A Weakly Supervised Dataset of Fine-Grained Emotions in Portuguese

Diogo Cortiz, Jefferson O. Silva, Newton Calegari, Ana Luísa Freitas3, Ana Angélica Soares, Carolina Botelho, Gabriel Gaudencio Rego, Waldir Sampaio, Paulo Sergio Boggio

OBJECTIVES

- This research aims to study the creation of a corpus of fine-grained emotions for low re- sourced languages, specifically Portuguese.
- Due to limited financial resources, a specific objective of this work is to study the use of the weak supervision strategy to construct our corpus.
- Weak supervision is a strategy when there is no human annotation of each data point, but the labels are attributed using noisy and limited sources or specific rules

RESEARCH QUESTIONS

- RQ1: Is the weak supervision strategy suitable for building an NLP corpus for the fine- grained Emotion Recognition task in a low resourced environment?

- RQ2: What is a proper weak supervision approach to construct a corpus for fine-grained Emotion Recognition tasks in NLP?

HYPOTHESIS

- H1: weak supervision could be a suitable strategy to build NLP corpus for emotion recognition
- H2: lexical- based approach can be an adequate strategy to collect samples for each of the categories of our dataset, using the Lexical Items (LI) as a criterion for defining the label in an ade- quate way for Portuguese
- H3: masking LI presented in the weakly supervised corpus can avoid the model to overfit.

DEFINING EMOTION CATEGORIES

- The emotion categories for this research were defined from a review of the GoEmotion work [Demszky et al. 2020]otion categories for this research were defined from a review of the GoEmotion work [Demszky et al. 2020].

- The first stage, the researchers (group of 7) discussed and reviewed each emotion in English during a working meeting. They propose a translation to portuguese.

- The second stage was reviewing the categories' definitions in Portuguese to check if they were consistent with the language

EMOTION CATEGORIES

CATEGORY IN PORTUGUESE	CORRESPONDING IN GOEMOTIONS	DEFINITION IN PORTUGUESE
ADMIRAÇÃO	ADMIRATION	Achar algo impressionante ou digno de respeito.
DIVERSÃO	AMUSEMENT	Achar algo engraçado ou se divertir.
RAIVA	ANGER	Forte sentimento de desprazer ou antagonismo.
ABORRECIMENTO	ANNOYANCE	Raiva leve, irritação.
APROVAÇÃO	APPROVAL	Ter ou expressar uma opinião favorável.
CONFUSÃO	CONFUSION	Falta de compreensão, incerteza.
CURIOSIDADE	CURIOSITY	Forte desejo de saber ou aprender algo.
DESEJO	DESIRE	Forte sentimento de querer algo ou desejar que algo aconteça.
DECEPÇÃO	DISAPPOINTMENT	desprazer causado pelo não cumprimento de expectativas.
DESAPROVAÇÃO	DISAPPROVAL	Ter ou expressar opinião desfavorável.
NOJO	DISGUST	Repulsa despertada por algo desagradável ou ofensivo.
VERGONHA	EMBARRASSMENT	Vergonha ou constrangimento.
ENTUSIASMO	EXCITEMENT	Sensação de grande empolgação e ansiedade.
MEDO	FEAR	Estar com medo ou preocupado.
GRATIDÃO	GRATITUDE	Sentimento de gratidão e apreciação.
LUTO	GRIEF	Tristeza intensa, especialmente causada pela morte de alguém.
ALEGRIA	JOY	Sensação de prazer e felicidade.
AMOR	LOVE	Forte emoção positiva de consideração e carinho.
NERVOSISMO	NERVOUSNESS	Apreensão, preocupação, ansiedade.
OTIMISMO	OPTIMISM	Esperança sobre o futuro ou sobre o sucesso de algo.
ORGULHO	PRIDE	Prazer devido às próprias conquistas ou de alguém
ALÍVIO	RELIEF	Tranquilidade e relaxamento após ansiedade ou angústia.
REMORSO	REMORSE	Arrependimento ou sentimento de culpa.
TRISTEZA	SADNESS	Dor emocional, tristeza.
SURPRESA	SURPRISE	Sentir-se surpreso, assustado com algo inesperado.
INVEJA	-	Desgosto provocado pela felicidade ou prosperidade alheia
SAUDADE	-	Lembrança grata de pessoa ausente ou um momento passado.
COMPAIXÃO	-	Sentimento piedoso de simpatia e de ajuda

LEXICAL ITENS FOR WEAK SUPERVISION

- For each of the emotions on the list, we initially look for related Lexical Items by synonyms. Only synonyms with a semantic relationship with the definition of emotion were considered.

- For each Verbal Lexical Item, we collect the different conjugations to cover all tenses and moods in Portuguese.

- To avoid the negation effect, we manipulated the data as follows: we searched for the combination of the word "nã̃o" (no/not in Portuguese) or "nem" (neither in Portuguese) followed by a Lexical Item in our list. If an example was found, we removed it from our dataset

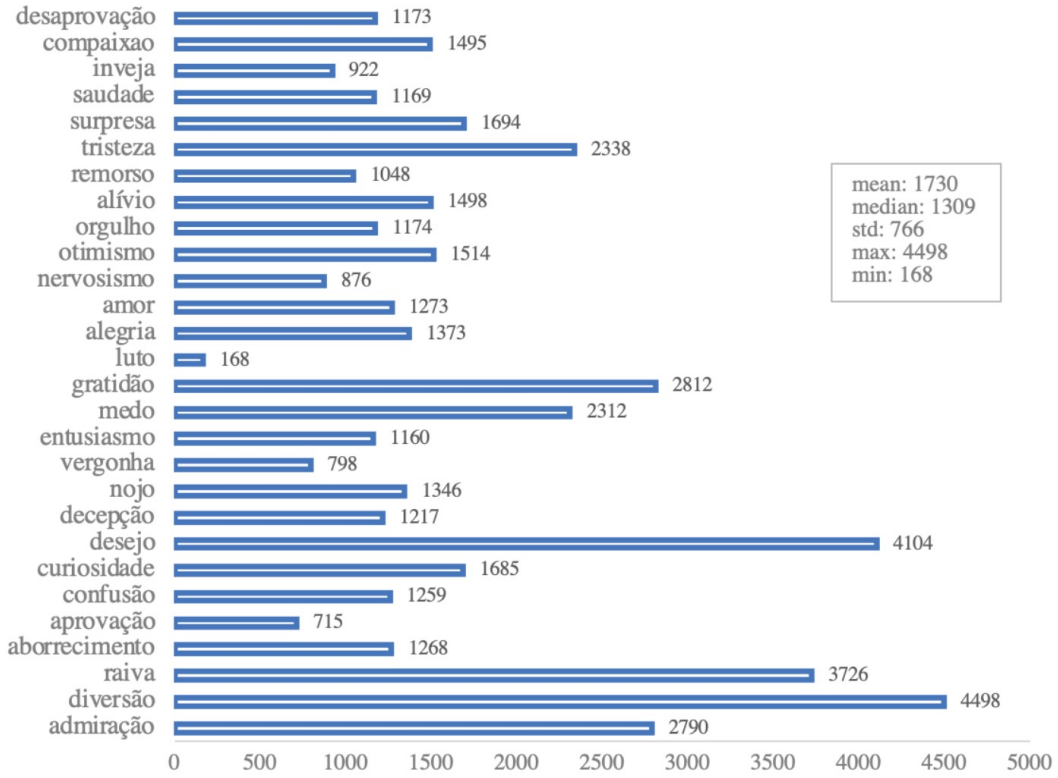
DATA

- We use Twitter as data source. The collection was made between the 23rd and 24th of June (2021) using the platform's official API.

- The filters used were the list of terms associated with each emotion. Retweets and replies were not considered, keeping only original tweets. Hashtags were removed, but emojis were kept.

- In total, 49179 tweets were collected using a weak supervision approach. Each example received the category label according to the Lexical Item used in the collection. For example, if a tweet was collected because it was filtered by a term associated with the emotion *amor*, it would be labeled to the *amor* category.

Figure 1. Examples per categories.



GOLD STANDARD DATA FOR VALIDATION

- we separated a set composed of 1773 examples from the dataset created earlier; we removed the labels assigned by the weak supervision approach so that a human could manually annotate them.
- Due to limited resources, it was not possible to cross-annotate the validation dataset.
- For this reason, it is not possible to present any measure of agreement between the annotators. We recognize the limitations of this procedure, which can reduce the quality of supervision and introduce bias.

MASKING LEXICAL ITENS IN TRAINNING DATASET

Original	tô indignada e não é pouco!
Masked	tô [MASK] e não é pouco!

- We masked Lexical Itens to investigate the impact in the model (overfiting, memorization). We ended up with three datasets for training three different models

- **NoMask Dataset**: without any masking techinque
- **30Mask Dataset**: applying the masking technique to 30% of examples for each category
- **FullMask Dataset:** masking all Lexical Items

MODELS

- To study the performance of our dataset, we fine-tuned a BERT language model (BERTimbau for portuguese).

- We use our three datasets to fine-tune three different models.

RESULTS

- **NoMask**: F1 = .64, Recall =.64, Precision = .70
- **30Mask:** F1 = .64, Recall =.63, Precision = .69
- **FullMask:** F1 = .22, Recall =.19, Precision = .35

CONCLUSIONS

- We found consistent results when evaluating our models, suggesting that weak supervision is an appropriate approach for initial work in the Emotion Recognition NLP task in Portuguese. The results supports our first hypotheses (H1).

- This research used a Lexical-based approach to collect, and weak supervise the dataset. The results help us to answer our RQ2 and validate the H2.

- We can neither validate nor refute our third hypothesis

THANK YOU

- Contact: dcortiz@pucsp.br / diogo@nic.br