Frontiers in Computing and Intelligent Systems
ISSN: 2832-6024 | Vol. 17, No. 2, 2026

A Study on Social Media Depression Detection Methods Based on Multi-Level User Representation Learning

Xu Zhang *

Future Intelligent Art Institute, Hebei Academy of Fine Arts, Shijiazhuang, Hebei, 050700, China
*** Corresponding author Email:** 3175931280@qq.com

Abstract: Social-media posts record not only what users say but also when and how they participate. These two sources of evidence can support computational screening for depression-related patterns, although the task is complicated by indirect language, overlapping class boundaries, and incomplete behavioral records. This study examines whether combining information at several levels produces more informative user representations than relying on isolated posts. The proposed workflow begins with BERT-based encoding of post content. Random downsampling is used in the pre-trained-model experiments to reduce the imbalance between the two labels. Posting frequency, activity-time distribution, and historical-post volume are then mapped into the same feature space as the text representation. Post-level vectors are aggregated for each user, and a graph neural network is applied to a similarity graph so that the final representation also reflects relationships among users. A supervised contrastive objective is included to encourage compact within-class representations and clearer separation between classes. Across the reported test results, user-level models were markedly stronger than tweet-level baselines. The user-level BERT + GNN configuration obtained an accuracy of 0.9404, an F1-score of 0.9399, and an AUC of 0.9703. These results suggest that historical context and relational structure are useful for this classification setting; they should not, however, be interpreted as a substitute for clinical assessment.

Keywords: Depression Detection; Social Media; Multi-Level User Representation Learning; Graph Neural Networks; Contrastive Learning.

1. Introduction

Social media has become a routine venue for recording everyday experiences, opinions, and emotional reactions [1–3]. For depression-related research, these records are attractive because they provide longitudinal text and activity traces at a scale that is difficult to obtain through questionnaires alone [2,3]. The same data are also difficult to interpret: distress may be expressed indirectly, normal and depression-related language often overlaps, and behavioral information is unevenly available across users.

Many earlier systems therefore concentrated on the text of individual posts. Pre-trained language models have improved contextual encoding [4,5], but a post-only representation omits patterns that emerge over time, such as changes in posting frequency or repeated themes. A second limitation is that a classification loss mainly separates labels at the decision boundary; it does not directly require a well-organized embedding space. Contrastive objectives provide one way to impose that additional structure [6].

This work treats the user, rather than the single post, as the main unit of analysis. Textual content, behavioral statistics, and similarity-based user relations are modeled in one pipeline, while supervised contrastive learning is used as an auxiliary objective. The resulting framework is intended to test how each information source changes classification performance and representation geometry.

The main contributions of the study are summarized below:

(1) A joint text-behavior representation is constructed by combining BERT embeddings with an MLP projection of user activity statistics.

(2) Historical posts are aggregated at the user level, reducing the influence of an unusually short, ambiguous, or noisy single post.

(3) A graph neural network is used to propagate information over a user-similarity graph, adding relational context to the user representation [7,11].

(4) A supervised contrastive term is added to the classification objective so that same-label samples are encouraged to remain close while different-label samples are pushed apart.

(5) Comparative, ablation, and t-SNE analyses are reported to examine the contribution of modeling granularity, behavioral features, graph structure, and contrastive learning.

2. Literature Review

2.1. Research on Depression Detection

Research on depression-related signals in social media has moved from manually designed features toward learned representations. Early systems commonly used word counts, sentiment lexicons, and conventional classifiers. Such features are transparent but have limited ability to represent context, negation, irony, or meanings that depend on a sequence of posts. More recent work has combined text with behavioral and relational information [8,9]. Sawhney et al. examined deep models for suicidal-ideation language [8], while the survey by Yang et al. identified multimodal evidence and long-term behavior as continuing challenges for social-media depression detection [9].

The broader trend is therefore not simply toward larger models, but toward representations that preserve context across several sources and time scales.

2.2. Text-Based Sentiment Analysis Methods

Text remains the most direct input because depressive signals are often discussed, implied, or narrated in language. BERT encodes each token using both left and right context,

which makes it better suited than sparse word-frequency features to expressions whose meaning depends on surrounding words [4].

RoBERTa retains the Transformer architecture while revising the pre-training procedure and data regime [5]. Findings from sentiment and mental-health classification studies indicate that these contextual encoders are strong baselines, although their output still reflects only the text supplied to them.

2.3. User Behavioral Features and User-Level Modeling

A user cannot always be characterized from wording alone. Posting frequency, time-of-day activity, and the amount of available history may provide complementary evidence, particularly when individual posts are short or semantically neutral.

User-level modeling addresses a related problem by pooling several posts from the same account. The pooled vector is less sensitive to one atypical message and can represent recurring language or activity patterns over a longer observation window.

2.4. Graph Neural Networks and Structural Information Modeling

Users may also resemble one another in content or behavior even when no explicit social link is available. A graph represents each user as a node and encodes selected similarities as edges. Message passing then allows a node representation to incorporate information from its neighborhood [7,11].

Graph convolutional networks are a common implementation for node classification. Their usefulness depends on the edge-construction rule: informative neighborhoods can support classification, whereas noisy similarities can spread misleading information.

2.5. Applications of Contrastive Learning in Representation Learning

Contrastive learning changes the geometry of a representation space by comparing samples rather than optimizing only a label prediction. SimCSE demonstrates this principle for sentence embeddings, where positive and negative pairs guide the organization of the encoded vectors [12].

With class labels available, supervised contrastive learning can treat other samples from the same class as positives. The objective therefore complements cross-entropy by rewarding within-class consistency and between-class separation, although improved geometry does not guarantee that every evaluation metric will increase.

2.6. Chapter Summary

The reviewed literature provides effective components for contextual text encoding, behavioral modeling, and graph-based learning, but these components are often evaluated separately.

The method in this paper combines them at the user level and uses contrastive learning as an auxiliary constraint. The experiments are designed to separate the effect of each component rather than treating the full pipeline as a single undifferentiated model.

3. Method

3.1. Overall Model Framework

The proposed framework represents a user through three linked views: the semantics of posted text, summary statistics of platform activity, and position in a similarity graph. Figure 1 shows how these views are encoded and merged before classification. The contrastive objective is applied to the learned user representation rather than to raw inputs.

The framework contains six stages:

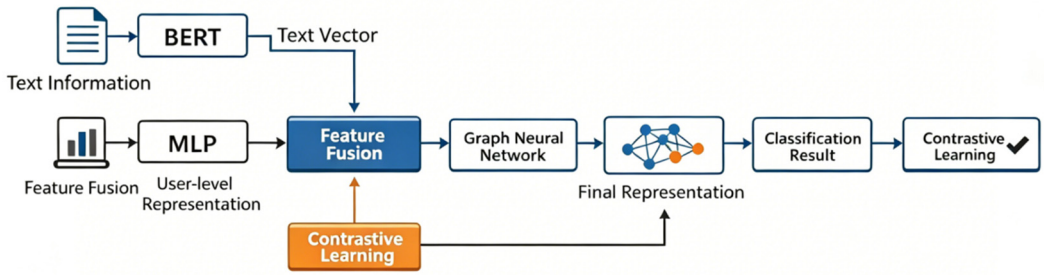


Figure 1. Multi-Level User Representation Learning Framework

(1) Text semantic representation module based on pre-trained language models;

(2) User behavior feature modeling module;

(3) Text and behavior feature fusion module;

(4) User-level representation aggregation module;

(5) Structural modeling module based on graph neural networks;

(6) Supervised contrastive learning optimization module.

The modules are trained as a connected pipeline. Text and behavior provide node features, user-level aggregation reduces post-specific noise, and graph propagation adds neighborhood context. The classifier and contrastive loss then shape the final representation for the depression-related label.

3.2. Text Semantic Representation

BERT is used to encode the text associated with each post. Its bidirectional Transformer layers produce a contextual vector for every token in the input sequence.

Let an input sequence be written as:

$x = \{w_1, w_2, \ldots, w_n\}$

The sequence is tokenized and passed through BERT, producing the hidden-state matrix:

$$H = BERT(x)$$

Here, H ∈ R^(n×d), where n is the token count and d is the hidden dimension. The hidden state at the [CLS] position is used as a fixed-length representation of the post:

$$h_text = H_[CLS]$$

The [CLS] vector is used in the subsequent fusion and aggregation stages. It is a compact summary of the encoded sequence rather than an explicit measurement of clinical state.

3.3. Modeling User Behavioral Features

Text does not capture all available user information, so the model also includes activity statistics. The feature vector contains posting frequency, the distribution of active hours, and the number of historical posts available for the user.

For a user, the behavioral feature vector is denoted by:

$$f_behav \in R\hat{}k$$

The k-dimensional vector is projected with a multilayer perceptron so that its output can be combined with the text embedding:

$$h_behav = MLP(f_behav)$$

The projected vector h_behav ∈ R^d has the same dimensionality as h_text, which allows the two representations to be combined in the next stage.

3.4. Multimodal Feature Fusion

The text and behavioral vectors are concatenated to retain information from both sources:

$$h = [h_text\ ;\ h_behav]$$

The operator [;] denotes concatenation. Unlike averaging, concatenation allows the next layer to learn separate weights for semantic and behavioral dimensions.

A fully connected layer then transforms the combined vector:

$$z = ReLU(W \cdot h + b)$$

W and b are learned during training. ReLU introduces a non-linear mapping before user-level aggregation.

3.5. User-Level Representation Modeling

A single post may contain too little context to characterize a user reliably. The model therefore groups posts by author and constructs one representation for each user.

For user u, the available historical posts are:

$$X_u = \{x_1, x_2, \dots, x_m\}$$

Each post x_i is encoded and fused with the relevant behavioral information. The resulting vectors are combined by mean pooling:

$$h_u = (1/m) \sum_{i=1}^{m} h_i$$

Mean pooling gives every available post equal weight and produces a stable fixed-size vector. It also has a limitation: highly informative and uninformative posts contribute equally, an issue that attention-based aggregation could address in later work.

3.6. Graph Neural Network Modeling

The user vectors are next used as node features in a graph neural network. The aim is to test whether similarity to other users supplies information that is absent from an isolated user vector.

Each user is represented by one node. Edges are created from the selected similarity rule, forming adjacency matrix A. A graph convolutional layer updates the node matrix as follows [11]:

$$H\hat{}(l+1) = \sigma(\, A\sim \cdot H\hat{}(l) \cdot W\hat{}(l)\,)$$

In this expression:

- H^(l) is the node-representation matrix at layer l;
- A-tilde is the normalized adjacency matrix;
- W^(l) is the learnable transformation at layer l; and
- σ(·) is the activation function.

After several propagation steps, a node contains information from its local neighborhood and, indirectly, from higher-order neighbors. This mechanism can enrich the user embedding when the graph connects genuinely similar accounts. It can also amplify errors when the edge set contains weak or spurious relationships, so graph quality is examined as a limitation rather than assumed to be uniformly beneficial.

3.7. Supervised Contrastive Learning Optimization

Cross-entropy trains the classifier to predict the label, while supervised contrastive learning additionally constrains distances among user embeddings.

For an anchor i, other samples with the same label form the positive set P(i). Positives should receive higher similarity scores than the remaining samples in the batch. The contrastive loss is written as:

$$L_con = -\Sigma\, log\,[\,\Sigma\, exp(sim(z_i, z_j)/\tau) \,/\, \Sigma\, exp(sim(z_i, z_k)/\tau)\,]$$

where:

P(i) is the set of positive samples for anchor i;

sim(·) is the selected similarity function (for example, cosine similarity);

τ is the temperature parameter; and

the joint training objective is:

$$L = L_cls + \lambda \cdot L_con$$

The coefficient λ controls the contribution of the contrastive term relative to classification loss.

This auxiliary objective changes how samples are arranged in the embedding space. Its value is assessed through both classification metrics and the t-SNE plots, because a cleaner visualization does not necessarily imply a higher AUC.

4. Experimental Design

The dataset contains 1,261,462 posts from users labeled as normal and 259,662 posts from users labeled as depressed. Random downsampling was used to create a balanced subset for the pre-trained-model experiments. Evaluation was conducted at two granularities: individual posts for the tweet-level baselines and aggregated histories for the user-level models.

The evaluated configurations were:

Tweet-level BERT, which classifies one post at a time;

User-level BERT, which averages representations across a user's posts;

User-level graph models, which propagate information over the similarity graph; and

User-level models with supervised contrastive learning.

The BERT-based models used the following settings:

Maximum sequence length: 96

Batch size: 8
Learning rate: 2e-5
Number of epochs: 2

Performance is reported using accuracy, precision, recall, F1-score, and area under the ROC curve (AUC). TF-IDF with logistic regression and TF-IDF with a support vector machine provide conventional baselines for the test-set comparisons in Table 1.

These settings define the experiments reported here. Because the BERT-based models were fine-tuned for only two epochs, the results are best read as a comparison of modeling choices under the stated setup rather than as a claim of exhaustive optimization.

5. Experimental Results and Analysis

5.1. Comparison of Overall Model Performance

Table 1 compares sparse-feature baselines with BERT models at tweet and user level. The table is discussed by examining both threshold-dependent measures and AUC, since accuracy alone can obscure class-specific behavior.

Table 1. Comparison of Model Performance on the Test Set

Model	Accuracy	Precision	Recall	F1-score	AUC
TF-IDF + LR	0.3259	0.1911	0.9123	0.3160	0.5819
TF-IDF + SVM	0.8320	0.6610	0.0322	0.0614	0.0000
BERT (tweet-level)	0.7350	0.7383	0.7350	0.7341	0.8175
BERT + Features	0.7492	0.7493	0.7492	0.7492	0.8361
User-level BERT	0.9339	0.9345	0.9339	0.9328	0.9719
User-level + Contrastive Learning	0.9359	0.9357	0.9359	0.9353	0.9680
User-level + GNN	**0.9404**	**0.9401**	**0.9404**	**0.9399**	**0.9703**

The logistic-regression baseline predicts many positive cases: recall is 0.9123, but precision is only 0.1911, which leads to an F1-score of 0.3160. The SVM shows the opposite failure mode. Its accuracy is 0.8320, yet recall falls to 0.0322 and the reported AUC is 0.0000. These values indicate that the apparently high SVM accuracy is not evidence of useful depression-related classification.

Tweet-level BERT produces a more balanced precision-recall profile, with an F1-score of 0.7341 and an AUC of 0.8175. Adding behavioral features at the tweet level raises accuracy from 0.7350 to 0.7492 and AUC from 0.8175 to 0.8361. The largest change occurs after aggregation by user: user-level BERT reaches 0.9339 accuracy and 0.9719 AUC. This pattern is consistent with the view that several posts provide a more stable signal than one post, although the experiment does not by itself establish why every individual prediction changes.

5.2. Analysis of the Effect of Introducing Behavioral Features

Table 2. Component Ablation Results

Model	Behavior	Contrastive	GNN	Accuracy
User-level BERT	✗	✗	✗	0.9339
Behavior only	✓	✗	✗	0.9492
Contrastive only	✗	✓	✗	0.9359
GNN only	✗	✗	✓	0.9404
Full fusion (not run)	✓	✓	✓	N/A

Table 2 isolates the components added to the user-level BERT baseline. Accuracy is used in this ablation table to keep the comparison compact; the fuller metric set appears in Tables 1 and 3. The “Full Fusion” row is reserved because a jointly trained configuration containing all three additions was not reported.

The baseline user-level BERT accuracy is 0.9339. With behavioral features, the value rises to 0.9492, a gain of 1.53 percentage points. This is the largest increase in the ablation table and suggests that activity statistics add information not fully captured by the pooled text representation.

Contrastive learning produces a smaller accuracy gain, from 0.9339 to 0.9359. Its main contribution in these experiments is therefore better interpreted through the organization of the representation space rather than through a large change in classification accuracy.

The GNN variant reaches 0.9404 accuracy, 0.65 percentage points above the user-level BERT baseline. This improvement supports the use of relational context, but it remains smaller than the gain reported for behavioral features in Table 2.

Taken together, the ablation results do not show that every component contributes equally. Behavioral features have the largest isolated accuracy effect, the GNN provides a moderate gain, and contrastive learning has a modest effect on scalar performance while visibly changing the embedding distribution. A complete joint-fusion experiment is still needed before claims about interaction among the modules can be made.

5.3. Analysis of User-Level and Graph Structural Modeling

Table 3 reorganizes the results by modeling granularity. It compares two tweet-level configurations with three user-level configurations, making the effect of aggregating historical posts visible before graph or contrastive components are considered.

Table 3. Performance at Different Modeling Granularities

Model granularity	Model	Accuracy	F1-score	AUC
tweet-level	BERT	0.7350	0.7341	0.8175
tweet-level	BERT+ Behavioral Features	0.7492	0.7492	0.8361
user-level	BERT	0.9339	0.9328	0.9719
user-level	BERT + Contrastive Learning	0.9359	0.9353	0.9680
user-level	BERT + GNN	**0.9404**	**0.9399**	**0.9703**

The change from tweet-level BERT to user-level BERT is the dominant result in Table 3. Accuracy increases from 0.7350 to 0.9339, F1-score from 0.7341 to 0.9328, and AUC from 0.8175 to 0.9719. The improvement is much larger than the changes produced by either the graph or contrastive module within the user-level setting.

One plausible explanation is that pooled histories expose recurring wording and longer-term activity that a single post cannot show [1,3]. The result should be described as an association within this dataset, because the experiment does not isolate every possible difference between post-level and user-level sampling.

Within the user-level group, the contrastive model has slightly higher accuracy and F1 than the baseline, whereas its AUC is lower (0.9680 versus 0.9719). The GNN obtains the highest accuracy and F1 in Table 3, but its AUC of 0.9703 also remains just below the user-level BERT value.

The graph result suggests that neighborhood information can refine some decisions even when the pooled text representation is already strong. At the same time, message passing is only as reliable as the graph used to define neighbors. Similarity edges that do not correspond to meaningful behavioral or linguistic relationships may mix incompatible users and reduce performance.

5.4. Analysis of the Effects of the Contrastive Learning Mechanism

The supervised contrastive term is intended to make label structure explicit in the embedding space. Cross-entropy is retained for prediction, while the auxiliary loss compares each anchor with same-label and different-label samples in the batch.

The numerical results are mixed rather than uniformly positive: accuracy rises from 0.9339 to 0.9359 and F1 from 0.9328 to 0.9353, but AUC declines from 0.9719 to 0.9680. Thus, the evidence supports a modest improvement at the selected decision threshold and a change in representation geometry, not a general improvement across all metrics [6,12].

5.5. Analysis of Representation Space Distribution

Figure 2 shows the t-SNE projection of the baseline user-level embeddings. The two labels occupy broad regions, but they overlap along the lower arc and near the central boundary. Several isolated points are also visible. Because t-SNE preserves local neighborhoods imperfectly, these patterns are descriptive rather than direct measures of classification quality.

In Figure 3, the contrastive embeddings form two more recognizable regions, especially on the right side of the projection. A mixed band remains near the upper and left portions of the plot, so the classes are not completely separable. The figure nevertheless illustrates the geometric effect that the auxiliary loss was designed to produce.

The visual change is consistent with greater within-class grouping, but it should be considered together with the metrics in Table 3. In particular, the clearer two-dimensional projection coexists with a slightly lower AUC, which cautions against treating a t-SNE plot as proof of better generalization.

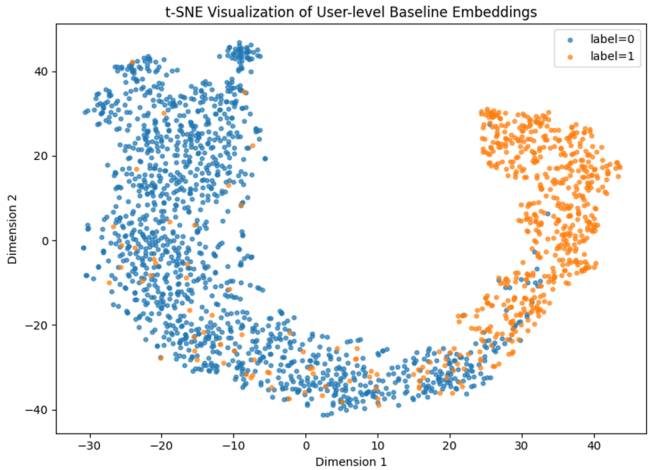


Figure 2. t-SNE Projection of User-Level BERT Embeddings

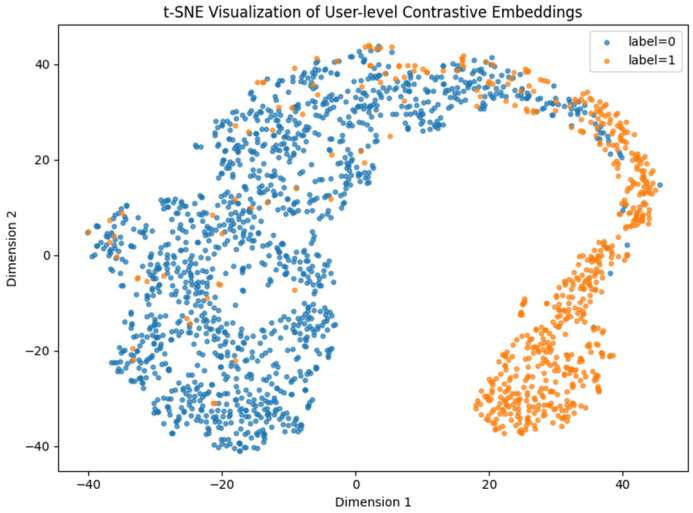


Figure 3. t-SNE Projection of User-Level Contrastive Embeddings

5.6. Error Analysis

Overall performance is high for the user-level models, but inspection of incorrect predictions reveals several recurring failure modes. The following categories summarize patterns observed in the reviewed errors; they are qualitative observations rather than a complete statistical taxonomy.

5.6.1. Misclassification Due to Implicit Expression

Some posts communicate distress through understatement, sarcasm, or indirect wording rather than through explicit negative terms. A representation based on surface sentiment can therefore miss the implied meaning, especially when the surrounding user history is limited.

For example:

"Nothing much has happened lately, but I just can't seem to muster the energy."

The sentence is grammatically neutral and contains no direct diagnostic term. Its reference to an inability to summon energy may nevertheless be relevant in context, which explains why a model can assign it to the normal class when the broader history is unavailable.

5.6.2. Insufficient Information in Short Texts

Very short posts provide few contextual cues and may support several interpretations. This problem is most visible in tweet-level classification, where there is no additional history to clarify whether a phrase is an isolated remark or part of a persistent pattern.

5.6.3. Missing or Unstable Behavioral Features

Behavioral variables are also unreliable for new, inactive, or irregular users. A low post count yields unstable frequency and time-of-day estimates, while sudden changes in activity can make a summary vector unrepresentative of the observation period.

5.6.4. Blurred Class Boundaries

Some user vectors lie near the boundary between the two labels. Their histories contain evidence associated with both classes, so small changes in sampling, pooling, or the decision threshold can alter the predicted label.

These boundary cases are important because social-media labels compress a continuous and context-dependent condition into two categories. Better embeddings may reduce overlap, but they cannot remove uncertainty created by noisy labels or incomplete histories.

5.6.5. Impact of Graph-Structure Noise

The similarity graph can introduce another source of error. An edge may connect users who share vocabulary or posting habits without sharing the same underlying condition. Message passing then transfers irrelevant information and can move a node toward the wrong region of the representation space.

6. Summary

The reviewed errors arise from four main sources: indirect language, insufficient textual context, unstable behavioral summaries, and noisy graph neighborhoods. These observations point to concrete improvements, including richer temporal modeling, more selective aggregation, and learned or confidence-weighted graph construction.

A subset of samples is misclassified by several different models. These cases usually combine more than one difficulty—for example, ambiguous language together with sparse activity data—so changing a single module is unlikely to resolve them. They are useful candidates for future case-level analysis and label review.

The error analysis therefore qualifies the headline accuracy results. Multi-level modeling reduces several weaknesses of tweet-level classification, but performance still depends on data coverage, label quality, and the assumptions used to construct user relationships.

7. Conclusion

This study evaluated a multi-level approach to classifying depression-related user labels from social-media data. BERT supplied contextual post embeddings, behavioral statistics added activity information, and mean pooling converted post representations into user vectors. A graph neural network then incorporated similarity-based neighborhood information.

The experiments identify user-level aggregation as the largest source of improvement: accuracy increased from 0.7350 for tweet-level BERT to 0.9339 for user-level BERT. The GNN produced the highest reported accuracy, 0.9404, while supervised contrastive learning modestly improved accuracy and F1 and produced a clearer t-SNE structure. Its AUC, however, was slightly below that of the user-level BERT baseline.

Several limitations restrict the interpretation of these

findings. The data come primarily from one platform, the graph is based on a relatively simple similarity rule, and multimodal content such as images is not modeled. The experiments also use a short fine-tuning schedule and do not report statistical significance across repeated runs.

Future work should evaluate cross-platform and cross-language transfer, replace mean pooling with temporal or attention-based aggregation, and learn graph edges jointly with node representations. Multimodal features may be added when they are available and ethically appropriate.

Finally, this system should be understood as a research classifier for platform data, not as a clinical diagnostic tool. External validation, privacy safeguards, and human review would be required before any practical use involving individual users.

References

[1] Shen, G., Jia, J., Nie, L., et al. (2017). Depression detection via harvesting social media: A multimodal dictionary learning solution. *Proceedings of the International Joint Conference on Artificial Intelligence (IJCAI)*.

[2] Orabi, A. H., Buddhitha, P., Orabi, M. H., et al. (2018). Deep learning for depression detection of Twitter users.*Proceedings of the ACL Workshop*.

[3] Zogan, H., Razzak, I., Wang, X., et al. (2021). Explainable depression detection with multi-aspect features using social media data. *IEEE Access*, *9*, 102154–102166.

[4] Devlin, J., Chang, M. W., Lee, K., & Toutanova, K. (2019). BERT: Pre-training of deep bidirectional transformers for language understanding. *Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies (NAACL-HLT)*, 4171–4186.

[5] Liu, Y., Ott, M., Goyal, N., Du, J., Joshi, M., Chen, D., et al. (2019). *RoBERTa: A robustly optimized BERT pretraining approach* [Preprint]. arXiv. https://arxiv.org/abs/1907.11692

[6] Chen, T., Kornblith, S., Norouzi, M., & Hinton, G. (2020). A simple framework for contrastive learning of visual representations. *Proceedings of the 37th International Conference on Machine Learning (ICML)*, 1597–1607.

[7] Hamilton, W. L., Ying, R., & Leskovec, J. (2017). Inductive representation learning on large graphs. *Advances in Neural Information Processing Systems (NeurIPS)*, 30, 1024–1034.

[8] Sawhney, R., Manchanda, P., Mathur, P., et al. (2018). Exploring and learning suicidal ideation connotations on social media with deep learning. *Proceedings of the 2018 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies (NAACL-HLT)*, 168–175.

[9] Yang, C., Chen, H., Li, Z., et al. (2023). A survey on depression detection via social media: Recent advances and challenges. *IEEE Transactions on Affective Computing*, *14*(4), 2745–2764.

[10] Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., et al. (2017). Attention is all you need. *Advances in Neural Information Processing Systems (NeurIPS)*, 30, 5998–6008.

[11] Kipf, T. N., & Welling, M. (2017). *Semi-supervised classification with graph convolutional networks* [Preprint]. arXiv. https://arxiv.org/abs/1609.02907

[12] Gao, T., Yao, X., & Chen, D. (2021). SimCSE: Simple contrastive learning of sentence embeddings. *Proceedings of the 2021 Conference on Empirical Methods in Natural Language Processing (EMNLP)*, 6387–6397.