SM-FEEL-BG - The First Bulgarian Datasets and Classifiers for Detecting Feelings, Emotions, and Sentiments of Bulgarian Social Media Texts

Irina Temnikova, Iva Marinova*, Silvia Gargova, Ruslana Margova, Stefan Minkov, Alexander Komarov, Nevena Grigorova, Yordan Mandevski, Tsvetelina Stefanova, Veneta Kireva, Dimana Vyatrova
Big Data for Smart Society Institute (GATE)
Institute of Information and Communication Technologies, Bulgarian Academy of Sciences*
Sofia, Bulgaria
{irina.temnikova, silvia.gargova, ruslana.margova}@gate-ai.eu,
iva.marinova@identrics.ai*, freelancer.neva@gmail.com, av.kom0411@gmail.com,
tsveti.stefanov@gmail.com, venkireva80@gmail.com, dimana.vqtrova@gmail.com,
danimandevski@gmail.com, stefanminkov2003@gmail.com

Abstract

This article introduces SM-FEEL-BG – the first Bulgarian-language package, containing 6 datasets with Social Media (SM) texts with emotion, feeling, and sentiment labels and 4 classifiers trained on them. All but one dataset from these are freely accessible for research purposes. The largest dataset contains 6000 Twitter, Telegram, and Facebook texts, manually annotated with 21 fine-grained emotion/feeling categories. The fine-grained labels are automatically merged into three coarse-grained sentiment categories, producing a dataset with two parallel sets of labels. Several classification experiments are run on different subsets of the fine-grained categories and their respective sentiment labels with a Bulgarian fine-tuned BERT. The highest Acc. reached was 0.61 for 16 emotions and 0.70 for 11 emotions (incl. 310 ChatGPT 4-generated texts). The sentiments Acc. of the 11 emotions dataset was also the highest (0.79). As Facebook posts cannot be shared, we ran experiments on the Twitter and Telegram subset of the 11 emotions dataset, obtaining 0.73 Acc. for emotions and 0.80 for sentiments. The article describes the annotation procedures, guidelines, experiments, and results. We believe that this package will be of significant benefit to researchers working on emotion detection and sentiment analysis in Bulgarian.

Keywords: emotion detection, sentiment analysis, Bulgarian, social media

1. Introduction

Detecting sentiments, feelings, and emotions in texts is useful for many practical tasks using Natural Language Processing (NLP). One of them is the automatic detection of disinformation – when intended as deception, as in European Commission's definition[1]). This is due to the fact, that several studies (Zuckerman et al., 1981; Larcker and Zakolyukina, 2012; Newman et al., 2003), consider unusual patterns of emotions as a sign which betrays deception. Detecting disinformation is especially important in social media, as many more people have access to it and can freely publish their texts. In psychology (Nandwani and Verma, 2021; Bakker et al., 2014; Ekman, 1992; Plutchik, 1982; Shaver et al., 1987; Lövheim, 2012), there are different views regarding the nature of emotions and their number (varying from 6 (Ekman, 1992; Shaver et al., 1987) to over 30 (Plutchik, 1982). According to some researchers (e.g. Plutchik, R. (1982)), emotions can have different intensities (e.g. "fear" is stronger than "apprehension"). Other researchers distinguish between emotions and feelings (Hansen, 2005; Damasio, 2004). However, there are no Bulgarian-language emotion detection tools, machine learning models, nor publicly available corpora and datasets annotated with emotions. The only available solution are works on sentiment analysis for Bulgarian. Such exist (Kraychev and Koychev, 2012; Petrova, 2021; Petrova and Bozhikova, 2022; Lazarova and Koychev, 2015; Kapukaranov and Nakov, 2015), including for social media texts (Smailović et al., 2015), but they do not allow detecting specific feelings or emotions (e.g. fear, distrust, and happiness). Previous work (Bianchi et al., 2022) proposed an easier alternative to manually annotating emotions in new languages, which consisted of translating existing emotions datasets using Machine Translation (MT). Studies of how well sentiments and emotions are preserved in datasets, translated by humans and machines, showed promising results (Öhman et al., 2016; Salameh et al., 2015; Kajava et al., 2020). However, using MT could cause losing the language- and culture-specific characteristics of the expressed emotions, as commented previously by other researchers (De Bruyne, 2023). For

[1] https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=COM%3A2020%3A790%3AFIN&qid=1607079662423

LREC-COLING 2024, pages 14954–14966
20-25 May, 2024.

this reason, we decided to create a dataset, with texts originally written in Bulgarian language and annotated by native speakers for the emotion(s) or feeling(s) expressed in the texts. Such a resource and the classifiers that are trained on it would allow a more precise emotion detection to recognize disinformation. Such resource would also reflect the specific social media language. It is already well-known that the language used in Internet communications, and especially in social media is different from the standard one. It may contain orthographic errors, words may have different meanings, and follow different syntactic rules. Abbreviations, mentions, and hashtags may also appear. This is also valid for Bulgarian. Additionally, in Bulgarian social media posts, the Cyrillic alphabet is mainly used, but users could transliterate the Bulgarian language into the Latin alphabet, and sometimes posts may contain code-switching between Bulgarian and English. For all the above reasons, in this article, we introduce a package, containing several versions of a new social media text dataset in Bulgarian language manually annotated for 21 categories of a mixture of emotions and feelings. The package includes the classifiers with the highest accuracies, trained on these datasets. To produce also new sentiment analysis resources, we automatically merged the emotions/feelings annotations into the three traditional sentiment categories (positive, negative, and neutral), and experimented with classifiers trained on them. The original dataset consists of a total of 6000 social media posts, collected from three platforms (Facebook, Telegram, and Twitter), and includes our publicly accessible dataset (Temnikova et al., 2023), enriched with additionally collected by us Facebook posts, and tweets. As Facebook posts cannot be shared, to allow future comparisons of our results with those of other researchers, we have created a version of the dataset with only Telegram and Twitter posts (3750 posts in total) with parallel emotion and sentiment labels. Additionally, we are releasing two emotions/feelings and sentiments classifiers, which reached the highest accuracies, and the versions of the datasets (with no Facebook posts) on which the respective classifiers were trained. The article presents the methods for collecting, cleaning, and pre-processing the data, the procedures followed for manual annotation, and the machine learning experiments. We believe that the findings and contributions of this article will be useful not only to researchers of Bulgarian but also to those working on emotion detection for other lower-resourced languages. In the rest of the article: Section 2 discusses related work, Section 3 provides details about the datasets, Section 4 presents the annotation procedures and results, Section 5 describes the machine learning experiments and their results, Section 6 provides a general discussion, Section 7 presents the conclusions, Section 8 discusses the ethical and legal aspects, Section 9 presents this work's limitations, and finally, Section 10 lists the acknowledgements.

2. Related Work

For the sake of simplicity, from now on, we will be denoting as "emotions" both emotions and feelings.

As in other NLP areas, in automatic emotion detection from texts, there is more work for English than for other languages. De Bruyne, L. (2023) observed that papers not exclusively centered on English constituted about one-third of all papers presented at the Workshop on Computational Approaches to Subjectivity, Sentiment & Social Media Analysis (WASSA) between 2011 and 2022.

We consider the works most similar to ours those that present:

- emotion detection datasets and classifiers for languages other than English (with more attention on languages close to Bulgarian - e.g. Slavic languages)
- sentiment analysis datasets and classifiers for Bulgarian

with a **special focus on those of them that contain (or are trained on) social media texts**.

Social media datasets, annotated with emotions and emotion detection classifiers trained on them have been created for several languages other than English. The closest languages to Bulgarian are: Russian (Sboev et al., 2021), Polish (Bogdanowicz et al., 2023), and Romanian (Ciobotaru and Dinu, 2021). Other languages include: Italian (Sprugnoli et al., 2020; Bianchi et al., 2021), Spanish (Mohammad et al., 2018), Indonesian (Savigny and Purwarianti, 2017; Saputri et al., 2018), Bangla (Tripto and Ali, 2018), Thai (Sarakit et al., 2015), Vietnamese, Filippino (Lapitan et al., 2016a), Tamil (Vasantharajan et al., 2022), Korean (Do and Choi, 2015), and Arabic (Mohammad et al., 2018). Similarly to our dataset, most of them contain between 3000 and 10000 texts (usually YouTube comments, but also tweets, and texts from other country-specific social media platforms). Such datasets also most frequently cover fewer categories of emotions than ours - usually Ekman's (1992) 6 basic ones (Anger, Disgust, Fear, Happiness, Sadness and Surprise) or Plutchik's 8. However, some researchers worked with 28 or 31 emotions (Vasantharajan et al., 2022; Liew et al., 2016). Due to the variety in the number of categories, IAA metrics, and classifier models used, it is hard to make a fair comparison. However, IAA results are usually low for this task, with very few exceptions. F1-Scores most frequently are around 0.50 or even lower, especially when there are around 30 categories.

While there is no work on emotion detection for Bulgarian, there are several articles from Bulgarian authors on Bulgarian sentiment analysis (Petrova, 2021; Petrova and Bozhikova, 2022; Kapukaranov and Nakov, 2015; Kraychev and Koychev, 2012; Kraychev, 2014; Raychev, 2009; Raychev and Nakov, 2009; Lazarova and Koychev, 2015), including a publicly accessible online tool[2]. Differently from our work, none of them were based on social media texts. Only Smailović et al. (Smailović et al., 2015) worked on sentiment analysis for social media texts in Bulgarian. However, while they collected a dataset of tweets from 2013 Bulgarian elections, their dataset is not publicly accessible.

3. Datasets

In this article, we introduce a new package, containing several new social media datasets with texts in Bulgarian language, written only in Cyrillic. Official Bulgarian language is written in Cyrillic letters, but younger Bulgarian generations tend to use at least 2 Internet variants of Latin transliterations of Bulgarian. The first and largest dataset contains 6000 texts and it is composed of two separately annotated datasets, collected from social media platforms. From now on we refer to them as *Emo-SM-BG2022* and *Emo-SM-BG2023*.

We assembled and annotated Emo-SM-BG2022 in 2022 (as its name suggests), with the specific aim to study emotions with respect to disinformation detection in social media posts, written in Bulgarian. This dataset contains in total 5000 social media posts from 3 platforms (Telegram, Twitter, and Facebook). Due to its specific purpose, the texts in this dataset are on topics usually related to misinformation and disinformation. Specifically, the Telegram and Twitter posts (in total 2750 texts) are a subset of our previous publicly available Twitter and Telegram datasets (Temnikova et al., 2023) on the topics of deception, manipulation, and Covid-19. We have expanded these 2750 texts, by adding 2250 Facebook posts of Bulgarian politicians and political influencers. The Facebook texts were collected by using Crowdtangle web dashboard (to which we were granted academic access). To ensure consistency with the Twitter and Telegram texts, we applied the same pre-processing (see details in (Temnikova et al., 2023). Specifically, we removed duplicates and also non-Bulgarian texts (for which we used FastText[3] and an additional round of manual review). After annotating Emo-SM-BG2022, we noticed that it contained mostly messages, annotated with negative emotions (probably due to the disinformation topics), which was making our dataset severely imbalanced. To alleviate this issue, we created Emo-SM-BG2023, which was collected from Twitter only, using a list of manually collected positive keywords. The positive keywords included expressions (with endings covering the different numbers, genders, and persons), like: горд (proud), възхищавам се (I admire), чудесен (wonderful), всичко ще е наред (everything will be fine), радвам се (I am glad/happy), обичам (I love), благодаря (thank you), and wishes, for example: добро утро (good morning), честит рожден ден (happy birthday), etc. We have additionally manually reviewed the tweets containing these keywords to ensure that they express truly positive emotions. During this process, we removed those tweets which used such keywords in a figurative way (for example ironically or sarcastically, expressing instead a negative emotion). The final version of Emo-SM-BG2023, which was annotated, contained 1000 tweets.

Table 1 shows the details of Emo-SM-BG2022 and Emo-SM-BG2023. As mentioned previously, they contained in total 6000 social media posts, out of which, we are releasing only 3750 posts (marked in the second column). Following CrowdTangle's academic access rules, we will not publish the Facebook texts from the Emo-SM-BG2022 dataset. From this first 6000-texts dataset, we produce smaller subsets, with texts, annotated from subsets of the fine-grained emotion categories for our classifier experiments. Additionally, we create parallel versions of Twitter and Telegram-only datasets to obtain reproducible results (see details in Section 5).

Emo-SM-BG2022		
Platform	Publ. Released?	Num. of Texts
Facebook	No	2250
Twitter	Yes (tweet IDs)	2250
Telegram	Yes (anon. texts)	500
Emo-SM-BG2023		
Platform	Publ. Released?	Num. of Texts
Twitter	Yes (tweet IDs)	1000

Table 1: Emo-SM-BG2022 and Emo-SM-BG2023 datasets details.

4. Annotation Procedures

We have followed slightly different procedures for annotating Emo-SM-BG2022 and Emo-SM-BG2023. For both datasets, we used the same original set of 21 categories of emotions/feelings,

[2] https://azbuki-ml.com/sentiment.

[3] https://fasttext.cc/docs/en/language-identification.html.

and the same annotation guidelines[4]. We updated the guidelines several times during the annotation of both datasets (based on annotators' feedback), which resulted in a slightly different version of the guidelines for Emo-SM-BG2023. For both datasets, we involved two non-overlapping groups of 5 annotators each. Both groups of annotators were native speakers of Bulgarian, who had knowledge of the Bulgarian reality, and experience using social media platforms. The only difference in this case was the number of annotations per social media post: each Emo-SM-BG2022 post was annotated by 3 annotators (with posts assigned randomly by the annotation tool), while each Emo-SM-BG2023 post was annotated by all 5 annotators. We also used the same web-based annotation tool for both datasets: GATE Teamware [5]. The annotation tool featured a collapsible window that, when expanded, would display a shorter version of the annotation guidelines, the social media post to be annotated, and radio buttons with emotion categories in Bulgarian. The interface also provided additional details, such as the number of posts that had been annotated, and how many were left to annotate. Figure 1 shows a screenshot of the annotation tool's interface, with annotators' instructions and emotion categories specifically translated into English for this article. As already discussed, some of the categories can be considered as *feelings* by some researchers (e.g. "Positive: Satisfaction/Approval", "Positive: Appreciation/Gratitude", "Positive: Sympathy/Compassion", and "Negative: Distrust", but they do also appear as emotions in the largest emotions classifications, e.g. in Plutchik, R. (1982)). We also added categories, which are rather *intentions*, such as "Positive: Offering Help/Support" and "Call for Action/Request/Call for Help", as we have frequently observed them in our data. During the annotation of both datasets, annotators were asked to assign either only one emotion or if they noticed that the post was expressing 2 or more emotions – to identify the Primary (principal, main) and the Secondary emotions[6]. The need to select in some cases two emotions was motivated by our previous observations, that 1) even if we made efforts to select the texts in a way to express just one emotion, this was not always possible, and 2) sometimes some of the annotators were noticing two or more emotions, while others would notice only one. It should be emphasized, however, that the selection of a Secondary emotion was not obligatory. Primary and Secondary emotions were chosen from the same set of 21 original emotions. One of them was "Other". When selected, annotators were asked to suggest a new more appropriate category in the "Comment" field. Due to the presence of this category, all social media posts were annotated. In both cases, the five annotators initially received the annotation guidelines to familiarize themselves. The subsequent stages varied slightly between Emo-SM-BG2022 and Emo-SM-BG2023. **Specifically, for Emo-SM-BG2022**, the annotators were trained by annotating 3 smaller batches of posts, each followed by a manual review of the most experienced annotator[7] and discussions. The 3 smaller batches (with 110, 115, and 117 texts), were manually selected in a way to contain the same number of social media posts for each of the 20 categories. After all annotators completed the annotations of each batch, the most experienced annotator reviewed the annotations, and discussed by voice with the whole group of annotators the posts for which: 1) annotations greatly differed; 2) annotation guidelines weren't adhered to; and 3) one or more annotators found category selection challenging. This process helped to remove inappropriate categories and merge some of them. As annotators were asked to suggest new categories, such reviews and discussions allowed adding new categories, especially if they were frequently proposed or suggested by more than one annotator. After completing the initial 3 batches, the annotators were randomly assigned the remaining 4658 posts. Each annotator received approximately 2794 social media posts. Inter-Annotator Agreement (IAA) was calculated only after all 5000 posts had been annotated (see details in Section 4.2).

As Emo-SM-BG2022's IAA had quite low values (see Table 2), we followed a slightly different procedure for annotating Emo-SM-BG2023, in an attempt to raise the IAA results. Emo-SM-BG2023 annotators were assigned 4 batches of a lower number of posts (100, 100, 200, 200), with the difference that each annotator had to annotate all posts within the same batch. After the completion of the annotation of each batch, we calculated the percentage IAA for the whole batch, and also the number of annotators, who agreed on the same Primary emotion for each annotated post. Then, an external reviewer[8] (who did not participate in the Emo-SM-BG2023 annotation) commented with the whole group only the

[4]The English version of the annotation guidelines can be accessed at: https://tinyurl.com/emot-bg.

[5]The annotation tool can be accessed at this link: https://gate.ac.uk/teamware/.

[6]Not to be confused with the "primary and secondary emotions" as in Shaver et al. (1987).

[7]During the annotation of Emo-SM-BG2022, the most experienced annotator who reviewed all the annotations was one of the authors of the annotation guidelines and also one of the five annotators.

[8]The external reviewer of the Emo-SM-BG2023 annotations did not participate in the actual Emo-SM-BG2023 annotation and was the same most experienced annotator from Emo-SM-BG2022 and one of the authors of the annotation guidelines.

posts in which fewer than 3 annotators[9] agreed on the same Primary emotion. We focused exclusively on disagreements regarding the Primary emotion for two reasons: 1) the annotation of a Secondary emotion was not obligatory, and 2) very few posts had more than 1 emotion. After each discussion, the annotators were assigned the next batch. Such manual reviews and discussions were done only for these first 4 batches. After the 4th batch, the annotators received the final 5th batch of 400 posts, which was not subject to a manual review. This procedure led to an increase in the percentage of cases where more than three annotators agreed on the same Primary emotion, moving from approximately 73% for batches 1 to 3, to 77.5% for batch 4, and then to 81.75% for batch 5. Overall, this meant that for 77.3% of the 1000 annotated posts, between three to five annotators agreed on the same Primary emotion.

We selected the 20 categories of emotions starting from the 6 basic ones of Ekman, P. (Ekman, 1992). We expanded this list by adding categories that two of the co-authors observed in the actual Emo-SM-BG2022 social media posts (including some from Plutchik, R. (1982)). Following a manual review of the posts in the initial stages of the Emo-SM-BG2022 annotation, we noticed that the dataset contained very few posts for some of the categories, so we merged them with other categories. This happened for example with "disgust", "anger", "outrage", and "hate", which were merged into one category (named "Negative: Anger/Outrage/Disgust/Hate").

4.1. Annotation Guidelines

As mentioned before, the annotation guidelines were gradually refined during the annotations of both Emo-SM-BG2022 and Emo-SM-BG2023, in accordance with the annotators' feedback. First, reflecting what is already known in the emotion detection field, the guidelines contain explanations on how to determine the source of the emotion in a post. Annotators are instructed to prioritize in the following order: 1) the emotion of the author of the post, 2) the emotion of somebody mentioned in the post 3) the known emotion with which the post would be perceived by a specific known audience 4) the personal interpretation of the annotator. Next, the guidelines list each of the 20 categories of concrete emotions. Each category is explained with a definition, examples in Bulgarian, and keywords that may signal the presence of such emotion. When keywords are given there is also a warning that the keywords may be used figuratively and the annotators should be carefully reading each post.

[9]We considered as acceptable agreement when 3, 4, or 5 annotators indicated the same Primary emotion.

A specific case is the category "Surprise", for which the guidelines specify that the annotators should write in the "Comment" field whether the surprise was positive or negative.

4.2. Annotation Results

As explained previously, annotation was done separately for Emo-SM-BG2022 and Emo-SM-BG2023. Each social media post in Emo-SM-BG2022 was annotated by 3 annotators, while each social media post in Emo-SM-BG2023 – by 5 annotators.

4.2.1. Qualitative Observations

Previous research has shown that annotating emotions is a highly subjective task and the more categories there are, the harder it becomes for annotators to agree (Öhman, 2020). We observed a similar trend in both Emo-SM-BG2022 and Emo-SM-BG2023 annotations. Sometimes 5 annotators would indicate 5 different Primary emotions, with each choice seeming well justified. Based on annotators' feedback, their selection of emotions was influenced by their personal experiences and their interpretation of social media posts' meaning. Another recurring pattern was that even if annotators identified the same two emotions, they frequently ordered them differently. In other words, they disagreed on which emotion was Primary and which was Secondary. We also noticed that an annotator's choice of emotions fluctuated from day to day. We hypothesize that these variations could be influenced by the annotators' moods. Finally, we noticed that the ability to recognize a social media post's emotion greatly depended on being familiar with the cultural and generation specifics of Bulgarian society. We could not run any comparative analysis on sociolinguistic dimensions of the social media posts, as, we did not store any information about their authors, to preserve their identities, as required by European laws. However, we observed interesting patterns distinguishing male from female annotators, which we will publish in future work.

4.2.2. Inter-Annotator Agreement

We calculated the Inter-Annotator Agreement (IAA) for both the primary and secondary emotions. The metrics used for calculating IAA were Fleiss' Kappa and Simple Percentage Agreement. Due to the substantial number of emotion categories (21 in total), the application of Krippendorff's Alpha was deemed unsuitable as it would introduce undue complexity into the calculations. Table 2 shows the IAA results of Primary Emotion and Secondary Emotion in Emo-SM-BG2022, Emo-SM-BG2023, and the separate subsets of Emo-SM-BG2022 from the three different social media platforms. Emo-SM-BG2022

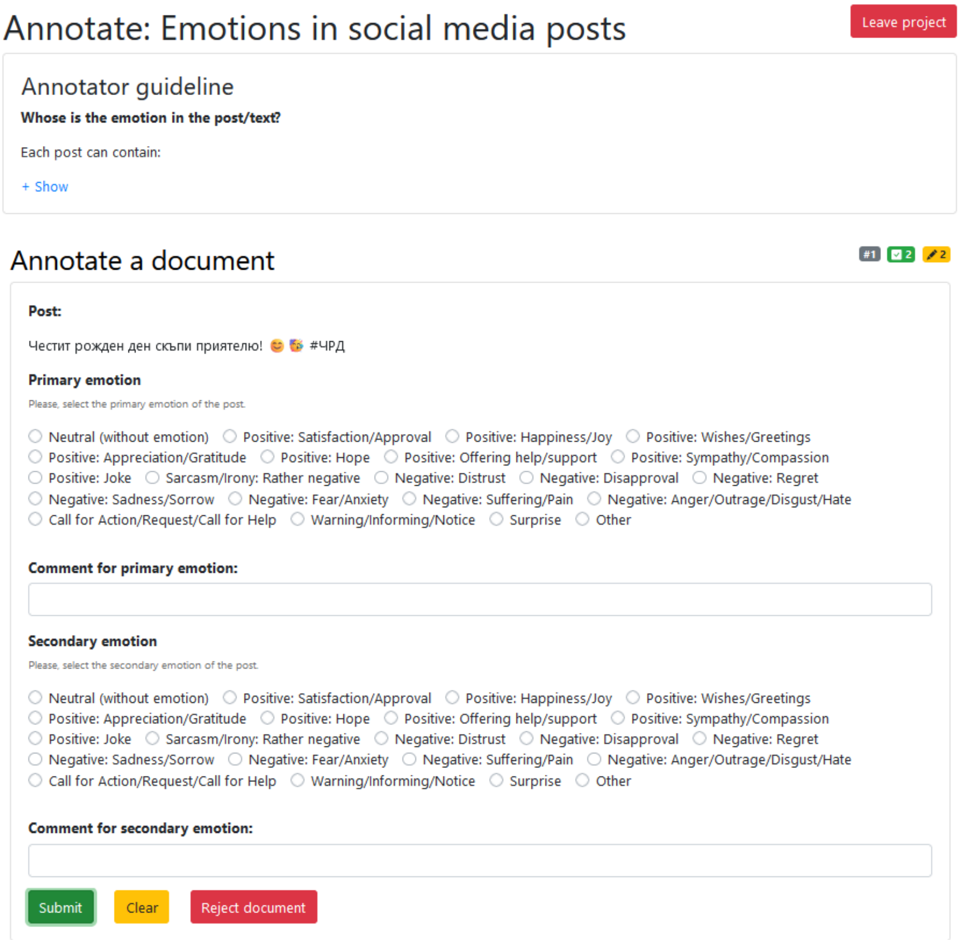

Figure 1: Screenshot of the annotation tool. The displayed social media post is a mock-up example.

Dataset				Fleiss' Kappa			Simpl.Per.Agr.	
		Annotators	Posts	Kappa	z	p-value	Posts	%-agree
Emo-SM-BG2022	Primary Emot.	3	5000	**0.317**	122.768	0	5000	**20.5**
	Second. Emot.	3	1579	-0.136	-19.975	0	25	12.0
Facebook	Primary Emot.	3	2250	0.348	89.667	0	2250	24.04
	Second. Emot.	3	726	-0.142	-13.515	0	10	10.0
Telegram	Primary Emot.	3	500	0.272	32.737	0	500	17.8
	Second. Emot.	3	262	-0.112	-6.96	0	8	0.0
Twitter	Primary Emot.	3	2250	0.276	69.5	0	2250	17.56
	Second. Emot.	3	591	-0.153	-13.498	0	7	28.57
Emo-SM-BG2023	Primary Emot.	5	1000	**0.469**	139.981	0	985	**25.482**
	Second. Emot.	5	85	0.306	24.489	0	85	10.588
	Both Emotions	5	1000	0.306	136.844	0	985	11.878

Table 2: IAA results for both Emo-SM-BG2022 and Emo-SM-BG2023.

"Both Emotions" was derived by considering both the Primary Emotion and the Secondary Emotion as a sorted list. By doing this, instances where two annotators indicated the same two emotions but in reverse order were counted as one. This approach was influenced by the external reviewer's observations that annotators frequently were hesitating whether to consider an emotion as Primary or Secondary.

Example 1 *Annotator 1 – Primary Emotion: "Positive: Happiness/Joy"; Secondary Emotion: "Positive: Hope"*

Annotator 2 – Primary Emotion: "Positive: Hope"; Secondary Emotion: "Positive: Happiness/Joy"

Table 2 shows that the IAA results were low for both Emo-SM-BG2022 and Emo-SM-BG2023 (with a slight increase for Emo-SM-BG2023, due to the

Exper.	Category	Acc.	Prec.	Rec.	F1	Test
Emotions-Exper. 2 **+FB** (16 labels)	**All categories**	**0.61**	**0.62**	**0.61**	**0.61**	**454**
	Negative: Fear/Anxiety		0.5	0.69	0.58	13
	Call for Action/Request/Call for Help		0.64	0.61	0.63	44
	Negative: Disapproval		0.3	0.39	0.34	28
	Sarcasm/Irony: Rather Negative		0.49	0.57	0.53	72
	Positive: Satisfaction/Approval		0.62	0.61	0.62	46
	Neutral (without Emotion)		0.75	0.71	0.73	90
	Negative: Distrust		0.33	0.17	0.22	12
	Negative: Sadness/Sorrow/Regret, Suffering/Pain		0.8	0.36	0.5	11
	Negative: Anger/Outrage/Disgust/Hate		0.66	0.57	0.61	65
	Surprise		0.5	0.5	0.5	2
	Positive: Joke		0.38	0.38	0.38	8
	Positive: Happiness/Joy		0.56	0.62	0.59	8
	Positive: Wishes/Greetings		0.91	0.91	0.91	32
	Positive: Hope		0.58	0.88	0.7	8
	Positive: Appreciation/Gratitude		0.73	0.67	0.7	12
	Positive: Offering Help/Support, Sympathy/Compassion		0	0	0	3
Emot.-Exper. 3 **+FB** (11 l.)	**All categories**	**0.67**	**0.68**	**0.67**	**0.67**	**366**
	Negative: Fear/Anxiety		0.57	0.63	0.6	**19**
	Call for Action/Request/Call for Help		0.71	0.71	0.71	52
	Negative: Anger/Outrage/Disgust/Hate/Disapproval		0.68	0.77	0.72	77
	Positive: Satisfaction/Approval/Happiness/Joy		0.57	0.62	0.6	50
	Neutral (without Emotion)		0.62	0.68	0.65	74
	Negative: Distrust		0.56	0.38	0.45	**13**
	Negative: Sadness/Sorrow/Regret, Suffering/Pain		0.5	0.31	0.38	**13**
	Positive: Wishes/Greetings		1	0.94	0.97	36
	Positive: Hope		1	0.36	0.53	**11**
	Positive: Appreciation/Gratitude		0.73	0.57	0.64	**14**
	Positive: Offering Help/Support, Sympathy/Compassion		0.6	0.43	0.5	**7**
Emot.-Exper. 4 **+FB** (11 l.)	**All categories**	**0.70**	**0.72**	**0.70**	**0.70**	**408**
	Negative: Fear/Anxiety		0.58	0.58	0.58	19
	Call for Action/Request/Call for Help		0.63	0.72	0.67	47
	Negative: Anger/Outrage/Disgust/Hate/Disapproval		0.7	0.73	0.72	83
	Positive: Satisfaction/Approval/Happiness/Joy		0.83	0.56	0.67	63
	Neutral (without Emotion)		0.72	0.69	0.71	72
	Negative: Distrust		0.42	0.47	0.44	17
	Negative: Sadness/Sorrow/Regret, Suffering/Pain		0.56	0.74	0.64	19
	Positive: Wishes/Greetings		0.9	0.9	0.9	30
	Positive: Hope		0.68	0.89	0.77	19
	Positive: Appreciation/Gratitude		0.83	0.86	0.84	22
	Positive: Offering Help/Support, Sympathy/Compassion		0.73	0.65	0.69	17
Emot.-Exper. 4 **-FB** (11 l.)	**All categories**	**0.73**	**0.75**	**0.73**	**0.73**	**273**
	Negative: Fear/Anxiety		0.58	0.47	0.52	15
	Call for Action/Request/Call for Help		0.60	0.50	0.55	18
	Negative: Anger/Outrage/Disgust/Hate/Disapproval		0.81	0.71	0.76	55
	Positive: Satisfaction/Approval/Happiness/Joy		0.63	0.84	0.72	49
	Neutral (without Emotion)		0.64	0.83	0.72	42
	Negative: Distrust		0.67	0.57	0.62	14
	Negative: Sadness/Sorrow/Regret, Suffering/Pain		0.91	0.71	0.80	14
	Positive: Wishes/Greetings		0.95	0.78	0.86	23
	Positive: Hope		0.82	0.60	0.69	15
	Positive: Appreciation/Gratitude		0.93	0.93	0.93	15
	Positive: Offering Help/Support, Sympathy/Compassion		1.00	0.77	0.87	13

Table 3: Fine-tuned BERT-WEB-BG emotions experiments results. We report the weighted average for Precision, Recall, and F1-score for all categories.

modified annotation procedure). The IAA of Emo-SM-BG2022's Primary emotion was 20.5% for Simple Percentage Agreement and 0.31 (indicating *fair agreement*) for Fleiss' Kappa. Emo-SM-BG2022's IAA was even lower for the Secondary emotion as it was optional and the annotated examples were significantly less. Emo-SM-BG2023's IAA for Primary emotion increased to 25.482% for Simple Percentage Agreement and to 0.469 (*moderate agreement*) for Fleiss' Kappa. We hypothesize that this higher value of Fleiss' Kappa was most probably due to our modified annotation procedure. The IAA for Emo-SM-BG2023's Secondary emotion was quite low. We do not observe higher results when a sorted list of both emotions is taken into account.

5. Classifiers

We run several experiments with a fine-tuned version of the recently published BERT-WEB-BG[10] (Marinova et al., 2023), trained on subsets of the new annotated dataset. We experimented both with the emotion categories and with the three coarse-grained sentiment categories into which the emotion categories were automatically merged. To produce the three sentiment categories, we merged all emotion categories with the word "Negative" into the *Negative* sentiment; the emotion categories with the word "Positive" into the *Positive* sentiment; and the *Neutral* sentiment was produced by merging the following three emotion categories: *Neutral (without emotion)*, *Warning/Informing/Notice*, and *Call for Action/Request/Call for Help*. In all experiments we split the data into train, validation, and test in this way: 80:10:10. When publishing the datasets, we are sharing the actual splits used for the respective classifiers.

5.1. Emotion Detection Classifiers

Our experiment settings concerned the method for selecting a unique label for each annotated social media post. For each annotator, we considered both the Primary and Secondary emotions (if such were provided). In this way, the maximum number of labels per post was 6 for Emo-SM-BG2022 (2 emotions per annotator, 3 annotators per post) and 10 for Emo-SM-BG2023 (2 emotions per annotator, 5 annotators per post). As not all annotators indicated a Secondary emotion, the empty values were removed. Selecting just posts with full agreement on the Primary emotion resulted in only 250 posts left.

In **Emotions-Experiment 1**, we used the original dataset, including Facebook, and all 21 categories, and selected as post's label the most frequently repeated category with a Python max function. However, this approach was inapplicable when two or more categories shared the same highest count. With these settings, the fine-tuned BERT-WEB-BG's accuracy was below 0.50. Due to these shortcomings, we do not report this experiment's results in Table 3.

To raise the accuracy, we run **Emotions-Experiment 2**, in which we: 1) merged the categories with the lowest number of occurrences into coarser categories; 2) Merged *Warning/Informing/Notice* into *Neutral (without emotion)*, 3) modified the post's label selection to consider only the posts with only one most frequently assigned category, 4) created a list of ignored posts, including those which didn't have only 1 most frequently assigned category and posts with category *Other*. This approach reduced the 6000 original posts to 4536, divided into 16 categories. The Accuracy slightly raised to 0.61 (as visible in Table 3). In **Emotions-Experiment 3** we further raised the Accuracy to 0.67 with the following modifications: 1) merged *Negative: Disapproval* with *Negative: Anger/Outrage/Disgust/Hate* into a coarser category, as the former is a lighter version of the latter. 2) did the same with *Positive: Satisfaction/Approval* and *Positive: Happiness/Joy*; 3) added to the list of ignored posts those, annotated with the categories *Surprise*, *Positive: Joke*, and *Sarcasm/Irony: Rather negative*, as such categories could be both positive and negative. We ended up with a total of 3663 posts, labeled with 11 categories.

Emotions-Experiment 4 including Facebook texts reached the highest Accuracy of 0.70. In it, we took the same label selection algorithm and the 11 categories from Emotions-Experiment 3, and added ChatGPT 4-generated examples to the categories with fewer than 20 posts in the test split (indicated in Table 3 for Experiment 3 in bold). This was done by asking ChatGPT 4 to generate similar examples, by providing a list of manually selected unambiguous posts clearly carrying only the specified category. This resulted in a total of 4080 posts. We also experimented on a subset of this data, from which we removed the Facebook posts, resulting in 2674 texts. Surprisingly, the Acc. raised to 0.73, with results rising also for specific categories (e.g. "Negative: Anger/Outrage/Disgust/Hate/Disapproval", "Negative: Sadness/Sorrow/Regret, Suffering/Pain", and "Positive: Offering Help/Support, Sympathy/Compassion" – see Table 3).

5.2. Sentiment Analysis Classifiers

We run **Sentiments-Experiment 1** by merging the 11 emotion categories obtained in Emotions-Experiment 3. We obtained 0.77 accuracy, as visible in Table 4. We additionally run **Sentiments-Experiment 2** in which we merged the human and ChatGPT-generated posts from Emotions-Experiment 4 into 3 sentiments. By doing this, the accuracy was further raised to 0.79. As this was the highest Acc. including Facebook texts, we ex-

[10]https://huggingface.co/usmiva/bert-web-bg.

Experiment	Category	Acc.	Prec.	Rec.	F1
Sentiments-Experiment 1 **+FB**	**all categories**	**0.77**	**0.77**	**0.77**	**0.77**
	negative		0.76	0.80	0.78
	neutral		0.72	0.72	0.72
	positive		0.82	0.79	0.81
Sentiments-Experiment 2 **+FB**	**all categories**	**0.79**	**0.79**	**0.79**	**0.79**
	negative		0.72	0.83	0.77
	neutral		0.80	0.71	0.76
	positive		0.85	0.80	0.83
Sentiments-Experiment 2 **-FB**	**all categories**	**0.80**	**0.80**	**0.80**	**0.80**
	negative		0.76	0.81	0.78
	neutral		0.68	0.68	0.68
	positive		0.90	0.85	0.88

Table 4: Fine-tuned BERT-WEB-BG results for sentiments. We report the weighted average for Precision, Recall, and F1-score for all categories.

perimented also after removing them. This gave us a further rise to 0.80 Acc. (these results are added to Table 4).

6. Discussion

Table 2 showed that the modified annotation procedure in Emo-SM-BG2022 raised Fleiss' Kappa from 0.317 (fair agreement) to 0.469 (moderate agreement). We hypothesize that the IAA is still low because of the high number of emotion categories, the possible partial overlap of some of them, and due to some posts containing 2 or more emotions. An example is: "Благодаря Ви за подкрепата!" ("Thank you for your support!"), which could be categorized as both *Positive: Appreciation/Gratitude* and *Positive: Offering Help/Support, Sympathy/Compassion*. Table 3 shows that in all emotion detection experiments, *Positive: Wishes/Greetings* obtained very high results. This is due to the fact that such posts contained easily recognizable keywords, like "Добро утро!" ("Good morning!"), "Честит имен ден!" ("Happy Name Day!"), and "Успех!" ("Good luck!"). A surprising discovery was also that removing Facebook texts increased the overall accuracies and the results for specific categories of Emotions-Experiment 4 and Sentiments-Experiment 2. The error analysis of the test data of both the emotions and the sentiment experiments with no Facebook texts showed that the classifiers assigned a label different from the original one in the following cases: 1) when the text has more than 1 emotion 2) when the text was ironic - this usually resulted in assigning a label with the opposite polarity; 3) difficulty with the category "Neutral (without Emotion)", as it usually contains some emotion. Finally, we hypothesize that the Accuracy would further increase if more ChatGPT-4 examples are added (due to most of our ChatGPT 4-generated examples expressing quite unambiguously only one emotion).

7. Conclusions

This article presented SM-FEEL-BG – the first package of Bulgarian language social media emotion and sentiment text datasets[11], and classifiers[12] (fine-tuned versions of a Bulgarian BERT) trained on them. A modified annotation procedure increased the IAA results. The obtained IAA and the BERT results are comparable with previous research on smaller monolingual datasets, especially with a large number of emotions (Sprugnoli et al., 2020; Liew et al., 2016). The package will be of significant benefit to researchers of Bulgarian emotions, for multilingual emotion detection (Bianchi et al., 2022), and also for researchers, working on other lower-resourced languages.

8. Ethical and Legal Considerations

The annotators, who participated in this research, are all part of the same research team, and annotation was part of their daily obligations. Sharing their exact salaries would be a breach of the rules of GATE Institute, but the average rate was 8 euros per hour. Additionally, the younger annotators have received NLP training and were added as co-authors to this article. As described in Section 3, part of the texts were a subset of already published datasets Temnikova et al. (2023a,c,b) with a Creative Commons Attribution-NonCommercial-ShareAlike 4.0 International license and additional restrictions, which we copy in Zenodo's page for

[11] SM-FEEL-BG datasets: https://zenodo.org/records/10870509; https://zenodo.org/records/10870526; https://zenodo.org/records/10870536

[12] SM-FEEL-BG models: https://zenodo.org/records/10870559; https://zenodo.org/records/10870569; https://zenodo.org/records/10870909; https://zenodo.org/records/10870958

the newly released datasets. The Telegram texts are already carefully anonymized, to prevent the reconstruction of users' identities. We follow Twitter rules and will share only the tweet ids of the 1000 positive tweets.

9. Discussion of Limitations

This work has the following limitations: 1) Emo-SM-BG2022 is focused on specific topics (lies, manipulation, and Covid-19). Ideally, it would be good to enrich the dataset with more topics. 2) The emotions categories have been created on the basis of the Emo-SM-BG2022 dataset. 3) According to annotators' observations, some frequently encountered categories are still missing (e.g. shame and envy). There are examples of posts, which are hard to assign to the existing categories. Finally, 4) SM-FEEL-BG contains only posts written in Cyrillic, while there are also Bulgarian social media posts, written with different Latin transliterations. 5) The Facebook posts cannot be publicly shared. To alleviate this issue, we provide datasets with no Facebook posts and classifiers, trained on them. Additionally, in future work, we plan to include other more accessible types of texts, such as the comments section of news articles, news articles, movies, and product reviews.

10. Acknowledgements

The work, presented in this article is partially supported by the GATE project, funded by Operational Programme Science and Education for Smart Growth under Grant Agreement No. BG05M2OP001-1.003-0002-C01.

Part of the work was supported by the project TRACES[13], and indirectly received funding from the European Union's Horizon 2020 Research and Innovation Action Programme, via the AI4Media Open Call #1 issued and executed under the AI4Media project (Grant Agreement no. 951911).

We are also especially thankful to the LREC-Coling 2024 reviewers for providing us with very helpful suggestions and to the LREC-Coling 2024 organisers and the area chairs for giving us the opportunity to publish this work. We are also grateful to Dr. Milena Dobreva for inspiring us to create several first contributions for Bulgarian language, and to Institute GATE and its Director, prof. Silvia Ilieva for giving us the opportunity to work on this interesting research.

[13]https://traces.gate-ai.eu/.

11. Bibliographical References

References

Iris Bakker, Theo Van Der Voordt, Peter Vink, and Jan De Boon. 2014. Pleasure, arousal, dominance: Mehrabian and russell revisited. *Current Psychology*, 33:405–421.

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.

Federico Bianchi, Debora Nozza, Dirk Hovy, et al. 2021. Feel-it: Emotion and sentiment classification for the italian language. In *Proceedings of the eleventh workshop on computational approaches to subjectivity, sentiment and social media analysis*. Association for Computational Linguistics.

Stanisław Bogdanowicz, Hanna Cwynar, Aleksandra Zwierzchowska, Cezary Klamra, Witold Kieraś, and Łukasz Kobyliński. 2023. Twitteremo: Annotating emotions and sentiment in polish twitter. In *International Conference on Computational Science*, pages 212–220. Springer.

Alexandra Ciobotaru and Liviu P Dinu. 2021. Red: A novel dataset for romanian emotion detection from tweets. In *Proceedings of the International Conference on Recent Advances in Natural Language Processing (RANLP 2021)*, pages 291–300.

Antonio R Damasio. 2004. Emotions and feelings. In *Feelings and emotions: The Amsterdam symposium*, volume 5, pages 49–57. Cambridge University Press Cambridge.

Luna De Bruyne. 2023. The paradox of multilingual emotion detection. In *Proceedings of the 13th Workshop on Computational Approaches to Subjectivity, Sentiment, & Social Media Analysis*, pages 458–466. Association for Computational Linguistics.

Petko Dimov and Elislav Ivanov. 2020. *Fake news*, pages 38–50. Military Academy "Georgi Stojkov Rakovski", Sofia.

Hyo Jin Do and Ho-Jin Choi. 2015. Korean twitter emotion classification using automatically built emotion lexicons and fine-grained features. In *Proceedings of the 29th Pacific Asia Conference on Language, Information and Computation: Posters*, pages 142–150.

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

Flemming Hansen. 2005. Distinguishing between feelings and emotions in understanding communication effects. *Journal of business research*, 58(10):1426–1436.

Sabit Hassan, Shaden Shaar, and Kareem Darwish. 2021. Cross-lingual emotion detection. *arXiv preprint arXiv:2106.06017*.

Vong Anh Ho, Duong Huynh-Cong Nguyen, Danh Hoang Nguyen, Linh Thi-Van Pham, Duc-Vu Nguyen, Kiet Van Nguyen, and Ngan Luu-Thuy Nguyen. 2020. Emotion recognition for vietnamese social media text. In *Computational Linguistics: 16th International Conference of the Pacific Association for Computational Linguistics, PACLING 2019, Hanoi, Vietnam, October 11–13, 2019, Revised Selected Papers 16*, pages 319–333. Springer.

Kaisla Kajava, Emily Öhman, Piao Hui, and Jörg Tiedemann. 2020. Emotion preservation in translation: Evaluating datasets for annotation projection. *Proceedings of Digital Humanities in Nordic Countries (DHN 2020)*.

Borislav Kapukaranov and Preslav Nakov. 2015. Fine-grained sentiment analysis for movie reviews in bulgarian. In *Proceedings of the international conference recent advances in natural language processing*, pages 266–274.

Boris Kraychev. 2014. *Extraction and Analysis of Opinions and Sentiments from Online Text*. Ph.D. thesis, Sofia University. In Bulgarian.

Boris Kraychev and Ivan Koychev. 2012. Computationally effective algorithm for information extraction and online review mining. In *Proceedings of the 2nd International Conference on Web Intelligence, Mining and Semantics*, pages 1–4.

Fermin Roberto Lapitan, Riza Theresa Batista-Navarro, and Eliezer Albacea. 2016a. Crowdsourcing-based annotation of emotions in Filipino and English tweets. In *Proceedings of the 6th Workshop on South and Southeast Asian Natural Language Processing (WSSANLP2016)*, pages 74–82, Osaka, Japan. The COLING 2016 Organizing Committee.

Fermin Roberto Lapitan, Riza Theresa Batista-Navarro, and Eliezer Albacea. 2016b. Crowdsourcing-based annotation of emotions in filipino and english tweets. In *Proceedings of the 6th Workshop on South and Southeast Asian Natural Language Processing (WSSANLP2016)*, pages 74–82.

David F Larcker and Anastasia A Zakolyukina. 2012. Detecting deceptive discussions in conference calls. *Journal of Accounting Research*, 50(2):495–540.

Gergana Lazarova and Ivan Koychev. 2015. Semi-supervised multi-view sentiment analysis. In *Computational Collective Intelligence: 7th International Conference, ICCCI 2015, Madrid, Spain, September 21-23, 2015, Proceedings, Part I*, pages 181–190. Springer.

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.

Hugo Lövheim. 2012. A new three-dimensional model for emotions and monoamine neurotransmitters. *Medical hypotheses*, 78(2):341–348.

Iva Marinova, Kiril Simov, and Petya Osenova. 2023. Transformer-based language models for bulgarian. In *Proceedings of the International RANLP Conference*.

Saif Mohammad, Felipe Bravo-Marquez, Mohammad Salameh, and Svetlana Kiritchenko. 2018. Semeval-2018 task 1: Affect in tweets. In *Proceedings of the 12th international workshop on semantic evaluation*, pages 1–17.

Pansy Nandwani and Rupali Verma. 2021. A review on sentiment analysis and emotion detection from text. *Social network analysis and mining*, 11(1):81.

Matthew L Newman, James W Pennebaker, Diane S Berry, and Jane M Richards. 2003. Lying words: Predicting deception from linguistic styles. *Personality and social psychology bulletin*, 29(5):665–675.

Emily Öhman. 2020. Emotion annotation: Rethinking emotion categorization. In *DHN post-proceedings*, pages 134–144.

Emily Öhman, Timo Honkela, and Jörg Tiedemann. 2016. The challenges of multi-dimensional sentiment analysis across languages. In *Proceedings of the workshop on computational modeling of people's opinions, personality, and emotions in social media (PEOPLES)*, pages 138–142.

Daniela Petrova. 2021. Automatic sentiment analysis on hotel reviews in bulgarian—basic approaches and results. In *Integrated Emerging Methods of Artificial Intelligence & Cloud Computing*, pages 48–56. Springer.

Daniela Petrova and Violeta Bozhikova. 2022. Sentiment and complexity analysis on two databases in bulgarian language–final estimation. In *2022 International Conference on Electrical, Computer, Communications and Mechatronics Engineering (ICECCME)*, pages 01–04. IEEE.

Robert Plutchik. 1982. A psychoevolutionary theory of emotions.

Veselin Raychev. 2009. *Automatic Recognition of Positive/Negative Subjectivity in Text*. Ph.D. thesis, Sofia University. In Bulgarian.

Veselin Raychev and Preslav Nakov. 2009. Language-independent sentiment analysis using subjectivity and positional information. *arXiv preprint arXiv:1911.12544*.

Victoria L Rubin. 2017. *Deception detection and rumor debunking for social media*. Sage London, UK.

Mohammad Salameh, Saif Mohammad, and Svetlana Kiritchenko. 2015. Sentiment after translation: A case-study on arabic social media posts. In *Proceedings of the 2015 conference of the North American chapter of the association for computational linguistics: Human language technologies*, pages 767–777.

Mei Silviana Saputri, Rahmad Mahendra, and Mirna Adriani. 2018. Emotion classification on indonesian twitter dataset. In *2018 International Conference on Asian Language Processing (IALP)*, pages 90–95. IEEE.

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.

Alexander Sboev, Aleksandr Naumov, and Roman Rybka. 2021. Data-driven model for emotion detection in russian texts. *Procedia Computer Science*, 190:637–642.

Phillip Shaver, Judith Schwartz, Donald Kirson, and Cary O'Connor. 1987. Emotion knowledge: further exploration of a prototype approach. *Journal of personality and social psychology*, 52(6):1061.

Paweł Skórzewski. 2019. Using book dialogs to extract emotions from texts in polish. *dimension*, 1:1.

Jasmina Smailović, Janez Kranjc, Miha Grčar, Martin Žnidaršič, and Igor Mozetič. 2015. Monitoring the Twitter sentiment during the Bulgarian elections. In *2015 IEEE international conference on data science and advanced analytics (DSAA)*, pages 1–10. IEEE.

Rachele Sprugnoli et al. 2020. Multiemotions-it: A new dataset for opinion polarity and emotion analysis for italian. In *Proceedings of the Seventh Italian Conference on Computational Linguistics (CLiC-it 2020)*, pages 402–408. Accademia University Press.

Irina Temnikova, Silvia Gargova, Ruslana Margova, Veneta Kireva, Ivo Dzhumerov, Tsvetelina Stefanova, and Hristiana Krasteva. 2023. New bulgarian resources for studying deception and detecting disinformation. In *Proceedings of the 10th Language & Technology Conference: Human Language Technologies as a Challenge for Computer Science and Linguistics*.

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.

Charangan Vasantharajan, Ruba Priyadharshini, Prasanna Kumar Kumarasen, Rahul Ponnusamy, Sathiyaraj Thangasamy, Sean Benhur, Thenmozhi Durairaj, Kanchana Sivanraju, Anbukkarasi Sampath, and Bharathi Raja Chakravarthi. 2022. Tamilemo: Fine-grained emotion detection dataset for tamil. In *International Conference on Speech and Language Technologies for Low-resource Languages*, pages 35–50. Springer.

Miron Zuckerman, Bella M DePaulo, and Robert Rosenthal. 1981. Verbal and nonverbal communication of deception. In *Advances in experimental social psychology*, volume 14, pages 1–59. Elsevier.

12. Language Resource References

Irina Temnikova and Silvia Gargova and Veneta Kireva and Tsvetelina Stefanova. 2023a. *TRACES Bulgarian Telegram Dataset Annotated with Linguistic Markers of Lies*. Zenodo.

Irina Temnikova and Silvia Gargova and Veneta Kireva and Tsvetelina Stefanova. 2023b. *TRACES Bulgarian Twitter Dataset on Covid-19 Annotated with Linguistic Markers of Lies*. Zenodo.

Irina Temnikova and Silvia Gargova and Veneta Kireva and Tsvetelina Stefanova. 2023c. *TRACES Bulgarian Twitter Dataset on Famous Bulgarian Political Cases of Suspected Lies, Annotated with Linguistic Markers of Lies*. Zenodo.