

Identifying Depressive Content in Decentralized Social Networks: A Case Study on Bluesky

Alicia Marzola Chaves¹, Gabriel Lemos¹, Gabriel Rolla Ferreira¹,
Lucas Pimenta Braga¹, Wagner Meira Jr.¹, Gisele L. Pappa¹

¹Computer Science Department
Federal University of Minas Gerais (UFMG) – Belo Horizonte, MG

{aliciachaves, gabrielemos, gabrielrolla, lucaspimenta
meira, glpappa}@dcc.ufmg.br

Abstract. *Depression is a global mental health concern, yet scalable screening remains challenging. While prior work focuses on centralized platforms, little is known about depressive signals in newer networks like Bluesky. We present a manually annotated dataset of 4,898 public English Bluesky posts (14.5% labeled as depressive). Evaluating a hybrid pipeline combining XGBoost and fine-tuned transformers (BERT-base, RoBERTa-base, and MentalBERT), our best model achieves an F1-score of 80.5% on Bluesky and 90.8% on Twitter. This highlights that core linguistic cues remain informative across decentralized platforms. The proposed system is intended for decision support rather than diagnosis, requiring clinical validation and ethical oversight.*

1. Introduction

Depressive disorder, or depression, affects approximately 5% of adults worldwide and remains one of the leading causes of disability [Organization 2023]. Beyond mood fluctuations, it can severely impair daily functioning, relationships, and overall quality of life [Organization 2023, Bains and Abdijadid 2023]. Despite its prevalence, the diagnosis of Major Depressive Disorder (MDD) still relies heavily on subjective self-reports and clinical observation. Field trials associated with the DSM-5 report low inter-rater reliability (with Cohen’s kappa values as low as 0.28), exposing inconsistencies in clinical assessment and underscoring the need for more objective and scalable decision-support tools [Regier et al. 2013, Lieblich et al. 2015].

Social media has emerged as a promising source of behavioral signals for early depression detection at scale. Prior studies, particularly on Twitter, have linked linguistic markers, such as first-person singular pronouns, negative emotion words, and cognitive rigidity, to depressive symptoms [De Choudhury et al. 2013b, Coppersmith et al. 2014]. Large-scale shared tasks and evaluation initiatives, such as CLEF eRisk, have further advanced early risk prediction by providing benchmark datasets and standardised experimental protocols for mental health-related NLP [Losada et al. 2019]. However, most evidence still concentrates on centralized platforms, leaving open questions about how these methods translate to newer environments with different interaction dynamics.

From a clinical and public-health perspective, these signals are not meant to replace diagnosis. Instead, they can support mental health professionals by prioritising cases for follow-up, monitoring changes in expressed risk over time, and informing outreach strategies when traditional screening is limited. In practice, such systems can be

integrated as triage or alert mechanisms that surface potentially concerning content for human review, enabling earlier engagement while preserving clinician oversight, patient safety, and ethical safeguards.

Bluesky, a decentralised social network launched in 2024, introduces a federated architecture, an open protocol ecosystem, and community-based moderation, all of which may influence how users communicate and engage online [Silberling 2024, Matei 2025]. Its user base also skews young, with the 18–24 age group representing a prominent demographic segment [Topics 2025]. Despite its rapid growth, there is still limited empirical evidence on mental health expression in this ecosystem, and to our knowledge, no prior work has explored depression detection in Bluesky.

This study addresses this gap by examining depression-related language in an emerging platform whose technical design and community norms may shape online expression. Rather than assuming that decentralization alone determines user behavior, we investigate whether Bluesky’s sociotechnical context, including its governance model, moderation style, and audience profile, is associated with differences in how depressive signals appear compared to a large centralized platform such as Twitter. Accordingly, we hypothesise that:

Bluesky’s platform context is associated with measurable differences in the linguistic expression of depressive tendencies compared to Twitter.

To test this hypothesis, we curated and manually annotated a dataset of Bluesky posts guided by a clinical psychologist. Using a lexicon grounded in clinical frameworks such as the DSM-5 and PHQ-9 [Association 2013, Spitzer et al. 2001], we captured both explicit and implicit depressive signals. We then proposed a hybrid detection pipeline combining: (i) a metadata-based classifier using XGBoost, and (ii) fine-tuned transformer architectures (BERT-base, RoBERTa-base, and MentalBERT). Unlike prior work, we integrated engagement metadata and conducted ablation studies to assess its predictive value. We further evaluated model generalization through external validation on a Twitter dataset.

Our main contributions are: (i) the first manually annotated Bluesky dataset for depressive content detection, created with support from clinical psychologists; (ii) a comparative analysis of linguistic and engagement patterns in a decentralized social platform; (iii) a hybrid detection pipeline combining transformer-based models and metadata through stacked ensembling; and (iv) an external validation on Twitter, showing both generalization capacity and platform-specific differences.

2. Related Work

Over the past decade, depression detection from social media has progressed from traditional feature-based classifiers to deep learning approaches, driven by larger datasets, stronger representation learning, and improved evaluation practices [Liu et al. 2022, Cao et al. 2025].

2.1. Traditional Machine Learning Approaches

Early work relied on handcrafted features combined with supervised algorithms. Ahmed et al. applied sentiment features with Naïve Bayes and NBTree on Twit-

ter data, reporting high accuracy but with strong dependence on TextBlob polarity scores [Ahmed et al. 2025]. D’Cruz et al. evaluated Naïve Bayes, Random Forest, and SVMs on Reddit and Twitter, achieving balanced precision and recall [D’Cruz et al. 2023]. Despite their simplicity and interpretability, these models often struggle with informal language, figurative expressions, and sarcasm, limiting robustness in real-world social media settings.

2.2. Transformer-Based Architectures

Transformer models such as BERT and RoBERTa have reshaped depression detection by capturing contextual and semantic cues beyond surface lexical features [Bokolo and Liu 2024]. MentalBERT, pre-trained on mental health forums, can improve performance in domain-specific corpora [Ji et al. 2021], but may underperform on general social media due to domain and style mismatch. Newer architectures such as DeBERTa incorporate disentangled attention and have shown strong general NLP performance, yet remain less explored in mental health tasks [He et al. 2021]. Ensemble strategies that combine multiple transformer signals have also been reported to improve robustness and predictive power [Tavchioski et al. 2023].

2.3. Metadata-Enhanced and Multimodal Models

Recent studies augment text representations with auxiliary signals such as posting time, emoji usage, and engagement metrics (e.g., likes and replies), often improving performance by capturing behavioral context [Kerasiotis et al. 2024]. However, metadata can encode platform-specific interaction dynamics and other confounds (e.g., audience size or moderation effects), requiring careful interpretation and bias-aware evaluation [Cao et al. 2025].

2.4. Datasets and Annotation Practices

Most work relies on datasets from Twitter, Reddit, DAIC-WOZ, or the CLEF eRisk initiative. Labels are frequently obtained via distant supervision (keywords, self-disclosure, or community membership), which can introduce noise and demographic bias. Expert-led manual annotation remains comparatively rare but is important for capturing subtler cues such as metaphor, temporality, and indirect self-disclosure. Moreover, many datasets omit inter-annotator agreement, limiting reproducibility and comparability across studies [Grabb et al. 2024].

2.5. Recent Advances and Algorithmic Bias

Recent directions include explainability and broader label spaces (e.g., multi-class mental health categorization). Cha et al. explored lexicon-driven approaches for Korean social media, highlighting opportunities for broader applications in mental health [Cha et al. 2022]. Hasan et al. compared transformers and LSTMs on Reddit, reporting high F1-scores under controlled experimental settings [Hasan and Kumar 2025]. At the same time, reviews continue to highlight persistent challenges, including sampling bias, inconsistent preprocessing, class imbalance, and unclear reporting, which can inflate results and hinder reproducibility [Cao et al. 2025].

2.6. Summary of Gaps Addressed

Despite substantial progress on centralized platforms such as Twitter and Reddit, depression detection has not been systematically studied in decentralized social networks. Platforms like Bluesky introduce different governance and moderation structures, as well as distinct audience profiles, which may shape how mental health-related language is produced and engaged with. Our work addresses this gap by analyzing depressive expression in Bluesky and by evaluating whether linguistic markers transfer across platforms with divergent sociotechnical contexts.

3. Dataset Characterization

3.1. Data Collection and Ethics

We collected public posts from Bluesky using Apify’s scraping tool, restricting retrieval to publicly available content to align with ethical standards in digital mental health research [Conway and O’Connor 2016]. During preprocessing, we removed personally identifying information (e.g., usernames and URLs) and retained only anonymized text and aggregated metadata.

Nonetheless, complete anonymization of linguistic data remains challenging due to potential re-identification through unique writing patterns and stylistic cues [Zimmer 2010]. To mitigate this risk, we adopted privacy-preserving practices, including data minimization and restricted sharing to anonymized and/or aggregated representations, thereby reducing potential harm to individuals.

Initial post retrieval was guided by a lexicon of depression-related expressions (Table 1), informed by clinical instruments such as the PHQ-9 [Spitzer et al. 2001] and the DSM-5, as well as linguistic patterns described in prior work (e.g., DEPTWEET) [Kabir et al. 2022]. This hybrid approach aimed to balance clinical relevance with coverage of informal expressions common in social media. To avoid unnecessary exposure to highly sensitive content in the paper, we report representative terms rather than providing explicit examples.

Category	Words/Expressions
General depressive cues	depress, alone, stress, desperate, gonna cry, solitude, emptiness, hate myself, feel useless, helpless, tired of everything, have no future, deep sadness, meaningless, dead inside, mentally exhausted, guilt, fear, emotional fatigue
High-severity cues	hopeless, worthless, no way out, no one needs me, I’m a burden, I wish I could disappear, I give up, end of the line

Table 1. Lexicon of depression-related terms used for initial data retrieval.

3.2. Manual Annotation

The research team, in conjunction with a clinical psychologist, independently annotated posts using a binary labeling scheme (*depressive* vs. *non-depressive*). Labels indicate the presence of depression-related linguistic cues in the post content (risk signals) and do not

represent a clinical diagnosis. A calibration phase involved jointly reviewing 200 posts to align labeling criteria, resolve ambiguities, and standardize edge cases such as irony, metaphor, and third-person references. Disagreements were resolved through discussion until consensus.

Inter-annotator agreement was calculated on a random sample of 100 posts, yielding Cohen’s Kappa of 0.81, indicating substantial agreement. For comparison, Atapattu et al. [Atapattu et al. 2022] report a Fleiss’s Kappa of 0.82 in a similar mental health corpus.

3.3. Dataset Statistics

Following preprocessing, we excluded posts with ambiguous intent (e.g., irony/metaphor without clear self-disclosure) to reduce label noise and improve interpretability. The final dataset includes **4,568** public English-language posts, with **514** (11.2%) labeled as depressive.

4. Methodology

This section presents the full pipeline for detecting depressive content in Bluesky posts, with emphasis on transparency, reproducibility, and alignment with best practices in NLP and digital mental health research. The methodology comprises five stages: (i) preprocessing and feature engineering, (ii) metadata-based baseline, (iii) transformer-based models, (iv) stacked ensemble learning, and (v) evaluation and error analysis.

4.1. Preprocessing and Feature Engineering

The anonymized dataset was loaded from CSV format and underwent a series of preprocessing steps inspired by prior work on social media-based depression detection [Guntuku et al. 2017, De Choudhury et al. 2013b]:

- Text entries were cast to string, and posts containing only whitespace or null values were removed.
- Emojis were removed during preprocessing, which may limit the model’s sensitivity to affective signals often conveyed through symbolic expression.
- Platform-specific artifacts such as URLs, mentions (@), and hashtags were removed.
- Contractions were expanded (e.g., “I’ll” → “I will”), glued terms separated (e.g., `mentalhealth` → `mental health`), and common abbreviations normalized (e.g., “kms” → “kill myself”).
- Tokenization and lowercasing were consistently applied across all models.

These steps follow prior work and were validated in preliminary experiments.

4.2. Metadata-Based Baseline: XGBoost

Given the predictive potential of user engagement features [Chancellor et al. 2016], we trained an XGBoost classifier using only the eight meta-features listed in Table 2. To mitigate class imbalance (14.5% depressive posts), random oversampling was applied to the training set only.

Hyperparameters were tuned via grid search based on validation performance. The selected configuration included a learning rate of 0.05, 200 estimators, maximum tree depth of 4, and a subsampling rate of 0.8.

Repost Count	Like Count	Reply Count	Hour of Posting
Month	Has Image?	Has Link?	Has Hashtag?

Table 2. Metafeatures used in the XGBoost baseline.

4.3. Transformer-Based Language Models

For text-based classification, we fine-tuned three Transformer models using the Hugging Face Transformers library, leveraging pre-trained weights and adapting them to our dataset. A classification head was added to each model and trained using cross-entropy loss on binary labels. All models were trained under identical experimental settings to enable fair comparison.

- **BERT-base uncased** [Devlin et al. 2018], a widely used general-purpose model for contextual embeddings.
- **MentalBERT** [Ji et al. 2021], pre-trained on mental health forums, tailored for domain-specific nuance.
- **RoBERTa-base** [Liu et al. 2019], an optimized BERT variant with superior training strategies.

To handle class imbalance, oversampling was applied only to the training set. Table 3 summarizes hyperparameter choices.

Parameter	Value
Optimizer	AdamW
Learning Rate	2×10^{-5}
Batch Size	8
Epochs	5 (early stopping on validation loss)

Table 3. Hyperparameters used for transformer fine-tuning.

4.4. Stacked Ensemble

To combine the strengths of different models, we implemented a stacked ensemble, described in Algorithm 1. Stacking has been shown to improve robustness in depression detection tasks by integrating heterogeneous signals [Tadesse et al. 2019]. This approach leverages the idea that different models capture complementary patterns in the data. While tree-based models (e.g., XGBoost) are effective at exploiting structured meta-features, transformer-based models capture contextual and semantic nuances in text.

In our setup, predictions from XGBoost, BERT, MentalBERT, and RoBERTa are combined into a low-dimensional feature vector, which is used as input to a Logistic Regression meta-learner. By learning how to optimally combine these predictions, the meta-learner can correct individual model biases and improve generalization.

Algorithm 1 Model Stacking Procedure

- 1: Train the baseline model using XGBoost with eight meta-features (Table 2).
 - 2: Independently train three transformers models: BERT, MentalBERT, and RoBERTa.
 - 3: Build a 4-dimensional feature vector using the predicted probability from each base model (XGBoost + the three transformers).
 - 4: **Meta-learner training:** Train a Logistic Regression meta-classifier on this new feature vector.
 - 5: Obtain predictions from the base models.
 - 6: Aggregate their outputs through the meta-learner to produce the final prediction.
-

4.5. Evaluation Metrics and Error Analysis

We split the dataset into 80% training, 10% validation, and 10% test sets using a stratified split to preserve the class distribution. Hyperparameters were selected on the validation set, and all final results are reported on the held-out test set. Random oversampling (when used) was applied only to the training split.

All models were evaluated on a held-out test set using standard classification metrics: accuracy, precision, recall, F1-score, and Area Under the Receiver Operating Characteristic Curve (AUROC). Given the inherent class imbalance in our dataset, we placed particular emphasis on metrics that reflect performance on the minority class.

5. Results

5.1. Engagement, Temporal, and Linguistic Patterns

Engagement Analysis. Figure 1 shows that over **60% of depressive posts received zero likes**, while non-depressive posts attracted higher engagement. Prior work suggests that reduced interaction may be consistent with social withdrawal [De Choudhury et al. 2013a] and stigma-related inhibition of emotional expression [Rai et al. 2024], although such associations should be interpreted cautiously given platform-specific dynamics.

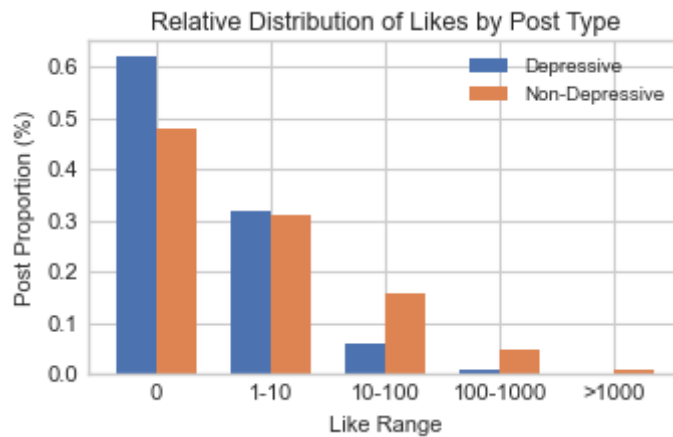


Figure 1. Distribution of like counts for depressive vs. non-depressive posts.

Temporal Analysis. Figure 2 shows that posting activity peaks between **11 PM and 1 AM** (user-local time). This late-evening concentration suggests that a portion of users

5.3. Language Models Evaluation

To explore the advantages of contextualized language understanding, we fine-tuned three transformer-based models: **BERT**, **MentalBERT**, and **RoBERTa**, under identical experimental conditions. Their performance was evaluated against a **stacked ensemble** that combined both text-based and metadata-based signals.

Table 4 presents the evaluation results across standard metrics, with the best-performing values per column highlighted in bold. Overall, transformer models showed substantial improvements in recall and F1-score, particularly when combined with oversampling and ensemble strategies.

<i>With oversampling</i>				
Models	Acc.	Prec.	Rec.	F1
BERT	0.946	0.830	0.775	0.802
MentalBERT	0.915	0.643	0.867	0.738
RoBERTa	0.935	0.757	0.790	0.773
Stacking	0.945	0.821	0.791	0.805

<i>Without oversampling</i>				
Models	Acc.	Prec.	Rec.	F1
BERT	0.937	0.839	0.681	0.752
MentalBERT	0.923	0.754	0.659	0.704
RoBERTa	0.924	0.668	0.920	0.774
Stacking	0.936	0.812	0.712	0.759

Table 4. Performance of transformer models evaluated with and without oversampling. Best scores per column are highlighted.

With oversampling, BERT achieved the best F1 among single models (0.802), balancing precision and recall. RoBERTa achieved the highest recall but incurred more false positives, thereby reducing precision. MentalBERT underperformed, likely due to a domain mismatch with the informal and sarcastic language observed on Bluesky. Overall, the stacked ensemble (BERT + XGBoost) achieved the highest F1 (0.805), supporting the usefulness of metadata, although it may encode platform-specific biases.

5.4. Ablation Study: Metadata Contribution

To assess the value of engagement metadata, we retrained the stacked ensemble using only textual model outputs, resulting in a **6% drop in F1-score**. This suggests that metadata provides complementary signals, but gains should be interpreted cautiously since engagement may reflect platform-specific dynamics rather than psychological states.

5.5. Cross-Platform Linguistic Insights

A comparative linguistic analysis highlighted differences between platforms:

- **Bluesky:** More explicit, self-reflective depressive disclosures (e.g., “I’m empty tonight”), likely influenced by a younger, niche user base and less algorithmic curation.
- **Twitter:** More sarcasm, meme culture, and indirect expressions shaped by broader demographics and content virality.

These observations are consistent with our hypothesis on platform-specific linguistic styles.

5.6. External Validation on Twitter

We tested the stacked ensemble on the Kaggle Twitter Depression dataset [InfamousCoder 2022], achieving an accuracy of 90%, a precision and recall of 0.869 and 0.951, respectively, and an F1 score of 0.908. These results outperform the Kaggle leaderboard baseline (F1 = 0.87), demonstrating strong cross-platform generalization. Still, sarcasm-rich posts remained a challenge, highlighting a key avenue for future research on explicit irony detection.

6. Conclusions and Future Work

This work presented the first large-scale study on depression detection in **Bluesky**, expanding mental health NLP research to decentralized social media. We introduced an expert-annotated dataset and a multi-stage detection pipeline combining transformers, metadata-based models, and stacking techniques.

Our results show that transformer-based architectures substantially outperform classical baselines, while metadata provides complementary gains that may encode platform-specific biases. External validation on Twitter indicates strong cross-platform generalization, although qualitative analysis suggests that platform architecture and audience characteristics shape how depressive language is expressed.

Limitations include the removal of emoji information, which was done to reduce noise and variability, as emoji usage is highly context-dependent and not explicitly modeled by our textual pipeline. Additionally, the task is framed as binary classification without modeling symptom severity. Future work should explore multimodal signals (e.g., emojis and temporal patterns), severity-aware or ordinal formulations grounded in clinical frameworks, domain-adaptive pretraining for informal and sarcastic discourse, and broader cross-platform evaluations across decentralized networks, with particular attention to failure cases such as somatic complaints, sarcasm, and retrospective disclosures.

Overall, our results suggest that depression detection models can generalize across platforms, while still benefiting from awareness of platform-specific linguistic styles and ethical safeguards.

To promote reproducibility, we make the repository containing our depression dataset publicly available.¹

References

Ahmed, M. F. et al. (2025). Depression detection using machine learning techniques on x/twitter data. *International Journal of Innovative Technology and Exploring Engineering (IJITCE)*.

¹<https://github.com/aliciamarzola/Bluesky-Depression-Dataset>

- Association, A. P. (2013). *Diagnostic and Statistical Manual of Mental Disorders (DSM-5)*.
- Atapattu, T. et al. (2022). Emoment: Multilingual mental health corpus with fine-grained emotion annotations. In *Proceedings of the 29th International Conference on Computational Linguistics (COLING 2022)*.
- Bains, N. and Abdijadid, S. (2023). Major depressive disorder. In *StatPearls [Internet]*.
- Bokolo, A. and Liu, C. (2024). Comparative performance of transformer-based models for detecting mental health conditions. *Electronics*.
- Cacheda, F. et al. (2019). Early detection of depression: social network analysis and random forest techniques. *Journal of Medical Internet Research*.
- Cao, L., Wang, R., and Zhou, Y. (2025). A systematic review of machine learning approaches for depression detection on social media. *Journal of Big Data Science*.
- Cha, J., Park, S., and Sim, J. (2022). A lexicon-based approach to examine depression detection in social media: the case of twitter and university community. *Palgrave Communications*.
- Chancellor, S. et al. (2016). Quantifying and predicting mental illness severity in online pro-eating disorder communities. *Proceedings of the 19th ACM Conference on Computer-Supported Cooperative Work & Social Computing (CSCW)*.
- Conway, M. and O'Connor, D. (2016). Social media, big data, and mental health: current advances and ethical implications. *Current Opinion in Psychology*.
- Coppersmith, G., Dredze, M., and Harman, C. (2014). Quantifying mental health signals in twitter. In *Proceedings of the Workshop on Computational Linguistics and Clinical Psychology: From Linguistic Signal to Clinical Reality*.
- D'Cruz, L., Dubey, V., and Thakur, P. (2023). Depression prediction from combined reddit and twitter data using machine learning. In *2023 2nd International Conference for Innovation in Technology (INOCON)*.
- De Choudhury, M., Counts, S., and Horvitz, E. (2013a). Social media as a measurement tool of depression in populations. *Proceedings of the 5th Annual ACM Web Science Conference*.
- De Choudhury, M. et al. (2013b). Predicting depression via social media. In *Proceedings of the International AAAI Conference on Weblogs and Social Media (ICWSM)*.
- Devlin, J. et al. (2018). Bert: Pre-training of deep bidirectional transformers for language understanding.
- Germain, A. and Kupfer, D. J. (2008). Circadian rhythm disturbances in depression. *Human Psychopharmacology: Clinical and Experimental*.
- Grabb, D., Lamparth, M., and Vasan, N. (2024). Risks from language models for automated mental healthcare: Ethics and structure for implementation. *arXiv preprint arXiv:2406.11852*.
- Guntuku, S. C. et al. (2017). Detecting depression and mental illness on social media: an integrative review. *Current Opinion in Behavioral Sciences*.

- Hasan, A. and Kumar, R. (2025). Benchmarking transformer and lstm models for depression detection on reddit. *arXiv preprint arXiv:2507.19511*.
- He, P., Gao, J., and Chen, W. (2021). Deberta: Decoding-enhanced bert with disentangled attention. *arXiv preprint arXiv:2006.03654*.
- InfamousCoder (2022). Depression: Twitter dataset + feature extraction.
- Ji, S. et al. (2021). Mentalbert: Publicly available pretrained language models for mental healthcare.
- Kabir, M. et al. (2022). Deptweet: A typology for social media texts to detect depression severities. *Computers in Human Behavior*.
- Kerasiotis, M., Ilias, L., and Askounis, D. (2024). Depression detection in social media posts using transformer-based models and auxiliary features. *Social Network Analysis and Mining*.
- Lieblich, S. M. et al. (2015). High heterogeneity and low reliability in the diagnosis of major depression will impair the development of new drugs. *BJPsych Open*.
- Liu, D. et al. (2022). Detecting and measuring depression on social media using a machine learning approach: Systematic review. *JMIR Mental Health*.
- Liu, Y. et al. (2019). Roberta: A robustly optimized bert pretraining approach.
- Losada, D. E., Crestani, F., and Parapar, J. (2019). Overview of erisk 2019: Early risk prediction on the internet. In *Experimental IR Meets Multilinguality, Multimodality, and Interaction (CLEF 2019)*.
- Matei, H. (2025). Bluesky vs twitter: Which platform is the future of social media?
- Organization, W. H. (2023). Depressive disorder (depression).
- Rai, S. et al. (2024). Key language markers of depression on social media depend on race. *Proceedings of the National Academy of Sciences of the United States of America*.
- Regier, D. A. et al. (2013). Dsm-5 field trials in the united states and canada, part ii: Test-retest reliability of selected categorical diagnoses. *American Journal of Psychiatry*.
- Silberling, A. (2024). Bluesky is now open for anyone to join — techcrunch.
- Spitzer, R. L., Kroenke, K., and Williams, J. B. (2001). Validity of a brief depression severity measure (phq-9). *Journal of General Internal Medicine*.
- Tadesse, M. M. et al. (2019). Detection of depression-related posts in reddit social media forum.
- Tavchioski, I., Robnik-Šikonja, M., and Pollak, S. (2023). Detection of depression on social networks using transformers and ensembles.
- Topics, E. (2025). Bluesky users: Age, gender, and demographic insights (2025).
- Zimmer, M. (2010). “but the data is already public”: On the ethics of research in facebook. *Ethics and Information Technology*.