

REPUBLIQUE ALGERIENNE DEMOCRATIQUE ET POPULAIRE
MINISTRE DE L'ENSEIGNEMENT SUPERIEUR ET DE LA
RECHERCHE SCIENTIFIQUE

Université SAAD DAHLAB Blida



Faculté des sciences

Département : Informatique

Option : Ingénierie de logiciel

MEMOIRE DE MASTER

THEME

DETECTION DE LA DEPRESSION UTILISANT E-RISK DATASET

Présenté par :

Foulani. Nacera

Mezaourou Boumediene

Encadré par : Phd Madani .A

President : Mr Derar

Examineur : Mr Kammache

Promotion : 2018/2019

Remerciement

Remerciement

Je tiens d'abord à remercier notre conseillère de thèse, Mme Amina Madani pour sa remarque contribution dans la réalisation de ce travail, je remercie Mme Fatima Boumahdi pour tout son soutien, sa patience et son aide pendant toute la durée de la recherche.

J'aimerais également remercier les enseignants qui ont offert leur aide et les experts qui ont préparé le terrain pour que ce projet devienne une réalité grâce à leurs efforts extraordinaires pour contribuer à la communauté de l'informatique.

Enfin, je dois exprimer ma profonde gratitude à mon époux, mes parents et à mes enfants me fournir un soutien indéfectible et un encouragement continu tout au long de ce projet. Cette réalisation n'aurait pas été possible sans eux. Merci.

Nacera Foulani.

Résumé

Résumé

La santé mentale est considérée comme l'un des fléaux les plus importants dans le monde d'aujourd'hui. Par conséquent, notre travail vise à utiliser le potentiel des plateformes de médias sociaux pour résoudre l'un des plus grands problèmes de santé mentale, soit l'identification de la dépression. Nous proposons un nouveau modèle d'apprentissage profond que nous formons sur un ensemble de données destinées à la dépression afin de détecter une telle maladie mentale à partir d'un tweet individuel. Nos principales contributions à ce travail se trouvent dans les points suivants : Nous avons formé notre propre mot intégration à l'aide d'un ensemble de données destinées à la dépression. Ensuite Nous avons utilisé un modèle qui s'adapte à n'importe quel type de données concernant les maladies mentaux, ce modèle fonctionne avec plusieurs méthodes d'apprentissage profond les réseaux de neurone convulatif (CNN), La méthode LSTM, BiLStM et Naive Bayes. Afin d'améliorer le processus d'extraction des caractéristiques et d'améliorer les performances du modèle. Nous avons analysé à travers différentes expériences la performance des quatre modèles d'apprentissage profond afin de fournir plus de perspectives et d'idées pour les recherches sur la dépression. Notre modèle a atteint une précision de 95 %, surpassant tous les modèles statistiques ou d'apprentissage approfondi que l'on trouve actuellement dans la littérature.

Mots clés : Analyse des sentiments, Apprentissage profond (deep learning), Santé mentale, Classification des textes.

Sommaire

Sommaire

Remerciement	02
Résumé.....	03
Table des matières.....	04
Table des figures.....	08
Liste des tableaux.....	09
Introduction Générale.....	10
Problématique.....	10
Objectif.....	11
Chapitre 1 Analyse des Sentiments.....	12
1.1. Introduction.....	13
1.2. Analyse des Sentiment et l’exploration d’opinions.....	13
1.2.1. Domaines d’applications de l’Analyse des Sentiments.....	14
1.2.1.1. Détection.....	14
1.2.1.2. Prédiction.....	16
1.2.2. Exploration du Texte (Texte Mining).....	18
1.3. Niveaux de l’Analyse de sentiment.....	19
1.3.1. L’analyse des sentiments au niveau du document.....	19
1.3.2. L’analyse des sentiments au niveau de la phrase.....	19
1.3.3. L’analyse des sentiments au niveau de la sous-phrase.....	19
1.4. Approches et techniques d’analyse de sentiments.....	20
1.4.1. L’exploitation des Emoticon.....	20
1.4.2. L’utilisation de Linguistic Inquiry and Word Count.....	20
1.4.3. Senticnet.....	18
1.4.4. SentiStrength.....	18
1.5. Les défis de l’analyse de Sentiment.....	21
1.6. Traitement Automatique Du langage (TAL).....	21
1.7. Conclusion.....	22
Chapitre 2 Deep Learning	23

Sommaire

2.1. Introduction	24
2.2. Définitions	24
2.2.1. Intelligence artificielle (IA)	24
2.2.2. Apprentissage Automatique (Machine Learning)	25
2.2.3. Méthodes d'apprentissage automatique	25
2.2.4. L'apprentissage supervisé	25
2.2.5. L'apprentissage non supervisé	26
2.2.6. Réseaux de Neurones Artificiels	26
2.2.7. L'apprentissage profond (Deep Learning)	26
2.2.8. Les classes de l'apprentissage profond	28
2.2.8.1. Apprentissage supervisé	28
2.2.8.2. Apprentissage non supervisé	29
2.2.8.3. Apprentissage Hybride (Hybrid Learning)	29
2.3. Analyse de Sentiment utilisant les techniques d'apprentissage profond	30
2.3.1. Réseaux de Neurones convolutifs (CNN)	31
2.3.2. Réseaux De Neurones Récurrents(RNN)	31
2.3.3. Réseaux De Neurones Profond (DNN)	32
2.3.4. Réseaux De Croyance Profond (deepbelief networks)DBN	33
2.3.5. Réseaux De Neurones Hybrides	34
2.4. Conclusion	35
Chapitre 3 Détection de la dépression dans les réseaux sociaux : état de l'art	36
3.1. Introduction	37
3.2. Les troubles Dépressifs	37
3.3. Détection de la dépression sur les réseaux sociaux	40
3.4. Discussion	44
3.5. Conclusion	47
Chapitre 4 Modèle de détection de dépression	48
4.1 Introduction	49
4.2 Architecture du Modèle	49
4.2.1 Extraction des données (train data, Test data & sentiments Users)	51
4.2.2 Prétraitement & Vectorisation	52

Sommaire

4.2.2.1 Prétraitement "Preprocessing "	52
4.2.2.2 Vectorisation	54
4.2.2.2.1 Tokénisation	54
4.2.2.2.2 Embeddings	54
4.2.2.2.3 Glove	55
4.2.3 Apprentissage des Modèles (Model Training)	56
4.2.3.1 Les Réseaux de Neurones Convolutifs(CNN)	56
4.2.3.2 Modèle Naive- Bayes	61
4.2.3.3 Le modèle LSTM	63
4.2.3.4 Modèle BiLSTM	67
4.2.4 Test et évaluation	68
4.3 Conclusion	68
Chapitre 5 Expérimentation et résultats	69
5.1. Introduction	70
5.2. Environnement de travail	70
5.2.1. Environnement matériel	70
5.2.2. Environnement logiciel	70
5.2.2.1. Python	71
5.2.2.2. Spyder 3.5	71
5.2.2.3. Anaconda	71
5.3. Package	71
5.3.1. Package CSV	71
5.3.2. Package Numpy	71
5.3.3. Package Nltk	71
5.3.4. Package pandas	71
5.3.5. Package TensorFlow	71
5.3.6. Package Keras	72
5.3.7. Package Gensim	72
5.4. Implémentation et Mise en œuvre	73
5.4.1. Le traitement des données(fichiers)	73
5.4.1.1. Regroupement des Fichiers et Prétraitement	73
5.4.1.2. Création des fichiers CSV	73
5.4.1.3. Structure des fichiers XML	74

Sommaire

5.4.1.4. Procédure de traitement des fichiers.....	75
5.4.1.5. Procédure de Prétraitement « Preprocessing »	78
5.5. Stockage binaire de toutes les étapes dans un fichier binaire «.pkl ».....	78
5.6. Evaluation.....	78
5.7. Mesures et performances	78
5.8. Phase d'apprentissage « Training ».....	81
5.8.1. Les modèles.....	82
5.8.2. Evaluation	82
5.9. Conclusion	82
6. Bibliographie	83

Liste des Figures

Liste des Figures

Figure 4-1: Notre Modèle d'architecture de Détection De Dépression -----	50
Figure 4-2: Structure interne des fichiers Train et Test dans la collection Erisk 2019----	51
Figure 4-3: Structure de la collection Erisk2019 (Dossiers et fichiers)-----	51
Figure 4-4: Représentation vectorielle dense-----	54
Figure 4-5: Représentation Vectorielle d'une phrase -----	55
Figure 4-6: Model CBOW [131] -----	55
Figure 4-7: Model skip-gram [131]-----	56
Figure 4-8 : Les deux(02) parties de Convolutional Neural Network [134]-----	56
Figure 4-9 : Les quatre types de couches pour un réseau de neurones Convolutif [134]-----	57
Figure 4-10 : Ensemble de Données de Formation sur la Météo -----	58
Figure 4-11: Architecture CNN pour la classification de phrases [134]. -----	60
Figure 4-12: Ensemble de Données de Formation sur la Météo -----	62
Figure 4-13: Le module de répétition dans un RNN [136] -----	63
Figure 4-14: Le module de répétition dans un LSTM [136]-----	63
Figure 4-15: Notation LSTM-BiLSTM [136]-----	64
Figure 4-16: L'état cellulaire du Module LSTM [136]-----	64
Figure 5-17 Exemple de perceptron [136]-----	65
Figure 5-18 : la première étape de Module LSTM [136]-----	65
Figure 5-19 : Deuxième étape du modèle LSTM[136] -----	66
Figure 4.20 : La troisième étape du modèle LSTM -----	66
Figure 4.22 :Structure générale de réseaux BiLSTM -----	67
Figure 5.1 Phase de développement du projet de notre modèle -----	73
Figure 5.2 Classification des utilisateurs -----	74
Figure 5.8 : Représentation graphique « histogramme » -----	81

Liste des Tableaux

Liste des Tableaux

Tableau 1-1 : Récapitulatif des études d'analyse des sentiments dans différents domaine -----	15
Tableau 1-2 : Récapitulatif des études d'analyse des sentiments prédictives dans différents ----- domaine -----	16
Tableau 3-1 : Tableau comparatif des travaux de détection de dépression. -----	44
Tableau 5-1 Les résultats obtenus après traitement des fichiers -----	79
Tableau 5-2 Etat de comparatif des résultats obtenus -----	81

Contexte

Le but de cette étude est d'utiliser les techniques d'apprentissage profond de l'intelligence artificielle pour surmonter l'un des défis les plus importants de la santé mentale, qui est d'identifier automatiquement la dépression du comportement d'un individu. Nous utilisons les tweets comme un moyen de suivre ce comportement en raison de la culture pensée-expressive de la plate-forme et la disponibilité des données.

Problématique

Malgré les efforts investis contre les troubles mentaux, la santé mentale demeure toujours l'un des problèmes de santé les plus dangereux au monde. Ce qui fait qu'une telle question mérite plus d'attention, ce sont les menaces qui découlent de sa négligence. La recherche a montré que les personnes souffrant d'une ou plusieurs maladies mentales subiront probablement un effet boule de neige vers d'autres troubles, entraînant des conséquences dégradantes pour leur vie, et dans certains cas, des conséquences fatales.

L'un des troubles mentaux les plus populaires parmi les populations du monde est le trouble dépressif majeur, communément connu sous le nom de dépression clinique, avec près de 300 millions de personnes qui en souffrent dans le monde [127]. Des études ont montré que 3 à 5 % des hommes et 8 à 10 % des femmes de l'ensemble de la population mondiale sont susceptibles de connaître un épisode dépressif majeur en un an [127]. Ce qui fait de la dépression les troubles les plus connus dans le monde, c'est la probabilité qu'elle soit déclenchée par d'autres problèmes de santé, car elle est souvent associée à d'autres maladies et troubles mentaux. C'est l'une des principales causes du suicide car elle a un retentissement important sur la vie des malades, conséquences individuelles, sociales, professionnelles, familiales, risque de suicides. Des pensées autour de la mort et suicidaires font partie des symptômes de la dépression, et ils sont considérés comme la conséquence la plus dramatique de la dépression d'où la détection de la dépression n'est pas toujours évidente. De nombreux gens souffrant de dépression, ayant tendance à se cacher de ce qu'ils traversent, de toute façon, ils ne demandent pas de soutien dans la plupart des cas ce qui peut entraîner l'empirement de leur état et même parfois le suicide.

A l'heure actuelle de nouvelles études qui consiste à prédire la dépression utilisant des méthodes plus avancé en parallèle que les méthodes classiques ces dernières se basent sur l'apprentissage profond et l'intelligence artificielle.

Introduction

Objectif

Une fois la dépression est identifiée et traitée, cela entrainera la guérison. À cet effet, notre principal objectif est la détection et la prédiction de la dépression en se basant principalement sur l'apprentissage profond (deep learning) et en utilisant des sources de données concrètes extraites des réseaux sociaux. Le partage des pensées et les désires les plus intimes sur les différents réseaux permet le suivi du comportement naturel des utilisateurs, ce qui donne la possibilité d'identifier les utilisateurs présentant des troubles psychologiques. Dans notre étude nous nous sommes basés sur des tweets comme une source de données, ces derniers enregistrent l'auto-expression des gens de la manière la plus naturelle. Notre Modèle consiste à détecter la dépression en analysant les tweets des utilisateurs, et la sélection de ces derniers dans deux classes positive dépression et négative dépression. La performance de notre modèle de classifications est basée sur plusieurs paramètres, dans le cas de notre modèle on a donné la priorité au paramètre Accuracy pour mesurer le pourcentage de fiabilité de notre modèle de détection de la dépression des utilisateurs.

La structure du mémoire

Ce mémoire est composé de 5 chapitres:

1. Le premier chapitre présente l'analyse des sentiments, ses domaines d'application et domaines connexes, ses différents types, et enfin certaines de ses techniques les plus utilisées.
2. Dans le deuxième chapitre, nous présentons l'apprentissage profond et ses notions de base, Après cela, nous avons illustré les différentes techniques d'apprentissage en profondeur utilisées dans l'analyse des sentiments.
3. Le troisième chapitre est consacré aux travaux liés à la détection de la dépression.
4. Dans le quatrième chapitre, nous présentons notre modèle de détection de dépression et nous explorons toutes ses étapes.
5. Dans le dernier chapitre, nous suivons notre processus d'expérimentation étape par étape pour élaborer, former et évaluer notre modèle. À la fin, nous partageons les résultats que nous avons obtenus.

Enfin, nous terminerons notre mémoire par une conclusion générale, dans laquelle nous partagerons les points d'apprentissage clés de cette initiative de recherche et les perspectives futures pour notre modèle proposé et comment il peut être amélioré et étendu pour éventuellement fournir de meilleurs résultats.

Analyse des Sentiments

Chapitre 1

Analyse des Sentiments

1.1 Introduction

« Ce que les autres pensent » a toujours été un élément d'information important pour la plupart d'entre nous pendant le processus de prise de décision. Dans notre vie quotidienne, nos comportements et décisions sont fortement influencés par les opinions des autres, dans la psychologie sociale ce phénomène s'appelle la conformité sociale. Consciemment ou non, chaque choix que nous faisons est basé sur une opinion antérieure.

Avec la naissance du web 2.0, les gens ont commencé à exprimer leurs opinions librement, publiquement et sous différentes formes, ce qui a augmenté le niveau d'intérêt, des organisations et des entreprises dans ces données en raison de sa valeur antérieure. Ces idées peuvent être très profitables aux entreprises et avoir un grand impact politique et économique sur la société. À partir de là, un nouveau domaine de recherche est né et est devenu le centre d'attention de la communauté de la recherche en informatique, il est appelé analyse des sentiments.

Dans ce chapitre, nous introduisons l'analyse des sentiments (également connu sous le nom d'exploration d'opinion), ses domaines d'application, ses différents niveaux et méthodes, et à la fin nous concluons avec les principaux défis rencontrés dans ce domaine.

1.2 Analyse des Sentiment et l'exploration d'opinions

Avant de passer à sa définition scientifique, analysons le mot « analyse des sentiments » lui-même et voyons sa définition littérale. Dans le dictionnaire, « sentiment » est défini comme une pensée, une opinion ou une idée fondée sur un sentiment au sujet d'une situation, ou une façon de penser de quelque chose [1] « analyse » est définie comme l'acte d'analyser quelque chose [2]. La combinaison de deux définitions, définira l'analyse des sentiments comme l'acte d'analyser des pensées, des opinions, ou des idées qui sont basés sur les sentiments, ou la façon de penser.

Bing Liu[3], un pionnier dans le domaine, définit l'analyse des sentiments comme suit :*L'analyse des sentiments, également appelé l'exploration d'opinion, est le domaine d'étude qui analyse les opinions, les sentiments des gens, les évaluations, les attitudes et les émotions envers des entités comme les produits, les services, les organisations, les personnes, les enjeux, les événements, les sujets et leurs attributs*[3].

Bien que les chercheurs utilisent plus souvent le terme « analyse des sentiments » que « recherche d'opinion », ils peuvent être utilisés de façon interchangeable parce qu'ils font référence au même domaine d'études. Les deux termes sont d'abord apparus vers 2003 [4] [5] mais la recherche dans ce domaine a commencé vers 2000 [6] [7] [9] [10] [11] et depuis, de nombreuses publications ont abordé le sujet [3].

1.2.1 Domaines d'applications de l'Analyse des Sentiments

Nous, les humains, avons tendance à penser que nous sommes libres dans nos choix et décisions, mais la vérité est que chaque décision que nous prenons est fortement influencée par les opinions des autres, consciemment ou inconsciemment.

Avec l'émergence du web participatif (Web 2.0), tous les utilisateurs d'Internet ont la possibilité de donner leurs avis sur une infinité de sujets (produits, services, prestations d'entreprise) et d'en discuter sur des forums, blogs, ou des groupes de discussions. Toutes ces opinions sont cruciales pour les entreprises, tout d'abord afin de développer des tâches de veille (technologique, marketing, concurrentielle, sociétale) mais aussi pour détecter les consommateurs mécontents, déterminer la perception des gens sur de nouveaux produits ou se renseigner sur la réputation de leur société. Toutes ces données peuvent aussi servir aux acteurs du monde politiques afin de connaître les intentions des citoyens.

Comme mentionné ci-dessus, en raison de sa relation et son importance dans tous les domaines de nos vies, l'analyse des sentiments est appliquée dans de nombreux domaines. Après avoir examiné différents documents de recherche et avoir vu des travaux antérieurs dans le domaine, nous nous sommes rendu compte que la plupart des applications d'analyse des sentiments appartiennent aux deux catégories suivantes : la détection et la prédiction.

1.2.1.1 Détection

La première catégorie d'analyse des sentiments est la détection cela veut dire qu'elle est utilisée pour détecter un problème, une maladie ou un besoin, selon l'industrie dans laquelle il est appliqué. Cette application consiste à étudier certains phénomènes ou de comprendre quelque chose qui s'est passé dans le monde et de trouver les raisons derrière elle, c'est pourquoi ils se concentrent principalement sur le passé et le présent des événements.

Nous citons quelques exemples qui sont liés à cette catégorie dans le tableau suivant :

Tableau 1.1 : Récapitulatif des études d'analyse des sentiments dans différents domaines.

Domaine	Etude	Explication
Commercial	[12], les revues ont été utilisées pour classer les produits et les marchands	Les sondages auprès des consommateurs montrent que les gens citent les examens de produits comme un facteur déterminant dans les décisions d'achat. Selon Terre-Neuve-et-Labrador, recommandations des consommateurs sont la forme de publicité la plus crédible parmi 78 % des répondants au sondage 2007[12].
Sportif	[13]les relations entre la ligne de pari de la NFL (The National Football League) avec l'opinion publique dans les blogs et Twitter ont été étudiées.	Dans cette étude il a été examiné les jeux de la NFL se sont déroulés pendant l'année 2006 à l'année 2009. Il a été constaté que le score sentimental a une corrélation significative avec les paris, tout comme les statistiques de performance objective. Il s'est également montré que les médias sociaux, comme les blogs et twitter, sont aussi informatifs comme la presse professionnelle écrite.
Politique	[14], les auteurs ont étudié les points de vue politiques.	modèle appliqué au domaine politique. Premièrement, basé sur la visualisations, des similitudes et les différences entre les sénateurs républicains et démocrates en ce qui concerne divers sujets .Deuxièmement, la comparaison de la performance des modèles de notation d'opinion avec 14 types de méthodes pour trouver les meilleurs. Il a été constaté que les phrases extraites par ces modèles de notation d'opinion peuvent exprimer efficacement les points de vue des opiniâtres.

Médiatique	[15] le flux de sentiments dans les réseaux sociaux a été étudié.	Explorer le flux de sentiments à travers des expériences sur un graphique contenant près de 8 millions de nœuds et 15 millions de bords. cette analyse indique que (1) les nœuds sont fortement influencés par leurs voisins immédiats.
Politique	[16]le sentiment de Twitter était lié aux sondages d'opinion publique.	Etablir un lien entre les mesures de l'opinion publique mesurées à partir des sondages et le sentiment mesuré à partir du texte, les corrélations peuvent atteindre 80 % et saisir d'importantes tendances à grande échelle. Les rapports soulignent le potentiel des flux de texte en tant que sous-ensemble et supplément pour le scrutin traditionnel

A savoir plusieurs d'autres travaux lié à cette catégorie ont été élaboré, notre sélection s'est basée sur les domaines d'applications les plus importants et les plus pertinents dans chaque domaine.

1.2.1.2 Prédiction

Bien que la détection joue un rôle très important dans l'application des prévisions, cette catégorie se distingue par le fait qu'elle met l'accent sur l'utilisation d'opinions détectées pour prédire les résultats futurs, ce qui est évidemment très nécessaire dans le monde des affaires et politique, et même dans le domaine de la santé. Nous citons quelques exemples qui appartiennent à cette catégorie dans le tableau ci-dessous:

Tableau : 1.2 Récapitulatif des études d'analyse des sentiments prédictives dans différents domaine

Domaine	Etude	Explication
Politique	[23], le sentiment Twitter a également été appliqué pour prédire les résultats des élections.	Dans cette étude était analysé plus de 100 000 messages Twitter mentionnant des partis ou des politiciens avant les élections fédérales allemandes de 2009. Dans l'ensemble, il a été constaté que Twitter est effectivement utilisé comme

		<p>plateforme pour les délibérations politiques. Le simple nombre de gazouillis reflète les préférences des électeurs et se rapproche des sondages électoraux traditionnels, tandis que le sentiment des messages sur Twitter correspond étroitement aux programmes politiques, aux profils des candidats, et la couverture médiatique de la campagne électorale.</p>
Politique	<p>[24], une méthode a été rapportée pour prédire les volumes de commentaires des blogs politiques.</p>	<p>Dans cette étude, il a été visionné de modéliser la relation entre le texte d'un billet de blogue politique et le volume de commentaires, il a été examiné la tâche de prédire quels post-blog auront un volume de réponse supérieur à la moyenne des lecteurs, mesuré en commentaires ou mots dans les commentaires. Les résultats montrent que la modélisation des sujets peut améliorer les rappels lorsque l'on publie des messages « à grand volume »</p>
Economique	<p>Dans [25], un modèle de sentiment a été proposé pour prédire le rendement des ventes.</p> <p>[26], [27] et [28], les données de Twitter, les critiques de films et les blogues ont été utilisés pour prédire les recettes au guichet des films</p>	<p>En raison de sa popularité élevée, Weblogs (ou blogs en bref) présentent une richesse d'informations qui peuvent être très utiles pour évaluer les sentiments et les opinions du grand public. Dans cette étude, il a été étudié le problème de l'information de sentiment minier à partir de blogs et d'étudier des façons d'utiliser ces informations pour prédire la performance des ventes de produits, Les expériences confirment l'efficacité et la supériorité de l'approche proposée.</p>

Industriel	[29], les humeurs de Twitter ont été utilisées pour prédire le marché boursier.	L'économie comportementale nous dit que les émotions peuvent affecter profondément le comportement individuel et la prise de décision. des états d'humeur collective dérivée des flux Twitter à grande échelle sont corrélés à la valeur de la moyenne industrielle Dow Jones (DJIA) au fil du temps.
Economique	[30], le sentiment du blogue et des nouvelles a été utilisé pour étudier les stratégies de négociation.	Le sentiment des nouvelles a une incidence sur le marché financier. Toutefois, trouver un prédicteur clair des rendements du marché à l'aide du sentiment des nouvelles demeure une tâche difficile. Cette étude examine la relation entre le sentiment d'actualité, les rendements cumulatifs du marché et la volatilité. Les résultats montrent que des niveaux anormaux de sentiment d'actualité sont importants pour prédire le rendement cumulé futur du marché et la volatilité.

Bien qu'il soit commencé principalement sur le langage écrit, l'analyse des sentiments est utilisée de nos jours sur des données visuelles et auditives ainsi. L'analyse des sentiments peut être considéré comme un sous-domaine de Traitement automatique du Langage Naturel (TAL) et même du Data Mining et il est récemment devenu l'un des domaines de recherche les plus actifs.

1.2.2 Exploration du Texte (Texte Mining)

L'exploration du texte (Texte Mining) a gagné du terrain en tant que domaine d'application important de l'extraction de données au cours des dernières années. L'industrie de l'extraction de textes est encore petite [31], mais elle croît rapidement, avec la prise de conscience des connaissances commerciales cachées dans des données non structurées. Le but de l'extraction de textes est de rassembler d'énormes quantités de données non structurées disponibles dans les entreprises et à l'extérieur, comme sur Internet, et de les utiliser pour obtenir des informations pour résoudre des problèmes du monde réel. UN des aspects centraux du textemining est de transformer

ces données textuelles peu structurées – si ce n’est par la langue utilisée – en données exploitables par les algorithmes classiques de data mining[32].

1.3 Niveaux de l’Analyse de sentiment

Trois différents niveaux sur lesquels l’analyse des sentiments doit être effectuée en fonction des granularités requises sont :

1.3.1 L’analyse des sentiments au niveau du document

Le niveau du document considère qu’un document est un avis sur une entité ou un aspect de celle-ci. Ce niveau est associé à la tâche appelée classification des sentiments au niveau du document [33], [34], [35]. Les tâches permettent de classer si une opinion complète documente un sentiment positif ou négatif. Le système détermine si l’examineur exprime une opinion globalement positive ou négative sur le produit. Cette tâche n’est connue que sous le nom de classification des sentiments au niveau du document. [36][37].

1.3.2 L’analyse des sentiments au niveau de la phrase

L’analyse du sentiment au niveau de la phrase est l’analyse la plus détaillée du document. Dans ce cas, la polarité est calculée pour chaque phrase, car chaque phrase est considérée comme une unité distincte et chaque phrase peut avoir une opinion différente. L’analyse des sentiments au niveau de la phrase comporte deux tâches :

Classification de la subjectivité : Une phrase peut être subjective ou objective. La phrase objective contient les faits. Il n’a pas de jugement ou d’opinion sur l’objet ou l’entité alors que la phrase subjective a des opinions [38]

Classification de sentiment : La phrase peut être classée comme positive, négative ou neutre selon les mots d’opinion qu’elle contient [38].

1.3.3 L’analyse des sentiments au niveau de la sous-phrase

Les analyses au niveau des documents et des phrases ne permettent pas de déterminer exactement ce que les gens aimaient ou n’aimaient pas. Le niveau d’aspect effectue une analyse plus fine. Le niveau d’aspect était auparavant appelé niveau d’élément (analyse et résumé de l’opinion fondée sur les éléments). Au lieu de regarder les concepts linguistiques (documents, paragraphes, phrases, clauses ou phrases), le niveau d’aspect regarde directement l’opinion elle-même. Elle repose sur l’idée qu’une opinion consiste en un sentiment (positif ou négatif) et en une

cible (d'opinion). Il est étroitement lié à des tâches telles que l'extraction d'opinions et la synthèse d'opinions [39], [40], [41].

1.4 Approches et techniques d'analyse de sentiments

Vu la croissance rapide et la popularité de l'analyse des sentiments dans les différentes études et domaines, plusieurs méthodes et techniques ont été mises en place. Dans cette section nous essayons de décrire brièvement quelques techniques et méthodes les plus pertinentes et les plus populaires utilisées dans notre étude.

1.4.1 L'exploitation des Emoticon

Les émoticônes sont devenues populaires ces dernières années, au point que certains sont maintenant inclus dans le dictionnaire English Oxford [1]. Les émoticônes sont principalement basés sur le visage et représentent des sentiments heureux ou tristes, La façon la plus simple de détecter la polarité (effet positif et négatif) d'un message. Par conséquent, les émoticônes ont souvent été utilisés en combinaison avec d'autres techniques pour construire un ensemble de données de formation sur les techniques d'apprentissage supervisé.

1.4.2 L'utilisation de Linguistic Inquiry and Word Count

LIWC (Linguistic Inquiry and Word Count) [42] est un outil d'analyse de texte qui évalue les composantes émotionnelles, cognitives et structurelles d'un texte donné en fonction de l'utilisation d'un langage contenant des mots et de leurs catégories. La LIWC ne permet pas de détecter les effets positifs et négatifs dans un contexte donné, mais fournit d'autres ensembles de catégories de sentiments il s'agit notamment de permettre aux utilisateurs d'inclure des indicateurs personnalisés plutôt que des indicateurs standard

1.4.3 Senticnet

Senticnet [43] est une méthode d'exploration de l'opinion et d'analyse du temps qui explore l'intelligence artificielle et les techniques Web. Le but de Senticnet est d'inférer la polarité des concepts de bon sens à partir de textes naturels à un niveau sémantique, plutôt qu'à la syntaxique. La méthode utilise les techniques de traitement du langage naturel (TAL) Senticnet a été testé et évalué comme un outil permettant de mesurer le niveau de polarité dans les opinions des patients sur le Na-National Health Service en Angleterre [44]. Les auteurs ont également recommandé Senticnet avec des données de blogs Live journal, où les messages ont été étiquetés par les auteurs avec plus de 130 humeurs, Puis catégorisé comme positif ou négatif [45], [46].

1.4.3 SentiStrength

Sentistrength est une technique qui met en œuvre une combinaison de méthodes d'apprentissage automatique [47]. La classification de base s'appuie sur l'ensemble des mots du dictionnaire LIWC [45] ajouté des fonctionnalités dédiées au contexte des réseaux sociaux en ligne (OSN). Ces fonctionnalités comprennent essentiellement une liste de mots négatifs et positifs avec une liste de boosters (comme « très » pour renforcer et « quelque peu » à affaiblir). L'utilisation d'émoticônes et de ponctuations répétées a été utilisée pour le même but ainsi.

1.5 Les défis de l'analyse de Sentiment

Aujourd'hui l'analyse de sentiments se concentre sur l'attribution d'une polarité à des expressions subjectives (les mots et les phrases qui expriment des opinions, des émotions, des sentiments, etc.) afin de décider de l'orientation d'un document ou de la polarité positive/négative/neutre d'une opinion dans un document. Des travaux allant au-delà ont mis l'accent sur la force d'une opinion d'expression où chaque proposition dans une phrase peut avoir un fond neutre, faible, moyen ou élevé .Cependant, L'analyse de sentiments exprimés dans les groupes de discussions, les blogs, les forums et autres sites spécialisés heurte à quelques difficultés:

Une des difficultés de la classification en « positif et négatif » réside dans la nécessité d'une bonne analyse syntaxique du texte, analyse qui peut se révéler particulièrement difficile dans des cas de coordination entre plusieurs parties d'une phrase, d'anaphore ou de coréférence.

Une autre difficulté du langage naturel pour l'analyse automatique de sentiments est les contextes intentionnels, pour lesquels l'expression d'opinion n'est pas un vrai sentiment. C'est le cas dans une phrase comme : « Je croyais que l'Algérie était un beau pays. » ont montré le lien qui existe entre les structures syntaxiques et sémantiques d'une phrase et l'expression de l'opinion qu'elle véhicule. Ainsi l'analyse de la phrase par « paquets de mots » donne des résultats peu satisfaisants alors qu'une analyse syntaxique du texte peut aider à trouver les expressions qui contiennent des opinions. Les deux phrases suivantes contiennent les mêmes paquets de mots sans pour autant exprimer les mêmes sentiments. En effet, la première phrase contient un sentiment positif alors que la deuxième est négative : « Je l'ai apprécié pas ».

1.6 Traitement Automatique Du langage (TAL)

L'analyse des sentiments est un sous-domaine du traitement automatique du langage (TAL), par conséquent, tous les problèmes du(TAL), que la plupart d'entre eux ne sont pas encore résolus, s'appliquent à lui et apportent plus de difficulté sur le terrain [41]. Certains de ces problèmes sont, la résolution de coréférence, la manipulation de négation, et les dés ambigüité de sens de mot, synonyme. Toutefois, il est également utile de se rendre compte que l'analyse des sentiments est

très restreinte le problème de TAL et que le système n'a pas besoin de comprendre pleinement la sémantique de chaque phrase ou document, mais seulement le besoin de comprendre certains aspects de celui-ci, c.-à-d., les sentiments positifs ou négatifs et leurs entités ou sujets cibles.

Les chercheurs ont maintenant une meilleure compréhension de l'ensemble du problème, sa structure et les questions fondamentales. De nombreux nouveaux modèles et méthodes (formels) ont été proposés.

La recherche antérieure dans le domaine principalement axé sur la classification du sentiment ou de la subjectivité exprimé dans des documents ou des phrases, qui sont insuffisant pour la plupart des applications réelles.

1.7 Conclusion

Dans ce chapitre, nous avons présenté les bases fondamentales de l'analyse des sentiments et de l'exploration d'opinion ainsi que les différentes techniques et méthodes utilisées et leurs principaux obstacles qui se posent sur le terrain.

Dans le chapitre suivant nous allons introduire le Deep Learning et les différents travaux liés à ce domaine d'étude.

Deep Learning

Chapitre2 Deep Learning

2.1 Introduction

Depuis toujours, l'homme a cherché à réduire ses efforts et à améliorer sa productivité dans tous les domaines en se faisant aider de machines. La première invention marquante ayant pour beaucoup révolutionner l'humanité en ce sens était la roue, c'est-à-dire une reproduction mécanique du roulement. La première pièce d'un robot de bois en quelque sorte.

Depuis l'invention des ordinateurs programmables, les humains rêvent de robots et de machines capables de penser et de faire des activités quotidiennes pour faciliter leur vie. Aujourd'hui, avec l'intelligence artificielle (IA), de tels rêves sont devenus une réalité et le domaine est devenu l'un des domaines de recherche les plus populaires qui est florissant jour après jour, tout en contestant toutes les barrières précédentes et établissant une nouvelle barre pour l'avenir de l'humanité. Aujourd'hui, l'IA se déploie partout et s'impose dans de nombreux domaines (jeux vidéo, automobile, aéronautique, conquête spatiale, techniques médicales, etc.).

Dans ce chapitre, nous introduisons l'apprentissage profond (Deep Learning, l'un des domaines de recherche les plus populaires de l'IA. Nous passons ensuite aux différentes classes et techniques en s'appuyant sur des exemples d'application réelle.

2.2 Définitions

Nous débuterons par la définition des termes les plus importants et les plus utilisés durant le reste de ce chapitre afin de faciliter la compréhension des prochaines séquences.

2.2.1 Intelligence artificielle (IA)

Bien que le terme Intelligence artificielle puisse sembler explicite, il existe de nombreuses définitions. Nous utiliserons celle de Poole, Mack Worth & Goebel dans notre travail, où elle se réfère à l'IA comme intelligence computationnelle et la définit comme suit : L'intelligence computationnelle est l'étude de la conception d'agents intelligents [49].

Le terme « agent » peut désigner tout ce qui agit dans un environnement particulier, tout ce qui fait quelque chose, que ce soit les humains, les animaux, les machines... etc.

Un agent intelligent d'autre part est un système qui agit intelligemment. Comme l'expliquent les chercheurs : il est approprié pour son objectif, il est flexible face à des environnements

changeants, des objectifs changeants, il apprend de l'expérience, et il fait des choix appropriés compte tenu des limitations et des calculs finis [49]

2.2.2 Apprentissage Automatique (Machine Learning).

L'apprentissage automatique est un sous-domaine de l'intelligence artificielle (IA) [50]. En général, l'objectif de l'apprentissage automatique est de comprendre la structure des données et de les intégrer dans des modèles qui peuvent être compris et utilisés par tout le monde [51].

Bien que l'apprentissage automatique soit un domaine de l'informatique, il diffère des approches informatiques traditionnelles. En effet dans cette dernière, les algorithmes sont des ensembles d'instructions explicitement programmées utilisées par les ordinateurs pour calculer ou résoudre des problèmes. Les algorithmes d'apprentissage automatique permettent aux ordinateurs de s'entraîner sur les entrées de données et utilisent l'analyse statistique pour produire des valeurs qui se situent dans une plage spécifique. Pour cette raison, l'apprentissage automatique facilite l'utilisation des ordinateurs dans la construction de modèles à partir de données d'échantillonnage afin d'automatiser les processus de prise de décision en fonction des données saisies.

2.2.3 Méthodes d'apprentissage automatique

Dans l'apprentissage automatique, les tâches sont généralement classées en grandes catégories. Deux méthodes d'apprentissage automatique les plus adoptées sont **l'apprentissage supervisé** qui forme des algorithmes basés sur des données d'entrée et de sortie étiquetées par l'homme et **l'apprentissage non supervisé** qui ne fournit pas à l'algorithme des données étiquetées pour lui permettre de trouver une structure et de découvrir une logique dans données entrées. Explorons donc ces méthodes plus en détail.

2.2.4 L'apprentissage supervisé

Dans l'apprentissage supervisé, l'ordinateur est fourni avec des exemples d'entrées qui sont étiquetés avec les sorties souhaitées. Le but de cette méthode est que l'algorithme puisse «apprendre» en comparant sa sortie réelle avec les sorties «enseignées» pour trouver des erreurs et modifier le modèle en conséquence. L'apprentissage supervisé utilise donc des modèles pour prédire les valeurs d'étiquettes sur des données non étiquetées supplémentaires.

2.2.5 L'apprentissage non supervisé

Dans l'apprentissage non supervisé, les données sont non étiquetées, de sorte que l'algorithme d'apprentissage trouve tout seul des points communs parmi ses données d'entrée. Les données non étiquetées étant plus abondantes que les données étiquetées, les méthodes d'apprentissage automatique qui facilitent l'apprentissage non supervisé sont particulièrement utiles.

2.2.6 Réseaux de Neurones Artificiels

De nos jours, on ne peut pas parler de machine Learning sans parler de réseau de neurones artificiels, un ensemble d'algorithmes dont le fonctionnement est inspiré des neurones biologiques et qui s'appuie également sur les méthodes statistiques. L'usage de réseaux de neurones est une **technique de l'Intelligence Artificielle**, optimisant les méthodes d'apprentissage automatiques [52].

Il existe plusieurs types de réseaux neuronaux, dont chacun vient avec ses propres cas d'utilisation spécifiques et niveaux de complexité. Le type le plus basique de réseau neuronal est appelé réseau neuronal feedforward, dans lequel l'information voyage dans une seule direction de l'entrée à la sortie. Un type de réseau plus largement utilisé est le réseau neuronal récurrent, dans lequel les données peuvent circuler dans plusieurs directions.

Ces réseaux neuronaux possèdent de plus grandes capacités d'apprentissage et sont largement utilisés pour des tâches plus complexes comme l'apprentissage de l'écriture ou de la reconnaissance linguistique. Il existe également d'autres réseaux neuronaux convolutionnels, Boltzmann réseaux de machines, réseaux Hopfield, et une variété d'autres.

Le bon choix du réseau de neurone dépend des données à former avec, et l'application spécifique. Dans certains cas, il peut être souhaitable d'utiliser des approches multiples, comme ce serait le cas avec une tâche difficile comme la reconnaissance vocale [52].

2.2.7 L'apprentissage profond (Deep Learning)

Adopté à partir de l'apprentissage de représentation, l'apprentissage profond est un ensemble de méthodes qui permet à une machine d'apprendre les données brutes qui lui sont transmises et extraire automatiquement les représentations nécessaires à la détection ou la classification, et elle le fait par le biais de multiples niveaux de représentation. La représentation est transformée d'un niveau à un niveau supérieur, plus abstrait, à travers des modules composés,

simples, mais non linéaires. Avec la composition d'assez de telles transformations, des fonctions très complexes peuvent être apprises [53].

Le deep learning est associé au fonctionnement des réseaux neuronaux artificiels. En principe, c'est un humain qui doit apprendre à la machine à reconnaître les caractéristiques communes à un objet. Mais ce n'est plus trop le cas avec la technique d'apprentissage profond. En effet, c'est la machine qui va analyser les objets avant de décider, afin d'identifier leurs points communs. L'homme valide ensuite.

Nous prenons une image comme exemple, une image se présente sous la forme d'un tableau de valeurs de pixels et les caractéristiques apprises.

1. Dans la première couche de représentation représentent généralement la présence ou l'absence de bords à des orientations et des emplacements particuliers dans l'image.
2. La deuxième couche détecte généralement les motifs en repérant des dispositions particulières des bords, indépendamment des petites variations dans les positions des bords.
3. La troisième couche peut assembler des motifs de plus grandes combinaisons qui correspondent à des parties d'objets familiers, et les couches suivantes détecteraient des objets comme des combinaisons de ces parties.

L'aspect clé de l'apprentissage profond est que ces couches de caractéristiques ne sont pas conçues par des ingénieurs humains : ils sont appris de données en utilisant une procédure d'apprentissage générale [53]

Depuis ses débuts, l'apprentissage profond réalise et réalisera de nouveau progrès record dans la résolution de problèmes qui n'ont pu être résolus pendant de nombreuses années par la communauté de l'intelligence artificielle. Il s'est avéré comme moyen important pour découvrir des structures complexes dans les données de haute dimension, il est applicable dans de nombreux domaines de la science, des affaires et du gouvernement. En plus de battre des records de reconnaissance d'image [54] [55] [56][57] et de reconnaissance vocale [58] [59] [60], elle a battu d'autres techniques d'apprentissage automatique pour prédire l'activité des molécules médicamenteuses potentielles [61], en analysant des données sur les accélérateurs de particules[62] [63], la reconstruction des circuits cérébraux [64], et la prévision des effets des mutations de l'ADN non codant sur l'expression des gènes et les maladies [65] [66]. Plus surprenant peut-être,

l'apprentissage profond a produit des résultats extrêmement prometteurs pour diverses tâches de compréhension du langage naturel [67], en particulier la classification des sujets, l'analyse des sentiments, la réponse aux questions [68] et la traduction du langage [69] [70].

Les chercheurs et les praticiens prédisent un brillant avenir pour l'apprentissage profond avec beaucoup plus de succès, et cela est dû à sa capacité de tirer parti de l'augmentation de la quantité de données et de calculs disponibles. De nouveaux algorithmes et architectures d'apprentissage sont constamment développés pour les réseaux neuronaux profonds afin d'accélérer ce progrès.

2.2.8 Les classes de l'apprentissage profond

Dans cette section, nous présenterons les trois classes d'apprentissage profond par la suite nous partagerons les types de réseaux neuronaux utilisés avec une brève explication ainsi que quelques exemples d'applications.

2.2.8.1 Apprentissage supervisé

Les réseaux profonds d'apprentissage supervisé, également connus sous le nom de réseaux profonds discriminatoires, sont destinés à fournir une puissance discriminante à des fins de classification de modèle, principalement en caractérisant les distributions postérieures des classes conditionnées par les données visibles. Les données d'étiquettes cibles sont toujours la perspective de régularisation où les modèles d'apprentissage non supervisés peuvent effectivement fournir une priorité sur l'ensemble des fonctions représentées par le modèle [71].

Un CRF (conditional random field) est intrinsèquement une architecture discriminante superficielle, caractérisée par la relation linéaire entre les fonctions d'entrée et les fonctions de transition. Son caractère superficiel est dû à l'équivalence établie entre le Trésor lui-même et les modèles gaussiens et les HMM (hiddenMarkov model) ayant reçu une formation discriminatoire [72]. Ces derniers temps, des CRF structurés en profondeur ont été développés en empilant la sortie de chaque couche inférieure du CRF, ainsi que la date d'entrée originale, sur la couche supérieure [73]. Il a été appliqué avec succès à la reconnaissance téléphonique [74], à l'identification de la langue parlée [73] et au langage naturel [73].

2.2.8.2 Apprentissage non supervisé

Les réseaux profonds pour l'apprentissage non supervisé, également connu sous le nom d'apprentissage génératif, visent à saisir la corrélation de haut niveau des données observées, lorsqu'aucune information sur les étiquettes de classe cible n'est disponible, à des fins d'analyse de

modèle. Lorsqu'ils sont utilisés en mode génératif, ces réseaux ont la capacité de caractériser les distributions statistiques conjointes des données visibles et de leurs classes associées, lorsqu'elles sont disponibles, et d'en faire partie intégrante.

Formellement, l'apprentissage non supervisé est défini comme n'utilisant aucune supervision spécifique à la tâche d'information dans le processus d'apprentissage. De nombreux réseaux profonds dans cette catégorie, comme les RBM (Restrict Boltzmann machines), les DBN (Deep Belief network) et les auto-encodeurs de dénouement généralisés [75], peuvent être utilisés pour de manière significative en produisant des échantillons à partir des réseaux. D'autre part, certains réseaux dans cette même catégorie, comme les réseaux de codage clairsemés et la forme originale des auto-encodeurs profonds, ne peuvent être échantillonnés facilement et ne sont donc pas de nature générative.

Parmi les différentes sous-classes de réseaux profonds génératifs ou non supervisés, les modèles profonds à base d'énergie, comme la forme originale de l'auto-encodeur profond [76] [77] [78], sont les plus courants [76] [79] [80] [81]. La plupart des autres formes d'auto-encodeurs profonds ne sont pas supervisées dans la nature, mais avec des propriétés et des implémentations très différentes.

2.2.8.3 Apprentissage Hybride (Hybrid Learning)

Réseaux profonds hybrides, où l'objectif est la discrimination qui est aidée, souvent de manière significative, avec les résultats de réseaux profonds génératifs ou non supervisés. Ceci peut être accompli par une meilleure optimisation ou/et régularisation des réseaux profonds de la deuxième catégorie (apprentissage non supervisé). Le terme « hybride » pour cette troisième catégorie fait référence à l'architecture profonde qui comprend ou utilise à la fois des composantes génératives et discriminatoires du modèle [82].

Un exemple de réseaux profonds hybrides est l'utilisation de modèles génératifs de DBN (Deep Belief Network) pour préformer des réseaux neuronaux convolutionnels profonds (CNN profonds) [84][85] [86]. Comme le DNN (Deep Neurone Network) entièrement connecté discuté plus tôt, la préformation contribue également à améliorer les performances des CNN profonds sur l'initialisation aléatoire. La formation préalable des DNN ou des CNN à l'aide d'un ensemble d'auto-encodeurs profonds régularisés [87], y compris les auto-encodeurs de dénouement, les auto-encodeurs contractifs et les auto-encodeurs épars, est également un exemple similaire de la catégorie des réseaux profonds hybrides.

2.3 Analyse de Sentiment utilisant les techniques d'apprentissage profond

Dans cette partie nous décrivons le maximum de techniques d'apprentissage profond couvrant l'analyse des sentiments.

2.3.1 Réseaux de Neurones convolutifs (CNN)

Les CNN (réseaux neuronaux convolutifs) [87] sont l'un des types de neurones les plus populaires, ils comprennent la mise en commun des couches et la sophistication car il donne une architecture standard pour cartographier les phrases de longueur variable en phrases de vecteurs dispersés de taille fixe. Cette étude [88] a proposé un nouveau cadre de réseau neuronal convolutifs (CNN) pour l'analyse visuelle des sentiments, prédire les sentiments du contenu visuel. L'approche d'apprentissage par transfert et l'hyper-paramètre ont été utilisée dans Googlenet. Pour le travail expérimental, un ensemble de données de twitter contenant 1269 images est sélectionné et la propagation arrière est appliquée.

Amazon Mechanical Turk (Mturk) et l'intelligence populaire de foule [89] est utilisé pour étiqueter les images. Cinq travailleurs ont été impliqués pour générer un label sentiment en faveur de chaque image. Le modèle proposé a été évalué sur cet ensemble de données et a acquis de meilleures performances que les systèmes existants. Les résultats montrent que le système proposé est de haute performance sans réglage précis sur le jeu de données Flickr. Cependant, Alexnet a été utilisé dans des travaux précédents et GoogleNet a fourni presque 9% de progrès de performance qu'AlexNet. Pour convertir Googlenet dans un cadre d'analyse du sentiment visuel, une meilleure extraction des caractéristiques a été atteinte. Des états stables et fiables ont été obtenus en utilisant des paramètres hyper.

La recherche détaillée de [89] a présenté un aperçu de l'analyse des sentiments liés aux micro-blogs. Le but de cette dernière était d'obtenir les opinions et les attitudes des utilisateurs sur les événements chauds en utilisant le réseau neuronal convolutifs (CNN).

L'utilisation de CNN résout les problèmes de l'extraction explicite des caractéristiques et apprend implicitement grâce aux données de formation. Pour collecter les données de la cible, l'URL d'imputation et le crawler ciblé ont été utilisés, 1000 commentaires de micro-blog ont été recueillis comme un corpus et divisé en trois étiquettes, c.-à-d., 274 émotions neutres, 300 émotions négatives et 426 émotions positives.

Dans l'étude[91], un modèle comprenant RNTN (Recursive Neural Tensor Network) et Sentiment Treebank a été proposé pour clarifier correctement les effets de composition des phrases à différents niveaux, c.-à-d. les phrases positives et négatives modèles existants. Dans les modèles

existants, la signification des phrases longues ne peut pas être exprimée efficacement par espaces de mots sémantiques, donc pour la détection des sentiments, plus riche et supervisé l'évaluation et la formation

ressources sont nécessaires car il faut des modèles de composition plus influents. Le RNTN atteint 80,7% précision dans la prédiction du sentiment en effectuant un étiquetage à grain fin sur toutes les phrases et surperformé les modèles précédents.

L'étude [92] a apporté un cadre généralisé pour reconnaître les vendeurs de logiciels malveillants et de cartes à cartes. Le modèle repose sur un apprentissage approfondi pour l'analyse des sentiments et est utilisé dans la classification des fils et l'échantillonnage en boule de neige pour évaluer la qualité du service/produit des vendeurs en analysant les commentaires des clients. L'évaluation du modèle proposé a été réalisée sur un forum de cardage russe et un crawler web a été utilisé pour approfondir la conversation à partir du forum.

Un sentiment Treebank a été utilisé et il a été formé en utilisant le réseau de tenseur neural récurrent sur un corpus de revue en ligne. Pour évaluer la validité et l'efficacité, deux expériences ont été menées dans le cadre desquelles le modèle proposé a été comparé aux modèles basés sur Naive Bayes, KNN et SVM. Cette étude a cherché les vendeurs qui sont fortement notés pour les services/produits malveillants et l'efficacité de l'apprentissage approfondi dans la reconnaissance de ces vendeurs. Les résultats ont indiqué que les techniques d'apprentissage profond obtiennent des résultats supérieurs à ceux des classificateurs peu profonds et il a été établi que les vendeurs de cartes ont moins de cotes que les vendeurs de logiciels malveillants

2.3.2 Réseaux De Neurones Récurrents(RNN)

Le réseau neuronal récurrent (RNN) [86] est un modèle influent dans la modélisation du langage car il ne représente pas le contexte de longueur fixe qui contamine tous les mots d'histoire. Dans cette étude [96], le HBRNN (hiérarchique bidirectionnel récurrents neurone) réseau neuronal bidirectionnel hiérarchique récurrent) a été développé pour extraire les avis des clients sur différents hôtels d'une manière complète et concise. Pour modéliser l'information séquentielle à long terme, HBRNN a utilisé la terminologie du RNN et le processus de production a été effectué au niveau de l'examen par HBRNN. Les données expérimentales ont été tirées de DBS textmining Challenge 2015[96]. Le rendement de HBRNN a été amélioré grâce aux paramètres fin et le modèle a été comparé avec LSTM (mémoire à court terme) et BLSTM (LSTM bidirectionnel). Après avoir effectué les expériences, le rappel d'évaluation, les scores, F1 et données très biaisées. Le

développement, l'ensemble d'essais et la comparaison des résultats avec les systèmes de référence, dix fois la validation croisée utilisée pour présenter la performance de HBRNN.

Les principaux défis qui ont été résolus sont le manque d'examens en ligne de grande qualité et le manque de forte asymétrie dans les données examinées. Les résultats expérimentaux de l'ensemble de données ont prouvé que l'HBRNN donnait de meilleurs résultats que d'autres méthodes. Ce modèle peut être appliqué à d'autres activités d'exploration d'opinion qui consiste en un énorme volume de données.

Cette contribution [86] a été faite pour surmonter la question d'un ensemble de données de Bangla car il est standard et large pour les tâches SA (Sentiment Analysis). Le problème a été résolu en fournissant un ensemble de données important pour l'analyse des sentiments de 10 000 BRBT (Bangla et Romanized Bangla Text) [86] a été utilisé pour tester l'ensemble de données en utilisant deux fonctions de perte, c'est-à-dire l'entrée croisée binaire et catégorique. Les données ont été recueillies à partir de différents sites comme YouTube, Facebook, Twitter et d'autres. Les expériences ont été menées pour préparer l'ensemble de données d'une marque pour une autre (et l'inverse) pour examiner le fait qu'il contribue à de meilleurs résultats.

Les chercheurs [97] ont proposé un modèle séquentiel pour mettre l'accent sur l'intégration des examens ayant une nature temporelle aux produits, puisque ces examens portaient moins sur les études existantes. La combinaison d'unités récurrentes synchronisées avec des neurones récurrents unités avec réseau neuronal récurrent est utilisée pour apprendre des représentations dispersées des produits et des utilisateurs. Pour la classification des sentiments, ces représentations sont intégrées dans le classificateur d'apprentissage automatique. L'approche a été évaluée à partir de trois ensembles de données recueillis auprès de Yelp et de la BMDI. Chaque examen a été étiqueté selon la cote attribuée. Pour former le réseau, l'algorithme de rétro-propagation avec la méthode d'optimisation stochastique d'Adam a été utilisé. Les résultats montrent la modélisation séquentielle de la représentation dispersée des produits et des utilisateurs, l'apprentissage améliore le rendement, la classification du sentiment au niveau des documents et l'approche proposée présentée sous forme de tableaux.

2.3.3 Réseaux De Neurones Profond (DNN)

L'étude [93], l'auteur a proposé un modèle d'analyse des sentiments tenant compte contenus textuels des réseaux sociaux. Ce nouveau schéma a utilisé un modèle de réseau neuronal profond tel quels encodeurs automatiques et le skip-gram. La base du schéma était le modèle CBOW

(Continuous Bag-Of Words). Le modèle proposé se composait de deux parties CBOW-LR (régression logistique) pour le texte et a été élargi comme le CBOW-DA-LR. La classification a été faite selon la polarité de l'information visuelle et textuelle. Quatre ensembles de données ont été évalués, soit Sanders Corpus, Sentiment140, Semeval- 2013 et Sentibank Twitter. Le modèle proposé a surpassé le CBOW+SVM et le FSLM (modèle linguistique probabiliste entièrement supervisé). Peut-être que l'ESLAM (modèle linguistique probabiliste entièrement supervisé) en termes de données de formation de petite taille avait surpassé le modèle actuel. La fonction d'apprentissage et de saut grammes ont tous deux nécessité de grands ensembles de données pour une meilleure performance.

L'étude [94], une architecture de réseau neuronal profond a été proposée pour évaluer la similitude des documents. L'architecture a été formée en utilisant plusieurs nouvelles du marché pour produire des vecteurs pour les articles. Le T&C nouvelles a été utilisé comme un ensemble de données. La similitude cosinus a été calculée entre les articles étiquetés et la polarité des documents a été prise en considération, mais le contenu n'a pas été pris en considération. La méthode proposée a permis d'obtenir un rendement supérieur en termes d'estimation de la similarité des articles selon la polarité.

2.3 .4 Réseaux De Croyance Profond (deepbelief networks)DBN

L'analyse automatique des sentiments exprimés dans les évaluations en ligne à grande échelle est très importante pour les applications commerciales intelligentes [98]. La classification par sentiment est la tâche la plus populaire de l'analyse des sentiments, ce qui est plus difficile que la classification traditionnelle par sujet. Les caractéristiques de base, comme les mots de vocabulaire, ne suffisent pas à bien classer les sentiments.

Deep Belief Network (DBN) est présenté pour découvrir des caractéristiques plus abstraites des sentiments. Pour saisir toute l'information sur les caractéristiques, il est possible de construire un réseau de grande taille, mais en même temps, le réseau de grande taille a tendance à trop s'adapter aux données de formation et même au bruit, ce qui réduira la capacité de généralisation du réseau. L'étude [98], L2-norme Deep Belief Network (L2DBN) est proposé, qui utilise la

régularisation L2-norme pour optimiser les paramètres de réseau de DBN. L2DBN est d'abord initialisé par un algorithme d'entraînement non supervisé, puis affiné par une procédure supervisée. Les paramètres du réseau sont optimisés en utilisant à la fois la perte de classification et la complexité du réseau. Les résultats expérimentaux montrent que le L2DBN proposé surpasse la méthode de pointe et le DBN de base sur des ensembles de données dorés, bruyants et hétérogènes.

L2DBN est proposé pour résoudre la tâche de classification des sentiments. Il peut obtenir des caractéristiques abstraites de haut niveau à partir des données d'entrée et obtenir une meilleure capacité de généralisation. Après une formation préalable et une mise au point discriminante avec la régularisation de la norme L2, le dernier réseau profond est plus puissant pour la tâche de classification des sentiments. Les résultats expérimentaux montrent que le L2DBN surpasse le DBN de base et le classificateur de pointe sur quatre ensembles de données d'examen de produits. L'amélioration est beaucoup plus importante pour les données bruyantes et hétérogènes. Nous pouvons conclure que L2DBN est plus approprié pour résoudre des problèmes dans des applications réelles avec des données bruyantes et hétérogènes. La bonne performance de L2DBN sur les données hétérogènes indique également le potentiel de L2DBN pour l'apprentissage du transfert de domaine, [63].

2.3.5 Réseaux De Neurones Hybrides

L'étude [101] a proposé deux techniques d'apprentissage profond pour la classification des sentiments des données Twitter thaïlandaises, c.-à-d., le réseau neuronal convolutifs (DCNN) et la mémoire à court terme (LSTM). Le traitement des données a été effectué correctement. Les données ont été recueillies auprès des utilisateurs et de leurs abonnés de Thai Twitter. Après avoir filtré les données, seuls les utilisateurs ayant des tweets thaïlandais et des tweets avec des caractères thaïlandais ont été sélectionnés. Cinq expériences ont été menées pour obtenir les meilleurs paramètres pour l'apprentissage profond, pour comparer l'apprentissage profond avec les techniques classiques et pour atteindre l'importance de la séquence des mots. Une triple validation croisée a été utilisée pour vérifier le processus. Les résultats ont conclu que la précision est élevée en DNN par rapport à la LSTM et les deux techniques d'apprentissage profond sont plus précises que SVM et Naïve Bayes mais inférieure à l'entropie maximale. Des précisions plus élevées ont été trouvées dans les phrases originales plutôt que de mélanger des phrases pour que la séquence de mots soit importante.

La recherche [102], un modèle hybride a été proposé qui consiste en un réseau neuronal probabiliste (PNN) et RBM (Restrictive Boltzmann machines) à deux couches. Le but de ce modèle

hybride d'apprentissage profond est d'obtenir une meilleure précision de la classification des sentiments. La polarité, c'est-à-dire les critiques négatives et positives varient selon le contexte afin de résoudre ce type de problème, ce modèle fonctionne bien, les examens neutres ne sont pas considérés. Des expériences ont été faites avec des ensembles de données de Pang et Lee et Blitzer, la classification binaire mise en œuvre sur chaque ensemble de données. La précision a été améliorée pour cinq ensembles de données en le comparant à l'état actuel de la technologie [97]. Il n'y a pas d'autres ressources dans l'approche proposée telle que POS tagger et dictionnaire de sentiment, etc. par conséquent, il est plus rapide que les concurrents. Pour atteindre un nombre réduit de caractéristiques, la réduction de la dimension a été mise en œuvre, car une étude précédente a utilisé une stratégie complexe pour la sélection des caractéristiques.

2.4 Conclusion

Dans ce chapitre nous avons décrit les principaux fondements, méthodes et techniques de classification des sentiments. Par conséquent, nous examinerons plus précisément dans le prochain chapitre certains des travaux réalisés à ce sujet ainsi que des études intéressantes sur l'apprentissage en profondeur qui nous ont aidés à proposer une nouvelle approche dans notre recherche.

Détection de la dépression dans les réseaux sociaux : état de l'art

Chapitre 3 Détection de la dépression dans les réseaux sociaux : Etat de l'art

3.1 Introduction

Comme dans tous les projets de recherches, l'étude des travaux antérieurs effectués dans le même domaine et pour résoudre le même problème joue un rôle très important dans l'utilisation de la bonne approche ainsi que la recherche de la solution la plus appropriée. Par conséquent, dans le présent chapitre, nous avons essayé de présenter les points saillants de notre analyse documentaire.

Nous avons commencé par une illustration de quelques statistiques sur la santé mentale et la dépression afin de mettre en évidence notre thème de recherche. Nous sommes approfondies dans notre problématique en partageant les principaux obstacles trouvés dans la détection de la dépression et les méthodes classiques de détection.

Dans la deuxième partie de ce chapitre, nous avons présenté les différents travaux similaires élaborés pour résoudre le problème de détections de dépression.

3.2 Les troubles Dépressifs

La dépression est une affection courante dans le monde qui concerne 300 millions [103] de personnes selon les estimations. Elle diffère des sautes d'humeur habituelles et des réactions émotionnelles passagères face aux problèmes du quotidien. Quand elle perdure et que son intensité est modérée ou sévère, la dépression peut devenir une maladie grave. Elle peut entraîner une grande souffrance, altérer la vie professionnelle, scolaire et familiale de la personne concernée. Dans le pire des cas, la dépression peut conduire au suicide. Chaque année, près de 800 000 [103] personnes meurent en se suicidant. Le suicide est la deuxième cause de mortalité chez les 15-29 ans [103].

Bien qu'il existe des traitements connus et efficaces pour combattre la dépression, moins de la moitié des personnes affectées dans le monde (dans certains pays, moins de 10%) bénéficient de tels traitements. Le manque de ressources, la pénurie de soignants qualifiés et la stigmatisation sociale liée aux troubles mentaux sont autant d'obstacles à l'administration de soins efficaces.

Dans tous les pays, quel que soit leur niveau de revenu, les personnes déprimées ne sont pas souvent diagnostiquées correctement alors que d'autres, non affectées par ce trouble, se voient parfois prescrire des antidépresseurs par suite d'une erreur de diagnostic. La charge de la dépression et des autres pathologies mentales est en augmentation dans le monde. Une résolution adoptée en mai 2013 par l'Assemblée mondiale de la Santé préconisait une réponse globale coordonnée au niveau des pays pour faire face aux troubles mentaux [104].

On se référant à [104] près de 300 millions de personnes souffrent de dépression clinique dans le monde, ce qui est encore plus effrayant. Selon [104] c'est que 3 à 5 % des hommes et 8 à 10 % des femmes de la population mondiale totale sont susceptibles de vivre un épisode dépressif majeur dans une période d'un an. La dépression coexiste souvent avec d'autres maladies et troubles mentaux. Selon [105], elle touche un quart des patients atteints de cancer, un tiers des survivants d'une crise cardiaque et jusqu'à 75 % des personnes diagnostiquées d'un trouble de l'alimentation. Encore plus absurdes conditions et les facteurs ont été associés à la dépression, et combinés ensemble, ils pèsent généralement lourdement sur les patients qui améliorent leur qualité de vie et celle de leurs pairs lourds [106].

Ces dernières années, les communautés de santé mentale en ligne sont devenues des ressources de premier plan pour le soutien en santé mentale [106]. En fait, le soutien provenant de ces communautés a permis d'améliorer de façon causale le bien-être mental comme une réduction de la probabilité de pensées suicidaires [107]. Ce soutien peut aller du soutien affectif (SE) au soutien informationnel (SI), en prenant souvent la forme d'empathie, de reconnaissance, de conseils ou d'une évaluation situationnelle sur divers enjeux comme la maladie mentale, la crise, la toxicomanie et la violence [108, 109, 110]. En outre, en raison de la grande qualité du soutien fourni par ces OMHCS (Online Mental Health Communities), Ils sont également considérés comme un « refuge » : ils permettent aux personnes d'exprimer des émotions désinhibantes, de s'auto-dévoiler d'expériences stigmatisées et d'établir des relations de confiance avec leurs pairs [107, 109, 110,111].

Les gens obtiennent souvent une conformité énorme avec une communauté par des moyens linguistiques : « Lorsque les gens interagissent, ils adaptent leur discours, leurs schémas vocaux et leurs gestes, pour s'adapter aux autres ». Plus précisément, les conventions d'une communauté sont généralement adoptées au fil du temps et se manifestent par des règles de communication ou des styles linguistiques établis par ses membres [112, 113, 115, 116]. De tels accommodements sont liés à l'amélioration de la rétroaction sociale, à une solidarité accrue, à de meilleurs échanges sociaux et à des sentiments réciproques d'intimité [114]. Bien que ces observations aient été validées dans les communautés en ligne en général [115, 117, 118], il existe un manque de preuves empiriques similaires indiquant que dans les OMSC, l'adaptation linguistique se traduit par un soutien social pour comprendre les processus sociaux des communautés, OMSS répondent aux besoins et aux demandes psychologiques critiques des personnes vulnérables.

Selon l'étude [119] Les médias sociaux sont devenus extrêmement populaires et, ces dernières années, les troubles mentaux chez les jeunes adultes sont devenus plus fréquents. Cela ne veut pas dire qu'ils sont liés, mais les chiffres sont ahurissants et méritent qu'on s'y attarde. En 2016, on estime que 44,7 millions d'adultes âgés de 18 ans et plus aux États-Unis avaient une maladie mentale. Les jeunes adultes âgés de 18 à 25 ans affichaient la prévalence la plus élevée de toutes les maladies mentales, soit 22,1 %, comparativement aux adultes âgés de 26 à 49 ans (21,1 %) et de 50 ans et plus (14,5 %) [119].

Une autre étude [120] portait sur l'utilisation des médias sociaux et l'isolement social chez les jeunes adultes américains. L'étude a utilisé un échantillon représentatif à l'échelle nationale de 1 787 jeunes de 19 à 32 ans. Il a évalué l'utilisation par les participants de 11 plateformes de médias sociaux : Facebook, Twitter, Google+, Youtube, LinkedIn, Instagram, Pinterest, Tumblr, Vine, Snapchat et Reddit. L'étude a révélé que ceux qui ont visité n'importe quelle plateforme au moins 58 fois par semaine étaient trois fois plus susceptibles de se sentir socialement isolés que ceux qui ont utilisé les médias sociaux moins de 9 fois par semaine. Les chercheurs ont conclu que les jeunes adultes se considéraient comme socialement isolés de leurs pairs, que ce soit vrai ou non. Ce n'est pas parce qu'ils croient qu'ils manquent d'amis qu'ils le font. En plus des sentiments d'isolement social et de dépression, les médias sociaux ont également été associés à l'image de soi. Une étude a révélé qu'une plus grande utilisation d'Instagram était associée à une plus grande auto-objectation et préoccupation au sujet de l'image corporelle.

3.3 Détection de la dépression sur les réseaux sociaux

Dans l'étude [121], un ensemble de données créé par Chaudronnier et al. Pour la linguistique informatique et clinique Psychologie (CLPsych) 2015 a été utilisé pour étudier la possibilité d'utiliser Twitter comme outil pour mesurer et prédire le trouble dépressif majeur. Plusieurs classificateurs statistiques ont été développés et comparés: Les arbres de décision, support linéaire Vector classificateur, Naive Bayes 1-gramme, Naive Bayes 2-gramme, la régression logistique, et la crête de classificateur. En ce qui concerne la précision, l'approche basée Naive Bayes-unigramme a excellé avec une précision de 86% Accuracy (paramètre de performance), mais il est tombé à court quand il s'agit de précision, ra et F-score.

D'autre part, le modèle de régression logistique a marqué un pourcentage plus élevé en termes de précision 86%, surpassant tous les autres classificateurs. Quand il vient recall (paramètre de performance) concernant le modèle Linear Support VectorMachine(linSVM) atteint le pourcentage (83%) le plus élevé par rapport aux autres, mais a pris du retard en précision. L'étude montre que si l'on doit établir des priorités le paramètre rappel est plus important que le paramètre précision, l'étude vise une approche autoritaire par rapport à une approche sous-portante, car l'étude porte sur l'identification potentielle de la dépression plutôt que sur le traitement. Le choix est basé sur un modèle de contrôle accordant la priorité au paramètre rappel plutôt qu'à la précision. En termes simples, l'insistance et sur la capacité d'identifier la plupart des personnes déprimées à risque d'identifier quelques faux positifs., accordons la priorité à la précision par rapport à un score F1 pour la même raison un modèle qui identifie bien la dépression est plus important qu'un modèle qui devient peu fiable à travers une myriade de faux positifs., Enfin, en prenant tout cela en considération, et en ajoutant le fait que les ressources de calcul et le temps sont à considérer, une approche multinomiale vers un classificateur Naïf Bayes a été choisie pour explorer de nouvelles recherches en raison des résultats qu'elle a montrés dans les expériences.

L'étude [122] qui visent à prédire la maladie mentale, sans s'y limiter à la dépression, en utilisant les médias sociaux. Les participants sont soit recrutés pour répondre à un sondage sur la dépression et partager leurs données sur Facebook ou Twitter (section A), soit les données sont recueillies auprès de sources publiques existantes en ligne (sections B, C et D). Ces sources comprennent la recherche de mots-clés dans les gazouillis publics pour identifier (et obtenir tous les gazouillis) les utilisateurs qui ont partagé leur diagnostic de santé mentale (section B), le langage de l'utilisateur sur les forums liés à la maladie mentale (section C), ou en recueillant des gazouillis publics qui mentionnent les mots clés de la maladie mentale pour l'annotation (section D).

L'analyse automatisée des médias sociaux est réalisée en établissant des modèles prédictifs qui utilisent des « caractéristiques » ou des variables qui ont été extraites des données des médias sociaux. Par exemple, les fonctions couramment utilisées comprennent le langage de l'utilisateur codé comme la fréquence de chaque mot, l'heure des messages et d'autres variables. Les caractéristiques sont ensuite traitées comme des variables indépendantes dans un algorithme (.Régression linéaire avec sélection de variable intégrée ou Support Vector Machines (SVM)) pour prédire la variable dépendante d'un résultat d'intérêt (p. ex., la santé mentale des utilisateurs).

Les modèles prédictifs sont formés, à l'aide d'un algorithme, sur une partie des données (l'ensemble de formation), puis sont évalués sur l'autre partie (l'ensemble de test) pour éviter le débordement — un processus appelé validation croisée. Les performances de prédiction sont ensuite présentées comme l'une des nombreuses mesures possibles

L'étude met l'accent sur le potentiel d'utilisation des médias sociaux pour la détection de la maladie mentale et comment il peut combler les lacunes existantes dans d'autres sources de données utilisées dans les expériences, cependant, il met en évidence le principal obstacle à ces plates-formes, ce qui est la difficulté de détecter la maladie chez les personnes qui sont pas au courant de leur état de santé mentale. Il aborde ainsi l'un des points les plus importants qui apparaît dans de telles études, ce qui est de savoir comment est-il éthique et juridique d'utiliser les données publiques à de telles fins.

Dans l'étude [123], les données ont été recueillies à partir des messages Twitter (Tweets) à l'aide du #Bellletsthalhashtagun (Bell une campagne créée par Bell Canada pour stigmatiser et promouvoir la sensibilisation et la compréhension des problèmes de santé mentale) une source où se trouve des tweets pertinents lié à la santé mentale. Le meilleur classificateur niveau utilisateur obtenu est 0,8750 avec une précision de 0,7778. Ce classificateur a été formé à l'aide de l'algorithme SVM. le modèle utilise les fonction du texte extraites comprenant le nombre de mots positifs et négatifs ,le nombre de pronoms de la première et deuxième personne ,et le sac de mots ces caractérisés par la polarité mot donne quatre caractéristiques ,le mot déprimant donne une caractéristiques et le module pronom retourne deux caractéristiques ,le nombre de fonctions retournées par le module Sac de mots (BOW) ne sont pas fixes. Utilisant le nombre de tweets déprimés, le nombre total de tweets et le pourcentage de déprimés. Ils 'appuie sur une règle pour supprimer les utilisateurs qui auto-divulguer à la dépression, mais ont moins plus de 10% de tweets de détresse afin d'améliorer les performances.

Les deux classificateurs ont été testés par différentes expériences et les résultats ont montré que les modèles de niveau utilisateur font beaucoup mieux encore avec un petit nombre de caractéristiques.

L'étude [124] de détection de la dépression dans les médias sociaux : Une solution d'apprentissage par dictionnaire multimodal. La base de données est conçue des données de dépression et de non dépressions bien étiquetées sur Twitter en ligne par la méthode heuristique. Pour représenter chaque utilisateur dans l'ensemble de données, l'extraction est basée sur des caractéristiques axées sur la dépression, inspirées non seulement par les symptômes hors ligne des critères de dépression, mais aussi par les comportements en ligne sur les médias sociaux. Ces caractéristiques sont composées de six groupes, à savoir les caractéristiques du réseau social, les caractéristiques du profil de l'utilisateur, les caractéristiques visuelles, les caractéristiques émotionnelles, les caractéristiques au niveau du sujet et le domaine spécifique.

Un modèle d'apprentissage de dictionnaire dépressif multimodal (MDL) a été conçu pour apprendre les représentations d'utilisateur clairement .plusieurs expériences ont été effectuées sur les ensembles de données étiquetés pour valider le modèle (+3% à +10% de mieux que les références en termes de F1-Measure) et analyser les contributions des fonctionnalités. Enfin, une analyse a été menée sur les contributions de fonctionnalité et les comportements en ligne de la dépression afin de révéler ceux qui ne sont pas couverts dans les critères de dépression et qui sont importants pour de telles recherches.

Dans l'étude [125], la base de données a été construite par l'intermédiaire d'échantillons dérivés automatiquement à partir d'une grande quantité de données Twitter. L'étude examine quatre troubles de santé mentale en particulier: la dépression, le trouble bipolaire, le trouble de stress post-traumatique et le trouble affectif saisonnier. Un classificateur statistique a été utilisé tout en construisant l'ensemble de données pour différencier entre les utilisateurs avec ces troubles. Un LIWC (Linguistic Inquiry Word Count).LIWC fournit aux cliniciens un outil de collecte de données quantitatives concernant l'état d'un patient par écrit (Pennebaker et al) L'analyse de chaque trouble a été menée pour mesurer les écarts de chaque maladie à partir d'un groupe de contrôle. Différentes expériences ont été réalisées au cours du processus, les résultats indiquent qu'il existe divers ensembles de signaux quantifiables pertinents pour la santé mentale observable sur Twitter.

La méthode [126] utilise l'apprentissage automatique de la machine pour classer automatiquement les patients anxieux, sentiment intérieur et émotions en utilisant des classificateurs en fonction des modèles syntaxiques n-grammes, le sentiment caractéristiques du lexique et mot distribués incorporations. L'ensemble de données utilisé par des psychologues experts spécialement destiné à l'analyse des sentiments pour la santé mentale. Dans ce travail, la classe traditionnelle de polarité « neutre » a été divisée en deux un sentiment double polarité (à la fois positive et négatif) et un sentiment ni positif ni négatif. Une polarité à quatre classes et binaire

Les expériences de prédiction et les résultats ont montré que ces dernières ont donné de meilleurs résultats, précision de 90% alors que la première a échoué avec seulement 78%.

Dans l'étude [127] l'ensemble de données utilisé a été généré par la collecte des tweets de différentes pages. Différents modèles d'apprentissage profond ont été expérimentés dans cette étude (CNN, RNN et GRU). D'autres critères ont été pris en considération pour comparer les performances, il a été examiné le modèle basé sur le caractère contre le modèle basé sur le mot et incorporations contre incorporations pré-entraîné. Les résultats ont montré que sur la base de mots GRU a surclassé tous les autres modèles avec une précision de 98%. Une CNN basée sur les mots a été considérée comme l'un des modèles les plus efficaces, car il a donné lieu à une précision de 97%.

L'étude [128] utilise l'apprentissage en profondeur pour étudier l'état psychologique. A CNN, un LSTM et un GRU ont été utilisés pour les expériences et les résultats ont montré que, même si CNN a surperformé les deux autres modèles en termes de précision, sa sortie est déraisonnable comparant aux sentiments de l'homme, ainsi, il a été conclu que la précision de modèle ne peut pas refléter l'état psychologique d'une personne. GRU d'autre part, a montré plus de résultats raisonnables.

L'étude [129] utilise la base de données CLEF Erisk 2018. L'exploration a été faite sur Earl risque de détection, lutter contre la dépression et Earl risque de détection de l'anxiété. Le défi est de détecter des traces de ces maladies Earl dans les textes publiés par les utilisateurs dans les médias sociaux. Deux modèles distincts, ont été utilisés sac distribué des mots et le modèle de mémoire distribuée, sur un ensemble de données de formation de l'anorexie Erisk 2018. Les sorties de ces deux modèles sont concaténées, Les résultats obtenues avec 0.85 F1-score, 0.91 Précision et 0.88 pour rappel.

Ce travail [130] présente une approche de détection précoce du risque de dépression et l'anxiété sur les réseaux sociaux dans CLEF eRisk 2018. Pour les deux maladies mentales, les modèles combinent l'information TF-IDF (frequency-inverse document frequency) TF-IDF est calculé pour chaque mot, et les mots avec un faible score TF-IDF seront supprimés dans la phrase. Enfin, le classificateur de phrase est formé avec les phrases raffinées. Le reste de la séquence sera suivi en utilisant l'encodeur de phrases basé sur réseaux de neurone convulsifs (CNN) pour identifier les articles écrits par des patients potentiels. L'évaluation officielle montre que le modèle atteint un F-score de 0,67 dans la détection de l'anxiété

3.4 Discussion

Nous avons constaté au cours de notre revue littérature que de nombreux travaux ont abordé le thème de l'analyse des sentiments tout en se basant sur les modèles statistiques traditionnels comme les classificateurs. Les différentes études cités dans le chapitre deux, étaient consacré pour l'analyse des sentiments dans les domaines beaucoup plus industriel, commercial, médical et politique mais en terme d'analyse de sentiment concernant Les maladies mentaux et les troubles de comportement comme l'anxiété et la dépression c.à.d. a des fin de santé mental sont récentes .

A travers les travaux partagés, nous avons remarqué que la source la plus riche de données en termes de quantité, diversité, et en contenu est Twitter. Il est prouvé que ces données peuvent être utilisées pour étudier les questions cliniques, précisément ce qui concerne les maladies mentaux, de plus, toutes les études ont montré l'importance et l'influence des données utilisées dans les résultats finaux. L'utilisation des mêmes techniques et méthodes sur différents ensembles de données produit toujours des résultats différents, ce qui prouve que l'ensemble de données est l'un des facteurs, les plus important, dans la réussite ou l'échec d'une étude.

Nous citons dans le tableau ci-dessous les études qu'ont contribuées à notre approche et conclusions,

Tableau 3-1 : Tableau comparatif des travaux de détection de dépression.

Etude	Année	But	Algorithme de classification	Dataset	Performances
[121]	2016	Identifier la dépression sur Twitter	Décision Trees	(CLPsych) 2015(dataset created by Coppersmith et al. for the Computational Linguistics and ClinicalPsycho	Precision :67% Recall :68% F-score :75% Accurancy :67%
			Linear Support Vector Classifier	logy) (CLPsych) 2015	Precision :83% Recall :83% F-score :83%

Détection de la dépression dans les réseaux sociaux :

État de l'art

					Accuracy :82%
			Naïve Bayes 1-gram		Precision :81% Recall :82% F-score :81% Accuracy :86%
			Naïve Bayes2-gram		Precision :82% Recall :82% F-score :82% Accuracy :82%
			Logistic Regression		Precision :86% Recall :82% F-score :84% Accuracy :82%
			Ridge Classifier		Precision :81% Recall :79% F-score :78% Accuracy :79%
[122]	2015	Detecting depression Andmental innless	Regresse-linear+SVM	Tweeter dataset	Accuracy :69%
			svm-Orignal	CLPsych2015	Precision : 50% Recall :50% F-score :50% Accuracy :80%
[123]	2017	Detecting at risks-user	Svm-Smote	BellLetsTalk data,	Precision : 66% Recall :25% F-score :35% Accuracy :65%
			Svm-Linear	BellLetsTalk	Precision : 58%

Détection de la dépression dans les réseaux sociaux :

État de l'art

				data,	Recall :77% F-score :66% Accurancy :60%
[124]	2017	Depression detection	MDL(multi model depressive dictionary)	Depression DatasetD1.Based on the tweets (2009-2016) DatasetD2 Based On tweet 2016	Precision : 85% Recall :85% F-score :85% Accurancy :85%
[126]	[2019]	Depression detection	SVM,RNN et CNN	Crowddist dataset MH-C dataset	Precision : 44% Recall :44% F-score :68% Precision : 44% Recall :44% F-score :45%
[127]	2018	Depression detection	RNN GRU	Twitter depression dataset	Precision : 99% Recall :97% F-score :98% Accurancy :98%
			GRU with char CNN	Twitter depression dataset	Precision : 99% Recall :98% F-score :99% Accurancy:98%
[128]	[2017]	Comparaison analyse sentiment et les résultats psychologiques	CNN, LSTM GRU (gated recurrent units)	Kummar et al 2015	Accurancy : 91% Accurancy : 90% Accurancy : 89%

[130]	[2018]	Depression detection	CNN	e-Risk 2018	Precision : 64% Recall :95% F-score :64%
-------	--------	-------------------------	-----	-------------	--

Suite à notre approfondissement dans la documentation, ainsi que la comparaison entre les différents travaux réalisés dans le même thème lié à notre étude. Notre modèle proposé consiste à effectuer

La classification des utilisateurs sur le plan de leur état mental (dépressif ou non dépressif). Nous avons utilisé la base de donnée e-Risk 2019 comme source de données, notre modèle consiste à extraire les données de cette dataset et les traitées par la suite.

Le modèle sera conçu sur la base des réseaux de neurones, le choix de la conception sera basé sur les résultats déjà obtenues dans d'autres travaux similaires l'approche consiste à concevoir un modèle de plusieurs couches qui traite les tweets des utilisateurs pour pouvoir les classifier à la fin. Nous utiliserons les réseaux de neurone convulatif (CNN) et d'autres méthodes comme LSTM (long short-term Memory) et naïve bayes ces derniers ont prouvé leur performance dans plusieurs travaux, la spécification de notre approche et de pouvoir utiliser plusieurs tests sur plusieurs réseaux et obtenir plusieurs résultats sur le même modèle. À la fin nous utiliserons une fonction soft max qui permettra la prise de décision.

Ce sont les points les plus importants pris en considération après le parcours de la documentation, nous partageons dans le chapitre suivant les différents détails concernant le modèle adopté.

3.5 Conclusion

Dans ce chapitre, nous avons présenté les œuvres les plus pertinentes liées à notre étude, ainsi que les plus influencées concernant la conception de notre modèle. Sur la base de l'analyse et de comparaison des résultats déjà obtenues nous avons inspiré notre conception afin de proposer une approche qui donnera de meilleurs résultats.

Modèle de détection de dépression

Chapitre 4 Modèle de détection de dépression

4.1 Introduction

Après avoir parcouru la littérature et avoir vu quelques travaux réalisés dans ce domaine, nous présentons dans ce chapitre l'architecture de notre modèle. Ce dernier est structuré de plusieurs parties, la première traite les données stockées dans les fichiers « utilisateurs », la deuxième traite les données de formation « train data » et de test « test data », la troisième traite l'apprentissage des modèles (CNN, Naive Bayes, LSTM, BiLSTM) et la dernière traite le test et l'évaluation des modèles

4.2 Architecture du Modèle

Le modèle proposé consiste à effectuer la classification des utilisateurs sur le plan de leur état mental (dépressif ou non dépressif) en se basant sur les données de Dataset e-Risk2019. Les principales étapes du modèle sont les suivantes:

- Étape 1: Extraction des données (train data, Test data & sentiments Users).
- Étape 2: Prétraitement et vectorisation (Preprocessing and vectorization).
- Étape 3: Apprentissage des modèles.
- Étape 4 : Test et évaluation.

Différents travaux ont utilisé les méthodes CNN, LSTM, Naive bayes ou BiLSTM séparément pour des études de recherche similaires, mais selon nos recherches, aucun n'a utilisé ces méthodes d'apprentissage en profondeur simultanément dans le même but que la nôtre. La combinaison de ces méthodes constitue la contribution majeure de notre modèle. Dans la figure 4.1, nous schématisons les étapes constituant notre modèle de détection de dépression.

Modèle de détection de dépression

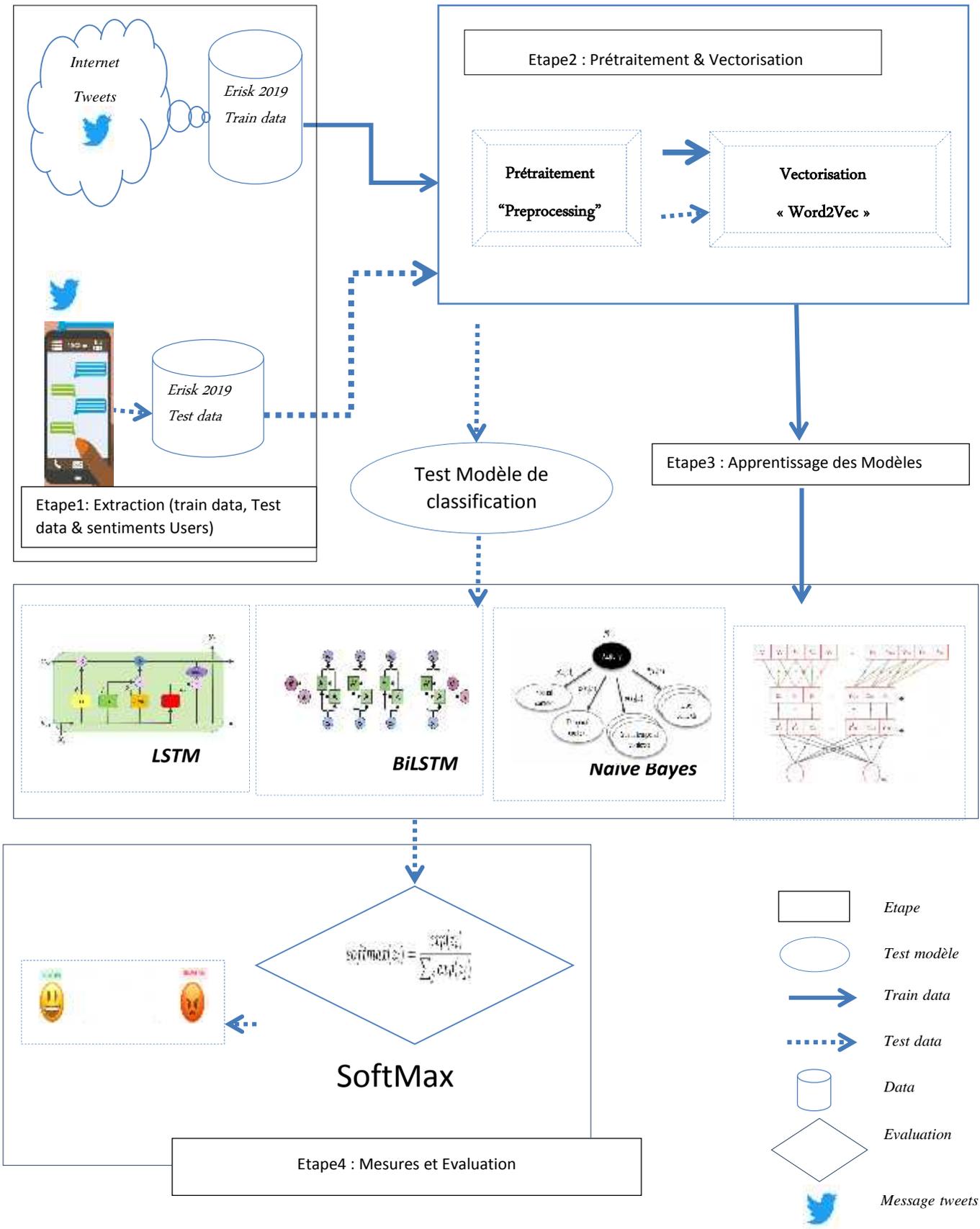


Figure 4.1 Architecture du modèle de détection de dépression

4.2.1 Extraction des données (train data, Test data & sentiments Users)

Cette section donne un aperçu de l'ensemble de données utilisées dans notre étude, ces dernières sont stockées dans des fichiers d'extension « xml » (Figure 4.2) regroupées par utilisateur.

Suivant le dataset e-Risk 2019, nous avons deux ensembles de données. L'ensemble d'apprentissage (Train data) contient 486 utilisateurs dont 83 sont déprimés et 403 sont non déprimés. L'ensemble de test (Test data) contient 401 utilisateurs dont 52 déprimés et 349 non déprimés).

Le traitement des fichiers commencera par les fichiers de la collection de Dataset Erisk 2019 « risk_golden_truth.txt » et « T1_erisk_golden.txt ». Le but de cette étape est la création d'un fichier d'extension CSV qui regroupe tous les textes de tous les utilisateurs du dossier chunk(i) pour le dossier train et un fichier d'extension « CSV » par utilisateur des dossiers chunk test.

```
<?xml version="1.0"?>
- <INDIVIDUAL>
  <ID>train_subject416</ID>
  - <WRITING>
    <TITLE></TITLE>
    <DATE>2012-10-07 21:38:08</DATE>
    <INFO>reddit post</INFO>
    <TEXT>Something I do which has really helped
      was only small things, that I took a shower
      sleep I looked at the list and tried to think a
    </TEXT>
  </WRITING>
</INDIVIDUAL>
```

Figure 4.2 : Structure interne des fichiers Train et Test dans la collection Erisk 2019

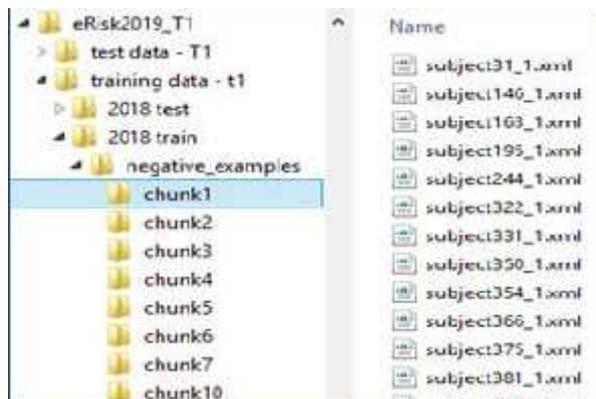


Figure 4.3 : les fichiers Utilisateur dans la collection Erisk2019

4.2.2 Prétraitement & Vectorisation

4.2.2.1 Prétraitement "Preprocessing "

Dans le traitement automatique de (TAL), la plupart des textes et documents contiennent de nombreux mots redondants pour la classification des textes, tels que mots vides, orthographes, etc. Dans cette section, nous expliquons brièvement certaines techniques et méthodes de conversion, nettoyage et traitement des documents texte.

- Supprimer les tweets non anglais, tweeter permet plus de 60 langues. Cependant, ce travail ne concerne actuellement que les tweets anglais. Nous supprimons les tweets de nature non anglaise.
- Remplacement des émoticônes. Les émoticônes jouent un rôle important dans la détermination du sentiment du tweet. Par conséquent, nous remplaçons les émoticônes par leur polarité de sentiment en recherchant dans le dictionnaire des émoticônes.
- Supprimer l'URL. Les URL présentés dans le tweet sont raccourcies à l'aide de TinyUrl en raison de la limitation du texte du tweet. Ces URL raccourcies ne contenaient pas beaucoup d'informations sur le sentiment du tweet. Ainsi, ceux-ci sont supprimés.
- Remplacer les mentions négatives avant de supprimer les mots vides. Les tweets consistent en diverses notions de négation. En général, les mots finissant par « n't » sont ajoutés avec un non comme « Won't » remplacer par « will not », « can't » remplacer par « can not » .
- Hashtags. Les hashtags sont essentiellement des résumés du tweet et sont donc très critiques. Ces hashtags ne contenaient pas beaucoup d'informations sur le sentiment du tweet. Ainsi, ceux-ci sont supprimés.
- Séquence de caractères répétés tweets fournit aux utilisateurs une plate-forme pour exprimer leur opinion de manière informelle tels que « woowooow »,.. Les tweets sont écrits de manière aléatoire, sans aucun souci de corriger la structure et l'orthographe. La correction orthographique joue un rôle important dans l'analyse des sentiments du contenu généré par l'utilisateur. Les gens utilisent des mots tels que "coooool" et "hunnnnngry". Afin de capturer de telles expressions, nous remplaçons la séquence de plus de trois caractères similaires par trois caractères. Par exemple, woowooow est remplacé par woow. Nous remplaçons par trois caractères afin de distinguer des mots tels que "cool" et "coooool".

- Nous supprimons les nombres et les chiffres parce qu'ils ne servent à rien pour mesurer le sentiment.
- lemmatisation : La lemmatisation est l'étape qui fait référence à l'analyse lexicale qui permet de regrouper des mots d'une même famille qui partagent le même suffixe lexical.
- La négation : Dans l'analyse du sentiment, le cas de la négation reste un problème non résolu. C'est un cas important dans ce type de tâches qui peut changer complètement le sentiment d'un tweet. Nous citons l'exemple suivant "Dont" et "Don't" sera "Do Not", "Couldn't" sera "Could Not".
- Suppression des mots vides : Les mots vides jouent un rôle négatif dans la tâche de classification des sentiments. Les mots vides se produisent à la fois dans les ensembles d'entraînement positifs et négatifs, ajoutant ainsi plus d'ambiguïté dans la formation du modèle. Exemple {'ourselves', 'hers', 'between', 'yourself', 'but', 'again', 'there', 'about', 'once', 'during', 'out', 'very', 'having', 'with', 'they', 'own', 'an', 'be', 'some', 'for', 'do', 'its', 'yours', 'such', 'into', 'of', 'most', 'itself', 'other', 'off', 'is', 's', 'am', 'or', 'who', 'as', 'from', 'him', 'each', 'the', 'themselves', 'until', 'below', 'are', 'we', 'these', 'your', 'his', 'through', 'don', 'nor', 'me', 'were', 'her', 'more', 'himself', 'this', 'down', 'should', 'our', 'their', 'while', 'above', 'both', 'up', 'to', 'ours', 'had', 'she', 'all', 'no', 'when', 'at', 'any', 'before', 'them', 'same', 'and', 'been', 'have', 'in', 'will', 'on', 'does', 'yourselves', 'then', 'that', 'because', 'what', 'over', 'why', 'so', 'can', 'did', 'not', 'now', 'under', 'he', 'you', 'herself', 'has', 'just', 'where', 'too', 'only', 'myself', 'which', 'those', 'i', 'after', 'few', 'whom', 't', 'being', 'if', 'theirs', 'my', 'against', 'a', 'by', 'doing', 'it', 'how', 'further', 'was', 'here', 'than' }

4.2.2.2 Vectorisation

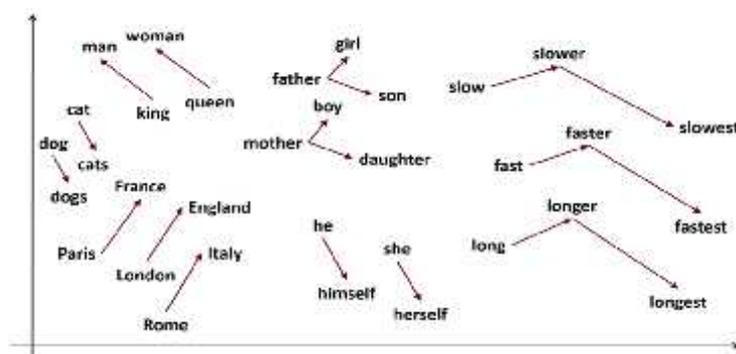
4.2.2.2.1 Tokénisation : La tokénisation est le processus consistant à décomposer un texte en mots, appelés jetons. La tokénisation repose principalement sur des méthodes heuristiques simples afin de séparer les jetons en suivant quelques étapes:

- Les jetons ou les mots sont séparés par des espaces, des signes de ponctuation ou des sauts de ligne
- Un espace ou des signes de ponctuation peuvent être inclus ou non en fonction des besoins
- Tous les caractères contenus dans des chaînes contiguës font partie du jeton. Les jetons peuvent être constitués de tous les caractères alpha, alphanumériques ou numériques uniquement

4.2.2.2.2 Embeddings

L'incorporation de mots est une technique utilisée pour représenter des documents avec une représentation vectorielle dense (Figure 4.4). Le vocabulaire de ces documents est mappé sur des vecteurs de nombres réels [131]. Les mots sémantiquement similaires sont mappés les uns aux autres dans l'espace vectoriel. Il existe des modèles d'intégration de mots prêts à être utilisés, tel que la bibliothèque Gensim.

Les représentations distribuées des mots dans un espace vectoriel aident les algorithmes d'apprentissage à obtenir de meilleures performances dans les tâches de traitement du langage naturel en regroupant des mots similaires.



4.2.2.2.3 Glove est un algorithme d'apprentissage non supervisé permettant d'obtenir des représentations vectorielles de mots.

Exp: phrase = "My dreams Are Fire

	My	Dreams	Are	Fire		Mot(v)
My	1	0	0	0	0	0	0
Dream	0	1	0	0	0	0	0
Are	0	0	1	0	0	0	0
Fire	0	0	0	1	0	0	0

Figure 4.5 : Représentation Vectorielle d'une phrase

CBOw (Figure 4.5) tente de prédire le mot cible actuel (le mot central) en fonction des mots de contexte source (mots environnants).

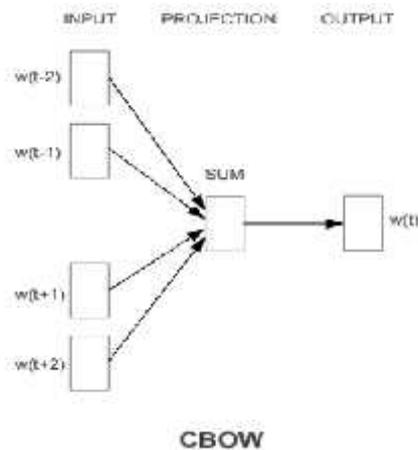


Figure 4.6 : Model CBOW [131]

Skip-gram tente généralement d'obtenir l'inverse de ce que fait le modèle CBOW. Il essaie de prédire les mots du contexte source (mots environnants) à partir d'un mot cible (le mot central).

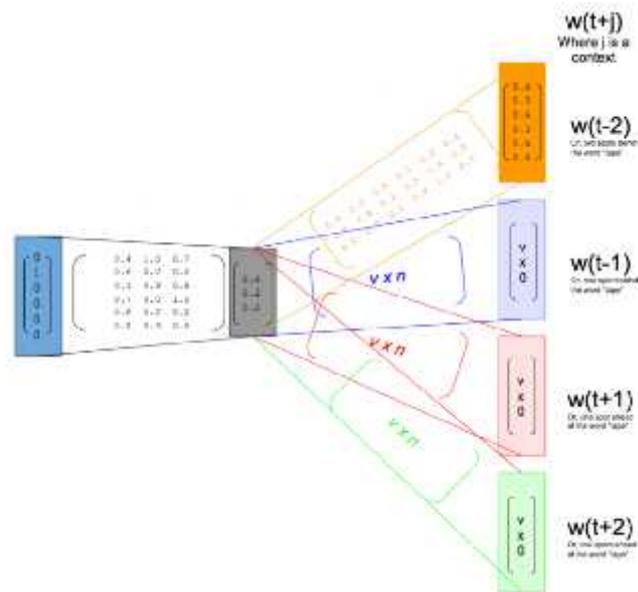


Figure 4.7 : Model skip-gram[131]

4.2.3 Apprentissage des Modèles (Model Training)

Notre démarche d'analyse de sentiments s'inscrit dans l'approche d'apprentissage automatique supervisé. Nous avons utilisé les modèles suivants « CNN », « Naive bayes », « LSTM » et « BiLSTM » qui seront utilisés dans l'étape « Test et évaluation ».

4.2.3.1 Les Réseaux de Neurones Convolutifs(CNN)

Aujourd'hui, les réseaux de neurones Convolutif, aussi appelés CNN ou ConvNet pour Convolutional Neural Network [132][133], sont toujours les modèles les plus performants pour la classification d'images. On distingue deux parties, une première partie que l'on appelle la partie convolutive du modèle et la seconde partie, que l'on appelle la partie classification du modèle qui correspond à un modèle MLP (Multi Layers Perceptron)(Figure4.8).

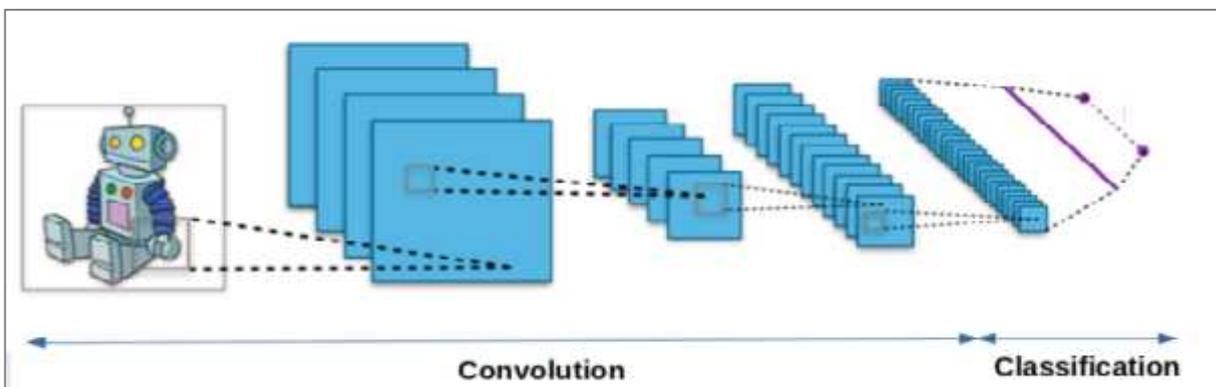


Figure4.8 : Les deux(02) parties de Convolutional Neural Network [134]

CNN est un réseau de neurones multicouches et plus précisément c'est un réseau profond composé de multiples couches qui en générale sont organisées en blocs (CONV RELU POOL). Il existe quatre types de couches pour un réseau de neurone Convolutif (Figure 4.9):

- La couche de convolution.
- La couche de pooling.
- La couche de correction ReLU .
- La couche fully-connected.

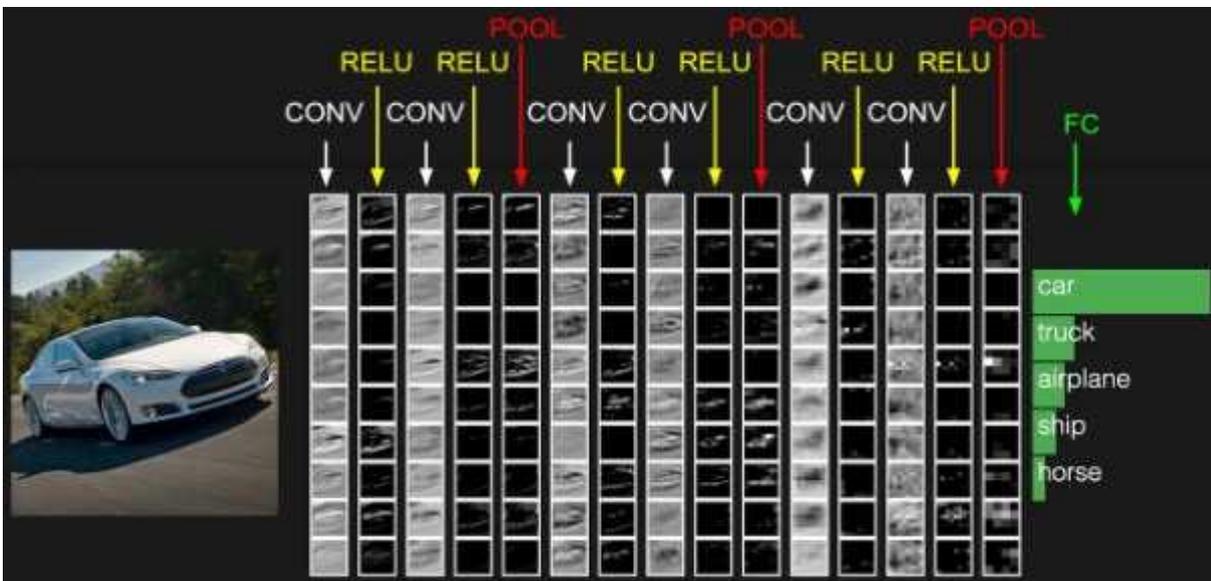


Figure 4.9 : Les quatre types de couches pour un réseau de neurones Convolutif [134]

La couche de convolution : La couche de convolution est la composante clé des réseaux de neurones convolutifs, et constitue toujours au moins leur première couche. Son but est de repérer la présence d'un ensemble de composants dans les images reçues en entrée (Figure 4.10). Pour cela, on réalise un filtrage par convolution, le principe est de faire "glisser" une fenêtre représentant le composant sur l'image, et de calculer le produit de convolution entre le composant et chaque portion de l'image balayée.

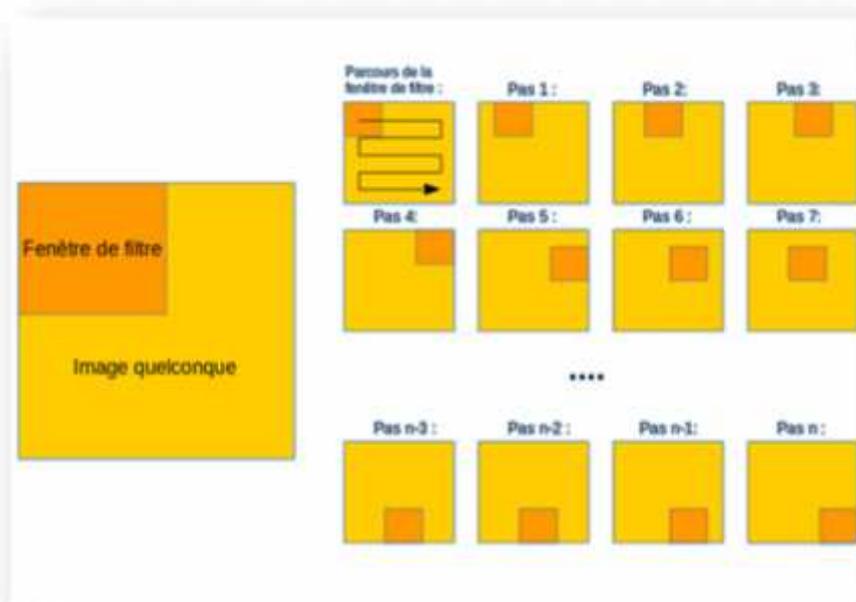


Figure 4.10 : Schéma du parcours de la fenêtre de filtre sur l'image [134]

Couche de mise en commun: pooling : La couche de pooling permet de réduire le nombre de paramètres et de calculs dans le réseau. On améliore ainsi l'efficacité du réseau et on évite le sur-apprentissage. Ce type de couche est souvent placé entre deux couches de convolution : elle reçoit en entrée plusieurs composants maps, et applique à chacune d'entre elles l'opération de pooling.

La couche ReLU : ReLU (Rectified Linear Units) désigne la fonction réelle non-linéaire définie par $\text{ReLU}(x) = \max(0, x)$. Elle remplace toutes les valeurs négatives reçues en entrées par des zéros. Elle joue le rôle de fonction d'activation [134]. Souvent, la correction Relu est préférable, mais il existe d'autres formes :

- La correction par tangente hyperbolique $f(x) = \tanh(x)$,
- La correction par la tangente hyperbolique saturante: $f(x) = |\tanh(x)|$,
- La correction par la fonction sigmoïde : $f(x) = (1 + e^{-x})^{-1}$

Couche fully-connected(FC) : La couche fully-connected constitue toujours la dernière couche d'un réseau de neurones, convolutif. La couche fully-connected permet de classifier le texte en entrée du réseau : Elle reçoit un vecteur en entrée et produit un nouveau vecteur en sortie. Pour cela, elle applique une combinaison linéaire puis éventuellement une fonction d'activation aux valeurs reçues en entrée.

CNN et le traitement automatique de la langue naturelle (TAL) :

Au lieu de pixels d'image, la plupart des fonctions TAL ont comme entrées des phrases ou des documents. L'idée pour appliquer un CNN est de transformer les mots et les documents sous une forme matricielle. Chaque ligne de la matrice correspond à un jeton, généralement un mot, mais il peut s'agir aussi d'un caractère. C'est-à-dire que chaque ligne est un vecteur qui représente un mot. Par exemple, pour une phrase de 20 mots utilisant une incorporation de 100 dimensions, nous aurions une matrice de 20×100 comme entrée.

Un filtrage par convolution dans le traitement des images consiste à faire "glisser" une fenêtre représentant le composant sur l'image, et de calculer le produit de convolution entre le composant et chaque portion de l'image balayée. Cependant en TAL, nous utilisons généralement des filtres qui glissent sur des lignes complètes de la matrice (les mots). Ainsi, la "largeur" des filtres est généralement la même que la largeur de la matrice d'entrée. La hauteur, ou la taille de la région, peut varier, mais les fenêtres glissantes de plus de 2 à 5 mots à la fois sont typiques

Au lieu de pixels d'image, la plupart des fonctions TAL ont comme entrées des phrases ou des documents. L'idée pour appliquer un CNN est de transformer les mots et les documents sous une forme matricielle. Chaque ligne de la matrice correspond à un jeton, généralement un mot, mais il peut s'agir aussi d'un caractère. C'est-à-dire que chaque ligne est un vecteur qui représente un mot. Par exemple, pour une phrase de 20 mots utilisant une incorporation de 100 dimensions, nous aurions une matrice de 20×100 comme entrée.

Dans (Figure 4.11) l'illustration d'une architecture CNN pour la classification de phrases. Nous décrivons trois tailles de région de filtre 2, 3 et 4 chacune ayant 2 filtres. Ces derniers effectuent des convolutions sur la matrice de phrases et génèrent des composants maps. Le regroupement 1-max est effectué sur chaque map, c'est-à-dire que le plus grand nombre de chaque feature map est enregistré. Ainsi, un vecteur de caractéristiques invariable est généré à partir des six maps et ces six composants, sont concaténés pour former un vecteur de caractéristiques pour l'avant-dernière couche. La variable finale soft max reçoit ensuite ce vecteur de caractéristiques en entrée et l'utilise pour classifier la

phrase; nous supposons ici une classification binaire et décrivons donc deux états de sortie possibles.

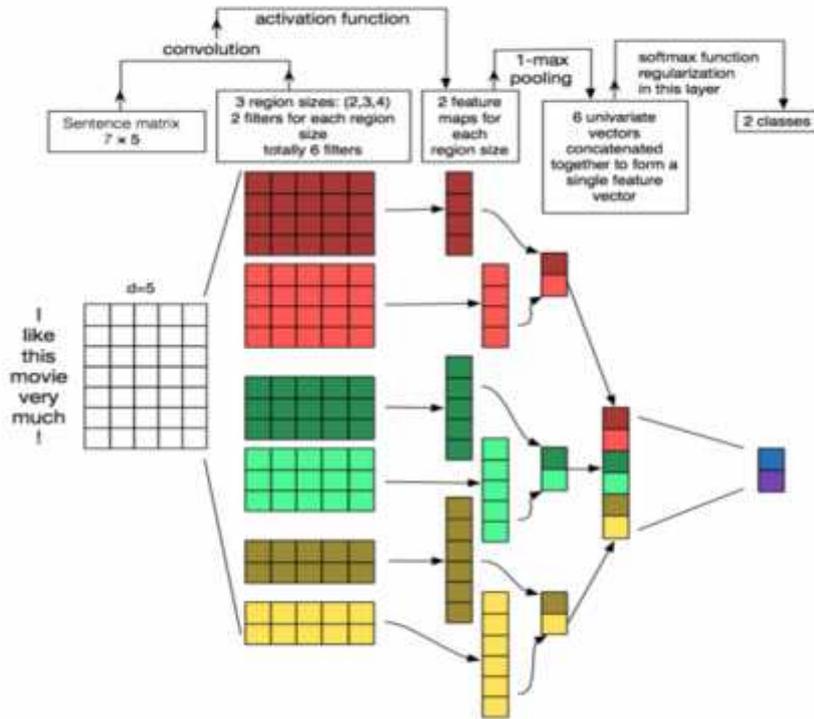


Figure 4.11 : Architecture CNN pour la classification de phrases [134].

4.2.3.2 Modèle Naive- Bayes

Naive-Bayes est un modèle de classification basée sur le théorème de Bayes. C'est l'un des algorithmes d'apprentissage supervisé les plus simples. Un modèle naïf bayésien est facile à construire, sans aucune estimation de paramètre itérative compliquée, ce qui le rend particulièrement utile pour les très grands ensembles de données.

$$P(h|D) = \frac{P(D|h)P(h)}{P(D)}$$

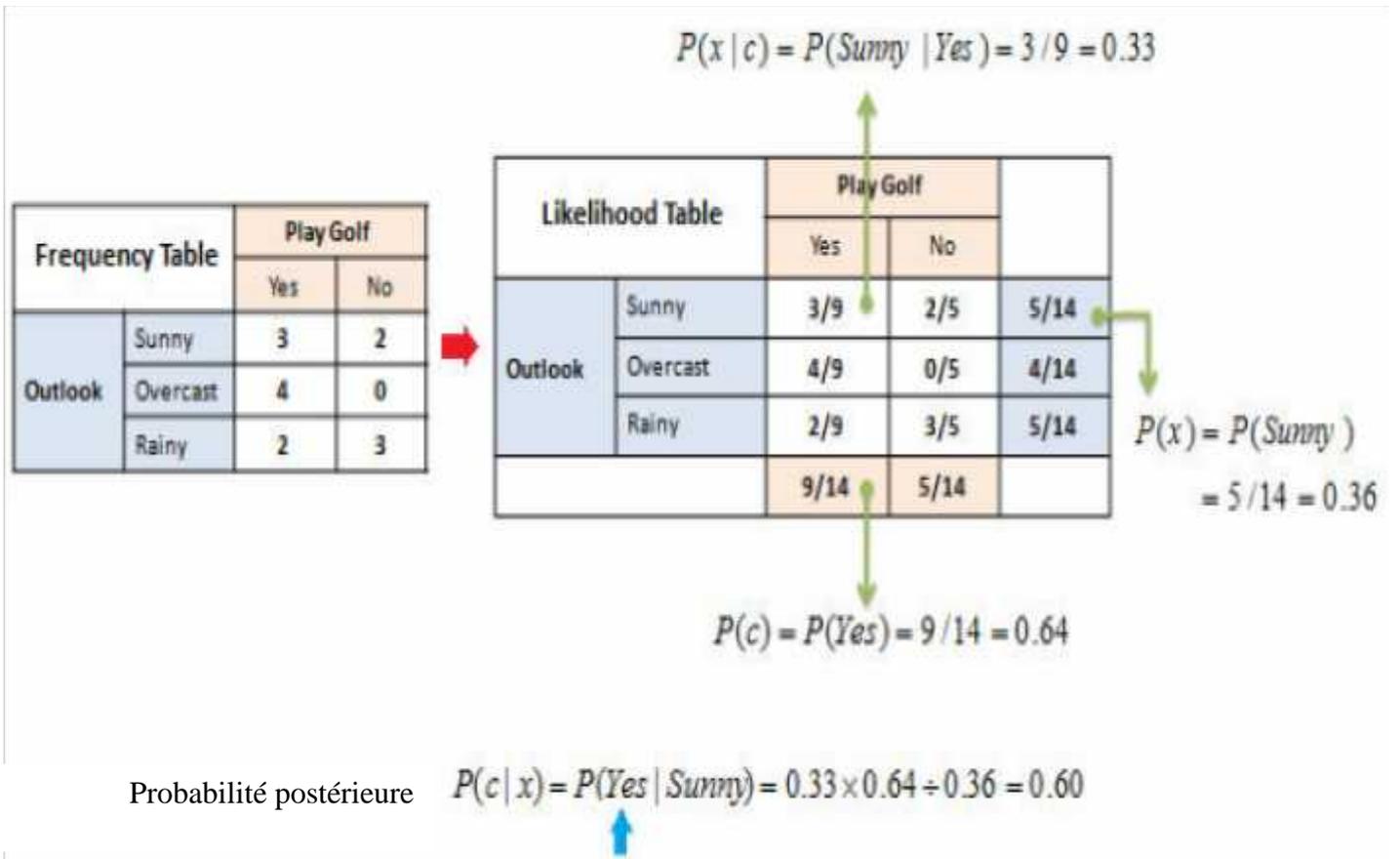
- $P(h)$: probabilité que l'hypothèse h soit vraie (quelles que soient les données). Ceci est connu comme la probabilité antérieure de h .
- $P(D)$: la probabilité des données (quelle que soit l'hypothèse). Ceci est connu comme la probabilité antérieure.
- $P(h | D)$: probabilité de l'hypothèse h à partir des données D . On parle de probabilité postérieure.
- $P(D | h)$: probabilité des données d étant donné que l'hypothèse h est vraie. Ceci est connu sous le nom de probabilité postérieure

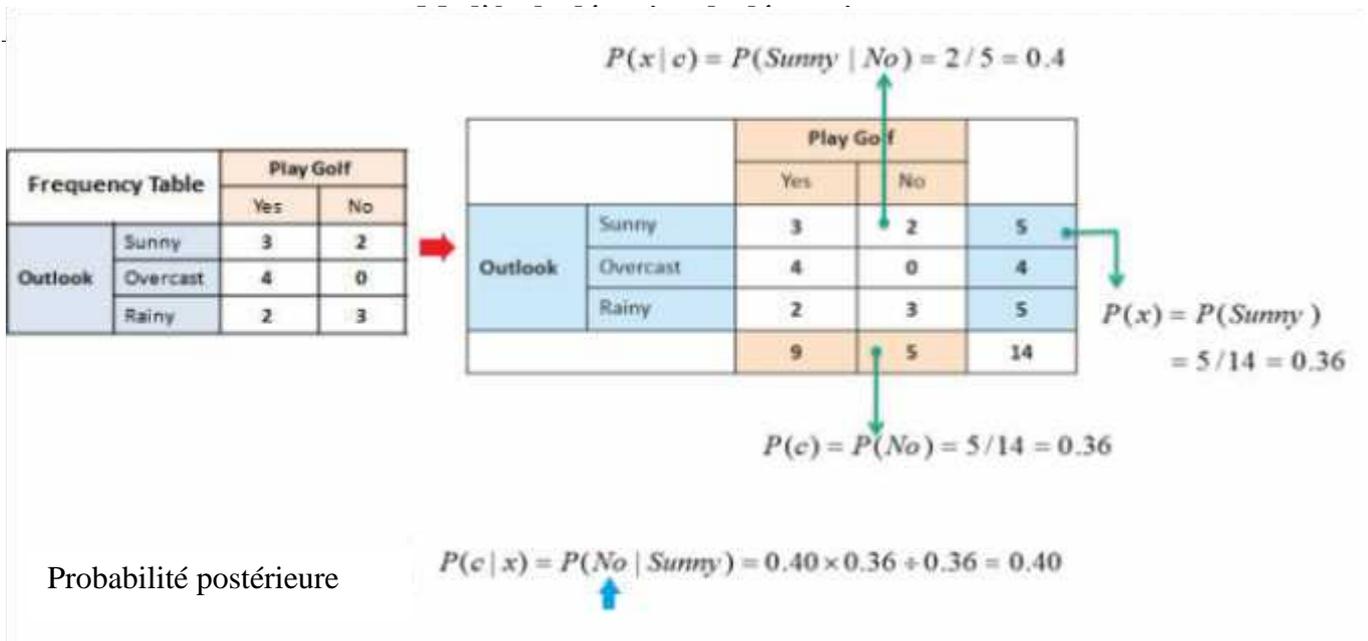
Afin de mieux comprendre son utilisation nous proposons l'exemple suivant : Nous avons un ensemble de données de formation sur la météo (Figure 4.12) et la variable cible correspondante "Play Golf". Ensuite, nous devons classer si les joueurs vont jouer ou non en fonction des conditions.

La probabilité postérieure peut être calculée en commençant par construire un tableau de fréquences pour chaque attribut par rapport à la cible. Ensuite, transformez les tables de fréquences en tables de vraisemblance et utilisez finalement l'équation naïve bayésienne pour calculer la probabilité a posteriori de chaque classe. La classe avec la probabilité postérieure la plus élevée est le résultat de la prédiction.

Outlook	Temp	Humidity	Windy	Play Golf
Rainy	Hot	High	False	No
Rainy	Hot	High	True	No
Overcast	Hot	High	False	Yes
Sunny	Mild	High	False	Yes
Sunny	Cool	Normal	False	Yes
Sunny	Cool	Normal	True	No
Overcast	Cool	Normal	True	Yes
Rainy	Mild	High	False	No
Rainy	Cool	Normal	False	Yes
Sunny	Mild	Normal	False	Yes
Rainy	Mild	Normal	True	Yes
Overcast	Mild	High	True	Yes
Overcast	Hot	Normal	False	Yes
Sunny	Mild	High	True	No

Figure 4.12 Ensemble de Données de Formation sur la Météo





4.2.3.3 Modèle LSTM

Les réseaux de mémoire LSTM constituent un type particulier de RNN, capable d'apprendre des dépendances à long terme [135]. Les LSTM sont explicitement conçus pour éviter le problème de dépendance à long terme. Tous les réseaux de neurones récurrents ont la forme d'une chaîne de modules répétitifs de réseau de neurones (Figure 4.13).

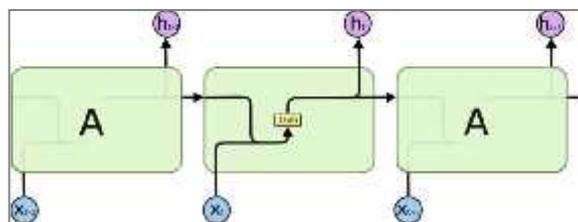


Figure 4.13. Le module de répétition dans un RNN [136]

Les LSTM ont également cette chaîne comme structure, mais le module répétitif à une structure différente. Au lieu d'avoir une seule couche de réseau neuronal, il y en a quatre, qui interagissent de manière très particulière (Figure 4.14).

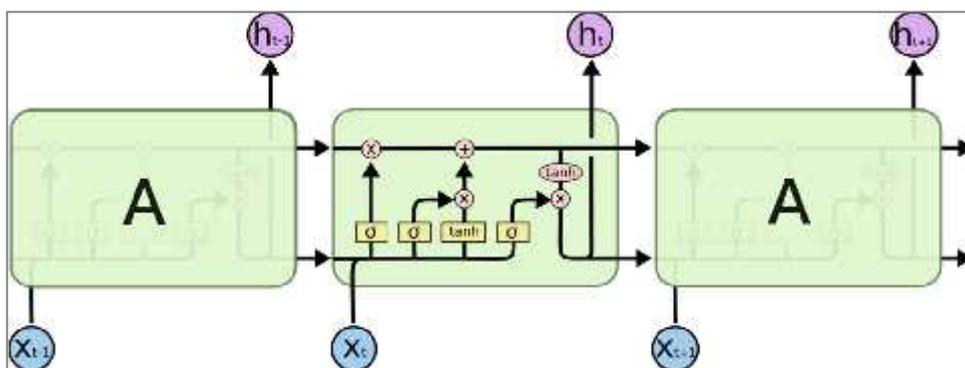


Figure 4.14 Le module de répétition dans un LSTM [136]

Afin de donner une illustration détaillée du fonctionnement du réseau LSTM, nous décrivons le diagramme LSTM, étape par étape. Pour le début, essayons simplement d'introduire les notations que nous utilisons au cœur de notre explication (Figure 4.15).

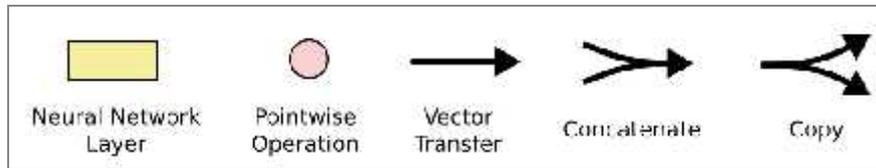


Figure 4.15 Notation LSTM-BiLSTM [136]

Dans le diagramme ci-dessus, chaque ligne porte un vecteur entier, de la sortie d'un nœud aux entrées des autres. Les cercles roses représentent des opérations ponctuelles, telles que l'addition de vecteurs, tandis que les zones jaunes sont des couches de réseaux de neurones apprises. La fusion de lignes indique une concaténation, tandis qu'une ligne de bille indique son contenu en cours de copie et les copies envoyées à des emplacements différents.

L'idée principale derrière les LSTM : La clé des LSTM est l'état de la cellule, la ligne horizontale qui traverse le haut du diagramme. L'état de la cellule est un peu comme un tapis roulant. Il court tout au long de la chaîne, avec seulement quelques interactions linéaires mineures. Il est très facile pour une information de circuler sans la modifier (Figure 4.16).

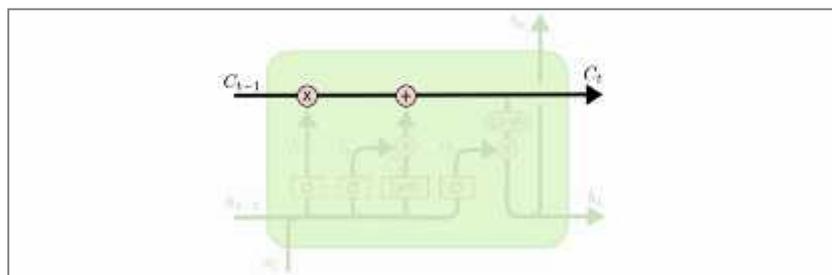


Figure 4.16 L'état cellulaire du Module LSTM [136]

Le LSTM a la capacité de supprimer ou d'ajouter des informations à l'état de la cellule, soigneusement régulées par des structures appelées portes. Ces dernières sont un moyen de laisser éventuellement passer l'information. Ils sont composés d'une couche de réseau neural sigmoïde et d'une opération de multiplication ponctuelle (Figure 4.17).

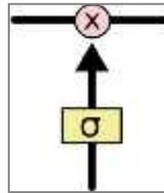


Figure 4.17 Exemple de perceptron [136]

La couche sigmoïde génère des nombres compris entre zéro et un, décrivant la quantité de chaque composant à laisser passer. Une valeur de zéro signifie "ne laisse rien passer", alors qu'une valeur de un(1) signifie "laisse tout passer!". Un LSTM a trois de ces portes, pour protéger et contrôler l'état de la cellule.

Les étapes du LSTM : La première étape de LSTM consiste à décider quelles informations nous stockons à l'état de cellule. Cette décision est prise par une couche sigmoïde appelée « oublier couche de grille. » (Figure 4.18) Il regarde h_{t-1} et x_t , et affiche un nombre compris entre 0 et 1 pour chaque numéro dans l'état de la cellule C_{t-1} . Un 1 représente « garder complètement cela » alors qu'un 0 représente "complètement se débarrasser de cela."

Revenons à notre exemple de modèle de langage essayant de prédire le mot suivant sur la base de tous les mots précédents. Dans un tel problème, l'état de la cellule pourrait inclure le sexe du sujet en question, afin que les pronoms corrects puissent être utilisés. Quand nous voyons un nouveau sujet, nous voulons oublier le genre de l'ancien sujet.

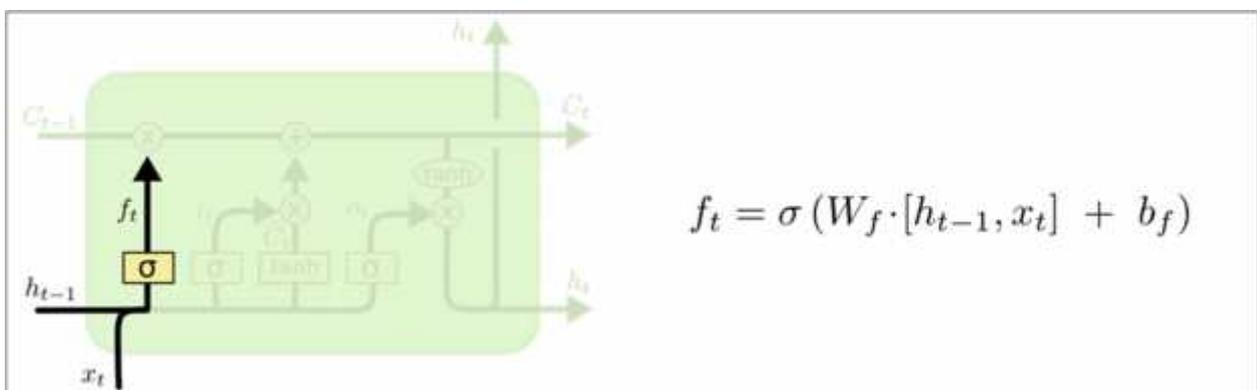


Figure 4.18 la première étape de Module LSTM [136]

L'étape suivante consiste à décider de nouvelles informations que nous stockerons dans l'état de la cellule. Cela se fait en deux parties (Figure 4.19).

1. Tout d'abord, une couche sigmoïde appelée «couche de grille d'entrée» décide des valeurs à mettre à jour.
2. Ensuite, une couche de tangente crée un vecteur de nouvelles valeurs candidates, \tilde{C}_t , cela pourrait être ajouté à l'état.

Dans la prochaine étape, nous combinons ces deux pour créer une mise à jour de l'état.

Dans l'exemple de notre modèle de langage, nous ajouter le sexe du nouveau sujet à l'état de la cellule, pour remplacer l'ancien que nous avons oublié.

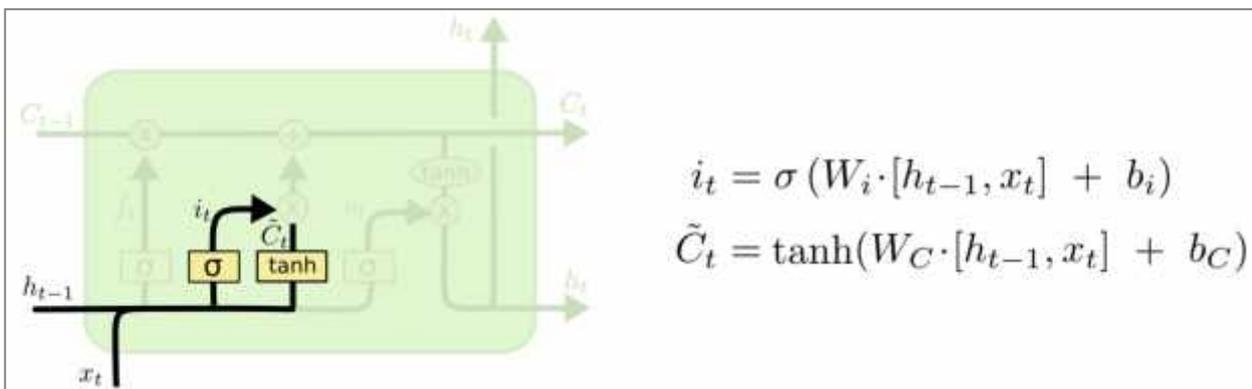


Figure 4.19 Deuxième étape du modèle LSTM[136]

Nous mettons à jour l'ancien état de la cellule, C_{t-1} , dans le nouvel état de cellule C_t .

Nous multiplions le vieil état par f_t dans l'ordre d'oublier les décisions dans la première étape. Puis on ajoute $i_t * \tilde{C}_t$. (Figure 4.20) Ce sont les nouvelles valeurs candidates, mises à l'échelle en fonction de la mesure dans laquelle nous avons décidé de mettre à jour chaque valeur d'état. Dans le cas du modèle linguistique, c'est à cet endroit que l'on supprime les informations sur le sexe de l'ancien sujet et on ajoute les nouvelles informations, déjà décidé au préalable.

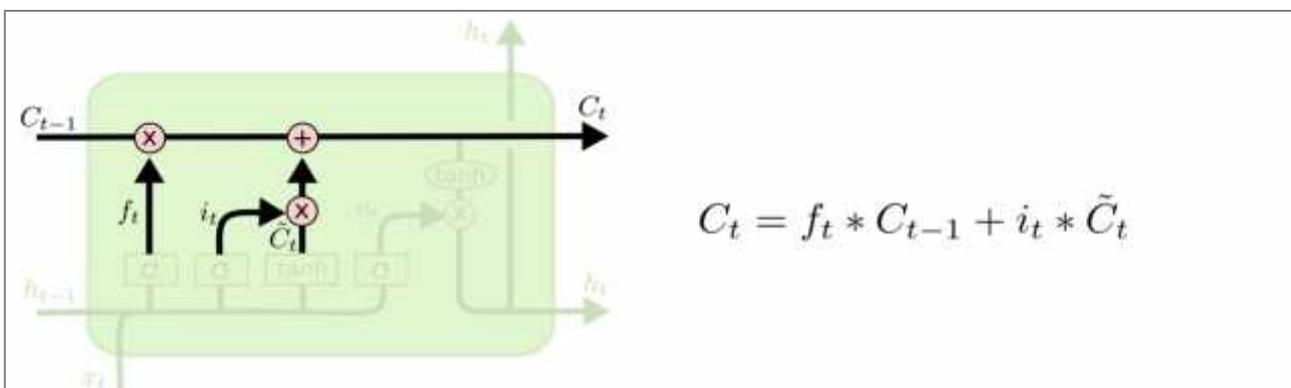


Figure 4.20 La troisième étape du modèle LSTM [136]

A la fin, nous devons décider de ce que nous allons avoir en sortie. Cette dernière sera basée sur notre état de cellule, mais sera une version filtrée. Tout d'abord, nous exécutons une couche sigmoïde qui détermine les parties de l'état de la cellule que nous allons afficher. Ensuite, nous mettons l'état de la cellule à travers \tanh (figure 4.21) (pour que les valeurs soient comprises entre -1 et 1) et multipliez-le par la sortie de la porte sigmoïde, de sorte que nous ne produisions que les parties que nous avons décidé. Pour l'exemple de modèle de langage, vu qu'il vient de voir un sujet, il peut vouloir sortir des informations pertinentes pour un verbe, au cas où ce serait la prochaine étape. Par exemple, il peut indiquer si le sujet est singulier ou pluriel, de sorte que nous sachions sous quelle forme un verbe devrait être conjugué si c'est ce qui suit par la suite.

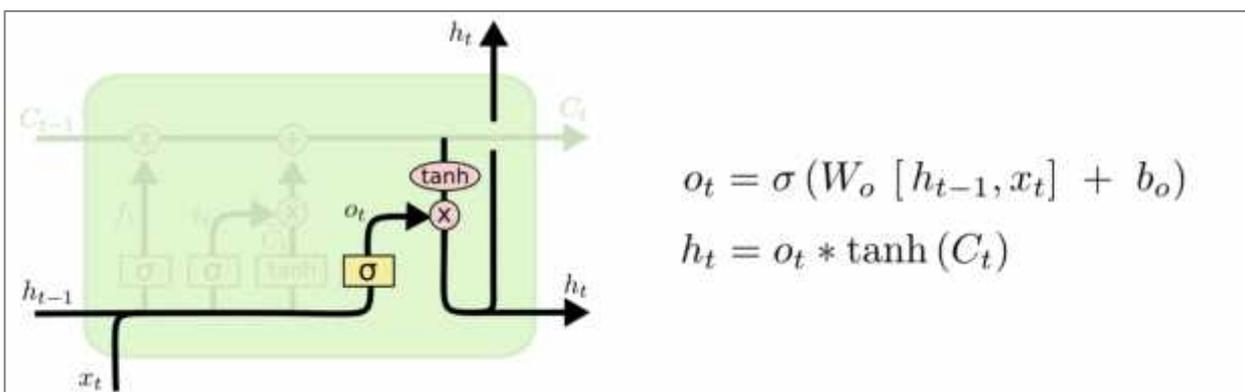


Figure 4.21 la dernière étape du model LSTM [136]

4.2.3.4 -Modèle BiLSTM :

LSTM dans son noyau, préserve l'information des entrées qui ont déjà passé à travers ,elle en utilisant l'état caché. Un BiLSTM calcule la séquence d'entrée de la direction opposée à une séquence cachée en avant et à une séquence cachée en arrière(Figure4.22). Le vecteur codé est formé par la concaténation des sorties finales avant et arrière, où est la séquence de sortie de la première couche cachée. Certains travaux antérieurs ont montré qu'en empilant plusieurs BiLSTM dans des réseaux de neurones, les performances de classification ou de régression peuvent être encore améliorées [137]

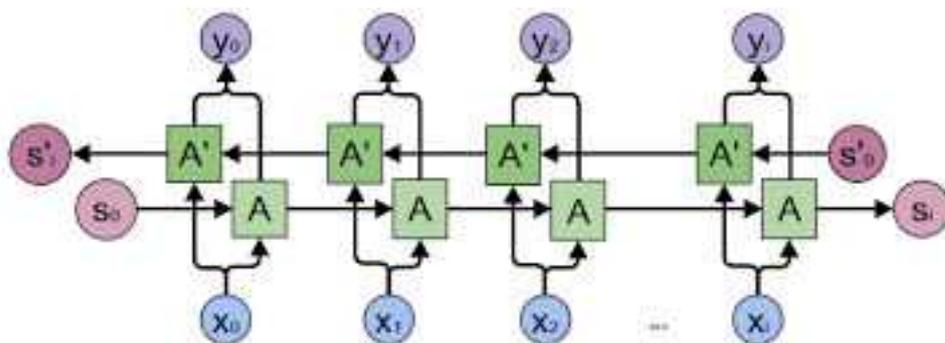


Figure 4.22 : Structure générale des réseaux BiLSTM [137]

4.2.4 Test et évaluation : Dans la couche finale, la représentation est transmise à la couche finale Soft max entièrement connectée en tant que vecteur de caractéristiques utilisé pour la classification. Le résultat sera une distribution de probabilité sur toutes les classes.

4.3 Conclusion

Pour conclure, nous avons passé en revue les différentes phases et les différentes couches de notre modèle proposé. Plus important encore, nous avons creusé profondément dans chaque partie pour donner une meilleure compréhension des bases de tous les modèles utilisés. Dans le chapitre suivant, nous décrirons le processus que nous avons suivi pour mettre en œuvre ce modèle et les résultats que nous avons obtenus à la fin.

Expérimentation et résultats

Chapitre 5

Expérimentation et résultats

5.1 Introduction

Maintenant que notre modèle a été conçu, il est temps de mettre la théorie en pratique et d'évaluer l'architecture proposée. Nous commencerons par présenter la configuration expérimentale que nous avons utilisée pour développer notre modèle et réaliser nos tests, puis nous partagerons plus de détails sur les données utilisées dans nos expériences et le processus suivi, et à la fin, nous partageons les résultats que nous avons obtenus.

5.2 Environnement de travail

5.2.1 Environnement matériel

Dans de telles études, le matériel joue un rôle crucial, en particulier lorsqu'il s'agit de savoir combien de données peuvent être traitées pour le temps donné et combien de temps il faut pour former le réseau neuronal profond et l'évaluer. Par conséquent, nous ressentons le besoin d'indiquer quel type de machine a été utilisé pour les expériences qui ont été faites et quelles sont ses caractéristiques.

Toutes les expériences décrites dans ce chapitre nous avons utilisé un serveur HP ProLiant ML350e Gen8 - Serveur - tour - 5U - 2 voies - 1 x Xeon E5-2407 / 2.2 GHz - RAM 32 Go - SATA - hot-swap 3.5" - HDD 2 x 500 Go - DVD - Matrox G200 - Gigabit LAN - Le serveur HP ProLiant ML350e Gen8 est une plate-forme double processeur au format tour.

5.2.2 Environnement logiciel

Pour mettre en œuvre que les modèles conçus et exécuter tous les programmes et tests, nous avons utilisé JetBrains Pycharm Community Edition 2018.1.2 environnement de développement [168]. Le lanceur est un IDE Python intelligent (Integrated Development Environment) avec une énorme collection d'outils sortis de la boîte, y compris un débogueur intégré et un coureur de test, profileur Python, un terminal intégré, et bien plus encore.

5.2.1 Python :

Le Python est un langage de programmation interprété, interactif, orienté objet. Il incorpore des modules, des exceptions, un typage dynamique et des classes. Python combine une puissance incroyable et une syntaxe très claire. Il dispose de liens avec de nombreuses interfaces de

programmation pour systèmes d'exploitation et bibliothèques, ainsi que divers systèmes d'interfaces graphiques.

5.2.2 Spyder 3.5

: Spyder est un environnement scientifique puissant écrit en Python. Il offre une combinaison unique des fonctionnalités avancées d'édition, d'analyse et de débogage.

5.2.3 Anaconda :

Anaconda est la plate-forme standard de la science des données Python, leader de l'innovation open source pour l'apprentissage automatique.

5.3 Package :

Comme il a été mentionné dans la section précédente, l'une des raisons pour lesquelles Python a été Utilisé pour ce projet est la quantité de bibliothèques importantes qui existent, par conséquent, nous sentons le besoin de présenter les plus importants qui ont été utilisés pour faire de notre modèle une réalité.

5.3.1 Package CSV

CSV (Comma Separated Values) module pour lire et écrire des données au format CSV. [145]

5.3.2 Package Numpy

Numpy(Numeric Python) est une bibliothèque numérique apportant aux expressions régulières. [138]

5.3.3 Package Nltk :

Nltk(Natural Language Toolkit) est une plate-forme pour la création de programmes Python pour travailler avec des données de langage humain. [139]

5.3.5 Package pandas

Pandas est une bibliothèque écrite pour le langage de programmation Python permettant la manipulation et l'analyse des données. Elle propose en particulier des structures de données et des opérations de manipulation de tableaux numériques et de séries temporelles. [140]

5.3.8 TensorFlow

[141] Un Framework de programmation pour le calcul numérique créé par Google et devenu un Framework Open Source en novembre 2015. Depuis, TensorFlow a continué à gagner une grand

importance et popularité et rapidement devenu l'un des Framework le plus utilisés pour le Deep Learning. Son nom s'inspire notamment du fait que les opérations actuelles des réseaux de neurones s'effectuent principalement via une table multidimensionnelle de données, appelée Tenseurs (Tensor) qui est l'équivalent d'une matrice. Aujourd'hui, ce framework est utilisé par les grandes entreprises [51] comme Intel, Twitter, Contribution 58 Lenovo, Nvidia, etc.

5.3.9 Keras : Keras

Un framework open source d'apprentissage profond pour le Python, capable de s'exécuter sur TensorFlow. Il est écrit par Francis Chollet, membre de l'équipe Google. Keras est utilisé dans un grand nombre de startups, de laboratoires de recherche [62] (dont le CERN, Microsoft Research et la NASA) et de grandes entreprises telles que Netflix, Yelp, Square, Uber, Google, etc. Il a été développé dans le cadre de l'effort de recherche du projet ONEIROS (Open-ended Neuro Electronic Intelligent Robot Operating System). En 2017, l'équipe TensorFlow de Google a pris la décision de fournir un support pour Keras et de l'intégrer dans la bibliothèque principale de TensorFlow. Il présente un ensemble d'abstractions de plus haut niveau et plus intuitives qui facilitent la configuration des réseaux neuronaux [144].

5.3.10 Gensim

[142] Gensim est une bibliothèque Python pour la modélisation de sujets, l'indexation de documents et la récupération de similarités avec les grands corpus. Le public cible est la communauté de traitement du langage naturel (NLP) et de récupération d'informations (IR).

5.4 Implémentation et Mise en œuvre

La Figure 5.1 présente les différentes phases de développement de notre modèle de détection de dépression.

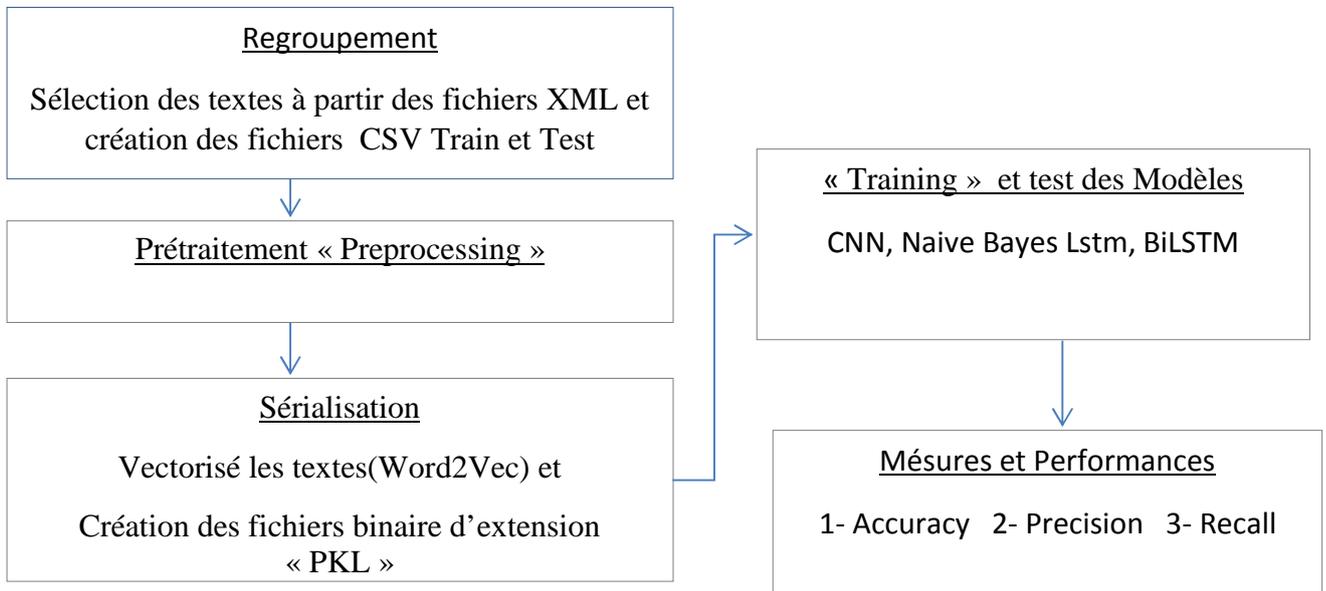


Figure 5.1 : Phases de développement du projet de notre modèle

5.4.1 Le traitement des données (fichiers)

5.4.1.1 Regroupement des Fichiers et Prétraitement

Dans ce qui suit nous allons présenter les procédures de regroupement et de prétraitement suivi dans notre travail, dont le but de cette étape est de nettoyer les tweets et leur rendre le plus proche possible à un langage formel.

D’abord nous avons commencé par le traitement des fichiers car les données sont stockées dans des fichiers d’extension « xml » et sont regroupées par utilisateur.

L’ensemble de formation contenait 486 utilisateurs (83 déprimés, 403 non déprimés) et l’ensemble de test contenait 401 utilisateurs (52 déprimés, 349 non déprimés)

5.4.1.2 Création des fichiers CSV

La structures des dossiers et des fichiers dans eRisk2019 est composée de :

Test data –T1\ T1_data.zip\data”: contient des fichiers d’extension « xml »

« training data – T1 » est divisé en deux(02) sous dossiers -

Negative-examples : contient les fichiers d'extension « xml » des données classées négativement par chunk de priorité (chunk1, chunk2...Chunk10)

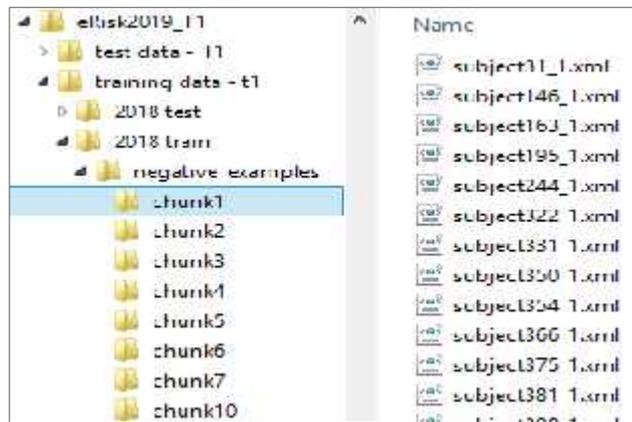


Figure 5.2 Classification des utilisateurs (fichiers.xml / chunk)

Positive-exemples : contient les fichiers d'extension « xml » des données classées négativement par chunk de priorité (chunk1, chunk2.... Chunk10).

5.4.1.3 La structure des Fichiers « xml »

La structure des fichiers « xml » est la suivante

<INDIVIDUAL>

<ID>subject758</ID>

<WRITING>

<DATE> 2013-10-27 13:26:17 </DATE>

<INFO> reddit post </INFO>

<TEXT> SMASH AND BURN PLS </TEXT>

</WRITING>

</INDIVIDUAL>

5.4.1.4 Procédure de traitement des fichiers

La procédure de traitement des fichiers consiste à faire :

- La sélection d'un dossier chunk
- La création d'un fichier csv nommé « *chunk + numéro de chunk+negative ou positive* »
- Insertion des textes des balises <TEXT> de tous les fichiers « xml » de dossier chunk sélectionné

```

#créer les chunk_positive_negative dans train
path_pos="C:\PXML\Train\positive/"
path_neg = "C:/PXML/Train/negative/"
for element in os.listdir(path_pos):
    if os.path.isdir(path_pos + element):
        f = open("C:/PXML/train/posneg/" + element + '.csv', 'w', newline='')
        fnames = ['text', 'sentiments']
        fout = csv.DictWriter(f, fieldnames=fnames)
        for fichier in glob.glob(path_pos + element + "/*.xml"):
            tree = etree.parse(fichier)
            for user in tree.xpath("/INDIVIDUAL/WRITING/TEXT"):
                # Travailler uniquement sur des chaines non vides
                if len(user.text) > 1:
                    text_trait = pretraitement(user.text)
                    # pas de sauvegarde si la liste est vide
                    if len(text_trait) > 1:
                        print(text_trait)
                        # prd=swn_polarity(text_trait)
                        label = 1 #sentiment positive
                        fout.writerow({'text': text_trait, 'sentiments': label})
        if os.path.isdir(path_neg + element):

```

Figure 5.3 Code source de la procédure de regroupement et de prétraitement des trains

```

path_test_posneg="C:\PXML/test/Positive_Negative/"

path_test = "C:\PXML/test/posneg/"

for element in os.listdir(path_test_posneg):

    if os.path.isdir(path_test_posneg + element):

        for fichier in glob.glob(path_test_posneg + element + "/*.xml"):

            fileName, fileExtension = os.path.splitext(basename(fichier))

            print(fileName)

```

Figure 5.4 Code source de la procédure de regroupement et de prétraitement des tests

5.4.1.5 Procédure de Prétraitement « Preprocessing »

Une fois les fichiers créés nous avons procédé à un prétraitement qui suit les étapes suivantes

- 1- Remplacer les émotions

- 2- Remplacer les abréviations
- 3- Supprimer les liens web(URL)
- 4- Supprimer les Hashtages(TAG)
- 5- Eliminer les caractères répétés
- 6- Supprimer les chiffres
- 7- Eliminer les ponctuations
- 8- Supprimer les mots vides.

```
Stopword= {'ourselves', 'hers', 'between', 'yourself', 'but', 'again', 'there', 'about', 'once',
'during', 'out', 'very', 'having', 'with', 'they', 'own', 'an', 'be', 'some', 'for', 'do', 'its', 'yours', 'such',
'into', 'of', 'most', 'itself', 'other', 'off', 'is', 's', 'am', 'or', 'who', 'as', 'from', 'him', 'each', 'the',
'themselves', 'until', 'below', 'are', 'we', 'these', 'your', 'his', 'through', 'don', 'nor', 'me', 'were', 'her',
'more', 'himself', 'this', 'down', 'should', 'our', 'their', 'while', 'above', 'both', 'up', 'to', 'ours', 'had',
'she', 'all', 'no', 'when', 'at', 'any', 'before', 'them', 'same', 'and', 'been', 'have', 'in', 'will', 'on', 'does',
'yourselves', 'then', 'that', 'because', 'what', 'over', 'why', 'so', 'can', 'did', 'not', 'now', 'under', 'he',
'you', 'herself', }
```

```
def pretraitement(text):
    TextRet= text #"liste de retour"
    print("Texte avant traintement",TextRet)
    TextRet=unicodedata.normalize('NFKD', TextRet).encode('ascii', 'ignore').decode(
        'utf-8', 'ignore')
    # La mise en minuscule de toutes nos données texte
    TextRet = TextRet.lower()
    # HTML tags
    TextRet = re.sub(r'<[^>]+>', '', TextRet)
    # @-mentions
    TextRet = re.sub(r'(?:@[\w_]+)', '', TextRet)
    # hash-tags
    TextRet = re.sub(r'(?:\#[\w_]+[\w\'\_\-]*[\w_]+)', '', TextRet)
    # URLs
    TextRet = re.sub(r'http[s]?://(?:[a-z]|[0-9]|[$-@.&+][!*\(\),]|(?:%[0-9a-f][0-9a-f]))+', '', TextRet)
    # numbers
    TextRet = re.sub(r'(?:(?:\d+,?)+(?:\.?\d+)?)', '', TextRet)
    # Parenthèses
    TextRet = re.sub(r'[\(\)]', '', TextRet)

    TextRet = TextRet.strip()
    print("Texte apres traintement", TextRet)
    return TextRet
```

Figure 5.4 Code source de la fonction prétraitement

5.4.1.6 Tokenizer

```
vocab_size = 20000
tokenizer = Tokenizer(num_words=vocab_size) # Setup tokenizer
tokenizer.fit_on_texts(texts)
sequences = tokenizer.texts_to_sequences(texts) # Generate sequences
word_index = tokenizer.word_index
```

Figure 5.5 Code source de la fonction Tokenizer

5.4.1.7 Vectorisation

GloVe : est un algorithme d'apprentissage non supervisé permettant d'obtenir des représentations vectorielles de mots.

```
glove_dir = 'C:/spyder/pkl/' # This is the folder with the dataset
embeddings_index = {} # We create a dictionary of word -> embedding
f = open(os.path.join(glove_dir, 'glove.6B.50d.txt'), encoding="utf8") # Open file
for line in f:
    values = line.split()
    word = values[0] # The first value is the word, the rest are the values of the embedding
    embedding = np.asarray(values[1:], dtype='float32') # Load embedding
    embeddings_index[word] = embedding # Add embedding to our embedding dictionary
f.close()
all_embs = np.stack(embeddings_index.values())
emb_mean = all_embs.mean() # Calculate mean
emb_std = all_embs.std() # Calculate standard deviation
print(emb_mean, emb_std)
embedding_dim = 100 # We use 100 dimensional glove vectors
word_index = tokenizer.word_index
nb_words = min(vocab_size, len(word_index)) # How many words are there actually
embedding_matrix = np.random.normal(emb_mean, emb_std, (nb_words, embedding_dim))
```

Figure 5.6 Code source de la fonction Vectorisation

5.5. Stockage binaire de toutes les étapes dans un fichier binaire « pickle »

Les fichiers dont l'extension est .PKL sont des fichiers de type (de format) : Exchange Migration Wizard Packing. Après avoir finalisé notre traitement, le stockage de cette phase est important pour l'utilisation dans l'apprentissage des modèles.

5.6 Evaluation

La conférence CLEF 2019 fournit une collection des données Dataset appelée eRisk (train et test) sur le site « early.irlab.org »

La collection de tests pour la tâche pilote consiste en une collection de messages d'un ensemble d'utilisateurs de médias sociaux. Il existe deux catégories d'utilisateurs, déprimés et non déprimés, et la collection contient, pour chaque utilisateur, une séquence d'écritures dans un ordre chronologique. Afin de simuler un scénario de détection précoce, le jeu de données a été divisé en 10 morceaux. Le premier bloc contient les 10% les plus anciens des messages, le deuxième bloc contient les 10% les plus anciens, et ainsi de suite. Deux fichiers ont été diffusés avec la collection eRisk 2019.

- Test_golden_truth.txt : pour les utilisateurs test e-Risk
- Risk_golden_truth.txt : pour les utilisateurs train e-Risk

```
import csv
import glob
import unicodedata
from os.path import basename, isfile

train_sentiment={}
def User_sentiments(fichier_user,sp):
    dic = {}
    # charnement du fichier
    if isfile(fichier_user):
        f_user = open(fichier_user, "r")
        lines = f_user.readlines()

        for line in lines:
            line4=line.replace("\n", "").split(sp)
            #print(line4)
            if len(line4)==2:
                lineKey.append(line4[0].strip())
                lineatt.append(line4[1])
                if line4[1]=="1" or line4[1]==1:
                    testpos+=1
                testtotal+=1
```

En résumé, l'ensemble de formation contenait 486 utilisateurs (83 déprimés, 403 non déprimés) et l'ensemble de test contenait 401 utilisateurs (52 déprimés, 349 non déprimés).

	Training data	Test data
Utilisateur	486	401
Positif	83	52
Négatif	434	349
Documents	531,394	544,447
Commentaire(sans titre)	367,439	366,845
Une seule ligne (texte vide)	141,849	147,197
Document vide	91	219
La moyenne de document par utilisateur	599,09	663,96
Moyenne de caractères par document	174,54	197,47

Tableau 5.1 : Les résultats obtenus après traitement des fichiers.

5.7 Mesures et performances

Maintenant que notre architecture est prête, avant de la mettre en œuvre, nous devons définir les paramètres qui nous permettront de mesurer les résultats de notre travail et de le comparer à d'autres travaux.

Pour ce faire, nous devons introduire ce qu'on appelle la matrice de confusion et les différentes variables que nous utiliserons pour calculer toutes nos mesures de rendement [144].

TP : Un **vrai positif** est un résultat où le modèle prédit correctement la **classe positive**.

TN : Un **vrai négatif** est un résultat où le modèle prédit correctement la **classe négative**.

FP : Un **faux positif** est un résultat où le modèle prédit incorrectement la **classe positive**.

FN : Un **faux négatif** est un résultat où le modèle prédit incorrectement la **classe négative**.

5.7.1 Précision

La précision est le rapport entre les observations positives correctement prédites et le total des observations positives prédites

$$\text{Précision} = \text{TP} / \text{TP} + \text{FP}$$

5.7.2 Recall

Le rappel est le rapport entre les observations positives correctement prédites et toutes les observations de la classe réelle

$$\text{Recall} = \text{TP} / \text{TP} + \text{FN}$$

5.7.3 F1-Score

La moyenne pondérée de la précision et du rappel. Par conséquent, ce score prend en compte à la fois les faux positifs et les faux négatifs. Intuitivement, il n'est pas aussi facile à comprendre que l'exactitude, mais F1 est généralement plus utile que l'exactitude, en particulier si la répartition des classes est inégale. La précision fonctionne mieux si les faux positifs et les faux négatifs ont un coût similaire. Si le coût des faux positifs et des faux négatifs est très différent, il vaut mieux regarder à la fois Précision et Recall.

$$\text{Score F1} = 2 * (\text{Rappel} * \text{Précision}) / (\text{Rappel} + \text{Précision})$$

5.7.4 Accuracy

La précision est la mesure de performance la plus intuitive. Il s'agit simplement d'un ratio d'observations correctement prédites par rapport au nombre total d'observations.

$$\text{Accuracy} = \text{TP} + \text{TN} / \text{TP} + \text{FP} + \text{FN} + \text{TN}$$

5.8 Phase d'apprentissage « training »

Dans cette section, nous allons passer en revue le processus détaillé que nous avons suivi pour former notre modèle. Puisque notre modèle inclut différents types de réseaux neuronaux, nous allons les parcourir séparément. Afin de tirer le meilleur parti de ce projet, étant donné que nous avons des ensembles de données axés sur la dépression, la meilleure option que nous pouvions adopter était d'utiliser une approche d'apprentissage supervisé.

5.8.1 Apprentissage de notre modèle

L'expérience la plus importante, et celle pour laquelle toute cette étude a été faite, nous avons utilisé le fichier pré-classifier de notre dataset pour alimenter notre première partie du modèle en sortie nous avons pris le fichier PKL a lui d'alimenté les quatre modèles cités auparavant, le modèle CNN, LSTM, BiLSTM et la fin Naïve bayes.

5.8.2 Résultats

Le tableau suivant présente les résultats de chaque expérience, en plus des résultats partagés avec qui utilisaient les mêmes ensembles de données que nous, mais avec des modèles statistiques à la place.

Modèles	Accuracy	Recall	F1-score
Naive Bayes	0.82	0.77	0.77
Modèle CNN	0.86	0.79	0.78
LSTM	0.90	0.89	0.84
BiLSTM	0.97	0.95	0.92

Tableau : 5.2 Etat de comparatif des résultats obtenus

Suivant les résultats affichés dans le tableau 5.2 et les graphes (Figure 5.2 et Figure 5.3), nous remarquons que le modèle BiLSTM surpasse les trois autres modèles avec un degré élevé (Accuracy =97 %, Recall=95 % et F1-score=92), nous concluons que le modèle BiLSTM est plus performant par rapport aux autres modèles selon notre étude et nos résultats.

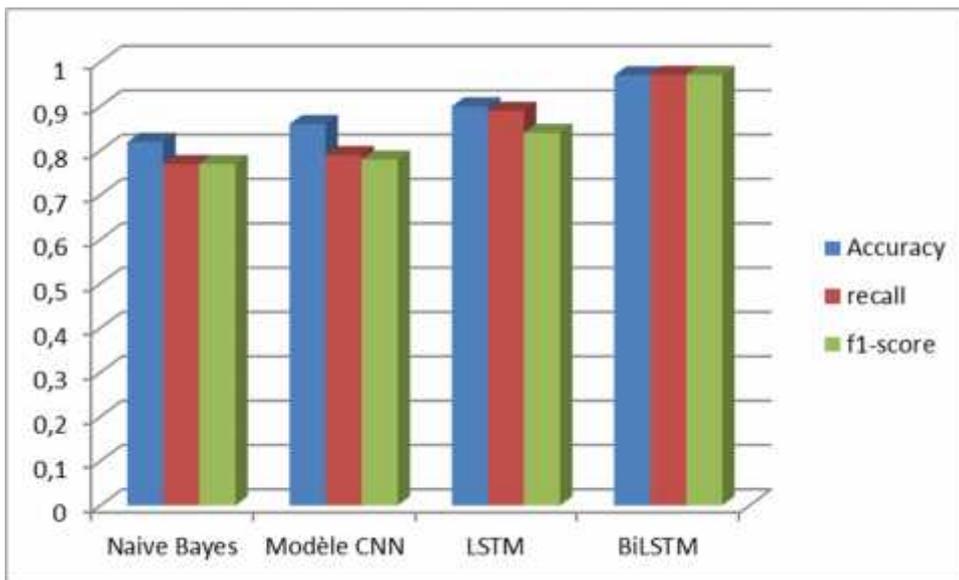


Figure 5.8 : Représentation graphique « Histogramme 3D » des résultats

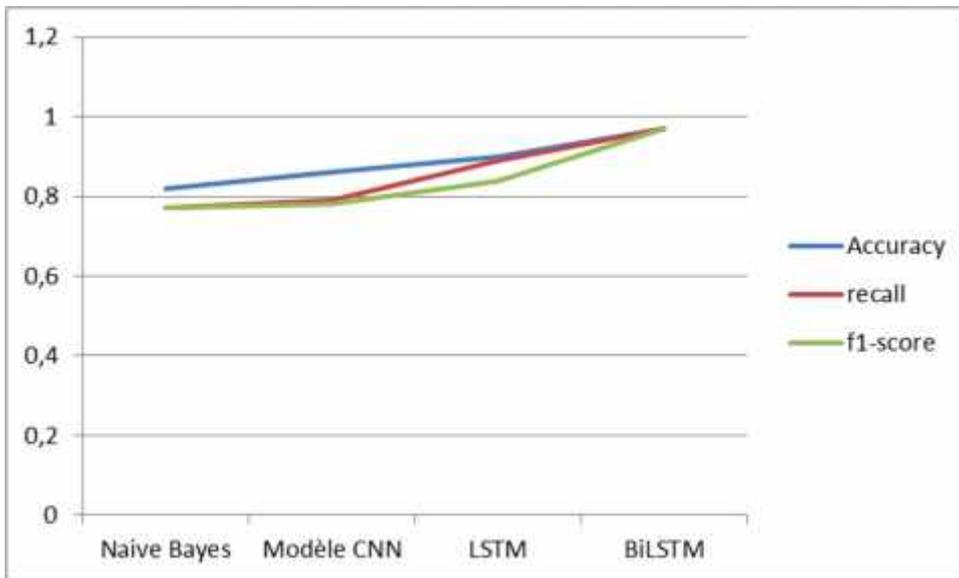


Figure 5.9 : Représentation graphique « Courbe » des résultats.

5.9 Conclusion

Dans ce chapitre, nous avons vu notre part de contribution au problème de l'analyse du sentiment, représentant les outils et les jeux de données utilisés, ainsi que les étapes que nous avons suivies pour obtenir les résultats que nous montrons également pour les différents modèles dont le but de faire la comparaison.

Bibliographie

Bibliographie

- [1] Cambridge Dictionary, "Meaning in the Cambridge English Dictionary," Cambridge Dictionary, 2018. [Online]. Available: <https://dictionary.cambridge.org/dictionary/english/sentiment>. [Accessed: 23-Jun-2018].
- [2] Cambridge Dictionary, "Meaning in the Cambridge English Dictionary," Cambridge Dictionary, 2018. [Online]. Available: <https://dictionary.cambridge.org/dictionary/english/sentiment>. [Accessed: 23-Jun-2018].
- [3] B. Liu, *Sentiment Analysis and Opinion Mining*, no. May. 2012.
- [4] K. Dave, K. Dave, S. Lawrence, S. Lawrence, D. M. Pennock, and D. M. Pennock, "Mining The peanut gallery: Opinion extraction and semantic classification of product reviews," *Proc. 12th Int. Conf. World Wide Web*, pp. 519–528, 2003.
- [5] T. Nasukawa, T. Nasukawa, J. Yi, and J. Yi, "Sentiment analysis: Capturing favorability using natural language processing," *Proc. 2nd Int. Conf. Knowl. Capture*, pp. 70–77, 2003.
- [6] S. Das and M. Chen, "Yahoo! for Amazon: Extracting market sentiment from stock message boards," *Proc. Asia Pacific Financ. ...*, vol. 35, pp. 43, 2001.
- [7] S. Morinaga, K. Yamanishi, K. Tateishi, and T. Fukushima, "Mining product reputations on the Web," in *Proceedings of the eighth ACM SIGKDD international conference on Knowledge discovery and data mining - KDD '02*, 2002, p. 341.
- [8] B. Pang, L. Lee, and S. Vaithyanathan, "Thumbs up?: sentiment classification using machine learning techniques," *Empir. Methods Nat. Lang. Process.*, vol. 10, no. July, pp. 79–86, 2002.
- [9] R. Tong, "An operational system for detecting and tracking opinions in on-line discussion," *Work. Notes ACM SIGIR 2001 Work. Oper. Text Classif.*, pp. 1–6, 2001.
- [10] P. D. Turney, "Thumbs up or thumbs down?," in *Proceedings of the 40th Annual Meeting on Association for Computational Linguistics - ACL '02*, 2001, no. July, p. 417.
- [11] J. M. Wiebe, "Learning subjective adjectives from corpora," *Proc. Natl. Conf. Artif. Intell.*, no. 1, pp. 735–741, 2000.
- [12] M. McGlohon, N. Glance, and Z. Reiter, "Star Quality : Aggregating Reviews to Rank Products and Merchants," *Wall Str. J.*, pp. 114–121, 2010.
- [13] Y. Hong and S. Skiena, "The Wisdom of Bookies? Sentiment Analysis Versus. the NFL Point Spread.," *Fourth Int. AAAI Conf. Weblogs Soc. Media*, pp. 251–254, 2010.
- [14] B. Chen, L. Zhu, D. Kifer, and D. Lee, "What Is an Opinion About? Exploring Political Standpoints Using Opinion Scoring Model," *Twenty-Fourth AAAI Conf. Artif. Intell.*, pp. 1007–1012, 2010.
- [15] M. Miller, C. Sathi, D. Wiesensthal, J. Leskovec, and C. Potts, "Sentiment Flow Through Hyperlink Networks.," *Fifth Int. AAAI Conf. Weblogs Soc. Media*, pp. 550–553, 2011.
- [16] B. O 'connor, R. Balasubramanyan, B. R. Routledge, and N. A. Smith, "From Tweets to Polls: Linking Text Sentiment to Public Opinion Time Series."
- [17] S. Mohammad and T. Yang, "Tracking Sentiment in Mail: How Genders Differ on Emotional Axes," *Proc. ACL 2011 Work. Comput. Approaches to Subj. Sentim. Anal.*, pp. 70–79, 2011.
- [18] S. M. Mohammad, "From once upon a time to happily ever after: Tracking emotions in mail and books," in *Decision Support Systems*, 2012, vol. 53, no. 4, pp. 730–741.

Bibliographie

- [19] R. Bar-Haim, E. Dinur, R. Feldman, M. Fresko, and G. Goldstein, "Identifying and following expert investors in stock microblogs," *Proc. Conf. Empir. Methods Nat. Lang. Process.*, pp. 1310–1319, 2011.
- [20] R. Feldman, B. Rosenfeld, R. Bar-Haim, and M. Fresko, "The Stock Sonar — Sentiment Analysis of Stocks Based on a Hybrid Approach," *Iaai*, no. c, pp. 1642–1647, 2011.
- [21] P. Sakunkoo and N. Sakunkoo, "Analysis of Social Influence in Online Book Reviews," *AAAI's ICWSM 2009*, pp. 1968–1970, 2009.
- [22] G. Groh and J. Hauffa, "Characterizing Social Relations Via NLP-based Sentiment Analysis," *Proc. Fifth Int. AAAI Conf. Weblogs Soc. Media*, pp. 502–505, 2011.
- [23] A. Tumasjan, T. Sprenger, P. Sandner, and I. Welpel, "Predicting elections with Twitter: What 140 characters reveal about political sentiment," *Proc. Fourth Int. AAAI Conf. Weblogs Soc. Media*, pp. 178–185, 2010.
- [24] T. Yano and N. A. Smith, "What 's Worthy of Comment ? Content and Comment Volume in Political Blogs *," *Fourth Int. AAAI Conf. Weblogs Soc. Media*, vol. i, pp. 359–362, 2010.
- [25] Y. Liu, J. Huang, A. An, and X. Yu, "ARSA: A sentiment-aware model for predicting sales performance using blogs," *Proc. ACM Spec. Interes. Gr. Inf. Retr. (SIGIR)*, 2007.
- [26] S. Asur and B. A. Huberman, "Predicting the future with social media," *Proc. Int. Conf. Web Intell. Intell. Agent Technol.*, pp. 492–499, 2010.
- [27] M. Joshi, D. Das, K. Gimpel, and a. N. Smith, "Movie Reviews and Revenues: An Experiment in Text Regression," *Hum. Lang. Technol. 2010 Annu. Conf. North Am. Chapter Assoc. Comput. Linguist.*, pp. 293–296, 2010.
- [28] E. Sadikov, A. Parameswaran, and P. Venetis, "Blogs as predictors of movie success," *Third Int. AAAI Conf. Weblogs Soc. Media*, pp. 304–307, 2009.
- [29] J. Bollen, H. Mao, and X.-J. Zeng, "[1010.3003] Twitter mood predicts the stock market." [Online]. Available: <http://arxiv.org/abs/1010.3003>. [Accessed: 23-Jun-2018].
- [30] W. Zhang and S. Skiena, "Trading Strategies to Exploit Blog and News Sentiment.," *Icwsn*, vol. 34, no. Chan 2003, pp. 375–378, 2010.
- [31] S. O'Dowd. Unstructured data and text analytics incapital markets, IDC research document, docfin212553/july, 2008
- [32] H. Takeuchi, L. V. Subramaniam, T. Nasukawa, andS. Roy. Getting insights from the voices of customers:Conversation mining at a contact center.*Inf. Sci.*,179(11):1584–1591, 2009
- [33] C. Zhang, D. Zeng, J. Li, F.-Y. Wang, W. Zuo, Sentiment analysis of Chinese documents: from sentence to document level, *J. Am. Soc. Inform. Sci. Technol.* 60 (12) (2009) 2474–2487
- [34]R. Moraes, J.F. Valiati, W.P. GaviãoNeto, Document-level sentiment classification: an empirical comparison between SVM and ANN, *Expert Syst. Appl.* 40 (2) (2013) 621–633
- [35] G. Li, F. Liu, Application of a clustering method on sentiment analysis, *J. Inform. Sci.* 38 (2) (2012) 127–139
- [36] Pang B., Lee, L., AndVaithyanathain. S. 2002.Thumbs up? Sentiment classification using machine learning techniques, In *Proceedings of the Conference on Empirical Methods in Natural Language Processing* , 79-86
- [37] Neha S. Joshi, Suhasini A. Itkat"A Survey on Feature Level Sentiment Analysis", (*IJCSIT*) *International Journal of Computer Science and Information Technologies*, Vol. 5 (4) , (2014),PP: 5422-5425 .

Bibliographie

- [38] J. M. Wiebe, R. F. Bruce, and T. P. O'Hara, "Development and use of a gold standard data set for subjectivity classifications," in Proceedings of the Association for Computational Linguistics, 1999, pp. 246–253.
- [39] B.A. Ojokoh, O. Kayode, A feature-opinion extraction approach to opinion mining, *J. Web Eng.* 11 (1) (2012) 51–63
- [40] O. Vechtomova, Facet-based opinion retrieval from blogs, *Inform. Process. Manage.* 46 (1) (2010) 71–88
- [41] Neha S. Joshi, Suhasini A. Itkat "A Survey on Feature Level Sentiment Analysis", (IJCSIT) International Journal of Computer Science and Information Technologies, Vol. 5 (4) , (2014), PP: 5422-5425.
- [42] Y. R. Tausczik and J. W. Pennebaker. The psychological meaning of words: Liwc and computerized text analysis methods. *Journal of Language and Social Psychology*, 29(1):254, 2010 .
- [43] E. Cambria, R. Speer, C. Havasi, and A. Hussain. Senticnet: A publicly available semantic resource for opinion mining. In *AAAI Fall Symposium Series*, 2010
- [44] E. Cambria, R. Speer, C. Havasi, and A. Hussain. Senticnet: A publicly available semantic resource for opinion mining. In *AAAI Fall Symposium Series*, 2010
- [45] J. Read. Using emoticons to reduce dependency in machine learning techniques for sentiment classification. In *ACL Student Research Workshop*, pages 43–48, 2005.
- [46] Y. R. Tausczik and J. W. Pennebaker. The psychological meaning of words: Liwc and computerized text analysis methods. *Journal of Language and Social Psychology*, 29(1):24–54, 2010.
- [47] M. Thelwall, "Heart and soul: Sentiment strength detection in the social web with sentiment strength," *Proc. CyberEmotions*, vol. 5, pp. 1–14, 2013.
- [48] B. Liu, *Sentiment Analysis and Opinion Mining*, no. May. 2012.
- [49] D. L. Poole, A. Mackworth, and R. G. Goebel, "Computational Intelligence and Knowledge," *Comput. Intell. A Log. Approach*, no. Ci, pp. 1–22, 1998.
- [50] T. Hastie, R. Tibshirani, and J. Friedman, *The elements of statistical learning: data mining, inference and prediction*, vol. 9. Springer, 2017 .
- [51] . S. Breese, D. Heckerman, and C. Kardie, Empirical analysis of predictive algorithms for collaborative filtering, *Proceedings of the 14th Conference on Uncertainty in Artificial Intelligence*, 1998, pp. 43-52.

Bibliographie

- [52] L. Dormehl, “What is an artificial neural network? Here’s everything you need to know,” Tech Radar, 2018. [Online]. Available: <https://www.techradar.com/news/what-is-5-everything-you-need-to-know>. [Accessed: 23-Jun-2018].
- [53] Y. Lecun, Y. Bengio, and G. Hinton, “Deep learning,” *Nature*, vol. 521, no. 7553, pp. 436–444, 2015.
- [54] A. Krizhevsky, I. Sutskever, and G. E. Hinton, “ImageNet Classification with Deep Convolutional Neural Networks,” *Adv. Neural Inf. Process. Syst.*, vol. 60, no. 6, pp. 1–9, 2012.
- [55] C. Faret, C. Couprie, L. Najman, and Y. Lecun, “Learning Hierarchical Features for Scene Labeling,” *Pattern Anal. Mach. Intell. IEEE Trans.*, vol. 35, no. 8, pp. 1915–1929, 2013.
- [56] J. Tompson, A. Jain, Y. LeCun, and C. Bregler, “Joint Training of a Convolutional Network and a Graphical Model for Human Pose Estimation,” *Adv. Neural Inf. Process. Syst.*, pp. 1799–1807, 2014.
- [57] C. Szegedy et al., “Going deeper with convolutions,” in *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, 2015, vol. 07–12–June, pp. 1–9.
- [58] T. Mikolov, A. Deoras, D. Povey, L. Burget, and J. ernocký, “Strategies for training large scale neural network language models,” 2011 IEEE Work. Autom. Speech Recognit. Understanding, ASRU 2011, Proc., pp. 196–201, 2011.
- [59] G. Hinton et al., “Deep Neural Networks for Acoustic Modeling in Speech Recognition,” *IEEE Signal Process. Mag.*, vol. 29, no. 6, pp. 82–97, 2012.
- [60] T. N. Sainath, A. Mohamed, B. Kingsbury, B. Ramabhadran, I. B. M. T. J. Watson, and Y. Heigley, “Deep Convolutional Neural Networks for LVCSR,” *Icassp 2013*, pp. 8614–8618, 2013.
- [61] J. Ma, R. P. Sheridan, A. Liaw, G. E. Dahl, and V. Svetnik, “Deep neural nets as a method for quantitative structure-activity relationships,” *J. Chem. Inf. Model.*, vol. 55, no. 2, pp. 263–274, 2015.
- [62] T. Ciodaro, D. Deva, J. M. De Seixas, and D. Damazio, “Online particle detection with neural networks based on topological calorimetry information,” in *Journal of Physics: Conference Series*, 2012, vol. 368, no. 1.
- [63] “Higgs Boson Machine Learning Challenge | Kaggle.” [Online]. Available: <https://www.kaggle.com/c/higgs-boson>. [Accessed: 23-Jun-2018].
- [64] M. Helmstaedter, K. L. Briggman, S. C. Turaga, V. Jain, H. S. Seung, and W. Denk, “Connectomic reconstruction of the innerplexiform layer in the mouse retina,” *Nature*, vol. 500, no. 7461, pp. 168–174, 2013.
- [65] M. K. K. Leung, H. Y. Xiong, L. J. Lee, and B. J. Frey, “Deep learning of the tissue-regulated splicing code,” *Bioinformatics*, vol. 30, no. 12, pp. 121–129, 2014.
- [66] H. Y. Xiong et al., “The human splicing code reveals new insights into the genetic determinants of disease,” *Science (80-.)*, vol. 347, no. 6218, 2015.

Bibliographie

- [67] R. Collobert, J. Weston, L. Bottou, M. Karlen, K. Kavukcuoglu, and P. Kuksa, “Natural Language Processing (Almost) from Scratch,” *J. Mach. Learn. Res.*, vol. 12, pp. 2493–2537, 2011.
- [68] A. Bordes, S. Chopra, and J. Weston, “Question Answering with Subgraph Embeddings,” pp. 615–620, 2014.
- [69] S. Jean, K. Cho, R. Memisevic, and Y. Bengio, “On Using Very Large Target Vocabulary for Neural Machine Translation,” pp. 1–10, 2014.
- [70] I. Sutskever, O. Vinyals, and Q. V Le, “Sequence to Sequence Learning with Neural Networks,” in *Proceedings of the 27th International Conference on Neural Information Processing Systems*, 2014, pp. 3104–3112.
- [71] G. Zweig and P. Nguyen, “A segmental CRF approach to large vocabulary continuous speech recognition,” *IEEE Work. Autom. Speech Recognit. Underst.*, pp. 152–157, 2009.
- [72] G. Heigold, H. Ney, P. Lehen, T. Gass, and R. Schlüter, “Equivalence of Generative and Log-Linear Models,” *IEEE Trans. Audio, Speech Lang. Process.*, vol. 19, no. 5, pp. 1138–1148, 2011.
- [73] D. Yu, S. Wang, and L. Deng, “Sequential labeling using deep-structured conditional random fields,” *IEEE J. Sel. Top. Signal Process.*, vol. 4, no. 6, pp. 965–973, 2010.
- [74] D. Yu and L. Deng, “Deep-structured hidden conditional random fields for phonetic recognition,” *Interspeech*, vol. 0, no. September, pp. 2986–2989, 2010.
- [75] Y. Bengio, N. Boulanger-Lewandowski, and R. Pascanu, “Advances in optimizing recurrent networks,” in *ICASSP, IEEE International Conference on Acoustics, Speech and Signal Processing - Proceedings*, 2013, pp. 8624–8628.
- [76] Y. Bengio, P. Lamblin, D. Popovici, and H. Larochelle, “Greedy layer-wise training of deep networks,” in *Adv. in Neural Inf. Proc. Syst.* 19, 2006, pp. 153–160.
- [77] L. Deng, M. Seltzer, D. Yu, A. Acero, A.-R. Mohamed, and G. Hinton, “Binary Coding of Speech Spectrograms Using a Deep Autoencoder,” *Interspeech*, no. September, pp. 1692–1695, 2010.
- [78] G. E. Hinton and R. R. Salakhutdinov, “Reducing the dimensionality of data with neural networks,” *Science (80-.)*, 2006.
- [79] Y. Bengio, “Learning Deep Architectures for AI,” *Found. Trends@ Mach. Learn.*, vol. 2, no. 1, pp. 1–127, 2009.
- [80] Y. LeCun, S. Chopra, M. A. Ranzato, and F. J. Huang, “Energy-based models in document recognition and computer vision,” in *Proceedings of the International Conference on Document Analysis and Recognition, ICDAR*, 2007, vol. 1, pp. 337–341.

Bibliographie

- [81] J. Ngiam, Z. Chen, P. W. Koh, and A. Y. Ng, "Learning Deep Energy Models," Proc. 28th Int. Conf. Mach. Learn. (ICML 11), pp. 1105–1112, 2011.
- [82] D. Erhan, A. Courville, and P. Vincent, "Why Does Unsupervised Pre-training Help Deep Learning?," J. Mach. Learn. Res., vol. 11, pp. 625–660, 2010.
- [83] A.~Mohamed, D.~Yu, and L.~Deng, "Investigation of full-sequence training of deep belief networks for speech recognition," in Proc. Interspeech, 2010..
- [84] H. Lee, R. Grosse, R. Ranganath, and A. Y. Ng, "Convolutional deep belief networks for scalable unsupervised learning of hierarchical representations," in Proceedings of the 26th Annual International Conference on Machine Learning - ICML '09, 2009, pp. 1–8.
- [85] H. Lee, R. Grosse, R. Ranganath, and A. Y. Ng, "Unsupervised learning of hierarchical representations with convolutional deep belief networks," Commun. ACM, vol. 54, no. 10, p. 95, 2011.
- [86] H. Lee, P. T. Pham, Y. Largman, and A. Y. Ng, "Unsupervised feature learning for audio classification using convolutional deep belief networks.," Nips, vol. 9, pp. 1096–1104, 2009.
- [87] Y. Bengio, A. Courville, and P. Vincent, "Representation learning: a review and new perspectives.," IEEE Trans. Pattern Anal. Mach. Intell., vol. 35, no. 8, pp. 1798–1828, 2013.
- [88] Y. Bengio, R. De Mori, G. Flammia, and R. Kompe, "Global Optimization of a Neural Network - Hidden Markov Model Hybrid," IEEE Trans. Neural Networks, vol. 3, no. 2, pp. 252–259, 1992.
- [89] Y. Bengio et al., "A Neural Probabilistic Language Model," J. Mach. Learn. Res., vol. 3, pp. 1137–1155, 2003.
- [90] E. Battenberg and D. Wessel, "Analyzing Drum Patterns using Conditional Deep Belief Networks," Int. Soc. Music Inf. Retr. Conf., pp. 37–42, 2012.
- [91] Y. Bengio, É. Thibodeau-Laufer, G. Alain, and J. Yosinski, "Deep Generative Stochastic Networks Trainable by Backprop," 2013.
- [92] Y. Bengio, L. Yao, G. Alain, and P. Vincent, "Generalized Denoising Auto-Encoders as Generative Models," 2013.
- [93] J. Bergstra and U. YoshuaBengio, "Random Search for Hyper-Parameter Optimization," J. Mach. Learn. Res., vol. 13, pp. 281–305, 2012.
- [94] A. Biem, S. Katagiri, E. McDermott, and B. H. Juang, "An application of discriminative feature extraction to filter-bank-based speech recognition," IEEE Trans. Speech Audio Process., vol. 9, no. 2, pp. 96–109, 2001.

Bibliographie

- [95] J. Bilmes, "Dynamic graphical models," *IEEE Signal Process. Mag.*, vol. 27, no. 6, pp. 29–42, 2010.
- [96] J. A. Bilmes and C. Bartels, "Graphical model architectures for speech recognition," *IEEE Signal Process. Mag.*, vol. 22, no. 5, pp. 89–100, 2005.
- [97] X. Glorot, A. Bordes, J. Weston, and Y. Bengio, "A Semantic Matching Energy Function for Learning with Multi-relational Data," 2013.
- [98] A. Bordes and J. Weston, "Learning Structured Embeddings of Knowledge Bases," *Artif. Intell.*, no. Bengio, pp. 301–306, 2009.
- [99] L. Bottou, "From machine learning to machine reasoning: An essay," *Mach. Learn.*, vol. 94, no. 2, pp. 133–149, 2014.
- [100] Y. Bengio, "Artificial neural networks and their application to sequence recognition," no. June, 1991.
- [101] E. Arisoy et al., "Deep Neural Network Language Models," *NAACL-HLT 2012 Work. Will We Ever Really Replace N-gram Model. Futur. Lang. Model. HLT*, pp. 20–28, 2012.
- [102] L. B. Y. Le Cun and L. Bottou, "Large scale online learning," *Adv. Neural Inf. Process. Syst.*, vol. 16, p. 217, 2004.
- [103] Eva Sharma and Munmun De Choudhury "Mental Health Support and its Relationship to Linguistic Accommodation in Online Community" ,pp 1-16 2018.
- [104] Albert Park, Andrea L Hartzler, Jina Huh, David WMcDonald, and Wanda Pratt. 2015.
- [105] . Tom Postmes, Russell Spears, and Martin Lea. 2000. Theformation of group norms in computer-mediatedcommunication. *Human communication research* 26, 3(2000), 341–371.
(2000), 341–371.
- [106] Jenny Preece. 2000. *Online communities: Designing usability and supporting socialbilty*. John Wiley & Sons, Inc.
- [107]. Jenny Preece, Diane Maloney-Krichmar, and ChadiaAbrás. 2003. History of emergence of onlinecommunities. *Encyclopedia of Community*. BerkshirePublishing Group, Sage (2003), 1–11.
- [108.] r/modnews. 2016. Moderators: Sticky Comments is nowavailable to all mods. https://www.reddit.com/r/modnews/comments/3x8vzl/moderators_sticky_comments_is_now_available_to/?st=j7qv8sla&sh=b5bd21b6. (2016). [Online;accessed 16-September-2017].

Bibliographie

- [109] Catherine Schaefer, James C Coyne, and Richard SLazarus. 1981. The health-related functions of socialsupport. *Journal of behavioral medicine* 4, 4 (1981), 381–406.
- [110] Thomas J Scheff. 1967. *Mental illness and social processes*. Harper & Row.
- [111]. Hannah R Snyder, Natalie Hutchison, Erika Nyhus, TimCurran, Marie T Banich, Randall C O'Reilly, and YukoMunakata. 2010. Neural inhibition enables selection; During language processing. *Proceedings of the National Academy of Sciences* 107, 38 (2010), 16483–16488
- .
- [112] Claire F Sullivan. 1996. Recipients' perceptions ofsupport attempts across various stressful life events. *Communication Research Reports* 13, 2 (1996), 183–190.
- [113]. Yla R Tausczik and James W Pennebaker. 2010. The psychological meaning of words: LIWC and computerized text analysis methods. *JLS* 29, 1 (2010), 24–54.
- [114]. Paul J Taylor and Sally Thomas. 2008. Linguistic stylematching and negotiation outcome. *NCMR* 1, 3 (2008), 263–281.
- [115] Catalina L Toma. 2014. Towards conceptual convergence: An examination of interpersonal adaptation. *Communication Quarterly* 62, 2 (2014), 155–178.
- [116] R Jay Turner, B Gail Frankel, and Deborah M Levin. 1983. *Social support: Conceptualization, measurement, and implications for mental health*. Research in community & mental health (1983).
- [117]. Tatiana A Vlahovic, Yi-Chia Wang, Robert E Kraut, and John M Levine. 2014. Support matching and satisfaction in an online breast cancer support community. In *CHI*. ACM, 1625–1634.
- [118] Joseph B Walther and Shawn Boyd. 2002. Attraction to computer-mediated social support. *Communication technology and society: Audience adoption and uses* 153188 (2002).
- [119] Q. Tul et al., "Sentiment Analysis Using Deep Learning Techniques: A Review," *Int. J. Adv. Comput. Sci. Appl.*, vol. 8, no. 6, 2017.
- [120] L. Arnold, S. Rebecchi, S. Chevallier, and H. Paugam-Moisy, "An Introduction to Deep Learning," *Esann*, no. April, p. 12, 2011.
- [121] R. Goebel and W. Wahlster, "Integrated Uncertainty in Knowledge Modelling and Decision Making," vol. 7027, 2011.
- [122] A. Graves, N. Jaitly, and A.-R. Mohamed, "Hybrid Speech Recognition With Deep Bidirectional Lstm," 2013.
- [123] C. Dos Santos and M. Gatti, "Deep convolutional neural networks for sentiment analysis of short texts," ... 25th *Int. Conf.*, pp. 69–78, 2014.

Bibliographie

- [124] K. Ravi and V. Ravi, "Sentiment classification of Hinglish text," 2016 3rd Int. Conf. Recent Adv. Inf. Technol. RAIT 2016, pp. 641–645, 2016.
- [54] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "ImageNet Classification with Deep Convolutional Neural Networks," *Adv. Neural Inf. Process. Syst.*, vol. 60, no. 6, pp. 1–9, 2012.
- [51] C. Farabet, C. Couprie, L. Najman, and Y. Lecun, "Learning Hierarchical Features for Scene Labeling," *Pattern Anal. Mach. Intell. IEEE Trans.*, vol. 35, no. 8, pp. 1915–1929, 2013.
- [56] J. Tompson, A. Jain, Y. LeCun, and C. Bregler, "Joint Training of a Convolutional Network and a Graphical Model for Human Pose Estimation," *Adv. Neural Inf. Process. Syst.*, pp. 1799–1807, 2014.
- [57] C. Szegedy et al., "Going deeper with convolutions," in *Proceedings of the IEEE Computer Society Conference on Computer Vision*.
- [58] T. Mikolov, A. Deoras, D. Povey, L. Burget, and J. ernocký, "Strategies for training large scale neural network language models," 2011 IEEE Work. Autom. Speech Recognit. Understanding, ASRU 2011, Proc., pp. 196–201, 2011.
- [59] G. Hinton et al., "Deep Neural Networks for Acoustic Modeling in Speech Recognition," *IEEE Signal Process. Mag.*, vol. 29, no. 6, pp. 82–97, 2012.
- [61] T. N. Sainath, A. Mohamed, B. Kingsbury, B. Ramabhadran, I. B. M. T. J. Watson, and Y. Heigley, "Deep Convolutional Neural Networks for LVCSR," *Icassp 2013*, pp. 8614–8618, 2013.
- [62] J. Ma, R. P. Sheridan, A. Liaw, G. E. Dahl, and V. Svetnik, "Deep neural nets as a method for quantitative structure-activity relationships," *J. Chem. Inf. Model.*, vol. 55, no. 2, pp. 263–274, 2015.
- [63] T. Ciodaro, D. Deva, J. M. De Seixas, and D. Damazio, "Online particle detection with neural networks based on topological calorimetry information," in *Journal of Physics Conference Series*, 2012, vol. 368, no. 1.
- [64] "Higgs Boson Machine Learning Challenge | Kaggle." [Online]. Available: <https://www.kaggle.com/c/higgs-boson>. [Accessed: 23-Jun-2018].
- [65] M. Helmstaedter, K. L. Briggman, S. C. Turaga, V. Jain, H. S. Seung, and W. Denk, "Connectomic reconstruction of the innerplexiform layer in the mouse retina," *Nature*, vol. 500, no. 7461, pp. 168–174, 2013.
- [66] M. K. K. Leung, H. Y. Xiong, L. J. Lee, and B. J. Frey, "Deep learning of the tissue-regulated splicing code," *Bioinformatics*, vol. 30, no. 12, pp. 121–129, 2014.

Bibliographie

- [67] H. Y. Xiong et al., “The human splicing code reveals new insights into the genetic determinants of disease,” *Science* (80-.), vol. 347, no. 6218, 2015.
- [68] R. Collobert, J. Weston, L. Bottou, M. Karlen, K. Kavukcuoglu, and P. Kuksa, “Natural Language Processing (Almost) from Scratch,” *J. Mach. Learn. Res.*, vol. 12, pp. 2493–2537, 2011.
- [69] A. Bordes, S. Chopra, and J. Weston, “Question Answering with Subgraph Embeddings,” pp. 615–620, 2014.
- [70] S. Jean, K. Cho, R. Memisevic, and Y. Bengio, “On Using Very Large Target Vocabulary for Neural Machine Translation,” pp. 1–10, 2014.
- [71] I. Sutskever, O. Vinyals, and Q. V Le, “Sequence to Sequence Learning with Neural Networks,” in *Proceedings of the 27th International Conference on Neural Information Processing Systems*, 2014, pp. 3104–3112.
- [72] R. Chenga Ivarayan and L. Deng, “Speech trajectory discrimination using the minimum classification error learning,” *IEEE Trans. Speech Audio Process.*, vol. 6, no. 6, pp. 505–515, 1998. *Speech Audio Process.*, vol. 6, no. 6, pp. 505–515, 1998.
- [73] M. Gibson and T. Hain, “Error Acoustic Model Estimation,” vol. 18, no. 6, pp. 1269–1279, 2010.
- [74] X. He, L. Deng, and W. Chou, “Discriminative learning in sequential pattern recognition: A unifying review for optimization-oriented speech recognition,” *IEEE Signal Processing Magazine*, vol. 25, no. 5, pp. 14–36, 2008.
- [75] H. Jiang and X. Li, “Parameter estimation of statistical models using convex optimization,” in *IEEE Signal Processing Magazine*.
- [76] B. Juang, W. Chou, and C. Lee, “Minimum Classification Error Rate Methods For Speech Recognition - Speech and Audio Processing, *IEEE Transactions on*,” *Speech Audio Process. IEEE Trans.*, vol. 5, no. 3, pp. 257–265, 1997. 2010, vol. 27, no. 3, pp. 115–127.
- [77] D. Povey and P. C. Woodland, “Minimum Phone Error and I-Smoothing for Improved Discriminative Training,” *Proc.ICASSP*, vol. 1, no. 1, pp. 105–108, 2002.
- [78] L. Xiao and L. Deng, “A geometric perspective of large-margin training of gaussian models,” *IEEE Signal Process. Mag.*, vol. 27, no. 6, pp. 118–123, 2010.
- [79] D. Yu, L. Deng, X. He, and A. Acero, “2007 IEEE International Conference on Acoustics, Speech and Signal Processing - ICASSP '07,” *2007 IEEE Int. Conf. Acoust. Speech Signal Process. - ICASSP '07*, vol. 4, p. IV-1137-IV-1140, 2007.
- [80] I. Heintz, E. Fosler-Lussier, and C. Brew, “Discriminative input stream combination for conditional random field phone recognition,” *IEEE Trans. Audio, Speech Lang. Process.*, vol. 17, no. 8, pp. 1533–1546, 2009.

Bibliographie

- [81] Y. Hifny and S. Renals, "Speech Recognition Using Augmented Conditional Random Fields," *IEEE Trans. Audio. Speech. Lang. Processing*, vol. 17, no. 2, 2009.
- [82] J. Peng, L. Bo, and J. Xu, "Conditional Neural Fields," *Adv. Neural Inf. Process. Syst.*, vol. 9, pp. 1–9, 2009.
- [83] D. Yang, P. Dixon, Y.-C. Pan, T. Oonishi, M. Nakamura, and S. Furui, "Combining a Two-step Conditional Rando.
- [84] D. Yu, S. Wang, Z. Karam, and L. Deng, "Language recognition using deep-structured conditional random fields," *Acoust. Speech Signal Process. (ICASSP), 2010 IEEE Int. Conf.*, pp. 5030–5033, 2010.
- [85] G. Zweig and P. Nguyen, "A segmental CRF approach to large vocabulary continuous speech recognition," *IEEE Work. Autom. Speech Recognit. Underst.*, pp. 152–157, 2009.
- [86] G. Heigold, H. Ney, P. Lehnen, T. Gass, and R. Schlüter, "Equivalence of Generative and Log-Linear Models," *IEEE Trans. Audio, Speech Lang. Process.*, vol. 19, no. 5, pp. 1138–1148, 2011.
- [87] D. Yu, S. Wang, and L. Deng, "Sequential labeling using deep-structured conditional random fields," *IEEE J. Sel. Top. Signal Process.*, vol. 4, no. 6, pp. 965–973, 2010.
- [88] D. Yu and L. Deng, "Deep-structured hidden conditional random fields for phonetic recognition.," *Interspeech*, vol. @, no. September, pp. 2986–2989, 2010.
- [89] Y. Bengio, P. Lamblin, D. Popovici, and H. Larochelle, "Greedy layer-wise training of deep networks," in *Adv. in Neural Inf. Proc. Syst.* 19, 2006, pp. 153–160.
- [90] L. Deng, M. Seltzer, D. Yu, A. Acero, A.-R. Mohamed, and G. Hinton, "Binary Coding of Speech Spectrograms Using a Deep Autoencoder," *Interspeech*, no. September, pp. 1692–1695, 2010.
- [91] G. E. Hinton and R. R. Salakhutdinov, "Reducing the dimensionality of data with neural networks," *Science (80-.)*, 2006.
- [92] Y. Bengio, "Learning Deep Architectures for AI," *Found. Trends® Mach. Learn.*, vol. 2, no. 1, pp. 1–127, 2009.
- [93] Y. LeCun, S. Chopra, M. A. Ranzato, and F. J. Huang, "Energy-based models in document recognition and computer vision," in *Proceedings of the International Conference on Document Analysis and Recognition, ICDAR, 2007*, vol. 1, pp. 337–341.
- [94] J. Ngiam, Z. Chen, P. W. Koh, and A. Y. Ng, "Learning Deep Energy Models," *Proc. 28th Int. Conf. Mach. Learn. (ICML 11)*, pp. 1105–1112, 2011.

Bibliographie

- [95] Y. Bengio, A. Courville, and P. Vincent, "Representation learning: a review and new perspectives.," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 35, no. 8, pp. 1798–1828, 2013.
- [96] Y. Bengio, R. De Mori, G. Flammia, and R. Kompe, "Global Optimization of a Neural Network - Hidden Markov Model Hybrid," *IEEE Trans. Neural Networks*, vol. 3, no. 2, pp. 252–259, 1992..
- [97] Y. Bengio et al., "A Neural Probabilistic Language Model," *J. Mach. Learn. Res.*, vol. 3, pp. 1137–1155, 2003
- [98] Y. Bengio et al., "A Neural Probabilistic Language Model," *J. Mach. Learn. Res.*, vol. 3, pp. 1137–1155, 2003.
- [99] Y. Bengio, P. Lamblin, D. Popovici, and H. Larochelle, "Greedy layer-wise training of deep networks," in *Adv. in Neural Inf. Proc. Syst.* 19, 2006, pp. 153–160.
- [100] Y. Bengio, A. Courville, and P. Vincent, "Representation learning: a review and new perspectives.," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 35, no. 8, pp. 1798–1828, 2013.
- [101] J. Bilmes, "Dynamic graphical models," *IEEE Signal Process. Mag.*, vol. 27, no. 6, pp. 29–42, 2010.
- [102] J. A. Bilmes and C. Bartels, "Graphical model architectures for speech recognition," *IEEE Signal Process. Mag.*, vol. 22, no. 5, pp. 89–100, 2005.
- [103] L. Andrade et al., "The epidemiology of major depressive episodes: Results from the International Consortium of Psychiatric Epidemiology (ICPE) Surveys," *Int. J. Methods Psychiatr. Res.*, vol. 12, no. 1, pp. 3–21, 2003
- [104] M. Hewitt and J. H. Rowland, "Mental health service use among adult cancer survivors: Analyses of the National Health Interview Survey," *J. Clin. Oncol.*, vol. 20, no. 23, pp. 4581–4590, 2002.
- [105] M. Nadeem, M. Horn, and G. Coppersmith, "Identifying Depression on Twitter," *CoRR*, pp. 1–9, 2016.
- [106] Steve Whittaker, Ellen Isaacs, and Vicki O'Day. 1997. Widening the net: workshop report on the theory and practice of physical and network communities. *SIGCHI BULLETIN* 29(1997), 27–30.
- [107] Munmun De Choudhury and Emre Kıcıman. 2017. The language of social support in social media and its effect on suicidal ideation risk. In *ICWSM*, Vol. 2017. NIH Public Access, 32.
- [108] Scott E Caplan and Jacob S Turner. 2007. Bringing theory to research on computer-mediated comforting communication. *Computers in human behavior* 23, 2(2007), 985–998.
- [109] Munmun De Choudhury and Sushovan De. 2014. Mental Health Discourse on reddit: Self-Disclosure, Social Support, and Anonymity. In *ICWSM*.

Bibliographie

- [110]Munmun De Choudhury and EmreKıcımın. 2017. Thelanguage of social support in social media and its effecton suicidal ideation risk. In ICWSM, Vol. 2017. NIHPublic Access, 32.
- [111]Christopher E Beaudoin and Chen-Chao Tao. 2007. Benefiting from social capital in online support groups:An empirical study of cancer patients. *CyberPsychology& Behavior* 10, (2007)587–590.
- [112]Barbara Becker and Gloria Mark. 2002. Socialconventions in computer mediated communication: Acomparison of three online shared virtual environments. In *The social life of avatars*. Springer, 19–39.
- [113]Gunther Eysenbach, John Powell, Marina Englesakis,Carlos Rizo, and Anita Stern. 2004. Health related virtual communities and electronic support groups: systematicreview of the effects of online peer to peer interactions.*Bmj* 328, 7449 (2004), 1166.
- [114] Gunter Eysenbach, John Powell, Marina Englesakis,CarlosRizo, and Anita Stern. 2004. Health related virtualcommunities and electronic support groups: systematicreview of the effects of online peer to peer interactions.*Bmj* 328, 7449 (2004), 1166.
- [115]Tom Postmes, Russell Spears, and Martin Lea. 2000. Theformation of group norms in computer-mediatedcommunication. *Human communication research* 26, 3(2000), 341–371.
- [116]Kathleen Ferrara. 1991. Accommodation in therapy.Contexts of accommodation: Developments in appliedsociolinguistics (1991), 187–222.
- [117] Amy L Gonzales, Jeffrey T Hancock, and James WPennebaker. 2010. Language style matching as apredictor of social dynamics in small groups.*Communication Research* 37, 1 (2010), 3–19.
- [118]Cristian Danescu-Niculescu-Mizil, Robert West, DanJurafsky, Jure Leskovec, and Christopher Potts. 2013. Nocountry for old members: User lifecycle and linguisticchange in online communities. In *WWW*. ACM, 307–318.
- [119]L. Andrade et al., “The epidemiology of major depressive episodes: Results from the International Consortium of Psychiatric Epidemiology (ICPE) Surveys,” *Int. J. Methods Psychiatr. Res.*, vol. 12, no. 1, pp. 3–21, 2003.
- [120]W.Akram1*, R.Kumar2“A Study on Positive and Negative Effects of Social Media on Society “Volume-5, Issue-10
- [121] M. Nadeem, M. Horn, and G. Coppersmith, “Identifying Depression on Twitter,” *CoRR*, pp. 1–9, 2016.
- [122] S. C. Guntuku, D. B. Yaden, M. L. Kern, L. H. Ungar, and J. C. Eichstaedt, “Detecting depression and mental illness on social media: an integrative review,” *Current Opinion in Behavioral Sciences*, vol. 18. pp. 43–49, 2017.

Bibliographie

- [123] Z. Jamil, D. Inkpen, P. Buddhitha, and K. White, “Monitoring Tweets for Depression to Detect At-risk Users,” in *CLPsych*, 2017, pp. 32–40.
- [124] G. Shen et al., “Depression detection via harvesting social media: A multimodal dictionary learning solution,” in *IJCAI International Joint Conference on Artificial Intelligence*, 2017, pp. 3838–3844.
- [125] G. Coppersmith, M. Dredze, and C. Harman, “Quantifying Mental Health Signals in Twitter,” *Proc. Work. Comput. Linguist. Clin. Psychol. From Linguist. Signal to Clin. Real.*, pp. 51–60, 2014.
- [126] B. Shickel, M. Heesacker, S. Benton, A. Ebadi, P. Nickerson, and P. Rashidi, “Self-Reflective Sentiment Analysis,” in *Proceedings of the Third Workshop on Computational Linguistics and Clinical Psychology*, 2016, pp. 23–32.
- [127] A. Wang and D. Singh, “Detecting Depression Through Tweets,” pp. 1–9, 2018.
- [128] H. Jo, S.-M. Kim, and J. Ryu, “What we really want to find by Sentiment Analysis: The Relationship between Computational Models and Psychological State,” 2017.
- [129] IRIT à e-risque Farah Benamara Farah Benamara et V'eronique Moriceau PP 1-12 2018.
- [130] A Neural Network Approach to Early Risk Detection of Depression and Anorexia on Social Media Text Yu-Tseng Wang¹, Hen-Hsen Huang¹, and Hsin-Hsi Chen pp 1-8.
- [131] T. Mikolov, I. Sutskever, K. Chen, G. Corrado, and J. Dean, “Distributed Representations of Words and Phrases and their Com- positionality,” 2013.
- [132] J.S. Hochreiter and J. Uergen Schmidhuber, “LONG SHORT-TERM MEMORY,” *Neural Comput.*, vol. 9, no. 8, pp. 1735–1780, 1997.
- [133] Y. Kim, “Convolutional Neural Networks for Sentence Classification,” 2014
- [134]] wildml, “Recurrent Neural Networks Tutorial, Part 1 – Introduction to RNNs WildML.” [Online]. Available: <http://www.wildml.com/2015/09/recurrent-neural-networks-tutorial-part-1-introduction-to-rnns/>. [Accessed: 27-Jun-2018].
- [135] S. Hochreiter and J. Uergen Schmidhuber, “LONG SHORT-TERM MEMORY,” *Neural Comput.*, vol. 9, no. 8, pp. 1735–1780, 1997.
- [136] C. Ola, “Understanding LSTM Networks -- colah’s blog,” 2015-08-27, 2015. [Online]. Available: <http://colah.github.io/posts/2015-08-Understanding-LSTMs/>. [Accessed: 27-Jun-2018]
- [137] C. Baziotis, N. Pelekis, and C. Doulkeridis, “DataStories at SemEval-2017 Task 4: Deep LSTM with Attention for Message-level and Topic-based Sentiment Analysis,” in *Proceedings of the 11th International Workshop on Semantic Evaluation (SemEval-2017)*, 2017, pp. 747–754.
- [138]] “NumPy — NumPy.” [Online]. Available: <http://www.numpy.org/>. [Accessed:
- [139]] “Natural Language Toolkit — NLTK 3.3 documentation.” [Online]. Available: <https://www.nltk.org/>. [Accessed: 27-Jun-2018].
- [140]] “Pandas — Pandas.” [Online]. Available: <http://www.pandas.org/>. [Accessed: 05-Jun-2019].
- [141] TensorFlow.” [Online]. Available: <https://www.tensorflow.org/>. [Accessed: 27-Jun-2018].

Bibliographie

[142]]“gensim: Topic modelling for humans.” [Online]. Available: <https://radimrehurek.com/gensim/>. [Accessed: 27-Jun-2018].

[143] “Python Data Analysis Library — pandas: Python Data Analysis Library.” [Online]. Available: <https://pandas.pydata.org/>. [Accessed: 27-Jun-2018].

[R. 144] Joshi, “Accuracy, Precision, Recall & F1 Score: Interpretation of Performance Measures - Exsilio Blog,” 2016-09-09, 2016. [Online]. Available: <http://blog.exsilio.com/all/accuracy-precision-recall-f1-score-interpretation-of-performance-measures/>. [Accessed: 27-Jun-2018].

[145]]“CSV — CSV.” [Online]. Available: <https://commons.apache.org/proper/commons-csv/>. [Accessed:

Bibliographie
