



HAL
open science

Repérage et caractérisation automatique des émotions dans des textes : traiter aussi leurs modes d'expression indirects

Aline Étienne, Delphine Battistelli, Gwéno­lé Lecorvé

► To cite this version:

Aline Étienne, Delphine Battistelli, Gwéno­lé Lecorvé. Repérage et caractérisation automatique des émotions dans des textes : traiter aussi leurs modes d'expression indirects. 35èmes Journées d'Études sur la Parole (JEP 2024) 31ème Conférence sur le Traitement Automatique des Langues Naturelles (TALN 2024) 26ème Rencontre des Étudiants Chercheurs en Informatique pour le Traitement Automatique des Langues (RECITAL 2024), Jul 2024, Toulouse, France. pp.650-676. hal-04623046

HAL Id: hal-04623046

<https://inria.hal.science/hal-04623046>

Submitted on 1 Jul 2024

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution 4.0 International License

Repérage et caractérisation automatique des émotions dans des textes : traiter aussi leurs modes d’expression indirects

Aline Étienne¹

Delphine Battistelli¹

Gwénolé Lecorvé²

(1) Univ. Paris-Nanterre, CNRS, MoDyCo – Nanterre, France (2) Orange – Lannion, France

aem.etienne@gmail.com, del.battistelli@gmail.com, gwenole.lecorve@orange.com

RÉSUMÉ

Cet article présente un modèle capable de prédire (A) si une phrase contient l’expression d’une émotion, (B) selon quel(s) mode(s) cette émotion est exprimée, (C) si elle est basique ou complexe, et (D) quelle est sa catégorie exacte. Notre principale contribution est d’intégrer le fait qu’une émotion puisse s’exprimer selon différents modes : depuis un mode direct, essentiellement lexicalisé, jusqu’à un mode plus indirect, où des émotions vont être seulement suggérées, mode dont les approches en TAL ne tiennent généralement pas compte. Nos expériences sur des textes en français pour les enfants mènent à des résultats tout à fait acceptables en comparaison de ce sur quoi des annotateurs humains experts en psycholinguistique s’accordent et à des résultats meilleurs que ceux produits par GPT-3.5 via du *prompting*. Ceci offre une perspective intéressante de prise en compte des émotions comme facteur d’analyse automatique de la complexité dans les textes, cadre plus général de nos travaux.

ABSTRACT

Spotting and characterization emotions in texts : also addressing indirect modes of expression.

This paper presents a model able to predict whether a sentence expresses an emotion, (B) the mode(s) in which it is expressed, (C) whether it is basic or complex, and (D) its emotional category. One of our major contributions is to integrate the fact that an emotion can be expressed in different modes : from a direct mode, essentially lexicalized, to a more indirect mode, where emotions will only be suggested, a mode that NLP approaches generally don’t take into account. Experiments on French texts intended for children show acceptable results compared to the human annotators’ agreement, and the lower results obtained by a large language model without fine-tuning. They offer an interesting perspective on how to take emotions into account as a factor in the automatic analysis of complexity in texts, which is the more general framework of our work.

MOTS-CLÉS : Émotions, mode d’expression, complexité, étiquetage multi-tâches.

KEYWORDS: Emotions, expression mode, complexity, multi-task labeling.

1 Introduction

En traitement automatique des langues (TAL), la tâche de repérage des émotions est assez souvent abordée à propos de textes produits dans un cadre d’interactions ou conversations (par ex. (Poria *et al.*, 2019)), orales comme écrites (chats, forums, tweets) avec des jeux de données souvent multimodaux (par ex. (Busso *et al.*, 2008; Poria *et al.*, 2018; Chen *et al.*, 2018)). La visée est généralement de pouvoir

identifier les émotions ressenties par les locuteurs en situation dialogique. L'analyse des émotions dans des textes autres que de type conversationnel, comme par exemple des textes journalistiques et encyclopédiques ou encore des romans est un domaine moins développé en TAL. Elle relève de fait d'un autre type de visée qui n'est plus celui de caractériser l'état émotionnel du locuteur mais plutôt des personnages qui composent ces textes. Comme le soulignent certains travaux en psycholinguistique, les émotions sont dédiées dans ces types de textes - de manière plus ou moins contrôlée par le scripteur - à capter l'attention du lecteur. Elles permettent en outre de créer du lien entre les situations décrites et donc sont un facteur clé dans la compréhension (par ex. pour les enfants dans Davidson *et al.*, 2001). Encore faut-il que ces émotions elles-mêmes soient identifiées et comprises. Là se joue une question qui devient alors celle d'envisager les émotions comme un facteur de complexité, au moins relative pour reprendre la terminologie de Ehret *et al.* (2023), c'est-à-dire qui tient compte de la difficulté perçue par les locuteurs en termes d'apprentissage ou de compréhension de la langue. Un texte sera ainsi d'autant plus complexe qu'il contient d'émotions considérées comme complexes par un type de locuteur donné. Dans le cas d'enfants par exemple, on sait que certaines catégories émotionnelles ne sont pas accessibles avant un certain âge et que leur mode d'expression (direct vs. indirect ou implicite) joue également un rôle dans l'accessibilité à leur signification.

De ces réflexions sur la question des émotions comme facteur de complexité, nous avons choisi d'orienter nos travaux vers une meilleure prise en compte de la diversité des modes d'expression des émotions. Nous présentons un modèle qui introduit ainsi la notion de mode d'expression en plus des informations habituelles sur les catégories émotionnelles (par ex., joie, peur, etc.). En pratique, le modèle classe les émotions dans les textes à travers quatre tâches : (A) prédire si une phrase contient ou non une émotion ; (B) si oui, comment elle est exprimée (le *mode*) ; (C) s'il s'agit d'une catégorie d'émotion basique ou complexe ; et (D) dans quelle catégorie émotionnelle elle se situe. Le modèle est construit à partir du modèle CamemBERT (Martin *et al.*, 2020) et de données dérivées d'un schéma d'annotation psycho-linguistiquement motivé comprenant différents types de sources. L'évaluation montre que le modèle proposé surpasse des approches fondées sur des ressources expertes, des architectures non neuronales (SVM et XGBoost), et de l'inférence en-contexte (*in-context learning* d'un grand modèle de langue (GPT-3.5)). De plus, l'évaluation humaine menée en complément montre que les erreurs de prédiction faites par le modèle proposé se situent généralement dans les mêmes proportions que celles faites par les humains.

Dans la suite, la section 2 dresse un panorama de la littérature dédiée à l'identification des émotions dans les textes, notamment en TAL. Les sections 3, 4 et 5 détaillent respectivement les tâches traitées, les données associées, puis le modèle proposé. La section 6 rapporte enfin les expériences et résultats.

2 Cadre d'analyse des émotions et travaux connexes

Cette section fournit un bref aperçu du cadre d'analyse des émotions dans lequel nous nous situons et qui justifie le choix du schéma et des données (annotées avec ce schéma). Elle positionne également notre travail parmi les études en TAL sur l'identification automatique des émotions.

2.1 L'analyse des émotions en tant que facteur de complexité d'un texte

En psycho-linguistique, le rôle-clé des émotions des personnages sur la compréhension des textes est un sujet bien documenté (par ex. Dijkstra *et al.*, 1995; Dyer, 1983). Parmi les travaux récents,

deux facteurs d'influence ont été mis en évidence dans la compréhension des émotions par les enfants, et donc des textes eux-mêmes : le *type d'émotion* exprimé, basique ou complexe – les émotions complexes (par ex. la fierté, la honte) étant plus difficiles à saisir car elles nécessitent une connaissance des normes sociales – (Davidson, 2006; Blanc & Quenette, 2017); ainsi que *la manière dont les émotions sont exprimées* (Creissen & Blanc, 2017)), directement *via* une étiquette émotionnelle, indirectement à travers la mention d'un comportement émotionnel, ou à travers la description d'une situation émotionnelle, cette dernière étant la plus difficile à comprendre. Bien sûr, la notion de catégorie émotionnelle est également abordée en psycho-linguistique, et il a été montré que certaines catégories prennent plus de temps à être maîtrisées par les enfants (par ex. Baron-Cohen *et al.*, 2010).

Du côté du TAL, plusieurs travaux (voir par ex. Bostan & Klinger, 2018; Acheampong *et al.*, 2020; Öhman, 2020) soulignent la grande hétérogénéité des schémas d'annotation des émotions – et des corpus annotés –, mettant ainsi clairement en évidence la difficulté de modéliser les émotions et *in fine* de les analyser. Cette hétérogénéité concerne tous les aspects de ces travaux : depuis les notions (par ex. le nombre et les types de catégories émotionnelles) et le type de données étudiées (journaux, tweets. . .) jusqu'aux procédures d'annotation (*crowdsourcing*, annotation par des experts) et méthodes d'évaluation mises en œuvre (par ex. avec ou sans accord entre annotateurs). Bien que certains travaux s'efforcent de prendre en compte des ensembles plus larges de notions et d'indices linguistiques pour analyser les émotions (par ex. Casel *et al.*, 2021; Kim & Klinger, 2019), le concept le plus couramment utilisé reste la notion de *catégorie émotionnelle*, souvent abordée à travers une liste d'émotions de base introduite soit par Ekman (1992) (colère, dégoût, peur, joie, tristesse et surprise) ou Plutchik (1980) (catégories d'Ekman, anticipation et confiance), avec un accent sur une manière d'exprimer les émotions : le lexique émotionnel. Comme souligné dans (Klinger, 2023) et dans (Troiano *et al.*, 2023), quelques approches très récentes en TAL visent cependant à acquérir une compréhension plus profonde des unités textuelles qui soutiennent l'évocation d'émotions en dehors des termes lexicaux directement émotionnels (par ex. «heureux», «colère»). Ces approches s'inspirent alors de modèles psychologiques et/ou linguistiques des émotions. Nous adoptons ici la même approche car nous visons à la fois les modes d'expression directs et indirects des émotions dans les textes. Comme Troiano *et al.* (2023), nous cherchons à évaluer à quel point les modèles computationnels peuvent capter des émotions exprimées indirectement (par ex., *via* la description de situations qui sont associées à des émotions eu égard à des normes sociales et à des conventions). Plus précisément, notre travail fait le choix du cadre proposé par Etienne *et al.* (2022). Ce cadre repose sur un schéma d'annotation détaillé des émotions et propose un corpus annoté manuellement dont la taille est compatible avec les expériences d'apprentissage automatique. À notre connaissance, c'est le seul travail qui correspond à l'objectif explicite d'analyser les émotions dans les textes en traitant à la fois les modes d'expression directs et indirects.

2.2 Identification automatique des émotions

En TAL, l'analyse des émotions dans les textes est généralement traitée comme une tâche de classification. L'hétérogénéité précédemment évoquée des schémas d'annotation et des corpus annotés se reflète alors dans la diversité des classes prédites, de la granularité des éléments à classifier et des méthodes pour développer et évaluer les classificateurs. La manière dont les résultats sont présentés varie donc également d'un article à l'autre, ce qui rend la comparaison des performances plus difficile.

L'accent est souvent mis sur la classification en émotions de base (Strapparava & Mihalcea, 2007; Mohammad, 2012; Abdaoui *et al.*, 2017; Demszky *et al.*, 2020; Öhman *et al.*, 2020; Bianchi *et al.*,

2021), bien que certains travaux utilisent un mélange d'émotions de base et complexes (Balahur *et al.*, 2012; Fraïsse & Paroubek, 2015; Abdaoui *et al.*, 2017; Mohammad *et al.*, 2018; Liu *et al.*, 2019; Demszky *et al.*, 2020). De plus, il existe une longue histoire de construction et d'utilisation de lexiques émotionnels et la diversité des marqueurs linguistiques des émotions n'est pas systématiquement prise en compte, bien qu'elle soit mentionnée dans plusieurs travaux (Alm *et al.*, 2005; Mohammad, 2012; Kim & Klinger, 2018; Demszky *et al.*, 2020)). Certains travaux étudient malgré tout d'autres moyens d'expression. par ex., Kim & Klinger (2019) analysent les expressions non verbales des émotions par les personnages dans un corpus de fanfictions (par ex. les regards, les gestes). Balahur *et al.* (2012) visent à détecter les émotions indirectes. Ces travaux ont pour limite de ne se focaliser à chaque fois que sur un seul mode d'expression, laissant ainsi de côté les complémentarités entre modes. Pour leur part, sur la base du modèle de processus de composants émotionnels de Scherer (2005), Casel *et al.* (2021) ont annoté puis prédit plusieurs composantes des émotions, tels que les symptômes physiologiques et les expressions motrices des émotions, ou l'évaluation cognitive des événements. Bien que (Casel *et al.*, 2021) traite d'un ensemble plus large d'indices, ceux-ci ne sont pas rigoureusement motivés linguistiquement. Par conséquent, en s'appuyant sur (Etienne *et al.*, 2022), la véritable originalité de notre travail réside dans la prise en compte de différents modes d'expression des émotions.

Historiquement, les modèles *Support Vector Machine* (SVM) ont été largement utilisés pour classer des phrases (Aman & Szpakowicz, 2007; Mohammad, 2012)) ou des textes (Abdaoui *et al.*, 2017; Balahur *et al.*, 2012; Fraïsse & Paroubek, 2015; Mohammad, 2012) selon la catégorie émotionnelle qu'ils expriment. Jusqu'à l'avènement des plongements, les entrées étaient principalement symboliques : sacs de mots ou n -grammes, caractéristiques basées sur des ressources émotionnelles telles que WordNetAffect (Aman & Szpakowicz, 2007; Balahur *et al.*, 2012; Strapparava & Mihalcea, 2007) ou lexiques émotionnels (Strapparava & Mihalcea, 2007; Abdaoui *et al.*, 2017; Kim & Klinger, 2018). Aujourd'hui, les réseaux neuronaux (Kim & Klinger, 2018) et les architectures Transformer (Liu *et al.*, 2019; Demszky *et al.*, 2020; Öhman *et al.*, 2020; Bianchi *et al.*, 2021) dominent évidemment l'état de l'art. En ce qui concerne le français, aucun modèle basé sur l'architecture Transformer n'a encore été proposé à notre connaissance.

3 Tâches

Construit dans la perspective globale de permettre l'analyse des émotions en tant que facteur de complexité, notre travail tient ainsi compte de deux éléments clé pour aborder la complexité d'une émotion : sa catégorie et son mode d'expression. L'objectif est de proposer un modèle Transformer pour 4 tâches de classification (notés A, B, C et D) au niveau de la *phrase*, par opposition au niveau du texte (cela peut, par exemple, permettre d'étudier comment la présence d'émotions évolue le long d'un texte). Les phrases peuvent contenir plusieurs émotions et les classifications sont donc multi-étiquettes. Toutes les tâches sont apprises ensemble, menant à un modèle unique.

Tâche A : Présence d'émotion La première tâche vise à prédire la présence d'informations émotionnelles dans une phrase donnée (prédiction binaire).

Tâche B : Mode d'expression Le mode d'expression se concentre sur les moyens linguistiques utilisés pour transmettre la présence d'une émotion dans un texte. Suivant Etienne *et al.* (2022), 4 modes sont considérés : les **émotions désignées** directement indiquées par un terme du lexique émotionnel (par ex. *heureux*, *effrayé*); les **émotions comportementales** qui s'appuient sur la

| | Phrase à classer (+ phrases voisines) | | | | | | | | | | | | | | | | | | |
|-----|---------------------------------------|----------|----------|---------|----------|---------|----------|------------|-------|--------|-------------|--------|----------|---------|----------|------|------|----------|-----------|
| | émot. | comport. | désignée | montrée | suggérée | basique | complexe | admiration | autre | colère | culpabilité | dégoût | embarras | fiereté | jalousie | joie | peur | surprise | tristesse |
| (1) | | | | | | | | | | | | | | | | | | | |
| (2) | ✓ | | ✓ | | | ✓ | | | | | | | | | | | ✓ | | |
| (3) | ✓ | ✓ | | | ✓ | ✓ | | | | ✓ | | | | | | | | | |
| (4) | ✓ | | ✓ | | | ✓ | ✓ | | | | ✓ | ✓ | | | | | | | |
| (5) | ✓ | | | ✓ | ✓ | ✓ | ✓ | | | | | | ✓ | | | ✓ | | ✓ | |

TABLE 1 – Exemples de phrases en contexte et étiquettes de référence pour les tâches A, B, C et D.

description d'un comportement émotionnel, telles que des manifestations physiologiques (par ex. *pleurer, sourire*) ou d'autres comportements (par ex. *gifter quelqu'un*); les **émotions montrées** qui sont exprimées par des caractéristiques linguistiques de surface très hétérogènes des énoncés qui reflètent principalement l'état émotionnel du scripteur (par ex. interjections, phrases courtes); les **émotions suggérées** qui émanent de la description d'une situation généralement associée à un sentiment émotionnel d'après les normes sociales et conventions (par ex. *voir un bon ami après une longue période* suggère la joie).

Tâche C : Type d'émotion La tâche C vise à prédire la présence de types d'émotions *basique* et *complexe* (2 prédictions binaires simultanées). À notre connaissance, cette notion n'a encore jamais été étudiée en tant que telle dans l'analyse automatique des émotions (bien que les catégories émotionnelles *basiques* et *complexes* aient été utilisées en TAL (cf. section 2.2)). Cela est probablement dû au fait que le type d'une émotion exprimée est directement lié à sa catégorie émotionnelle. Cependant, le type d'émotion est en lui-même un marqueur de complexité, comme nous l'avons vu.

Tâche D : Catégorie émotionnelle Conformément à [Etienne et al. \(2022\)](#), la tâche D est conçue pour étiqueter 11+1 catégories émotionnelles, à savoir les 6 émotions de base d'Ekman (*colère, dégoût, peur, joie, tristesse* et *surprise*) et 5 émotions complexes (*admiration, embarras, culpabilité, jalousie* et *fiereté*). Une dernière catégorie, nommée *autre*, est utilisée pour capturer les marqueurs qui expriment toute autre émotion (par ex. haine, mépris, amour, etc.).

4 Données

Notre corpus est dérivé de celui fourni dans [\(Etienne et al., 2022\)](#). Il se compose de 1 594 textes en français (28K phrases, 515K mots) destinés aux enfants de 6 à 14 ans, répartis en 3 types : très principalement des textes journalistiques (91 % des phrases), des articles encyclopédiques (9 %) et des romans (1 %). Originellement, le corpus est accompagné d'annotations d'experts délimitant des unités émotionnelles (segments) dans les textes. Chaque unité est décrite par son mode d'expression et sa catégorie émotionnelle. Nous avons fusionné ces annotations depuis le niveau des segments vers le niveau de la phrase. Ainsi, une phrase donnée peut couvrir plusieurs unités émotionnelles. La présence d'émotions et les types d'émotions ont été dérivés des étiquettes de mode d'expression et

| Tâche | Étiquettes | Prop. phr. (%) | | |
|----------------------------|-------------|----------------|------|------|
| | | entr. | dév. | test |
| (A) Prés. émotion | émotion | 20,2 | 15,8 | 17,6 |
| | comport. | 4,6 | 3,6 | 4,3 |
| (B) Mode d'expression | désigné | 5,3 | 5,2 | 5,7 |
| | montré | 3,6 | 2,3 | 3,5 |
| | suggéré | 7,1 | 5,8 | 6,3 |
| (C) Type d'émotion | basique | 15,4 | 12,6 | 13,9 |
| | complexe | 2,0 | 2,1 | 2,3 |
| | admiration | 0,6 | 1,1 | 1,0 |
| (D) Catégorie émotionnelle | autre | 5,0 | 3,2 | 3,7 |
| | colère | 4,6 | 3,2 | 3,4 |
| | culpabilité | 0,1 | 0,0 | 0,1 |
| | dégoût | 0,2 | 0,3 | 0,2 |
| | embarras | 0,6 | 0,6 | 0,6 |
| | fierté | 0,7 | 0,4 | 0,9 |
| | jalousie | 0,0 | 0,0 | 0,0 |
| | joie | 3,2 | 2,3 | 3,6 |
| peur | 3,8 | 3,3 | 3,8 | |
| surprise | 3,0 | 3,1 | 2,5 | |
| tristesse | 2,5 | 2,0 | 2,5 | |

TABLE 2 – Répartition des étiquettes

de catégorie émotionnelle. Au final, chaque phrase est associée à un vecteur de 19 booléens. Des exemples de phrases en contexte sont fournis dans le tableau 1.

Les données sont divisées en ensembles d'entraînement, de développement et de test (70/10/20% des phrases, respectivement), tel que toutes les phrases d'un texte se trouvent dans le même sous-ensemble, ceci afin d'éviter un biais d'entraînement sur les particularités des textes (par ex., le nom d'un personnage). Le tableau 2 présente la proportion des étiquettes au sein du corpus. Globalement, les proportions sont comparable d'un sous-ensemble à un autre. Plusieurs déséquilibres apparaissent au sein des tâches. **(A)** Seulement 15-20% des phrases sont émotionnelles. **(B)** Les modes d'expression sont assez uniformément répartis, 'désigné' étant le moins fréquent (3% des phrases) et *suggéré* le plus courant (6%). Les sommes des pourcentages de chaque mode sont supérieures aux pourcentages de l'étiquette *émotionnelle* car une phrase peut certaines émotions sont véhiculés par plusieurs modes et une phrase peut aussi contenir plusieurs unités émotionnelles dont les modes respectifs diffèrent. **(C)** Les étiquettes des types d'émotions sont très déséquilibrées, avec une nette dominance des émotions basiques. La catégorie émotionnelle 'autre' (tâche D) n'est associée à aucun type d'émotion, d'où le fait que la somme des pourcentages *basique* et *complexe* est inférieure à celle des phrases émotionnelles. **(D)** Les étiquettes des catégories émotionnelles sont déséquilibrées, avec des pourcentages toujours en dessous de 5% des phrases. Les catégories *colère*, *peur*, *joie*, *tristesse*, *surprise* et *autre* sont dominantes, alors que d'autres sont très rares (*dégoût*, *culpabilité* et *jalousie*).

5 Modèle proposé

Le modèle proposé résulte d'un affinage (*fine-tuning*) de la version de base du modèle pré-entraîné CamemBERT (Martin *et al.*, 2020). Il s'agit d'un modèle de type encodeur (BERT) de 110 millions de paramètres et 12 couches BERT. Il a été pré-entraîné sur 138 Go de textes français (Suárez *et al.*, 2019). Bien que des modèles de langue plus récents et plus grands (génératifs) comme Llama2 ou Mistral conduiraient probablement à de meilleurs résultats, le choix d'un modèle de taille raisonnable

| Tâche | Modèle | Macro R | Macro P | Macro F1 |
|----------------------------|--------------|--------------|--------------|--------------|
| (A) Présence d'une émotion | SVM | 0,481 | 0,659 | 0,556 |
| | XGBoost | 0,223 | 0,700 | 0,338 |
| | GPT-3.5 | 0,622 | 0,443 | 0,518 |
| (B) Mode d'expression | notre modèle | 0,764 | 0,741 | 0,752 |
| | SVM | 0,267 | 0,721 | 0,368 |
| | XGBoost | 0,218 | 0,730 | 0,314 |
| | GPT-3.5 | 0,513 | 0,101 | 0,152 |
| (C) Type d'émotion | notre modèle | 0,626 | 0,665 | 0,645 |
| | SVM | 0,211 | 0,343 | 0,261 |
| | XGBoost | 0,120 | 0,659 | 0,200 |
| | GPT-3.5 | 0,756 | 0,123 | 0,199 |
| | notre modèle | 0,557 | 0,662 | 0,601 |
| (D) Catégorie émotionnelle | SVM | 0,125 | 0,487 | 0,186 |
| | XGBoost | 0,192 | 0,565 | 0,272 |
| | GPT-3.5 | 0,697 | 0,109 | 0,174 |
| | notre modèle | 0,397 | 0,463 | 0,420 |

TABLE 3 – Performances des modèles (moyennes sur 3 exécutions, tous les écarts-types sont inférieurs à 0,02).

est motivé par deux raisons. Premièrement, notre objectif est de montrer que, contrairement à plusieurs autres tâches en TAL, la caractérisation fine des émotions dans les textes ne peut pas être réalisée en sollicitant des modèles de langage génériques (c.-à-d. non spécialisés) de grande taille *via* de l'apprentissage en-contexte (c.-à-d. sans affinage). Ensuite, notre travail vise une solution légère, de sorte que la caractérisation des émotions puisse être intégrée comme un processeur pour l'analyse de la complexité des textes dans une collection massive de textes d'un moteur de recherche public. Ainsi, même si l'affinage de modèles plus grands fait partie de nos perspectives, l'article ne l'aborde pas.

Nous affinons le modèle CamemBERT en remplaçant sa dernière couche de prédiction de jetons par une couche de classification binaire de la taille du nombre d'étiquettes, avec l'entropie croisée binaire comme fonction de perte. L'affinage porte sur tous les poids du modèle, c.-à-d. qu'aucune couche n'est gelée. Suite à des travaux de prototypage sur l'ensemble de développement, le modèle final n'est pas directement appris à partir de CamemBERT. Un premier affinage est conduit sur la seule tâche A pendant 3 époques (couche de classification de taille 1), puis le modèle final est affiné sur toutes les tâches en partant de ce modèle intermédiaire pendant 6 époques supplémentaires (la couche de classification finale est remplacée par une couche vierge de taille 19). L'optimiseur est Adam avec un taux d'apprentissage de 10^{-5} (sans décroissance) et des lots de 8 exemples.

D'autres expérimentations (non rapportés dans cet article) ont été conduites sur l'ensemble de développement, par exemple sur le choix d'une fenêtre ou non autour des phrases, la pondération des classes ou non, ou encore le choix d'un premier affinage uniquement sur la tâche A ou non. Au final, les résultats présentés sont ceux de la meilleure stratégie obtenues sur l'ensemble de développement en moyenne les résultats sur 3 exécutions de l'apprentissage avec des initialisation aléatoire différentes. Notamment, une pondération entre classes est adoptée afin que ne favorise pas trop les classes majoritaires. Le facteur de pondération maximal est borné à 50 pour, à l'inverse, ne pas donner non plus trop d'importance aux classes très rares. Enfin, le modèle prend en entrée un triplet de phrases où la phrase cible à étiqueter est entourée de sa phrase précédente et suivante sous la forme avant : {précédente}</s>actuelle : {cible}</s>après : {suivante}</s>. Pour plus de détails, le lecteur est invité à consulter (Etienne, 2023).

6 Évaluations automatique et humaine

6.1 Comparaison avec d'autres modèles

Le modèle proposé est comparé à trois autres types de modèles. Des modèles **SVM** ont été entraînés car c'est une approche historique dans le domaine. Deux types de descripteurs d'entrée ont été utilisés : (i) des sac-de-jetons où les jetons proviennent du tokeniseur de CamemBERT, restreints à ceux de l'ensemble d'entraînement, résultant en des vecteurs d'entrée de dimension 18 437 ; (ii) des plongements de phrases de taille 768 obtenus avec SentenceTransformer (Reimers & Gurevych, 2019) et CamemBERT¹. Des modèles **XGBoost** ont été entraînés car c'est une technique plus récente, légère et compétitive pour de nombreuses tâches de classification, en particulier avec des données déséquilibrées (Chen & Guestrin, 2016). Les descripteurs d'entrée sont les mêmes que pour les SVM. Notre approche est comparée à **GPT-3.5** (Ouyang *et al.*, 2022). Pour un échantillon d'entrée donné, GPT-3.5 est sollicité de manière incrémentale pour l'annoter avec des étiquettes binaires (oui/non). Consécutivement pour chaque tâche et étiquette, une description en langage naturel de ce qui est

1. <https://huggingface.co/dangvantuan/sentence-camembert-base>

| Réf. | Lg. | Étiquettes | Modèle | Lexique | Granularité | Macro-F1 du meill. modèle |
|---------------------------|-----|--|-------------|---------|------------------|---------------------------|
| notre modèle | Fra | colère, dég., joie, peur, surpr., trist. | Transformer | aucun | triplets de phr. | 0,52 |
| (Öhman et al., 2020) | Ang | <i>idem</i> anticipation | Transformer | aucun | phr. | 0,54 |
| (Kim and Klinger, 2018) | Fra | <i>idem</i> + conf., anticipation | symb. | NRC | triplets de phr. | 0,31 |
| | | | MLP | aucun | triplets de phr. | 0,31 |
| (Fraisse, Paroubek, 2015) | Fra | colère, peur, tristesse | SVM | perso. | paragr. | 0,31 |

TABLE 4 – Éléments de comparaison avec des travaux proches.

| Tâche | Étiquette | Approche | Macro-F1 |
|--|--------------|--------------|--------------|
| (A) Présence d’une émotion | émotionnelle | notre modèle | 0.752 |
| | | TextBlob | 0.299 |
| | | Emotaix | 0.445 |
| (B) Mode d’expression | comport. | notre modèle | 0.626 |
| | | Emotaix | 0.041 |
| | | notre modèle | 0.807 |
| (D) Catégories émot. (mode <i>désigné</i> seul.) | désignée | notre modèle | 0.466 |
| | | Emotaix | 0.425 |
| | | notre modèle | 0.575 |
| Polarité émotionnelle | positive | TextBlob | 0.163 |
| | | notre modèle | 0.678 |
| | | TextBlob | 0.168 |

TABLE 5 – Éléments de comparaison avec des outils disponibles pour le français.

attendu est fournie au modèle avant de demander de répondre, accompagnée d’exemples issus de l’ensemble d’entraînement pour chaque étiquette. Différentes amorces ont été testées (*cf.* détails dans l’annexe A). Celle retenue reporte de 2 à 4 exemples positifs par étiquette. Contrairement à SVM, XGBoost et notre modèle, cette approche n’est pas économe mais elle se passe d’entraînement.

Le tableau 3 résume les performances sur l’ensemble de test des meilleurs modèles pour chaque tâche – pour SVM et XGBoost, les descripteurs sac-de-jetons ; pour GPT-3.5, les amorces sans exemples négatifs – et les compare à notre modèle. Les modèles sont évalués à travers les scores de rappel (R), de précision (P) et de score F1. Dans l’ensemble, il apparaît que notre modèle proposé surpasse significativement les SVM, XGBoost et GPT-3.5 en termes de scores F1 pour toutes les tâches, avec des valeurs qui sont presque le double du modèle le mieux classé pour les tâches B, C et D. Il semble surtout que tous les autres modèles tendent à favoriser soit le rappel (GPT-3.5) soit la précision (SVM, XGBoost), tandis que notre modèle est équilibré. Enfin, les faibles résultats de GPT-3.5 montrent que la tâche est difficile et nécessite un affinage.

6.2 Comparaison avec les travaux connexes

À défaut de travaux véritablement similaires aux nôtres, cette section rapporte des résultats complémentaires pour donner une meilleure intuition de la performance de notre modèle.

Travaux comparables les plus proches Le tableau 4 résume les performances des trois travaux les plus proches que nous avons pu trouver dans la littérature. Ils ont été choisis parce qu’ils prédisent tous des étiquettes sur une granularité proche de celui de la phrase. (Öhman *et al.*, 2020) permet une comparaison avec un autre modèle Transformer ; (Fraisse & Paroubek, 2015) avec un autre travail en français ; et (Kim & Klinger, 2018) avec un méthode qui travaille au niveau des marqueurs linguistiques (par opposition au niveau phrastique ou textuel). Tous se concentrent uniquement sur les catégories émotionnelles. Les résultats montrent que notre modèle est compétitif.

Implémentations fondées sur des ressources existantes Deux ressources disponibles en français sont intéressantes pour l’identification des émotions : TextBlob (<https://textblob.readthedocs.io/>), une bibliothèque d’analyse des sentiments qui intègre un lexique français où

les termes sont associés à un poids négatif et positif reflétant leur polarité ; Emotaix (Piolat & Bannour, 2009), un autre lexique comprenant des associations (i) de termes avec des catégories émotionnelles pour le seul mode désigné, et (ii) d'autres termes avec le mode comportemental (mais cette fois sans information sur la catégorie émotionnelle). Plusieurs tâches gérées par notre modèle ont été répliquées *via* TextBlob et Emotaix. Pour tenir compte des différences entre ces ressources et notre modèle proposé, la tâche B a été limitée aux seuls modes comportemental et désigné et la tâche D au mode désigné. De plus, notre modèle a été testé sur une tâche de prédiction de la polarité émotionnelle sur notre ensemble de test puisque TextBlob est conçu pour cet usage. Pour prédire la polarité *via* notre modèle, les catégories ont été prédites et empiriquement projetés vers la polarité positive ou négative (par ex., *colère est négative, joie est positive*). Comme le montre le tableau 5, notre modèle se comporte nettement mieux que TextBlob et Emotaix, y compris dans la tâche de polarité émotionnelle pour laquelle il n'a pas été spécifiquement conçu. La seule tâche pour laquelle la concurrence demeure est la prédiction des catégories lorsque le mode est désigné, ce qui est la situation la plus facile par rapport à la prise en compte de tous les modes.

6.3 Résultats par étiquette

Le tableau 6 présente les résultats de notre classifieur sur toutes les étiquettes de toutes les tâches. Des observations supplémentaires peuvent être faites comme suit. En ce qui concerne les modes d'expression (B), les émotions désignées sont très bien reconnues ($F1 > 0,8$), contrairement aux émotions suggérées ($F1 < 0,5$). Cela n'est pas surprenant car les émotions désignées sont les plus faciles à identifier pour un annotateur humains, tandis que les émotions suggérées ont la part d'interprétation la plus grande. La performance pour les types d'émotions (C) semble pour sa part liée aux résultats sur les catégories émotionnelles, puisque l'étiquette *basique* est, conformément à l'intuition, mieux reconnue que l'étiquette *complexe*. Enfin, concernant les catégories émotionnelles (D), trois d'entre elles ne sont jamais prédites (*culpabilité, dégoût et jalousie*). Ce sont les étiquettes les plus rares de l'ensemble d'entraînement, probablement trop pour que le modèle apprenne à les prédire. De fait, les catégories émotionnelles les mieux prédites sont les émotions de base, plus fréquentes, à savoir les étiquettes *surprise, peur et colère* (cf. tableau 2). Cependant, alors que *surprise* est l'étiquette la mieux prédite de la tâche D, ce n'est pas la plus représentée dans l'ensemble d'entraînement. Au contraire, *tristesse* n'est pas bien reconnue, même si c'est l'une des catégories émotionnelles les plus fréquentes. De nos analyses complémentaires, cela semblerait s'expliquer par des interactions parfois fortes entre les notions de mode d'expression et de catégorie émotionnelle. Par exemple, la catégorie *surprise*, qui est principalement *montrée* dans le corpus, est en moyenne 14 fois mieux reconnue lorsqu'elle est exprimée par ce mode par rapport aux autres modes. De même, la *colère*, principalement *comportementale* dans le corpus, est 4 fois mieux prédite dans ce mode.

6.4 Évaluation humaine

Étant donné la difficulté des tâches considérées, il est opportun de recouper l'évaluation automatique avec une analyse humaine, notamment pour donner une intuition de ce que représentent les erreurs de prédiction observées. Une expérience de validation perceptive a ainsi été menée avec trois experts en complexité textuelle et en émotions. Chacun d'eux a été informé des tâches et des définitions des étiquettes en psycho-linguistique et en linguistique. Ils ont ensuite été chacun confrontés à 150 phrases de l'ensemble de test et à leurs étiquettes de catégorie émotionnelle et de mode d'expression. Ces

| Tâche | MacroR | MacroP | MacroF1 | Étiquettes | R | P | F1 |
|----------------------------|--------|--------|---------|-------------|-------|-------|-------|
| (A) Présence d'une émotion | 0,764 | 0,741 | 0,752 | émotion | 0,764 | 0,741 | 0,752 |
| | | | | comportem. | 0,601 | 0,653 | 0,626 |
| (B) Mode d'expression | 0,626 | 0,665 | 0,645 | désigné | 0,811 | 0,803 | 0,807 |
| | | | | montré | 0,667 | 0,726 | 0,695 |
| | | | | suggéré | 0,426 | 0,479 | 0,451 |
| (C) Type d'émotion | 0,557 | 0,662 | 0,601 | basique | 0,705 | 0,733 | 0,719 |
| | | | | complexe | 0,409 | 0,591 | 0,484 |
| | | | | admiration | 0,281 | 0,457 | 0,348 |
| | | | | autre | 0,745 | 0,592 | 0,660 |
| | | | | colère | 0,670 | 0,685 | 0,677 |
| | | | | culpabilité | 0,000 | 0,000 | 0,000 |
| | | | | dégoût | 0,000 | 0,000 | 0,000 |
| | | | | embarras | 0,364 | 0,600 | 0,453 |
| (D) Catégorie émotionnelle | 0,397 | 0,463 | 0,420 | fiereté | 0,333 | 0,615 | 0,432 |
| | | | | jalousie | 0,000 | 0,000 | 0,000 |
| | | | | joie | 0,530 | 0,709 | 0,606 |
| | | | | peur | 0,717 | 0,661 | 0,688 |
| | | | | surprise | 0,697 | 0,739 | 0,717 |
| | | | | tristesse | 0,428 | 0,504 | 0,463 |

TABLE 6 – Performances détaillées de notre modèle

étiquettes provenaient soit des annotations humaines de référence, soit des prédictions de notre modèle. Pour chaque étiquette, les experts devaient dire s'ils étaient d'accord ou non avec l'annotation proposée. Bien sûr, ils n'avaient pas connaissance de l'origine des étiquettes.

Le tableau 7 rapporte les taux d'accord des experts avec les étiquettes proposées, en fonction de la source de l'étiquette. Bien que l'accord le plus fort soit lorsque les étiquettes humaines et celles du modèle concordent (*humain & modèle*), les scores d'accord sont globalement très élevés, en particulier pour le mode d'expression. Ces résultats tendent donc à montrer que, même lorsque le modèle prédit différemment de la référence, la prédiction est généralement considérée comme pertinente par les experts humains. Cela démontre que notre modèle est capable de généraliser correctement et que les scores F1 des expériences précédentes sous-estiment la qualité perçue des prédictions du modèle.

7 Conclusion et perspectives

Nous avons proposé un modèle d'analyse des émotions dans des textes qui est original en TAL car il prend en compte leurs modes d'expression directs mais aussi indirects. De plus, les expériences montrent que ce modèle a de bonnes performances par rapport à d'autres approches, des travaux comparables et des solutions à partir de ressources sur étagère. L'évaluation humaine a montré que ce niveau est presque équivalent à ce que les humains peuvent faire.

À l'avenir, des prédictions intra-phrastiques, délimitant des unités, permettant l'inclusion d'autres notions, comme celle d'expérimenteur, sont des pistes pour sophistiquer encore l'analyse. L'affinage de modèles génératifs semble alors nécessaire. Sur un autre plan, une application directe de notre modèle est l'analyse de la complexité – contexte plus large de notre travail –, puisque les étiquettes prédites reflètent des marqueurs de complexité. Plus largement, notre travail pourrait contribuer à la recherche en psychologie pour étudier le lien entre le langage émotionnel et l'état psychologique du scripteur/locuteur, dans la lignée des études rapportées dans (Tausczik & Pennebaker, 2010).

| Source de l'étiq. | Opinion de l'évaluateur | Proportion (nb. d'étiquettes) | |
|-------------------|-------------------------|-------------------------------|--------------|
| | | catég. émot. | mode d'expr. |
| Humain & Modèle | Accord | 95,5 % (105) | 97,7 % (129) |
| | Désaccord | 4,5 % (5) | 2,3 % (3) |
| Modèle | Accord | 92,1 % (58) | 91,1 % (41) |
| | Désaccord | 7,9 % (5) | 8,9 % (4) |
| Modèle | Accord | 76,5 % (39) | 90,2 % (37) |
| | Désaccord | 23,5 % (12) | 9,8 % (4) |

TABLE 7 – Accord des experts vis-à-vis des prédictions communes à l'annotateur humain et à notre modèle, ou spécifiques à chacun.

Remerciements

Ce travail a en partie été financé par l'Agence Nationale de la Recherche (ANR) dans le cadre du projet TextToKids (ANR-19-CE38-0014).

A Détails de l'inférence avec GPT-3.5

GPT-3.5 a été utilisé en mode conversationnel. Les amorces sont donc une alternance de messages entre l'*utilisateur* et l'*assistant*, précédés d'un message global dit du *système*. Les messages de l'utilisateur parcourent toutes les étiquettes de toutes les tâches A à D en expliquant le sens de chaque étiquette, ceux de l'assistant sont une réponse binaire (« oui » / « non ») pour signaler la présence ou absence de la classe en question. Deux types de messages sont considérés pour l'utilisateur : soit les explications de chaque étiquette sont accompagnées d'exemple positifs, soit elles sont accompagnées d'exemple positifs *et* négatifs (c.-à-d. des contre-exemples). Seule la tâche A (présence ou non d'une information émotionnelle) fait exception puisqu'elle est, en effet, toujours accompagnée de contre-exemples, quelque soit le type d'amorce. La section A.1 donne le détail des résultats sur l'ensemble de test pour chaque approche en comparaison de notre modèle. Les sections A.2 et A.3 montrent ensuite les détails des deux types d'amorces. Nous utilisons la version 0311 de GPT-3.5 pour toutes les expériences.

A.1 Détails des résultats de GPT-3.5

La table 8 donne le détail des résultats sur l'ensemble de test pour chaque approche en comparaison de notre modèle. Dans l'ensemble ces résultats montrent que notre modèle s'en sort mieux et que l'approche sans contre-exemples est meilleure que celle avec contre-exemples. Le problème principal de GPT-3.5 semble être qu'il prédit trop d'étiquettes (rappel élevé mais précision faible). Nous pouvons néanmoins noter que GPT-3.5 semble mieux s'en sortir sur les classes rares car notre modèle ne les prédit pas.

A.2 Sans contre-exemples

Système :

Tu joues le rôle d'un expert linguiste qui annote des phrases en t'intéressant à leur dimension émotionnelle.

L'annotation porte au niveau de la phrase et prend la forme de questions successives. Pour comprendre le contexte, la phrase à annoter est donnée avec sa phrase précédente et sa phrase suivante, mais la réponse à chaque question doit uniquement porter sur la seule phrase à annoter, et non sur la phrase précédente ou suivante.

- Phrase précédente: Nicolas Hulot n'appartient à aucun parti politique.
- Phrase à annoter: Il a refusé trois fois le poste de ministre de l'Ecologie avant d'accepter la proposition d'Emmanuel Macron.
- Phrase suivante: Mais ça ne s'est pas très bien passé.

Utilisateur :

Définition: une phrase est dite "émotionnelle" si elle exprime explicitement ou

| Tâche | Notre modèle | | | | GPT3.5 sans contre-exemples | | | | GPT3.5 avec contre-exemples | | | | Étiquettes |
|-------|--------------|-------|-------|-------|-----------------------------|-------|-------|-------|-----------------------------|-------|-------|-------|----------------|
| | macro-F1 | R | P | F1 | macro-F1 | R | P | F1 | macro-F1 | R | P | F1 | |
| A | 0,752 | 0,764 | 0,741 | 0,752 | 0,518 | 0,622 | 0,443 | 0,518 | 0,382 | 0,744 | 0,257 | 0,382 | Émotionnel |
| B | 0,645 | 0,601 | 0,653 | 0,626 | 0,152 | 0,218 | 0,154 | 0,181 | 0,132 | 0,416 | 0,069 | 0,118 | Comportemental |
| | | 0,811 | 0,803 | 0,807 | | 0,404 | 0,095 | 0,154 | | 0,442 | 0,126 | 0,196 | Désigné |
| | | 0,667 | 0,726 | 0,695 | | 0,897 | 0,066 | 0,122 | | 0,656 | 0,054 | 0,100 | Montré |
| C | 0,601 | 0,426 | 0,479 | 0,451 | 0,199 | 0,534 | 0,089 | 0,153 | 0,220 | 0,372 | 0,066 | 0,112 | Suggéré |
| | | 0,705 | 0,733 | 0,719 | | 0,709 | 0,203 | 0,315 | | 0,466 | 0,286 | 0,354 | Base |
| D | 0,420 | 0,409 | 0,591 | 0,484 | 0,174 | 0,803 | 0,043 | 0,082 | 0,125 | 0,260 | 0,051 | 0,085 | Complexe |
| | | 0,281 | 0,457 | 0,348 | | 0,825 | 0,036 | 0,069 | | 0,526 | 0,028 | 0,052 | Admiration |
| | | 0,745 | 0,592 | 0,660 | | 0,798 | 0,045 | 0,085 | | 0,606 | 0,040 | 0,074 | Autre |
| | | 0,670 | 0,685 | 0,677 | | 0,665 | 0,234 | 0,346 | | 0,590 | 0,268 | 0,369 | Colère |
| | | 0,000 | 0,000 | 0,000 | | 1,000 | 0,222 | 0,364 | | 1,000 | 0,003 | 0,006 | Culpabilité |
| | | 0,000 | 0,000 | 0,000 | | 0,800 | 0,073 | 0,133 | | 0,800 | 0,131 | 0,225 | Dégoût |
| | | 0,364 | 0,600 | 0,453 | | 0,424 | 0,110 | 0,175 | | 0,758 | 0,025 | 0,048 | Embarras |
| | | 0,333 | 0,615 | 0,432 | | 0,958 | 0,023 | 0,045 | | 0,771 | 0,020 | 0,039 | Fierté |
| | | 0,000 | 0,000 | 0,000 | | 0,000 | 0,000 | 0,000 | | 0,000 | 0,000 | 0,000 | Jalousie |
| | | 0,530 | 0,709 | 0,606 | | 0,837 | 0,110 | 0,194 | | 0,683 | 0,148 | 0,243 | Joie |
| | | 0,717 | 0,661 | 0,688 | | 0,731 | 0,153 | 0,253 | | 0,741 | 0,104 | 0,182 | Peur |
| | | 0,697 | 0,739 | 0,717 | | 0,873 | 0,058 | 0,109 | | 0,789 | 0,050 | 0,094 | Surprise |
| | | 0,428 | 0,504 | 0,463 | | 0,449 | 0,247 | 0,319 | | 0,399 | 0,110 | 0,173 | Tristesse |

TABLE 8 – Comparaison des résultats pour notre modèle et les deux approches testées avec GPT-3.5 pour les tâches A, B, C et D.

implicitement une émotion, qu'elle soit exprimée par le narrateur ou un personnage.
Par exemple:

- émotionnelle: "Cette information a beaucoup énervé Marie."
- émotionnelle: "Andrée a sautillé partout en chantant."
- émotionnelle: "Oh, non... C'est vraiment dommage !"
- émotionnelle: "Ces deux amis se retrouvent après une longue séparation."
- non émotionnelle: "Avant d'arriver devant une salle de classe, les enseignants, eux aussi, sont sur les bancs de l'école."
- non émotionnelle: "De 2007 à 2012, il a été le Premier ministre de l'ancien président Nicolas Sarkozy."
- non émotionnelle: "Récemment, une nouvelle autorisation a été délivrée pour un deuxième test dans le courant de l'année 2019."
- non émotionnelle: "Avant de sortir, Billy prépare un dîner orange : une soupe de potiron, des cuisses de canard à l'orange avec une purée de carottes et une tarte à la citrouille."

Question: La phrase à annoter est-elle **émotionnelle** ?

Réponse (oui/non):

Assistant:

<réponse du modèle>

Utilisateur:

Définition: La catégorie émotionnelle "colère" recouvre les émotions suivantes: agacement, colère, contestation, désaccord (si émotion suggérée), désapprobation, énervement, fureur/rage, indignation, insatisfaction, irritation, mécontentement, réprobation et révolte. Par exemple :

- "C'est notamment pour cette raison que des "gilets jaunes", les personnes qui manifestent et bloquent des routes dans le pays depuis plusieurs semaines, sont en colère."
- "- Ton commentaire est déplacé, jeune homme ! a-t-elle dit d'un air pincé."

Question: Si la phrase à annoter est émotionnelle, est-ce que la catégorie

émotionnelle ****colère**** est présente ?

Réponse (oui/non):

Assistant:

<réponse du modèle>

Utilisateur:

Définition: La catégorie émotionnelle "dégoût" recouvre les émotions suivantes: dégoût, lassitude et répulsion. Par exemple :

- "Beurk !"
- "Ça peut paraître dégoûtant, mais on peut manger des insectes."

Question: Si la phrase à annoter est émotionnelle, est-ce que la catégorie émotionnelle ****dégoût**** est présente ?

Réponse (oui/non):

Assistant:

<réponse du modèle>

Utilisateur:

Définition: La catégorie émotionnelle "joie" recouvre les émotions suivantes: amusement, enthousiasme, exaltation, joie et plaisir. Par exemple :

- "Pour fêter ses buts, il lui arrive souvent de danser."
- "- Je suis bien aise de vous voir, me dit le roi sur un ton amical."

Question: Si la phrase à annoter est émotionnelle, est-ce que la catégorie émotionnelle ****joie**** est présente ?

Réponse (oui/non):

Assistant:

<réponse du modèle>

Utilisateur:

Définition: La catégorie émotionnelle "peur" recouvre les émotions suivantes: angoisse, appréhension, effroi, horreur, inquiétude, méfiance, peur, stress et timidité. Par exemple :

- "Le Front national, qui est d'extrême droite, faisait peur, à cause des idées qu'il défendait."
- "Il y avait un grand silence dans la maison."

Question: Si la phrase à annoter est émotionnelle, est-ce que la catégorie émotionnelle ****peur**** est présente ?

Réponse (oui/non):

Assistant:

<réponse du modèle>

Utilisateur:

Définition: La catégorie émotionnelle "surprise" recouvre les émotions suivantes: étonnement, stupeur, surprise. Par exemple :

- "Finalement, ils ont été pris en charge... par les agriculteurs locaux, dans un camion benne !"
- "Tous, étonnés, se taisent."

Question: Si la phrase à annoter est émotionnelle, est-ce que la catégorie émotionnelle ****surprise**** est présente ?

Réponse (oui/non):

Assistant:

<réponse du modèle>

Utilisateur:

Définition: La catégorie émotionnelle "tristesse" recouvre les émotions suivantes: blues, chagrin, déception, désespoir, peine, souffrance et tristesse. Par exemple :

- "Sa mère venait de mourir et son père était au front."
- "L'âne continuait à examiner la peinture d'un regard plutôt attristé."

Question: Si la phrase à annoter est émotionnelle, est-ce que la catégorie émotionnelle ****tristesse**** est présente ?

Réponse (oui/non):

Assistant:

<réponse du modèle>

Utilisateur:

Définition: La catégorie émotionnelle "admiration" recouvre les émotions suivantes: admiration. Par exemple :

- "De nos jours, ce site exceptionnel permet de montrer toute la richesse de la civilisation romaine et la façon dont les villes et la société étaient organisées."
- "- Tes enfants sont vraiment merveilleux, ma chérie, dit-elle à sa fille."

Question: Si la phrase à annoter est émotionnelle, est-ce que la catégorie émotionnelle ****admiration**** est présente ?

Réponse (oui/non):

Assistant:

<réponse du modèle>

Utilisateur:

Définition: La catégorie émotionnelle "culpabilité" recouvre les émotions suivantes: culpabilité. Par exemple :

- "Et je l'avais bien mérité."
- "Surtout, il ne faut pas se sentir coupable de ne pas avoir réagi."

Question: Si la phrase à annoter est émotionnelle, est-ce que la catégorie émotionnelle ****culpabilité**** est présente ?

Réponse (oui/non):

Assistant:

<réponse du modèle>

Utilisateur:

Définition: La catégorie émotionnelle "embarras" recouvre les émotions suivantes: embarras, gêne, honte, humiliation et timidité. Par exemple :

- "Après cette humiliante défaite, Napoléon abdique une nouvelle fois, ce qui marque définitivement la fin de l'Empire et de sa période de retour appelée "les Cent jours"."
- "Légèrement décontenancée, la prof s'est raclé la gorge et commencé la lecture."

Question: Si la phrase à annoter est émotionnelle, est-ce que la catégorie émotionnelle ****embarras**** est présente ?

Réponse (oui/non):

Assistant:

<réponse du modèle>

Utilisateur:

Définition: La catégorie émotionnelle "fierté" recouvre les émotions suivantes: fierté et orgueil. Par exemple :

- "Flavia entre dans la cour comme une conquérante, entourée de ses supporters."

- "Magawa peut être fier de lui, car il vient de recevoir une médaille d'or."

Question: Si la phrase à annoter est émotionnelle, est-ce que la catégorie émotionnelle ****fierté**** est présente ?

Réponse (oui/non):

Assistant:

<réponse du modèle>

Utilisateur:

Définition: La catégorie émotionnelle "jalousie" recouvre les émotions suivantes: jalousie. Par exemple :

- "Mais quand Flavia découvre le jeune génie du piano, elle se sent comme écrasée."
- "On dirait presque qu'il fait partie de l'instrument."

Question: Si la phrase à annoter est émotionnelle, est-ce que la catégorie émotionnelle ****jalousie**** est présente ?

Réponse (oui/non):

Assistant:

<réponse du modèle>

Utilisateur:

Définition: La catégorie émotionnelle "autre" recouvre les émotions suivantes: amour, courage, curiosité, désir, détermination, envie, espoir, haine, impuissance, mépris et soulagement. Par exemple :

- "Dans chaque camp, ils se sont mobilisés pour donner envie aux gens de voter comme eux."
- "Ils n'apprécient pas du tout l'attitude des dirigeants, notamment celle du président, "qu'ils jugent méprisant, déconnecté de la réalité, du quotidien", note le sociologue Alexis Spire."

Question: Si la phrase à annoter est émotionnelle, est-ce que la catégorie émotionnelle ****autre**** est présente ?

Réponse (oui/non):

Assistant:

<réponse du modèle>

Utilisateur:

Définition: Les émotions suivantes sont dites "de base" : Colère, Dégoût, Joie, Peur, Surprise, Tristesse.

Question: Si la phrase à annoter est émotionnelle, contient-elle une ****émotion de base**** ?

Réponse (oui/non):

Assistant:

<réponse du modèle>

Utilisateur:

Définition: Les émotions suivantes sont dites "complexes": Admiration, Culpabilité, Embarras, Fierté, Jalousie.

Question: Si la phrase à annoter est émotionnelle, contient-elle une ****émotion complexe**** ?

Réponse (oui/non):

Assistant:

<réponse du modèle>

Utilisateur:

Définition: Une émotion est dite du mode "désigné" lorsqu'elle est exprimée par un terme du lexique émotionnel. Par exemple :

- "Pierre est heureux d'être bientôt à la retraite.", où la joie de Pierre est désignée par le terme "heureux".
- "Cette information a beaucoup énervé Marie.", où la colère de Marie est désignée par le terme "énervé".

Question: Si la phrase à annoter est émotionnelle, est-ce que le mode ****désigné**** est utilisé ?

Réponse (oui/non):

Assistant:

<réponse du modèle>

Utilisateur:

Définition: Une émotion est dite du mode "comportemental" lorsqu'elle est exprimée par la description d'une manifestation physique (physiologique ou comportementale) de l'émotion. Par exemple :

- "Paul sanglote.", où la tristesse de Paul est exprimée par le comportement "sanglote".
- "Andrée a sautillé partout en chantant.", où la joie de Andrée est exprimée par le comportement "sautillé partout en chantant".

Question: Si la phrase à annoter est émotionnelle, est-ce que le mode ****comportemental**** est utilisé ?

Réponse (oui/non):

Assistant:

<réponse du modèle>

Utilisateur:

Définition: Une émotion est dite du mode "montré" lorsqu'elle est exprimée par des caractéristiques linguistiques de l'énoncé qui traduisent l'état émotionnel dans lequel se trouvait l'énonciateur au moment de l'énonciation. Par exemple :

- "Oh, chouette ! Quelle bonne idée !", car la joie de l'énonciateur est traduite au sein de l'énoncé par les interjections "oh" et "chouette", les énoncés averbaux et les points d'exclamations.
- "Oh, non... C'est vraiment dommage !", car la tristesse de l'énonciateur est traduite au sein de l'énoncé par l'interjection "oh", l'énoncé averbal, les points de suspension et le point d'exclamation.

Question: Si la phrase à annoter est émotionnelle, est-ce que le mode ****montré**** est utilisé ?

Réponse (oui/non):

Assistant:

<réponse du modèle>

Utilisateur:

Définition: Une émotion est dite du mode "suggéré" lorsqu'elle est exprimée par la description d'une situation associée de manière conventionnelle à un ressenti émotionnel. Par exemple :

- "Le père de Jeanne est mort hier à cause d'un cancer.", où la tristesse de Jeanne est suggérée par la description du décès, il y a peu de temps, de son père (une personne proche d'elle).
- "Ces deux amis se retrouvent après une longue séparation.", où la joie des deux amis est suggérée par la description de leurs retrouvailles après un temps long.

Question: Si la phrase à annoter est émotionnelle, est-ce que le mode ****suggéré**** est utilisé ?

Réponse (oui/non):

Assistant:

<réponse du modèle>

A.3 Avec contre-exemples

Systeme:

Tu joues le rôle d'un expert linguiste qui annote des phrases d'après leurs dimensions émotionnelle.

Les différentes annotations sont toute binaires (absence ou présence d'une propriété). Elles vont porter sur la nature émotionnelle ou non des phrases et, si oui, le mode d'expression de la ou des émotions présentes (désignée, comportementale, montrée ou suggérée), la ou les catégories émotionnelles (joie, peur, colère, tristesse, etc.) et le ou les types d'émotion ("de base" ou "complexe"). Chaque propriété est décrite par une définition et des exemples.

L'annotation La phrase à annoter est entourée des balises <annotate>...</annotate>.

Utilisateur:

Définition : une phrase est dite "émotionnelle" si elle exprime explicitement ou implicitement une émotion, qu'elle soit exprimée par le narrateur ou un personnage.

Question : La phrase à annoter est-elle **émotionnelle** ?

Exemples :

- <annotate>Avant de sortir, Billy prépare un dîner orange : une soupe de potiron, des cuisses de canard à l'orange avec une purée de carottes et une tarte à la citrouille.</annotate> -> non
- <annotate>Cette information a beaucoup énervé Marie.</annotate> -> oui
- <annotate>Andrée a sautillé partout en chantant.</annotate> -> oui
- <annotate>Récemment, une nouvelle autorisation a été délivrée pour un deuxième test dans le courant de l'année 2019.</annotate> -> non - <annotate>Oh, non... C'est vraiment dommage !</annotate> -> oui
- <annotate>De 2007 à 2012, il a été le Premier ministre de l'ancien président Nicolas Sarkozy.</annotate> -> non
- <annotate>Ces deux amis se retrouvent après une longue séparation. -> oui
- <annotate>Avant d'arriver devant une salle de classe, les enseignants, eux aussi, sont sur les bancs de l'école.</annotate> -> non

Annotation (oui/non) :

- Nicolas Hulot n'appartient à aucun parti politique. <annotate>Il a refusé trois fois le poste de ministre de l'Ecologie avant d'accepter la proposition d'Emmanuel Macron.</annotate> Mais ça ne s'est pas très bien passé. ->

Assistant:

réponse du modèle

Utilisateur:

Définition : La catégorie émotionnelle "colère" recouvre les émotions suivantes: agacement, colère, contestation, désaccord (si émotion suggérée), désapprobation, énervement, fureur/rage, indignation, insatisfaction, irritation, mécontentement, réprobation et révolte.

Question : Si la phrase à annoter est émotionnelle, est-ce que la catégorie émotionnelle **colère** est présente ?

Exemples :

- <annotate>De 2007 à 2012, il a été le Premier ministre de l'ancien président

Nicolas Sarkozy.</annotate> -> non

- <annotate>C'est notamment pour cette raison que des "gilets jaunes", les personnes qui manifestent et bloquent des routes dans le pays depuis plusieurs semaines, sont en colère.</annotate> -> oui.

- <annotate>Tous, étonnés, se taisent.</annotate> -> non.

- <annotate>- Ton commentaire est déplacé, jeune homme ! a-t-elle dit d'un air pincé.</annotate> -> oui.

- <annotate>Après cette humiliante défaite, Napoléon abdique une nouvelle fois, ce qui marque définitivement la fin de l'Empire et de sa période de retour appelée "les Cent jours".</annotate> -> non.

Annotation (oui/non) :

- Nicolas Hulot n'appartient à aucun parti politique. <annotate>Il a refusé trois fois le poste de ministre de l'Ecologie avant d'accepter la proposition d'Emmanuel Macron.</annotate> Mais ça ne s'est pas très bien passé. ->

Assistant:

réponse du modèle

Utilisateur:

Définition : La catégorie émotionnelle "dégoût" recouvre les émotions suivantes: dégoût, lassitude et répulsion.

Question : Si la phrase à annoter est émotionnelle, est-ce que la catégorie émotionnelle ****dégoût**** est présente ?

Exemples :

- <annotate>Ça peut paraître dégoûtant, mais on peut manger des insectes.</annotate> -> oui.

- <annotate>Beurk !</annotate> -> oui.

- <annotate>Finalement, ils ont été pris en charge... par les agriculteurs locaux, dans un camion benne !</annotate> -> non.

- <annotate>Le Front national, qui est d'extrême droite, faisait peur, à cause des idées qu'il défendait.</annotate> -> non.

- <annotate>Avant d'arriver devant une salle de classe, les enseignants, eux aussi, sont sur les bancs de l'école.</annotate> -> non

Annotation (oui/non) :

- Nicolas Hulot n'appartient à aucun parti politique. <annotate>Il a refusé trois fois le poste de ministre de l'Ecologie avant d'accepter la proposition d'Emmanuel Macron.</annotate> Mais ça ne s'est pas très bien passé. ->

Assistant:

réponse du modèle

Utilisateur:

Définition : La catégorie émotionnelle "joie" recouvre les émotions suivantes: amusement, enthousiasme, exaltation, joie et plaisir.

Question : Si la phrase à annoter est émotionnelle, est-ce que la catégorie émotionnelle ****joie**** est présente ?

Exemples :

- <annotate>Dans chaque camp, ils se sont mobilisés pour donner envie aux gens de voter comme eux.</annotate> -> non.

- <annotate>- Je suis bien aise de vous voir, me dit le roi sur un ton amical.</annotate> -> oui.

- <annotate>Beurk !</annotate> -> non.

- <annotate>Avant d'arriver devant une salle de classe, les enseignants, eux aussi, sont sur les bancs de l'école.</annotate> -> non

- <annotate>Pour fêter ses buts, il lui arrive souvent de danser.</annotate> -> oui.

Annotation (oui/non) :

- Nicolas Hulot n'appartient à aucun parti politique. <annotate>Il a refusé trois fois le poste de ministre de l'Écologie avant d'accepter la proposition d'Emmanuel Macron.</annotate> Mais ça ne s'est pas très bien passé. ->

Assistant:

réponse du modèle

Utilisateur:

Définition : La catégorie émotionnelle "peur" recouvre les émotions suivantes: angoisse, appréhension, effroi, horreur, inquiétude, méfiance, peur, stress et timidité.

Question : Si la phrase à annoter est émotionnelle, est-ce que la catégorie émotionnelle **peur** est présente ?

Exemples :

- <annotate>Le Front national, qui est d'extrême droite, faisait peur, à cause des idées qu'il défendait.</annotate> -> oui.
- <annotate>Dans chaque camp, ils se sont mobilisés pour donner envie aux gens de voter comme eux.</annotate> -> non.
- <annotate>Ça peut paraître dégoûtant, mais on peut manger des insectes.</annotate> -> non.
- <annotate>Récemment, une nouvelle autorisation a été délivrée pour un deuxième test dans le courant de l'année 2019.</annotate> -> non
- <annotate>Il y avait un grand silence dans la maison.</annotate> -> oui.

Annotation (oui/non) :

- Nicolas Hulot n'appartient à aucun parti politique. <annotate>Il a refusé trois fois le poste de ministre de l'Écologie avant d'accepter la proposition d'Emmanuel Macron.</annotate> Mais ça ne s'est pas très bien passé. ->

Assistant:

réponse du modèle

Utilisateur:

Définition : La catégorie émotionnelle "surprise" recouvre les émotions suivantes: étonnement, stupeur, surprise.

Question : Si la phrase à annoter est émotionnelle, est-ce que la catégorie émotionnelle **surprise** est présente ?

Exemples :

- <annotate>Finalement, ils ont été pris en charge... par les agriculteurs locaux, dans un camion benne !</annotate> -> oui.
- <annotate>Avant d'arriver devant une salle de classe, les enseignants, eux aussi, sont sur les bancs de l'école.</annotate> -> non
- <annotate>Mais quand Flavia découvre le jeune génie du piano, elle se sent comme écrasée.</annotate> -> non.
- <annotate>Beurk !</annotate> -> non.
- <annotate>Tous, étonnés, se taisent.</annotate> -> oui.

Annotation (oui/non) :

- Nicolas Hulot n'appartient à aucun parti politique. <annotate>Il a refusé trois fois le poste de ministre de l'Écologie avant d'accepter la proposition d'Emmanuel Macron.</annotate> Mais ça ne s'est pas très bien passé. ->

Assistant:

réponse du modèle

Utilisateur:

Définition : La catégorie émotionnelle "tristesse" recouvre les émotions suivantes: blues, chagrin, déception, désespoir, peine, souffrance et tristesse.

Question : Si la phrase à annoter est émotionnelle, est-ce que la catégorie émotionnelle ****tristesse**** est présente ?

Exemples :

- <annotate>Avant d'arriver devant une salle de classe, les enseignants, eux aussi, sont sur les bancs de l'école.</annotate> -> non
- <annotate>Le Front national, qui est d'extrême droite, faisait peur, à cause des idées qu'il défendait.</annotate> -> non.
- <annotate>Sa mère venait de mourir et son père était au front.</annotate> -> oui.
- <annotate>Légèrement décontenancée, la prof s'est raclé la gorge et commencé la lecture.</annotate> -> non.
- <annotate>L'âne continuait à examiner la peinture d'un regard plutôt attristé.</annotate> -> oui.

Annotation (oui/non) :

- Nicolas Hulot n'appartient à aucun parti politique. <annotate>Il a refusé trois fois le poste de ministre de l'Ecologie avant d'accepter la proposition d'Emmanuel Macron.</annotate> Mais ça ne s'est pas très bien passé. ->

Assistant:

réponse du modèle

Utilisateur:

Définition : La catégorie émotionnelle "admiration" recouvre les émotions suivantes: admiration.

Question : Si la phrase à annoter est émotionnelle, est-ce que la catégorie émotionnelle ****admiration**** est présente ?

Exemples :

- <annotate>Tous, étonnés, se taisent.</annotate> -> non.
- <annotate>De nos jours, ce site exceptionnel permet de montrer toute la richesse de la civilisation romaine et la façon dont les villes et la société étaient organisées.</annotate> -> oui.
- <annotate>Magawa peut être fier de lui, car il vient de recevoir une médaille d'or.</annotate> -> non.
- <annotate>Avant de sortir, Billy prépare un dîner orange : une soupe de potiron, des cuisses de canard à l'orange avec une purée de carottes et une tarte à la citrouille.</annotate> -> non
- <annotate>- Tes enfants sont vraiment merveilleux, ma chérie, dit-elle à sa fille.</annotate> -> oui.

Annotation (oui/non) :

- Nicolas Hulot n'appartient à aucun parti politique. <annotate>Il a refusé trois fois le poste de ministre de l'Ecologie avant d'accepter la proposition d'Emmanuel Macron.</annotate> Mais ça ne s'est pas très bien passé. ->

Assistant:

réponse du modèle

Utilisateur:

Définition : La catégorie émotionnelle "culpabilité" recouvre les émotions suivantes: culpabilité.

Question : Si la phrase à annoter est émotionnelle, est-ce que la catégorie émotionnelle ****culpabilité**** est présente ?

Exemples :

- <annotate>Et je l'avais bien mérité.</annotate> -> oui.
- <annotate>Tous, étonnés, se taisent.</annotate> -> non.
- <annotate>Surtout, il ne faut pas se sentir coupable de ne pas avoir réagi.</annotate> -> oui.
- <annotate>Tous, étonnés, se taisent.</annotate> -> non.

- <annotate>Avant d'arriver devant une salle de classe, les enseignants, eux aussi, sont sur les bancs de l'école.</annotate> -> non

Annotation (oui/non) :

- Nicolas Hulot n'appartient à aucun parti politique. <annotate>Il a refusé trois fois le poste de ministre de l'Écologie avant d'accepter la proposition d'Emmanuel Macron.</annotate> Mais ça ne s'est pas très bien passé. ->

Assistant:

réponse du modèle

Utilisateur:

Définition : La catégorie émotionnelle "embarras" recouvre les émotions suivantes: embarras, gêne, honte, humiliation et timidité.

Question : Si la phrase à annoter est émotionnelle, est-ce que la catégorie émotionnelle ****embarras**** est présente ?

Exemples :

- <annotate>Le Front national, qui est d'extrême droite, faisait peur, à cause des idées qu'il défendait.</annotate> -> non.
- <annotate>- Tes enfants sont vraiment merveilleux, ma chérie, dit-elle à sa fille.</annotate> -> non.
- <annotate>Avant d'arriver devant une salle de classe, les enseignants, eux aussi, sont sur les bancs de l'école.</annotate> -> non
- <annotate>Après cette humiliante défaite, Napoléon abdique une nouvelle fois, ce qui marque définitivement la fin de l'Empire et de sa période de retour appelée "les Cent jours".</annotate> -> oui.
- <annotate>Légèrement décontenancée, la prof s'est raclé la gorge et commencé la lecture.</annotate> -> oui.

Annotation (oui/non) :

- Nicolas Hulot n'appartient à aucun parti politique. <annotate>Il a refusé trois fois le poste de ministre de l'Écologie avant d'accepter la proposition d'Emmanuel Macron.</annotate> Mais ça ne s'est pas très bien passé. ->

Assistant:

réponse du modèle

Utilisateur:

Définition : La catégorie émotionnelle "fierté" recouvre les émotions suivantes: fierté et orgueil.

Question : Si la phrase à annoter est émotionnelle, est-ce que la catégorie émotionnelle ****fierté**** est présente ?

Exemples :

- <annotate>Avant de sortir, Billy prépare un dîner orange : une soupe de potiron, des cuisses de canard à l'orange avec une purée de carottes et une tarte à la citrouille.</annotate> -> non
- <annotate>On dirait presque qu'il fait partie de l'instrument.</annotate> -> non.
- <annotate>Magawa peut être fier de lui, car il vient de recevoir une médaille d'or.</annotate> -> oui.
- <annotate>Flavia entre dans la cour comme une conquérante, entourée de ses supporters.</annotate> -> oui.
- <annotate>Il y avait un grand silence dans la maison.</annotate> -> non.

Annotation (oui/non) :

- Nicolas Hulot n'appartient à aucun parti politique. <annotate>Il a refusé trois fois le poste de ministre de l'Écologie avant d'accepter la proposition d'Emmanuel Macron.</annotate> Mais ça ne s'est pas très bien passé. ->

Assistant:

réponse du modèle

Utilisateur:

Définition : La catégorie émotionnelle "jalousie" recouvre les émotions suivantes: jalousie.

Question : Si la phrase à annoter est émotionnelle, est-ce que la catégorie émotionnelle ****jalousie**** est présente ?

Exemples :

- <annotate>On dirait presque qu'il fait partie de l'instrument.</annotate> -> oui.
- <annotate>Et je l'avais bien mérité.</annotate> -> non.
- <annotate>Et je l'avais bien mérité.</annotate> -> non.
- <annotate>Mais quand Flavia découvre le jeune génie du piano, elle se sent comme écrasée.</annotate> -> oui.
- <annotate>Avant d'arriver devant une salle de classe, les enseignants, eux aussi, sont sur les bancs de l'école.</annotate> -> non

Annotation (oui/non) :

- Nicolas Hulot n'appartient à aucun parti politique. <annotate>Il a refusé trois fois le poste de ministre de l'Ecologie avant d'accepter la proposition d'Emmanuel Macron.</annotate> Mais ça ne s'est pas très bien passé. ->

Assistant:

réponse du modèle

Utilisateur:

Définition : La catégorie émotionnelle "autre" recouvre les émotions suivantes: amour, courage, curiosité, désir, détermination, envie, espoir, haine, impuissance, mépris et soulagement.

Question : Si la phrase à annoter est émotionnelle, est-ce que la catégorie émotionnelle ****autre**** est présente ?

Exemples :

- <annotate>De nos jours, ce site exceptionnel permet de montrer toute la richesse de la civilisation romaine et la façon dont les villes et la société étaient organisées.</annotate> -> non.
- <annotate>L'âne continuait à examiner la peinture d'un regard plutôt attristé.</annotate> -> non.
- <annotate>Récemment, une nouvelle autorisation a été délivrée pour un deuxième test dans le courant de l'année 2019.</annotate> -> non
- <annotate>Ils n'apprécient pas du tout l'attitude des dirigeants, notamment celle du président, "qu'ils jugent méprisant, déconnecté de la réalité, du quotidien", note le sociologue Alexis Spire.</annotate> -> oui.
- <annotate>Dans chaque camp, ils se sont mobilisés pour donner envie aux gens de voter comme eux.</annotate> -> oui.

Annotation (oui/non) :

- Nicolas Hulot n'appartient à aucun parti politique. <annotate>Il a refusé trois fois le poste de ministre de l'Ecologie avant d'accepter la proposition d'Emmanuel Macron.</annotate> Mais ça ne s'est pas très bien passé. ->

Assistant:

réponse du modèle

Utilisateur:

Définition : Les émotions suivantes sont dites "de base" : Colère, Dégoût, Joie, Peur, Surprise, Tristesse.

Question : Si la phrase à annoter est émotionnelle, contient-elle une ****émotion de base**** ?

Annotation (oui/non) :

- Nicolas Hulot n'appartient à aucun parti politique. <annotate>Il a refusé trois fois le poste de ministre de l'Ecologie avant d'accepter la proposition d'Emmanuel

Macron.</annotate> Mais ça ne s'est pas très bien passé. ->

Assistant:

réponse du modèle

Utilisateur:

Définition : Les émotions suivantes sont dites "complexes": Admiration, Culpabilité, Embarras, Fierté, Jalousie.

Question : Si la phrase à annoter est émotionnelle, contient-elle une **émotion complexe** ?

Annotation (oui/non) :

- Nicolas Hulot n'appartient à aucun parti politique. <annotate>Il a refusé trois fois le poste de ministre de l'Ecologie avant d'accepter la proposition d'Emmanuel Macron.</annotate> Mais ça ne s'est pas très bien passé. ->

Assistant:

réponse du modèle

Utilisateur:

Définition : Une émotion est dite du mode "désigné" lorsqu'elle est exprimée par un terme du lexique émotionnel.

Question : Si la phrase à annoter est émotionnelle, est-ce que le mode **désigné** est utilisé ?

Exemples :

- <annotate>Pierre est heureux d'être bientôt à la retraite.</annotate> -> oui (car la joie de Pierre est désignée par le terme "heureux").
- <annotate>Oh, non... C'est vraiment dommage !</annotate> -> non.
- <annotate>Avant d'arriver devant une salle de classe, les enseignants, eux aussi, sont sur les bancs de l'école.</annotate> -> non
- <annotate>Oh, non... C'est vraiment dommage !</annotate> -> non.
- <annotate>Cette information a beaucoup énervé Marie.</annotate> -> oui (car la colère de Marie est désignée par le terme "énervé").

Annotation (oui/non) :

- Nicolas Hulot n'appartient à aucun parti politique. <annotate>Il a refusé trois fois le poste de ministre de l'Ecologie avant d'accepter la proposition d'Emmanuel Macron.</annotate> Mais ça ne s'est pas très bien passé. ->

Assistant:

réponse du modèle

Utilisateur:

Définition : Une émotion est dite du mode "comportemental" lorsqu'elle est exprimée par la description d'une manifestation physique (physiologique ou comportementale) de l'émotion.

Question : Si la phrase à annoter est émotionnelle, est-ce que le mode **comportemental** est utilisé ?

Exemples :

- <annotate>Cette information a beaucoup énervé Marie.</annotate> -> non.
- <annotate>Paul sanglote.</annotate> -> oui (car la tristesse de Paul est exprimée par le comportement "sanglote").
- <annotate>Avant d'arriver devant une salle de classe, les enseignants, eux aussi, sont sur les bancs de l'école.</annotate> -> non
- <annotate>Le père de Jeanne est mort hier à cause d'un cancer.</annotate> -> non.
- <annotate>Andrée a sautillé partout en chantant.</annotate> -> oui (car la joie de Andrée est exprimée par le comportement "sautillé partout en chantant").

Annotation (oui/non) :

- Nicolas Hulot n'appartient à aucun parti politique. <annotate>Il a refusé trois

fois le poste de ministre de l'Écologie avant d'accepter la proposition d'Emmanuel Macron.</annotate> Mais ça ne s'est pas très bien passé. ->

Assistant:

réponse du modèle

Utilisateur:

Définition : Une émotion est dite du mode "montré" lorsqu'elle est exprimée par des caractéristiques linguistiques de l'énoncé qui traduisent l'état émotionnel dans lequel se trouvait l'énonciateur au moment de l'énonciation.

Question : Si la phrase à annoter est émotionnelle, est-ce que le mode ****montré**** est utilisé ?

Exemples :

- <annotate>Andrée a sautillé partout en chantant.</annotate> -> non.
- <annotate>Paul sanglote.</annotate> -> non.
- <annotate>Oh, chouette ! Quelle bonne idée !</annotate> -> oui (car la joie de l'énonciateur est traduite au sein de l'énoncé par les interjections "oh" et "chouette", les énoncés averbaux et les points d'exclamations).
- <annotate>Oh, non... C'est vraiment dommage !</annotate> -> oui (car la tristesse de l'énonciateur est traduite au sein de l'énoncé par l'interjection "oh", l'énoncé averbal, les points de suspension et le point d'exclamation.)
- <annotate>Avant d'arriver devant une salle de classe, les enseignants, eux aussi, sont sur les bancs de l'école.</annotate> -> non

Annotation (oui/non) :

- Nicolas Hulot n'appartient à aucun parti politique. <annotate>Il a refusé trois fois le poste de ministre de l'Écologie avant d'accepter la proposition d'Emmanuel Macron.</annotate> Mais ça ne s'est pas très bien passé. ->

Assistant:

réponse du modèle

Utilisateur:

Définition : Une émotion est dite "suggérée" lorsqu'elle est exprimée par la description d'une situation associée de manière conventionnelle à un ressenti émotionnel.

Question : Si la phrase à annoter est émotionnelle, est-ce que le mode ****suggéré**** est utilisé ?

Exemples :

- <annotate>Oh, chouette ! Quelle bonne idée !</annotate> -> non.
- <annotate>Le père de Jeanne est mort hier à cause d'un cancer.</annotate> -> oui (car où la tristesse de Jeanne est suggérée par la description du décès, il y a peu de temps, de son père, une personne proche d'elle).
- <annotate>Ces deux amis se retrouvent après une longue séparation.</annotate> -> oui (car la joie des deux amis est suggérée par la description de leurs retrouvailles après un temps long).
- <annotate>De 2007 à 2012, il a été le Premier ministre de l'ancien président Nicolas Sarkozy.</annotate> -> non
- <annotate>Andrée a sautillé partout en chantant.</annotate> -> non.

Annotation (oui/non) :

- Nicolas Hulot n'appartient à aucun parti politique. <annotate>Il a refusé trois fois le poste de ministre de l'Écologie avant d'accepter la proposition d'Emmanuel Macron.</annotate> Mais ça ne s'est pas très bien passé. ->

Assistant:

réponse du modèle

Références

- ABDAOUI A., AZÉ J., BRINGAY S. & PONCELET P. (2017). Feel : a french expanded emotion lexicon. *Language Resources and Evaluation*, **51**(3), 833–855. Publisher : Springer.
- ACHEAMPONG F. A., WENYU C. & NUNOO-MENSAH H. (2020). Text-based emotion detection : Advances, challenges, and opportunities. *Engineering Reports*, **2**(7), e12189. Publisher : Wiley Online Library.
- ALM C. O., ROTH D. & SPROAT R. (2005). Emotions from Text : Machine Learning for Text-based Emotion Prediction. In *Proceedings of the Human Language Technology Conference and Conference on Empirical Methods in Natural Language Processing (HLT-EMNLP)*, p. 579–586, Vancouver, British Columbia, Canada : Association for Computational Linguistics.
- AMAN S. & SZPAKOWICZ S. (2007). Identifying expressions of emotion in text. In *Proceedings of the International Conference on Text, Speech and Dialogue (TSD)*, p. 196–205 : Springer.
- BALAHUR A., HERMIDA J. M. & MONTORO A. (2012). Detecting implicit expressions of emotion in text : A comparative analysis. *Decision support systems*, **53**(4), 742–753. Publisher : Elsevier.
- BARON-COHEN S., GOLAN O., WHEELWRIGHT S. & GRANADER Y. (2010). Emotion Word Comprehension from 4 to 16 Years Old : A Developmental Survey. *Frontiers in Evolutionary Neuroscience*, **0**. Publisher : Frontiers, DOI : [10.3389/fnevo.2010.00109](https://doi.org/10.3389/fnevo.2010.00109).
- BIANCHI F., NOZZA D. & HOVY D. (2021). Feel-it : Emotion and sentiment classification for the italian language. In *Proceedings of the Workshop on Computational Approaches to Subjectivity, Sentiment and Social Media Analysis*, p. 76–83.
- BLANC N. & QUENETTE G. (2017). La production d'inférences émotionnelles entre 8 et 10 ans : quelle méthodologie pour quels résultats ? *Enfance*, **4**(4), 503–511. Place : Paris Publisher : NecPlus, DOI : [10.3917/enf1.174.0503](https://doi.org/10.3917/enf1.174.0503).
- BOSTAN L.-A.-M. & KLINGER R. (2018). An Analysis of Annotated Corpora for Emotion Classification in Text. In *Proceedings of the International Conference on Computational Linguistics (COLING)*, p. 2104–2119, Santa Fe, New Mexico, USA : Association for Computational Linguistics.
- BUSSO C., BULUT M., LEE C.-C., KAZEMZADEH A., MOWER E., KIM S., CHANG J. N., LEE S. & NARAYANAN S. S. (2008). Iemocap : Interactive emotional dyadic motion capture database. *Language resources and evaluation*, **42**, 335–359.
- CASEL F., HEINDL A. & KLINGER R. (2021). Emotion recognition under consideration of the emotion component process model. In *Proceedings of the Conference on Natural Language Processing*, p. 49–61, Düsseldorf, Germany : KONVENS 2021 Organizers.
- CHEN S.-Y., HSU C.-C., KUO C.-C., KU L.-W. *et al.* (2018). Emotionlines : An emotion corpus of multi-party conversations. *arXiv preprint arXiv :1802.08379*.
- CHEN T. & GUESTRIN C. (2016). Xgboost : A scalable tree boosting system. In *Proceedings of the ACM SIGKDD international conference on knowledge discovery and data mining*, p. 785–794.
- CREISSEN S. & BLANC N. (2017). Quelle représentation des différentes facettes de la dimension émotionnelle d'une histoire entre l'âge de 6 et 10ans ? apports d'une étude multimédia. *Psychologie Française*, **62**(3), 263–277. Cognition et multimédia : les atouts du numérique en situation d'apprentissage, DOI : <https://doi.org/10.1016/j.psfr.2015.07.006>.
- DAVIDSON D. (2006). The Role of Basic, Self-Conscious and Self-Conscious Evaluative Emotions in Children's Memory and Understanding of Emotion. *Motivation and Emotion*, **30**(3), 232–242. DOI : [10.1007/s11031-006-9037-6](https://doi.org/10.1007/s11031-006-9037-6).

DAVIDSON D., LUO Z. & BURDEN M. J. (2001). Children's recall of emotional behaviours, emotional labels, and nonemotional behaviours : Does emotion enhance memory? *Cognition and Emotion*, **15**(1), 1–26. Place : United Kingdom Publisher : Taylor & Francis, DOI : [10.1080/0269993004200105](https://doi.org/10.1080/0269993004200105).

DEMSZKY D., MOVSHOVITZ-ATTIAS D., KO J., COWEN A. S., NEMADE G. & RAVI S. (2020). GoEmotions : A Dataset of Fine-Grained Emotions. *CoRR*, **abs/2005.00547**. arXiv : 2005.00547.

DIJKSTRA K., ZWAAN R. A., GRAESSER A. C. & MAGLIANO J. P. (1995). Character and reader emotions in literary texts. *Poetics*, **23**(1-2), 139–157. Publisher : Elsevier.

DYER M. G. (1983). The role of affect in narratives. *Cognitive science*, **7**(3), 211–242. Publisher : Wiley Online Library.

EHRET K., BERDICEVSKIS A., BENTZ C. & BLUMENTHAL-DRAMÉ A. (2023). Measuring language complexity : challenges and opportunities. *Linguistics Vanguard*, **9**(s1), 1–8.

EKMAN P. (1992). An argument for basic emotions. *Cognition and Emotion*, **6**(3-4), 169–200. Publisher : Routledge _eprint : <https://doi.org/10.1080/02699939208411068>, DOI : [10.1080/02699939208411068](https://doi.org/10.1080/02699939208411068).

ETIENNE A. (2023). *Analyse automatique des émotions dans les textes : contributions théoriques et applicatives dans le cadre de l'étude de la complexité des textes pour enfants*. Thèse de doctorat, Université de Nanterre-Paris X.

ETIENNE A., BATTISTELLI D. & LECORVÉ G. (2022). A (Psycho-)Linguistically Motivated Scheme for Annotating and Exploring Emotions in a Genre-Diverse Corpus. In *Proceedings of the Conference on Language Resources and Evaluation (LREC)*, Marseille, France.

FRAISSE A. & PAROUBEK P. (2015). Utiliser les interjections pour détecter les émotions. In *Actes de la conférence sur le Traitement Automatique des Langues Naturelles*, p. 279–290.

KIM E. & KLINGER R. (2018). Who feels what and why? annotation of a literature corpus with semantic roles of emotions. In *Proceedings of the International Conference on Computational Linguistics (COLING)*, p. 1345–1359.

KIM E. & KLINGER R. (2019). An Analysis of Emotion Communication Channels in Fan-Fiction : Towards Emotional Storytelling. In *Proceedings of the Workshop on Storytelling*, p. 56–64, Florence, Italy : Association for Computational Linguistics. DOI : [10.18653/v1/W19-3406](https://doi.org/10.18653/v1/W19-3406).

KLINGER R. (2023). Where are we in event-centric emotion analysis? bridging emotion role labeling and appraisal-based approaches.

LIU C., OSAMA M. & ANDRADE A. D. (2019). DENS : A Dataset for Multi-class Emotion Analysis. *CoRR*, **abs/1910.11769**. arXiv : 1910.11769.

MARTIN L., MULLER B., ORTIZ SUÁREZ P. J., DUPONT Y., ROMARY L., DE LA CLERGERIE , SEDDAH D. & SAGOT B. (2020). CamemBERT : a Tasty French Language Model. In *Proceedings of the Annual Meeting of the Association for Computational Linguistics (ACL)*, p. 7203–7219, Online : Association for Computational Linguistics.

MOHAMMAD S. (2012). #Emotional Tweets. In *Proceedings of the Joint Conference on Lexical and International Workshop on Semantic Evaluation (SemEval)*, p. 246–255, Montréal, Canada : Association for Computational Linguistics.

MOHAMMAD S., BRAVO-MARQUEZ F., SALAMEH M. & KIRITCHENKO S. (2018). SemEval-2018 Task 1 : Affect in Tweets. In *Proceedings of the International Workshop on Semantic Evaluation (SemEval)*, p. 1–17, New Orleans, Louisiana : Association for Computational Linguistics. DOI : [10.18653/v1/S18-1001](https://doi.org/10.18653/v1/S18-1001).

OUYANG L., WU J., JIANG X., ALMEIDA D., WAINWRIGHT C., MISHKIN P., ZHANG C., AGARWAL S., SLAMA K., RAY A. *et al.* (2022). Training language models to follow instructions

with human feedback. *Proceedings of the Advances in Neural Information Processing Systems*, **35**, 27730–27744.

PIOLAT A. & BANNOUR R. (2009). An example of text analysis software (emotax-tropes) use : The influence of anxiety on expressive writing. *Current psychology letters. Behaviour, brain & cognition*, **25**(2), 2009).

PLUTCHIK R. (1980). A general psychoevolutionary theory of emotion. In *Theories of emotion*, p. 3–33. Elsevier.

PORIA S., HAZARIKA D., MAJUMDER N., NAIK G., CAMBRIA E. & MIHALCEA R. (2018). Meld : A multimodal multi-party dataset for emotion recognition in conversations. *arXiv preprint arXiv :1810.02508*.

PORIA S., MAJUMDER N., MIHALCEA R. & HOVY E. (2019). Emotion recognition in conversation : Research challenges, datasets, and recent advances. *IEEE Access*, **7**, 100943–100953.

REIMERS N. & GUREVYCH I. (2019). Sentence-bert : Sentence embeddings using siamese bert-networks. In *Proceedings of the Conference on Empirical Methods in Natural Language Processing and the International Joint Conference on Natural Language Processing (EMNLP-IJCNLP)*, p. 3982–3992.

SCHERER K. R. (2005). What are emotions ? And how can they be measured ? *Social science information*, **44**(4), 695–729. Publisher : Sage Publications Sage CA : Thousand Oaks, CA.

STRAPPARAVA C. & MIHALCEA R. (2007). SemEval-2007 Task 14 : Affective Text. In *Proceedings of the International Workshop on Semantic Evaluations (SemEval)*, p. 70–74, Prague, Czech Republic : Association for Computational Linguistics.

SUÁREZ P. J. O., SAGOT B. & ROMARY L. (2019). Asynchronous pipeline for processing huge corpora on medium to low resource infrastructures. In *Proceedings of the Workshop on the Challenges in the Management of Large Corpora* : Leibniz-Institut für Deutsche Sprache.

TAUSCZIK Y. R. & PENNEBAKER J. W. (2010). The psychological meaning of words : LIWC and computerized text analysis methods. *Journal of language and social psychology*, **29**(1), 24–54. Publisher : Sage Publications Sage CA : Los Angeles, CA.

TROIANO E., OBERLÄNDER L. & KLINGER R. (2023). Dimensional modeling of emotions in text with appraisal theories : Corpus creation, annotation reliability, and prediction. *Computational Linguistics*, **49**(1).

ÖHMAN E. (2020). Emotion annotation : Rethinking emotion categorization. *Proceedings of the CEUR Workshop*, **2865**, 134–144. Publisher : CEUR-WS.

ÖHMAN E., PÀMIES M., KAJAVA K. & TIEDEMANN J. (2020). XED : A Multilingual Dataset for Sentiment Analysis and Emotion Detection. In *Proceedings of the International Conference on Computational Linguistics (COLING)*, p. 6542–6552, Barcelona, Spain (Online) : International Committee on Computational Linguistics. DOI : [10.18653/v1/2020.coling-main.575](https://doi.org/10.18653/v1/2020.coling-main.575).