

Artificial intelligence-based suicide prevention and prediction: A systematic review (2019–2023)

Original

Artificial intelligence-based suicide prevention and prediction: A systematic review (2019–2023) / Atmakuru, A., Shahini, A., Chakraborty, S., Seoni, S., Salvi, M., Hafeez-Baig, A., Rashid, S., Tan, R.S., Barua, P.D., Molinari, F., Acharya, U.R..
- In: INFORMATION FUSION. - ISSN 1566-2535. - STAMPA. - 114:(2025). [10.1016/j.inffus.2024.102673]

Availability:

This version is available at: 11583/2992563 since: 2024-09-18T07:26:30Z

Publisher:

Elsevier

Published

DOI:10.1016/j.inffus.2024.102673

Terms of use:

This article is made available under terms and conditions as specified in the corresponding bibliographic description in the repository

Publisher copyright

Elsevier postprint/Author's Accepted Manuscript

© 2025. This manuscript version is made available under the CC-BY-NC-ND 4.0 license
<http://creativecommons.org/licenses/by-nc-nd/4.0/>. The final authenticated version is available online at:
<http://dx.doi.org/10.1016/j.inffus.2024.102673>

(Article begins on next page)

Artificial Intelligence-based Suicide Prevention and Prediction: A Systematic Review (2019-2023)

Anirudh Atmakuru

Manning College of Information and Computer Sciences

University of Massachusetts, Amherst

Massachusetts, USA

Email: aatmakuru@umass.edu

ORCID: 0000-0003-1448-7334

Alen Shahini

Biolab, PolitoBIOMedLab, Department of Electronics and Telecommunications

Politecnico di Torino, Corso Duca degli Abruzzi 24, 10129 Turin, Italy

Email: alen.shahini@polito.it

ORCID: 0009-0000-4905-5668

Subrata Chakraborty

Faculty of Science, Agriculture, Business and law, University of New England, Armidale, NSW, 2531

Australia Centre for Advanced Modelling and Geospatial Information Systems (CAMGIS)

Faculty of Engineering and IT, University of Technology Sydney, Sydney, NSW, 2007, Australia

Email: Subrata.chakraborty@une.edu.au

ORCID: 0000-0002-0102-5424

Silvia Seoni

Biolab, Department of Electronics and Telecommunications

Politecnico di Torino, Corso Duca degli Abruzzi 24, 10129 Turin, Italy

Email: silvia.seoni@polito.it

ORCID: 0000-0001-5372-2654

Massimo Salvi

Department of Electronics and Telecommunications

Politecnico di Torino, Corso Duca degli Abruzzi 24, 10129 Turin, Italy

Email: massimo.salvi@polito.it

ORCID: 0000-0001-7225-7401

Abdul Hafeez-Baig

School of Management and Enterprise

University of Southern Queensland

Toowoomba, Queensland, Australia
Email: Abdul.Hafeez-Baig@unisq.edu.au
ORCID: 0000-0003-3848-8008

Sadaf Rashid

Sunshine Coast Hospital and Health Service
Queensland, Australia
Email: Sadaf.Rashid@health.qld.gov.au

Ru San Tan

Department of Cardiology, National Heart Centre Singapore, Singapore
Duke-NUS Medical School, Singapore
Email: tanrsnhc@gmail.com
ORCID: 0000-0003-2086-6517

Prabal Datta Barua

School of Business (Information System), University of Southern Queensland, Australia
Email: Prabal.Barua@unisq.edu.au
ORCID: 0000-0001-5117-8333

Filippo Molinari

Department of Electronics and Telecommunications
Politecnico di Torino, Corso Duca degli Abruzzi 24, 10129 Turin, Italy
Email: filippo.molinari@polito.it
ORCID: 0000-0003-1150-2244

U Rajendra Acharya

School of Mathematics, Physics and Computing
University of Southern Queensland
Springfield, Australia
Email: rajendra.acharya@usq.edu.au
ORCID: 0000-0003-2689-8552

CORRESPONDENCE TO: Subrata Chakraborty

Abstract

Suicide is a major global public health concern, and the application of artificial intelligence (AI) methods, such as natural language processing (NLP), machine learning (ML), and deep learning (DL), has shown promise in advancing suicide prediction and prevention efforts. Recent advancements in AI – particularly NLP and DL have opened up new avenues of research in suicide prediction and prevention. While several papers have reviewed specific detection techniques like NLP or DL, there has been no recent study that acts as a one-stop-shop, providing a comprehensive overview of all AI-based studies in this field. In this work, we conduct a systematic literature review to identify relevant studies published between 2019 and 2023, resulting in the inclusion of 156 studies. We provide a comprehensive overview of the current state of research conducted on AI-driven suicide prevention and prediction, focusing on different data types and AI techniques employed. We discuss the benefits and challenges of these approaches and propose future research directions to improve the practical application of AI in suicide research. AI is highly capable of improving the accuracy and efficiency of risk assessment, enabling personalized interventions, and enhancing our understanding of risk and protective factors. Multidisciplinary approaches combining diverse data sources and AI methods can help identify individuals at risk by analyzing social media content, patient histories, and data from mobile devices, enabling timely intervention. However, challenges related to data privacy, algorithmic bias, model interpretability, and real-world implementation must be addressed to realize the full potential of these technologies. Future research should focus on integrating prediction and prevention strategies, harnessing multimodal data, and expanding the scope to include diverse populations. Collaboration across disciplines and stakeholders is essential to ensure that AI-driven suicide prevention and prediction efforts are ethical, culturally sensitive, and person-centered.

Keywords: Artificial intelligence; natural language processing; machine learning; mental health; suicide prevention; suicide prediction.

1. Introduction

Suicide is one of the most important causes of unnatural deaths and mortality across the globe [1]. Over 800,000 people die by suicide every year, with many more attempting suicide or experiencing suicidal thoughts [2]. The devastating impact of suicide on individuals, families, and societies renders it a significant threat to public health. Hence, there is an urgent need for effective suicide prediction and prevention strategies [3, 4]. Suicide prevention and prediction are two interconnected approaches that leverage screening, assessment, and data analytics to reduce suicide risk [5]. Suicide prevention refers to evidence-based population-level strategies designed to minimize risk factors and enhance protective factors across groups. Common prevention initiatives include public awareness campaigns, gatekeeper training, means restriction, and improving access to mental health services [6]. The goal is to implement broad

evidence-based programs to make suicidal behaviors less likely [7]. On the other hand, suicide prediction focuses on estimating individual risk to identify high-risk persons for targeted intervention. Prediction techniques analyze risk factors, warning signs, and data patterns to generate personalized risk probabilities [8].

Multilevel prevention initiatives and individual prediction techniques are complementary. While prevention programs work to minimize factors contributing to risk across populations, prediction models strive to classify suicide risk accurately [9]. Current screening programs often combine aspects of both prevention and prediction. Emergency department protocols allow prompt identification and intervention for those presenting with suicidal thoughts or behaviors (STBs) [10]. Screening through electronic health records (EHR) also helps early detection and monitoring over time. It is important that these systems are fair and unbiased, and sensitive to differences in mental healthcare needs [11]. Suicide prediction also relies on clinical evaluations to understand mental health status, trauma impacts, substance use disorders, and experiences like bullying that elevate risk. Incorporating biopsychosocial assessments and emerging data analytics enhances suicide prediction and the ability to provide timely, patient-centered prevention [12].

Artificial intelligence (AI) techniques, including machine learning (ML) and deep learning (DL), have shown great potential to enhance suicide prediction and prevention [2, 13]. ML is a subset of AI that involves algorithms and statistical models that can automatically learn complex patterns from large datasets. Instead of relying only on human-written rules, ML algorithms self-optimize by analyzing large amounts of data over time [14]. In simple terms, DL can be considered a specialized ML method based on artificial neural networks (ANNs), which can learn from multiple layers of representations [15]. Another key AI technique commonly used in suicide research is natural language processing (NLP). NLP enables computers to process and analyze human language data automatically [16], and allows the mining of unstructured textual data like clinical notes, online posts and speech transcripts to extract insights related to suicide risk [17]. For instance, NLP can identify emotional distress cues, interpersonal concerns, hopelessness and other warning signs in a patient's clinical history or social media content [18]. These give the possibility to study the context history, to analyze suicidal thoughts, to know the location and their contacts, and it is possible to build models that can offer support and assistance to such individuals [<https://doi.org/10.1007/s10209-023-00991-2>].

AI systems hold great promise in suicide research [19]. They can detect subtle data patterns, relationships, and variable interactions that are difficult for humans to discern. Diverse data types like clinical records, questionnaires, biological markers, and social media can be synthesized [20]. There is scalability to large populations while allowing the model to continuously improve by learning from new data. Predictions can be made more objectively, without individual biases or limited clinical judgment [21]. Thanks to developments in AI in recent years, new research perspectives in suicide prevention and prediction have emerged. Articles analyzing specific suicide detection techniques based on NLP or DL are now present and widespread in the literature. However, a comprehensive and unified analysis of all AI

techniques is lacking in this field. In this regard, the proposal of our article is to offer a comprehensive review of all AI techniques aimed at suicide prevention and prediction, with a focus on the last five years (2019-2023).

2. Material and Methods

In this section, we describe the process adopted to select the different papers for our review and specify how these considered papers are grouped and analyzed in the rest of the study.

2.1 Related Reviews

In this work, we used different combinations of the words “suicide”, “machine learning”, “deep learning”, “natural language processing”, “review”, and “survey” to search Science Direct, PubMed, Web of Science, IEEE Xplore, PsychINFO, and Google Scholar databases and understand the recent review studies published in this field. This search resulted in a total of 141 unique papers, out of which 32 were systematic review studies published in Q1 journals and related to our work. Table A. 12 lists these studies while Figure 1: Comparison of the focus and coverage of our systematic review with five existing literature reviews on the topic of suicide prevention/prediction. pictorially shows five representative review papers that best summarize these findings. Arowosegbe et al. [17] specifically examined the use of NLP in detecting and preventing suicide ideation (SI). Combining structured and unstructured data in NLP models produces more accurate outcomes than using either type alone. Passive monitoring of individuals with mental health issues is essential for suicide prevention. The conclusion highlights that AI and ML offer promising new directions for improving risk prediction and suicide prevention strategies, with NLP providing potentially cost-effective and efficient alternatives to current methods. While their study covered SI prediction studies from textual data, they did not cover other modes of data, such as audio, video, and images. Bernert et al. [22] examined the use of ML in suicide prevention. They highlight traditional risk factors for suicide, such as mood disorders, male gender, and unemployment, while also identifying new variables like sleep patterns and NLP-derived speech data. These emerging factors underscore the need for further research in suicide prevention. Despite the challenges posed by the rarity of suicide, AI and ML offer significant potential for enhancing precision in risk prediction and prevention. The integration of AI, particularly NLP applied to social media, could play a crucial role in early detection, risk assessment, and treatment development, with important methodological considerations. This review did not consider DL-based methods. Franco-Martín et al. [23] examined AI-based suicide preventive interventions in the decade up to early 2017. The authors point to a lack of studies on social networks in this area. They found only one paper related to ML. Their study focuses on technology-based data like that obtained from social media, and hence the studies reviewed cover NLP-based techniques in general. Hopkins et al. [24] reviewed a small number of studies and compared structured and unstructured data for ML-based prediction of suicidal behaviors. The review found that ML algorithms achieved accuracy scores equal to or higher than traditional suicide risk assessments and clinician judgment. Both structured and unstructured data sources yielded

similar accuracy in predicting suicide risk, despite differences in data organization. Given the global challenge of suicide as a leading cause of preventable death and the limited progress in risk detection, exploring innovative, technology-based methods is crucial for advancing clinical practice. They did not include DL studies in their review. Finally, Lejeune et al. [25] reviewed studies that used AI to predict suicide risk at individual and group levels. While AI has been successfully implemented in other medical fields and could complement existing suicide prevention tools, it is not yet ready for clinical use. Its effectiveness in clinical practice remains unclear. Only limited studies were included in their review. Our current work proposes to cover ML, DL, and NLP-based methods as well as address dual aspects of suicide prediction and prevention, thus providing a broader scope than previously published reviews. This review not only encompasses the insights from these previous studies but also extends the scope to include the most recent developments (2019-2023) across all AI domains in suicide research, providing a more comprehensive and up-to-date overview of the field.

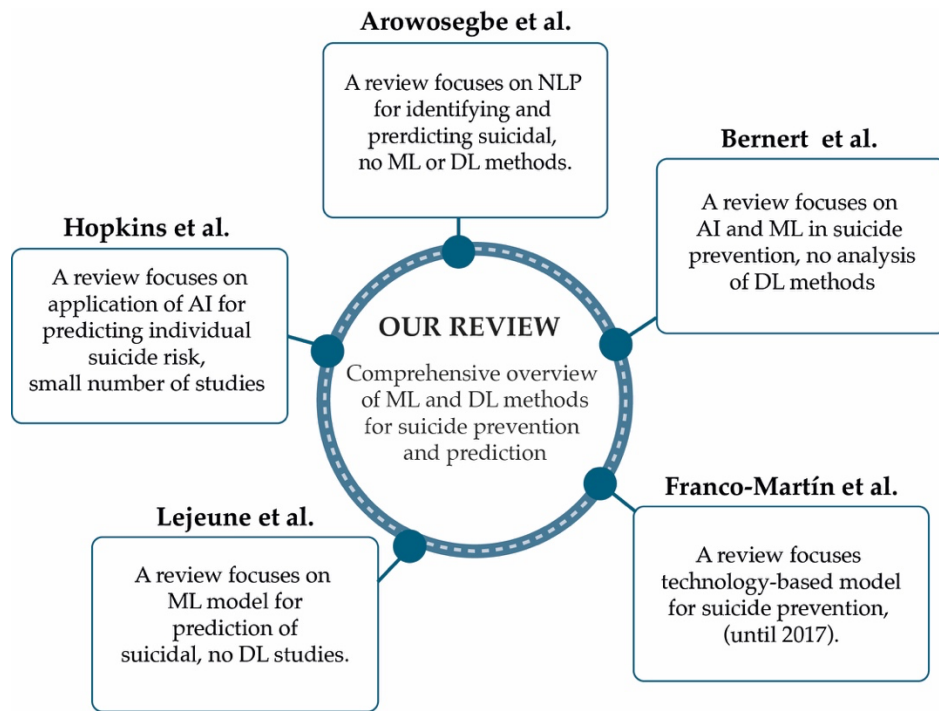


Figure 1: Comparison of the focus and coverage of our systematic review with five existing literature reviews on the topic of suicide prevention/prediction.

In our systematic review, we aim to provide a comprehensive overview of how AI, ML, DL, and NLP techniques have been applied to suicide prevention and prediction over the past 5 years. We provide a contemporary overview of the performance of predictive modeling studies and the effectiveness of AI-based preventive interventions, as well as a discussion of ethical considerations and translational challenges. We discuss different publicly available databases for suicide research, present several limitations and challenges faced by this field today, and list specific future directions to overcome these, specifically focusing on studies that use multimodal data.

2.2 Literature Search

We performed a systematic search of Science Direct, PubMed, Web of Science, IEEE Xplore, and PsychINFO databases for Q1-ranked peer-reviewed journal articles on AI-enabled suicide prediction and/or prevention that had been published in the English language between 2019 and March 2024 using the Boolean search string: (self-harm OR suicide OR suicidal OR self-destruction) AND (machine learning OR natural language processing OR artificial intelligence OR deep learning OR neural network OR automated) AND (prevention OR detection OR identification OR prediction). Articles were searched for the presence of these keywords in their titles and abstracts. The initial search yielded 1421 records. After excluding 531 duplicates, 92 review papers, 295 conference papers, 6 book chapters, 21 abstracts, and 144 non-Q1-ranked journal articles, 332 studies remained. The latter underwent full-text review for relevance, and a further 116 non-suicide-related, 33 non-suicide prediction/prevention-related, and 27 non-ML-related articles were excluded, leaving 156 articles for inclusion in the analysis (Figure 2: Study search strategy by PRISMA guidelines.).

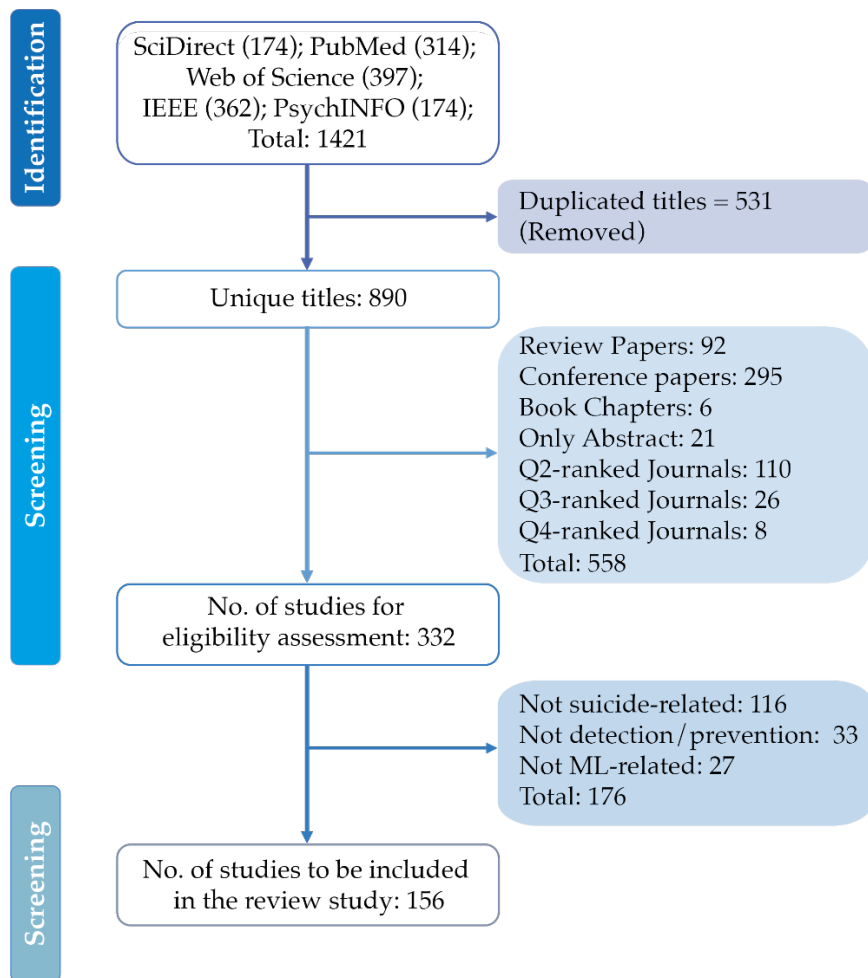


Figure 2: Study search strategy by PRISMA guidelines.

2.3 Applications

The 156 studies considered for review were first grouped based on their primary objective: suicide risk prediction or suicide prevention. Although both categories focus on suicide, there are important differences in their scope and target populations. Suicide prediction refers to estimating an individual's risk of suicide attempt or death in the future. In contrast, suicide prevention involves strategies, interventions and programs aimed at reducing suicide risk and suicidal behaviors. While papers in the suicide prediction category primarily work on predicting the chances of any given individual committing suicide, the papers in the suicide prevention category deal largely with predicting SI in more vulnerable, high-risk individuals, and these studies also discuss strategies to prevent cases of suicide. Furthermore, we identified the most common data sources used in the reviewed studies and categorized them into 5 types:

1. Clinical Data (Table A. 1, Table A. 6) - Includes EHR, hospital data, medical records, clinical notes and assessments, physiological data such as EEG/MRI/fMRI, and biomarkers and genetic data.
2. Behavioral and Psychosocial Data (Table A. 2, Table A. 7) - Contains questionnaires, surveys, interviews, self-reports collecting direct inputs, and counseling and consultation notes.
3. Digital and Technology Data (Table A. 3, Table A. 8) - Comprises social media posts, chats, messages, wearable sensors, digital footprints, and audio/phone calls.
4. Administrative and Public Data (Table A. 4, Table A. 9) - Includes government and public records.
5. Multimodal Data (Table A. 5, Table A. 10)- Refers to combinations of the above data types.

Figure 3 shows these five categories along with some examples of each data type. Finally, we also categorized the works by geographic area and classification technique employed (NLP, ML, or DL). These interrelated methods have different strengths and limitations for suicide research, from detecting signals in unstructured text to multivariate pattern recognition. while Section 4 reports key insights and statistics on suicide research using AI techniques.

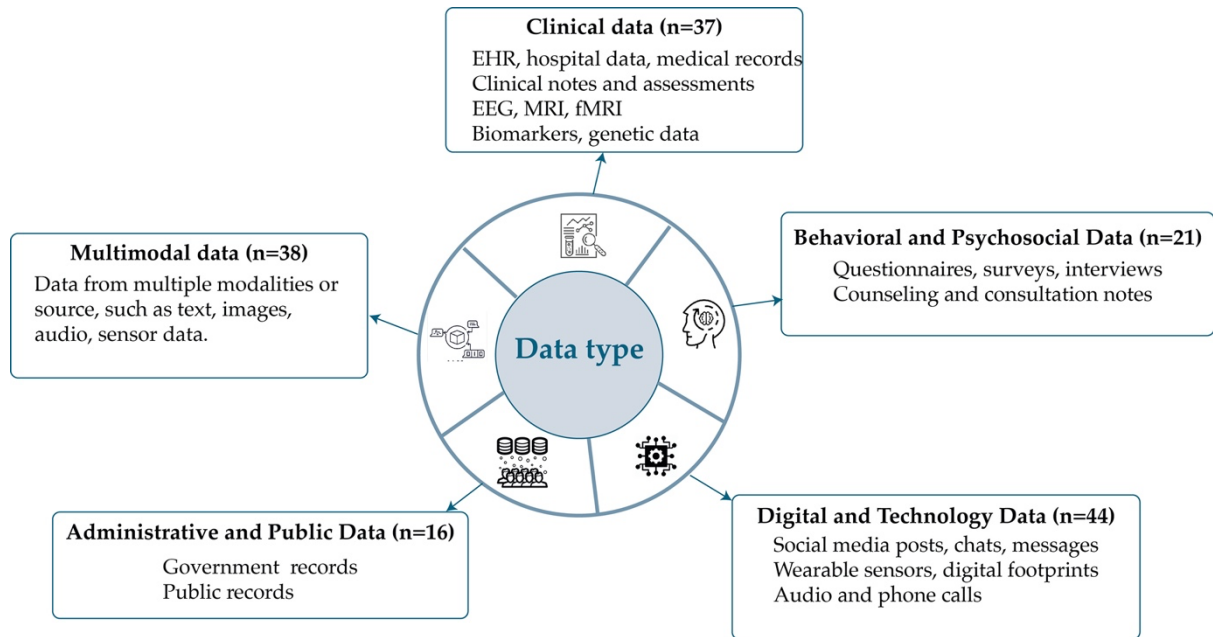


Figure 3: Categories of data sources used in studies on AI techniques for suicide prevention and prediction.

3. Results and Discussion

In this section, we present the important results of the reviewed studies. First, we perform a statistical analysis to understand different trends in this domain. Following that, we discuss in-depth all reviewed articles published on suicide prevention and prediction categories and the type of data used in the studies. Finally, we discuss the publicly available datasets for suicide research.

3.1 Statistical analysis of the reviewed studies

Figure 4: Reviewed papers categorized by task (suicide prediction/prevention), divided by publication year. presents a bar plot showing the trend in the number of studies focusing on suicide prediction and prevention from 2019 to 2024. The graph reveals a rapid increase in research on this topic from 2020 onwards. This is likely influenced by the COVID-19 pandemic and associated lockdowns. Across all years, the number of studies on prediction is consistently higher than those on prevention. This suggests that researchers may find prediction a more targeted and well-defined task, while prevention efforts can be more diffuse and multifaceted, spanning diverse populations and strategies.

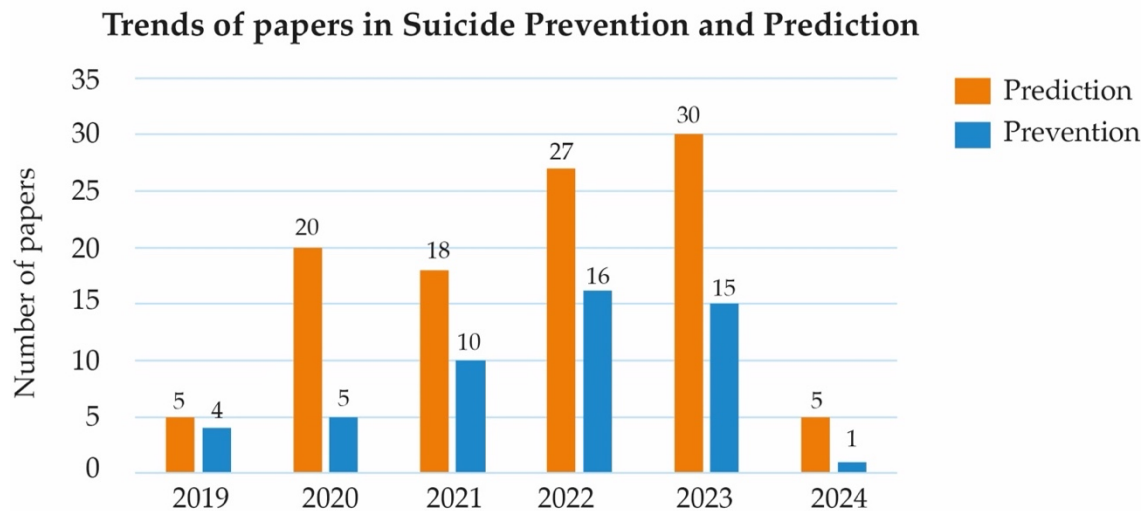


Figure 4: Reviewed papers categorized by task (suicide prediction/prevention), divided by publication year.

Figure 5: Distribution of data types used in studies of suicide prevention and prediction. provides a high-level categorization of the data types used in the reviewed studies. Most studies leveraged digital and technology data (n=44), such as social media posts and online content, followed by multimodal data (n=38), highlighting the potential of integrating diverse data streams for a more comprehensive understanding of suicide risk factors. Digital and technology data includes any data from sources like social media sites, messages, phone calls, videos, wearables, body sensors, etc. Notably, a significant proportion of studies employed clinical data (n=37) from sources like EHRs and medical notes. Data from medical imaging techniques like EEG and MRI, along with data involving biomarkers and genetic information, fall in this category. Behavioral and psychosocial data (n=21) from sources such as surveys, interviews and questionnaires, as well as administrative and public records data (n=16), were also utilized to a lesser extent. With online platforms and social media acting as a rich source of almost real-time data on individuals' thoughts, emotions, and behaviors, this breakdown highlights the growing role of digital data in suicide research. The prevalence of social media data points to its potential for early warning signs and real-time monitoring but also raises important considerations around data quality, representativeness, and ethics. Clinical data remains indispensable for getting insights into professional mental health expertise. Questionnaires, interviews, and clinical assessments remain gold standards that provide essential diagnostic and phenomenological context. Emerging data types like medical imaging, digital phenotyping, and omics data hold promise for novel biological and behavioral insights but are still in the early stages. The increasing use of multimodal data is promising, as combining digital footprints with clinical, behavioral, and socio-environmental factors may give a more holistic picture of suicide risk.

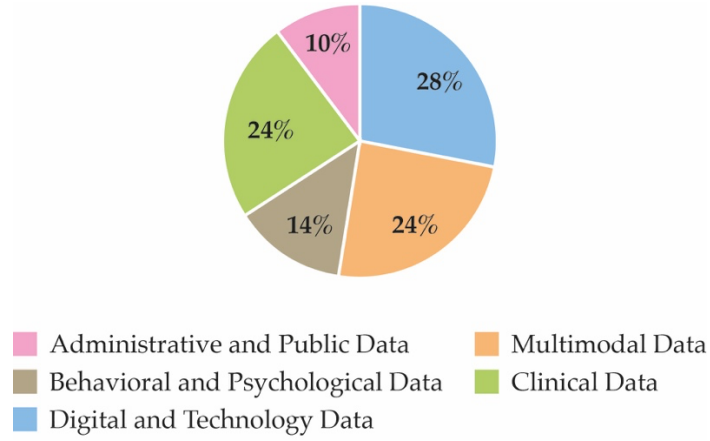


Figure 5: Distribution of data types used in studies of suicide prevention and prediction.

Several papers focused on data retrieved from a specific country. The reviewed articles were categorized by the countries in which the studies were conducted, which was identified based on where the data was obtained from. Figure 6: Continent-wise geographical distribution of the reviewed studies. illustrates the geographical distribution of the reviewed studies based on the continent, while Figure 7: Country-wise geographical distribution of the reviewed studies. highlights the countries where suicide-related research has taken place the most. The study and analysis of suicide prevention and prediction strategies find most space in Asia and North America. These are the most represented continents, accounting for over 75% of the reviewed studies. The overwhelming part of the contribution to North America, the most represented continent, is provided by the United States. As for Asia, the largest contribution is made by China, followed by Korea and Taiwan. In Europe, the Nordic countries, including the United Kingdom, Denmark, Netherlands, and Norway, have the largest number of studies. Oceania contributes to 9% of the reviewed research. This indicates the importance given to mental health, effective steps taken by the governments, and high-level infrastructure provided for suicide research in these countries, along with growing awareness and openness, encouraging proactive interventions. At the same time, this is also a reflection of high suicide rates in these countries, while different countries from the African continent lack representation, despite high levels of suicide [26].

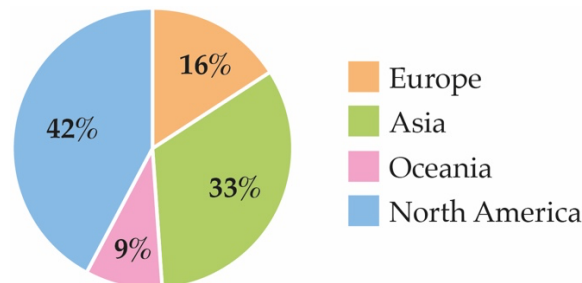


Figure 6: Continent-wise geographical distribution of the reviewed studies.

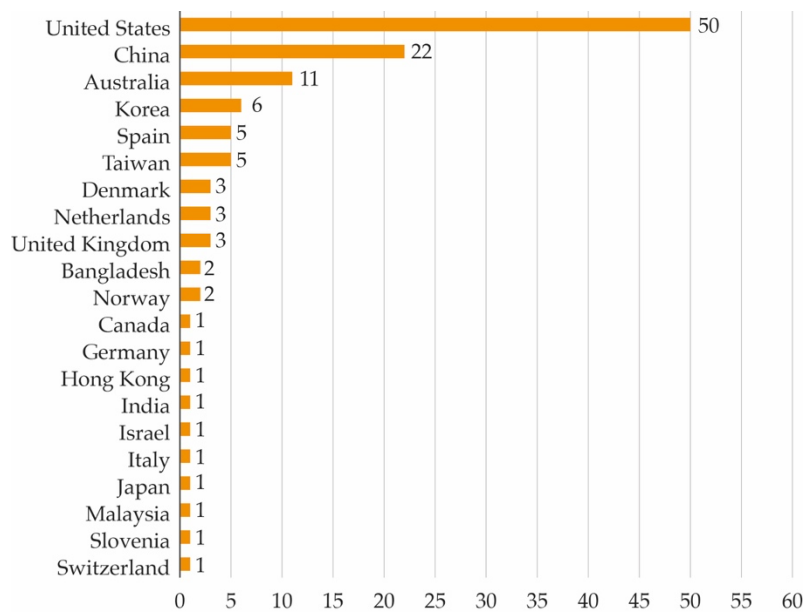


Figure 7: Country-wise geographical distribution of the reviewed studies.

Figure 8: Distribution of the studies based on age group. and **Error! Reference source not found.**offer valuable insights into the populations most affected by suicide risk, as studied in the reviewed literature. The data reveals that adults, young people, and adolescents are the age groups most frequently investigated. This aligns with epidemiological data showing that suicide is a leading cause of death among young and middle-aged individuals globally. The figure also highlights key risk factors and subpopulations studied, including people with psychiatric disorders, military personnel and veterans, and individuals impacted by the COVID-19 pandemic. The high representation of psychiatric populations underscores the well-established link between mental health conditions and suicide risk. Military and veteran populations also emerge as a key focus, likely due to the unique stressors, traumas, and access to lethal means associated with these groups. The presence of COVID-19 as a notable factor reflects the profound mental health toll of the pandemic, which disrupted social connections, livelihoods, and support systems. These findings emphasize the need for targeted, culturally sensitive interventions that address the unique risk factors and needs of these diverse subpopulations. AI-driven prediction and prevention efforts should strive to be inclusive and equitable, accounting for the intersectionality of risk factors. Once again, it is to be noted that the counts in Figure 8: Distribution of the studies based on age group. and **Error! Reference source not found.**do not add up to 156 (the number of papers reviewed in this study) because several papers do not target any such group of the population.

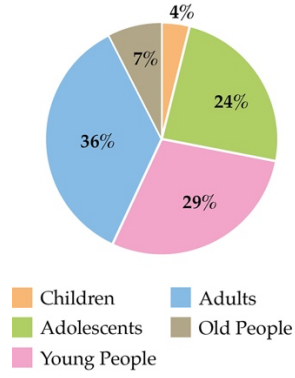


Figure 8: Distribution of the studies based on age group.

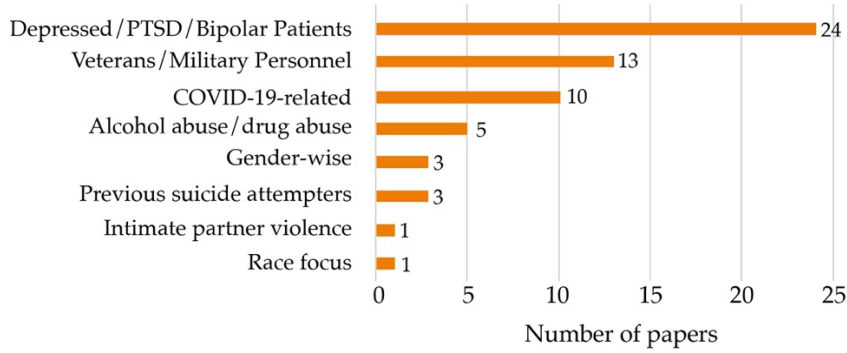


Figure 9: Distribution of studies based on primary focus.

3.2 Suicide prevention

This section provides an exhaustive description of the reviewed works on suicide prevention strategies categorized by the type of data used.

3.2.1 Clinical data

This section summarizes recent works applying ML, DL, and NLP to analyze clinical data for insights related to STBs. These works utilize techniques like ANNs, Support Vector Machine (SVM), Random Forest (RF), etc., to extract meaningful patterns from EHR, emergency department notes, and other clinical text. Table A. 1 summarizes the works reported in this section. For instance, McManus et al. [27] came up with the implementation of a national clinical text processing infrastructure within the Department of Veterans Affairs. They recorded the various operational challenges and solutions in integrating text processing for suicide and overdose prevention at a large scale. Similarly, Robinson et al. [7] developed a state-based self-harm monitoring system in Australia. Using ML-based techniques, they were able to identify episodes of self-harm from emergency department records successfully. Despite such studies, it is commonly observed that relying solely on diagnostic codes is not a good practice because they often tend to underreport SI [16, 28]. Hence, they focused on deploying NLP techniques to extract actionable insights from unstructured clinical notes. Using such state-of-the-art NLP techniques allowed them to identify hidden patterns indicative of suicide risk. One key similarity across the studies is the utilization of large

datasets, such as EHRs, to refine predictive accuracy for suicide risks. For example, Zhu et al. [16] and Goldman-Mellor et al. [28] applied different ML methods to analyze patterns associated with suicidal behaviors. In their works, models like SVM and RF performed exceedingly well in identifying individuals at risk from textual data in EHRs. They reached an F1-score of up to 0.9889. In a similar study, Bejan et al. [29] also highlighted the effectiveness of NLP in processing large-scale EHRs. Measuring the performance using the area under the receiver operating characteristic curve (AUROC) score, they achieved a score of 98.6 for SI and 97.3 for suicide attempts. These findings suggest that combining NLP with diagnostic codes enhances the identification of suicidal behaviors, enabling more timely and accurate interventions.

Apart from being able to utilize patterns and features that are not obvious to the human eye, AI-based models have also gained popularity for being portable and scalable. Being able to perform well across different clinical settings for a wide range of patient cohorts is very important. NLP methods developed independently at different institutions have the ability to detect suicidality across diverse healthcare systems, though performance may vary slightly [32]. Similar to NLP-based methods, ML-based studies have also been able to identify predictors of suicide attempts in different settings. Edgcomb et al. [31] achieved an AUROC of 0.84, predicting SI among women with serious mental illness in both single and multi-institutional datasets.

Such studies collectively highlight the significant potential of AI in enhancing suicide prevention efforts [32, 33, 34]. They illustrate a shift towards data-driven, real-time monitoring systems that can dynamically inform public health responses. However, challenges remain in quantitatively measuring and comparing AI's impact on suicide prevention across different health systems.

3.2.2 Behavioral and Psychosocial Data

Recent studies have utilized AI and ML to enhance suicide prevention strategies using behavioral and psychological data. Behavioral and psychological data refers to information gathered about individual's actions, habits, mental processes, and emotional states from sources such as questionnaires, surveys, interviews, self-reports, and counseling notes. Motivated to explore the potential of AI in early intervention, Kearns et al. [35] developed a risk calculator for U.S. Army soldiers, achieving an AUROC of 0.85. Haghish et al. [36] also achieved an area under curve (AUC) score of over 0.85. However, their important contribution was the study of the significance of false positives in suicide prediction, suggesting that these individuals may represent a genuine at-risk population. Meanwhile, Prelog et al. [37] explored a novel approach using COVID-19-related data to assess SI. With Biernesser et al. [38] working towards SI in adolescents, these studies highlight the two-fold capability of AI. First, they're adaptable to various data sources and contexts. Second, the importance of considering false positives not as mere statistical errors but as potentially at-risk individuals enhances the effectiveness of suicide prevention programs. Table A. 2 summarizes the works reported in this section.

3.2.3 Digital and Technology Data

Several papers [39, 40, 41, 42] use ML to predict and mitigate suicide risk through social media and other digital platforms (Table A. 3). When digital and technology data involves posts, comments, and chats from various social media platforms, NLP and linguistic feature extraction come into the picture. Linguistic markers like first-person pronouns and negations often have significant associations with suicide risk [43]. Similarly, tweets related to suicide prevention often correlate positively with help-seeking behavior, while they correlate negatively with suicide rates [44]. Tweets about suicide cases tend to have the opposite effect. Caicedo et al. [45] systematically analyzed various classifiers on a bilingual corpus from social networks, achieving the highest performance with the KStar algorithm (AUROC 0.81036, F-measure 0.7148). Using diverse social media platforms, they highlighted the advantages of multilingual corpora and systematic classifier evaluations. Their work highlights the importance of corpus quality and demonstrates the feasibility of automatic detection of SI across various digital environments. Meanwhile, Bello et al. [46] focused on classifying suicide-related news articles using convolutional neural networks (CNNs). This suggests that apart from social media platforms like Twitter and Reddit, there is a high impact of media coverage on public perception and the thematic consistency of suicide news.

Research on suicide prediction using NLP in digital data (Table A. 3) presents diverse methodologies and findings. Yang et al. [47] implemented a knowledge graph-based AI system to monitor and intervene with high-risk users on Weibo. By analyzing patterns in user comments and providing timely interventions, they prevented 3,629 potential suicides. For effective intervention and prevention, it is crucial to identify suicide risk at an early stage. Distant supervision and DL methods have proven effective in building early warning systems [48]. For example, BERT-based models have been effective in analyzing tweets to validate the alignment of social media-identified suicide risks with actual suicide statistics [48]. This supports the utility of social media as a tool for public surveillance and intervention. Levkovich et al. [49] compared the performance of ChatGPT-3.5 and ChatGPT-4 in assessing suicide risk. Unsurprisingly, ChatGPT-4 closely matched professional evaluations in severe cases, while ChatGPT-3.5 tended to underestimate risks. Deas et al. [50] studied anti-mattering in online suicide support communities, using NLP to analyze posts from the r/SuicideWatch subreddit. They concluded that SI had a significant correlation with feelings of insignificance. Salmi et al. [51] investigated changes in help-seeker conversations on a Dutch suicide prevention helpline during COVID-19. Like [48], they also used BERT for topic modeling and found significant shifts in conversation topics post-lockdown. Such a temporal analysis emphasizes the need for real-time data for timely interventions. The integration of NLP in such studies highlights the potential of AI to enhance traditional suicide prevention methods by providing real-time, scalable, and accurate risk assessments [52, 53, 54, 55, 56]. However, the studies also emphasize the need for ongoing refinement and validation of these models to address cultural and linguistic nuances and to ensure their reliability across different contexts.

3.2.4 Administrative and Public Data

Table A. 4 presents six suicide prevention studies using administrative and public data along with ML and NLP techniques across diverse contexts and populations. These papers emphasize the use of large-scale databases like the National Violent Death Reporting System (NVDRS) to train models better to predict intricate community characteristics like intimate partner violence (IPV).

Haroz et al. [57] specifically targeted Native American communities, developing a model that achieved an AUC of 0.87, significantly enhancing the prediction accuracy for suicide risks within this group. Turner et al. [58] described the implementation of the Zero Suicide Framework within a mental health service in Australia, including steps taken and outcomes. On the other hand, Wang et al. [59] focused on a broader demographic. They created a case-control study to predict the overall population risk of suicide using administrative data. In such settings, targeted prevention strategies are the need of the hour [60, 61]. 7.1% of suicides were IPV-related [60]. NLP tools have achieved high performance (specificity and precision of 0.98 and 0.72, respectively) in identifying IPV circumstances in suicide cases [61].

Among the studies, Haroz et al. [57] reported very good results presenting a targeted approach within a high-risk community. This illustrates the effectiveness of tailored predictive models considering unique demographic and community factors. This also addresses the ongoing need to refine and adapt suicide prevention models to enhance their applicability in predicting broader social and community indicators.

3.2.5 Multimodal Studies

Recent studies have used multimodal data sources, such as mobile devices, wearables, EHRs, clinical notes, and social media. These approaches aim to identify at-risk individuals and deliver interventions by detecting patterns from an amalgamation of data, such as an individual's behavior, emotions, and environment. Table A. 5 summarize the works reported in this section.

Kessler et al. [62] improved post-discharge suicide prediction among psychiatric inpatients by using EHR, clinical notes, and social determinants of health (SDoH). Their NLP-based model achieved AUROC scores between 0.747-0.780. Similarly, Heckler et al. [63] utilized sensor data, electronic medical assessments, and social media to develop a perming SI-prediction model over 3, 5, and 7-day periods.

Multimodal data has been extremely useful in preventing suicide among young adults and adolescents [64, 65, 66]. Mouchabac et al. [64] developed a Bayesian network-based smartphone application for adolescents to prevent suicide relapses, demonstrating a notable improvement in crisis management through personalized, algorithm-driven advice. Liu et al. [65] implemented the Proactive Suicide Prevention Online approach using ML to identify high-risk individuals on social media, achieving high-performance metrics (precision: 0.86, recall: 0.78, F-measure: 0.86, accuracy: 0.88). Finally, Balbuena et al. [66] also used ML models to analyze both the general population and clinical samples, highlighting the complexity and limitations in predicting imminent suicide risk. While Mouchabac et al. [64] focused on

preclinical interventions and Liu et al. [65] on real-time identification and management through social media, Balbuena et al. [66] provided a broader analysis of risk factors in large datasets by highlighting the challenges of low predictive value despite extensive data.

SVM and logistic regression (LR) were often used in multimodal data. Real-time monitoring of at-risk individuals is important for personalized interventions. Berrouiguet et al. [67] used continuous data from smartphone sensors and ecological momentary assessment to predict suicide risk. Similarly, Sels et al. [68] combined self-report data and passive mobile sensing to monitor and predict SI, achieving accurate predictions using SVM. Kyron et al. [69] focused on using big data techniques to analyze media coverage of suicide during the COVID-19 pandemic. Increased media attention on COVID-19 and suicide-related topics seemed to have a significant correlation. They achieved high precision and recall scores in classifying news articles, highlighting the potential of AI in understanding and mitigating the impacts of media on public health. Finally, Bello et al. [70] explored the causal relationships between COVID-19 media coverage and suicide rates using SVM. Achieving an F1 score of 0.73, they found that SVM is highly capable of analyzing large datasets (from news articles and social media posts), and performing efficient multi-class classification.

These studies highlight the efficacy of SVM and LR in suicide prevention using multimodal data. They also emphasize integrating multimodal data, such as passive sensor data, self-reports, and media content, to enhance predictive accuracy for timely interventions [71].

3.3 Suicide prediction

This section provides an exhaustive description of the reviewed works on suicide prediction strategies categorized by the type of data used.

3.3.1 Clinical data

Traditionally, ML-based models have performed exceedingly well on structured data. Clinical data is one example [72, 73, 74, 75, 76, 77, 78]. Methods like RF and SVM have been applied to EHRs, doctors' notes, medical checkup data, and other measurements to identify crucial patterns for SI [79].

Table A. 6 summarizes the works reported in this section. Like how administrative and public data gave us insights into IPV (section 3.1.4), clinical data helps us identify factors like alcohol consumption, healthcare uptake, and depression severity for SI prediction [80, 81, 82, 83]. RF was the most commonly used ML model to predict SI from clinical data. Mens et al. [80] applied RF to general practice data, achieving an AUC of 0.82 and a sensitivity of 0.39. They noted that many patients had an increase in healthcare uptake right before suicide attempts. Rozova et al. [81] also used RF on emergency department triage notes, showing strong performance with an AUC of 0.87. Nordin et al. [82] went one step further in building an ensemble model. They combined RF with Gradient Boosting (GB) for explainability and achieved an accuracy of 0.84. Like [80], they found that past suicide attempts and the severity of depression were key predictors of SI. Jiang et al. [83] compared RF with Lasso regression, finding that RF performed well in predicting suicide attempts among depressed patients. In

their study, poisoning and psychiatric disorders were identified as significant predictors. Cho et al. [84] focused on using RF for suicide prediction in a general population cohort. They identified physical activity and alcohol consumption as important variables.

Apart from RF, SVM was the most commonly used technique. For example, Hong et al. [3] used SVM to differentiate between suicide attempters and patients with SI but without attempts among adolescents with MDD, achieving a balanced accuracy of 78.59%. Similarly, Zhu et al. [85] used SVM to classify bipolar disorder patients with and without a history of suicide attempts. They achieved a slightly higher accuracy of 84% for their task using features from resting-state functional connectivity. Bajaj et al. [86] applied SVM on MRI data of adolescents to distinguish those at suicide risk from typically developing peers with an AUC of 87.18%. Finally, Xu et al. [87] used SVM to analyze dynamic functional network connectivity features to stratify MDD patients with diverse suicide risk levels. They were able to distinguish between suicide attempters and people with ideation 80% of the time.

Several studies [11, 88] used NLP on clinical data to predict SI in adolescents and to identify past suicide attempts. However, the performance of NLP is not restricted to the task of identifying past suicide attempts. Cusick et al. [89] employed a weakly supervised DL approach to detect current SI from clinical notes and achieved an accuracy of 94% in identifying at-risk patients. Similarly, NLP helped Levis et al. [90] to predict SI successfully from unstructured EHR data.

Yet another advantage of using NLP-based metrics on clinical data is domain-specific vocabularies. Compared to structured data, these features provide more information and help predict SI better [91]. They are also very efficient in identifying SI from clinical notes in large healthcare systems [92].

Although less common as compared to NLP, transfer learning and DL have also been used on clinical data. Chen et al. [93] used CNNs and DenseNet on MRI data to distinguish depressive patients with SI from healthy controls. The DenseNet model achieved an accuracy of 93.7% in identifying suicide attempters using isotropic value data, outperforming the simpler CNN models. Xu et al. [2] incorporated a comorbidity network called Dx2Vec to predict self-harm within 12 months post-discharge among 2,323 patients. This model outperformed traditional approaches with significant improvements in recall for positive cases. These studies highlight the potential of NLP and DL to enhance early intervention and suicide prevention efforts.

3.3.2 Behavioral and Psychosocial Data

The application of ML and data mining techniques to behavioral and psychosocial data shows promise for identifying individuals at risk of suicide (Table A. 7). Behavioral and psychological data include surveys, clinical notes, social media posts, and crisis counseling messages. Studies here have utilized supervised learning algorithms like decision trees (DT), LR, SVM, and different NLP and ensemble techniques. For example, Li et al. [94] and Broadbent et al. [95] highlight the integration of RF and NLP to analyze clinical interviews and crisis counseling data to distinguish SI from depression. Bantilan et al. [96] also use NLP and therapist interventions to develop a suicide alert system for telemedicine. Finally, Morales-

Rodríguez et al. [97] apply weak supervision-based labeling to enhance emotion classification in suicide notes.

Across the studies, RF was consistently used alongside other ML algorithms. One of the best examples is Lin et al. [98], who achieved over 98% accuracy in predicting SI based on psychological stress domains using RF. Similarly, Jung et al. [98] found RF to be highly accurate (77.8%), although slightly less than extreme gradient boosting (XGB) and SVM. Additionally, RF was seen to perform well in settings with imbalanced data [100] and was also superior to traditional LR models in predicting suicide-related outcomes [101]. Finally, Zuromski et al. [102] and Mahmud et al [103] also established the robust predictive power of RF by evaluating it on metrics like ROC, balanced sensitivity, and specificity. These studies showcase the promise of RF models in accurately predicting suicidal behavior, particularly when combined with techniques to handle class imbalance and optimize feature selection.

3.3.3 Digital and Technology Data

Several studies have leveraged ML [104, 105, 106], DL [107, 108, 109, 110], and NLP [111, 112, 113, 114] models to analyze patterns in social media posts, EHRs, and other digital text data to identify signs of suicide risk and provide timely interventions. Table A. 8 summarizes the works reported in this section.

Schoene et al. [115] investigated the use of hierarchical multiscale RNNs for detecting suicide notes from other social media texts. Yao et al. [116] focused on detecting suicidality among opioid users on Reddit through an ML-based approach, highlighting the integration of NLP techniques with social media data. Ragheb et al. [117] introduced a novel ensemble architecture called Negatively Correlated Noisy Learners for identifying at-risk users on social networks, addressing issues like depression, anorexia, self-harm, and suicide with enhanced classification accuracy. Talaat et al. [118] explored the use of social media for personality and emotion recognition through ML, which indirectly aids in identifying potential suicidal tendencies. Across these studies, the emphasis on text analysis and the application of advanced algorithms like RNNs and ensemble methods are to be noted. The best results were observed in Ragheb et al.'s study [117], where their ensemble model demonstrated superior accuracy and robustness in detecting at-risk users, showcasing the potential of ensemble techniques in mental health monitoring.

Shen et al. [5] developed an RF model to predict suicide attempts among Chinese medical students. The model achieved an AUC of 0.9255 and an accuracy of 90.1%, highlighting its robustness in a high-risk population. Rabani et al. [119] proposed an enhanced feature engineering approach for suicidal risk identification using data from Twitter and Reddit. Once again, their Extreme Gradient-Boosting ensemble model achieved the highest accuracy, significantly outperforming classical feature extraction methods. Lastly, Ramírez-Cifuentes et al. [120] performed a comparative study of different DL techniques such as LSTM and CNN, to detect depressive states in social media posts, emphasizing the effectiveness of word embeddings in enhancing prediction accuracy.

When dealing with social media, Twitter has been a widely sought-after choice for collecting data. Baghdadi et al. [121] focused on detecting suicidal tendencies in Arabic tweets, achieving an impressive weighted sum metric of 95.26% using BERT models. Schoene et al. [122] also worked on Twitter but introduced a dataset called TWISCO and used graph convolutional networks to classify suicide-related content on Twitter. Their work stands out because of its ability to distinguish actual suicidal intent from sarcastic, humorous mentions. Buddhitha et al. [123] explored multi-task learning to detect SI and mental disorders. Their analysis of data collected from Twitter highlights the impact of post-traumatic stress disorder (PTSD) on predictability. Additionally, their multi-task learning model demonstrated enhanced cross-platform performance with an AUC of 0.96 when data was combined from both PTSD and depression-diagnosed users.

DL-based methods have also been very successful in this field. Renjith et al. [124] employed a combination of LSTMs and CNNs with the attention mechanism to detect SI from social media posts. Their model achieved an accuracy of 90.3% and an F1-score of 92.6%, demonstrating significant improvements over other baselines. Adarsh et al. [125] proposed an ensemble model incorporating semantic network analysis and clustering algorithms to identify early warning signs of suicidal thoughts. They achieved excellent results post-label correction and bias removal. Lastly, Zhang et al. (2021) [15] developed a transformer-based model called TransformerRNN that outperformed traditional ML and other DL models by effectively capturing both local and global semantic features. TransformerRNN achieved a precision, recall, and F1-score of 95.0%, 94.9%, and 94.9%, respectively.

3.3.4 Administrative and Public Data

Recent studies [126, 127, 128] have used administrative and public data to develop ML and NLP models for suicide risk prediction and analysis. Because of their large scale, administrative and public data provide great insights at the population level. Initial studies have shown promise in predicting the risk of SI using features extracted from such healthcare, government, and survey data. Table A. 9 summarize the works reported in this section.

Lin et al. [129] employed DTs, SVMs, and ANNs on psychiatric patient data from Taiwan. Compared to individual models, they found that an ensemble of Adaboost and DT achieved the highest prediction accuracy (97.1%). Kwon et al. [130] worked on similar lines, developing a model using XGB on Korean adolescent data. Achieving an AUROC of 86.5%, depression and stress levels were the key SI-predicting features. Horvath et al. [131] focused on prison populations, comparing various ML models and found gradient tree boosting most effective with an AUC of 0.875.

Four papers [10, 132, 133, 134] used NLP-based approaches for suicide prediction. Rahman et al. [132] employed NLP to enhance the classification of undetermined intent deaths. To improve SI prediction, they found that considering racial differences in narrative texts is very important. Combining structured and text-based data, they achieved an AUC of 0.987. Wang et al. [133] developed a transformer-based classifier to distinguish between SDoH related to suicide crises and narratives of death investigations. They achieved AUC scores of over 0.94

for both circumstances and crises. Shin et al. [134] and Kim et al. [10] focused on predicting SI in Korean adolescents using various algorithms. They used CNN and BERT-based models to achieve an accuracy of over 90%. The hierarchical attention network by Kim et al. [10] on emergency department data to predict self-harm diagnoses achieved great results. These studies demonstrated significant predictive capabilities of ensemble and NLP-based approaches, each tailored to specific data and populations. They also highlight the importance of combining structured and unstructured data for better prediction.

3.3.5 Multimodal Studies

Multimodal approaches integrating various data types have shown promise in predicting suicidal thoughts. They combine data from sources like clinical assessments and biomarkers [135, 136, 137, 138], mobile sensors [139], and social media inputs [140, 141, 142, 143] with rule-based approaches [144], ML [145, 146, 147, 148], and DL [149, 150] methods to provide a comprehensive view of risk and protective factors. Since they operate at different levels over time, they allow the capturing of a fuller picture of risk and protective factors. In this section, we discuss and summarize some of the best works published using multimodal data while Table A. 10 further enlists all the works considered here.

Rozek et al. [151] utilized the MondoBrain AI platform to predict suicide attempts among military personnel. They found that worst-point SI, history of multiple suicide attempts, and gender (male) were key predictors of suicide. They also found that LR was not the most effective algorithm in multimodal settings. Miranda et al. [152] employed the DeepBiomarker2 model on electronic medical records of PTSD/alcohol disorder patients to predict adverse events like suicide. Moving ahead of [151], their work suggests that DL methods perform exceedingly well in multimodal and clinical settings. Cai et al. [153] developed a depression detection model using social media data, leveraging linguistic and behavioral features to enhance predictive accuracy. Their model achieved the highest performance with an F1-score of 0.9202, demonstrating the efficacy of multimodal feature fusion in mental health surveillance. Finally, Hao et al. [154] took a rather unconventional approach using ML to identify suicide attempters among patients with severe depression. They analyzed functional connectivity between the insula and amygdala and achieved a significant differentiation accuracy of 85%.

Cao et al. [13] developed a personal suicide-oriented knowledge graph for detecting SI on social media by identifying key risk factors from users' posts, personality traits, and experiences. Supporting the claim from [151] about the ability of DL-based models on multimodal data, they achieved a classification accuracy of over 93% using a two-layered attention mechanism. Wilimitis et al. [155] demonstrated that combining face-to-face screening HER data improved suicide risk prediction. With an AUROC of 0.879 for SI, such a combination outperformed standalone models.

Kim et al. [156] evaluated the use of RF and K-Nearest Neighbors (KNNs) for predicting SI and attempts in college students, achieving accuracy of 92.9% and 95% with the two models, respectively. Their study highlighted the robustness of the Minnesota Multiphasic Personality

Inventory-2 combined with ML for reliable suicide risk prediction. Li et al. [157] proposed a Text-CNN with a multi-feature and label association model for detecting SI from Chinese social media posts. They used text data, posting time, and social interaction features, outperforming previous models in terms of accuracy. Lastly, Ghanadian et al. [158] focused on synthetic data generation to enhance SI detection. They used large language models to generate socially aware synthetic data that can be used to train more effective ML models. Using synthetic data would help researchers overcome privacy concerns and data scarcity issues.

3.4 The significance of AI techniques and multimodal data

Error! Reference source not found. describes the number of papers that use different AI-based approaches for studying suicide prediction and prevention. A sudden rise in interest in this field can be noted from 2020 onwards – resulting from both the gaining popularity of AI and the influence of the COVID-19 pandemic and lockdowns, which had significant mental health impacts worldwide [159]. The result of lockdowns also spiked an increase in the use of social media platforms during this period. The breakdown in Figure 5: Distribution of data types used in studies of suicide prevention and prediction. highlights the growing role of digital data in suicide research, as online platforms provide a rich source of real-time data on individuals' thoughts, emotions, and behaviors [160].

Error! Reference source not found. describes the grouping of papers based on their use of ML (n=107), DL (n=20), and NLP-based (n=29) techniques for the classification task. Clearly, ML-based models are more prevalent than NLP and DL-based techniques. This could be because, compared to DL and NLP-based approaches, ML techniques are well-suited for handling structured data and offer more interpretability. Many studies using ML focus on identifying the most important predictors of suicide risk using various feature selection techniques. This interpretability is crucial and valuable for informing clinical decision-making and intervention design. Among ML-based models, RF and SVM have been the most sought-after techniques, accounting for over 40% of the ML-based studies. While RF reduces overfitting and improves performance by averaging multiple decision trees, SVM excels by maximizing the margin between classes, offering robustness, especially in high-dimensional spaces [161]. Among ML-based models, LR has also been popular among researchers, accounting for over 20% of the studies, primarily because of its simplicity and interpretability in binary classification problems. In contrast, DL and NLP techniques are more commonly employed with unstructured data like images, signals, and textual data from social media sites. Their ability to learn complex, hierarchical representations is powerful but comes with challenges in interpretability and transparency. NLP-based methods are used on textual data, primarily on social media posts and messages. BERT and LSTMs are among the most common techniques being used to encode, understand, and decode this data. Over the last couple of years, the use of large language models (LLMs) has slowly started gaining popularity with models like ChatGPT performing well in several real-world tasks, primarily

because of its large model size and huge training data. Among the classical DL models, transformer-based models that use the attention mechanism account for many studies (14%). But even today, most DL-based models are built using CNNs because of their ability to capture spatial hierarchies and features in images automatically. These CNN architectures account for over 30% of the DL-based studies. Table A. 11 presents a complete list of studies utilizing each of these commonly used AI-based techniques.

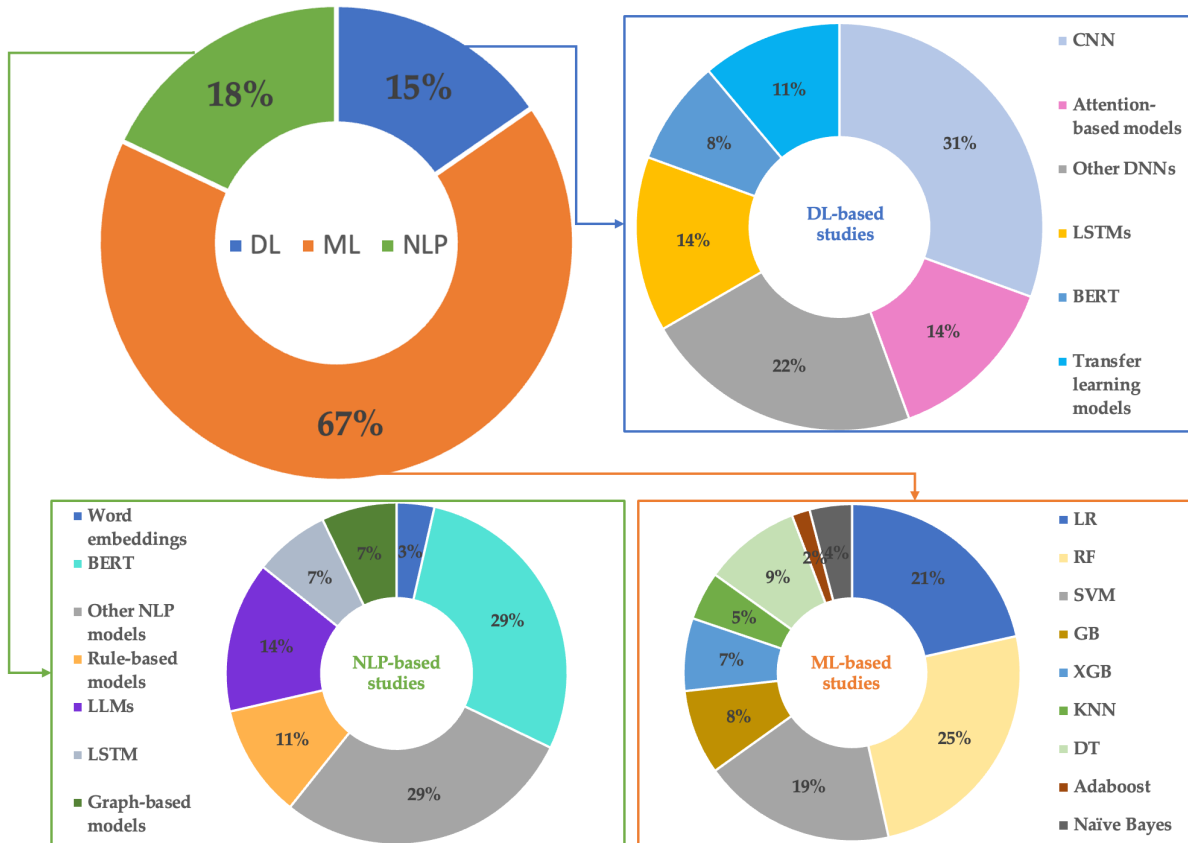


Figure 10: Papers utilizing ML, DL, and NLP techniques for suicide research.

Figure 11: Distribution of AI techniques used for suicide prediction and prevention tasks. delves into the specific AI techniques employed across the reviewed studies, broken down by prediction and prevention focus. With 42 studies utilizing the model, RF emerges as the most widely used technique for suicide prediction, followed by LR (26 studies) and SVM (24 studies). This popularity of RF could be due to its ability to capture complex interactions, handle high-dimensional data, and provide feature importance measures [162]. LR and SVM are well-established techniques for binary classification tasks like suicide risk prediction. In the context of suicide prevention, similar trends are observed, with RF, LR, and SVM being the most common ML techniques. This suggests that these methods are versatile and effective for prediction and prevention tasks. Among DL-based techniques, CNNs (16 studies) and LSTM networks (6 studies) are the most frequently used. CNNs are particularly adept at processing grid-like data such as images and matrices, while LSTMs are designed for sequential data like time series and text. It is to be noted that any ML-based paper, for example, could use both

RF and SVM techniques and perform a comparative study. This implies that while the numbers in **Error! Reference source not found.** would add up to 156 (the number of papers considered for review); the numbers in Figure 11: Distribution of AI techniques used for suicide prediction and prevention tasks. would exceed 156.

Finally, the use of multimodal information that incorporates numerous data sources, including clinical documentation, social media, and biometric information, gives a more holistic understanding of a person's suicide risk. These approaches leverage the strengths of different data modalities and AI techniques to capture complex patterns and interactions that may not be apparent when analyzing a single data source [163]. However, multimodal fusion also presents challenges, such as data heterogeneity, temporal alignment, and the need for advanced AI techniques to effectively combine and analyze diverse data types [164].

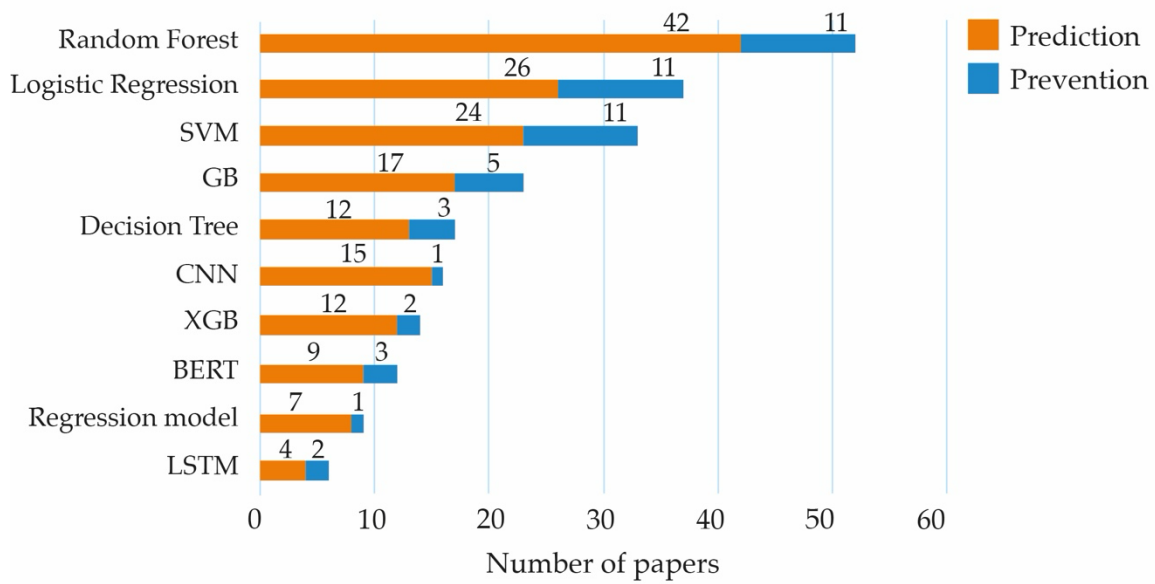


Figure 11: Distribution of AI techniques used for suicide prediction and prevention tasks.

3.5 Analysis of the best models obtained in each data modality

Our analysis in Sections 3.2 and 3.3 observed that RF and SVM were the most commonly used algorithms for predicting SI from clinical data. RF also performed exceedingly well in dealing with behavioral or psychological data. To assess digital and technology data or administrative and public data that are primarily textual, NLP-based methods like BERT and transformers proved to be the best. Finally, DL-based models and ensemble classifiers outperformed other models on multimodal data. To better understand the best models for each of these data modalities, we analyze the performance metrics of the reviewed studies in this section. For comparison, the three most commonly used metrics have been presented.

The three most commonly used metrics in the reviewed studies were accuracy, AUC, and F1 scores (Table 1: Number of prevention-based papers presenting a quantitative metric. Table 4: Number of prediction-based papers presenting a quantitative metric.). However, these metrics are not featured in more than 50% of the studies within most categories (datatypes), thereby

complicating comparing models within the same or across different data types. It is pertinent to note that a few studies report multiple metrics, which implies that the aggregate of values indicated in the first four rows of Table 1: Number of prevention-based papers presenting a quantitative metric. Table 4: Number of prediction-based papers presenting a quantitative metric. do not necessarily correspond to the total number of studies for a given column representing the data type. This variability highlights the challenge of establishing a consistent framework for comparative analysis due to the diverse methodological approaches adopted in the literature.

Table 1: Number of prevention-based papers presenting a quantitative metric.

Metric Presented	Clinical Data	Behavioral and Psychosocial Data	Digital and Technology Data	Administrative, Public Data	Multimodal Data
Accuracy	3	0	4	0	1
AUC	2	2	3	0	2
F1	2	0	3	2	2
Other/No Metric	5	2	12	4	5
Total Papers	12	4	19	6	10

Table 2: Maximum accuracies achieved on the suicide prevention task for each data modality.

Dataset Type	Maximum Accuracy	Reference
Clinical Data	0.933	Castillo-Sánchez et al. [23], 2022
Digital and Technology Data	0.991	Bello et al. [46], 2023
Multimodal Data	0.888	Liu et al. [65], 2019

Table 3: Maximum AUC achieved on suicide prevention task for each data modality.

Dataset Type	Maximum AUC	Reference
Clinical Data	0.986	Bejan et al. [29], 2022
Behavioural, Psychosocial Data	0.850	Kearns et al. [35], 2023
Digital and Technology Data	0.937	Iyer et al. [40], 2022
Administrative, Public Data	0.810	Haroz et al. [57], 2020
Multimodal Data	0.850	Kyron et al. [69], 2021

The methodologies and results of Nestsiarovich et al. [33] and Bejan et al. [29] offer insights into the evolving landscape of leveraging ML techniques to enhance suicide detection in clinical data. While Nestsiarovich et al. [33] focus on the comparative risk analysis of pharmacotherapies to infer self-harm events from structured health data, Bejan et al. [29] focus

on extracting nuanced, unstructured information regarding suicidal behavior through NLP. Bejan et al. [29] developed a scalable NLP method to improve SI prediction from clinical notes in a vast EHR repository comprising over 200 million notes. They extracted suicide query terms using word2vec and applied a weakly supervised model for labeling cases of suicidal outcomes, demonstrating high performance in the NLP validation of identified patients. This addresses the limitations posed by relying solely on diagnostic codes, which often result in the under-coding and under-reporting of suicide-related outcomes. Nestsiarovich et al. [33] and Bejan et al. [29] achieved AUC scores of 0.994 and 0.986, respectively.

Kearns et al. [35] present a practical risk calculator for identifying U.S. Army soldiers at high risk of suicidal behavior as they transition to civilian life. Utilizing an ML model incorporating only a short series of self-report survey questions. They achieved high performance (AUROC of 0.85) with only 17 key predictors significantly aiding in targeting interventions for those at greatest risk.

Bello et al. [46] and Iyer et al. [40], explored AI approaches to suicide risk assessment on digital and technological data. Bello et al. [46] focused on analyzing media texts, including press articles and tweets with suicide-related hashtags, to identify suicide-related topics using a neural model and achieved an AUC of 0.991. Iyer et al. [40] targeted the automatic classification of suicide risk in short segments of speech, from telephone calls to counseling services. They carefully preprocessed the voice signals, considered gender a crucial factor, and extracted biomarkers through Lasso regression, achieving a classification accuracy of 99.85% of speech frames.

Finally, Bello et al. [70] employed ML and Big Data techniques, including web scraping of news articles and tweets, to analyze media coverage of suicide in relation to COVID-19 using NLP. They extracted over 100K news items and 20M tweets, focusing on the presence of suicide and COVID-19 topics. They found a significant relationship between increased coverage of COVID-19 and subsequent rise in suicide-related media content. They also identified a shift in media reporting, with suicide-related news increasingly categorized under health topics since the pandemic outbreak, suggesting a growing awareness of suicide and mental health issues.

Table 4: Number of prediction-based papers presenting a quantitative metric.

Metric Presented	Clinical Data	Behavioural and Psychosocial Data	Digital and Technology Data	Administrative, Public Data	Multimodal Data
Accuracy	13	4	9	2	11
AUC	3	0	12	1	1
F1	8	9	2	7	6
Other/No Metric	4	4	4	2	11
Total Papers	25	17	25	10	28

Table 5: Maximum accuracy achieved on suicide prediction task for each data modality.

Dataset Type	Maximum Accuracy	Reference
Clinical Data	0.94	Cusick et al. [89], 2021
Behavioural, Psychosocial Data	0.91	Weller et al. [165], 2021
Digital and Technology Data	0.99	Talaat et al. [118], 2023
Administrative, Public Data	0.983	Lin et al. [129], 2022
Multimodal Data	0.95	Kim et al. [156], 2021

Table 6: Maximum AUC achieved on suicide prediction task for each data modality.

Dataset Type	Maximum AUC	Reference
Clinical Data	0.9	Song et al. [74], 2022
Behavioural, Psychosocial Data	0.92	Li et al. [157], 2023
Digital and Technology Data	0.925	Shen et al. [5], 2020
Administrative, Public Data	0.987	Rahman et al. [132], 2022
Multimodal Data	0.94	Miranda et al. [152], 2024

Cusick et al. [89] explored the use of CNNs for detecting current SI from unstructured clinical notes within EHRs to produce one of the best works (94% accuracy) in suicide detection using clinical data. Utilizing a rule-based approach, they labeled a large dataset of clinical notes drawn from a cohort of 600 patients at risk for SI. They also identified additional patients and encounters not previously captured by structured diagnosis codes. When tested on a random sample of 5,000 clinical notes, they were able to identify cases of current SI with high accuracy, suggesting its potential utility in enhancing suicide prevention efforts and research through large-scale screening of clinical documentation.

Li et al. [157], and Weller et al. [165], employed ML techniques on behavioral and psychosocial data to identify predictors of STB. Li et al. [157] focused on detecting SI among individuals with depression using NLP to analyze language features in clinical interviews. They translated verbal responses into psychological categories to identify SI based on language patterns, achieving AUC scores between 0.76 and 0.92. While Li et al. 's [157] the approach introduced a novel method for suicide risk detection in clinical settings without relying on explicit patient disclosure, Weller et al. [165] emphasized the utility of ML in understanding and predicting adolescent STB at a population level. They targeted a broader population, analyzing survey responses from over 179,000 high school students to predict STB using survey data and demographic information. Focusing on factors related to familial life, drug consumption, demographics, and peer acceptance, they were able to identify SI with an accuracy of 91%.

Talaat et al. [118] focused on the development of an enhanced recognition system called PERS for identifying personality traits and emotional states from textual data on social media. PERS

incorporated data acquisition, preprocessing, personality recognition, and emotion recognition modules, utilizing ML algorithms for multiclass classification. Gray Wolf Optimization (GWO) for hyperparameter optimization and Group GWO for feature selection achieved accuracy rates of 99.99% in personality detection using the RF and 88.06% in emotion recognition using DTs.

Wang et al. [133] addressed the global public health challenge of suicide by leveraging NLP to identify SDoH-related social risks from narratives of death investigations. Utilizing data from the NVDRS, they adapted a suicide-specific SDoH ontology (Suicide-SDoHO) and developed a transformer-based model. Unlike traditional structured data coding, they were able to identify circumstances and crises related to SDoH in the narratives. They achieved AUC scores of 0.966 and 0.942 for classifying circumstances and crises, respectively, establishing the ability of NLP to classify SDoH-related social risks from unstructured data.

Finally, Miranda et al. [152] focused on predicting adverse events in patients with co-morbid PTSD and alcohol use disorder using EHRs and a DL model called DeepBiomarker2 to achieve an AUROC score of 0.94. They used multimodal information, including lab tests, medications, co-morbidities, individual and neighborhood-level SDoH, psychotherapy, and veteran status, to predict adverse events such as opioid use disorder, suicide-related events, depression, and death within three months post-diagnosis. The study highlighted medications like vilazodone, dronabinol, and tenofovir for potentially reducing adverse event risk. It also indicated the influence of SDoH parameters, where, for instance, cognitive behavioral therapy and trauma-focused psychotherapy were linked to lowered risk, whereas veteran status and other factors might cause increased risk.

3.6 Public datasets available for suicide prediction

A list of all publicly available datasets related to suicide prediction or prevention is available in Table 7: Summary of publicly available datasets.. Numerous datasets gathered from social media sites like Twitter [166] and Reddit [167, 168, 169] are being used to develop systems detecting SI. These datasets are rich sources of user-generated content. Sinha et al. [166] developed a manually annotated dataset from Twitter by employing a lexicon of suicidal phrases and social engagement data from real-time and historical tweets.

Compared to other social media sites, Reddit offers a unique advantage. It is a platform that allows for extensive linguistic expression due to its generous character limit. Gaur et al. [167] compiled and annotated a suicide risk severity assessment dataset from Reddit drawn from the SuicideWatch subreddit and annotated by four practicing clinical psychiatrists to ensure accuracy and reliability. The University of Maryland Reddit Suicidality Dataset (UMD) [168, 169] is another dataset drawn from the same subreddit and annotated by experts. The collection spans several years and includes details like the content, location, and timing of each post and comment. By providing annotated datasets [167, 168, 169], they address common issues with social media data, such as the accuracy of ground truth labels, which can often be unreliable [43]. The annotation process includes several measures to enhance reliability: initial selection of at-risk users based on social media activity, the use of mental

health experts and crowdsourced annotators, and the implementation of consensus mechanisms, such as having multiple annotators for each user, to ensure accuracy.

Compared to platforms like Reddit and Twitter, Instagram is less commonly used as a data source for studying SI. Lekkas et al. [104] collected data from a study examining the relationships between acute suicidality, language usage, and activity on Instagram. This dataset focuses on German adolescents and includes instances of non-suicidal self-injury identified on Instagram. Public user data and post content from subjects were collected for four weeks following a personal interview conducted via Instagram messenger with individuals who reported a lifetime history of SI. Despite these datasets gaining popularity recently, they do not provide access to and only include users' tweeting or posting activities for brief periods. They fail to capture users' full tweeting and posting histories. To the best of our knowledge, no publicly available depression dataset exists that comprehensively collects users' entire tweeting and posting histories, posing a considerable obstacle to research in depression and suicide detection.

The Dealer's Record of Sale (DROS) database [128] offers a unique perspective on suicide research. It contains records of handgun transactions for approximately 1.95 million individuals in California. It can be used to identify individuals at a higher risk of firearm suicide and to explore critical features influencing these predictions.

Obtaining publicly available data for questionnaire responses is extremely hard. NVDRS [137] encompasses 267,804 suicide death incidents across all 50 U.S. states, Puerto Rico, and the District of Columbia, spanning from 2003 to 2019. Each record includes narratives describing the suicide death from two different perspectives: one based on medical examiner records and another based on law enforcement reports. This dataset thus provides in-depth SDoH information concerning violent deaths. The NHIRD dataset [170], obtained from the National Health Insurance program, is a nationwide collection of data about 23 million people in Taiwan. It contains complete information about clinical visits, including prescription details and diagnostic codes from suicide attempters with at least one psychiatric visit.

Roden et al. [171] were one of the first to develop a clinical research repository called the Vanderbilt Research Derivative. It is a DNA biobank linked to phenotypic data from an electronic medical record system.

Zingmond [172] collected discharge data from the California Office of Statewide Health Planning and Development on all deliberate self-harm visits from January 2009 to December 2011 by individuals aged ten or more years and includes information on their date, cause, and manner of death. Similarly, the UCLA-xDR repository [173] contains 10 years of outpatient and inpatient EHR data collected from 2006 through 2016 from 2 academic medical hospitals within the UCLA Health System. Another similar dataset is the NYC-CDRN [32] which contains outpatient and inpatient EHR data collected between 2009 and 2017 from 7 health systems across New York City. Finally, the GEO databases, GSE102556 and GSE101521 [174], constitute two mRNA expression profiling datasets collected by experienced psychiatrists and care providers. Along with information on gene expression, the database also consists of information about the patient's demographic data and suicide information.

Finally, we briefly discuss a couple of depression-related public databases. Cai et al. [175] constructed and published the first large-scale depression dataset. With information about the tweeting histories of 23237 users, it is a well-labelled social media-based dataset from the most popular Chinese social network platform, Sina Weibo. The Longitudinal Study of Australian Children (LSAC) [176] is a nationally representative longitudinal study that includes self-completed questionnaires, face-to-face interviews, and computer-assisted interviews, with participants including the children studied, their parents, caregivers, and teachers of infants and children.

Table 7: Summary of publicly available datasets.

Dataset Name	Source Social Media Platform	Description	Size (Records/Users)	Labels/Features	Reference
Sinha et al. Twitter Dataset	Twitter	Manually annotated with suicidal phrases and social engagement data.	34,306 tweets	Suicidal, Non-Suicidal	[166]
Gaur et al. Reddit Dataset	Reddit	Annotated for suicide risk severity by clinical psychiatrists from SuicideWatch subreddit.	500 posts	SI, Suicidal Behavior, Actual Attempt, Suicide Indicator, Supportive	[167]
University of Maryland Reddit Suicidality Dataset (UMD)	Reddit	Over 100,000 posts/comments annotated to assess suicide risk level.	>100,000 posts/comments	Risk levels of none, low, moderate, and severe.	[168, 169]
Instagram German Adolescents Dataset	Instagram	Focused on acute suicidality, language usage, and activity among German adolescents.	Interviews of over 50 users via Instagram.	Presence/absence of acute SI	[104]
Dealer's Record of Sale	N/A	Contains records of handgun transactions used	Over 5 million records	Detailed transaction details that	[128]

(DROS Database)		to identify firearm suicide risk.		include purchaser and seller identifiers along with the firearm information.	
National Violent Death Reporting System (NVDRS)	N/A	Comprehensive suicide death incident data across the U.S.	267,804 incidents	Two death narratives - summary of events from the perspective of the Coroner or Medical Examiner record and Law Enforcement report.	[137]
UCLA-xDR and NYC-CDRN Datasets	N/A	Inpatient and outpatient EHR data from academic medical hospitals and health systems in NYC.	Data on ~12 million patients	Outpatient and inpatient EHR data, along with suicide information.	[173, 31]
GEO Databases GSE10255 6 and GSE10152 1	N/A	mRNA expression profiling datasets, including demographic and suicide information.	107 human subjects	Gene expression, demographic data, and patient's suicide information.	[174]
Depression Dataset on Sina Weibo	Sina Weibo	A social media-based dataset labeled for depressive and non-depressive users from China.	23,237 users	Depressed, non-depressed.	[175]
Longitudinal Study of Australian	N/A	Nationally representative longitudinal study data collected biennially	5,107 infants between the ages of 3 and 19 months, and 4,983 children.	Self-completed questionnaires and interviews given by the children	[176]

Children (LSAC)		excluding COVID-19 impacted years.		studied, their parents, caregivers, and teachers.	
Vanderbilt Research Derivative	N/A	DNA biobank linked to electronic medical records focusing on clinical research	384 patients.	N/A ICD-9 codes, structured laboratory data, unstructured patient data, allele identification, medications	[171]
OHSPD Dataset	N/A	Discharge data from the California Office of Statewide Health Planning and Development on all deliberate self-harm visits.	Information about 1,59,713 women.	Features including the manner, cause, and date of death.	[175]
NHIRD Dataset	N/A	Nationwide dataset from Taiwan's National Health Insurance (clinical information, prescription details, etc.), focusing on psychiatric patients and suicide attempters.	23 million people.	Many-time/single-time suicide attempters.	[170]

4. Benefits, limitations and future directions of using AI for suicide prediction and prevention

The application of ML, DL, and NLP offers several significant benefits in suicide prevention and prediction. They allow us to discover patterns that may not be obvious to humans. Some of the key benefits include:

- *Early identification of at-risk individuals:* AI algorithms can analyze huge amounts of data from diverse sources to identify individuals who may be at high risk of suicidal behavior.

This can facilitate timely interventions and support by detecting subtle warning signs and risk factors early on, especially on various social media platforms.

- *Improved accuracy and efficiency:* ML and DL models have demonstrated high levels of accuracy in predicting STB, often outperforming traditional risk assessment tools. They have also achieved this by being highly scalable. They can process large volumes of data quickly and consistently, reducing the burden on mental health professionals.
- *Personalized risk assessment:* AI techniques can help create personalized risk profiles by integrating multiple data sources and considering risk factors specific to an individual, such as demographic characteristics, medical history, and psychosocial stressors. This can lead to more targeted and effective prevention strategies.
- *Real-time monitoring and intervention:* AI-powered tools can continuously monitor an individual's digital footprint, such as their social media activity or smartphone usage patterns, for signs of distress or suicidal intent. When high-risk patterns are detected, these systems can trigger real-time alerts and facilitate rapid intervention by mental health professionals or crisis support services.
- *Enhancing understanding of risk factors and protective factors:* AI methods can help researchers identify novel risk factors and protective factors for suicide by uncovering complex relationships and interactions within large datasets. This understanding can inform the development of more effective prevention programs and public health policies.
- *Efficient resource allocation:* By prioritizing cases based on risk levels and optimizing resource allocation within mental health services, psychiatrists can ensure that high-risk individuals receive the support and intervention they need promptly.

Despite ML, DL, and NLP techniques showing promising results, several challenges remain. One of the main challenges is the limited interpretability and explainability of AI models, particularly in the context of high-stakes mental health decisions. Ensuring transparency and building trust among clinicians and patients is crucial for the successful implementation of AI-driven interventions.

Another challenge lies in the ethical considerations surrounding the use of sensitive personal data, such as EHRs and social media posts, for suicide prediction and prevention. Developing robust data governance frameworks and engaging in ongoing dialogue with stakeholders could help to address these ethical concerns. Utilizing sensitive clinical data for training ML models raises privacy and confidentiality concerns [178]. ML-based strategies using data from smartphones, wearables, and smart home devices enhance short-term, personalized prediction of STBs in adolescents, detecting risk factors like emotional distress and sleep issues. Ensuring compliance with regulations such as HIPAA [179] and maintaining patient confidentiality while extracting valuable insights from the data presents a significant challenge. Incorporating ML-driven tools and predictions into everyday clinical practice can also be challenging. Biases present in the data, such as the underrepresentation of certain demographic groups or clinical characteristics, can lead to inaccurate predictions and potentially exacerbate disparities in mental health care [22]. The complexity of some ML and

DL models can make it difficult to understand how they arrive at specific predictions or recommendations [180]. It is crucial to remember that predictive models should complement, not replace, clinical judgment and human interaction. Over-reliance on algorithmic predictions without considering individual patient nuances and the clinician's expertise could harm patient care. There is a risk of clinicians becoming de-skilled and not using their instinct and clinical judgment, which may ultimately endanger the workforce developing capable and confident clinicians in clinical assessments of suicide risk. Psychiatrists and mental health clinicians need to understand how to interpret and utilize these technologies effectively while considering their limitations and integrating them with existing treatment approaches. Navigating these challenges requires collaboration between psychiatrists, data scientists, ethicists, and policymakers to develop responsible and effective approaches to integrating ML, DL, and NLP technologies into mental health practice.

Furthermore, the generalizability and fairness of AI models across diverse populations and settings remain important challenges. Many existing studies have focused on specific subgroups or regions, and there is a need to validate and adapt these models to ensure their effectiveness in different cultural and societal contexts.

To advance the field of AI-driven suicide research, future studies (Figure 12) should focus on several key directions:

1. Connecting prediction and prevention: While current research has emphasized prediction, it is essential to recognize the interrelationship between prediction and prevention. Accurate predictive models can enable more targeted and personalized prevention strategies. Future studies should aim to develop a greater balance and synergy between prediction and prevention, leveraging AI techniques to optimize limited mental healthcare resources and ultimately reduce suicide rates.
2. Harnessing multimodal data: Integrating multiple data sources, such as clinical records, social media, wearables, and biometric data, can provide a more comprehensive understanding of suicide risk factors and behaviors. Future research should continue to explore advanced AI techniques for multimodal data integration while prioritizing responsible and privacy-preserving practices.
3. Addressing understudied and underserved populations: There is a critical need to expand suicide research to include diverse and underrepresented populations, such as racial and ethnic minorities and people in low and middle-income countries. Future studies should focus on understanding the unique risk factors, protective factors, and cultural considerations for these populations and developing tailored AI-driven interventions to address their specific needs.
4. Advancing DL applications: As AI technologies evolve, DL techniques are expected to play an increasingly important role in suicide research, particularly with the growing availability of multimodal and digital data. However, researchers must carefully navigate the trade-offs between predictive performance and interpretability and the ethical implications of using complex, black-box models in mental health contexts.

5. Fostering interdisciplinary collaboration: Suicide research benefits from the collaboration of experts from diverse fields, including psychology, psychiatry, computer science, and public health. Encouraging interdisciplinary teams and promoting knowledge sharing across domains can lead to innovative solutions and more comprehensive approaches to suicide prevention and prediction.
6. Emphasizing real-world implementation and impact: As AI-driven suicide research advances, it is essential to focus on translating these findings into real-world clinical practice and public health initiatives. Future studies should prioritize the development and evaluation of implementable, scalable, and culturally sensitive AI-based interventions, considering acceptability, feasibility, and cost-effectiveness.

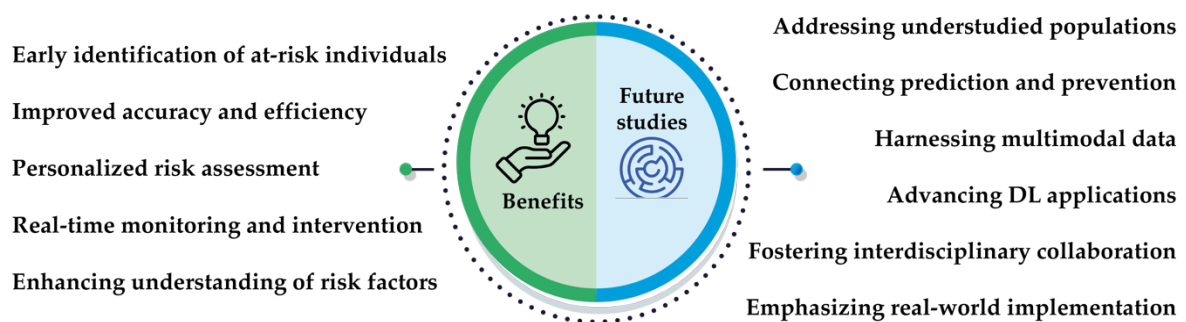


Figure 12: Benefits and future studies for suicide prediction and prevention.

Moreover, the limitations inherent in the current body of research highlight the critical need for adopting standardized quantitative metrics within ML-based suicide detection. First, researchers need to include at least one quantitative metric, as this allows for clear evaluation and comparison of study results. Without these metrics, judging the effectiveness and usefulness of proposed models becomes unclear, making it hard to advance toward reliable detection methods. Furthermore, the creation of a uniform metric, adopted by all researchers, is crucial. A standard measure would not only make results easier to reproduce, increasing the reliability of ML applications in detecting suicide, but also simplify the process of verifying these results. This consistency would enable easier comparison across different studies, creating a unified body of research that helps identify the most effective methods faster. Ultimately, the adoption of a common metric would serve as a linchpin in the development of more robust and universally applicable suicide detection systems, marking a significant stride toward the optimization of intervention strategies and the reduction of suicide rates.

5. Conclusion

Suicide prevention and prediction are major global public health priorities. Recent advances in AI techniques, including ML, DL and NLP, show promise for advancing suicide prevention and risk assessment efforts. This systematic review comprehensively summarizes the current literature on AI applications for suicide prevention and prediction. The most common data sources were social media posts, clinical records, self-report questionnaires, and multimodal

data combining diverse digital traces. Classifiers like random forests and support vector machines produced high performance in detecting SI from all of these data sources. Further, NLP techniques like BERT are highly efficient in learning representations from unstructured textual data that best identify SI and emotional distress, especially on platforms like Twitter and Reddit. The reviewed studies demonstrate the potential of AI to enable more accurate, scalable and personalized suicide risk screening and prevention initiatives. We further highlight the need and potential of using multimodal data that combines information from different sources to predict SI at early stages accurately. ML and DL methods show particular promise for integrating multimodal data, including social media, smartphones, wearables and medical records, to uncover insights and patterns predictive of suicide risk. The review also highlights several challenges and limitations that must be addressed to realize the full potential of AI in suicide prevention and prediction. These include concerns around data privacy and security, algorithmic bias and fairness, model interpretability and explainability, and the need for rigorous evaluation and validation of AI-driven interventions in real-world settings. Future research should focus on developing more integrated approaches that combine accurate prediction with effective prevention strategies, harnessing the power of multimodal data, and expanding the scope of research to include diverse and underrepresented populations.

6. Funding Statement

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

References

- [1] N. B. Allen, B. W. Nelson, D. Brent, and R. P. Auerbach, "Short-term prediction of suicidal thoughts and behaviors in adolescents: Can recent developments in technology and computational science provide a breakthrough?," *J Affect Disord*, vol. 250, pp. 163–169, May 2019, doi: 10.1016/J.JAD.2019.03.044.
- [2] Z. Xu, Q. Zhang, and P. S. F. Yip, "Predicting post-discharge self-harm incidents using disease comorbidity networks: A retrospective machine learning study," *J Affect Disord*, vol. 277, pp. 402–409, Dec. 2020, doi: 10.1016/J.JAD.2020.08.044.
- [3] S. Hong *et al.*, "Identification of suicidality in adolescent major depressive disorder patients using sMRI: A machine learning approach," *J Affect Disord*, vol. 280, pp. 72–76, Feb. 2021, doi: 10.1016/j.jad.2020.10.077.
- [4] Y. Liang, X. Zheng, and D. D. Zeng, "A survey on big data-driven digital phenotyping of mental health," *Information Fusion*, vol. 52, pp. 290–307, Dec. 2019, doi: 10.1016/J.INFFUS.2019.04.001.

- [5] Y. Shen *et al.*, "Detecting risk of suicide attempts among Chinese medical college students using a machine learning algorithm," *J Affect Disord*, vol. 273, pp. 18–23, Aug. 2020, doi: 10.1016/j.jad.2020.04.057.
- [6] V. Menon, K. Subramanian, N. Selvakumar, and S. Kattimani, "Suicide prevention strategies: An overview of current evidence and best practice elements," *International Journal of Advanced Medical and Health Research*, vol. 5, no. 2, p. 43, 2018, doi: 10.4103/IJAMR.IJAMR_71_18.
- [7] J. Robinson *et al.*, "Development of a self-harm monitoring system for victoria," *Int J Environ Res Public Health*, vol. 17, no. 24, pp. 1–12, Dec. 2020, doi: 10.3390/ijerph17249385.
- [8] A. Roy, K. Nikolitch, R. McGinn, S. Jinah, W. Klement, and Z. A. Kaminsky, "ARTICLE OPEN A machine learning approach predicts future risk to suicidal ideation from social media data", doi: 10.1038/s41746-020-0287-6.
- [9] C. G. Walsh, J. D. Ribeiro, and J. C. Franklin, "Predicting Risk of Suicide Attempts Over Time Through Machine Learning," <https://doi.org/10.1177/2167702617691560>, vol. 5, no. 3, pp. 457–469, Apr. 2017, doi: 10.1177/2167702617691560.
- [10] D. Kim, S. Cogill, and S. Yang, "11 An Interpretable Deep Learning Model for the Prevention of Self-Harm and Suicide," *Ann Emerg Med*, vol. 74, no. 4, p. S6, Oct. 2019, doi: 10.1016/j.annemergmed.2019.08.014.
- [11] N. J. Carson *et al.*, "Identification of suicidal behavior among psychiatrically hospitalized adolescents using natural language processing and machine learning of electronic health records," *PLoS One*, vol. 14, no. 2, Feb. 2019, doi: 10.1371/journal.pone.0211116.
- [12] R. Shao *et al.*, "Multimodal Neural Evidence on the Corticostriatal Underpinning of Suicidality in Late-Life Depression," *Biol Psychiatry Cogn Neurosci Neuroimaging*, vol. 7, no. 9, pp. 905–915, Sep. 2022, doi: 10.1016/j.bpsc.2021.11.011.
- [13] L. Cao, H. Zhang, and L. Feng, "Building and Using Personal Knowledge Graph to Improve Suicidal Ideation Detection on Social Media," *IEEE Trans Multimedia*, vol. 24, pp. 87–102, 2022, doi: 10.1109/TMM.2020.3046867.
- [14] A. Burnett *et al.*, "Machine learning algorithms to classify self-harm behaviours in New South Wales Ambulance electronic medical records: A retrospective study," *Int J Med Inform*, vol. 161, May 2022, doi: 10.1016/j.ijmedinf.2022.104734.
- [15] T. Zhang, A. M. Schoene, and S. Ananiadou, "Automatic identification of suicide notes with a transformer-based deep learning model," *Internet Interv*, vol. 25, Sep. 2021, doi: 10.1016/j.invent.2021.100422.
- [16] H. Zhu, X. Xia, J. Yao, H. Fan, Q. Wang, and Q. Gao, "Comparisons of different classification algorithms while using text mining to screen psychiatric inpatients with suicidal behaviors," *J Psychiatr Res*, vol. 124, pp. 123–130, May 2020, doi: 10.1016/j.jpsychires.2020.02.019.
- [17] A. Arowosegbe and T. Oyelade, "Application of Natural Language Processing (NLP) in Detecting and Preventing Suicide Ideation: A Systematic Review," *International Journal of Environmental Research and Public Health*, vol. 20, no. 2. MDPI, Jan. 01, 2023. doi: 10.3390/ijerph20021514.

- [18] T. Zhang, K. Yang, S. Ji, and S. Ananiadou, "Emotion fusion for mental illness detection from social media: A survey," *Information Fusion*, vol. 92, pp. 231–246, Apr. 2023, doi: 10.1016/J.INFFUS.2022.11.031.
- [19] Heckler, W. F., de Carvalho, J. V., & Barbosa, J. L. V. (2022). Machine learning for suicidal ideation identification: A systematic literature review. *Computers in Human Behavior*, 128, 107095. Doi: <https://doi.org/10.1016/j.chb.2021.107095>
- [20] A. Thieme, D. Belgrave, and G. Doherty, "Machine Learning in Mental Health: A Systematic Review of the HCI Literature to Support the Development of Effective and Implementable ML Systems," *ACM Transactions on Computer-Human Interaction*, vol. 27, no. 5, p. 34, 2020, doi: 10.1145/3398069.
- [21] G. Berkelmans, L. Schwaren, S. Bhulai, R. van der Mei, and R. Gilissen, "Identifying populations at ultra-high risk of suicide using a novel machine learning method," *Compr Psychiatry*, vol. 123, May 2023, doi: 10.1016/j.comppsy.2023.152380.
- [22] R. A. Bernert, A. M. Hilberg, R. Melia, J. P. Kim, N. H. Shah, and F. Abnoui, "Artificial intelligence and suicide prevention: A systematic review of machine learning investigations," *Int J Environ Res Public Health*, vol. 17, no. 16, pp. 1–25, Aug. 2020, doi: 10.3390/IJERPH17165929.
- [23] M. A. Franco-Martín, J. L. Muñoz-Sánchez, B. Sainz-de-Abajo, G. Castillo-Sánchez, S. Hamrioui, and I. de la Torre-Díez, "A Systematic Literature Review of Technologies for Suicidal Behavior Prevention," *J Med Syst*, vol. 42, no. 4, Apr. 2018, doi: 10.1007/s10916-018-0926-5.
- [24] D. Hopkins, Danielle Hopkins, David J. Hallford, and L. B. Sander, "Structured data vs. unstructured data in machine learning prediction models for suicidal behaviors: A systematic review and meta-analysis."
- [25] A. Lejeune *et al.*, "Artificial intelligence and suicide prevention: A systematic review," *European Psychiatry*, vol. 65, no. 1, 2022, doi: 10.1192/j.eurpsy.2022.8.
- [26] World Health Organization. (2021). Suicide worldwide in 2019: global health estimates.
- [27] K. F. McManus *et al.*, "Deploying a national clinical text processing infrastructure," *Journal of the American Medical Informatics Association*, vol. 31, no. 3, pp. 727–731, Mar. 2024, doi: 10.1093/jamia/ocad249.
- [28] S. J. Goldman-Mellor, H. S. Bhat, M. H. Allen, and M. Schoenbaum, "Suicide Risk Among Hospitalized Versus Discharged Deliberate Self-Harm Patients: Generalized Random Forest Analysis Using a Large Claims Data Set," *Am J Prev Med*, vol. 62, no. 4, pp. 558–566, Apr. 2022, doi: 10.1016/j.amepre.2021.08.028.
- [29] C. A. Bejan *et al.*, "Improving ascertainment of suicidal ideation and suicide attempt with natural language processing," *Sci Rep*, vol. 12, no. 1, Dec. 2022, doi: 10.1038/s41598-022-19358-3.
- [30] M. Cusick *et al.*, "Portability of natural language processing methods to detect suicidality from clinical text in US and UK electronic health records," *J Affect Disord Rep*, vol. 10, Dec. 2022, doi: 10.1016/j.jadr.2022.100430.

- [31] J. B. Edgcomb, R. Thiruvalluru, J. Pathak, and J. O. Brooks Iii, "Machine Learning to Differentiate Risk of Suicide Attempt and Self-harm After General Medical Hospitalization of Women With Mental Illness," 2021. [Online]. Available: www.lww-medicalcare.com.
- [32] J. Gong, G. E. Simon, and S. Liu, "Machine learning discovery of longitudinal patterns of depression and suicidal ideation," *PLoS One*, vol. 14, no. 9, Sep. 2019, doi: 10.1371/journal.pone.0222665.
- [33] A. Nestsiarovich *et al.*, "Using machine learning imputed outcomes to assess drug-dependent risk of self-harm in patients with bipolar disorder: A comparative effectiveness study," *JMIR Ment Health*, vol. 8, no. 4, Apr. 2021, doi: 10.2196/24522.
- [34] G. Castillo-Sánchez, M. J. Acosta, B. Garcia-Zapirain, I. De la Torre, and M. Franco-Martín, "Application of Machine Learning Techniques to Help in the Feature Selection Related to Hospital Readmissions of Suicidal Behavior," *Int J Ment Health Addict*, vol. 22, no. 1, pp. 216–237, Feb. 2024, doi: 10.1007/s11469-022-00868-0.
- [35] J. C. Kearns *et al.*, "A practical risk calculator for suicidal behavior among transitioning U.S. Army soldiers: Results from the Study to Assess Risk and Resilience in Servicemembers-Longitudinal Study (STARRS-LS)," *Psychol Med*, vol. 151, no. 7, Mar. 2023, doi: 10.1017/S0033291723000491.
- [36] E. F. Haghighi, B. Laeng, and N. Czajkowski, "Are false positives in suicide classification models a risk group? Evidence for 'true alarms' in a population-representative longitudinal study of Norwegian adolescents," *Front Psychol*, vol. 14, 2023, doi: 10.3389/fpsyg.2023.1216483.
- [37] P. Rus Prelog, T. Matic, P. Pregelj, and A. Sadikov, "A pilot predictive model based on COVID-19 data to assess suicidal ideation indirectly," *J Psychiatr Res*, vol. 163, pp. 318–324, Jul. 2023, doi: 10.1016/j.jpsychires.2023.05.008.
- [38] C. Biernesser, J. Zelazny, D. Brent, T. Bear, C. Mair, and J. Trauth, "Automated Monitoring of Suicidal Adolescents' Digital Media Use: Qualitative Study Exploring Acceptability within Clinical Care," *JMIR Ment Health*, vol. 8, no. 9, Sep. 2021, doi: 10.2196/26031.
- [38] V. Purushothaman, J. Li, and T. K. Mackey, "Detecting Suicide and Self-Harm Discussions Among Opioid Substance Users on Instagram Using Machine Learning," *Front Psychiatry*, vol. 12, May 2021, doi: 10.3389/fpsyg.2021.551296.
- [40] R. Iyer, M. Nedeljkovic, and D. Meyer, "Using Voice Biomarkers to Classify Suicide Risk in Adult Telehealth Callers: Retrospective Observational Study," *JMIR Ment Health*, vol. 9, no. 8, Aug. 2022, doi: 10.2196/39807.
- [41] R. Iyer, M. Nedeljkovic, and D. Meyer, "Using Vocal Characteristics to Classify Psychological Distress in Adult Helpline Callers: Retrospective Observational Study," *JMIR Form Res*, vol. 6, no. 12, Dec. 2022, doi: 10.2196/42249.
- [42] M. Chau, T. M. H. Li, P. W. C. Wong, J. J. Xu, P. S. F. Yip, and H. Chen, "Finding people with emotional distress in online social media: A design combining machine learning and rule-BASED classification," *MIS Q*, vol. 44, no. 2, pp. 933–956, Jun. 2020, doi: 10.25300/MISQ/2020/14110.
- [43] C. Lao, J. Lane, and H. Suominen, "Analyzing Suicide Risk from Linguistic Features in Social Media: Evaluation Study," *JMIR Form Res*, vol. 6, no. 8, Aug. 2022, doi: 10.2196/35563.

- [44] T. Niederkrotenthaler *et al.*, "Association of 7 million+ tweets featuring suicide-related content with daily calls to the Suicide Prevention Lifeline and with suicides, United States, 2016–2018," *Australian and New Zealand Journal of Psychiatry*, vol. 57, no. 7, pp. 994–1003, Jul. 2023, doi: 10.1177/00048674221126649.
- [45] R. W. Acuña Caicedo, J. M. Gómez Soriano, and H. A. Melgar Sasieta, "Assessment of supervised classifiers for the task of detecting messages with suicidal ideation," *Heliyon*, vol. 6, no. 8, Aug. 2020, doi: 10.1016/j.heliyon.2020.e04412.
- [46] H. J. Bello, N. Palomar-Ciria, E. Baca-García, and C. Lozano, "Suicide Classification for News Media Using Convolutional Neural Networks," *Health Commun*, vol. 38, no. 10, pp. 2178–2187, 2023, doi: 10.1080/10410236.2022.2058686.
- [47] B. X. Yang *et al.*, "A Suicide Monitoring and Crisis Intervention Strategy Based on Knowledge Graph Technology for 'Tree Hole' Microblog Users in China," *Front Psychol*, vol. 12, Oct. 2021, doi: 10.3389/fpsyg.2021.674481.
- [48] S. Wang *et al.*, "Public Surveillance of Social Media for Suicide Using Advanced Deep Learning Models in Japan: Time Series Study From 2012 to 2022," *J Med Internet Res*, vol. 25, 2023, doi: 10.2196/47225.
- [49] I. Levkovich and Z. Elyoseph, "Suicide Risk Assessments Through the Eyes of ChatGPT-3.5 Versus ChatGPT-4: Vignette Study," *JMIR Ment Health*, vol. 10, no. 1, Jan. 2023, doi: 10.2196/51232.
- [50] N. Deas *et al.*, "I just want to matter: Examining the role of anti-mattering in online suicide support communities using natural language processing," *Comput Human Behav*, vol. 139, Feb. 2023, doi: 10.1016/j.chb.2022.107499.
- [51] S. Salmi, S. Mérelle, R. Gilissen, R. van der Mei, and S. Bhulai, "Detecting changes in help seeker conversations on a suicide prevention helpline during the COVID- 19 pandemic: in-depth analysis using encoder representations from transformers," *BMC Public Health*, vol. 22, no. 1, Dec. 2022, doi: 10.1186/s12889-022-12926-2.
- [52] M. A. Moreno, A. D. Gower, H. Brittain, and T. Vaillancourt, "Applying Natural Language Processing to Evaluate News Media Coverage of Bullying and Cyberbullying," *Prevention Science*, vol. 20, no. 8, pp. 1274–1283, Nov. 2019, doi: 10.1007/s11121-019-01029-x.
- [53] G. Fu *et al.*, "Distant supervision for mental health management in social media: Suicide risk classification system development study," *J Med Internet Res*, vol. 23, no. 8, Aug. 2021, doi: 10.2196/26119.
- [54] S. K. Kar, S. Tripathy, and S. M. Y. Arafat, "Voice assists (Alexa) and suicidal behavior: A potential area of suicide prevention," *Asian J Psychiatr*, vol. 72, p. 103117, Jun. 2022, doi: 10.1016/J.AJP.2022.103117.
- [55] R. Kumar, S. Mukherjee, T. M. Choi, and L. Dhamotharan, "Mining voices from self-expressed messages on social-media: Diagnostics of mental distress during COVID-19," *Decis Support Syst*, vol. 162, Nov. 2022, doi: 10.1016/j.dss.2022.113792.
- [56] A. M. Lake *et al.*, "Lifeline Crisis Chat: Coding form development and findings on chatters' risk status and counselor behaviors," *Suicide Life Threat Behav*, vol. 52, no. 3, pp. 452–466, Jun. 2022, doi: 10.1111/sltb.12835.

- [57] E. E. Haroz, C. G. Walsh, N. Goklish, M. F. Cwik, V. O’Keefe, and A. Barlow, “Reaching Those at Highest Risk for Suicide: Development of a Model Using Machine Learning Methods for use With Native American Communities,” *Suicide Life Threat Behav*, vol. 50, no. 2, pp. 422–436, Apr. 2020, doi: 10.1111/sltb.12598.
- [58] K. Turner *et al.*, “Implementing a systems approach to suicide prevention in a mental health service using the Zero Suicide Framework,” *Australian and New Zealand Journal of Psychiatry*, vol. 55, no. 3. SAGE Publications Inc., pp. 241–253, Mar. 01, 2021. doi: 10.1177/0004867420971698.
- [59] J. L. Wang *et al.*, “A case-control study on predicting population risk of suicide using health administrative data: a research protocol,” *BMJ Open*, vol. 13, no. 2, Feb. 2023, doi: 10.1136/bmjopen-2022-066423.
- [60] J. M. Kafka, K. E. Moracco, B. W. Pence, P. J. Trangenstein, M. D. Fliss, and L. McNaughton Reyes, “Intimate partner violence and suicide mortality: A cross-sectional study using machine learning and natural language processing of suicide data from 43 states,” *Injury Prevention*, Oct. 2023, doi: 10.1136/ip-2023-044976.
- [61] J. M. Kafka, M. D. Fliss, P. J. Trangenstein, L. McNaughton Reyes, B. W. Pence, and K. E. Moracco, “Detecting intimate partner violence circumstance for suicide: Development and validation of a tool using natural language processing and supervised machine learning in the National Violent Death Reporting System,” *Injury Prevention*, vol. 29, no. 2, pp. 134–141, Apr. 2023, doi: 10.1136/ip-2022-044662.
- [62] R. C. Kessler *et al.*, “Evaluation of a Model to Target High-risk Psychiatric Inpatients for an Intensive Postdischarge Suicide Prevention Intervention,” *JAMA Psychiatry*, vol. 80, no. 3, pp. 230–240, Mar. 2023, doi: 10.1001/jamapsychiatry.2022.4634.
- [63] Weslei Felipe Heckler, L. P. Feijó, J. V. de Carvalho, and J. L. V. Barbosa, “Thoth: An intelligent model for assisting individuals with suicidal ideation,” *Expert Syst Appl*, vol. 233, Dec. 2023, doi: 10.1016/j.eswa.2023.120918.
- [64] S. Mouchabac, P. Leray, V. Adrien, F. Gollier-Briant, and O. Bonnot, “Prevention of suicidal relapses in adolescents with a smartphone application: Bayesian network analysis of a preclinical trial using in silico patient simulations,” *J Med Internet Res*, vol. 23, no. 9, Sep. 2021, doi: 10.2196/24560.
- [65] X. Liu *et al.*, “Proactive suicide prevention online (PSPO): Machine identification and crisis management for Chinese social media users with suicidal thoughts and behaviors,” *J Med Internet Res*, vol. 21, no. 5, May 2019, doi: 10.2196/11705.
- [66] L. D. Balbuena *et al.*, “Identifying long-term and imminent suicide predictors in a general population and a clinical sample with machine learning,” *BMC Psychiatry*, vol. 22, no. 1, Dec. 2022, doi: 10.1186/s12888-022-03702-y.
- [67] S. Berrouguet, M. L. Barrigón, J. L. Castroman, P. Courtet, A. Artés-Rodríguez, and E. Baca-García, “Combining mobile-health (mHealth) and artificial intelligence (AI) methods to avoid suicide attempts: The Smartcrises study protocol,” *BMC Psychiatry*, vol. 19, no. 1, Sep. 2019, doi: 10.1186/s12888-019-2260-y.

- [68] L. Sels *et al.*, "SIMON: A Digital Protocol to Monitor and Predict Suicidal Ideation," *Front Psychiatry*, vol. 12, Jul. 2021, doi: 10.3389/fpsyt.2021.554811.
- [69] M. J. Kyron, G. R. Hooke, and A. C. Page, "Prediction and network modelling of self-harm through daily self-report and history of self-injury," *Psychol Med*, vol. 51, no. 12, pp. 1992–2002, Sep. 2021, doi: 10.1017/S0033291720000744.
- [70] H. J. Bello, N. Palomar-Ciria, C. Lozano, C. Gutiérrez-Alonso, and E. Baca-García, "Examining the relationship between COVID-19 and suicide in media coverage through Natural Language Processing analysis," *European Journal of Psychiatry*, vol. 38, no. 1, Jan. 2024, doi: 10.1016/j.ejpsy.2023.100227.
- [71] S. I. Ranapurwala *et al.*, "Innovations in suicide prevention research (INSPIRE): a protocol for a population-based case-control study," *Injury Prevention*, vol. 28, no. 5, pp. 483–490, Jun. 2022, doi: 10.1136/injuryprev-2022-044609.
- [72] G. Hasey *et al.*, "Detection of Suicidal Ideation in Depressed Subjects Using Resting Electroencephalography Features Identified by Machine Learning Algorithms," *Biol Psychiatry*, vol. 87, no. 9, pp. S380–S381, May 2020, doi: 10.1016/j.biopsych.2020.02.974.
- [73] Z. X. Huang *et al.*, "Development and validation of the risk score for estimating suicide attempt in patients with major depressive disorder," *Soc Psychiatry Psychiatr Epidemiol*, 2023, doi: 10.1007/s00127-023-02572-3.
- [74] S. Il Song, H. T. Hong, C. Lee, and S. B. Lee, "A machine learning approach for predicting suicidal ideation in post stroke patients," *Sci Rep*, vol. 12, no. 1, Dec. 2022, doi: 10.1038/s41598-022-19828-8.
- [75] S. Peng, Y. Zhou, L. Xiong, and Q. Wang, "Identification of novel targets and pathways to distinguish suicide dependent or independent on depression diagnosis," *Sci Rep*, vol. 13, no. 1, Dec. 2023, doi: 10.1038/s41598-023-29101-1.
- [76] J. C. Weng, T. Y. Lin, Y. H. Tsai, M. T. Cheok, Y. P. E. Chang, and V. C. H. Chen, "An autoencoder and machine learning model to predict suicidal ideation with brain structural imaging," *J Clin Med*, vol. 9, no. 3, Mar. 2020, doi: 10.3390/jcm9030658.
- [77] J. P. Stange *et al.*, "Using resting-state intrinsic network connectivity to identify suicide risk in mood disorders," *Psychol Med*, vol. 50, no. 14, pp. 2324–2334, Oct. 2020, doi: 10.1017/S0033291719002356.
- [78] M. Malgaroli, T. D. Hull, N. Bantilan, B. Ray, and N. Simon, "Suicide Risk Automated Detection Using Computational Linguistic Markers From Patients' Communication With Therapists," *Biol Psychiatry*, vol. 87, no. 9, p. S444, May 2020, doi: 10.1016/j.biopsych.2020.02.1131.
- [79] M. H. McCullough, M. Small, B. Jayawardena, and S. Hood, "Mapping clinical interactions in an Australian tertiary hospital emergency department for patients presenting with risk of suicide or self-harm: Network modeling from observational data," *PLoS Med*, vol. 21, no. 1, Jan. 2024, doi: 10.1371/journal.pmed.1004241.
- [80] K. van Mens *et al.*, "Applying machine learning on health record data from general practitioners to predict suicidality," *Internet Interv*, vol. 21, Sep. 2020, doi: 10.1016/j.invent.2020.100337.

- [81] V. Rozova, K. Witt, J. Robinson, Y. Li, and K. Verspoor, "Detection of self-harm and suicidal ideation in emergency department triage notes," *Journal of the American Medical Informatics Association*, vol. 29, no. 3, pp. 472–480, Mar. 2022, doi: 10.1093/jamia/ocab261.
- [82] N. Nordin, Z. Zainol, M. H. Mohd Noor, and L. F. Chan, "An explainable predictive model for suicide attempt risk using an ensemble learning and Shapley Additive Explanations (SHAP) approach," *Asian J Psychiatr*, vol. 79, Jan. 2023, doi: 10.1016/j.ajp.2022.103316.
- [83] T. Jiang, D. Nagy, A. J. Rosellini, ; Erzsébet Horváth-Puhó, K. M. Keyes, and T. L. Lash, "Prediction of Suicide Attempts Among Persons with Depression: A Population-Based Case Cohort Study", doi: 10.1093/aje/kwad237/7459183.
- [84] S. E. Cho, Z. W. Geem, and K. S. Na, "Prediction of suicide among 372,813 individuals under medical check-up," *J Psychiatr Res*, vol. 131, pp. 9–14, Dec. 2020, doi: 10.1016/j.jpsychires.2020.08.035.
- [85] R. Zhu *et al.*, "Discriminating Suicide Attempters and Predicting Suicide Risk Using Altered Frontolimbic Resting-State Functional Connectivity in Patients With Bipolar II Disorder," *Front Psychiatry*, vol. 11, Nov. 2020, doi: 10.3389/fpsy.2020.597770.
- [86] S. Bajaj *et al.*, "Machine learning based identification of structural brain alterations underlying suicide risk in adolescents," *Discover Mental Health*, vol. 3, no. 1, Dec. 2023, doi: 10.1007/s44192-023-00033-6.
- [87] M. Xu *et al.*, "Identification of suicidality in patients with major depressive disorder via dynamic functional network connectivity signatures and machine learning," *Transl Psychiatry*, vol. 12, no. 1, Dec. 2022, doi: 10.1038/s41398-022-02147-x.
- [88] J. S. Obeid *et al.*, "Identifying and predicting intentional self-harm in electronic health record clinical notes: Deep learning approach," *JMIR Med Inform*, vol. 8, no. 7, Jul. 2020, doi: 10.2196/17784.
- [89] M. Cusick *et al.*, "Using weak supervision and deep learning to classify clinical notes for identification of current suicidal ideation," *J Psychiatr Res*, vol. 136, pp. 95–102, Apr. 2021, doi: 10.1016/j.jpsychires.2021.01.052.
- [90] M. Levis, J. Levy, V. Dufort, G. T. Gobbel, B. V. Watts, and B. Shiner, "Leveraging unstructured electronic medical record notes to derive population-specific suicide risk models," *Psychiatry Res*, vol. 315, Sep. 2022, doi: 10.1016/j.psychres.2022.114703.
- [91] D. Morrow *et al.*, "A case for developing domain-specific vocabularies for extracting suicide factors from healthcare notes," *J Psychiatr Res*, vol. 151, pp. 328–338, Jul. 2022, doi: 10.1016/j.jpsychires.2022.04.009.
- [92] F. Xie, D. S. Ling Grant, J. Chang, B. I. Amundsen, and R. C. Hechter, "Identifying Suicidal Ideation and Attempt from Clinical Notes within a Large Integrated Health Care System," *Permanente Journal*, vol. 26, no. 1, pp. 85–93, 2022, doi: 10.7812/TPP/21.102.
- [93] V. Chin-Hung Chen *et al.*, "Convolutional Neural Network-Based Deep Learning Model for Predicting Differential Suicidality in Depressive Patients Using Brain Generalized q-Sampling Imaging," *Journal of Clinical Psychiatry*, vol. 82, no. 2, Mar. 2021, doi: 10.4088/JCP.19M13225.

- [94] T. M. H. Li *et al.*, "Detection of Suicidal Ideation in Clinical Interviews for Depression Using Natural Language Processing and Machine Learning: Cross-Sectional Study," *JMIR Med Inform*, vol. 11, no. 1, Jan. 2023, doi: 10.2196/50221.
- [95] M. Broadbent *et al.*, "A machine learning approach to identifying suicide risk among text-based crisis counseling encounters," *Front Psychiatry*, vol. 14, 2023, doi: 10.3389/fpsy.2023.1110527.
- [96] N. Bantilan, M. Malgaroli, B. Ray, and T. D. Hull, "Just in time crisis response: suicide alert system for telemedicine psychotherapy settings," *Psychotherapy Research*, vol. 31, no. 3, pp. 302–312, 2021, doi: 10.1080/10503307.2020.1781952.
- [97] F. M. Morales-Rodríguez, J. P. Martínez-Ramón, J. M. Giménez-Lozano, and A. M. Morales Rodríguez, "Suicide Risk Analysis and Psycho-Emotional Risk Factors Using an Artificial Neural Network System," *Healthcare (Switzerland)*, vol. 11, no. 16, Aug. 2023, doi: 10.3390/healthcare11162337.
- [98] G. M. Lin, M. Nagamine, S. N. Yang, Y. M. Tai, C. Lin, and H. Sato, "Machine Learning Based Suicide Ideation Prediction for Military Personnel," *IEEE J Biomed Health Inform*, vol. 24, no. 7, pp. 1907–1916, Jul. 2020, doi: 10.1109/JBHI.2020.2988393.
- [99] J. S. Jung, S. J. Park, E. Y. Kim, K. S. Na, Y. J. Kim, and K. G. Kim, "Prediction models for high risk of suicide in Korean adolescents using machine learning techniques," *PLoS One*, vol. 14, no. 6, Jun. 2019, doi: 10.1371/journal.pone.0217639.
- [100] R. Su, J. R. John, and P. I. Lin, "Machine learning-based prediction for self-harm and suicide attempts in adolescents," *Psychiatry Res*, vol. 328, Oct. 2023, doi: 10.1016/j.psychres.2023.115446.
- [101] A. K. Littlefield *et al.*, "Machine Learning to Classify Suicidal Thoughts and Behaviors: Implementation Within the Common Data Elements Used by the Military Suicide Research Consortium," *Clinical Psychological Science*, vol. 9, no. 3, pp. 467–481, May 2021, doi: 10.1177/2167702620961067.
- [102] K. L. Zuromski *et al.*, "Pre-deployment predictors of suicide attempt during and after combat deployment: Results from the Army Study to Assess Risk and Resilience in Servicemembers," *J Psychiatr Res*, vol. 121, pp. 214–221, Feb. 2020, doi: 10.1016/j.jpsychires.2019.12.003.
- [103] S. Mahmud *et al.*, "Medicine ® Machine learning approaches for predicting suicidal behaviors among university students in Bangladesh during the COVID-19 pandemic A cross-sectional study," 2023, doi: 10.21203/rs.3.rs-2069873/v1.
- [104] D. Lekkas, R. J. Klein, and N. C. Jacobson, "Predicting acute suicidal ideation on Instagram using ensemble machine learning models," *Internet Interv*, vol. 25, Sep. 2021, doi: 10.1016/j.invent.2021.100424.
- [105] X. Du and Y. Sun, "Linguistic features and psychological states: A machine-learning based approach," *Front Psychol*, vol. 13, Jul. 2022, doi: 10.3389/fpsyg.2022.955850.
- [106] A. Mbarek, S. Jamoussi, and A. Ben Hamadou, "An across online social networks profile building approach: Application to suicidal ideation detection," *Future Generation Computer Systems*, vol. 133, pp. 171–183, Aug. 2022, doi: 10.1016/j.future.2022.03.017.

- [107] Y. Ophir, R. Tikochinski, C. S. C. Asterhan, I. Sisso, and R. Reichart, "Deep neural networks detect suicide risk from textual facebook posts," *Sci Rep*, vol. 10, no. 1, Dec. 2020, doi: 10.1038/s41598-020-73917-0.
- [108] U. Naseem, M. Khushi, J. Kim, and A. G. Dunn, "Hybrid Text Representation for Explainable Suicide Risk Identification on Social Media," *IEEE Trans Comput Soc Syst*, 2022, doi: 10.1109/TCSS.2022.3184984.
- [109] S. Onie *et al.*, "Understanding and detecting behaviours prior to a suicide attempt: A mixed-methods study," *Australian and New Zealand Journal of Psychiatry*, vol. 57, no. 7, pp. 1016–1022, Jul. 2023, doi: 10.1177/00048674231152159.
- [110] T. Ghosh, M. H. Al Banna, M. J. Al Nahian, M. N. Uddin, M. S. Kaiser, and M. Mahmud, "An attention-based hybrid architecture with explainability for depressive social media text detection in Bangla," *Expert Syst Appl*, vol. 213, Mar. 2023, doi: 10.1016/j.eswa.2022.119007.
- [111] D. Liu *et al.*, "Suicidal ideation cause extraction from social texts," *IEEE Access*, vol. 8, pp. 169333–169351, 2020, doi: 10.1109/ACCESS.2020.3019491.
- [112] Z. Xu *et al.*, "Detecting suicide risk using knowledge-aware natural language processing and counseling service data," *Soc Sci Med*, vol. 283, Aug. 2021, doi: 10.1016/j.socscimed.2021.114176.
- [113] H. Metzler, H. Baginski, T. Niederkrotenthaler, and D. Garcia, "Detecting Potentially Harmful and Protective Suicide-Related Content on Twitter: Machine Learning Approach," *J Med Internet Res*, vol. 24, no. 8, Aug. 2022, doi: 10.2196/34705.
- [114] H. A. Burkhardt, X. Ding, A. Kerbrat, K. A. Comtois, and T. Cohen, "From benchmark to bedside: transfer learning from social media to patient-provider text messages for suicide risk prediction," *Journal of the American Medical Informatics Association*, vol. 30, no. 6, pp. 1068–1078, Jun. 2023, doi: 10.1093/jamia/ocad062.
- [115] A. M. Schoene, A. P. Turner, G. De Mel, and N. Dethlefs, "Hierarchical Multiscale Recurrent Neural Networks for Detecting Suicide Notes," *IEEE Trans Affect Comput*, vol. 14, no. 1, pp. 153–164, Jan. 2023, doi: 10.1109/TAFFC.2021.3057105.
- [116] H. Yao, S. Rashidian, X. Dong, H. Duanmu, R. N. Rosenthal, and F. Wang, "Detection of suicidality among opioid users on reddit: Machine learning-based approach," *J Med Internet Res*, vol. 22, no. 11, Nov. 2020, doi: 10.2196/15293.
- [117] W. Ragheb, J. Aze, S. Bringay, and M. Servajean, "Negatively Correlated Noisy Learners for At-Risk User Detection on Social Networks: A Study on Depression, Anorexia, Self-Harm, and Suicide," *IEEE Trans Knowl Data Eng*, vol. 35, no. 1, pp. 770–783, Jan. 2023, doi: 10.1109/TKDE.2021.3078898.
- [118] F. M. Talaat, E. M. El-Gendy, M. M. Saafan, and S. A. Gamel, "Utilizing social media and machine learning for personality and emotion recognition using PERS," *Neural Comput Appl*, vol. 35, no. 33, pp. 23927–23941, Nov. 2023, doi: 10.1007/s00521-023-08962-7.
- [119] S. T. Rabani, A. M. Ud Din Khanday, Q. R. Khan, U. A. Hajam, A. S. Imran, and Z. Kastrati, "Detecting suicidality on social media: Machine learning at rescue," *Egyptian Informatics Journal*, vol. 24, no. 2, pp. 291–302, Jul. 2023, doi: 10.1016/j.eij.2023.04.003.

- [120] D. Ramirez-Cifuentes, C. Largeron, J. Tissier, R. Baeza-Yates, and A. Freire, "Enhanced Word Embedding Variations for the Detection of Substance Abuse and Mental Health Issues on Social Media Writings," *IEEE Access*, vol. 9, pp. 130449–130471, 2021, doi: 10.1109/ACCESS.2021.3112102.
- [121] N. A. Baghdadi, A. Malki, H. M. Balaha, Y. AbdulAzeem, M. Badawy, and M. Elhosseini, "An optimized deep learning approach for suicide detection through Arabic tweets," *PeerJ Comput Sci*, vol. 8, 2022, doi: 10.7717/PEERJ-CS.1070.
- [122] A. M. Schoene, L. Bojanic, M. Q. Nghiem, I. M. Hunt, and S. Ananiadou, "Classifying Suicide-Related Content and Emotions on Twitter Using Graph Convolutional Neural Networks," *IEEE Trans Affect Comput*, vol. 14, no. 3, pp. 1791–1802, Jul. 2023, doi: 10.1109/TAFFC.2022.3221683.
- [123] P. Buddhitha and D. Inkpen, "Multi-task learning to detect suicide ideation and mental disorders among social media users," *Front Res Metr Anal*, vol. 8, 2023, doi: 10.3389/frma.2023.1152535.
- [124] S. Renjith, A. Abraham, S. B. Jyothi, L. Chandran, and J. Thomson, "An ensemble deep learning technique for detecting suicidal ideation from posts in social media platforms," *Journal of King Saud University - Computer and Information Sciences*, vol. 34, no. 10, pp. 9564–9575, Nov. 2022, doi: 10.1016/j.jksuci.2021.11.010.
- [125] V. Adarsh, P. Arun Kumar, V. Lavanya, and G. R. Gangadharan, "Fair and Explainable Depression Detection in Social Media," *Inf Process Manag*, vol. 60, no. 1, Jan. 2023, doi: 10.1016/j.ipm.2022.103168.
- [126] T. Jiang *et al.*, "Suicide prediction among men and women with depression: A population-based study," *J Psychiatr Res*, vol. 142, pp. 275–282, Oct. 2021, doi: 10.1016/j.jpsychires.2021.08.003.
- [127] J. L. Gradus *et al.*, "Predicting Sex-Specific Nonfatal Suicide Attempt Risk Using Machine Learning and Data From Danish National Registries," *Am J Epidemiol*, vol. 190, no. 12, pp. 2517–2527, Dec. 2021, doi: 10.1093/aje/kwab112.
- [128] H. S. Laqueur, C. Smirniotis, C. McCort, and G. J. Wintemute, "Machine Learning Analysis of Handgun Transactions to Predict Firearm Suicide Risk," *JAMA Netw Open*, p. E2221041, 2022, doi: 10.1001/jamanetworkopen.2022.21041.
- [129] I. L. Lin, J. Y. C. Tseng, H. T. Tung, Y. H. Hu, and Z. H. You, "Predicting the Risk of Future Multiple Suicide Attempt among First-Time Suicide Attempters: Implications for Suicide Prevention Policy," *Healthcare (Switzerland)*, vol. 10, no. 4, Apr. 2022, doi: 10.3390/healthcare10040667.
- [130] R. Kwon, H. Lee, M. S. Kim, J. Lee, and D. K. Yon, "Machine learning-based prediction of suicidality in adolescents during the COVID-19 pandemic (2020–2021): Derivation and validation in two independent nationwide cohorts," *Asian Journal of Psychiatry*, vol. 88. Elsevier B.V., Oct. 01, 2023. doi: 10.1016/j.ajp.2023.103704.
- [131] A. Horvath, M. Dras, C. C. W. Lai, and S. Boag, "Predicting Suicidal Behavior Without Asking About Suicidal Ideation: Machine Learning and the Role of Borderline Personality Disorder Criteria," *Suicide Life Threat Behav*, vol. 51, no. 3, pp. 455–466, Jun. 2021, doi: 10.1111/sltb.12719.

- [132] N. Rahman, R. Mozer, R. K. McHugh, I. R. H. Rockett, C. M. Chow, and G. Vaughan, "Using natural language processing to improve suicide classification requires consideration of race," *Suicide Life Threat Behav*, vol. 52, no. 4, pp. 782–791, Aug. 2022, doi: 10.1111/sltb.12862.
- [133] S. Wang *et al.*, "An NLP approach to identify SDoH-related circumstance and suicide crisis from death investigation narratives," *J Am Med Inform Assoc*, vol. 30, no. 8, pp. 1408–1417, Jul. 2023, doi: 10.1093/jamia/ocad068.
- [134] S. Shin and K. Kim, "Prediction of suicidal ideation in children and adolescents using machine learning and deep learning algorithm: A case study in South Korea where suicide is the leading cause of death," *Asian J Psychiatr*, vol. 88, Oct. 2023, doi: 10.1016/j.ajp.2023.103725.
- [135] Z. Xu *et al.*, "A Predictive Model of Risk Factors for Conversion From Major Depressive Disorder to Bipolar Disorder Based on Clinical Characteristics and Circadian Rhythm Gene Polymorphisms," *Front Psychiatry*, vol. 13, Jul. 2022, doi: 10.3389/fpsyt.2022.843400.
- [136] Y. Y. Joo *et al.*, "Association of Genome-Wide Polygenic Scores for Multiple Psychiatric and Common Traits in Preadolescent Youths at Risk of Suicide," *JAMA Netw Open*, vol. 5, no. 2, Feb. 2022, doi: 10.1001/jamanetworkopen.2021.48585
- [137] M. Lozupone *et al.*, "Apolipoprotein E genotype, inflammatory biomarkers, and non-psychiatric multimorbidity contribute to the suicidal ideation phenotype in older age. The Salus in Apulia Study," *J Affect Disord*, vol. 319, pp. 202–212, Dec. 2022, doi: 10.1016/j.jad.2022.09.046.
- [138] K. Qin *et al.*, "Transcriptional Patterns of Brain Structural Covariance Network Abnormalities Associated with Suicidal Thoughts and Behaviors in Major Depressive Disorder," *Biol Psychiatry*, Feb. 2024, doi: 10.1016/j.biopsych.2024.01.026.
- [139] A. Haines-Delmont *et al.*, "Testing suicide risk prediction algorithms using phone measurements with patients in acute mental health settings: Feasibility study," *JMIR Mhealth Uhealth*, vol. 8, no. 6, Jun. 2020, doi: 10.2196/15901.
- [140] S. Lyu, X. Ren, Y. Du, and N. Zhao, "Detecting depression of Chinese microblog users via text analysis: Combining Linguistic Inquiry Word Count (LIWC) with culture and suicide related lexicons," *Front Psychiatry*, vol. 14, Feb. 2023, doi: 10.3389/fpsyt.2023.1121583.
- [141] M. Chatterjee, P. Kumar, P. Samanta, and D. Sarkar, "Suicide ideation detection from online social media: A multi-modal feature based technique," *International Journal of Information Management Data Insights*, vol. 2, no. 2, Nov. 2022, doi: 10.1016/j.jjime.2022.100103.
- [142] D. Ramírez-Cifuentes *et al.*, "Detection of suicidal ideation on social media: Multimodal, relational, and behavioral analysis," *J Med Internet Res*, vol. 22, no. 7, Jul. 2020, doi: 10.2196/17758.
- [143] A. Fulginiti *et al.*, "Getting to the Root of the Problem: A Decision-Tree Analysis for Suicide Risk Among Young People Experiencing Homelessness," *J Soc Social Work Res*, vol. 13, no. 2, pp. 327–352, Jun. 2022, doi: 10.1086/715211.
- [144] J. B. Edgcomb, C. H. Tseng, M. Pan, A. Klomhaus, and B. T. Zima, "Assessing Detection of Children With Suicide-Related Emergencies: Evaluation and Development of Computable Phenotyping Approaches," *JMIR Ment Health*, vol. 10, 2023, doi: 10.2196/47084.

- [145] W. Yang *et al.*, "Establishment of a risk prediction model for suicide attempts in first-episode and drug naïve patients with major depressive disorder," *Asian J Psychiatr*, vol. 88, Oct. 2023, doi: 10.1016/j.ajp.2023.103732.
- [146] D. Vogt, A. J. Rosellini, S. Borowski, A. E. Street, R. W. O'Brien, and N. Tomoyasu, "How well can U.S. military veterans' suicidal ideation be predicted from static and change-based indicators of their psychosocial well-being as they adapt to civilian life?," *Soc Psychiatry Psychiatr Epidemiol*, vol. 59, no. 2, pp. 261–271, Feb. 2024, doi: 10.1007/s00127-023-02511-2.
- [147] M. Pavicic *et al.*, "Using iterative random forest to find geospatial environmental and Sociodemographic predictors of suicide attempts," *Front Psychiatry*, vol. 14, 2023, doi: 10.3389/fpsy.2023.1178633.
- [148] B. J. H. Yarborough *et al.*, "Opioid-related variables did not improve suicide risk prediction models in samples with mental health diagnoses," *J Affect Disord Rep*, vol. 8, Apr. 2022, doi: 10.1016/j.jadr.2022.100346.
- [149] S. Barzilay *et al.*, "Real-Time Real-World Digital Monitoring of Adolescent Suicide Risk During the Six Months Following Emergency Department Discharge: Protocol for an Intensive Longitudinal Study," *JMIR Res Protoc*, vol. 12, 2023, doi: 10.2196/46464.
- [150] D. Choi *et al.*, "Development of a Machine Learning Model Using Multiple, Heterogeneous Data Sources to Estimate Weekly US Suicide Fatalities," *JAMA Netw Open*, vol. 3, no. 12, Dec. 2020, doi: 10.1001/jamanetworkopen.2020.30932.
- [151] D. C. Rozek *et al.*, "Using Machine Learning to Predict Suicide Attempts in Military Personnel," *Psychiatry Res*, vol. 294, Dec. 2020, doi: 10.1016/j.psychres.2020.113515.
- [152] O. Miranda *et al.*, "Prediction of adverse events risk in patients with comorbid post-traumatic stress disorder and alcohol use disorder using electronic medical records by deep learning models," *Drug Alcohol Depend*, vol. 255, Feb. 2024, doi: 10.1016/j.drugalcdep.2023.111066.
- [153] Y. Cai, H. Wang, H. Ye, Y. Jin, and W. Gao, "Depression detection on online social network with multivariate time series feature of user depressive symptoms," *Expert Syst Appl*, vol. 217, May 2023, doi: 10.1016/j.eswa.2023.119538.
- [154] Z. Hao, H. Li, L. Ouyang, F. Sun, X. Wen, and X. Wang, "Pain avoidance and functional connectivity between insula and amygdala identifies suicidal attempters in patients with major depressive disorder using machine learning," *Psychophysiology*, vol. 60, no. 1, Jan. 2023, doi: 10.1111/psyp.14136.
- [155] D. Wilimitis *et al.*, "Integration of Face-to-Face Screening with Real-time Machine Learning to Predict Risk of Suicide among Adults," *JAMA Netw Open*, vol. 5, no. 5, p. E2212095, May 2022, doi: 10.1001/jamanetworkopen.2022.12095.
- [156] S. Kim, H. K. Lee, and K. Lee, "Detecting suicidal risk using MMPI-2 based on machine learning algorithm," *Sci Rep*, vol. 11, no. 1, Dec. 2021, doi: 10.1038/s41598-021-94839-5.
- [157] Z. Li, W. Cheng, J. Zhou, Z. An, and B. Hu, "Deep learning model with multi-feature fusion and label association for suicide detection," *Multimed Syst*, vol. 29, no. 4, pp. 2193–2203, Aug. 2023, doi: 10.1007/s00530-023-01090-1.

- [158] H. Ghanadian, I. Nejadgholi, and H. Al Osman, "Socially Aware Synthetic Data Generation for Suicidal Ideation Detection Using Large Language Models," *IEEE Access*, vol. 12, pp. 14350–14363, 2024, doi: 10.1109/ACCESS.2024.3358206.
- [159] L. Sher, J. J. Peters, and L. Sher, "The impact of the COVID-19 pandemic on suicide rates From the", doi: 10.1093/qjmed/hcaa202.
- [160] Y. Jiang, W. Li, M. Shamim Hossain, M. Chen, A. Alelaiwi, and M. Al-Hammadi, "A snapshot research and implementation of multimodal information fusion for data-driven emotion recognition," *Information Fusion*, vol. 53, pp. 209–221, 2020, doi: 10.1016/j.inffus.2019.06.019.
- [161] Sheykhmousa, M., Mahdianpari, M., Ghanbari, H., Mohammadimanesh, F., Ghamisi, P., & Homayouni, S. (2020). Support vector machine versus random forest for remote sensing image classification: A meta-analysis and systematic review. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 13, 6308-6325. doi: 10.1109/JSTARS.2020.3026724
- [162] B. Xu, J. Z. Huang, G. Williams, Q. Wang, and Y. Ye, "Classifying very high-dimensional data with random forests built from small subspaces," *International Journal of Data Warehousing and Mining*, vol. 8, no. 2, pp. 44–63, Apr. 2012, doi: 10.4018/jdwm.2012040103.
- [163] M. Salvi *et al.*, "Multi-modality approaches for medical support systems: A systematic review of the last decade," 2023, doi: 10.1016/j.inffus.2023.102134.
- [164] T. Baltrušaitis, C. Ahuja, and L.-P. Morency, "Challenges and applications in multimodal machine learning," *The Handbook of Multimodal-Multisensor Interfaces: Foundations, User Modeling, and Common Modality Combinations - Volume 2*, pp. 17–48, Oct. 2018, doi: 10.1145/3107990.3107993.
- [165] O. Weller, L. Sagers, C. Hanson, M. Barnes, Q. Snell, and E. Shannon Tass, "Predicting suicidal thoughts and behavior among adolescents using the risk and protective factor framework: A large-scale machine learning approach," *PLoS One*, vol. 16, no. November, Nov. 2021, doi: 10.1371/journal.pone.0258535.
- [166] P. P. Sinha *et al.*, "#suicidal-A Multipronged Approach to Identify and Explore Suicidal Ideation in Twitter", doi: 10.1145/3357384.3358060.
- [167] M. Gaur *et al.*, "Knowledge-aware Assessment of Severity of Suicide Risk for Early Intervention," 2019, doi: 10.1145/3308558.3313698.
- [168] A. Zirikly, P. Resnik, " Ozlem Uzuner, and K. Hollingshead, "CLPsych 2019 Shared Task: Predicting the Degree of Suicide Risk in Reddit Posts," *Proceedings of the Sixth Workshop on Computational Linguistics and Clinical Psychology*, pp. 24–33, 2019, Accessed: May 21, 2024. [Online].
- [169] Shing, H. C., Nair, S., Zirikly, A., Friedenber, M., Daumé III, H., & Resnik, P. (2018, June). Expert, crowdsourced, and machine assessment of suicide risk via online postings. In *Proceedings of the fifth workshop on computational linguistics and clinical psychology: from keyboard to clinic* (pp. 25-36).
- [170] I. L. Lin, J. Y. C. Tseng, H. T. Tung, Y. H. Hu, and Z. H. You, "Predicting the Risk of Future Multiple Suicide Attempt among First-Time Suicide Attempters: Implications for Suicide Prevention Policy," *Healthcare (Switzerland)*, vol. 10, no. 4, Apr. 2022, doi: 10.3390/healthcare10040667.

- [171] D. M. Roden *et al.*, "Development of a Large-Scale De-Identified DNA Biobank to Enable Personalized Medicine," *Clin Pharmacol Ther*, vol. 84, no. 3, pp. 362–369, Sep. 2008, doi: 10.1038/CLPT.2008.89.
- [172] Zingmond, D. (2010). Linkage Documentation: Death Statistical Master File linkage to OSHPD databases (PDD, EDD, and ASD) with three year mortality outcomes for all eligible records; years 2005 to 2009. *Los Angeles, CA*.
- [173] R. Kaushal *et al.*, "Changing the research landscape: the New York City Clinical Data Research Network," *Journal of the American Medical Informatics Association*, vol. 21, no. 4, pp. 587–590, Jul. 2014, doi: 10.1136/AMIAJNL-2014-002764.
- [174] S. Peng, Y. Zhou, L. Xiong, and Q. Wang, "Identification of novel targets and pathways to distinguish suicide dependent or independent on depression diagnosis," *Scientific Reports* 2023 13:1, vol. 13, no. 1, pp. 1–13, Feb. 2023, doi: 10.1038/s41598-023-29101-1.
- [175] Y. Cai, H. Wang, H. Ye, Y. Jin, and W. Gao, "Depression detection on online social network with multivariate time series feature of user depressive symptoms," *Expert Syst Appl*, vol. 217, p. 119538, May 2023, doi: 10.1016/J.ESWA.2023.119538.
- [176] A. Sanson, J. Nicholson, and J. Ungerer, "Introducing the Longitudinal Study of Australian Children: LSAC DISCUSSION PAPER N0.1," *Australian Institute of Family Studies*, 2002.
- [177] Goldman-Mellor, S., & Margerison, C. E. (2019). Maternal drug-related death and suicide are leading causes of postpartum death in California. *American journal of obstetrics and gynecology*, 221(5), 489-e1.
- [178] N. N. Gomes de Andrade, D. Pawson, D. Muriello, L. Donahue, and J. Guadagno, "Ethics and Artificial Intelligence: Suicide Prevention on Facebook," *Philosophy and Technology*, vol. 31, no. 4. Springer Netherlands, pp. 669–684, Dec. 01, 2018. doi: 10.1007/s13347-018-0336-0.
- [179] Anderson, T. L. (2013). Navigating HIPAA's Hidden Minefields: A Leader's Guide to Using HIPAA Correctly to Decrease Suicide and Homicide in the Military. *Army Law.*, 15.
- [180] O. J. Kirtley, K. van Mens, M. Hoogendoorn, N. Kapur, and D. de Beurs, "Translating promise into practice: a review of machine learning in suicide research and prevention," *Lancet Psychiatry*, vol. 9, no. 3, pp. 243–252, Mar. 2022, doi: 10.1016/S2215-0366(21)00254-6.
- [181] T. Costa *et al.*, "The burden associated with, and management of, difficult-to-treat depression in patients under specialist psychiatric care in the United Kingdom," *Journal of Psychopharmacology*, vol. 36, no. 5, pp. 545–556, May 2022, doi: 10.1177/02698811221090628.
- [182] B. X. Yang *et al.*, "Characteristics of High Suicide Risk Messages From Users of a Social Network—Sina Weibo 'Tree Hole,'" *Front Psychiatry*, vol. 13, Feb. 2022, doi: 10.3389/fpsy.2022.789504.
- [183] J. Pestian *et al.*, "A Machine Learning Approach to Identifying Changes in Suicidal Language," *Suicide Life Threat Behav*, vol. 50, no. 5, pp. 939–947, Oct. 2020, doi: 10.1111/sltb.12642.
- [184] K. S. Choi *et al.*, "Deep graph neural network-based prediction of acute suicidal ideation in young adults," *Sci Rep*, vol. 11, no. 1, Dec. 2021, doi: 10.1038/s41598-021-95102-7.
- [185] E. E. Haroz *et al.*, "Designing a clinical decision support tool that leverages machine learning for suicide risk prediction: Development study in partnership with native american care providers," *JMIR Public Health Surveill*, vol. 7, no. 9, Sep. 2021, doi: 10.2196/24377.

- [186] S. Ghosh, A. Ekbal, and P. Bhattacharyya, "A Multitask Framework to Detect Depression, Sentiment and Multi-label Emotion from Suicide Notes," *Cognit Comput*, vol. 14, no. 1, pp. 110–129, Jan. 2022, doi: 10.1007/s12559-021-09828-7.
- [187] D. Shin *et al.*, "Detection of Depression and Suicide Risk Based on Text From Clinical Interviews Using Machine Learning: Possibility of a New Objective Diagnostic Marker," *Front Psychiatry*, vol. 13, May 2022, doi: 10.3389/fpsyt.2022.801301.
- [188] J. Wright-Berryman, J. Cohen, A. Haq, D. P. Black, and J. L. Pease, "Virtually screening adults for depression, anxiety, and suicide risk using machine learning and language from an open-ended interview," *Front Psychiatry*, vol. 14, 2023, doi: 10.3389/fpsyt.2023.1143175.
- [189] S. Li, W. Pan, P. S. F. Yip, J. Wang, W. Zhou, and T. Zhu, "Uncovering the heterogeneous effects of depression on suicide risk conditioned by linguistic features: A double machine learning approach," *Comput Human Behav*, vol. 152, Mar. 2024, doi: 10.1016/j.chb.2023.108080.
- [190] G. S. Malhi, P. Das, T. Outhred, D. Gessler, J. John Mann, and R. Bryant, "Cognitive and emotional impairments underpinning suicidal activity in patients with mood disorders: an fMRI study," *Acta Psychiatr Scand*, vol. 139, no. 5, pp. 454–463, May 2019, doi: 10.1111/acps.13022.
- [191] Hossain, E., Rana, R., Higgins, N., Soar, J., Barua, P. D., Pisani, A. R., & Turner, K. (2023). Natural language processing in electronic health records in relation to healthcare decision-making: a systematic review. *Computers in biology and medicine*, 155, 106649. Doi: <https://doi.org/10.1016/j.compbio.2023.106649>
- [192] Saffar, A. H., Mann, T. K., & Ofoghi, B. (2023). Textual emotion detection in health: Advances and applications. *Journal of Biomedical Informatics*, 137, 104258. Doi: <https://doi.org/10.1016/j.jbi.2022.104258>
- [193] Garg, M. (2023). Mental health analysis in social media posts: a survey. *Archives of Computational Methods in Engineering*, 30(3), 1819-1842. Doi: <https://doi.org/10.1007/s11831-022-09863-z>
- [194] Lopez-Castroman, J., Moulahi, B., Azé, J., Bringay, S., Deninotti, J., Guillaume, S., & Baca-Garcia, E. (2020). Mining social networks to improve suicide prevention: A scoping review. *Journal of neuroscience research*, 98(4), 616-625. Doi: <https://doi.org/10.1002/jnr.24404>
- [195] Malhotra, A., & Jindal, R. (2022). Deep learning techniques for suicide and depression detection from online social media: A scoping review. *Applied Soft Computing*, 130, 109713. Doi: <https://doi.org/10.1016/j.asoc.2022.109713>
- [196] Celedonia, K. L., Corrales Compagnucci, M., Minssen, T., & Lowery Wilson, M. (2021). Legal, ethical, and wider implications of suicide risk detection systems in social media platforms. *Journal of Law and the Biosciences*, 8(1), 1sab021. Doi: <https://doi.org/10.1093/jlb/lsab021>
- [197] Milne-Ives, M., Selby, E., Inkster, B., Lam, C., & Meinert, E. (2022). Artificial intelligence and machine learning in mobile apps for mental health: A scoping review. *PLOS digital health*, 1(8), e0000079. Doi: <https://doi.org/10.1371/journal.pdig.0000079>
- [198] Onie, S., Li, X., Liang, M., Sowmya, A., & Larsen, M. E. (2021). The use of closed-circuit television and video in suicide prevention: narrative review and future directions. *JMIR mental health*, 8(5), e27663. doi:10.2196/27663

- [199] Wislocki, K., Jager-Hyman, S., Brady, M., Weiss, M., Schaechter, T., Khazanov, G., ... & Becker-Haimes, E. (2023). Freely available training videos for suicide prevention: scoping review. *JMIR mental health*, 10, e48404. doi:10.2196/48404
- [200] Kleiman, E. M., Glenn, C. R., & Liu, R. T. (2023). The use of advanced technology and statistical methods to predict and prevent suicide. *Nature reviews psychology*, 2(6), 347-359.
- [201] Lejeune, A., Le Glaz, A., Perron, P. A., Sebti, J., Baca-Garcia, E., Walter, M., ... & Berrouiguet, S. (2022). Artificial intelligence and suicide prevention: a systematic review. *European psychiatry*, 65(1), e19. DOI: <https://doi.org/10.1192/j.eurpsy.2022.8>
- [202] Bhatt, P., Liu, J., Gong, Y., Wang, J., & Guo, Y. (2022). Emerging artificial intelligence–empowered mhealth: scoping review. *JMIR mHealth and uHealth*, 10(6), e35053. doi:10.2196/35053
- [203] 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. Doi: <https://doi.org/10.1007/s10462-022-10290-6>
- [204] Ji, S., Pan, S., Li, X., Cambria, E., Long, G., & Huang, Z. (2020). Suicidal ideation detection: A review of machine learning methods and applications. *IEEE Transactions on Computational Social Systems*, 8(1), 214-226. Doi: 10.1109/TCSS.2020.3021467
- [205] McHugh, C. M., & Large, M. M. (2020). Can machine-learning methods really help predict suicide?. *Current opinion in psychiatry*, 33(4), 369-374. Doi: 10.1097/YCO.0000000000000609
- [206] Nordin, N., Zainol, Z., Noor, M. H. M., & Chan, L. F. (2022). Suicidal behaviour prediction models using machine learning techniques: A systematic review. *Artificial intelligence in medicine*, 132, 102395. Doi: <https://doi.org/10.1016/j.artmed.2022.102395>
- [207] Kerr, J. A., Gillespie, A. N., O'Connor, M., Deane, C., Borschmann, R., Dashti, S. G., ... & Patton, G. C. (2023). Intervention targets for reducing mortality between mid-adolescence and mid-adulthood: a protocol for a machine-learning facilitated systematic umbrella review. *BMJ open*, 13(10), e068733. Doi: <https://doi.org/10.1136/bmjopen-2022-068733>
- [208] Cox, C. R., Moscardini, E. H., Cohen, A. S., & Tucker, R. P. (2020). Machine learning for suicidology: A practical review of exploratory and hypothesis-driven approaches. *Clinical Psychology Review*, 82, 101940. Doi: <https://doi.org/10.1016/j.cpr.2020.101940>
- [209] Kusuma, K., Larsen, M., Quiroz, J. C., Gillies, M., Burnett, A., Qian, J., & Torok, M. (2022). The performance of machine learning models in predicting suicidal ideation, attempts, and deaths: A meta-analysis and systematic review. *Journal of psychiatric research*, 155, 579-588. Doi: <https://doi.org/10.1016/j.jpsychires.2022.09.050>
- [210] Parsaei, M., Taghavizanjani, F., Cattarinussi, G., Moghaddam, H. S., Di Camillo, F., Akhondzadeh, S., ... & Delvecchio, G. (2023). Doi: <https://doi.org/10.1016/j.jad.2018.11.073> Classification of suicidality by training supervised machine learning models with brain MRI findings: A systematic review. *Journal of affective disorders*. Doi: <https://doi.org/10.1016/j.jad.2023.08.034>
- [211] Burke, T. A., Ammerman, B. A., & Jacobucci, R. (2019). The use of machine learning in the study of suicidal and non-suicidal self-injurious thoughts and behaviors: A systematic review. *Journal of affective disorders*, 245, 869-884. Doi: <https://doi.org/10.1016/j.jad.2018.11.073>

- [212] Auerbach, R. P., Pagliaccio, D., Allison, G. O., Alqueza, K. L., & Alonso, M. F. (2021). Neural correlates associated with suicide and nonsuicidal self-injury in youth. *Biological psychiatry*, 89(2), 119-133. Doi: <https://doi.org/10.1016/j.biopsych.2020.06.002>
- [213] Bentley, K. H., Maimone, J. S., Kilbury, E. N., Tate, M. S., Wisniewski, H., Levine, M. T., ... & Kleiman, E. M. (2021). Practices for monitoring and responding to incoming data on self-injurious thoughts and behaviors in intensive longitudinal studies: A systematic review. *Clinical psychology review*, 90, 102098. Doi: <https://doi.org/10.1016/j.cpr.2021.102098>
- [214] Borges do Nascimento, I. J., Marcolino, M. S., Abdulazeem, H. M., Weerasekara, I., Azzopardi-Muscat, N., Gonçalves, M. A., & Novillo-Ortiz, D. (2021). Impact of big data analytics on people's health: Overview of systematic reviews and recommendations for future studies. *Journal of medical Internet research*, 23(4), e27275. doi:10.2196/27275
- [215] Dev, A., Roy, N., Islam, M. K., Biswas, C., Ahmed, H. U., Amin, M. A., ... & Mamun, K. A. (2022). Exploration of EEG-based depression biomarkers identification techniques and their applications: a systematic review. *IEEE Access*, 10, 16756-16781. Doi: 10.1109/ACCESS.2022.3146711
- [216] Wulz, A. R., Law, R., Wang, J., & Wolkin, A. F. (2022). Leveraging data science to enhance suicide prevention research: a literature review. *Injury prevention*, 28(1), 74-80. Doi: <https://doi.org/10.1136/injuryprev-2021-044322>

Appendix

Table A. 1: Summary of studies (n=12) involving suicide prevention using clinical data.

Author, year	Dataset type	Aim of the Study	Study Population	Methods	Findings/Results (%)
Robinson et al. [7], 2020	Clinical data	To characterize the creation of a state-wide self-harm surveillance system in Victoria	Australia: Children, adolescents	NLP: SVM	-
Bejan et al. [29], 2022	Clinical notes	To enhance the determination of suicidal outcomes	America: Adolescents, Adults	NLP: Word2Vec	Scalable and precise NLP techniques can be formulated to identify suicidal behavior in EHRs to boost prevention efforts. AUROC: Suicidal ideation of 0.986 and suicide attempt of 0.973
Zhu et al. [17], 2020	Electronic health records	To contrast the performance of techniques based on text mining to detect suicidal behaviors	China: Adults, Depressed patients	ML: SVM, RF, AdaBoost	F1-score: SVM (0.9889); RF (0.9838); AdaBoost (0.9828)
McManus et al. [27], 2023	Electronic health records	Deployment of a national clinical text processing infrastructure within the Veterans Affairs Department	United States: Veterans, Adults, Old people	NLP: Nvidia BioMegatron large language model	This approach was effective for the radiology use case, where the pre-highlighted terms were precise, and the vocabulary was constrained.
Burnett et al. [15], 2022	Electronic health records	To categorize self-harm associated behaviors	New South Wales: Adolescents, Young Adults	ML: SVM, RF, DT, LR	Accuracy: SVM (0.868); RF (0.855); DT (0.844); LR (0.867)
Cusick et al. [30], 2022	Electronic health records	To share NLP approaches among different institutions	United Kingdom: Adolescents, Autism disorder	NLP: Rule-based	NLP models detect suicidality better at their development organization than other healthcare institutions, achieving lower

performance elsewhere.F1-score 0.85 vs. 0.68

Patients with treatment-resistant depression have a much higher illness burden, including worse depressive episodes, suicidality, comorbidities, unemployment, hospitalizations, and specialist treatment time.

An elevated self-harm risk was noticed during adolescence.AUC: 0.994; MCC: 0.225; sensitivity: 0.930; specificity: 0.984

The main factors linking to suicide were benzodiazepine overdose, suicidal ideation, medication overdose, antipsychotic overdose, and self-harm by jumping. Accuracy: NN (0.933); SVM (0.851)

The main predictors of suicide readmission were previous medical condition, past suicidal actions, older age, and history of pregnancy-associated mental health issue. Accuracy: 0.84

In simple analyses, suicide risk was notably higher in hospitalized versus discharged patients.

- Results showed that patients' first 8 questions about symptoms and last question about suicidal ideation

Costa et al. [181], 2022	Electronic health records	To identify patients with probable DTD and analyze demographic, disease and treatment data compared to 'non-DTD' MDD patients	United Kingdom: MDD, Adults	NLP: NN	
Nestsiarovich et al. [33], 2021	Electronic health records	To evaluate all frequently used bipolar disorder medications, as well as psychotherapy for the risk of self-harm	Unites States: Bipolar patients	ML: Regression	
Castillo-Sánchez et al. [34], 2022	Hospital data	To assist in feature selection pertaining to hospital readmissions for suicidal behavior	Spain: Bipolar, Alcohol depressed patients, Personality disorder	ML: SVM	
Edgcomb et al. [31], 2021	Hospital data	To discriminate the risk of suicide attempts and self-harm after general medical hospitalization among females	United States: Women, Depressed patients, Bipolar, Adults, Old people	ML: DT	
Goldman-Mellor et al. [28], 2022	Hospital data	Impacts of hospitalization	California: Recurrence prediction kind	ML: RF	
Gong et al., 2019 [32]	Questionnaire	Understanding of temporal correlations between depression symptoms over time	America: Depressed patients, Adults	ML: ANN	

scores displayed strong temporal correlations.

AUROC: Area Under the Receiver Operator Curve; AUC: Area Under the Curve; ANN: Artificial Neural Network; DL: Deep Learning; DTD: Difficult-to-Treat Depression; DT: Decision Tree; EHR: Electronic Health Records; LR: Logistic Regression; MDD: Major Depressive Disorder; MCC: Matthews Correlation Coefficient; NLP: Natural Language Processing; ML: Machine Learning; NN: Neural Network; RF: Random Forest; SA: Suicide Attempt; SI: Suicide Ideation; SVM: Support Vector Machine.

Table A. 2: Summary of studies (n=4) involving suicide prevention using behavioral and psychosocial data.

Author, year	Dataset type	Aim of the Study	Study Population	Methods	Findings/Results (%)
Biernesser et al. [38], 2021	Interview	To examine the acceptance of automated monitoring of suicidal teenagers' digital media use within clinical care	United States: Adolescents, Adults, Clinicians	NLP: thematic data analysis	While participants saw potential benefits of automated monitoring, adolescents, parents, and clinicians had concerns about privacy losses, false alerts, and liability issues in responding to alerts of suicidal risk.
Haghish et al. [36], 2023	Survey	False positive study in suicide classification models	Norway: Adolescents	ML: GB	Individuals incorrectly identified as high suicide risk by a screening tool had a much higher rate of suicide attempts versus those correctly identified as low risk.
Kearns et al. [35], 2023	Survey	To assess a suicide-prevention intervention for high-risk transitioning soldiers	United States: Army soldiers, Young Adults, Adults	ML: LR, RF, SVM, XGB	A brief self-report survey can accurately identify transitioning service members at high suicide risk for targeted intervention before discharge to prevent post-transition attempts. AUROC of 0.85
Prelog et al. [37], 2023	Survey	To estimate the presence of SI	Slovenia: Adults, COVID	ML: LR, RF, SVM, XGB	The examined indicators have the potential to be made into a rapid screening tool that could evaluate suicidal thoughts and

behaviors indirectly. AUROC: 0.83 on
unseen data

AUROC: Area Under the Receiver Operator Curve; LR: Logistic Regression; GB: Gradient Boosting; ML: Machine Learning; NLP: Natural Language Processing; RF: Random Forest; SA: Suicide Attempt; SI: Suicide Ideation; SVM: Support Vector Machine; XGB: Extreme Gradient Boosting.

Table A. 3: Summary of studies (n=19) involving suicide prevention using digital and technology data.

Author, year	Dataset type	Aim of the Study	Study Population	Methods	Findings/Results (%)
Moreno et al. [52], 2019	Articles	To study the impact and extent of news coverage about online and in-person bullying.	United States of America	NLP: Linguistic inquiry and word count	Cyberbullying articles are more fear-based than bullying articles.
Iyer et al. [40], 2022	Audio/calls	To automatically classify short speech segments (40 milliseconds) as low or high suicide risk.	Australia	ML: GB	Accuracy: 0.9985.
Iyer et al. [41], 2022	Audio/calls	To automatically classify short speech segments from suicide hotline callers as high or low distress using vocal characteristics and ML.	Australia: adults	ML: GB	AUC: 0.9739.
Kar et al. [54], 2022	Audio/calls	-	India	DL	-
Lake et al. [56], 2022	Chats	To analyze crisis chat transcripts on suicide risk, counselor responses, and pre-chat thoughts v/s later risk.	Adolescents, young adults, adults	ML: statistical analysis	Crisis chat serves young, high-risk users, most endorsing suicide thoughts.
Salmi et al. [51], 2022	Chats	To see if COVID-19 related to changes in suicide problems reported to a helpline.	Netherlands: gender-based, COVID	NLP: BERT	Younger, male, or lone-living help seekers showed changes relating to increased suicidality - warranting close monitoring.

Caicedo et al. [45], 2020	Social media	To evaluate ML models for identifying self-injury in messages.	Miscellaneous	ML: k-star algorithm	AUROC: 0.81036; F1 score: 0.7148.
Chau et al. [42], 2020	Social media	To assess a search tool for professionals to identify emotional distress in blogs.	China: stressed patients	ML: SVM	F1-score: 0.7216.
Fu et al. [53], 2021	Social media	To automatically detect comments with a high risk of suicide.	China	NLP: BERT	Accuracy: 0.7868.
Purushothaman et al. [39], 2021	Social media	To detect conversations about suicide, self-harm and substance use on Instagram.	Drug addicts	DL: LSTM	Substance use disorders and mental health posts are more likely to discuss suicide.
Yang et al. [181], 2021	Social media	To detect SI early for crisis intervention.	China: adults	NLP: Knowledge graphs	"Tree Hole Action" is effective for suicide monitoring and crisis response on social media.
Kumar et al. [55], 2022	Social media	To identify new signs of depression and suicide risk during COVID-19.	COVID-19	ML: LR	COVID-19 lockdowns led to poor mental health due to anxiety.
Lao et al. [43], 2022	Social media	To assess suicide risk on social media.	Miscellaneous	ML: RF, SVM, GB	Linguistic traits linked to suicide risk included self-expression, first-person pronouns, and negation. AUC: 0.66-0.68. F1-score: RF (0.65); GB (0.62); SVM (0.52)
Yang et al. [47], 2022	Social media	To explore time, content, and suicide behavior traits in high-risk comments.	China	ML: knowledge graphs	Negative emotions and thoughts are top reasons given for suicide. Women and city dwellers are more likely to discuss suicide online.
Bello et al. [46], 2023	Social media	To identify if a text discusses suicide.	Spain	ML: CNN	Accuracy: 99.91%.

Deas et al. [50], 2023	Social media	To examine anti-mattering and SI in online peer groups.	Miscellaneous	ML: LR	Monthly Twitter risks temporally aligned with records (correlation=0.533, higher at 0.652 for following month records).
Niederkröten thaler et al. [44], 2023	Social media	To assess relationships between Twitter post topics and calls to Lifeline/completed suicides.	United States of America	ML: LR	Daily social media suicide prevention posts linked to more calls/fewer deaths.
Wang et al. [48], 2023	Social media	To confirm if social media suicide risks align with actual behaviors.	Japan	NLP: BERT	Monthly Twitter risks temporally aligned with records (correlation=0.533, higher at 0.652 for following month records).
Levkovich et al. [49], 2023	Vignette	To evaluate ChatGPT-4's suicide risk assessment vs ChatGPT-3.5.	Adults	NLP: ChatGPT-4, ChatGPT-3.5	ChatGPT-4 estimated risk similar to experts, with better SI recognition than 3.5.

AUC: Area Under the Curve; AUROC: Area Under the Receiver Operator Curve; CNN: Convolutional Neural Network; BERT: Bidirectional Encoder Representations from Transformers; DL: Deep Learning; LDA: Latent Dirichlet Allocation; GB: Gradient Boosting; LIWC: Linguistic Inquiry and Word Count; LSTM: Long-Short Term Memory; NLP: Natural Language Processing; RF: Random Forest; LR: Logistic Regression; ML: Machine Learning; SI: Suicide Ideation; SVM: Support Vector Machine.

Table A. 4: Summary of studies (n=6) involving suicide prevention using administrative and public data.

Author, year	Dataset type	Aim of the Study	Study Population	Methods	Findings/Results (%)
Haroz et al. [57], 2020	Govt data	Determine highest risk subgroups in this community	United States: Adolescents, Young Adults, Adults	ML: LR, DT	AUC: DT (0.81); LR (0.87)
Berkelmans et al. [21], 2023	Govt data	Identify those at extremely high risk for suicide due to combined factors.	Netherlands: Adolescents, Young Adults, Adults, Old people	ML: LR	ML effectively identified extremely high-risk groups for intervention. The LR interaction model had an AUC of 0.76.
Kafka et al. [60], 2023	Govt data	Intimate partner violence is associated with SI	United States: Intimate partner violence,	ML: LR	7.1% of suicides were Intimate partner violence related

Kafka et al. [61], 2023	Govt data	Develop and validate a tool identifying intimate partner violence mentions	Adolescents, Young adults, Adults, Old people United States: Intimate partner violence, Adolescents, Young adults, Adults, Old people	ML: RF	Model had robust sensitivity (0.70), specificity (0.98), precision (0.72) and kappa values (0.69).
Wang et al. [59], 2023	Govt data	Outline rationale and methods for creating models predicting population suicide risk. Implementing a suicide prevention program within a large Australian health service. It outlined the process and reported early outcomes and lessons learned.	Canada	ML: LR	-
Turner et al. [58], 2021	Public records		Australia	ML: Evolutionary algorithm	The model detects suicidal/self-harm presentations with high accuracy (95% sensitivity, 92% specificity).

AUC: Area Under the Curve; DT: Decision Tree; IPV: Intimate Partner Violence; LR: Logistic Regression; ML: Machine Learning; NLP: Natural Language Processing; RF: Random Forest; SI: Suicide Ideation.

Table A. 5: Summary of studies (n=10) involving suicide prevention using multimodal data.

Author, year	Dataset type	Aim of the Study	Study Population	Methods	Findings/Results (%)
Berrouiguet et al. [67], 2019	Body sensors/wearables	Determine the link between suicide risk, sleep changes and appetite issues	Spain, France: Adults	ML: LR	This protocol helps shift from typical risk identification to tailored prevention strategies per patient characteristics.

Ranapurwala et al. [71], 2022	Govt data, Public records, Electronic health records	Estimate factor-outcome ties, develop predictors via workflow linking long-term data to ensure continuous monitoring	United States	ML	-
Balbuena et al. [66], 2022	Hospital data	Identify risk/protective factors in the general population signaling near-term suicide risk	Norway: Young Adults, Adults	ML: RF	ML can identify early prevention targets. 0.57AUC
Kessler et al. [62], 2023	Hospital data, Govt data	Determine if adding clinical notes and public records can boost prediction	United States: Veterans, Adults, Old people	ML: Shapely scores for XAI	ML model's ability to predict suicide after psychiatric discharge improved significantly with predictors from clinical notes using NLP and social determinants public records.
Heckler et al. [63], 2023	Questionnaire, Body sensors/wearables	Thoth, an intelligent model helps those with suicidal thoughts by analyzing context like sensor, psychological and sociodemographic info	Brazil	ML: ANN	F1-scores of 94.12%, 92.44%, 92.61% for predicting next-day suicidal thoughts given histories of 3, 5, and 7 days.
Mouchabac et al. [64], 2021	Questionnaires	Develop autonomous care for suicide prevention in patients with prior SA.	Adolescents	ML: Bayes Net	Bayes Net efficiently builds a digital assistant algorithm for suicide crisis management
Kyron et al. [69], 2021	Self-reports, Questionnaires, Hospital data	How prevention could use short-term modeling and prediction of self-harm factors	Australia: Adults	ML: LR	The model detected 77.7% of self-harm (sensitivity). Psychological distress, will to live/die and relationships reciprocally linked that day. AUC: 0.85
Sels et al. [68], 2021	Self-reports, Sensors	Design digital metrics using self-reports and passive mobile data to predict suicidal thoughts and psychiatric readmissions in high-risk individuals	Switzerland	ML: RF, SVM	-

Bello et al. [70], 2024	Social media, Newspapers	Study COVID-19 and suicide link in media through NLP	COVID	NLP: Naive Bayes, SVM	Strong relationship between the appearance of subjects' "health" and "covid", and also between "covid" and "suicide". F1-score: Suicide (1); Covid (0.99); Multilabel (0.73)
Liu et al. [65], 2019	Social media, Questionnaires	Test feasibility and acceptability of Proactive Suicide Prevention Online	Chinese: Young adults	ML: LR, RF, DT, SVM	10.32% of comments accurately classified as suicidal, with 88% accuracy. 89.73% reported suicidal thoughts in surveys.
AUROC: Area Under the Receiver Operator Curve; ANN: Artificial Neural Network; AUC: Area Under the Curve; Decision Tree; LR: Logistic Regression; ML: Machine Learning; RF: Random Forest; NLP: Natural Language Processing; SI: Suicide Ideation; SVM: Support Vector Machine.					

Table A. 6: Summary of studies (n=25) involving suicide prediction using clinical data.

Author, year	Dataset type	Aim of the Study	Study Population	Methods	Findings/Results (%)
Nordin et al. [82], 2023	Clinical data	To create an explainable predictive model for self-harm and analyze important features.	Malaysia: depressed patients	ML: RF, GB	Accuracy: RF (0.855); GB (0.86).
Huang et al. [73], 2022	Clinical data	To develop and validate a dynamic nomogram predicting SA personalized for depression patients.	China: MDD, young adults, adults	ML: LR	The ML model showed good overall accuracy (Harrell's C-index = 0.823) and clinical applicability by decision curve analysis.
Margaroli et al. [78], 2020	Clinical notes	To detect suicide content in therapy communication.	Miscellaneous	ML: LR, RF, DT, XGB, MLP	AUC: 0.8278.
Cusick et al. [89], 2021	Clinical notes	To detect "current" SI from unstructured clinical notes.	United States of America: adolescents, young adults, adults	DL: CNN	Accuracy: 0.94; F1-score: 0.82.
Hasey et al. [72], 2020	EEG	To detect SI using EEG data.	-	NLP: RF	AUC: 0.768; Accuracy: 0.705.

Morrow et al. [91], 2022	EHRs	To identify life events in notes more than coded events.	United States of America: veterans, young adults, adults, old people	ML: LR	Adding NLP variables to the REACH VET model increased life event variable concentration in higher risk percentiles when discriminating suicide deaths from natural deaths with no mental health history.
Mens et al. [80], 2020	EHRs	To recognize suicidal behavior in primary care using routine data.	Netherlands: adults	ML: RF	Sensitivity: 0.39; AUC: 0.85.
Carson et al. [11], 2019	EHRs	To identify suicidal behavior in hospitalized teens.	United States of America: adolescents, depressed patients	NLP: RF	Accuracy: 0.47.
Obeid et al. [88], 2020	EHRs	To improve detection of self-harm patients and predict future events.	United States of America	NLP: CNN	AUC: 0.882; F1-score: 0.769.
Xie et al. [92], 2023	EHRs	To develop an NLP algorithm to identify SI/SA.	United States of America	NLP: Rule-based	Sensitivity: 88.5%.
Levis et al. [90], 2022	EHRs	To analyze clinical notes and create a data-driven model predicting suicide risk.	United States of America: veterans, adults, old people	NLP: DT, RF, XGB	NLP-derived models demonstrated strong predictive accuracy.
Xu et al. [2], 2020	EHRs	To improve individual self-harm prediction.	China, Hong Kong	DL: LSTM, DNN	PPV: 0.54; NPV: 0.98; Recall: 0.72; Specificity: 0.96.
Jiang et al. [83], 2023	EHRs	To predict SA.	Denmark: depressed patients, adolescents, young adults, adults	ML: LR, RF	AUC: RF (0.8-0.84); LR (0.66-0.68).
Peng et al. [75], 2023	Genetic data	To identify novel targets and pathways differentiating depression, suicide from comorbid cases.	Adults, old people	ML: RF	The study identified novel gene targets and patterns differentiating depression subgroups and suicide.

Song et al. [74], 2022	Hospital data	To develop and validate a model to predict SI using clinical data.	Korea: post-stroke patients, young adults, adults, old people	ML: XGB, CatBoost, LGBM NLP: Naive Bayes, LR, KNN, RF, GB, LSTM	AUC: 0.900. Rehabilitation motivation is closely associated with SI.
Rozova et al. [81], 2022	Hospital data	To automate emergency-note screening directly for self-harm.	Australia	ML: Naive Bayes, LR, KNN, RF, GB, LSTM	Best performance: GB. F1-score: 0.764.
McCullough et al. [79], 2024	Hospital data	To develop a data-driven framework analyzing information flow in emergency psychiatric care to evaluate clinical practice and models.	Australia	ML: Naive Bayes	Balanced accuracy: 0.82.
Cho et al. [84], 2020	Hospital data	To identify high-risk groups in the general public.	Korea: adults	ML: RF	AUC: 0.818.
Weng et al. [76], 2020	MRI	To predict people with SI using brain imaging.	Taiwan: MDD, young adults, adults	ML: LR, XGB, CNN	AUC: 0.84; Accuracy: 0.73.
Chen et al. [93], 2021	MRI	To create an algorithm to detect SI and SA.	Taiwan: previous suicide attempters, MDD, adults	DL: CNN, DenseNet	Accuracy: 0.886 for CNN; 0.912 for DenseNet.
Zhu et al. [85], 2020	MRI	To identify suicidal BD-II patients.	China: MDD, BD, adults	ML: SVM	Accuracy: 0.84.
Hong et al. [3], 2021	MRI	To classify suicide risk.	China: adolescents, young adults, MDD	ML: SVM	Structural MRI is useful for classifying suicide risk. Prediction balanced accuracy: 0.7859.
Xu et al. [87], 2022	MRI	To use dFNC for detecting suicidality in depressed patients.	China: MDD, young adults, adults	ML: SVM	-
Bajaj et al. [86], 2023	MRI	To identify brain structural changes underlying suicide risk in adolescents.	United States of America: adolescents, MDD	ML: SVM	Accuracy: 0.7479.

Stange et al. [77], 2020	MRI	To identify suicide risk in mood disorders.	United States of America: young adults, mood disorder patients	DL: CNN	Extracted data accurately classified suicidal behavior groups (79-88% measures).
-----------------------------	-----	---	--	---------	--

AUC: Area Under the Curve; BD: Bipolar Disorder; CNN: Convolutional Neural Network; MLP: Multilayer Perceptron; DNN: Deep Neural Network; DT: Decision Tree; EEG: electroencephalography; EHRs: Electronic Health Records; GB: Gradient Boosting; LGBM: Light Gradient Boost Model; XGB: Extreme Gradient Boosting LR: Logistic Regression; MDD: Major Depressive Disorder; LSTM: Long-Short Term Memory; ML: Machine Learning; NLP: Natural Language Processing; NS: without suicide risk; PPV: Positive Predicted Value; RF: Random Forest; SA: Suicide Attempt; SI: Suicide Ideation; SVM: Support Vector Machine.

Table A. 7: Summary of studies (n=17) involving suicide prediction using behavioral and psychosocial data.

Author, year	Dataset type	Aim of the Study	Study Population	Methods	Findings/Results (%)
Broadbent et al. [95], 2023	Counselling data	To identify client suicide risk.	United States of America (Utah, Idaho, and Nevada)	NLP: BERT	AUC: 0.9037.
Pestian et al. [183], 2020	Interview	To monitor SI up to 30 days after care.	United States of America: adolescents, young adults, adults	ML: SVM	AUC: 0.89.
Haroz et al. [185], 2021	Interview	To create a clinical decision tool reservation-based suicide surveillance.	United States of America	ML	Clinical use of ML-based suicide prediction algorithms is elusive.
Shin et al. [187], 2022	Interview	To identify high-risk suicide groups through spoken words in semi-structured interviews.	United States of America: young adults, adults, mood disorder	ML: Naive Bayes	AUC: 0.800.
Li et al. [94], 2023	Interview	To determine if self-harm acts can be detected from interview	China: depressed patients, adults	ML: RF	AUC for high suicide risk identification was 0.76-0.89. AUC for general risk detection was 0.83-0.92.

		language features using NLP/ML.			
Wright-Berryman et al. [188], 2023	Interview	To evaluate model performance identifying depression, anxiety, suicide risk from brief interviews.	United States of America: adults	ML: LR, SVM, XGB	AUC: 0.70 for SVM.
Ghosh et al. [186], 2022	Personal notes	To focus on learning depression detection, sentiment, emotion from unstructured text.	Children, adolescents, young adults, adults, old people	DL: Attention-based model	Recall: 0.5647.
Choi et al. [184], 2021	Questionnaire	To detect self-harm.	Korea: young adults, COVID	DL: Graph NNs	AUC: 0.878.
Bantilan et al. [96], 2021	Self-report	To detect suicide content from therapy communications.	United States of America: depressed patients	ML: LR	AUC: 0.8278.
Lin et al. [98], 2020	Survey	To predict SI in military using six psychological stress domains by sex.	Taiwan: military personnel, adults	ML: LR, DT, RF, SVM, MLP, XGB	The proposed algorithms improved accuracy for predicting presence of suicide ideation (SI) ≥ 1 by up to 5.7% compared to the conventional BSRS-5 score ≥ 7 criterion. For predicting more severe SI ≥ 2 , accuracy improved by 6.1%.
Littlefield et al. [101], 2021	Survey	To classify military thoughts and behaviors using consortium data elements.	United States of America: military personnel, young adults, adults	ML: RF, LR	Models had similar classification using different features.
Mahmud et al. [103], 2023	Survey	To predict suicidal behaviors and key predictors among Bangladeshi students during COVID-19.	Bangladesh: university students, young adults, COVID	ML: SVM, RF, DT, Naive Bayes, KNN	SVM: Accuracy (0.79), specificity (0.81), receiver operating characteristic (0.89), Kappa (0.59), and sensitivity (0.81).

Zuromski et al. [102], 2020	Survey	To detect self-harm.	United States of America: army soldiers, young adults, adults	LR, RF, GB, Naive Bayes	AUC scores between 0.75 and 0.77.
Weller et al. [165], 2021	Survey	To develop an algorithm predicting youth suicidal thoughts/acts using risk framework and social determinants.	United States of America: adolescents	ML: LR, DT, KNN, Naive Bayes, XGB, LGBM	Accuracy of LGBM model: 0.91.
Jung et al. [99], 2019	Survey	To develop a model to identify high-risk suicidal Korean teens.	Korea: adolescents	ML: LR, RF, SVM, ANN, XGB	Accuracy: XGB (0.79); SVM (0.787); LR (0.779); RF (0.778); ANN (0.775).
Morales-Rodríguez et al. [97], 2023	Survey	To examine variables linked to suicide risk and predictive ability.	Spain: students	DL: ANN	Emotional attention contributed most to ANN prediction, followed by problem solving and perfectionism.
Su et al. [100], 2023	Survey, interview	To predict self-harm and suicide act risk in adolescents.	Australia: adolescents	ML: RF	The most important factors across models included depressed feelings, SDQ scores, self-perceptions, and school/parent factors. AUROC for self-harm prediction was 0.7397 and SA prediction was 0.7220.

AUC: Area Under the Curve; MLP: Multilayer Perceptron; ANN: Artificial Neural Network; DT: Decision Tree; AUROC: Area Under the Receiver Operator Curve; ML: Machine Learning; XGB: Extreme Gradient Boosting; BERT: Bidirectional Encoder Representations from Transformers; GB: Gradient Boosting; LGBM: Light Gradient Boost Model; LR: Logistic Regression; NLP: Natural Language Processing; NN: Neural Network; RF: Random Forest; SA: Suicide Attempt; SVM: Support Vector Machine.

Table A. 8: Summary of studies (n=25) involving suicide prediction using digital and technology data.

Author, year	Dataset type	Aim of the Study	Study Population	Methods	Findings/Results (%)
--------------	--------------	------------------	------------------	---------	----------------------

Onie et al. [109], 2023	CCTV videos	To examine acceptability and feasibility of automated crisis detection.	Australia	DL: computer vision models	The algorithm correctly identified crisis clips 80% of the time and rejected non-crisis clips 90% of the time.
Xu et al. [112], 2021	Online Conversations	To detect SI in online counseling systems.	China, Hong Kong, young adults, adults	NLP: Graph-encoders	Baseline LSTM.
Liu et al. [111], 2020	Social media	To summarize the research on SI causes using psychological and sociological analysis.	Chinese	NLP: CRF, Char-BiLSTM-CRF, BERT	The experiment showed word features worked best. F1-score: 0.79.
Ophir et al. [107], 2020	Social media	To predict SI from tone of posts and chats on social media.	Adults	DL: DNN	AUC: 0.697-0.746.
Shen et al. [5], 2020	Social media	To predict SI in medical students.	China: students, adolescents	ML: RF	AUC: 0.9255; Accuracy: 0.901.
Yao et al. [116], 2020	Social media	To extract suicidal posts among opioid Reddit users.	Drug addicts	ML: LR, RF, SVM, RNN, CNN	F1-score: LR (0.940); RF (0.927); SVM (0.943); RNN (0.952); CNN (0.961).
Lekkas et al. [104], 2021	Social media	To predict acute SI on Instagram.	Germany: adolescents	ML: XGB, DT, KNN, NN	Accuracy: 0.702.
Ragheb et al. [117], 2021	Social media	To detect at-risk users with depression, eating disorder, self-harm, suicidal thoughts.	Miscellaneous	DL: BERT, XLNet	F1-score: 0.79.
Ramírez-Cifuentes et al. [120], 2021	Social media	To analyze text distinguishing Reddit posts on depression, eating disorders, SI, alcohol from controls.	Miscellaneous	ML: RF, LR	Accuracy: 0.9723.
Schoene et al. [115], 2023	Social media	To predict SI from depression notes.	Miscellaneous	NLP: LSTM	F1-score: Distinguish suicide posts (0.8266); Suicide classification (0.961)

Zhang et al. [15], 2021	Social media	To detect SI online.	Miscellaneous	DL: Transformers, LSTM	F1-score: 0.949.
Baghdadi et al. [121], 2022	Social media	To detect SI online.	Arabic	NLP: BERT	Best-weighted sum metric.
Du et al. [105], 2022	Social media	To classify texts of various psychological states.	Miscellaneous	ML: SVM, ANN	Accuracy: SI: (0.882 with DL); texts of anxiety (0.863 on DL); texts of depression (0.835 on SVM).
Mbarek et al. [106], 2022	Social media	To detect profiles with SI to prevent suicide.	Miscellaneous	ML: RF, Bayes Net, SVM, DT, AdaBoost	F1-score: Bayes Net (0.748); SVM (0.829); RF (0.854); AdaBoost (0.748); DT (0.785).
Metzler et al. [117], 2022	Social media	Classify social media data by characteristics linked to suicide effects in research	Miscellaneous	DL: BERT, XLNet	F1-score was 0.55 for the suicidal ideation/acts category. Accuracy of 88% for classifying suicide and non-suicide related tweets.
Naseem et al. [108], 2022	Social media	Explain social media suicide risk identification	Miscellaneous	DL: Transformer	F1-score: 0.79.
Renjith et al. [124], 2022	Social media	To detect individuals with suicidal thoughts	Miscellaneous	DL: LSTM, Attention CNN	Accuracy: 0.903; F1- score: 0.926.
Schoene et al. [122], 2023	Social media	To classify expressions of suicidality online	Miscellaneous	DL: LSTM, BERT, CNN	With text and NLP features CNN F1-score: 0.91.
Adarsh et al. [125], 2023	Social media	To distinguish between depression symptoms and suicidal ideas	young adults, adults	ML: KNN, SVM	Accuracy: 98.05%.
Buddhitha et al. [123], 2023	Social media	To detect SI online.	America: young adults, adults	ML: CNN	F1-score: 0.889.
Burkhardt et al. [114], 2023	Social media	To detect SI online.	America: young adults, adults	NLP: BERT	F1-score: 0.797.

Ghosh et al. [110], 2023	Social media	To detect SI online on Bangla social media texts.	Bangladesh	DL: CNN	Accuracy: 0.943.
Rabani et al. [119], 2023	Social media	To develop multi-class model identifying social media suicide risk levels.	Miscellaneous	ML: SVM, RF, XGB	Highest accuracy: XGB 0.9633.
Talaat et al. [118], 2023	Social media	To propose enhanced system recognizing personality and emotion.	Miscellaneous	ML: LR, RF, SVM, DT, kNN	Accuracy: personality detection on RF: 0.9999; emotion recognition on DT: 0.8806.
Li et al. [189], 2024	Social media	To estimate depression's impact on suicide risk.	Depressed patients	ML: DT	Depressed individuals had 18% higher suicide risk than non-depressed.

AUC: Area Under the Curve; ANN: Artificial Neural Network; CCTV: closed-circuit television; XGB: Extreme Gradient Boosting; CNN: Convolutional Neural Network; DL: Deep Learning; DNN: Deep Neural Network BERT: Bidirectional Encoder Representations from Transformers;; DT: Decision Tree; KARA: Knowledge-Aware Risk Assessment; LR: Logistic Regression; LSTM: Long-Short Term Memory; ML: Machine Learning; RF: Random Forest; NLP: Natural Language Processing; CRF: Conditional Random Field; RNN: Recurrent Neural Network; SA: Suicide Attempt; SI: Suicide Ideation; SVM: Support Vector Machine.

Table A. 9: Summary of studies (n=10) involving suicide prediction using administrative and public data.

Author, year	Dataset type	Aim of the Study	Study Population	Methods	Findings/Results (%)
Kim et al. [10], 2019	Govt data	Develop accurate, interpretable model identifying patients at high SA risk.	United States	ML: Attention NN	The model identified patients diagnosed with self-harm or suicide on next visit with AUC of 0.89
Gradus et al. [127], 2021	Govt data	Examine sex-specific Danish registry profiles for suicide act risk	Danish: gender-based, Adolescents, Young Adults, Adults	ML: RF	Substance use, psych meds, prior poisoning diagnoses and stress disorders predicted SA risk for both sexes. The top 5% of predicted risk accounted for 44.7% of male and 43.2% of female SA
Horvath et al. [131], 2021	Govt data	To identify suicidal behavior in a prison population	United States: prisoners	ML: GB	AUC: 0.875; F1-score: 0.846

Jiang et al. [126], 2021	Govt data	Predict suicide among depressed persons using ML	Denmark: gender-wise, Depressed patients, Adults	ML: DT, RF	AUC RF (0.75 women, 0.70 men)
Lin et al. [129], 2022	Govt data	Predict future multiple suicide acts in psychiatric patients.	Taiwan: Psychiatric patients	ML: DT, SVM, ANN	Accuracy: 0.983
Rahman et al. [132], 2022	Govt data	Predict suicide in undetermined cases using known homicide and suicide data.	United States: Race-focus	NLP, ML: LIWC, LR, RF, NN	Findings showed comparable Black and White suicide rates, indicating excess Black misclassification. This could be mitigated using race-specific data. AUC: 0.987
Kwon et al. [130], 2023	Govt data	Develop ML model predicting adolescent SA	Korea: Adolescents, COVID	ML: RF, XGB, LGBM, AdaBoost	Balanced Accuracy – AUC: RF (0.786 - 0.85); AdaBoost (0.788 - 0.86); LGBM (0.784 - 0.86); XGB (0.78 - 0.87)
Shin et al. [134], 2023	Govt data	Quickly predict suicidal thoughts and suggest countering social risks.	Korea: Children, Adolescents	DL: LR, RF, XGB, MLP, CNN	Best performance: CNN. Accuracy: 0.8833; AUC: 0.9370
Wang et al. [133], 2023	Govt data	Identify individual social risks from death narratives.	United States, Puerto Rico: Adults	NLP: BERT	AUC: classifying circumstances (0.966); classifying crises (0.942)
Laqueur et al. [128], 2022	Handgun purchase records	Test if handgun records plus ML forecast firearm suicide risk	California: Adults	ML: RF	In a study of nearly 2 million with handgun records, close to 40% of the top 5% highest risk transactions led to purchaser firearm suicide within 1 year. Over two-thirds of transactions over 0.95 RF score ended in purchaser firearm suicide within 1 year (24 of 35).

AUC: Area Under the Curve; ANN: Artificial Neural Network; MLP: Multilayer Perceptron; DT: Decision Tree; GB: Gradient Boosting; XGB: Extreme Gradient Boosting; LGBM: Light Gradient Boosting Model; LIWC: Linguistic Inquiry Word Count; LR: Logistic Regression; ML: Machine Learning; NLP: Natural language Processing; NN: Neural Network; RF: Random Forest; SA: Suicide Attempt; SDoH: Social Determinants of Health; SVM: Support Vector Machine; BERT: Bidirectional Encoder Representations from Transformers.

Table A. 10: Summary of studies (n=28) involving suicide prediction using multimodal data.

Author, year	Dataset type	Aim of the Study	Study Population	Methods	Findings/Results (%)
Lozupone et al. [137], 2022	Biomarkers, questionnaires, hospital data	To investigate biomarker, social, psychological, and metabolic relationships to SI in older subjects.	Italy: old age people	ML: LR	The AUC for detecting SI in LLD and its subtypes (LOD and EOD) were 0.8058, 0.6362, 0.6693 respectively.
Allen et al. [1], 2019	Body sensors/wearables	To predict SI using body sensors and wearables.	Adolescents	ML: SVM	Highlights opportunities for innovation through smartphone, wearable, and smart home technology to enable early identification and intervention for suicide.
Miranda et al. [152], 2024	EHRs	To predict adverse events for patients with comorbid PTSD or alcohol addiction.	United States: PTSD, alcohol addicts	DL: Transformer, DNN	Medications and therapies impacted risk, while veteran status, segregation, limited parks/English, younger age, and income increased risk. AUC: 0.94.
Yarborough et al. [148], 2022	EHRs	To evaluate if opioids improve suicide risk models from Mental Health Research Network.	United States: drug addicts	ML: Regression	Opioid variables do not improve suicide risk prediction beyond mental health predictors for diagnosed patients.
Shao et al. [12], 2022	EHRs, MRI	To understand key neural mechanisms for suicidal thoughts and actions.	Taiwan: old people	ML: SVM	Neural features predict SI better than clinical data or questionnaires alone.
Wilimitis et al., 2022 [155]	EHRs, Questionnaires	To evaluate ML model and C-SSRS abilities to individually and jointly predict SA, SI.	United States: adults	ML: LR, RF	Combined models had better AUROC (above 0.86) than EHR or questionnaire-based models.
Malhi et al. [190], 2019	fMRI, clinical interview	To study neural factors behind emotional cognitive flexibility in MDD patients and their association with SI.	Australia: mood disorder patients, adults	ML: ANOVA	Cognitive–emotional deficits may cause suicidal activity.

Hao et al. [154], 2023	fMRI, interviews	To predict SI using multimodal data.	China: MDD, adolescents, young adults, adults	ML: KNN, LR, RF, SVM	Multimodality classifiers (63-92% accuracy) significantly outperformed brain-only (56-85%) or behavior-only (64-73%) models.
Joo et al. [136], 2022	Genetic data, self-reported questionnaire	To assess if polygenic scores and stress associate with youth suicide risk.	United States of America: children	ML: LR, RF	Accuracy: STB - 0.67; SI - 0.66.
Kim et al. [156], 2021	Government data, Interview	To evaluate MMPI-2's utility in suicide risk assessment using MMPI-2 and risk evaluation results.	United States of America: students, young adults	ML: RF, KNN	Accuracy SI, SA: KNN (0.916, 0.947); RF (0.929, 0.95).
Pavicic et al. [147], 2020	Government data, Survey	To predict SA using data from the Million Veteran Program.	United States of America: veterans, young adults, adults, old people	ML: RF	Married men living with spouses have lower suicide attempt rates.
Edgcomb et al. [144], 2023	Hospital data	To assess ICD codes and chief complaints vs clinician review to detect self-injury in children.	United States of America: children	ML: LR, RF	Applying ML to health records may strengthen detecting self-injury in children v/s codes alone.
Xu et al. [135], 2022	Hospital data, genetic data	To use clinical and genetic factors to predict MDD to BD conversion.	China: MDD, bipolar disorder, adults	ML: LR	SA and hospitalizations associated with MDD to BD conversion.
Haines-Delmont et al. [139], 2020	Interviews, body sensors/wearables	To study suicide risk prediction.	United Kingdom	ML: KNN	Accuracy: 0.68.
Qin et al. [138], 2024	MRI, genetic data	To study the association of structural covariance network and STB in individuals with MDD.	China: MDD, adults, young adults	ML: SVM	Balanced accuracy: 0.625; AUC: 0.676.

Vogt et al. [146], 2023	Public data, self-reports	To predict US veterans' suicidal ideation.	United States: veterans, adults	ML: RF	Well-being factors predicted, though not as well as psychopathology by self-report, for veterans' SI prediction.
Yang et al. [145], 2023	Questionnaire, biochemical sample, hospital data	To accurately detect suicide attempt risk in MDD patients.	China: MDD, adults, young adults	ML: Regression	AUROC: 0.720.
Barzilay et al. [149], 2023	Self-reports, body sensors/wearables	To examine passive mobile data and clinical/self-reported suicide link over 6 months.	Israel: adolescents	DL: LR, RF, VAE	-
Rozek et al. [151], 2020	Self-reports, clinical data	To predict SI in soldiers.	United States of America: military personnel, young adults, adults	ML	Sensitivity: 0.308; Specificity: 0.992; Accuracy: 0.873.
Ramírez-Cifuentes et al. [142], 2020	Social media	To assess Spanish social media users' suicide risk.	Spain	ML: LR, RF, SVM, MLP, CNN	Accuracy: SVM (0.84).
Cao et al. [13], 2022	Social media	To establish key individual SI risk factors.	China: depressed patients, bipolar, young adults	DL: CNN	Accuracy: 0.93.
Chatterjee et al. [141], 2022	Social media	To provide early SI detection by evaluating social media.	Miscellaneous	ML: SVM, RF, DT, LR	AUC: LR (0.86); RF (0.83); SVM (0.87); DT (0.77).
Cai et al. [153], 2023	Social media	To detect depression.	China	DL: CNN	F1: 0.92.
Li et al. [157], 2023	Social media	To predict SI using social media data.	China	DL: RF, KNN, SVM, CNN	Accuracy: 0.8875 for CNN.
Ghanadian et al. [158], 2024	Social media	Generation of synthetic data about SI.	Miscellaneous	NLP: ChatGPT, Flan-T5, and Llama	F-1 score: 0.82.

Choi et al. [150], 2023	Social media, economic data.	To estimate weekly US suicide deaths in near real time.	United States of America	ML: NN	By combining multiple data streams, the ML method estimated weekly suicide deaths with high correlation (0.811) and annual suicide rates with low error (0.55%) to actual counts and trends.
Fulginiti et al. [143], 2022	Social media, interview	To predict SI and SA.	United States: young people experiencing homelessness	ML: DT	Accuracy: SI (0.68); SA (0.71).
Lyu et al. [140], 2023	Social media, questionnaires	To use linguistic features in English social media for depression detection.	China: adolescents, young adults, adults	ML: LR, SVM, RF, GB	Best performance: LR. Pearson correlation coefficient: 0.33; R-squared: 0.10; split-half reliability was 0.75.

DNN: Deep Neural Network; BP: Bipolar Disorder; CNN: Convolutional Neural Network; DL: Deep Learning; DT: Decision Tree; GB: Gradient Boosting; LR: Logistic Regression; MDD: Major Depressive Disorder; ML: Machine Learning; SVM: Support Vector Machine; MMPI-2: Minnesota Multiphasic Personality Inventory-2; NLP: Natural Language Processing; PTSD: Post-Traumatic Stress Disorder; RF: Random Forest; SA: Suicide Attempt; SDoH: Social Determinants of Health; SI: Suicide Ideation; STB: suicidal thoughts and behaviors; VAE: Variational Autoencoder; MLP: Multilayer Perceptron; ICD-10-CM: related International Classification of Diseases, Tenth Revision, Clinical Modification; NN: Neural Network.

Table A. 11: List of studies utilizing different AI-based techniques.

AI Technique	Model	Number of Papers	References
DL	CNN	16	[13, 46, 76, 77, 88, 89, 93, 110, 115, 116, 123, 124, 134, 142, 153, 157]
	Attention-based models	5	[10, 15, 108, 124, 152, 186]
	Transfer learning models	4	[93, 109, 113, 117]
DL/NLP	BERT	12	[48, 51, 53, 95, 111, 113, 114, 115, 117, 121, 133]
	LSTM	6	[2, 15, 39, 81, 111, 115, 124]
ML	LR	37	[14, 22, 35, 37, 44, 50, 55, 57, 59, 60, 65, 67, 69, 73, 75, 78, 83, 91, 96, 98, 99, 101, 102, 116, 118, 120, 135, 136, 137, 140, 141, 142, 144, 154, 155, 165, 188]

NLP	RF	53	[5, 11, 14, 16, 28, 35, 37, 43, 61, 65, 66, 68, 72, 75, 78, 80, 81, 82, 83, 84, 90, 94, 99, 100, 101, 102, 103, 106, 111, 116, 118, 119, 120, 126, 127, 128, 130, 132, 134, 136, 140, 141, 142, 144, 146, 147, 149, 154, 155, 156, 157]
	SVM	35	[1, 3, 7, 12, 14, 16, 33, 35, 37, 42, 43, 65, 68, 70, 85, 86, 87, 98, 99, 103, 105, 106, 116, 118, 119, 125, 129, 138, 140, 141, 142, 154, 157, 183, 188]
	XGB	14	[35, 37, 74, 76, 78, 90, 98, 99, 104, 119, 130, 134, 165, 188]
	KNN	8	[103, 104, 118, 125, 139, 154, 156, 165]
	DT	15	[31, 57, 65, 78, 90, 98, 104, 106, 118, 126, 129, 141, 143, 189]
	Ada Boost	3	[16, 106, 130]
	Naïve Bayes	7	[64, 79, 102, 103, 104, 165, 187]
	Word embeddings	1	[29]
	Rule-based models	2	[30, 52, 92]
	LLMs	3	[27, 49, 158]
	Graph-based models	2	[112, 182]

Table A. 12: List of related systematic review studies.

Title	Reference	Year of Publication	Representative Paper(s) in Figure 1
Natural Language Processing in Electronic Health Records in relation to healthcare decision-making: A systematic review	[191]	2023	[17]
Textual emotion detection in health: Advances and applications	[192]	2023	[17]
Mental Health Analysis in Social Media Posts: A Survey	[193]	2023	[17, 23]
Mining social networks to improve suicide prevention: A scoping review	[194]	2020	[17, 23]
Deep learning techniques for suicide and depression detection from online social media: A scoping review	[195]	2022	[17, 23]
Legal, ethical, and wider implications of suicide risk detection systems in social media platforms	[196]	2021	[17, 23]

Artificial intelligence and machine learning in mobile apps for mental health: A scoping review	[197]	2022	[22]
The Use of Closed-Circuit Television and Video in Suicide Prevention: Narrative Review and Future Directions	[198]	2021	[23]
Freely Available Training Videos for Suicide Prevention: Scoping Review	[199]	2023	[23]
The use of advanced technology and statistical methods to predict and prevent suicide	[200]	2023	[23]
Artificial intelligence and suicide prevention: a systematic review	[201]	2022	[22, 24]
Emerging Artificial Intelligence-Empowered mHealth: Scoping Review	[202]	2022	[22, 24]
Artificial intelligence for suicide assessment using Audiovisual Cues: a review	[203]	2023	[22, 23, 24]
Suicidal Ideation Detection: A Review of Machine Learning Methods and Applications	[204]	2021	[25]
Can machine-learning methods really help predict suicide?	[205]	2020	[25]
Suicidal behaviour prediction models using machine learning techniques: A systematic review	[206]	2022	[25]
Intervention targets for reducing mortality between mid-adolescence and mid-adulthood: a protocol for a machine-learning facilitated systematic umbrella review	[207]	2023	[25]
Machine learning for suicidology: A practical review of exploratory and hypothesis-driven approaches	[208]	2020	[22, 25]
Machine learning for suicidal ideation identification: A systematic literature review	[19]	2022	[22, 25]
The performance of machine learning models in predicting suicidal ideation, attempts, and deaths: A meta-analysis and systematic review	[209]	2022	[22, 25]
Classification of suicidality by training supervised machine learning models with brain MRI findings: A systematic review	[210]	2023	[22, 25]
The use of machine learning in the study of suicidal and non-suicidal self- injurious thoughts and behaviors: A systematic review	[211]	2019	[22, 25]

Neural Correlates Associated With Suicide and Non suicidal Self-injury in Youth	[212]	2021	[24]
Practices for monitoring and responding to incoming data on self-injurious thoughts and behaviors in intensive longitudinal studies: A systematic review	[213]	2021	[23]
Impact of Big Data Analytics on People's Health: Overview of Systematic Reviews and Recommendations for Future Studies	[214]	2021	[24]
Exploration of EEG-Based Depression Biomarkers Identification Techniques and Their Applications: A Systematic Review	[215]	2022	[24]
Leveraging data science to enhance suicide prevention research: a literature review	[216]	2022	[22, 24]
