



RAPPORT DE STAGE, DÉPARTEMENT GÉNIE INFORMATIQUE (TN10)

Identification automatique des émotions dans des récits narratifs en combinant psychologie et apprentissage profond

Stage de recherche au sein du Laboratoire des Méthodes Formelles (Université Paris-Saclay, CNRS, ENS Paris-Saclay) du 7 février au 22 juillet 2022

école —————
normale —————
supérieure —————
paris-saclay —————

université
PARIS-SACLAY



Adresse : 4 Av. des Sciences, 91190 Gif-sur-Yvette
Auteur : Gustave Cortal

Suiveur entreprise : Alain Finkel
Suiveur UTC : Domitille Lourdeaux

Semestre de printemps 2022

Remerciements

Pour mon stage d'ingénieur de recherche, j'ai eu l'opportunité de travailler au sein du Laboratoire des Méthodes Formelles de l'École Normale Supérieure de Paris-Saclay. Je tiens à remercier Alain pour m'avoir fait confiance et m'avoir permis de travailler sur la méthode qu'il a développée, l'Analyse Cognitive des Émotions. Alain était très souvent présent pour me conseiller sur les différents aspects de la recherche, notamment sur la méthodologie, les objectifs et les enjeux de celle-ci. Je remercie Patrick pour sa supervision dans le domaine du traitement automatique des langues et pour son expertise sur l'analyse du sentiment et la fouille d'opinion. Je remercie Lina pour sa supervision sur le côté apprentissage automatique de ce stage. Les compétences complémentaires de mes superviseurs ont permis une agréable communication des idées provenant de la psychologie cognitive et du traitement automatique des langues. Je tiens à souligner la part importante qui est attribuée à la créativité et l'autonomie, ce qui m'a permis de m'épanouir pleinement durant le stage. C'est avec plaisir que je continuerai à travailler avec eux en effectuant un doctorat sous leur supervision. Je remercie Domitile Lourdeaux, ma tutrice UTC, pour son encadrement durant le stage et pour avoir pris le temps de venir dans les locaux du laboratoire.

Sommaire

Résumé technique	3
1 Présentation des domaines de recherche	4
1.1 Analyse Cognitive des Émotions	4
1.2 Analyse du sentiment	8
1.3 Identification des causes d'une émotion	10
2 Jeu de données composé de scènes émotionnelles	15
2.1 Construction du jeu de données	15
2.2 Modèles de classification	20
3 Identification automatique des émotions et de leurs rôles sémantiques	26
3.1 Jeu de données français pour l'analyse des émotions	26
3.2 Nouveau schéma d'annotation	26
3.3 Méthode basée sur les règles	27
3.4 Structure de graphe	28
Conclusion	30
Acronymes	41

Résumé technique

L'analyse des émotions dans les textes souffre de deux limitations majeures : les corpus de référence sont pour la plupart petits et homogènes, et l'identification des émotions est souvent simplifiée comme un problème de classification au niveau de la phrase. Pour résoudre ces problèmes, nous présentons un nouveau schéma d'annotation pour explorer les émotions et leurs raisons, ainsi qu'un nouveau jeu de données en français composé de récits autobiographiques d'une scène émotionnelle. Les textes ont été recueillis en appliquant l'Analyse Cognitive des Émotions développée par Alain Finkel pour aider les personnes à améliorer leur gestion des émotions. La méthode nécessite l'analyse manuelle d'un événement émotionnel par un entraîneur formé à l'Analyse Cognitive.

Après avoir présenté l'Analyse Cognitive des Émotions et les différents axes de recherche qui nous intéressent, nous montrons les différentes étapes pour la construction d'un jeu de données en français pour l'analyse des émotions. Ensuite, nous proposons un nouveau schéma d'annotation ainsi qu'une approche basée sur des règles pour annoter automatiquement les émotions et leurs rôles sémantiques (par exemple, les raisons de l'émotion) afin de faciliter l'identification des aspects pertinents par l'entraîneur. Pour finir, nous étudions les orientations futures pour l'analyse des émotions à l'aide d'une nouvelle structure de graphe.

Chapitre 1

Présentation des domaines de recherche

1.1 Analyse Cognitive des Émotions

1.1.1 Présentation

L'Analyse Cognitive des Emotions (ACE) (FINKEL, 2022) est une théorie des émotions discrètes puisqu'elle étudie comment les quatre émotions primaires (joie, tristesse, colère et peur) apparaissent dans des récits autobiographiques décrivant de brèves scènes avec les émotions ressenties par l'auteur. L'ACE est présentée dans un récent manuel². Elle fournit une méthode pour comprendre les émotions et permet de modifier les représentations négatives d'une scène émotionnelle. Elle analyse le récit d'une scène catégorisée selon la grille Faits, Émotions, Raisons, Actions (FERA), qui sera présentée dans la prochaine section. Le but de l'analyse est d'identifier les faits, les interprétations, les émotions et les raisons des émotions. Grâce à cela, il est possible de suggérer de nouvelles actions qui auraient pu être effectuées suite aux émotions vécues. La méthode permet d'améliorer sa santé psychologique puisqu'elle permet aux personnes de mieux réguler leurs émotions.

L'ACE considère qu'une émotion a pour fonction la survie de l'individu en favorisant les meilleures actions à prendre. Plus particulièrement, une émotion arrivant à notre conscience est un message pour résoudre un problème associé à cette émotion. La douleur a pour objet de nous protéger en fermant nos frontières alors que le plaisir les ouvre et nous oriente vers le monde, tous deux veillant à notre survie. Une hypothèse fondamentale de l'ACE est de comprendre les émotions comme un raffinement du plaisir et de la douleur. Ainsi, nous obtenons des comportements plus complexes que les comportements basiques « aller vers » et « éviter ».

L'ACE est une théorie de l'évaluation cognitive (LAZARUS & FOLKMAN, 1984). Elle considère qu'une émotion survient lorsque nous évaluons une situation qui infirme ou confirme notre modèle du monde. Notre modèle du monde est composée de nos territoires et objets importants qui sont associés à nos besoins à satisfaire.

1.1.2 Objet, territoire et fait

Les problèmes universels, signalés par les quatre émotions primaires, sont formalisés à travers les notions de territoire et d'objet. Les territoires sont multiples, changent avec le temps et varient selon les individus. La genèse et le développement d'un territoire dépendent du milieu associé à l'individu, comme le contexte culturel, social et familial. Les territoires d'un individu peuvent être

2. <https://www.dunod.com/sciences-humaines-et-sociales/manuel-d-analyse-cognitive-emotions-theorie-et-applications>

son corps, son logement, ses valeurs professionnelles, ses croyances, son image de soi, sa fiabilité pour une tâche, son temps, sa liberté de mouvement, son droit à exister, etc. La notion d'objet est prise au sens large et correspond à une partie d'un territoire comme une certaine image de soi, une croyance ou une personne.

La colère et la peur sont des signaux qu'au moins un de nos territoires est attaqué. C'est à nous de défendre (colère) ou de fuir (peur) par rapport à l'attaquant (qui peut être autrui, une partie de soi, etc.). La joie et la tristesse sont des signaux indiquant qu'un changement s'est produit dans notre lien avec un territoire. La tristesse arrive chez un sujet qui évalue une situation comme représentant une séparation avec un objet important de son monde. La joie est le résultat de l'évaluation d'une situation qui représente l'acquisition ou le renforcement d'un objet important. Par exemple, je peux être joyeux si j'obtiens une mention à mon master, car mon objet important *image de soi positive* sera renforcé. Si je décide de travailler à l'étranger, je peux être triste, car je serai physiquement séparé de mon objet important *famille*.

Les territoires et les objets sont associés à des besoins humains qui doivent être satisfaits. Par exemple, selon la pyramide de Maslow (MASLOW, 1943), les êtres humains ont des besoins physiologiques, de sécurité, d'amour et d'appartenance, d'estime et d'accomplissement de soi.

Un fait est un énoncé portant sur des objets et des actions présents dans une scène. Un fait peut se passer dans le monde extérieur (perception) ou dans le monde intérieur (évocation). Un enjeu important de l'ACE est de distinguer les faits des interprétations.

1.1.3 Système de traitement de l'information émotionnelle

Similaire à de nombreuses théories en psychologie (par exemple, la psychanalyse de Freud, la Gestalt-thérapie de Perl, la thérapie centrée sur les émotions de Greenberg et la plupart des théories en psychologie sur les émotions comme la théorie de l'évaluation cognitive), l'ACE considère que l'esprit, dans une scène donnée, traite les émotions et les cognitions associées selon un cycle. La plupart du temps, ce processus n'est pas conscient et commence avec l'identification d'une situation et de ses enjeux. Puis, il est suivi par une réflexion concernant les bénéfices et les désavantages des possibles choix d'actions. Une décision est prise et l'action choisie est exécutée. Pour finir, le cycle se termine avec un retour à un état de préparation pour traiter une prochaine scène.

L'ACE est une théorie computationnelle qui considère que l'esprit est une boucle sans fin qui traite les différents objets de notre monde dans le but de satisfaire ses besoins. Il arrive que l'esprit ne traite pas correctement l'information et boucle sur des éléments qu'il n'arrive pas à digérer. L'ACE permet de faciliter le traitement de l'information émotionnelle en envisageant des actions possibles pour débloquer des situations. Souvent, il faudra au préalable modifier son modèle du monde (c.-à-d. sa vision de soi, des autres et du monde).

Selon (FRISTON, 2010), l'esprit compare constamment les sensations du monde perçues et évoquées avec son propre modèle du monde afin de minimiser les surprises. L'esprit vérifie en permanence si son modèle du monde est efficace pour prédire et expliquer correctement les sensations qu'il perçoit. Il est satisfait si son modèle du monde coïncide avec le monde perçu. Néanmoins, si l'écart est trop important, l'esprit est insatisfait et essaie de modifier son modèle du monde et le monde qu'il perçoit. L'esprit est donc considéré comme un système de traitement de l'information complexe où l'*a posteriori* devient *a priori*. Son modèle du monde ainsi que ses objectifs sont en devenir. L'ACE permet d'aider les personnes à mieux réguler leur émotion en proposant de nouvelles manières de voir le monde ainsi que de nouvelles actions pour agir dans le monde.

1.1.4 Exemple

Prenons un exemple simple. Je suis dans la file d'attente pour prendre un ticket à la caisse d'un cinéma lorsque quelqu'un me double. Je me sens en colère puisque je pense que je ne suis pas respecté, ni les règles sociales. J'évalue mes actions possibles et leurs conséquences : protester

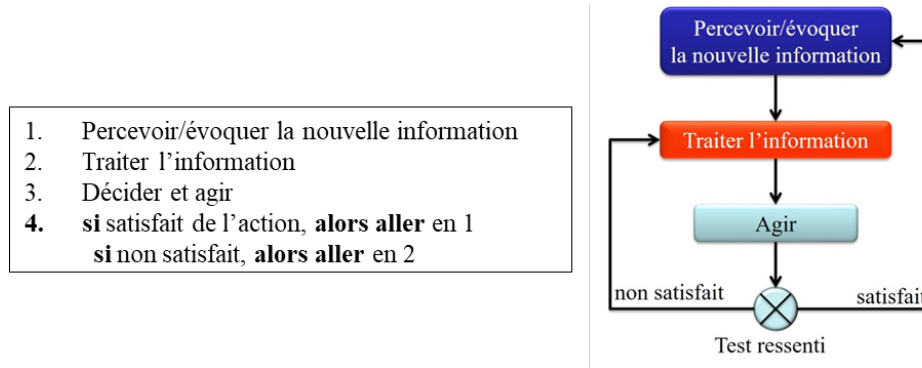


FIGURE 1.1: L'esprit comme un système de traitement de l'information émotionnelle. Image provenant de (FINKEL, 2022).

verbalement, pousser la personne physiquement, ne rien faire ou s'enfuir. Comme ma peur d'un conflit est plus grande que ma colère, je décide de rester silencieux et de ne rien faire.

Le conflit que j'ai évité dans le monde extérieur est intériorisé dans mon esprit. Je suis peut-être en colère contre la partie de moi-même qui ne défend pas ses droits, ou je suis peut-être triste, car je me retrouve séparé de ma vision d'un monde équitable. Dans cet exemple, le cycle de traitement de l'émotion ne s'est pas correctement exécuté comme il le faudrait. Je reste mentalement préoccupé après la scène. J'ai des regrets et des doutes. Je rejoue mentalement la scène différemment.

La scène précédente met en scène une attaque contre mes deux territoires suivants : mon temps libre (contraint par la file d'attente du cinéma) et ma vision rassurante d'un monde organisé et prévisible avec des lois acceptées par presque tout le monde.

Dans ce rapport, nous proposons de comprendre les émotions et leurs raisons en identifiant automatiquement les territoires et objets pertinents impliqués dans une scène émotionnelle.

1.1.5 Grille FERA

Dans une session ACE, les personnes qui souhaitent mieux gérer leurs émotions écrivent un récit autobiographique d'une scène émotionnelle passée qu'elles ont vécue, dans un lieu et un temps donnés, avec des personnages identifiés. L'entraîneur ACE impose des instructions pour l'écriture des scènes. Ces instructions représentent des métadonnées importantes pour la construction d'un jeu de données d'analyse des émotions. Nous décrivons dans le chapitre 2 la construction d'un tel jeu de données. La grille Faits, Émotions, Raisons, Actions (FERA) structure les récits autobiographiques. Elle est composée de quatre parties majeures :

- *Faits* décrit les faits qui sont observés par tout le monde dans la scène. Cette partie inclut également les pensées et les ressentis expérimentés par l'auteur, car les événements internes ne sont pas observables et sont considérés comme vrai, c.-à-d. non réfutable.
- *Émotions* identifie les émotions expérimentées par l'auteur. Notons que les émotions observées des autres participants sont considérées comme des *Faits*.
- *Raisons* identifie les raisons de l'émotion par rapport à la théorie des territoires de l'ACE. Les territoires et objets pertinents pour l'auteur sont identifiés.
- *Actions* analyse les actions passées, rejoue mentalement la scène dans le présent et teste les possibles actions pour le futur. Le critère pour stopper la grille FERA est de trouver les actions adaptées à la situation.

Notons qu'au sein des faits, il existe les faits extérieurs (objets perçus) et les faits intérieurs (objets évoqués). Les faits intérieurs sont composés des ressentis et des pensées de l'auteur de la grille. Contrairement aux faits extérieurs, ils sont souvent peu visibles de l'extérieur.

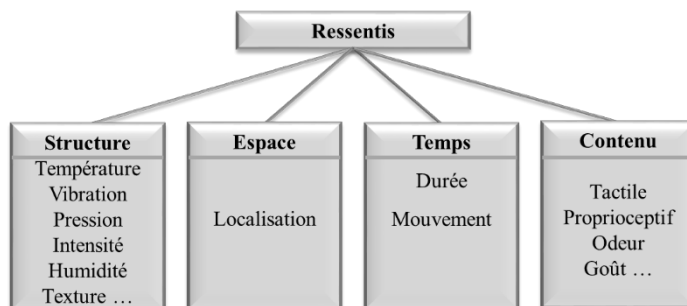


FIGURE 1.2: Description des ressentis par FINKEL (2022).

Un exemple d'une grille FERA complétée est disponible en Annexe (Figure 3.1). Nous verrons, dans le chapitre 2, que le format des grilles a changé au cours du temps ainsi que les consignes données. Néanmoins, peu importe le format, les quatre parties majeures représentent la colonne vertébrale de ces grilles.

Au sein des quatre parties majeures, nous trouvons des catégories qui seront importantes pour la construction de notre jeu de données. Elles sont les métadonnées que nous voulons extraire pour chaque grille. Cela nous permettra de relier pour chaque réponse d'un auteur d'une grille sa catégorie correspondante. Voici la description des différentes catégories :

- FAIT EXTÉRIEUR : l'auteur décrit les comportements observables de lui-même et des autres. Ils sont identifiés en répondant aux deux questions : « qui a fait (ou n'a pas fait) quoi (avec son corps) ? » et « qui a dit (ou n'a pas dit) quoi ? ». Les faits extérieurs sont captés par nos sens et sont les comportements verbaux, non verbaux, physiologiques et de préparation à l'action. Les déductions mentales sont prosrites, il faut seulement s'attacher aux perceptions sensorielles sans essayer d'interpréter.
- FAIT INTÉRIEUR RESENTIS : l'auteur décrit les sensations physiques de son corps, qui peuvent être localisée. Une description des ressentis physiques est illustrée figure 1.2.
- FAIT INTÉRIEUR PENSÉES : l'auteur rapporte ce qu'il a pensé durant la scène émotionnelle. Ces pensées contiennent souvent des interprétations qui sont intéressantes pour l'analyse, car celles-ci nous informent sur son modèle du monde.
- VÉRIFICATION FAIT : l'auteur analyse si ce qu'il a écrit précédemment est factuel ou non.
- IDENTIFICATION ÉMOTION : l'auteur identifie une des quatre émotions primaires qu'il a vécue.
- TERRITOIRE : l'auteur décrit, selon l'émotion vécue, les besoins non satisfaits, les territoires attaqués ainsi que les objets acquis ou dont on a été séparé.
- ACTION INTÉRIEURE : l'auteur analyse les actions qu'il a effectuées dans la scène passée et propose éventuellement une mise à jour de son modèle du monde à travers les nouvelles ressources qu'il a acquises.
- ACTION EXTÉRIEURE : l'auteur teste les actions possibles qu'il n'avait pas envisagées auparavant ou qu'il s'était interdit de faire.
- ACTION TEST : l'auteur décrit le résultat de ces nouvelles actions.

L'entraîneur ACE aide l'auteur à identifier, à partir de l'analyse guidée de son récit, les territoires et objets pertinents qui sont en jeu dans une scène émotionnelle. L'analyse de l'entraîneur vise à mieux comprendre les raisons des émotions et à proposer des actions correctives pour mieux gérer les situations. Dans le chapitre 3, nous décrivons le modèle que nous avons développé pour identifier automatiquement les rôles sémantiques (comme le SUJET, TERRITOIRE, OBJET, etc.) dans un texte. La solution proposée vise à automatiser une étape importante de l'analyse, à savoir l'identification des émotions et de leurs raisons. Nous présentons dans la prochaine section une tâche de TAL importante pour nos objectifs, l'analyse du sentiment.

1.2 Analyse du sentiment

1.2.1 Présentation

Depuis son introduction par PANG et al. (2002) il y a deux décennies, l'analyse du sentiment (ou fouille d'opinion) est devenu un champ influent de la recherche en Traitement Automatique des Langues (TAL) avec de nombreuses applications dans l'industrie (comme l'analyse du sentiment du marché en finance ou l'analyse des réactions émotionnelles des clients face à un produit en marketing). Néanmoins, la majorité des recherches sur l'analyse du sentiment considère seulement la tâche de catégorisation du texte (PORIA et al., 2020), c.-à-d. classifier en deux ou trois catégories les sentiments (positif, négatif ou neutre). L'analyse de sentiment prend rarement en compte l'aspect psychologique pour véritablement comprendre les sentiments et leurs raisons. Au contraire, l'analyse des émotions cherche à identifier les différents types d'émotions exprimés dans un texte, un audio ou une vidéo (VARNI et al., 2020). En plus d'étudier les quatre émotions primaires, la détection des émotions considère également de plus grandes échelles, comme les émotions universelles d'Ekman (tristesse, joie, colère, peur, dégoût, surprise) (EKMAN & FRIESEN, 1971) et même des modèles en trois dimensions comme la roue des émotions de Plutchik, illustrée sur la figure 1.3. Une étude sur les jeux de données pour l'analyse des émotions dans des textes a été réalisée par L.-A.-M. BOSTAN et KLINGER (2018).

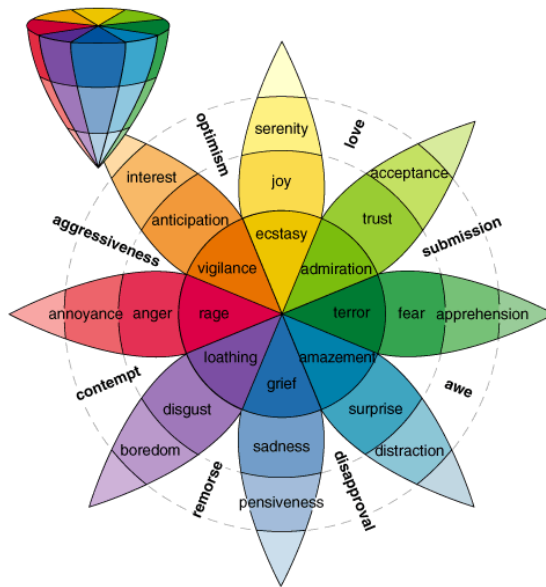


FIGURE 1.3: Le modèle tridimensionnel décrit les relations entre les émotions en faisant une analogie avec la roue des couleurs. La dimension verticale du cône représente l'intensité. Le cercle représente les degrés de similarité entre les émotions. Les huit secteurs indiquent qu'il existe huit dimensions primaires, correspondant aux huit émotions primaires définies par la théorie. Les émotions situées dans les espaces vides sont les mélanges de deux émotions primaires. Image provenant de PLUTCHIK (2001).

1.2.2 Directions de recherche

Aujourd'hui, la majorité des jeux de données sur la détection de la polarité globale d'une phrase sont arrivées à saturation, car les méthodes de classification basées sur l'apprentissage profond obtiennent des scores presque parfaits sur ces jeux de données simples. Nous voulons mettre l'attention sur

des domaines clés de recherche et identifier les directions qui restent à explorer. Particulièrement, nous verrons que l'analyse du sentiment ne peut se réduire seulement à la tâche d'identification de la polarité globale d'une phrase.

Voici les directions de recherche identifiées par PORIA et al. (2020) :

- **Avoir des représentations latentes du contenu textuel qui prennent en compte le sentiment.** Par exemple, il est possible d'incorporer l'information des sentiments dans la fonction de perte d'un réseau de neurones pour prendre en compte les régularités des sentiments. Deux phrases ayant le même sentiment vont être proches lors du plongement dans l'espace latent, alors que deux phrases ayant des sentiments opposés seront éloignées l'une de l'autre.
- **Identifier les sentiments au niveau des aspects.** Les analyses à l'échelle du document ou de la phrase supposent que l'empan de texte contient seulement un sentiment envers une entité. Il est intéressant d'avoir une analyse plus fine des sentiments en se plaçant au niveau des aspects : qui porte l'opinion et pourquoi ? Qu'est-ce qui pousse une personne à exprimer un sentiment sur un sujet particulier ?
L'analyse au niveau des aspects pose de nouveaux enjeux : quelles sont les catégories des aspects (par exemple, dans la phrase « J'adore la 5G », l'aspect 5G est relié au concept de RÉSEAU qui est relié au concept de TÉLÉPHONE), comment les identifier lorsqu'ils ne sont pas explicitement mentionnés dans un texte, comment les extraire simultanément, comment exploiter les relations entre eux, etc.
- **Avoir des jeux de données plus larges et plus riches.** Les jeux de données existants sont petits et l'annotation s'opère souvent à l'échelle de la phrase sans prendre en compte les différents aspects qui existent au sein d'une même phrase (cf. *Semeval-2016* et *Amazon product reviews*).
- **Prendre en compte les différentes modalités.** Par exemple, le son peut être utilisé pour capter l'intonation de la voix ou les images peuvent être utilisées pour capter les expressions faciales. Ces données peuvent représenter une information non négligeable pour la prédiction du sentiment.
- **Introduire la connaissance du sens commun dans les modèles.** Il faut aller au-delà de la simple reconnaissance de motifs. Par exemple, il est possible d'utiliser des réseaux sémantiques pour augmenter la connaissance du monde d'un modèle. Un exemple de réseau sémantique est illustré en Annexe (Figure 3.8).
- **Comprendre le flux séquentiel des sentiments.** Par exemple, une phrase peut être représentée par un circuit électronique où des opérateurs logiques contrôlent le flux du sentiment (CAMBRIA & HUSSAIN, 2015).

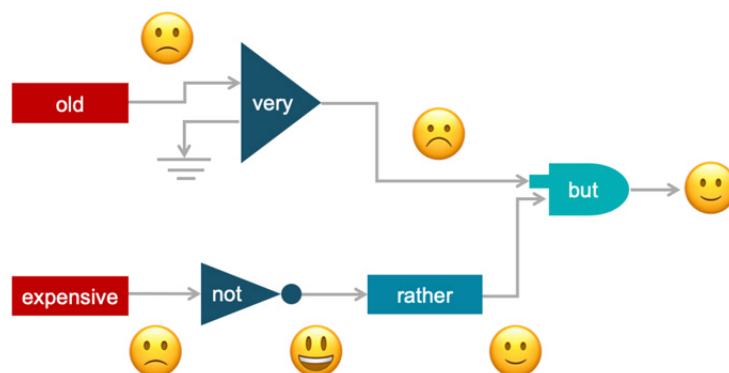


FIGURE 1.4: Représentation par un circuit électronique de la phrase *The car is very old but it is rather not expensive.* Image provenant de CAMBRIA et HUSSAIN (2015).

- **Incorporer l’information des sentiments pour d’autres tâches de TAL.** Par exemple, prendre en compte les sentiments pour la traduction automatique, la génération de texte, la détection de l’ironie, du sarcasme et des discours haineux.
- **Adapter à plusieurs domaines.** L’analyse du sentiment utilise souvent des données venant des réseaux sociaux comme Twitter. La valence des mots affectifs varie selon les domaines. Il est possible d’apprendre des représentations partagées qui sont invariantes aux domaines en utilisant des connaissances extérieures (par exemple, provenant de graphes de connaissance représentant les relations sémantiques entre les concepts).
- **Réduire les biais.** Il est nécessaire d’identifier les biais présents dans les jeux de données d’entraînement ainsi que dans l’architecture des modèles prédictifs. Il faut créer de nouveaux corpus d’évaluation comme *The Equity Evaluation Corpus*¹.
Il est courant d’observer des corrélations fallacieuses dans les modèles d’analyse de sentiment. De nombreux biais sont présents sur les genres, les métiers et les lieux, alors que la plupart du temps ces éléments ne devraient pas influencer le sentiment global d’une phrase. Par exemple, en utilisant un modèle d’apprentissage profond² affiné pour prédire le sentiment d’une phrase, nous observons que la phrase « Les films en France » possède un score de négativité de 0.261 alors que la phrase « Les films en Iraq » possède un score de négativité de 0.459. Il est possible de réduire les biais des modèles en générant des exemples contrefactuels dans les données d’entraînement pour rendre les modèles plus robustes (L. YANG et al., 2021). Les modèles apprendraient les vrais termes causaux importants pour la prédiction des sentiments.

Notre travail de recherche prend en compte les limites présentes dans la littérature sur l’analyse du sentiment et s’inscrit dans certaines directions de recherche décrites plus haut. Nous mettons l’accent sur l’analyse au niveau des aspects et nous travaillons sur un nouveau jeu de données en français avec une annotation plus fine et plus riche par rapport aux jeux de données existants.

1.3 Identification des causes d’une émotion

1.3.1 Extraction de la causalité

Présentation

L’extraction d’information dans un corpus est composée en général de trois sous-tâches : la reconnaissance des entités nommées, l’extraction des relations et l’extraction d’événements. L’extraction de la causalité (J. YANG et al., 2021), qui est un sous-domaine de l’extraction des relations, consiste à identifier les relations de cause à effet entre des pairs de noms labellisés dans un texte.

Les relations causales sont explicitement ou implicitement présentes dans un texte. Elles peuvent être présentes au sein d’une même phrase (intra-phrastique) ou entre différentes phrases (inter-phrastique). Les différentes formes de relations causales sont illustrées dans le tableau 1.1.

Phrases	Formes de la causalité
Le stress financier est l’une des causes principales du divorce .	Explicite et intra-phrastique
Le stress financier peut accélérer une séparation .	Implicite et intra-phrastique
Vous pourriez entendre que l’indivertité peut mener au divorce .	Explicite et inter-phrastique
Néanmoins, le stress financier est un autre facteur important.	

TABLEAU 1.1: Les formes de relations causales.

1. <https://saifmohammad.com/WebPages/Biases-SA.html>

2. <https://huggingface.co/cardiffnlp/twitter-xlm-roberta-base-sentiment>

Différentes approches

Les systèmes basés sur les connaissances sont construits à partir de la création manuelle de règles par des experts d'un domaine. Ils ne nécessitent donc pas de données labellisées. Les systèmes basés sur l'apprentissage statistique sont construits en entraînant des modèles (comme un arbre de décision, une régression logistique ou un bayésien naïf) à partir de caractéristiques manuellement sélectionnées. Les systèmes basés sur l'apprentissage profond (comme un réseau de neurones convolutif ou un réseau de neurones récurrents) apprennent à extraire leurs propres caractéristiques. L'apprentissage de ces derniers modèles nécessite des données labellisées pour effectuer différentes tâches comme la classification d'un sentiment.

Les systèmes basés sur les connaissances et les systèmes basés sur l'apprentissage statistique sont performants pour des domaines restreints, comme le domaine médical, mais ont des difficultés à découvrir les relations causales implicites entre différentes phrases (inter-phrastique). Les approches basées sur l'apprentissage profond offrent une meilleure généralisation à d'autres domaines et peuvent explorer les relations implicites grâce aux bénéfices du pré-entraînement et à la modélisation de dépendances à longue distance, notamment grâce aux mécanismes d'attention. Par exemple, JIN et al. (2020) utilisent des réseaux de neurones convolutifs pour extraire les caractéristiques essentielles d'un texte, puis utilisent un réseau de neurones récurrents pour obtenir l'information sémantique à longue distance entre une cause et un effet.

Notion de cause et d'effet complexe

Il reste des problèmes ouverts. La réalité est plus complexe. Plusieurs causes peuvent produire un unique effet et une cause peut produire plusieurs effets. Des boucles de rétroaction peuvent exister entre des causes et des effets. De plus, pour notre analyse des récits d'une scène émotionnelle, nous travaillons à l'échelle du document, il faut donc extraire des relations causales à cette échelle, parfois à travers plusieurs dizaines de phrases. L'extraction à l'échelle du document est facilitée par l'émergence de nouveaux jeux de données (YAO et al., 2019). Des réseaux de neurones pour graphes sont notamment utilisés pour effectuer cette tâche (ZENG et al., 2020). Les récits autobiographiques comportent des relations souvent implicites et nécessitent des connaissances en psychologie pour les décoder. Automatiser cette tâche est donc un défi excitant.

1.3.2 Extraction des causes d'une émotion

Présentation

Nous présentons la tâche d'extraction des causes d'une émotion telle que présentée dans la littérature. Un document D est considéré comme un ensemble ordonné de clauses $D = [c_1, c_2, \dots, c_d]$. Une clause est un groupe de mots contenant un sujet et un prédicat et fonctionnant comme le membre d'une phrase composée. Par exemple, « Quand je suis venu, je l'ai vu » est une phrase à deux clauses alors que « J'ai reçu mon diplôme l'année dernière » est une phrase à une seule clause.

Différentes tâches

La tâche d'Extraction de la Cause d'une Émotion (ECE) a pour la première fois été définie comme un problème d'extraction à l'échelle du mot par CHEN et al. (2010). GUI et al. (2016) proposent d'effectuer l'ECE en effectuant une extraction à l'échelle de la clause, tout en publiant un nouveau corpus en chinois.

XIA et DING (2019) introduisent la tâche d'Extraction des Pairs d'Émotion-Cause (EPEC). L'objectif de la tâche EPEC est d'extraire les paires émotion-cause présentes dans un texte, $P = \{\dots, (c_i, c_j), \dots\}$ où $c_i \in D$ est une clause d'émotion et $c_j \in D$ est une clause de cause. La tâche EPEC extrait à la fois la clause d'émotion et la clause de cause, alors que dans la tâche ECE, l'annotation de la clause de l'émotion est donnée à l'avance. Cette dernière tâche est donc plus simple

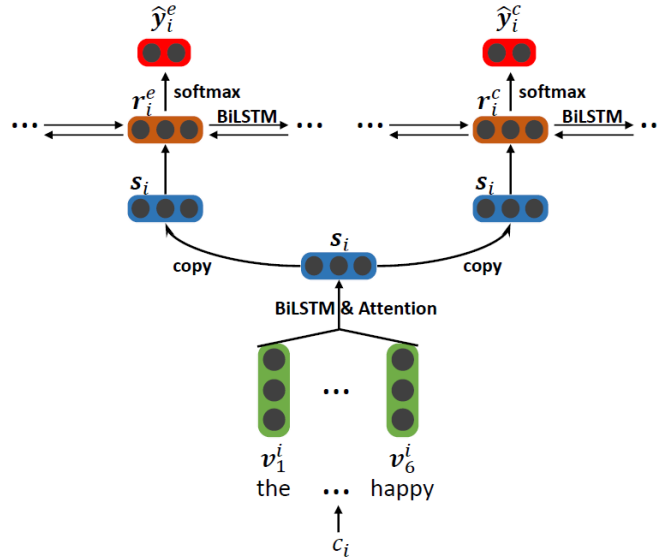


FIGURE 1.5: L’architecture pour l’extraction des clauses émotions et des clauses causes dans un texte. Image provenant de XIA et DING (2019).

et ignore les influences mutuelles entre une émotion et une cause. De plus, les émotions doivent être annoté avant l’extraction des causes dans le jeu de test, ce qui limite les applications de l’ECE dans le monde réel scénarios.

Description de l’architecture

Une partie de l’architecture (la première étape) du modèle pour l’EPEC est décrite sur la figure 1.5. Nous décrivons l’architecture de ce modèle en détail puisqu’il a inspiré de nombreux papiers dans la littérature. Une clause c_i est représentée par une suite de mots v_j^i . Un *Bi-directional Long Short-Term Memory* (Bi-LSTM) est utilisé pour obtenir une représentation latente s_i de la clause en modélisant les relations entre les mots qui la constituent. Les différentes représentations latentes s_i des clauses sont ensuite passées dans un deuxième Bi-LSTM qui modélise les relations entre les différentes clauses d’un texte pour obtenir une nouvelle représentation latente des clauses r_i^e , utilisée pour prédire la cause. En parallèle, un autre Bi-LSTM est utilisé de la même façon pour obtenir une représentation latente des clauses r_i^c pour prédire l’émotion. À partir de cette représentation latente, un réseau de neurones entièrement connecté prédit si la clause est une émotion y_i^e ou une cause y_i^c .

Dans la deuxième partie du modèle, non décrit sur la figure 1.5, un produit cartésien entre toutes les clauses émotions et les clauses causes est utilisé pour obtenir l’ensemble de toutes les paires possibles. Les représentations latentes des clauses émotions et des clauses causes sont concaténées $[r_i^e, r_i^c]$ et passent dans une régression logistique qui prédit s’il existe une relation causale ou non entre l’émotion et la cause.

C’est donc une approche en deux étapes ou approche *pipeline*. Tout d’abord, nous apprenons une représentation latente des clauses pour prédire respectivement si ce sont des émotions, des causes ou ni l’un ni l’autre. Puis, nous apprenons à prédire parmi les différents candidats retenus les bonnes paires d’émotion-cause. Les modèles utilisant une approche *pipeline* souffre de la propagation d’erreur. En effet, si nous ne prédisons pas correctement qu’une clause est véritablement une cause ou une émotion, l’erreur se propagera lors de la prédiction des paires possibles dans la deuxième étape. C’est pourquoi les modèles récents privilégient une approche « bout-à-bout » en apprenant simultanément l’extraction et le filtrage des paires d’émotion-cause. SINGH et al. (2021) montrent

l'efficacité d'un entraînement joint pour la tâche EPEC avec une prise en compte des influences mutuelles entre les clauses émotions et les clauses causes. En effet, les émotions peuvent aider à mieux extraire les causes et inversement. L'architecture du modèle est disponible en Annexe (Figure 3.3).

1.3.3 Extraction des stimuli : études anglaises et françaises

Notion de stimulus

Dans certaines études, provenant majoritairement de sources non-chinoises, la notion de cause d'une émotion est remplacée par la notion de stimulus. Un stimulus est un événement, un objet ou une personne qui déclenche une émotion chez une personne. La tâche d'extraction des stimuli s'opère à l'échelle du token plutôt qu'à l'échelle de la clause. En effet, OBERLÄNDER et KLINGER (2020) ont étudié les différentes manières de détecter un stimulus dans un texte, à l'échelle de la clause (« est-ce que cette clause contient un stimulus ? ») ou à l'échelle du token (« quelles sont le ou les tokens qui décrivent un stimulus ? »). La classification à l'échelle de la clause est majoritairement effectuée sur des jeux de données en chinois. Cependant, l'étude montre que la clause n'est pas une unité appropriée pour détecter le stimulus en anglais. Nous faisons l'hypothèse que la même observation peut s'effectuer pour le français, c'est pourquoi nous choisissons de travailler à l'échelle du token. La classification des clauses est à plus gros grain et donc possède de meilleures performances par rapport à la classification des tokens. Cependant, elle ne peut correctement identifier le début et la fin exacts d'un empan de texte correspondant à un stimulus. De plus, il faut au préalable identifier ce qu'est une clause dans un texte, l'approche souffre donc de la propagation d'erreur. D. GHAZI (2015) construit un jeu de données avec une annotation des émotions et de leurs stimuli, pouvant être un ou plusieurs tokens, en s'inspirant du cadre Émotions-dirigées de la base de données lexicale FrameNet ¹.

Format IOB

PAS DE STIMULUS	STIMULUS
[Elle est contente]	[que les choses se passent bien.]

TABLEAU 1.2: Classification du stimulus à l'échelle la clause.

O	O	O	B	I	I	I	I	I	O
Elle	est	contente	que	les	choses	se	passent	bien	.

TABLEAU 1.3: Classification du stimulus à l'échelle du token.

Le format *Inside-Outside-Beginning* (IOB) est un format commun pour labelliser les tokens d'une phrase. Il s'agit, pour chaque token, d'identifier si celui-ci est un token marquant le début (B pour *Beginning*) ou l'intérieur (I pour *Inside*) de l'empan de texte qui représente un stimulus. Si le token est en dehors de l'empan de texte, il est marqué O pour *Outside*. Un modèle de classification à l'échelle du token devra donc prédire, pour chaque token, s'il appartient à la classe I, O ou B. Notre jeu de données, présentée dans le chapitre 2, supporte ce format.

Études anglaises et françaises

Il existe plusieurs travaux sur l'annotation des rôles sémantiques liés aux émotions sur des textes en anglais. S. MOHAMMAD et al. (2013) ont annoté 4058 tweets durant les élections présidentielles américaines de 2012 en considérant le SUJET qui ressent l'émotion, le STIMULUS ainsi que la CIBLE de l'émotion. KIM et KLINGER (2018) ont annoté un corpus littéraire constitué de 1720 paragraphes

1. <https://framenet.icsi.berkeley.edu/fndrupal/>

provenant du Projet Gutenberg¹. Les INDICES (marqueurs montrant la présence d’une émotion dans un texte) sont annotés ainsi que les SUJETS, les CIBLES et les STIMULI liés à l’émotion.

DANG et al. (2021) introduisent un corpus constitué de titres de journaux allemands. Ils identifient les INDICES, les SUJETS et les STIMULI avec une annotation à l’échelle du token. De plus, ils prennent en compte la perception du lecteur : quelle(s) émotion(s) ressentent les lecteurs lorsqu’ils lisent un titre de journaux ? Un travail similaire a été effectué pour l’annotation des titres de journaux anglais (L. A. M. BOSTAN et al., 2020).

À notre connaissance, les seuls travaux en français qui propose d’identifier les rôles sémantiques liés aux émotions sont le corpus pour reconnaître les émotions dans les livres pour enfant (ÉTIENNE et al., 2020) et la campagne Défi Fouille de Textes (DEFT) de 2018 (PAROUBEK et al., 2018). Il s’agit d’identifier la polarité globale (positive et négative) des tweets français faisant référence aux transports en Île-de-France. Plusieurs rôles sont annotés comme le SUJET, la CIBLE, les MODIFIEURS qui indiquent l’intensité d’une émotion ainsi que les INDICES.

Nous voyons que les jeux de données pour l’analyse des causes d’une émotion sont majoritairement en chinois, un peu en anglais et n’existent quasiment pas en français. De plus, les schémas d’annotation proposés sont rarement à grain fin. Pour finir, les théories en psychologie sont seulement prises en compte pour définir les différentes classes à prédire, mais aucune théorie n’est utilisée pour réellement comprendre les rôles sémantiques liés aux émotions (SUJET, INDICE, STIMULI, etc.). Nous revenons sur la définition des différents rôles sémantiques dans le chapitre 3.

Pour remédier à ces limites, dans le prochain chapitre, nous proposons un nouveau jeu de données en français basé sur des récits autobiographiques d’une scène émotionnelle en s’appuyant sur une nouvelle méthode en psychologie des émotions, l’Analyse Cognitive des Émotions développée par Alain Finkel. Dans le dernier chapitre, nous présentons notre nouveau schéma d’annotation.

1. <https://www.gutenberg.org>

Chapitre 2

Jeu de données composé de scènes émotionnelles

2.1 Construction du jeu de données

Les affiliations culturelles d'un individu peuvent grandement influencer l'émotion exprimée dans une situation donnée. Nous sommes donc intéressés par la construction d'un jeu de données en français annoté au niveau des aspects, en observant que le français est une langue peu considérée pour l'analyse des émotions. De plus, nous voulons avoir un matériel textuel adapté aux problèmes spécifiques de l'ACE.

2.1.1 Extraction des grilles

Nous travaillons à partir de 598 documents Word pouvant contenir chacun plusieurs grilles FERA (Tableau 2.1). Chaque grille est remplie par une unique personne et correspond à une seule émotion à traiter (colère, joie, tristesse ou peur). Ces documents ont été récoltés manuellement sur une période allant de 2005 à 2022 par des psychologues entraînés à l'ACE, et donc à l'analyse des grilles FERA. Durant ces années, le format des grilles a changé plusieurs fois ainsi que les consignes données. L'enjeu est de convertir toutes les grilles dans un format commun pour pouvoir ensuite les traiter plus facilement. De plus, les grilles comportent des informations personnelles comme des noms, des dates et des lieux qu'il faut supprimer pour le respect de la vie privée des auteurs.

Le tableau 2.1 illustre les différents formats des grilles. Certaines grilles ont des dimensions différentes qui peuvent parfois être reliées à une année et à des consignes précises. Par exemple, les grilles ayant 7 lignes et 4 colonnes correspondent aux grilles de l'année 2019. Les grilles ayant 3 lignes et 4 colonnes contiennent seulement les FAITS EXTÉRIEURS, les FAITS INTÉRIEURS et la VÉRIFICATION DES FAITS.

Nous montrons l'extraction et la conversion d'une grille dans un format commun à travers un exemple simple sur une grille du document « 239.docx ». Le texte brut de la grille est mis dans un tableau selon son type (s'il correspond à une consigne ou à la réponse d'un auteur). Le tableau 2.2 contient trois colonnes correspondant aux consignes qui ont été extraites. Le tableau 2.3 contient une colonne correspondant aux réponses de l'auteur de la grille. Le but est d'assigner pour chaque réponse de l'auteur une catégorie en analysant les colonnes contenant les consignes. Pour faciliter l'analyse, les colonnes correspondant aux consignes sont concaténées pour obtenir une seule nouvelle colonne. Les deux tableaux sont ainsi unifiés pour obtenir un nouveau tableau normalisé contenant une colonne pour les consignes et une colonne pour les réponses (Tableau 2.4). Le nombre de

nom	nb_grille	nb_ligne	nb_col	texte
142.docx	2	7	4	[Aéroport lors du passage de la sécurité, la pe...
432.docx	2	12	4	[Vendredi dernier au bureau. Nous...
168.docx	2	8	2	[Colère L'examen du permis de conduire...
558.docx	4	11	4	[Un matin, j'arrive au travail, je croise une c...
481.docx	2	8	4	[Un vendredi après-midi, je fais passer un part...
504.docx	2	8	4	[Un lundi matin, j'arrive dans le service de ma...
297.docx	1	9	4	[Le lundi 3 mai, je devais me rendre à un collo...
...
453.docx	2	11	4	[Durant les vacances de février, Fabrice et moi...
371.docx	2	9	4	[Il y a 3 semaines, la fille de mon ami est ven...
147.docx	2	8	2	[Colère Les faits se sont passés un samedi soir...
396.docx	2	11	4	[Formation Acta. Alain fait un tour de ...
276.docx	4	5	4	[Mardi dernier, j'ai retrouvé dans le restauran...

TABLEAU 2.1: Statistiques des documents Word.

FAITS	EXTÉRIEURS	\nLe contexte : lieu, temps, personnes, situation, actions. . .
FAITS	INTÉRIEURS	Mes ressentis physiques : \nlocalisations et pré...
FAITS	INTÉRIEURS	Le film de mes pensées : \n images, sons, mots . . .
	Vérification des faits	Contexte et ressentis \nsont-ils exprimés factu...
	Identification de l'émotion	Une ou plusieurs ? \nLaquelle est traitée ici ?
RAISONS	colère/peur	Quel «territoire» attaqué ? \nPar qui/quoi ?
RAISONS	Vérifications \n colère/peur	Le territoire est-il à moi ? \nL'attaque est ext...
		Ce que m'apporte cette analyse au moment de l'é...
ACTIONS	INTÉRIEURES	Mes actions \n intérieures \n qui ont permis de ...
ACTIONS	EXTÉRIEURES	Mes actions \nextérieures \n qui ont permis de s...

TABLEAU 2.2: Colonnes correspondant aux consignes.

Je suis au cabinet, seul, nous sommes lundi et...
J'ai une sensation d'oppression dans le thorax,...
A chaque fois que je pense à cette situation et...
Contexte : exprimé factuellement \nRessentis : f...
Peur/ dégoût
Le territoire attaqué est mon territoire de con...
Il s'agit bien de mon territoire, l'attaque est...
Rien, car je suis dans l'attente et dans l'expe...
L'émotion n'est pas diminuée dans le sens où il...
Le seul moyen de stopper l'attaque est la fuite...

TABLEAU 2.3: Colonne correspondant aux réponses de l'auteur d'une grille.

colonnes obtenu est le même pour toutes les grilles, nous n'avons plus le problème de travailler sur des grilles ayant des dimensions différentes.

catégorie	réponse
FAITS EXTÉRIEURS \nLe contexte : lieu, temps, p...	Je suis au cabinet, seul, nous sommes lundi et ...
FAITS INTÉRIEURS Mes ressentis physiques : \nloc...	J'ai une sensation d'oppression dans le thorax,...
FAITS INTÉRIEURS Le film de mes pensées : \n ima...	A chaque fois que je pense à cette situation et...
Vérification des faits Contexte et ressentis ...	Contexte : exprimé factuellement \nRessentis : f...
Identification de l'émotion Une ou plusieurs ?...	Peur/ dégoût
RAISONS colère/peur Quel «territoire» attaqué ?\...	Le territoire attaqué est mon territoire de con...
RAISONS Vérifications \ncolère/peur Le territor...	Il s'agit bien de mon territoire, l'attaque est...
Ce que m'apporte cette analyse au moment de l...	Rien, car je suis dans l'attente et dans l'expe...
ACTIONS INTÉRIEURES Mes actions \n intérieures \...	L'émotion n'est pas diminuée dans le sens où il...
ACTIONS EXTÉRIEURES Mes actions \nextérieures \n...	Le seul moyen de stopper l'attaque est la fuite...

TABLEAU 2.4: Groupement des consignes en une seule colonne.

2.1.2 Identification des catégories

Nous rappelons les différentes catégories que nous avons retenues : FAIT EXTÉRIEUR, FAIT INTÉRIEUR, RESENTIS, FAIT INTÉRIEUR PENSÉES, IDENTIFICATION ÉMOTION, TERRITOIRE, VÉRIFICATION FAIT, ACTION INTÉRIEURE, ACTION EXTÉRIEURE, ACTION TEST. Les catégories communes entre toutes les grilles sont identifiées avec un filtrage par expression régulière sur chaque ligne de la colonne contenant les consignes. Par exemple, il suffit de trouver l'occurrence de « faits extérieurs » dans une consigne pour savoir que la réponse correspondante est un FAIT EXTÉRIEUR. Nous obtenons un tableau normalisé avec des catégories communes à toutes les grilles (Tableau 2.5). Certaines catégories sont ensuite regroupées entre elles pour une meilleure cohérence. Par exemple, les catégories TERRITOIRE (COLÈRE/PEUR) et TERRITOIRE (JOIE/TRISTESSE) sont regroupées sous la catégorie TERRITOIRE. Les consignes qui n'ont pas pu être identifiées avec le filtrage par expression régulière sont regroupées sous la catégorie RIEN.

catégorie	réponse
fait extérieur	Je suis au cabinet, seul, nous sommes lundi et ...
fait intérieur ressentis	J'ai une sensation d'oppression dans le thorax,...
fait intérieur pensées	A chaque fois que je pense à cette situation et...
vérification fait	Contexte : exprimé factuellement \nRessentis : f...
identification émotion	Peur/ dégoût
territoire (colère/peur)	Le territoire attaqué est mon territoire de con...
vérification émotion	Il s'agit bien de mon territoire, l'attaque est...
apport analyse	Rien, car je suis dans l'attente et dans l'expe...
action intérieure	L'émotion n'est pas diminuée dans le sens où il...
action extérieure	Le seul moyen de stopper l'attaque est la fuite...

TABLEAU 2.5: Tableau normalisé final avec les classes retenues.

2.1.3 Désindentification des grilles

Le texte en entête et en bas de page des documents Word est supprimé puisqu'il contient généralement des informations personnelles comme « Gustave - module 2 » et « ACTA 2015 à rendre le lundi 15 décembre midi au plus tard ». Le nom des documents est aussi remplacé par un chiffre : « grille-gustave.docx » devient « 256.docx ».

Nous utilisons un modèle français de reconnaissance des entités nommées¹ pour identifier automatiquement dans les grilles les occurrences de noms, dates et lieux que nous remplaçons par des tokens spécifiques [NOM], [DATE] et [LIEU].

2.1.4 Prétraitement des grilles

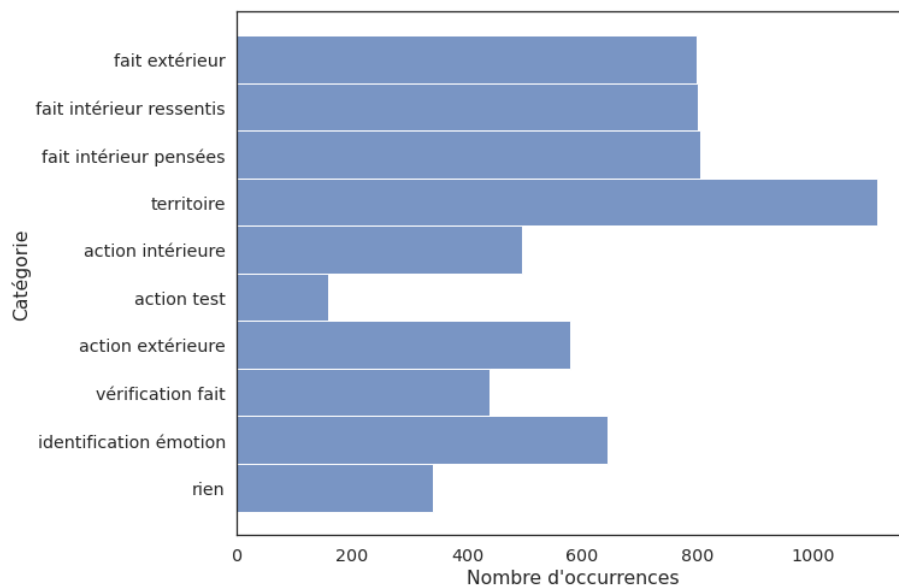


FIGURE 2.1: Distribution des catégories.

Nous décrivons le prétraitement du contenu textuel des grilles. Les réponses vides, les réponses contenant peu de tokens (inférieur à 3 tokens) et les réponses faisant référence à une absence de réponse comme « x » sont supprimées. Les grilles dupliquées sont également supprimées. Un identifiant unique est attribué pour chaque grille. Le jeu de données final est composé de 888 grilles uniques, soit un total de 6174 réponses.

La figure 2.1 montre la distribution des catégories que nous avons extraites. Environ 400 réponses n’ont pas pu être catégorisées (c.-à-d. elles sont dans la catégorie RIEN) avec la méthode des expressions régulières. Les FAITS EXTÉRIEURS et les FAITS INTÉRIEURS sont en nombre égal puisque le remplissage d’une grille nécessite au moins de répondre à ces deux consignes. L’ACTION TEST est remplie peu de fois puisque c’est la dernière réponse à remplir. Généralement, les personnes remplissant les grilles n’ont pas le temps de répondre à cette dernière catégorie.

La figure 2.2 montre le nombre moyen de tokens par catégorie. Nous voyons que la longueur des FAITS EXTÉRIEURS est importante, puisque les personnes prennent du temps pour expliquer le contexte de la scène émotionnelle vécue. Le contenu textuel d’IDENTIFICATION ÉMOTION est très petit puisque généralement les personnes écrivent seulement l’émotion qu’elles ont ressentie.

Un filtrage par expression régulière est utilisé sur chaque réponse de la catégorie IDENTIFICATION DE L’ÉMOTION pour identifier l’émotion correspondant à chaque grille. Nous avons pu identifier avec cette méthode l’émotion de 644 grilles. Bien qu’une grille corresponde à une seule émotion, parfois l’auteur précise d’autres émotions qu’il a vécues. 563 (87%) grilles possèdent une seule émotion, 78 (12%) grilles possèdent deux émotions et 3 (0.5%) grilles possèdent trois émotions. La distribution des émotions n’est pas équilibrée, 312 (48%) grilles concernent la colère, 204 (31%) la peur, (12%)

1. https://gitlab.com/MAPA-EU-Project/mapa_project

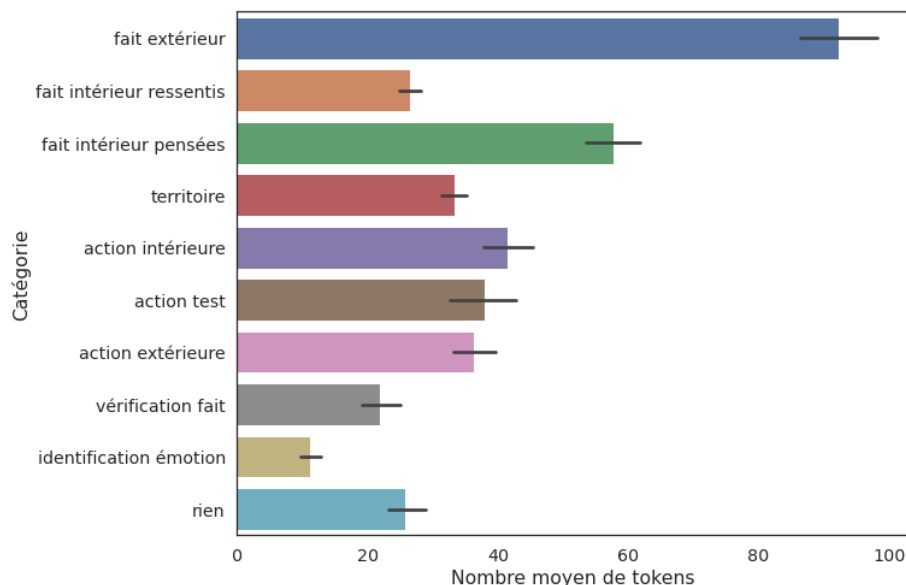


FIGURE 2.2: Nombre moyen de tokens par catégorie.

la tristesse et 69 (10%) la joie. La dominance des émotions négatives est normale, une personne remplit généralement une grille lorsqu'elle souhaite mieux traiter une situation qu'elle a vécue.

La bibliothèque de TAL SpaCy¹ est utilisée pour tokeniser et lemmatiser les réponses. La tokenisation est le fait de segmenter une phrase en plusieurs tokens. Par exemple, la version tokenisée de « Elle est contente que les choses se passent bien. » est la suite de tokens : « elle, est, contente, que, les, choses, se, passent, bien, . ». Chaque token peut être encodé sous la forme d'un vecteur pour pouvoir ensuite être utilisé dans un modèle d'apprentissage automatique. La lemmatisation consiste à retrouver la forme canonique des différents mots, que l'on désigne sous le terme de lemme. Par exemple, le lemme des différents tokens « petit, petite, petits, petites » est « petit ». Une fois la lemmatisation effectuée, il est parfois important de supprimer les mots (*stop words*) qui ne sont pas intéressants pour l'analyse. En français, nous pouvons par exemple retirer les pronoms et les articles².

À partir de la lemmatisation des réponses, nous obtenons les lemmes le plus fréquent pour chaque catégorie. Une illustration des lemmes les plus fréquents selon les catégories est disponible en Annexe (Figure 3.4 et 3.5). Nous observons l'importance des verbes pour distinguer les différentes catégories (comme *ressentir* et *sentir* pour les FAITS INTÉRIEURS RESENTIS, *faire* et *être* pour les FAITS EXTÉRIEURS). Chaque catégorie possède ses propres champs lexicaux. Par exemple, pour les FAITS EXTÉRIEURS, les nombreuses occurrences des mots *travail*, *collègue*, *étudiant* et *bureau* indiquent que les grilles sont spécifiques au domaine du travail et de l'enseignement. Ceci est logique puisque la plupart des auteurs des grilles proviennent de l'enseignement supérieur. Les modèles d'apprentissage construits à partir de ces données seront donc spécifiques à ces domaines et auront du mal à généraliser à toutes les scènes émotionnelles disponibles « dans la nature ». Notons la présence du champ lexical du corps (*cœur*, *tête*, *main*, etc.) pour les FAITS INTÉRIEURS RESENTIS. Pour IDENTIFICATION ÉMOTION, les 4 émotions primaires sont présents en grande majorité, mais aussi d'autres émotions secondaires comme la *surprise* et la *honte*. Parmi les différentes catégories, nous repérons le vocabulaire des faits, défini par FINKEL (2022). Il contient les verbes décrivant des

1. <https://spacy.io>

2. Notons que la pertinence d'un mot dépend du problème, certains mots considérés généralement comme un *stopword* peuvent jouer un rôle important dans une analyse.

sensations et des perceptions (*voir, entendre, sentir*), les paramètres de sens (*chaud, froid, rouge, agréable*), les émotions primaires ainsi que les expressions des préférences de l’auteur. Nous repérons également, dans la catégorie FAIT INTÉRIEUR RESENTIS, le vocabulaire associé à la physiologie des émotions. Par exemple, la peur peut être liée à un afflux de sang dans les jambes ou une paralysie du corps, la colère peut être liée à une élévation de la température et du rythme cardiaque.

Il est important d’examiner les différents composants de l’émotion (comme les pensées et les ressentis physiques) et vérifier s’ils sont cohérents avec l’émotion identifiée. Dans la prochaine section, nous entraînons des modèles de classification pour prédire une émotion à partir des réponses des différentes catégories de la grille FERA. Des modèles sont également entraînés pour prédire à partir d’une réponse la catégorie correspondante.

2.2 Modèles de classification

2.2.1 Bibliothèques utilisées

Pour l’entraînement des modèles, nous utilisons les bibliothèques Python suivantes :

- Transformers¹, développée par l’entreprise Huggingface, qui permet d’entraîner des modèles de TAL atteignant l’état de l’art. De nombreux modèles pré-entraînés et des jeux de données sont disponibles gratuitement. Les modèles entraînés avec cette bibliothèque peuvent fonctionner avec PyTorch. Nous l’utilisons pour entraîner nos modèles CamemBERT ;
- Pytorch qui permet d’effectuer du calcul tensoriel, notamment pour de l’apprentissage profond ;
- Scikit-learn² pour effectuer de l’apprentissage automatique. Nous l’utilisons pour diviser l’ensemble de données entre les données d’entraînement et les données de test, pour entraîner les régressions logistiques, les forêts aléatoires et les séparateurs à vaste marge, ainsi que pour calculer certaines métriques comme le score F1 ;
- Pandas³ qui permet la manipulation des données dans un format tabulaire. Nous l’utilisons pour construire et manipuler le jeu de données.

2.2.2 Modèles d’apprentissage automatique

Méthode TF-IDF

Pour la classification, les modèles sont entraînés pour prédire les 9 catégories que nous avons retenues : FAIT EXTÉRIEUR, FAIT INTÉRIEUR RESENTIS, FAIT INTÉRIEUR PENSÉES, IDENTIFICATION ÉMOTION, TERRITOIRE, VÉRIFICATION FAIT, ACTION INTÉRIEURE, ACTION EXTÉRIEURE, ACTION TEST. C’est un problème de multiclassification monolabel puisque l’on prédit une seule classe (parmi les neufs) pour chaque réponse.

Les caractéristiques des réponses sont extraites avec la méthode *Term Frequency – Inverse Document Frequency* (TF-IDF). C’est une méthode de pondération couramment utilisée dans la recherche d’informations, en particulier dans l’exploration de texte. Cette mesure statistique permet d’évaluer l’importance des termes contenus dans les documents (ici, les réponses des auteurs) liés à un jeu de données. Le poids augmente proportionnellement au nombre de fois qu’un mot apparaît dans la réponse d’un auteur (Fréquence des Termes). Le poids diminue aussi proportionnellement au nombre de fois qu’un mot apparaît dans le jeu de données entier (Fréquence Inverse des Documents). Ainsi, plus un terme est rare au sein du jeu de données, plus il aura un poids important et sera considéré comme discriminant. Voici les différentes formules pour calculer la pondération d’un terme t parmi D documents :

1. <https://huggingface.co>
2. <https://scikit-learn.org/>
3. <https://pandas.pydata.org>

$$\text{TF}(t) = \frac{\text{Nombre d'apparitions du terme } t \text{ dans le document}}{\text{Nombre total de termes dans le document}} \quad (2.1)$$

$$\text{IDF}(D) = \frac{\text{Nombre de documents } D}{\text{Nombre de documents où le terme } t \text{ apparaît}} \quad (2.2)$$

$$\text{TF_IDF}(t, D) = \text{TF}(t) \times \text{IDF}(D) \quad (2.3)$$

La méthode TF-IDF est utilisée sur les tokens des réponses obtenus précédemment. Nous obtenons une matrice de dimension (5833, 10523) qui contient les vecteurs d'entraînement de nos modèles de classification. La première dimension (5833) correspond au nombre de réponses présentes dans le jeu de données. La deuxième dimension (10523) correspond à la taille de notre vocabulaire (c.-à-d. le nombre de tokens uniques dans le jeu de données). Chaque réponse est donc représentée par un vecteur de dimension 10523. Par exemple, pour la phrase « Boule au ventre, respiration forte », les tokens, après la suppression des *stop words*, sont « boule, ventre, respiration, forte ». Cette phrase est représentée par un vecteur rempli de 0 sauf pour la 1404^{ème}, 4250^{ème}, 8202^{ème} et 9937^{ème} valeurs contenant les poids respectifs des tokens : 0.47, 0.48, 0.58 et 0.45. Le troisième token est le plus important puisqu'il a le plus grand poids (0.58).

Entraînement des modèles

À partir de la matrice de dimension (5833, 10523) obtenue avec la méthode TF-IDF, nous entraînons trois modèles d'apprentissage automatique : la régression logistique, la forêt aléatoire et le séparateur à vaste marge. Pour l'évaluation des modèles, nous effectuons une validation croisée à 5 paquets et calculons le score F1, le support et la précision avec une micro-moyenne¹.

modèle	précision	support	score F1
Régression logistique	0.766 ± 0.010	0.758 ± 0.013	0.759 ± 0.012
Forêt aléatoire	0.745 ± 0.008	0.735 ± 0.006	0.733 ± 0.007
Séparateur à vaste marge	0.779 ± 0.010	0.762 ± 0.012	0.762 ± 0.012

TABLEAU 2.6: Scores obtenus sur la régression logistique, la forêt aléatoire et le séparateur à vaste marge.

Nous observons, sur le tableau 2.6, que les trois modèles ont des performances similaires, bien que le séparateur à vaste marge possède des performances sensiblement supérieures par rapport aux deux autres modèles.

Analyse des caractéristiques

Nous pouvons analyser les caractéristiques importantes permettant de discriminer les différentes classes. Pour cela, la bibliothèque eli5² est utilisée pour l'explication des caractéristiques en utilisant la méthode de l'importance de la permutation. Celle-ci consiste à mesurer de combien le score (ici, le score F1) diminue lorsqu'une caractéristique n'est pas disponible (comme la caractéristique correspondant au mot « travail »). Cela demande beaucoup de moyens de calcul de supprimer à chaque fois une caractéristique de l'ensemble de données et d'entraîner un nouveau modèle pour calculer le score obtenu sur les données de validation. En effet, il faudrait entraîner 10523 modèles puisque nous avons 10523 caractéristiques différentes. Pour éviter d'entraîner à nouveau le modèle à chaque fois, nous pouvons supprimer une caractéristique uniquement sur les données de validation et donc calculer le score sans celle-ci. Au lieu de supprimer complètement la caractéristique, on la

1. La micro-moyenne permet de prendre en compte le déséquilibre des classes à prédire lors de la pondération.

2. <https://github.com/eli5-org/eli5>

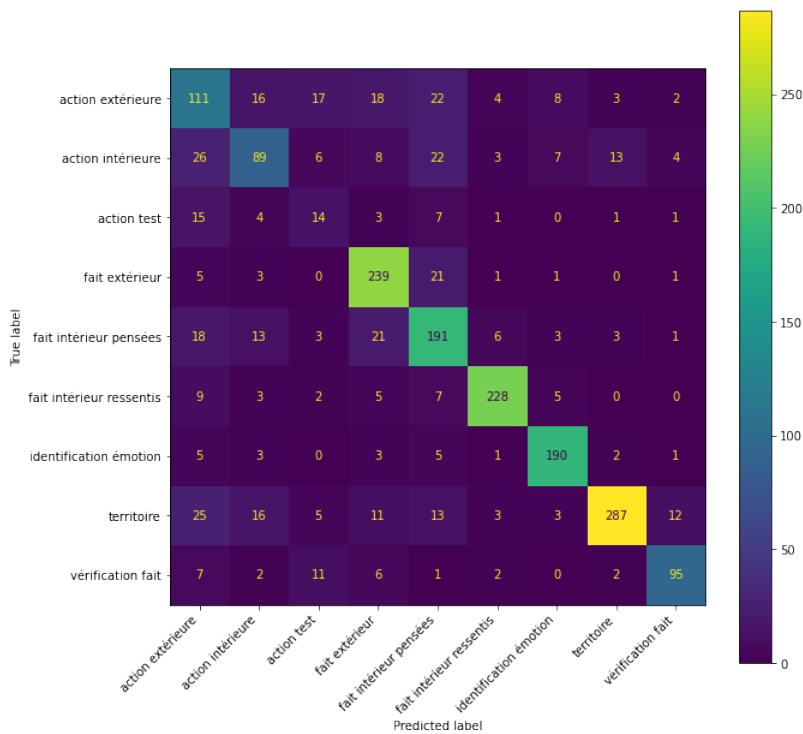


FIGURE 2.3: Matrice de confusion pour la régression logistique.

remplace par un bruit aléatoire en utilisant les valeurs des autres caractéristiques, d'où l'importance de la permutation. Cette méthode peut également être utilisée pour supprimer les caractéristiques non importantes d'un modèle. L'analyse des caractéristiques est disponible en Annexe (Figure 3.6).

Avec la matrice de confusion, illustrée figure 2.3, nous voyons que certaines classes n'ont pas pu être classifiées correctement. Par exemple, le modèle de classification a du mal à discriminer entre les différentes actions (ACTION EXTÉRIEURE, ACTION INTÉRIEURE et ACTION TEST). Nous pensons que ce problème est en partie dû au fait que nous nous concentrons seulement sur les mots d'une phrase sans essayer de comprendre celle-ci dans sa globalité. Les auteurs des grilles utilisent les mêmes mots pour décrire les différentes actions, mais la contextualité change. Il est donc intéressant d'entraîner un modèle qui prend en compte la contextualité. Pour cela, dans la prochaine section, nous affinons un modèle CamemBERT.

2.2.3 Modèles de langue pré-entraînés

Présentation de BERT

Bidirectional Encoder Representations from Transformers (BERT) est un modèle de langue développé par Google en 2018. BERT utilise l'architecture Transformer (VASWANI et al., 2017) et permet d'apprendre, grâce au mécanisme d'attention, les relations contextuelles entre les mots d'un texte. Contrairement à certains modèles comme les *Long Short-Term Memory* (LSTM) qui lisent le texte séquentiellement en prenant seulement en compte le contexte à gauche, l'encodeur du Transformer est bidirectionnel puisqu'il prend en compte le contexte à gauche et à droite. Lors de l'entraînement, BERT utilise deux stratégies :

- *Masked Language Modeling* (MLM). Avant d'envoyer les séquences de tokens dans BERT, une partie des tokens de chaque séquence est remplacée aléatoirement par le token spécial

[MASK]. Le modèle essaie de prédire la vraie valeur de [MASK] en se basant sur les autres tokens qui constituent la séquence. Par exemple, pour la phrase *I [MASK] a student*, BERT doit prédire $[MASK] = am$;

- *Next Sentence Prediction* (NSP). Durant l’entraînement, 50% des entrées arrivent de pair avec la séquence qui les suit dans le document original, alors que 50% est une séquence aléatoirement choisie dans le corpus. Le but est de prédire si la seconde séquence est connectée à la première.

Lors de l’entraînement d’un modèle BERT, MLM et NSP sont entraînés ensemble, le but est de minimiser la fonction de perte de ces deux stratégies.

A Robustly Optimized BERT Pretraining Approach (RoBERTa) (LIU et al., 2019), développé par Facebook, est construit à partir de l’architecture de BERT, mais le pré-entraînement a été optimisé. La valeur des hyperparamètres a été modifiée comme le taux d’apprentissage et la taille du lot qui sont désormais plus élevés. Le modèle est également pré-entraîné sur un plus large corpus de texte (160Go). Pour finir, une seule stratégie d’entraînement a été gardée, qui est le MLM.

Présentation de CamemBERT

CamemBERT (MARTIN et al., 2020), développée par Institut national de recherche en sciences et technologies du numérique (INRIA), est un modèle basé sur l’architecture RoBERTa, entraîné sur 138Go de texte français. Les données françaises proviennent du corpus multilingue *Open Super-large Crawled ALMAAnaCH coRpus* (OSCAR)¹, un corpus filtré provenant de Common Crawl². Nous affinons le modèle pré-entraîné CamemBERT-base, contenant 110 millions de paramètres, sur notre jeu de données. Nous utilisons également une version distillée de CamemBERT nommée DistilCamemBERT (DELESTRE & AMAR, 2022), contenant deux fois moins de paramètres et ayant un temps d’inférence deux fois plus rapide comparé à CamemBERT-base.

Affinement de CamemBERT

L’affinement d’un modèle pré-entraîné permet de bénéficier d’une riche information externe à notre jeu de données, ce qui peut améliorer les performances de notre modèle. L’affinement consiste à geler les poids du modèle pré-entraîné CamemBERT (c.-à-d. aucun poids ne sera modifié après la descente du gradient) et à ajouter un classifieur (comme une régression logistique ou un réseau de neurones entièrement connecté) qui prend en entrée les sorties du modèle CamemBERT et qui prédit les classes qui nous intéressent (ici, les 9 catégories de la grille FERA). L’affinement demande moins de moyens de calcul que l’entraînement d’un modèle entier puisque l’on entraîne seulement le classifieur ajouté.

Le modèle est affiné sur dix passages (*epoch*) du jeu de données entier. Un planificateur est utilisé pour contrôler le taux d’apprentissage (*learning rate*) à chaque itération. Pendant les 10% premières itérations (équivalent à un passage sur le jeu de données entier), le taux d’apprentissage passe de 0 à $5 * 10^{-5}$ linéairement.

Le modèle CamemBERT entraîné sur notre jeu de données est correctement affiné au bout de 8 passages sur le jeu de données entier. En effet, après le 8^{ème} passage, le score F1 de notre modèle sur le jeu de données de validation diminue, ce qui montre que le modèle commence à sur-apprendre.

Nous effectuons la même chose pour l’affinement de DistilCamemBERT. Les poids des modèles affinés sur notre jeu de données ainsi que les détails de l’affinement sont disponibles sur HuggingFace : <https://huggingface.co/gustavecortal/camembert-base-cae> (CamemBERT) et <https://huggingface.co/gustavecortal/distilcamembert-base-cae> (DistilCamemBERT)

Nous voyons sur le tableau 2.7 que DistilCamemBERT et CamemBERT obtiennent respectivement 12.1 et 13.3 points de score F1 de plus que le meilleur modèle d’apprentissage automatique de la

1. <https://oscar-corpus.com>

2. Common Crawl est un répertoire ouvert de données récupérées sur des pages web : <https://commoncrawl.org>

modèle	précision	support	score F1
DistilCamemBERT	0.884	0.884	0.883
CamemBERT	0.897	0.895	0.895

TABLEAU 2.7: Scores obtenus sur DistilCamemBERT et CamemBERT.

section précédente, le séparateur à vaste marge. Ce score est impressionnant et montre l'intérêt du pré-entraînement et de la prise en compte de la contextualité des phrases pour notre problème. Nous pouvons utiliser ces modèles pour prédire les réponses qui n'ont pas pu être catégorisées avec le filtrage par expression régulière, les réponses ayant pour catégorie RIEN.

2.2.4 Classification des émotions

Nous effectuons maintenant le même raisonnement (c.-à-d. entraînement des régressions logistiques, forêts aléatoires, séparateurs à vaste marge et des modèles CamemBERT) pour la prédiction des émotions (peur, tristesse, joie et colère) à partir des réponses des différentes catégories. Nous voulons savoir qu'elle est le composant de l'émotion (ressentis physiques, pensées, etc.) le plus pertinent pour identifier une émotion. Chaque modèle est entraîné sur les réponses d'une catégorie seulement. Nous excluons les réponses des grilles ayant plusieurs émotions identifiées. C'est un problème de multiclassification monolabel puisque chaque réponse correspond à une émotion seulement. Il serait également intéressant de prendre en compte les grilles ayant plusieurs émotions, et donc se placer dans un problème de multiclassification multilabel.

catégorie	modèle	précision	support	score F1
TERRITOIRE	Régression logistique	0.732 ± 0.019	0.731 ± 0.022	0.73 ± 0.021
TERRITOIRE	Forêt aléatoire	0.718 ± 0.026	0.712 ± 0.023	0.706 ± 0.025
TERRITOIRE	Séparateur à vaste marge	0.731 ± 0.021	0.712 ± 0.017	0.702 ± 0.017
VÉRIFICATION FAIT	Régression logistique	0.476 ± 0.036	0.372 ± 0.041	0.364 ± 0.026
VÉRIFICATION FAIT	Forêt aléatoire	0.425 ± 0.096	0.365 ± 0.053	0.342 ± 0.053
VÉRIFICATION FAIT	Séparateur à vaste marge	0.465 ± 0.082	0.361 ± 0.048	0.346 ± 0.042
FAIT INTÉRIEUR PENSÉES	Régression logistique	0.674 ± 0.068	0.677 ± 0.04	0.657 ± 0.046
FAIT INTÉRIEUR PENSÉES	Forêt aléatoire	0.702 ± 0.045	0.638 ± 0.038	0.594 ± 0.035
FAIT INTÉRIEUR PENSÉES	Séparateur à vaste marge	0.705 ± 0.058	0.663 ± 0.042	0.631 ± 0.036
ACTION INTÉRIEURE	Régression logistique	0.615 ± 0.048	0.619 ± 0.033	0.597 ± 0.029
ACTION INTÉRIEURE	Forêt aléatoire	0.675 ± 0.085	0.604 ± 0.021	0.533 ± 0.028
ACTION INTÉRIEURE	Séparateur à vaste marge	0.621 ± 0.074	0.586 ± 0.023	0.526 ± 0.034
FAIT EXTÉRIEUR	Régression logistique	0.633 ± 0.012	0.614 ± 0.013	0.59 ± 0.021
FAIT EXTÉRIEUR	Forêt aléatoire	0.619 ± 0.026	0.618 ± 0.03	0.574 ± 0.034
FAIT EXTÉRIEUR	Séparateur à vaste marge	0.663 ± 0.065	0.597 ± 0.029	0.545 ± 0.042
FAIT INTÉRIEUR RESENTIS	Régression logistique	0.749 ± 0.026	0.738 ± 0.028	0.736 ± 0.024
FAIT INTÉRIEUR RESENTIS	Forêt aléatoire	0.756 ± 0.042	0.721 ± 0.035	0.705 ± 0.034
FAIT INTÉRIEUR RESENTIS	Séparateur à vaste marge	0.737 ± 0.041	0.705 ± 0.033	0.684 ± 0.03
ACTION TEST	Régression logistique	0.54 ± 0.091	0.6 ± 0.069	0.543 ± 0.045
ACTION TEST	Forêt aléatoire	0.636 ± 0.115	0.69 ± 0.037	0.608 ± 0.055
ACTION TEST	Séparateur à vaste marge	0.615 ± 0.12	0.67 ± 0.039	0.578 ± 0.042
ACTION EXTÉRIEURE	Régression logistique	0.591 ± 0.077	0.604 ± 0.051	0.574 ± 0.068
ACTION EXTÉRIEURE	Forêt aléatoire	0.598 ± 0.145	0.575 ± 0.041	0.477 ± 0.075
ACTION EXTÉRIEURE	Séparateur à vaste marge	0.574 ± 0.109	0.57 ± 0.034	0.474 ± 0.067

TABLEAU 2.8: Scores obtenus à partir des réponses des différentes catégories sur la régression logistique, la forêt aléatoire et le séparateur à vaste marge.

Nous voyons, sur le tableau 2.8, que la catégorie qui permet le mieux d’identifier les émotions est FAIT INTÉRIEUR RESENTIS avec un score F1 de 0.736 avec la régression logistique. Analyser ses ressentis physiques aide donc à identifier les émotions que l’on ressent, comme préconisé par FINKEL (2022). La catégorie TERRITOIRE vient en deuxième position avec un score F1 de 0.73 avec la régression logistique. Analyser l’état d’un territoire important (par exemple, s’il subit une attaque ou non) permet d’identifier une émotion vécue, c’est l’un des objectifs principaux de l’ACE. Il serait intéressant d’entraîner un modèle qui prend en compte plusieurs catégories *en même temps* (par exemple, TERRITOIRE, FAIT INTÉRIEUR RESENTIS et FAIT INTÉRIEUR PENSÉES) et voir si les performances sont meilleures.

Nous entraînons également un modèle CamemBERT à partir des FAITS INTÉRIEURS RESENTIS et nous obtenons un score F1 de 0.775, soit 3.9 points de plus que le meilleur modèle précédent (0.736). L’écart entre les modèles est moins important par rapport à la prédiction des catégories des grilles dans la section précédente. Cela montre que la prise en compte de la contextualité est, peut-être, moins importante pour la bonne prédiction des émotions. L’analyse des caractéristiques, obtenue avec la méthode de l’importance des permutations, pour la prédiction des différentes émotions à partir des FAITS INTÉRIEURS RESENTIS est disponible en Annexe (Figure 3.7).

2.2.5 Méthodes d’optimisations des modèles

Si dans le futur nous voulons intégrer nos modèles dans une application, il faudra essayer de réduire le plus possible le temps d’inférence tout en préservant les performances, surtout si nous voulons utiliser des modèles demandant beaucoup de moyens de calcul comme CamemBERT. C’est pourquoi la distillation couplée à d’autres optimisations comme la quantification ou l’élagage peuvent être de bonnes solutions pour la mise en production.

Distillation

La distillation des connaissances consiste à compresser la connaissance d’un large modèle professeur (CamemBERT), dans un plus petit modèle élève (DistilCamemBERT). Le modèle élève possède une architecture moins complexe par rapport à l’architecture du modèle professeur, ce dernier possédant plus de paramètres. Le modèle élève est entraîné à partir du modèle professeur déjà entraîné. Le but du modèle élève, lors de l’entraînement, est d’imiter le comportement du modèle professeur en essayant de reproduire les sorties de ses différentes couches, et pas seulement la sortie finale. Nous voyons que sur notre jeu de données, DistilCamemBERT possède des performances très légèrement inférieures à CamemBERT tout en ayant une vitesse d’inférence 2 fois plus rapide.

Quantification

La quantification consiste à approximer les poids d’un modèle. Il existe deux sortes de quantification : la quantification durant l’entraînement et la quantification après l’entraînement. Nous allons parler seulement de la quantification après entraînement puisqu’elle est la plus simple à implémenter. La quantification d’entiers est une stratégie d’optimisation qui convertit les nombres à virgule flottante 32 bits, tels que les poids et les sorties d’activation d’un réseau de neurones, en nombres à virgule fixe 8 bits les plus proches. Cela se traduit par un modèle quatre fois plus petit ainsi qu’une vitesse d’inférence trois fois plus rapide¹. Il existe également la conversion des poids en valeurs à virgule flottante 16 bits qui permet de diviser par deux la taille du modèle. Bien entendu, il faut s’assurer que la dégradation est acceptable et ne diminue pas trop les performances du modèle.

1. https://www.tensorflow.org/lite/performance/post_training_quantization

Chapitre 3

Identification automatique des émotions et de leurs rôles sémantiques

3.1 Jeu de données français pour l’analyse des émotions

Certains corpus avec une annotation des émotions existent pour le français, par exemple le jeu de données de la campagne Défi Fouille de Textes (DEFT) de 2018 (PAROUBEK et al., 2018) ou le corpus de reconnaissance des émotions dans les livres pour enfants (ÉTIENNE et al., 2020). Cependant, ils ne peuvent pas être utilisés pour l’ACE pour plusieurs raisons : le matériel textuel ou le modèle d’émotion est incompatible, il n’y a pas assez de données pour l’entraînement du modèle, etc.

Le jeu de données, composé des récits autobiographiques d’une scène émotionnelle extraits dans le chapitre précédent, devra indiquer les émotions et leurs rôles sémantiques : INDICE (un marqueur qui indique l’expression d’une émotion, pouvant être un seul mot), SUJET (l’auteur qui ressent une émotion), CIBLE (une entité ou une personne ciblée par une émotion) et STIMULUS (un événement qui déclenche une émotion). Ces rôles sont utilisés par CAMPAGNANO et al. (2022) pour unifier plusieurs ensembles de données hétérogènes qui contiennent des annotations à la fois pour les émotions et leurs rôles sémantiques. Ils adoptent la roue des émotions de Plutchik (PLUTCHIK, 2001) pour la définition des émotions à prédire.

Au lieu de considérer l’analyse des émotions comme un problème de classification au niveau de la phrase, nous nous concentrons au niveau de l’aspect. Nous proposons de comprendre en profondeur un texte donné décrivant une scène émotionnelle, en identifiant automatiquement qui ressent une émotion, ce qui pousse une entité à exprimer une émotion envers un certain aspect et pourquoi. Par exemple, dans cette phrase, « Gustave aime les plantes carnivores parce qu’elles sont belles », Gustave (SUJET) expose sa joie (INDICE) envers les plantes carnivores (CIBLE) parce qu’elles sont belles (STIMULUS).

3.2 Nouveau schéma d’annotation

Nous proposons d’étendre le schéma d’annotation avec de nouveaux rôles sémantiques basés sur l’ACE afin de mieux comprendre les raisons d’une émotion. Nous introduisons de nouveaux rôles sémantiques, TERRITOIRE et OBJET, correspondant respectivement à la notion de territoire et d’objet dans l’ACE. Nous introduisons ATTAQUE correspondant aux expressions liées à l’acte d’attaquer ou d’être attaqué (par exemple, attaque, assaut, agression, etc.) et ATTAQUANT correspondant à

l'entité qui attaque un TERRITOIRE. L'identification préalable d'ATTAQUANT et ATTAQUE facilite l'identification de TERRITOIRE. Par exemple, dans la phrase « Mes compétences sont attaquées par Marc », « Mes compétences » est un TERRITOIRE lié aux valeurs professionnelles de l'auteur et à son image de soi compétente attaquée par l'ATTAQUANT « Marc ». Ces nouveaux rôles sémantiques peuvent être considérés comme un raffinement de la notion de STIMULUS présentée ci-dessus. Nous introduisons également deux rôles complémentaires : MODIFICATEUR pour prendre en compte l'intensité d'une émotion (par exemple, « Je suis **un peu** triste ») et NÉGATION pour préserver le sens original des expressions utilisant des marqueurs de négation (par exemple, « Elle **n'était pas** en colère »).

3.3 Méthode basée sur les règles

Nous proposons une méthode basée sur la création manuelle de règles pour automatiquement identifier les émotions et leurs rôles sémantiques dans des récits autobiographiques d'une scène émotionnelle. Nous tirons parti des caractéristiques linguistiques en utilisant l'analyse syntaxique des dépendances, la résolution des coréférences et le marquage des parties du discours avec la bibliothèque de TAL SpaCy. Pour la résolution des coréférences, la bibliothèque crosslingual coreference¹ est utilisée. Celle-ci rajoute à la *pipeline* SpaCy une analyse des coréférences en utilisant des modèles entraînés sur des données anglaises, mais avec des représentations multilingues. Les liens de coréférence sont utilisés pour connecter différentes expressions se référant à un unique référent (par exemple, « Je suis en colère contre **lui**. **Il** a cassé mon vélo »). En effet, il est utile d'identifier les multiples occurrences d'un même SUJET ou d'une même CIBLE pour mieux comprendre le flux émotionnel d'un texte.

Nous utilisons WordNet (MILLER, 1994), une base de données lexicale de relations sémantiques comprenant des synonymes, des hyponymes et des méronymes, pour identifier les INDICES et les mots liés à une ATTAQUE d'un TERRITOIRE. Pour la langue française, nous choisissons le WordNet français appelé WOLF (SAGOT & FIŠER, 2008). Les lexiques de sentiments et d'émotions sont également utilisés pour améliorer l'identification des INDICES. SentiWordNet (BACCIANELLA et al., 2010) est construit à partir de WordNet. Dans ce lexique, chaque sens d'un mot est associé à un degré de positivité, de négativité et de neutralité. Le lexique NRC (S. M. MOHAMMAD & TURNEY, 2013) est un autre lexique populaire, où chaque mot est associé à huit émotions de base (colère, peur, anticipation, confiance, surprise, tristesse, joie et dégoût) et à deux polarités (négatif et positif).

Comme nous travaillons avec des textes autobiographiques, l'auteur d'un récit se décrit souvent avec des pronoms à la première personne. Il est donc facile de détecter le SUJET avec un filtrage à l'aide d'expressions régulières. Pour identifier les rôles sémantiques complexes tels que le TERRITOIRE, nous définissons manuellement plusieurs règles en utilisant des caractéristiques linguistiques. Par exemple, un TERRITOIRE est trouvé s'il est le sujet d'une ATTAQUE à la voix passive : « **Mes compétences** sont attaquées par Marc ».

Les méthodes basées sur des règles ne nécessitent pas de données d'entraînement. L'interprétation des résultats est l'un de leurs principaux avantages. Cependant, il est parfois difficile de formuler des règles et une tâche peut nécessiter un grand nombre de règles, ce qui rend la méthode très spécifique au domaine. Le temps de vérification de la cohérence et de l'homogénéité augmente considérablement avec le nombre de règles. Notons qu'au moment de la rédaction de ce rapport, l'annotation du jeu de données avec notre nouveau schéma d'annotation ainsi que l'évaluation du premier prototype basé sur des règles n'a pas encore été effectuées, mais elle le sera dans un avenir proche.

Dans de futurs travaux, nous combinerons notre méthode basée sur les règles avec des techniques récentes d'apprentissage profond pour prendre le meilleur des deux approches. Par exemple, l'architecture Transformer (VASWANI et al., 2017) est efficace pour la modélisation du contexte. RAY

1. <https://github.com/pandora-intelligence/crosslingual-coreference>

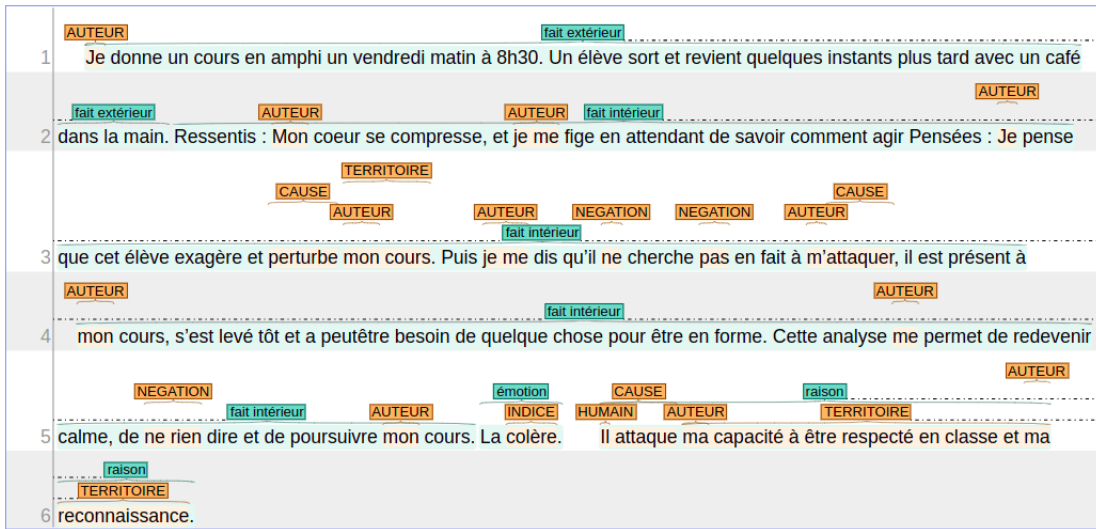


FIGURE 3.1: Identification automatique des rôles sémantiques avec la méthode basée sur les règles. Affichage des annotations dans la plateforme INCEpTION.

et CHAKRABARTI (2022) proposent de combiner une méthode à base de règles avec un réseau de neurones convolutif pour améliorer les performances de l'extraction des aspects. LI et al. (2021) montrent qu'un réseau de neurones peut générer de nouvelles règles basées sur des règles manuellement construites de haute qualité (*seed rules*). Dans notre cas, les règles apprises pourraient expliquer les rôles sémantiques prédits, et donc découvrir de nouvelles façons d'identifier le contenu émotionnel d'un texte en le formalisant.

3.4 Structure de graphe

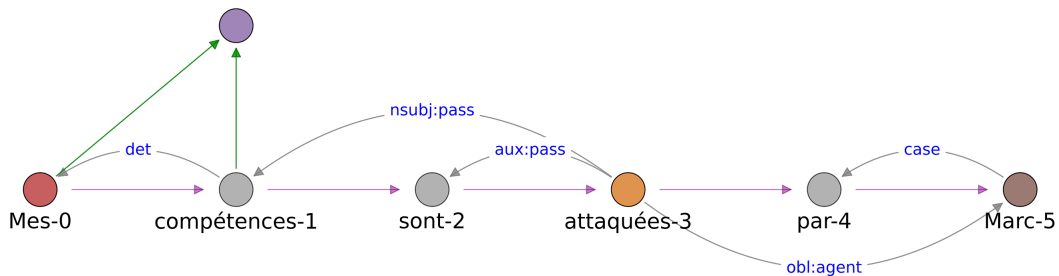


FIGURE 3.2: Visualisation de la phrase : « Mes compétences sont attaquées par Marc ». **Les couleurs des arêtes** indiquent les différents types de relations, par exemple, l'appartenance à un groupe nominale est en vert et les relations séquentielles sont en violet (par exemple, de *Mes-0* à *compétences-1*). **Les couleurs des nœuds** indiquent les rôles sémantiques, par exemple le SUJET est en rouge (*Mes-0*), le TERRITOIRE est en violet (*Mes-0 compétences-1*), l'ATTAQUE (*attaquées-3*) est en jaune et l'ATTAQUANT (*Marc-5*) est en marron.

Une émotion est un phénomène complexe qui résonne dans plusieurs niveaux d'analyse à travers différentes échelles. Nous proposons de représenter l'expression d'une émotion par une structure de graphe qui peut être visualisée. Une phrase ou un texte entier correspond à un graphe dans lequel les nœuds sont des mots et les arêtes indiquent des relations de différents types entre les mots.

Nous incorporons notre méthode à base de règles dans la structure du graphe. La figure 3.2 illustre l'application de visualisation construite avec la bibliothèque de graphe NetworkX¹. L'application peut afficher différents niveaux d'analyse du texte (comme notre analyse des émotions et l'analyse des dépendances) sur un plan unique. Par exemple, des liens de coréférence relient certains aspects des émotions entre eux. Le visualiseur de relations peut donc être utilisé pour faciliter la création manuelle de règles.

Nous prévoyons d'augmenter la structure du graphe avec de nouvelles relations sémantiques en extrayant les chemins de connaissance de ConceptNet (SPEER et al., 2017). Il s'agit d'un réseau sémantique multilingue qui fournit des concepts reliés par de grandes quantités de relations sémantiques (par exemple, relation d'appartenance, de lieu, d'utilité, etc.). Pour rappel, un exemple de réseau sémantique est illustré en Annexe (Figure 3.8). Par exemple, YAN et al. (2021) incorporent des connaissances du sens commun, provenant de ConceptNet, pour réduire le biais de position dans les modèles d'Extraction de la Cause d'une Émotion (ECE). Nous pensons que les réseaux sémantiques peuvent être utiles pour mieux capturer les dépendances entre une émotion et ses rôles sémantiques, car ils tirent parti de la connaissance du sens commun.

Notre prochain objectif, pour l'amélioration des performances, est de concevoir des réseaux de neurones pour graphes pour capturer les éléments structurels spécifiques à la structure que nous proposons. Par exemple, MARCHEGGIANI et TITOV (2017) exploitent les informations syntaxiques en utilisant des réseaux convolutifs sur les graphes pour encoder les phrases, car les représentations sémantiques sont similaires aux représentations syntaxiques. Les méthodes proposées peuvent être combinées et étendues pour exploiter d'autres informations que la structure de graphe offre.

1. <https://networkx.org>

Conclusion

Pour remédier à certaines limitations présentes dans l'analyse des émotions, nous proposons de comprendre en profondeur une scène émotionnelle en effectuant une analyse fine d'une émotion et de ses rôles sémantiques au niveau de l'aspect. L'Analyse Cognitive des Émotions ainsi que les différents courants de recherche en TAL qui nous intéressent ont été présentés. Nous introduisons un nouveau schéma d'annotation, basé sur l'Analyse Cognitive des Émotions, ainsi qu'un nouveau jeu de données en français composé de récits autobiographiques d'une scène émotionnelle. Comme l'analyse manuelle des récits prend du temps, nous fournissons un assistant automatisé pour l'entraîneur, qui peut ainsi se concentrer sur les aspects qui ne peuvent pas être soumis au traitement automatique. Notre méthode basée sur des règles identifie automatiquement les émotions et leurs rôles sémantiques dans un texte. À l'avenir, après avoir annoté les récits autobiographiques avec notre schéma d'annotation et effectué l'évaluation quantitative de notre méthode à base de règles, nous prévoyons de la combiner avec des modèles récents d'apprentissage profond (comme les réseaux de neurones pour graphes) grâce à la structure de graphe que nous avons développée pour améliorer les performances.

Les travaux effectués durant ce stage font l'objet d'un papier nommé *Natural Language Processing for Cognitive Analysis of Emotions*, accepté pour la conférence *Semantics, Memory and Emotion*¹ (2022). Un poster a également été accepté pour la conférence *European Mathematical Psychology Group*² (2022). Je suis heureux de pouvoir poursuivre mon travail sur l'identification automatique des émotions et de leurs causes en effectuant un doctorat au sein de la même équipe!

1. <https://sme-2022.llf-paris.fr>

2. <https://event.unitn.it/empg2022/>

Annexe

Faits extérieurs	Je suis dans ma chambre en Angleterre, j'ai devant moi une assiette de tomates et œufs durs et je m'apprête à verser dessus la nouvelle sauce de salade que je viens d'acheter. Je verse, je goûte, cette sauce est infâme.
Faits intérieurs	Ressentis : Tous mes muscles se relâchent, je m'avachis sur mon siège et je regarde dans le vague. Pensées : Je revois mon ancienne sauce de salade qui était si bonne, je me revois acheter cette nouvelle sauce en me disant « Tiens, on va goûter ».
Identifier les émotions vécues et en choisir une	La tristesse et la colère. Je choisis la colère.
Quels territoires sont attaqués (colère et peur), perdus (tristesse), retrouvés ou conquis (joie) ?	Mon sens du goût attaqué par cette sauce infâme, mes habitudes et la sensation de sécurité qu'elles procurent ont été attaquées ainsi que la connaissance de ce qui m'entoure, la sécurité de la salade tomate-œuf-sauce que j'ai l'habitude de manger.
Êtes-vous satisfait (tests des pensées et ressentis) de vos actions dans la scène ?	NON
Actions intérieures : analyser le passé et revivre au présent. Quelle mise à jour éventuelle de votre base de connaissances ?	Je n'essaierai plus les aliments bizarres qu'ils mangent ici.
Actions extérieures : s'il ne vous reste rien à faire dans le monde extérieur alors ARRÊT. Sinon tester plusieurs actions possibles, choisissez-en une (ou pas) et faites-la.	J'ai arrêté la salade et j'ai mangé des petits suisses que j'avais déjà goûtés auparavant (importés de France).

TABLEAU 3.1: Exemple d'une grille FERA complétée.

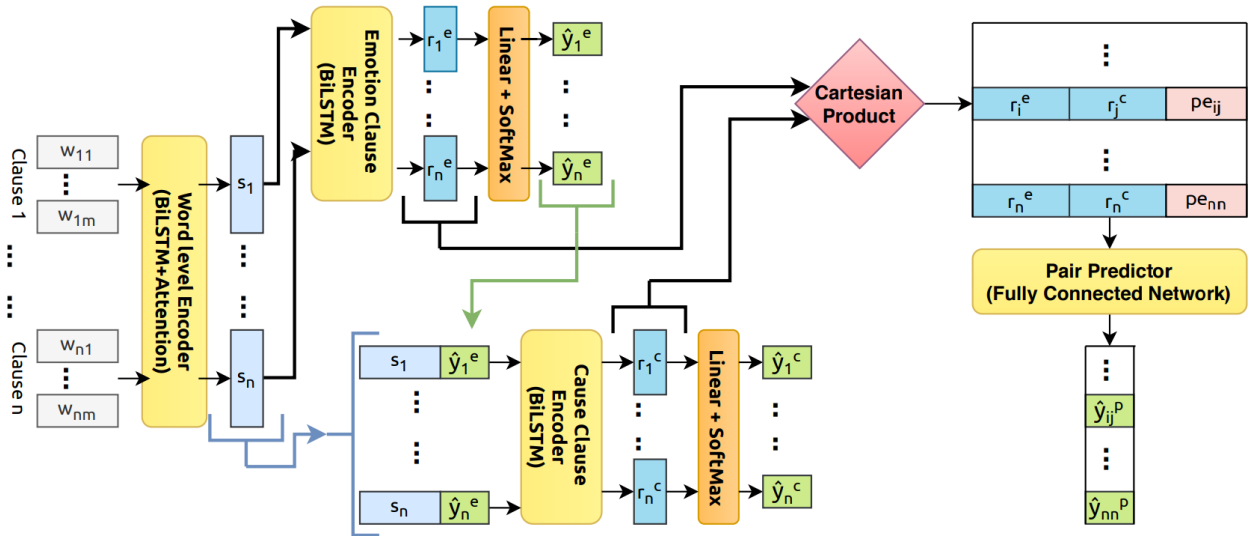


FIGURE 3.3: Modèle « bout-à-bout » proposé par SINGH et al. (2021) pour l'EPEC.

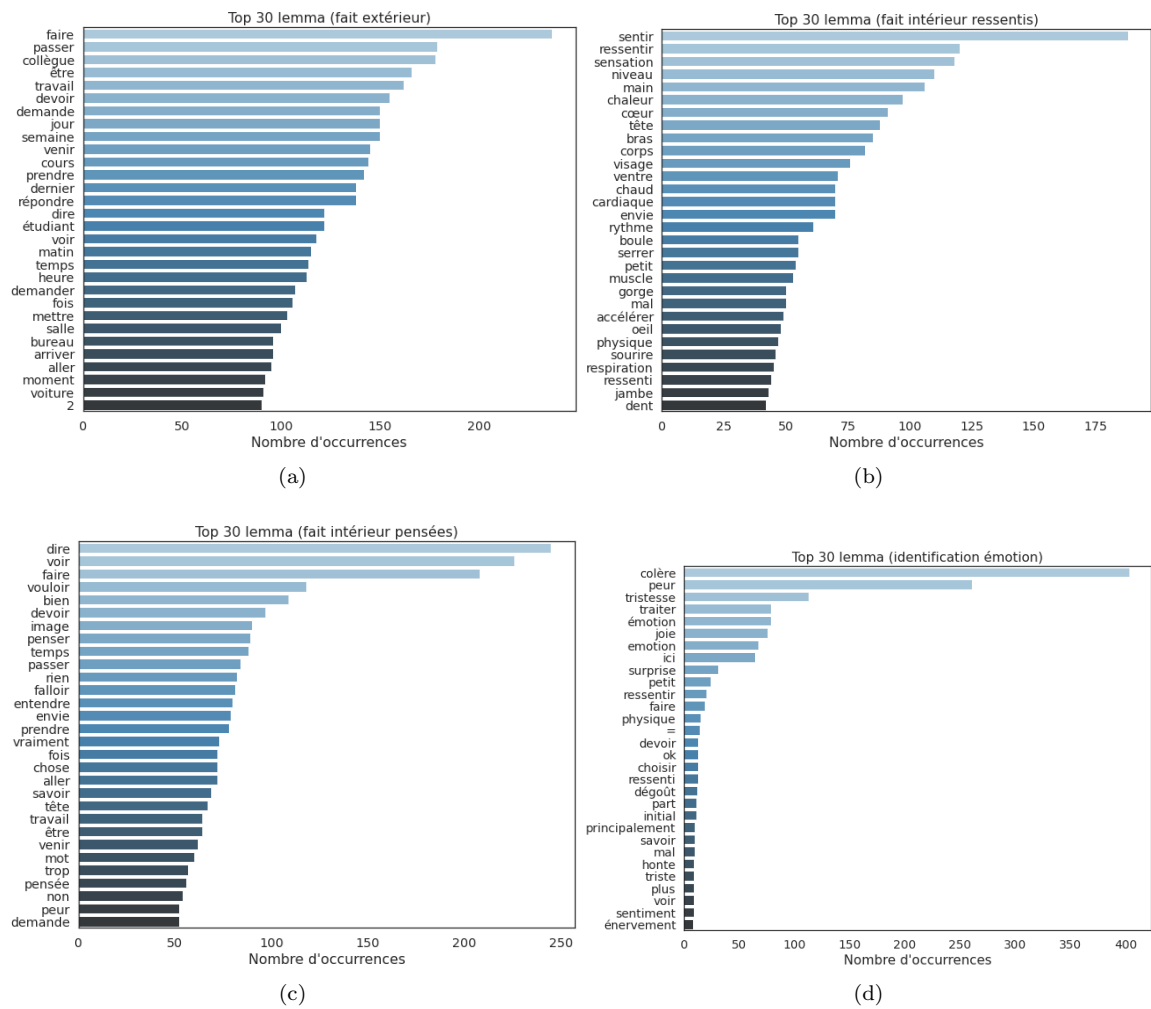


FIGURE 3.4: Distribution des lemmes les plus fréquents selon les catégories FAIT EXTÉRIEUR, FAIT INTÉRIEUR RESENTIS, FAIT INTÉRIEUR PENSÉES et IDENTIFICATION ÉMOTION.

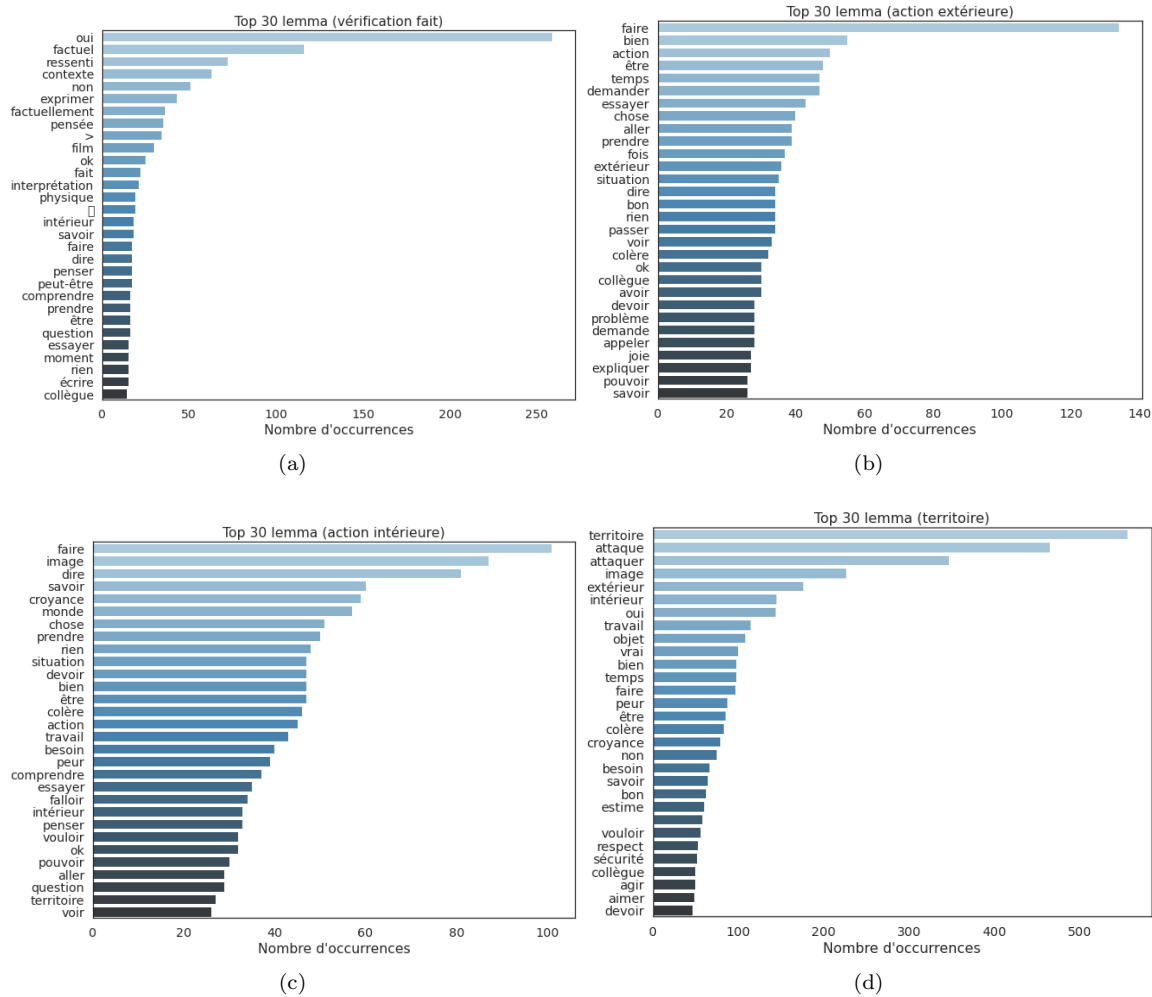


FIGURE 3.5: Distribution des lemmes les plus fréquents selon les catégories VÉRIFICATION FAIT, ACTION EXTÉRIEUR, ACTION INTÉRIEUR, TERRITOIRE.

y=fait extérieur top features		y=fait intérieur pensées top features		y=fait intérieur ressentis top features		y=identification émotion top features		y=territoire top features		y=vérification fait top features	
Weight ²	Feature	Weight ²	Feature	Weight ²	Feature	Weight ²	Feature	Weight ²	Feature	Weight ²	Feature
+3.118	matin	+3.285	dis	+3.892	cœur	+7.029	colère	+6.428	territoire	+5.317	oui
+2.444	collègue	+2.890	vois	+3.358	chaleur	+6.091	peur	+5.367	attaque	+5.262	factuel
+2.440	midi	+2.780	qu	+3.244	sens	+4.720	tristesse	+4.646	attaqué	+3.939	contexte
+2.427	soir	+2.749	envie	+3.198	mains	+3.466	joie	+4.243	image	+3.480	factuellement
+2.414	bureau	+2.503	imagine	+3.121	sensation	+2.668	emotion	+3.867	objet	+3.274	non
+2.292	semaine	+2.412	images	+3.010	bras	+2.345	émotion	+3.259	attaquée	+2.858	ok
+2.258	ans	+2.400	mots	+2.991	chaud	+2.167	traitée	+2.621	sécurité	+2.576	interprétation
+2.249	salle	+2.224	vraiment	+2.980	corps	+2.132	ici	+2.607	extérieure	+2.567	ressentis
+2.235	dernier	+2.061	imaginé	+2.931	niveau	+1.980	surprise	+2.394	confiance	+2.485	factuelle
+2.089	voiture	+2.012	pensé	+2.850	ventre	+1.584	dégoût	+2.335	estime	+2.393	faits
+1.965	jours	+1.994	revois	+2.681	boule	+1.571	agacement	+2.180	travail	+2.260	pensées
+1.929	répond	+1.847	voyais	+2.403	yeux	+1.505	triste	+2.009	intérieure	+1.729	exprimé
+1.915	dernière	+1.805	merde	+2.376	ressens	+1.377	petite	+1.948	attaqués	+1.628	émotion
+1.889	jour	+1.706	insultes	+2.312	cardiaque	+1.331	emotions	+1.925	respect	+1.583	essayé
+1.743	vendredi	+1.621	faut	+2.287	dents	+1.300	frustration	+1.921	séparation	+1.541	note
+1.695	mois	+1.608	veut	+2.283	gorge	+1.252	stress	+1.912	territoires	+1.526	film
+1.678	cours	+1.581	content	+2.259	tête	+1.209	inquiétude	+1.910	aimé	+1.450	factuelles
+1.669	semaines	+1.507	oh	+2.146	rythme	+1.174	impatience	+1.881	crédibilité	+1.443	exprimés
+1.655	train	+1.493	espère	+2.116	visage	+1.127	émotions	+1.868	oui	+1.430	sorti
+1.614	entreprise	+1.468	scène	+2.052	sourcils	+1.126	énervement	+1.852	compétences	+1.418	vérifier
+1.607	dimanche	+1.440	veux	+2.022	sourire	+1.123	angoisse	+1.797	capacité	+1.389	succèdent
+1.577	étudiants	+1.430	exagère	+1.958	envie	+1.116	penseront	+1.751	temps	+1.361	précédemment
+1.524	passé	+1.365	pourvu	+1.930	ressenti	+1.051	ressentie	+1.494	séparé	+1.316	factuels
+1.508	fin	+1.343	jamais	+1.903	plexus	+1.044	euphorie	+1.485	extérieur	+1.297	parti
+1.507	stage	+1.337	falloir	+1.888	ressentis	+1.009	collère	+1.445	liberté	+1.280	changerais
+1.503	personnes	+1.314	bien	+1.857	muscles	+1.004	frayeur	+1.444	présent	+1.236	ressenti
+1.487	paris	+1.302	ca	+1.775	respiration	+0.946	mépris	+1.333	intégrité	+1.233	soucis
+1.463	maison	+1.246	pensées	+1.740	serre	+0.940	honte	+1.322	tranquillité	+1.199	frisson
+1.451	soirée	+1.150	mauvais	+1.711	poitrine	+0.920	exaspération	+1.296	présence	+1.194	ordre
+1.430	sommes	+1.120	quelqu	+1.691	épaules	+0.898	béatitude	+1.241	croissance	+1.147	yes
+1.421	ami	+1.085	disais	+1.658	jambes	+0.898	colère	+1.230	groupe	+1.118	vérification
+1.412	amie	+1.075	devoir	+1.653	estomac	+0.893	colere	+1.217	vie	+1.104	inquiète
+1.410	mail	+1.065	prend	+1.603	tension	+0.882	dégout	+1.210	agit	+1.076	nulle
+1.386	examen	+1.052	musique	+1.558	fatigue	+0.877	stupeur	+1.187	intérieur	+1.072	exagéré
+1.381	lieu	+1.047	moments	+1.539	accélération	+0.857	bonheur	+1.178	reconnaissance	+1.070	superpose
+1.366	lundi	+1.043	ensuite	+1.507	coeur	+0.785	tract	+1.133	physique	+1.070	tri
+1.320	mars	+1.020	rentrer	+1.489	température	+0.781	uniquement	+1.106	manque	+1.055	75
+1.306	téléphone	+1.016	mince	+1.477	crispé	+0.780	probablement	+1.093	besoin	+1.054	weekend
+1.306	moment	+1.015	mauvaise	+1.455	souffle	+0.777	fierté	+1.073	physiquement	+1.009	grande
+1.296	arrive	+1.001	sons	+1.399	énergie	+0.738	principalement	+1.046	biens	+0.994	haute
+1.288	minutes	+0.997	revoit	+1.372	bouche	+0.728	panique	+1.027	espace	+0.991	habitude
+1.280	demande	+0.987	tête	+1.371	mâchoire	+0.712	déçu	+0.969	droit	+0.988	10
+1.265	internet	+0.982	perdu	+1.366	accélère	+0.707	rougis	+0.952	mort	+0.984	prononcé
+1.261	demandé	+0.980	idem	+1.359	tremblements	+0.692	forcément	+0.950	idée	+0.981	exprimée
+1.254	site	+0.971	aime	+1.341	mal	+0.680	bases	+0.939	vraie	+0.961	raconte
+1.249	déjeuner	+0.948	voir	+1.332	sentais	+0.678	anxiété	+0.934	externe	+0.957	déception
+1.246	commence	+0.903	aie	+1.309	poings	+0.662	physiques	+0.932	argent	+0.953	raconter
+1.246	laboratoire	+0.895	con	+1.267	légère	+0.661	permanente	+0.925	compétence	+0.924	intérieurs
+1.231	samedi	+0.893	reproches	+1.245	rouge	+0.661	culpabilité	+0.902	bon	+0.920	reparle
+1.226	tour	+0.878	visualise	+1.245	froid	+0.660	jalousie	+0.884	lol	+0.908	partie

FIGURE 3.6: Importance des caractéristiques pour la prédiction des catégories. Le modèle entraîné est la régression logistique. Le tableau est tronqué.

y=['colère'] top features		y=['joie'] top features		y=['peur'] top features		y=['tristesse'] top features	
Weight ²	Feature	Weight ²	Feature	Weight ²	Feature	Weight ²	Feature
+1.627	chaleur	+3.485	sourire	+1.340	ventre	+1.554	triste
+1.346	dents	+1.541	joie	+1.291	cœur	+1.409	poids
+1.285	bras	+1.207	rire	+1.154	moites	+1.332	lourd
+0.962	sourcils	+1.156	souris	+1.110	jambes	+1.165	réchauffement
+0.933	mâchoire	+1.105	muscles	+0.916	froides	+1.146	yeux
+0.907	besoin	+1.088	agréable	+0.806	coeur	+1.130	larme
+0.895	colère	+1.036	détente	+0.780	ressenti	+1.097	pleurer
+0.808	poings	+1.023	content	+0.713	oppression	+1.052	larmes
+0.801	augmentation	+0.999	détendue	+0.686	ressentis	+1.036	pincement
+0.783	mains	+0.990	duvet	+0.634	peur	+1.035	sol
+0.779	visage	+0.957	beaucoup	+0.631	fort	+1.032	soupirs
+0.743	serre	+0.850	souriais	+0.611	tombent	+1.027	déglutition
+0.732	crispé	+0.822	trépigne	+0.610	noue	+1.027	niveau
+0.690	mâchoires	+0.821	lever	+0.610	noue	+0.978	émotion
+0.654	bouger	+0.811	relâchement	+0.604	panique	+0.912	rien
+0.637	enervement	+0.810	grande	+0.596	tremblements	+0.908	sol
+0.627	pression	+0.776	profonde	+0.560	stress	+0.861	pointe
+0.606	agite	+0.776	profonde	+0.534	vite	+0.854	fatigué
+0.602	moment	+0.764	profondément	+0.530	tachycardie	+0.829	regardant
+0.588	aucun	+0.749	détendu	+0.524	tremblent	+0.826	picotent
+0.577	température	+0.741	pouffe	+0.522	sueurs	+0.826	douloureuse
+0.576	crispée	+0.741	euphorique	+0.510	coupé	+0.820	lourde
+0.565	serré	+0.741	facilement	+0.508	rapides	+0.804	tristesse
+0.524	voix	+0.689	légèreté	+0.499	accélération	+0.802	perte
+0.514	fatigue	+0.688	accélèrent	+0.491	tremble	+0.791	explose
+0.503	froncés	+0.679	pétitement	+0.481	frissons	+0.791	sanglot
+0.490	serrent	+0.629	chaude	+0.455	malaise	+0.777	corps
+0.467	fortement	+0.629	bonheur	+0.447	resserre	+0.763	manque
+0.462	serrées	+0.629	papillons	+0.447	nœud	+0.757	sternum
+0.457	monte	+0.621	agréable	+0.446	bougent	+0.750	épaules
+0.443	poing	+0.621	reguliere	+0.445	vocales	+0.733	tête
+0.436	contractés	+0.613	parlai	+0.444	coup	+0.723	gorge
+0.434	crispation	+0.607	gestes	+0.444	ordinateur	+0.713	œil
+0.416	hausse	+0.587	sensation	+0.420	seconde	+0.710	fragilité
+0.412	serrés	+0.583	calme	+0.415	bouffée	+0.691	plomb
+0.394	cherche	+0.557	vite	+0.410	souffle	+0.691	chape
+0.382	crispe	+0.550	douloureux	+0.410	souffle	+0.677	noué
+0.380	énergé	+0.545	technologie	+0.408	sèche	+0.675	tranquille
+0.371	chauffe	+0.545	fête	+0.406	court	+0.672	faible
+0.371	table	+0.542	complètement	+0.404	creux	+0.646	alourdir
+0.370	aise	+0.542	petit	+0.404	négatif	+0.646	enthousiasme
+0.366	pouls	+0.541	impatience	+0.402	bouche	+0.645	assois
+0.366	joues	+0.520	poil	+0.398	coupe	+0.641	thorax
+0.365	crane	+0.520	altitude	+0.396	main	+0.636	soutien
+0.362	sortir	+0.520	sommes	+0.394	frisson	+0.636	accolade
+0.359	afflux	+0.520	3400m	+0.391	verbal	+0.636	faciales
+0.357	chose	+0.518	grand	+0.387	monde	+0.614	dés
+0.353	fermé	+0.515	crier	+0.385	normale	+0.605	traits
+0.353	accélération	+0.515	difficulté	+0.385	maladresse	+0.605	tombe
+0.350	point	+0.508	cage	+0.383	légèrement	+0.605	organes
...	702 more positive ...	+0.499	réalisé	+0.378	figé	+0.592	perds
...	776 more negative	230 more positive ...	+0.369	bloqué	...	154 more positive ...
-0.572	sternum	...	1248 more negative ...	-0.625	muscles	-0.673	rythme
-0.637	tombent	-0.627	mal	-0.681	lourd	-0.904	vite
-0.722	moites	-0.632	niveau	-0.700	sens	-0.907	chaleur
-1.160	cœur	-0.750	tête	-0.729	sang	-0.949	chaud
-1.237	sourire	-0.798	gorge	-1.198	sourire	-1.050	sourire
		-1.018	bras				

FIGURE 3.7: Importance des caractéristiques pour la prédiction des émotions à partir des ressentis physiques. Le modèle entraîné est la régression logistique.

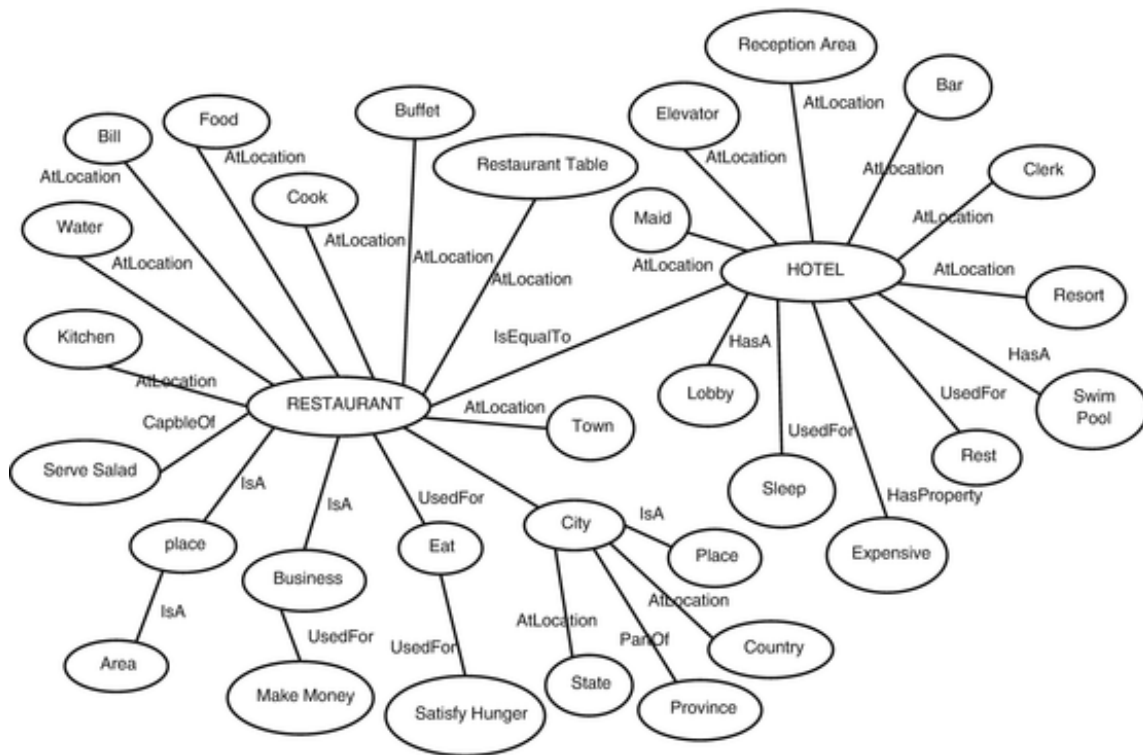


FIGURE 3.8: Réseau sémantique extrait de ConceptNet à partir de deux concepts : RESTAURANT et HOTEL. Image provenant de AGARWAL et MITTAL (2015).

Table des matières

Résumé technique	3
1 Présentation des domaines de recherche	4
1.1 Analyse Cognitive des Émotions	4
1.1.1 Présentation	4
1.1.2 Objet, territoire et fait	4
1.1.3 Système de traitement de l'information émotionnelle	5
1.1.4 Exemple	5
1.1.5 Grille FERA	6
1.2 Analyse du sentiment	8
1.2.1 Présentation	8
1.2.2 Directions de recherche	8
1.3 Identification des causes d'une émotion	10
1.3.1 Extraction de la causalité	10
1.3.2 Extraction des causes d'une émotion	11
1.3.3 Extraction des stimuli : études anglaises et françaises	13
2 Jeu de données composé de scènes émotionnelles	15
2.1 Construction du jeu de données	15
2.1.1 Extraction des grilles	15
2.1.2 Identification des catégories	17
2.1.3 Désidentification des grilles	17
2.1.4 Prétraitement des grilles	18
2.2 Modèles de classification	20
2.2.1 Bibliothèques utilisées	20
2.2.2 Modèles d'apprentissage automatique	20
2.2.3 Modèles de langue pré-entraînés	22
2.2.4 Classification des émotions	24
2.2.5 Méthodes d'optimisations des modèles	25
3 Identification automatique des émotions et de leurs rôles sémantiques	26
3.1 Jeu de données français pour l'analyse des émotions	26
3.2 Nouveau schéma d'annotation	26
3.3 Méthode basée sur les règles	27
3.4 Structure de graphe	28
Conclusion	30
Acronymes	41

Table des figures

1.1	L'esprit comme un système de traitement de l'information émotionnelle. Image provenant de (FINKEL, 2022).	6
1.2	Description des ressentis par FINKEL (2022).	7
1.3	Le modèle tridimensionnel décrit les relations entre les émotions en faisant une analogie avec la roue des couleurs. La dimension verticale du cône représente l'intensité. Le cercle représente les degrés de similarité entre les émotions. Les huit secteurs indiquent qu'il existe huit dimensions primaires, correspondant aux huit émotions primaires définies par la théorie. Les émotions situées dans les espaces vides sont les mélanges de deux émotions primaires. Image provenant de PLUTCHIK (2001).	8
1.4	Représentation par un circuit électronique de la phrase <i>The car is very old but it is rather not expensive</i> . Image provenant de CAMBRIA et HUSSAIN (2015).	9
1.5	L'architecture pour l'extraction des clauses émotions et des clauses causes dans un texte. Image provenant de XIA et DING (2019).	12
2.1	Distribution des catégories.	18
2.2	Nombre moyen de tokens par catégorie.	19
2.3	Matrice de confusion pour la régression logistique.	22
3.1	Identification automatique des rôles sémantiques avec la méthode basée sur les règles. Affichage des annotations dans la plateforme INCEPTION.	28
3.2	Visualisation de la phrase : « Mes compétences sont attaquées par Marc ». Les couleurs des arêtes indiquent les différents types de relations, par exemple, l'appartenance à un groupe nominale est en vert et les relations séquentielles sont en violet (par exemple, de <i>Mes-0</i> à <i>compétences-1</i>). Les couleurs des nœuds indiquent les rôles sémantiques, par exemple le SUJET est en rouge (<i>Mes-0</i>), le TERRITOIRE est en violet (<i>Mes-0 compétences-1</i>), l'ATTAQUE (<i>attaquées-3</i>) est en jaune et l'ATTAQUANT (<i>Marc-5</i>) est en marron.	28
3.3	Modèle « bout-à-bout » proposé par SINGH et al. (2021) pour l'EPEC.	32
3.4	Distribution des lemmes les plus fréquents selon les catégories FAIT EXTÉRIEUR, FAIT INTÉRIEUR RESENTIS, FAIT INTÉRIEUR PENSÉES et IDENTIFICATION ÉMOTION.	33
3.5	Distribution des lemmes les plus fréquents selon les catégories VÉRIFICATION FAIT, ACTION EXTÉRIEUR, ACTION INTÉRIEUR, TERRITOIRE.	34
3.6	Importance des caractéristiques pour la prédiction des catégories. Le modèle entraîné est la régression logistique. Le tableau est tronqué.	35
3.7	Importance des caractéristiques pour la prédiction des émotions à partir des ressentis physiques. Le modèle entraîné est la régression logistique.	36
3.8	Réseau sémantique extrait de ConceptNet à partir de deux concepts : RESTAURANT et HOTEL. Image provenant de AGARWAL et MITTAL (2015).	37

Liste des tableaux

1.1	Les formes de relations causales.	10
1.2	Classification du stimulus à l'échelle la clause.	13
1.3	Classification du stimulus à l'échelle du token.	13
2.1	Statistiques des documents Word.	16
2.2	Colonnes correspondant aux consignes.	16
2.3	Colonne correspondant aux réponses de l'auteur d'une grille.	16
2.4	Groupement des consignes en une seule colonne.	17
2.5	Tableau normalisé final avec les classes retenues.	17
2.6	Scores obtenus sur la régression logistique, la forêt aléatoire et le séparateur à vaste marge.	21
2.7	Scores obtenus sur DistilCamemBERT et CamemBERT.	24
2.8	Scores obtenus à partir des réponses des différentes catégories sur la régression logis- tique, la forêt aléatoire et le séparateur à vaste marge.	24
3.1	Exemple d'une grille FERA complétée.	31

Acronymes

ACE Analyse Cognitive des Emotions. 4–6, 15, 25, 26

BERT *Bidirectional Encoder Representations from Transformers*. 22, 23

Bi-LSTM *Bi-directional Long Short-Term Memory*. 12

DEFT Défi Fouille de Textes. 14, 26

ECE Extraction de la Cause d'une Émotion. 11, 29

EPEC Extraction des Pairs d'Émotion-Cause. 11–13, 32, 39

FERA Faits, Émotions, Raisons, Actions. 4, 6, 7, 15, 20, 23, 31, 38, 40

INRIA Institut national de recherche en sciences et technologies du numérique. 23

IOB *Inside-Outside-Beginning*. 13

LSTM *Long Short-Term Memory*. 22

MLM *Masked Language Modeling*. 22, 23

NSP *Next Sentence Prediction*. 23

OSCAR *Open Super-large Crawled ALMAnaCH coRpus*. 23

RoBERTa *A Robustly Optimized BERT Pretraining Approach*. 23

TAL Traitement Automatique des Langues. 7, 8, 10, 19, 20, 27, 30

TF-IDF *Term Frequency – Inverse Document Frequency*. 20, 21

Bibliographie

- AGARWAL, B. & MITTAL, N. (2015). Sentiment Analysis Using ConceptNet Ontology and Context Information. *Socio-Affective Computing* (p. 63-75). Springer International Publishing. (Cf. p. 37).
- BACCIANELLA, S., ESULI, A. & SEBASTIANI, F. (2010). SentiWordNet 3.0 : An Enhanced Lexical Resource for Sentiment Analysis and Opinion Mining. *Proceedings of the Seventh International Conference on Language Resources and Evaluation (LREC'10)* (cf. p. 27).
- BOSTAN, L. A. M., KIM, E. & KLINGER, R. (2020). GoodNewsEveryone : A Corpus of News Headlines Annotated with Emotions, Semantic Roles, and Reader Perception. *Proceedings of the 12th Language Resources and Evaluation Conference*, 1554-1566 (cf. p. 14).
- BOSTAN, L.-A.-M. & KLINGER, R. (2018). An Analysis of Annotated Corpora for Emotion Classification in Text. *Proceedings of the 27th International Conference on Computational Linguistics*, 2104-2119 (cf. p. 8).
- CAMBRIA, E. & HUSSAIN, A. (2015). *Sentic Computing : A Common-Sense-Based Framework for Concept-Level Sentiment Analysis* (1st). Springer Publishing Company, Incorporated. (Cf. p. 9).
- CAMPAGNANO, C., CONIA, S. & NAVIGLI, R. (2022). SRL4E – Semantic Role Labeling for Emotions : A Unified Evaluation Framework. *Proceedings of the 60th Annual Meeting of the Association for Computational Linguistics (Volume 1 : Long Papers)*, 4586-4601 (cf. p. 26).
- CHEN, Y., LEE, S. Y. M., LI, S. & HUANG, C.-R. (2010). Emotion Cause Detection with Linguistic Constructions. *Proceedings of the 23rd International Conference on Computational Linguistics (Coling 2010)*, 179-187 (cf. p. 11).
- D. GHAZI, e. S. S., D. Inkpen. (2015). Detecting Emotion Stimuli in Emotion-Bearing Sentences. (Cf. p. 13).
- DANG, B. M. D., OBERLÄNDER, L. & KLINGER, R. (2021). Emotion Stimulus Detection in German News Headlines. (Cf. p. 14).
- DELESTRE, C. & AMAR, A. (2022). DistilCamemBERT : une distillation du modèle français CamemBERT. *CAP (Conférence sur l'Apprentissage automatique)* (cf. p. 23).
- EKMAN, P. & FRIESEN, W. V. (1971). Constants across cultures in the face and emotion. *Journal of personality and social psychology*, 17 2, 124-9 (cf. p. 8).
- ÉTIENNE, A., BATTISTELLI, D. & LECORVÉ, G. (2020). L'expression des émotions dans les textes pour enfants : constitution d'un corpus annoté (Expressing emotions in texts for children : constitution of an annotated corpus). *Actes de la 6e conférence conjointe Journées d'Études sur la Parole (JEP, 33e édition), Traitement Automatique des Langues Naturelles (TALN, 27e édition), Rencontre des Étudiants Chercheurs en Informatique pour le Traitement Automatique des Langues (RÉCITAL, 22e édition). Volume 2 : Traitement Automatique des Langues Naturelles*, 205-212 (cf. p. 14, 26).
- FINKEL, A. (2022). *Manuel d'analyse cognitive des émotions : Théorie et applications*. Dunod. (Cf. p. 4, 6, 7, 19, 25).
- FRISTON, K. (2010). The free-energy principle : a unified brain theory ? *Nature Reviews Neuroscience*, 11 (2), 127-138 (cf. p. 5).
- GUI, L., WU, D., XU, R., LU, Q. & ZHOU, Y. (2016). Event-Driven Emotion Cause Extraction with Corpus Construction. *Proceedings of the 2016 Conference on Empirical Methods in Natural Language Processing*, 1639-1649 (cf. p. 11).
- JIN, X., WANG, X., LUO, X., HUANG, S. & GU, S. (2020). Inter-sentence and Implicit Causality Extraction from Chinese Corpus. In H. W. LAUW, R. C.-W. WONG, A. NTOULAS, E.-P. LIM, S.-K. NG & S. J. PAN (Éd.), *Advances in Knowledge Discovery and Data Mining* (p. 739-751). Springer International Publishing. (Cf. p. 11).
- KIM, E. & KLINGER, R. (2018). Who Feels What and Why ? Annotation of a Literature Corpus with Semantic Roles of Emotions. *Proceedings of COLING 2018, the 27th International Conference on Computational Linguistics* (cf. p. 13).
- LAZARUS, R. S. & FOLKMAN, S. (1984). *Stress, appraisal, and coping / Richard S. Lazarus, Susan Folkman*. Springer Pub. Co New York. (Cf. p. 4).
- LI, J., DING, H., SHANG, J., MCAULEY, J. & FENG, Z. (2021). Weakly Supervised Named Entity Tagging with Learnable Logical Rules. *Proceedings of the 59th Annual Meeting of the Association for Computational Linguistics and the 11th International Joint Conference on Natural Language Processing (Volume 1 : Long Papers)*, 4568-4581 (cf. p. 28).
- LIU, Y., OTT, M., GOYAL, N., DU, J., JOSHI, M., CHEN, D., LEVY, O., LEWIS, M., ZETTMLOYER, L. & STOYANOV, V. (2019). *RoBERTa : A Robustly Optimized BERT Pretraining Approach*. (Cf. p. 23).

- MARCHEGGIANI, D. & TITOV, I. (2017). Encoding Sentences with Graph Convolutional Networks for Semantic Role Labeling. *Proceedings of the 2017 Conference on Empirical Methods in Natural Language Processing*, 1506-1515 (cf. p. 29).
- MARTIN, L., MULLER, B., ORTIZ SUÁREZ, P. J., DUPONT, Y., ROMARY, L., de la CLERGERIE, É., SEDDAH, D. & SAGOT, B. (2020). CamemBERT : a Tasty French Language Model. *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics*, 7203-7219 (cf. p. 23).
- MASLOW, A. H. (1943). A theory of human motivation. *Psychological Review*, 50, 370-396 (cf. p. 5).
- MILLER, G. A. (1994). WordNet : A Lexical Database for English. *Human Language Technology : Proceedings of a Workshop held at Plainsboro, New Jersey, March 8-11, 1994* (cf. p. 27).
- MOHAMMAD, S., KIRITCHENKO, S. & MARTIN, J. (2013). Identifying Purpose Behind Electoral Tweets (cf. p. 13).
- MOHAMMAD, S. M. & TURNEY, P. D. (2013). Crowdsourcing a Word-Emotion Association Lexicon. *Computational Intelligence*, 29(3), 436-465 (cf. p. 27).
- OBERLÄNDER, L. & KLINGER, R. (2020). Token Sequence Labeling vs. Clause Classification for English Emotion Stimulus Detection. (Cf. p. 13).
- PANG, B., LEE, L. & VAITHYANATHAN, S. (2002). Sentiment Classification using Machine Learning Techniques. *Proceedings of the 2002 Conference on Empirical Methods in Natural Language Processing, EMNLP 2002, Philadelphia, PA, USA, July 6-7, 2002*, 79-86 (cf. p. 8).
- PAROUBEK, P., GROUIN, C., BELLOT, P., CLAVEAU, V., ESHKOL-TARAVELLA, I., FRAISSE, A., JACKIEWICZ, A., KAROUI, J., MONCEAUX, L. & TORRES-MORENO, J.-M. (2018). DEFT2018 : recherche d'information et analyse de sentiments dans des tweets concernant les transports en Île de France (DEFT2018 : Information Retrieval and Sentiment Analysis in Tweets about Public Transportation in Île de France Region). *Actes de la Conférence TALN. Volume 2 - Démonstrations, articles des Rencontres Jeunes Chercheurs, ateliers DeFT*, 219-230 (cf. p. 14, 26).
- PLUTCHIK, R. (2001). The Nature of Emotions : Human emotions have deep evolutionary roots, a fact that may explain their complexity and provide tools for clinical practice. *American Scientist*, 89(4), 344-350 (cf. p. 8, 26).
- PORIA, S., HAZARIKA, D., MAJUMDER, N. & MIHALCEA, R. (2020). Beneath the Tip of the Iceberg : Current Challenges and New Directions in Sentiment Analysis Research. *IEEE Transactions on Affective Computing*, 1-1 (cf. p. 8, 9).
- RAY, P. & CHAKRABARTI, A. (2022). A Mixed approach of Deep Learning method and Rule-Based method to improve Aspect Level Sentiment Analysis. *Applied Computing and Informatics*, 18(1), 163-178 (cf. p. 27).
- SAGOT, B. & FIŠER, D. (2008). Building a free French wordnet from multilingual resources. *OntoLex* (cf. p. 27).
- SINGH, A., HINGANE, S., WANI, S. & MODI, A. (2021). An End-to-End Network for Emotion-Cause Pair Extraction. (Cf. p. 12, 32).
- SPEER, R., CHIN, J. & HAVASI, C. (2017). ConceptNet 5.5 : An Open Multilingual Graph of General Knowledge. *Proceedings of the Thirty-First AAAI Conference on Artificial Intelligence*, 4444-4451 (cf. p. 29).
- VARNI, G., HUPONT, I., CLAVEL, C. & CHETOUANI, M. (2020). Computational Study of Primitive Emotional Contagion in Dyadic Interactions. *IEEE Transactions on Affective Computing*, 11(2), 258-271 (cf. p. 8).
- VASWANI, A., SHAZEER, N., PARMAR, N., USZKOREIT, J., JONES, L., GOMEZ, A. N., KAISER, L. & POLOSUKHIN, I. (2017). Attention is All You Need. *Proceedings of the 31st International Conference on Neural Information Processing Systems*, 6000-6010 (cf. p. 22, 27).
- XIA, R. & DING, Z. (2019). Emotion-Cause Pair Extraction : A New Task to Emotion Analysis in Texts. (Cf. p. 11, 12).
- YAN, H., GUI, L., PERGOLA, G. & HE, Y. (2021). Position Bias Mitigation : A Knowledge-Aware Graph Model for Emotion Cause Extraction. *arXiv preprint arXiv : Arxiv-2106.03518* (cf. p. 29).
- YANG, J., HAN, S. C. & POON, J. (2021). A Survey on Extraction of Causal Relations from Natural Language Text. (Cf. p. 10).
- YANG, L., LI, J., CUNNINGHAM, P., ZHANG, Y., SMYTH, B. & DONG, R. (2021). Exploring the Efficacy of Automatically Generated Counterfactuals for Sentiment Analysis. *Proceedings of the 59th Annual Meeting of the Association for Computational Linguistics and the 11th International Joint Conference on Natural Language Processing (Volume 1 : Long Papers)*, 306-316 (cf. p. 10).
- YAO, Y., YE, D., LI, P., HAN, X., LIN, Y., LIU, Z., LIU, Z., HUANG, L., ZHOU, J. & SUN, M. (2019). DocRED : A Large-Scale Document-Level Relation Extraction Dataset. (Cf. p. 11).
- ZENG, S., XU, R., CHANG, B. & LI, L. (2020). Double Graph Based Reasoning for Document-level Relation Extraction. *Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing (EMNLP)*, 1630-1640 (cf. p. 11).