



RS Global
Journals

Scholarly Publisher
RS Global Sp. z O.O.
ISNI: 0000 0004 8495 2390

Dolna 17, Warsaw, Poland 00-773
Tel: +48 226 0 227 03
Email: editorial_office@rsglobal.pl

JOURNAL	World Science
p-ISSN	2413-1032
e-ISSN	2414-6404
PUBLISHER	RS Global Sp. z O.O., Poland
ARTICLE TITLE	A REVIEW ON MACHINE LEARNING APPROACHES FOR THE DETECTION OF SUICIDAL TENDENCIES
AUTHOR(S)	Kazi Golam Rabbany, Aisultan Shoiynbek, Darkhan Kuanyshbay, Assylbek Mukhametzhanov, Akbayan Bekarystankyzy, Temirlan Shoiynbek
ARTICLE INFO	Kazi Golam Rabbany, Aisultan Shoiynbek, Darkhan Kuanyshbay, Assylbek Mukhametzhanov, Akbayan Bekarystankyzy, Temirlan Shoiynbek. (2024) A Review on Machine Learning Approaches for the Detection of Suicidal Tendencies. <i>World Science</i> . 3(85). doi: 10.31435/rsglobal_ws/30092024/8222
DOI	https://doi.org/10.31435/rsglobal_ws/30092024/8222
RECEIVED	12 August 2024
ACCEPTED	16 September 2024
PUBLISHED	19 September 2024
LICENSE	 This work is licensed under a Creative Commons Attribution 4.0 International License .

© The author(s) 2024. This publication is an open access article.

A REVIEW ON MACHINE LEARNING APPROACHES FOR THE DETECTION OF SUICIDAL TENDENCIES

Kazi Golam Rabbany

Narxoz University

ORCID ID: 0009-0007-4549-0815

Aisultan Shoiynbek

PhD, Professor, Narxoz University

ORCID ID: 0000-0002-9328-8300

Darkhan Kuanyshbay

PhD, Assistant Professor, SDU University

ORCID ID: 0000-0001-5952-8609

Assylbek Mukhametzhano

Master's student, SDU University

ORCID ID: 0009-0009-8528-9985

Akbayan Bekarystankyzy

PhD, Senior lecturer, Narxoz University

ORCID ID: 0000-0003-3984-2718

Temirlan Shoiynbek

Ms, Senior lecturer, Narxoz University

DOI: https://doi.org/10.31435/rsglobal_ws/30092024/8222

ARTICLE INFO

Received: 12 August 2024

Accepted: 16 September 2024

Published: 19 September 2024

KEYWORDS

Suicide Prevention, Depression Detection, Machine Learning, Natural Language Processing, Speech Analysis, Social Media Data, Clinical Data Analysis.

ABSTRACT

With the increasing prevalence of mental health issues, particularly suicidal behaviors, the need for early and accurate detection has become critical. This paper explores the current landscape of machine learning approaches used for the detection of suicidal tendencies. It examines a wide range of machine learning techniques applied to various data sources, including social media, clinical records, psychological assessments, self-reported forms like PHQ-9, audio speech recordings, and multimodal data integrating speech and visual information. This comprehensive review aims to reveal the types of existing research based on these varied datasets, highlighting the nuances of data collection, significant features identified, and the results obtained by different studies. Additionally, the review discusses the challenges and limitations associated with these approaches, providing researchers and practitioners with valuable insights into the potential and pitfalls of machine learning applications in diagnosing individuals at risk of suicide. The goal is to inform future research and improve early detection methods to ultimately reduce suicide rates.

Citation: Kazi Golam Rabbany, Aisultan Shoiynbek, Darkhan Kuanyshbay, Assylbek Mukhametzhano, Akbayan Bekarystankyzy, Temirlan Shoiynbek. (2024) A Review on Machine Learning Approaches for the Detection of Suicidal Tendencies. *World Science*. 3(85). doi: 10.31435/rsglobal_ws/30092024/8222

Copyright: © 2024 Kazi Golam Rabbany, Aisultan Shoiynbek, Darkhan Kuanyshbay, Assylbek Mukhametzhano, Akbayan Bekarystankyzy, Temirlan Shoiynbek. This is an open-access article distributed under the terms of the **Creative Commons Attribution License (CC BY)**. The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

1. Introduction.

Suicide is a critical public health challenge that claims over 700,000 deaths annually worldwide. This phenomenon is not uniformly distributed across all age groups; it is particularly pronounced among the youth. Globally, suicide ranks as the fourth leading cause of death among individuals aged 15 – 29, highlighting its significant impact on the most productive years of human life.

The complexity of suicide arises from a myriad of interrelated factors. It's not solely mental health disorders that contribute, but a broader spectrum including substance abuse, social isolation, history of violence or abuse, loss, conflict, disaster, and societal stigma. High-risk groups include individuals facing acute crises like financial troubles or relationship issues, which can lead to impulsive actions. This complexity highlights the significance of the issue, underscoring the need for diverse and in-depth understanding and research. It's essential to recognize the multifaceted nature of suicide to grasp its impact fully and to inform the development of comprehensive strategies for addressing this critical public health concern.

2. Suicidal Tendency.

2.1 Warning Signs and Risk Factors. For the researchers, it is crucial to understand and recognize the signs that might indicate suicidal tendencies. These warning signs are diverse and multifaceted, often manifesting in behavioral, emotional, and environmental patterns.

The risk factors for suicide encompass a broad range of individual, relationship, community, and societal aspects. These factors, while not deterministic, significantly contribute to the likelihood of suicidal thoughts or behaviors. At the individual level, factors such as a history of previous suicide attempts, mental health conditions like depression, bipolar disorder, schizophrenia, and substance use disorders, and persistent feelings of despair are key contributors. Chronic illnesses and pain also play a significant role in increasing suicide risk.

2.2 Defining “Suicidal Tendency”. Suicide risk is defined as the probability of suicide attempt or death within a specific timeframe, focusing on immediate and future risks [1]. Suicidal ideation can be defined, in the context of social media, as thoughts about ending one's life, emphasizing a range from passive thoughts to active planning [2].

3. Machine Learning Approaches for Suicidal Tendency Detection. Numerous researchers have made significant strides, particularly concerning suicidal tendencies, suicidal ideation, and depression. These studies have demonstrated the promising potential of machine learning as a tool to address these critical mental health issues. Even though there are many simpler traditional statistical analytical studies available [3], this paper focuses on the studies using machine learning approaches.

3.1 Contemporary Model Evaluation Techniques.

AUC-ROC is a performance metric used to evaluate classification models in machine learning. It measures a model's ability to distinguish between classes, especially useful in imbalanced scenarios like predicting rare events. AUC values range from 0 to 1. A higher AUC indicates better model performance, with 1 being perfect classification and 0.5 equating to random guessing. [4], [5], [6].

Other metrics, like Precision, Sensitivity, Specificity etc. are also important factors for mental health as it may have serious consequence to not correctly diagnosing a person with mental illness. Different performance metrics may require balancing at varying thresholds, depending on the clinical context and the relative costs of false negatives and false positives [4], [7].

3.2 Studies with Popular Machine Learning Algorithms.

The possibility of diagnosing depression using interview text was confirmed; regarding suicide risk, the diagnosis accuracy increased when demographic variables were incorporated. Therefore, participants' words during an interview show significant potential as an objective and diagnostic marker through machine learning. [5] The Patient Health Questionnaire-9 (PHQ-9) was used for assessing subjective depression in participants. Speech-to-text conversion followed by analysis using the Naive Bayes Classifier & Demographics Ensemble Model was conducted. The model's performance in predicting suicide risk in depressive patients showed a sensitivity of 74.4% and a specificity of 47.7% when using text alone. This was improved to a sensitivity of 81.6%, specificity of 64.7%, and AUC of 0.800 when demographic data were included in the ensemble model.

In another study, the researchers [6] recruited survey participants internationally from various online forums dedicated to mental health, self-injury, and suicide topics to predict short-term suicide ideation and attempts. The participants were all adults, being at least 18 years old and proficient in English, and included in follow-up surveys. The study found that machine learning algorithms outperformed traditional statistical methods. They suggested that an optimized combination of numerous, but not necessarily excessive, suicide predictors could effectively predict suicide ideation and, to a lesser extent, non-fatal attempts. This study has limitations due to its focus solely on individuals with severe suicidal tendencies and self-injuring behaviors. Additionally, the reliance on participants' proficiency in English, which is not their native language, is a critical factor, particularly since the study employs text analysis algorithms in English. This could potentially lead to inaccuracies in the text analysis, thereby affecting the study's generalizability.

Another study aimed to determine whether suicidal ideation can be detected via language features in clinical interviews for depression using natural language processing (NLP) and machine learning (ML). [8] The study utilized ordinal logistic regression and Random Forest machine learning techniques to analyze responses to the Hamilton Depression Rating Scale (HAMD) questions. It focused on differentiating suicidal ideation from depression by examining language features specific to suicidality, independent of the confounding effects of depressive symptoms. This approach aimed to identify unique language markers associated with suicidal thoughts. The study's findings indicate that there are other subtle markers, potentially in the way people use language, that can indicate a risk of suicide. The effectiveness of ML models in this study in discriminating between different levels of suicide risk (high, low, and non-suicidal) indeed suggests that suicide risk can manifest in various groups, not just those who are clearly at high risk or those with explicit suicidal ideation. This finding is significant because it underscores the complexity of suicide risk. It's not a binary state (at risk or not at risk), but rather a spectrum where individuals may exhibit different levels of risk. This spectrum approach allows for a more nuanced understanding and potentially more targeted interventions.

Another interesting study explores the relationship between psychological stress and suicide ideation in military personnel, acknowledging their higher stress and suicide risk compared to the general population. [9] It challenged traditional statistical methods' moderate correlation findings by employing machine learning techniques – logistic regression, decision tree, random forest, gradient boosting regression tree, support vector machine, and multilayer perceptron. These methods are used to predict suicide ideation based on six key psychological stress domains in both male and female military members.

Another study introduced a new network feature for detecting suicidal ideation from clinical texts. [10] This study tackled some of the known issues by experimenting with statistical text features and constructing networks for feature extraction and classification. The research highlights the potential of both logistic classifiers and deep learning methods, emphasizing the need for further experiments and data before clinical application.

On the other hand, another group of researchers [11] analyzed US national survey data in two waves in separating time periods to identify risk factors for suicide attempts. The survey in this study ensured population-level generalizability by incorporating a complex survey design and sampling weights. The strongest risk factors of future suicide attempts were related to previous suicidal behaviors. Other important novel risk factors identified were related to socioeconomic disadvantage. Lower educational level and experiencing a financial crisis in the last year were among the 10 most important variables. This research uncovered a link at the individual level between economic hardship and the risk of suicide attempts. In this study, Balance Random Forest (BRF) performed better than regular random forest plotting for classification models with class-imbalanced data. The study evaluated how model accuracy was impacted by using fewer features. This was done by constructing new Balanced Random Forest (BRF) models using only the top 5 and 10 most significant variables identified from the full feature set. The results showed that reducing the number of features led to a decrease in model accuracy, suggesting that limiting features might not be the best approach. A limitation noted was the study's focus on individuals 18 years or older, not addressing younger individuals' suicide risks.

Table 1. Comparison of the studies.

Paper	Data Sources	Algorithm Used	Results
Detection of Depression and Suicide Risk Based on Text From Clinical Interviews Using Machine Learning: Possibility of a New Objective Diagnostic Marker [5]	Speech to text interview (Patient Health Questionnaire-9)	Naive Bayes Classifier Demographics Ensemble Model	Predicting depression, anxiety, suicidal ideation, and impulsivity: area under the curve (AUC) of 0.905, a sensitivity of 0.699, and a specificity of 0.964 Predicting high-suicide-risk group: the AUC of the ensemble model incorporating demographic variables was 0.800
Predicting Imminent Suicidal Thoughts and Nonfatal Attempts: The Role of Complexity [6]	Recruited adult survey participants online with follow-up surveys	Comparison of traditional statistical methods with machine learning algorithms	Suicide Ideation: AUC 0.87, Sensitivity 0.98 Precision 0.94 (28 days from baseline) Suicide Attempts: AUC 0.83, Sensitivity 0.67 Precision 0.98 (28 days from baseline)
Detection of Suicidal Ideation in Clinical Interviews for Depression Using Natural Language Processing and Machine Learning: Cross-Sectional Study [8]	Clinical interviews for depression (Hamilton Depression Rating Scale)	Natural Language Processing Random Forest	AUC 0.76-0.89; P<.001 (high suicide risk) AUC 0.83-0.92; P<.001(both low and high suicide risk)
Machine Learning Based Suicide Ideation Prediction for Military Personnel [9]	Self-reported questionnaire (Brief Symptom Rating Scale 5) Formal health examination Historical cohort of 3,546 military men and women	Traditional statistical methods Logistic Regression Decision Tree Random Forest Gradient Boosting Regression Tree Support Vector Machine Multilayer Perceptron	ROC-AUC: 100% for both SVM & Multilayer Perceptron
Toward Suicidal Ideation Detection with Lexical Network Features and Machine Learning [10]	Each of the three data collections comprises transcribed interviews conducted by clinical experts	Logistics Regression Multilayer Perceptron Convolutional Neural Network Excess Weight Density	AUC 95%: high suicidal ideation AUC 69%: generalized
Identification of Suicide Attempt Risk Factors in a National US Survey Using Machine Learning [11]	US national survey data in two waves	Missing-indicator method Balanced Random Forest Random Forest	ROC-AUC 0.857 Sensitivity 85.3% Specificity of 73.3%

Table 1 provides a comprehensive overview of various studies focused on the detection and prediction of suicidal tendencies and related mental health issues, primarily using vocal speech and related data sources. Each study employs distinct methodologies, leveraging a range of algorithms from Naive Bayes Classifiers to advanced machine learning techniques like Multilayer Perceptron and Convolutional Neural Networks. The studies vary in their data sources, ranging from speech-to-text interviews and clinical interviews to self-reported questionnaires and national survey data.

3.3 Machine Learning Studies for Adolescents. In a study focusing on adolescents, the researchers [13] conducted a retrospective, longitudinal cohort analysis using data from the Vanderbilt Synthetic Derivative. This study spanned from January 1998 to December 2015 and included a cohort of 974 adolescents who had nonfatal suicide attempts. In all comparisons made, Random Forests demonstrated superior performance over logistic regression, underscoring the effectiveness of using machine learning on longitudinal clinical data. This approach shows promise as a scalable method to enhance the screening for the risk of nonfatal suicide attempts among adolescents.

In another study, Classification Tree Analysis (CTA) was used to analyze a dataset of adolescents to create decision trees that help in identifying those at risk for suicide ideation. [7] The study identified three unique solutions, each with different levels of sensitivity and specificity, for detecting suicide ideation in adolescents. Sensitivity scores of the classification trees varied from 44.6% to 77.6%. The tree that prioritized specificity over sensitivity focused on past suicide ideation history. Another tree, offering moderate sensitivity and high specificity, considered depressive symptoms, suicide attempts or history in family and friends, and social support. The most sensitive, but least specific, tree incorporated these factors along with gender, ethnicity, hours of sleep, school-related factors, and future outlook. However, the study encountered challenges with notably low sensitivity or specificity.

In another research, machine learning algorithms were employed to enhance the prediction of suicide attempt risk, as demonstrated in the Early Developmental Stages of Psychopathology (EDSP) study. [14] This longitudinal research spanned 10 years, focusing on adolescents and young adults aged 14 – 24. The study involved 2797 participants who underwent at least one of three follow-up assessments. Using sixteen baseline predictors identified from prior literature, the study evaluated the risk of follow-up SAs. Four ML algorithms – logistic regression, lasso, ridge, and random forest – were compared using repeated nested 10-fold cross-validation. Model performance, measured by the area under the curve (AUC), showed mean AUCs of 0.828, 0.826, 0.829, and 0.824, respectively, for each model.

In another study [12], three primary findings emerged regarding the prediction of suicide attempts using machine learning models. Firstly, indicators of previous self-harm were identified as the most significant predictors across all three models, highlighting their critical role. Secondly, the predictors were categorized, revealing a hierarchy of importance: mental health, emotion and motivation, drug use, sexuality, demography, victimization, physical health, personality, attitudes, and behavior. Notably, within the emotion and motivation category, feelings of being unloved and having low self-esteem were among the top predictors. The majority of these significant predictors, except for two in the first model, were based on adolescents' self-reports. Lastly, the study found an evolution in the importance of these variables over time. Different categories of predictors gained or lost significance depending on the developmental stage, with mental health and emotion/motivation variables becoming more pivotal in later models, reflecting the dynamic nature of these factors in assessing suicide risk.

Table 2. Comparison of the studies with adolescents.

Paper	Data Sources	Algorithm Used	Results
Predicting suicide attempts in adolescents with longitudinal clinical data and machine learning [13]	Longitudinal cohort analysis using data from the Vanderbilt Synthetic Derivative Spanned from January 1998 to December 2015 and included a cohort of 974 adolescents who had nonfatal suicide attempts Clinical records, providing insights into this demographic	Logistics Regression Random Forest	Random forests significantly outperformed logistic regression in every comparison. Best results: AUC 0.94 [0.92-0.96] at 720 days; 0.97 [0.95-0.98] at 7 days
Prospective identification of adolescent suicide ideation using classification tree analysis: Models for community-based screening [7]	4,799 youth completed both Waves 1 and 2 of the National Longitudinal Study of Adolescent to Adult Health	Classification Tree Analysis	Findings revealed 3 distinct solutions with varying sensitivity and specificity for identifying adolescents who reported suicide ideation. Sensitivity of the classification trees ranged from 44.6% to 77.6%.
Prospective prediction of suicide attempts in community adolescents and young adults, using regression methods and machine learning [14]	2797 participants who underwent at least one of three follow-up assessments in the Early Developmental Stages of Psychopathology (EDSP) study	Logistic Regression Lasso Ridge Random Forest	The mean AUCs of the four predictive models, logistic regression, lasso, ridge, and random forest, were 0.828, 0.826, 0.829, and 0.824, respectively
Predicting Lifetime Suicide Attempts in a Community Sample of Adolescents Using Machine Learning Algorithms [12]	7-year-olds in the Millennium Cohort Study combining a large set of self- and other-reported variables from different categories	Logistic Regressions Elastic Net Regressions Gradient Boosting Machines	Elastic net regressions and GBM models achieved similar averaged balanced accuracies (ABAs) (.76 and .76 for Model 1, .83 and .82 for Model 2, and .84 and .85 for Model 3, respectively)

Table 2 presents a succinct comparison of various studies focusing on the prediction of suicide risk and ideation among adolescents. These studies utilize a range of data sources, including longitudinal cohort analyses, national surveys, and large-scale cohort studies, encompassing diverse adolescent populations. The algorithms employed range from Logistic Regression and Random Forests to more complex methods like Elastic Net Regressions and Gradient Boosting Machines.

In essence, Table 2 underscores the complexity and necessity of employing robust and nuanced analytical methods to accurately predict suicide risk among adolescents, a task that is both challenging and crucial for early intervention and prevention efforts.

3.4 Usages of NLP using Social Media Data. One study analyzed suicidal tweets from Twitter by using several sets of word embedding and tweet features, and twelve classifiers' models. [15] The study's model aims to establish a standard for predicting suicidal ideation using active social media platforms like Twitter. Many more studies have been conducted on Twitter data. [16] However, the challenge lies in verifying the accuracy and effectiveness of this model due to limitations in publicly available datasets from social media sites, which raises questions about the reliability of such data for predictive analysis. Another study addressed this concern. [17] This research developed a framework for annotating a mental-health-related textual dataset from Reddit, focusing on identifying posts and comments with suicide attempts and ideations. The study used an active machine learning method, starting with a small dataset and expanding it by incorporating expert judgments on challenging samples and automated annotations for more straightforward cases. However, the dataset used is quite small, posing a challenge for developing an effective machine learning model. Additionally, while using expert opinions for the most ambiguous samples can be helpful, it may also bring in subjective biases. Thus, the degree of reliance on expert input and its influence on the model's performance warrants careful evaluation.

Numerous studies have employed AI to classify text related to suicidal ideation as either positive or negative. Yet, there's a lack of clarity on how the texts specifically influence ML/DL model outcomes. A particular study is initiating efforts to unravel how the words and sentences in the collected texts affect the results of these classification models. [18] In this study, classical ML algorithms like C-Support Vector Classification (SVC), Extra Trees, and Random Forest were trained and explained. The ELI5 method was used for both local and global interpretability of these algorithms, analyzing features with respective weight values. It helps in understanding the decisions made by complex models by breaking them down into simpler, more understandable terms. Notably, Extra Trees and Random Forest classifiers showed similar influence from single and dual-term features, while SVC was more influenced by dual-term features, with terms like "suicide" and "sadness" being significant for indicating positive suicidal ideation.

Table 3. Comparison of the studies using data from social media.

Paper	Data Sources	Algorithm Used	Results
Analysis of Suicidal Tweets from Twitter Using Ensemble Machine Learning Methods [15]	Suicidal tweets from Twitter	NLP Ensemble classifier (twelve)	Best results from 12 classifiers: Voting Classifier (Accuracy: 0.904, Precision: 0.913, Recall: 0.895, F1 Score: 0.90)
Using Machine Learning Algorithms to Detect Suicide Risk Factors on Twitter [16]	12,066 public tweets from 3,873 Twitter users	Latent Semantic Analysis, Latent Dirichlet Allocation, Non-negative Matrix Factorization, Decision Tree and K-means Clustering	Decision tree classification model achieved 0.844 in precision, 0.912 in sensitivity, and 0.829 in specificity in classifying users into "HighRisk" and "AtRisk" groups.
Enriching an online suicidal dataset with active machine learning [17]	Text data from "SuicideWatch" Reddit channel 1000 manually annotated samples	K Nearest Neighbors (KNN), Decision Tree (DT), Random Forest (RF), Logistic Regression (LR), Stochastic Gradient Descent (SGD) Classifier, Naive Bayes (NB), and Support Vector Machine (SVM with the linear kernel)	Suicide Ideation: Accuracy 0.86, Sensitivity 0.93, Specificity 0.77 Suicide Attempt: Accuracy 0.87, Sensitivity 0.68, Specificity 0.91
How can machine learning identify suicidal ideation from user's texts? Towards the explanation of the Boamente system [18]	5,699 tweets from Twitter users (non-clinical text)	C-Support Vector Classification, Extra Trees, and Random Forest	The features "suicide", "desire to kill oneself" and "sadness" had a higher importance value.

Table 3 briefly compares studies that utilize social media data for predicting suicide risk and ideation. These studies analyze data from platforms like Twitter and Reddit, employing various algorithms including NLP, Decision Tree, Random Forest, and Support Vector Machine. Key findings include establishing a standard for social media text analysis, classifying Twitter users into risk groups, learning from small datasets with expert input, and identifying important features like 'suicide' and 'sadness' in tweets. These studies collectively demonstrate the potential of social media data in identifying and analyzing suicidal tendencies.

3.5 Remaining Challenges & Limitations. While machine learning and deep learning methods have shown promise in detecting mental illnesses using NLP, a study has highlighted several remaining challenges in this field. [19] It suggests that future research should focus on addressing these key issues to enhance the effectiveness of NLP in mental health diagnostics. The study highlights key challenges in using machine learning for mental illness detection: 1) Data quality issues, with the need for extensive, unbiased, and clinically-validated datasets; 2) Model performance concerns, such as instability due to data source variability and class imbalance; 3) A focus on enhancing model interpretability for clinical utility; and 4) Ethical considerations in handling sensitive mental health data. It emphasizes exploring semi-supervised and unsupervised learning, leveraging multi-modal data, and ensuring ethical compliance in research.

3.6 Speech Analysis for Depression & Suicide Risk Detection. Major depression is the key indicator of suicide risk. [12] A study focused on exploring the potential of using voice as a biomarker to distinguish between minor and major depression. [20] The research involved analyzing voice features extracted from semi-structured interviews. In total, 21 voice features were evaluated, and a three-group comparison (not depressed, minor depression, and major depression) was conducted using analysis of variance. Seven voice indicators were found to significantly differ among the groups, even after adjusting for factors like age, body mass index, and non-psychiatric medication use. In the study, various machine learning techniques were evaluated, with the multi-layer processing approach emerging as the most effective. This method demonstrated a significant performance with an Area Under the Curve (AUC) of 65.9%, along with a sensitivity of 65.6% and a specificity of 66.2%. The research successfully highlighted distinct voice patterns in depressive episodes, validating the capability of machine learning in accurately differentiating between individuals who are not depressed and those with minor or major depression.

In the study, voice features were extracted and analyzed from four distinct aspects: glottal, tempo-spectral, formant, and other physical attributes. [20] The glottal features, which provide information about how sound is produced by the vocal cords, were obtained by analyzing and parameterizing the waveform of each utterance. Key parameters like the Glottal Closure Instance (GCI) were calculated and refined through iterative adaptive inverse filtering, leading to the extraction of three specific parameters: the opening phase, closing phase, and closed phase. Tempo-spectral features, typically used in music information retrieval, were extracted using the "Librosa" audio processing toolkit. These encompassed temporal elements (such as the length of utterances) and tempo (periodicity of onset), as well as spectral features like spectral centroid, bandwidth, roll-off frequency, and root mean square energy. Formant features, crucial in phonetics, were derived using linear prediction coefficients (LPCs). These features represent the resonance of the vocal tract, focusing on the first to third formants and their bandwidths. Finally, other physical attributes like mean and variance of pitch and magnitude, Zero-Crossing Rate (ZCR), and voice portions were analyzed. ZCR, in particular, gauged the intensity of voice utterance and frequency of occurrence, helping to distinguish between voiced and silent frames in an utterance. This comprehensive approach to voice feature extraction allowed for a nuanced analysis of vocal characteristics in the context of detecting depression.

The study's analysis of the 21 extracted voice features identified eight features that significantly differed across the groups. Among these, seven features showed statistical significance after adjusting for age, BMI, and medication use. These were the spectral centroid, spectral roll-off, formant bandwidth 2 (BW2), squared mean pitch, standard deviation of pitch, Zero-Crossing Rate (ZCR), and voice portion. Notably, in comparing the normal (ND) and minor depression (mDE) groups, features like the spectral centroid, spectral roll-off, squared mean pitch, standard deviation of pitch, mean magnitude, ZCR, and voice portion exhibited differences. However, between the mDE and major depression (MDE) groups, only the standard deviation of pitch was significantly different.

Interestingly, the voice features did not demonstrate a consistent trend correlating with the increasing severity of depressive episodes. The Jonckheere–Terpstra test further confirmed a sequence for each group, where all seven significant variables either increased or decreased in the order of ND, MDE, and mDE, illustrating the nuanced role these features play in distinguishing between different levels of depression severity.

In the study, the best performing algorithm for predicting the severity of depressive episodes was the Multi-Layer Perceptron (MLP). This model used all 21 voice characteristics for constructing the prediction model. Comparatively, other algorithms such as Logistic Regression (LR) and Gaussian Naive Bayes (GNB) showed lower performance, with AUCs ranging from 58.8 to 64.7 and sensitivity and specificity between 41.6 and 57.2. These models did not exhibit any improvement in performance with increased training data, unlike the MLP. The overall accuracy of the MLP in predicting the severity of depressive episodes was 60.0% with an 8:2 train-test split in 93 cases, indicating a promising but not definitive predictive capability.

Another study introduced DEPAC, a new audio dataset focused on mental distress analysis. [21] This dataset is uniquely labeled using established thresholds derived from standard screening tools for depression and anxiety. Alongside the dataset, they presented a set of hand-curated features, encompassing both acoustic and linguistic aspects, which have proven effective in detecting indicators of mental illnesses in human speech.

DEPAC is distinctively labeled with scores from two standard scales: the Patient Health Questionnaire-9 (PHQ-9) for depression and the Generalized Anxiety Disorder-7 (GAD-7) for anxiety assessment. This dataset boasts a larger sample size compared to other available public corpora, offering a rich diversity in speech tasks and participants, including varied education levels, genders, and age groups. A unique contribution of this work is the introduction of a hand-curated set of acoustic and linguistic features, developed with insights from both clinical and machine learning (ML) experts. These features serve as predictors in models quantifying depression severity. The study also presents baseline model performances for predicting depression severity levels, providing a benchmark for future research. When compared to baseline models from the AVEC 2016 and AVEC 2019, the results from DEPAC demonstrate competitive performance, underscoring the quality of the dataset and the efficacy of the proposed feature set in assessing depression severity.

In another research, [22] investigated the efficiency of machine learning models in detecting depression from speech. They compared models trained on conventional hand-curated acoustic features with those based on deep representation features. Their findings revealed that the conventional feature-based models are not only equally effective or even superior in predicting depression severity, but also significantly more cost-effective in terms of computational resources. These models performed consistently well regardless of variables like speech content and length, speaker's gender, and severity of depression. The study highlighted the influence of speaker's gender and score distribution on model performance, emphasizing the need for balanced training data. They noted that speech content and length had minimal impact, especially for short samples under one minute. Conclusively, the study recommends using conventional feature-based models in real-time, resource-limited applications, while reserving deep models for scenarios requiring more detailed analysis and greater computational power.

Previously, in the systematic review, [23] explored the use of speech as an automated biomarker for a range of psychiatric disorders, with a focus on depression and schizophrenia. They highlight the potential of machine learning technology in utilizing speech samples, either obtained clinically or remotely, as a tool for improving diagnosis and treatment of mental health issues. This review is the first of its kind to assess speech for automated assessments across various psychiatric disorders, building on earlier work that mainly concentrated on depression and schizophrenia. The authors note that speech patterns, such as pitch, monotony, intensity, and rate, along with hesitations and stuttering, have been long recognized as indicators of mental disorders. These acoustic features, including variations in fundamental frequency (f0), jitter, shimmer, and f0 variability, correlate with the severity of disorders like major depressive disorder. The authors emphasize the advantages of speech as a biomarker: it's difficult to conceal symptoms through speech, it directly expresses emotions and thoughts, and it's relatively easy and cost-effective to obtain using modern digital devices. This approach holds promise for broad application, especially in low-resource languages and settings where advanced natural language processing technologies are unavailable.

In another study, the researchers developed an MFCC (Mel Frequency Cepstral Coefficients)-based Recurrent Neural Network (RNN), specifically utilizing Long Short-Term Memory (LSTM) layers, to detect depression and assess its severity levels from speech. The approach involved preprocessing audio recordings, extracting, and normalizing MFCC features, which were then input into the deep RNN. To address the challenges of limited training data and potential overfitting, the study employed strategies like augmenting training data and transferring knowledge from related tasks. Evaluated on the DAIC-WOZ database, this model achieved a notable accuracy of 76.27% and a low root mean square error in depression assessment. Additionally, the study explored the impact of incorporating other modalities, finding that adding visual features to the audio-based model significantly enhanced its performance, achieving an accuracy of 95.6% and improved F1 scores for depression detection. This research highlights the potential of combining audio and visual data in creating more effective tools for mental health assessment [24].

Another research introduced DEPA, a novel self-supervised, pre-trained depression audio embedding technique, designed to enhance depression detection. Utilizing an encoder-decoder network, DEPA is extracted from both in-domain datasets focused on depression (like DAIC and Major Depression Disorder datasets) and out-domain datasets (such as Switchboard, Alzheimer's). This method significantly improved performance in downstream tasks, especially in sparse datasets. DEPA, extracted at the response level, demonstrated superior performance compared to traditional features like LMS, STFT, HCVP, and x-vector. For instance, when pretrained on the in-domain DAIC dataset, DEPA yielded notably higher F1 scores (DAIC F1 0.90, MDD F1 0.94) compared to other methods. The study also showed the benefits of pretraining on larger datasets, such as Alzheimer's Disease (AD), indicating that using additional out-domain data can be advantageous for depression detection. The results, validated on the MDD dataset, underscore DEPA's efficacy in summarizing long sequences, suggesting its potential as a generalized method for medical applications in audio classification [25].

Table 4. Comparison of the studies using speech analysis.

Paper	Data Sources	Algorithm Used	Results
1	2	3	4
Detection of Minor and Major Depression through Voice as a Biomarker Using Machine Learning [20]	Recruited from the patient population who visited the outpatient clinic of Seoul National University Hospital for depressive symptoms. Participants' ages ranged from 19 to 65 years, classified into 3 labeled groups	Different statistical analysis, Logistic Regression, Gaussian Naive Bayes, Support Vector Machine, and Multilayer Perceptron	Best result: Multilayer Perceptron AUC of 65.9%, sensitivity of 65.6%, and specificity of 66.2%
DEPAC: a Corpus for Depression and Anxiety Detection from Speech [21]	The participants located in USA and Canada recruited via crowdsourcing and consists of a variety of self-administered speech tasks	N/A	A novel mental distress analysis audio dataset DEPAC is introduced.
Cost-effective Models for Detecting Depression from Speech [22]	DEPAC dataset	Support Vector Machine Random Forest Feedforward Neural Network	SVM and FNN models performed better on conventional features than on VGG-16, while performance of RF is marginally better (0.0004%) on VGG-16

Table 4. Continuation.

1	2	3	4
Automated assessment of psychiatric disorders using speech: A systematic review [23]	Studies from the last 10 years using speech		Variations in fundamental frequency (f0), jitter, shimmer, and f0 variability, correlate with the severity of disorders like major depressive disorder.
MFCC-based Recurrent Neural Network for automatic depression recognition and assessment from speech [24]	Self-report depression test of the Patient Health Questionnaire of eight questions (the PHQ-8) DAIC-WOZ corpus RAVDESS dataset AVi-D dataset	Gaussian Mixture Models (GMM), Support Vector Machines (SVM) with raw data, Support Vector Machines with GMM, Multilayer Perceptron neural networks (MLP), and Hierarchical Fuzzy Signature (HFS) Deep Convolutional Neural Network, Deep Convolutional Neural Network followed by a Deep Network (DCNNDNN) and Long Short Term Memory (LSTM) MFCC-based RNN	MFCC-based RNN (LSTM): RMSE = 0.4 Accuracy = 76.27% F1=85% (Non Depression) F1=46% (Depression)
DEPA: Self-Supervised Audio Embedding for Depression Detection [25]	In-domain depressed datasets (DAIC and MDD) and out-domain (Switchboard, Alzheimer's) datasets	Bidirectional Long Short Term Memory (BLSTM)	DEPA pretrained on In-domain DAIC suggests a significantly better result on detection presence detection using STFT features (DAIC F1 0.90, MDD F1 0.94) compared to LMS features (DAIC F1 0.68, MDD F1 0.71) as well as other approaches without DEPA. Pretraining on large datasets, e.g. DEPA on AD reached F1 0.94 & MAE 4.75

Table 4 shows a concise overview of studies using speech analysis for predicting depression, a relevant factor in suicidal risk assessment. These studies employ a range of algorithms like Logistic Regression, Support Vector Machine, and neural networks to analyze speech features. Key findings include the identification of voice indicators correlating with depression severity, introduction of novel datasets like DEPAC for mental distress analysis, and the importance of considering factors like gender and balanced training data for effective model performance. The studies demonstrate the potential of speech analysis in mental health assessment, with implications for detecting suicidal risks.

3.7 Multimodal Data Analysis. A group of researchers explored the emerging field of using audiovisual features for detecting suicide ideation and behavior. They highlighted that while much research has focused on depression, specific non-verbal cues related to suicidal tendencies, such as voice/speech acoustics and visual indicators, are distinct and warrant separate investigation. The study emphasized the importance of recording settings, ranging from controlled environments to more naturalistic, free settings using smartphones, each posing unique challenges in data quality and feature extraction. Notably, features like Mel Frequency Cepstral Coefficients (MFCC) and fundamental frequency (F0) in speech, as well as dynamic facial expressions, had been strongly correlated with mental states and show promise in suicide assessment. The review also discussed the utility of collaborative tools like COVAREP and GeMAPS for acoustic feature processing, and OpenFace for visual feature analysis. Despite the potential of these methods, the field faces challenges such as the lack of large-scale datasets for training machine learning and deep learning models. The authors concluded that automatic suicide assessment using audiovisual cues is a promising yet nascent area of research, highlighting the need for more extensive data. [26]

4. Discussion. By synthesizing findings from recent studies, this paper aims to guide future research and enhance the development of effective and reliable machine learning models for early suicidal tendency detection. It has highlighted the significant advancements and ongoing challenges in applying machine learning techniques for detecting suicidal tendencies and the diversity of data sources and techniques employed. The findings reveal the importance of data nuances and feature selection in improving model accuracy and reliability.

However, the variability in data sources and the heterogeneity of features and outcomes across studies further complicate the generalizability of findings. There is a clear need for standardized protocols and larger, more diverse datasets to enhance the robustness of these machine learning models.

Future research should focus on addressing these challenges by improving data collection and standardization methods. Additionally, integrating multimodal data will be crucial for their practical application in clinical settings.

5. Conclusion. While significant progress has been made in using machine learning for the detection of suicidal tendencies, ongoing efforts are required to refine these techniques to more accurate early detection and intervention strategies. By addressing the challenges of data variability and feature heterogeneity, and by integrating multimodal data, future research can contribute to more robust and generalizable models, ultimately improving outcomes for individuals at risk of suicide.

6. Funding. This research has been/was/is funded by the Science Committee of the Ministry of Science and Higher Education of the Republic of Kazakhstan (Grant No. AP22786670).

REFERENCES

1. Boudreaux, E. D. (2021). Applying machine learning approaches to suicide prediction using healthcare data: Overview and future directions. *Frontiers in Psychiatry*. <https://doi.org/10.3389/fpsy.2021.707916>.
2. Su, C. (2020). Machine learning for suicide risk prediction in children and adolescents with electronic health records. *Translational Psychiatry*. <https://doi.org/10.1038/s41398-020-01100-0>.
3. Ribeiro, J. D., et al. (2016). Self-injurious thoughts and behaviors as risk factors for future suicide ideation, attempts, and death: A meta-analysis of longitudinal studies. *Psychological Medicine*, 46(2), 225-236. <https://doi.org/10.1017/S0033291715001804>.
4. Chen, Q., et al. (2020). Predicting suicide attempt or suicide death following a visit to psychiatric specialty care: A machine learning study using Swedish national registry data. *PLoS Medicine*, 17(11), e1003416. <https://doi.org/10.1371/journal.pmed.1003416>.
5. Shin, D., et al. (2022). Detection of depression and suicide risk based on text from clinical interviews using machine learning: Possibility of a new objective diagnostic marker. *Frontiers in Psychiatry*, 13. <https://doi.org/10.3389/fpsy.2022.801301>.
6. Ribeiro, J. D., Huang, X., Fox, K. R., Walsh, C. G., & Linthicum, K. P. (2019). Predicting imminent suicidal thoughts and nonfatal attempts: The role of complexity. *Clinical Psychological Science*, 7(5), 941-957. <https://doi.org/10.1177/2167702619838464>.

7. Hill, R. M., Oosterhoff, B., & Kaplow, J. B. (2017). Prospective identification of adolescent suicide ideation using classification tree analysis: Models for community-based screening. *Journal of Consulting and Clinical Psychology*, 85(7), 702-711. <https://doi.org/10.1037/ccp0000218>.
8. Li, T. M. H., et al. (2023). Detection of suicidal ideation in clinical interviews for depression using natural language processing and machine learning: Cross-sectional study. *JMIR Medical Informatics*, 11, e50221. <https://doi.org/10.2196/50221>.
9. Lin, G.-M., Nagamine, M., Yang, S.-N., Tai, Y.-M., Lin, C., & Sato, H. (2020). Machine learning-based suicide ideation prediction for military personnel. *IEEE Journal of Biomedical and Health Informatics*, 24(7), 1907-1916. <https://doi.org/10.1109/JBHI.2020.2988393>.
10. Bayram, U., et al. (2022). Toward suicidal ideation detection with lexical network features and machine learning. *Northeast Journal of Complex Systems*, 4(1). <https://doi.org/10.22191/nejcs/vol4/iss1/2>.
11. García de la Garza, Á., Blanco, C., Olfson, M., & Wall, M. M. (2021). Identification of suicide attempt risk factors in a national US survey using machine learning. *JAMA Psychiatry*, 78(4), 398. <https://doi.org/10.1001/jamapsychiatry.2020.4165>.
12. Jankowsky, K., Steger, D., & Schroeders, U. (2023). Predicting lifetime suicide attempts in a community sample of adolescents using machine learning algorithms. *Assessment*. <https://doi.org/10.1177/10731911231167490>.
13. Walsh, C. G., Ribeiro, J. D., & Franklin, J. C. (2018). Predicting suicide attempts in adolescents with longitudinal clinical data and machine learning. *Journal of Child Psychology and Psychiatry*, 59(12), 1261-1270. <https://doi.org/10.1111/jcpp.12916>.
14. Miché, M., et al. (2020). Prospective prediction of suicide attempts in community adolescents and young adults, using regression methods and machine learning. *Journal of Affective Disorders*, 265, 570-578. <https://doi.org/10.1016/j.jad.2019.11.093>.
15. Sakib, T. H., Ishak, M., Jhumu, F. F., & Ali, M. A. (2021). Analysis of suicidal tweets from Twitter using ensemble machine learning methods. In 2021 International Conference on Automation, Control and Mechatronics for Industry 4.0 (ACMI), IEEE, July (pp. 1-7). <https://doi.org/10.1109/ACMI53878.2021.9528252>.
16. Sakib, T. H., Ishak, M., Jhumu, F. F., & Ali, M. A. (2021). Analysis of suicidal tweets from Twitter using ensemble machine learning methods. In 2021 International Conference on Automation, Control and Mechatronics for Industry 4.0 (ACMI), IEEE, July (pp. 1-7). <https://doi.org/10.1109/ACMI53878.2021.9528252>.
17. Fodeh, S., et al. (2019). Using machine learning algorithms to detect suicide risk factors on Twitter. In 2019 International Conference on Data Mining Workshops (ICDMW), IEEE, November (pp. 941-948). <https://doi.org/10.1109/ICDMW.2019.00137>.
18. Liu, T., Zheng, Z., Zhou, Y., Yang, Y., & Song, Y. (2022). Enriching an online suicidal dataset with active machine learning. In Proceedings of the ACM Southeast Conference, New York, NY, USA: ACM, April (pp. 196-200). <https://doi.org/10.1145/3476883.3520213>.
19. de Oliveira, A. C., Diniz, E. J. S., Teixeira, S., & Teles, A. S. (2022). How can machine learning identify suicidal ideation from user's texts? Towards the explanation of the Boamente system. *Procedia Computer Science*, 206, 141-150. <https://doi.org/10.1016/j.procs.2022.09.093>.
20. Zhang, T., Schoene, A. M., Ji, S., & Ananiadou, S. (2022). Natural language processing applied to mental illness detection: A narrative review. *NPJ Digital Medicine*, 5(1), 46. <https://doi.org/10.1038/s41746-022-00589-7>.
21. Shin, D., et al. (2021). Detection of minor and major depression through voice as a biomarker using machine learning. *Journal of Clinical Medicine*, 10(14), 3046. <https://doi.org/10.3390/jcm10143046>.
22. Tasnim, M., Ehghaghi, M., Diep, B., & Novikova, J. (2022). DEPAC: A corpus for depression and anxiety detection from speech. In Proceedings of the Eighth Workshop on Computational Linguistics and Clinical Psychology, Stroudsburg, PA, USA: Association for Computational Linguistics (pp. 1-16). <https://doi.org/10.18653/v1/2022.clpsych-1.1>.
23. Tasnim, M., & Novikova, J. (2022). Cost-effective models for detecting depression from speech. In 2022 21st IEEE International Conference on Machine Learning and Applications (ICMLA), IEEE, December (pp. 1687-1694). <https://doi.org/10.1109/ICMLA55696.2022.00259>.
24. Low, D. M., Bentley, K. H., & Ghosh, S. S. (2020). Automated assessment of psychiatric disorders using speech: A systematic review. *Laryngoscope Investigative Otolaryngology*, 5(1), 96-116. <https://doi.org/10.1002/lio2.354>.
25. Rejaibi, E., Komaty, A., Meriaudeau, F., Agrebi, S., & Othmani, A. (2022). MFCC-based recurrent neural network for automatic clinical depression recognition and assessment from speech. *Biomedical Signal Processing and Control*, 71, 103107. <https://doi.org/10.1016/j.bspc.2021.103107>.

26. Zhang, P., Wu, M., Dinkel, H., & Yu, K. (2021). DEPA: Self-supervised audio embedding for depression detection. In Proceedings of the 29th ACM International Conference on Multimedia, New York, NY, USA: ACM, October (pp. 135-143). <https://doi.org/10.1145/3474085.3479236>.
27. Dhelim, S., Chen, L., Ning, H., & Nugent, C. (2023). Artificial intelligence for suicide assessment using audiovisual cues: A review. *Artificial Intelligence Review*, 56(6), 5591-5618. <https://doi.org/10.1007/s10462-022-10290-6>.