

Dédicace

Je dédie ce travail à tous ceux qui me sont chers,

A

Ma mère,

Qui m'a soutenu toutes les fois quand je n'étais pas à l'aise. Elle est d'une importance primordiale puisque son amour est inconditionnel et inébranlable. J'espère ne jamais te décevoir ni trahir ta confiance et tes sacrifices.

A

Mon père,

De tous les pères. Tu es le meilleur tu as été et tu seras toujours un exemple pour moi par tes qualités humaines, ta persévérance et perfectionnisme. La relation amicale entre nous n'a pas de prix. . En ce jour, j'espère réaliser l'un de tes rêves.

A

Mon cher frère,

Tu étais toujours là pour me soutenir et m'encourager. Je vous souhaite une vie pleine de bonheur et de succès et que Dieu vous protège.

Au final aucune dédicace ne saurait exprimer mes respects, ma reconnaissance et tout Mon amour. Puisse dieu vous préserver et vous procurer santé et bonheur.

Rana 

Remerciements

La construction de ce mémoire n'aurait été possible sans l'intervention de certaines personnes.

Je tiens à exprimer ici mes plus sincères remerciements pour leurs précieux conseils.

Je souhaite tout d'abord adresser mes vifs et chaleureux remerciements à Madame Souhir BOUAZIZ, mon encadrante académique, pour l'accompagnement tout au long de ce travail, qui n'aurait pu être mené à bien sans son soutien et ses encouragements. Ses conseils avisés, ses critiques constructives et ses qualités humaines m'ont été d'une aide précieuse pour mener à terme ce projet.

Je remercie également les membres du jury pour le temps qu'ils ont consacré à l'examen de ce mémoire et j'espère qu'ils y trouveront les qualités de clarté et de motivation attendues.

Enfin, j'exprime ma profonde gratitude envers toutes les personnes qui m'ont aidé, de près ou de loin, à accomplir ce travail.

Tables des matières

Introduction générale.....	1
Chapitre 1 : Les fausses nouvelles dans les médias sociaux :	
Définitions et approches.....	4
1. Introduction.....	5
2. Les fausses nouvelles.....	5
3. Les fausses nouvelles dans les médias sociaux	6
4. Les méthodes de détection des fausses nouvelles	8
4.1. Les méthodes supervisées	9
4.1.1. Méthodes fondées sur le contenu des nouvelles	10
4.1.1.1. Modalité unique.....	10
4.1.1.2. Multimodalité	11
4.1.2. Méthodes fondées sur le contexte social	12
4.1.2.1. Crédibilité de l'utilisateur	12
4.1.2.2. Mode de propagation	13
4.1.3. Méthodes externes fondées sur les connaissances	15
4.1.4. Méthodes hybrides	16
4.2. Les méthodes faiblement supervisées	19
4.2.1. Faible supervision du contenu.....	19
4.2.2. Faible supervision sociale	21
4.3. Les méthodes non supervisées	22
4.3.1. Méthodes de détection des anomalies	22
4.3.2. Méthodes probabilistes basées sur des modèles graphiques	23
4.3.3. Méthodes graphiques	23
4.3.4. Méthodes génératives basées sur l'apprentissage contradictoire	23
4.3.5. Méthodes d'apprentissage par transfert	24
5. Conclusion.....	25
Chapitre 2 : La détection des fausses nouvelles basée sur l'analyse	
des sentiments : Etat de l'art.....	26
1. Introduction.....	27
2. Analyse des sentiments dans le traitement du langage naturel	27

3. Approches et techniques d'analyse de sentiment dans le traitement du langage naturel.....	29
3.1. Les techniques basées sur les lexiques	30
3.2. Les techniques basées sur l'apprentissage automatique	31
3.2.1. les modèles traditionnels	31
3.2.2. les modèles d'apprentissage profond	32
3.3. Les approches hybrides	36
3.4. Les techniques basés sur les Transformateurs.....	37
4. Les méthodes de détection des fausses nouvelles basée sur l'analyse des sentiments	40
4.1. Systèmes de détection de fausses nouvelles basés sur l'analyse de sentiment (SA).....	40
4.2. SA comme caractéristique pour les systèmes de détection de fausses nouvelles	42
4.3. Étude comparative des systèmes de détection de fausses nouvelles utilisant l'analyse de sentiment.....	45
5. Conclusion.....	48
Chapitre 3 : Analyse des sentiments pour la détection des fausses nouvelles : Système proposé	50
1. Introduction	51
2. Objectifs du système proposé pour la détection des fausses nouvelles 51	
3. Architecture générale du Système Proposé	52
3.1. L'ensemble de données.....	53
3.2. Étapes de prétraitement de texte	53
3.3. Unité de classification de texte	53
3.3.1. Modèle basé sur l'apprentissage profond : Bi-LSTM et CNN.....	54
3.3.2. Modèle basé sur les transformateurs : BERT	60
3.4. Unité d'analyse de sentiment.....	63
4. Méthodologie d'évaluation.....	64
5. Conclusion.....	66
Chapitre 4 : Résultats expérimentaux.....	67
1. Introduction	68

2. Configuration des modèles du système proposé.....	68
2.1. Environnement d'exécution	68
2.2. Paramétrage du modèle Bi-LSTM.....	69
2.3. Paramétrage du modèle CNN	70
2.4. Paramétrage du modèle BERT	71
3. Résultats et discussions.....	72
3.1. Visualisation des données	73
3.1.1. Ensemble de données ISOT	73
3.1.2. Ensemble de données FakeNewsNet	76
3.2. Prétraitement de texte.....	78
3.3. Analyse des sentiments.....	79
3.4. Classification de texte.....	81
3.4.1. Résultats de classification pour l'ensemble de données ISOT	81
3.4.2. Résultats de classification pour l'ensemble de données GossipCop	85
3.5. Comparaison des performances des modèles et discussions	88
4. Etude comparatives.....	89
5. Conclusion.....	91
Conclusion générale	93
Bibliographie.....	95

Liste des figures

Figure 1-1 : Les raisons d'utiliser les médias sociaux selon le rapport Digital 2021 Global Digital Overview (Hamed, et al., 2023)	6
Figure 1-2 : Le rôle des médias sociaux dans la diffusion de fausses nouvelles (Hamed, et al., 2023).....	7
Figure 1-3 : Taxonomie des méthodes de détection des fausses nouvelles.....	9
Figure 2-1 : La roue de l'émotion Plutchik (Hamed, et al., 2023).....	28
Figure 2-2 : Taxonomie des techniques d'analyse de sentiment	30
Figure 2-3 : Différences entre deux approches de classification de la polarité des sentiments, apprentissage automatique et apprentissage profond (Dang, et al., 2020).....	33
Figure 2-4 : Réseaux Neuronaux Profonds (DNN) (Zhu, et al., 2022).....	34
Figure 2-5 : Réseaux de Neurones Convolutifs (CNN) (Dang, et al., 2020)	34
Figure 2-6 : Réseaux à Mémoire à Long Court Terme (LSTM) (Dang, et al., 2020).....	36
Figure 2-7 : La structure des modèles PLMs (Alghamdi, et al., 2023)	38
Figure 3-1: Détection des fausses nouvelles basée sur l'analyse des sentiments	52
Figure 3-2 : Unité de classification de texte des modèles d'apprentissage profond	54
Figure 3-3 : L'architecture de base de Bi-LSTM utilise Word embedding.....	55
Figure 3-4 : Diagramme en couche de Bi-LSTM.....	56
Figure 3-5 : L'architecture de base de CNN utilise Word embedding.....	57
Figure 3-6 : Diagramme en couche de CNN	59
Figure 3-7 : Unité de classification de texte du modèle basé sur les transformateurs : BERT	60
Figure 3-8 : Pre-training et fine-tuning du modèle BERT par (Devlin, et al., 2019).....	61
Figure 3-9 : Diagramme en couche de BERT	62
Figure 3-10 : Unité d'analyse de sentiment pour les modèles d'apprentissage profond	63

Figure 3-11 : Unité d'analyse de sentiment pour les modèles basés sur les transformateurs : BERT	64
Figure 3-12 : Formule de l'exactitude (Accuracy)	65
Figure 3-13 : Formule de la précision (Precision)	65
Figure 3-14 : Formule du rappel (Recall)	65
Figure 3-15 : Formule du score F1	66
Figure 4-1 : Google Colab	68
Figure 4-2 : Bibliothèques python	69
Figure 4-3 : Répartition des articles des nouvelles selon la véracité	74
Figure 4-4 : Distribution temporelle des articles selon la véracité	74
Figure 4-5 : Comparaison de la longueur moyenne des titres entre les vraies et fausses nouvelles dans ISOT	75
Figure 4-6 : Nuage de mots de l'ensemble de données ISOT	75
Figure 4-7 : Ensemble de données GossipCop	77
Figure 4-8 : Longueur moyenne des titres pour les vraies et fausses nouvelles dans GossipCop	77
Figure 4-9 : Nuage de mots de l'ensemble de données GossipCop	78
Figure 4-10 : Les mots fréquents dans les titres des vraies et fausses nouvelles avant le prétraitement de texte dans GossipCop	78
Figure 4-11 : Les mots fréquents dans les titres des vraies et fausses nouvelles après le prétraitement de texte dans GossipCop	79
Figure 4-12 : analyse des sentiments avec TextBlob dans ISOT	79
Figure 4-13 : Classification des sentiments dans ISOT	80
Figure 4-14 : Distribution des sentiments pour les vraies et fausses nouvelles dans ISOT	80
Figure 4-15 : Conversion des sentiments en valeurs numériques	81
Figure 4-16 : Performance d'entraînement et de validation du modèle Bi-LSTM dans ISOT	82
Figure 4-17 : Matrice de confusion du modèle Bi-LSTM dans ISOT	82
Figure 4-18 : Performance d'entraînement et de validation du modèle CNN dans ISOT	83
Figure 4-19 : Matrice de confusion du modèle CNN dans ISOT	83
Figure 4-20 : Performance d'entraînement et de validation du modèle BERT dans ISOT	84

Figure 4-21 : Matrice de confusion du modèle BERT dans ISOT	84
Figure 4-22 : Performance d'entraînement et de validation du modèle Bi-LSTM dans GossipCop	85
Figure 4-23 : Matrice de confusion du modèle Bi-LSTM dans GossipCop	86
Figure 4-24 : Performance d'entraînement et de validation du modèle CNN dans GossipCop	86
Figure 4-25 : Matrice de confusion du modèle CNN dans GossipCop	87
Figure 4-26 : Performance d'entraînement et de validation du modèle BERT dans GossipCop	87
Figure 4-27 : Matrice de confusion du modèle BERT dans GossipCop	88
Figure 4-28 : Comparaison des Performances d'Entraînement et de validation des Modèles CNN, Bi-LSTM et BERT	89

Liste des Tableaux

Tableau 1-1 : Les méthodes supervisées pour la détection des Fausses Nouvelles	19
Tableau 1-2 : Les méthodes non supervisées pour la détection des Fausses Nouvelles	25
Tableau 2-1 : Principales caractéristiques des systèmes de détection de fausses nouvelles utilisant SA : systèmes fournissant des résultats de performance quantitatifs sur la tâche.	47
Tableau 2-2 : Principales caractéristiques des systèmes de détection de fausses nouvelles utilisant l'analyse de sentiment : systèmes fournissant des résultats de performance quantitatifs sur la tâche.	48
Tableau 4-1 : Hyperparamètres du modèle Bi-LSTM dans ISOT.....	69
Tableau 4-2 : Hyperparamètres du modèle Bi-LSTM dans GossipCop	70
Tableau 4-3 : Hyperparamètres du modèle CNN dans ISOT	71
Tableau 4-4 : Hyperparamètres du modèle CNN dans GossipCop	71
Tableau 4-5 : Hyperparamètres du modèle BERT dans ISOT	72
Tableau 4-6 : Hyperparamètres du modèle BERT dans GossipCop	72
Tableau 4-7 : La répartition de l'ensemble de données ISOT.	73
Tableau 4-8: Les statistiques de l'ensemble de données FakeNewsNet.....	76
Tableau 4-9 : Performance des modèles sur l'ensemble de données ISOT	85
Tableau 4-10 : Performance des modèles sur l'ensemble de données GossipCop	88
Tableau 4-11 : Résultats des modèles de détection de fausses nouvelles sur l'ensemble de données ISOT	90
Tableau 4-12 : Résultats des modèles de détection de fausses nouvelles sur l'ensemble de données GossipCop	91

Liste des abréviations

AE-GCN	AutoEncoder Graph-Convolutional Network
BERT	Bidirectional Encoder Representation from Transformer
Bi-LSTM	Bidirectional Long Short-Term Memory
BOW	Bag Of Words
CBOW	Continuous Bag-Of-Words
CNN	Convolutional Neural Network
CRF	Conditional Random Field
CV	Count Vectorizer
DDGCN	Dual Dynamic Graph Convolutional Network
DL	Deep Learning
DNN	Deep Neural Network
FND	Fake News Detection
FNR	Fake News Revealer
GAE	Graph AutoEncoder
GAL	Generative Adversarial Learning
GAN	Generative Adversarial Network
GCN	Graph Convolutional Network
GLAN	Graph-based Linear Assignment Network

GLOVE	Global Vectors for Word Representation
GRU	Gated Recurrent Unit
HAN	Hierarchical Attention Networks
H-MCAN	Hierarchical Multi-modal Contextual Attention Network
HMM	Hidden Markov Models
KG	knowledge Graphs
KMAGCN	Knowledge-aware Multi-modal Adaptive Graph Convolutional Network
k-NN	k-Nearest Neighbors
LSTM	Long Short-Term Memory
MCAN	Multi-modal Co-Attention Network
ML	Machine Learning
MLM	Masked Language Model
MSCCNN	Memristor-Based Sparse Compact Convolutional Neural Network
NLP	Natural Language Processing
NLTK	Natural Language Toolkit
NSP	Next Sentence Prediction
PLM	Pre-trained Language Models
RNN	Recurrent Neural Network
RoBERTa	Robustly optimized BERT approach
SA	Sentiment Analysis

SMAN	Multi-head Structure-aware Attention Network
SVM	Support Vector Machine
TF-IDF	Term Frequency-Inverse Document Frequency
UGC	User-generated Content
UPFD	User Preference-aware Fake Detection
VAE-GCN	Variational AutoEncoder- Graph Convolutional Network
VGAE	Variational Graph Auto-Encoder
WMD	Word Mover's Distance

Introduction générale

Les fausses nouvelles sont des informations délibérément créées pour être trompeuses ou mensongères. Elles sont conçues pour diffuser de la désinformation, et leur succès repose presque entièrement sur les lecteurs, qui ont le pouvoir de les partager ou de les arrêter (Özgöbek & Gulla, 2017). Les réseaux sociaux représentent actuellement l'allié le plus puissant des fausses nouvelles : un espace non supervisé et non réglementé où l'information est accessible à l'échelle mondiale et instantanément, et où les gens passent généralement seulement quelques secondes sur chaque publication. Ils constituent un moyen de diffusion parfaitement adapté pour propager la désinformation (Tacchini, et al., 2017) (Vedova, et al., 2018). Bien sûr, en raison de leur taille, les réseaux sociaux resteront non supervisés dans un avenir proche, du moins pas par des humains. Ainsi, les solutions envisagées sont (1) éduquer les utilisateurs à distinguer les fausses nouvelles des vraies ou (2) développer un système de détection des fausses nouvelles capable de bloquer ces dernières ou d'avertir les utilisateurs en cas de possible manque de véracité.

Cependant, avec la popularité croissante des médias sociaux largement utilisés à des fins politiques, le problème des fausses nouvelles a pris plus d'importance ces dernières années, posant ainsi un grand défi de détection. La diffusion rapide de fausses nouvelles peut causer des dommages considérables aux individus, à la société, voire aux pays. Dans ce contexte, le sentiment public exerce une influence croissante sur la société. Comprendre l'opinion publique est souvent nécessaire dans de nombreux domaines : de la stratégie de marque à la prédiction d'événements tels que les résultats d'élections. La capacité à mesurer précisément le sentiment public en temps quasi réel est donc très souhaitable. L'analyse des sentiments est l'une des approches permettant d'évaluer le sentiment public.

L'analyse des sentiments consiste à examiner les opinions, perceptions, attitudes, pensées et émotions des individus partagées sur différentes plateformes de médias sociaux. Plus précisément, l'analyse des sentiments vise à classer un texte particulier comme étant un sentiment neutre, positif ou négatif. Dang et al. (Dang, et al., 2020) ont identifié trois principales approches dans l'analyse des sentiments : basée sur l'apprentissage automatique, basée sur des lexiques et l'approche hybride. La détection des fausses nouvelles est difficile, et les études sur ce sujet en sont encore à leurs débuts.

Cependant, la prolifération des fausses nouvelles est en pleine expansion. Cela contribue à développer et à explorer des pistes de recherches permettant d'actualiser et d'améliorer les techniques d'identification de ces fausses nouvelles. De très nombreuses études sur l'état de l'art consacrées à l'identification des fausses nouvelles sur les réseaux sociaux exploitent soit une, soit plusieurs des caractéristiques d'une nouvelle comme le contenu, la propagation en réseau ou l'utilisateur. Les sentiments véhiculés dans les titres des nouvelles pourraient, en revanche, être au cœur de l'identification des fausses nouvelles en apportant un indice de la crédibilité de l'information. Dans ce sens, les techniques d'apprentissage profond récentes contribuent à l'amélioration, la classification, la prédiction ou l'analyse du contenu textuel grâce à leur capacité d'apprentissage efficace, d'extraction de caractéristiques et de modélisation des structures complexes.

Dans ce travail, nous proposons un modèle pour la détection des fausses nouvelles qui s'appuie sur une analyse de sentiment prise comme un des points principaux permettant d'améliorer le

modèle de détection proposé. Ce modèle basé sur une approche d'apprentissage profond permet de traiter un vaste corpus de fausses nouvelles pour en extraire, grâce à l'analyse des sentiments, des caractéristiques susceptibles d'enrichir le modèle proposé par une information complémentaire basée sur l'analyse des sentiments des articles d'actualité. Ces caractéristiques viennent compléter les caractéristiques du contenu des nouvelles au sein du modèle de détection proposé pour permettre une plus grande performance du détecteur.

Cette approche se distingue de l'existant en ce qu'elle permet une prise en compte plus fine des émotions et sentiments, souvent ignorés dans les méthodes classiques de détection, apportant ainsi une dimension supplémentaire à la classification. De plus, grâce à l'utilisation de modèles avancés, il devient possible de traiter efficacement de grandes quantités de données, un aspect crucial dans la détection des fausses nouvelles où les ensembles de données sont massifs et diversifiés.

Dans le cadre de l'analyse des sentiments, nous avons calculé les polarités du texte. Les résultats de cette analyse sont ensuite fusionnés avec ceux de la tâche de classification, qui repose sur des modèles d'apprentissage profond tels que CNN, Bi-LSTM et BERT permettent l'extraction des caractéristiques complexes et contextuelles des articles afin de renforcer le modèle de détection avec des informations à la fois sémantiques et émotionnelles.

Dans ce projet, nous utilisons deux ensembles de données tel que ISOT et GossipCop qui reflètent respectivement les vraies et fausses nouvelles et les ressources officielles (vraies).

Les résultats expérimentaux confirment l'efficacité de notre approche : notre modèle basé sur BERT a atteint des performances élevées, notamment une précision de 99,34 % sur l'ensemble de données ISOT et de 97,54 % sur GossipCop. Ces résultats témoignent de la robustesse de BERT pour saisir les nuances complexes des contenus textuels, renforcée par l'ajout de l'analyse de sentiment, qui enrichit la compréhension contextuelle et émotionnelle des informations.

La structure de ce mémoire est organisée comme suit :

- Nous commençons par le chapitre 1 qui présente un aperçu des méthodes utilisées pour identifier les fausses nouvelles, en mettant l'accent sur celles qui se propagent sur les réseaux sociaux.
- Au chapitre 2, nous explorons l'état actuelle de la technologie permettant de détecter les fausses nouvelles et ainsi un accent particulier y est mis sur l'analyse de sentiments.
- Nous exposons au chapitre 3 notre principale contribution à la résolution du problème à l'étude.
- Enfin, au chapitre 4 nous présentons les résultats expérimentaux de notre système de détection de fausses informations, en explorant, analysant et comparant les performances de différents modèles d'apprentissage profond.