Multistream Modelling for Mental Health: Modelling Linguistic and Temporal Contexts with Mutual and Self-Excitation in Social Media

Anthony Hills[1], **Talia Tseriotou**[1], **Mahmud Elahi Akhter**[1], **Junyu Mao**[5], **Iqra Ali**[1], **Xenia Miscouridou**[3,4], **Maria Liakata**[1,2]
[1]Queen Mary University of London, [2]The Alan Turing Institute, [3]University of Cyprus, [4]Imperial College London, [5]University of Southampton
{a.r.hills,t.tseriotou,m.liakata}@qmul.ac.uk

Abstract

We present MHRoBERT (Multistream HEAT over Recurrence over BERT), a hierarchical transformer architecture for longitudinal mental health monitoring that models self- and mutual excitation patterns in linguistic and temporal data across multivariate event streams relating to an individual's mental health. To supply the model with complementary perspectives on each post, we apply a Large Language Model (LLM) based annotation to extract three streams from social media posts: emotional states, personal life events, and mental health symptoms. A central finding is that multi-task learning with these automatically-generated stream labels provides substantial, consistent improvements across all model architectures evaluated. Multistream information further consistently benefits simpler models not explicitly designed to exploit it: LLM baselines incorporating stream annotations improve macro F1 by 12.6% over text-only prompting. These results have direct implications on Moments of Change detection: multistream auxiliary supervision yields consistent, substantial gains regardless of architecture, suggesting it is a simple and portable strategy that future systems can readily adopt with minimal architectural changes. MHRoBERT additionally produces interpretable learned parameters across streams, revealing temporal interaction patterns between mental health indicators.

1 Introduction

Mental health conditions manifest through multiple, interconnected signals – from emotional expressions and behavioural changes to physical symptoms and life events. Despite that, traditional approaches to analysing and modeling social media for mental health monitoring examine individual behavioural signals in isolation, without modelling the cross-stream dependencies and temporal interactions between them (Bao et al., 2024). Network theory of psychopathology posits that mental health conditions emerge from systems of mutually reinforcing symptoms that interact causally over time (Borsboom and Cramer, 2013; Borsboom, 2017), while subsequent work demonstrates that emotions form dynamic, mutually influencing networks over time (Bringmann et al., 2013) and that the interplay between positive and negative affect carries clinically meaningful information about mental health trajectories (Wichers et al., 2012). On social media, prior work has largely focused on individual behavioural signals without modelling how different streams co-evolve and mutually interact over time (Yazdavar et al., 2020; Garg, 2023), leaving important cross-stream temporal dynamics unaddressed.

In this work, we leverage the stream taxonomy inspired by Mao et al. (2025) to obtain three complementary perspectives on each post – emotional states, personal life events, and mental health symptoms – via LLM-based annotation, and investigate how these streams can be integrated into temporal modelling architectures for Moments of Change (MoC) in mood detection (Tsakalidis et al., 2022b,a; Tseriotou et al., 2025). A widely applicable finding is that multi-task learning with these automatically-generated stream labels provides substantial, consistent performance improvements across all evaluated model architectures. Multistream information widely benefits diverse architectures, with LLM baselines showing substantial improvements when provided with stream annotations. These results suggest that incorporating multistream auxiliary supervision is a broadly applicable strategy for improving MoC detection. Building on this, we present MHRoBERT (Multistream HEAT over Recurrence over BERT), a hierarchical transformer architecture that captures cross-stream temporal interactions through multistream HEAT layers and general temporal dynamics through unistream aggregation, while maintaining direct access to sequential information via residual connections. The main contributions of

Proceedings of the 11th Workshop on Computational Linguistics and Clinical Psychology (CLPsych 2026), pages 56–75

this work are:

- We demonstrate that multi-task learning with automatically-generated multistream labels – derived using LLM annotation inspired by the taxonomy of Mao et al. (2025) consistently and substantially improves MoC detection across diverse model architectures, including those not explicitly designed to use this information.
- MHRoBERT, a hierarchical architecture that models both self-excitation within individual streams and mutual excitation between streams, capturing nuanced temporal interactions between longitudinal mental health event data.
- Interpretable learned parameters (ϵ, β, μ) that reveal temporal interactions between streams.
- A recommendation for future CLPsych shared task submissions: our results suggest multistream auxiliary supervision is a practical, architecture-agnostic strategy for improving performance on Moments of Change detection.

Through this multistream approach, we identify patterns in how different aspects of mental health interact and evolve over time, providing a more comprehensive view of mental health progression than is possible from text alone.

2 Related Work

Multivariate Hawkes Processes Multivariate Hawkes processes (Hawkes, 1971) are the theoretical foundation for modeling mutual excitation between event streams. Zhou et al. (2013) developed efficient learning algorithms for large-scale applications, while more recently, Zhang et al. (2020) proposed the self-attentive Hawkes process, demonstrating how attention mechanisms could capture complex dependencies between different event types. In the social media domain, Rizoiu et al. (2017) used Hawkes processes to model information diffusion, demonstrating how different types of social interactions could mutually reinforce each other.

LLMs for Automated Annotation The use of LLMs for textual annotation is an emerging area and shows significant potential to create large-scale, consistently labeled datasets at low cost (Tseng et al., 2025; Meng et al., 2022). Recent work shows that LLMs can match or exceed performance of crowd workers in text annotation tasks across multiple domains (Gilardi et al., 2023; Törnberg, 2023). Li and Conrad (2024) showed that LLMs are effective when annotating explicit stances in social media posts, achieving high agreement with human annotators. However, discrepancies arise in cases where stances are implicit or ambiguous, suggesting hybrid approaches may be required for nuanced annotation scenarios. Additionally, Kristensen-McLachlan et al. (2023) raised concerns about transparency and reproducibility, highlighting significant variability in LLM performance across different tasks and noting that traditional supervised classifiers often outperform LLMs in scenarios requiring nuanced understanding. This is particularly important in mental health, where even subtle distinctions can be clinically significant.

Longitudinal Modeling for Mental Health LLMs have been used for mental health classification (Amin et al., 2023), data augmentation (Liyanage et al., 2023), and reasoning (Xu et al., 2024), demonstrating promise in detecting psychological indicators (Yang et al., 2023), extracting relevant evidence from text (Xu et al., 2024), and generating clinically informed summaries (Song et al., 2024). LLMs using instruction fine-tuning and Chain-of-Thought (CoT) prompting (Yang et al., 2023) have also been employed, though such approaches risk incorrect predictions and flawed reasoning, especially in complex conversations (Li et al., 2023). Longitudinal modeling with LLMs has emerged as a pivotal paradigm for capturing language, behavior, and mental states over time. Unlike conventional NLP tasks that treat inputs as isolated instances, longitudinal modeling enable continuous monitoring and nuanced understanding of dynamic processes such as user stance changes on social media or psychological state progression in health narratives (Park and Conway, 2017; Tsakalidis et al., 2022a; Tseriotou et al., 2023, 2024; Hills et al., 2024; Morini et al., 2025), In dialogue systems (Kwak et al., 2023), longitudinal context is essential for maintaining personalization (Chen et al., 2023), and psychological grounding across multiple conversational turns (Wang et al., 2023; Zheng et al., 2023). Current approaches have not been integrated with LLMs and focus on a single stream of data (unistream).

3 Methodology

We first discuss how we obtain multiple interacting data streams from a single data stream (§3.1). These are expected to influence each other but follow different temporal patterns. We then present a multistream temporal architecture, with cross-excitation for capturing interactions between streams (§3.2)

3.1 Stream Extraction using LLMs

We use LLMs to extract three complementary perspectives from social media posts (for the full taxonomies and prompts used to extract these, see Appendix A), namely:

Emotional States: For emotions, we adopt the taxonomy of the fine-grained labels for Affect proposed in the MIND framework (Slonim, 2024), used to label the same dataset we make use of in this work in the CLPSych 2025 shared task (Tseriotou et al., 2025). These consist of twelve intensive emotions (e.g. Happy, Proud, Sad) spanning adaptive and maladaptive affective states.

Personal Life Events: To capture significant life changes that may impact mental health, we employed the broad categories proposed in Mao et al. (2025). These categories, designed specifically for mental health transitions, were inspired by the Major Life Events Taxonomy of Haimson et al. (2021), which identifies 121 life event types across 12 categories, including health, relationships, career, and identity changes.

Symptoms: We draw on a subset of the symptom categories introduced in Mao et al. (2025) to capture both physical and psychological labels related to mental health changes. The taxonomy encompasses categories that commonly manifest during periods of mental health change – including substance use patterns, trauma responses, physical symptoms, sleep disturbances, and appetite changes

This overall categorization of the streams was designed to identify warning signs and behavioral changes that could indicate changes in mood and mental health, while remaining broad enough to capture diverse conditions.

3.1.1 Prompt Design and Language Model

We implement our automated stream annotation using Qwen 2.5-32B-Instruct (Yang et al., 2025) in a chain-of-thought (CoT) few-shot classification setting. Our selection of Qwen 2.5 was informed by comparative evaluations in Mao et al. (2025) demonstrating superior performance over other open-source models for mental health annotation tasks. A key consideration was the model's ability to run locally – enabling all processing to be performed on-premises without sending sensitive user data to external servers. This is in contrast to cloud-based language models that require data to be sent to remote services, making Qwen 2.5 particularly suitable for applications involving sensitive mental health data. For each stream, we used CoT prompting strategies (detailed in Appendix A) that guide the model through explicit reasoning before producing final annotations. Our multi-label prompt design was adapted from the approach of Mao et al. (2025), in which the model provides its reasoning for each category along with an explicit "yes" or "no" judgment before predicting the final answer label, thereby improving annotation quality and interpretability.

3.2 Multistream HEAT Architecture

Our model extends HEAT (Historical Emotional AggregaTion), a temporal aggregation mechanism inspired by the Hawkes process. Originally proposed by Sawhney et al. (2021b) for BERT encodings and adapted by Hills et al. (2024) for LSTM hidden states with learnable self-excitation (ϵ) and decay (β) parameters, HEAT weights historical representations by temporal proximity. The self-excitation parameter ϵ controls how much previous events influence future representations, while the decay parameter β determines how quickly this influence diminishes over time. We modify this approach to handle multiple interacting streams of temporally sensitive linguistic data, by processing user timelines through three main stages:

Linguistic Encoding Posts are first encoded using BERT (`bert-base-uncased`) to obtain contextual embeddings. For each post i in a timeline, BERT produces a representation $e^{(i)} \in \mathbb{R}^{768}$ by extracting the [CLS] token embedding.

Sequential Modeling The BERT embeddings are processed through an LSTM layer to capture sequential dependencies:

$$h^{(i)} = \text{LSTM}(e^{(i)}, h^{(i-1)}) \tag{1}$$

where $h^{(i)} \in \mathbb{R}^{d_{lstm}}$ represents the hidden state encoding both linguistic content and sequential context up to post i.

Multistream HEAT Representation For each target stream l, we compute a multistream HEAT representation that captures cross-stream temporal interactions:

$$H_{\text{multi}}^{(i,l)} = \mu_l + \sum_m \sum_{j<i} h^{(j)} \cdot \mathbb{I}_m^{(j)} \cdot \epsilon_{(l,m)} \cdot e^{-\beta_{(l,m)}\Delta\tau_j} \tag{2}$$

where: μ_l is the learned baseline intensity for target stream l, $\mathbb{I}_m^{(j)} \in \{0, 1\}$ indicates whether post j

belongs to stream m, $\epsilon_{(l,m)} \in \mathbb{R}^+$ captures the excitation from stream m to stream l, $\beta_{(l,m)} \in \mathbb{R}^+$ controls the temporal decay rate of this excitation and $\Delta\tau_j = t_i - t_j$ is the time elapsed since post j. The learned ϵ and β matrices reveal interpretable patterns of how different mental health indicators influence each other temporally.

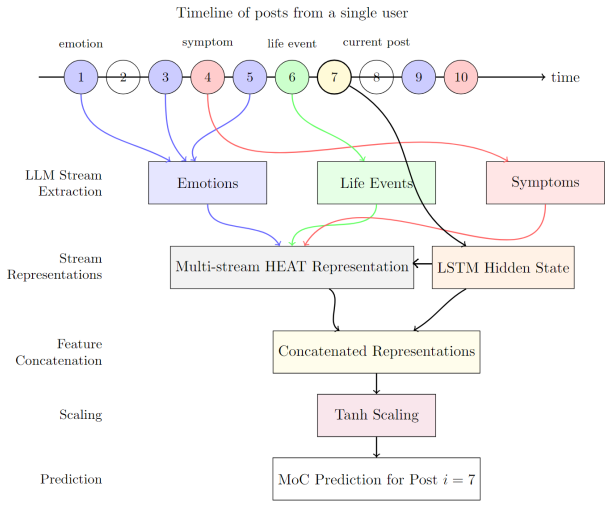

Figure 1: MHRoBERT architecture for MoC prediction. Posts are encoded with BERT and processed through an LSTM to capture sequential context. An LLM annotates posts with emotion, life event, and symptom labels. The final representation concatenates three components: (1) a residual connection from the current post's LSTM state, (2) a unistream representation aggregating all LSTM states via HEAT, and (3) a multistream representation where HEAT pools LSTM states masked by their stream labels. This combined representation predicts whether a given post indicates a MoC.

Unistream HEAT Representation In addition to the multistream representation, we compute a unistream HEAT aggregation over all posts, regardless of stream assignment. This uses a Markovian emphasis (Hills et al., 2024) that weights recent context more heavily:

$$H_{\text{uni}}^{(i)} = h^{(i-1)} + \sum_{j<i} h^{(j)} \cdot \epsilon_{\text{uni}} \cdot e^{-\beta_{\text{uni}}\Delta\tau_j} \quad (3)$$

where ϵ_{uni} and β_{uni} are separate learnable parameters which aggregate historical representations regardless of stream labels, in the unistream fashion.

Combined Representation The final representation concatenates three components:

$$\mathbf{r}^{(i)} = [H_{\text{multi}}^{(i,1)}; H_{\text{multi}}^{(i,2)}; H_{\text{multi}}^{(i,3)}; H_{\text{uni}}^{(i)}; h^{(i)}] \quad (4)$$

where the semicolon denotes concatenation. This design captures: (a) Stream-specific temporal patterns (multistream HEAT); (b) General temporal dynamics (unistream HEAT); (c) Direct sequential information (residual LSTM connection) (See Fig. 1). The concatenated representation is scaled with $\tanh$ activation and passed through a linear classifier:

$$\hat{y}^{(i)} = \text{softmax}(W_{\text{out}} \cdot \tanh(\mathbf{r}^{(i)}) + b_{\text{out}}) \quad (5)$$

4 Experiments

Dataset and Preprocessing We make use of the Reddit dataset introduced in Tsakalidis et al. (2022a), containing longitudinal social media timelines labelled for Moments of Change (MoC). The dataset comprises 186 users across 255 timelines and 6,195 posts, with a median of 18 posts per timeline and a median inter-post gap of 22.72 hours (Hills et al., 2024). Given a historical sequence of posts, the task is to identify whether a post denotes a Switch (S), an Escalation (E), or Neither (O). Posts were annotated by four annotators following the scheme of Tsakalidis et al. (2022b), with final labels determined by majority vote; inter-annotator agreement, measured via Intersection over Union, was 0.264 (S), 0.309 (E), and 0.832 (O) (Tsakalidis et al., 2022a). The dataset is highly imbalanced, with Switch (S: 6.6%) and Escalation (E: 15.8%) events representing a small minority compared to No Change posts (O: 77.6%), motivating our use of macro-averaged F1 as the primary evaluation metric and focal loss during training. Training and implementation details are described in Section C.

Model Configuration We implement and evaluate three variants of hierarchical architectures:
RoBERT: "Recurrence over BERT". Baseline with BERT encoding and LSTM sequential modeling, without temporal aggregation layers. The classifier operates directly on LSTM hidden states.
HoRoBERT: "HEAT over Recurrence over BERT". Hierarchical model with a single Markovian HEAT layer (Hills et al., 2024). This applies temporal aggregation over all LSTM hidden states using learnable ϵ and β parameters, with emphasis on recent context through Markovian weighting.
MHRoBERT: "Multistream HEAT over Recurrence over BERT". Our proposed multistream architecture that extends HoRoBERT with three key components (See Fig. 1):

- *Multistream HEAT*: Stream-specific temporal aggregations with cross-stream excitation matrices $\epsilon_{(l,m)}$ and $\beta_{(l,m)}$ and per-stream baseline intensities μ_l

- *Unistream HEAT*: A Markovian HEAT layer aggregating all posts regardless of stream labels, using independent parameters ϵ_{uni} and β_{uni}
- *Residual connection*: Direct access to the current post's LSTM hidden state

The final representation concatenates all three components before classification, enabling the model to leverage both stream-specific interactions and general temporal patterns.

Central to our architecture is HEAT (Historical Emotional AggregaTion) (Sawhney et al., 2021a; Hills et al., 2024), which produces a temporally-weighted aggregation of historical LSTM hidden states using two learnable parameters: self-excitation $\epsilon \in \mathbb{R}^+$, controlling how strongly past posts amplify the current representation, and decay $\beta \in \mathbb{R}^+$, governing how quickly that influence diminishes with elapsed time $\Delta\tau_j = t_i - t_j$:

$$H^{(i)} = h^{(i-1)} + \sum_{j<i} h^{(j)} \cdot \epsilon \cdot e^{-\beta \Delta\tau_j} \quad (6)$$

MHRoBERT extends this to the multistream setting with stream-specific and cross-stream $\epsilon_{(l,m)}$, $\beta_{(l,m)}$ parameters (Section 3.2). In the fine-tuned setting, all BERT parameters are updated end-to-end; in the frozen setting only the LSTM and subsequent layers are trained. Multi-task variants add an auxiliary head predicting stream labels from the same shared representation.

4.1 LLMs for Moments of Change (MoC) Detection

We compared the performance of our specialized models developed for temporal modelling with LLMs applied in a zero-shot setting. To assess the capabilities of LLM models for automatic Moments of Change (MoC) detection in longitudinal data, we utilised a 32-billion parameter instruction-tuned model **Qwen2.5-32B-Instruct** (Reategui-Rivera et al., 2025; Nie et al., 2024). Evaluation was performed on **Reddit** data, through prompting-based inference under two settings:

(a) QwenLLM: In this setting, the model was prompted to classify each post within a user's timeline as one of S (Switch), E (Escalation), or O (None), depending on the past posts as context. Each post was evaluated independently, given only the past textual history of the user's posts.

(b) QwenLLMMultistream: The QwenLLM-Multistream prompting gives the model richer contextual signals that mimic a structured psychological assessment framework. The model used parallel cues pertaining to fine-grained emotions, mental health symptoms and personal life events to identify transitions (S), intensification (E), or stability (O) in the user's emotional trajectory. See Appendix B for the complete prompt templates and configuration details for both LLM models.

5 Results

Table 1 shows the performance of different architectures for predicting Moments of Change in mood. When fine-tuning BERT, HoRoBERT achieves the best overall macro-averaged F1 score (0.694), with MHRoBERT achieving competitive performance (0.689). All three hierarchical architectures demonstrate strong performance when BERT is fine-tuned end-to-end.

5.1 Model Performance Across Settings

The **fine-tuned BERT** setting yields the strongest results overall. HoRoBERT shows particular strength in recall (0.718) and achieves the best F1 for Switches (0.531) and Escalations (0.638). MHRoBERT achieves the highest precision (0.689) and performs well on Switches (F1 = 0.515), though slightly behind HoRoBERT. The baseline RoBERT (LSTM without temporal modeling) also demonstrates competitive performance (macro F1 = 0.685), suggesting that sequential modeling with fine-tuned representations captures substantial signal for Moments of Change detection.

With **frozen BERT**, all models show reduced performance (macro F1 = 0.600–0.609), as expected when linguistic representations cannot adapt to the task. However, the **multi-task learning** setting, which also uses frozen BERT but incorporates auxiliary tasks predicting stream-specific labels (emotions, life events, symptoms), achieves remarkably competitive performance (macro F1 = 0.673–0.674). This represents only a 2% drop compared to fine-tuned BERT, despite using fixed representations. This demonstrates that **multistream auxiliary supervision provides substantial benefits**, effectively compensating for the lack of linguistic fine-tuning by encouraging the model to learn representations sensitive to multiple aspects of mental health.

Under multi-task learning, MHRoBERT achieves the best Switch detection (F1 = 0.514), while all three models show strong and balanced performance. The consistently high performance across all three architectures in this setting

Model	Macro-Avg			Switch (S)			Escalation (E)			No Change (O)		
	P	R	F1	P	R	F1	P	R	F1	P	R	F1
Fine-tuned BERT												
MHRoBERT	**.689**	.696	.689	**.551**	.485	.515	**.579**	.697	.632	.936	**.906**	**.921**
HoRoBERT	.680	**.718**	**.694**	.532	**.531**	**.531**	.564	**.738**	**.638**	**.944**	.886	.913
RoBERT	.675	.700	.685	.516	.512	.514	.572	.685	.623	.937	.901	.918
Single-task (Frozen BERT)												
MHRoBERT	.567	.703	.600	.293	**.639**	.402	.449	.712	.551	.958	.759	.847
HoRoBERT	.574	**.713**	.607	**.310**	.629	**.414**	.446	**.760**	**.562**	**.967**	.752	.846
RoBERT	**.575**	.703	**.609**	.307	.616	.408	**.462**	.717	.561	.958	**.776**	**.857**
Multi-task (Frozen BERT)												
MHRoBERT	.689	.662	**.674**	.559	**.476**	**.514**	.599	.586	.592	.909	.923	.916
HoRoBERT	.688	**.663**	**.674**	.533	.467	.497	**.620**	.595	.606	.911	**.927**	**.919**
RoBERT	**.696**	.660	.673	**.585**	.432	.496	.590	**.628**	**.608**	**.913**	.919	.917
LLM Baselines												
QwenLLM	.447	.496	.427	.119	.496	.192	.343	.357	.350	.879	.636	.738
QwenLLMMultistream	**.517**	**.570**	**.481**	**.141**	**.650**	**.232**	**.468**	**.367**	**.411**	**.941**	**.694**	**.799**

Table 1: Per-class and macro-averaged results, for predicting MoCs. For each metric, **bold** denotes the best score within that experimental condition, while **bold+underline** denotes the best score across all conditions. P = Precision, R = Recall, F1 = F1-score.

underlines that the benefit is driven by the auxiliary supervision signal itself rather than by any particular model design. However, it is worth noting, that macro-averaged F1 masks important precision-recall tradeoffs: in clinical monitoring contexts, high recall for Switch events may be preferable, as a missed sudden mood change could signal an undetected emerging crisis whereas false alarms can be filtered by a human reviewer, favouring settings such as single-task frozen BERT (Switch recall: 0.639) over the multi-task variant (0.476) despite the latter's higher F1.

5.2 Benefits of Multistream Information

A key finding of this work is that automatically-generated multistream labels – derived via LLM annotation inspired by the taxonomy of Mao et al. (2025) – **improve performance across different model architectures**, including those not explicitly designed to handle this information. The LLM results provide compelling evidence: while the baseline QwenLLM achieves only 0.427 macro F1, incorporating multistream information in prompts (QwenLLMMultistream) improves performance to 0.481 (12.6% relative improvement). This improvement is particularly notable for rare classes, with Switch recall improving from 0.496 to 0.650.

These results highlight the value of multistream supervision beyond any specific architecture. The stream labels provide a consistent benefit whether the model explicitly models stream interactions (MHRoBERT), uses them as auxiliary supervision (multi-task setting), or incorporates them in prompts (LLMs). These findings have direct relevance for the CLPsych shared task on Moments of Change detection: across all architectures evaluated on the Reddit dataset used in that shared task, models trained with multistream auxiliary supervision consistently and substantially outperform their single-task counterparts. We therefore recommend that future CLPsych submissions for MoC detection explore multistream auxiliary supervision as a practical and scalable strategy for improving performance.

Furthermore, the multistream framework enhances model **interpretability at multiple levels**. At the data level, individual posts can be inspected for their stream-specific labels, allowing clinicians or researchers to understand which aspects of mental health (emotions, life events, symptoms) are present in the discourse.

At the model level, architectures like MHRoBERT produce interpretable learned parameters (Section 5.5) that reveal how different streams interact temporally. Even for models without explicit multistream architectures, the presence of stream labels enables more transparent analysis – practitioners can examine not just whether a mood change was predicted, but which combination of emotional expressions, life circumstances, and symptom mentions contributed to that prediction. This interpretability is particularly valuable in

mental health applications where understanding the reasoning behind predictions is as important as prediction accuracy itself.

5.3 Comparison of Temporal Modeling Approaches

The improvements from adding temporal HEAT layers (HoRoBERT) or multistream HEAT (MHRoBERT) over the LSTM baseline (RoBERT) in the fine-tuned setting suggest several considerations. The strong performance of RoBERT indicates that sequential modeling with fine-tuned BERT embeddings captures much of the relevant temporal structure. While MHRoBERT's multistream architecture provides interpretable cross-stream interaction parameters (Section 5.5), the performance gains over simpler approaches are modest in the fine-tuned setting. However, the architecture's ability to leverage multistream information proves particularly valuable in the multi-task setting, where it achieves the best Switch detection. This suggests that explicit multistream modeling becomes more beneficial when combined with auxiliary supervision that encourages learning stream-specific patterns.

5.4 Analysis of Model Components

The results demonstrate that both the unistream temporal component (HoRoBERT) and residual LSTM representations (present in all models) contribute meaningfully to performance. HoRoBERT consistently achieves high recall on rare events, suggesting that the unistream HEAT component effectively captures temporal patterns associated with mood changes. MHRoBERT's additional multistream HEAT component provides interpretable interaction parameters while maintaining competitive performance across all settings.

The importance of end-to-end BERT fine-tuning is evident, with fine-tuned models substantially outperforming frozen BERT variants in single-task learning. However, the multi-task results reveal that **multistream auxiliary supervision can largely bridge this gap**, achieving near-equivalent performance with frozen representations. This finding has practical implications, as it enables effective mood change detection without expensive fine-tuning, making the approach more accessible for resource-constrained applications or domains where labeled data for fine-tuning is limited.

5.5 Stream Interaction Analysis

The resulting model produces interpretable parameters, including excitation (ϵ) and decay (β) matrices, along with baseline intensities (μ), which provide insights into the temporal relationships between streams. These relationships are visualized in Figure 2.

Our multistream approach demonstrates several advantages over the baseline approaches. The integration of emotional states, life events, and symptoms provides a more comprehensive view of mental health changes, with the learned interaction parameters (Figure 2) offering insights into temporal dynamics between streams.

The **excitation patterns** reveal differential influences between streams. Emotions exhibit the highest self-excitation ($\epsilon = 0.01$), suggesting emotional states tend to reinforce themselves. Cross-stream excitation from emotions to symptoms ($\epsilon = 0.009$) is notably strong, indicating that emotional states may influence symptom manifestation. Life events show more balanced excitation across streams ($\epsilon = 0.004$–0.006), while symptoms demonstrate weaker cross-stream influence overall ($\epsilon = 0.002$–0.004).

The **decay rates** (β) show interesting temporal patterns. Emotions display very slow self-decay ($\beta = 0.0002$), suggesting persistence of emotional states, while their cross-stream decay is faster ($\beta = 0.004$–0.007). Life events exhibit consistently fast decay across all streams ($\beta = 0.01$–0.009), potentially reflecting their more transient nature. The decay from emotions to symptoms is relatively slow ($\beta = 0.001$), which may indicate lasting emotional influence on symptoms.

The **baseline intensities** show positive values for emotions ($\mu = 0.0062$) and symptoms ($\mu = 0.0056$), with life events near zero ($\mu = -0.00093$). This pattern suggests that while emotional and symptom streams maintain some baseline activity, life events are primarily driven by external occurrences rather than intrinsic baseline patterns. These learned parameters provide a foundation for understanding how different aspects of mental health interact temporally, though further investigation would be valuable to validate these patterns against clinical observations.

5.6 Stream Interpretability Analysis

The decomposition of the multistream HEAT component of MHRoBERT provides key interpretabil-

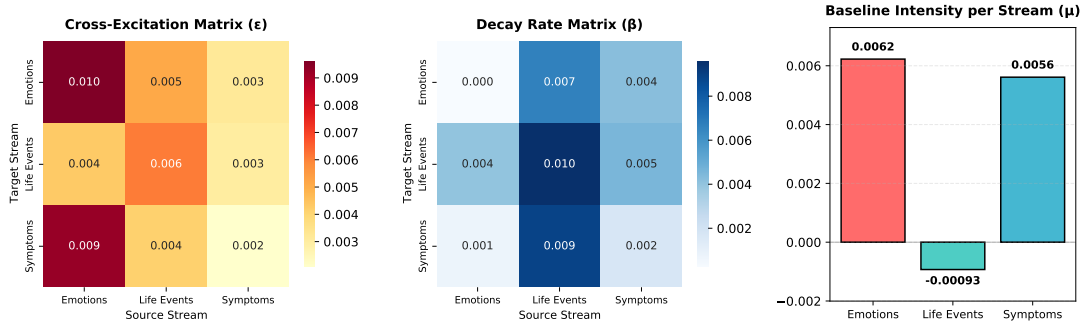

Figure 2: Learned interaction parameters showing excitation (ϵ) and decay (β) values between different streams, where higher values indicate stronger relationships. Learned baseline intensities (μ) for each individual stream are presented, where higher values indicate stronger contribution to the associated stream representations - regardless of event occurrence.

ity insights. As shown in Figure 3, the *Emotion* stream contributes the largest share of multistream attribution ($\sim 41\%$), followed by *Symptom* ($\sim 39\%$) and *Life Event* ($\sim 19\%$). This ordering, consistent across classes, indicates a coherent internal structure within multistream HEAT.

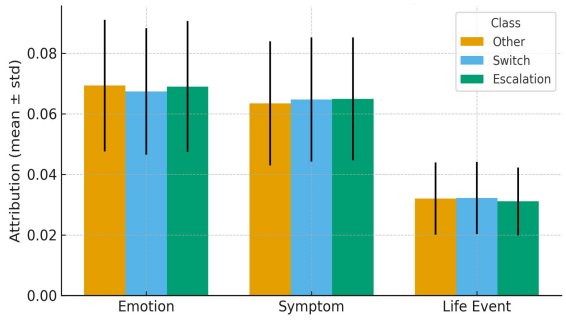

Figure 3: Class wise multistream attribution with multistream HEAT component (fine-tuned MHRoBERT).

Table 2 summarizes the representational geometry results using Nearest-Centroid (global) and k-Nearest Neighbors (local) evaluations. The unistream space exhibits stronger class separability and contributes more to gradient attributions, whereas the multistream space, though less discriminative, remains semantically informative. Its lower F1 reflects an expected entanglement of cross-stream embeddings that jointly encode emotional, symptomatic, and life-event cues.

While this representation is less separable, it captures richer inter-stream relationships, enhancing interpretability and potential transferability. In contrast, the unistream representation emphasizes temporal discrimination, producing distinct clusters aligned with mental state transitions. Together, these complementary inductive biases suggest that multistream HEAT prioritizes semantic alignment and interpretability, while unistream HEAT maintains geometric clarity and predictive precision as reflected in the fused stream results.

Method	UHEAT	MHEAT	Fused
Nearest-Centroid (cosine)			
Macro-F1	0.4413	0.3368	**0.4493**
Precision	0.4388	0.3812	0.4388
Recall	0.5091	0.4229	0.5091
k-NN (k=5, cosine)			
Macro-F1	0.4032	0.3882	**0.4057**
Precision	0.3961	0.3994	0.4014
Recall	0.4106	0.3907	0.4106

Table 2: Representation geometry analysis.

5.7 Future Research Directions

This work opens several promising avenues for future research in automated mental health monitoring and detecting changes more generally, by considering multiple interrelated temporal streams of data. In terms of stream extraction, more sophisticated prompt engineering techniques could improve the quality and reliability of automated annotations. Integration with external clinical knowledge bases could provide more psychologically validated stream categories, while systematic comparison with expert human annotation would help validate the LLM-based approach.

The modeling framework itself could be extended in several ways. Incorporating more advanced, fine-tuned linguistic representations could better capture the nuances of mental health-related language. Additional streams, such as social interaction patterns or environmental factors, could provide even more comprehensive monitoring capabilities. The temporal modeling could be enhanced to capture more complex dynamics, including seasonal patterns and long-term trends.

From a clinical application perspective, this lays the groundwork for developing and improving existing mental health monitoring frameworks. The

interpretable nature of the stream interactions could help clinicians understand the progression of mental health states over time, particularly by highlighting which combinations of emotional, behavioral, and life event patterns most strongly indicate impending changes in mental health.

6 Conclusion

This paper has demonstrated the value of multistream supervision for mental health monitoring, leveraging LLM-annotated emotional states, personal life events, and mental health symptoms as auxiliary signals, inspired by the taxonomy of Mao et al. (2025). Building on this, we introduced MHRoBERT, a novel hierarchical architecture that integrates stream-specific temporal interactions, general temporal dynamics, and direct sequential information through a unified representation. Its interpretable learned parameters (ϵ, β, μ) reveal interaction patterns between mental health indicators with potential clinical value. Crucially, our empirical results show that multistream auxiliary supervision benefits diverse architectures, establishing it as a practical, architecture-agnostic strategy for improving MoC detection.

A particularly important finding is that multistream auxiliary supervision provides consistent and substantial benefits across all evaluated architectures on the Reddit dataset from the CLPsych 2022 shared task on Moments of Change detection, while also being architecture-agnostic, scalable, and demonstrably effective.

Future work could focus on validating and refining stream annotations through clinical evaluation, exploring additional streams such as social interaction patterns or environmental factors, and investigating temporal modelling approaches that more explicitly account for time intervals in training. These directions could lead to more reliable systems for monitoring mental health through social media, supporting better understanding and outcomes for individuals experiencing mental health difficulties.

Limitations

While this work demonstrates the potential of multistream temporal modelling for mental health monitoring, several important limitations should be acknowledged.

Performance Gains The performance improvements of MHRoBERT over simpler baselines are relatively modest in the fine-tuned setting, with macro F1 scores differing by only a few percentage points. While the multistream information provides interpretable cross-stream interaction parameters and shows stronger benefits in multi-task learning scenarios, the gains over the LSTM baseline (RoBERT) in single-task fine-tuned settings are limited. Additionally, fine-tuned BERT with multi-task learning exhibited gradient instability during training (NaN gradients in a substantial proportion of Reddit batches), which we attribute to the interaction between the multi-task loss and BERT's deeper layers; stabilising this setting is left to future work.

Stream Annotation Quality The multistream annotations are generated automatically using LLMs without human validation, which introduces potential noise and inconsistencies. While LLMs have shown promise for annotation tasks (Gilardi et al., 2023; Törnberg, 2023), they can struggle with implicit or ambiguous content (Li and Conrad, 2024) and show significant variability across tasks (Kristensen-McLachlan et al., 2023). In mental health contexts, where subtle distinctions have significant clinical implications, the lack of expert validation of our automatically generated stream labels represents a notable limitation. The three-stream taxonomy (emotions, life events, symptoms), while grounded in psychological frameworks, has not been validated by mental health professionals for this specific application. Despite the noisy imperfect annotations, we still demonstrate several benefits for making use of such annotations in this paper.

Furthermore, while greedy decoding (temperature $= 0.0$, `do_sample=False`) ensures deterministic and reproducible annotations, prompt sensitivity was not systematically evaluated, the conceptual boundaries between streams are not always clear-cut (e.g., *Depressed* and *Detachment* may reflect overlapping psychological states), and whether results generalise to annotations from other LLMs beyond Qwen 2.5-32B-Instruct (Mao et al., 2025) remains an open question for future work.

Temporal Modeling Assumptions The HEAT mechanism assumes that temporal influence decays exponentially over time, which may not accurately reflect all psychological processes. Mental health trajectories can involve complex patterns including seasonality, and long-term trends that may not be well-captured by exponential decay functions. The current implementation uses learned static baseline

intensities, whereas more sophisticated approaches might benefit from dynamic, context-dependent baselines that adapt to individual users' temporal patterns.

Clinical Validity and Deployment This work has not been validated in clinical settings or with mental health professionals. The interpretable parameters produced by MHRoBERT (excitation and decay matrices) require clinical evaluation to determine whether they align with established understanding of mental health dynamics. Furthermore, deploying such systems for real-world mental health monitoring raises critical concerns about false positives/ negatives, privacy, informed consent, and the potential for harm if predictions are acted upon without appropriate clinical oversight. The models are intended as research tools to understand mental health patterns in social media data, not as diagnostic or intervention tools.

Computational Requirements The hierarchical architecture with end-to-end BERT fine-tuning requires substantial computational resources, which may limit accessibility for researchers or organizations with limited infrastructure. While the multi-task frozen BERT setting offers a more resource-efficient alternative, it still requires significant compute for training on longitudinal data.

Ethical Considerations Automated monitoring of mental health through social media raises important ethical questions about privacy, consent, and appropriate use of such systems. While this work uses publicly available Reddit data with institutional review board approval, the ability to identify mood changes from user posts could potentially be misused for surveillance or discrimination. Users posting on public forums may not anticipate or consent to their content being analyzed for mental health indicators.

Acknowledgments

This work was supported by a UKRI/EPSRC Turing AI Fellowship to Maria Liakata (grant EP/V030302/1), Keystone grant funding from Responsible AI UK (grants EP/Y009800/1 and KP0016), the Alan Turing Institute (grant EP/N510129/1), and an MRC grant (grant MR/X030725/1).

Ethics Statement

Before starting this research, approval was secured from the Institutional Review Board of the lead university. This study considers ethical considerations when dealing with the analysis of user-generated content on social media platforms, specifically Reddit. The ethical implications of our research, in particular the ability to identify changes in mood within user timelines, share similar concerns to that of prior research focused on identifying personal events through social media, and recognizing signs of suicidal thoughts. To help mitigate these risks, measures were taken such as the limited and regulated access to the developed software and the annotations that were used in this study.

References

Mostafa M. Amin, Erik Cambria, and Björn W. Schuller. 2023. Will affective computing emerge from foundation models and general artificial intelligence? a first evaluation of chatgpt. *IEEE Intelligent Systems*, 38(2):15–23.

Eliseo Bao, Anxo Pérez, and Javier Parapar. 2024. Explainable depression symptom detection in social media. *Health Information Science and Systems*, 12(1):47.

Denny Borsboom. 2017. A network theory of mental disorders. *World psychiatry*, 16(1):5–13.

Denny Borsboom and Angélique OJ Cramer. 2013. Network analysis: an integrative approach to the structure of psychopathology. *Annual review of clinical psychology*, 9(1):91–121.

Laura F Bringmann, Nathalie Vissers, Marieke Wichers, Nicole Geschwind, Peter Kuppens, Frenk Peeters, Denny Borsboom, and Francis Tuerlinckx. 2013. A network approach to psychopathology: New insights into clinical longitudinal data. *PloS one*, 8(4):e60188.

Ruijun Chen, Jin Wang, Liang-Chih Yu, and Xuejie Zhang. 2023. Learning to memorize entailment and discourse relations for persona-consistent dialogues. In *Proceedings of the AAAI conference on artificial intelligence*, volume 37, pages 12653–12661.

Muskan Garg. 2023. Mental health analysis in social media posts: A survey: M. garg. *Archives of Computational Methods in Engineering*, 30(3):1819–1842.

Fabrizio Gilardi, Meysam Alizadeh, and Maël Kubli. 2023. Chatgpt outperforms crowd workers for text-annotation tasks. *Proceedings of the National Academy of Sciences*, 120(30):e2305016120.

Oliver L. Haimson, Albert J. Carter, Shanley Corvite, Brookelyn Wheeler, Lingbo Wang, Tianxiao Liu, and Alexxus Lige. 2021. The major life events taxonomy: Social readjustment, social media information sharing, and online network separation during times of life transition. *Journal of the Association for Information Science and Technology*, 72(7):933–947.

Alan G Hawkes. 1971. Point spectra of some mutually exciting point processes. *Journal of the Royal Statistical Society: Series B (Methodological)*, 33(3):438–443.

Anthony Hills, Talia Tseriotou, Xenia Miscouridou, Adam Tsakalidis, and Maria Liakata. 2024. Exciting mood changes: A time-aware hierarchical transformer for change detection modelling. In *Findings of the Association for Computational Linguistics: ACL 2024*, pages 12526–12537, Bangkok, Thailand. Association for Computational Linguistics.

Ross Deans Kristensen-McLachlan, Miceal Canavan, Márton Kardos, Mia Jacobsen, and Lene Aarøe. 2023. Chatbots are not reliable text annotators. *arXiv preprint arXiv:2311.05769*.

Jin Myung Kwak, Minseon Kim, and Sung Ju Hwang. 2023. Context-dependent instruction tuning for dialogue response generation.

Mao Li and Frederick Conrad. 2024. Advancing annotation of stance in social media posts: A comparative analysis of large language models and crowd sourcing. *arXiv preprint arXiv:2406.07483*.

Yucheng Li, Bo Dong, Chenghua Lin, and Frank Guerin. 2023. Compressing context to enhance inference efficiency of large language models.

Chandreen Liyanage, Muskan Garg, Vijay Mago, and Sunghwan Sohn. 2023. Augmenting Reddit posts to determine wellness dimensions impacting mental health. In *The 22nd Workshop on Biomedical Natural Language Processing and BioNLP Shared Tasks*, pages 306–312, Toronto, Canada. Association for Computational Linguistics.

Junyu Mao, Anthony Hills, Talia Tseriotou, Maria Liakata, Aya Shamir, Dan Sayda, Dana Atzil-Slonim, Natalie Djohari, Arpan Mandal, Silke Roth, et al. 2025. Automated data enrichment using confidence-aware fine-grained debate among open-source LLMs for mental health and online safety. *arXiv preprint arXiv:2512.06227*.

Yu Meng, Jiaxin Huang, Yu Zhang, and Jiawei Han. 2022. Generating training data with language models: Towards zero-shot language understanding. In *Advances in Neural Information Processing Systems*, volume 35, page 462–477.

Virginia Morini, Salvatore Citraro, Elena Sajno, Maria Sansoni, Giuseppe Riva, Massimo Stella, and Giulio Rossetti. 2025. Online posting effects: Unveiling the non-linear journeys of users in depression communities on reddit. *Computers in Human Behavior Reports*, 17:100542.

Jingping Nie, Hanya Shao, Yuang Fan, Qijia Shao, Haoxuan You, Matthias Preindl, and Xiaofan Jiang. 2024. Llm-based conversational ai therapist for daily functioning screening and psychotherapeutic intervention via everyday smart devices. *arXiv preprint arXiv:2403.10779*.

Albert Park and Mike Conway. 2017. Longitudinal changes in psychological states in online health community members: Understanding the long-term effects of participating in an online depression community. *Journal of Medical Internet Research*, 19:e71.

C Mahony Reategui-Rivera, Aref Smiley, and Joseph Finkelstein. 2025. Llm-based chatbot to reduce mental illness stigma in healthcare providers. In *2025 IEEE 15th Annual Computing and Communication Workshop and Conference (CCWC)*, pages 00001–00007. IEEE.

Marian-Andrei Rizoiu, Young Lee, Swapnil Mishra, and Lexing Xie. 2017. Hawkes processes for events in social media. In *Frontiers of multimedia research*, pages 191–218.

Ramit Sawhney, Shivam Agarwal, Arnav Wadhwa, and Rajiv Shah. 2021a. Exploring the Scale-Free Nature of Stock Markets: Hyperbolic Graph Learning for Algorithmic Trading. In *Proceedings of the Web Conference 2021*, pages 11–22, Ljubljana Slovenia. ACM.

Ramit Sawhney, Harshit Joshi, Rajiv Ratn Shah, and Lucie Flek. 2021b. Suicide Ideation Detection via Social and Temporal User Representations using Hyperbolic Learning. In *Proceedings of the 2021 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*, pages 2176–2190, Online. Association for Computational Linguistics.

Dana Atzil Slonim. 2024. Self-other dynamics (sod): A transtheoretical coding manual.

Jiayu Song, Jenny Chim, Adam Tsakalidis, Julia Ive, Dana Atzil-Slonim, and Maria Liakata. 2024. Combining hierachical VAEs with LLMs for clinically meaningful timeline summarisation in social media. In *Findings of the Association for Computational Linguistics: ACL 2024*, pages 14651–14672, Bangkok, Thailand. Association for Computational Linguistics.

Adam Tsakalidis, Jenny Chim, Iman Munire Bilal, Ayah Zirikly, Dana Atzil-Slonim, Federico Nanni, Philip Resnik, Manas Gaur, Kaushik Roy, Becky Inkster, Jeff Leintz, and Maria Liakata. 2022a. Overview of the CLPsych 2022 shared task: Capturing moments of change in longitudinal user posts. In *Proceedings of the Eighth Workshop on Computational Linguistics and Clinical Psychology*, pages 184–198, Seattle, USA. Association for Computational Linguistics.

Adam Tsakalidis, Federico Nanni, Anthony Hills, Jenny Chim, Jiayu Song, and Maria Liakata. 2022b. Identifying moments of change from longitudinal user text. In *Proceedings of the 60th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, pages 4647–4660, Dublin, Ireland. Association for Computational Linguistics.

Yu-Min Tseng, Wei-Lin Chen, Chung-Chi Chen, and Hsin-Hsi Chen. 2025. Evaluating large language models as expert annotators. *arXiv preprint arXiv:2508.07827*.

Talia Tseriotou, Jenny Chim, Ayal Klein, Aya Shamir, Guy Dvir, Iqra Ali, Cian Kennedy, Guneet Singh Kohli, Anthony Hills, Ayah Zirikly, Dana Atzil-Slonim, and Maria Liakata. 2025. Overview of the CLPsych 2025 shared task: Capturing mental health dynamics from social media timelines. In *Proceedings of the 10th Workshop on Computational Linguistics and Clinical Psychology (CLPsych 2025)*, pages 193–217, Albuquerque, New Mexico. Association for Computational Linguistics.

Talia Tseriotou, Adam Tsakalidis, Peter Foster, Terence Lyons, and Maria Liakata. 2023. Sequential path signature networks for personalised longitudinal language modeling. In *Findings of the Association for Computational Linguistics: ACL 2023*, pages 5016–5031, Toronto, Canada. Association for Computational Linguistics.

Talia Tseriotou, Adam Tsakalidis, and Maria Liakata. 2024. TempoFormer: A transformer for temporally-aware representations in change detection. In *Proceedings of the 2024 Conference on Empirical Methods in Natural Language Processing*, pages 19635–19653, Miami, Florida, USA. Association for Computational Linguistics.

Petter Törnberg. 2023. Chatgpt-4 outperforms experts and crowd workers in annotating political twitter messages with zero-shot learning. *arXiv preprint arXiv:2304.06588*.

Hongru Wang, Rui Wang, Fei Mi, Yang Deng, Zezhong Wang, Bin Liang, Ruifeng Xu, and Kam-Fai Wong. 2023. Cue-CoT: Chain-of-thought prompting for responding to in-depth dialogue questions with LLMs. In *Findings of the Association for Computational Linguistics: EMNLP 2023*, pages 12047–12064, Singapore. Association for Computational Linguistics.

Marieke Wichers, Claudia Lothmann, Claudia JP Simons, Nancy A Nicolson, and Frenk Peeters. 2012. The dynamic interplay between negative and positive emotions in daily life predicts response to treatment in depression: a momentary assessment study. *British Journal of Clinical Psychology*, 51(2):206–222.

Xuhai Xu, Bingsheng Yao, Yuanzhe Dong, Saadia Gabriel, Hong Yu, James Hendler, Marzyeh Ghassemi, Anind K. Dey, and Dakuo Wang. 2024. Mental-llm: Leveraging large language models for mental health prediction via online text data. *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.*, 8(1).

An Yang, Baosong Yang, Beichen Zhang, Binyuan Hui, Bo Zheng, Bowen Yu, Chengyuan Li, Dayiheng Liu, Fei Huang, Haoran Wei, Huan Lin, Jian Yang, Jianhong Tu, Jianwei Zhang, Jianxin Yang, Jiaxi Yang, Jingren Zhou, Junyang Lin, Kai Dang, Keming Lu, Keqin Bao, Kexin Yang, Le Yu, Mei Li, Mingfeng Xue, Pei Zhang, Qin Zhu, Rui Men, Runji Lin, Tianhao Li, Tianyi Tang, Tingyu Xia, Xingzhang Ren, Xuancheng Ren, Yang Fan, Yang Su, Yichang Zhang, Yu Wan, Yuqiong Liu, Zeyu Cui, Zhenru Zhang, and Zihan Qiu. 2025. Qwen2.5 technical report.

Kailai Yang, Shaoxiong Ji, Tianlin Zhang, Qianqian Xie, Ziyan Kuang, and Sophia Ananiadou. 2023. Towards interpretable mental health analysis with large language models. In *Proceedings of the 2023 Conference on Empirical Methods in Natural Language Processing*. Association for Computational Linguistics.

Amir Hossein Yazdavar, Mohammad Saeid Mahdavinejad, Goonmeet Bajaj, William Romine, Amit Sheth, Amir Hassan Monadjemi, Krishnaprasad Thirunarayan, John M Meddar, Annie Myers, Jyotishman Pathak, et al. 2020. Multimodal mental health analysis in social media. *Plos one*, 15(4):e0226248.

Qiang Zhang, Aldo Lipani, Omer Kirnap, and Emine Yilmaz. 2020. Self-attentive Hawkes process. In *Proceedings of the 37th International Conference on Machine Learning*, volume 119 of *Proceedings of Machine Learning Research*, pages 11183–11193. PMLR.

Wen Zheng, Natasa Milic-Frayling, and Ke Zhou. 2023. Contextual knowledge learning for dialogue generation. In *Proceedings of the 61st Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, pages 7822–7839, Toronto, Canada. Association for Computational Linguistics.

Ke Zhou, Hongyuan Zha, and Le Song. 2013. Learning triggering kernels for multi-dimensional hawkes processes. In *Proceedings of the 30th International Conference on Machine Learning*, volume 28 of *Proceedings of Machine Learning Research*, pages 1301–1309, Atlanta, Georgia, USA. PMLR.

A LLM Prompting Templates

We use chain-of-thought (CoT) prompting with Qwen 2.5-32B-Instruct to extract three streams of mental health-relevant information from social media posts. Below are the complete prompt templates used for each stream.

A.1 Model Configuration

We implement our automated stream annotation using Qwen 2.5-32B-Instruct with the following configuration:

- **Model**: Qwen 2.5-32B-Instruct
- **Quantization**: 4-bit quantization
- **Precision**: bfloat16
- **Decoding Strategy**: Greedy decoding (deterministic) with `do_sample=False`
- **Max New Tokens**: 1024 tokens
- **Temperature**: 0.0 (greedy decoding)

A.2 Emotional States Prompt

```
You are tasked with classifying the
    emotional content of a social media
    post.

Emotion Categories:
- Happy: Content, joyful, hopeful,
    excited, glad, cheerful
- Proud: Feeling accomplished,
    satisfied with achievements,
    confident in success
- Sad: Emotional pain, grieving, hurt,
    heartbroken over loss
- Anxious: Fearful, tense, worried,
    nervous, scared, panicked
- Depressed: Despair, hopeless,
    worthless, empty, suicidal thoughts
- Apathetic: Numb, indifferent, don't
    care, emotionally blunted
- Angry: Aggression, hate, rage,
    disgust, contempt
- Ashamed: Guilty, embarrassed,
    humiliated, self-critical
- Loneliness: Feeling alone, isolated,
    lacking meaningful connection
- None: The post does not clearly
    reflect any of the above emotion
    categories.

Instructions:
1. Based on the provided examples,
    analyze the post by evaluating each
    emotion category.
2. For each category, explain your
    reasoning and clearly state "yes"
    or "no" regarding its presence.
3. Finally, list all applicable emotion
    categories that are present in the
    post. If more than one applies,
    separate them with commas.

Below are some examples:
{few_shot_examples}

Post to Analyze:
Post:
"{post}"

Please strictly follow the output
    format exactly as shown below. Do
    not use bold, markdown, or extra
    formatting.

Output Format:
Explanation:
- Happy: [reason]. So the answer is yes
    (or is no).
- Proud: [reason]. So the answer is yes
    (or is no).
- Sad: [reason]. So the answer is yes
    (or is no).
- Anxious: [reason]. So the answer is
    yes (or is no).
- Depressed: [reason]. So the answer is
    yes (or is no).
- Apathetic: [reason]. So the answer is
    yes (or is no).
- Angry: [reason]. So the answer is yes
    (or is no).
- Ashamed: [reason]. So the answer is
    yes (or is no).
- Loneliness: [reason]. So the answer
    is yes (or is no).
- None: [reason]. So the answer is yes
    (or is no).

Answer:
Only output exact category names:
    Happy, Proud, Sad, Anxious,
    Depressed, Apathetic, Angry,
    Ashamed, Loneliness, or None. Use
    commas to separate multiple labels,
    if any.
```

A.3 Personal Life Events Prompt

```
You are provided with a social media
    post and must identify any personal
    life events based on the life event
    categories defined below.

Life events are experiences that have a
    major personal impact on an
    individual. They must involve a
    clearly identifiable occurrence or
    change. These events may have
    occurred in the past (explicitly
    stated, or inferred from context if
    the impact is clear), be occurring
    in the present (explicitly
    described or clearly implied), or
    be expected in the near future
    (only if explicitly stated).

- Mental Health:
Mental health life events cover
    discussions about receiving a
    formal mental health diagnosis;
    starting or adjusting psychiatric
    medication; beginning therapy;
    recovery or significant symptom
    improvement; or experiencing acute
    psychological episodes such as
    manic episodes, psychotic breaks,
    panic attacks, dissociative
    episodes, self-harm, suicide
    attempts, or acute suicidal
    ideation requiring crisis
    intervention. It does not include
    general low mood, stress, anxiety,
    or depression unless explicitly
    tied to one of these major events.
- Physical Health:
Physical health life events cover
    accidents; injuries; diagnoses or
    survival of serious illnesses;
    chronic conditions that have a
    notable impact on daily life; or
    cases where the individual or
    someone important to them has been
    hospitalized, undergone surgery,
    received emergency medical
    treatment, or experienced other
    major medical interventions.
    Pregnancy-related experiences,
    including becoming pregnant,
```

```
    pregnancy loss, abortion, or
    complications, are also included.
    Common aches, headaches, fatigue,
    or other symptoms without being
    part of a major physical health
    event are not included.
- Abuse & Addiction:
Covers major life events involving
    abuse or substance-related issues.
    Abuse may occur in childhood or
    adulthood and may take physical,
    sexual, emotional, psychological,
    or other forms. Addiction-related
    events include the onset of heavy
    drug or alcohol use; acknowledgment
    of addiction; overdoses; withdrawal
    symptoms; relapses; or recovery
    (including participation in
    treatment programs such as
    rehabilitation or support groups).
- Relationship & Loss:
Covers specific relationship changes or
    disruptions in meaningful personal
    connectionsfamily or non-familythat
    result in strong emotional impact.
    Includes beginning or ending
    friendships or romantic
    relationships; marriage or divorce;
    parental separation or divorce;
    serious arguments or conflicts;
    relationships becoming abusive or
    toxic; the addition of new family
    members (e.g., through birth or
    adoption); parenting difficulties;
    emotionally distressing health
    issues involving family members,
    partners, close friends, or other
    meaningful connections; or the
    death of a person or pet who was
    important to the individual.
- Career & Education:
Career-related life events cover
    starting or losing a job;
    promotions; demotions; troubles at
    work; being unable to find a job;
    retirement; or becoming a business
    owner. Education-related life
    events include starting or
    graduating from school;
    transferring schools; leaving
    school without graduating; being
    denied entry into school; taking
    major exams; experiencing
    significant academic challenges; or
    achieving major academic milestones
    (e.g., earning a certification).
- Financial & Legal & Societal:
Financial life events include major
    purchases (e.g., a house or car);
    financial gains or milestones
    (e.g., paying off debt); financial
    losses (e.g., stolen or damaged
    property); or ongoing financial
    difficulties. Legal life events
    include law violations; arrests;
    court appearances; lawsuits or
    legal actions; or major
    interactions with the justice
    system, whether the individual is
    subject to legal proceedings or
    initiating them. Societal life
    events with clear personal impact
    include natural disasters;
    pandemics; major political events;
    societal issues; or notable public
    encounters (e.g., meeting a
    celebrity).
- Lifestyle & Identity & Environment:
Covers major life events involving
    changes in lifestyle habits,
    personal identity, or living
    environment. Lifestyle changes
    include adopting new health- and
    well-being-related routines, such
    as adjustments in diet; exercise;
    sleep management; stress management
    (e.g., through meditation);
    reducing substance use; fostering
    positive social connections; or
    getting a pet. It does not include
    clinical treatment for mental
    health issues. Identity transitions
    include identifying or changing
    one's gender or sexual orientation
    (e.g., coming out as LGBTQ+);
    discussing a new sexual experience;
    or exploring, adopting, or changing
    political or religious beliefs
    (including changes in religious
    practices). Relocation life events
    include moving to a new place;
    moving in or out of the family
    home; family members moving into or
    out of the household; or
    significant travel experiences.
- None:
Applies when the post does not clearly
    reflect any of the above life event
    categories.

Instructions:
1. Read the post carefully and evaluate
    whether it matches each defined
    life event category.
2. For each life event category,
    explain your reasoning. If a
    category applies, support your
    answer with direct evidence from
    the post. If it does not apply,
    explain why there is insufficient
    or no evidence. Clearly state "yes"
    or "no" for each category.
3. Finally, list all broad life event
    categories that apply. If more than
    one applies, separate them with
    commas.

Below are some examples:
{few_shot_examples}

Post to Analyze:
Post:
"{post}"

Please strictly follow the output
    format exactly as shown below. Do
    not use bold, markdown, or extra
    formatting.

Output Format:
```

```
Explanation:
- Mental Health: [reason]. So the
    answer is yes (or is no).
- Physical Health: [reason]. So the
    answer is yes (or is no).
- Abuse & Addiction: [reason]. So the
    answer is yes (or is no).
- Relationship & Loss: [reason]. So the
    answer is yes (or is no).
- Career & Education: [reason]. So the
    answer is yes (or is no).
- Financial & Legal & Societal:
    [reason]. So the answer is yes (or
    is no).
- Lifestyle & Identity & Environment:
    [reason]. So the answer is yes (or
    is no).
- None: [reason]. So the answer is yes
    (or is no).

Answer:
Output exact category names: Mental
    Health, Physical Health, Abuse &
    Addiction, Relationship & Loss,
    Career & Education, Financial &
    Legal & Societal, Lifestyle &
    Identity & Environment, or None.
    Use commas to separate multiple
    labels, if any."""
```

A.4 Mental Health Symptoms Prompt

```
You are provided with a social media
    post and must identify any symptoms
    of mental health issues according
    to the symptom categories defined
    below.

Symptom Categories:
- Suicidal Thoughts:
Distinct from general distress, this
    involves specific thoughts about
    ending one's life, including
    considering methods, making plans,
    or expressing a desire to die. This
    includes passive thoughts ("I wish
    I wouldn't wake up") and active
    plans, as well as references to
    past suicide attempts.

- Somatoform & Substance Abuse:
Somatoform refers to physical symptoms
    that cause significant distress but
    lack clear medical explanations.
    This includes various types of pain
    (e.g., headaches, muscle aches);
    unexplained physical sensations;
    physical fatigue; and physical
    complaints affecting different body
    systems. They are often accompanied
    by cognitive symptoms such as
    impaired memory, reduced processing
    speed, and poor planning. Substance
    abuse covers problematic use of
    alcohol, drugs, or other
    substances. This includes excessive
    drinking or drug use; failed
    attempts to cut down; spending a
    disproportionate amount of time
    obtaining, using, or recovering
    from substances; and continued use
    despite negative consequences.

- Eating Pathology:
Includes problematic cognitions and
    behaviors related to food, eating
    patterns, body image, and weight.
    This includes fear of weight gain
    (e.g., frequent body-checking,
    obsessive weighing); body image
    disturbance (e.g., negative
    fixation on weight or shape);
    self-worth tied to appearance
    (e.g., harsh self-criticism,
    comparisons); denial of low weight;
    and guilt/shame around eating
    (e.g., anxiety about food,
    avoidance of eating situations).
    Other features include using food
    as emotional coping and obsessive
    focus on food, cooking, or
    weight-related content. Behavioral
    manifestations include restrictive
    eating (e.g., smaller portions,
    skipping meals, eliminating food
    groups, extreme diets, calorie
    counting); fasting to compensate;
    and chewing and spitting (chewing
    food without swallowing). Binge
    eating involves consuming
    abnormally large amounts of food
    with feelings of loss of control,
    eating rapidly or until
    uncomfortably full, eating alone
    due to embarrassment, and
    experiencing negative emotions
    afterward. Compensatory behaviors
    include self-induced vomiting,
    misuse of laxatives or diuretics,
    use of diet pills or "slimming
    teas," fasting as a non-purging
    compensatory behavior and excessive
    exercise (e.g., working out despite
    fatigue/injury, exercise
    interfering with daily
    functioning). Secretive eating
    behaviors, including concealing
    eating patterns or using coded
    terminology online, may also occur.

- Sexual:
Involves difficulties with sexual
    functioning and satisfaction. This
    includes problems with sexual
    desire, arousal, or performance;
    pain during sexual activity;
    distress related to sexual
    experiences; and addictive sexual
    behaviors that impair various other
    functions.

- Thought Disorder:
Relates to disruptions in perception,
    beliefs, and thought organization.
    This includes unusual sensory
    experiences (e.g., hearing voices);
    extreme and fixed, false beliefs;
    disorganized thinking or speech;
```

```
    and difficulty distinguishing
    between what is real and what is
    not.

- Antisocial and Antagonistic Behavior:
Refers to actions that violate societal
    norms and infringe upon the rights
    of others. This includes
    rule-breaking or illegal behavior,
    physical aggression, bullying,
    deceitfulness or theft, blame
    externalization, disregard for
    social norms or obligations, and
    property destruction. It also
    covers behaviors such as
    deliberately provoking or annoying
    others, attention-seeking,
    rudeness, manipulating others for
    personal gain, and other
    deliberately over-extroverted
    behaviors (physical or emotional).
    These behaviors may be accompanied
    by impulsivity, distractibility,
    and risk-taking.

- Detachment:
Characterized by withdrawal from social
    and emotional experiences. This
    includes avoidance of social
    interaction, limited emotional
    expression, lack of close
    relationships, reduced capacity to
    experience pleasure in activities
    (particularly in social contexts),
    and social isolation.

- None:
Applies when the post does not clearly
    reflect any of the above symptom
    categories.

Instructions:
1. Read the post carefully and evaluate
    whether it matches each defined
    symptom category.
2. For each category, explain your
    reasoning. If a category applies,
    support your answer with direct
    evidence from the post. If it does
    not apply, explain why there is
    insufficient or no evidence.
    Clearly state "yes" or "no" for
    each category.
3. Finally, list all applicable symptom
    categories that are present in the
    post. If more than one applies,
    separate them with commas.

Below are some examples:
{few_shot_examples}

Post to Analyze:
Post:
"{post}"

Please strictly follow the output
    format exactly as shown below. Do
    not use bold, markdown, or extra
    formatting.

Output Format:
Explanation:
- Suicidal Thoughts: [reason]. So the
    answer is yes (or is no).
- Somatoform & Substance Abuse:
    [reason]. So the answer is yes (or
    is no).
- Eating Pathology: [reason]. So the
    answer is yes (or is no).
- Sexual: [reason]. So the answer is
    yes (or is no).
- Thought Disorder: [reason]. So the
    answer is yes (or is no).
- Antisocial and Antagonistic Behavior:
    [reason]. So the answer is yes (or
    is no).
- Detachment: [reason]. So the answer
    is yes (or is no).
- None: [reason]. So the answer is yes
    (or is no).

Answer:
Output exact category names: Suicidal
    Thoughts, Somatoform & Substance
    Abuse, Eating Pathology, Sexual,
    Thought Disorder, Antisocial and
    Antagonistic Behavior, Detachment,
    or None. Use commas to separate
    multiple labels, if any."""
```

A.5 Chain-of-Thought Examples

To illustrate the reasoning process required by our prompts, we provide synthetic in context learning examples demonstrating the chain-of-thought approach. These examples are fictional and created for demonstration purposes to protect the privacy of the individuals in our dataset. The actual prompts used during annotation contained 5-10 similar examples per stream, all following the same structured reasoning format shown below. Each example demonstrates explicit evaluation of every category before producing a final classification.

A.5.1 Emotion Classification Example

```
Post:
"Finally submitted my thesis after
    months of work! Feeling accomplished
and ready for what's next."

Explanation:
- Happy: The post explicitly states
    "Feeling accomplished" which
    directly
  indicates pride and satisfaction. The
      forward-looking tone suggests
  hopefulness. So the answer is yes.
- Proud: While accomplishment is
    present, this aligns more with
    general
  happiness than specific pride in
      achievement. So the answer is no.
- Sad: The post expresses positive
    emotions with no grief or emotional
    pain.
```

```
  So the answer is no.
- Anxious: No fear, worry, or tension
    is expressed. So the answer is no.
- Depressed: The post conveys
    accomplishment and optimism,
    opposite to
  despair. So the answer is no.
- Apathetic: Clear emotional engagement
    is present, not numbness.
  So the answer is no.
- Angry: No aggression or anger is
    expressed. So the answer is no.
- Ashamed: The post expresses positive
    feelings, not guilt or
    embarrassment.
  So the answer is no.
- Loneliness: No isolation or lack of
    connection is mentioned.
  So the answer is no.
- None: Clear positive emotion is
    present. So the answer is no.

Answer:
Happy
```

A.5.2 Life Events Classification Example

```
Post:
"Started therapy last week and got
    prescribed medication for my
    anxiety.
Hoping this helps."

Explanation:
- Mental Health: The post explicitly
    mentions starting therapy and
    receiving
  psychiatric medication for anxiety,
      which are major mental health
      treatment
  events. So the answer is yes.
- Physical Health: No accidents,
    injuries, illnesses, or physical
    medical
  interventions are mentioned. So the
      answer is no.
- Abuse & Addiction: No mention of
    abuse or substance-related issues.
  So the answer is no.
- Relationship & Loss: No specific
    relationship changes are described.
  So the answer is no.
- Career & Education: No work or
    school-related events are mentioned.
  So the answer is no.
- Financial & Legal & Societal: No
    financial, legal, or societal events
  are discussed. So the answer is no.
- Lifestyle & Identity & Environment:
    The post focuses on clinical
  intervention rather than lifestyle
      modifications. So the answer is
      no.
- None: A clear mental health life
    event is present. So the answer is
    no.

Answer:
Mental Health
```

A.5.3 Symptoms Classification Example

```
Post:
"Constant headaches all week and
    exhausted no matter how much I
    sleep.
Also been drinking more than usual to
    cope."

Explanation:
- Suicidal Thoughts: No death wishes or
    suicidal ideation are expressed.
  So the answer is no.
- Somatoform & Substance Abuse: The
    post mentions "constant headaches"
    and
  chronic exhaustion (somatic symptoms)
      plus "drinking more than usual to
  cope" (substance use for emotional
      regulation). So the answer is yes.
- Eating Pathology: No eating, weight,
    or body image concerns are
    mentioned.
  So the answer is no.
- Sexual: No sexual symptoms are
    discussed. So the answer is no.
- Thought Disorder: The thinking is
    coherent with no perceptual
    disturbances.
  So the answer is no.
- Antisocial and Antagonistic Behavior:
    No rule-breaking or aggression is
  described. So the answer is no.
- Detachment: While fatigue is present,
    there's no indication of social
  withdrawal. So the answer is no.
- None: Clear symptoms are present. So
    the answer is no.

Answer:
Somatoform & Substance Abuse
```

B Baseline LLM Prompting Templates for Mood Change Detection

This section details the complete prompt templates and model configurations used for mood change detection (MoC) experiments with LLMs.

B.1 Model Configuration

For automatic mood change detection, we used the following configuration:

- **Model:** Qwen2.5-32B-Instruct
- **Quantization:** 4-bit (NF4)
- **Precision:** `bfloat16`
- **Decoding Strategy:** Greedy decoding (`do_sample=False`)
- **Attention Implementation:** Default (no custom optimization)

- **Tokenizer Padding:** Right-side padding with EOS token substitution
- **Hardware:** Single NVIDIA GPU with 4-bit quantized loading via BitsAndBytes
- **Max New Tokens:** 32
- **Temperature:** 0.0

B.2 Mood Change Detection Prompt (QwenLLM)

```
You are an expert assistant for
    identifying self-disclosed mood
    changes in user timelines on social
    media.

Your task is to classify the current
    post into exactly one of these
    three labels based on the context
    of previous posts:

- S (Switch): A sudden shift in mood
    direction (positive  negative)
    compared to prior posts.
- E (Escalation): A gradual or clear
    intensification of mood in the same
    direction.
- O (None): No significant mood change
    or mood remains neutral/consistent.

### Instructions:
- Read all previous posts in the
    timeline as context to understand
    the user's prior mood.
- Read the current post carefully.
- Determine if the current post
    represents a sudden mood switch
    (S), a gradual escalation (E), or
    no significant mood change (O).
- **Output only the label (S, E, or O)
    on a single line no extra text,
    explanation, or commentary**
- If there are **no previous posts**,
    assume the baseline mood is neutral.

Previous posts:
{prev_posts if prev_posts else "No
    previous posts"}

Current post:
{current_post}

Answer:
```

Listing 1: QwenLLM prompt template for timeline-based MoC classification.

B.3 Mood Change Detection (QwenLLMMultistream) Prompt

```
You are an assistant for identifying a
    user's self-disclosed mood changes
    over time.
Each post belongs to a timeline of user
    posts. Your task is to classify the
    current
post based on prior posts as one of the
    following:

- S (Switch): Sudden shift in mood
    direction (positive  negative)
- E (Escalation): Gradual or clear
    intensification of mood in the same
    direction
- O (None): No significant mood change
    or mood remains consistent

AUTOMATED MULTISTREAM LABEL CONTEXT:
Each post includes automated labels
    from three streams:

1. FINE-GRAINED EMOTION LABELS (12
    categories organized by
    adaptiveness):

   Positive Adaptive Emotions:
   - Happy: Content, joyful, hopeful,
       excited, glad, cheerful
   - Proud: Feeling accomplished,
       satisfied with achievements,
       confident in success
   - Calm: Peaceful, relaxed, serene,
       tranquil emotional states
   - Vigor: High energy, vitality,
       enthusiasm, active engagement
   - Feeling loved: Experiencing
       affection, care, or emotional
       support from others

   Negative Adaptive Emotions:
   - Sad: Emotional pain, grieving,
       hurt, heartbroken over loss
       (adaptive response to loss)
   - Justifiable anger: Appropriate
       anger in response to unfair
       treatment or injustice

   Negative Maladaptive Emotions:
   - Anxious: Fearful, tense, worried,
       nervous, scared, panicked
   - Depressed: Despair, hopeless,
       worthless, empty, suicidal
       thoughts
   - Apathetic: Numb, indifferent,
       don't care, emotionally blunted
   - Angry: Aggression, hate, rage,
       disgust, contempt (non-justified)
   - Ashamed: Guilty, embarrassed,
       humiliated, self-critical
   - Loneliness: Feeling alone,
       isolated, lacking meaningful
       connection


2. MENTAL HEALTH SYMPTOM LABELS (7
    categories):

   - Suicidal Thoughts: Specific
       thoughts about ending one's
       life, including considering
       methods, making plans,
       expressing desire to die,
       passive thoughts ("I wish I
```

wouldn't wake up"), active plans, references to past suicide attempts

- Somatoform & Substance Abuse: Physical symptoms causing significant distress without clear medical explanations (headaches, muscle aches, unexplained sensations, physical fatigue, cognitive symptoms like impaired memory/processing). Substance abuse includes problematic alcohol/drug use, failed attempts to cut down, excessive time obtaining/using/recovering from substances, continued use despite negative consequences

- Eating Pathology: Problematic food/eating behaviors and cognitions including fear of weight gain, body image disturbance, self-worth tied to appearance, guilt/shame around eating, using food as emotional coping, restrictive eating patterns, binge eating with loss of control, compensatory behaviors (vomiting, laxatives, excessive exercise), secretive eating behaviors

- Sexual: Difficulties with sexual functioning including problems with desire/arousal/performance, pain during sexual activity, distress related to sexual experiences, addictive sexual behaviors

- Thought Disorder: Disruptions in perception, beliefs, thought organization including unusual sensory experiences (hearing voices), extreme fixed false beliefs, disorganized thinking/speech, difficulty distinguishing reality

- Antisocial and Antagonistic Behavior: Actions violating societal norms and others' rights including rule-breaking, illegal behavior, physical aggression, bullying, deceitfulness, blame externalization, property destruction, deliberately provoking others, attention-seeking, rudeness, manipulation for personal gain

- Detachment: Withdrawal from social and emotional experiences including avoidance of social interaction, limited emotional expression, lack of close relationships, reduced capacity for pleasure (especially social contexts), social isolation

3. PERSONAL LIFE EVENT LABELS (7 categories):

- Mental Health: Receiving formal mental health diagnosis, starting/adjusting psychiatric medication, beginning therapy, recovery/significant symptom improvement, acute psychological episodes (manic episodes, psychotic breaks, panic attacks, dissociative episodes, self-harm, suicide attempts, acute suicidal ideation requiring crisis intervention)

- Physical Health: Accidents, injuries, diagnoses/survival of serious illnesses, chronic conditions with notable daily life impact, hospitalization, surgery, emergency medical treatment, major medical interventions, pregnancy-related experiences (becoming pregnant, pregnancy loss, abortion, complications)

- Abuse & Addiction: Major events involving abuse (physical, sexual, emotional, psychological) in childhood or adulthood; addiction-related events including onset of heavy drug/alcohol use, acknowledgment of addiction, overdoses, withdrawal symptoms, relapses, recovery (treatment programs, rehabilitation, support groups)

- Relationship & Loss: Specific relationship changes/disruptions in meaningful personal connections with strong emotional impact including beginning/ending friendships or romantic relationships, marriage/divorce, parental separation/divorce, serious arguments/conflicts, relationships becoming abusive/toxic, addition of new family members, parenting difficulties, emotionally distressing health issues involving family/partners/close friends, death of important person or pet

- Career & Education: Career events (starting/losing job, promotions, demotions, work troubles, unable to find job, retirement, becoming business owner); Education events (starting/graduating school,

```
        transferring schools, leaving
        without graduating, denied
        school entry, major exams,
        significant academic challenges,
        major academic milestones)

    - Financial & Legal & Societal:
        Financial events (major
        purchases like house/car,
        financial gains/milestones,
        financial losses, ongoing
        financial difficulties); Legal
        events (law violations, arrests,
        court appearances, lawsuits,
        major justice system
        interactions); Societal events
        with personal impact (natural
        disasters, pandemics, major
        political events, societal
        issues, notable public
        encounters)

    - Lifestyle & Identity &
        Environment: Lifestyle changes
        (adopting new health routines,
        diet/exercise/sleep/stress
        management adjustments, reducing
        substance use, fostering
        positive social connections,
        getting pet - not clinical
        mental health treatment);
        Identity transitions
        (gender/sexual orientation
        identification/changes, coming
        out as LGBTQ+, new sexual
        experiences,
        exploring/adopting/changing
        political/religious beliefs);
        Relocation events (moving to new
        place, moving in/out of family
        home, family members moving
        in/out of household, significant
        travel experiences)

INTERPRETATION GUIDELINES:
- Fine-grained emotions (especially
    transitions between
    Positive/Negative
    Adaptive/Maladaptive categories)
    are primary indicators of mood
    switches and escalations
- Life events often trigger mood
    changes - look for temporal
    relationships between events and
    emotional shifts
- Symptom patterns can indicate
    escalating mental health states or
    underlying mood deterioration
- Cross-stream interactions (e.g.,
    relationship loss + depressed
    emotion + detachment symptoms)
    provide rich context for mood
    trajectory assessment
- Labels are automated and noisy but
    provide valuable psychological
    state indicators for detecting mood
    changes over time

You will ONLY output one of the three
    labels: S, E, or O. No additional
    explanation.

Previous posts in timeline:
{prev_context if prev_context else "No
    previous posts."}

Current post to classify:
{current_formatted}

Classification:
```

Listing 2: QwenLLMMultistream Prompt template for multistream mood change detection (S, E, O).

C Training and Implementation

C.1 Infrastructure

All models (excluding LLMs) were implemented with PyTorch and trained on 2 NVIDIA A100-PCIE-40GB GPUs using Focal Loss to address class imbalance in labels of Moments of Change - for single-task models. LLM models were all implemented with 2 NVIDIA H100 GPUs.

Multi-task models used a combined loss of focal loss and binary cross entropy (BCE) loss, giving equal weighting to both losses and equal weighting to the 3 streams in the BCE loss. The optimization process used the Adam optimizer with gradient accumulation of 4 steps. Training was carried out with a maximum of 10 epochs, selecting the best model in terms of macro-average F1 score on a validation set using 5-fold cross validation. The hyperparameter search space is described in Table C.2.

C.2 Grid-search Used in Experimental Procedure

We train and evaluate all our models on 3 seeds, taking the average scores on the resulting test sets - aside from the LLM baselines which were evaluated with 1 seed. We evaluate on the same test set proposed in the CLPsych 2022 shared task for Reddit (Tsakalidis et al., 2022a). We performed a grid-search on over the enlisted hyperparameters – selecting the best performing model based on macro-average F1 score on the validation set, and optimizing the model using focal loss with a gamma of 2.0, training for 10 epochs, Learning rate: 0.00005, LSTM/ Transformer hidden dimension: 512, $\epsilon_{\text{prior}} = \beta_{\text{prior}} = 0.001$ for MHRoBERT and $\epsilon_{\text{prior}} = \beta_{\text{prior}} = 0.01$ for HoRoBERT, chunk size: 16, stride: 8.