Chronic pain patient narratives allow for the estimation of current pain intensity

Diogo A.P. Nunes* (diogo.p.nunes@inesc-id.pt; corresponding author), Joana Ferreira-Gomes† (jogomes@med.up.pt), Daniela Oliveira‡ (danielasoff@gmail.com), Carlos Vaz‡ (carlosjcvaz@gmail.com), Sofia Pimenta‡ (sofiadsp@sapo.pt), Fani Neto† (fanineto@med.up.pt), and David Martins de Matos* (david.matos@inesc-id.pt)

* Instituto de Engenharia de Sistemas e Computadores - Investigação e Desenvolvimento, Portugal
† Universidade do Porto, Portugal
‡ Centro Hospitalar Universitário de São João, Portugal

***Abstract*— We demonstrate a proof-of-concept for the analysis of the language of chronic pain for pain intensity estimation. Importantly, we show that focus on specific words/themes is especially correlated with specific pain intensity categories. We interviewed chronic pain patients and collected demographic and clinical data. 65 patients (40 females), averaging 56.4 ± 12.7 years of age, participated in the study. Patients reported their current pain intensity on a Visual Analogue Scale, which we discretized into 3 classes: mild, moderate, and severe pain. We extracted language features from the transcribed interview of each patient and used them to classify their pain intensity category. We measured performance with the weighted F_1 score. Finally, we analyzed potential confounding variables for internal validity. The best performing model was the Support Vector Machine with an Early Fusion of select language features, with an F_1 of 0.60, improving 39.5% upon the baseline. Patients with mild pain focused more on verbs, whilst moderate and severe pain patients focused on adverbs, and nouns and adjectives, respectively. We show that language features from patient narratives indeed convey information relevant for pain intensity estimation, and that our models can take advantage of that.**

Keywords—chronic musculoskeletal pain, pain intensity estimation, patient narratives, natural language processing

I. Introduction

Subjective reports of pain intensity, such as a Visual Analogue Scale (VAS), are susceptible to inconsistencies, impression or deception biases, socially suggestive cues, and others [1]. Thus, research has focused on automatic estimation of patient pain intensity. Most are based on facial expressions [1]–[5], but others have also used electroencephalograms [6], brain imaging [7], [8], and autonomic features [9]. To the best of our knowledge, there is no research work attempting to estimate current pain intensity from spontaneous natural language utterances, specifically chronic pain patient narratives. However, language is key for clinical chronic pain assessment and management [10], [11]. The narrative of the personal experience often includes valuable information about the bodily distribution of pain, temporal patterns, intensity, emotional and psychological impacts, and others. Additionally, the choice of words may reflect the underlying mechanisms of the causal agent(s) [10] (if any), which can be used to redirect therapeutics.

The language of pain has been previously researched, such as in the structuring of the grammar of pain [12] and the study of its lexical profile, which resulted in the McGill Pain Questionnaire (MPQ) [13] that is widely used to characterize pain from a verbal standpoint in clinical settings [14], [15]. Other works have focused on computationally studying this language, providing evidence for the valuable information it conveys, although not to directly estimate pain intensity. One study explored what people report on social media about their experience of chronic pain to automatically update MPQ descriptors [16]. Another developed a pipeline to extract and interpret biomedical entities and relations, also from social media posts [17]. However, not many works have attempted to use patient narratives directly obtained from registered chronic pain patients, thus, lacking demographic and clinical parameters essential to contextualize their results. One work analyzed language features, directly obtained from patient narratives, to predict chronic pain placebo responders, with 79% accuracy [18]. We also used patient narratives to classify chronic pain patients to one of two rheumatologic base-pathologies, with 79% accuracy [19]. These works highlight the importance of the linguistic expression of chronic pain for its computational assessment. We raise the hypothesis that Natural Language Processing (NLP) techniques can extract useful information from chronic pain patient narratives, and that this insightful information can be used to estimate the current pain intensity and better understand the language of chronic pain.

II. Materials

A. Participants and Data Collection

94 persons participated in this study. These are ≥ 18 years of age, diagnosed with osteoarthritis, rheumatoid arthritis, or spondyloarthritis (including psoriatic arthritis), and pain symptoms, for at least 3 months. All data were collected under a protocol approved by the Ethics Commission of Centro Hospitalar Universitário de São João (CHUSJ), in which data confidentiality is explicitly protected. All participants signed an informed consent form before data collection and after being made aware of the study's purpose. All data are anonymous. Data collection took place at the rheumatology department of CHUSJ from Oct. 2019 to Oct. 2020. At the end of the scheduled clinical appointment, participants were informed of the study, its objectives, and which data would be collected. This included the verbal description of the chronic pain experience (i.e., patient narrative resulting from an interview, recorded with a smartphone only used for this purpose) and demographic and clinical data. The whole collection process was performed by the clinician, in Portuguese. Since language, sociocultural, and psychosocial variables are intertwined, the context in which the patient is describing the experience of chronic pain affects both the vocabulary and how it is reported [20]. Thus, data collection took place after the scheduled appointment with the (already familiar) clinician, in the appropriate clinical context. To maximize the number of participants, three clinicians participated in the collection. Since psychosocial variables might affect chronic pain perception and description, including who the interviewers were, how they were presented, and how they talked to the patient [21], we consider the interviewer as an independent variable.

2372-9198/23/$31.00 ©2023 IEEE
DOI 10.1109/CBMS58004.2023.00134

B. Interview

Due to time constraints, there was no test pilot to investigate possible interview limitations. However, it was designed in conjunction with two pain clinicians, from the same rheumatology department, who have multiple years of experience in communicating with chronic pain patients, and in clinically assessing them. The script was as follows (translated from Portuguese):

1. Where does it hurt?
2. How would you describe your pain? How do you feel it/which sensations does it cause?
3. How has pain intensity evolved in the past month?
4. How would you consider pain to affect your day-to-day, namely, your physical, professional, and social activities and your emotional state?
5. What do you believe to be the cause of your pain?
6. How would you say your pain has evolved, considering the current treatment?
7. How do you expect your pain to develop in the coming months?

C. Demographic and Clinical Data

Demographic data consisted of age, gender, highest completed education, current (or last) professional occupation, and whether the participant was professionally active. Clinical data consisted of the number of years since the initial diagnostic, since the first reports of pain, medication, and values of Erythrocyte Sedimentation Rate (ESR) and C-Reactive Protein (CRP). Pain intensity was self-reported using a VAS, which was discretized into three classes, mild (1-4), moderate (5-6), severe (7-10) [22], due to the limited number of participants. These categories are already commonly used between clinicians and patients [23].

III. Methods

A. Data Preprocessing

Each interview recording was semi-automatically segmented into clinician and patient audio segments. Each patient segment was manually transcribed, maintaining all speech disfluencies [24], without meta-annotations, to keep the textual data as close as possible to the natural utterance. Punctuation was inferred from the pauses and natural flow of the dialog. Patient text was automatically Part-of-Speech (POS) tagged and lemmatized using STRING [25] (using a total of 13 tags: noun, verb, adjective, determinant, pronoun, article, adverb, preposition, conjunction, numeral, interjection, past participle, and relation).

B. Confounding Variables on Pain Intensity

Due to the limited availability of participants, possibly confounding variables, such as age, gender, and pathology, could not be controlled a priori, nor during experimentation. Thus, to maintain internal validity, all demographic and clinical variables were analysed to understand to which extent they explain pain intensity differences in our population.

C. Language Features

Each participant's interview was used to extract a set of features. These were: verbosity (word count, interview length, and word rate), Term Frequency / Inverse Document Frequency (TF-IDF) on the lemmatized text concatenated with the corresponding POS tag, TF-IDF on the POS tags of each lemma, and topic distribution (12 topics, using CluWords [26] – obtained in other experiments with the same data [27]).

D. Pain Intensity Estimation

We considered two classifiers, due to their ease of use, interpretability, and low complexity, which is important for a limited sample size: Decision Tree Classifier (DT), and Support Vector Machine (SVM) (implementation and default parameters from the Sci-kit Learn package [28]). They have been extensively used in similar works [1], [6], [7]. As a baseline, we used the Zero Rate Baseline (ZRB) classifier, which deterministically predicts the most prevalent target class in the training set. We measured model performance according to the weighted F_1 score as implemented in Sci-kit. We focused the evaluation on the weighted F_1 score and not others, such as accuracy, precision, or recall, because it already provides an aggregate view of the last two (equally weighted), and accuracy is not well-suited for imbalanced multi-class settings. We trained and validated our models in a Leave One Out Validation (LOOV) setting, which is used in similar works [2]–[4]. We tested our models with 6 sets of features, effectively obtaining 12 distinct models. The sets of features are: (1) verbosity, (2) TF-IDF, (3) POS TF-IDF, (4) topic distribution, (5) Early Fusion (EF), and (6) Late Fusion (LF). In the case of EF [29], all features were pre-concatenated into a single feature-vector for classification. In the case of LF [29], a different model was fitted with each set of features, using a majority voting scheme for prediction [30]. Given one of the 6 sets of features to perform pain intensity estimation, we used only the top-k features, based on the SelectKBest method provided by Sci-kit. This ranking is given by their ANOVA F-value statistical significance, and k ($2 \leq k \leq 20$) is given by the GridSearchCV method, from the same package. The number of features was limited to 20 to reduce the feature space due to the limited sample size.

E. Influence of Confounding Variables on Language

Language is the dependent variable of our study. However, it may depend on sociocultural, demographic, and clinical parameters. To account for this dependency and modulation of language, for each demographic and clinical parameter, we discarded all language features that lead us to reject the null hypothesis of independence with that parameter.

IV. Results

A. Participants

Of the 94 participants, 65 (56.4 ± 12.7 years of age) completed the full interview and provided information for all data fields. Table I shows their demographic and clinical parameter distributions per pain intensity category.

B. Quality Control of Data Preprocessing

We randomly sampled 10 participants, including 86 transcription files and 1914 POS-tagged and lemmatized tokens. 7 transcription files had at least one error (8.2% error rate), of which some were due to inaudible words. 46 tokens were incorrectly tagged or lemmatized (2.4% error rate). We deemed it unnecessary to check all events for possible errors.

C. Influence of Confounding Variables on Pain Intensity

No statistically significant differences were observed in the three pain intensity categories regarding age (one-way ANOVA, mild, moderate, and severe, respectively: 55.8 ± 13.0, 54.7 ± 12.5, 59.1 ± 10.9; p = 0.63), the number of years since the initial diagnostics (one-way ANOVA: 11.3 ± 9.5, 11.9 ± 9.1, 15.4 ± 8.6; p = 0.37), the number of years since the first reports of pain (one-way ANOVA: 14.6 ± 11.2, 14.8 ± 10.7, 20.1 ± 9.0; p = 0.25), ESR (one-way ANOVA: 18.5 ± 16.5, 21.9 ± 16.9, 21.1 ± 17.2; p = 0.79), and CRP (one-way ANOVA: 6.8 ± 11.9, 6.0 ± 6.3, 5.9 ± 6.0; p = 0.95). However, a chi-square test for independence led us to reject the null hypothesis that pain intensity and gender (p < 0.001), education level (p < 0.001), whether the participant is professionally active (p = 0.002), pathology (p < 0.001), and who the interviewer was (p < 0.001) are independent.

TABLE I. PARAMETER DISTRIBUTION PER PAIN INTENSITY

Pain [a]	**Age**	**Gender**	**Edu. level**
mild (n=38)	55.8 ± 13.0	21×F, 17×M	27×B, 11×H
moderate (n=12)	54.7 ± 12.5	6×F, 6×M	10×B, 2×H
severe (n=15)	59.1 ± 10.9	13×F, 2×M	12×B, 3×H
Pain	**Prof. active**	**Interviewer**	**Diag.**
mild (n=38)	20×A, 18×NA	4×A, 29×B, 5×C	11.3 ± 9.5
moderate (n=12)	7×A, 5×NA	2×A, 10×B	11.9 ± 9.1
severe (n=15)	8×A, 7×NA	3×A, 11×B, 1×C	15.4 ± 8.6
Pain	**Time pain**	**ESR**	**CRP**
mild (n=38)	14.6 ± 11.2	18.5 ± 16.5	6.8 ± 11.9
moderate (n=12)	14.8 ± 10.7	21.9 ± 16.9	6.0 ± 6.3
severe (n=15)	20.1 ± 9.0	5.9 ± 6.0	5.9 ± 6.0
Pain	**Pathology**		
mild (n=38)	20×RA, 16×S, 2×O, 1×PA		
moderate (n=12)	4×RA, 7×S, 3×O		
severe (n=15)	5×RA, 9×S, 1×PA		

a. Variables Age, Time diagnostics (Diag.), and Time pain are reported in years. Legend: F = female, M = male, B = basic education, H = high-school education, A = professionally active, NA = not professionally active, A, B, C = interviewers, RA = rheumatoid arthritis, S = spondyloarthritis, O = osteoarthritis, PA = psoriatic arthritis.

D. Classification Results

Table II shows the classification results. The best performance was obtained when using the technique of EF with the SVM model, with an F_1 score of 0.58. Taking this model, we discarded all features for which we did not have sufficient evidence to reject correlation with demographic and clinical parameters. Performance increased to 0.60 (+0.02). Table III shows the confusion matrix of this score.

TABLE II. WEIGHTED F_1 SCORES FOR EACH TYPE OF FEATURES

Feature set	**ZRB**	**DT**	**SVM**
Verbosity	0.43	0.44	0.43
TF-IDF	0.43	0.55	0.51
POS TF-IDF	0.43	0.36	0.40
Topic distribution	0.43	0.39	0.39
Early Fusion (EF)	0.43	0.56	0.58
Late Fusion (LF)	0.43	0.46	0.40

TABLE III. CONFUSION MATRIX

Model: EF+SVM with select features		**Predicted label** [b]		
		mild	***moderate***	***severe***
Real label	***mild***	0.92 (1.0)	0.00 (0.0)	0.08 (0.0)
	moderate	0.50 (1.0)	0.42 (0.0)	0.08 (0.0)
	severe	0.73 (1.0)	0.13 (0.0)	0.13 (0.0)

b. The confusion matrix for the ZRB is shown in parentheses for comparison.

E. Language Differences Between Intensity Categories

In the LOOV setting we trained n = 65 models on n − 1 samples, each of which selecting k features. In total, 117 unique features were chosen and all of them belonged to the TF-IDF set. Because we preemptively tagged each individual word in the TF-IDF set with its POS tag, we could observe the syntactic distribution of these features: 36% are nouns (average importance ($\times 10^{-3}$) given by each pain intensity category, respectively, mild, moderate, severe: 6, 6, 7), 21% verbs (7, 5, 5), 17% adjectives (3, 3, 4), 11% adverbs (6, 13, 7), and 15% others (5, 7, 5). Table IV shows the topmost weighted words (on average), for each of these POS tags.

TABLE IV. TOP-10 MOST WEIGHTED WORDS FOR EACH POS TAG.

POS	**Top-10 weighted words** [c]
Noun	parte (part; 28, 19, 40), tratamento (treatment; 23, 22, 17), mês (month; 21, 14, 17), noite (night; 5, 53, 25), problema (problem; 16, 32, 12), casa (home; 17, 3, 30), ano (year; 6, 15, 35), posição (position; 6, 17, 16), sentido (sense; 9, 0, 19), um_bocado (a_bit; 4, 0, 19)
Verb	sentir (feel; 44, 38, 34), passar (go_away; 28, 16, 14), acordar (wake; 19, 11, 0), suportar (endure; 8, 12, 7), prender (lock; 14, 0, 0), preocupar (worry; 14, 0, 0), sofrer (suffer; 6, 5, 11), mandar (order; 4, 5, 13), ultrapassar (surpass; 7, 0, 5), pedir (ask; 3, 0, 6)
Adjective	social (social; 3, 14, 13), mínimo (minimum; 5, 0, 15), preso (locked; 11, 0, 0), nervoso (nervous; 9, 0, 0), principal (main; 9, 0, 0), parado (stopped; 5, 0, 4), maluco (crazy; 6, 0, 0), seco (dry; 5, 0, 0), sensação (sensation; 0, 16, 0), tranquilo (tranquil; 4, 0, 0)
Adverb	mais_ou_menos (more_or_less; 24, 47, 17), mal (bad; 12, 19, 18), principalmente (mainly; 13, 6, 17), pouco (a_little; 15, 0, 3), na_totalidade (in_sum; 3, 26, 0), talvez (maybe; 2, 0, 19), sinceramente (sincerely; 2, 5, 9), n_vezes (n_times; 0, 22, 0), sobretudo (above_all; 0, 20, 0), tanto (so_much; 0, 19, 0)

c. English translation is shown in parentheses, followed by the average weight ($\times 10^{-3}$) given per pain intensity category separated by commas, respectively: mild, moderate, severe.

V. DISCUSSION

Although not very large, our sample size compares to other similar works [1]–[5], [18]. Still, it has limited representativeness (participants are mostly above 50 years of age, female, with a basic education level, and with more than 10 years of reported pain). We tried to circumvent this by testing the null hypothesis of independence between all demographic and clinical parameters, and the target pain intensity. However, there is not sufficient statistical evidence to reject the possibility of correlation between pain intensity and gender, education level, whether the participant was professionally active, pathology, and who the interviewer was.

The best results were obtained with the SVM model, using EF, with select features. Specifically, it obtained a weighted F_1 score of 0.60, which represents a 39.5% performance increase over the ZRB. Moreover, it accurately classifies 92% of mild, 42% of moderate, and 13% of severe pain cases. The corresponding confusion matrix allows us to conclude that there is confusion with mild pain, and that the most difficulty arises when distinguishing mild and severe pain patients.

Our results indicate that, for our sample, the most important words that allow to distinguish between pain intensity categories are, in order, nouns, verbs, adjectives, and adverbs. Indeed, nouns allow for the best distinction of severe pain, adverbs for moderate pain, and verbs for mild pain. This is expected, since the remaining POS, such as pronouns, determinants, and prepositions, carry very little semantic value, instead playing a more syntactic role in the construction of sentences. Most words directly relate to some dimension of the experience of chronic pain, e.g., treatment, periods of time, feelings and sensations, suffering, enduring, and surpassing pain. These are the words found to be most relevant for the task of pain intensity classification. More specifically, nouns seem mostly used to refer to treatments, time periods, and parts of the body. Verbs seem to refer mostly to feeling pain and fearing its outcomes. Multiple adjectives are used, of which we highlight the importance of social activity, nervousness, tranquility, and pain-related sensations, such as dryness. Finally, adverbs are used to specify other words, such as adjectives and verbs. Regarding time periods, mild pain associates more importance to the month, while moderate and severe pain associate more importance to the night and to the year, respectively (although not all, some of these differences are indeed striking). It also tells us that the verb "to suffer" is mostly important to severe pain, whilst "to endure" and "to go away" are mostly important to moderate and mild pain, respectively. Expectedly, all pain categories give high importance to the verb "to feel". Our feature extraction did not capture negation, thus not allowing us to judge the use of these specific verbs. The adjective social is important for moderate and severe pain, but not to mild pain, most likely due to the impact caused by pain on the day-to-day life. Finally, moderate pain participants seem to use more descriptive adverbs than patients with mild and severe pains, which suggests a need to detail a possibly ambiguous experience.

We have shown that focus on specific words/themes is especially correlated with specific pain intensity categories. Our approach does not limit chronic pain patients to close-ended questions and answers, allowing them to freely discuss the concerns that they find most relevant at the time of reporting. This inverts the standard assessment methods by letting the patient guide the narrative to their important factors, instead of the other way around, which could help reduce the impact of social biases in self-reports and not force the patient to translate the personal experience of pain into impersonal molds. Future research should focus on expanding the sample size, especially to other cultures and languages, and automating the whole preprocessing pipeline to reduce costs.

Acknowledgments

Diogo A.P. Nunes is supported by a scholarship granted by Fundação para a Ciência e Tecnologia (FCT), with reference 2021.06759.BD. This work was supported by Portuguese national funds through FCT, with reference UIDB/50021/2020. This research was supported by the Portuguese Recovery and Resilience Plan (RRP) through project C645008882-00000055 (Responsible.AI).

References

[1] Z. Hammal and J. F. Cohn, "Automatic Detection of Pain Intensity," in *ICMI'12: Proc of the ACM Intl Conf on Multimodal Interaction*, 2012, pp. 47–52.

[2] G. Bargshady et al., "Enhanced deep learning algorithm development to detect pain intensity from facial expression images," *Expert Syst Appl*, vol. 149, 2020.

[3] X. Xin, X. Lin, S. Yang, and X. Zheng, "Pain intensity estimation based on a spatial transformation and attention CNN," *PLoS One*, vol. 15, p. e0232412, 2020.

[4] M. Tavakolian and A. Hadid, "A Spatiotemporal Convolutional Neural Network for Automatic Pain Intensity Estimation from Facial Dynamics," *Int J Comput Vis*, vol. 127, no. 10, pp. 1413–1425, 2019.

[5] J. Zhou, X. Hong, F. Su, and G. Zhao, "Recurrent Convolutional Neural Network Regression for Continuous Pain Intensity Estimation in Video," in *Proc of the IEEE Conf on Comput Vis and Pattern Recognition (CVPR) Workshops*, 2016, pp. 84–92.

[6] A. Kimura et al., "Objective characterization of hip pain levels during walking by combining quantitative electroencephalography with machine learning," *Sci Rep*, vol. 11, no. 1, 2021.

[7] J. Lee et al., "Machine learning-based prediction of clinical pain using multimodal neuroimaging and autonomic metrics," *Pain*, vol. 160, no. 3, pp. 550–560, 2019.

[8] A. Marquand et al., "Quantitative prediction of subjective pain intensity from whole-brain fMRI data using Gaussian processes," *Neuroimage*, vol. 49, no. 3, pp. 2178–2189, 2010.

[9] D. Lopez-Martinez and R. Picard, "Continuous Pain Intensity Estimation from Autonomic Signals with Recurrent Neural Networks," in *40th Ann Intl Conf of the IEEE Engineering in Medicine and Biology Society (EMBC)*, 2018.

[10] D. Wilson, "Language and the pain experience," Physiotherapy Research International Physiother. *Res. Int*, vol. 14, no. 1, pp. 56–65, 2009.

[11] R. Melzack and Torgerson WS, "On the Language of Pain," *Anesthesiology*, vol. 34, no. 1, pp. 50–59, 1971.

[12] Halliday M.A.K., "On the grammar of pain," *Functions of Language*, vol. 5, no. 1, pp. 1–32, 1998.

[13] R. Melzack, "The McGill Pain Questionnaire: Major Properties and Scoring Methods," *Pain*, vol. 1, pp. 277–299, 1975.

[14] J. Katz and R. Melzack, "Measurement of Pain," *Surgical Clinics of North America*, vol. 79, no. 2, pp. 231–252, 1999.

[15] M. D. Sullivan, "Pain in language: from sentience to sapience," *Pain Forum*, vol. 4, no. 1, pp. 3–14, 1995.

[16] M. Y. Tan, C. E. Goh, and H. H. Tan, "Contemporary english pain descriptors as detected on social media using artificial intelligence and emotion analytics algorithms: Cross-sectional study," *JMIR Form Res*, vol. 5, no. 11, 2021.

[17] V. Foufi, T. Timakum, C. Gaudet-Blavignac, C. Lovis, and M. Song, "Mining of textual health information from Reddit: Analysis of chronic diseases with extracted entities and their relations," *J Med Internet Res*, vol. 21, no. 6, 2019.

[18] S. E. Berger et al., "Quantitative language features identify placebo responders in chronic back pain," *Pain*, vol. 162, no. 6, pp. 1692–1704, 2021.

[19] D. A.P. Nunes et al., "Analysis of chronic pain descriptions for base-pathology prediction: the case of rheumatoid arthritis versus spondylitis pathology prediction based on pain descriptions," *DOR*, vol. 28, no. 2, pp. 78–87, 2022.

[20] G. R. Hansen and J. Streltzer, "The psychology of pain," *Emergency Medicine Clinics of North America*, vol. 23, no. 2. W.B. Saunders, pp. 339–348, 2005.

[21] T. Hechler, D. Endres, and A. Thorwart, "Why harmless sensations might hurt in individuals with chronic pain: About heightened prediction and perception of pain in the mind," *Front Psychol*, vol. 7, no. OCT, 2016.

[22] R. C. Serlin, T. R. Mendoza, Y. Nakamura, K. R. Edwards, and C. S. Cleeland, "When is cancer pain mild, moderate or severe? Grading pain severity by its interference with function," *Pain*, vol. 61, no. 2, pp. 277–284, 1995.

[23] K. O. Anderson, "Role of cutpoints: Why grade pain intensity?," *Pain*, vol. 113, no. 1. pp. 5–6, 2005.

[24] E. Shriberg, "To 'errrr' is human: Ecology and acoustics of speech disfluencies," *J Int Phon Assoc*, vol. 31, no. 1, pp. 153–169, 2001.

[25] N. Mamede, J. Baptista, C. Diniz, and V. Cabarrão, "STRING: A hybrid statistical and rule-based natural language processing chain for portuguese," in *10th Intl Conf on Computational Processing of Portuguese, Demo Session*, 2012.

[26] F. Viegas et al., "Cluwords: Exploiting semantic word clustering representation for enhanced topic modeling," in WSDM 2019 - *Proc of the 12th ACM Intl Conf on Web Search and Data Mining*, 2019, pp. 753–761.

[27] D. A.P. Nunes, J. Ferreira Gomes, F. Neto, and D. Martins de Matos, "Chronic Pain and Language: A Topic Modelling Approach to Personal Pain Descriptions," Sep. 2021, Accessed: Feb. 22, 2023. [Online]. Available: http://arxiv.org/abs/2109.00402

[28] F. Pedregosa et al., "Scikit-learn: Machine Learning in Python," *J of Machine Learning Research*, vol. 12, pp. 2825–2830, 2011, [Online]. Available: http://scikit-learn.sourceforge.net.

[29] J. Sebastian and P. Pierucci, "Fusion techniques for utterance-level emotion recognition combining speech and transcripts," in *Proc of the Ann Conf of the Intl Speech Communication Association, INTERSPEECH*, 2019, vol. 2019-September, pp. 51–55.

[30] R. Atallah and A. Al-Mousa, "Heart Disease Detection Using Machine Learning Majority Voting Ensemble Method," in *2nd Intl Conf on New Trends in Computing Sciences (ICTCS)*, 2019