Mapping Long-term Causalities in Psychiatric Symptomatology and Life Events from Social Media

Siyuan Chen[1], **Meilin Wang**[1], **Minghao Lv**[1], **Zhiling Zhang**[2], **Qianqian Ju**[3], **Dejiyangla**[3], **Yujia Peng**[3,4], **Kenny Q. Zhu**[5*], **Mengyue Wu**[1*]

[1] X-LANCE, MoE Key Lab of Artificial Intelligence, AI Institute, Shanghai Jiao Tong University, China

[3] Beijing Key Laboratory of Behavior and Mental Health,
School of Psychological and Cognitive Sciences, Peking University, China

[4]National Key Laboratory of General Artificial Intelligence,
Beijing Institute for General Artificial Intelligence, China

[5]University of Texas at Arlington, USA

Abstract

Social media is a valuable data source for exploring mental health issues. However, previous studies have predominantly focused on the semantic content of these posts, overlooking the importance of their temporal attributes, as well as the evolving nature of mental disorders and symptoms. In this paper, we study the causality between psychiatric symptoms and life events, as well as among different symptoms from social media posts, which leads to better understanding of the underlying mechanisms of mental disorders. By applying these extracted causality features to tasks such as diagnosis point detection and early risk detection of depression, we notice considerable performance enhancement. This indicates that causality information extracted from social media data can boost the efficacy of mental disorder diagnosis and treatment planning.

1 Introduction

Mental health, a critical facet of overall well-being, remains a global challenge affecting individuals from diverse backgrounds (Dreisbach et al., 2019). Traditional methods for studying mental health have often relied on clinical assessments, surveys, and interviews (Beck et al., 1996), providing valuable but limited snapshots of an individual's mental state. Mental disorders often manifest as long-term conditions, unfolding gradually over extended periods (Collingridge et al., 2010). This intrinsic characteristic of mental health issues presents a challenge to current clinical diagnosis and assessment methods, which predominantly concentrate on symptom duration within a narrow two-week window (American Psychiatric Association et al., 2013). Such an approach may not fully capture the nuanced and evolving nature of mental illnesses, potentially overlooking the progressive development of symptoms and the broader impact on an individual's life.

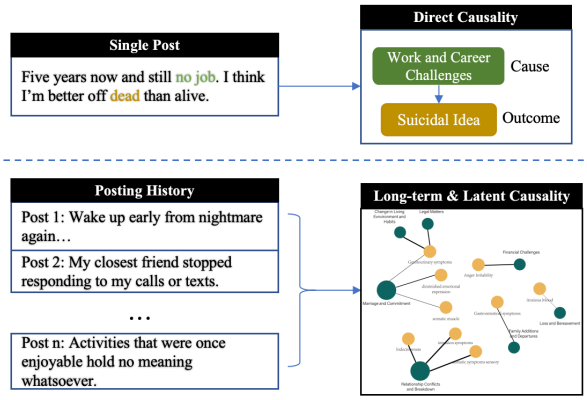

Figure 1: Direct causal relationships from a single post (above) versus long-term and latent causality from a multi-post history (below).

This discrepancy highlights the need for a more longitudinal perspective in mental health assessment, one that considers the full spectrum of symptom development and progression over time, thereby providing a more accurate and holistic view of an individual's mental health state. To this end, social media platforms serve as a valuable resource for tracking mental health, with users frequently documenting their thoughts and emotions over extended periods. In contrast to traditional clinical methods that focus on short-term symptoms, social media posts offer a continuous, candid narrative of an individual's mental state. This user-generated content provides insights into the evolving nature of mental health, capturing subtle changes and patterns over time that might be overlooked in brief clinical assessments. Discussions about symptoms and life stressors on social media are closely tied to mental well-being (Charles et al., 2013; Harandi et al., 2017). Monitoring the evolution of users' posts over time enables a more comprehensive analysis of the origins and progression of their condi-

*Corresponding authors.
[2]Independent researcher.

Proceedings of the 2024 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies (Volume 1: Long Papers), pages 5472–5487
June 16-21, 2024

tions, facilitating early interventions when signs of mental health issues emerge.

Previous studies (Shen and Rudzicz, 2017; Zhang et al., 2022a) have consistently focused on detecting mental disorders through the textual content of social media posts, neglecting the analysis of chronological attributes. While achieving high detection accuracy (Chancellor and De Choudhury, 2020; Chen et al., 2023), we argue that a singular outcome is insufficient for a profound comprehension of mental disorder development and the relationships among various factors (e.g., psychiatric symptoms). Recently, some pioneering studies (Garg et al., 2022; Saxena et al., 2023) began to explore the causes of mental health issues in social media posts. Nevertheless, their focus remains solely on extracting *direct* causal relationships from the semantic information within one post (Luo et al., 2016), as illustrated in Figure 1. This approach can only capture limited causality, as a substantial number of *long-term*, *latent* causal relationships may not necessarily manifest within a single post.

Therefore, our work seeks to address these limitations by revealing latent causes behind psychiatric symptoms through a computational method encompassing users' entire posting history. Building upon existing literature that indicates reciprocal influences among symptoms (Agirman et al., 2021; Shah et al., 2023) and the potential for stressful life events to cause symptoms (Radell et al., 2021; Ruengorn et al., 2021), we endeavor to explore both "*symptom-to-symptom*" and "*life-event-to-symptom*" causal relationships in this research.

We conduct our analysis on a large-scale dataset of Reddit posts with users diagnosed with various mental disorders (Chen et al., 2023). Our initial step involves training models to identify psychiatric symptoms and life events[1] from our text dataset. Then, we employ the classical causal discovery method, propensity score matching (PSM) (Rosenbaum and Rubin, 1983), to unveil the causal relationships between the past symptoms or life events and the future symptoms. To illustrate the efficacy of causality unveiled from social media posts, we anchor our findings in authoritative psychiatry literature, and find that our results aligns with many clinically controlled experiments. Furthermore, we integrate these identified causal relationships as additional features into two chronological disease detection tasks, namely diagnosis point detection and early risk detection of depression. The enhanced detection performance also underscores the importance of causality. The main contributions of this work are:

- We propose to mine implicit, subtle, and long-term causal relationships between factors related to mental disorders from the enormous and evolving social media stream, which can overcome the limitation of single text extraction methods (Garg et al., 2022).
- We discover various reliable *"symptom-to-symptom"* and *"life event-to-symptom"* causal effects with Propensity Score Matching, which can be supported by existing clinical evidence.
- We achieve a significant performance boost in the Early Risk Detection and Diagnosis Point Detection task by applying these extracted causal relationships, which further verify their reliability and efficacy.

2 Approach

Our objective is to elucidate the causal relationships between symptoms and life events, so it is imperative to respectively identify these elements from social media data.

2.1 Symptom and Life Event Identification

Psychiatric Symptom Building upon prior research that thoughtfully outlined 38 psychiatric symptoms across 7 mental disorders, as well as proposed a symptom identification dataset named PsySym containing 83K annotated sentences from Reddit posts (Zhang et al., 2022b). We adopt their symptom definition[2] and leverage their supervised symptom identification model, trained on this annotated dataset. The model incorporates a Mental BERT-based encoder (Ji et al., 2022) and a linear classifier.

Life Event Given the relatively limited scope of previous research on detecting life events, we refer to the Holmes-Rahe Stress Inventory (Noone,

[1]We identify 38 symptoms, such as depressed mood, poor memory, etc., and 11 life events, including financial challenges (Noone, 2017; Zhang et al., 2022b). The complete list is provided in Appendix A.

[2]These symptoms (e.g., anxious mood, sleep disturbance, poor memory) are carefully extracted from DSM-5 (American Psychiatric Association et al., 2013), so that there is as little semantic overlap as possible between them. We list all the symptoms in Table 6 (Appendix A).

2017), which encompasses 43 stressful life events. While the inventory is comprehensive, the multitude of categories poses challenges for annotation and model training. Hence, we consolidate these 43 life events into 11 groups based on similarity[3]. Then, we annotated a life event identification dataset[4] using the same procedure as Zhang et al. (2022b) and trained a supervised model on this dataset using the same model architecture (i.e., Mental BERT with linear classifier).

Utilizing these two classifiers, we can deduce a 38-dimensional symptom vector and an 11-dimensional life event vector for each post, where each dimension signifies the probability of a specific symptom or life event. We present the detailed identification results of these models in Appendix C, to show that using these classifiers can help us automatically and accurately extract psychiatric symptoms and life events on Reddit corpus.

2.2 Causality Inference

In this section, we first provide formal definition of our task, followed by the specified approach we used to extract causal relations.

2.2.1 Preliminaries

Exploring causal relationships involves a primary question about:

> What would the *outcome* be if the *treatment* is given[5]?

Therefore, if we want to find out the causal relationship between symptom s_o and s_t, the question becomes "What would s_o (outcome) progresses if a person has s_t (treatment)?" Intuitively, we can measure this "progression" by calculating the difference in outcomes between the treated and untreated (i.e., control) groups. To quantify this difference and assess the causal relationship between treatment and outcome, the Average Treatment Effect (ATE) (Rosenbaum and Rubin, 1983) is introduced:

$$ATE = E[Y(1) - Y(0)] \quad (1)$$

Here, $Y(0)$ represents the outcome for a unit without the treatment, and $Y(1)$ denotes the outcome for the same unit with the treatment.

However, the relationship between s_o and s_t might be a spurious correlation rather than a causal one, induced by other variables, known as *confounders*, which are correlated with both the treatment and the outcome (Feder et al., 2022). Therefore, to establish trustworthy causal relationships, it is crucial to minimize the impact of confounding effects. This can be achieved by thoughtfully selecting treated and control groups, ensuring their similarity on other attributes apart from the treatment variable.

2.2.2 Propensity Score Matching

To enhance the selection of treated and control groups, we apply Propensity Score Matching (PSM) (Rosenbaum and Rubin, 1983), which is widely used in observational studies to reduce bias and the influence of confounding variables (Imbens and Rubin, 2015). The main idea of PSM is to find groups of Treatment and Control posts whose covariates are statistically similar to one another, where the former group has received treatment and the latter has not. The PSM model matches posts based on their *likelihood* of receiving the treatment, represented as the propensity score. The PSM methodology entails two key stages:

- **Estimating Propensity Scores:** We build logistic regression model to predict a post's treatment likelihood based on their covariates vector X. The estimated propensity score is given by:

$$e(X) = \frac{1}{1 + e^{-X\beta}}$$

- **Matching:** Then, treated and control groups are paired 1:1 based on similar propensity scores using a nearest-neighbor matching technique.

Causality between Symptoms To measure the causality between symptom s_o and s_t, we apply PSM to compute the propensity score for each post. In this process, we consider symptoms other than s_o and s_t as the covariates, ensuring that the matched Treatment and Control pairs exhibit high similarity in these other symptoms. Subsequently, the posts are classified into two groups: Treatment group and Control group, based on whether the post has referenced symptom s_t. Then, we can measure the difference in the outcome of these two groups. For a post i mentioning symptom s_t, if there exists another post mentioning symptom s_o within

[3]The corresponding relationship between the original definition and our merged grouping is detailed in Appendix A.

[4]The detailed annotation procedure of the life event dataset can be found in Appendix B.

[5]We use the term "treatment" in accordance with the causal inference terminology, which means a binary variable that may affect the outcome.

a certain time window w[6], we consider outcome $Y_i = 1$ for this post. With matched pairs established, we estimate the average treatment effect as the difference in means of the matched pairs:

$$ATE(s_t, s_o) = \frac{1}{N_m} \sum_{(i,j)} (Y_i - Y_j)$$

where N_m symbolizes the number of matched pairs, (i, j) is a matched pair of posts.

Causality between Life events and Symptoms Similar to assessing causality between symptoms, we use PSM to match the Treatment group (users with posts mentioning life event l_t) with the Control group (users without posts mentioning l_t). Covariates, in this case, include other life events except for the treatment life event l_t. The outcome symptom of a life event l_t is determined by whether symptom s_o is mentioned within the time window w.

3 Applications of Causality

In this section, we show that we can discover reliable causal relationships that can also be supported with established clinical findings (Section 3.1), and how these causal features can be effectively utilized for mental disorder detection (Section 3.2).

3.1 Inferred Causal Relationships

We can automatically discover various causal relationships from social media with the method mentioned above. To validate their reliability, we also conduct literature reviews, and find that many of them can be supported by existing studies.

We show some examples in Table 1. We can see that some of the conclusions are intuitive, like breakdown will increase the risk of future depression. However, others are more subtle, such as irritability can cause weight change. We then find that, according to Vanzhula et al. (2019), the relationship between eating and irritability stands out a crucial pathway influencing comorbidity between Post-Traumatic Stress Disorder (PTSD) and Eating Disorders (EDs).

To highlight discrepancies between our findings and existing literature, we further examine the top eight causal relationships with the highest ATE scores in Table 2. Among these, five relationships are supported by existing literature. Two relationships show disagreement to some extent, proposing alternative associations. An example is Seinsche et al. (2023)'s finding suggesting a connection between social anxiety and a clearer memory about distasteful social situations, which is contradict to our causal link between fear of being negatively evaluated and poor memory. Moreover, one finding got mixed results, depending on studies and samples. Overall, existing literature mostly focuses on symptom-disease causality, providing limited evidence to direct causal examinations between pairs of symptoms. Our current findings have the promise of inspiring future studies focusing on direct causal relationship examinations on pairs of symptoms.

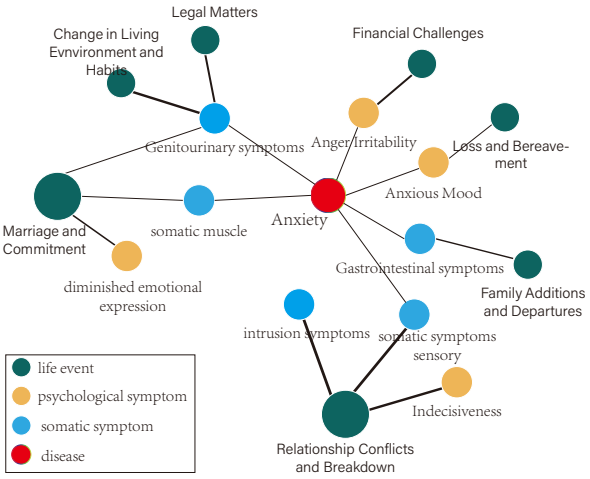

Figure 2: Visualization results for the causality between life events and symptoms with the time window of 1 year. The size of LEs nodes reflects the number of related symptoms, and the thickness of the lines indicates the strength of the causal relationship (ATE).

To make the intricate relationships more intuitive, Figure 2 illustrates a visual representation of the causal relationships between life events (LEs) and symptoms, shedding light on the intricate connections within the context of anxiety. Note that previous works predominantly focus on establishing connections between diseases and life events, while the association between symptoms and LEs is less explored in existing literature. Therefore, our examination of these relationships can provide valuable insights into the intricate web of factors contributing to mental health outcomes.

Moreover, our method can not only find casual relationships qualitatively, but also measure their effects quantitatively with ATE. This enables us to incorporate these findings as numerical features for mental disorder detection algorithms.

[6]For the sake of clarity, we will refer to this time window as **"causal window"** in the following part.

Cause Symptom/Life Events	Result Symptom	Causal Window	ATE	Support
Anger Irritability	Weight and appetite change	30	0.686	Vanzhula et al. (2019)
Fear of gaining weight	Sleep disturbance	30	0.692	Vanzhula et al. (2019)
Hyperactivity agitation	Depressed Mood	90	0.761	Boschloo et al. (2015)
Relationship Conflicts and Breakdown	Depressed Mood	365	0.527	Konac et al. (2021)

Table 1: Example of discovered causal relationships and their corresponding literature supports.

Cause Symptom	Result Symptom	Supported/Disagreed	ATE	References
fears of being negatively evaluated	poor memory	Disagreed	0.894	Seinsche et al. (2023)
fears of being negatively evaluated	Depressed Mood	Supported	0.886	Jacobson and Newman (2017) Kim et al. (2019)
fears of being negatively evaluated	Suicidal ideas	Supported	0.874	Klumpp et al. (2023)
Hyperactivity agitation	Impulsivity	Disagreed	0.869	Grandjean et al. (2021)
fears of being negatively evaluated	do things easily get painful consequences	Supported	0.857	Moscovitch et al. (2018) Chu et al. (2016)
panic fear	somatic symptoms sensory	Mixed	0.855	Ehlers (1993) Kang et al. (2024)
fears of being negatively evaluated	Autonomic symptoms	Supported	0.855	Alvares et al. (2013) Weeks and Zoccola (2016)
panic fear	Decreased energy tiredness fatigue	Supported	0.845	Pasquini et al. (2015)

Table 2: Systematic literature review of the top eight causal relationships with the highest ATE scores.

3.2 Causality as Feature

Given the chronological nature of causal relationships, we leverage them in two temporal tasks: Diagnosis Point Detection and Early Risk Detection, both of which involve the detection of mental disorders along a continuum. The two tasks are illustrated in Figure 3, and we briefly introduce them here:

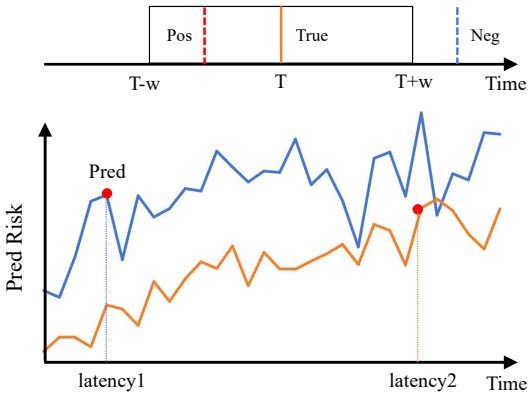

Figure 3: Illustration of the DPD and ERD task. For DPD, we want the predicted time to fall in a window w of the actual diagnosis point T. Prediction with in $[T-w, T+w]$ will be considered positive, otherwise negative. For ERD, we want to predict the risk of a patient as soon as possible (the lower latency, the better), while not making false alert for non-patients.

Task 1. Diagnosis Point Detection (DPD) refers to the identification of the specific time window when a mental disorder is diagnosed in an individual according to their social medial posts. This temporal insight into the diagnosis is valuable because one's mental health state is not static. For this task, MacAvaney et al. (2018) proposed a dataset named RSDD-Time, which includes 598 manually annotated self-reported depression diagnosis Reddit posts that include temporal information about the diagnosis. The complete posting history can be found in the original RSDD dataset (Yates et al., 2017).

Task 2. Early Risk Detection (ERD) aims to detect mentor disorders in early stage (Losada and Crestani, 2016; Zhang et al., 2022a). It can enable early interventions to half the effort and double the results. Here we focus on the ERD of Depression.

In the ERD setting, for a user U_i with posts $[P_{i,1}, P_{i,2}, ..., P_{i,n}]$ in their posting history (where n is the total number of posts, and $P_{i,j}$ is the j-th user-generated post of U_i), posts come one by one. Therefore, only $[P_{i,1}, P_{i,2}, ..., P_{i,t}]$ is available to the model at the t-th time. The model can make an early prediction of y_i at $t(t \leq n)$ once it is confident enough, such that the prediction can make a good tradeoff between accuracy and earliness.

The ERD task doesn't require additional temporal annotations in the dataset, as we care about earliness rather than the exact diagnosis point. Thus, experiments can be conducted using any self-reported depression diagnosis dataset, such as RSDD.

Method of applying causality To incorporate causal relationships into these two tasks, our primary motivation can be summarized as "constructing a more comprehensive daily symptom sequence". For user U, their daily symptom sequence

can be denoted as $[S_1, S_2, ..., S_{n_d}]$, where n_d is the total number of days during the posting history, and S_i means the symptom vectors inferred from the i-th day's user-generated posts of U.

As Figure 4 shows, social media posts may not capture the entirety of an individual's symptom evolution, as users may not share their experiences at all times when symptoms occur. Therefore, the symptom sequence identified from a single post will be incomplete. However, the extracted causal relationships can serve as a universal feature to bridge the gaps in these incomplete symptom sequences. As the example shown in Figure 4, when we have obtained several "life-event-to-symptom" causal relationships (e.g., "Relationship Conflicts and Breakdown" causing "Indecisiveness" with an ATE value of 0.593), we can infer that the user is likely to experience the symptom of "Indecisiveness" even when the user posts nothing. Now we can formalize the method of applying these causal relationships, taking "symptom-to-symptom" relationships as an example.

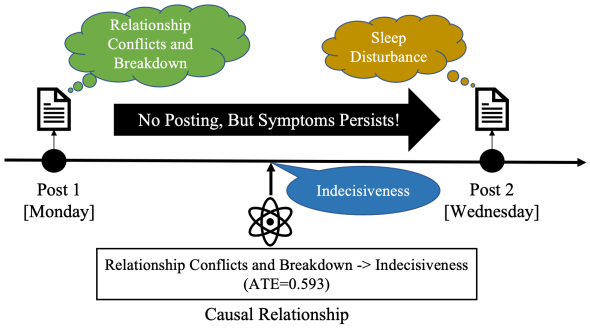

Figure 4: Illustration of incomplete symptom sequence in social media posts, while causal relationships can help to replenish missing symptoms.

Let S_D denote the original symptom vector of a user on day D. The adjusted symptom vector S'_D, considering the causal relationship, is calculated as follows:

$$S'_D = S_D + \frac{1}{W} C_{symp}^T \sum_{i=1}^{W} S_{D-i} \quad (2)$$

Here, C_{symp} represents the causal matrix, with $C_{symp}[i][j]$ indicating the Average Treatment Effect (ATE) of the i-th symptom causing the j-th symptom. The variable W represents the time window. Therefore, the causal matrix can help us predict how much the probabilities of other symptoms will increase or decrease based on the symptom sequences $[S_{D-W}, ..., S_{D-1}]$ from the previous W days,

Similarly, we can adjust the original symptom vector using "life-events-to-symptom" causal relationship:

$$S'_D = S_D + \frac{1}{W} C_{LE}^T \sum_{i=1}^{W} L_{D-i} \quad (3)$$

Here, C_{LE} represents the causal matrix, with $C_{LE}[i][j]$ indicating the ATE of the i-th life event causing the j-th symptom. L_D denotes the life event vector on day D.

4 Experiments

In this section, we conducted experiments to evaluate the effectiveness of applying causal relationships to two tasks: Early Detection of Depression and Diagnosis Point Detection, comparing the results with baseline models.

4.1 Early Risk Detection of Depression (ERD)

Dataset We utilize an ERD dataset proposed by Chen et al. (2023). Users and posts were extracted from a publicly available Reddit corpus. The dataset select depression users by detection patterns which consist of two components: one that matches a self-reported diagnosis (e.g., "diagnosed with"), and another that maps relevant keywords to the depression (e.g., major depressive disorder). Control users (i.e., healthy persons) are randomly sampled from those who never posted or commented in mental health related subreddits. The dataset consists of 3,105 users with depression and 17,209 control users.

Baseline We employ **PsySym** (Zhang et al., 2022b) as baseline, which utilizes CNN of various kernel sizes as backbone, and the inputs are extracted psychiatric symptom features the same as this work. This symptom-based baseline can outperform lots of pure-text methods including BERT-based ones (Nguyen et al., 2022).

Evaluation Metric We use the official metrics $ERDE_5$ and $ERDE_{50}$ for Early Detection task proposed by Losada and Crestani (2016). The lower $ERDE_5$ and $ERDE_{50}$, the better model performs early detection, and $ERDE_5$ has a higher penalty than $ERDE_{50}$ for late detection. Detailed introduction of these metrics can be found in Appendix E.

Experiment Results We conduct three runs for each method using different seed values, and the

results of ERD task is demonstrated in Table 3. We implement *+symp* by adjusting users' original symptom sequences based on "symptom-to-symptom" causality, and *+symp&LE* incorporates both types of causal relationships. Generally, we can see that the early detection result can benefit from our methods that applies causality to fulfill the incomplete symptom sequences.

For the two model variants, they perform comparably on $ERDE_{50}$, while *+symp&LE* perform better on the more latency-sensitive metric $ERDE_5$. For different causal window size, shorter ones (e.g., 30, 90) are more effective in this task emphasizing low latency. This may be attributed to the fact that causality inferred from a shorter window size is more immediate, enabling the timely identification of potential indicators for early detection.

Method	CW	$ERDE_5$	$ERDE_{50}$
baseline	-	13.62±.006	6.72±.105
+symp	30 days	13.49±.013	**6.07**±.092*
	90 days	13.29±.020**	6.36±.220
	180 days	13.58±.028	7.41±.126
+symp&LE	30 days	13.42±.014*	6.27±.158
	90 days	**13.20**±.020**	6.60±.273
	180 days	13.63±.047	6.72±.137

Table 3: Results of ERD Task. "CW" means "causal window", whose definition[6] can be found in Section 2.2.2. *symp* is short for "symptom", and *LE* is short for "life event". The p-values indicating the significance of the differences between the baseline and our method is demonstrated as (*):p<0.1, (**):p<0.05

Case Study To assess the effectiveness of incorporating the causal relationships in ERD, we visualize the predicted risk score of one user before and after the inclusion of both two type of causal relationships in Figure 5. It clearly demonstrates that after incorporation the causal relations, the predicted risk score surpass the detection threshold (0.5) at an earlier stage compared to the baseline model, which means the model can recognize the depression risk earlier. The reason is that the proposed method can recognize indicative life events before symptom manifestations. For example, before point A, the user posted:

"Recently, I left my retail job."

which matches the "*Work and Career Challenges*" life event. The LE-symptom causal relationship will indicate higher risk of depression in the future, facilitating earlier detection. Therefore, even when users do not explicitly express depressive symptoms, our approach, leveraging the association between life events and symptoms, enables us to sensitively capture latent signs of depression.

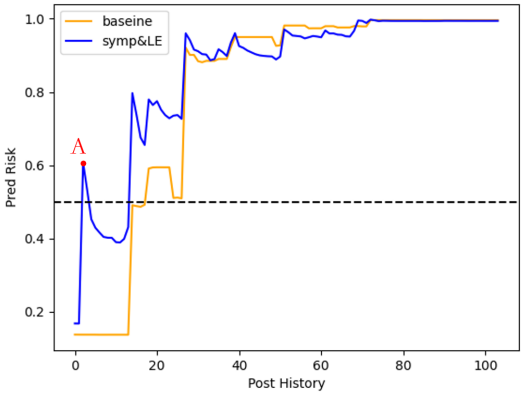

Figure 5: Comparison of the predicted risk along time from two methods on a depression patient in Early Risk Detection. The user has no explicit symptom expression before A, but the symp&LE method can capture earlier sign from LE.

4.2 Diagnosis Point Detection (DPD)

Dataset The dataset of this task is RSDD-Time (MacAvaney et al., 2018), which has been mentioned in Section 3. The dataset is an annotated corpus sourced from Reddit. It encompasses two kinds of text spans: diagnoses (e.g., "I was diagnosed") and temporal expressions associated with the diagnosis (e.g., "today"). Consequently, the diagnosis point label can be derived through the utilization of these annotations.

Baseline The essence of our DPD task is the conventional Change Point Detection Problem (Truong et al., 2020), which aims to identify points in a time series indicative of a significant shift in the underlying data distribution. Therefore, we employ **RuLSIF** (Liu et al., 2013), a well-established CPD model widely recognized for its performance (Hushchyn and Ustyuzhanin, 2021), as our baseline. RuLSIF[7] utilizes a least-squares fitting approach to gauge the dissimilarity between the distributions of successive segments within a time series. When a substantial difference is observed in the distribution between two consecutive segments, a change point is identified. The method offers a non-parametric solution, making minimal assumptions about the underlying data.

[7]Relative unconstrained Least Squares Importance Fitting

Evaluation Metric The DPD task aims to identify the exact diagnosis time of an individual, which is quite challenging by analysing the symptom sequences in the posting history. Consequently, applying strict metrics like accuracy and F1 score directly is impractical in this context. To address this, we calculate the F1 score using a smooth time window, defining true positive (TP) samples as those within a specific temporal proximity to the actual diagnosis time. In this study, we set the time window to 30 days, and we refer to this metric as $F_1(w = 30)$.

Experiment Results Figure 6 illustrates the experiment results of DPD task. We compare the results of baseline (red line) with causality-enhanced methods in various settings. "symp" denotes the inclusion of "symptom-to-symptom" causality, "LE" indicates the inclusion of "life-event-to-symptom" causality, and "symp&LE" encompasses both causality. We can find that causal relationships can significantly improve the detection accuracy of diagnosis point. Interestingly, the "LE" causality generally outperforms "symp" across all window sizes, and the combination of both causalities shows a substantial improvement, especially with a causal window of 180.

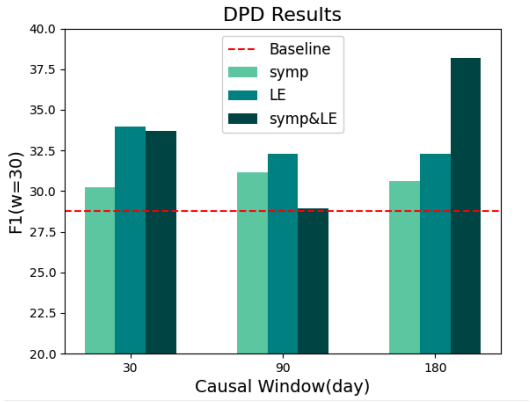

Figure 6: Results of DPD task. The definition of "causal window"[6] can be found in Section 2.2.2. *symp* is short for "symptom", and *LE* is short for "life event".

5 Related Work

In this section, we present relevant literature on causality analysis in the context of social media and mental health.

Causal Analysis Methods Generally, there are two ways to establish causal relationships. The first is Randomized Controlled Trial (RCT), typically applied in strict clinical settings (Cipriani et al., 2018); the second one is observational (Gianicolo et al., 2020), including methods like Propensity Score Matching (PSM) (Rosenbaum and Rubin, 1983) and Regression Discontinuity Design (RDD) (Cattaneo et al., 2019)—strategies that mimic RCT conditions by meticulously controlling numerous covariates.

RCT provides controlled and stringent conditions, ensuring a high level of internal validity. However, their drawback lies in potential limited generalizability (Kennedy-Martin et al., 2015), as the strict conditions and carefully selected participants may not fully mirror real-world diversity. Moreover, RCTs can be resource-intensive and ethically challenging in specific situations, constraining their feasibility for certain research inquiries.

In this study, we adopt the observational method to infer causality from social media data. We do acknowledge that observational studies are weaker than RCTs in making conclusive causal claims, but their validity is supported by statistical literature (Caliendo and Kopeinig, 2005), and they can provide complementary advantages since the analysis is conducted in large population.

Mental-health-related Causality Inference on Social Medial Exploring mental-health-related causal relationships is common in clinical studies (i.e. the theoretical and mechanism research in psychiatric diseases). For example, the Network Theory suggests that mental disorders arise from the causal interplay between symptoms (Borsboom and Cramer, 2013). Additionally, research on the network analysis of depression and anxiety symptoms has revealed potential causal relationships among them, with findings empirically supporting the idea that certain symptoms may act as central hubs, influencing the dynamics of the entire network (Beard et al., 2016).

However, there is a scarcity of prior research utilizing computational methods to infer the causes of specific psychiatric symptoms or mental disorders on social media. Some prior works, such as Saha et al. (2019), utilize PSM to infer causal relationships between psychiatric medication use and symptom outcomes on Reddit Corpus. Additionally, Yuan et al. (2023) also utilize similar method to mine the causality between mental health coping and the severity of mental disorders. However, these studies mainly made qualitative conclusions about the inferred causality, while our work makes

further exploration by utilizing them quantitatively as features for downstream tasks like mental disease detection. What's more, other works (Garg et al., 2022; Saxena et al., 2023) use information retrieval methods to extract causal relationships between stressful events and mental disorders from social media posts. However, these studies focus on extracting direct causal relationships from the semantic information within one post, which may overlook the various long-term, subtle causal relationships that may be absent in a single post.

6 Conclusion

Mental disorders, situated as chronologically evolving diseases, highlight the importance of considering the entire spectrum of symptom development and progression over time. In our study, we delved into such chronological aspects of social media posts, uncovering significant causal relationships between symptoms and life events through an observational causal method, the Propensity Scoring Matching analysis. We identified causal links among 38 symptoms and their connections to 11 life event categories. By corroborating our results on *symptom-to-symptom* and *life event-to-symptom* with existing clinical literature, we provided a direct analysis of the causal relationships identified. These findings were then applied to two practical tasks, namely the Diagnosis Point Detection and Early Risk Detection of Depression. Enhanced performance when incorporating causal features on both tasks suggested the effectiveness and necessity of long-term causing relations. Our research underscores the critical importance of causal relations in understanding the complex interplay between symptoms, life events and mental disorders, thus advancing the science of mental disorder prevention and early detection.

7 Ethical Statement

In this work, we make every effort to minimize the risk of personal privacy leakage during the data collection process. We replaced usernames with random identifiers to prevent identification of users without external information. All datasets used in our study are either publicly available or adhere to their respective licenses. We sign and comply with the data use agreement to prevent privacy infringement or other potential misuses. All posts in examples were de-identified and paraphrased for anonymity. What's more, we carefully considered the application of social media for the detection of mental illnesses. The purpose of this work is not to replace psychiatrists. Instead, we hope our model will be used as an effective auxiliary tool by experienced psychiatrists in the future.

8 Limitations

In our study, there are some limitations that could be addressed in future research:

1. Although the causal relationships between life events and symptoms we identified achieved good results in downstream tasks, and we considered as many common and impactful life events as possible, the 11 categories life events we selected might not cover all events that could potentially affect mental health in life.

2. In addition to studying the causal relationships between life events and symptoms, as well as between symptoms themselves, we could also consider other factors and their causal relationships with mental disorders and symptoms.

3. Exploration of other downstream tasks involving temporal analysis of mental disorders is necessary. We identified diagnosis point detection and early risk detection here while more tasks can benefit from causal relations.

References

Gulistan Agirman, Kristie B. Yu, and Elaine Y. Hsiao. 2021. Signaling inflammation across the gut-brain axis. *Science*, 374(6571):1087–1092.

Gail A Alvares, Daniel S Quintana, Andrew H Kemp, Anita Van Zwieten, Bernard W Balleine, Ian B Hickie, and Adam J Guastella. 2013. Reduced heart rate variability in social anxiety disorder: associations with gender and symptom severity. *PloS one*, 8(7):e70468.

DSMTF American Psychiatric Association, American Psychiatric Association, et al. 2013. *Diagnostic and statistical manual of mental disorders: DSM-5*, volume 5. American psychiatric association Washington, DC.

Courtney Beard, Alex J Millner, Marie JC Forgeard, Eiko I Fried, Kean J Hsu, Michael T Treadway, Chelsea V Leonard, SJ Kertz, and Thröstur Björgvinsson. 2016. Network analysis of depression and anxiety symptom relationships in a psychiatric sample. *Psychological medicine*, 46(16):3359–3369.

Aaron T Beck, Robert A Steer, and Gregory K Brown. 1996. *Beck depression inventory (BDI-II)*. Pearson.

Denny Borsboom and Angélique OJ Cramer. 2013. Network analysis: an integrative approach to the structure of psychopathology. *Annual review of clinical psychology*, 9:91–121.

Lynn Boschloo, Claudia D van Borkulo, Mijke Rhemtulla, Katherine M Keyes, Denny Borsboom, and Robert A Schoevers. 2015. The network structure of symptoms of the diagnostic and statistical manual of mental disorders. *PloS one*, 10(9):e0137621.

Marco Caliendo and Sabine Kopeinig. 2005. Some practical guidance for the implementation of propensity score matching. *IZA Institute of Labor Economics Discussion Paper Series*.

M. D. Cattaneo, Nicolás Idrobo, and Rocío Titiunik. 2019. A practical introduction to regression discontinuity designs.

Stevie Chancellor and Munmun De Choudhury. 2020. Methods in predictive techniques for mental health status on social media: a critical review. *NPJ digital medicine*, 3(1):43.

Susan T Charles, Jennifer R Piazza, Jacqueline Mogle, Martin J Sliwinski, and David M Almeida. 2013. The wear and tear of daily stressors on mental health. *Psychological science*, 24(5):733–741.

Siyuan Chen, Zhiling Zhang, Mengyue Wu, and Kenny Zhu. 2023. Detection of multiple mental disorders from social media with two-stream psychiatric experts. In *Proceedings of the 2023 Conference on Empirical Methods in Natural Language Processing*, pages 9071–9084, Singapore. Association for Computational Linguistics.

Carol Chu, Jennifer M Buchman-Schmitt, Fallon B. Moberg, and Thomas E. Joiner. 2016. Thwarted belongingness mediates the relationship between fear of negative evaluation and suicidal ideation. *Cognitive Therapy and Research*, 40:31–37.

Andrea Cipriani, Toshi A Furukawa, Georgia Salanti, Anna Chaimani, Lauren Z Atkinson, Yusuke Ogawa, Stefan Leucht, Henricus G Ruhe, Erick H Turner, Julian PT Higgins, et al. 2018. Comparative efficacy and acceptability of 21 antidepressant drugs for the acute treatment of adults with major depressive disorder: a systematic review and network meta-analysis. *The Lancet*, 391(10128):1357–1366.

Graham L Collingridge, Stephane Peineau, John G Howland, and Yu Tian Wang. 2010. Long-term depression in the cns. *Nature reviews neuroscience*, 11(7):459–473.

Caitlin Dreisbach, Theresa A. Koleck, Philip E. Bourne, and Suzanne Bakken. 2019. A systematic review of natural language processing and text mining of symptoms from electronic patient-authored text data. *International Journal of Medical Informatics*, 125:37–46.

Anke Ehlers. 1993. Interoception and panic disorder. *Advances in Behaviour Research and Therapy*, 15(1):3–21.

Amir Feder, Katherine A. Keith, Emaad Manzoor, Reid Pryzant, Dhanya Sridhar, Zach Wood-Doughty, Jacob Eisenstein, Justin Grimmer, Roi Reichart, Margaret E. Roberts, Brandon M. Stewart, Victor Veitch, and Diyi Yang. 2022. Causal Inference in Natural Language Processing: Estimation, Prediction, Interpretation and Beyond. *Transactions of the Association for Computational Linguistics*, 10:1138–1158.

Muskan Garg, Chandni Saxena, Veena Krishnan, Ruchi Joshi, Sriparna Saha, Vijay Mago, and Bonnie J Dorr. 2022. CAMS: an annotated corpus for causal analysis of mental health issues in social media posts. *arXiv preprint arXiv:2207.04674*.

Emilio A L Gianicolo, Martin Eichler, Oliver Muensterer, Konstantin Strauch, and Maria Blettner. 2020. Methods for evaluating causality in observational studies. *Deutsches Arzteblatt international*, 116(7):101—107.

Aurélie Grandjean, Isabel Suarez, Elisa Diaz, Laure Spieser, Boris Burle, Agnès Blaye, and Laurence Casini. 2021. Stronger impulse capture and impaired inhibition of prepotent action in children with adhd performing a simon task: An electromyographic study. *Neuropsychology*, 35(4):399.

Tayebeh Fasihi Harandi, Maryam Mohammad Taghinasab, and Tayebeh Dehghan Nayeri. 2017. The correlation of social support with mental health: A meta-analysis. *Electronic physician*, 9(9):5212.

Mikhail Hushchyn and Andrey Ustyuzhanin. 2021. Generalization of change-point detection in time series data based on direct density ratio estimation. *Journal of Computational Science*, 53:101385.

Guido W. Imbens and Donald B. Rubin. 2015. *Causal Inference for Statistics, Social, and Biomedical Sciences: An Introduction*. Cambridge University Press.

Nicholas C Jacobson and Michelle G Newman. 2017. Anxiety and depression as bidirectional risk factors for one another: A meta-analysis of longitudinal studies. *Psychological bulletin*, 143(11):1155.

Shaoxiong Ji, Tianlin Zhang, Luna Ansari, Jie Fu, Prayag Tiwari, and Erik Cambria. 2022. MentalBERT: Publicly available pretrained language models for mental healthcare. In *Proceedings of the Thirteenth Language Resources and Evaluation Conference*, pages 7184–7190, Marseille, France. European Language Resources Association.

Sukjae J Kang, Jong-Hyun Kim, Dong-Il Kim, Benjamin Z Roberts, and Sung Han. 2024. A pontomesencephalic pacapergic pathway underlying panic-like behavioral and somatic symptoms in mice. *Nature Neuroscience*, pages 1–12.

Tessa Kennedy-Martin, Sarah E. Curtis, Douglas E Faries, Susan Robinson, and Joseph A. Johnston. 2015. A literature review on the representativeness of randomized controlled trial samples and implications for the external validity of trial results. *Trials*, 16.

Jacqueline HJ Kim, William Tsai, Tamar Kodish, Lam T Trung, Anna S Lau, and Bahr Weiss. 2019. Cultural variation in temporal associations among somatic complaints, anxiety, and depressive symptoms in adolescence. *Journal of psychosomatic research*, 124:109763.

Heide Klumpp, Fini Chang, Brian W Bauer, and Helen J Burgess. 2023. Objective and subjective sleep measures are related to suicidal ideation and are transdiagnostic features of major depressive disorder and social anxiety disorder. *Brain sciences*, 13(2):288.

Deniz Konac, Katherine S Young, Jennifer Lau, and Edward D Barker. 2021. Comorbidity between depression and anxiety in adolescents: Bridge symptoms and relevance of risk and protective factors. *Journal of psychopathology and behavioral assessment*, 43:583–596.

Song Liu, Makoto Yamada, Nigel Collier, and Masashi Sugiyama. 2013. Change-point detection in time-series data by relative density-ratio estimation. *Neural Networks*, 43:72–83.

David E Losada and Fabio Crestani. 2016. A test collection for research on depression and language use. In *International conference of the cross-language evaluation forum for European languages*, pages 28–39. Springer.

Zhiyi Luo, Yuchen Sha, Kenny Q. Zhu, Seung-won Hwang, and Zhongyuan Wang. 2016. Commonsense causal reasoning between short texts. In *Fifteenth international conference on the principles of knowledge representation and reasoning*.

Sean MacAvaney, Bart Desmet, Arman Cohan, Luca Soldaini, Andrew Yates, Ayah Zirikly, and Nazli Goharian. 2018. RSDD-time: Temporal annotation of self-reported mental health diagnoses. In *Proceedings of the Fifth Workshop on Computational Linguistics and Clinical Psychology: From Keyboard to Clinic*, pages 168–173, New Orleans, LA. Association for Computational Linguistics.

David A Moscovitch, Vanja Vidovic, Ariella P Lenton-Brym, Jessica R Dupasquier, Kevin C Barber, Taylor Hudd, Nick Zabara, and Mia Romano. 2018. Autobiographical memory retrieval and appraisal in social anxiety disorder. *Behaviour Research and Therapy*, 107:106–116.

Thong Nguyen, Andrew Yates, Ayah Zirikly, Bart Desmet, and Arman Cohan. 2022. Improving the generalizability of depression detection by leveraging clinical questionnaires. In *Proceedings of the 60th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, pages 8446–8459, Dublin, Ireland. Association for Computational Linguistics.

Peter Noone. 2017. The holmes–rahe stress inventory. *Occupational Medicine*, 67.

Massimo Pasquini, Daria Piacentino, Isabella Berardelli, Valentina Roselli, Annalisa Maraone, Lorenzo Tarsitani, and Massimo Biondi. 2015. Fatigue experiences among ocd outpatients. *Psychiatric Quarterly*, 86:615–624.

Milen Radell, Eid Abo Hamza, Wid Daghustani, Asma Perveen, and Ahmed Moustafa. 2021. The impact of different types of abuse on depression. *Depression Research and Treatment*, 2021:1–12.

Paul R. Rosenbaum and Donald B. Rubin. 1983. The central role of the propensity score in observational studies for causal effects. *Biometrika*, 70:41–55.

Chidchanok Ruengorn, Ratanaporn Awiphan, Nahathai Wongpakaran, Tinakon Wongpakaran, Surapon Nochaiwong, and for the Health Outcomes and Mental Health Care Evaluation Survey Research Group (HOME-Survey) . 2021. Association of job loss, income loss, and financial burden with adverse mental health outcomes during coronavirus disease 2019 pandemic in thailand: A nationwide cross-sectional study. *Depression and Anxiety*, 38(6):648–660.

Koustuv Saha, Benjamin Sugar, John Torous, Bruno Abrahao, Emre Kıcıman, and Munmun De Choudhury. 2019. A social media study on the effects of psychiatric medication use. *Proceedings of the International AAAI Conference on Web and Social Media*, 13(01):440–451.

Chandni Saxena, Muskan Garg, and Gunjan Ansari. 2023. Explainable causal analysis of mental health on social media data. In *Neural Information Processing: 29th International Conference, ICONIP 2022, Virtual Event, November 22–26, 2022, Proceedings, Part II*, page 172–183, Berlin, Heidelberg. Springer-Verlag.

Rosa J Seinsche, Susanne Fricke, Marie K Neudert, Raphaela I Zehtner, Bertram Walter, Rudolf Stark, and Andrea Hermann. 2023. Memory representation of aversive social experiences in social anxiety disorder. *Journal of Anxiety Disorders*, 94:102669.

Anish Shah, Viola Vaccarino, Yi-An Ko, Zakaria Almuwaqqat, Mariana Garcia, Kasra Moazzami, Maggie Wang, Oleksiy Levantsevych, an young, Laura Ward, Jonathon Nye, Paolo Raggi, David S. Sheps, Rachel J. Lampert, Doug Bremner, Yan Sun, Ernest V. Garcia, Arshed A. Quyyumi, and Amit J. Shah. 2023. Mental stress-induced autonomic dysfunction is associated with cardiovascular mortality. *Journal of the American College of Cardiology*, 81(8_Supplement):1146–1146.

Judy Hanwen Shen and Frank Rudzicz. 2017. Detecting anxiety through reddit. In *Proceedings of the Fourth Workshop on Computational Linguistics and Clinical Psychology—From Linguistic Signal to Clinical Reality*, pages 58–65.

Charles Truong, Laurent Oudre, and Nicolas Vayatis. 2020. Selective review of offline change point detection methods. *Signal Process.*, 167.

Irina A Vanzhula, Benjamin Calebs, Laura Fewell, and Cheri A Levinson. 2019. Illness pathways between eating disorder and post-traumatic stress disorder symptoms: Understanding comorbidity with network analysis. *European Eating Disorders Review*, 27(2):147–160.

Justin W Weeks and Peggy M Zoccola. 2016. Fears of positive versus negative evaluation: Distinct and conjoint neuroendocrine, emotional, and cardiovascular responses to social threat. *Journal of Experimental Psychopathology*, 7(4):632–654.

Andrew Yates, Arman Cohan, and Nazli Goharian. 2017. Depression and self-harm risk assessment in online forums. In *Proceedings of the 2017 Conference on Empirical Methods in Natural Language Processing*, pages 2968–2978, Copenhagen, Denmark. Association for Computational Linguistics.

Yunhao Yuan, Koustuv Saha, Barbara Keller, Erkki Tapio Isometsä, and Talayeh Aledavood. 2023. Mental health coping stories on social media: A causal-inference study of papageno effect. In *Proceedings of the ACM Web Conference 2023*, WWW '23, page 2677–2685, New York, NY, USA. Association for Computing Machinery.

Zhiling Zhang, Siyuan Chen, Mengyue Wu, and Kenny Q. Zhu. 2022a. Psychiatric scale guided risky post screening for early detection of depression. In *Proceedings of the Thirty-First International Joint Conference on Artificial Intelligence, IJCAI-22*, pages 5220–5226. International Joint Conferences on Artificial Intelligence Organization. AI for Good.

Zhiling Zhang, Siyuan Chen, Mengyue Wu, and Kenny Q. Zhu. 2022b. Symptom identification for interpretable detection of multiple mental disorders on social media. In *Proceedings of the 2022 Conference on Empirical Methods in Natural Language Processing*, page 9970–9985. Association for Computational Linguistics.

A Symptom and Life Event Definition

We consider a total of 7 mental disorders and 38 symptoms following Zhang et al. (2022b), which are deduced and integrated from DSM-5 (American Psychiatric Association et al., 2013). The 7 disorders and their representative symptoms are listed in the Table 4. And all of the 38 symptoms are listed in the Table 6. Additionally, we merge the 43 life events in Holmes-Rahe Stress Inventory (Noone, 2017) into 11 classes of life events, as illustrated in Table 7.

Disease	Typical Symptoms
Anxiety	anxious mood;panic fear
ADHD	inattention;hyperactivity;impulsivity
Bipolar Disorder	drastic shift in mood and energy
Depression	depressed mood;suicidal ideas
Eating Disorder	compensatory behaviors to prevent weight gain
OCD	obsession;compulsions
PTSD	intrusion symptoms;sleep disturbance

Table 4: 7 Diseases and their Representative Symptoms

Life events	kappa
Loss and Bereavement	0.91
Marriage and Commitment	0.91
Relationship Conflicts and Breakdown	0.89
Family Additions and Departures	0.37
Health and Well-being	0.91
Work and Career Challenges	0.98
Financial Challenges	0.89
Education Transitions	0.92
Change in Living Environment and Habits	0.82
Vacations and Holidays	0.92
Legal Matters	0.93
Average	0.86

Table 5: Agreement (Fleiss' Kappa) of three annotator in the annotation of Life Event Dataset

B Life Events Dataset

We adopted an annotation similar to that of previous work (Zhang et al., 2022b). Our life event dataset contains 2643 posts related to one or more life events, alongside 5000 control posts (i.e., posts unrelated to any life event). We engaged three experienced annotator for the task, and their agreement (Fleiss' Kappa) is showed in Table5.

C Detailed Symptom and Life Event Identification Results

The detailed identification results of Symptom and Life Event Identification Models are illustrated in Figure 7 and Figure 8 respectively. The high auc and F1 scores show that using these classifiers can help us automatically and accurately extract psychiatric symptoms and life events on Reddit corpus.

D Causality Results

Here, we present the results of two types of causality in the form of a heatmap. Figure 10 shows the causality between symptoms and Figure 11 shows the causality between LEs and symptoms. These numbers in figures indicate ATE values of the causal relationships.

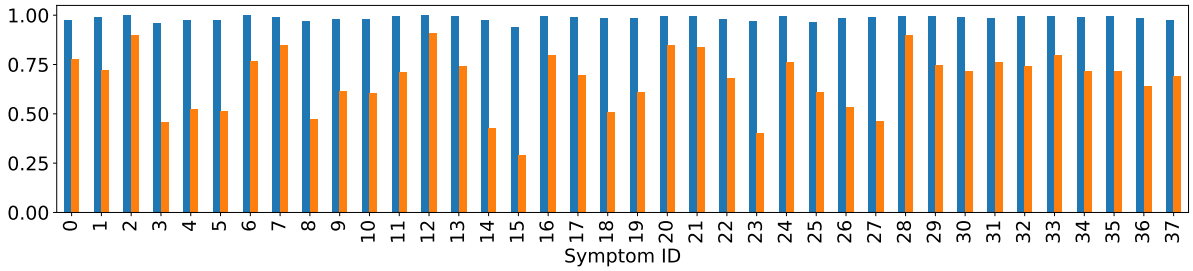

Figure 7: Identification performance of each symptom. The blue bar shows the AUC while the orange bar shows F1, and Symptom ID follows the order of Table 6.

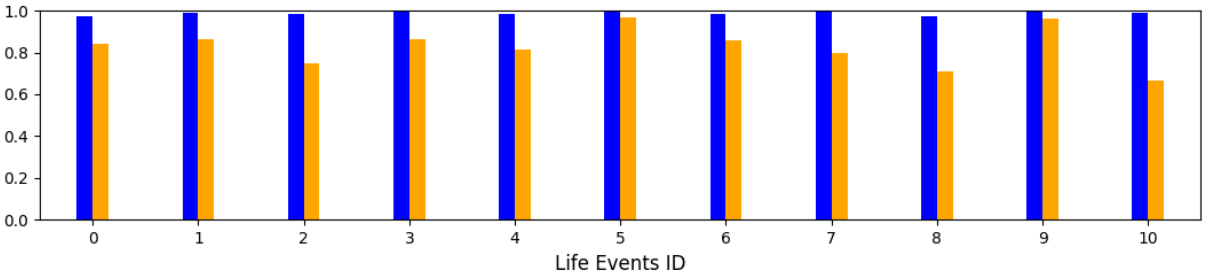

Figure 8: Identification performance of each life event. The blue bar shows the AUC while the orange bar shows F1, and Life Event ID follows the order of Table 7.

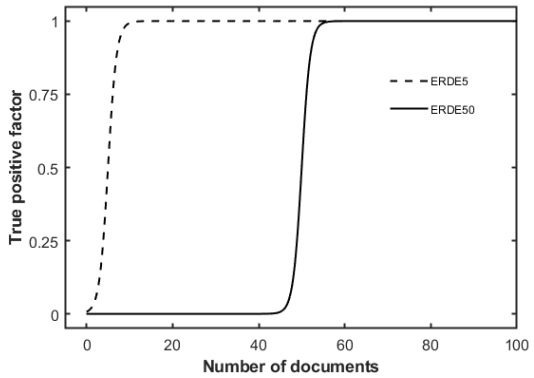

Figure 9: the cost factor $lc_o(k)$ for $ERDE_5$ and $ERDE_{50}$.

E Evaluation Metrics

The following will introduce evaluation metrics for two downstream tasks.

$$Precision = \frac{TP}{TP + FP}$$

$$Recall = \frac{TP}{TP + FN}$$

where TP represents true positives (samples correctly predicted as positive),TN represents true negatives (samples correctly predicted as negative), FP represents false positives (samples incorrectly predicted as positive), and FN represents false negatives (samples incorrectly predicted as negative). Diagnosis Point Detection uses the F1 score as a metric, which balances precision and recall.

$$F1 - Score = 2 \times \frac{Precision \times Recall}{Precision + Recall}$$

The metric used for early detection of depression, Early Risk Detection Error (ERDE)measure, is defined as follows:

$$ERDE_o(d, k) = \begin{cases} c_{fp} & for\ FP \\ c_{fn} & for\ FN \\ lc_o(k) \cdot c_{tp} & for\ TP \\ 0 & for\ TN \end{cases}$$

c_{fp} and c_{fn} are used to adjust the severity of false positives (FP) and false negatives (FN). c_{fn} was set to 1, while c_{fp} is set to the ratio of the number of positive cases in the data to the total number of users. $lc_o(k)$ ($\in [0, 1]$) encodes the cost of delaying the detection of true positives (TP), and c_{tp} defines the level of penalty for delaying TP. Setting c_{tp} to 1 and it means that delaying detection is equivalent to not detecting the case. The function $lc_o(k)$ determines after how many posts k the cost of true positives starts to increase and is defined as follows:

$$lc_o(k) = 1 - \frac{1}{1 + e^{k-o}}$$

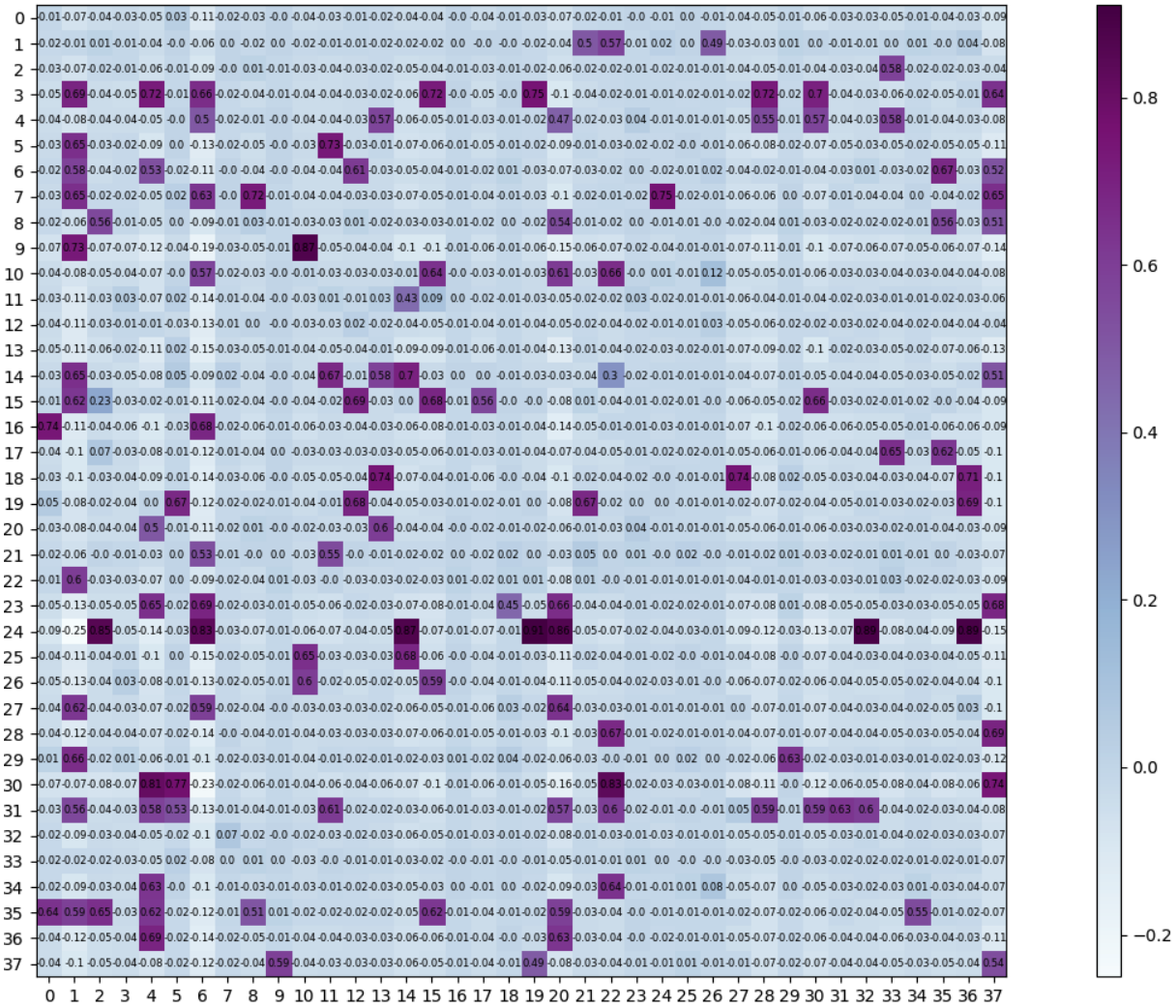

Figure 10: The causal matrix between symptoms with the time window of 180 days. Symptom ID follows the order of Table 6.

The parameter o controls the position on the x-axis where the cost increases more rapidly.We use $ERDE_5$ and $ERDE_{50}$ as evaluation metrics for the results of early detection of depression,shown in Figure 9.

F Experiment Setting

In our ERD experiment, we trained the baseline model with the dataset proposed by Zhang et al. (2022b) that originated from a publicly available Reddit corpus. The training process employs a batch size of 64 and learning rate of 0.01. We used symptom features only and the posting list will be limited to a maximum of 256. To prevent over-fitting, we implement early-stopping based on validation performance, with a patience of 4 epochs.

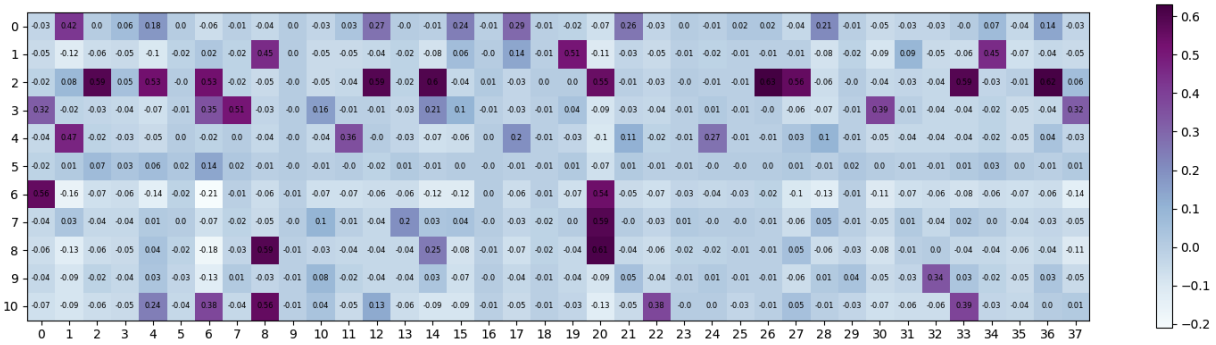

Figure 11: The causal matrix between LEs and symptoms with the time window of 1 year. Symptom ID and LE ID follow the order of Table 6 and Table 7.

id	Symptom
1	Anger Irritability
2	Anxious Mood
3	Autonomic symptoms
4	Cardiovascular symptoms
5	Catatonic behavior
6	Decreased energy tiredness fatigue
7	Depressed Mood
8	Gastrointestinal symptoms
9	Genitourinary symptoms
10	Hyperactivity agitation
11	Impulsivity
12	Inattention
13	Indecisiveness
14	Respiratory symptoms
15	Suicidal ideas
16	Worthlessness and guilty
17	Avoidance of stimuli
18	Compensatory behaviors to prevent weight gain
19	Compulsions
20	Diminished emotional expression
21	Do things easily get painful consequences
22	Drastic shift in mood and energy
23	Fear about social situations
24	Fear of gaining weight
25	Fears of being negatively evaluated
26	Flight of ideas
27	Intrusion symptoms
28	Loss of interest or motivation
29	More talkative
30	Obsession
31	Panic fear
32	Pessimism
33	Poor memory
34	Sleep disturbance
35	Somatic muscle
36	Somatic symptoms others
37	Somatic symptoms sensory
38	Weight and appetite change

Table 6: Id and its corresponding symptoms

id	Life Event Categories	Original Life Events
1	Loss and Bereavement	Death of a spouse; Death of a close family member; Death of a close friend
2	Marriage and Commitment	Marriage; Marital reconciliation
3	Relationship Conflicts and Breakdown	Divorce; Marital separation; Change in number of arguments with spouse; Trouble with in-laws
4	Family Additions and Departures	Son or daughter leaving home; Gain of new family member
5	Health and Well-being	Personal injury or illness; Sex difficulties; Pregnancy; Change in health of family member
6	Work and Career Challenges	Fired at work;Retirement; Change in responsibilities at work; Change to a different line of work; Spouse begins or stops work; Trouble with boss; Change in work hours or conditions; Business readjustment
7	Financial Challenges	Change in financial state; A large mortgage or loan; Foreclosure of mortgage or loan; A moderate loan or mortgage
8	Education Transitions	Begin or end school/college; Change in school/college
9	Change in Living Environment and Habits	Change in living conditions; Revision of personal habits; Change in sleeping habits; Change in eating habits; Change in church activities; Change in residence; Change in recreation; Change in social activities; Change in number of family get-togethers
10	Vacations and Holidays	Vacation; Christmas
11	Legal Matters	Jail term; Minor violations of the law

Table 7: All life events and the 11 major categories