MultiEmotions-It: a New Dataset for Opinion Polarity and Emotion Analysis for Italian

Rachele Sprugnoli
CIRCSE Research Centre, Università Cattolica del Sacro Cuore
Largo Agostino Gemelli 1, 20123 Milano
rachele.sprugnoli@unicatt.it

Abstract

English. This paper[1] presents a new linguistic resource for Italian, called MultiEmotions-It, containing comments to music videos and advertisements posted on YouTube and Facebook. These comments are manually annotated according to four different dimensions: i.e., relatedness, opinion polarity, emotions and sarcasm. For the annotation of emotions we adopted the Plutchik's model taking into account both basic and complex emotions, i.e. dyads.

1 Introduction

Emotions play an influential role in consumer behaviour affecting the decision to purchase goods and services of different types, including music (Mizerski and White, 1986; Lacher, 1989). Both positive and negative emotions have an influence and this is why marketing strategies have always focused on both rational and emotional aspects (Cotte and Ritchie, 2005).

With the advent of social media, platforms such as YouTube and Facebook have gained importance in the marketing industry because they allow to connect and engage consumers (Kujur and Singh, 2018). The progressive consolidation of social media as marketing spaces has highlighted the need to monitor unstructured data written by social media users. In this context, the application of Sentiment Analysis techniques have flourished with the aim of tracking customers' opinions and attitudes by analysing comments or reviews posted on social media channels (Micu et al., 2017).

In this paper we present a new linguistic resource for Italian, called MultiEmotions-It, containing comments to music videos and advertisement posted on YouTube and Facebook. Comments are manually annotated according to four different dimensions: relatedness, opinion polarity, emotions and sarcasm. Particular attention is devoted to the annotation of emotions for which we adopted the model proposed by Plutchik (1980). Following Plutchik, we take into consideration both the eight basic emotions (`joy, sadness, fear, anger, trust, disgust, surprise, anticipation`) and the dyads, that is feelings composed of two basic emotions (e.g., `love` is a blend of `joy` and `Trust`). At the time of writing, MultiEmotions-It is the only freely available manually annotated dataset for emotion analysis for Italian.[2]

2 Related Works

The computational study of opinions and emotions falls within the scope of the Sentiment Analysis research field (Liu, 2012). Opinion polarity identification is a task aiming at understanding whether a text is expressing positive, negative or neutral sentiment towards the subject of the text. As for emotions, their analysis follows two main approaches (Buechel and Hahn, 2017): in the first one emotions are classified into discrete categories based on the theories of psychologists such as those of Ekman (Ekman, 1992) and Plutchik whereas in the second approach emotions are represented in a dimensional form using continuous values such as *valence*, *arousal* and *dominance* (the so called VAD model).

Survey papers like the ones by Hakak et al. (2017), Bostan & Klinger (2018) and Kim & Klinger (2019) report on studies that focus on different text genres, mainly news (Strapparava and Mihalcea, 2007), social media (Mohammad, 2012) and literary works (Alm et al., 2005).

[2] https://github.com/RacheleSprugnoli/Esercitazioni_SA/tree/master/dataset

Among social media, Twitter is the most studied platform and datasets of annotated tweets are available for different Sentiment Analysis tasks. For emotion analysis see, among others, EmpaTweet (Roberts et al., 2012) and EmoTweet (Liew et al., 2016). The literature also reports works on Facebook posts and YouTube comments with corpora and systems developed for various languages such as English (Preoţiuc-Pietro et al., 2016), Thai (Sarakit et al., 2015), Bangla (Tripto and Ali, 2018) and Indonesian (Savigny and Purwarianti, 2017). As for Italian, there are several emotion lexicons, for example (Araque et al., 2019; Passaro and Lenci, 2016; Mohammad and Turney, 2013; Mohammad, 2018), but, at the moment, no dataset with annotated emotions has been released yet.[3]

Similarly to SenTube (Uryupina et al., 2014), MultiEmotions-It includes YouTube comments and contains the annotation of opinion polarity: however, we also include comments to Facebook posts and we pay particular attention to the categorical annotation of emotions. More specifically, our emotion annotation is inspired by that proposed by Phan et al. (2016) that goes beyond the classification of only the basic emotions to include Plutchik's dyads so to better capture the spectrum of human emotional experience.

	YT MVs	YT Ads	FB Ads	Avg
unrelated	0.43	0.43	0.41	0.42
neutral	0.30	0.50	0.34	0.38
positive	0.59	0.78	0.77	0.71
negative	0.49	0.71	0.64	0.61
joy	0.50	0.61	0.63	0.58
trust	0.33	0.42	0.35	0.37
sadness	0.45	0.28	0.43	0.39
anger	0.47	0.67	0.49	0.54
fear	0.13	0.11	0.10	0.11
disgust	0.48	0.53	0.27	0.43
surprise	0.33	0.12	0.12	0.19
anticipation	0.35	0.11	0.15	0.20
sarcasm	0.49	0.34	0.24	0.36

Table 1: Inter-annotator agreement in terms of Krippendorff's Alpha for YouTube music videos (YT MVs), YouTube advertisements (YT Ads), Facebook advertisements (FB Ads). Last column reports the average across the three categories of comments.

3 Dataset Development

3.1 Data Collection

Comments were scraped from YouTube and Facebook around mid-April 2020 using "Web Scraper"[4], an extension for browsers. We focused on two genres of media contents: music videos (MVs) on YouTube and advertisements (Ads) both in the form of short videos (on YouTube and Facebook) and pictures (only on Facebook).

We chose 9 music videos of the songs presented during Sanremo Music Festival 2020 selecting both songs that reached the top of the chart in the contest and those that ranked in the last positions. All those videos had thousands of comments: we downloaded the most recent ones, at least one hundred comments per video. Finding advertising videos with lots of comments on YouTube was more complicated because many brands disable the possibility of adding comments to their channel. In the end, we managed to select 20 videos of various products, mostly of food and services, such as telecommunication and banking. Similar products and services were also chosen on Facebook by downloading the comments from 13 different posts.

3.2 Data Annotation

The annotation was performed in the context of the "Sentiment Analysis" seminar held within the 'Comunicazione per l'impresa, i media e le organizzazioni complesse"[5] master's degree at Università Cattolica del Sacro Cuore in Milan. The annotation process lasted 1 week and involved thirty six students: each student annotated 30 comments for each category (i.e., YouTube MVs, YouTube Ads, Facebook Ads) for a total of 90 comments. Each comment was annotated by two students. It is important to note that students had no previous experience in linguistic annotation but had specific training in the strategic management of communication flows on various media platforms.

Annotation Guidelines. Students were required to annotate the following four dimensions for each comment; a comment may consist of more than one sentence but was analysed as a single unit:

1. Relatedness: does the comment refer to the media content? Is the comment written in a

[3]Annotated datasets for emotion analysis have been mainly developed in enterprises and are not public, see for example (Bolioli et al., 2013).

[4]`https://webscraper.io/`

[5]EN: "Communication for the enterprise, the media and complex organizations"

COMMENT	UNR	NEU	POS	NEG	JOY	TRU	SAD	ANG	FEA	DIS	SUR	ANT	SAR	EMOTIONS
Saludos desde Argentina!!!	1	0	0	0	0	0	0	0	0	0	0	0	0	
Ha superato diodato con le vius	0	1	0	0	0	0	0	0	0	0	0	0	0	
Frizzante come una Coca light scaduta	0	0	0	1	0	0	0	0	0	1	0	0	1	`disgust`
Questa canzone mi ha rovinato l'esistenza	0	0	0	1	0	0	1	0	1	0	0	0	0	`despair`
Idea interessante! Meno l'esecuzione...	0	0	1	1	0	1	1	0	0	0	1	0	0	`trust - disappointment`

Table 2: Examples of annotation.

	YT MVs	YT Ads	FB Ads
# COMMENTS	1,080	1,080	1,080
# WORDS	17,762	19,603	20,844
`unrelated`	65	59	51
`neutral`	84	96	115
`positive`	896	509	597
`negative`	46	434	373
`joy`	397	172	149
`trust`	797	463	574
`sadness`	188	213	292
`anger`	30	211	117
`fear`	5	29	32
`disgust`	32	155	48
`surprise`	174	140	199
`anticipation`	28	56	42
`sarcasm`	7	30	8

Table 3: Dataset Statistics.

language other than Italian? Comments that are not related to the media content or that are not written in Italian are to be annotated as `unrelated`.

2. Opinion Polarity: is the comment positive, negative or neutral with respect to the media content? Positive and negative polarities are not mutually exclusive: a comment can have a mixed polarity containing both positive and negative opinions on different aspects of the media content.

3. Emotions: what emotions are expressed in the comment? This dimension applies only to comments with positive or negative opinion polarity. Each comment can be annotated with one or more emotions at the same time: the list of emotions to assign includes Plutchik's basic emotions and dyads. Conflict or mixed emotions can appear in the same comment.

4. Sarcasm: are emotions expressed using sarcasm? Following Gibbs (2000), we define sarcasm as a language device that conveys the opposite of its literal meaning (Cignarella et al., 2018).

Annotation was carried on using spreadsheets where the aforementioned dimensions were converted into 13 fields: `unrelated`, `neutral`, `positive`, `negative`, `joy`, `trust`, `sadness`, `anger`, `fear`, `disgust`, `surprise`, `anticipation`, `sarcasm`. Each field had to be filled in with a binary value: 0 (the dimension is absent) or 1 (the dimension is present). Spreadsheets contained 4 additional metadata fields: `type`, `title`, `URL`, `comment`. For the annotation, students were provided with the images of Plutchik's "Wheel of Emotions" [6] and of the combination of emotions in dyads [7].

Inter-Annotator Agreement. Table 1 reports the results of the inter-annotator agreement (IAA): we measured the Krippendorff's Alpha for each label and for each pair of annotators and then we computed the average for each type of comment. The average across the three type of comments is reported in the table as well. For all the labels, IAA is below the 0.8 threshold usually considered as good reliability for content analysis research (Klaus, 1980; Artstein and Poesio, 2008), however these results are in line with the ones obtained in similar works presenting a multi-label annotation of emotions or the annotation of mixed emotions (Aman and Szpakowicz, 2007; Phan et al., 2016). The analysis of the cases of disagreement revealed several interesting issues: i) labels `unrelated` and `neutral` tended to be confused with each other. For example, the comment *Qualcuno mi sa dire dove si trova il porticato della quinta immagine?* (Can anyone tell me where the portico in

[6] `https://commons.wikimedia.org/wiki/File:Plutchik-wheel.svg`

[7] `https://i.pinimg.com/originals/83/93/d6/8393d660082c3124a684edc3cade4607.jpg`

DYADS	`love`	*amo questa musica* EN: I love this music
	`disappointment`	*Io non capisco come faccia ad essere fra le ultime questa canzone.* EN: I don't understand how this song is ranked so low.
	`sentimentality`	*Mi veniva da piangere.... Ricordavo la vecchia pubblicità* EN: It makes me want to cry...I remembered the old advertisement
MIX	`trust - disappointment`	*Bellissima!!! Come possa essere ultima! Mah...* EN: Gorgeous!!! How can it be the last! Mah ...
	`trust - sentimentality`	*Io ho pianto. Complimenti a Barilla* EN: I cried. Congratulations to Barilla
	`love - sentimentality`	*A te la manina tremava e io piangevo..* ♥ EN: Your hand was shaking and I was crying ..♥

Table 4: Top 3 dyads and mixes of emotions in the dataset with associated examples. Dyads mentioned in the table are composed by two basic emotions as follows: `love = joy + trust; disappointment = surprise + sadness; sentimentality = trust + sadness`.

the fifth image is located?) is related to the content of the video but it is neutral; ii) sarcasm was confused with other forms of figurative language such as metaphors, e.g. *È l'Ibrahimovic dei biscotti: perfetto* (EN: it is the Ibrahimovic of biscuits: perfect); iii) the assignment of `positive` and `negative` labels registered the highest scores (average Alpha across the 3 categories: 0.71 for `positive` and 0.61 for `negative`). Nevertheless, sometimes annotators failed to distinguish between the annotation of opinion polarity and the annotation of emotions by assigning a negative polarity to comments containing negative emotions. However, the two dimensions do not always match: for example, the comment *sta canzone meritava molto di più* (EN: this song deserved much more) expresses disappointment but also an implicit appreciation for the song and thus a positive opinion polarity. iv) the IAA on the single emotion labels varies greatly: a similar wide variability is reported also in previous works even when dealing with non multi-label annotation (Strapparava and Mihalcea, 2008; Aman and Szpakowicz, 2007).

Creation of the Ground Truth. All comments were manually revised and disagreement were reconciled so to assign gold labels. In this way, we generated a ground truth dataset where the noise coming from the annotation of non-expert annotators was minimized. Moreover, the field `emotions` was added to the spreadsheets so to make explicit the name of the emotions conveyed by the comments. Table 2 shows the structure of the final dataset (metadata fields are not displayed due to space limitation) and some examples of annotation. In particular, the table reports: an unrelated comment, a neutral comment, a comment with a negative polarity, a basic emotion (i.e. `disgust`) and sarcasm, a comment with a negative polarity and a dyad (i.e., `disgust` which is made of `sadness` and `fear`), a comment with mixed polarity and mixed emotions.

4 Dataset Analysis

Table 3 summarizes the statistics of our final dataset showing the distribution of labels in the three categories of media content. MultiEmotions-It contains 3,240 comments for a total of more than 58,000 tokens. Only 470 comments (14.5% of the whole dataset) have no associated emotions because annotated as `unrelated` or `neutral`. Comments with positive opinion polarity are more than those with negative polarity: this is especially evident for YouTube MVs that are mostly commented by supporters of the artists performing in the video. Sarcasm is not a pervasive phenomenon: the number of comments annotated with the corresponding label is marginal, covering 1.6% of the total number of comments with an affective content, i.e. annotated with at least one emotion. More specifically, sarcasm co-occurs with two basic emotions: that is, `anger` (10 comments) and `disgust` (9 comments).

As for emotions, `trust` is the most frequent one: indeed, many comments express admiration towards the media content in different ways, for example by thanking the brand, declaring loyalty to a product or expressing appreciation for a specific feature of the media content (e.g. the location of the video). The emotion `trust` does not appear in the dataset only as a basic emotion but also in several combinations: indeed, 36.5% of the comments with an affective content

are annotated with a dyad and 18.3% with a mix of emotions. Table 4 reports the 3 most frequent dyads and mixes of emotions in the dataset together with an example. As shown in the table, `sentimentality` (that is a combination of `trust` and `sadness`) plays an important role in Ads that try to induce a deep, overwhelming emotional response. Indeed, `sentimentality` is an emotion that marketing research has identified as a fundamental purchase decision variable (Morton et al., 2013).
`Optimism` (`anticipation` + `joy`) and `pessimism` (`anticipation` + `sadness`) are not very frequent in the dataset with 65 and 16 occurrences respectively. However, it is interesting to note that they are mainly associated with comments on advertisements related to the COVID-19 pandemic, for example:

- `optimism`: *All'Italia che, ancora una volta, resiste!* EN: To Italy that, once again, resists!
- `pessimism`: *mamma mia quanta retorica spicciola ....finita l'epidemia staremo tutti ad odiarci e ad insultarci come sempre ....un paese che non ha senso più di esistere* EN: oh my gosh, how much rhetoric .... once the epidemic is over we will all be hating and insulting each other as always .... a country that no longer makes sense to exist

5 Baseline System

To establish a baseline on our data, we developed a simple multi-label classification model using the fastText library (Joulin et al., 2016).[8] The aim of the model is to assign the correct emotion labels to comments. To this end, we randomly split comments and their annotated emotion labels into train and validation following an 80:20 ratio, thus having 2,592 comments for training and the remaining 648 for testing the performance of the learned classifier on new data. Texts have been lowercased and punctuation removed. We trained the model with the following parameters:

- learning rate: 0.5
- epochs: 25
- word n-grams: 2
- loss function: one-vs-all

With the previous setting, we obtained 0.57 Precision, 0.43 Recall and 0.49 F-measure. Only four labels registered a F-measure above 0.5: i.e., `trust` (0.68), `love` (0.54), `delight` (0.53), `sentimentality` (0.50).[9]

6 Conclusion and Future Work

This paper describes MultiEmotions-It, a new manually annotated dataset for opinion polarity and emotion analysis made of more than 3,000 comments on music videos and advertisements published on YouTube and Facebook.
As for future work, we plan to: (i) extend the annotation guidelines to distinguish the specific object towards which the opinion is directed (e.g. the product, the actor, the location of the video) following the work by Severyn et al. (2016), (ii) extend the dataset with new comments taken also from Instagram and Twitter, (iii) extract a new word-emotion association lexicon from MultiEmotions-It using vector space models (Passaro et al., 2015) in order to cover complex emotions.

Acknowledgements

The author wants to thank the students of the "Sentiment Analysis" seminar held within the "Comunicazione per l'impresa, i media e le organizzazioni complesse" master's degree at Università Cattolica del Sacro Cuore (Milan) for the annotation they performed.

References

Cecilia Ovesdotter Alm, Dan Roth, and Richard Sproat. 2005. Emotions from text: machine learning for text-based emotion prediction. In *Proceedings of human language technology conference and conference on empirical methods in natural language processing*, pages 579–586.

Saima Aman and Stan Szpakowicz. 2007. Identifying expressions of emotion in text. In *International Conference on Text, Speech and Dialogue*, pages 196–205. Springer.

Oscar Araque, Lorenzo Gatti, Jacopo Staiano, and Marco Guerini. 2019. Depechemood++: a bilingual emotion lexicon built through simple yet powerful techniques. *IEEE transactions on affective computing*.

[8] `https://fasttext.cc/docs/en/supervised-tutorial.html`

[9] `Love` is a blend of `joy` and `trust`; `delight` is a dyad made of `love` and `surprise`; `sentimentality` is made of `trust` and `sadness`.

Ron Artstein and Massimo Poesio. 2008. Inter-coder agreement for computational linguistics. *Computational Linguistics*, 34(4):555–596.

Andrea Bolioli, Federica Salamino, and Veronica Porzionato. 2013. Social media monitoring in real life with blogmeter platform. *ESSEM@ AI* IA*, 1096:156–163.

Laura-Ana-Maria Bostan and Roman Klinger. 2018. An analysis of annotated corpora for emotion classification in text. In *Proceedings of the 27th International Conference on Computational Linguistics*, pages 2104–2119, Santa Fe, New Mexico, USA, August. Association for Computational Linguistics.

Sven Buechel and Udo Hahn. 2017. Emobank: Studying the impact of annotation perspective and representation format on dimensional emotion analysis. In *Proceedings of the 15th Conference of the European Chapter of the Association for Computational Linguistics: Volume 2, Short Papers*, pages 578–585.

Alessandra Teresa Cignarella, Simona Frenda, Valerio Basile, Cristina Bosco, Viviana Patti, Paolo Rosso, et al. 2018. Overview of the EVALITA 2018 task on irony detection in italian tweets (IronITA). In *Sixth Evaluation Campaign of Natural Language Processing and Speech Tools for Italian (EVALITA 2018)*, volume 2263, pages 1–6. CEUR-WS.

June Cotte and Robin Ritchie. 2005. Advertisers' theories of consumers: Why use negative emotions to sell? *ACR North American Advances*.

Paul Ekman. 1992. An argument for basic emotions. *Cognition & emotion*, 6(3-4):169–200.

Raymond W Gibbs. 2000. Irony in talk among friends. *Metaphor and symbol*, 15(1-2):5–27.

Nida Manzoor Hakak, Mohsin Mohd, Mahira Kirmani, and Mudasir Mohd. 2017. Emotion analysis: A survey. In *2017 International Conference on Computer, Communications and Electronics (COMPTELIX)*, pages 397–402. IEEE.

Armand Joulin, Edouard Grave, Piotr Bojanowski, and Tomas Mikolov. 2016. Bag of tricks for efficient text classification. *arXiv preprint arXiv:1607.01759*.

Evgeny Kim and Roman Klinger. 2019. A survey on sentiment and emotion analysis for computational literary studies. *Zeitschrift fuer Digitale Geisteswissenschaften*, 4.

Krippendorff Klaus. 1980. Content analysis: An introduction to its methodology.

Fedric Kujur and Saumya Singh. 2018. Emotions as predictor for consumer engagement in youtube advertisement. *Journal of Advances in Management Research*.

Kathleen T Lacher. 1989. Hedonic consumption: Music as a product. *ACR North American Advances*.

Jasy Suet Yan Liew, Howard R Turtle, and Elizabeth D Liddy. 2016. EmoTweet-28: a fine-grained emotion corpus for sentiment analysis. In *Proceedings of the Tenth International Conference on Language Resources and Evaluation (LREC'16)*, pages 1149–1156.

Bing Liu. 2012. *Sentiment analysis and opinion mining*, volume 5 of *Synthesis lectures on human language technologies*. Morgan & Claypool Publishers.

Adrian Micu, Angela Eliza Micu, Marius Geru, and Radu Constantin Lixandroiu. 2017. Analyzing user sentiment in social media: Implications for online marketing strategy. *Psychology & Marketing*, 34(12):1094–1100.

Richard W Mizerski and J Dennis White. 1986. Understanding and using emotions in advertising. *Journal of Consumer Marketing*.

Saif M Mohammad and Peter D Turney. 2013. Crowdsourcing a word–emotion association lexicon. *Computational Intelligence*, 29(3):436–465.

Saif Mohammad. 2012. # emotional tweets. In * *SEM 2012: The First Joint Conference on Lexical and Computational Semantics–Volume 1: Proceedings of the main conference and the shared task, and Volume 2: Proceedings of the Sixth International Workshop on Semantic Evaluation (SemEval 2012)*, pages 246–255.

Saif M. Mohammad. 2018. Obtaining reliable human ratings of valence, arousal, and dominance for 20,000 english words. In *Proceedings of The Annual Conference of the Association for Computational Linguistics (ACL)*, Melbourne, Australia.

Anne-Louise Morton, Cheryl Rivers, Stephen Charters, and Wendy Spinks. 2013. Champagne purchasing: the influence of kudos and sentimentality. *Qualitative Market Research: an international journal*.

Lucia C. Passaro and Alessandro Lenci. 2016. Evaluating context selection strategies to build emotive vector space models. In *Proceedings of the Tenth International Conference on Language Resources and Evaluation LREC 2016*. European Language Resources Association (ELRA).

Lucia Passaro, Laura Pollacci, and Alessandro Lenci. 2015. ItEM: A vector space model to bootstrap an Italian emotive lexicon. In *Second Italian Conference on Computational Linguistics CLiC-it 2015*, pages 215–220. Academia University Press.

Duc-Anh Phan, Hiroyuki Shindo, and Yuji Matsumoto. 2016. Multiple emotions detection in conversation transcripts. In *Proceedings of the 30th Pacific Asia Conference on Language, Information and Computation: Oral Papers*, pages 85–94.

Robert Plutchik. 1980. A general psychoevolutionary theory of emotion. In *Theories of emotion*, pages 3–33. Elsevier.

Daniel Preoţiuc-Pietro, H Andrew Schwartz, Gregory Park, Johannes Eichstaedt, Margaret Kern, Lyle Ungar, and Elisabeth Shulman. 2016. Modelling valence and arousal in facebook posts. In *Proceedings of the 7th workshop on computational approaches to subjectivity, sentiment and social media analysis*, pages 9–15.

Kirk Roberts, Michael A Roach, Joseph Johnson, Josh Guthrie, and Sanda M Harabagiu. 2012. EmpaTweet: Annotating and Detecting Emotions on Twitter. In *Proceedings of the Eighth International Conference on Language Resources and Evaluation (LREC'12)*, volume 12, pages 3806–3813. Citeseer.

Phakhawat Sarakit, Thanaruk Theeramunkong, Choochart Haruechaiyasak, and Manabu Okumura. 2015. Classifying emotion in thai youtube comments. In *2015 6th International Conference of Information and Communication Technology for Embedded Systems (IC-ICTES)*, pages 1–5. IEEE.

Julio Savigny and Ayu Purwarianti. 2017. Emotion classification on youtube comments using word embedding. In *2017 International Conference on Advanced Informatics, Concepts, Theory, and Applications (ICAICTA)*, pages 1–5. IEEE.

Aliaksei Severyn, Alessandro Moschitti, Olga Uryupina, Barbara Plank, and Katja Filippova. 2016. Multi-lingual opinion mining on youtube. *Information Processing & Management*, 52(1):46–60.

Carlo Strapparava and Rada Mihalcea. 2007. Semeval-2007 task 14: Affective text. In *Proceedings of the Fourth International Workshop on Semantic Evaluations (SemEval-2007)*, pages 70–74.

Carlo Strapparava and Rada Mihalcea. 2008. Learning to identify emotions in text. In *Proceedings of the 2008 ACM symposium on Applied computing*, pages 1556–1560.

Nafis Irtiza Tripto and Mohammed Eunus Ali. 2018. Detecting multilabel sentiment and emotions from bangla youtube comments. In *2018 International Conference on Bangla Speech and Language Processing (ICBSLP)*, pages 1–6. IEEE.

Olga Uryupina, Barbara Plank, Aliaksei Severyn, Agata Rotondi, and Alessandro Moschitti. 2014. Sentube: A corpus for sentiment analysis on youtube social media. In *LREC*, pages 4244–4249.