Exploring Gender Differences in Chronic Pain Discussions on Reddit

Ancita Maria Andrade[1], Tanvi Banerjee[1], and Ramakrishna Mundugar[2]

[1] Wright State University, Dayton OH 45431, USA,
andrade.12@wright.edu, tanvi.banerjee@wright.edu
[2] Manipal Institute of Technology, Manipal Academy of Higher Education Manipal, KA, India, 576104
ramakrishna.m@manipal.edu

Abstract. Pain, both physical and emotional, is a fundamental human experience categorized as acute or chronic. While research has explored its causes and treatments across disciplines, the influence of gender has often been overlooked in earlier studies. In this study, we utilized Natural Language Processing (NLP) to analyze and gain deeper insights into individuals' pain experiences, with a particular focus on gender differences. We successfully classified posts into male and female corpora using the Hidden Attribute Model-Convolutional Neural Network (HAM-CNN), achieving an F1 score of 0.86 by aggregating posts based on usernames. Our analysis found that female posts were more emotionally focused, with conditions like migraine and sinusitis more common among women. It also revealed gender-based differences in responses to pain medication.

Keywords: Chronic pain, Gender identity, Reddit, Medication

1 Introduction

Social media platforms have become powerful tools for communication, enabling individuals worldwide to share experiences, seek support, and access information within diverse virtual communities. Their broad reach, accessibility, and anonymity foster open discussion of sensitive or stigmatized topics. Reddit, in particular, stands out due to its community-driven structure with topic-specific subreddits and active moderation, fostering focused and supportive environments. This makes social media, and Reddit specifically, an invaluable resource for capturing authentic, patient-driven narratives that offer unique insights beyond clinical data.

Chronic pain is a widespread public health challenge affecting millions globally, with notable differences in how men and women experience and express pain. Studying these gender differences is essential to understanding these variations in pain perception and expression, which can influence diagnosis, treatment, and support. Analyzing online narratives helps researchers understand how gender influences chronic pain experiences and supports the development of gender-sensitive pain management approaches. In this study, we focused on

gender identity and limited our analysis to a binary classification of male and female as non-binary users make up only 1.3% of the total reddit users.

This study addresses the following research questions.

1. How effectively can supervised machine learning techniques predict an author's gender based on features extracted from their Reddit posts or comments?
2. To what extent can we use data-driven techniques to elicit common ways in which men and women differ in expressing their pain?
3. How effective are techniques such as topic modeling in identifying gender-based differences in discussions related to chronic conditions and the experiences shared by individuals?
4. What are the most important differences in patterns of medication usage between men and women using Named Entity Recognition (NER) and sentiment analysis?

2 Related Work

This study is grounded in gender prediction, which necessitates a labeled dataset for supervised machine learning. We utilized the RedDust dataset [1], which compiles user-disclosed attributes from Reddit, including gender, age, profession, and hobbies. Among the predictive models evaluated, the Hidden Attribute Model–Convolutional Neural Network (HAM-CNN) demonstrated the highest accuracy for gender classification.

Author profiling has also been extensively explored on Twitter, with prior work employing a range of supervised learning techniques, from traditional models such as logistic regression and support vector machines to more advanced deep learning architectures, including gated recurrent units (GRUs) [4].

Reddit has previously been used to study chronic pain; one study [3] analyzed 12 related subreddits, identifying key themes and overlaps, especially in r/backpain and r/sciatica.

3 Methods

3.1 Data Collection and Preprocessing

To support gender classification, we used the RedDust dataset [1], which contains Reddit user attributes, including gender. We collected posts and comments from 36 chronic pain-related subreddits using the Python Reddit API Wrapper (PRAW), covering both general (e.g., r/ChronicPain, r/Pain) and condition-specific communities (e.g., r/Migraine, r/Sinusitis). Data was collected up to March 31, 2025, resulting in 28,067 posts from approximately 58,000 users. We filtered the dataset to include only English-language posts and comments containing at least three words. Entries authored by bots or moderators (e.g., 'Automoderator') were excluded. Duplicate posts by the same user across multiple subreddits were also removed to reduce redundancy.

3.2 Gender Classification

To address RQ1, we first trained gender classification models using the RedDust dataset, which contains 2.49 million posts from 54.8K users. Each post/comment was treated as an independent instance (post-level classification), suitable for Reddit's prevalent use of throwaway accounts.

Table 1: Metrics for the gender classifier for different models

Model	**Gender**	**Precision**	**Recall**	**F1 Score**
Logistic Regression (Post Level)	Male	0.64	0.66	0.65
	Female	0.65	0.63	0.64
Bi GRU (Post Level)	Male	0.67	0.68	0.67
	Female	0.67	0.66	0.67
HAM-CNN (Post Level)	Male	0.66	0.72	0.69
	Female	0.69	0.63	0.66
HAM-CNN (User Level)	Male	**0.86**	**0.86**	**0.86**
	Female	**0.86**	**0.87**	**0.87**

We evaluated both traditional and deep learning models, including Logistic Regression, bi-directional GRU [4], and HAM-CNN [1]. To handle class imbalance, we applied undersampling and used a stratified train-test split (80:20). HAM-CNN achieved the best performance at the post level (F1-score = 0.67).

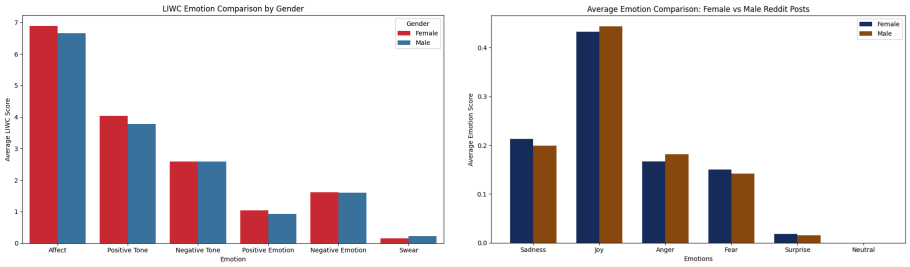
Following prior work [6], we also trained the HAM-CNN model at the user level by aggregating all posts/comments per user. This approach substantially improved performance, yielding an F1-score of **0.86** and an AUC of **0.94**.

For gender prediction in our collected Reddit dataset, we grouped content by username and applied the user-level HAM-CNN model. To ensure high precision, only users with confidence scores above 0.75 (female) or below 0.25 (male) were retained, resulting in a final set of approximately 43,000 users.

3.3 Text Analysis Methods

Reddit users often share a wide range of experiences, from personal achievements to difficult challenges. To analyze the emotional content of these discussions, we employed LIWC (Linguistic Inquiry and Word Count) to examine psychological and linguistic attributes, including positive and negative tone and the use of swear words. We used Hugging Face Transformers to detect specific emotions such as sadness, joy, anger, fear, surprise, disgust, and neutrality in posts from both male and female users.

To uncover latent themes in user discussions, we applied Latent Dirichlet Allocation (LDA) separately to male and female corpora. The optimal number of topics was selected based on coherence scores, resulting in 8 topics for the female dataset and 3 for the male dataset.



(a) Comparing the Sentiments across female and male posts

(b) Comparing the Emotions across female and male posts

Fig. 1: Comparative Analysis of Sentiments and Emotions by Gender

NER was applied using the Python Drug NER library to extract high-confidence drug mentions from Reddit posts. While effective in identifying full drug names (e.g., "Methadone," "Lidocaine," "Gabapentin"), the tool missed abbreviations or short forms (e.g., "oxy" for oxycodone).

We focused on frequently mentioned drugs—Amitriptyline, Lidocaine, Morphine, Pregabalin, and Tramadol—excluding inconsistent data on dosages or specific conditions. A rule-based classifier categorized posts into four groups: effective, ineffective, side effects, and neutral, based on keyword patterns and symptom references. Subsequent descriptive and statistical analyses examined drug efficacy and emotional sentiment across genders.

4 Results and Discussion

We analyzed LIWC affect features by gender to understand emotional expression in chronic pain discussions. Female users showed higher fear and sadness, especially when discussing symptoms and medication side effects—aligning with prior research [9,10] noting greater emotional intensity and frustration among women and anger often directed inward or at healthcare providers.

Our analysis showed that male users expressed more anger compared to females. Their posts frequently conveyed frustration and irritation, often aimed at themselves or close family members. Interestingly, many male users also expressed feelings of guilt or remorse for experiencing such emotions. In contrast, female users demonstrated a broader range of emotional expression, including more intensely positive and negative experiences. This is supported by prior research [5], which suggests that women may have stronger emotional memory due to interactions between the amygdala and hippocampus. Neurological studies [8] have found that females generally have larger orbital frontal cortex volumes, which may help regulate anger and aggression. These biological differences could help explain the observed gender variation in emotional expression and the use of strong language.

Table 2: LDA Topics Extracted from Male Posts

Topic No	Theme and Associated Words	Example Posts/Comments
1	**Exercises & Musculoskeletal Health** nerve, leg, exercise, disk, pt, muscle, stretch, diet, mri, supplement	" I've seen several guys focus way to hard on deadlift, and not prioritize core, and get really bad anterior pelvic tilt and way too much curve." "I wouldn't stop lifting all together, this could actually make things worse"
2	**Social coping** need, doctor, life, sorry, best, post, diagnose, experience, hope, ask	" I want to manage tolerance as best as possible." "I suppose it's natural to start seeing the world and life from a very different perspective than most people."
3	**Temporal chronic pain experience** bad, night, work, week, sleep, year, hour, migraine, hour, month	"I had longer period of back pain after lifting about 1 year ago." "ive been experiencing some weird symptoms that feels like they originate from my upper stomach."

LDA was employed to analyze topic distributions separately for male and female corpora, with the results presented in Table 2 and Table 3. The analysis revealed two key themes: gender-based differences in commonly discussed chronic conditions and variations in the expression of lived experiences.

In the female dataset, six out of eight topics were primarily associated with specific chronic conditions, including vulvodynia, endometriosis, irritable bowel syndrome (IBS), sinusitis, lower back pain, and migraine. These findings align with prior literature indicating that such conditions are more prevalent among females [11]. Additionally, topics relating to the personal impact of chronic pain and the significance of social support emerged. The frequent use of terms such as "helpful," "appreciate," and "share" suggests a higher degree of emotional expression and a propensity for seeking social connection, consistent with emotion-focused coping strategies reported in previous studies [7].

Conversely, the male corpus revealed topics emphasizing physical activity, musculoskeletal health, and coping mechanisms. Topic 1 was largely centered around physical wellness, while Topic 2 reflected self-management strategies, including strength training, rehabilitation programs, medication use, and active rest. This pattern indicates a tendency among male users to engage in problem-focused coping and behavioral distraction, favoring solution-oriented approaches over emotional expression. These results are consistent with existing research suggesting that men often prioritize the physical aspects of pain management [2].

The gender-based descriptive analysis of drug outcomes is summarized in Table 4. Z-tests (p=0.05) were conducted to assess differences in drug efficacy

Table 3: LDA Topics Extracted from Female Posts

Topic No	Theme and Associated Words	Example Posts/Comments
1	**Vulvodynia** test, symptom, doctor, infection, burning, cream, uti, vulvodynia, treatment, birth, bladder	"I have been experiencing vestibulodynia for the past two years." "She did conclude that this pain was at my vulva and mentioned vulvodynia"
2	**Endometriosis** endo, surgery, endometriosis, period, doctor, cyst, symptom, surgeon, year, ultrasound	"It is very scary to navigate endometriosis". "How frequently do they deal with cysts 10+ cm?"
3	**Impact of chronic pain** work, life, care, understand, chronic, friend, hope, able, job, illness	"Accepting that I was no longer able to be 100% independent and self-sufficient was EXTREMELY difficult for me"
4	**IBS& Diet** eat, food, diet, ibs, stomach, drink, bad, trigger, food, help, sugar	"I have IBS-M and had to cut out foods with fiber." "I struggled with IBS for the past 10 years. Wherever I went, I had to make sure a bathroom was nearby."
5	**Sinusitis** sinus, nose, nasal, infection, ent, allergy, rinse, eye, love, post	"But that led me to a CT scan which showed all 8 sinuses infected and completely blocked."
6	**Pelvic & lower back pain** pelvic, nerve, muscle, floor, pt, therapy, exercise, adhesion, stretch, physical	"Lower Back Pain - Ive had sciatic pain since I was around 18 years old "
7	**Social support/Emotional glue** thank, share, good, great, appreciate, wow, glad, helpful, advice, awesome	"I'm glad everyone has chimed in with their two cents." "Thank you. I hope you get some relief from your pain as well."
8	**Migraine & Headache** migraine, headache, bad, stop, medication, sleep, med, fatigue, start, work	"I'm new to this migraine rollercoaster. I've been struggling with what I now know are migraines recurrently"

Table 4: Efficacy and Emotion Distribution Percentages of Drugs by Gender

Drug	Gender	Total Mentions	Effective (%)	Ineffective (%)	Side Effect (%)	Anger (%)	Fear (%)	Joy (%)	Sadness (%)
Amitriptyline	Female	180	**68.3**	1.1	6.7	23.89	**16.67**	43.33	16.11
	Male	6	66.7	0.0	0.0	16.67	33.33	**33.33**	16.67
Lidocaine	Female	471	64.1	0.9	4.5	17.41	16.14	41.83	22.51
	Male	81	59.3	**3.7**	7.4	13.58	8.64	46.91	29.63
Morphine	Female	325	49.8	1.2	6.2	**22.77**	22.15	31.38	21.85
	Male	52	63.5	3.8	1.9	7.69	13.46	**55.77**	23.08
Pregabalin	Female	435	58.2	1.4	7.1	18.85	**16.75**	36.32	26.90
	Male	111	56.8	0.9	4.5	11.71	8.11	**57.66**	20.72
Tramadol	Female	296	57.1	1.0	6.4	20.61	12.50	41.22	**23.65**
	Male	84	58.3	1.2	1.2	19.05	**21.43**	46.43	11.90

and emotional responses across genders. Key findings for each drug are outlined below.

Amitriptyline: A significantly higher proportion of females reported the drug as effective ($p = 0.04$), though no significant differences were found for ineffec-

tiveness ($p = 0.18$) or side effects ($p = 0.086$). Prior studies suggest amitriptyline may be safer for women [13], and men may experience side effects like restless leg syndrome more frequently [12].

Lidocaine: More males reported lidocaine as ineffective. Other outcomes showed no significant gender differences. Previous research has reported mixed results, with some studies suggesting men and women respond differently depending on the method of administration [14].

Morphine: While efficacy did not differ significantly, females expressed more anger ($p = 0.012$) and males more joy ($p = 0.001$). This may relate to women's higher likelihood of adverse effects such as nausea and dizziness [17].

Pregabalin: Females expressed significantly more fear ($p = 0.02$), while males expressed more joy ($p = 0.00$). Preclinical studies suggest higher susceptibility to pregabalin abuse in female mice [15], though human evidence remains limited.

Tramadol: Gender differences emerged in emotional responses: males expressed more fear ($p = 0.04$), and females more sadness ($p = 0.019$). Studies show women are more prone to tramadol-associated side effects, including vomiting and other adverse reactions [16], possibly explaining the emotional variation.

5 Conclusion

In this study, we used NLP to analyze Reddit posts on chronic pain, revealing gender-based differences in emotional expression and medication responses. These insights can enhance patient-provider communication, support demographic-specific care, and bring attention to important concerns such as medication side effects and inequities in care. Our future work will involve collaborating with clinical specialists to examine how medications impact individuals of different genders.

Acknowledgements We would like to express our gratitude to Anna Tigunova for providing the RedDust dataset to support the development of our gender classifier.

References

1. A. Tigunova, A. Yates, P. Mirza, and G. Weikum, "c European Language Resources Association (ELRA), licensed under CC," 2020. Available: http://www.lrec-conf.org/proceedings/lrec2020/pdf/2020.lrec-1.751.pdf
2. J. Strong, T. Mathews, R. Sussex, F. New, S. Hoey, and G. Mitchell, "Pain language and gender differences when describing a past pain event," Pain, vol. 145, no. 1, pp. 86–95, Sep. 2009, doi: https://doi.org/10.1016/j.pain.2009.05.018.
3. D. A. P. Nunes, J. Ferreira-Gomes, F. Neto, and D. M. de Matos, "Modeling chronic pain experiences from online reports using the Reddit Reports of Chronic Pain dataset," Information, vol. 14, no. 4, p. 237, Apr. 2023, doi: https://doi.org/10.3390/info14040237.

4. C. Suman, R. S. Chaudhary, S. Saha, and P. Bhattacharyya, “An attention based multi-modal gender identification system for social media users,” Multimedia Tools and Applications, Sep. 2021, doi: https://doi.org/10.1007/s11042-021-11256-6.
5. L. Brizendine, Female Brain, Phoenix Publishing Project, 2014.
6. E. Vasilev, ”People and Knowledge Networks: Inferring Gender of Reddit Users,” Master’s thesis, University of Koblenz-Landau, 2016. [Online]. Available: https://kola.opus.hbz-nrw.de/frontdoor/deliver/index/docId/1619/file/Master_thesis_Vasilev.pdf. Accessed: Apr. 17, 2025.
7. E. J. Bartley and R. B. Fillingim, “Sex differences in pain: a brief review of clinical and experimental findings,” British Journal of Anaesthesia, vol. 111, no. 1, pp. 52–58, Jul. 2013, doi: https://doi.org/10.1093/bja/aet127.
8. E. Güvendir, “Why are males inclined to use strong swear words more than females? An evolutionary explanation based on male intergroup aggressiveness,” Language Sciences, vol. 50, pp. 133–139, Jul. 2015, doi: https://doi.org/10.1016/j.langsci.2015.02.003.
9. J. L. Riley, M. E. Robinson, J. B. Wade, C. D. Myers, and D. D. Price, “Sex Differences in Negative Emotional Responses to Chronic Pain,” The Journal of Pain, vol. 2, no. 6, pp. 354–359, Dec. 2001, doi: https://doi.org/10.1054/jpai.2001.27000.
10. A. Okifuji, D. C. Turk, and S. L. Curran, “Anger in chronic pain,” Journal of Psychosomatic Research, vol. 47, no. 1, pp. 1–12, Jul. 1999, doi: https://doi.org/10.1016/s0022-3999(99)00006-9.
11. “Gender Differences in Chronic Pain Conditions,” International Association for the Study of Pain (IASP). https://www.iasp-pain.org/resources/fact-sheets/gender-differences-in-chronic-pain-conditions/
12. K. R. Baughman, C. C. Bourguet, and S. K. Ober, “Gender differences in the association between antidepressant use and restless legs syndrome,” Movement Disorders, vol. 24, no. 7, pp. 1054–1059, Apr. 2009, doi: https://doi.org/10.1002/mds.22517.
13. L. W.-H. Chen et al., “Amitriptyline and Sexual Function: A Systematic Review Updated for Sexual Health Practice,” American Journal of Men’s Health, vol. 12, no. 2, pp. 370–379, Oct. 2017, doi: https://doi.org/10.1177/1557988317734519.
14. F. Salis, S. Sardo, G. Finco, Gian Luigi Gessa, F. Franconi, and R. Agabio, “Sex-Gender Differences Are Completely Neglected in Treatments for Neuropathic Pain,” Pharmaceuticals, vol. 17, no. 7, pp. 838–838, Jun. 2024, doi: https://doi.org/10.3390/ph17070838.
15. H. O. Alsaab et al., “Sex differences in pregabalin-seeking like behavior in a conditioned place preference paradigm,” Saudi Pharmaceutical Journal, vol. 28, no. 12, pp. 1749–1755, Dec. 2020, doi: https://doi.org/10.1016/j.jsps.2020.11.001.
16. J. Muriel et al., “Long-term deprescription in chronic pain and opioid use disorder patients: Pharmacogenetic and sex differences,” Acta pharmaceutica, vol. 73, no. 2, pp. 227–241, Jun. 2023, doi: https://doi.org/10.2478/acph-2023-0018.
17. R. B. Fillingim et al., “Morphine responses and experimental pain: Sex differences in side effects and cardiovascular responses but not analgesia,” The Journal of Pain, vol. 6, no. 2, pp. 116–124, Feb. 2005, doi: https://doi.org/10.1016/j.jpain.2004.11.005.