Exploring Optimism and Pessimism in Online Discourse of Individuals with Depression

Anonymous ACL submission

Abstract

The relationship between depression and the concepts of optimism and pessimism has been extensively researched by psychologists. In this paper, we use computational approaches to study how optimism and pessimism are expressed in the online discourse of people diagnosed with depression. Publicly available datasets are used for the development of an optimism/pessimism detection model, as well as for the analyses performed on social media posts of individuals with depression, as measured by BDI-II, a validated questionnaire for assessing depression. To analyze the optimistic and pessimistic posts by individuals with depression, we use LIWC features and perform topic modeling. Our results show that while there might not be significant differences between the amount of optimistic versus pessimistic posts depressed and control individuals have, the content of the posts differ meaningfully, both in terms of linguistic features and approached topics.

1 Introduction

Depression is one of the most prevalent mental disorders and has been extensively researched (Lim et al., 2018; Xu et al., 2021). Many studies focus on understanding how depression manifests and its relationship with mood and emotions (Rottenberg, 2005). In addition to emotions, previous research has also investigated the connection between depression and the concepts of optimism and pessimism. Karhu et al. (2024) demonstrate a bidirectional relationship: optimism not only buffers against depressive symptoms but is also eroded by them, while pessimism both predicts and is intensified by depression. Complementary studies by Korn et al. (2014) and Hobbs et al. (2022) reveal that, unlike healthy individuals who display an optimistic bias when updating beliefs about the future, those with depression tend to weigh negative information more heavily. In addition, optimism is associated with better psychological well-being and more effective coping (Scheier et al., 2001), as well as better treatment outcomes, including reduced rehospitalization (Tindle et al., 2012). Prior research also highlights a reduced risk of work disability and an enhanced likelihood of returning to work following a depression-related disability (Kronström et al., 2011).

In recent years, computational analyses of social media data have provided insights into the relationship between psychological constructs and mental health. De Choudhury et al. (2013) identified Twitter users with depression through language and engagement analysis, indicating that social media can help spot depression onset, enabling proactive mental health interventions. Chen et al. (2020) found that mood shifts in social media correlate with symptom scores, suggesting language reflects actual mood in depression.

Depression detection is a prominent topic in NLP, with traditional methods such as Support Vector Machines (SVMs), logistic regression, and random forests being used (Gan et al., 2024). More recently, there has been a transition to modern methods that use attention, deep learning, and pretrained models (De Santana Correia and Colombini, 2022), demonstrating significant performance increases. However, in addition to identifying mental health disorders, language can offer insights into broader psychological states, such as optimism and pessimism, which are often associated with conditions like depression (Herwig et al., 2009). Previous research from NLP has explored the manifestations of emotions (Uban et al., 2021; Aragon et al., 2021) and even happy moments using social media data from individuals with depression (Bucur et al., 2024). Although research from NLP has focused on developing more effective models for detecting optimism and pessimism (Ruan et al., 2016; Caragea et al., 2018; Alshahrani et al., 2021), to our knowledge, there has been no analysis of optimism and pessimism in the social media language used by individuals with depression.

This study explores the correlation between optimism and pessimism on social media and depressive symptoms, a link supported by psychological literature. The aim is to develop optimism-pessimism detection systems using advanced transformer-based architecture and conduct studies on the relationship between optimism and mental health issues like depression. Detecting optimism and pessimism in social media is considered a first step towards understanding and detecting mental health issues. To the best of our knowledge, this is the first computational analysis of the correlation between optimistic and pessimistic social media language in people with depression. Thus, we aim to answer the following research questions:

- **RQ1**: In what proportions are optimism and pessimism respectively manifested in the discourse of individuals with depression?

- **RQ2**: How is optimism manifested in the social media language of individuals with depression?

Quantifying optimism and pessimism in depressive discourse (RQ1) challenges the traditional view that depression is solely characterized by negative affect. Analyzing social media language for manifestations of optimism (RQ2) uncovers subtle linguistic cues that conventional assessments might miss. This approach can inform the development of digital tools for early detection and personalized intervention strategies.

2 Related Work

Though it is still in its early stages, research on detecting optimism and pessimism in social media is expanding, partly because of the COVID-19 epidemic. A deep-learning technique was presented by Blanco and Lourenço (2022) to examine the expression of optimistic and pessimistic sentiments in COVID-19-related Twitter conversations. They examined several network configurations using a pre-trained transformer embedding for semantic feature extraction and found that bi-LSTM systems produced the most successful models. According to the study, optimistic interactions tended to stay positive, whereas conversations with strong pessimistic signals showed little emotional change.

In order to improve prediction accuracy for optimism and pessimism, Alshahrani et al. (2020) employed XLNet, a network that combines several auto-regressive language models, to capture semantic relationships and negations. On the benchmark dataset OPT (Ruan et al., 2016), the study's significant 63.32% error reduction increased the state-of-the-art accuracy from 90.32% to 96.45%. Accuracy at the tweet and user levels for two defined thresholds—0 and 1/-1—was one of the assessment measures.

Cobeli et al. (2022) introduce a Multi-Task Knowledge Distillation architecture, achieving an accuracy of 86.60% on the OPT dataset. The research found that certain POS tags, such as nouns, are consistently prevalent throughout all optimism ranges. Other tags, like hashtags, are associated with optimism levels. The use of emoticons, punctuation, and user remarks also influenced optimism. As tweets became more positive, first-person singular pronouns were less frequent, suggesting a connection between pessimism and depression. The architecture outperformed earlier setups for the 1/-1 threshold definition of optimism.

The concept of computational analyses in the field of mental health detection correlations in social media speech has been investigated to an extent in the study by Bucur et al. (2021), which looks into the relationship between offensive language and depression by examining how people with depression use offensive speech in their social media posts. According to the authors' data, there is a greater prevalence of derogatory language in the online speech of individuals who have been diagnosed with depression.

In our research, we use computational methods to analyze the online discourse of individuals with depression. We aim to explore the impact of optimism and pessimism, motivated by existing psychological research and advancements in NLP models designed to detect these two mental attitudes.

3 Data

We use two data collections in our experiments: the OPT dataset (Ruan et al., 2016) with annotations for optimism and pessimism, and the eRisk 2021 dataset (Parapar et al., 2021) with social media individuals with depression.

The most popular dataset for optimism/pessimism identification was introduced by Ruan et al. (2016). It contains 7,475 randomly chosen tweets from 500 pessimistic individuals and 500 who were considered optimists. To select the texts, tweets containing optimism or pessimism-related keywords were found, highlighting both optimistic and pessimistic users. Each tweet was evaluated and classified by five human annotators using Amazon Mechanical Turk on a scale. To guarantee accuracy, quality control procedures were put in place, such as defining optimism and pessimism precisely, excluding commentators who answered "check" questions incorrectly, and comparing annotations to the average score to spot anomalies. Human annotators rated tweets on a disposition scale from 3 (extremely optimistic) to -3 (very pessimistic); this scale made it possible to distinguish between tweets in a complex way, allowing different levels of optimism and pessimism to be identified within the text. The average of all the evaluations for the acquired annotations is the final score. The rigorous quality control procedure resulted in a high final inter-annotator agreement (Krippendorff's alpha of 0.731).

In our experiments, we consider the three possible classes: posts with an average annotation below -1 are labeled as pessimistic, those with a score of -1 to 1 belong to the neutral class, and the remaining posts are optimistic. This three-class setting provides greater granularity and intuitiveness.

Our approach is different from the direction taken in the studies mentioned in the previous section; both works identify the need to address posts with average scores between -1 and 1 separately, as they are the most ambiguous in the given context, even for human interpretation. In one of their approaches, Cobeli et al. (2022) choose to eliminate the specific group of posts, and consider the two classes, optimistic and pessimistic, so as to have a clearer distinction between the two attitudes. Alshahrani et al. (2020) employed the same method of ignoring the respective posts to address the ambiguity, calling it the -1/1 threshold. In both studies, this approach significantly improved model performance; however, for our work, we chose not to use a similar technique but rather keep the ambiguous data and create an additional class for it, for two main reasons:

1. We believe retaining this data ensures preserving the complexity and authenticity of real-world social media posts, as realistically, not all posts are and should be classified as either optimistic or pessimistic

2. Eliminating the respective posts would mean reducing the data to almost half of the original size (3,847).

The eRisk 2021 dataset related to depression (Losada and Crestani, 2016; Parapar et al., 2021) contains social media users who were asked to fill in the BDI-II questionnaire (Beck et al., 1996) for the assessment of their depression status. Following this, their Reddit social media data was collected with their consent. The BDI-II questionnaire contains 21 questions related to depression symptoms, and the answers are used to calculate an overall score that indicates the level of depression. The training dataset consists of 90 users with ground truth BDI-II scores and 46,502 posts from Reddit. The test dataset contains 80 users with a total of 32,237 posts. In our experiments, we use the data from all 170 users in the eRisk dataset. Because BDI-II is used by mental health professionals to diagnose depression, we consider users with a score above the established cut-off of 19 (Subica et al., 2014; von Glischinski et al., 2019) as having depression, while those with scores below this threshold are considered control users.

To ensure the model reliably detects optimism and pessimism in Reddit posts, we are including a performance evaluation using a manually labeled sample. In this sense, we have selected a total of 150 posts in equal amounts from all possible subgroups: optimistic/pessimistic/neutral posts from depressed individuals, as well as optimistic/pessimistic/neutral posts from the control group. The texts were then manually rated by 3 human annotators, following the procedure from Ruan et al. (2016), resulting in an average score in the -3/3 interval. The obtained score is then mapped according to the three available classes, the same as the original OPT data. These labels are used to validate the results obtained by our optimism/pessimism detection model, with results being presented in a later section (Section 5.1).

4 Methodology

4.1 Detection of optimism and pessimism

Due to its good downstream performance across a great variety of tasks (Liu et al., 2019; Guo et al., 2022; Amin et al., 2023), we use in our experiments a RoBERTa-based model fine-tuned on the OPT dataset, which is then used to predict optimism, pessimism and neutral labels on the eRisk depression data.

The model, which we will refer to as RoBERTa-OPT-3Labels from now on, was trained using

Optimism		Pessimism		Neutral	
Control	**Depression**	**Control**	**Depression**	**Control**	**Depression**
I'm happy that everything turned out rather well for you in the end, and that gives me a lot of hope for my future.	I graduated [...] and got my driver's license! [...] I know what the next goal to work for is. [...] I honestly value my friendships more.	It is sad to think that the life that we will live in is set for imminent destruction.	Something must always [...] remind me how painful life is and that it will never GENUINELY get better. [...] Everyone would be better off without me [...] I will never be good enough.	Beagles are usually listed as a breed that tends to get along well with cats [...]	I only consume great, but lesser-known media. Are you familiar with Steins;Gate and Morrowind? Thought so.

Table 1: Selected examples that were predicted as optimistic, pessimistic, or neutral from the depression and control groups.

the HuggingFace platform, with twitter-roberta-base-sentiment-latest serving as the base model (Camacho-Collados et al., 2022). The base model was refined for sentiment analysis using the TweetEval benchmark (Barbieri et al., 2020) after being trained on about 124 million tweets. In our training, we set a learning rate of 5e-5, three epochs, a maximum sequence length of 128 tokens, an 8-batch size, and a warmup ratio of 0.1. To reduce overfitting, the optimizer employed was AdamW, a variation of the Adam optimizer with weight decay. The learning rate was decreased linearly from the starting value to zero using the "linear" learning rate scheduler. In order to avoid exploding gradient problems, the maximum gradient norm was fixed at 1. To guarantee consistency of outcomes, the seed was set to 42. If, after five successive evaluations, there was no progress in the validation metric, early stopping was employed by setting the early stopping patience to 5. The early stopping threshold, which denotes the minimum significant change in the tracked metric needed for it to be deemed an improvement, was set at 0.01.

4.2 LIWC

LIWC 22 (Boyd et al., 2022) is an advanced text analysis tool that categorizes language into different dimensions, including psychologically meaningful ones, enabling the detection of cognitive, emotional, and social cues within the written content. In our study, we focus on the most context-significant LIWC-derived features to analyze optimistic and pessimistic posts by individuals with depressive symptoms. We quantify these differences using z-scores derived from the Mann–Whitney U test, a nonparametric statistical method that assesses whether one group systematically ranks higher or lower than another on a given variable, being particularly suited for analyzing linguistic features that may not follow a normal distribution. Specifically, we use the test to compare how the linguistic features (as categorized by LIWC) differ between the optimistic and pessimistic posts within the depression and control groups. The z-scores reflect the magnitude of these differences, allowing us to quantify how strongly specific language patterns (such as references to future focus, negative emotions, or social behavior) are associated with either optimistic or pessimistic contexts in each group. We also apply the Benjamini–Hochberg procedure at a nominal alpha=0.05, which orders the p-values and computes adaptive thresholds to control the expected proportion of false discoveries. Features with FDR-adjusted p-values below 0.05 were deemed significant, ensuring that our inferences maintain high sensitivity to true effects while limiting the rate of false positives across all tests.

4.3 Topic Modeling

We implemented a robust topic modeling framework using BERTopic (Grootendorst, 2022) to uncover themes within social media posts, and to explore their associations with sentiment and mental health indicators. Our approach leveraged a customized BERTopic pipeline, which integrates text representation, dimensionality reduction, and clustering techniques.

First, we generated dense text embeddings with SentenceTransformer ('all-MiniLM-L6-v2') and reduced dimensionality using UMAP, opting for a reduced n_neighbors from 15 to 10, while preserving intrinsic data structure. Clustering was achieved with HDBSCAN, with min_cluster_size increased from 10 to 80 (for more robust topic clusters), following text preprocessing with a CountVectorizer configured for bi-grams that included the standard English stopwords, extended with common internet noise words: 'http', 'https', 'amp', 'com', 'www', 'r/'.

To enhance interpretability, topics were refined using a custom representation that leverages KeyBERT (Grootendorst, 2020), combined with Part-of-Speech filtering (via SpaCy's "en_core_web_sm") and Maximal Marginal Relevance (MMR), yielding high-quality, contextually relevant keywords. The final model assigned topics to each post, which were aggregated by sentiment (optimism, neutral, pessimism) and depression status (depressed vs. control). Chi-squared tests of independence were then employed to statistically assess differences in topic distributions across the target groups. The most representative words for each topic can be seen in Appendix A Table 4.

5 Results and Discussions

5.1 Model Performance

We have performed experiments with various other models, including Naïve Bayes, SVM, and CNN, which each yielded the following test accuracies: 62.12%, 65.24%, and 65.59%, respectively. We used GridSearchCV for hyperparameter tuning for both the Naïve Bayes model and the SVM classifier. The CNN model uses a sequential model with embedding, convolutional, global max pooling, and dense layers, as well as dropout for regularization, and is trained over ten epochs to avoid overfitting, with early stopping We have selected the best model based on our experiments, which was the RoBERTa-based one. The reported results can be seen in Appendix A Table 5.

The RoBERTa-OPT-3Labels model shows consistent and competitive performance, with an accuracy of 71.65%, a weighted F1 score of 71.23%, and nearly matching precision and recall values on the test set. The weighted AUC of 0.8452 further underlines its ability to effectively distinguish among the three classes. As this is, to the best of our knowledge, the first work to consider a 3-class approach, it would be interesting to see the results of the state-of-the-art models that interpreted the 1/-1 scenario by eliminating the neutral/ambiguous posts (Caragea et al. (2018); Alshahrani et al. (2020); Alshahrani et al. (2021); Cobeli et al. (2022)). We present selected predicted samples in Table 1.

Our classifier was tested on the constructed gold validation set described in Section 3, where it achieved an overall accuracy of 83%, demonstrating reliable alignment with human annotations. Class-specific F1-scores were uniformly high—0.83 for neutral, 0.83 for optimistic, and 0.85 for pessimistic—indicating balanced performance across categories. Notably, the model exhibits perfect precision for neutral texts and perfect recall for pessimistic ones, while capturing 90% of optimistic instances. Class-specific results can also be seen in the form of the confusion matrix in Figure 1. These results support RoBERTa-OPT-3Labels's suitability for automated sentiment analysis.

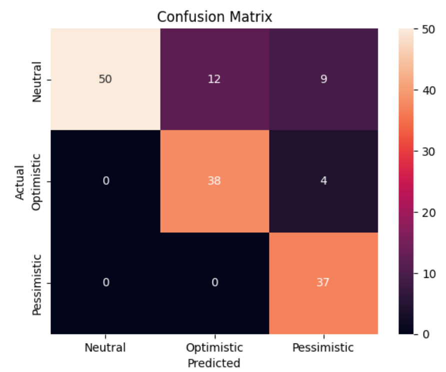

Figure 1: Confusion Matrix for results predicted by RoBERTa-OPT-3Labels on the gold validation data.

5.2 General Statistics Interpretation

After running the predictions for optimism and pessimism using the RoBERTa-OPT-3Labels model, we find that users in the depression group have, on average, fewer optimistic posts than the control group, but a similar number of pessimistic posts. In addition, users in the control group have more posts labeled as neutral. The exact descriptive statistics can be found in Appendix A Tables 6 and 7.

To test for statistical significance, we compare the number of optimistic, pessimistic, and neutral posts between the two groups using Mann–Whitney U test, Cohen's d, and Pearson correlation (Table 2). The Mann–Whitney U test yields non-significant z-scores and p-values for both optimistic (-1.23, p = 0.22) and pessimistic (-0.20, p = 0.84) posts, suggesting that both groups produce similar amounts of content in these categories. In addition, the small effect sizes (Cohen's d = -0.18 for optimism, 0.06 for pessimism) and weak Pearson correlations further support this lack of meaningful distinction.

However, a more significant difference can be seen in the number of neutral posts for the performed tests, with a small to moderate effect size (d = -0.35). This suggests that individuals with depression post significantly fewer neutral statements than people not diagnosed with depression, potentially reflecting a tendency to engage more with emotionally valenced (optimistic or pessimistic) language rather than neutral discourse (Broome et al., 2015).

	Mann–Whitney U test (z, p)	Cohen's d	Pearson Correlation (r, p)
Optimistic	(-1.23, 0.22)	-0.18	(-0.09, 0.26)
Pessimistic	(-0.20, 0.84)	0.06	(0.03, 0.68)
Neutral	(-2.21, 0.03)	-0.35	(-0.17, 0.03)

Table 2: Statistical Test Results for Optimism and Pessimism

While the statistical tests indicate no significant differences in the number of optimistic or pessimistic posts between depression and control individuals, our subsequent analyses will demonstrate that the content of these posts may vary substantially. We will proceed to show that the way optimism and pessimism are expressed in language differs between depressed and non-depressed users in a meaningful way.

5.3 LIWC Analysis Results

Figure 2 presents a side-by-side comparison of statistically significant ($p<0.05$, as measured by the Mann–Whitney U test) LIWC feature usage across optimistic (left panel) and pessimistic (right panel) posts by individuals with and without depression, measured via z-scores. The categories marked by (*) are significant according to the FDR-adjusted p-values. By analyzing these scores, we have outlined several key patterns. In optimistic posts by individuals with depression, increased use of assent and impersonal pronouns suggests a more detached or externally directed expression of optimism, possibly reflecting a coping mechanism. This aligns with research showing that depressed individuals often display reduced self-focus in positive contexts, such as using fewer first-person pronouns when recalling positive memories—an indication of difficulty integrating positive experiences into the self-concept (Himmelstein et al., 2018). In contrast, optimistic posts from control individuals are characterized by greater use of the "family" category, suggesting that their expressions of optimism are more socially anchored and relational. This may reflect a healthier integration of social connectedness and support into positive emotional experiences. When looking at pessimistic posts, the gap between depressed and control users is pronounced. Depressed individuals exhibit a significant increase in words related to negative affect (e.g., general negative emotion and tone, sadness), suggesting a tendency toward more negative and critical thought processes (Mor and Winquist, 2002). The elevated authenticity score suggests that these expressions of pessimism are likely perceived as more honest and self-revealing. Additionally, greater use of cognitive processing (cogproc) words may reflect an effort to make sense of negative experiences, a pattern common in depressive cognition. Depressed individuals also display the presence of more adverbs and a generally higher linguistic score, pointing to increased verbal complexity, possibly signaling ruminative thought patterns. Control users, while also expressing negative content in pessimistic posts, tend to do so with fewer markers of pervasive distress, and the scope of their pessimism seems to orbitate more around external, leisure-related topics. The statistical differences measured with the Mann–Whitney U test and supported by the FDR validation reveal how depressed individuals use their language differently, firstly in comparison to the control group, but also based on the sentiment of the content, with pessimistic posts exhibiting a more pronounced negative linguistic profile.

5.4 Topic Modeling Results

The chi-squared results across the target (depression versus control) groups reveal significant thematic differences in how individuals communicate optimism, pessimism, and neutrality. We will be addressing results for six distinct subgroups, based on the depression label and the optimism/pessimism/neutral associations, with visualizations available in Figure 3 as a heatmap. We present in Table 3 the most overrepresented and underrepresented topics for each target group. In Appendix A Tables 8 and 9, we present the top 10 topics for the depression and the control group, respectively. Also in Appendix A, Figure 4 displays the standardized residuals, calculated from the observed and expected topic frequencies across the three sentiment classes (neutral, optimism, pessimism).

The disparities suggest that psychological states influence topic preferences in online discourse. The pronounced engagement of Depression-Neutral posts in online debate and artificial intelligence contrasts with the avoidance of these topics in Control-Neutral posts, highlighting a potential association between depression and increased argumentative or analytical engagement when not expressing opti-

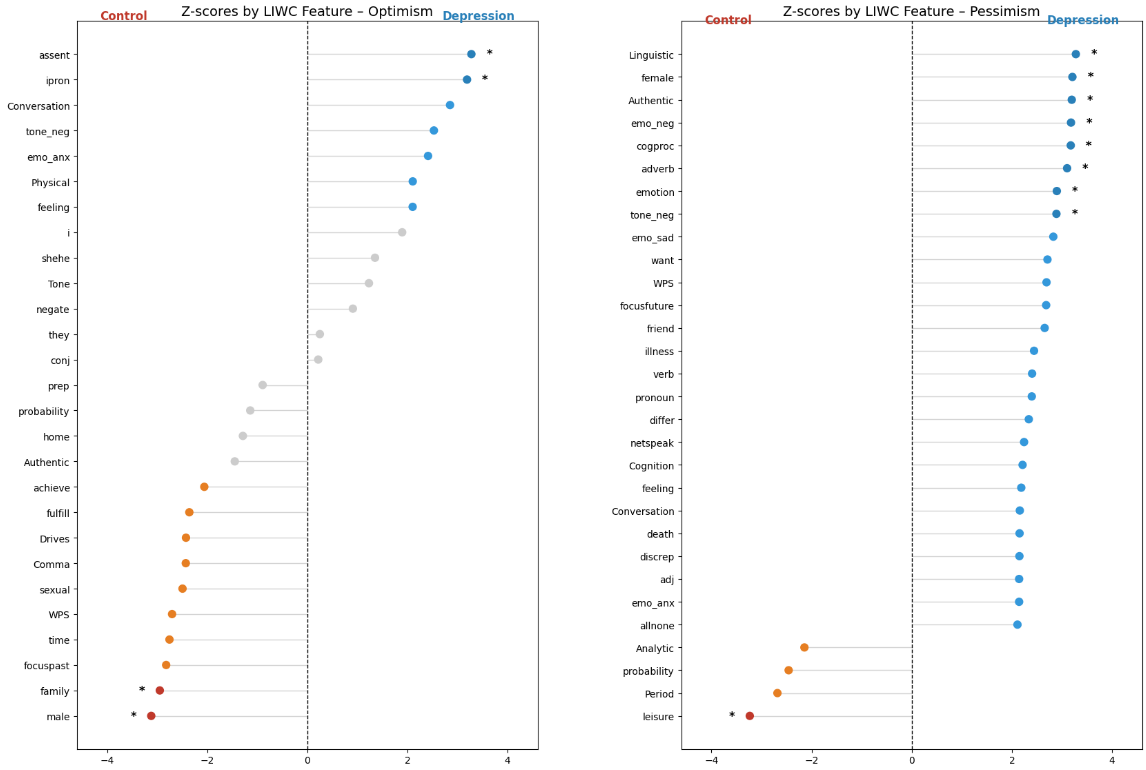

Figure 2: Statistically significant (p<0.05, Mann-Whitney U test) z-scores for the differences between the depression and control groups for posts labeled as optimistic and pessimistic by the RoBERTa model. Results with (*) are statistically significant according to the FDR-adjusted p-values.

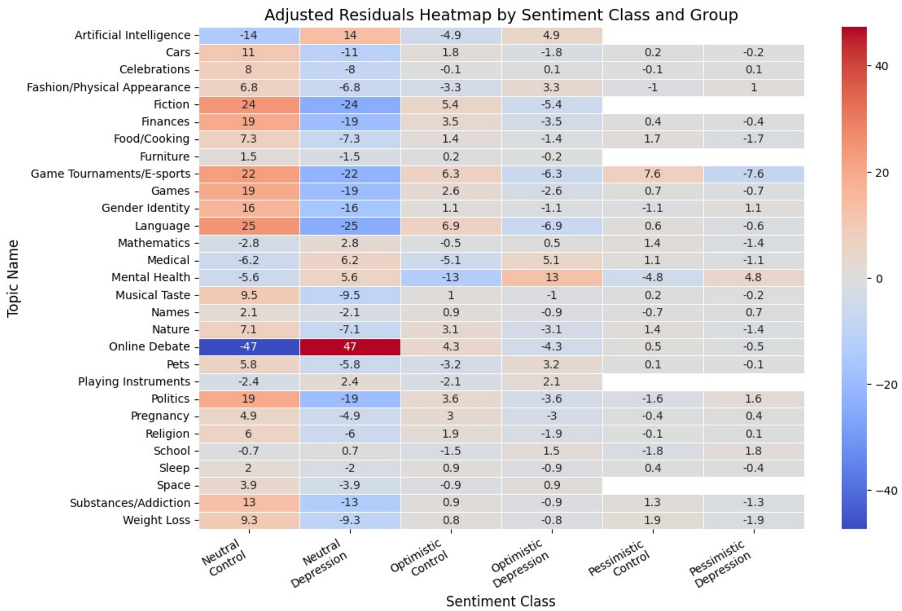

Figure 3: Heatmap of standardized residuals. The colors indicate which topics are significantly overrepresented (red) or underrepresented (blue) in each group.

mism or pessimism. On the other hand, individuals with depression seem overall more comfortable engaging in mental health-related discourse in all sentiment settings, even in optimistic posts. The significant engagement with e-sports in Control-Pessimism posts may indicate a preference for structured, competitive digital interactions in this category, perhaps as a coping mechanism or an outlet for engagement that does not necessitate personal disclosure. The control group also seems

Group	Category	Overrepresented Topics	Underrepresented Topics
Depressed	**Neutral**	Medical, AI, Online Debate	Fiction, Language, E-sports
	Optimism	Mental Health, Medical, AI	Language, E-sports, Fiction
	Pessimism	Mental Health, School, Politics	Online Debate, AI, Pets
Control	**Neutral**	Fiction, Language, E-sports	Online Debate, AI, Medical
	Optimism	Language, E-sports, Fiction	Mental Health, AI, Medical
	Pessimism	E-sports, Weight Loss, Food	Mental Health, School, Politics

Table 3: Top three topic overrepresentation and underrepresentation across depression and control groups

to be engaged in talks about fictional works and general leisure/lifestyle topics, which don't seem as prevalent in the depression group, a theory also supported by literature that suggests reduced engagement in such activities by people diagnosed with depression (Eisemann, 1984). This result is consistent with the observations from the LIWC feature analysis.

To be noted that the missing values seen in the Pessimistic category (both for depression and control groups) were intentionally excluded, as there were no posts of the respective topics belonging to that specific subgroup, thus not being statistically significant.

Coherence scores were also calculated for the generated topics, utilizing gensim (Rehurek and Sojka, 2011), both with the c_v and u_mass formulas. The obtained scores were 0.795 using the c_v formula (where the range is from 0 to 1, values closer to 1 being considered better) and -0.583, respectively, when we used the u_mass score (values around 0 being considered good). Based on the c_v score, our model is generating coherent and interpretable topics, with the u_mass score supporting this view as a secondary interpretation.

5.5 Revisiting Research Questions

Addressing **RQ1**, our analyses reveal that the overall proportions of optimistic and pessimistic posts among individuals with depression are statistically similar to those of the control group. This indicates that, in terms of frequency, individuals in the depression group do not necessarily exhibit a reduced tendency to express optimism compared to control users, though the control group moderately engages more in neutral content. However, while the quantity of such expressions appears consistent, the qualitative content differs markedly.

In response to **RQ2**, our findings indicate that optimism in the social media language of individuals with depression is manifested in a more nuanced and complex manner. Although optimistic posts are present at comparable rates, the linguistic features and thematic content of these posts suggest a distinct expression of optimism that is intertwined with elements of resilience and coping. Specifically, while their optimistic posts are not marked by a significant negative tone, a more detached or externally directed expression of optimism, potentially reflecting a distancing strategy from personal agency, can be seen. Notably, even within contexts that are ostensibly positive, individuals with depression demonstrate less engagement with cultural, lifestyle, and leisure topics, maintaining a great focus on mental health discussions.

6 Conclusions and Future Work

Our study investigated the expressions of optimism and pessimism in the social media discourse of individuals with depression using computational methods. Although no significant differences were observed in the actual amounts of optimistic versus pessimistic posts between the depression and control groups, our analyses revealed meaningful differences in the linguistic content and thematic focus of these posts. Notably, while pessimistic posts from individuals with depression exhibited a pronounced negative linguistic profile, the expressions of optimism—though subtler—appear to represent a complex interplay of resilience and coping mechanisms. These findings might provide insights into adaptive strategies within this target group. Overall, our results not only corroborate existing psychological theories regarding language, psychological and depressive states, but also highlight the potential of transformer-based models, topic modeling and LIWC features in capturing nuanced variations in online discourse related to mental health.

Subsequent investigations may benefit from a longitudinal approach to examine how expressions of optimism and pessimism evolve over time in relation to depressive symptoms. Additionally, integrating multimodal data—such as images, user interactions, and metadata—may provide a more comprehensive understanding of online expressions of optimism and pessimism.

Limitations

In our experiments, we used the OPT dataset obtained from Twitter/X to train a transformer-based model for predicting optimism and pessimism labels in depression-related content sourced from Reddit. This choice was made due to the limited availability of datasets from the same domain. The OPT dataset is the most commonly used dataset for this specific task (Caragea et al., 2018; Cobeli et al., 2022). Additionally, we selected the eRisk 2021 dataset because it includes social media users who have completed the validated BDI-II questionnaire, which provides a reliable assessment of depression. Prior research suggests that transformer-based models are effective for transfer learning across different platforms (Uban et al., 2022), and we have validated this statement through the construction of our manually annotated gold standard subset, as well as conducting statistical tests on our results.

Ethical Considerations

This paper uses OPT, a publicly available dataset with annotations for optimism and pessimism. In addition, the eRisk 2021 dataset was made available to us after signing a data usage agreement form. We have adhered to the data agreement, and we did not make any attempt to contact the users or to de-anonymize the data. The sample of posts presented in this paper has been paraphrased to ensure the anonymity of the users. Annotation of a small subset of data was done consensually, with full annotator anonymity. Our primary focus is on quantifying and analyzing optimistic and pessimistic sentiments within the texts of the mental health dataset. We do not aim to predict mental health status or conditions based on this dataset.

References

A. Alshahrani, M. Ghaffari, K. Amirizirtol, and X. Liu. 2020. Identifying optimism and pessimism in twitter messages using xlnet and deep consensus. In *International Joint Conference on Neural Networks (IJCNN)*.

A. Alshahrani, M. Ghaffari, K. Amirizirtol, and X. Liu. 2021. Optimism/pessimism prediction of twitter messages and users using bert with soft label assignment. In *International Joint Conference on Neural Networks (IJCNN)*.

Mostafa M. Amin, E. Cambria, Björn Schuller, and E. Cambria. 2023. Will affective computing emerge from foundation models and general artificial intelligence? a first evaluation of chatgpt. *IEEE Intelligent Systems*, 38:15–23.

Mario Ezra Aragon, Adrian Pastor Lopez-Monroy, Luis Carlos González-Gurrola, and Manuel Montes-y Gómez. 2021. Detecting mental disorders in social media through emotional patterns-the case of anorexia and depression. *IEEE transactions on affective computing*, 14(1):211–222.

Fabio Barbieri, José Camacho-Collados, Lisa Ellen Anke, and Leandro Neves. 2020. Tweeteval: Unified benchmark and comparative evaluation for tweet classification. In *Findings of the Association for Computational Linguistics: EMNLP 2020*. Association for Computational Linguistics.

Aaron T Beck, Robert A Steer, Gregory K Brown, et al. 1996. Beck depression inventory.

G. Blanco and A. Lourenço. 2022. Optimism and pessimism analysis using deep learning on covid-19 related twitter conversations. *Information Processing and Management*, 59(3):102918.

Ryan L Boyd, Ashwini Ashokkumar, Sarah Seraj, and James W Pennebaker. 2022. The development and psychometric properties of liwc-22. *Austin, TX: University of Texas at Austin*, 10.

M. R. Broome, K. E. A. Saunders, P. J. Harrison, and S. Marwaha. 2015. Mood instability: Significance, definition and measurement. *British Journal of Psychiatry*, 207(4):283–285.

A. Bucur, M. Zampieri, and L. P. Dinu. 2021. An exploratory analysis of the relation between offensive language and mental health. In *Findings of the Association for Computational Linguistics: ACL-IJCNLP 2021*, pages 3600–3606.

Ana-Maria Bucur, Berta Chulvi, Adrian Cosma, and Paolo Rosso. 2024. The expression of happiness in social media of individuals reporting depression. *IEEE Transactions on Affective Computing*.

J. Camacho-Collados, K. Rezaee, T. Riahi, A. Ushio, D. Loureiro, D. Antypas, J. Boisson, L. E. Anke, F. Liu, and E. M. Camara. 2022. Tweetnlp: Cutting-edge natural language processing for social media. In *Proceedings of the 2022 Conference on Empirical Methods in Natural Language Processing: System Demonstrations*.

C. Caragea, L. P. Dinu, and B. Dumitru. 2018. Exploring optimism and pessimism in twitter using deep learning. In *Proceedings of the Conference on Empirical Methods in Natural Language Processing (EMNLP)*.

Lushi Chen, Walid Magdy, Heather Whalley, and Maria Klara Wolters. 2020. Examining the role of mood patterns in predicting self-reported depressive symptoms. In *Proceedings of the 12th ACM Conference on Web Science*, WebSci '20, page 164–173, New York, NY, USA. Association for Computing Machinery.

Ş. Cobeli, I. Iordache, S. Yadav, C. Caragea, L.P. Dinu, and D. Iliescu. 2022. Detecting optimism in tweets using knowledge distillation and linguistic analysis of optimism. In *International Conference on Language Resources and Evaluation*.

Munmun De Choudhury, Michael Gamon, Scott Counts, and Eric Horvitz. 2013. Predicting depression via social media. 7(1):128–137.

A. De Santana Correia and E. L. Colombini. 2022. Attention, please! a survey of neural attention models in deep learning. *Artificial Intelligence Review*, 55(8):6037–6124.

M. Eisemann. 1984. Leisure activities of depressive patients. *Acta Psychiatrica Scandinavica*, 69(1):45–51.

L. Gan, Y. Guo, and T. Yang. 2024. Machine learning for depression detection on web and social media. *International Journal on Semantic Web and Information Systems*, 20(1):1–28.

Maarten Grootendorst. 2020. KeyBERT: Minimal keyword extraction with BERT. https://doi.org/10.5281/zenodo.4461265.

Maarten Grootendorst. 2022. Bertopic: Neural topic modeling with a class-based tf-idf procedure.

Yuting Guo, Yao Ge, Yuan-Chi Yang, Mohammed Ali Al-Garadi, and Abeed Sarker. 2022. Comparison of pretraining models and strategies for health-related social media text classification. In *Healthcare*, volume 10, page 1478. MDPI.

U. Herwig, A. B. Brühl, T. Kaffenberger, T. Baumgärtner, H. Boeker, and L. Jäncke. 2009. Neural correlates of 'pessimistic' attitude in depression. *Psychological Medicine*, 40(5):789–800.

P. Himmelstein, S. Barb, M. A. Finlayson, and K. D. Young. 2018. Linguistic analysis of the autobiographical memories of individuals with major depressive disorder. *PloS one*, 13(11):e0207814.

Catherine Hobbs, Petra Vozarova, Aarushi Sabharwal, Punit Shah, and Katherine Button. 2022. Is depression associated with reduced optimistic belief updating? *Royal Society Open Science*, 9(2):190814.

Jutta Karhu, Juha Veijola, and Mirka Hintsanen. 2024. The bidirectional relationships of optimism and pessimism with depressive symptoms in adulthood–a 15-year follow-up study from northern finland birth cohorts. *Journal of Affective Disorders*, 362:468–476.

Christoph W Korn, Tali Sharot, Hendrik Walter, Hauke R Heekeren, and Raymond J Dolan. 2014. Depression is related to an absence of optimistically biased belief updating about future life events. *Psychological medicine*, 44(3):579–592.

Kim Kronström, Hasse Karlsson, Hermann Nabi, Tuula Oksanen, Paula Salo, Noora Sjösten, Marianna Virtanen, Jaana Pentti, Mika Kivimäki, and Jussi Vahtera. 2011. Optimism and pessimism as predictors of work disability with a diagnosis of depression: a prospective cohort study of onset and recovery. *Journal of affective disorders*, 130(1-2):294–299.

Grace Y Lim, Wilson W Tam, Yanxia Lu, Cyrus S Ho, Melvyn W Zhang, and Roger C Ho. 2018. Prevalence of depression in the community from 30 countries between 1994 and 2014. *Scientific reports*, 8(1):2861.

Y. Liu, M. Ott, N. Goyal, J. Du, M. Joshi, D. Chen, O. Levy, M. Lewis, L. Zettlemoyer, and V. Stoyanov. 2019. Roberta: A robustly optimized bert pretraining approach. *OpenReview*.

D. E. Losada and F. Crestani. 2016. A test collection for research on depression and language use. In *Lecture notes in computer science*, pages 28–39.

N. Mor and J. Winquist. 2002. Self-focused attention and negative affect: A meta-analysis. *Psychological Bulletin*, 128(4):638–662.

Javier Parapar, Patricia Martín-Rodilla, David E Losada, and Fabio Crestani. 2021. Overview of erisk at clef 2021: Early risk prediction on the internet (extended overview). *CLEF (Working Notes)*, 1:864–887.

Radim Rehurek and Petr Sojka. 2011. Gensim–python framework for vector space modelling. *NLP Centre, Faculty of Informatics, Masaryk University, Brno, Czech Republic*, 3(2).

Jonathan Rottenberg. 2005. Mood and emotion in major depression. *Current Directions in Psychological Science*, 14(3):167–170.

X. Ruan, S. R. Wilson, and R. Mihalcea. 2016. Finding optimists and pessimists on twitter. In *Proceedings of the 54th Annual Meeting of the Association for Computational Linguistics (Volume 2: Short Papers)*.

Michael F Scheier, Charles S Carver, and Michael W Bridges. 2001. Optimism, pessimism, and psychological well-being. *Optimism & pessimism: Implications for theory, research, and practice.*, pages 189–216.

Andrew M Subica, J Christopher Fowler, Jon D Elhai, B Christopher Frueh, Carla Sharp, Erin L Kelly, and Jon G Allen. 2014. Factor structure and diagnostic validity of the beck depression inventory–ii with adult clinical inpatients: Comparison to a gold-standard diagnostic interview. *Psychological assessment*, 26(4):1106.

Hilary Tindle, Bea Herbeck Belnap, Patricia R Houck, Sati Mazumdar, Michael F Scheier, Karen A Matthews, Fanyin He, and Bruce L Rollman. 2012. Optimism, response to treatment of depression, and rehospitalization after coronary artery bypass graft surgery. *Psychosomatic medicine*, 74(2):200–207.

Ana-Sabina Uban, Berta Chulvi, and Paolo Rosso. 2021. An emotion and cognitive based analysis of mental health disorders from social media data. *Future Generation Computer Systems*, 124:480–494.

Ana Sabina Uban, Berta Chulvi, and Paolo Rosso. 2022. Multi-aspect transfer learning for detecting low resource mental disorders on social media. In *Proceedings of the Thirteenth Language Resources and Evaluation Conference*, pages 3202–3219.

Michael von Glischinski, Ruth von Brachel, and Gerrit Hirschfeld. 2019. How depressed is "depressed"? a systematic review and diagnostic meta-analysis of optimal cut points for the beck depression inventory revised (bdi-ii). *Quality of Life Research*, 28:1111–1118.

Dong Xu, Yi-Lun Wang, Kun-Tang Wang, Yue Wang, Xin-Ran Dong, Jie Tang, and Yuan-Lu Cui. 2021. A scientometrics analysis and visualization of depressive disorder. *Current neuropharmacology*, 19(6):766–786.

A Appendix

A.1 Characteristics Of Annotators

There were three annotators that rated the optimism/pessimism labels for the gold Reddit subset, two female and one male. All annotators have a minimum C1 English language level and were given verbal basic instructions, following the same labeling setup as Ruan et al. (2016). It was explained that their identities remained anonymous, as only the ratings were of interest in this study.

Topic	**Representative Words**
Online Debate	changemyview, appeal, wiki, removed, moderation, rules, comments see, ban, review, discussion
Mental Health	therapy, depression, psychiatrist, self harm, list, mental health, hope something, suicidal, trauma
Games	games, pc, gaming, minecraft, tanks, sims, screen, ios, deck, mobile
Politics	communism, capitalism, ideology, communist, fascist, hk, ideologies, oppose, political, protests
Musical Taste	song, songs, bastille, lyrics, lost, band
Pets	dogs, meat, animals, animal, service, cats, trained, environment, ethical, pets
Gender Identity	gender, men, sexuality, bisexual, queer, lgbtq, feminine, dysphoria, transgender, lesbian
Celebrations	birthday, christmas, children, celebrate, parents, santa, teenagers
Fashion/Physical Appearance	wear, skin, jeans, acne, dress, makeup, outfit, masks, shirts, curl
Language	english, spain, spanish, dutch, languages, translated, native language, speak english, hebrew, fluent
Food/Cooking	pizza, eggs, recipe, rice, ice cream, breakfast, recipes, chips, peanut butter
Game Tournaments/E-sports	twitter, smash, tournament, losers, league, players, baseball, bracket, teams, civil war
Substances/Addiction	weed, drink, alcohol, drugs, smoked, alcoholic, toilet, wash, heroin, spiritual
Nature	flag, blue, flowers, colors, lights, purple, pink, weather, trees, autumn
Religion	god, bible, philosophy, christianity, religion, belief, faith, exist, atheist, universe
Medical	pharmacy, hospital, residency, covid, clinical, med, health care, pandemic, script, jobs
Names	names, last name, named, first name, pronounced, like name, husband, japanese, change name, gender
Pregnancy	pregnancy, birth control, periods, pcos, symptoms, bleeding, knee, pills, trimester, shoulders
Weight Loss	calories, eat, diet, keto, insulin, fat, carbs, weight loss, workout, foods
Cars	cars, engines, f1, driving, miles, germany, roads, rust, wheel, driven
Fiction	agata, menem, empire, gods, rebellion, union, wars, ancient, river, folk
Mathematics	equation, 3x, formula, derivative, sin, values, solve, slope, graph, triangle
Sleep	sleep, dreams, sleeping, bed, nap, waking, slept, every night, fall asleep, nightmares
Finances	ira, taxes, income, owe, loan, retirement, debt, credit card, monthly, file
Artificial Intelligence	ai, intelligence, artificial, intelligent, robot, machines, neural, cognitive, technology, agent
Playing Instruments	guitar, frequency, tune, pitch, instrument, mic, tones, speakers, modes, barrel
School	high school, schools, transfer, gpa, berkeley, graduation, applied, grades, email, colleges
Furniture	sofa, furniture, ikea, couch, drawer, desk, living room, curtains, cabinets, pillows
Space	black hole, planets, solar, radius, space, gravity, lens, moons, stars, galaxy

Table 4: Topics and their Representative Words

Model	Val. Acc.	Test Acc.
Naïve Bayes	63.42	62.12
SVM	67.52	65.24
CNN	64.67	65.59
RoBERTaOPT3Labels	71.54	71.65

Table 5: Baselines and comparison with RoBERTa-OPT-3Labels

	Neutral	Optimistic	Pessimistic	Total
mean	303.59	65.52	12.76	381.88
std	325.32	76.75	16.81	391.66
min	9.00	2.00	0.00	16.00
25%	47.00	14.00	1.00	66.00
50%	150.00	33.00	5.00	199.00
75%	590.00	88.00	18.00	702.00
max	1132.00	416.00	81.00	1208.00

Table 6: Descriptive Statistics for Depression Group

	Neutral	Optimistic	Pessimistic	Total
mean	422.40	79.33	11.68	513.42
std	362.08	75.55	15.79	428.46
min	21.00	2.00	0.00	26.00
25%	57.00	14.75	2.00	66.00
50%	317.50	64.50	6.00	396.00
75%	784.25	117.50	14.00	969.50
max	1258.00	334.00	92.00	1478.00

Table 7: Descriptive Statistics for Control Group

Rank	Topic	Count
1	Online Debate	8,923
2	Mental Health	1,992
3	Politics	1,127
4	Language	822
5	Game Tournaments/E-sports	733
6	Games	590
7	Gender Identity	574
8	Musical Taste	567
9	Pets	525
10	Substances/Addiction	393

Table 8: Control Group – Top 10 Topics by Count

Rank	Topic	Count
1	Online Debate	27,241
2	Mental Health	5,430
3	Politics	1,174
4	Pets	903
5	Musical Taste	722
6	Fashion/Physical Appearance	506
7	Gender Identity	506
8	Artificial Intelligence	499
9	Language	450
10	Game Tournaments/E-sports	402

Table 9: Depressed Group – Top 10 Topics By Count

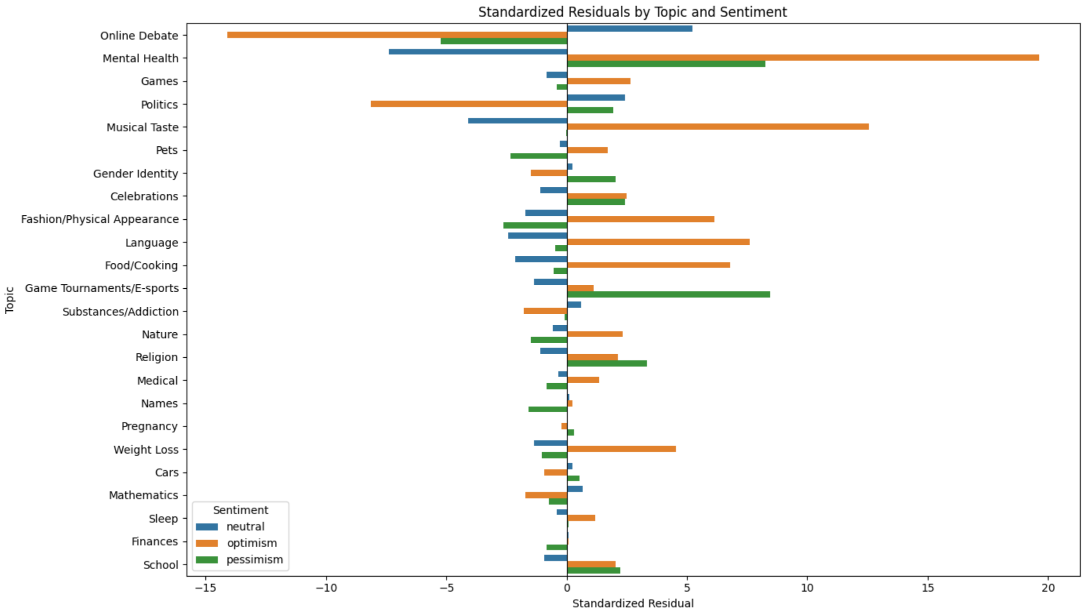

Figure 4: Standardized residuals - observed and expected topic frequencies across the three sentiment classes (neutral, optimism, pessimism).