ACUMEN: INTERNATIONAL JOURNAL OF MULTIDISCIPLINARY RESEARCH
ISSN: 3060-4745 IF(Impact Factor)10.71 / 2025 Volume 3, Issue 5

SHADOWS IN YOUNG MINDS: IDENTIFYING SUICIDE RISK IN ADOLESCENTS

ANBARASU , Medical Student, Samarkand state medical university, Uzbekistan, anbarasujayaganesh29@gmail.com

DINESH , Medical Student, Samarkand state medical university, Uzbekistan, dineshkanna678@gmail.com

SAFA, Medical Student, Samarkand state medical university, Uzbekistan, ningjingsafa@gmail.com

HAMNA , Medical Student, Samarkand state medical university, Uzbekistan, hamna3383@gmail.com

Abstract

The escalating rate of adolescent suicide represents a profound global public health crisis that demands immediate and innovative intervention strategies. As traditional clinical assessments often fail to capture the dynamic and multifaceted nature of suicidal ideation, computational methods have emerged as a vital supplementary tool. This paper proposes a comprehensive, multimodal machine learning framework designed to identify suicide risk in adolescents by integrating clinical health records, acoustic speech features, and linguistic markers from digital platforms. By synthesizing diverse data modalities and applying semi-supervised learning techniques, this approach aims to overcome the limitations of isolated datasets and provide a more robust, real-time assessment of patient vulnerability. Ultimately, the integration of advanced predictive models into psychiatric care holds the potential to facilitate timely interventions, reduce the cognitive burden on crisis responders, and save young lives.

Introduction

Adolescent suicide has tragically become one of the leading causes of mortality among young people worldwide, representing an urgent crisis that overwhelms contemporary healthcare systems. The transition through adolescence is fraught with complex psychological, social, and physiological changes, which can be exacerbated by economic pressures, academic stress, and external crises. While historical efforts to assess suicide risk have relied heavily on direct clinical interviews and self-reported questionnaires, predictive capabilities using these standard methods are often only slightly better than chance (Bhat & Goldman-Mellor, 2017). Consequently, the problem definition of this paper centers on designing a robust, automated methodology capable of identifying early markers of suicide risk in adolescents across multiple behavioral and physiological modalities before acute crises manifest.

Despite decades of psychiatric research, existing approaches to suicide risk prediction remain highly insufficient for several structural reasons. First, traditional self-report assessments and clinical interviews are subject to significant misclassification, as they rely entirely on the patient's immediate self-awareness and willingness to disclose their most vulnerable thoughts (Marie et al., 2025). Second, many contemporary machine learning efforts remain siloed in their data sources; models trained exclusively on electronic health records (EHRs) lack the real-time behavioral context of the patient, while NLP-based social media models frequently suffer from severe class imbalances and a chronic lack of reliable, ground-truth labeled data (Lovitt et al., 2024). Because risk factors fluctuate rapidly in youth, static and unimodal assessments consistently fail to capture the precipitous decline in a patient's mental state.

To address these critical gaps in computational psychiatry, this paper introduces a novel approach to early risk detection. Specifically, the paper's contributions are as follows:

- First, we propose a comprehensive, multimodal suicide risk assessment framework that strategically fuses structured clinical history with unstructured speech and linguistic data to generate a holistic patient profile.
- Second, we present a semi-supervised evaluation strategy designed to mitigate the persistent challenge of imbalanced datasets and sparse annotations in computational mental health research.

Related Work

Clinical Data and Health Behaviors

The foundational category of suicide risk prediction leverages structured clinical records and epidemiological data. Researchers have successfully utilized anonymized electronic health records, encompassing hundreds of thousands of adolescents, to build neural network models capable of predicting suicide attempts based on previous hospital visits and demographic variables (Bhat & Goldman-Mellor, 2017). Similarly, cross-sectional studies have applied data-driven models like Bayesian tree ensembles to link specific health behaviors—such as substance abuse, early sexual activity, and feelings of hopelessness—to elevated suicide risk among minority and vulnerable youth populations (Wei & Mukherjee, 2020). The primary strength of this approach is its reliance on verifiable, high-impact clinical history. However, its weakness lies in its static nature; relying on past hospital admissions or survey data inherently misses the acute, real-time development of suicidal ideation. In contrast to these strictly clinical models, our work supplements historical data with dynamic acoustic and linguistic features to capture immediate emotional volatility.

Speech and Acoustic Analysis

A growing body of literature focuses on non-invasive suicide risk assessment through computational paralinguistics and speech analysis. Recent studies have demonstrated that deep learning-based spectral representations, such as wav2vec, and handcrafted acoustic features like Mel-frequency cepstral coefficients (MFCCs) can successfully differentiate high-risk patients from low-risk patients (Amiriparian et al., 2024). Furthermore, research highlights the necessity of gender-specific modeling, as speech characteristics linked to suicide risk manifest differently across genders; for instance, agitation correlates with increased risk in males but points the opposite way in females (Gerczuk et al., 2024). While highly innovative, these speech models often struggle with generalization across different recording environments and face performance drops between development and test datasets (Marie et al., 2025). Our proposed framework addresses this weakness by treating speech as a complementary module that is continuously contextualized by the patient's clinical metadata, rather than relying on it as an isolated diagnostic metric.

Social Media and Conversational Agents

The third category encompasses the application of natural language processing (NLP) to social media text and the deployment of conversational AI in clinical workflows. Platforms like Reddit have become valuable resources for studying mental health, prompting researchers to develop weakly-supervised frameworks and pseudo-labeling techniques to overcome the scarcity of annotated suicide risk data (Yang et al., 2021). Additionally, advanced conversational assistants, powered by domain-adapted transformer models, have been engineered to assist frontline crisis responders by automatically identifying issue tags during youth support chats (Obadinma et al., 2024). While these text-based models excel at identifying distress in real-time, they often lack the clinical rigor and patient history necessary to confirm long-term risk (Lovitt et al., 2024). By integrating semi-supervised linguistic analysis with formal clinical EHR frameworks, our approach bridges the gap between anonymous digital cries for help and actionable psychiatric intervention.

Method/Approach

To effectively capture the multifaceted nature of adolescent suicide risk, we propose a multimodal machine learning framework consisting of three distinct processing modules: a Clinical History Module, an Acoustic Processing Module, and a Linguistic & Social Module. The Clinical History Module processes structured tabular data, including demographics, past emergency department visits, and documented health behaviors such as substance use or physical fighting (Wei & Mukherjee, 2020). The Acoustic Processing Module ingests short, non-invasive voice recordings, applying a fine-tuned wav2vec 2.0 architecture alongside gender-specific normalization layers to extract deep spectral features (Gerczuk et al., 2024). Finally, the Linguistic Module utilizes an ensemble of transformer models to analyze conversational transcripts and social media text, employing pseudo-labeling techniques derived from related mental health domains like anxiety and depression to bolster its predictive power (Yang et al., 2021).

The key design choices in this architecture revolve around the sophisticated fusion of these disparate modalities. We utilize a weighted attention mechanism with mixup regularization for late fusion, allowing the model to dynamically assign importance to different data streams based on the available inputs (Marie et al., 2025). This choice is highly intentional; in emergency medicine, a patient might present with extensive clinical records but refuse to speak, or conversely, provide ample conversational data but lack a formal psychiatric history. By incorporating metadata such as biological sex and previous suicide attempts into the fusion layer, the model contextualizes the unstructured speech and text embeddings, which has been shown to yield significant absolute improvements in balanced accuracy (Amiriparian et al., 2024).

The evaluation plan for this proposed methodology relies on a hypothetical, large-scale, multimodal dataset constructed from partnering youth psychiatric wards and crisis hotlines. We will frame the assessment as a binary classification problem to predict the presence or absence of acute suicide attempts within a designated followup period (Bhat & Goldman-Mellor, 2017). Given the inherent scarcity of positive labels, we will implement a semi-supervised learning paradigm, manually verifying a subset of pseudo-labeled text data to ensure quality control across multiple training trials (Lovitt et al., 2024). Model performance will be rigorously measured using precision, recall, F-score metrics, and the Area Under the Receiver Operating Characteristic Curve (ROC-AUC), ensuring that the framework prioritizes high sensitivity to minimize catastrophic false negatives in clinical deployment.

Discussion

The practical implications of deploying this multimodal framework in real-world healthcare settings are substantial. If integrated into emergency department triage systems, the model could function as a rapid, non-invasive screening tool that automatically flags high-risk adolescents for immediate psychiatric evaluation. Furthermore, this predictive framework could be synergized with Computational Health Economic Models (CHEMs) to help policymakers simulate the cost-effectiveness of implementing AI-augmented early intervention programs at a national level (Hamilton et al., 2023). By embedding these insights into tools utilized by frontline crisis responders, mental health agencies can drastically reduce the cognitive burden on staff, streamline administrative issue identification, and optimize the allocation of severely limited clinical resources (Obadinma et al., 2024).

Despite its potential, the proposed framework is constrained by several critical limitations and failure modes. First, the reliance on fused modalities introduces a vulnerability to missing data; if a patient does not engage with social media or is selectively mute during an assessment, the absence of linguistic or acoustic features may severely degrade the model's predictive confidence. Second, the framework faces significant domain shift challenges; models trained on pristine clinical speech recordings often fail to generalize to the noisy, chaotic acoustic environments typical of emergency departments (Marie et al., 2025). Third, the incorporation of semi-supervised learning introduces a risk of confirmation bias, wherein poorly generated pseudo-labels from unlabeled digital text might propagate classification errors deep into the model's architecture, inadvertently misclassifying low-risk distress as high-risk ideation (Yang et al., 2021).

The deployment of such technology also necessitates a rigorous examination of ethical considerations and inherent risks. A primary ethical risk is the potential for algorithmic bias, particularly against marginalized minority groups who are already at a disproportionately higher risk for attempting suicide (Wei & Mukherjee, 2020). If the training data lacks sufficient representation of diverse dialects, socioeconomic backgrounds, or cultural expressions of grief, the model may systematically fail the most vulnerable adolescents. Another critical consideration involves patient privacy and transparency; qualitative research indicates that young people demand to know "what's under the hood" of AI agents, fearing that overly automated systems might dehumanize their care or compromise their confidentiality without consent (Poulsen et al., 2026).

To advance this critical field of study, future work must explore several new directions. First, researchers should pivot toward longitudinal tracking frameworks that utilize continuous, passive sensing (such as wearable devices) to monitor physiological shifts over time, rather than relying on cross-sectional snapshots of clinical visits. Second, future iterations of this technology should focus on co-designing conversational interfaces directly with youth consumers, ensuring that AI-driven assessments act as personalized, empathetic agents that humanize the clinical experience while operating strictly within safe, governed boundaries (Poulsen et al., 2026).

Conclusion

The rising tide of adolescent suicide demands interventions that transcend the limitations of traditional, unimodal psychiatric assessments. This paper has outlined a comprehensive, multimodal machine learning framework that synthesizes clinical history, deep acoustic speech representations, and semi-supervised linguistic analysis to accurately identify suicide risk in youth populations. By capturing the complex interplay between static demographic risk factors and dynamic emotional states, this methodology offers a more resilient and nuanced diagnostic tool for emergency medical personnel and crisis responders.

Ultimately, while the integration of artificial intelligence into youth mental health services presents profound ethical and technical challenges, it also offers a beacon of hope for overstretched healthcare systems. By prioritizing transparent, generalizable, and ethically governed predictive models, the psychiatric community can harness computational power to detect the earliest shadows in young minds. Doing so ensures that vulnerable adolescents are met with the right tools, in the right place, at the right time, fundamentally transforming early intervention and saving lives.

References

Bhat, Harish S., & Goldman-Mellor, Sidra J. (2017). *Predicting Adolescent Suicide Attempts with Neural Networks*. https://arxiv.org/pdf/1711.10057v2 https://arxiv.org/pdf/1711.10057v2

Marie, Ambre, Maoudj, Ilias, Dardenne, Guillaume, & Quellec, Gwenolé (2025). *Suicide Risk Assessment Using Multimodal Speech Features: A Study on the SW1 Challenge Dataset*. https://arxiv.org/pdf/2505.13069v1 https://arxiv.org/pdf/2505.13069v1

Lovitt, Max, Ma, Haotian, Wang, Song, & Peng, Yifan (2024). *Suicide Risk Assessment on Social Media with Semi-Supervised Learning*. https://arxiv.org/pdf/2411.12767v2 https://arxiv.org/pdf/2411.12767v2

Wei, Zhiyuan, & Mukherjee, Sayanti (2020). *Health-behaviors associated with the growing risk of adolescent suicide attempts: A data-driven cross-sectional study*. https://arxiv.org/pdf/2009.03966v1 https://arxiv.org/pdf/2009.03966v1

Amiriparian, Shahin, Gerczuk, Maurice, Lutz, Justina, Strube, Wolfgang, Papazova, Irina, Hasan, Alkomiet, Kathan, Alexander, & Schuller, Björn W. (2024). *Non-Invasive Suicide Risk Prediction Through Speech Analysis*. https://arxiv.org/pdf/2404.12132v3 https://arxiv.org/pdf/2404.12132v3

Gerczuk, Maurice, Amiriparian, Shahin, Lutz, Justina, Strube, Wolfgang, Papazova, Irina, Hasan, Alkomiet, & Schuller, Björn W. (2024). *Exploring Gender-Specific Speech Patterns in Automatic Suicide Risk Assessment*. https://arxiv.org/pdf/2407.11012v1 https://arxiv.org/pdf/2407.11012v1

Yang, Chenghao, Zhang, Yudong, & Muresan, Smaranda (2021). *Weakly-Supervised Methods for Suicide Risk Assessment: Role of Related Domains*. https://arxiv.org/pdf/2106.02792v2 https://arxiv.org/pdf/2106.02792v2

Obadinma, Stephen, Lachana, Alia, Norman, Maia, Rankin, Jocelyn, Yu, Joanna, Zhu, Xiaodan, Mastropaolo, Darren, Pandya, Deval, Sultan, Roxana, & Dolatabadi, Elham (2024). *FAIIR: Building Toward A Conversational AI Agent Assistant for Youth Mental Health Service Provision*. https://arxiv.org/pdf/2405.18553v4 https://arxiv.org/pdf/2405.18553v4

Hamilton, Matthew P, Gao, Caroline X, Wiesner, Glen, Filia, Kate M, Menssink, Jana M, Plencnerova, Petra, Baker, David G, McGorry, Patrick D, Parker, Alexandra, Karnon, Jonathan, Cotton, Sue M, & Mihalopoulos, Cathrine (2023). *A prototype software framework for transferable computational health economic models and its early application in youth mental health*. PharmacoEconomics (2024). https://doi.org/10.1007/s40273-024-01378-8 https://doi.org/10.1007/s40273-024-01378-8

Poulsen, Adam, Hickie, Ian B., Gorban, Carla, Haan, Zsofi de, Capon, William, Eyeson-Annan, Ebenezer, Radwan, Jalal, Scott, Elizabeth M., Iorfino, Frank, & LaMonica, Haley M. (2026). *Young people's perceptions and recommendations for conversational generative artificial intelligence in youth mental health*. https://arxiv.org/pdf/2604.13381v1 https://arxiv.org/pdf/2604.13381v1