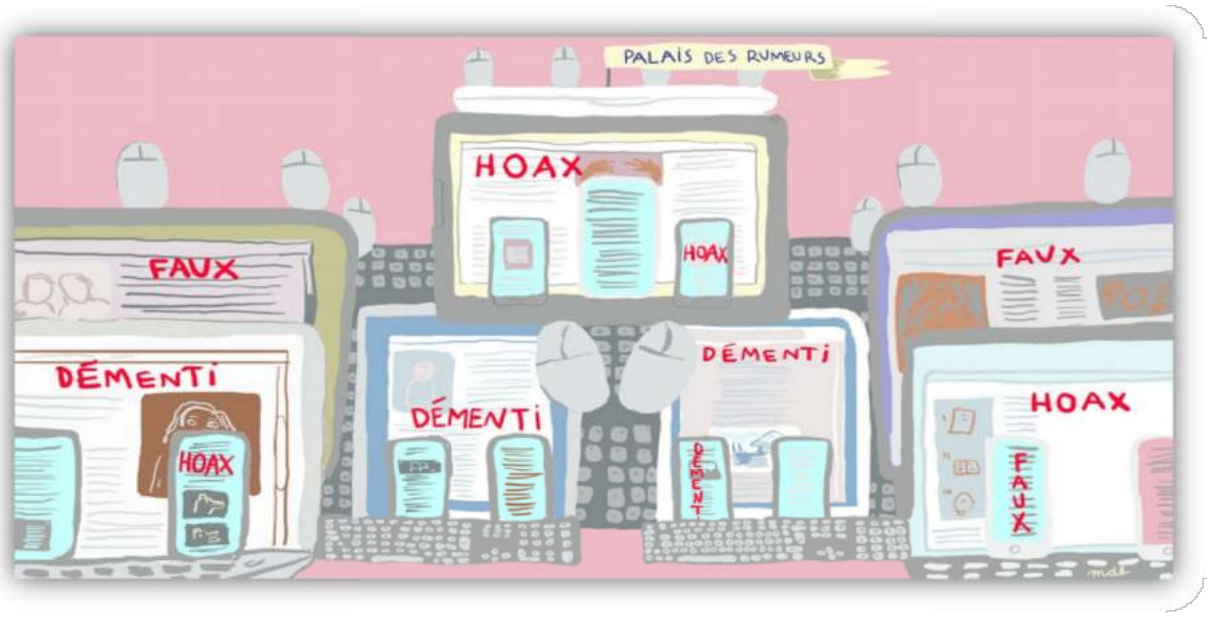

Figure 1.3 : La propagation des informations erronées ou invérifiables.[1]

6. Mécanismes de diffusion sur les plateformes sociales modernes

Les rumeurs se diffusent sur les plateformes sociales modernes par divers mécanismes influencés par des facteurs cognitifs, sociaux et technologiques. Voici quelques points :

- **Mécanismes cognitifs** : Les rumeurs se propagent souvent en raison de biais cognitifs. Les utilisateurs ont tendance à partager des informations qui confirment leurs croyances ou émotions. Cela est accentué par la nature émotionnelle et souvent alarmiste des rumeurs, qui captent l'attention et incitent au partage rapide sans vérification.[14]
- **Propagation virale** : La viralité des informations sur les réseaux sociaux joue un rôle crucial. Les algorithmes des plateformes favorisent les contenus qui génèrent de l'engagement, ce qui peut inclure des rumeurs sensationnalistes. Les rumeurs peuvent ainsi atteindre un large public rapidement.[15]
- **Réseaux sociaux** : Les réseaux sociaux eux-mêmes structurent la diffusion des rumeurs. Les liens sociaux et les groupes d'entre-soi favorisent la circulation rapide de l'information. Les rumeurs peuvent traverser des communautés homogènes, renforçant leur crédibilité perçue.[16]
- **Technologies et algorithmes** : Les plateformes sociales utilisent des algorithmes qui peuvent amplifier la diffusion des rumeurs. Les informations fausses mais virales sont souvent mises en avant, augmentant leur visibilité et leur

impact.[17]Les plateformes comme Facebook et Twitter utilisent des algorithmes qui favorisent la visibilité des contenus engageants, y compris les rumeurs.[18]

- **Analyse de la crédibilité** : L'analyse de la crédibilité des rumeurs est complexe et nécessite une compréhension approfondie du contenu et de son contexte. Des systèmes de vérification et des outils d'intelligence artificielle sont de plus en plus utilisés pour détecter et contrer les rumeurs.[2]

Ces points illustrent comment les rumeurs peuvent se propager et avoir un impact significatif sur les sociétés modernes via les réseaux sociaux.

7. *Système de détection des rumeurs dans les réseaux sociaux :*

Les systèmes de détection des rumeurs sur les réseaux sociaux se basent sur plusieurs étapes cruciales. Voici un aperçu :

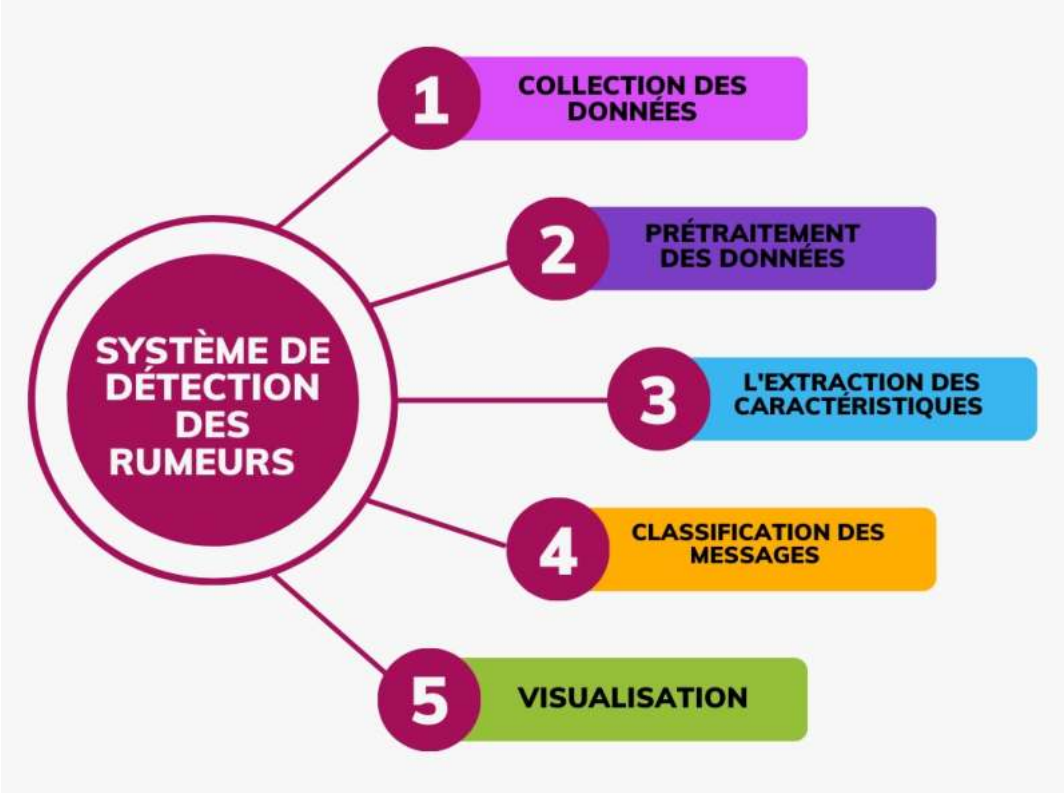

Figure 1.4:Architecture des systèmes de détection des rumeurs dans les réseaux publics.

7.1 Collecte de données:

Cette étape implique la collecte d'un vaste ensemble de messages et de publications à partir de diverses plateformes de réseaux sociaux. Ces données peuvent inclure des tweets, des posts Facebook, des commentaires, etc. Les techniques de web scraping et d'API sont couramment utilisées pour cette tâche.[2]

7.2 Prétraitement des données:

Les données collectées sont nettoyées pour éliminer les bruits, les doublons et les éléments non pertinents. Cela inclut le filtrage des publicités, des messages automatisés et des informations non textuelles. Cette étape est cruciale pour assurer la qualité des données analysées.[17]

7.3 L'extraction des caractéristiques:

L'extraction des caractéristiques est une étape cruciale pour la détection des rumeurs dans les réseaux sociaux. Elle consiste à identifier et à extraire des éléments spécifiques des messages et des interactions pour permettre une classification précise des rumeurs. Voici les principales caractéristiques extraites dans ce contexte :

7.3.1 Caractéristiques linguistiques :

- N-grams : Séquences de mots ou de caractères utilisés pour capturer les motifs fréquents dans le texte.
- Fréquence des mots : Analyse de la fréquence des mots clés associés à des rumeurs.
- Sentiment analysis : Détection des sentiments (positif, négatif, neutre) pour comprendre le ton des messages.[2]

7.3.2 Caractéristiques liées aux utilisateurs :

- Historique des publications : Analyse des publications antérieures pour identifier les utilisateurs susceptibles de propager des rumeurs.
- Influence sociale : Mesure du nombre de followers, de likes, et de partages pour déterminer l'impact potentiel des messages.[19]

7.3.3 Caractéristiques structurelles :

- Analyse du réseau : Étude des connexions entre les utilisateurs pour comprendre la propagation des informations.
- Géolocalisation : Localisation des messages pour identifier les zones géographiques où les rumeurs se propagent le plus.[20]

- Analyse des sentiments et des émotions : Les systèmes analysent le ton et le sentiment des messages pour identifier les rumeurs potentiellement négatives ou alarmantes. Cette analyse peut aider à prioriser les contenus à examiner de plus près.[21]

7.3.4 Caractéristiques temporelles:

- Heure de publication : Analyse des heures de publication pour identifier les pics de propagation des rumeurs.
- Durée de vie des messages : Mesure du temps pendant lequel les messages restent actifs et se propagent.[22]

7.3.5 Métadonnées:

- Informations sur le dispositif : Type de dispositif utilisé pour publier les messages (mobile, ordinateur).
- Type de contenu : Analyse des médias attachés aux messages (images, vidéos, liens).[23]

L'extraction de ces caractéristiques permet de nourrir des modèles de machine learning pour classifier les messages et identifier les rumeurs de manière plus efficace.

7.4 Classification des messages:

Les messages sont classifiés comme vrais ou faux en utilisant des modèles d'apprentissage automatique. Ces modèles sont formés sur des ensembles de données annotées, où chaque message est étiqueté comme vérifié ou faux.[2]

Un système de classification des rumeurs se compose de quatre éléments : la détection des rumeurs, le suivi des rumeurs, la classification des prises de position sur les rumeurs et la classification de la véracité des rumeurs. Lorsqu'il s'agit de concevoir un système de classification des rumeurs, le facteur qui détermine largement les approches à utiliser est la temporalité, c'est-à-dire les nouvelles rumeurs qui émergent pendant des nouvelles de dernière minute. Les rumeurs qui apparaissent dans le contexte des nouvelles de dernière minute sont généralement celles qui n'ont pas été observées auparavant. Par conséquent, les rumeurs doivent être détectées automatiquement et un système de classification des rumeurs doit être capable de traiter de nouvelles rumeurs jamais vues auparavant, en tenant compte du fait que les données d'entraînement disponibles pour le système peuvent différer de ce qui sera observé par la suite. Dans ces cas, où la détection et la résolution rapides des rumeurs sont cruciales, un flux de publications doit être traité en temps réel.[24]

7.4.1 Détection des rumeurs:

Un classificateur binaire reçoit un flux continu de publications et attribue à chaque publication une étiquette, soit "rumeur" soit "non-rumeur". C'est essentiel pour combattre la propagation des rumeurs.[25]

7.4.2 Suivi des rumeurs:

Les médias sociaux sont recherchés pour des publications contenant les mots-clés de la rumeur ou une phrase définissant le sujet de la rumeur.[25]

Identification et suivi des rumeurs détectées inclut le traçage de la propagation des rumeurs et l'analyse des réseaux pour comprendre comment et par qui elles sont diffusées.[11]

7.4.3 Classification de la position:

La classification de la position dans les réseaux sociaux en ligne consiste à déterminer le type d'orientation que chaque message individuel exprime à l'égard de la véracité contestée d'une rumeur. Un classifieur de position prend en entrée un ensemble R ={ r_1 ,r_2 ,r_n}de rumeurs, où chaque rumeur r_i est composée d'une collection de taille variable de messages M ={ m_1,m_2,m_j }discutant de cette rumeur. L'objectif consiste à déterminer la position (pour ou contre, par exemple) de chaque message m_j traitant de r_i.[2]

Chaque publication liée à la rumeur générée par le module de surveillance des rumeurs reçoit une position, telle que "soutien", "démenti" ou "interrogation", par ce module.[25]

Mendoza et al. (2010) visent à comprendre les positions des utilisateurs de Twitter via une analyse manuelle. Ils constatent que la majorité des tweets qui sont liés à de vraies rumeurs supportent ces rumeurs, alors que la moitié des tweets qui sont associés à de fausses rumeurs questionnent ou s'opposent à ces rumeurs. Cela tend à montrer que les positions exprimées vis-à-vis des rumeurs sont une indication précieuse pour déterminer la véracité d'une rumeur.[2]

7.4.4 Classification de la véracité

La classification de la véracité d'une rumeur est la finalité et l'étape cruciale du processus d'analyse de la rumeur. Cette tâche consiste à déterminer une valeur de véracité. Formellement, une rumeur r est définie comme un ensemble de messages M ={ m_1,m_2.....,m_n }. Détecter la véracité de la rumeur consiste à déterminer si la rumeur r est confirmée comme vraie, prouvée fausse ou que sa valeur de véracité demeure non vérifiée, par une fonction de prédiction :

f(r)⟶ {vraie, fausse, non vérifiée}, telle que :

$$f(r) = \begin{cases} \text{vraie} & \text{si } r \text{ est confirmée comme vraie,} \\ \text{fausse} & \text{Si } r \text{ est prouvée fausse} \\ \text{non vérifiée} & \text{sinon} \end{cases}$$

Cette section évalue la légitimité d'une rumeur en combinant les résultats des deux premières sections avec des données provenant d'autres sources web.[25] Certaines rumeurs peuvent circuler pendant longtemps sans que leur véracité soit établie avec certitude. Ces rumeurs suscitent un intérêt continu important malgré la difficulté à établir la vérité réelle. De plus, le système peut utiliser les discussions historiques sur la rumeur pour classer les discussions en cours, où le vocabulaire est beaucoup moins susceptible de varier, et donc le classificateur basé sur des données anciennes peut encore être utilisé pour des données nouvelles. Contrairement aux rumeurs émergentes, pour les rumeurs de longue date, le traitement est généralement rétrospectif, de sorte que les publications n'ont pas nécessairement besoin d'être traitées en temps réel.[24]

7.5 Visualisation et alerte

- Création de tableaux de bord pour visualiser les rumeurs détectées et leur propagation.
- Systèmes d'alerte pour notifier les administrateurs ou les utilisateurs lorsque des rumeurs sont identifiées.[17][11]

Ces étapes permettent de créer un système robuste pour identifier et gérer les rumeurs sur les réseaux sociaux.

8. Techniques et méthodologie de détection des rumeurs

La détection des rumeurs est un domaine de recherche complexe, particulièrement pertinent dans le contexte des réseaux sociaux où les fausses informations peuvent se propager rapidement. Voici quelques techniques et méthodologies qui peuvent être classées en trois principales catégories : l'analyse textuelle, l'analyse visuelle et l'analyse des réseaux.

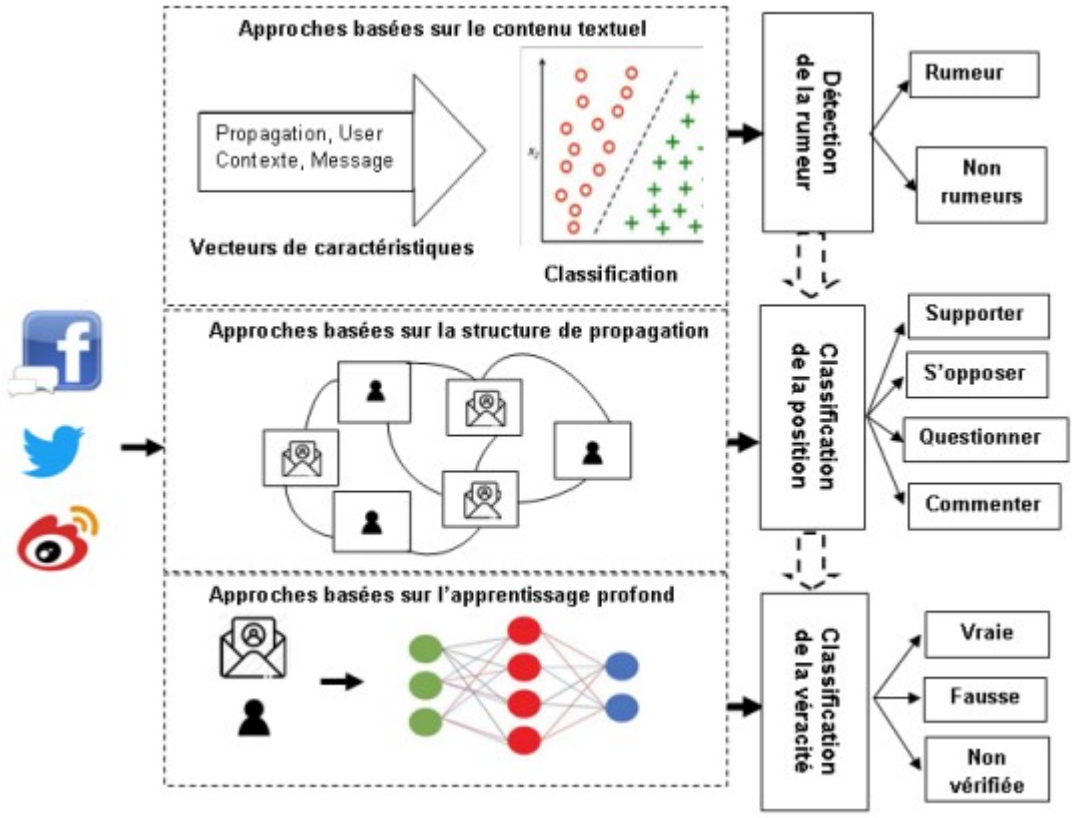

Figure 1. 3:Cartographie des tâches et des familles de méthodes pour prédire la véracité des rumeurs.[2]

8.1 Analyse textuelle

Cette méthode se concentre sur le contenu des messages pour identifier des indices de rumeurs.[2].On peut examiner des éléments comme :

- Le style d'écriture : présence de points d'exclamation excessifs, majuscules abusives, ton accusateur.[26]
- Le vocabulaire utilisé : mots chargés d'émotion, formulation vague ou sensationnaliste [27]
- La présence de citations et de sources fiables : absence de sources ou sources inconnues pouvant indiquer un manque de crédibilité.[26]

Les algorithmes de traitement du langage naturel (NLP) sont couramment utilisés pour détecter des mots-clés, des expressions et des structures de phrases typiques des rumeurs[2]. L'analyse textuelle permet également d'évaluer la crédibilité de la source et de détecter les hoax[20].

8.1.1 Utilisation des algorithmes de traitement du langage naturel

Le traitement automatique du langage naturel (NLP) combine la linguistique informatique (modélisation du langage humain basée sur des règles) avec des modèles statistiques et des modèles de machine learning pour permettre aux ordinateurs et aux appareils numériques de reconnaître, de comprendre et de générer du texte et de la parole.
[28]

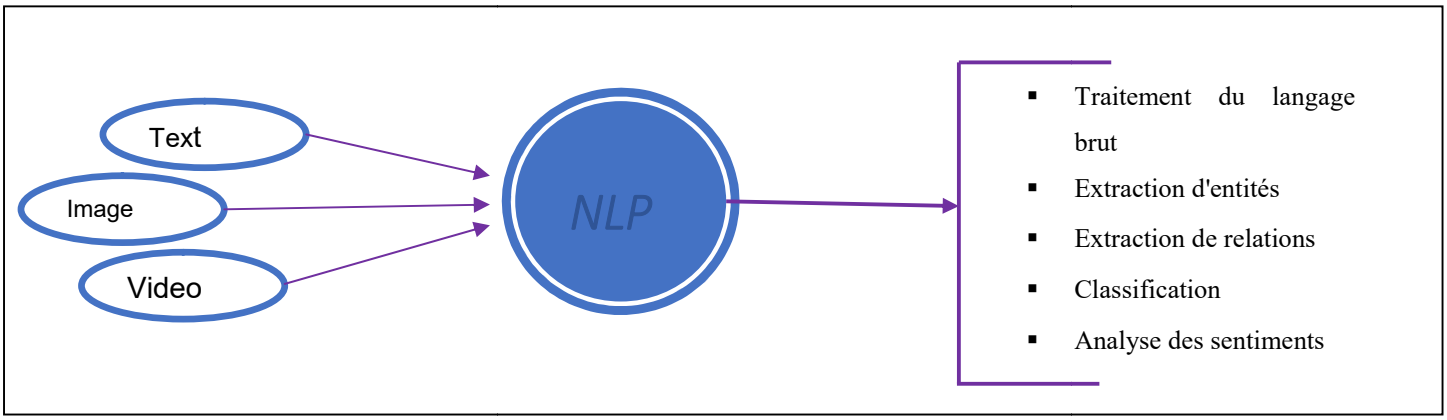

Figure 1. 4:Les applications du traitement de langage naturel.

Le traitement du langage naturel (NLP) est utilisé pour analyser les textes afin de détecter les rumeurs dans les réseaux sociaux. Voici comment ces algorithmes sont appliqués :

1. **Prétraitement des Données :** Les données textuelles sont nettoyées et normalisées pour être analysées de manière efficace. Cela inclut la suppression des stop words, la tokenisation et la lemmatisation [29]

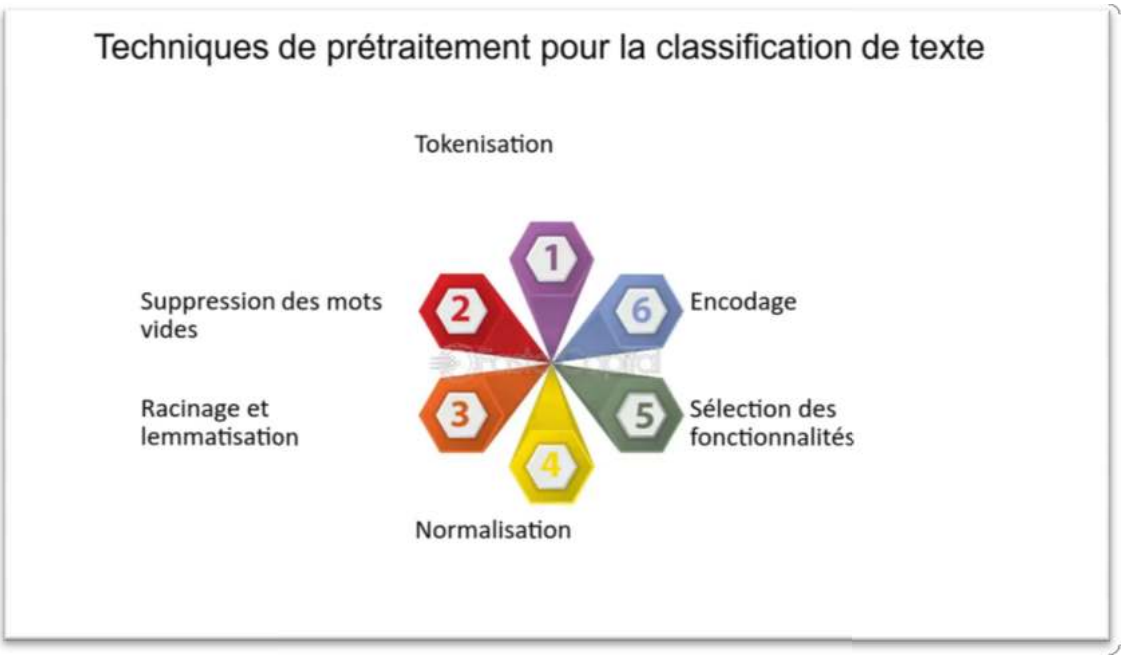

Figure 1. 5:Techniques de prétraitement pour la classification de texte.[30]

***Tokenisation* :** Consiste à diviser le texte donné en éléments distincts. Il s'agit de mots ou d'autres symboles séparés par des espaces et des signes de ponctuation et qui forment une phrase. Les algorithmes de tokenisation peuvent être classés en fonction de leurs entrées et de leurs sorties : certains se limitent à des phrases simples, d'autres acceptent des blocs de textes arbitraires, et d'autres encore décomposent les données textuelles en mots individuels uniquement. Pour les langues alphabétiques telles que l'anglais, les scripts déterministes suffisent généralement à obtenir une tokenisation précise.[31]

***Détermination de la racine et la lemmatisation* :** Un autre processus informatique important pour la normalisation de texte consiste à éliminer les affixes flexionnels, tels que les suffixes -ed et -s en anglais. La recherche de radicaux est le processus qui consiste à trouver le même concept sous-jacent pour plusieurs mots, de sorte qu'ils doivent être regroupés en une seule caractéristique en éliminant les affixes.[31]

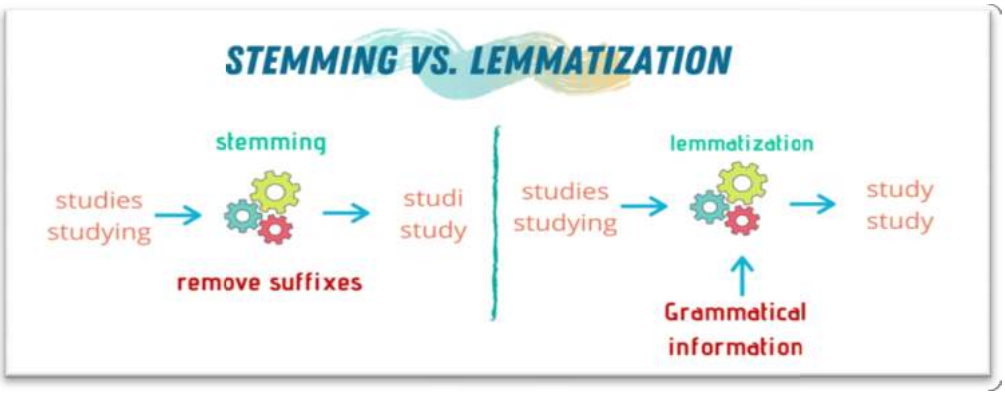

Figure 1. 6:Différence entre stemming et lemmatization.[32]

2. **Extraction d'Information** :Les algorithmes de NLP identifient des motifs et extraient des informations pertinentes des textes, comme les entités nommées et les relations entre elles.[33]
3. **Analyse des Sentiments** : Les algorithmes évaluent le ton des messages pour déterminer s'ils sont positifs, négatifs ou neutres. Cette analyse aide à identifier des messages susceptibles de contenir des rumeurs.[34]
4. **La reconnaissance d'entités nommées :** identifie et classe les entités nommées telles que les personnes, les lieux, les dates et les organisations.[29]
5. **Détection des Anomalies** : Les modèles de NLP peuvent détecter des anomalies dans les flux de messages, signalant des sujets ou des informations inhabituels qui pourraient indiquer la propagation de rumeurs.[28]

8.1.2 Avantages du traitement du langage naturel

- Élévation de la communication : le NLP permet une communication plus naturelle avec les applications de recherche.[29]
- Automatisation des tâches répétitives : Le NLP offre la possibilité d'automatiser des opérations comme le tri des courriels, la réponse à des questions couramment posées, ou encore la transcription de conversations.
- Curation de contenus : le NLP peut identifier les informations les plus pertinentes pour les utilisateurs individuels en fonction de leurs préférences.[29]

8.1.3 Difficultés et limites du traitement du langage naturel

- Compréhension et mémoire contextuelles limitées : les modèles de NLP ont souvent des difficultés à interpréter ou retenir le sens des mots ou des phrases en fonction du contexte dans lequel ces mots ou phrases sont utilisés. Ce problème peut donner lieu à des erreurs d'interprétation ou des analyses de données textuelles erronées.[34]
- Ambiguïté et polysémie :La signification des mots et des phrases peut varier (ambiguïté lexicale et syntaxique), ce qui rend la compréhension précise par les machines difficile.
- La variation du langage : Les dialectes, les jargons et les styles des langues naturelles sont abondants. Il est difficile pour les modèles TLN de gérer cette diversité.

8.1.4 Les types de classification

Les classifications sont des systèmes utilisés pour organiser des éléments en catégories basées sur des critères spécifiques. Il existe plusieurs types de classification selon le type d'apprentissage utilisé :

1. L'apprentissage supervisé (supervised Learning)

L'apprentissage supervisé consiste à utiliser des ensembles de données étiquetées pour former un algorithme de machine Learning. Ladite supervision tient au fait que les étiquettes sur les données aident le modèle à prédire des résultats avec précision. Autrement dit, les données de formation contiennent déjà les réponses correspondantes aux sorties attendues.[35]

L'exploration de données peut être divisée en deux catégories de problèmes lors de l'apprentissage supervisé :

a. La classification :

Les problèmes de classification consistent à utiliser un algorithme pour attribuer avec précision des données de test à des catégories spécifiques. Par exemple, les algorithmes d'apprentissage supervisé peuvent être utilisés pour classer les spams dans un dossier distinct de la boîte de réception. Les classifieurs linéaires, les machines à vecteurs de support, les arbres de décision et les forêts d'arbres décisionnels sont tous des types courants d'algorithmes de classification.[36]

b. La régression :

La régression est un autre type de méthode d'apprentissage supervisé qui utilise un algorithme pour comprendre la relation entre les variables dépendantes et indépendantes. Les modèles de régression sont utiles pour prédire des valeurs numériques sur la base de différents points de données. Les algorithmes de régression sont par exemple la régression linéaire, la régression logistique et la régression polynomiale.[36]

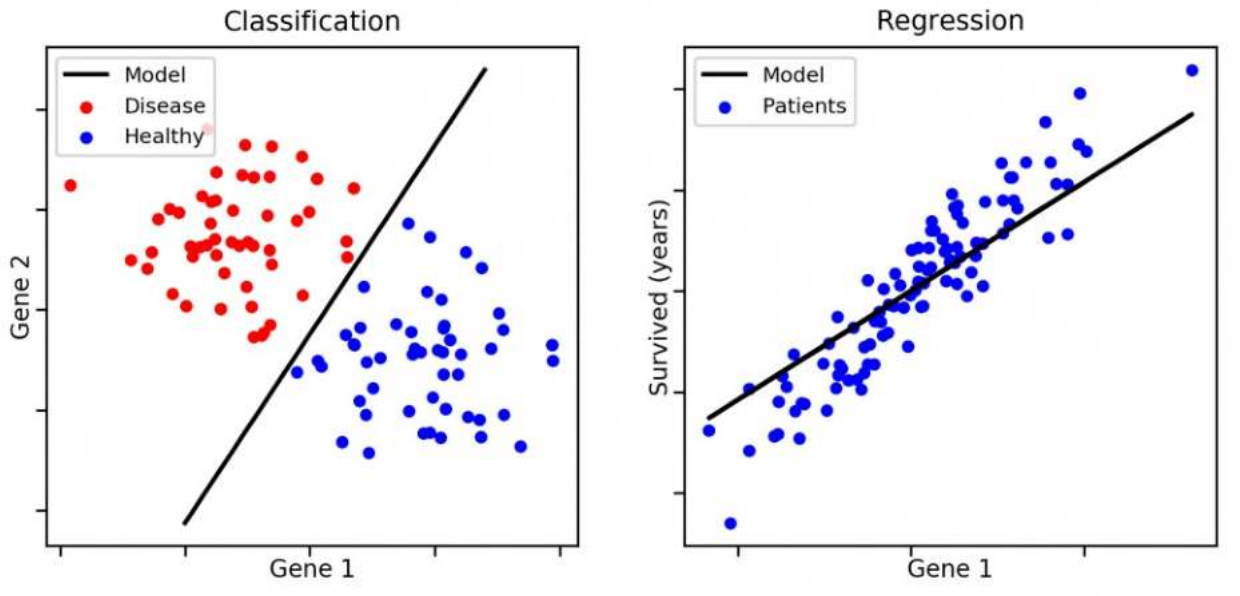

Figure 1. 7:La classification et la régression[37]

2. L'apprentissage non supervisé (unsupervised Learning)

L'apprentissage non supervisé consiste à entraîner des modèles, sans réaliser d'étiquetage manuel ou automatique des données au préalable. Les algorithmes regroupent les données en fonction de leur similitude, sans aucune intervention humaine.[38]

a. Le clustering ou partitionnement de données :

Le clustering est une technique d'exploration de données permettant de regrouper des données non étiquetées en fonction de leurs similitudes ou de leurs différences. Cette technique est utile pour la segmentation de marché, la compression d'images, etc.[36]

b. L’association :

L’association est un autre type de méthode d’apprentissage non supervisé qui utilise différentes règles pour trouver des relations entre les variables d’un ensemble de données donné.[36]

c. La réduction de la dimensionnalité :

Tâches de dimensionnalité où l'algorithme cherche à réduire le nombre de variables, de caractéristiques ou de fonctionnalités dans un ensemble de données. Étant donné que certaines de ces dimensions sont corrélées, des informations redondantes ou répétées peuvent augmenter le bruit de l'ensemble de données et avoir un impact sur la formation et performant de modèle. Cette technique est souvent utilisée dans l'étape de prétraitement des données, par exemple lorsque le bruit est supprimé des données visuelles pour améliorer la qualité de l'image.[39]

3. L’apprentissage par renforcement

L’apprentissage par renforcement ou Reinforcement Learning est une méthode de Machine Learning de plus en plus utilisée. Elle consiste à laisser les ordinateurs apprendre de leurs expériences grâce à un système de récompense ou de pénalité. Il pourrait même s’agir de la clé permettant l’avènement d’une intelligence artificielle générale comparable à celle de l’humain.[40]

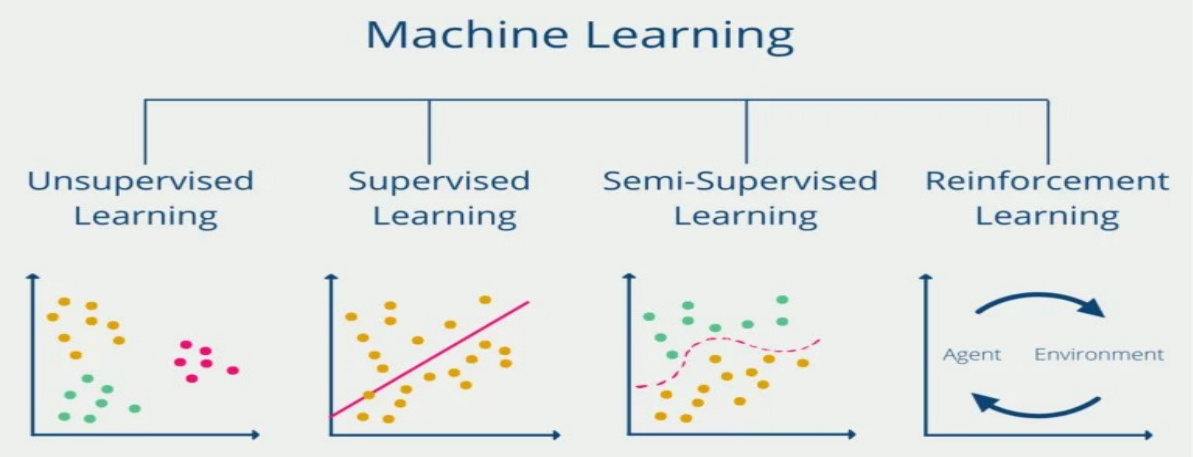

Figure 1. 8 Les types de classification.[41]

8.2 Analyse de contenu visuelle

Cette technique consiste à visualiser les réseaux de propagation de l'information pour mieux comprendre comment les rumeurs se diffusent. L'analyse visuelle des réseaux (visual network analysis) aide à identifier les nœuds clés et les chemins de diffusion des rumeurs[42]. Cette approche permet de repérer rapidement les points d'origine et les principales voies de propagation.

8.2.1 Méthodes de vérification des images et vidéos partagées en ligne

La vérification des images et vidéos partagées en ligne est cruciale pour lutter contre la désinformation. Voici quelques techniques courantes utilisées pour authentifier ces contenus :

a. Analyse du contenu de l'image

Les méthodes décrites dans cette section exploitent uniquement le contenu des images. Il existe principalement deux types d'algorithmes de détection des images contrefaites : les algorithmes actifs et passifs.[43]

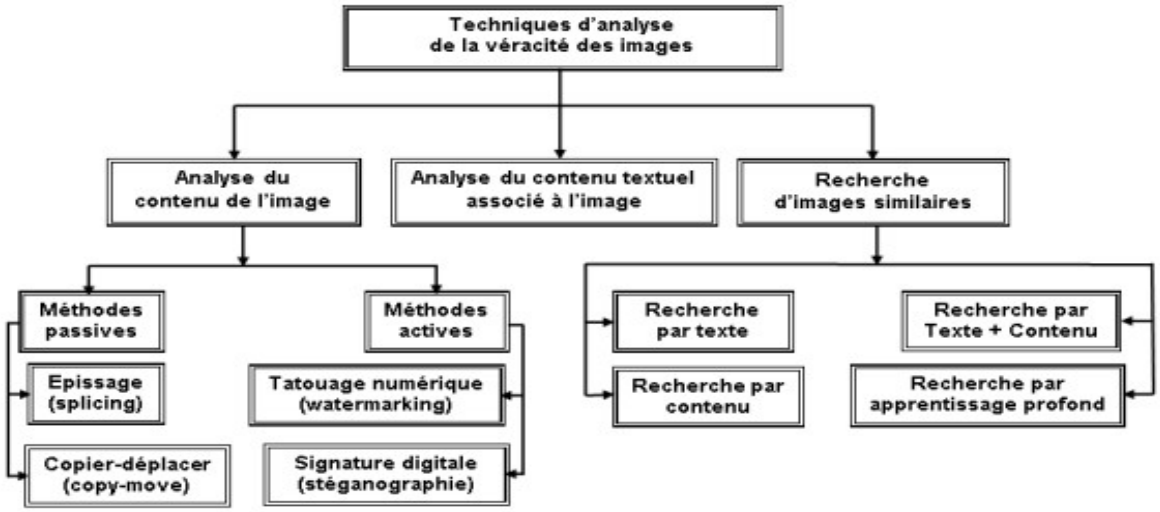

Figure 1. 9: Les approches de prédiction du rumeur base sur les images.[43]

Les algorithmes actifs exploitent une signature, comme un filigrane, la double compression des images JPEG ou la signature laissée par les appareils de capture. Cette signature est mise en correspondance avec la signature de l'image originale pour détecter toute altération.[2]

Les algorithmes passifs exploitent le contenu même de l'image. Il existe deux familles principales d'algorithmes passifs : les algorithmes de détection d'épissage (splicing) et de copier-déplacer (copy. move). Dans une falsification par épissage, des parties de deux images ou plus sont assemblées pour former une nouvelle image. Dans une falsification par copier-déplacer, une partie d'une image est copiée et collée dans une autre partie de la même image.[2]

b. Analyse des métadonnées

Les métadonnées d'une image ou d'une vidéo fournissent des informations sur la date, l'heure et l'appareil utilisé pour capturer le contenu. Des outils comme ExifTool peuvent être utilisés pour extraire et analyser ces données.[44]

c. Outils de Recherche d'Images Inversée

La recherche d'images inversée est une technique permettant de trouver des informations sur une image en utilisant l'image elle-même comme point de départ. Voici quelques méthodes et outils pour effectuer une recherche d'images inversée :

1. Google Images : Utilisez Google Images pour effectuer une recherche inversée en téléchargeant une image ou en collant l'URL de l'image dans la barre de recherche.[45]
2. Google Lens : Disponible sur les smartphones, Google Lens permet d'effectuer des recherches d'images inversées directement depuis l'application en scannant l'image ou en appuyant longuement dessus.[46]

La recherche d'images inversée est un outil puissant pour vérifier la véracité des images en ligne et pour obtenir des informations contextuelles utiles.

8.3 Analyse des Réseaux et Diffusion

L'analyse des réseaux sociaux (Social Network Analysis - SNA) est une approche permettant de comprendre la structure et la dynamique des interactions au sein de groupes sociaux. Voici quelques éléments clés concernant l'analyse des réseaux et la diffusion :

8.3.1 Théorie des Réseaux Sociaux

C'est une méthode qui analyse les relations entre les individus ou les groupes. Elle étudie l'impact de ces relations sur les comportements, les échanges d'informations et les interactions sociales.

8.3.2 Diffusion dans les Réseaux

- Topologie des réseaux : La forme et la structure des réseaux influencent la manière dont les informations se diffusent. Une topologie bien connectée favorise une diffusion rapide et étendue.[47]
- Études empiriques : Les recherches montrent que l'irrégularité des courbes de diffusion est due à la complexité de la topologie des réseaux sociaux.[47]

8.3.3 Étude des Graphes Sociaux pour Comprendre la Propagation des Rumeurs

- Comportements et rôles dans les réseaux sociaux : Identifier les rôles des différents acteurs et leurs comportements pour comprendre les dynamiques de propagation.[48]
- Fouille de graphes et réseaux sociaux : Techniques pour analyser les interactions dans les réseaux sociaux, comprendre les schémas de diffusion des rumeurs et des innovations.[49]

8.3.4 Techniques de Détection des Anomalies et des Points de Diffusion Critique

***Définition d'anomalie* :** une valeur aberrante est un élément d'un jeu de données qui s'écarte distinctement dureste des données. Il existe trois différents types d'anomalies :

1) **Anomalies ponctuelles :** Si une donnée seule peut être considérée comme anormale en comparaison avec le reste des données, alors celle-ci est définie comme une anomalie ponctuelle.[50]

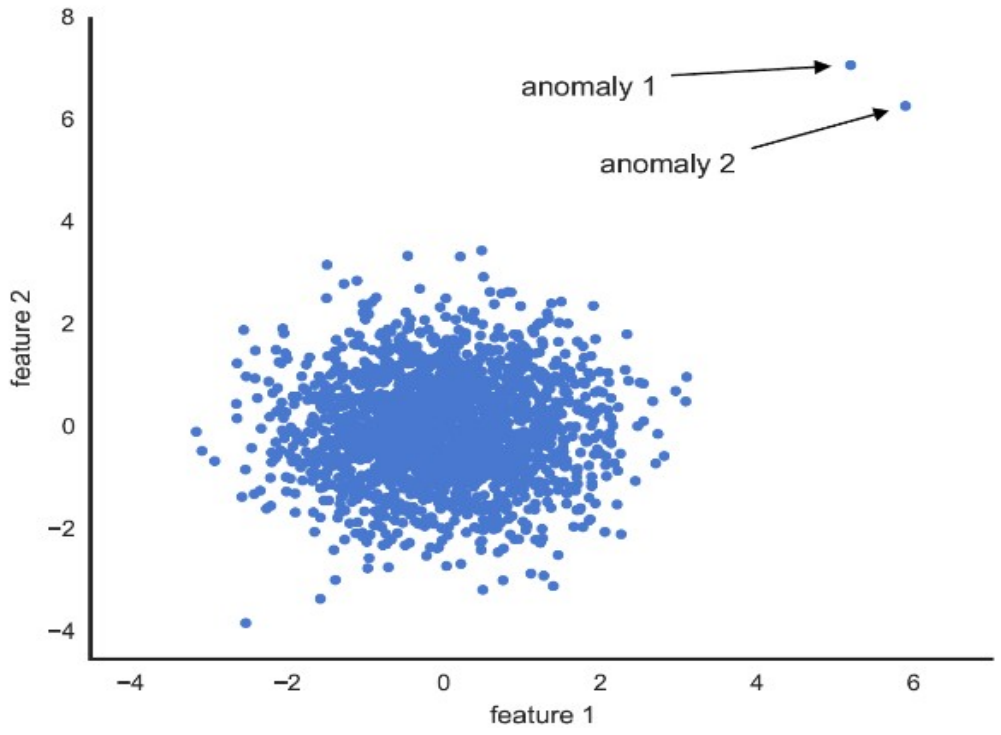

Figure 1. 10: 2 Deux anomalies ponctuelles dans un espace bi-dimensionnel.[50]

2) **Les anomalies contextuelles :**définissent une instance de données individuelles qui sont anormales dans un contexte particulier : si elles s´écartent de manière significative en fonction d'un contexte sélectionnée.[48]

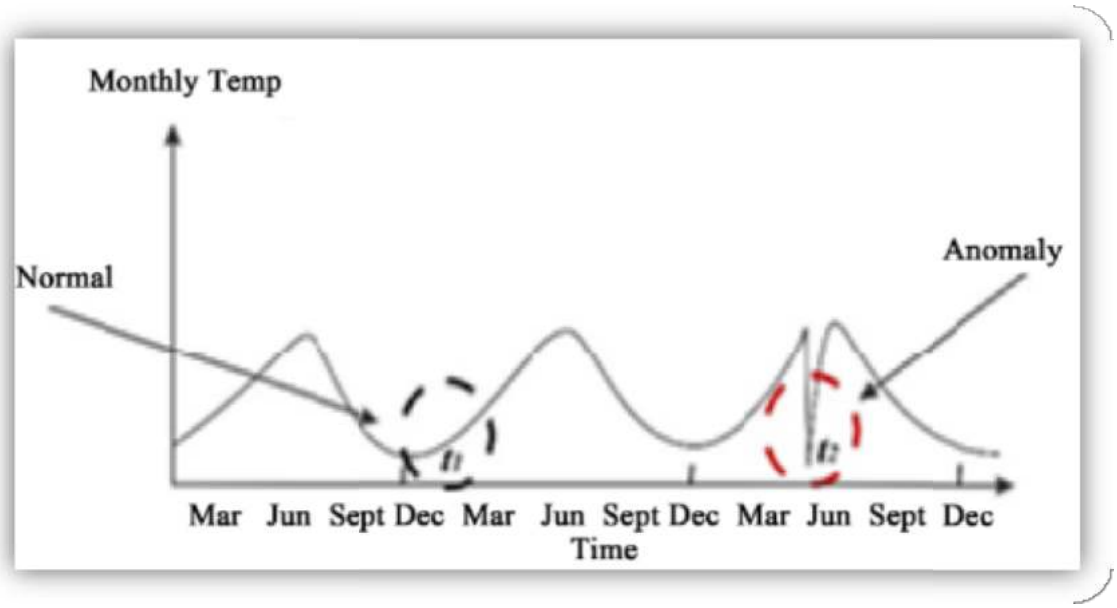

Figure 1. 11:Illustration d'une anomalie contextuelle.[48]

3) **Les anomalies collectives :** définissent un comportement collectif inattendu pendant un certain temps en comparaison avec l'ensemble des données.[48]

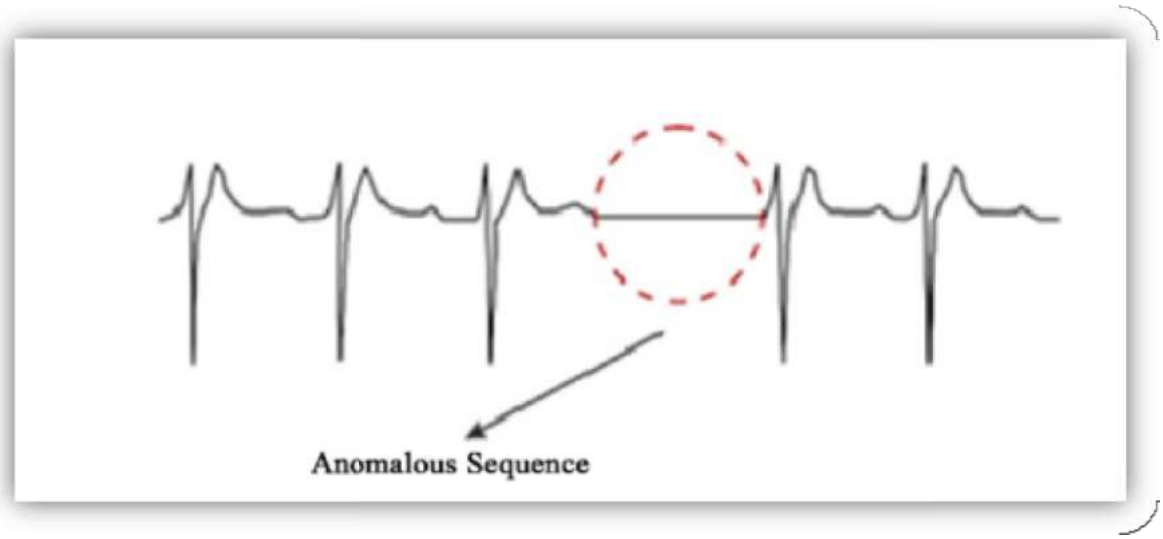

Figure 1. 12:Illustration d'une anomalie collective.[48]

Apprentissage automatique pour la détection d'anomalies **:** Utilisation de techniques de machine learning pour identifier des anomalies dans les données réseau, crucial pour détecter les points de diffusion critiques et les comportements suspects.[51]

Ces sources offrent une base solide pour explorer les méthodes et théories pertinentes à l'analyse des réseaux et à la diffusion d'informations et de rumeurs.

9. *Impact des rumeurs*

Les rumeurs ont un impact considérable sur les plans social, politique et économique :

9.1 Conséquences sociales

Les rumeurs reflètent souvent une défiance envers les institutions sociales qui valident l'information. Leur propagation peut détériorer la confiance au sein des communautés, créant des tensions et des divisions entre les groupes sociaux.[52]

En 2018, les Indiens ont reçu une vidéo qui montre le corps mutilé d'un enfant sur "WhatsApp". On ne sait pas exactement d'où la vidéo provient, ni si elle a été trafiquée.

Une voix implore les gens de la transmettre à d'autres et de rester vigilants. Les message sont poussé les parents à garder leurs enfants à l'intérieur. Les enseignants ont signalé une baisse énorme de la fréquentation scolaire.[53]

9.2 Conséquences politiques

Dans le domaine politique, les rumeurs sont fréquemment utilisées comme outils de manipulation pour influencer l'opinion publique ou discréditer des adversaires. Elles émergent dans des environnements hautement compétitifs où la vérification des faits peut être négligée, renforçant ainsi le scepticisme envers les autorités politiques.[54][55]

Plus de la moitié des Européens pourraient avoir été témoins d'une forme de désinformation promue par des acteurs russes sur les réseaux sociaux à l'approche des élections parlementaires en 2019, selon une analyse faite par 'POLITICO'. L'objectif de ces efforts, selon l'analyse, est d'amplifier les questions qui divisent les pays européens afin de saper les institutions démocratiques et de créer des tensions internes d'une manière qui, en fin de compte, favorise l'état russe.[53]

9.3 Conséquences économiques

Les rumeurs peuvent affecter les marchés financiers et la stabilité économique des entreprises. Par exemple, des rumeurs d'acquisition ou de faillite peuvent provoquer des fluctuations importantes des cours boursiers, affectant la valeur des actions et la confiance des investisseurs. Les entreprises doivent donc gérer soigneusement leur communication pour atténuer ces risques.[56][57]

10. Enjeux et Défis de la Détection des Rumeurs

10.1 Enjeux de la détection des rumeurs

- Sécurité de l'information : Les rumeurs peuvent causer des dommages significatifs, en particulier en ce qui concerne la cybersécurité. Elles peuvent mener à des failles de sécurité et nécessitent des mesures proactives pour prévenir et guérir ces impacts[58].
- Confiance des utilisateurs : Les utilisateurs des réseaux sociaux et autres plateformes en ligne doivent avoir confiance dans la véracité des informations qu'ils consomment. La propagation de fausses informations érode cette confiance, rendant crucial le développement d'outils de détection efficaces.[20]

10.2 Défis de la détection des rumeurs

- Volume et rapidité de l'information : La quantité massive d'informations diffusées sur les réseaux sociaux rend la vérification rapide et précise extrêmement difficile. Les rumeurs se propagent souvent plus vite que les vérifications ne peuvent être effectuées, compliquant ainsi la tâche des outils de détection.[59]
- Complexité des contenus : Les rumeurs peuvent être camouflées dans des contenus complexes et variés, nécessitant des outils sophistiqués pour les identifier et les vérifier efficacement.[60]

11. Description de différents travaux ultérieurs

L'objectif des recherches récentes sur la reconnaissance des rumeurs dans les médias sociaux est d'améliorer l'identification et la réduction des informations trompeuses en se concentrant sur plusieurs aspects importants. Des recherches ont montré à quel point les utilisateurs des médias sociaux sont importants pour repérer et arrêter la propagation des rumeurs, en particulier en cas d'urgence. Les rumeurs peuvent être considérablement atténuées par les efforts des utilisateurs pour vérifier les faits et signaler les informations trompeuses.

11.1 Analyse des méthodes de détection des rumeurs

11.1 .1 Le travail de D. Rajasekar et L. Robert (2023)

L'article "Context A ware Deep Ensemble Learning Model for Rumor Detection" présente un modèle de détection des rumeurs basé sur des techniques avancées d'apprentissage profond et d'ensembles. Ce modèle utilise des informations contextuelles pour améliorer la précision de la détection des rumeurs sur les réseaux sociaux.

Les méthodes du modèle utilisé :

- TF-IDF (Term Frequency-Inverse Document Frequency) pour la représentation vectorielle des textes.
- Utilisation de Word Embeddings comme Word2Vec et GloVe pour améliorer la compréhension contextuelle.
- Utilisation de classifiers tels que le Support Vector Machine (SVM) et les Arbres de Décision pour la classification finale des rumeurs.

Les résultats expérimentaux montrent que le modèle proposé surpasse les méthodes traditionnelles de détection des rumeurs en termes de précision et de rappel. Les points saillants incluent :Une précision améliorée grâce à l'intégration des informations contextuelles et une robustesse accrue grâce à l'approche d'ensemble qui combine plusieurs modèles de deep learning.[61]

11.2 Méthodes basées sur l'apprentissage automatique

11.2.1 Le travail du C. Maigrot et al (2018)

L'article "Fusion par apprentissage pour la détection de fausses informations dans les réseaux sociaux" par C. Maigrot et al. Explore l'application de techniques d'apprentissage automatique pour identifier les fausses informations circulant sur les plateformes de médias sociaux. Les auteurs se concentrent sur la fusion de plusieurs méthodes de classification afin d'améliorer la précision de la détection de fausses informations.

L'article présente une approche basée sur la fusion par apprentissage, combinant différentes techniques de classification pour obtenir de meilleurs résultats. Les principales méthodes utilisées incluent : Apprentissage Supervisé, Boosting sur des Arbres de Décision, Ensembles de Données Variés. Les classifiers employés dans l'étude comprennent : Arbres de Décision Algorithmes de Fusion.

Les résultats de l'étude montrent que la fusion par apprentissage, et en particulier l'utilisation du boosting sur des arbres de décision, permet d'améliorer de manière significative la détection des fausses informations. Cette méthode démontre une meilleure performance en termes de précision et de rappel par rapport aux classifiers individuels. Les auteurs concluent que la fusion de multiples approches de classification constitue une voie prometteuse pour lutter contre la propagation de fausses informations sur les réseaux sociaux.[20]

11.2.2 Le travail de *Borui Pan* (2022)

L'article "A Two-step Rumor Detection and Classification Method Using Machine Learning" présente une méthode innovante pour détecter et classifier les rumeurs sur les réseaux sociaux. Cette approche repose sur une analyse en deux étapes, combinant plusieurs techniques d'apprentissage automatique pour améliorer la précision et l'efficacité de la détection des rumeurs.

Les classifiers utilisés : Support Vector Machine (SVM), Réseaux de Neurones Artificiels (ANN), Arbres de Décision.

Les résultats montrent que cette méthode en deux étapes améliore significativement la précision de la détection et de la classification des rumeurs par rapport aux méthodes traditionnelles. La réduction de la dimensionnalité a permis d'augmenter l'efficacité du modèle tout en maintenant une haute précision. L'utilisation combinée des SVM pour la détection et des ANN pour la classification a montré des performances supérieures par rapport à l'utilisation isolée de chaque méthode.[62]

11.3 Techniques avancées et études de cas

11.3.1 Le travail de A.R. Pathak et ses collègues (2020)

L'article "Analyse des Techniques de Détection de Rumeurs sur les Médias Sociaux "par A.R. Pathak et ses collègues offre une revue complète des approches de pointe utilisées pour détecter les rumeurs sur les plateformes de médias sociaux. Les auteurs discutent de divers ensembles de données utilisés pour la détection de rumeurs et présentent un examen détaillé des différentes méthodologies employées pour identifier et classer les rumeurs. L'article vise à fournir des insights sur les défis et les opportunités dans le domaine de la détection de rumeurs.

L'article décrit plusieurs méthodes utilisées pour la détection de rumeurs, en se concentrant à la fois sur les techniques de machine learning traditionnelles et les approches plus récentes de deep learning. Certaines des méthodes abordées incluent :(Méthodes Basées sur les Caractéristiques, Méthodes Basées sur la Propagation, détection de la position).

Plusieurs classificateurs sont, notamment :(Machines à Vecteurs de Support (SVM), Forêts Aléatoires, Réseaux Neuronaux Profonds (DNN), Réseaux de Neurones de Graphes (GNN)). Les résultats discutés dans l'article indiquent que les approches de deep learning, en particulier celles utilisant les GNN, offrent des améliorations significatives en termes de précision de détection des rumeurs par rapport aux méthodes traditionnelles. L'utilisation de la détection de la position améliore également la capacité à vérifier les rumeurs en incorporant les opinions et interactions des utilisateurs. La combinaison des méthodes basées sur les caractéristiques, la propagation et la détection de la position offre un cadre robuste pour identifier et atténuer efficacement les rumeurs sur les plateformes de médias sociaux.[24]

11.3.2 Le travail de Chen et al (2017)

L'article "IKM at SemEval-2017 Task 8: Convolutional Neural Networks for Stance Detection and Rumor Verification" de Chen et al. (2017) présente une approche utilisant des réseaux de neurones convolutifs (CNN) pour détecter la position (stance) des tweets et vérifier la véracité des rumeurs. L'objectif principal de cette étude est de classer les tweets en fonction de leur position par rapport à une rumeur donnée et de déterminer si la rumeur est vraie ou fausse.

Les méthodes employées dans cet article incluent : Réseaux de Neurones Convolutifs (CNN), Classification Multi-Classes, Multi-Task Learning. Les classifiers principaux utilisés comprennent : Softmax Classifier, Apprentissage Multi-Tâches.

Les résultats montrent que l'approche basée sur les réseaux de neurones convolutifs améliore significativement les performances par rapport aux bases line classifiers. L'utilisation de différentes tailles de filtres permet de capturer diverses caractéristiques des tweets, et l'apprentissage multi-tâches renforce la robustesse des prédictions. Cette méthode dépasse les performances de référence dans les tâches de détection de la position et de vérification des rumeurs.[63]

11.3.3 Le travail de S. Vanitha et R. Prabahari (*2023*)

L'article "Analysis of Rumour Detection using Deep Learning Methods on Social Media" par S. Vanitha et R. Prabahari se concentre sur l'utilisation des techniques d'apprentissage profond pour la détection des rumeurs sur les réseaux sociaux.

Les méthodes de détection des rumeurs utilisées par le prétraitement des données, incluant le nettoyage, la tokenisation, et la réduction des mots à leurs formes de base. L'extraction des caractéristiques utilise des techniques comme Word2Vec et BERT pour représenter les mots. Les modèles de deep learning employés comprennent les CNN pour identifier des caractéristiques locales, les RNN pour capturer les dépendances séquentielles, et les transformers pour les relations complexes à longue distance. Les classifiers incluent des ensembles de modèles combinant plusieurs techniques pour améliorer la robustesse, ainsi que des SVM et Random Forests pour la classification finale. Les résultats montrent une précision, une robustesse et une efficacité accrues, validées par des tests sur des données de réseaux sociaux comme Twitter et Weibo.[25]

12. Comparaison entre les travaux antérieurs

Ce tableau représente une comparaison entre les travaux antérieurs :

Titre	Auteur(s)	Année	Base de données	Classifieur	Précision
Analysis of rumour detection using deep learning methodes on social media[25]	*S. Vanitha, R. Prabahari*	*2023*	*Twitter, Weibo*	*CNN, RNN*	*85%*
Context Aware Deep Ensemble Learning Model for Rumor Detection [61]	*D.Rajasekar, L. Robert*	*2023*	*Twitter*	*EGCN (TT-EGCN)*	*87%*
A Two-step Rumor Detection and Classification Method Using Machine Learning [62]	*Borui Pan*	*2022*	*COVID-19 datasets*	*Supernetwork, Random Forest, SVM*	*80%*
IKM at SemEval-2017 Task 8: Convolutional Neural Networks for Stance Detection and Rumor Verification [63]	*Yi-Chin Chen, Zhao-Yang Liu*	*2017*	*Twitter*	*CNN*	*78%*
Fusion par apprentissage pour la détection de fausses informations dans les réseaux sociaux [22]	*CédricMaigrot, EwaKijak*	*2018*	*Twitter, Facebook*	*Fusion de modèlesd'apprentissage*	*82%*
Analyse des Techniques de Détection de Rumeurs sur les Médias Sociaux [24]	*Ajeet Ram Pathak,Aditee Mahajan*	*2020*	*Twitter, Reddit*	*CNN, RNN*	*81%*
Learning to detect misleading content on twitter	*Christina Boididou, Symeon Papadopoulos,*	*2022*	*Twitter*	*BERT*	*88%*

Tableau 1: Comparaison entre les travaux ultérieurs

13. Conclusion

Dans ce chapitre, nous avons présenté un état de l'art des différents aspects et techniques de la détection des rumeurs dans les réseaux sociaux. Les avancées technologiques et les approches innovantes jouent un rôle crucial dans la lutte contre la désinformation, assurant ainsi un environnement en ligne plus sûr et plus fiable. Dans le deuxième chapitre, nous aborderons les principales méthodes et techniques de l'apprentissage automatique (« Machine Learning»).

Chapitre 02

Système de détection des rumeurs basé sur l'apprentissage automatique

1. Introduction

La diffusion rapide des rumeurs sur les médias sociaux représente un défi majeur pour la société moderne. Devant ce phénomène, l'apprentissage automatique (Machine Learning) émerge comme une solution prometteuse pour détecter et contrer ces fausses informations. Ce chapitre examine les diverses approches de l'apprentissage automatique utilisées pour détecter les rumeurs, en soulignant leur efficacité et leurs utilisations concrètes.

2. Apprentissage automatique

Apprentissage automatique est un sous-domaine de l'intelligence artificielle. Cela fait référence à un ensemble de techniques mathématiques et de transformation de données permettant à un logiciel d'apprendre à partir de données historiques afin de s'ajuster automatiquement pour parfaire l'exécution d'une tâche spécifique.[64]

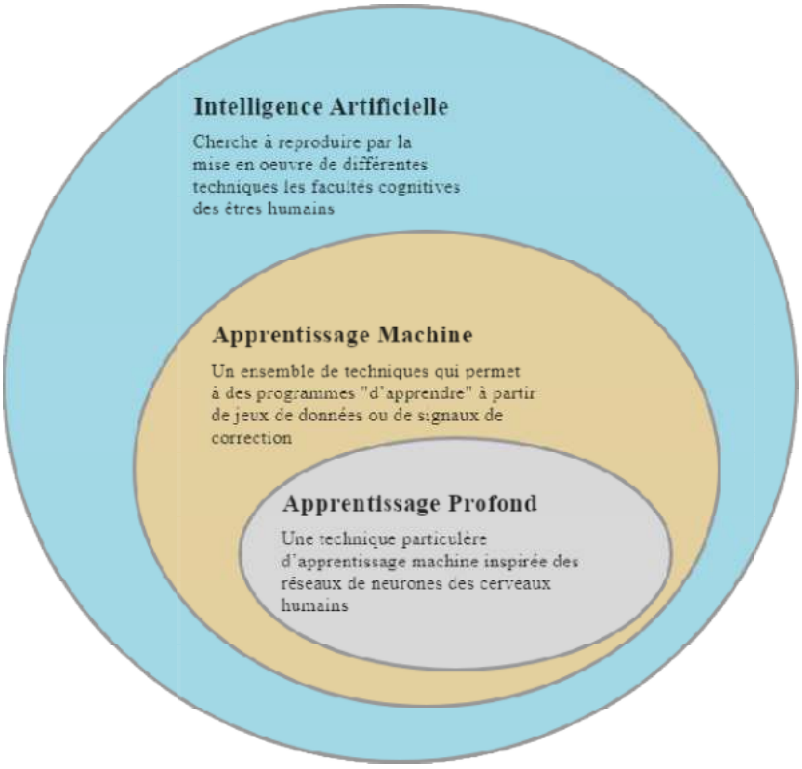

Figure 2. 1:L'apprentissage automatique et l'apprentissage profond.[64]

3. Apprentissage profond

L'apprentissage profond (Deep Learning), l'une des sous-disciplines les plus prometteuses dans le domaine de l'apprentissage machine, réfère aux différentes techniques d'apprentissage machine qui se font par le biais des réseaux de neurones artificiels profonds (Deep Neural Network).[64]

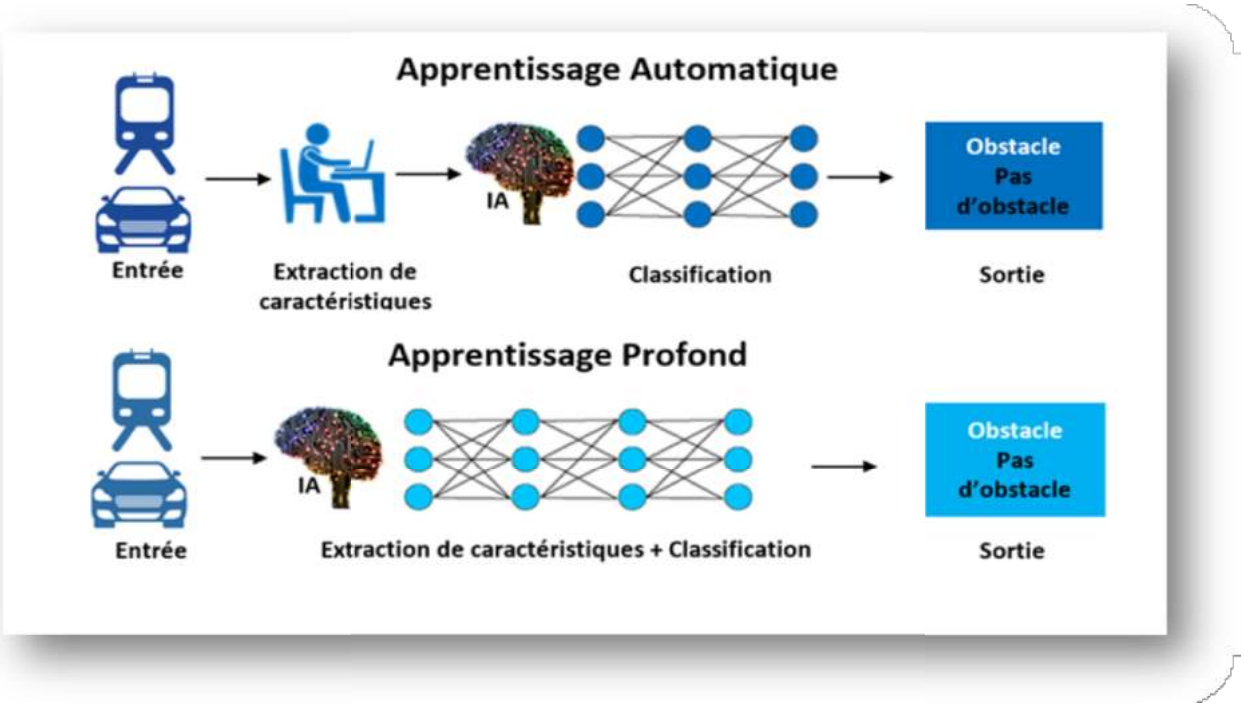

Figure 2. 2: Différence entre l'apprentissage automatique et l'apprentissage profond.[65]

3. Apprentissage profond vs Apprentissage automatique

Comparaison entre l'apprentissage automatique et l'apprentissage profond dans différents facteurs :

Facteurs	**Apprentissage profond**	**Apprentissage automatique**
Données d'exigence	Nécessite des données volumineuses	Peut être réglé de différentes manières
Précision	Fournir une grande précision	Donne moins de précision
Temps d'entrainement	Prend du temps à s'entraîner	Prend moins de temps à s'entraîner
Dépendance matérielle	Nécessite un GPU pour s'entraîner correctement	Entraîner sur GPU
Réglage des hyperparamètres	Peut être réglé de différentes manières	Capacité de réglage limitées

Figure 2. 3:Apprentissage profond vs Apprentissage automatique.[66]

5. Approches traditionnelles de détection des rumeurs :

5.1 La machine à vecteurs de support (SVM)

La machine à vecteurs de support (ou SVM pour Support Vector Machine en anglais) est un algorithme d'apprentissage automatique supervisé. L'objectif principal de la SVM est de créer une frontière de décision qui sépare les données en deux classes différentes de manière optimale. Cette frontière est appelée hyperplan, qui est déterminé en maximisant la marge entre les deux classes, c'est-à-dire la distance entre les points les plus proches de chaque classe. [67]

La figure 2.4 montre un diagramme schématique du processus de classification utilisant la méthode SVM, qui comprend un processus de formation et de test. Les noyaux SVM utilisent des polynômes et une fonction de base radiale (RBF). Ces choix sont considérés comme un mécanisme de classification efficace puisque ces noyaux montrent une séparation non linéaire entre les classes. Pour garantir une grande précision de prédiction, le processus de validation croisée est effectué. Cela permet d'obtenir les meilleurs paramètres du noyau.[68]

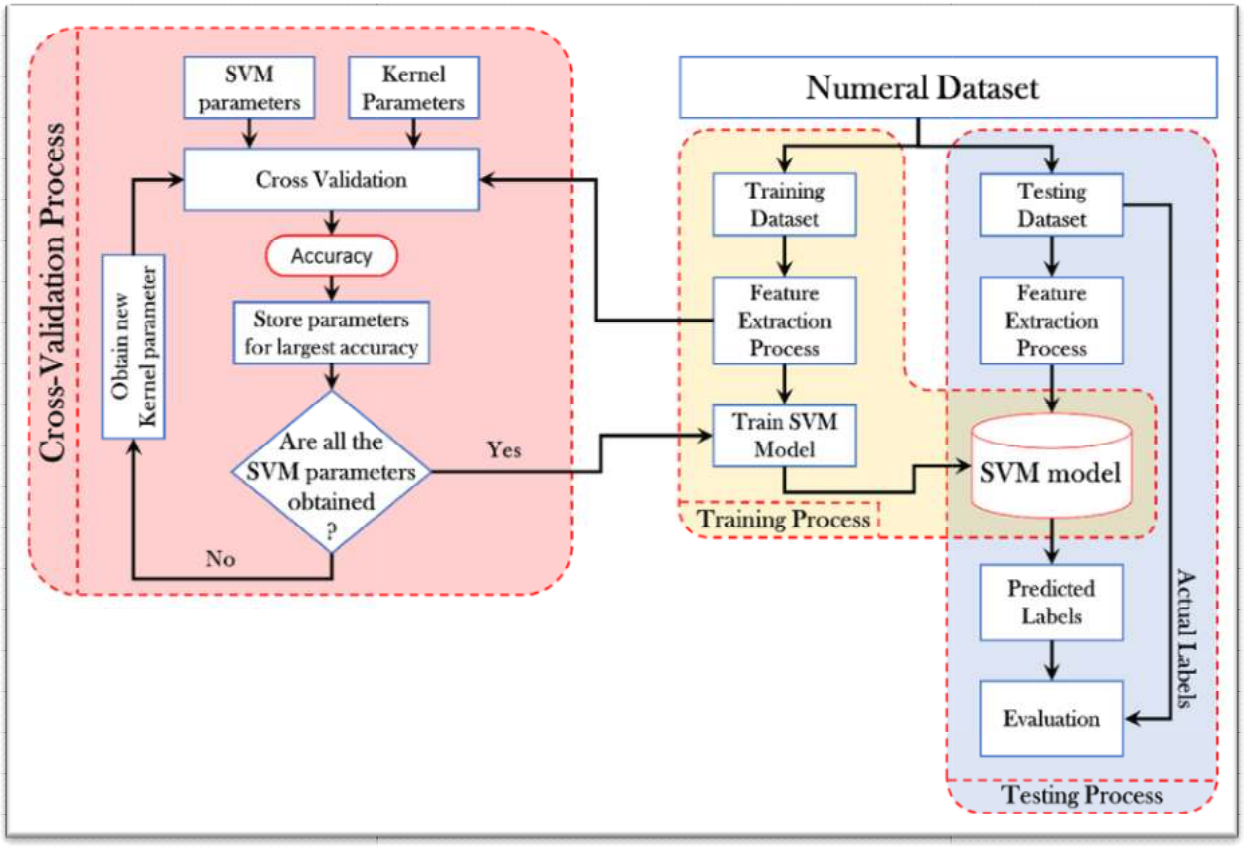

Figure 2. 4:Diagramme schématique du processus de classification des SVM.[68]

L'algorithme SVM n'est pas adapté aux grands ensembles de données et ne fonctionne pas très bien lorsque l'ensemble de données a plus de bruit, c'est-à-dire que les classes cibles se chevauchent. Comme le classificateur de vecteur de support fonctionne en plaçant des points de données, au-dessus et en dessous de l'hyperplan de classification, il n'y a pas d'explication probabiliste pour la classification.[68]

5.2 K-means

L'algorithme Kmeans est un algorithme itératif qui tente de partitionner l'ensemble de données en K sous-groupes (clusters) distincts et non superposés prédéfinis, où chaque point de données appartient à un seul groupe. Il essaie de rendre les points de données intra-cluster aussi similaires que possible tout en gardant les clusters aussi différents que possible. [69]

Il attribue des points de données à un cluster de sorte que la somme de la distance au carré entre les points de données et le centroïde du cluster (moyenne arithmétique de tous les points de données appartenant à ce cluster) soit au minimum. Moins nous avons de variation au sein des clusters, plus les points de données sont homogènes (similaires) au sein du même cluster. Voir la figure [69]

Le fonctionnement de l'algorithme kmeans est le suivant [69] :

1-Spécifiez le nombre de clusters K.

2-Initialisez les centroïdes en mélangeant d'abord l'ensemble de données, puis en sélectionnant au hasard K points de données pour les centroïdes sans remplacement.

3-Continuez à itérer jusqu'à ce qu'il n'y ait aucun changement dans les centroïdes. c'està-dire que l'affectation des points de données aux clusters ne change pas.

4-Calculez la somme de la distance au carré entre les points de données et tous les centroïdes.

5-Attribuez chaque point de données au cluster le plus proche (centre de gravité).

6-Calculez les centroïdes des clusters en prenant la moyenne de tous les points de données appartenant à chaque cluster.

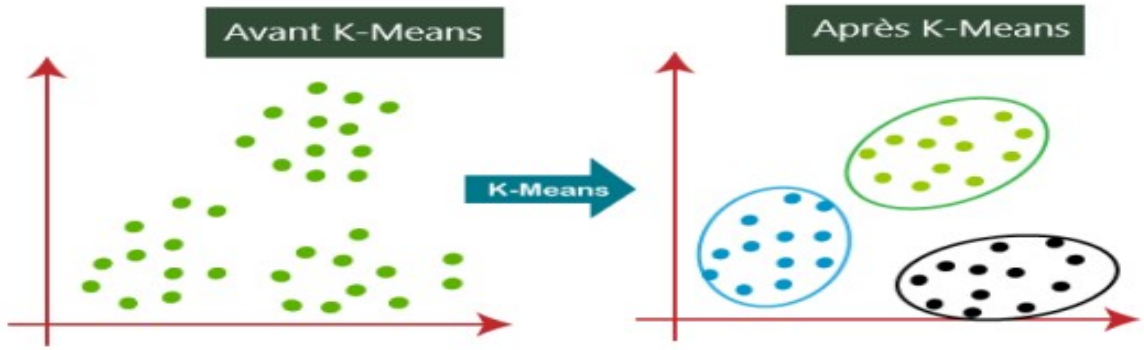

Figure 2. 5:Exemple de K-means.[69]

5.3 Naive Bayes

Naive Bayes est une méthode d'apprentissage automatique traditionnelle, qui a été largement appliquée sur le terrain de classification de texte. La formule de Naive Bayes est la suivante:[62]

$$p(A|\mathrm{B}) = \frac{\mathrm{P(AB)}}{\mathrm{P(B)}} = \frac{\mathrm{P(B|A)} * \mathrm{P(A)}}{\mathrm{P(B)}}$$

Où cette formule représente la probabilité que l'événement A se produise lorsque l'événement B s'est déjà produit. Dans la tâche de classification des rumeurs, l'événement A correspond à la catégorie de cette information et l'événement B aux caractéristiques de l'information, de sorte que l'objectif est de trouver la plus grande probabilité de catégorie, ce qui est également le résultat de la classification. Naive Bayes présente des avantages en termes de vitesse de calcul et de bonnes performances dans la classification de texte.[62]

6. Apprentissage profond pour la détection des rumeurs

6.1 Les réseaux de neurones:

Les réseaux de neurones artificiels, une famille de techniques d'apprentissage machine, permettent de capturer des patrons par un processus de propagation de données à travers un ensemble d'unités liées entre elles (neurones) et de mécaniques automatiques de correction d'erreur (p. ex. corriger l'erreur de prédiction du réseau de neurones). Les neurones de ces réseaux sont généralement organisés en couches ; l'information se propage d'une couche à une autre en commençant par la couche d'entrée (input layer) et en finissant par la couche de sortie (output layer), les couches du milieu étant les couches intermédiaires.[64]

Un neurone est une entité qui permet de recevoir les données d'entrée et produire une donnée en sortie. Un ensemble de neurones peut être vu comme une couche (entrées, cachées ou sorties).[66]

6.2 Les types des réseaux de neurones artificiels

Les réseaux de neurones artificiels se déclinent en plusieurs types, chacun ayant des caractéristiques et des applications spécifiques :

6.2.1 Perceptrons multicouches (MLP)

Le perceptron multicouche (MLP) est un type de réseau de neurones artificiels qui comprend plusieurs couches de neurones : une couche d'entrée, une ou plusieurs couches cachées, et une couche de sortie. Cette structure permet au MLP de surmonter les limitations des perceptrons simples, qui ne peuvent traiter que des données linéaires. Le MLP est particulièrement utile

pour traiter des données non linéaires, car il peut modéliser des relations complexes entre les variables d'entrée et la sortie.[70]

Dans un MLP, chaque neurone de la couche cachée est connecté à tous les neurones de la couche suivante. Les connexions entre les neurones sont pondérées, et ces poids sont ajustés au cours du processus d'apprentissage, généralement par rétro propagation. Cette technique permet au réseau de minimiser l'erreur en ajustant les poids en fonction de la différence entre la sortie prédite et la sortie réelle.[71]

$$Y=\varphi_{out}\left(\sum_i w_i^{(2)} h_i^{(2)} + b^{(2)}\right)$$

$$\forall i, h_i^{(2)}=\varphi\left(\sum_j w_{ij}^{(1)} h_j^{(1)} + b_i{}^{(1)}\right)$$

$$\forall i, h_i^{(1)}=\varphi\left(\sum_j w_{ij}^{(0)} x_j + b_i{}^{(0)}\right)$$

Cette représentation graphique correspond au modèle :[71]

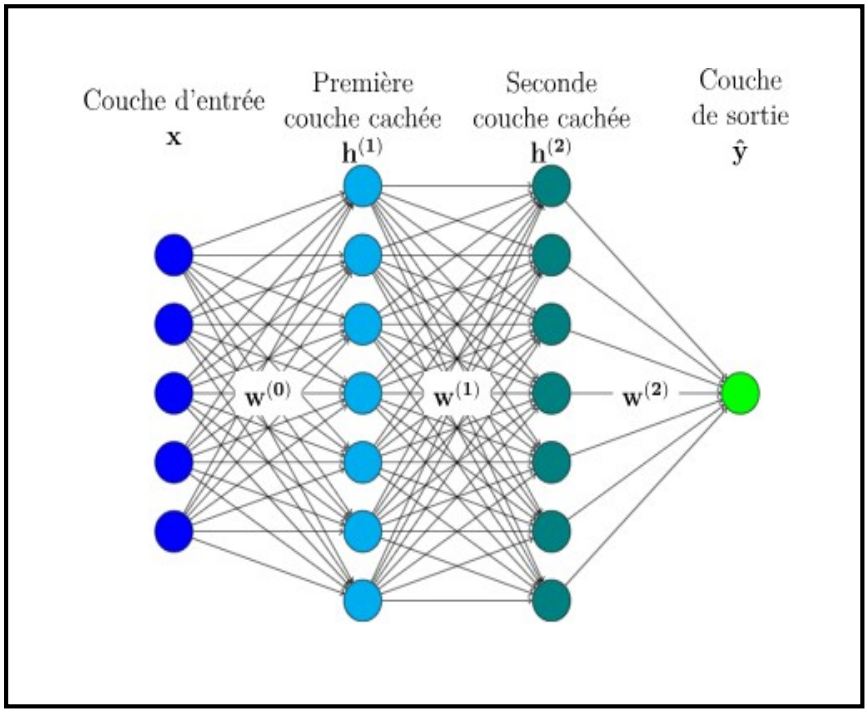

Figure 2. 6:Perceptrons multicouches.[71]

Les fonctions d'activation, telles que la fonction sigmoïde ou ReLU, jouent un rôle essentiel dans les MLP en introduisant la non-linéarité nécessaire pour modéliser des données complexes. Sans ces fonctions d'activation, le MLP ne pourrait pas capturer les nuances des relations non linéaires entre les entrées et les sorties.[72]

6.2.2 Les réseaux de neurones convolutifs (CNN)

Les réseaux de neurones convolutifs (CNN) sont une catégorie de réseaux de neurones artificiels largement utilisés pour l'analyse d'images et de vidéos. Ces réseaux sont particulièrement efficaces dans la reconnaissance d'objets, la classification d'images, et d'autres tâches de vision par ordinateur.

Structure et Fonctionnement

Les réseaux de neurones convolutifs (CNN) sont composés de plusieurs types de couches qui travaillent ensemble pour extraire des caractéristiques des données d'entrée, principalement des images. La structure typique inclut :

1. Couches de Convolution : Elles appliquent des filtres (ou noyaux) à l'image d'entrée pour détecter des motifs locaux comme les bords et les textures. Chaque filtre produit une carte de caractéristiques qui met en évidence un certain type de motif.[73]

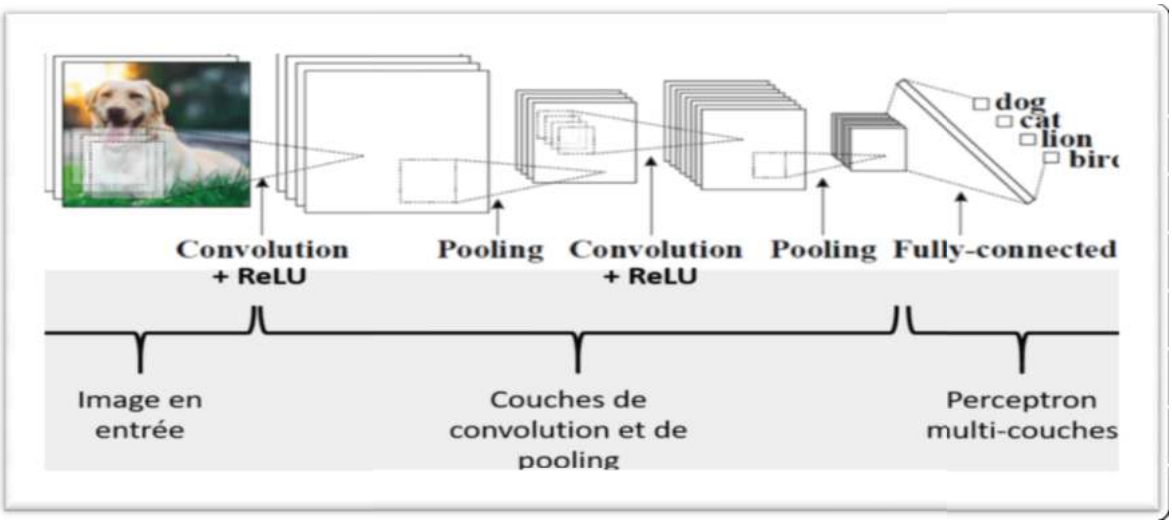

Figure 2. 7:Réseaux de neurones convolutifs.[74]

2. Couches de Pooling : Après la convolution, les couches de pooling réduisent la dimensionnalité des cartes de caractéristiques, tout en conservant les informations les plus importantes. Le pooling max est souvent utilisé, prenant la valeur maximale dans une région spécifique pour diminuer la sensibilité aux translations.[75]

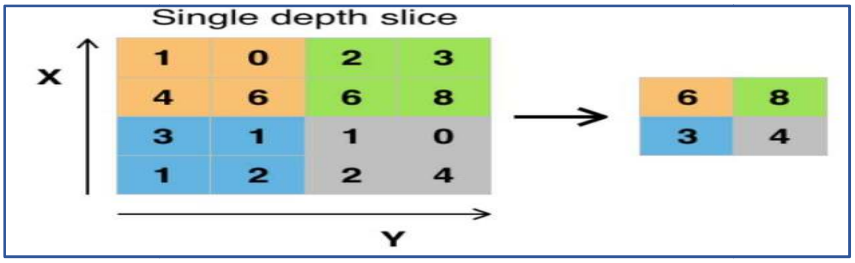

Figure 2. 8:Réseaux de neurones convolutifs – Le Pooling.[74]

3. Couches Entièrement Connectées : En fin de réseau, ces couches combinent les cartes de caractéristiques en une sortie qui représente la classification finale ou une autre tâche spécifique.[74]

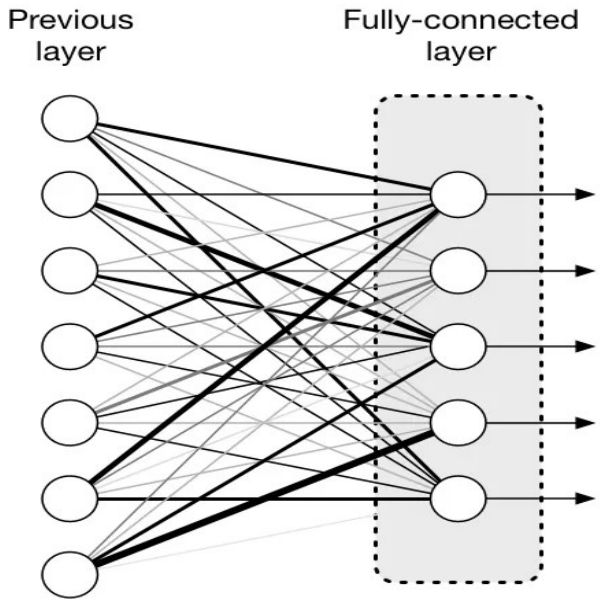

Figure 2. 9:Couche entièrement connectée (Fullyconnected).[76]

Applications

Les CNN sont principalement utilisés dans des domaines tels que la reconnaissance d'images, la détection d'objets, la segmentation d'images, et même le traitement du langage naturel. Ils sont particulièrement efficaces pour les tâches de vision par ordinateur en raison de leur capacité à apprendre directement des données brutes, sans besoin de caractéristiques définies manuellement.[77]

Limites

Cependant, les CNN ont aussi des limites :

Dépendance aux Données : Ils nécessitent un grand volume de données annotées pour un apprentissage efficace.

Complexité Computationnelle : L'entraînement des CNN, surtout pour les réseaux profonds, est extrêmement coûteux en termes de temps de calcul et de ressources matérielles.[76]

6.2.3 Réseaux de neurones récurrents (RNN)

Les réseaux de neurones récurrents (RNN) sont un type de NN largement utilisé pour effectuer le processus d'analyse de séquence car le RNN est conçu pour extraire les informations contextuelles en définissant les dépendances entre les différents horodatages.

RNN se compose de nombreuses couches récurrentes successives, et ces couches sont séquentiellement modélisées afin de cartographier la séquence avec d'autres séquences.

RNN a une forte capacité à capturer les données contextuelles de la séquence Cependant, les indices contextuels dans la structure du réseau sont stables et sont efficacement utilisés pour réaliser le processus de classification des données.[69]

La caractéristique principale et la plus importante du RNN est son état caché, qui mémorise certaines informations sur une séquence. L'état est également appelé état mémoire car il mémorise l'entrée précédente du réseau. Il utilise les mêmes paramètres pour chaque entrée car il exécute la même tâche sur toutes les entrées ou couches cachées pour produire la sortie. Cela réduit la complexité des paramètres, contrairement aux autres réseaux neuronaux.[78]

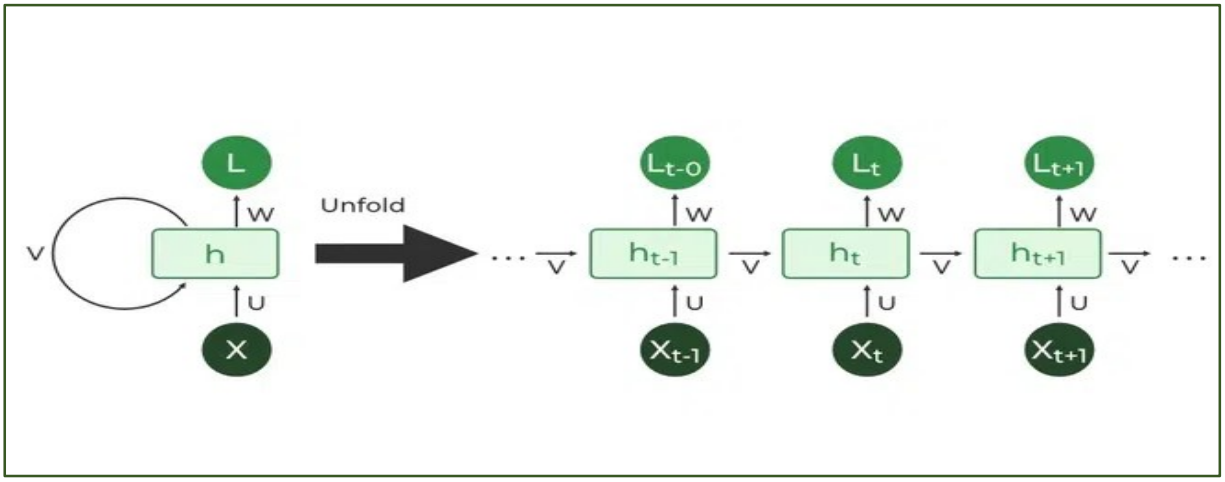

Figure 2. 10:Réseau neuronal récurrent.[78]

Les composants du réseau neuronal récurrent RNN :

- Couche d'entrée : cette couche reçoit l'élément initial des données de séquence.[79]
- Couche cachée : Au cœur du RNN, la couche cachée contient un ensemble de neurones interconnectés. Chaque neurone traite l'entrée actuelle ainsi que les informations de l'état de la couche cachée précédente.[79]
- Fonction d'activation : cette fonction introduit la non-linéarité dans le réseau, lui permettant d'apprendre des modèles complexes. Elle transforme l'entrée combinée de la couche d'entrée actuelle et de l'état de la couche cachée précédente avant de la transmettre.[79]

- Types de fonctions d'activation :[66]

  Sigmoïde : produit une courbe en forme de S. Bien que de nature non linéaire, il ne tient toutefois pas compte des légères variations des entrées, ce qui entraîne des résultats similaires.

  Fonctions de tangente hyperbolique (tanh) : est une fonction lente à converger.

  Unité linéaire rectifiée (ReLu) : Cette fonction converge plus rapidement, optimise et produit la valeur souhaitée plus rapidement. Cette fonction est utilisée dans les couches cachées.

  Softmax : utilisée dans la couche de sortie dans le cas d'une classification multiclasse.
- Couche de sortie : la couche de sortie génère la prédiction du réseau en fonction des informations traitées.[79]
- Connexion récurrente : une des principales différences entre les réseaux neuronaux rénaux est la connexion récurrente au sein de la couche cachée. Cette connexion permet au réseau de transmettre les informations d'état cachées (la mémoire du réseau) à l'étape temporelle suivante.[79]

Types de réseau neuronal récurrent RNN :

Il existe quatre types de RNN en fonction du nombre d'entrées et de sorties du réseau.

1. Un à un : Ce type de RNN se comporte de la même manière que n'importe quel réseau neuronal simple. Il est également connu sous le nom de réseau neuronal Vanilla. Dans ce réseau neuronal, il n'y a qu'une seule entrée et une seule sortie.[78]
2. Un à plusieurs : Une forme de RNN connue sous le nom d'un-à-plusieurs produit plusieurs sorties à partir d'une seule entrée. Il accepte une taille d'entrée fixe et génère une série de données.[80]
3. Plusieurs à un : Une série d'entrées est nécessaire pour qu'une sortie fixe soit affichée. Lorsqu'une seule sortie provenant de nombreuses unités d'entrée ou d'une série d'entre elles est requise, plusieurs-à-un est utilisé. Une illustration typique de ce type de réseau neuronal récurrent dans l'apprentissage en profondeur est l'analyse des sentiments.[80]
4. Plusieurs à plusieurs :pour plusieurs vers plusieurs, il existe de nombreuses options. Deux entrées donnent trois sorties. Les systèmes de traduction automatique, tels que les systèmes de traduction de l'anglais vers le français ou vice versa, utilisent des réseaux plusieurs vers plusieurs.[79]

Architecture RNN :

Réseaux neuronaux récurrents bidirectionnels (BRRN) :

Cette variante traite les données dans les deux sens. Cela lui permet de capturer le contexte des deux côtés d'une séquence, ce qui est utile pour des tâches comme l'analyse des sentiments où la compréhension de la phrase entière est cruciale.[79]

Mémoire à long terme et à court terme (LSTM) :

LSTM est une architecture RNN populaire, qui a été introduite par Sepp Hochreiter et Juergen Schmidhuber comme solution au problème du gradient nul. Dans leur article (lien externe à ibm.com), ils s'efforcent de résoudre le problème des dépendances à long terme. Autrement dit, si l'état précédent qui influence la prédiction actuelle ne se situe pas dans un passé récent, le modèle RNN peut ne pas être en mesure de prédire avec précision l'état actuel.

Pour remédier à cela, les LSTM ont des « cellules » dans les couches cachées du réseau neuronal, qui ont trois portes : une porte d'entrée, une porte de sortie et une porte d'oublie.

Ces portes contrôlent le flux d'informations nécessaire pour prédire la sortie dans le réseau. Par exemple, si des pronoms de genre, tels que « elle », ont été répétés plusieurs fois dans des phrases précédentes, vous pouvez les exclure de l'état de la cellule.[81]

Unités récurrentes fermées (GRU) :

Un GRU est similaire à un LSTM car il permet également de résoudre le problème de mémoire à court terme des modèles RNN. Au lieu d'utiliser un « état de cellule » pour réguler les informations, il utilise des états cachés et au lieu de trois portes, il en a deux : une porte de réinitialisation et une porte de mise à jour. Similaires aux portes des LSTM, les portes de réinitialisation et de mise à jour contrôlent la quantité et la nature des informations à conserver.[81]

Avantages et inconvénients du réseau neuronal récurrent :

- *Avantages:[78]*

-Un RNN mémorise chaque élément d'information au fil du temps. Il est utile dans la prévision des séries chronologiques uniquement en raison de sa capacité à mémoriser également les entrées précédentes. C'est ce qu'on appelle la mémoire à long terme.

-Les réseaux neuronaux récurrents sont même utilisés avec des couches convolutives pour étendre le voisinage effectif des pixels.

- *Inconvénients :[78]*

-Problèmes de disparition et d'explosion de gradient.

-Former un RNN est une tâche très difficile.

-Il ne peut pas traiter de très longues séquences si vous utilisez tanh ou relu comme fonction d'activation.

7. Conclusion

Les réseaux de neurones apportent une contribution précieuse à la lutte contre la désinformation, mais ils doivent être intégrés dans des systèmes plus larges de vérification des faits et d'éducation numérique pour être pleinement efficaces. Dans ce qui suit, nous présenterons la conception de notre projet.

Chapitre 03

Modélisation et Mise en œuvre d'un système de détection des rumeurs.

1. Introduction

Ce chapitre présente une approche pour élaborer et mettre en place un système de détection des rumeurs.

La première section de ce chapitre souligne l'importance cruciale du prétraitement dans l'amélioration de la précision de la détection des rumeurs. Des techniques telles que la Tokenisation et Séquencement des données textuelles et le padding ont été mises en avant. La section suivante a présenté la conception de notre modèle, avec un choix d'architecture populaire le plus précise, que nous avons entraînée en utilisant la base de données kaggle via le langage de programmation Python. Ensuite, Nous avons comparé les types optimiseurs en évaluant leurs performances respectives et une comparaison entre le nombre d'époques.

Enfin, La dernière section de ce chapitre on a une discussion du résultat obtenu.

2. L'environnement du développement

Pour le développement de notre système; nous avons utilisé les outils matériels et logiciels suivants :

2.1 Outils matériels

Pour réaliser notre projet, nous avons utilisé un PC portable marque HP avec les caractéristiques suivantes :

- Processeur : Intel(R) Core (TM) i5-8250U CPU @ 1.60GHz 1.80 GHz.
- Capacité Mémoire (RAM) : 8.00Go.
- Capacité disque dur : 256Go.
- Type de système : Système d'exploitation 64bits, processeur x64.
- Spécifications du Windows : Windows 10 Professionnel.
- nous avons choisi le langage de programmation Python, et les différents bibliothèques (Tensorflow, Keras, Numpy…).

2.2 Outils Logiciels

- Visual Studio Code

Visual Studio Code est un éditeur de code source léger et disponible pour Windows, macOS et Linux. Il est doté d'une prise en charge intégrée de JavaScript, TypeScript et Node.js et dispose d'un d'extensions pour d'autres langages et environnements d'exécution (tels que C++, C#, Java, Python, PHP, Go, .NET). [82]

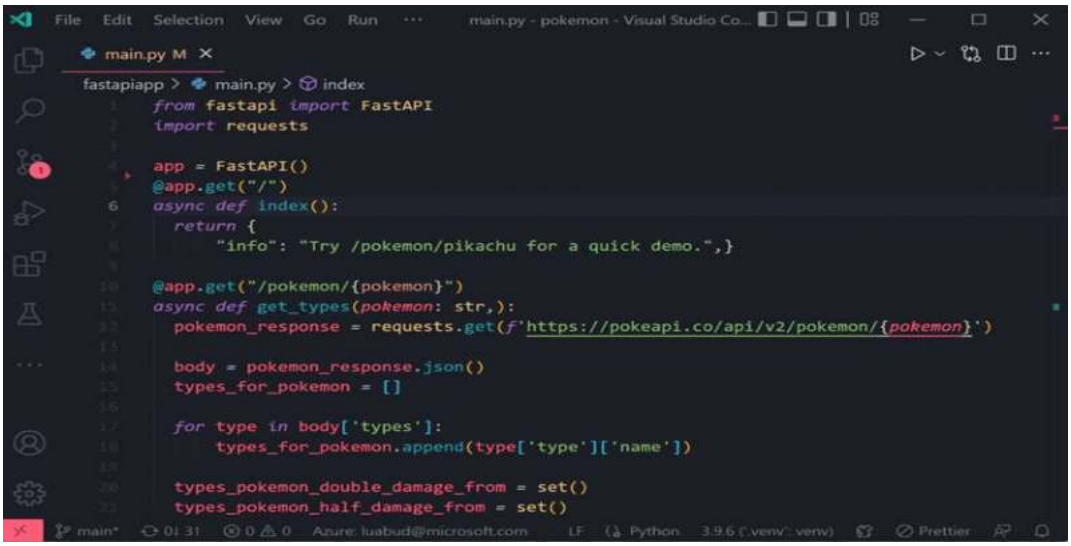

Figure 3. 1:L'environnement de travail de Visual studio code.[82]

- Google Colab (Colaboratory)

Google Colab, ou "Colaboratory," est un service cloud gratuit basé sur les notebooks Jupyter. Il permet aux utilisateurs d'écrire et d'exécuter du code Python directement depuis leur navigateur sans nécessiter d'installation. Colab est particulièrement apprécié pour sa capacité à fournir un accès facile à des ressources informatiques puissantes, telles que les GPU et TPU, sans frais supplémentaires, ce qui en fait un outil idéal pour l'apprentissage automatique, la science des données et les réseaux neuronaux.[83]

Les principales caractéristiques de Google Colab incluent :[83]

- Aucune configuration requise : Fonctionne entièrement dans le cloud.
- Accès à des ressources de calcul : Utilisation gratuite de GPU et TPU pour des tâches complexes.
- Intégration avec Google Drive : Partage et collaboration simplifiés via Google Drive.
- Environnement interactif : Permet d'inclure du texte, des images, des graphiques, et même du LaTeX dans les notebooks.

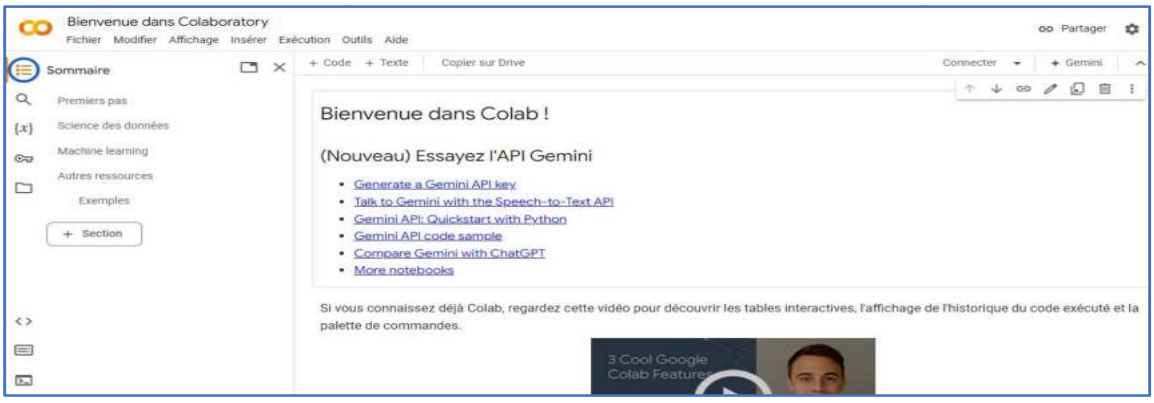

Figure 3. 2:Google Colab (Colaboratory).

- Python 3.9.13

  Python est un langage de programmation interprété, orienté objet et de haut niveau, doté d'une sémantique dynamique. Ses structures de données intégrées de haut niveau, combinées à un typage et une liaison dynamique, le rendent très attractif pour le développement rapide d'applications, ainsi que pour une utilisation en tant que langage de script ou de liaison pour connecter des composants existants entre eux. La syntaxe simple et facile à apprendre de Python met l'accent sur la lisibilité et réduit donc le coût de maintenance du programme. Python prend en charge les modules et les packages, ce qui encourage la modularité du programme et la réutilisation du code. L'interpréteur Python et la vaste bibliothèque standard sont disponibles sous forme source ou binaire gratuitement pour toutes les principales plates-formes et peuvent être distribués gratuitement.[84]

2.3 Les bibliothèques

- **Tensorflow**

TensorFlow est une bibliothèque open-source très populaire, développée par Google, utilisée pour construire et déployer des modèles d'apprentissage automatique (machine learning) et d'intelligence artificielle. Elle offre un cadre flexible pour la création de réseaux de neurones et autres modèles d'apprentissage, s'adaptant à divers environnements, tels que le cloud, les appareils mobiles, le web, et les ordinateurs de bureau. Cette plateforme end-to-end est dotée d'une architecture qui facilite l'intégration de nombreuses autres bibliothèques et extensions, permettant ainsi aux utilisateurs de créer des modèles complexes avec des données structurées sous forme de graphes ou d'autres formats avancés.[85][86]

TensorFlow propose également des interfaces conviviales pour les développeurs, telles que TensorFlowKeras, qui simplifient la création et l'entraînement de modèles d'apprentissage profond. Il dispose également d'une vaste communauté de développeurs qui partagent des ressources, des exemples de code et des modèles pré-entraînés, ce qui facilite l'apprentissage et l'utilisation de cette bibliothèque.[86]

- **Keras**

Keras est une API d'apprentissage en profondeur écrite en Python et capable de s'exécuter sur JAX ,TensorFlow ou PyTorch .[87]

Keras est :[87]

Simple :Keras réduit la charge cognitive du développeur pour vous permettre de vous concentrer sur les parties du problème qui comptent vraiment.

Flexible :Keras adopte le principe de divulgation progressive de la complexité : les flux de travail simples doivent être rapides et faciles, tandis que les flux de travail arbitrairement avancés doivent être possibles via un chemin clair qui s'appuie sur ce que vous avez déjà appris.

Puissant : Keras offre des performances et une évolutivité de premier ordre : il est utilisé par des organisations telles que la NASA, YouTube ou Waymo.

- **Numpy**

  NumPy est le package fondamental pour le calcul scientifique en Python. Il s'agit d'une bibliothèque Python qui fournit un objet tableau multidimensionnel, divers objets dérivés (tels que des tableaux masqués et des matrices) et un assortiment de routines pour des opérations rapides sur des tableaux, notamment mathématiques, logiques, manipulation de formes, tri, sélection, E/S, transformées de Fourier discrètes, algèbre linéaire de base, opérations statistiques de base, simulation aléatoire et bien plus encore.[88]

- **Matplotlib**[89]

  Matplotlib est un outil de visualisation de données utilisé dans le langage Python. On l'utilise principalement pour réaliser des graphiques en 2D, mais elle peut aussi servir à réaliser des graphiques en 3D avec certaines spécifications.

-Créez des tracés de qualité de publication.

-Créez des figures interactives qui peuvent zoomer, se déplacer, se mettre à jour.

- le style visuel et la mise en page.

-Exporter vers de nombreux formats de fichiers.

- dans Jupyter Lab et les interfaces utilisateur graphiques.

-Utilisez une riche gamme de packages tiers construits sur Matplotlib.

3. Base de données utilisée

La collection SMS Spam est un ensemble de messages SMS étiquetés qui ont été collectés pour la recherche de spams SMS. Elle contient un ensemble de messages SMS en anglais de 5 574 messages, étiquetés selon qu'ils sont légitimes ou indésirables.

Les fichiers contiennent un message par ligne. Chaque ligne est composée de deux colonnes : v1 contient le libellé (ham ou spam) et v2 contient le texte brut.[90]

Ce corpus a été collecté à partir de sources gratuites ou accessibles gratuitement à des fins de recherche sur Internet :[90]

- Une collection de 425 messages SMS de spam a été extraite manuellement du site Web Grumbletext. Il s'agit d'un forum britannique dans lequel les utilisateurs de téléphones portables font des réclamations publiques sur les messages SMS de spam, la plupart d'entre eux sans signaler le message de spam reçu. L'identification du texte des messages de spam dans les réclamations est une tâche très difficile et longue, et elle implique une analyse minutieuse de centaines de pages Web. [90]
- Un sous-ensemble de 3 375 messages SMS de radioamateur choisis au hasard dans le corpus SMS de la NUS (NSC), qui est un ensemble de données d'environ 10 000 messages légitimes collectés pour la recherche au département d'informatique de l'université nationale de Singapour. Les messages proviennent en grande partie de Singapouriens et principalement d'étudiants fréquentant l'université. Ces messages ont été collectés auprès de volontaires qui ont été informés que leurs contributions allaient être rendues publiques. [90]
- Une liste de 450 messages SMS de radioamateur collectés à partir de la thèse de doctorat de Caroline Tag.[90]
- Enfin, ils ont incorporé le SMS Spam Corpus v.0.1 Big. Il contient 1 002 messages SMS ham et 322 messages spam.[90]

Voici quelques messages de la base de données :

Courrier indésirable	Légitimes
GAGNANT !! En tant que client du réseau, vous avez été sélectionné pour recevoir une récompense de 900 € ! Pour réclamer...	Je vais bientôt rentrer à la maison et je ne veux plus parler de ça ce soir, ok ? J'ai assez pleuré...
Merci pour votre abonnement à Ringtone UK, votre mobile sera facturé 5 €/mois. Veuillez confirmer en r...	Bonjour ! Comment vas-tu et comment s'est passé ton samedi ? Je t'envoyais juste un message pour savoir si tu avais décidé de faire quelque chose...
SMS.ac Sptv : Les Devils du New Jersey et les Red Wings de Détroit jouent au hockey sur glace. Correct ou incorrect ?...	Désolé, je vous appellerai plus tard lors de la réunion.

Tableau 2 : Quelques messages des bases de données.

4. *Expérimentation*

4.1 Le modèle propose

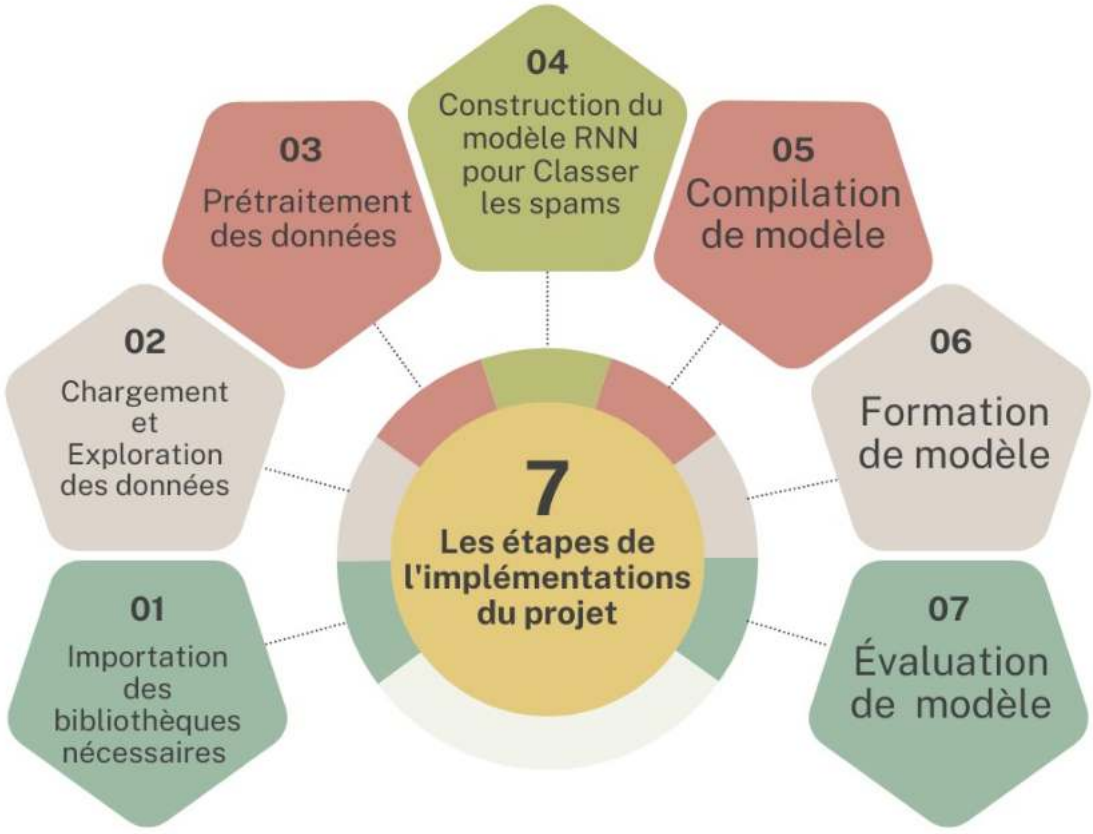

Figure 3. 3:Les étapes de l'implémentation du projet.

1. **Importation des bibliothèques nécessaires :**

La première étape consiste à importer les bibliothèques Python nécessaires pour le projet. Typiquement, ces bibliothèques incluent :

Pandas : pour la manipulation et l'analyse de données.

NumPy : pour les opérations mathématiques.

TensorFlow/Keras : pour la création du modèle de réseau neuronal.

scikit-learn : pour les outils de prétraitement et d'évaluation des données.

2. **Chargement et Exploration des Données Kaggle**

Les données sont généralement téléchargées depuis Kaggle, puis chargées dans un DataFrame pour exploration.

```
[ ]  data = pd.read_csv('SPAM - Data.csv')
```

```
 0   Category  5572 non-null   object
 1   Message   5572 non-null   object
dtypes: object(2)
```

Cette étape permet de comprendre la structure des données, vérifier les valeurs manquantes, et visualiser les différentes classes (spam ou non-spam).

```
  Category                                            Message
0      ham  Go until jurong point, crazy.. Available only ...
1      ham                      Ok lar... Joking wif u oni...
2     spam  Free entry in 2 a wkly comp to win FA Cup fina...
3      ham  U dun say so early hor... U c already then say...
4      ham  Nah I don't think he goes to usf, he lives aro...
```

3. Prétraitement des Données

Le prétraitement des données est essentiel pour préparer les données textuelles pour le modèle RNN :

- Encodage des étiquettes : Convertir les étiquettes de texte (par exemple, 'spam' et 'ham') en valeurs numériques.

```
[10]  texts = []
      labels = []

      for i,label in enumerate(data['Category']):
          texts.append(data['Message'][i])
          if label == 'ham':
              labels.append(0)
          else:
              labels.append(1)

      texts = np.asarray(texts)
      labels = np.asarray(labels)
```

❖ Tokenisation et Séquencement : Transformer le texte en séquences d'entiers, où chaque entier représente un mot ou un jeton unique.

```
[ ]  # Tokenization
     tokenizer=Tokenizer()
     tokenizer.fit_on_texts(texts)
     sequences = tokenizer.texts_to_sequences(texts)
```

❖ Padding : Rendre toutes les séquences de la même longueur pour pouvoir les passer dans le réseau neuronal.

```
[ ]  # Padding
     max_len = 500
     data = tf.keras.preprocessing.sequence.pad_sequences(sequences,maxlen=max_len)
```

❖ Divisez les données en ensembles d'entrainement et de test :

```
[ ]  len(X_train)
     4457

[ ]  len(X_test)
     1115
```

4 .Construction du Modèle RNN

On construit ensuite le modèle de réseau neuronal récurrent (RNN). Les RNN sont bien adaptés aux données séquentielles comme le texte.

```
[ ]  max_features = 10000
     num_epochs = 10
     batch_size = 60

     model = Sequential()
     model.add(Embedding(max_features,32))
     model.add(SimpleRNN(32))
     model.add(Dense(1,activation = 'sigmoid'))
```

5 .Compilation du modèle

La compilation du modèle inclut la définition de la fonction de perte, de l'optimiseur et des métriques que l'on souhaite suivre pendant l'entraînement.

```
[ ]  model.compile(optimizer='rmsprop',loss='binary_crossentropy',metrics=['acc']
```

6 .Entraînement des Données

Après la compilation du modèle, on entraîne le modèle sur les données d'entraînement.

```
history = model.fit(X_train,y_train,epochs=num_epochs,batch_size=batch_size,validation_split=0.2)
```

Sortie :

```
Epoch 1/10
60/60 ━━━━━━━━━━━━━━━━━━━━ 11s 137ms/step - acc: 0.8666 - loss: 0.4202 - val_acc: 0.9170 - val_loss: 0.2111
Epoch 2/10
60/60 ━━━━━━━━━━━━━━━━━━━━ 11s 147ms/step - acc: 0.9709 - loss: 0.1213 - val_acc: 0.9709 - val_loss: 0.0894
Epoch 3/10
60/60 ━━━━━━━━━━━━━━━━━━━━ 9s 123ms/step - acc: 0.9833 - loss: 0.0696 - val_acc: 0.9854 - val_loss: 0.0521
Epoch 4/10
60/60 ━━━━━━━━━━━━━━━━━━━━ 9s 150ms/step - acc: 0.9887 - loss: 0.0456 - val_acc: 0.9865 - val_loss: 0.0437
Epoch 5/10
60/60 ━━━━━━━━━━━━━━━━━━━━ 10s 139ms/step - acc: 0.9908 - loss: 0.0344 - val_acc: 0.9787 - val_loss: 0.0812
Epoch 6/10
60/60 ━━━━━━━━━━━━━━━━━━━━ 10s 136ms/step - acc: 0.9943 - loss: 0.0254 - val_acc: 0.9865 - val_loss: 0.0513
Epoch 7/10
60/60 ━━━━━━━━━━━━━━━━━━━━ 9s 113ms/step - acc: 0.9851 - loss: 0.0455 - val_acc: 0.9697 - val_loss: 0.0964
Epoch 8/10
60/60 ━━━━━━━━━━━━━━━━━━━━ 10s 111ms/step - acc: 0.9968 - loss: 0.0203 - val_acc: 0.9854 - val_loss: 0.0520
Epoch 9/10
60/60 ━━━━━━━━━━━━━━━━━━━━ 12s 135ms/step - acc: 0.9969 - loss: 0.0156 - val_acc: 0.9787 - val_loss: 0.0631
Epoch 10/10
60/60 ━━━━━━━━━━━━━━━━━━━━ 10s 139ms/step - acc: 0.9985 - loss: 0.0091 - val_acc: 0.9843 - val_loss: 0.0568
```

7 .Évaluation du modèle

Finalement, on évalue le modèle sur un ensemble de test pour vérifier ses performances.

```
[ ]  predict = model.predict(X_test)
     accuracy = model.evaluate(X_test,y_test)

     print("Test loss is {0:.2f} accuracy is {1:.2f}".format(accuracy[0],accuracy[1]))
```

Sortie :

```
35/35 ━━━━━━━━━━━━━━━━━━━━ 2s 38ms/step
35/35 ━━━━━━━━━━━━━━━━━━━━ 1s 21ms/step - acc: 0.9843 - loss: 0.0672
Test loss is 0.08 accuracy is 0.98
```

On a utilisé matplotlib pour tracer la courbe qui montre la variation de la précision (Accuracy) et la fonction de perte (Loss) en fonction du nombre d'époques.

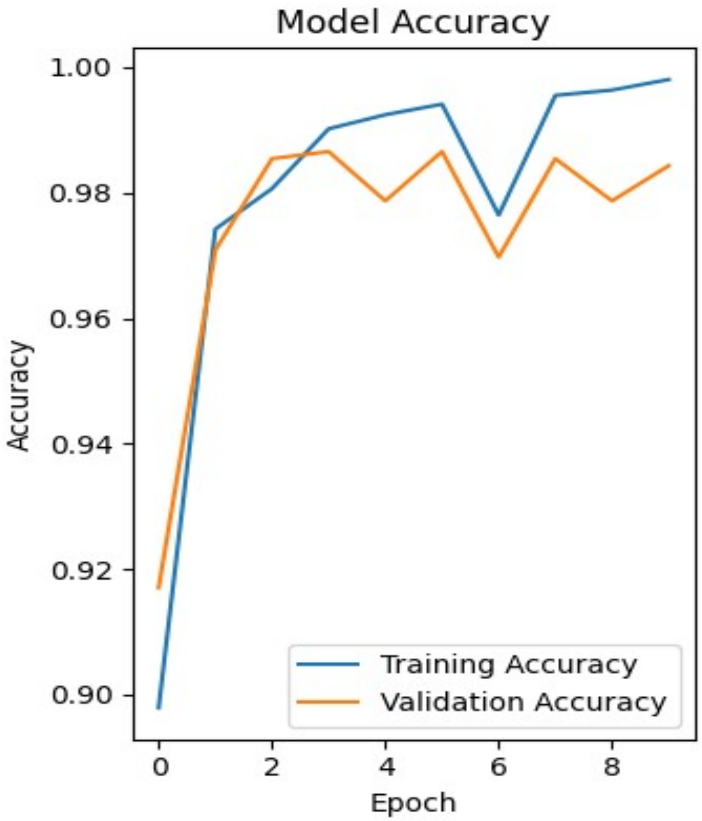

Figure 3. 4:La variation de la précision (Accuracy) en fonction du nombre d'époques.

Durant l'entraînement, on évalue les performances du modèle sur un autre ensemble de validation afin de vérifier sa précision et de réaliser les ajustements d'hyper paramètres requis.

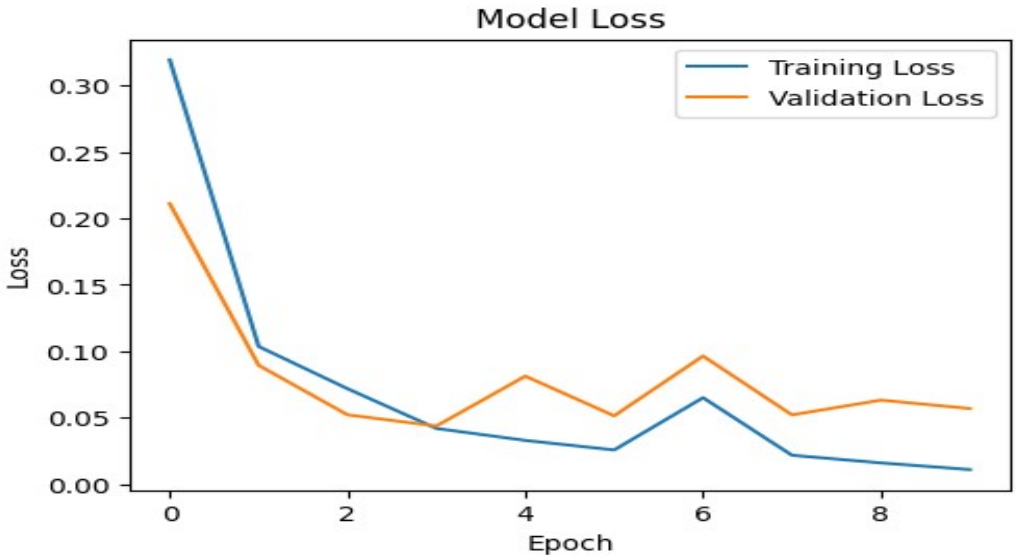

Figure 3. 5:La variation de la fonction de perte (Loss) en fonction du nombre d'époques.

Les performances du modèle sur le dataset de test ont été remarquables, avec une précision de 98%. Le modèle a été développé de manière efficace afin de différencier les différents messages présentés dans le groupe de données en fonction de leur niveau de précision.

Résumé du modèle d'entrainement

Model: "sequential"

Layer (type)	Output Shape	Param #
embedding (Embedding)	(None, 500, 32)	320,000
simple_rnn (SimpleRNN)	(None, 32)	2,080
dense (Dense)	(None, 1)	33

Total params: 644,228 (2.46 MB)
Trainable params: 322,113 (1.23 MB)
Non-trainable params: 0 (0.00 B)
Optimizer params: 322,115 (1.23 MB)

Proposition des prédictions

Nous examinons quelques exemples de tests, puis nous prévoyons la performance du modèle et les erreurs qu'il commet.

Message 1 = "Eh u remember how 2 spell his name... Yes i did. He v naughty make until i v wet."

```
1/1 ━━━━━━━━━━━━━━━━━━━━ 0s 49ms/step
Predicted label: ham
```

Message 2= "Hey, can we meet tomorrow to discuss the project?"

```
1/1 ━━━━━━━━━━━━━━━━━━━━ 0s 37ms/step
Predicted label: ham
```

Message 3="You've won a $1,000 gift card! Click here to claim your prize."

```
1/1 ━━━━━━━━━━━━━━━━━━━━ 0s 32ms/step
Predicted label: spam
```

Message 4= "Your account has been compromised. Please verify your account information by clicking this link immediately."

```
1/1 ━━━━━━━━━━━━━━━━━━━━ 0s 34ms/step
Predicted label: spam
```

Il est évident que notre modèle a pris la bonne prédiction du résultat et a correctement identifié le spam.

4.2 Comparaison

A. Comparaison entre différents types d'optimiseurs :

Ce tableau représente les résultats trouvés avec les différents optimiseurs :

Optimiseurs	precisions	la fonction de perte (Loss)	Certitude d'essai
Rmsprop	99,85%	0,0091	98%
Adam	99,98%	0.0038	98,21%
SGD	89%	0.2561	85%

Tableau 3 .Comparaison entre différents types d'optimiseurs

B. Comparaison entre le nombre d'époques :

Ce tableau représente les résultats de la précision et la certitude d'essai pendant les différentes époques :

Nombre d'époque	precisions	Certitude d'essai
5	86,73%	86%
10	99,98%	98,21%
100	**100%**	**98%**

Tableau 4 : Comparaison entre le nombre d'époques.

4.3 Discussion

a) Discussion des résultats obtenus pour les différents optimisateurs dans le tableau, qui comprend la précision, la fonction de perte (Loss) et la certitude d'essai :

1. RMSProp :

- Précision (99,85%) : RMSProp a obtenu une précision très élevée, ce qui indique que le modèle classifie correctement les échantillons dans la grande majorité des cas.
- Fonction de perte (0,0091) : La perte est très faible, ce qui signifie que le modèle est bien ajusté aux données d'entraînement. RMSProp est particulièrement efficace dans les tâches où les données sont fortement corrélées, comme dans les RNN.
- Certitude d'essai (98%) : La certitude d'essai est également élevée, confirmant que RMSProp offre une bonne stabilité et généralisation lors de la classification sur des données non vues.

2. Adam :

- Précision (99,98%) : Adam dépasse légèrement RMSProp en termes de précision, atteignant presque la perfection. Cela montre la puissance d'Adam pour trouver un minimum global plus efficacement.
- Fonction de perte (0.0038) : La perte extrêmement basse montre que le modèle a appris très efficacement avec Adam, ce qui est souvent le cas grâce à son ajustement dynamique du taux d'apprentissage.
- Certitude d'essai (98,21%) : Adam maintient une légère avance en termes de certitude d'essai, suggérant qu'il non seulement apprend mieux, mais qu'il généralise également légèrement mieux que RMSProp.

3. SGD (Stochastic Gradient Descent) :

- Précision (89%) : La précision de SGD est nettement inférieure à celle des autres optimisateurs, indiquant qu'il n'a pas aussi bien réussi à minimiser les erreurs de classification.
- Fonction de perte (0.2561) : La perte relativement élevée montre que SGD n'a pas pu bien ajuster le modèle aux données, ce qui est courant avec des données complexes où un taux d'apprentissage fixe ne suffit pas.
- Certitude d'essai (85%) : La certitude d'essai plus basse reflète le fait que le modèle entraîné avec SGD a plus de difficultés à généraliser correctement sur des données inconnues.

Résultat :

- Adam est clairement le meilleur optimiseur parmi ceux testés, offrant la meilleure précision, la plus faible perte, et la meilleure certitude d'essai. Cela le rend idéal pour des tâches complexes nécessitant une convergence rapide et efficace.
- RMSProp est également très performant, proche d'Adam, et peut être préféré dans certains cas où des réseaux plus simples ou spécifiques sont utilisés.
- SGD est nettement moins performant, ce qui le rend moins adapté pour les tâches complexes comme celles qui nécessitent des RNN, sauf dans des contextes où la simplicité et la compréhension des processus sont plus importantes que la précision maximale.

b) Discussion des résultats obtenus pour la performance d'un modèle de détection des rumeurs à différentes époques d'entraînement, avec des mesures de précision et de certitude d'essai.

➢ Analyse des Précisions :

- 5 Époques : La précision du modèle est de 86,73%. Ce chiffre indique que le modèle est capable de faire des prédictions correctes dans environ 87% des cas. Bien que cette précision soit relativement élevée, elle suggère que le modèle n'est pas encore totalement optimisé, probablement en raison du nombre insuffisant d'époques d'entraînement.
- 10 Époques : La précision atteint 99,98%, ce qui montre une amélioration significative par rapport à 5 époques. Cette augmentation suggère que le modèle s'améliore considérablement avec un entraînement plus long, atteignant presque une précision parfaite. Cela indique que le modèle apprend efficacement à partir des données d'entraînement et généralise bien sur les données de test.
- 100 Époques : La précision est maintenant de 100%, ce qui signifie que le modèle ne fait plus d'erreurs sur les données de test. Cela peut indiquer que le modèle est bien ajusté aux données d'entraînement, mais il est important de vérifier s'il s'agit d'un surajustement (overfitting), où le modèle pourrait avoir appris à trop bien s'adapter aux données d'entraînement, perdant ainsi sa capacité à généraliser à des données nouvelles ou non vues.

➢ Analyse de la Certitude d'Essai :

- 5 Époques : La certitude d'essai est de 86%, ce qui est cohérent avec la précision observée. Cela suggère que le modèle a un bon niveau de confiance dans ses prédictions, mais qu'il peut encore y avoir une certaine variabilité dans les résultats.

- 10 Époques : La certitude d'essai augmente à 98,21%, ce qui correspond à l'amélioration de la précision. Cette augmentation indique que le modèle devient plus confiant dans ses prédictions à mesure qu'il apprend mieux les caractéristiques des données.

- 100 Époques : La certitude d'essai reste élevée à 98%, malgré l'atteinte d'une précision parfaite. La légère diminution par rapport à 10 époques pourrait suggérer une légère perte de confiance dans les prédictions en raison du surajustement ou de la complexité accrue du modèle.

- Interprétation Globale :

  Le modèle affichant une précision de 100% à 100 époques suggère qu'il est capable de classifier de manière optimale les exemples de l'ensemble de test. Il est essentiel de vérifier que cette précision parfaite ne résulte pas d'un surajustement en validant le modèle sur un ensemble de données indépendant.

5. *Conclusion*

Dans ce chapitre, nous avons présenté les outils de développement et la base de données utilisés pour concevoir notre système de détection des rumeurs dans les messages. Nous avons détaillé le principe de fonctionnement et la mise en œuvre de notre système, en mettant particulièrement en lumière les trois optimiseurs employés : RMSProp, Adam et SGD.
Ce chapitre s'est focalisé sur l'évaluation des performances du modèle sur un ensemble de données spécifiques. Les résultats expérimentaux obtenus montrent que l'algorithme sélectionné et les techniques d'amélioration de précision mises en place se révèlent efficaces pour la détection des rumeurs.

Conclusion Générale

Les progrès de l'apprentissage profond ont permis de concevoir un système de détection des rumeurs extrêmement efficace et fiable, en utilisant notamment les réseaux de neurones récurrents (RNN). Ces systèmes présentent de multiples bénéfices pour l'analyse des échanges en ligne, aidant ainsi à prévenir la désinformation et à préserver l'intégrité de l'information sur les plateformes sociales.

Au sein de ce mémoire, nous avons élaboré un système de détection des rumeurs qui utilise les techniques d'apprentissage profond, notamment les RNN, afin d'analyser et de classer les messages qui pourraient contenir des rumeurs.

Par la suite, L'utilisation de l'algorithme RNN permet à notre système de détecter les dépendances temporelles présentes dans les données textuelles, ce qui permet d'améliorer la précision de la détection. Il a été prouvé que la modification adéquate des hyperparamètres favorise l'amélioration des performances des réseaux neuronaux récurrents.

L'évaluation de notre modèle, après entraînement et validation sur un jeu de données Kaggle, a montré une précision élevée dans la détection des rumeurs.

Malgré les succès obtenus, notre système de détection des rumeurs peut présenter certaines limitations, notamment en raison de l'ambiguïté des messages, de la variabilité du langage utilisé sur les réseaux sociaux, et des données d'entraînement limitées.

Il existe plusieurs perspectives pour l'avenir des systèmes de détection des rumeurs :

- Intégration de capacités multilingues pour traiter les rumeurs dans différentes langues et cultures, à mesure que les réseaux sociaux deviennent de plus en plus mondialisés.
- Amélioration de la reconnaissance en temps réel pour une détection plus rapide et plus réactive.
- Test du modèle sur d'autres bases de données pour évaluer sa généralisation.
- Enrichissement de la base de données avec plus de données variées pour renforcer la robustesse du système.

Réferences

[1] «Du Moyen Âge à Internet, les ressorts de la rumeur,» 02 mai 2016. [En ligne]. Available: https://larevuedesmedias.ina.fr/series/du-moyen-age-internet-les-ressorts-de-la-rumeur. [Accès le juin 2024].

[2] A. Abderrazek , . F. Cécile, H. Nouria et . D. Jerome, «Vers une analyse des rumeurs dans les réseaux sociaux,» Université de Lyon, Lyon 2, 2019.

[3] «Types de rumeurs dans l'entreprise,» 7 avril 2015. [En ligne]. Available: https://www.petite-entreprise.net/P-2442-81-G1-types-de-rumeurs-dans-l-entreprise.html.

[4] «GÉNÉRALITÉS RUMEURS,» 2020. [En ligne]. Available: ccousp.cm - Généralités Rumeurs.

[5] G. Claire , «Chapitre 3. Les rumeurs : sources de risque ?,» chez *Alerte marketing: Comprend anticiper, gérer les crises* , Louvain-la-Neuve: De Boeck Supérieur., 2005, pp. 49-61.

[6] P. Aldrin, «Chapitre 1. Formes et résonances sociales des rumeurs,» chez *Sociologie politique des rumeurs*, Paris cedex 14: Presses Universitaires de France, 2005, pp. 15-46.

[7] «Rumeur,» 14 janvier 2024. [En ligne]. Available: wikipedia.org - Rumeur — Wikipédia. [Accès le avril 2024].

[8] «La propagation des fausses informations sur les réseaux sociaux : étude de la plateforme Twitter,» 25 novembre 2020. [En ligne]. Available: https://www.csa.fr/Informer/Collections-du-CSA/Focus-Toutes-les-etudes-et-les-comptes-rendus-synthetiques-proposant-un-zoom-sur-un-sujet-d-actualite/La-propagation-des-fausses-informations-sur-les-reseaux-sociaux-etude-de-la-plateforme-Twitter. [Accès le AVRIL 2024].

[9] «Pourquoi croyons-nous aux fake news ?,» 21 sept 2022 . [En ligne]. Available: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC9558089/. [Accès le 2024].

[10] R. Zbienen, «Coronavirus et réseaux sociaux : premières réflexions stratégiques sur une "infodémie",» 6 mai 2020. [En ligne]. Available: https://www.frstrategie.org/publications/notes/coronavirus-reseaux-sociaux-premieres-reflexions-strategiques-sur-une-infodemie-2020. [Accès le AVRIL 2024].

[11] . J. Boatini et Tonin, «La rumeur à l'ère des réseaux sociaux numériques,» 2017.

[12] . B. Christina, P. Symeon , A. Lazaros et K. Yiannis , «Learning to Detect Misleading Content on Twitter,» Centre for Research and Technology Hellas, 2017.

[13] «Épisode 2/4 : Calomnies, rumeurs et murmures, comment circulaient les fausses informations ?,» 26 janvier 2021. [En ligne]. Available: https://www.radiofrance.fr/franceculture/podcasts/le-cours-de-l-histoire/calomnies-rumeurs-et-murmures-comment-circulaient-les-fausses-informations-6246304. [Accès le Mai 2024].

[14] J. Renard, «L'étude des rumeurs,» chez *La pensée sociale*, Érès , 2009, pp. 137-157.

[15] L. Vighier, «Viralité et information sur les réseaux sociaux : quels mécanismes, quelles menaces,» 24 nov 2023. [En ligne]. Available: https://enseignants.lumni.fr/parcours/1225/viralite-et-information-sur-les-reseaux-sociaux-quels-mecanismes-quelles-menaces.html. [Accès le JUIN 2024].

[16] P. Aldrin, «Rumeurs : il n'y a pas que la vérité qui compte,» Octobre 2005. [En ligne]. Available: https://www.scienceshumaines.com/rumeurs-il-n-y-a-pas-que-la-verite-qui-compte_fr_5199.html. [Accès le JUIN 2024].

[17] C. OUNOUGHI et . A. BENAOUDA, «Quantification de la propagation des rumeurs sur les,» 2018.

[18] E. Taïeb, «Persistance de la rumeur: Sociologie des rumeurs électroniques,» 2001.

[19] C. Maigrot, «Détection de fausses informations dans les réseaux sociaux,» 2019.

[20] Cédric Maigrot, Ewa Kijak et Vincent Claveau, «Détection de fausses informations dans les réseaux sociaux : l'utilité des fusions de connaissances,» chez *Document numérique*, 2018.

[21] K. SMAILI, «Suivi et détection des fake news et deepfakes dans les réseaux sociaux arabes – TRADEF,» 2022. [En ligne]. Available: https://anr.fr/Projet-ANR-22-ASGC-0003. [Accès le AVRIL 2024].

[22] Cédric Maigrot, Ewa Kijak et Vincent Claveau, «Fusion par apprentissage pour la détection de fausses informations dans les réseaux sociaux,» vol. 21, p. 55 à 80, 2018.

[23] M. SHEHATA, «La traçabilité pour détecter la mutation d'une rumeur en légende urbaine,» 2021. [En ligne]. [Accès le MAI 2024].

[24] A. R. Pathak, Aditee Mahajan, Keshav Singh et Aishwarya Patil, «Analysis of Techniques for Rumor Detection in Social Media,» Procedia Computer Science, 2020.

[25] S.Vanitha et R. Prabahari, «Analysis of Rumour Detection using Deep Learning Methods on Social media,» International Journal of Innovative Science and Research Technology, 2023 .

[26] «Les Etapes de l'Analyse d'un Texte,» 27 Feb 2023. [En ligne]. Available: https://www.superprof.fr/ressources/francais/francais-4eme/etude-travail-succession.html. [Accès le JUIN 2024].

[27] «Tout sur l'analyse textuelle et ses approches,» 1 Nov Nov 1, 2022. [En ligne]. Available: https://www.voxco.com/fr/blog/tout-sur-lanalyse-textuelle-et-ses-approches/. [Accès le JUIN 2024].

[28] «Qu'est-ce que le traitement automatique du langage naturel (NLP) ?,» JANUARY 2024. [En ligne]. Available: https://www.ibm.com/fr-fr/topics/natural-language-processing. [Accès le JUIN 2024].

[29] «Qu'est-ce que le traitement du langage naturel (NLP) ?,» [En ligne]. Available: https://www.elastic.co/fr/what-is/natural-language-processing. [Accès le JUIN 2024].

[30] «Classification De Texte Et Apprentissage Automatique Dans R,» [En ligne]. Available: https://fastercapital.com/fr/sujet/classification-de-texte-et-apprentissage-automatique-dans-r.html. [Accès le JUIN 2024].

[31] «Le guide définitif 2022 du traitement du langage naturel (NLP),» 15 novembre 2022. [En ligne]. Available: https://nexocode.com/blog/posts/definitive-guide-to-nlp/. [Accès le MAI 2024].

[32] D212digital, «What is Lemmatization and Stemming in NLP?,» 7 Nov 2022. [En ligne]. Available:https://212digital.medium.com/what-is-lemmatization-and-stemming-in-nlp-e25e142332c4. [Accès le JUIN 2024].

[33] J. Robert, «Natural Language Processing (NLP) : Définition et principes,» 22 Juil 2020. [En ligne]. Available: https://datascientest.com/introduction-au-nlp-natural-language-processing. [Accès le JUIN 2024].

[34] «Les secrets du traitement du langage naturel décryptés,» [En ligne]. Available: https://www.iso.org/fr/intelligence-artificielle/traitement-langage-naturel. [Accès le JUIN 2024].

[35] V. R, «Apprentissage supervisé et non supervisé : comment les différencier ?,» mars 2022. [En ligne]. Available: https://intelligence-artificielle.com/apprentissage-supervise-et-non-supervise/. [Accès le juin 2024].

[36] «Apprentissage supervisé vs apprentissage non supervisé,» mars 2021. [En ligne]. Available: https://mobiskill.fr/blog/conseils-emploi-tech/apprentissage-supervise-vs-apprentissage-non-supervise/.. [Accès le juin 2024].

[37] Z. ISMAILI, «Apprentissage Supervisé Vs. Non Supervisé,» 28 janvier 2019. [En ligne]. Available:https://brightcape.co/apprentissage-supervise-vs-non-supervise/. [Accès le JUIN 2024].

[38] «Apprentissage non-supervisé : définition et algorithmes populaires,» janvier 2022. [En ligne]. Available:https://www.journaldunet.fr/intelligence-artificielle/guide-de-l-intelligence-artificielle/1501309-apprentissage-non-supervise/. [Accès le JUIN 2024].

[39] G. Cuofano, «Apprentissage supervisé ou non supervisé en bref,» 14 novembre 2023. [En ligne]. Available:https://fourweekmba.com/fr/apprentissage-supervis%C3%A9-vs-non-supervis%C3%A9/. [Accès le JUIN 2024].

[40] L. Bastien, «Reinforcement Learning : qu'est-ce que l'apprentissage par renforcement ?,» juin 2021. [En ligne]. Available: https://www.lebigdata.fr/reinforcement-learning-definition.. [Accès le JUIN 2024].

[41] «L'apprentissage par renforcement,» 26 janv 2022. [En ligne]. Available: https://databasecamp.de/en/ml/reinforcement-learnings. [Accès le juin 2024].

[42] «Voir les réseaux. Petite introduction à l'analyse visuelle des réseaux,» 18 JUIN 2021. [En ligne]. Available:https://www.ens-lyon.fr/actualite/recherche/voir-les-reseaux-petite-introduction-lanalyse-visuelle-des-reseaux. [Accès le JUIN 2024].

[43] K. Abassia, «Détection de fausses informations dans les réseaux sociaux,» 2020.

[44] «Comment vérifier l'authenticité et l'origine de photos et vidéos,» 17 AVRIL 2024. [En ligne]. Available:https://www.kaspersky.fr/blog/real-or-fake-image-analysis-and-provenance/21787/. [Accès le JUIN 2024].

[45] «Google,» [En ligne]. Available: images.google.com .

[46] «Recherche d'image inversée : Vérifier des photos.,» JUIN 2024. [En ligne]. Available: https://newsinitiative.withgoogle.com/fr-fr/resources/trainings/reverse-image-search-verifying-photos/.

[47] D. Raynaud, «Chapitre 2. Sociologie des réseaux de diffusion,» chez *Sociologie fondamentale* , paris, Matériologiques, (2021, pp. 147-180).

[48] J. Debure, «Détection de comportements et identification de rôlesdans les réseaux sociaux,» theses.hal.science, Paris, 2021.

[49] M. Crucianu, «Cours - Fouille de graphes et réseaux sociaux (1 et 2),» 03 mai 2024. [En ligne]. Available:https://cedric.cnam.fr/vertigo/Cours/RCP216/coursFouilleGraphesReseauxSociaux.html. [Accès le JUIN 2024].

[50] «Les 3 différents types d'anomalies,» 20 fév 2017. [En ligne]. Available: https://leanbi.ch/fr/blog/les_3_differents_types_danomalies/. [Accès le JUIN 2024].

[51] K. Hoarau, «Apprentissage automatique pour la détection d'anomalies dans les graphes issus des données réseau,» Français, 2022..

[52] P. Aldrin, «Rumeurs : il n'y a pas que la vérité qui compte,» Octobre 2005. [En ligne]. Available: https://www.scienceshumaines.com/rumeurs-il-n-y-a-pas-que-la-verite-qui-compte_fr_5199.html. [Accès le juin 2024].

[53] F. Amina, «Détection des fausses nouvelles par apprentissage profond,» 2021.

[54] P. Aldrin, «Chapitre 3. Les rumeurs comme objet de science politique,» chez *Sociologie politique des rumeurs*, cedex 14: Presses Universitaires de France, 2005.

[55] P. Aldrin, «Chapitre 4 Jeux et enjeux politiques de l'information : les rumeurs dans l'agon politique,» chez *Sociologie politique des rumeurs*, Paris cedex 14: Presses Universitaires de France., 2005, pp. 135-188.

[56] P. Aldrin, «L'impensé social des rumeurs politiques. Sur l'approche dominocentrique du phénomène et son,» Mots Les langages du politique, 2010.

[57] V.-L. Bénabou et Céline Castets-Renard, «Le droit face au défi de la rumeur,» 2019.

[58] S. MEURANT et Rémi CARDON, «La cybersécurité des entreprises - Prévenir et guérir : quels remèdes contre les cyber virus ?,» 10 juin 2021. [En ligne]. Available: https://www.senat.fr/rap/r20-678/r20-678_mono.html. [Accès le JUIN 2024].

[59] G. P. M. Yemga, «RSSI : nouveaux enjeux et défis de toujours,» MAI 2023. [En ligne]. Available: https://www.journaldunet.com/cybersecurite/1522443-rssi-nouveaux-enjeux-et-defis-de-toujours/. [Accès le JUILLET 2024].

[60] S. Ghernaouti, «Chapitre 9. La sécurité des applications et des contenus,» chez *Cybersécurité*, Paris: Dunod., 2022, p. 265 à 297.

[61] D. Rajasekar et L. Robert, «Context Aware Deep Ensemble Learning Model for Rumor Detection,» Coimbatore, 2023.

[62] B. Pan, «A Two-step Rumor Detection and Classification Method Using,» Dallas, Texas, USA, 2022.

[63] Yi-Chin Chen, Zhao-Yang Liu et Hung-Yu Kao, «IKM at SemEval-2017 Task 8: Convolutional Neural Networks for Stance Detection and Rumor Verification,» Tainan, Taiwan, ROC, 2017.

[64] «Qu'est-ce que l'apprentissage automatique (Machine Learning) ?,» 2024. [En ligne]. Available: https://praedictia.com/page/lapprentissage-machine/quest-ce.html. [Accès le mai 2024].

[65] «https://www.segulatechnologies.com/fr/actualites/premco-predire-et-eviter-les-collisions-avec-vehicules/,»mai2023.[Enligne].Available: https://www.segulatechnologies.com/fr/actualites/premco-predire-et-eviter-les-collisions-avec vehicules/. [Accès le juin 2024].

[66] «Notions de base sur l'apprentissage profond,» [En ligne]. Available: https://rpubs.com/alexvezeau/SCI1035-CH1. [Accès le juin 2024].

[67] D. R. ,. K. &. S. Ritik, «Handwritten Digit Recognition using Machine learning,» 2021.

[68] sadik,h, basheera,m et marwah,a, «A robust handwritten numeral recognition using hybrid orthogonal polynomials and moments,» mars 2021. [En ligne].

[69] H. S. M. Abderrahim, «Développement d'un système de reconnaissance de caractères arabes manuscrits,» 2022.

[70] «Perceptron multicouche,» janvier 2024. [En ligne]. Available: https://datafranca.org/wiki/Perceptron_multicouche. [Accès le juin 2024].

[71] R. Tavenard, «Introduction au Deep Learning,» décembre 2023. [En ligne]. Available: https://rtavenar.github.io/deep_book/fr/content/fr/mlp.html. [Accès le juin 2024].

[72] «Comment configurer le nombre de couches et de nœuds dans un réseau neuronal,» décembre 2020. [En ligne]. Available: https://ressources.labomedia.org/comment_configurer_le_nombre_de_couches_et_de_n%C5%93uds_dans_un_reseau_neuronal. [Accès le juin 2024].

[73] É. Blent, «Réseaux convolutifs (CNN) : comment ça marche ?,» juin 2022. [En ligne]. Available: https://blent.ai/blog/a/cnn-comment-ca-marche. [Accès le juin 2024].

[74] «Comprendre les réseaux de neurones,» avril 2019. [En ligne]. Available: https://www.aspexit.com/reseau-de-neurones-on-va-essayer-de-demystifier-un-peu-tout-ca-3/. [Accès le juin 2024].

[75] J. Robert, «Convolutional Neural Network : Tout ce qu'il y a à savoir,» Juin 2020. [En ligne]. Available: https://datascientest.com/convolutional-neural-network. [Accès le juin 2024].

[76] Y. S. Obam, «Comprendre les Réseaux de Neurones Convolutifs (CNN),» Décembre 2019. [En ligne]. Available: https://yannicksergeobam.medium.com/comprendre-les-r%C3%A9seaux-de-neurones-convolutifs-cnn-d5f14d963714. [Accès le juin 2024].

[77] Daniella, «Comprendre les réseaux de neurones convolutifs (CNN),» juin 2024. [En ligne]. Available: https://www.innovatiana.com/post/convolutional-neural-network. [Accès le ouàt 2024].

[78] «Introduction to Recurrent Neural Network,» Jullet 2024. [En ligne]. Available: https://www.geeksforgeeks.org/introduction-to-recurrent-neural-network/. [Accès le ouàt 2024].

[79] D. Kalita, «Qu'est-ce que les réseaux neuronaux récurrents (RNN) ?,» juillet 2024. [En ligne]. Available: https://www.analyticsvidhya.com/blog/2022/03/a-brief-overview-of-recurrent-neural-networks-rnn/. [Accès le ouàt 2024].

[80] C. Jeeva, «Introduction to Recurrent Neural Network (RNN),» mai 2023. [En ligne]. Available: https://www.scaler.com/topics/deep-learning/rnn/. [Accès le juillet 2024].

[81] «Qu'est-ce qu'un réseau neuronal récurrent (RNN) ?,» [En ligne]. Available: https://www.ibm.com/topics/recurrent-neural-networks. [Accès le juilet 2024].

[82] visualstudio, [En ligne]. Available: https://code.visualstudio.com/docs.

[83] «Colaboratory,» [En ligne]. Available: https://colab.research.google.com/.

[84] «python,» [En ligne]. Available: https://www.python.org/doc/essays/blurb/.

[85] «tensorflow.org,» [En ligne]. Available: https://www.tensorflow.org/resources/libraries-extensions?hl=fr. [Accès le juillet 2024].

[86] «TensorFlow,» [En ligne]. Available: https://www.kalanda.net/tensorflow/. [Accès le juillet 2024].

[87] «keras,» keras.io, [En ligne]. Available: https://keras.io/api/.

[88] numpy.org, juillet 2024. [En ligne]. Available: https://numpy.org/doc/stable/.

[89] matplotlib.org, juillet 2024. [En ligne]. Available: https://matplotlib.org/.

[90] «Ensemble de données de collecte de spam par SMS,» kaggle.com, [En ligne]. Available: https://www.kaggle.com/datasets/uciml/sms-spam-collection-dataset.

[91] [En ligne].