

Académie de Montpellier
Université Montpellier II
Sciences et Techniques du Languedoc

MÉMOIRE DE STAGE DE MASTER 2

effectué au Laboratoire d'Informatique de Robotique
et de Micro-électronique de Montpellier

Spécialité : **DECOL**

**Analyse d'émotions dans les forums de
santé**

Date de soutenance : **05/07/2013**

Par **Soumia MELZI**

Sous la direction de
Sandra BRINGAY et Pascal PONCELET

Remerciement

Avant tout, je remercie mes encadrants, Sandra Bringay et Pascal Poncelet, qui m'ont aidé, pour leurs précieux conseils, leur patience, leur disponibilité et leur encouragement et qui m'ont fait profiter de leur grande expérience.

Je remercie également, les membres du Jury pour avoir bien voulu juger mon travail et Pierre Pompidor qui a collaboré avec nous dans mon travail.

A ceux qui ont participé de près ou de loin pour la réalisation de ce travail, en particulier l'annotateur professionnel Florence Galtier et tous les annotateurs non professionnels de santé et tout le personnel de l'équipe TATOO, un grand merci !

En fin, je dédie ce modeste travail à ma puce Sirine et à mon cher mari.

Abstract

Internet health forums are online spaces of discussion where patients express their feelings about their diseases, treatments, medicines, etc. with no restriction due to anonymity offered by these websites. This makes health forums a real source of information that can be very useful for professionals in order to better understand patient's feelings, problems and behaviours. This information can be mined using text mining techniques. Emotion analysis is a text mining task that deals with extracting emotions (*e.g.* joy, anger, surprise) from textual data. In this work, we focus on (1) emotion analysis on health forum's discussions and (2) detection of the medical object (disease, medicine) of extracted emotions. Our approach to emotion analysis in health forums was as follow : First, we have collected a corpus of messages on a health forum. Then, we have annotated this corpus by two strategies (automatically and manually). The annotation process revealed to us the difficulty of annotation task because of the low inter-annotator agreement between humans. Finally, we have deployed two different methods in order to extract emotions and identify relationships between emotions and medical objects.

Keywords : sentiment analysis, emotion analysis, annotation

Résumé

Les forums de discussion dans le domaine de santé sont des espaces d'échanges sur Internet où les patients partagent leurs expériences par rapport à leurs maladies, traitements, médecins, etc. Sous couvert d'anonymat, les patients expliquent très librement leurs expériences personnelles. Les forums de santé sont donc une source d'information qui peut être très utile pour les professionnels de santé pour mieux comprendre les problèmes, les comportements et les sentiments des patients. Ces informations peuvent être exploitées en utilisant des techniques de fouille de textes. L'analyse d'émotion est une tâche de fouille de texte qui s'intéresse à l'extraction des émotions (*e.g.* joie, colère, surprise, etc.) exprimées dans les textes. Dans ce travail, nous nous sommes intéressés à (1) l'analyse d'émotions émises par des patients dans les forums de santé et (2) à la détection de l'objet médical (maladie, médicament, etc.) sur lequel porte l'émotion. Notre processus d'analyse d'émotions dans les forums de santé est le suivant : premièrement, nous avons collecté un corpus de données sur un forum de santé. Deuxièmement, nous avons suivi deux stratégies (automatique et manuelle) pour l'annotation du corpus. Cette annotation nous a permis de confirmer la difficulté de la tâche d'analyse d'émotions car même les humains n'étaient pas d'accord sur la présence d'une émotion dans un message. Finalement, nous avons expérimenté différentes méthodes pour extraire les émotions exprimées dans les messages et pour identifier le lien entre les émotions et les objets médicaux.

Mots clés : analyse de sentiments, analyse d'émotion, annotation

Table des matières

1	Introduction	6
2	Analyse de sentiments	7
2.1	Tâches associées	7
2.2	Architecture générale d'un système d'analyse de sentiments	9
3	Etat de l'art des principales méthodes	10
3.1	Ressources utilisées	10
3.1.1	Lexiques de sentiments	10
3.1.2	Bases de connaissances	11
3.2	Méthodes d'analyse de sentiments	12
3.2.1	Méthodes symboliques	13
3.2.2	Méthodes statistiques	14
3.2.3	Méthodes hybrides	16
3.2.4	Discussion	16
4	Analyse d'émotions dans les forums de santé	18
4.1	Contexte et problématique	18
4.2	Verrous	18
4.3	Méthodologie	19
4.3.1	Phase 1 : Annotation automatique	19
4.3.2	Phase 2 : Annotation manuelle	21
4.3.3	Phase 3 : Classification automatique	22
4.3.4	Phase 4 : Identification des liens entre traces d'émotions et objets médicaux.	25
4.4	Constitution des corpus	25
5	Résultats expérimentaux	27
5.1	Evaluation du lexique	27
5.1.1	Répartition des mots du lexique sur les six catégories émo- tionnelles	27
5.1.2	Corrélation entre les six catégories émotionnelles	28
5.1.3	Adéquation du lexique d'émotion au corpus d'étude	28
5.1.4	Connotation émotionnelle des objets médicaux	29
5.2	Accord entre annotateurs	29
5.3	Statistiques sur les deux corpus CAA et CAM	31
5.3.1	Répartition des phrases annotées selon les catégories émo- tionnelles	32
5.3.2	Répartition des phrases selon le nombre d'émotions qui y sont présentes	32
5.3.3	Répartition des phrases dans le corpus final selon les six catégories émotionnelles après la ré-annotation	33
5.3.4	Nombre d'attributs pour chaque corpus	33
5.4	Évaluation de la classification automatique	33

TABLE DES MATIÈRES

5.4.1	Corpus CAA	34
5.4.2	Corpus CAM	36
5.4.3	Comparaisons du CAA vs. CAM	36
6	Conclusion et perspectives	37
A	Résultats complets	41
A.1	Corpus CAA	41
A.2	Corpus CAM	42

1 Introduction

Les forums de discussion dans le domaine de santé sont des espaces d'échanges sur Internet où les patients partagent leurs expériences en s'exprimant en toute liberté. Sous couvert d'anonymat, les patients relatent souvent très librement leurs expériences personnelles. Les forums de santé offrent aujourd'hui de nouvelles opportunités pour les professionnels de santé qui ont accès à des échanges entre patients, entre patients et professionnels et même entre professionnels. En effet, mieux connaître les problèmes rencontrés par les patients peut être très important pour les professionnels de santé. Le patient n'est pas forcément prompt à expliquer ses problèmes à son propre médecin alors qu'il le fera plus facilement à des inconnus via les forums. L'analyse de ces informations peut s'avérer très utile. Par exemple, une étude récente¹ a montré, en étudiant les forums traitant du Mediator, que les internautes avaient associés, dès 2003, ce médicament à la notion de risque. Une telle information aurait sans doute pu être utile pour les médecins pour anticiper le scandale. Analyser les forums de santé est donc un enjeu important pour les professionnels de santé car ils représentent une base volumineuse, riche et variée de connaissances et des perceptions qu'ont les patients de leur maladie et des soins qui leur sont éventuellement prodigués. Dans ce contexte éminemment subjectif, la caractérisation et la compréhension de ces perceptions est difficile mais néanmoins particulièrement intéressante dans la perspective de compléter et d'améliorer les programmes de santé publique.

Notre travail s'inscrit dans cette voie. Nous nous intéressons à une tâche particulière de fouille de texte qui porte sur l'analyse de sentiments et plus précisément à l'analyse d'émotions (joie, peur, surprise, *etc.*) dans les forums de santé. Notre objectif est de reconnaître les émotions émises par les patients par rapport à différentes maladies, médicaments ou traitements. Lors de l'étude bibliographique menée au début du stage, nous avons identifié un certain nombre de méthodes permettant d'associer automatiquement des émotions à des textes. Malgré les bonnes performances décrites par les auteurs de ces méthodes, nous avons remis en question la tâche d'extraction d'émotions en elle-même. Pour cela, nous avons réalisé une étude préalable sur un panel important d'annotateurs qui nous a permis de montrer qu'il était difficile pour des humains de s'accorder sur le type d'émotion(s) présente(s) dans un message de forum. Nous avons donc proposé une méthode qui permet de détecter la présence d'une émotion, de la caractériser (positive ou négative) et de proposer les étiquettes des émotions les plus probables. Pour mettre en place cette méthode, nous avons utilisé des techniques de classification classiques de la littérature. La principale difficulté a été d'identifier les meilleurs descripteurs de l'émotion et de construire un corpus pertinent afin de valider cette méthode. Nous avons également cherché à identifier la cible de l'émotion (*e.g.* de quoi une personne a peur). Ce travail en étant à ses prémices, nous le décrirons en tant que perspectives.

1. www.conversationnel.fr

Le reste de ce document est organisé comme suit : la section 2 propose une première catégorisation de l'analyse des sentiments. En effet, comme nous allons le constater, l'analyse de sentiments recouvre dans la littérature de nombreuses notions. Notre objectif ici est de clarifier les principales tâches mais également de présenter une architecture générale d'analyse de sentiments. Dans la section 3, nous présentons les principales ressources utilisées et les méthodes récentes associées à cette problématique. La section 4 décrit la démarche que nous proposons. La section 5 décrit les résultats obtenus. Enfin la dernière section conclut notre travail et résume les principales perspectives.

2 Analyse de sentiments

Depuis le début des années 2000, le domaine de l'analyse de sentiments a connu un intérêt croissant (*e.g.* [1, 21, 13]). Beaucoup de travaux de recherche se sont intéressés aux différentes problématiques de l'analyse de sentiments, impliquant des chercheurs de différentes communautés : psychologie, sciences politiques et sociales, linguistique computationnelle, TAL², fouille de données, *etc.* Étudiée par tant de communautés, l'analyse de sentiments est donc sujette à de nombreuses interprétations. Dans ce rapport, nous nous focalisons sur les méthodes développées dans la communauté fouille de données en ayant conscience qu'il pourrait être très intéressant de les croiser avec celles issues des autres domaines de recherche précédemment cités. Aussi, dans cette section, nous présentons les différentes tâches associées à l'analyse de sentiments. Nous décrivons également les fonctionnalités principales d'un système d'extraction de sentiments.

2.1 Tâches associées

L'analyse de sentiments, également souvent appelée fouille d'opinions (*opinion mining*), est une tâche qui s'intéresse à l'extraction de sentiments exprimés explicitement ou implicitement dans des textes. Liu dans [13] propose une hiérarchie de l'analyse des sentiments comme un domaine d'étude qui analyse les opinions, les sentiments, les attitudes et les émotions des gens par rapport à des produits, des services, des organisations, des personnes, des événements, *etc.* Plus généralement, l'analyse de sentiments englobe les tâches suivantes (cf. Figure 1) :

- L'analyse de subjectivité [26] se concentre sur la détection des expressions ou des mots subjectifs, c'est à dire qui expriment un sentiment, dans les textes ;
- L'analyse de polarité [1] s'intéresse à la classification des mots et des textes selon qu'ils soient de polarité positive (sentiment positif), négative ou neutre ;

2. Traitement Automatique du Langage

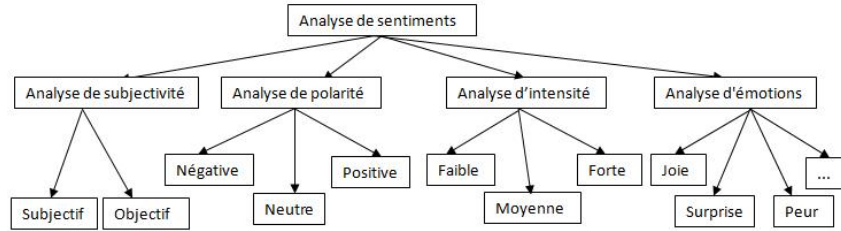


FIGURE 1 – Les principales tâches associées à l'analyse de sentiments

- L'analyse d'intensité [18] se focalise sur les différents niveaux d'intensité de la polarité (positif, très positif, négatif, très négatif, *etc.*). Ces approches offrent une granularité plus précise sur les opinions émises ;
- L'analyse d'émotions³ [21, 14, 2] s'intéresse à la catégorie émotionnelle du texte. Selon Izard [9], les émotions appartiennent à dix catégories : colère, dégoût, peur, culpabilité, intérêt, joie, tristesse, détresse, honte et surprise. Un ensemble plus réduit de ces catégories a été défini par Ekman [3] : colère, dégoût, peur, joie, tristesse, surprise.

De manière à illustrer ces différentes tâches, considérons l'exemple suivant :

Hier soir, je suis allé au cinéma avec mes copains. Le film était vraiment bien, j'ai adoré l'acteur principal qui était excellent, le scénario et le décor. Cependant, au début, j'avais un peu peur mais ça a rapidement disparu et on s'est bien amusé :)

FIGURE 2 – Un exemple de texte exprimant des sentiments

La première phrase "*Hier soir, je suis allé au cinéma avec mes copains*" est une phrase objective. Elle n'exprime aucune opinion ni aucun sentiment particulier. Généralement l'extraction de telles phrases est utile pour minimiser l'espace de recherche dans des textes afin de se concentrer sur les phrases qui expriment des sentiments. L'analyse de polarité s'intéresse à l'extraction de phrases du type : "*le film était vraiment bien*" qui reflète l'opinion exprimée par des personnes sur un film. L'analyse de l'intensité permet de donner plus de gradualité dans l'expression d'opinion et ainsi différencier "*le film était vraiment bien*" (polarité positive) de "*l'acteur principal qui était excellent*" (polarité très positive). Enfin, l'analyse d'émotions s'intéresse à des phrases exprimant l'émotion de peur comme "*j'avais un peu peur*" ou bien une émotion de joie "*on s'est bien amusé*". Cet exemple illustre qu'il est difficile de séparer les différentes tâches. En effet, alors que l'opinion peut être caractérisée en positive, neutre

3. Les premières émotions : "admiration, amour, haine, désir, joie et tristesse" ont été identifiées par le philosophe René Descartes dans son traité Les Passions de l'âme. Toutes les autres émotions définies par la suite sont un sous ensemble ou une spécification des ces émotions.

ou négative, quelles sont les interactions avec les émotions? Quel est l'inverse d'une émotion : l'inverse d'une opinion positive peut être neutre ou négative mais quid de l'inverse d'une émotion : "*je n'avais pas peur*" veut-il dire que la personne est joyeuse? surprise? qu'elle exprime une opinion négative pour l'émotion peur? Toutes ces questions resteront sous-jacentes dans ce rapport et nous y reviendrons plus particulièrement dans notre proposition.

2.2 Architecture générale d'un système d'analyse de sentiments

Cette section a pour objectif de présenter une approche générale d'analyse de sentiments. Elle vise à montrer les différentes étapes utilisées pour remplir les différentes tâches précédemment décrites. Elle ne se veut pas exhaustive mais permet d'offrir au lecteur les différents enchaînements qui sont réalisés par les approches existantes. Le processus d'analyse de sentiments se décompose généralement en plusieurs étapes (cf. Figure 3) : la première étape consiste à collecter les différentes données textuelles (titres de presse, réseaux sociaux, forums de discussion, *etc.*) généralement du Web. Dans la deuxième étape, les données collectées sont prétraitées pour pouvoir être analysés. Différents prétraitements peuvent être effectués :

- La lemmatisation qui consiste à transformer les différentes inflexions d'un mot en leur forme canonique (lemme). Ainsi, les verbes sont ramenés à l'infinitif et les noms au singulier masculin en supprimant les éventuels suffixes. Par exemple, les mots *bon*, *bonne*, *bonnes* sont transformés en *bon*;
- La suppression des mots vides ou des mots outils qui sont des mots communs du langage qui n'ont pas, ou peu, de pouvoir discriminant à l'égard du problème étudié. En français, des mots vides évidents pourraient être *le*, *la*, *de*, *du*, *ce*, *ça*, *etc.*

Enfin, pour que les textes pré-traités soient exploitables par les méthodes d'analyse de sentiments, une étape de représentation est nécessaire. Les textes sont généralement représentés sous la forme : (1) de sac de mots, i.e. par un vecteur de termes, (2) de n-grammes, i.e. des séquences de N mots adjacents extraits à partir d'une séquence de mots. Par exemple, les 2-grammes extraites à partir du texte "*coupe du monde*" sont "*coupe du*" et "*du monde*", ou (3) par des patrons (par exemple le patron < sujet, verbe, complément d'objet >), *etc.* Une fois les données prêtes à être analysées, plusieurs méthodes d'analyse de sentiments (statistiques, symboliques, hybrides) peuvent alors être appliquées pour extraire les sentiments. Ces méthodes utilisent des ressources externes, des corpus annotés ou font appel à des moteurs de recherche Web pour l'extraction de sentiments. Une intervention humaine est parfois nécessaire afin de définir des règles linguistiques, valider les résultats ou annoter les corpus. Nous reviendrons, plus en détail, sur ces méthodes dans la section suivante.

L'évaluation des résultats de l'analyse de sentiments est généralement faite par des mesures d'évaluation traditionnelles comme, par exemple, la précision, le rappel et le Fscore. En effet, l'analyse de sentiments peut être vue comme une tâche de classification où il s'agit d'attribuer les textes à des classes (joie, peur, surprise, *etc.*). Les mesures d'évaluation sont calculées pour chacune de ces classes c de la manière suivante :

$$\text{Précision } (c) = \frac{\text{nombre de textes correctement attribués à la classe } c}{\text{nombre de textes attribués à la classe } c} \quad (1)$$

$$\text{Rappel } (c) = \frac{\text{nombre de textes correctement attribués à la classe } c}{\text{nombre de textes de classe } c} \quad (2)$$

Fscore est une mesure qui combine les mesures de rappel et de précision, il s'agit de la moyenne harmonique des deux.

$$\text{Fscore } (c) = \frac{2 \cdot \text{précision } (c) \cdot \text{rappel } (c)}{\text{précision } (c) + \text{rappel } (c)} \quad (3)$$

L'évaluation finale se fait par la moyenne de ces mesures calculées sur chacune des classes séparément.

3 Etat de l'art des principales méthodes

Dans cette section, nous présentons les principales ressources habituellement utilisées pour l'analyse de sentiments ainsi qu'un état de l'art des différentes méthodes.

3.1 Ressources utilisées

Dans cette section, nous présentons les ressources linguistiques (i.e. les lexiques définissant des mots de sentiments et les bases de connaissances structurées) généralement utilisées dans l'analyse de sentiments.

3.1.1 Lexiques de sentiments

Des lexiques de mots, publiquement disponibles, ont été développés dans l'objectif d'analyse de sentiments. En anglais, le lexique de subjectivité MPQA [28] est composé de 8221 mots où chaque mot est étiqueté d'une polarité (positive, négative, neutre). Liu [13] propose un autre lexique composé de 6800 mots étiquetés de leurs polarités (positive, négative). Mohammad et al. [17] ont construit un lexique de 14.200 mots annotés par leurs polarités ainsi que leurs associations à des émotions (joie, tristesse, colère, peur, surprise, anticipation, confiance, dégoût).

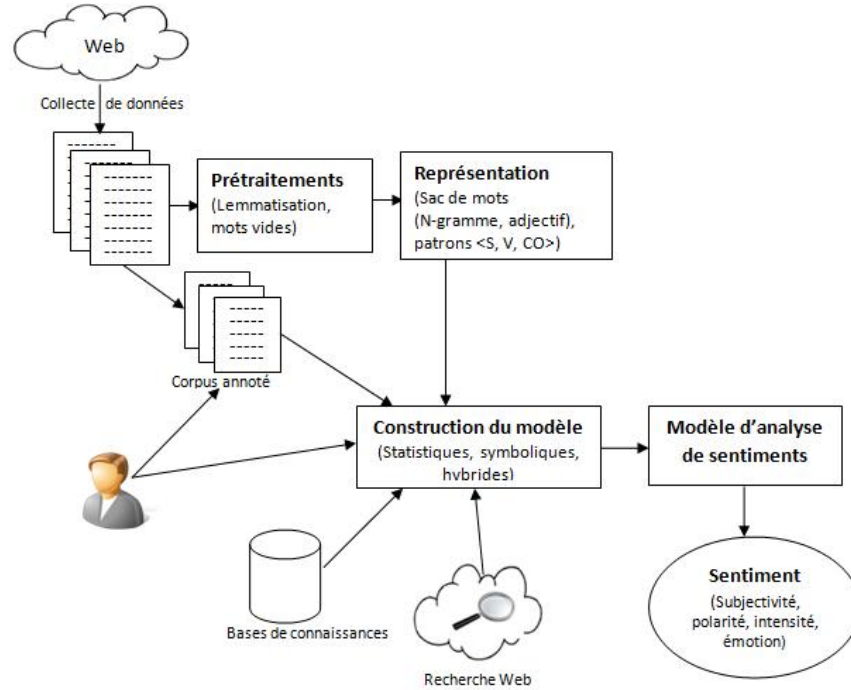


FIGURE 3 – Architecture générale d'un système d'analyse de sentiments

3.1.2 Bases de connaissances

WordNet [16] : WordNet est une base de données lexicale contenant des mots de la langue anglaise. Elle est organisée en synsets (ensembles de synonymes) où chaque synset contient un ensemble de mots sémantiquement proches, leurs catégories grammaticales (nom, verbe, adverbe, *etc.*). Les synsets sont reliés entre eux par des liens sémantiques (synonymie, antonymie, hyponymie, *etc.*) qui définit une structure de graphe de synsets. WordNet contient 147.249 mots et 207.016 synsets. Cette base est utilisée généralement pour l'analyse des relations de synonymie et d'antonymie entre les mots de sentiments, en considérant par exemple que deux mots synonymes ont la même polarité (*e.g.* [10]).

SentiWordNet [4] : SentiWordNet est une extension de la base de données WordNet où chaque synset s de WordNet est étiqueté par trois valeurs numériques $Obj(s)$, $Pos(s)$ et $Neg(s)$, qui représentent respectivement des scores d'objectivité, de positivité et de négativité. Ces valeurs sont comprises entre 0 et 1 et leur somme est égale à 1 pour chaque synset. Outre la prise en compte de la polysémie, SentiWordNet est très utile dans l'analyse de sentiments pour sa large couverture. En effet, tous les synsets de WordNet sont étiquetés en fonction de chacune des trois polarités (objectif, positif, négatif) et en raison de

leur évaluation fine (les étiquettes sont quantifiées par des valeurs numériques).

Par exemple, le mot estimable peut avoir deux sens différents :

- le sens " *may be computed or estimated*", dans ce cas , l'adjectif *estimable* a un score de 1 pour *Obj* (les scores *Pos* et *Neg* sont nuls) ;
- le sens " *deserving of respect or high regard*", dans ce cas le mot a un score de 0,75 pour *Pos*, un score de 0 pour *Neg*, et un score de 0,25 pour *Obj*.

WordNet-Affect [22] : Il s'agit d'une autre extension de la base de données WordNet pour représenter des concepts affectifs. En particulier, une ou plusieurs étiquettes affectives (appelés *a-labels*, *e.g.* émotion, attitude, sensation, comportement) sont associées à un certain nombre de synsets de WordNet. WordNet-Affect contient 1.903 termes dont 539 noms, 517 adjectifs, 238 verbes et 15 adverbes. Par exemple, le verbe *fear* et le nom *anger* sont étiquetés par "émotion" et le verbe *feel* est étiqueté par "sensation".

Thésaurus de Roget [15] : Le thésaurus de Roget est composé de six classes dont une classe contient des mots relatifs au sentiment. Chaque classe est composée de plusieurs divisions qui possèdent elles-mêmes plusieurs sections, ce qui donne une structure d'arbre qui couvre plus d'un millier de branches des mots sémantiquement liés.

Dictionnaire DAL [25] : En 1989, Whissell a introduit le dictionnaire DAL (Dictionary of Affect in Language) dans lequel les mots sont étiquetés avec trois scores : évaluation, activation et perception⁴. Le score d'évaluation varie de 1 (mauvais) à 3 (très agréable). Le score d'activité varie de 1 (passif) à 3 (actif). Le score de perception varie de 1 (difficulté d'imaginer mentalement le mot) à 3 (facilité à former une image mentale du mot). Par exemple, le mot *abandonner*, généralement considéré comme négatif, a une valeur de -0,86 pour le score d'évaluation, et -0.67 pour le score d'activité indiquant le sens passif et enfin le score de perception a une valeur de 1 pour indiquer la facilité de former une image mentale de ce mot.

3.2 Méthodes d'analyse de sentiments

Les méthodes d'analyse de sentiments sont généralement classés en deux catégories (cf. Figure 4) : méthodes symboliques et méthodes statistiques. Les méthodes symboliques utilisent traditionnellement des règles écrites manuellement et un lexique de sentiments alors que les méthodes statistiques construisent un modèle statistique à partir d'un corpus de textes et l'utilisent pour la détection des sentiments dans d'autres textes. Une troisième catégorie peut être considérée, il s'agit des méthodes hybrides (combinaison des deux approches statistiques et symboliques).

4. Le terme "perception" est utilisé ici comme traduction du terme anglais "imagery" employé dans l'article d'origine.

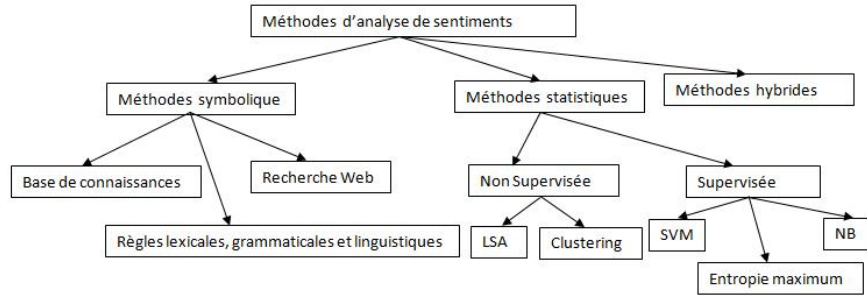


FIGURE 4 – Principales méthodes d'analyse de sentiments

3.2.1 Méthodes symboliques

Les méthodes symboliques, que l'on peut également qualifier de linguistiques, utilisent généralement un lexique de sentiments et un ensemble de règles fines définies manuellement pour chaque domaine. L'analyse de sentiments se fait au niveau du mot, de la phrase ou encore du texte.

Mulder et al. [18] proposent une approche lexicale et grammaticale pour analyser les sentiments dans les textes. Les auteurs utilisent l'expression "valeur affective" comme un moyen d'exprimer le sens affectif (polarité et intensité) d'un mot et la représentent par une valeur réelle entre -1 et 1 (-1 pour la polarité négative et 1 pour la polarité positive). Leur approche est basée sur cinq mécanismes dont les plus importants sont : l'intensification, la négation et la quantification.

Intensification : la valeur affective de certains adjectifs et verbes peut être intensifiée ou diminuée par d'autres mots, préfixes et suffixes. Par exemple l'expression "*very good*" montre que l'adverbe "*very*" intensifie la valeur affective de l'adjectif "*good*" (a priori positive) en la rendant plus positive.

Négation : la négation inverse la valeur affective d'une expression, elle est réalisée soit par le mot "*not*", soit par le préfixe "*un*". La négation d'un verbe ne change pas seulement sa valeur affective mais aussi la valeur affective de son complément d'objet direct.

Quantification : le mécanisme de quantification agit sur la cible du sentiment. Cette dernière peut représenter un ensemble d'entités. Un quantificateur sélectionne un sous-ensemble de l'ensemble complet et attribue aussi une valeur affective à ce sous-ensemble. Par exemple, "*skilled artists are scarce*" montre que la valeur affective positive de l'adjectif "*skilled*" est utilisée négativement à cause du sous-ensemble restreint des artistes désignés par l'adjectif "*scarce*".

Cette méthode obtient un rappel de 76.4% en utilisant la grammaire et les cinq mécanismes contre un rappel de 71,4% pour la méthode référence (baseline). Malgré cette amélioration, cette méthode reste difficile à mettre en œuvre à cause de la complexité des mécanismes et du travail manuel important.

Lu et al. [14] construisent un système qui utilise un étiqueteur de rôles sémantique et un moteur de recherche Web afin de détecter les émotions dans le texte. La démarche générale se décompose en trois étapes :

- Appliquer l'étiqueteur de rôles sémantique pour trouver le sujet, le verbe et le complément de la phrase ;
- Faire une recherche sur Google⁵ pour trouver les adjectifs qui décrivent le sujet et le complément de la phrase ;
- Combiner l'adjectif lié au sujet avec l'adjectif lié au complément en utilisant le verbe pour trouver l'émotion exprimée dans la phrase. La liste des règles émotionnelles pour les différentes combinaisons de certains adjectifs avec un verbe est créée manuellement.

Chaumartin [2] propose un système basé sur des règles linguistiques pour l'étiquetage d'émotions dans les titres de presse en utilisant les ressources lexicales suivantes : WordNet, SentiWordNet et WordNet-Affect. Le processus d'étiquetage se déroule en deux étapes :

- Détecter la valeur de l'émotion pour chaque mot en utilisant WordNet-Affect et SentiWordNet. Les émotions trouvées sont ensuite utilisées pour déduire la polarité du titre, par exemple, l'émotion "joie" représente une polarité positive ;
- Trouver le sujet principal du titre en utilisant l'analyseur syntaxique de Stanford [11] en considérant que le sujet est important pour la détection d'émotions. Les autres mots du titre sont évalués individuellement pour chaque émotion en utilisant SentiWordNet et WordNet-Affect.

Ce système basé sur des règles détecte les six émotions dans les titres avec une précision moyenne de 89,43%. Cependant, le rappel est faible (entre 0% et 22 %). L'émotion "dégout" donne de très mauvais résultats, 0 pour le rappel et 0 pour la précision.

Les méthodes symboliques utilisent des règles d'analyse de sentiment définies manuellement offrant ainsi une bonne précision. Par contre, ces règles sont définies manuellement ce qui peut être coûteux sachant qu'elles changent d'un domaine à un autre.

3.2.2 Méthodes statistiques

Les méthodes statistiques procèdent par la construction d'un modèle statistique à partir d'un corpus d'apprentissage. Ce modèle est utilisé dans un deuxième temps pour la classification de nouveaux textes.

Strapparava et al. [21] essayent de classer automatiquement les titres de presse selon l'émotion la plus appropriée parmi six émotions prédéfinies (colère, dégoût, peur, joie, tristesse et surprise). Pour cela, ils proposent deux méthodes statistiques : une méthode qui utilise des bases de connaissances et une méthode qui

5. <http://www.google.com>

utilise un corpus de textes annotés. La première méthode utilise la technique LSA (Analyse Sémantique Latente) qui est une technique de TAL qui permet de réduire l'espace de représentation en passant de l'espace de mots à l'espace de concepts (ensemble de mots qui apparaissent ensemble). L'émotion est représentée de trois façons différentes dans l'espace LSA : par le mot spécifique (*e.g.* "joie"), par tous ses synonymes WordNet (*e.g.* "bonheur", "contentement", *etc.*) ou encore par tous les mots étiquetés de l'émotion "joie" dans WordNet-Affect. Dans la deuxième méthode, un modèle Naïve Bayes (NB) est appris à partir d'un corpus de blogs annotés. Les résultats montrent que le système LSA avec la troisième représentation de l'émotion a le meilleur rappel. NB donne le meilleur résultat pour l'émotion "joie" sachant que celle-ci est la plus représentée dans le corpus d'apprentissage. Cela confirme l'inconvénient des méthodes d'apprentissage supervisé qui nécessitent beaucoup d'exemples d'entraînement pour construire un bon classifieur. Il est à noter que toutes les méthodes implémentées ont donné des résultats très mauvais pour l'émotion "dégout".

Hatzivassiloglou et McKeown [8] proposent une approche basée sur l'extraction des adjectifs liés par des conjonctions (*but, or, and*) à partir de corpus pour l'analyse de polarité. L'approche générale se décompose en quatre étapes :

- Extraction des adjectifs reliés par des conjonctions ;
- Utilisation des informations de conjonctions et un modèle de régression log-linéaire afin de déterminer si deux adjectifs en conjonction ont des polarités identiques ou différentes. Le résultat est un graphe dans lequel les nœuds sont les adjectifs et les arcs indiquent si la polarité est similaire ou différente entre deux adjectifs ;
- Utilisation d'un algorithme de clustering pour séparer les nœuds du graphe en deux sous-ensembles de polarités différentes (positifs et négatifs) ;
- Le sous-ensemble de la plus grande taille est étiqueté comme positif (selon les auteurs, les adjectifs de polarité positive ont tendance à être plus fréquents).

Le corpus utilisé dans les expérimentations est issu de *Wall Street Journal* et comporte 1336 adjectifs. Leur approche obtient un rappel de 78.08%.

L'avantage des méthodes statistiques est qu'elles construisent des modèles par apprentissage sans recours à des connaissances extérieures. Par contre, elles utilisent un corpus étiqueté qui n'est pas toujours disponible et beaucoup de temps est nécessaire pour étiqueter un corpus de taille importante. Il est également difficile de construire un ensemble pertinent d'attributs (mots, n-grammes, adjectifs, *etc.*).

3.2.3 Méthodes hybrides

Kozareva et al. [12] proposent une approche basée sur l'utilisation de trois moteurs de recherche Web (MyWay⁶, AlltheWeb⁷ et Yahoo⁸) pour déterminer la catégorie de l'émotion dans les titres de presse. Leur hypothèse est qu'un ensemble de mots qui apparaissent ensemble dans plusieurs documents avec une émotion donnée ont une probabilité forte d'exprimer la même émotion. Le Score de l'émotion est obtenu en utilisant la mesure de l'information mutuelle (MI) qui mesure la corrélation entre cette émotion et les mots qui composent le titre.

$$MI(mot_1, mot_2) = \frac{freq(mot_1 \text{ ET } mot_2)}{freq(mot_1) \cdot freq(mot_2)} \quad (4)$$

Où $freq(mot)$ est le nombre de résultats retournés par le moteur de recherche en utilisant la requête " mot ". Ensuite, le score final MI est pondéré par la moyenne de tous les scores MI entre chaque mot du titre et une émotion. Leur méthode détecte de manière significative l'émotion peur, et colère (la précision 40% et le rappel 26%). Le rappel et la précision sont nuls pour l'émotion dégout.

Harb et al. [7] proposent une méthode qui se décompose en trois étapes :

- La première étape consiste à collecter automatiquement des documents à partir du Web qui expriment des opinions sur un domaine donné en utilisant deux listes de mots positifs et négatifs suivantes [24] :
 $P = \{ good, nice, excellent, positive, fortunate, correct, superior \}$
 $N = \{ bad, nasty, poor, negative, unfortunate, wrong, inferior \}$
- Dans la deuxième étape, ils utilisent les corpus collectés afin d'enrichir automatiquement les deux ensembles P et N . L'enrichissement est basé sur le calcul de la corrélation entre les adjectifs du corpus et les adjectifs des listes P et N ;
- Le but de la dernière étape est de classer les nouveaux documents en utilisant les ensembles d'adjectifs construits dans l'étape précédente. Pour chaque document, ils calculent le nombre d'adjectifs positifs et négatifs présents dans le texte et dans les deux listes enrichies précédemment en même temps. Le document est affecté à la classe majoritaire.

La méthode proposée obtient une performance de 71.73% sur les textes positifs et 62.20% sur les textes négatifs en termes de Fscore.

3.2.4 Discussion

La figure 1 représente un tableau récapitulatif des différentes méthodes d'analyse de sentiments présentées ci-dessus. Chaque méthode est décrite par son principe de fonctionnement, la tâche associée (subjectivité, polarité, intensité et émotion), les ressources et les corpus utilisés.

6. www.myway.com
7. www.alltheweb.com
8. www.yahoo.com

3 ETAT DE L'ART DES PRINCIPALES MÉTHODES

<i>Réf.</i>	<i>Principes</i>	<i>Tâches</i>	<i>Ressources</i>	<i>Corpus</i>
Méthodes symboliques				
[18]	Règles lexicales et grammaticales	polarité		42 phrases
[14]	Etiqueteur des rôles sémantique et recherche Web	émotion	recherche Web (Google)	phrases
[2]	Règles linguistiques et bases de connaissances	émotion, polarité, intensité	WordNet, SentiWordNet, WordNet-Affect	Titres de presse
Méthodes statistiques				
[21]	LSA et Naïve Bayes	émotion	WordNet, WordNet-Affect	Titres de presse
[8]	clustering	polarité		Titres de presse
Méthodes hybrides				
[12]	Recherche Web, mesure de corrélation	émotion	recherche Web (MyWay, AlltheWeb et Yahoo)	Titres de presse
[7]	Recherche Web, mesure de corrélation	polarité	Web	corpus sur un domaine donné

TABLE 1 – Tableau récapitulatif des principales méthodes d’analyse de sentiments

Comme cela a été souligné précédemment (section 3.2.1), les méthodes symboliques utilisent des règles définies manuellement pour chaque domaine d’étude, ce qui nécessite beaucoup de travail manuel et l’intervention des experts. En revanche, ces méthodes ne nécessitent pas de corpus annotés pour faire l’apprentissage contrairement aux méthodes statistiques qui, elles, utilisent des corpus d’apprentissage annotés manuellement (ce qui peut être fastidieux). Les méthodes hybrides tentent à leur tour de tirer profit à la fois des méthodes symboliques et des méthodes statistiques en faisant un compromis entre elles.

Les résultats donnés par les méthodes symboliques sont généralement caractérisés par une bonne précision et un faible rappel. Par contre, les méthodes statistiques sont connues pour leur bon rappel et faible précision. Les méthodes d’analyse de sentiments sont différentes par le fonctionnement, les ressources utilisées et les résultats. Le choix d’une méthode adaptée à un problème donné est donc fait en fonction des données, des ressources à disposition et des critères que l’on souhaite optimiser (par exemple, rappel, précision).

Mise en œuvre

4 Analyse d'émotions dans les forums de santé

4.1 Contexte et problématique

Parmi les forums de discussion qui sont particulièrement propices à l'apparition de textes subjectifs on trouve les forums de santé. Doctissimo⁹, Same-story¹⁰, Allo Docteurs¹¹ ou encore Spine-health¹² sont des exemples de forums de santé où les patients s'expriment librement et partagent leurs expériences sur leurs maladie et les traitements qui leurs sont éventuellement prodigués. Nous nous intéressons, dans ce travail, à l'analyse d'émotions dans ces forums. Cette tâche consiste à étudier les émotions (*e.g.* joie, colère, peur, tristesse...) à partir de textes rédigés en langue naturelle (titres de presse, blogs, messages, *etc.*). Comme nous venons de le montrer dans la section 3, contrairement à l'analyse d'opinions qui a été largement étudiée dans la littérature [8, 7, 1, 18], l'analyse d'émotions n'a pas bénéficiée de cet intérêt blue, la raison principale étant qu'il est difficile d'annoter des corpus de documents manuellement selon les émotions qu'ils expriment. En outre, ces travaux ne se sont pas intéressés au domaine de la santé qui soulève à lui seul d'autres problématiques que nous décrivons dans la section suivante. Par ailleurs, nous soutenons dans ce manuscrit que la tâche d'extraction de sentiments, en elle même, est très complexe et nous montrons, en l'évaluant, l'accord observé entre plusieurs annotateurs réalisant cette tâche manuellement.

4.2 Verrous

L'analyse des émotions dans les données collectées sur les forums de santé présentent de multiples verrous liés au domaine de la santé, au type de corpus étudié et à la tâche d'analyse.

Dans ces forums de santé, les textes sont hétérogènes (le vocabulaire du médecin et le vocabulaire du patient sont différents) ce qui rend difficile l'utilisation de ressources médicales pour détecter l'objet médical sur lequel porte l'émotion (tout le monde dit "j'ai mal à la tête", ce qui est connu chez les médecins par "la céphalée").

Les obstacles liés au corpus résident dans la volumétrie des textes ainsi que dans leur structure et leur contenu (beaucoup d'erreurs d'orthographe, utilisation des abréviations et du langage informel). Une autre difficulté concerne la longueur des messages des forums de santé (plusieurs phrases) sachant que la plupart des méthodes d'analyse d'émotions ont été appliquées sur des titres de presse [12, 21, 2] qui sont généralement très courts. Cette longueur complique la tâche d'analyse car un long message peut contenir plusieurs émotions.

D'autres verrous sont liés à la tâche de l'analyse d'émotions : premièrement, l'émotion est exprimée de plusieurs façons : smileys (*e.g.* :-) :-S), mots d'émo-

9. <http://www.doctissimo.fr/>

10. <http://www.same-story.com/>

11. <http://www.allodocteurs.fr>

12. <http://www.spine-health.com>

tion (e.g. "*heureux*", "*étonné*", etc.), lettres répétées (e.g. "*ffffff*", "*nooon*", etc.), majuscules (e.g. "*NUL*"), etc. Par ailleurs, le traitement de la négation est plus facile dans l'analyse de polarité que dans l'analyse d'émotions car, dans le premiers cas, il s'agit de détecter la polarité (généralement positive ou négative) ce qui fait que la négation d'une polarité devient généralement la polarité inverse. Par exemple, dans la phrase "je n'aime pas le dentiste" on peut déduire aisément que le sentiment exprimé est négatif. Par contre, dans l'analyse d'émotions, la négation n'est pas évidente à traiter. Par exemple, dans la phrase "*je n'ai pas peur*", il n'est pas évident de connaître à quelle classe d'émotion affecter cette "négation de la peur". S'agit-t'il de la même classe ? s'agit-il d'une autre classe ? dans ce cas comment la déterminer de manière automatique ?

En outre, l'émotion n'est pas seulement exprimée par des mots de sentiment, elle dépend aussi, bien sûr, de son contexte. Par exemple, la phrase "*je ne pense plus prendre ce mauvais médicament*" peut être classée différemment selon la connaissance de l'état d'esprit du locuteur (e.g. dans la catégorie "*colère*" ou "*tristesse*"). Seul le contexte peut donner des éléments pour éventuellement trancher pour l'une ou l'autre catégorie.

4.3 Méthodologie

Notre objectif est de proposer une approche qui soit capable de prendre en compte les contraintes des forums de santé pour offrir à l'utilisateur un outil qui détecte la présence d'une émotion, qui détecte la polarité d'une émotion (positive pour joie et négative pour colère, tristesse, dégoût et peur) et qui propose les étiquettes les plus probables des émotions. A terme, notre approche permettra également de découvrir la cible de l'émotion comme nous le décrirons en perspective.

Notre démarche s'est articulée en quatre temps (cf. Figure 5). Tout d'abord, nous avons collecté un corpus sur le forum de santé Spine-health¹³, ce corpus est annoté automatiquement en utilisant un lexique d'émotion. L'annotation automatique est validée manuellement par deux types différents d'annotateurs professionnels de santé et non professionnels de santé. Une fois l'annotation manuelle réalisée, nous étudions l'accord entre les annotateurs en utilisant la mesure statistique Kappa et nous expérimentons deux méthodes de classification classiques (statistiques et symboliques) afin d'évaluer leurs performances et leurs limites vis à vis de notre objectif. L'objectif est de choisir une ou plusieurs de ces méthodes ou éventuellement d'adapter l'une de ces méthodes pour qu'elle convienne au mieux à nos données et nous permette de surmonter les verrous liés à notre étude. Nous allons dans la suite de cette section décrire précisément chacune de ces phases.

4.3.1 Phase 1 : Annotation automatique

Dans cette phase, nous avons annoté automatiquement les messages selon les émotions et les objets médicaux qu'ils contiennent. L'objectif de cette phase a

13. <http://www.spine-health.com/>

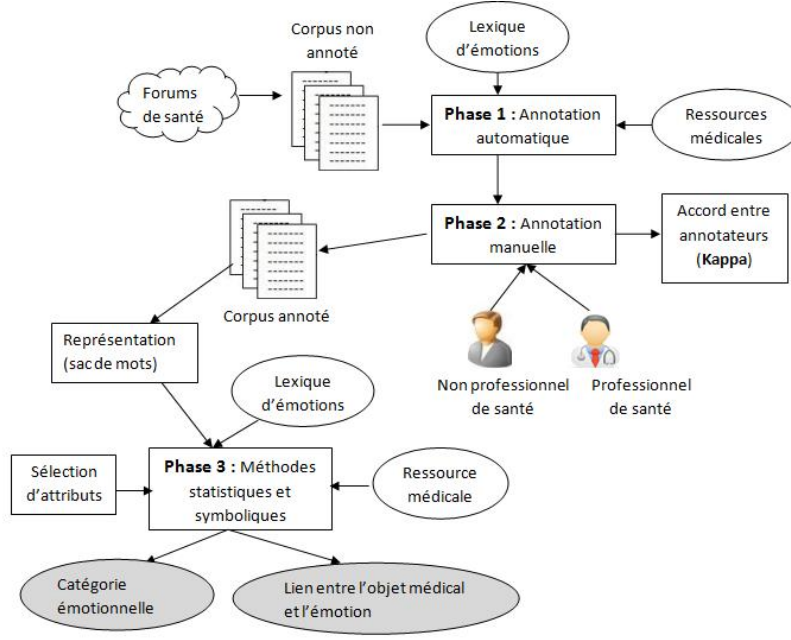


FIGURE 5 – Notre démarche d’analyse d’émotions dans les forums de santé

été de réaliser une analyse quantitative des traces d’émotions exprimées par les patients dans les messages issus des forums de santé. **Annotation des traces d’émotions.** Comme Strapparava et al. [21], pour détecter automatiquement les émotions exprimées dans les messages, nous nous sommes basés sur un lexique d’émotions. Nous avons choisi celui de Mohammad et al. [17] pour sa disponibilité et sa richesse (plus de 14000 mots d’émotions). Ce lexique est composé de 6 catégories émotionnelles e_i avec $i \in [1..6]$. Chaque mot du lexique w_j avec $j \in [1..14000]$ peut être associé à plusieurs émotions (e.g. le mot *abandoned* est associé aux émotions *peur*, *colère* et *tristesse*). Il est à noter que le lexique contient différentes classes grammaticales (nom, adjectifs, verbe, etc).

Nous annotons un message m composé de k mots $m = \{w_1, w_2, \dots, w_k\}$ avec l’émotion e_i si $\exists w_j \in m$ qui est associé à l’émotion e_i . Certains messages sont annotés avec plusieurs émotions. Par exemple, la phrase présentée dans la figure 6 contient deux mots *sufferer* et *injured* qui sont associés tous les deux aux émotions *peur* et *tristesse*. Cette phrase est donc annotée par deux émotions *peur* et *tristesse*. Les phrases ne contenant pas de mots du lexique sont considérées comme objectives.

However many chronic **sufferers** with CES have the pumps and work by directing the med directly to the **injured** nerves below the L5 S1

FIGURE 6 – Exemple d’une phrase de notre corpus annotée automatiquement

Comme nous le détaillerons dans la suite, cette annotation automatique permet de filtrer les messages objectifs (qui ne contiennent pas un mot d'émotion) et sert d'indication aux annotateurs pour l'annotation manuelle.

Annotation des objets médicaux. Dans cette étape, nous annotons automatiquement les entités médicales (noms de médicaments, maladies, virus, *etc.*) présents dans les messages du corpus. Afin de réaliser cette tâche, nous utilisons la ressource médicale MeSH (Medical Subject Headings)¹⁴. MeSH est un thésaurus développé par NLM (United States National Library of Medicine) et utilisé pour l'indexation des articles par le moteur de recherche médical PubMed¹⁵. Il est publiquement disponible et téléchargeable gratuitement¹⁶.

Nous avons choisi d'extraire les entités médicales à partir de MeSH pour sa disponibilité et sa richesse. En effet, MeSH est composée de plusieurs catégories médicales (maladie, médicament, symptôme, virus, *etc.*). Nous annotons un message s'il contient un terme médical de MeSH. Par exemple, la phrase présentée dans la figure 7 contient six objets médicaux (*pain, discs, herniated, arthritis, neck et head*).

I have been out of work 5 years due to low back **pain** but have **herniated discs** and severe **arthritis** in **neck** but sounds like your **head pain** is pretty severe.

FIGURE 7 – Exemple d'une phrase contenant des objets médicaux

Cette annotation permet de filtrer le jeu de données pour ne garder que les messages contenant au moins une entité médicale. En effet, notre objectif n'est pas seulement de trouver l'émotion exprimée dans le message mais aussi l'objet (médical) sur lequel porte l'émotion.

4.3.2 Phase 2 : Annotation manuelle

Dans un deuxième temps, ce corpus a été annoté manuellement via une interface Web. Cette annotation manuelle nous a permis de :

- Comparer l'annotation automatique à l'annotation manuelle censée être la plus juste, c'est-à-dire la plus représentative de l'émotion exprimée par le patient (même si celle-ci reste subjective et contextuelle) ;
- Mesurer si les annotateurs sont capables de s'accorder sur la présence d'une émotion ;
- Comparer l'annotation réalisée par un professionnel de santé et un non professionnel pour savoir si le professionnel détecte des éléments qui sont spécifiques à son domaine d'expertise en faisant une étude de l'accord entre les annotateurs ;
- Sélectionner un sous-ensemble du corpus initial afin d'obtenir un corpus équilibré (contenant le même nombre de messages par catégorie d'émotion).

14. <http://www.ncbi.nlm.nih.gov/mesh>

15. PubMed <http://www.ncbi.nlm.nih.gov/pubmed>

16. <http://bioportal.bioontology.org/ontologies/1351>

4 ANALYSE D'ÉMOTIONS DANS LES FORUMS DE SANTÉ

This is a Web application for annotating texts and sentences coming from health forums.
 Annotation must be performed according to the emotion category (joy, surprise, anger, fear, sadness, disgust ...) expressed in the sentence.
 The emotion category that is checked by default is detected using an automatic method and may be unchecked if you believe that it is not expressed in the sentence.
 In case where no emotion is expressed in the sentence (objective), all emotions must be unchecked.
 If you're not sure of your decision or you don't understand the meaning of a sentence, please check "I don't know"

<p>Well my work comp insurance company will not approve an epidural steroid injection since I do not have radiculopathy so my surgeon now wants to pursue medial branch blocks or intra-articular facet joint injections. Hopefully work comp insurance will approve these. Has anyone had either one of these type of injections? Anakala</p>	<input checked="" type="checkbox"/> joy <input type="checkbox"/> fear <input type="checkbox"/> surprise <input type="checkbox"/> disgust <input type="checkbox"/> anger <input type="checkbox"/> sadness <input type="checkbox"/> I don't know
--	--

<p>So a while back I had an episode which led to pain in my back as well as chest and upper stomach sharp pains, I couldn't breath and they took me to the emergency room which stated I had a really back stomach virus that caused me to have this pain. Now I have a back problem it's started as a little pain now I can't walk sit lie down nor laugh or cough or sneeze. The ER said I had a herniated disc keep in mind I'm only 16, I went to another doctor who said it was a herniated disc the ER had me on two dosages of morphine which didn't work the doctor had me on perks, Vicodin anything you think of even steroids nothing helped this pain and he sent me to the pain specialist who had me running around claiming it wasn't a herniated disc he thinks it was my coccyx he had me do another MRI and he was wrong while on my medication I was getting dizzy which was normal for me I always get dizzy I think it's because of my anemia, anyway at one point I passed out it was really scary, the doctor said to watch it and he said it may be the meds. Today though I had another episode much like the one when I apparently had a stomach bug, except the chest pains are far worse close to my heart my aunt called when it happened and she got so worried I couldn't talk I couldn't breathe and I was dizzy it hurt so much and again my back was hurting finally it subsided but it still was hurting a little and I was still dizzy. Basically I'm going to see an ortho for my spine on Tuesday because these other doctors have given me the run around I'm only 16, a female and I should not even be having these problems also does anyone know if I should see someone about my heart and the chest pains it felt like a mini heart attack if that's possible and also about my back do you think the two could be related nerve wise? Please help I'm worried. Crystal</p>	<input checked="" type="checkbox"/> joy <input type="checkbox"/> fear <input type="checkbox"/> surprise <input type="checkbox"/> disgust <input type="checkbox"/> anger <input type="checkbox"/> sadness <input type="checkbox"/> I don't know
---	--

FIGURE 8 – Plateforme d'annotation pour les professionnels de santé

Nous avons mis en place une plateforme Web pour l'annotation des phrases. Cette plateforme a été spécialisée en deux versions. La première est destinée aux professionnels de santé et permet de visualiser la phrase à annoter dans son contexte (message) (cf. Figure 8) et la deuxième est dédiée aux non professionnels (cf. Figure 9) où seule la phrase est affichée mais pas son contexte.

Le choix de présenter la phrase dans son contexte pour les professionnels de santé a pour objectif de les aider à mieux comprendre les problèmes des patients. Toutefois, afin de comparer l'annotation réalisée par un professionnel de santé et un non professionnel (accord entre annotateurs), ce dernier a utilisé l'interface destinée aux professionnels de santé pour ne pas biaiser les résultats.

L'annotation se fait selon l'émotion (joie, surprise, colère, tristesse, dégoût, peur) exprimée dans la phrase. L'annotateur doit cocher la ou les émotions exprimées dans la phrase. Chaque phrase ayant été étiquetée automatiquement en utilisant le lexique, l'émotion correspondant à cet étiquetage est présentée par défaut mais peut être décochée si l'annotateur estime qu'elle n'est pas exprimée dans la phrase. Dans le cas où la phrase n'exprime aucune émotion toutes les émotions doivent être décochées. Enfin, si l'annotateur ne comprend pas le sens de phrase, il doit cocher " je ne sais pas ". Pour faciliter l'annotation de plusieurs phrases, un annotateur obtient en se connectant 15 phrases exprimant la même émotion (*e.g.* uniquement des phrases annotées automatiquement par l'émotion colère). Si une phrase a été annotée automatiquement par plusieurs émotions, elle sera annotée manuellement autant de fois que le nombre d'émotions qu'elle exprime.

4.3.3 Phase 3 : Classification automatique

Nous choisissons de tester deux méthodes statistiques : machine à vecteurs supports (SVM) et Naïve Bayes (NB). Ces méthodes sont connues pour leurs

4 ANALYSE D'ÉMOTIONS DANS LES FORUMS DE SANTÉ

This is a Web application for annotating texts and sentences coming from health forums.
 Annotation must be performed according to the emotion category (joy, surprise, anger, fear, sadness, disgust ...) expressed in the sentence.
 The emotion category that is checked by default is detected using an automatic method and may be unchecked if you believe that it is not expressed in the sentence.
 In case where no emotion is expressed in the sentence (objective), all emotions must be unchecked.
 If you're not sure of your decision or you don't understand the meaning of a sentence, please check "I don't know"

Well my work comp insurance company will not approve an epidural steroid injection since I do not have radiculopathy so my surgeon now wants to pursue medial branch blocks or intra-articular facet joint injections ☒ joy ☐ fear ☐ surprise ☐ disgust ☐ anger ☐ sadness ☐ I don't know

Now I have a back problem it's started as a little pain now I can't walk sit lie down nor laugh or cough or sneeze ☒ joy ☐ fear ☐ surprise ☐ disgust ☐ anger ☐ sadness ☐ I don't know

FIGURE 9 – Plateforme d'annotation pour les non professionnels de santé

bonnes performances dans la problématique d'analyse d'opinions et d'émotions [1, 19, 21]. Nous présentons, dans cette section, la méthodologie de représentation des données (types d'attributs, de pondération, *etc.*).

Pour appliquer les méthodes statistiques, nous utilisons une représentation en sac de mots et des valeurs booléennes (l'attribut prend la valeur "vrai" s'il est présent dans le texte et la valeur "faux" sinon). Dans la représentation sac de mots, les données sont représentées sous forme vectorielle, où chaque texte est représenté par un vecteur d'attributs et où les attributs sont les mots composant les textes. Par exemple, les deux phrases P1 et P2 sont représentées en sac de mots avec des valeurs booléennes comme le montre le tableau 2 où les colonnes représentent les attributs (uni-grammes) et les lignes représentent les phrases.

P1 = *After I recoved and longer took the severe pain have been in*

P2 = *After I have been in severe pain*

	After	I	recoved	and	longer	took	the	severe	pain	have	been	in
P1	V	V	V	V	V	V	V	V	V	F	F	F
P2	V	V	F	F	F	F	F	V	V	V	V	V

TABLE 2 – Exemple de représentation en sac de mots avec les uni-grammes et des valeurs booléennes

Nous faisons varier les attributs utilisés afin de trouver les meilleurs descripteurs des émotions. Voici les attributs que nous avons utilisés :

- **Unigrammes** : il s'agit de l'approche classique de sélection d'attributs dans laquelle chaque document est représenté par ses mots ;
- **Unigrammes + lemmatisation** : l'avantage de la lemmatisation est que les mots sont généralisés par leur lemme ce qui permet de réduire la dimension mais cette généralisation peut engendrer une perte d'information. Nous utiliserons les deux configurations (avec et sans lemmatisation) pour savoir laquelle est la plus adapté à notre jeu de données ;
- **Unigrammes + POS** : nous enrichissons la représentation par l'étiquetage grammatical (POS) qui permet la désambiguïsation lexicale. Par exemple, il permet de désambiguïser le mot "love" dans les phrases "love story" et "I love my doctor". En effet, le mot "love" représente un nom dans la première phrase et un verbe dans la deuxième ;
- **Adjectifs** : nous utilisons également l'étiquetage grammatical pour extraire les adjectifs. Les adjectifs sont, selon Wiebe [27], de bons indica-

teurs de subjectivité dans les textes. Les résultats de [19] ont montré que l'utilisation de tous les mots du corpus est plus efficace que l'utilisation des adjectifs seuls ;

- **Bigrammes** : un bigramme est composé de 2 mots successifs qui permet de capturer plus de contexte et plus spécifiquement pour capter la négation ;
- **Mots d'émotion** : tous les mots de lexique qui expriment au moins une émotion parmi les six ;
- **Méta-attributs** : nous utilisons également des méta-attributs qui permettent de capturer les émotions dans les messages comme les smileys (:-), :(...), la ponctuation (!!, ??, ...), les lettres répétées (loooool) et les majuscules (NUL) ;
- **Contexte d'émotion** : pour prendre en compte le contexte des émotions nous utilisons deux attributs que nous appelons émotion voisine et émotion globale :
 - L'émotion voisine prend la valeur vrai si la phrase qui précède ou qui suit la phrase en question exprime la même émotion que celle-ci,
 - L'émotion globale prend la valeur vrai s'il existe une autre phrase dans le message qui exprime la même émotion que la phrase en question.
- **Patrons** : comme [23], nous enrichissons également nos attributs en utilisant les patrons syntaxiques. Dans le paragraphe suivant, nous présentons les étapes suivies pour extraire les patrons.

Nous avons utilisé les motifs séquentiels pour extraire les patrons fréquents de la manière suivante : nous commençons par un étiquetage des entités médicales en utilisant la ressource médicale MeSH (étiquette : EM) et des mots d'émotions en utilisant le lexique d'émotions (étiquette : Emt) dans les phrases, ensuite nous extrairons la catégorie grammaticale de chaque mot de la phrase. Enfin, chaque phrase est considérée comme une séquence où les itemsets sont les catégories grammaticales des mots. Exemple :

- Phrase : *chronic pain may cause secondary depression* ;
- Etiquetage des mots d'émotion + entités médicales : chronic/Emt/EM pain/Emt/EM may cause secondary depression/Emt ;
- Etiquetage grammatical : chronic/JJ/Emt/EM pain/NN/Emt/EM may/MD cause/VV secondary/JJ depression/NN/Emt ;
- Séquence : JJ/Emt/EM NN/Emt/E MD VV JJ NN/Emt.

Une fois ces séquences extraites, nous avons utilisé l'algorithme de recherche des motifs séquentiels GSP [20]. Nous avons obtenu quatre niveaux de motifs séquentiels fréquents (cf. Tableau 3). Enfin, les motifs fréquents qui contiennent au moins une étiquette d'un mot d'émotion sont utilisés comme des attributs (lors de la représentation de la phrase pour la classification, l'attribut patron prend la valeur vrai si la forme syntaxique de la phrase respecte le patron).

Nombre de sé- quences	Séquence les plus fréquentes	Fréquence
1	JJ/EMt	2515
2	JJ/EMt,NN/EMt/EM	1317
3	JJ/EMt,DT,NN/EMt/EM	237
4	JJ/EMt,PP,VBP,IN	123

TABLE 3 – Exemples des séquences fréquentes

4.3.4 Phase 4 : Identification des liens entre traces d'émotions et objets médicaux.

Une fois le message associé à une ou plusieurs émotions, nous cherchons à identifier l'objet de l'émotion. Par exemple, dans la phrase "J'ai peur de l'anesthésie", l'objet de la peur exprimée par le locuteur est l'anesthésie. Pour cela, nous avons collaboré avec Pierre Pompidor du LIRMM qui a développé une méthode symbolique basée sur les patrons syntaxiques pour trouver l'objet médical sur lequel porte l'émotion.

4.4 Constitution des corpus

La figure 10 montre le processus de construction de nos corpus. Tout d'abord, nous avons collecté 17000 messages issus du forum de santé en anglais Spine-health (spécialisé dans la thématique douleur et plus précisément la douleur de dos). Deuxièmement, nous avons filtré les messages objectifs (qui ne contiennent pas un mot de lexique). Le tableau 4 présente la répartition des messages sur les six catégories émotionnelles à l'issue de la phase 1 d'annotation automatique. Comme le montre le tableau, la plupart des messages sont subjectifs et expriment plus la tristesse (82%) le dégoût (78%). Ce résultat confirme l'intuition qui apparaît à la lecture des messages du forum où les personnes s'expriment sur les douleurs ou les problèmes qu'ils ont par rapport au mal de dos.

Objectif	subjectif					
22%	78%					
	Joie	surprise	peur	colère	tristesse	dégoût
	63%	48%	62%	51%	82%	74%

TABLE 4 – Répartition des messages sur les six catégories émotionnelles en utilisant une méthode automatique

Ensuite, les messages subjectifs sont segmentés en phrases. Dans un message, plusieurs émotions sont généralement mélangées et beaucoup d'entités médicales y sont présentes du fait de la longueur relative des messages par rapport aux phrases (les messages contiennent sept phrases en moyenne). Nous choisissons donc de travailler avec les phrases. Nous utilisons le point comme séparateur. Notre corpus contient alors des phrases dont nous avons gardé des informations

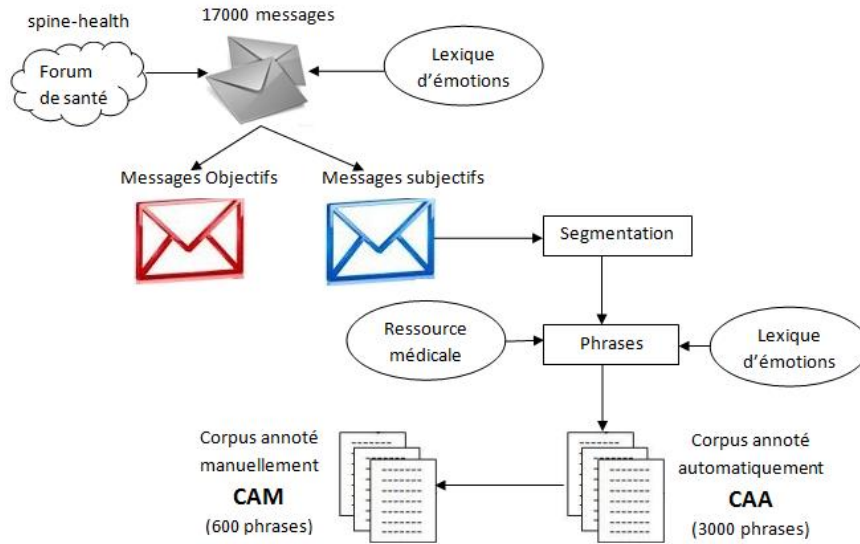


FIGURE 10 – Processus de construction des corpus

pour reconstruire ultérieurement les messages. Le tableau 5 présente la répartition des messages selon le nombre d'émotions qui y sont présentes (en utilisant la méthode automatique).

1	2	3	4	5	6
5%	9%	10%	16%	19%	41%

TABLE 5 – Répartition des messages selon le nombre d'émotions y présentes

Nous remarquons sur ce tableau que la majorité des messages contiennent plus d'une émotion. En effet, la moitié des messages contiennent plus de 5 émotions. En revanche, les phrases ne présentent pas la même répartition (tableau 6).

1	2	3	4	5	6
26%	35%	17%	13%	5%	4%

TABLE 6 – Répartition des phrases selon le nombre d'émotions y présentes

Nous avons finalement gardé 3000 phrases pour constituer un corpus annoté automatiquement (CAA)¹⁷ et qui respectent les conditions suivantes :

- Avoir un nombre de mots supérieur ou égal à 20 ;

17. Pour la simplicité, dans la suite de ce document, nous utilisons l'abréviation "CAA" pour désigner le corpus annoté automatiquement.

- Au moins 10% des mots de la phrase sont des entités médicales ;
- Au moins 10% des mots de la phrase se trouvent dans le lexique.

Un sous ensemble de ce corpus a été annoté manuellement. Nous le notons CAM¹⁸. Ce corpus contient 600 phrases qui ont été annotées par 60 annotateurs.

Pour plus de détails sur ces corpus, voir la section 5.3.

5 Résultats expérimentaux

Dans cette partie, nous avons réalisé une étude du lexique utilisé dans la phase 1 (cf. Section 4.3.1) afin de mieux comprendre son architecture et son adéquation à notre jeu de données. Par la suite, nous avons étudié l'accord entre les annotateurs (professionnels et non professionnels). Cette étude nous a permis de conclure sur la difficulté de la tâche d'annotation des émotions, même pour les humains. Enfin, nous avons évalué la classification afin de trouver le meilleur algorithme (SVM ou NB), le(s) meilleur(s) descripteur(s) de l'émotion et la différence existant entre un corpus annoté automatiquement à partir d'un lexique comme proposé dans la littérature et un corpus annoté manuellement.

5.1 Évaluation du lexique

Nous avons réalisé un certain nombre de statistiques sur le lexique. En particulier, nous avons étudié la distribution des mots du lexique sur les six classes d'émotions et la corrélation entre ces dernières, l'adéquation du lexique pour les messages de santé et la répartition des entités médicales présentes dans le lexique. Cette étude nous a aidé à commenter les résultats obtenus par la suite.

5.1.1 Répartition des mots du lexique sur les six catégories émotionnelles

Le tableau 7 représente la distribution des mots sur les six classes d'émotions : joie, surprise, peur, colère, tristesse et dégoût. Nous remarquons que les

Joie	Surprise	Peur	Colère	Tristesse	Dégout
689	534	1476	1247	1191	1058

TABLE 7 – Répartition des mots du lexique de Mohammad sur les six catégories émotionnelles

émotions négatives¹⁹ sont mieux représentées dans le lexique (plus d'un millier de mots pour chacune des catégories) contrairement aux deux émotions "joie" et "surprise".

¹⁸. Dans la suite de ce document, nous utilisons l'abréviation "CAM" pour désigner le corpus annoté manuellement.

¹⁹. Dans le reste de ce document, nous utilisons le terme "émotions négatives" pour désigner les émotions suivantes : peur, colère, tristesse et dégoût

5.1.2 Corrélation entre les six catégories émotionnelles

Nous avons également étudié la corrélation qui existe entre ces catégories dans le lexique car nous avons constaté que beaucoup de mots sont étiquetés par les mêmes catégories d'émotion (*e.g.* joie et surprise). Pour cela, nous utilisons l'information mutuelle (MI) qui mesure la dépendance mutuelle entre deux variables. Elle est calculée par la formule suivante :

$$MI(Emotion1, Emotion2) = \frac{nb(Emotion1 \text{ ET } Emotion2)}{nb(Emotion1) \cdot nb(Emotion2)} \quad (5)$$

Où $nb(e)$ représente le nombre de mots du lexique qui sont étiquetés par l'émotion 'e'. Le tableau 8 montre la valeur de l'information mutuelle (multipliée par 10^3) calculée par la formule (1) entre chaque couple d'émotions.

	Joie	Surprise	Peur	Colère	Tristesse	Dégout
Joie	X	0.549	0.049	0.043	0.046	0.027
Surprise		X	0.241	0.201	0.166	0.139
Peur			X	0.382	0.374	0.311
Colère				X	0.350	0.416
Tristesse					X	0.342
Dégout						X

TABLE 8 – Valeur de l'information mutuelle (multipliée par 10^3) calculée entre chaque couple d'émotions

Nous remarquons qu'il existe une corrélation relativement forte entre les émotions négatives (une information mutuelle supérieure à 0.3 pour les différentes combinaisons entre peur, colère, tristesse et dégoût). Il est à noter également que l'émotion surprise est plus corrélée avec l'émotion joie qu'avec les émotions négatives.

5.1.3 Adéquation du lexique d'émotion au corpus d'étude

Le lexique utilisé étant généraliste, il est important de vérifier s'il est utilisable dans notre contexte. Nous calculons pour cela le nombre de mots du lexique qui apparaissent au moins une fois dans les données brutes. Le tableau 9 représente ces nombres pour les six émotions. Comme nous pouvons le consta-

Joie	Surprise	Peur	Colère	Tristesse	Dégout
68%	69%	58%	56%	61%	57%

TABLE 9 – Pourcentage des mots du lexique qui apparaissent dans notre corpus

ter, la plupart des mots d'émotions des forums sont présents dans le lexique ce qui nous permet de considérer que ce dernier est en adéquation avec nos besoins même si le domaine est très orienté.

5.1.4 Connotation émotionnelle des objets médicaux

Le lexique que nous utilisons contient des mots qui sont aussi des entités médicales (*e.g. cancer, pain, etc.*). L'intersection entre le lexique et l'ensemble des entités médicales extraites à partir de MeSH montre que 585 entités médicales sont réparties sur les six catégories émotionnelles (cf. Tableau 10).

Joie	Surprise	Peur	Colère	Tristesse	Dégout
11%	15%	18%	13%	16%	17%

TABLE 10 – Entités médicales présents dans le lexique réparties sur les 6 catégories émotionnelles

Même si ces entités médicales possèdent des caractéristiques émotionnelles, de manière à ne pas biaiser les l'évaluation de l'émotion globale exprimée, nous choisissons d'éliminer ces 585 mots du lexique.

5.2 Accord entre annotateurs

Pour évaluer l'accord entre les annotateurs, nous avons utilisé la mesure statistique Kappa " k " [5] car cette mesure est souvent utilisée pour calculer l'accord entre annotateurs. L'accord k entre deux annotateurs A et B est basé sur le tableau 11 et calculé par l'équation 6.

$$k = \frac{Pr(a) - Pr(e)}{1 - Pr(e)} \quad (6)$$

où $Pr(a)$ est l'accord relatif entre les annotateurs (équation 7) et $Pr(e)$ la probabilité de l'accord aléatoire (équation 8).

Annotateur B	Annotateur A		Total
	–	+	
+	a	b	B^+
–	c	d	B^-
Total	A^+	A^-	n

TABLE 11 – Accord entre les annotateurs [5]

$$Pr(a) = \frac{a + b}{n} \quad (7)$$

$$Pr(e) = \frac{A^+A^- + B^+B^-}{n} \quad (8)$$

Si les annotateurs sont totalement en accord, $k = 1$. S'ils sont totalement en désaccord (ou en accord dû uniquement au hasard), $k < 0$ (voir tableau 12).

Afin d'étudier l'accord entre les annotateurs non professionnels, 150 phrases ont été annotées par deux annotateurs non professionnels de santé. Le tableau 15

5 RÉSULTATS EXPÉRIMENTAUX

k	Interpretation
< 0	Désaccord
0.0-0.20	Accord très faible
0.21-0.40	Accord faible
0.41-0.60	Accord modéré
0.61-0.80	Accord fort
0.81-1.00	Accord presque parfait

TABLE 12 – Interprétation de k de Kappa [5]

présente l'accord entre les deux annotateurs non professionnels de santé pour les six catégories émotionnelles ainsi que les deux classes "pas d'émotion" et "je ne sais pas". Chaque cellule de ce tableau représente le nombre de phrases annotées par les deux émotions précisées par la ligne et la colonne. La diagonale du tableau contient le nombre de phrases en accord. On remarque que les annotateurs étaient plus d'accord sur l'émotion dégoût et moins d'accord sur l'émotion colère. Nous avons demandé aux deux annotateurs de définir les quatre émotions pour lesquelles ils ont été le moins en accord : surprise, colère, dégoût et tristesse. Leur définitions sont données dans la table 13. Il est important de noter que leurs représentations mentales de ces émotions sont très divergentes. Ce constat est cohérent avec la valeur de la mesure kappa calculée qui est égale à 0.26 ce qui illustre bien que l'accord global entre les annotateurs reste faible.

	Non professionnels de santé A	Non professionnels de santé B
Surprise	Sentiment exprimé lors d'un événement inattendu	Des nouvelles dont on ne s'attendait pas
Colère	Mécontentement mais un peu plus prononcé que le dégoût	Quand on est énervé
Dégoût	Le dégoût exprime généralement la douleur	Pas d'espoir de trouver une solution à son problème
Tristesse	Contrairement au dégoût qui est instantané, la tristesse est un sentiment qui est là depuis longtemps et dont les causes remontent au passé	Pas d'espoir + un problème qui dure longtemps

TABLE 13 – Définition de quatre émotions par deux annotateurs

De même, pour évaluer l'accord entre les annotateurs professionnels et non professionnels, 150 phrases ont été annotées par les deux catégories. Le tableau 15 présente l'accord entre les deux types d'annotateurs par catégorie d'émotion.

5 RÉSULTATS EXPÉRIMENTAUX

Non prof de santé B	Non prof de santé A								
		Joie	Surprise	Colère	Peur	Dégout	Tristesse	Sais pas	Neutre
	Joie	13	0	0	0	0	2	0	2
	Surprise	0	2	0	1	4	1	0	1
	Colère	1	2	2	1	6	3	5	11
	Peur	0	11	1	9	15	4	0	0
	Dégout	0	9	4	1	30	3	2	10
	Tristesse	1	5	4	1	11	8	3	11
	Sais pas	1	1	2	1	0	0	1	4
	Neutre	0	0	0	1	1	1	0	11

TABLE 14 – Accord entre annotateurs non professionnels de santé pour les six catégories émotionnelles ainsi que les deux classes "pas d'émotion" et "je ne sais pas"

Dans ce cas, les annotateurs étaient plutôt d'accord pour les émotions "joie", "peur" et "je ne sais pas". L'accord global entre les annotateurs est modéré ($k=0.46$).

Malgré le peu de volumes de ces expérimentations, nous avons l'intuition qu'il peut y avoir plus de variabilité entre individus qu'entre professionnels et non professionnels. On constate que certaines émotions sont corrélées. Par exemple, "peur", "dégout", "colère" et "tristesse" peuvent être regroupées en une classe négative, la "joie" dans une classe positive et la "surprise" en neutre. Cette constatation est cohérente avec les résultats présentées dans la section 5.1 sur l'analyse du lexique. Nous utilisons cette information dans la section 5.4 pour l'étape de classification en opposant une tâche de classification bi-classe (émotions positives vs. négatives) à la tâche que nous nous étions initialement fixé (multi-classe). En conclusion, l'analyse d'émotion est une tâche difficile car mêmes les humains ne sont pas d'accord sur la présence d'une émotion.

Professionnel de santé	Non professionnel de santé								
		Joie	Surprise	Colère	Peur	Dégout	Tristesse	Sais pas	Neutre
	Joie	9	0	2	1	0	1	0	0
	Surprise	0	0	0	0	0	0	0	0
	Colère	0	1	5	1	1	1	0	0
	Peur	0	2	1	9	0	0	0	0
	Dégout	0	0	0	2	3	0	0	0
	Tristesse	0	1	4	2	6	2	0	0
	Sais pas	3	0	8	5	15	13	0	2
	Neutre	6	3	5	15	19	14	0	9

TABLE 15 – Accord entre annotateurs professionnel de santé et non professionnel pour les six catégories émotionnelles ainsi que les deux classes "pas d'émotion" et "je ne sais pas"

5.3 Statistiques sur les deux corpus CAA et CAM

Dans cette section, nous réalisons quelques statistiques sur chaque corpus (CAA et CAM) comme la répartition des phrases annotées selon les catégories

émotionnelles, *etc.* Nous rappelons que le corpus CAA contient 3000 parmi lesquelles 900 ont été annotées manuellement par 60 annotateurs. Pour les deux corpus, nous ré-annotons les phrases contenant plusieurs émotions par l'émotion majoritaire par laquelle la phrase a été annotée. D'une part notre objectif est de chercher l'émotion la plus exprimée et d'autre part le phénomène de sur-émotion peut affecter les performances des méthodes de classification. Pour les deux corpus, nous présentons les points suivants :

- Répartition des phrases annotées selon les catégories émotionnelles ;
- Répartition des phrases selon le nombre d'émotions qui y sont présentes ;
- Répartition des phrases dans le corpus final selon les catégories émotionnelles après la ré-annotation ;
- Nombre d'attributs pour chaque corpus.

Nous présentons simultanément les statistiques des deux corpus pour pouvoir les comparer.

5.3.1 Répartition des phrases annotées selon les catégories émotionnelles

Le tableau 17 représente la répartition des phrases dans le corpus CAA selon les six catégories émotionnelles. Nous remarquons que le corpus n'est pas parfaitement équilibré. Le tableau 17 présente la répartition des phrases dans le

Joie	Surprise	Peur	Colère	Tristesse	Dégout
22%	14%	39%	22%	39%	18%

TABLE 16 – Répartition des phrases dans le corpus CAA selon les six catégories émotionnelles

corpus CAM sur les six catégories émotionnelles ainsi que les deux classes "pas d'émotion" et "je ne sais pas". Nous remarquons que 45% des phrases ont été annotées comme neutres (elle n'expriment pas d'émotion) et 9% des phrases ne sont pas claires (l'annotateur coche "je ne sais pas").

Joie	Surprise	Peur	Colère	Tristesse	Dégout	pas d'émotion	je ne sais pas
10%	4%	17%	13%	19%	14%	45%	9%

TABLE 17 – Répartition des phrases dans le corpus CAM selon les catégories émotionnelles

5.3.2 Répartition des phrases selon le nombre d'émotions qui y sont présentes

Le tableau 18 représente la répartition des phrases de corpus CAA et de corpus CAM selon le nombre d'émotions qui y sont présentes. Nous remarquons dans ce tableau que pour les deux corpus, la majorité des phrases contiennent une seule émotion.

5 RÉSULTATS EXPÉRIMENTAUX

	1	2	3	4	5	6
CAA	54%	30%	11%	4%	1%	0%
CAM	53%	29%	11%	6%	1%	0%

TABLE 18 – Répartition des phrases de corpus CAA et de corpus CAM selon le nombre d’émotions qui y sont présentes

5.3.3 Répartition des phrases dans le corpus final selon les six catégories émotionnelles après la ré-annotation

Finalement, après la ré-annotation des phrases à plusieurs émotions par l’émotion majoritaire et la suppression des phrases annotés par "pas d’émotion" et "je ne sais pas" dans le corpus CAM, les deux corpus finaux sont réparti comme le montre le tableau 19. Nous remarquons que le corpus CAA est très déséquilibré et que les deux émotions peur et tristesse sont les mieux représentées. Nous remarquons également que le corpus CAM est mieux équilibré que le corpus CAA. Néanmoins, l’émotion surprise demeure toujours très mal représentée.

	Joie	Surprise	Peur	Colère	Tristesse	Dégout
CAA	13%	4%	33%	9%	35%	6%
CAM	14%	6%	23%	14%	25%	18%

TABLE 19 – Répartition des phrases de corpus CAA et de corpus CAM selon les six catégories émotionnelles après la ré-annotation

5.3.4 Nombre d’attributs pour chaque corpus

Le tableau 20 représente le nombre d’attributs dans les deux corpus CAA et CAM. Il est à noter que nous filtrons les attributs qui apparaissent une seule fois dans le corpus. Comme la taille de corpus CAM fait 1/5 de la taille de CAA, nous remarquons que les nombres des uni-grammes avec et sans lemmatisation, uni-grammes avec étiquetage grammaticales, bigrammes et adjectifs ont considérablement baissé dans le CAM.

5.4 Évaluation de la classification automatique

Dans cette partie, nous répondons aux questions suivantes :

- SVM est-il meilleur que NB ?
- Quelle sont les meilleurs descripteurs de l’émotion ?
- Est-ce que le corpus annoté manuellement est très différent que le corpus annoté automatiquement ?

Pour nos expérimentations sur la classification, nous utilisons les implémentations de Weka [6] pour SVM et NB. Les corpus utilisés pour l’apprentissage et les tests sont : le corpus CAA annoté automatiquement en utilisant le lexique d’émotions et celui annoté manuellement (CAM) par des non professionnels de santé.

Attributs		Nombre	
		CAA	CAM
Uni-grammes	sans lemmatisation	2860	637
	avec lemmatisation	2685	761
	+POS	3837	904
Méta-attributs		4	4
Contexte d'émotion		2	2
Patrons		4	2
Bigrammes		2130	743
Adjectifs		957	350
Mots d'émotion		3650	3650

TABLE 20 – Nombre d'attributs dans les deux corpus CAA et CAM

Le choix du premier est lié au fait qu'il est beaucoup plus volumineux (3000 phrases vs. 600 phrases pour le manuel). En outre, nous souhaitons évaluer si l'annotation automatique est suffisante pour classer les émotions. Pour les deux algorithmes, nous faisons la classification des différentes catégories émotionnelles (multi-classe) et la classification des émotions positives vs. négatives (bi-classe) où l'émotion positive représente la joie et les émotions négatives représentent la colère, la peur, la tristesse et le dégoût. L'émotion surprise est éliminée car elle peut appartenir aux deux classes d'émotion positive et négative (voir la figure 11). Nous utilisons les mesures de rappel, précision et F-mesure (cf. formules de la section 2.2) pour évaluer les résultats obtenus et une validation croisée (à 10 folds) et on fait varier le type d'attribut. Voici les combinaisons des attributs que nous avons utilisés :

- Uni-grammes
- Uni-grammes + lemmatisation
- Uni-grammes + Pos (étiquetage grammatical)
- Uni-grammes + méta-attributs (smileys, majuscule, *etc.*)
- Uni-grammes + le contexte d'émotion
- Uni-grammes + patrons syntaxiques
- Bigrammes
- Adjectifs
- Mots d'émotions

Dans chacun des tableaux montrant les résultats des expériences, nous affichons en gras les meilleurs résultats. Des expériences complémentaires sont données en annexe de ce mémoire.

5.4.1 Corpus CAA

Afin de répondre à nos questions, nous réalisons une comparaison entre les résultats obtenus avec les deux algorithmes SVM vs. NB avec la classification des différentes catégories émotionnelles (multi-classe) vs. la classification en positives/négatives. La comparaison est basée sur le calcul de la moyenne des

5 RÉSULTATS EXPÉRIMENTAUX

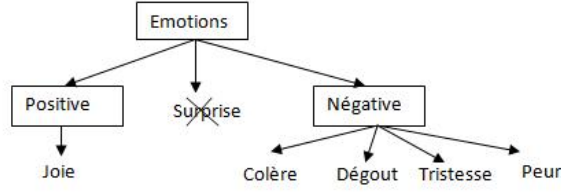


FIGURE 11 – Classement des émotions en polarité positive et négative

mesures rappel, précision et F-mesure entre les différentes classes (les six classes d'émotion dans le cas de classification multi-classe et entre la classe positive et négative dans le deuxième cas).

Le tableau 23 montre les résultats obtenus avec le corpus CAA et en faisant varier les différents types d'attributs. Nous remarquons que les mots d'émotions donnent les meilleurs résultats avec SVM et pour le NB, les uni-grammes avec et sans lemmatisation, les méta-attributs, l'émotion globale et les patrons donnent le meilleur rappel et le meilleur F-score et la meilleure précision est donné par les mots d'émotions pour la classification en bi-classes (émotions positives vs. négatives). En conclusion, nous pouvons constater que les meilleurs résultats sur le

Attributs		NB						SVM					
		Multi-Emotion			Pos/Neg			Multi-Emotion			Pos/Neg		
		P	R	F	P	R	F	P	R	F	P	R	F
Uni-grammes	sans lemmatisation	27.05	21.25	23.8	57.15	53.75	55.39	20.86	19.6	20.21	62.9	58.9	60.83
	avec lemmatisation	25.36	21.23	23.11	57.8	54.05	55.86	21.98	20.33	21.12	64.8	61.75	63.23
	+POS	22.08	19.03	20.44	56.25	53.4	54.78	23.45	22.43	22.93	61.75	59.25	60.45
	+Méta-attributs	27.11	21.25	23.82	56.9	53.6	55.2	23.08	21.66	22.35	63.85	61.25	62.52
	+Contexte d'émotion	29.08	23.86	26.21	75.3	53.75	55.46	30.68	29.45	30.05	63.35	60.95	62.12
	+patrons	27.01	21.16	23.73	57.5	54.05	55.75	23.16	61.65	22.38	62.8	58.85	60.76
Bigrammes		19.96	17.23	18.49	56.35	45.75	50.49	32.75	65.88	43.75	59.2	50.55	54.53
Adjectifs		25	18.56	21.36	58.25	48	52.63	24.03	20.11	21.90	59.75	50.95	55
Mots d'émotion		22.46	20.58	21.48	75.15	43.65	55.22	37.9	33.7	33.7	76.7	65.9	70.89

TABLE 21 – Comparaisons des résultats obtenus avec NB vs. SVM en utilisant le corpus CAA

corpus CAA sont obtenus par SVM en prenant comme descripteurs uniquement les mots d'émotions (cf. Tableau 22).

Attributs	SVM					
	Multi-Emotion			Pos/Neg		
	P	R	F	P	R	F
Mots d'émotion	37	33	34	76	65	65

TABLE 22 – Le meilleur algorithme est SVM et les meilleurs descripteurs de l'émotion sont les mots d'émotion (CAA)

5 RÉSULTATS EXPÉRIMENTAUX

5.4.2 Corpus CAM

Les meilleurs descripteurs des émotions en utilisant le corpus CAM avec les deux types de classification (multi-classe et bi-classe) et pour les deux algorithmes NB et SVM sont les uni-grammes + les patrons (cf. Tableau 21).

Attributs		NB						SVM					
		Multi-Emotion			Pos/Neg			Multi-Emotion			Pos/Neg		
		P	R	F	P	R	F	P	R	F	P	R	F
Uni-grammes	sans lemmatisation	15.53	15.85	15.69	58.5	53.9	56.1	15.7	15.91	15.8	60.55	58.65	59.58
	avec lemmatisation	15.68	15.51	15.59	57.6	54.85	56.19	15.08	15.53	15.3	59.1	54.35	56.62
	+POS	15.18	16.41	15.77	57	53.9	55.4	15.15	14.9	15.02	55.95	54.95	55.21
	+Méta-attributs	16.03	16.68	16.53	59.35	54.45	56.79	14.76	15	14.88	60	58.9	59.44
	+Contexte d'émotion	15.13	16.16	15.63	59.35	54.5	56.76	17.9	18.26	18.08	60	58.65	59.66
	+patrons	14.56	15.55	15.04	59.35	54.4	56.76	14.83	15	14.91	60.3	59.05	59.66
	Bigrammes	13.91	13.91	13.91	65.6	51.5	57.7	14.21	14.63	14.42	56.85	53.4	55.07
	Adjectifs	18.28	18.73	18.5	69.15	56.05	61.91	17.31	17.31	17.31	57.7	54.15	55.86
	Mots d'émotion	21.35	18.76	19.97	40	50	44.44	42.7	41.26	41.97	81.05	74.55	77.66

TABLE 23 – Comparaisons des résultats obtenus avec NB vs. SVM en utilisant le corpus CAM

Conclusion : En utilisant le corpus CAM, les uni-grammes + les patrons donnent les meilleurs résultats pour les deux algorithmes SVM et NB mais SVM est meilleur (cf. Tableau 24).

Attributs	SVM					
	Multi-Emotion			Pos/Neg		
	P	R	F	P	R	F
Patrons	49	55	55	82	83	83

TABLE 24 – Le meilleur algorithme est le SVM et les patrons sont les meilleurs descripteurs de l'émotion (corpus CAM)

5.4.3 Comparaisons du CAA vs. CAM

Afin d'identifier les différences entre corpus, nous comparons les meilleurs résultats obtenus pour chaque corpus c'est-à-dire SVM avec les mots d'émotion pour le corpus CAA et SVM avec les uni-grammes + les patrons pour le corpus CAM (cf. Tableau 25). Bien que le corpus CAM soit moins volumineux, il obtient les meilleurs résultats. Cela est dû sans doute à un meilleur équilibrage de ce corpus. Toutefois, en allant étudiés un échantillon de phrases mal classées dans le corpus CAA, on se rend compte que le problème est souvent lié à un étiquetage inadéquat des mots du lexique. Ce qui rend à confirmer que l'annotation automatique uniquement est insuffisante.

Pour conclure, comme on s'y attendait la tâche consistant à associer une émotion précise à un message ne donne pas des résultats très concluants : au mieux une F-mesure de 55. Par contre, l'association entre un message et un type d'émotion (positive vs. négative) semble plus facile à réaliser puisque nous obtenons au mieux une F-mesure de 83. Pour conclure cette section, voici les réponses aux questions que nous sommes posées :

- Meilleur algorithme : SVM ;

CAA (Mots d'émotions)						CAM(patrons)					
Multi-Emotion			Pos/Neg			Multi-Emotion			Pos/Neg		
P	R	F	P	R	F	P	R	F	P	R	F
37	33	34	76	65	65	49	55	55	82	83	83

TABLE 25 – Le meilleur algorithme est le SVM et les patrons sont les meilleurs descripteurs de l'émotion (corpus annoté manuellement)

- Meilleur descripteur de l'émotion : les uni-grammes + les patrons ;
- Meilleur corpus : le corpus annoté manuellement.

6 Conclusion et perspectives

Ce document présente une étude bibliographique sur les méthodes d'analyse de sentiments et la démarche adoptée pour analyser les émotions (joie, peur, surprise, *etc.*) dans les forums de santé. Dans la première partie, nous avons décrit les différentes tâches d'analyse de sentiments (analyse de subjectivité, analyse de polarité, analyse d'intensité et analyse d'émotions) ainsi qu'une architecture générale du système d'analyse de sentiments. Dans la partie état de l'art, les ressources et les méthodes couramment utilisées dans l'analyse de sentiments ont été détaillées. Ces dernières se décomposent en méthodes symboliques, statistiques et hybrides. L'état de l'art a été conclu par un tableau récapitulatif des méthodes présentées et une discussion sur les avantages et les inconvénients de chaque méthode.

Dans la deuxième partie, nous avons décrit notre démarche d'analyse d'émotions dans les forums de santé. Celle-ci se décompose en quatre étapes : annotation automatique, annotation manuelle, extraction des émotions et identification de liens entre l'émotion et les objets médicaux. La difficulté principale résidait dans l'acquisition et l'annotation des données. Pour l'extraction des émotions, nous avons utilisé deux algorithmes SVM et NB pour leurs performances reconnues dans cette problématique. Pour identifier les meilleurs descripteurs des émotions, nous avons fait varier plusieurs types d'attributs : bigrammes, adjectifs, patrons, *etc.* Nous avons testé les algorithmes en mode multi-classe et bi-classe. Le meilleur résultat a été obtenu en utilisant l'algorithme SVM en bi-classe avec les uni-grammes et les patrons comme descripteurs et le corpus annoté manuellement.

Une perspective immédiate consiste à trouver le lien entre l'émotion et les objets médicaux discutés dans les messages. Pour cela, un premier travail a été entrepris avec Pierre Pompidor. Il s'agit d'utiliser un moteur de reconnaissance de motifs syntaxiques qui met en lumière des associations entre mots d'une phrase. Par exemple, à partir d'une phrase telle que "*its now (nombr :6 months since that opp the pain is still there plus i now get really bad head aches almost daily*", le moteur est capable de relier bad à head. Or, *bad* est étiqueté comme émotion et *head* est étiqueté comme objet médical. Cette méthode est bien sur

à améliorer car les règles ne sont pas complètes pour l'anglais et trop peu de motifs sont découverts.

A plus long terme, les perspectives associées à ce travail sont nombreuses. En effet, il serait également intéressant de :

- Evaluer l'intensité de l'émotion (forte, moyenne, faible) en utilisant par exemple une liste d'adverbes qui augmentent ou diminuent l'intensité d'émotion comme : "peu, assez, très ou trop" ;
- Prendre en compte la négation. Par exemple, la phrase "je n'ai pas peur de me faire opérer " est annoté par l'émotion peur si on ne traite pas la négation.
- Étudier un corpus en français ;
- Le forum spine-health est plus spécialisé dans la thématique "douleur" et la pathologie discutée est une pathologie de personnes âgées. Afin d'étudier d'autres sentiments, de diversifier les sources de données et les locuteurs il faudra étudier d'autres thématiques ;
- Etudier d'autres phénomènes que l'émotion comme le remerciement, l'encouragement, la félicitation, le conseil, l'insulte, *etc.*
- Plus généralement, la recherche d'émotions peut être utilisée pour modéliser l'évolution des émotions dans le temps. Par exemple, est ce qu'on observe, dans le temps, des changements dans les émotions du même utilisateur de la forme : " surprise puis peur", "surprise puis colère", *etc.* ;
- Etudier l'influence des médias sur les changements des émotions et des avis des patients. Par exemple, récemment il y avait un débat sur la pilule nouvelle génération qui n'avait pas assez d'argument pour un témoin professionnel de santé alors qu'elle a influencé négativement les patientes. En effet, plusieurs femmes ont arrêté de prendre ce type de pilule, ce qui a augmenté par conséquent le taux de grossesse et d'avortement ²⁰.

20. www.terrafemina.com/forme/sante/articles/23088-pilules-dangereuses-le-nombre-divg-a-t-il-augmente-.html

Références

- [1] Erik Boiy, Pieter Hens, Koen Deschacht, and Marie-Francine Moens. Automatic sentiment analysis in on-line text. In *Proceedings of the 11th International Conference on Electronic Publishing*, pages 349–360, Vienna, Austria, 2007.
- [2] François-régis Chaumartin. UPAR7 : A knowledge-based system for headline sentiment tagging. In *SemEval '07 : 4th International Workshop on Semantic Evaluations*, number June, pages 422–425, Prague, Czech Republic, 2007. Association for Computational Linguistics.
- [3] Paul Ekman. An argument for basic emotions. *Cognition and emotion*, 6(3-4) :169–200, 1992.
- [4] Andrea Esuli and Fabrizio Sebastiani. Sentiwordnet : A publicly available lexical resource for opinion mining. *Proceedings of the 5th Conference on Language Resources and Evaluation (LREC'06)*, pages 417–422, 2006.
- [5] Kilem Gwet and D Ph. Inter-Rater Reliability : Dependency on Trait Prevalence and Marginal Homogeneity. (2) :1–9, 2002.
- [6] Mark Hall, Eibe Frank, Geoffrey Holmes, Bernhard Pfahringer, Peter Reutemann, and Ian H. Witten. The weka data mining software : an update. *SIGKDD Explor. Newsl.*, 11(1) :10–18, November 2009.
- [7] Ali Harb, Michel Plantié, Gerard Dray, Mathieu Roche, François Troussel, and Pascal Poncelet. Web Opinion Mining : How to extract opinions from blogs? Categories and Subject Descriptors. In *CSTST'08 : International conference on Soft Computing as Transdisciplinary Science and Technology*, 2008.
- [8] Vasileios Hatzivassiloglou and Kathleen R. McKeown. Predicting the semantic orientation of adjectives. In *ACL '98 Proceedings of the 35th Annual Meeting of the Association for Computational Linguistics and Eighth Conference of the European Chapter of the Association for Computational Linguistics*, pages 174–181, Stroudsburg, PA, USA, 1997. Association for Computational Linguistics.
- [9] Carroll E. Izard. *Human Emotions*. New York, NY, USA, plenum pre edition, 1977.
- [10] Jaap Kamps, Maarten Marx, Robert J. Mokken, and Maarten De Rijke. Using wordnet to measure semantic orientations of adjectives. In *Proceedings of the 4th International Conference on Language Resources and Evaluation LREC 2004, IV*, pages 1115–1118, Paris, France, 2004. European Language Resources Association.
- [11] Dan Klein and Christopher D. Manning. Fast exact inference with a factored model for natural language parsing. *Advances in Neural Information Processing Systems 15 (NIPS)*, 2002.
- [12] Zornitsa Kozareva, Borja Navarro, Sonia Vazquez, and Andrés Montoyo. UA-ZBSA : a headline emotion classification through web information. In

RÉFÉRENCES

- SemEval '07 Proceedings of the 4th International Workshop on Semantic Evaluations*, number June, pages 334–337, Stroudsburg, PA, USA, 2007. ACL.
- [13] Bing Liu. *Sentiment Analysis and Opinion Mining*, volume 5. Morgan & Claypool Publishers, May 2012.
- [14] Cheng-yu Lu, Jen-shin Hong, and Samuel Cruz-lara. Emotion Detection in Textual Information by Semantic Role Labeling and Web Mining Techniques. In *Third Taiwanese-French Conference on Information Technology - TFIT 2006*, Nancy, France, 2006.
- [15] Peter Mark. *Roget's Thesaurus of English Words and Phrases*. Available from Project Gutenberg, Illinois Benedictine College, Lisle IL (USA), 1852.
- [16] George A. Miller. WordNet : a lexical database for English. *Communications of the ACM*, 38(11) :39–41, November 1995.
- [17] Saif M. Mohammad and Peter D. Turney. Emotions Evoked by Common Words and Phrases : Using Mechanical Turk to Create an Emotion Lexicon. In *CAAGET '10 Proceedings of the NAACL HLT 2010 Workshop on Computational Approaches to Analysis and Generation of Emotion in Text*, pages 26–34, Stroudsburg, PA, USA, 2010. ACL.
- [18] Matthijs Mulder, Anton Nijholt, Marten Den Uyl, and Peter Terpstra. A lexical grammatical implementation of affect. In *7th International Conference on Text, Speech & Dialogue (TSD 2004)*, pages 171–177, Heidelberg, 2004. Springer Verlag.
- [19] Bo Pang, Lillian Lee, and Shivakumar Vaithyanathan. Thumbs up? : sentiment classification using machine learning techniques. In *EMNLP '02 Proceedings of the ACL-02 conference on Empirical methods in natural language processing - Volume 10*, volume 10, pages 79–86, Stroudsburg, PA, USA, 2002. ACL.
- [20] A.K. Pujari. *Data Mining Techniques*. Universities Press, 2001.
- [21] Carlo Strapparava and Reda Mihalcea. Learning to identify emotions in text. In *SAC '08 Proceedings of the 2008 ACM symposium on Applied computing*, pages 1556–1560, New York, NY, USA, 2008. ACM.
- [22] Carlo Strapparava and Alessandro Valitutti. WordNet-Affect : an affective extension of WordNet. In *Proc. of 4th International Conference on Language Resources and Evaluation*, pages 1083–1086, Lisbon, 2004.
- [23] Sabiha Tahrat, Eric Kergosien, and Sandra Bringay. Text2Geo : des données textuelles aux informations géospatiales Text2Geo : Vers un nouveau processus d' extraction d' information spatiale. (2) :2–7.
- [24] Peter D. Turney and Michael L. Littman. Measuring praise and criticism : Inference of semantic orientation from association. *ACM Transactions on Information Systems (TOIS)*, 21(4) :315–346, 2003.
- [25] Cynthia Whissell. *The dictionary of affect in language*. Academic Press, 1989.

- [26] Janyce Wiebe, T. Wilson, and C Cardie. Annotating expressions of opinions and emotions in language. *Language Resources and Evaluation*, 2005.
- [27] Janyce M Wiebe. Learning Subjective Adjectives from Corpora. (1), 1982.
- [28] Theresa Wilson, Janyce Wiebe, and Paul Hoffmann. Recognizing contextual polarity in phrase-level sentiment analysis. In *In Proceedings of Human Language Technologies Conference/Conference on Empirical Methods in Natural Language Processing (HLT/EMNLP 2005)*, pages 347–354, Morristown, NJ, USA, 2005. Association for Computational Linguistics.

A Résultats complets

A.1 Corpus CAA

Dans cette partie nous présentons les résultats obtenus avec le corpus CAA. Le tableau 26 représente les résultats obtenus avec NB pour les différentes catégories émotionnelles. Nous remarquons que les mots d’émotions donnent les meilleurs résultats pour les deux émotions peur et tristesse. Les uni-grammes sans lemmatisation, les méta-attributs et les patrons donnent les meilleurs résultats pour les deux émotions joie, colère et dégoût. Enfin les adjectifs donnent les meilleurs résultats pour la surprise.

Pour tous les types d’attributs et les mesures d’évaluation, les meilleurs résultats sont obtenus pour les émotions tristesse, peur et joie car elles sont les mieux représentées dans le corpus (35%, 33% et 13% respectivement).

Le tableau 27 montre les performances de NB pour émotions positives vs. négatives. Les uni-grammes avec et sans lemmatisation, l’étiquetage grammaticale (pos), les méta-attributs, l’émotion globale et les patrons donnent le meilleur rappel et le meilleur F-score pour les émotions positives (joie) et la meilleure précision pour les émotions négatives. Inversement, les mots d’émotion donnent le meilleur rappel et le meilleur F-score pour les émotions négatives et la meilleure précision pour les émotions positives.

Le tableau 28 montre les performances de SVM pour les six catégories d’émotion. Les mots d’émotion donnent les meilleurs résultats pour les émotions joie, peur, colère et dégoût. Les meilleurs résultats pour l’émotion tristesse sont obtenus avec l’émotion globale. De même, les émotions tristesse, peur et joie obtiennent toujours les meilleurs résultats car elles représentent les classes majoritaires du corpus.

Le tableau 29 montre les performances de SVM pour émotions positives vs. négatives. Les mots d’émotions donnent les meilleurs résultats pour les émotions positives et la meilleure précision pour les émotions négatives qui obtiennent leurs meilleurs rappels et le meilleur F-scores avec les bigrammes et les adjectifs.

A RÉSULTATS COMPLETS

Attributs		Joie			surprise			peur		
		P	R	F	P	R	F	P	R	F
Uni-grammes	sans lemmatisation	27.5	25.2	26.5	36	3.9	7.1	34.1	39.4	36.6
	avec lemmatisation	27.5	23.6	25.4	27.3	3.9	6.8	36	44.1	39.7
	+POS	22.9	19.8	21.2	25.7	3.9	6.8	33.4	39.9	36.4
	+Méta-attributs	27.7	25.1	26.3	36	3.9	7.1	34	39.3	36.5
	+Contexte d'émotion	26.3	24.3	25.3	36	3.9	7.1	39.6	47.6	43.2
	+patrons	28	25.2	26.6	36	3.9	7.1	33.8	39	36.2
Bigrammes		20.8	9.8	13.3	23.7	3.9	6.8	30	41.5	34.8
Adjectifs		30.9	13.5	18.8	37	4.4	7.8	33	51.6	40.2
Mots d'émotion		52.8	3.9	7.3	0	0	0	39.8	51.1	44.7
Attributs		colère			dégout			tristesse		
		P	R	F	P	R	F	P	R	F
Uni-grammes	sans lemmatisation	17.5	11.2	13.7	9.8	2.8	4.3	37.4	45	40.8
	avec lemmatisation	14.3	8.3	10.5	8.5	2.4	3.8	38.6	45.1	41.6
	+POS	11.8	7.6	9.3	2.7	0.7	1.1	36	42.3	38.9
	+Méta-attributs	18.4	11.9	14.4	9.5	2.8	4.3	37.1	44.5	40.4
	+Contexte d'émotion	16.8	11.2	13.5	9	2.8	4.3	46.8	53.4	49.9
	+patrons	17.6	11	13.5	9.4	2.8	4.3	37.3	45.1	40.9
Bigrammes		22	8.3	12	0	0	0	23.3	39.9	35.7
Adjectifs		9.1	.7	2.9	3.7	0.3	0.6	36.3	41.4	38.7
Mots d'émotion		0	0	0	0	0	0	42.2	68.5	52.2

TABLE 26 – Performances de NB pour les différentes catégories émotionnelles (corpus CAA)

Attributs		Positive			Négative		
		P	R	F	P	R	F
Uni-grammes	sans lemmatisation	26.9	20.1	23	87.4	91.1	89.2
	avec lemmatisation	28.1	20.6	23.8	87.5	91.3	89.4
	+POS	25.2	19.5	22	87.3	90.5	88.9
	+Méta-attributs	26.4	19.8	22.6	87.4	90.9	89.1
	+Contexte d'émotion	27.2	20.1	23.1	87.4	91.2	89.3
	+patrons	27.5	20.6	23.6	87.5	91.1	89.3
Bigrammes		26.4	5.2	8.7	86.3	97.6	91.6
Adjectifs		29.9	9.4	14.3	86.6	96.4	91.3
Mots d'émotion		64.3	1.3	2.5	86	99.9	92.5

TABLE 27 – Performances de NB pour émotions positives vs. négatives (Corpus CAA)

A.2 Corpus CAM

Dans cette partie nous présentons les résultats obtenus avec le corpus CAM. Le tableau 30 représente les résultats obtenus avec NB pour les différentes catégories émotionnelles. Nous remarquons que les patrons donnent les meilleurs

A RÉSULTATS COMPLETS

Attributs		Joie			surprise			peur		
		P	R	F	P	R	F	P	R	F
Uni-grammes	sans lemmatisation	27.3	27.3	27.3	9.6	5.3	6.8	26.5	29.4	27.9
	avec lemmatisation	26.9	26.8	26.8	12.6	6.6	8.6	28	30.4	29.1
	+POS	25.8	24.7	25.2	17.9	16.9	17.4	29.6	34.4	31.8
	+Méta-attributs	26	25.1	25.6	13.3	8.6	10.4	29.1	30.5	29.8
	+Contexte d'émotion	27.2	28.1	27.6	15.8	9.2	11.6	42.3	48	45
	+patrons	25.6	24.7	25.1	12.8	8.6	10.3	28.5	29.7	29.1
	Bigrammes	23.9	6.3	10	100	3.9	7.6	28.8	53.9	37.5
Adjectifs		30.1	28.3	29.2	30	3.9	7	33.1	51.1	40.2
Mots d'émotion		46.6	43	44.8	25	11.4	15.7	43.9	54.9	48.8
Attributs		colère			dégout			tristesse		
		P	R	F	P	R	F	P	R	F
Uni-grammes	sans lemmatisation	17.6	14.4	15.9	13.8	9.4	11.2	30.4	31.8	31.1
	avec lemmatisation	16.1	13.6	14.7	17.9	12.2	14.5	30.4	32.4	31.4
	+POS	21.6	20.1	20.8	12.8	6	8.2	33	32.5	32.8
	+Méta-attributs	18.6	16.2	17.3	16.7	11	13.3	34.8	38.6	36.6
	+Contexte d'émotion	21.9	18.6	20.1	13	9.4	10.9	63.9	63.4	63.6
	+patrons	18.3	16.2	17.1	18.8	12	14.6	35	38.7	36.8
	Bigrammes	12.5	1.7	3	0	0	0	31.3	329.53	30.4
Adjectifs		3.8	0.4	0.8	11.1	2.1	3.5	36.1	34.9	35.5
Mots d'émotion		31.5	26.9	29	32.2	19	23.9	48.2	47	47.6

TABLE 28 – Performances de SVM pour les différentes catégories émotionnelles (corpus CAA)

Attributs		Positive			Négative		
		P	R	F	P	R	F
Uni-grammes	sans lemmatisation	37.5	24.5	29.7	88.3	93.3	90.7
	avec lemmatisation	40.4	30.8	34.9	89.2	92.7	90.9
	+POS	35.2	26.7	30.4	88.2	91.8	90
	+Méta-attributs	38.6	30.3	34	89.1	92.2	90.6
	+Contexte d'émotion	37.7	29.9	33.3	89	92	90.5
	+patrons	37.3	24.4	29.5	88.3	93.3	90.7
	Bigrammes	32.4	1.7	3.2	86	99.4	92.3
Adjectifs		33.3	2.9	5.4	86.2	99	92.1
Mots d'émotion		63.2	35.1	45.1	90.2	96.7	93.4

TABLE 29 – Performances de SVM pour émotions positives vs. négatives(Corpus CAA)

résultats sauf pour la précision de l'émotion joie et le rappel de l'émotion dégoût dont les meilleurs résultats sont obtenus avec les mots d'émotions. De même pour NB pour émotions positives vs. négatives(tableau 31), les patrons donnent toujours les meilleurs résultats sauf pour le rappel des émotions négatives dont le meilleur résultat est aussi obtenu avec les mots d'émotions.

Pour SVM avec multi-classification, les patrons donnent les meilleurs résultats pour les six catégories d'émotion (cf. Tableau 32). Nous remarquons égale-

A RÉSULTATS COMPLETS

Attributs		Joie			surprise			peur		
		P	R	F	P	R	F	P	R	F
Uni-grammes	sans lemmatisation	28.9	28	28.4	16.7	7.2	10.1	13.5	10.1	11.6
	avec lemmatisation	26.6	10.5	15.1	16.1	6	8.8	16.3	11.3	13.2
	+POS	30.3	33	31.6	11.8	7.2	9	13	9.1	10.7
	+Méta-attributs	29.7	30	29.9	16.7	7.2	10.1	12.2	9.1	10.4
	+Contexte d'émotion	29	29	29	13.5	6	8.3	11.1	8.1	9.4
	+patrons	28	28	28	11.8	4.8	6.8	11.6	8.1	9.5
Bigrammes		21.9	7	10.6	33.3	2.4	4.5	4.2	1	1.6
Adjectifs		39.4	28	32.7	16.7	2.4	4.2	18.9	10.1	13.2
Mots d'émotion		72.7	4	7.6	0	0	0	0	0	0

Attributs		colère			dégout			tristesse		
		P	R	F	P	R	F	P	R	F
Uni-grammes	sans lemmatisation	6.7	2.7	3.8	21.4	41.1	28.1	6	6	6
	avec lemmatisation	5.3	2	2.9	20.8	56.8	30.5	9	6.5	7.5
	+POS	2.2	1.4	1.7	24.6	37.5	29.7	9.2	10.3	9.7
	+Méta-attributs	7.7	2.7	4	22.9	44.2	30.2	7	6.9	6.9
	+Contexte d'émotion	6.5	2.7	3.8	23	43.4	30	7.7	7.8	7.7
	+patrons	6.7	2.7	3.8	21.8	41.9	28.6	7.5	7.8	7.6
Bigrammes		0	0	0	19.5	70.5	30.6	4.3	2.6	3.2
Adjectifs		0	0	0	22.9	58.1	32.9	11.8	13.8	12.7
Mots d'émotion		0	0	0	22.1	92.2	35.7	33.3	16.4	22

TABLE 30 – Performances de NB pour les différentes catégories émotionnelles (corpus CAM)

Attributs		Positive			Négative		
		P	R	F	P	R	F
Uni-grammes	sans lemmatisation	35	14	20	82	93.8	87.5
	avec lemmatisation	32.8	19	24.1	82.4	90.7	86.3
	+POS	32	16	21.3	82	91.8	86.7
	+Méta-attributs	36.5	15	21.3	82.2	93.9	87.6
	+Contexte d'émotion	36.5	15	21.3	82.2	93.8	87.6
	+patrons	36.5	15	21.3	82.2	93.8	87.6
Bigrammes		50	4	7.4	81.2	99	89.2
Adjectifs		55.6	15	23.6	82.7	97.1	89.3
Mots d'émotion		0	0	0	80.7	100	89.3

TABLE 31 – Performances de NB pour émotions positives vs. négatives (Corpus CAM)

ment que les mots d'émotions donnent aussi des bons résultats. Le tableau 33 montre les performances de SVM pour émotions positives vs. négatives. Nous remarquons que les patrons et les mots d'émotions donnent les meilleurs résultats.

A RÉSULTATS COMPLETS

Attributs		Joie			surprise			peur		
		P	R	F	P	R	F	P	R	F
Uni-grammes	sans lemmatisation	29.6	32	30.8	19.2	18.1	18.6	16	16.2	16.1
	avec lemmatisation	27.3	13.5	18.1	19.2	8.4	11.7	11.7	6.1	8
	+POS	23.7	23	23.4	17.3	15.7	16.5	13.9	14.1	14
	+Méta-attributs	27.9	31	29.4	16.2	13.3	14.6	15.2	15.2	15.2
	+Contexte d'émotion	24.2	30	26.8	17.1	15.7	16.4	22.5	23.2	22.9
	+patrons	29.1	32	30.5	17.6	15.7	16.6	13	13.1	13.1
Bigrammes		27	24	25.4	20	10.8	14.1	9.8	6.1	7.5
Adjectifs		37.9	36	36.9	13	7.2	9.3	14.1	12.1	13
Mots d'émotion		54.2	57.5	55.8	50.3	47.6	48.9	35	32.3	33.6
Attributs		colère			dégout			tristesse		
		P	R	F	P	R	F	P	R	F
Uni-grammes	sans lemmatisation	4.9	5.4	5.1	17.6	17.8	17.7	6.9	6	6.4
	avec lemmatisation	6.8	2.7	3.9	20.7	59.9	30.8	4.8	2.6	3.4
	+POS	7.6	6.8	7.1	19.2	19.5	19.4	9.2	10.3	9.7
	+Méta-attributs	5.5	6.8	6.1	18.7	19.4	19	5.1	4.3	4.7
	+Contexte d'émotion	8.1	8.1	8.1	18.5	17.1	17.7	17	15.5	16.2
	+patrons	4.9	5.4	5.1	17.8	17.8	17.8	6.6	6	6.3
Bigrammes		4	1.4	2	18.6	40.3	25.4	5.9	5.2	5.5
Adjectifs		5.6	2.7	3.6	21.1	28.7	24.3	12.2	17.2	14.3
Mots d'émotion		46.9	31.1	37.4	39.9	48.1	43.6	29.9	31	30.4

TABLE 32 – Performances de SVM pour les différentes catégories émotionnelles (corpus CAM)

Attributs		Positive			Négative		
		P	R	F	P	R	F
Uni-grammes	sans lemmatisation	37.2	29	32.6	83.9	88.3	86
	avec lemmatisation	36.1	15	21.2	82.1	93.7	87.5
	+POS	29.6	21	24.6	82.3	88	85.1
	+Méta-attributs	36	31	33.3	84	86.8	86.5
	+Contexte d'émotion	36.1	30	32.8	83.9	87.3	85.6
	+patrons	36.5	31	33.5	84.1	87.1	85.5
Bigrammes		31.8	14	19.4	81.9	92.8	87
Adjectifs		33.3	16	21.6	82.1	92.3	86
Mots d'émotion		72.5	54	61.9	89.6	95.1	92.3

TABLE 33 – Performances de SVM pour émotions positives vs. négatives (Corpus CAM)