



HAL
open science

Emotion Identification for French in Written Texts: Considering Modes of Emotion Expression as a Step Towards Text Complexity Analysis

Aline Étienne, Delphine Battistelli, Gwéno   Lecorv  

► To cite this version:

Aline   tienne, Delphine Battistelli, Gw  no   Lecorv  . Emotion Identification for French in Written Texts: Considering Modes of Emotion Expression as a Step Towards Text Complexity Analysis. 14th Workshop on Computational Approaches to Subjectivity, Sentiment, & Social Media Analysis. ACL (Association for Computational Linguistics), ACL (Association for Computational Linguistics), Aug 2024, Bangkok, Thailand. hal-04718946

HAL Id: hal-04718946

<https://hal.science/hal-04718946v1>

Submitted on 8 Nov 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 destin  e au d  p  t et    la diffusion de documents scientifiques de niveau recherche, publi  s ou non,   manant des   tablissements d'enseignement et de recherche fran  ais ou   trangers, des laboratoires publics ou priv  s.

Emotion Identification for French in Written Texts: Considering Modes of Emotion Expression as a Step Towards Text Complexity Analysis

Aline Étienne

Univ. Paris-Nanterre, CNRS, MoDyCo – Nanterre, France
acm.etienne@gmail.com

Delphine Battistelli

del.battistelli@gmail.com

Gwénoél Lecorvé

Orange – Lannion, France
gwenole.lecorve@orange.com

Abstract

The objective of this paper is to predict (A) whether a sentence in a written text expresses an emotion, (B) the mode(s) in which the emotion is expressed, (C) whether it is basic or complex, and (D) its emotional category. One of our major contributions, in addition to a dataset and a model¹, 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. The scope is on written texts, i.e. it does not focus on conversational or multi-modal data. In this context, modes of expression are seen as a factor towards the automatic analysis of complexity in texts. Experiments on French texts show acceptable results compared to the human annotators' agreement to predict the mode and category, and outperforming results compared to using a large language model with in-context learning (i.e. no fine-tuning) on all tasks.

1 Introduction

In Natural Language Processing (NLP), emotion detection and classification are often addressed in the context of interactions or conversations (e.g., (Poria et al., 2019)), with either spoken, written (chats, forums, tweets) or multimodal datasets (e.g., (Busso et al., 2008; Poria et al., 2018; Chen et al., 2018)). The goal is usually to identify the emotions felt by speakers in dialogic situations. On the contrary, the analysis of emotions in non-conversational texts, like journalistic and encyclopedic texts or novels, is less developed in NLP. It indeed implies a different goal, which is no longer to characterize the emotional state of speakers but rather of characters/people in these texts. As pointed out in psycholinguistics, emotions in these

types of texts are used—with more or less control by the writer—to capture the reader's attention. They also help to create a connection between the described situations and, thus, are a key factor in understanding (e.g., for children in Davidson et al., 2001). However, it is crucial that these emotions themselves are identified and understood. This leads to the idea that emotions can be considered as a factor of complexity, at least relative complexity in the terminology of Ehret et al. (2023), meaning it takes into account the difficulty perceived by speakers in terms of language learning or understanding. A text will thus be all the more complex as it contains emotions considered complex by a given type of speaker. In the case of children, for example, it is known that certain emotional categories are not accessible in early ages, and that their mode of expression (direct vs. indirect or implicit) also plays a role in accessing their meaning.

From these reflections on the question of emotions as a factor of complexity, this paper is oriented towards a better consideration of the diversity of modes of expression of emotions. We present a model and dataset that introduces the notion of mode of expression in addition to the usual information on emotional categories (e.g., joy, fear, etc.). In practice, the model classifies emotions in texts through four tasks: (A) predict whether a sentence contains an emotion or not; (B) if yes, how it is expressed (the *mode*); (C) whether it is a basic or complex emotion category; and (D) in which emotional category it falls. Examples of these tasks on written texts from our dataset are given in Table 1. The model is a CamemBERT model (Martin et al., 2020) fine-tuned on data including different types of sources (newspapers, novels, encyclopedias) annotated with a psycho-linguistically motivated schema.¹ Evaluation shows that the proposed model outperforms approaches based on expert resources, non-neural architectures (SVM and XGBoost), and in-context learning using GPT-3.5. A

¹Dataset and model can be downloaded on HuggingFace: <https://huggingface.co/TextToKids>.

Sentence (+ surrounding sentences)	(A)	(B) Expr. mode				(C) Type		(D) Emotional category												
	is emotional	behavioral	labeled	displayed	suggested	basic	complex	admiration	other	anger	guilt	disgust	embarrassment	pride	jealousy	joy	fear	surprise	sadness	
How does the coronavirus spread? Epecially through respiratory droplets expelled by an infected person. Respiratory droplets are small droplets of saliva that are released into the air when we talk, cough, or sneeze.																				
It is mainly celebrated in the Anglo-Saxon world. Traditionally, children wear funny costumes. They dress up as often despised and feared creatures such as ghosts, vampires, or witches and go door-to-door in the neighborhood, asking for candies or pastries.	✓		✓			✓										✓				
— He succumbed after ingesting his herbal tea and a toxic substance, presumably cyanide. From there, it was only a small step for Angus’s mother to accuse the king of murder as she rushed towards her brother. — The herbal tea...	✓	✓		✓		✓				✓										
This summer, Nolita had to eat a sausage for the first time in a long time because there was nothing else. "I forced myself," she said. "It disgusted me, and I felt guilty," she recounted.	✓		✓			✓	✓				✓	✓								
At the Rome Olympics, the historic event takes place during the marathon: Ethiopian Abebe Bikila becomes the first athlete from black Africa to become an Olympic champion. What’s more, he achieved this feat... bare-foot! He had indeed developed the habit of running barefoot back home in Ethiopia.	✓			✓	✓	✓	✓							✓		✓			✓	

Table 1: Examples (translated from French) of sentences in context and reference labels for Tasks A (presence of an emotion), B (mode(s) of expression), C (type/complexity of the emotions(s)), and D (emotional category(ies)). Tasks B, C, and D are multi-label tasks.

complementary human evaluation shows that the prediction errors made by the proposed model are generally in the same proportions as those made by humans. Finally, the paper discusses interactions between expression modes and emotional categories. While complexity analysis is the motivation of our work, the paper is restricted to Tasks A-D. Application to complexity analysis are left for future work.

Section 2 browses the literature on emotion identification in written texts, particularly in NLP. Sections 3, 4 and 5 detail the tasks addressed, the associated data, and the proposed model, respectively. Section 6 reports the experiments and results.

2 Framework and Related Work

This section provides a brief overview of the framework for emotion analysis in which the paper is situated and which justifies the choice of schema and data (annotated with this schema). It also positions our work among studies in NLP.

2.1 The Analysis of Emotions as a Complexity Factor of a Text

In psycholinguistics, the key role of characters’ emotions on text comprehension is well-documented (e.g., [Dijkstra et al., 1995](#); [Dyer, 1983](#)). Among recent works, two influencing factors have been highlighted in children’s understanding of emotions, and thus of the texts themselves: the *type of emotion* expressed, basic or complex—the complex emotions (e.g., pride, shame) being more difficult to grasp as they require knowledge of social norms—([Davidson, 2006](#); [Blanc and Quenette, 2017](#)); as well as *the way emotions are expressed* ([Creissen and Blanc, 2017](#))), directly via an emotional label, indirectly through the mention of an emotional behavior, or through the description of an emotional situation, the latter being the most difficult to understand. Of course, the notion of emotional category is also addressed in psycholinguistics, and it has been shown that some categories take longer to be mastered by children (e.g., [Baron-Cohen et al., 2010](#)).

On the NLP side, several works (see e.g., [Bostan and Klinger, 2018](#); [Acheampong et al., 2020](#); [Öhman, 2020](#)) highlight the great heterogeneity of emotion annotation schemas—and annotated corpora—, thus clearly demonstrating the difficulty of modeling emotions and, in the end, of analyzing them. This heterogeneity ranges from the notions (e.g., the number and types of emotional categories) and the type of data studied (journals, tweets, etc.) up to the annotation procedures (*crowdsourcing*, annotation by experts) and evaluation methods implemented (e.g., with or without agreement between annotators). Although some works strive to take into account broader sets of notions and linguistic cues to analyze emotions (e.g., [Casel et al., 2021](#); [Kim and Klinger, 2019](#)), the most commonly used concept remains the notion of *emotional category*, often approached through a list of basic emotions introduced either by [Ekman \(1992\)](#) (anger, disgust, fear, joy, sadness, and surprise) or [Plutchik \(1980\)](#) (Ekman’s categories, anticipation, and trust), with a focus on one way of expressing emotions: the emotional lexicon. As highlighted in ([Klinger, 2023](#)) and in ([Troiano et al., 2023](#)), a few very recent approaches in NLP aim to acquire a deeper understanding of the textual units that support the evocation of emotions outside of directly emotional lexical terms (e.g., "happy", "anger"). These approaches are then inspired by psychological and/or linguistic models of emotions. We adopt the same approach here because we aim to capture both direct and indirect modes of expression of emotions in texts. Like [Troiano et al. \(2023\)](#), we seek to assess to what extent computational models can capture emotions expressed indirectly (e.g., via the description of situations that are/ associated with emotions with regard to social norms and conventions). More specifically, our work adopts the framework proposed by [Etienne et al. \(2022\)](#), which proposes a detailed annotation schema of emotions for French. To our knowledge, this is the only work with the explicit objective of analyzing emotions in texts by addressing both direct and indirect modes of expression in French.

2.2 Automatic Identification of Emotions

In NLP, the analysis of emotions in texts is generally treated as a classification task. The previously mentioned heterogeneity of annotation schemas and annotated corpora is then reflected in the diversity of predicted classes, the granularity of elements to be classified, and the methods for developing

and evaluating classifiers. The way results are presented thus also varies from one paper to another, making performance comparison more difficult.

The focus is often on the classification of basic emotions ([Strapparava and Mihalcea, 2007](#); [Mohammad, 2012](#); [Abdaoui et al., 2017](#); [Demszky et al., 2020](#); [Öhman et al., 2020](#); [Bianchi et al., 2021](#)), although some works use a mix of basic and complex emotions ([Balahur et al., 2012](#); [Fraisie and Paroubek, 2015](#); [Abdaoui et al., 2017](#); [Mohammad et al., 2018](#); [Liu et al., 2019](#); [Demszky et al., 2020](#)). Moreover, there is a long history of building and using emotional lexicons, and the diversity of linguistic markers of emotions (e.g., split syntactic structures, exclamation points) is not systematically taken into account, although it is mentioned in several works ([Alm et al., 2005](#); [Mohammad, 2012](#); [Kim and Klinger, 2018](#); [Demszky et al., 2020](#)) with even, in the case of [Balahur et al. \(2012\)](#) for example, the objective to detect implicit emotions. These works have the limitation of focusing each time only on one mode of expression, thus leaving aside the complementarities between modes. For their part, based on Scherer’s model of emotional components process ([2005](#)), [Casel et al. \(2021\)](#) annotated and then predicted several components of emotions, such as physiological symptoms and motor expressions of emotions, or the cognitive evaluation of events. Although [Casel et al. \(2021\)](#) deal with a broader set of cues, these are not rigorously motivated linguistically. Therefore, relying on [Etienne et al. \(2022\)](#), the originality of our work lies in taking into account different modes of expression of emotions.

Historically, *Support Vector Machine* (SVM) models have been widely used to classify sentences ([Aman and Szpakowicz, 2007](#); [Mohammad, 2012](#)) or texts ([Abdaoui et al., 2017](#); [Balahur et al., 2012](#); [Fraisie and Paroubek, 2015](#); [Mohammad, 2012](#)) according to the emotional category they express. Until the advent of embeddings, the inputs were mainly symbolic: bags of words or *n*-grams, features based on emotional resources such as WordNetAffect ([Aman and Szpakowicz, 2007](#); [Balahur et al., 2012](#); [Strapparava and Mihalcea, 2007](#)) or emotional lexicons ([Strapparava and Mihalcea, 2007](#); [Abdaoui et al., 2017](#); [Kim and Klinger, 2018](#)). Today, neural networks ([Kim and Klinger, 2018](#)) and Transformer architectures ([Liu et al., 2019](#); [Demszky et al., 2020](#); [Öhman et al., 2020](#); [Bianchi et al., 2021, 2022](#)) obviously dominate the state of the art. For French, to our knowl-

edge, the only one approach using Transformer models for emotion recognition that exists is (Cortal et al., 2023). But, as the authors point it out, they need to explore for future work more deeply emotion expression modes in order to improve their classifier. This is the point we are addressing as the modes of emotions are at the heart of our work.

3 Tasks

Built in the global perspective of enabling the analysis of emotions as a complexity factor, our work thus takes into account two key elements to address the complexity of an emotion: its category and its mode of expression. The goal is to propose a Transformer model for 4 classification tasks (noted A, B, C, and D) at the *sentence* level, as opposed to the text level (this can, for example, allow studying how the presence of emotions evolves along a text). Sentences can contain several emotions, as in Table 1. Therefore, the classifications are multi-label with no interdependency across labels within the same task.

3.1 Task A: Presence of Emotion

The first task aims to predict the presence of emotional information in a given sentence (binary prediction).

3.2 Task B: Mode of Expression

The mode of expression focuses on the linguistic means used to convey the presence of an emotion in a text. Following Etienne et al. (2022), 4 modes are considered: the **labeled emotions** directly indicated by a term from the emotional lexicon (e.g., *happy*, *scared*); the **behavioral emotions** which rely on the description of an emotional behavior, such as physiological manifestations (e.g., *crying*, *smiling*) or other behaviors (e.g., *slapping someone*); the **displayed emotions** which are expressed by very heterogeneous surface linguistic features of statements that mainly reflect the emotional state of the writer (e.g., interjections, short sentences); the **suggested emotions** which emanate from the description of a situation generally associated with an emotional feeling according to social norms and conventions (e.g., *seeing a good friend after a long period* suggests joy).

3.3 Task C: Type of Emotion

Task C aims to predict the presence of *basic* and *complex* emotion types (2 simultaneous binary predictions). To our knowledge, this notion has not yet

Subset	Texts	Sent.	Tokens	Emotional sent.
train	1,129	19,553	360K	3,952
dev	182	2,770	53K	438
test	283	5,588	102K	984
Total	1,594	27,911	515K	5,374

Table 2: Statistics over the dataset.

been studied as such in automatic emotion analysis (although the emotional categories *basic* and *complex* have been used in NLP (cf. section 2.2)). This is probably due to the fact that the type of an expressed emotion is directly related to its emotional category. However, the type of emotion is in itself a marker of complexity, as we have seen.

3.4 Task D: Emotional Category

In accordance with Etienne et al. (2022), Task D is designed to label 11+1 emotional categories, namely the 6 basic emotions of Ekman (*anger*, *disgust*, *fear*, *joy*, *sadness*, and *surprise*) and 5 complex emotions (*admiration*, *embarrassment*, *guilt*, *jealousy*, and *pride*). A last category, named *other*, is used to capture markers that express any other emotion (e.g., hate, contempt, love, etc.).

4 Data

As detailed in Table 2, our proposed corpus consists of 1,594 French texts (28K sentences, 515K words) intended for children aged 6 to 14 years, divided into 3 types: mainly journalistic texts (91% of the sentences), encyclopedic articles (9%), and novels (1%). Annotations conducted by 6 experts associate emotional units (segments) in the texts with their mode of expression and emotional category, following the annotation schema and guide in (Etienne et al., 2022). Inter-annotator agreements are presented in Appendix A. These annotations are then merged from the segment level to the sentence level. Thus, a given sentence may cover several emotional units (1.54 unit per sentence on average at least one unit is present). The presence of emotions and the types of emotions have been derived from the mode of expression and emotional category labels. In the end, each sentence is associated with a vector of 19 booleans (again, see Table 1). The data are divided into training, development, and test sets (70/10/20% of the sentences, respectively), such that all sentences from a text are in the same subset, in order to avoid a training bias on the peculiarities of the texts (e.g., the name of a character).

Table 3 presents the proportion of labels within the corpus. Overall, the proportions are comparable from one subset to another. Several imbalances appear within the tasks. **(A)** Only 15-20% of sentences are emotional. **(B)** Modes of expression are quite evenly distributed, *displayed* being the least frequent (3% of sentences) and *suggested* the most common (6%). The sums of the percentages of each mode are higher than the percentages of the *emotional* label because certain emotions are conveyed by several modes and a sentence can also contain several emotional units whose respective modes differ. **(C)** The labels of emotion types are very unbalanced, with a clear dominance of basic emotions. The emotional category 'other' (task D) is not associated with any type of emotion, hence the fact that the sum of the percentages *basic* and *complex* is lower than that of emotional sentences. **(D)** The labels of emotional categories are unbalanced, with percentages always below 5% of sentences. The categories *anger*, *fear*, *joy*, *sadness*, *surprise*, and *other* are dominant, while others are very rare (*disgust*, *guilt*, and *jealousy*).

Finally, Table 4 presents the co-occurrences of each label for Tasks B, C, and D on the test set. Overall, when an emotion is present, it appears several modes (B) are usually used. This is less true for types (C) and categories (D). Still, some emotions are rarely alone, like *disgust*, *embarrassment*, *guilt* or *pride*. Regarding categories, some associations seem more frequent than others, e.g., *joy+pride*, *fear+sadness*, *joy+pride*, and *joy+surprise*.

5 Proposed Model

All tasks are learned together, leading to a single proposed model. This model results from fine-tuning the base version of the pre-trained CamemBERT model (Martin et al., 2020). It is a BERT-type encoder model with 110 million parameters and 12 BERT layers. It was pre-trained on 138GB of French texts (Suárez et al., 2019). Although fine-tuning more recent and larger generative language models like Llama2 or Mistral would likely yield better results, the choice of a reasonably sized model is motivated by two reasons. Firstly, our goal is to demonstrate that, unlike several other tasks in NLP, fine-grained emotion characterization in texts cannot be achieved by leveraging large generic (i.e., non-specialized) language models via in-context learning (i.e., without fine-tuning). Secondly, our work aims for a lightweight solution, so that emo-

Task	Labels	Sent. Prop. (%)		
		train	dev	test
(A) Pres. of emotion	emotional	20.2	15.8	17.6
	behavioral	4.6	3.6	4.3
	labeled	5.3	5.2	5.7
(B) Expression mode	displayed	3.6	2.3	3.5
	suggested	7.1	5.8	6.3
	basic	15.4	12.6	13.9
(C) Emotion type	complex	2.0	2.1	2.3
	admiration	0.6	1.1	1.0
(D) Emotional category	anger	4.6	3.2	3.4
	disgust	0.2	0.3	0.2
	embarrass.	0.6	0.6	0.6
	fear	3.8	3.3	3.8
	guilt	0.1	0.0	0.1
	jealousy	0.0	0.0	0.0
	joy	3.2	2.3	3.6
	pride	0.7	0.4	0.9
	sadness	2.5	2.0	2.5
	surprise	3.0	3.1	2.5
other	5.0	3.2	3.7	

Table 3: Distribution of labels

(B) Expression mode(s)	behavioral	+	+	+	+	no other	(C) Emotion type(s)	basic	+	+	no other		
	labeled	+	+	+	+	basic		+	+	no other			
	displayed	+	+	+	+	complex		+	+	no other			
	suggested	+	+	+	+	basic		-	71	708			
						complex		71	-	56			
behavioral	-	49	23	59	130								
labeled	49	-	28	76	189								
displayed	23	28	-	61	97								
suggested	59	76	61	-	179								
(D) Emotion category(ies)	+ admiration												
	+ anger												
	+ disgust												
	+ embarrass.												
	+ fear												
	+ guilt												
	+ jealousy												
	+ joy												
	+ pride												
	+ sadness												
	+ surprise												
	+ other												
	+ none other												
admiration	-	1	0	0	0	0	0	21	14	4	15	3	23
anger	1	-	0	3	8	0	1	2	0	15	8	10	146
disgust	0	0	-	1	2	1	0	0	0	3	1	1	6
embarrass.	0	3	1	-	6	2	0	1	0	7	1	5	13
fear	0	8	2	6	-	1	0	4	0	35	12	14	145
guilt	0	0	1	2	1	-	0	0	0	0	0	0	1
jealousy	0	1	0	0	0	0	-	0	0	0	0	0	0
joy	21	2	0	1	4	0	0	-	34	7	20	21	124
pride	14	0	0	0	0	0	0	34	-	1	11	5	11
sadness	4	15	3	7	35	0	0	7	1	-	4	12	64
surprise	15	8	1	1	12	0	0	20	11	4	-	5	92
other	3	10	1	5	14	0	0	21	5	12	5	-	149

Table 4: Co-occurrences of the labels in the test set of multi-label tasks B, C, and D.

tion characterization can be used as a processor for analyzing text complexity in a massive collection of texts from a public search engine. Thus, while fine-tuning larger models is part of our future work, this paper does not address it.

We fine-tune the CamemBERT model by replacing its last token prediction layer with a binary classification layer of the size of the number of labels, using binary cross-entropy as the loss function. The fine-tuning involves all model weights, i.e., no layers are frozen. The final model is not directly learned from CamemBERT. An initial fine-tuning is conducted on Task A alone for 3 epochs (classification layer of size 1), then the final model is fine-tuned on all tasks starting from this intermediate model for an additional 6 epochs (the final classification layer is replaced by a fresh layer of size 19). The optimizer is Adam with a learning rate of 10^{-5} (no decay) and batches of 8 examples. The choice of the hyper-parameters was made based on prototyping experiments on the development set.

Other experiments were conducted on the development set², for example, on the choice of a window around sentences, class weighting or not, or the choice of initial fine-tuning only on Task A or not. Ultimately, the results presented are those of the best strategy obtained on the development set averaging results over 3 training runs with different random initializations. Notably, a weighting between classes is adopted so as not to overly favor the majority classes. The maximum weighting factor is capped at 50 to, conversely, not give too much importance to very rare classes. Finally, the model takes as input a triplet of sentences where the target sentence to be labeled is surrounded by its preceding and following sentence in the form before: {previous}</s>current: {target}</s>after: {next}</s>.

6 Automatic and Human Evaluations

6.1 Comparison with Other Models

The proposed model is compared to three other types of models. **SVM** models were trained as they are a historical approach in the field. Two types of input features were used: (i) bag-of-tokens where tokens come from the CamemBERT tokenizer, restricted to those from the training set, resulting in input vectors of dimension 18,437; (ii) sentence embeddings of size 768 obtained with SentenceTransformer (Reimers and Gurevych, 2019) and CamemBERT (the pretrained model, not our fine-tuned model) as the token encoder³. **XG-**

²Details are published in a research report, which cannot be cited here for anonymity reasons.

³<https://huggingface.co/dangvantuan/sentence-camembert-base>

Task	Model	Training variant	Macro R	Macro P	Macro F1
(A) Presence of emotion	Naive		0.20	0.18	0.19
	SVM	bag-of-tokens	0.48	0.66	0.56
		embeddings	0.50	0.67	0.57
	XGBoost	bag-of-tokens	0.22	0.70	0.34
		embeddings	0.19	0.67	0.57
	GPT3.5	pos. ex. only	0.62	0.44	0.52
		pos.+neg. ex.	0.74	0.26	0.38
	ours	all tasks	0.76	0.74	0.75
		only A	0.77	0.73	0.75
	(B) Expression mode	Naive		0.05	0.05
SVM		bag-of-tokens	0.27	0.72	0.37
		embeddings	0.20	0.67	0.29
XGBoost		bag-of-tokens	0.22	0.73	0.31
		embeddings	0.06	0.31	0.10
GPT3.5		pos. ex. only	0.51	0.10	0.15
		pos.+neg. ex.	0.47	0.08	0.13
ours		all tasks	0.63	0.67	0.64
		only A+B	0.63	0.67	0.65
(C) Emotion type		Naive		0.09	0.08
	SVM	bag-of-tokens	0.21	0.34	0.26
		embeddings	0.21	0.83	0.27
	XGBoost	bag-of-tokens	0.12	0.66	0.20
		embeddings	0.07	0.38	0.12
	GPT3.5	pos. ex. only	0.76	0.12	0.20
		pos.+neg. ex.	0.37	0.17	0.22
	ours	all tasks	0.56	0.66	0.60
		only A+C	0.28	0.27	0.27
	(D) Emotional category	Naive		0.02	0.02
SVM		bag-of-tokens	0.12	0.49	0.19
		embeddings	0.12	0.48	0.18
XGBoost		bag-of-tokens	0.19	0.57	0.27
		embeddings	0.04	0.21	0.07
GPT3.5		pos. ex. only	0.70	0.11	0.17
		pos.+neg. ex.	0.64	0.15	0.13
ours		all tasks	0.40	0.46	0.42
		only A+B	0.39	0.46	0.41

Table 5: Model performances (averages over 3 runs, all standard deviations are below 0.02).

Boost models were trained as it is a more recent, lightweight, and competitive technique for many classification tasks, especially with unbalanced data (Chen and Guestrin, 2016). The input features are the same as for the SVMs. Our approach is compared to **GPT-3.5**⁴ (Ouyang et al., 2022). For a given input sample, GPT-3.5 is incrementally solicited to annotate it with binary labels (yes/no). Consecutively for each task and label, a natural language description of what is expected is provided to the model before asking for a response,

⁴Version 0311, with a temperature of zero.

accompanied by examples from the training set for each label. Different prompts were tested exhibiting examples in a few-shot manner for each label. In one version, only 2-4 positive examples are used, whereas the other also reports negative ones (2-3). Details can be found in Appendix C. Unlike SVM, XGBoost, and our model, this approach does require any training. Finally, results for a **naive** approach are also reported to help measuring how far each model is from this lower bound. For each label, this naive approach randomly predicts the presence or absence of the label based on the Bernoulli law parametrized by the frequency of the label in the training set.

Table 5 summarizes the performances on the test set of the models for each task and compares them to our model. Models are evaluated through recall (R), precision (P), and F1 scores. Overall, it appears that our proposed model significantly outperforms SVMs, XGBoost, and GPT-3.5 in terms of F1 scores for all tasks, with values almost double those of the best-ranked model for tasks B, C, and D. It seems especially that all other models tend to favor either recall (GPT-3.5) or precision (SVM, XGBoost), while our model is balanced. Finally, the poor results of GPT-3.5 show that the task is difficult. This calls for either fine-tuning as we did, or more advanced inference techniques of large language models, e.g., by decomposing more the reasoning for each task using *-of-thought techniques or by exploring example selection methods.

Table 5 also reports the comparison with 4 variants of the model where Tasks B to D are trained separately on top of A. These results show that multi-task does not degrade the performance but it does not really improve them neither, except for Task C where guessing the category probably helps. This may lead one to consider that the interaction between mode and category is not very strong. Further discussions are exposed in Section 6.5.

6.2 Comparison with Related Work

In the absence of truly similar work to ours, this section reports additional results to give a better intuition of the performance of our model.

Closest Comparable Works Table 6 summarizes the performances of the three closest works we could find in the literature. They were chosen because they all predict labels at a granularity close to that of the sentence. (Öhman et al., 2020) allows a comparison with another Transformer model; (Fraissee and Paroubek, 2015) with another work

Ref.	Lg	Labels	Model	Lexicon	Granularity	Macro-F1
ours	Fr	anger, disg., joy, fear, surpr., sadn.	Transformer	none	sent. triplets	0.52
(Öhman et al., 2020)	En	same + trust, anticipation	Transformer	none	sent.	0.54
(Kim and Klinger, 2018)	En	same + trust, anticipation	symb.	NRC lexicon	sent. triplets	0.31
			MLP	none	sent. triplets	0.31
(Fraissee, Paroubek, 2015)	Fr	anger, fear, sadness	SVM	custom	paragr.	0.31

Table 6: Comparison elements with close works.

Task	Label	Approach	Macro-F1
(A) Pres. of emotion	emotional	ours	0.75
		TextBlob	0.30
		Emotaix	0.45
(B) Expression mode	behavioral	ours	0.63
		Emotaix	0.04
		labeled	ours
Emotaix	0.56		
(D) Emot. categories (labeled mode only)	all	ours	0.47
		Emotaix	0.43
Emotion polarity	positive	ours	0.58
		TextBlob	0.16
		negative	ours
TextBlob	0.17		

Table 7: Comparison with tools available for French.

in French; and (Kim and Klinger, 2018) with a method that works at the level of linguistic markers (as opposed to the phrasal or textual level). All focus solely on emotional categories. The results show that our model is competitive.

Implementations Based on Existing Resources In the absence of dedicated models for French, two resources are currently available in French if one wants to consider emotion identification in texts: TextBlob (<https://textblob.readthedocs.io/>), a sentiment analysis library that integrates a French lexicon where terms are associated with a negative and positive weight reflecting their polarity; Emotaix (Piolat and Bannour, 2009), another lexicon comprising associations (i)

Source of label	Evaluator’s opinion	Proportion (num. of labels)	
		emot. cat.	expr. mode
human	Agree	94.2% (163)	96.0% (170)
	Disagree	6.8% (10)	4.0% (7)
model	Agree	89.4% (144)	97.1% (166)
	Disagree	10.6% (17)	2.9% (7)
human	Agree	95.5% (105)	97.7% (129)
and model	Disagree	4.5% (5)	2.3% (3)
human but	Agree	92.1% (58)	91.1% (41)
not model	Disagree	7.9% (5)	8.9% (4)
model but	Agree	76.5% (39)	90.2% (37)
not human	Disagree	23.5% (12)	9.8% (4)

Table 8: Experts’ agreement regarding predictions of the human annotator and/or our model.

of terms with emotional categories for the labeled mode only, and (ii) other terms with the behavioral mode (but this time without information on the emotional category). While it is expected that our transformer model outperforms these tools, the experiments allow us to determine the extent of the significant gap and on which tasks (including unseen tasks during training). Several tasks managed by our model were replicated *via* TextBlob and Emotaix. To account for the differences between these resources and our proposed model, Task B was limited to only the behavioral and labeled modes and Task D to the labeled mode. Moreover, our model was tested on a task of predicting emotional polarity on our test set since TextBlob is designed for this use. To predict polarity *via* our model, categories were predicted and empirically projected towards positive or negative polarity (e.g., *anger* is *negative*, *joy* is *positive*). As shown in Table 7, our model performs significantly better than TextBlob and Emotaix, including in the emotional polarity task for which it was not specifically designed. The only task where the competition remains is the prediction of categories when the mode is labeled, which is the easiest situation compared to considering all modes.

6.3 Human Evaluation

Given the difficulty of the tasks considered, it is appropriate to cross-reference the automatic evaluation with a human analysis, particularly to give an intuition of what the observed prediction errors represent. A perceptual validation experiment was thus conducted with three experts in text complexity and emotions. Each of them was informed of the tasks and the definitions of labels in psycholinguistics and linguistics. They were then each

Task	Macro			Labels	R	P	F1
	R	P	F1				
(A) Pres. of emot.	0.76	0.74	0.75	emotional	.76	.74	.75
				behavioral	.60	.65	.63
(B) Express. mode	0.63	0.67	0.64	labeled	.81	.80	.81
				displayed	.67	.73	.70
				suggested	.43	.48	.45
				basic	.70	.73	.72
(C) Emot. type	0.56	0.66	0.60	complex	.41	.59	.48
				admiration	.28	.46	.35
				anger	.67	.68	.68
				disgust	.00	.00	.00
				embarass.	.36	.60	.45
				fear	.72	.66	.69
				guilt	.00	.00	.00
				jealousy	.00	.00	.00
				joy	.53	.71	.61
				pride	.33	.62	.43
	sadness	.43	.50	.46			
	surprise	.70	.74	.72			
(D) Emot. category	0.40	0.46	0.42	other	.75	.59	.66

Table 9: Detailed performances of our model.

confronted with 150 sentences from the test set and their labels for emotional category and mode of expression. These labels came either from the human reference annotations or from the predictions of our model. For each label, the experts had to say whether they agreed or not with the proposed annotation. Of course, they were not aware of the origin of the labels.

Table 8 reports the experts’ agreement rates with the proposed labels, depending on the source of the label.⁵ Although the strongest agreement is when the human and model labels match (*human & model*), the agreement scores are generally very high, especially for the mode of expression. These results thus tend to show that, even when the model predicts differently from the reference, the prediction is generally considered relevant by human experts. This demonstrates that our model is able to generalize correctly and that the F1 scores from previous experiments underestimate the perceived quality of the model’s predictions.

6.4 Results by Label

Table 9 presents the results of our classifier on all labels of all tasks from the test set. Additional observations can be made as follows. Regarding expression modes (B), labeled emotions are very

⁵We considered that the experts agree with a label when at least 2 out of 3 declared that they agreed with the label.

	admiration	anger	disgust	embarrass.	fear	guilt	jealousy	joy	pride	sadness	surprise
behavioral	14	48	12	16	15	7	0	19	8	14	3
labeled	38	12	17	36	41	47	0	25	12	29	23
displayed	15	11	17	16	4	20	33	16	17	11	52
suggested	33	29	55	32	41	27	67	41	62	47	22

(a) Frequencies (%) for each pair (mode, category)

behavioral	.04	.47	—	.00	.23	.00	.01	.17	—	.11	.03
labeled	.06	.18	.02	.07	.50	.01	.00	.27	.05	.18	.21
displayed	.07	.13	.01	.04	.02	—	—	.33	.03	.16	.53
suggested	.05	.18	.03	.03	.24	.00	—	.18	.06	.14	.08

(b) F1 of each mode w.r.t. the reference category

behavioral	.11	.60	—	.16	.30	.00	.00	.26	—	.19	.01
labeled	.45	.29	.00	.57	.63	.00	.00	.53	.45	.44	.41
displayed	.04	.08	.00	.00	.03	—	—	.26	.06	.04	.58
suggested	.16	.35	.00	.31	.34	.00	—	.28	.25	.28	.22

(c) F1 of each category w.r.t. the reference mode

Table 10: Interactions between mode and category: frequency of cooccurrence in the reference (a) ; impact of each on the prediction of the other (b and c).

well recognized ($F1 > 0.8$), unlike suggested emotions ($F1 < 0.5$). This is not surprising as labeled emotions are the easiest to identify for a human annotator, while suggested emotions have the largest part of interpretation (as shown by inter-annotator agreements in Appendix A). Performance for emotion types (C) seems in turn linked to the results on emotional categories, since the *basic* label is, as intuition would suggest, better recognized than the *complex* label. Finally, regarding emotional categories (D), three of them are never predicted (*guilt*, *disgust*, and *jealousy*). These are the rarest labels in the training set, probably too rare for the model to learn to predict them. Indeed, the best-predicted emotional categories are the basic emotions, more frequent, namely the labels *surprise*, *fear*, and *anger* (see Table 3). However, while *surprise* is the best-predicted label of Task D, it is not the most represented in the training set. Conversely, *sadness* is not well recognized, even though it is one of the most frequent emotional categories. From our additional analyses, this seems to be explained by sometimes strong interactions between the notions of expression mode and emotional category.

6.5 Correlation Between Mode and Category

This final section assesses how specific modes impact the prediction of emotional categories, and *vice-versa*. As background information, Table 10.a

reports the cooccurrence relative frequencies between modes and categories in the test set.

Table 10.b explores how the mode’s predictability varies with the emotional category expressed. Strong associations between emotional categories and expression modes enhance mode recognition, but the suggested mode remains challenging across categories, highlighting its complexity for both our model and human annotators.

Then, Table 10.c indicates that F1 scores are generally higher when emotions are expressed through the labeled mode. However, *anger* and *surprise* deviate from this pattern, performing better in *behavioral* and *displayed* modes, respectively. The effectiveness of our model in recognizing emotions like behavioral *anger* and displayed *surprise* is influenced by their strong association with these modes in the training data. However, factors such as the rarity of the emotion in the training set (e.g., *disgust*) and the linguistic characteristics of the mode also play significant roles. For instance, *joy* and *fear* are better recognized when labeled, despite being frequently suggested, due to the inherent challenge in recognizing the suggested mode.

7 Conclusion and Perspectives

In this paper, we have addressed the task of detecting and classifying emotions in written texts, as opposed to conversational data. Due to the applicative perspective towards complexity analysis, we introduced a dataset of French texts and a model (<https://huggingface.co/TextToKids>) which, additionnally to the usual notion of emotion categories, takes into account their direct but also indirect modes of expression. The experiments show that this model performs well compared to other approaches, comparable works, and solutions from off-the-shelf resources. Human evaluation has shown that this level is almost equivalent to what humans can do.

In the future, intra-sentential predictions should be investigated, delimiting units and including other notions such as the experiencers. This would require fine-tuning generative models (as opposed to the discriminative model used in this paper). Then, a direct application of our model is the analysis of complexity as the predicted labels reflect markers of complexity. More broadly, our work could contribute to research in psychology to study the link between emotional language and the psychological state of the writer/speaker.

8 Limitations

The main limitation of the presented article is the choice to work at the sentence level, contextualized by their preceding and following sentences, which restricts the task and may not accurately reflect how a reader analyzes a text. As shown in (Etienne et al., 2022) and (Battistelli et al., 2022), emotional units can be linked to text segments further in the text due to causal relationships (a given situation leads to an emotional state, or an emotional state triggers an action). This can lead to group sentences together, ultimately allowing for a structuring of passages at the text level. The restriction to a window of three sentences does not allow the model to leverage this information. Another consequence of working at the sentence level is the loss of the notion of emotional unit. If two emotions are present in a sentence, the current annotations no longer indicate which segments of the sentence each emotion respectively pertains to. We plan to provide another version of our corpus, annotated at the unit level and without contextual restrictions.

References

- Amine Abdaoui, Jérôme Azé, Sandra Bringay, and Pascal Poncelet. 2017. Feel: a french expanded emotion lexicon. *Language Resources and Evaluation*, 51(3):833–855. Publisher: Springer.
- Francisca Adoma Acheampong, Chen Wenyu, and Henry Nunoo-Mensah. 2020. Text-based emotion detection: Advances, challenges, and opportunities. *Engineering Reports*, 2(7):e12189. Publisher: Wiley Online Library.
- Cecilia Ovesdotter Alm, Dan Roth, and Richard Sproat. 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)*, pages 579–586, Vancouver, British Columbia, Canada. Association for Computational Linguistics.
- Saima Aman and Stan Szpakowicz. 2007. Identifying expressions of emotion in text. In *Proceedings of the International Conference on Text, Speech and Dialogue (TSD)*, pages 196–205. Springer.
- Alexandra Balahur, Jesús M Hermida, and Andrés Montoyo. 2012. Detecting implicit expressions of emotion in text: A comparative analysis. *Decision support systems*, 53(4):742–753. Publisher: Elsevier.
- Simon Baron-Cohen, Ofer Golan, Sally Wheelwright, and Yael Granader. 2010. Emotion Word Comprehension from 4 to 16 Years Old: A Developmental Survey. *Frontiers in Evolutionary Neuroscience*, 0. Publisher: Frontiers.
- Delphine Battistelli, Aline Étienne, and Gwénoél Lecorvé. 2022. L’émotion à un niveau textuel: la fonction structurante des émotions observée à partir d’annotations. *Discours. Revue de linguistique, psycholinguistique et informatique. A journal of linguistics, psycholinguistics and computational linguistics*, (30).
- Federico Bianchi, Debora Nozza, and Dirk Hovy. 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*, pages 76–83.
- Federico Bianchi, Debora Nozza, and Dirk Hovy. 2022. XLM-EMO: Multilingual emotion prediction in social media text. In *Proceedings of the 12th Workshop on Computational Approaches to Subjectivity, Sentiment & Social Media Analysis*, pages 195–203, Dublin, Ireland. Association for Computational Linguistics.
- Nathalie Blanc and Guy Quenette. 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.
- Laura-Ana-Maria Bostan and Roman Klinger. 2018. An Analysis of Annotated Corpora for Emotion Classification in Text. In *Proceedings of the International Conference on Computational Linguistics (COLING)*, pages 2104–2119, Santa Fe, New Mexico, USA. Association for Computational Linguistics.
- Carlos Busso, Murtaza Bulut, Chi-Chun Lee, Abe Kazemzadeh, Emily Mower, Samuel Kim, Jeanette N Chang, Sungbok Lee, and Shrikanth S Narayanan. 2008. Iemocap: Interactive emotional dyadic motion capture database. *Language resources and evaluation*, 42:335–359.
- Felix Casel, Amelie Heindl, and Roman Klinger. 2021. Emotion recognition under consideration of the emotion component process model. In *Proceedings of the Conference on Natural Language Processing*, pages 49–61, Düsseldorf, Germany. KONVENS 2021 Organizers.
- Sheng-Yeh Chen, Chao-Chun Hsu, Chuan-Chun Kuo, Lun-Wei Ku, et al. 2018. Emotionlines: An emotion corpus of multi-party conversations. *arXiv preprint arXiv:1802.08379*.
- Tianqi Chen and Carlos Guestrin. 2016. Xgboost: A scalable tree boosting system. In *Proceedings of the ACM SIGKDD international conference on knowledge discovery and data mining*, pages 785–794.
- Gustave Cortal, Alain Finkel, Patrick Paroubek, and Lina Ye. 2023. Emotion recognition based on psychological components in guided narratives for emotion regulation. In *Proceedings of the 7th Joint SIGHUM Workshop on Computational Linguistics for Cultural*

- Heritage, Social Sciences, Humanities and Literature*, pages 72–81, Dubrovnik, Croatia. Association for Computational Linguistics.
- S. Creissen and N. Blanc. 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.
- Denise Davidson. 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.
- Denise Davidson, Zupai Luo, and Matthew J. Burden. 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.
- Dorottya Demszky, Dana Movshovitz-Attias, Jeongwoo Ko, Alan S. Cowen, Gaurav Nemade, and Sujith Ravi. 2020. [GoEmotions: A Dataset of Fine-Grained Emotions.](#) *CoRR*, abs/2005.00547. ArXiv: 2005.00547.
- Katinka Dijkstra, Rolf A Zwaan, Arthur C Graesser, and Joseph P Magliano. 1995. Character and reader emotions in literary texts. *Poetics*, 23(1-2):139–157. Publisher: Elsevier.
- Michael G Dyer. 1983. The role of affect in narratives. *Cognitive science*, 7(3):211–242. Publisher: Wiley Online Library.
- Katharina Ehret, Aleksandrs Berdicevskis, Christian Bentz, and Alice Blumenthal-Dramé. 2023. Measuring language complexity: challenges and opportunities. *Linguistics Vanguard*, 9(s1):1–8.
- Paul Ekman. 1992. [An argument for basic emotions.](#) *Cognition and Emotion*, 6(3-4):169–200. Publisher: Routledge _eprint: <https://doi.org/10.1080/02699939208411068>.
- Aline Etienne, Delphine Battistelli, and Gwénoél Lecorvé. 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.
- Amel Fraisse and Patrick Paroubek. 2015. Utiliser les interjections pour détecter les émotions. In *Actes de la conférence sur le Traitement Automatique des Langues Naturelles*, pages 279–290.
- Evgeny Kim and Roman Klinger. 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)*, pages 1345–1359.
- Evgeny Kim and Roman Klinger. 2019. [An Analysis of Emotion Communication Channels in Fan-Fiction: Towards Emotional Storytelling.](#) In *Proceedings of the Workshop on Storytelling*, pages 56–64, Florence, Italy. Association for Computational Linguistics.
- Roman Klinger. 2023. [Where are we in event-centric emotion analysis? bridging emotion role labeling and appraisal-based approaches.](#)
- Chen Liu, Muhammad Osama, and Anderson de Andrade. 2019. [DENS: A Dataset for Multi-class Emotion Analysis.](#) *CoRR*, abs/1910.11769. ArXiv: 1910.11769.
- Louis Martin, Benjamin Muller, Pedro Javier Ortiz Suárez, Yoann Dupont, Laurent Romary, Éric de la Clergerie, Djamel Seddah, and Benoît Sagot. 2020. [CamemBERT: a Tasty French Language Model.](#) In *Proceedings of the Annual Meeting of the Association for Computational Linguistics (ACL)*, pages 7203–7219, Online. Association for Computational Linguistics.
- Saif Mohammad. 2012. [#Emotional Tweets.](#) In *Proceedings of the Joint Conference on Lexical and International Workshop on Semantic Evaluation (SemEval)*, pages 246–255, Montréal, Canada. Association for Computational Linguistics.
- Saif Mohammad, Felipe Bravo-Marquez, Mohammad Salameh, and Svetlana Kiritchenko. 2018. [SemEval-2018 Task 1: Affect in Tweets.](#) In *Proceedings of the International Workshop on Semantic Evaluation (SemEval)*, pages 1–17, New Orleans, Louisiana. Association for Computational Linguistics.
- Long Ouyang, Jeffrey Wu, Xu Jiang, Diogo Almeida, Carroll Wainwright, Pamela Mishkin, Chong Zhang, Sandhini Agarwal, Katarina Slama, Alex Ray, et al. 2022. Training language models to follow instructions with human feedback. *Proceedings of the Advances in Neural Information Processing Systems*, 35:27730–27744.
- Annie Piolat and Rachid Bannour. 2009. An example of text analysis software (emotaix-tropes) use: The influence of anxiety on expressive writing. *Current psychology letters. Behaviour, brain & cognition*, 25(2, 2009).
- Robert Plutchik. 1980. A general psychoevolutionary theory of emotion. In *Theories of emotion*, pages 3–33. Elsevier.
- Soujanya Poria, Devamanyu Hazarika, Navonil Majumder, Gautam Naik, Erik Cambria, and Rada Mihalcea. 2018. [Meld: A multimodal multi-party dataset for emotion recognition in conversations.](#) *arXiv preprint arXiv:1810.02508*.
- Soujanya Poria, Navonil Majumder, Rada Mihalcea, and Eduard Hovy. 2019. Emotion recognition in conversation: Research challenges, datasets, and recent advances. *IEEE Access*, 7:100943–100953.

Nils Reimers and Iryna Gurevych. 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)*, pages 3982–3992.

Klaus R Scherer. 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.

Carlo Strapparava and Rada Mihalcea. 2007. **SemEval-2007 Task 14: Affective Text**. In *Proceedings of the International Workshop on Semantic Evaluations (SemEval)*, pages 70–74, Prague, Czech Republic. Association for Computational Linguistics.

Pedro Javier Ortiz Suárez, Benoît Sagot, and Laurent Romary. 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.

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

Emily Öhman. 2020. Emotion annotation: Rethinking emotion categorization. *Proceedings of the CEUR Workshop*, 2865:134–144. Publisher: CEUR-WS.

Emily Öhman, Marc Pàmies, Kaisla Kajava, and Jörg Tiedemann. 2020. **XED: A Multilingual Dataset for Sentiment Analysis and Emotion Detection**. In *Proceedings of the International Conference on Computational Linguistics (COLING)*, pages 6542–6552, Barcelona, Spain (Online). International Committee on Computational Linguistics.

A Inter-Annotator Agreement

To maximize the number of annotations, each text has been annotated by one annotator. Then, the validity of the annotations has been evaluated by comparing the annotations of two productive annotators from the 6 ones in the whole campaign (referred to as A1 and A2) with another 7-th expert. The Cohen’s Kappa for each label is given in 11. For a comparison, in (Kim and Klinger, 2018), *joy* is annotated with a Kappa value of 0.4.

B Confusion matrices

Table 12 reports the confusion matrices for each task for our model.

These details first show that, inspite of the imbalance between emotional and non-emotional samples, the classification is not biased.

Label	Label	A1 vs. A7	A2 vs. A7
(A) Presence of emotion	emotional	0.65	0.68
	behavioral	0.65	0.75
(B) Expression mode	labeled	0.79	0.67
	displayed	0.61	0.74
	suggested	0.53	0.38
(C) Emotion type	basic	0.64	0.68
	complex	0.50	0.61
	admiration	0.36	0.70
	anger	0.73	0.69
(D) Emotion category	disgust	0.75	1.00
	embarrass.	0.53	0.50
	fear	0.65	0.63
	guilt	0.50	–
	jealousy	1.00	–
	joy	0.54	0.47
	pride	0.49	0.00
	sadness	0.55	0.28
surprise	0.60	0.85	

Table 11: Cohen’s Kappa for each label between annotators A1 and A2, and an additional one, A7

Regarding the modes (B), most frequent errors relates to guessing a mode where there is none or the contrary. Then, *behavioral* and *suggested* modes are those with the highest number of false positives. This is probably due to the fact that these mode are less direct and require interpretation. Finally, the most frequent confusions between two modes are *labeled* guessed as *suggested*, and the reciprocal. This maybe means that the definition of a *suggested* emotion falls back in the end to a valence consideration of lexical items.

On the side of types (C), the biggest confusion is with the *none* case. Then, there is no specific bias towards *basic* or *complex*.

Finally, our model primarily predicts the classes of Task D where they are expected. Mainly, no prevalent confusion between classes emerges. When no emotional category is expected, the classes *other*, *anger*, *fear*, and *sadness* are the most predicted. These false positives do not clash with the *other* class, which is very heterogeneous, but are surprising for the other classes, which are normally better defined.

C Prompts for GPT-3.5

GPT-3.5 was used in conversational mode. The prompts are thus an alternation of messages between the *user* and the *assistant*, preceded by a global message from the *system*. The user’s messages cover all the labels from all tasks A to D, explaining the meaning of each label, while the assistant’s responses are binary ("yes" / "no") to

		Predictions				Predictions		
Reference		emo.	not emo.	Reference		basic	comp.	none
	(A)	emo.	not emo.		basic	547	48	205
		emo.	not emo.		comp.	35	54	50
		258	4338			195	17	4536

		Predictions					
Reference		behav.	label.	displ.	sugg.	none	
	(B)	behav.	142	45	19	63	48
		label.	30	260	33	78	30
		displ.	15	25	131	36	53
		sugg.	42	79	44	145	132
		none	60	39	38	88	4536

		Predictions													
Reference		adm.	ang.	disg.	emb.	fear	guilt	jal.	joy	pride	sad.	surp.	other	none	
	(D)	adm.	16	0	0	0	0	0	7	7	1	10	5	23	
		ang.	1	125	0	5	10	0	0	3	0	6	6	12	44
		disg.	0	0	0	2	3	0	0	1	0	6	0	4	1
		emb.	0	2	0	12	2	0	0	0	0	6	4	6	9
		fear	1	11	0	3	151	0	0	3	0	20	8	17	41
		guilt	0	0	0	1	0	0	0	0	0	1	0	0	2
		jal.	0	1	0	0	0	0	0	0	0	0	0	0	0
		joy	9	2	0	1	4	0	0	109	15	6	20	26	52
		pride	4	0	0	1	0	0	0	8	16	0	5	4	20
		sad.	4	9	0	5	26	0	0	5	0	58	4	15	41
		surp.	6	7	0	1	8	0	0	9	5	2	100	6	29
		other	3	9	0	2	12	0	0	18	3	11	3	155	35
		none	9	45	0	1	60	0	0	29	5	35	21	74	4340

Table 12: Confusion matrices of our model for Tasks A, B, C, and D

indicate the presence or absence of the respective class. Two types of messages are considered for the user: either the explanations of each label are accompanied by positive examples, or they are accompanied by both positive and negative examples (i.e., counter-examples). Only Task A (presence or absence of emotional information) is an exception since it is always accompanied by counter-examples, regardless of the type of prompt. Sections C.2 and C.3 then show the details of the two types of prompts. We use version 0311 of GPT-3.5 for all experiments.

C.1 Detailed results

Table 13 provides detailed results on the test set for each approach compared to our model. Overall, these results show that our model performs better and that the approach without counter-examples is better than the one with counter-examples. The main problem with GPT-3.5 seems to be that it predicts too many labels (high recall but low precision). However, it is worth noting that GPT-3.5 seems to perform better on rare classes because our

model does not predict them.

C.2 With positive examples only

System:

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.

User:

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

User:

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>

User:

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>

User:

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>

User:

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."

Tasks	Modèle E				GPT3.5 with positive examples only				GPT3.5 with positive and negative examples				Labels
	macro-F1	R	P	F1	macro-F1	R	P	F1	macro-F1	R	P	F1	
(A)	0.75	0.76	0.74	0.75	0.52	0.62	0.44	0.52	0.38	0.74	0.26	0.38	emotional
(B)	0.64	0.60	0.65	0.63	0.15	0.22	0.15	0.18	0.13	0.42	0.07	0.12	behavioral
		0.81	0.80	0.81		0.40	0.10	0.15		0.44	0.13	0.20	labeled
		0.67	0.73	0.70		0.90	0.07	0.12		0.66	0.05	0.10	displayed
		0.43	0.48	0.45		0.53	0.09	0.15		0.37	0.07	0.11	suggested
(C)	0.60	0.70	0.73	0.72	0.20	0.71	0.20	0.32	0.22	0.47	0.29	0.35	basic
		0.41	0.59	0.48		0.80	0.04	0.08		0.26	0.05	0.09	complex
(D)	0.42	0.28	0.46	0.35	0.17	0.82	0.04	0.07	0.13	0.53	0.03	0.05	admiration
		0.67	0.68	0.68		0.66	0.23	0.35		0.59	0.27	0.37	anger
		0.00	0.00	0.00		0.80	0.07	0.13		0.80	0.13	0.23	disgust
		0.36	0.60	0.45		0.42	0.11	0.18		0.76	0.02	0.05	embarrassment
		0.72	0.66	0.69		0.73	0.15	0.25		0.74	0.10	0.18	fear
		0.00	0.00	0.00		1.00	0.22	0.36		1.00	0.00	0.01	guilt
		0.00	0.00	0.00		0.00	0.00	0.00		0.00	0.00	0.00	jealousy
		0.53	0.71	0.61		0.84	0.11	0.19		0.68	0.15	0.24	joy
		0.33	0.62	0.43		0.96	0.02	0.05		0.77	0.02	0.04	pride
		0.43	0.50	0.46		0.45	0.25	0.32		0.40	0.11	0.17	sadness
		0.70	0.74	0.72		0.87	0.06	0.11		0.79	0.05	0.09	surprise
		0.75	0.59	0.66		0.80	0.04	0.08		0.61	0.04	0.07	other

Table 13: Detailed comparison between our model and the two approaches based on GPT3.5

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

Réponse (oui/non):

Assistant:

<réponse du modèle>

User:

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>

User:

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>

User:

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>

User:

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>

User:

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>

User:

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>

User:

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>

User:

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 méprisante, 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>

User:

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>

User:

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>

User:

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>

User:

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>

User:

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>

User:

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>

C.3 With positive and negative examples

System:

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

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>.

User:

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

User:

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 abdiqua 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

User:

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

User:

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'Ecologie avant d'accepter la proposition d'Emmanuel Macron.</annotate> Mais ça ne s'est pas très bien passé. ->

Assistant:

réponse du modèle

User:

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'Ecologie avant d'accepter la proposition d'Emmanuel Macron.</annotate> Mais ça ne s'est pas très bien passé. ->

Assistant:

réponse du modèle

User:

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'Ecologie avant d'accepter la proposition d'Emmanuel Macron.</annotate> Mais ça ne s'est pas très bien passé. ->

Assistant:

réponse du modèle

User:

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

User:

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

User:

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'Ecologie avant d'accepter la proposition d'Emmanuel Macron.</annotate> Mais ça ne s'est pas très bien passé. ->

Assistant:

réponse du modèle

User:

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'Ecologie avant d'accepter la proposition d'Emmanuel Macron.</annotate> Mais ça ne s'est pas très bien passé. ->

Assistant:

réponse du modèle

User:

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'Ecologie avant d'accepter la proposition d'Emmanuel Macron.</annotate> Mais ça ne s'est pas très bien passé. ->

Assistant:

réponse du modèle

User:

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

User:

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

User:

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

User:

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

User:

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

User:

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'Ecologie avant d'accepter la proposition d'Emmanuel Macron.</annotate> Mais ça ne s'est pas très bien passé. ->

Assistant:

réponse du modèle

User:

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'Ecologie avant d'accepter la proposition d'Emmanuel Macron.</annotate> Mais ça ne s'est pas très bien passé. ->

Assistant:

réponse du modèle

User:

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'Ecologie avant d'accepter la proposition d'Emmanuel Macron.</annotate> Mais ça ne s'est pas très bien passé. ->

Assistant:

réponse du modèle