Journal of Emerging Intelligence and Engineering Technologies

Article

Mental Health Risk Detection Through Multimodal Analysis of Daily Digital Footprints

Mahesh Marri[1], Sojy John[2], Venkata Phaneendra Reddi[3], Pradeep Rao Vennamaneni[4], Sri Prudhvi Adusumilli[5], Naveen Shrivastava[6]
[1]Amazon Web Services
[2]Money Services, Inc.
[3]Atlantic Health System
[4]Inovalon Inc.
[5]Charter Communications
[6]AAA

Abstract

Despite the high prevalence of mental health disorders, including depression, anxiety and suicidality, and their burden, these conditions are widely underdiagnosed worldwide. Current clinical evaluation depends on periodic self-reported symptom checklists, which provide only a snapshot of the patient's functioning over a short period of time. The paper collates data from the 17 studies, which were published between 2013 and 2021, to explore the potential to integrate the three types of passive digital footprints — social media language/imagery, smartphone sensor logs and wearable physiological data — to enable continuous unobtrusive detection of mental health risks. The synthesis is structured around three modalities that are assessed separately in the literature: analysis of text-based social media, behavioural sensing from a smartphone, and physiological monitoring from wearables, and then how multimodal fusion architectures combine these signals. Results showed high variance in detection performance, ranging from a language-based model trained on Facebook data with an AUC score of 0.62 to the one trained on Twitter data with a score of 0.72, depending on the richness of the features; the best unimodal models were based on behavioral entropy from smartphone GPS and home-stay duration, showing AUC as high as 0.83; wearable-only models exhibited comparatively low discriminative power, ranging from 0.51 to 0.66 AUC; the highest F1-scores for cooperative multimodal fusion models were as high as 0.900, which corresponds to approximately a 30 percent relative error reduction over unimodal baselines. The synthesis also highlights common method or ethics challenges such as the use of self-reported or crowd-sourced ground truth, selection bias of active platform users, and unanswered privacy and consent issues in passive sensing. The results indicate that integration of multimodal data, as opposed to any particular digital footprint, may be best suited for scalable mental health risk screening in a clinically translatable way.

Keywords: Digital phenotyping, depression detection, multimodal fusion, social media analysis, smartphone sensing, wearable devices, passive sensing, natural language processing, machine learning, mental health screening, behavioral biomarkers, digital biomarkers.

Received: 10/08/2026
Revised: 14/09/2026
Accepted: 27/09/2026
Published: 13/10/2026

© 2026 The Author(s). Published by the Journal of Emerging Intelligence and Engineering Technologies. This is an open-access article distributed under the terms of the Creative Commons Attribution-NonCommercial 4.0 International License (CC BY-NC 4.0), which permits use, distribution, and reproduction in any medium for non-commercial purposes, provided the original work is properly cited. https://creativecommons.org/licenses/by-nc/4.0/ Deed - Attribution-NonCommercial 4.0 International - Creative Commons

Introduction

Depression and other mental disorders are among the most common causes of disability globally, but many of those with mental disorders are likely to be under-diagnosed or not treated, as traditional approaches to screening rely on brief encounters and self-report of depression retrospectively [1]. Personal digital devices have ushered in another paradigm, one in which the behaviour of humans is quantified incrementally, incidentally, and continuously via data gathered from their device interactions, called digital phenotyping [1, 2]. As smartphones, wearables and social media platforms are carried and used all day

long, the behavioural traces they leave behind—in the form of location trajectories, communication patterns, language preferences, sleep and activity rhythms etc.—may capture early or subclinical markers of psychiatric risk, that are recorded long before a formal diagnosis is sought, and therefore form a "digital footprint" [2, 3]. This paper explores how these digital footprints, alone and together, have been applied in studies published to date (2021). Literature shows that there are three general modalities: Textual and visual content on social media, passively recorded sensor and usage data from a smartphone, and physiological or activity data from consumer wearables. While each modality has shown some level of predictive signal, published results have widely different levels of performance depending on data type, feature richness, and validation technique, and single-modality systems often lack the level of reliability needed for clinical use [4], [5]. But multimodal methods which combine two or more of these data sources have started to show better discriminative results than unimodal methods, and there is a need to closely investigate the architectures of multimodal methods, and the circumstances in which they outperform single-source methods.

Background and The Diagnostic Gap

Digital Phenotyping and Passive Sensing

Digital phenotyping has been coined to refer to the continuous in-situ quantification of an individual's phenotype based on the data gathered from personal digital devices, especially smartphones, which are ubiquitous and have sensor arrays built into them [1]. Onnela and Rauch [2] found that there were times when data were active and warranted user input, such as in-app surveys, but also times when they were passive and gathered passively, via sensors such as global positioning system, accelerometer, microphone, and call and text logs. This separation is particularly important for mental health applications where repeated self-reports are reactive and burdensome and longitudinal observation at a temporal resolution not possible using traditional clinical interviews. Later, the digital phenotyping framework was expanded to the global arena where the assessment of this kind from the smartphone has been suggested as a scalable, low-cost alternative to the clinical diagnosis of mental disorders, especially in mental health scarce regions [3].

The Diagnostic Gap

Although continuous passive monitoring is attractive, the lack of a formal diagnosis for a large proportion of individuals with clinically meaningful depressive or anxiety symptoms that are detectable through such approaches is a reason for this line of research: even for those who receive a formal diagnosis, diagnosis often occurs years after symptoms are present. A variety of literatures on social media-based detection have been reviewed, with feature families identified, such as linguistic, sentiment, posting frequency and temporal regularity, ego-network and social-engagement structure, and recently, even the visual information provided with textual content [5]. This variety of feature types is assumed to be a reflection of the assumption that depression is present in various observable aspects of digital behaviour at a time, and that this is reflected in the multimodal framing used throughout this paper.

Modalities of Digital Footprint Analysis

Social Media Text and Language Markers

Initial research on the detection of depression on social media set the groundwork for the general feasibility of language-based screening. De Choudhury et al. created

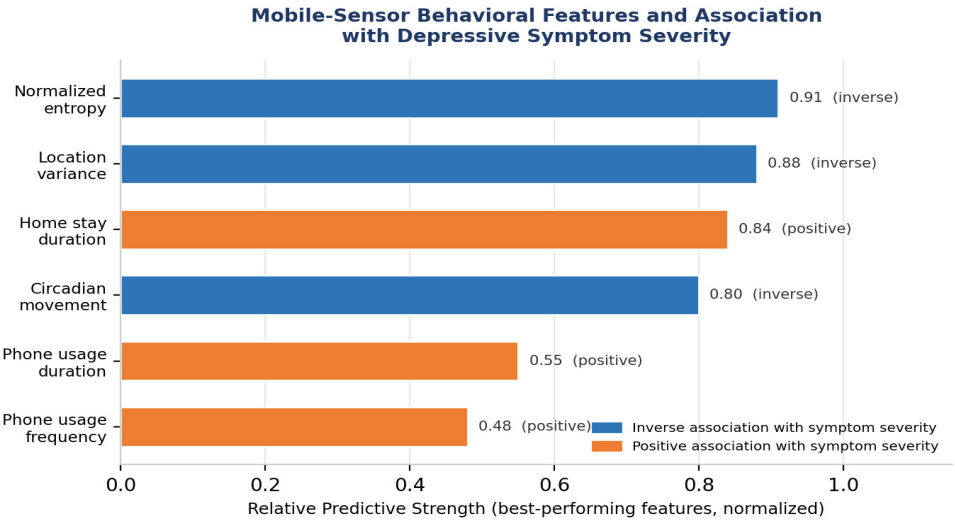


Figure 1: Mobile-sensor behavioral features and their reported direction and relative strength of association with depressive symptom severity, based on Saeb et al. [13].

a crowd-sourced set of Twitter users that reported having a clinical depression diagnosis, and, using a support vector machine classifier that learned on five behavioral dimensions—social engagement, emotional expression, linguistic style, ego-network structure, and mentions of antidepressant medication—they attained a classification accuracy of about 70 percent [7].

Based on electronic-health-record-linked data, rather than self-report, Eichstaedt et al. analysed the language of 683 ED patients, including 114 who had documented depression, and identified that language before the patients' diagnosis predicted depression (area under the curve 0.69); adding demographic variables did not increase the predictive value that could be gained from language alone (area under the curve 0.72 when the input window was limited to the six months before diagnosis); a detectable signal had an area under the curve of 0.62 and was present as early as three months before clinical documentation [8]. Reece et al. expanded language-based detection to Post-Traumatic Stress Disorder, in addition to depression, with 204 Twitter users and almost 280,000 tweets, finding that their computational models outperformed general-practitioner diagnostic accuracy benchmarks, which correctly identified depression in 42% of true cases, and PTSD in 2.5% of true cases [9].

In more recent feature-engineering studies, combining sentiment lexicons with content-based textual features has improved the accuracy of single-classifier classifiers beyond 96 percent on benchmark datasets [10] and ensemble-classifier classifiers beyond 98 percent [11] on a corpus of 54,412 posts from five mental health-related Reddit subreddits, where the difficulty lies in distinguishing between several co-occurring disorders instead of a single binary classification.

Smartphone-Based Behavioral Sensing

In addition to text-based approaches, passive smartphone sensing will provide behavioral signal. The StudentLife study tracked 48 students across Dartmouth College for a 10-week period, recording their location, activity, conversation, and screen and application use along with administering a Patient Health Questionnaire every other week, and found significant correlations between some measures of their behavior from the sensors with both the severity of their depression and their cumulative GPAs, and a recurring "term lifecycle" in which positive affect, conversational activity, and sleep regularity were highest and then

decreased as the term progressed, moving from workload to both workload and stress [12].

In a smaller sample of 28 analyzable adults who were monitored for two weeks, Saeb et al. identified location variance, normalized entropy of movement, home-stay duration, and circadian movement regularity as the top four correlates of depressive symptom severity (rated by the Patient Health Questionnaire) and determined that combining all the features did not increase, and in some cases, worsened, predictive accuracy by overfitting [13].

Wearable Physiological and Activity Sensing

Passive sensing from wearable, environmental, and smartphone signals have been reviewed, with physiological channels such as heart rate, activity through the accelerometer, and sleep-related rhythms cited as potential digital biomarkers for mental health monitoring, with evidence for their use to predict mental health outcomes being limited in comparison to text-based methods [14]. This caution is borne out by the work of Rykov et al. who compared the effectiveness of 290 working adults across 14 different days using the Fitbit Charge 2 and found that across the model variants a classification accuracy of 86 percent was achieved, but with discriminative metrics ranging from 0.51 to 0.66 (Fig. 2 and Table II) [15] with sensitivity between 3 per cent and 13 per cent.

Multimodal Fusion

In unimodal systems, results have been mixed, but when multiple digital footprints streams are combined, they can enhance system performance. Yazdavar et al. fused textual, visual and user-interaction features extracted from Twitter profiles in a statistical framework, increasing the average F1 scores by around five percentage points compared to the best unimodal baseline; they also created a social-media depression score, which was correlated with the population-level depression statistics released by public health officials [16].

Gui et al. [17] proposed a new cooperative multi-agent reinforcement-learning architecture (COMMA), that selects the most informative textual and visual posts from a user's history before classification, and the fused model achieved an F1 score of 0.900, compared to 0.865 for the strongest prior multimodal neural baseline, and 0.849 for the best feature-engineered unimodal model, representing a relative improvement of over 30 percent in classification error. Together, these findings indicate that the marginal information conveyed by each of these modalities is not completely shared, or redundant, so that the combination of behavioural, linguistic and visual data provides a partial representation of the presentation of depression that is more complete than any single modality.

Comparative Performance Analysis

Table III summarizes the modality, data source, scale, and headline reported outcome for the ten empirical studies reviewed in Section III, illustrating substantial heterogeneity in study design, sample size, and evaluation metric,

Table I: Mobile-Sensor Behavioral Features and Association with Depressive Symptom Severity

Feature	*Sensor*	*Direction*	*Relative Predictive Strength*
Normalized entropy	GPS	Inverse	Highest (lowest NRMSD)
Location variance	GPS	Inverse	High
Home-stay duration	GPS	Positive	High
Circadian movement regularity	GPS	Inverse	High
Phone usage duration	Usage logs	Positive	Moderate
Phone usage frequency	Usage logs	Positive	Moderate-low

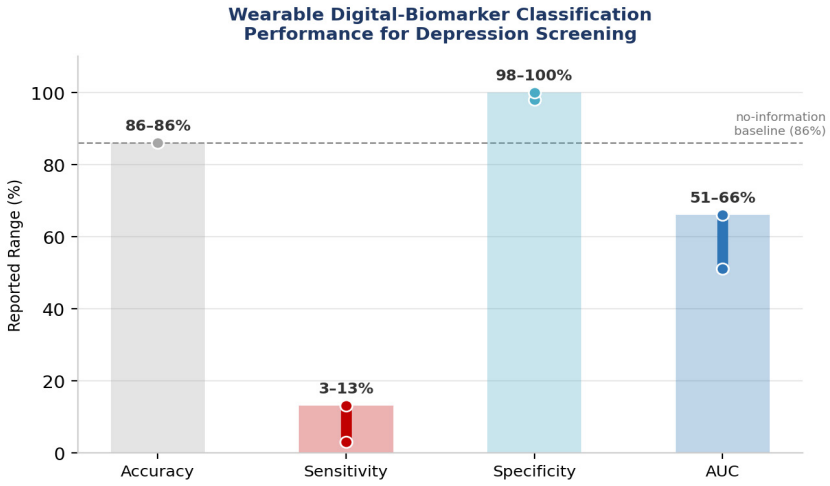


Figure 2: Reported classification performance ranges for wearable digital-biomarker depression screening, based on Rykov et al. [15].

and Fig. 3 visualizes these outcomes on a common, normalized scale for ease of comparison. Because studies report accuracy, F1-score, area under the curve, or correlation-based outcomes interchangeably, direct numerical comparison must be interpreted cautiously. Performance is frequently reported as the F1-score, the harmonic mean of precision and recall,

$$F1 = 2 \times \frac{(Precision \times Recall)}{(Precision + Recall)} \quad (1)$$

which weighs the disadvantages of failing to identify at risk individuals (false negative) against the disadvantages of unnecessary intervention (false positive). The range of performance values reported by the modality category is summarized in Table IV.

Text-based social media classifiers range widely in performance, from around 0.62 to 0.72 area under the curve on clinically validated, electronic-health-record-linked samples [8] to 96 to 98 percent accuracy on feature-rich, lexicon-augmented classifiers tested on benchmark (not clinically confirmed) social media samples [10], indicating that reported accuracy is dependent of the rigor with which ground-truth labels are set. The reviewed literature did not deliver a headline accuracy value for the GPS on the smartphones but instead suggested to take the behavioural entropy, location variance, home-stay duration and circadian regularity as individual correlates of symptom severity used to assess the smartphone. The weakest discriminative performance of the modalities reviewed was wearable-only physiological sensing (AUC slightly above chance and sensitivity below 15 percent [15]). The multimodal fusion architectures consistently surpassed their

Table II: Wearable Digital-Biomarker Categories and Predictive Performance

Category	*Sensor channel*	*Performance metric*	*Reported value*
Physical activity	Steps	Included in composite model	—
Cardiac	Heart rate	Included in composite model	—
Sleep	Sleep duration / timing	Included in composite model	—
Circadian rhythm	Rest-activity variability	Included in composite model	—
Composite classification	All channels combined	Accuracy / Sens. / Spec. / AUC	86% / 3–13% / 98–100% / 0.51–0.66

Table III: Comparative Overview of Reviewed Digital Footprint Modalities and Studies

Ref.	*Modality*	*Data Source*	*Scale*	*Key Reported Outcome*
[7]	Social text	Twitter (crowdsourced)	Self-reported diagnosed users	SVM accuracy ≈ 70%
[12]	Smartphone sensing	GPS, accelerometer, audio, usage	48 students, 10 weeks	Significant correlations with PHQ-9 and GPA
[13]	Smartphone sensing	GPS + usage logs	28 of 40 adults, 2 weeks	Entropy & location variance: lowest NRMSD
[8]	Social text	Facebook (EHR-linked)	683 patients (114 depressed)	AUC = 0.69 (0.72, 6-mo. window)
[9]	Social text	Twitter	204 users / 279,951 tweets	Outperformed GP diagnostic baseline
[10]	Social text	Lexicon + content features	90-feature benchmark set	Accuracy >96% (>98% ensemble)
[11]	Social text	Reddit (SWMH)	54,412 posts, 5 classes	F1 ≈ 0.65 (baseline)
[15]	Wearable	Fitbit Charge 2	267 of 290 adults, 14 days	AUC 0.51–0.66 (weak)
[16]	Multimodal	Twitter (text+image+interaction)	Benchmark dataset	+5 pts. F1 over unimodal SOTA
[17]	Multimodal	Twitter (text+image, COMMA)	Benchmark dataset	F1 = 0.900 (~30% error reduction)

unimodal counterparts (modality and publication year shown in Fig. 4) by as much as a five-percentage-point F1 gain [16] or a 30-percent relative error reduction [17]. Table V also places these computational results in perspective with a human diagnostic baseline: in benchmark studies, general practitioners correctly identified 42 percent of true depression cases and 2.5 percent of true PTSD cases, with high specificity in ruling out non-cases [9] — illustrating the limitations of the clinician baseline that originally inspired computational screening approaches. Even for relatively low levels of accuracy, these digital models can reach or surpass these baselines, suggesting that the threshold for clinical utility may not be as high as simply the level of laboratory benchmark accuracy.

Methodological and Ethical Challenges

The literature reviewed does have some common weaknesses, as described in Table VI. Ground-truth validity is often ambiguous, because multiple influential studies are based on diagnosis claims from crowdsourced data or self-reported diagnoses, and models may learn to predict self-diagnosis of depression instead of the clinical depression itself [6], [7]. Selection and sampling bias are ubiquitous because the populations studied typically include only those who are using the platform (as in Twitter, Facebook, and Fitbit-using populations) and who are more likely than the general population to seek out online platforms for expressing distress (due to age, socioeconomic status, or other factors) [5, 6]. Few studies reviewed investigated generalizability across platforms, languages and cultural contexts, with most of these being English-language samples from North America [5].

Passive sensing of location, communication and physiological data poses substantive privacy and informed consent issues, especially in the context of digital phenotyping which may provide sensitive information about mood, social relationships and routines that the users might not have intended to provide as evidence for clinical inference [1,2]. There is still also a considerable gap between technical performance and regulatory or clinical validation, as suggested by reviews of the wider digital psychiatry sector, where a multitude of apps, chatbots and sensing devices have flooded the market, leaving the field of mental health applications far from having been rigorously evaluated in a prospective way for clinical use [4]. Last but not least, as Section IV demonstrates, single-modality

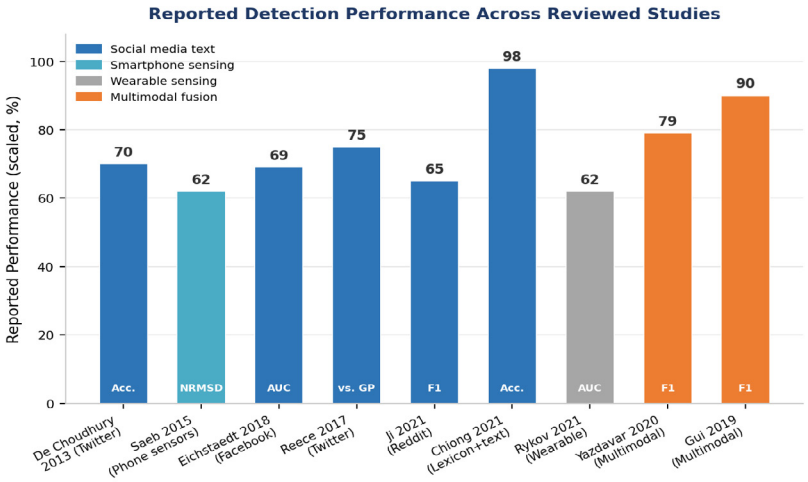


Figure 3: Reported detection performance (accuracy, F1, or AUC, scaled to a 0–100 axis) across nine representative studies, grouped by modality.

performance in particular for systems that are only wearable can be so poor that deployment without complementary data sources can lead to false reassurance and missed cases [15]. These obstacles all point to the need to consider technical performance indicators in conjunction with an explicit description of the quality of the labels, population representativeness, and consent architecture before any detection system is ready to be used in the field for screening.

Toward Integrated Multimodal Frameworks

To support the continued success of multimodal fusion demonstrated in Section III-D, Fig. 5 shows a conceptual four-layered architecture that combines the fusion strategies used throughout the literature in this study. Raw digital footprints are gathered initially from three different complementary sources: (1) text and imagery from social media, (2) sensor and usage logs from smartphones, and (3) physiological streams from wearables. Specific feature extraction is applied to each source to generate linguistic representation, sentiment identification, behavioral representation, mobility representation and physiological identification and activity representations respectively. In the latter strategy (late fusion) used by some of those reviewed works [16, 17] a combined risk representation for a specific user may be obtained by a weighted combination of modality-specific model outputs.

$$R=\Sigma\left(i=1 to n\right) wi \cdot fi\left(xi\right)\left(2\right)$$

Table IV: Methodological and Ethical Challenges Identified in the Literature

Category	*Description*	*Representative source(s)*
Ground-truth validity	Reliance on self-reported / crowdsourced diagnosis rather than clinician-confirmed records	[6], [7]
Sampling / selection bias	Active platform users may not represent the general population	[5], [6]
Generalizability	Limited testing across languages, platforms, and cultures	[5]
Privacy & consent	Passive sensing captures sensitive behavioral and location data	[1], [2]
Clinical validation gap	Rapid tool proliferation outpaces rigorous prospective evaluation	[4]
Single-modality weakness	Wearable-only systems show near-chance discriminative power	[15]

Table V: Reported Performance Range by Modality Category

Modality	*Metric Type*	*Reported Range*	*Representative Refs.*
Social media text	Accuracy / AUC / F1	62–98% acc.; 0.62–0.72 AUC; 0.65–0.90 F1	[7], [8], [9], [10], [11]
Smartphone sensing	Correlation-based	No single accuracy figure; best-feature NRMSD minimized	[12], [13]
Wearable sensing	AUC / Sens. / Spec.	0.51–0.66 AUC; 3–13% sens.; 98–100% spec.	[15]
Multimodal fusion	F1	0.86–0.90 (+5 to +30% relative gain)	[16], [17]

There, fi(xi) is the risk score generated by the classifier for modality i, xi is the modality i-specific feature vector, wi is a learned or empirically determined weight that takes into account the relative reliability of that modality, and n is the number of available modalities.

The generalisation of the cooperative multi-agent selection mechanism used by Gui et al. [17] is given by Equation (2) where textual and visual agents jointly and adaptively make a classification decision – not by simply averaging their respective decisions – and is consistent with the additive performance gains reported by Yazdavar et al. when interaction features were added to text and image features [16]. The final architectural layer consists of a classifier or risk estimation model that translates the fused representation into a mental health risk score or alert that can be used to guide clinical decisions. The utility of such an architecture is the ability to down-weight modalities that may be unreliable for that user for whom they wear, such as a user exhibiting low adherence to wearing the modality, and to still generate a reliable composite estimate from the other modalities.

Discussion: Trends and Outlook

This temporal progression highlighted in Fig. 6 is representative of a general trend in the field over the eight-year time frame examined here. Initial research (2013–2015) demonstrated feasibility, and generally involved single modalities and relatively small or self-reported samples [7, 12, 13]. Meanwhile, conceptualization of the digital phenotyping framework was consolidated in 2016–2018, along with the first clinical validation of the digital phenotyping paradigm with electronic health records [2, 3, 8] and a growing number of comparative and critical reviews of the emerging evidence base [5]. The center of gravity of the field shifted towards multimodal fusion and towards more in-depth methodological self-reflection, with

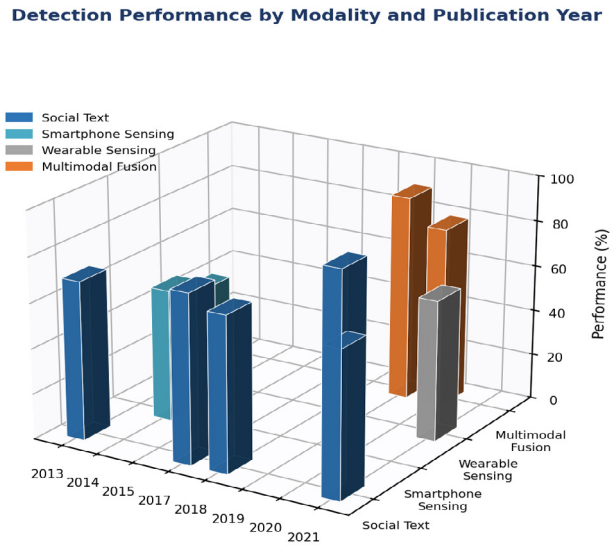


Figure 4: Three-dimensional comparison of reported detection performance by modality and publication year.

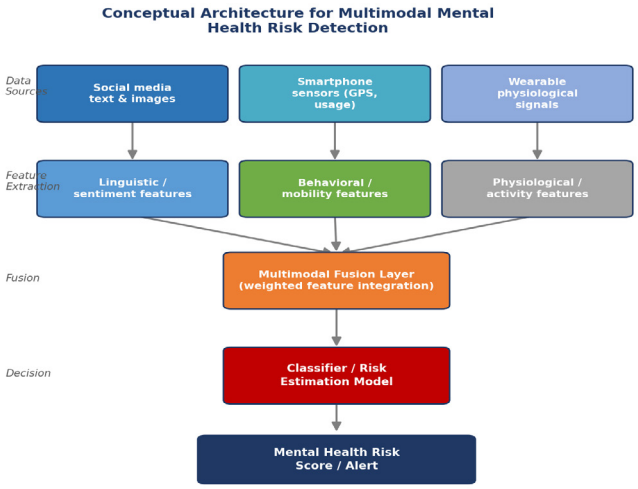


Figure 5: Conceptual four-layer architecture for multimodal mental health risk detection, integrating social media, smartphone, and wearable data sources.

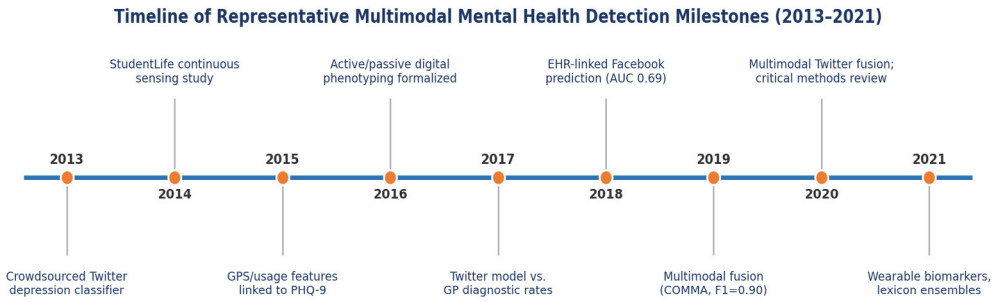


Figure 6: Timeline of representative multimodal mental health detection milestones, 2013–2021.

multimodal models that operate cooperatively [17], with extended multimodal benchmarks [16] and with a landmark critical review of predictive methodology [6] – as described in Table VII – as well as the continuation of the exploration of complementary data streams such as wearable [15] and passive environmental sensing [14].

Two trends will continue to influence future development that are not included in this review. First, the substantial gains in performance by combining the multimodal systems over unimodal systems indicate that future detection systems will adopt more of the layered fusion of classifiers in Fig. 5 than classifiers from a single source. Second, the methodological critiques summarised in Section V suggest that future advances will require more than just an increase in the accuracy of classification models; it will also require the expansion of ground-truth validation, consent frameworks, and population diversity, describing the form, size, and extent to which that model performs well.

Conclusion

This review of 17 studies published between 2013 and 2021 suggests that the digitally produced footprints of social media use, smartphone sensing, and wearable devices include unique, though not redundant, signals of mental health risk and that none of these modalities reviewed here has the combination of accuracy, generalizability, and ethical robustness needed for stand-alone clinical

deployment. The highest single performance was found among text-based social media models while behavioural entropy and home-stay duration derived from smartphones turned out to be the most informative passive-sensing correlates, and physiological data obtained from wearables alone had comparatively limited discriminative power. Two or more of these streams were combined in multimodal fusion architectures to outperform consistently unimodal baselines by up to 0.900 F1. Future work should focus on integrated multimodal architectures of this type, confirmed by clinician data and not self-reported ground truth, and be designed in the context of privacy and consent suitable for continuous passive data collection.

References

T. R. Insel, "Digital phenotyping: Technology for a new science of behavior," JAMA, vol. 318, no. 13, pp. 1215–1216, 2017.

J.-P. Onnela and S. L. Rauch, "Harnessing smartphone-based digital phenotyping to enhance behavioral and mental health," Neuropsychopharmacology, vol. 41, no. 7, pp. 1691–1696, 2016.

T. R. Insel, "Digital phenotyping: A global tool for psychiatry," World Psychiatry, vol. 17, no. 3, pp. 276–277, 2018.

J. Torous et al., "The growing field of digital psychiatry: Current evidence and the future of apps, social media, chatbots, and virtual reality," World Psychiatry, vol. 20, no. 3, pp. 318–335, 2021.

S. C. Guntuku, D. B. Yaden, M. L. Kern, L. H. Ungar, and J. C. Eichstaedt, "Detecting depression and mental illness on social media: An integrative review," Curr. Opin. Behav. Sci., vol. 18, pp. 43–49, 2017.

S. Chancellor and M. De Choudhury, "Methods in predictive techniques for mental health status on social media: A critical review," npj Digit. Med., vol. 3, no. 1, pp. 1–11, 2020.

M. De Choudhury, M. Gamon, S. Counts, and E. Horvitz, "Predicting depression via social media," in Proc. Int. AAAI Conf. Web Social Media, vol. 7, no. 1, 2013, pp. 128–137.

J. C. Eichstaedt et al., "Facebook language predicts depression in medical records," Proc. Natl. Acad. Sci. U.S.A., vol. 115, no. 44, pp. 11203–11208, 2018.

A. G. Reece, A. J. Reagan, K. L. M. Lix, P. S. Dodds, C. M. Danforth, and E. J. Langer, "Forecasting the onset and course of mental illness with Twitter data," Sci. Rep., vol. 7, art. 13006, 2017.

R. Chiong, G. S. Budhi, and S. Dhakal, "Combining sentiment lexicons and content-based features for depression detection," IEEE Intell. Syst., vol. 36, no. 6, pp. 99–105, 2021.

S. Ji, X. Li, Z. Huang, and E. Cambria, "Suicidal ideation and mental disorder detection with attentive relation networks," Neural Comput. Appl., vol. 34, pp. 10309–10319, 2021.

Syed, M. H., Ragireddy, U. K., Addepalli, P. V., Bandaru, S. R., Shyam, D., & Manoharan, P. (2025). Enhancing Retail Sector Customer Engagement Through the Application of Modern and Intelligent AI-Driven Analytical Techniques. *International Journal of Humanities and Information Technology*, *7*(01), 19-26.

Shyam, D., Manoharan, P., Syed, M. H., Ragireddy, U. K., Addepalli, P. V., & Bandaru, S. R. (2025). An AI-Driven Classification System for Early Detection of Customer Churn in E-Commerce Platforms. *Journal of Computer Science and Technology Studies*, *7*(11), 452-462.

Bandaru, S. R., Shyam, D., Manoharan, P., Syed, M. H., Ragireddy, U. K., & Addepalli, P. V. (2025). Customer Sentiment Analysis for Online Shopping Platforms Using NLP and Deep Learning Methods. *Journal of Cyber-Physical Security and Robotics*, *1*(02), 1-11.

Mylavarapu, K. B., Pilli, J., Jannu, P. K., Mohammad, J. A., Panchali, S. H., & Kavirayani, U. M. (2025). An Enhanced Intrusion Detection Architecture Using ML-Based Anomaly Detection on Large-Scale Data. *International Journal of Humanities and Information Technology*, *7*(01), 20-26.

Panchali, S. H., Mylavarapu, K. B., Pilli, J., Jannu, P. K., & Mohammad, J. A. (2025). Review of Secure Communication Systems for Cyber-Physical Systems and Industrial IoT. *Journal Of Engineering And Computer Sciences*, *4*(6), 349-362.

Pilli, J., Jannu, P. K., Mohammad, J. A., Panchali, S. H., Kavirayani, U. M., & Mylavarapu, K. B. (2025). Data-Driven Artificial intelligence (AI)-based System for Identifying Fraudulent Telephony Activities. *Journal of Cyber-Physical Security and Robotics*, *1*(02), 23-31.

Jannu, P. K., Mohammad, J. A., Panchali, S. H., Kavirayani, U. M., Mylavarapu, K. B., & Pilli, J. (2025). A Review of Sensor Data Acquisition Methods for Accurate and Timely Detection of Faults in Mechanical Systems. *Journal of Computer Science and Technology Studies*, *7*(9), 796-803.

Keshireddy, S. R., Nagumotu, V. T., Kavuluri, H. V. R., Pathani, A. K., Dasari, A., & Chilakapati, V. K. (2025). Emerging Communication Technologies for IoT: 5G and Edge Networking Perspectives. *International Journal of Humanities and Information Technology*, *7*(01), 19-28.

Dasari, A., Chilakapati, V. K., Keshireddy, S. R., Nagumotu, V. T., Kavuluri, H. V. R., & Pathani, A. K. (2025). Transforming Undergraduate Education Through Modern Pedagogical Innovations. *Journal of Computer Science and Technology Studies*, *7*(12), 547-556.

Pathani, A. K., Dasari, A., Chilakapati, V. K., Keshireddy, S. R., Nagumotu, V. T., & Kavuluri, H. V. R. (2025). Big Data and Cloud Computing Integration: A Review of Scalable Information Retrieval Techniques. *Journal of Cyber-Physical Security and Robotics*, *1*(02), 12-22.

Suresh, A. (2025). AI-Driven Business Intelligence Automation: Integrating Data Engineering, Auto-BI, and Large Language Models. International Journal of AI, BigData, Computational and Management Studies, 6(3), 97-108. https://doi.org/10.63282/3050-9416.IJAIBDCMS-V6I3P112

Suresh, A. (2025). Conversational Analytics Using LLMs: Transforming Enterprise Data Consumption through Natural Language Interfaces. American International Journal of Computer Science and Technology, 7(5), 92-102. https://doi.org/10.63282/3117-5481/AIJCST-V7I5P108

Syed, M. H., Ragireddy, U. K., Addepalli, P. V., Bandaru, S. R., Shyam, D., & Manoharan, P. (2026). AI-Powered Data Analytics for Strengthening Customer Loyalty and Personalization in Retail. *Journal of Data Analysis and Critical Management*, *2*(01), 28-36.

Manoharan, P., Syed, M. H., Ragireddy, U. K., Addepalli, P. V., Bandaru, S. R., & Shyam, D. (2026). A Review of Digital Tools and Technologies for Enhancing Customer Engagement in Financial Services. *Journal of Contemporary Science and Technology Management*, *2*(01), 13-25.

Ragireddy, U. K., Addepalli, P. V., Bandaru, S. R., Shyam, D., Manoharan, P., & Syed, M. H. (2026). Predictive Customer Segmentation in Retail Using Machine Learning and Behavioral Analytics. *European Journal of Applied Science, Engineering and Technology*, *4*(2), 268-278.

Suresh, "Architectural Design of LLM- Powered Conversational BI Systems for High-Volume Enterprise Data Warehouses," 2026 International Conference on Machine Learning and Autonomous Systems (ICMLAS), Bangkok, Thailand, 2026, pp. 1333-1341, doi: 10.1109/ICMLAS67792.2026.11483861.

Taluri, R. (2024). A Hybrid Data Engineering and Generative AI Architecture for Intelligent Data Governance, Metadata Management, and Automated Data Quality Assessment on AWS. International Journal of Artificial Intelligence,

Data Science, and Machine Learning, 5(2), 230-240. https://doi.org/10.63282/3050-9262.IJAIDSML-V5I2P126

Suresh, A. (2024). Auto-BI Frameworks Powered by Generative Artificial Intelligence for Scalable Self-Service Data Analytics in Large Organizations. International Journal of Artificial Intelligence, Data Science, and Machine Learning, 5(3), 191-201. https://doi.org/10.63282/3050-9262.IJAIDSML-V5I3P119

Taluri, R. (2024). A Cloud-Native Reference Architecture for Data Engineering, Generative AI, and Decision Intelligence Using AWS and Amazon Bedrock. International Journal of AI, BigData, Computational and Management Studies, 5(1), 218-227. https://doi.org/10.63282/3050-9416.IJAIBDCMS-V5I1P122

Suresh, A. (2025). AI-Driven Business Intelligence Automation: Integrating Data Engineering, Auto-BI, and Large Language Models. International Journal of AI, BigData, Computational and Management Studies, 6(3), 97-108. https://doi.org/10.63282/3050-9416.IJAIBDCMS-V6I3P112

Taluri, R. (2021). Cloud-Native Architectures for Enterprise Financial Data Management, Analytics, and Regulatory Reporting Compliance. International Journal of Emerging Trends in Computer Science and Information Technology, 2(2), 101-111. https://doi.org/10.63282/3050-9246.IJETCSIT-V2I2P112

Suresh, A. (2025). Conversational Analytics Using LLMs: Transforming Enterprise Data Consumption through Natural Language Interfaces. American International Journal of Computer Science and Technology, 7(5), 92-102. https://doi.org/10.63282/3117-5481/AIJCST-V7I5P108

Taluri, R. (2022). Cloud Data Engineering Strategies for Large-Scale Financial Data Integration and Intelligent Corporate Performance Reporting. International Journal of Emerging Research in Engineering and Technology, 3(4), 176-188. https://doi.org/10.63282/3050-922X.IJERET-V3I4P119

Jannu, P. K., Mohammad, J. A., Panchali, S. H., Kavirayani, U. M., Mylavarapu, K. B., & Pilli, J. (2026). Recent Advances in Anomaly Identification for IoT Devices Using Machine and Deep Learning Models. *Journal of Data Analysis and Critical Management*, *2*(01), 37-47.

Mylavarapu, K. B., Pilli, J., Jannu, P. K., Mohammad, J. A., Panchali, S. H., & Kavirayani, U. M. (2026). Early Diagnosis of Heart Conditions Using AI-Driven on Electronic Health Data. *Journal of Contemporary Science and Technology Management*, *2*(01), 26-37.

Pilli, J., Jannu, P. K., Mohammad, J. A., Panchali, S. H., Kavirayani, U. M., & Mylavarapu, K. B. (2026). A Survey on Power Optimization Techniques in Contemporary Mobile SoC Architectures. *European Journal of Applied Science, Engineering and Technology*, *4*(2), 321-333.

Chilakapati, V. K., Keshireddy, S. R., Nagumotu, V. T., Kavuluri, H. V. R., Pathani, A. K., & Dasari, A. (2026). A Survey on Generative AI Chatbots in Customer Service: Challenges, Benefits, and Use Cases. *Journal of Data Analysis and Critical Management*, *2*(01), 17-27.

Pathani, A. K., Dasari, A., Chilakapati, V. K., Keshireddy, S. R., Nagumotu, V. T., & Kavuluri, H. V. R. (2026). Transforming Customer Support with AI: Emerging Tools, Trends and Their Service Effectiveness. *Journal of Contemporary Science and Technology Management*, *2*(01), 1-12.

Dasari, A., Chilakapati, V. K., Keshireddy, S. R., Nagumotu, V. T., Kavuluri, H. V. R., & Pathani, A. K. (2026). Applications of Machine Learning in Predicting Customer Churn in Telecommunications: A Review Analysis. *European Journal of Applied Science, Engineering and Technology*, *4*(2), 257-267.

Kumar, A., Wadhwa, M., Kalla, D., Konduru, S. C., Nandawat, C., & Sharma, M. (2025, October). Benchmarking the Trade-Offs in Object Detection: Accuracy, Speed, and Energy Efficiency. In *International Conference on Artificial Intelligence and Networking* (pp. 410-422). Cham: Springer Nature Switzerland.

NR, A. R., Rajasri, T., Praveen, R., Kalla, D., Bendale, S. P., & Venu, N. (2025, April). CAC Training-A Unified Cybersecurity Training Program for Military Staff. In *2025 3rd International Conference on Communication, Security, and Artificial Intelligence (ICCSAI)* (Vol. 3, pp. 569-573). IEEE.

Kalla, D., Kuraku, D. S., & Samaah, F. (2021). Enhancing cyber security by predicting malwares using supervised machine learning models. *International Journal of Computing and Artificial Intelligence*, *2*(2), 55-62.

Zeeshan, M., Bhadauria, K., Pahal, L., Nagrath, P., & Kalla, D. (2025, June). Ensemble-Based Deep Learning for Automated Diabetic-Retinopathy Detection Using CNNs and Transfer Learning. In *International Conference on Data Analytics & Management* (pp. 216-228). Cham: Springer Nature Switzerland.

Aggarwal, A., Agarwal, L., Rella, B. P. R., Nagpal, N., Kalla, D., & Sharma, M. (2025, June). A Performance Comparison of Machine Learning Models for Rain Prediction. In *International Conference on Data Analytics & Management* (pp. 319-328). Cham: Springer Nature Switzerland.

Nagrath, P., Saini, I., Zeeshan, M., Komal, Komal, & Kalla, D. (2025, June). Predicting Mental Health Disorders with Variational Autoencoders. In *International Conference on Data Analytics & Management* (pp. 38-51). Cham: Springer Nature Switzerland.

NR, A. R., Rajasri, T., Praveen, R., Kalla, D., Bendale, S. P., & Venu, N. (2025, April). CAC Training-A Unified Cybersecurity Training Program for Military Staff. In *2025 3rd International Conference on Communication, Security, and Artificial Intelligence (ICCSAI)* (Vol. 3, pp. 569-573). IEEE.

Kalla, D., Smith, N., Samaah, F., & Polimetla, K. (2024). Hybrid scalable researcher recommendation system using azure data lake analytics. *Journal of Data Analysis and Information Processing*, *12*, 76-88.

Kumar[1], A., Wadhwa[1], M., Kalla, D., Konduru, S. C., Nandawat, C., & Sharma, M. (2026, May). Benchmarking the Trade-Offs in Object Detection: Accuracy, Speed, and Energy. In *Proceedings of International Conference on Artificial Intelligence and Networks: ICAIN 2025, Volume 4* (Vol. 4, p. 410). Springer Nature.

A. Suresh, "Large Language Models for Explainable and Self-Service Business Intelligence in Large Enterprises," 2026 6th International Conference on Expert Clouds and Applications (ICOECA), Bengaluru, India, 2026, pp. 1571-1578, doi: 10.1109/ICOECA68095.2026.11485058.

A.Suresh, "A Comprehensive Study on Auto - BI Systems using Generative AI for Scalable and Explainable Enterprise Analytics," 2026 6th International Conference on Expert Clouds and Applications (ICOECA), Bengaluru, India, 2026, pp. 1579-1585, doi: 10.1109/ICOECA68095.2026.11485569.

Krishnappa, K. H., R, M., & Chavan, L. (2026). ML-Driven Automated Functional Verification Framework for Digital Designs. *Electronics*, *15*(12), 2687.

Kalla, D., & Samaah, F. (2023). Exploring Artificial Intelligence And Data-Driven Techniques For Anomaly Detection In Cloud Security. *Available at SSRN 5045491*.

Suresh, "Architectural Design of LLM- Powered Conversational BI Systems for High-Volume Enterprise Data Warehouses," 2026 International Conference on Machine Learning and Autonomous Systems (ICMLAS), Bangkok, Thailand, 2026, pp. 1333-1341, doi: 10.1109/ICMLAS67792.2026.11483861.

Vadisetty, R., Polamarasetti, A., Varadarajan, V., Kalla, D., & Ramanathan, G. K. (2025, May). Cyber Warfare and AI Agents: Strengthening National Security Against Advanced Persistent Threats (APTs). In International Conference on Intelligence-Based Transformations of Technology and Business Trends (pp. 578-587). Cham: Springer Nature Switzerland.

R. Wang et al., "StudentLife: Assessing mental health, academic performance and behavioral trends of college students using smartphones," in Proc. 2014 ACM

Int. Joint Conf. Pervasive Ubiquitous Comput., 2014, pp. 3–14.
S. Saeb et al., "Mobile phone sensor correlates of depressive symptom severity in daily-life behavior: An exploratory study," J. Med. Internet Res., vol. 17, no. 7, art. e175, 2015.
P. A. Kyriacou, "Wearable, environmental, and smartphone-based passive sensing for mental health monitoring," Front. Digit. Health, vol. 3, art. 662811, 2021.
Y. Rykov, T.-Q. Thach, I. Bojic, G. Christopoulos, and J. Car, "Digital biomarkers for depression screening with wearable devices: Cross-sectional study with machine learning modeling," JMIR mHealth uHealth, vol. 9, no. 10, art. e24872, 2021.
A. H. Yazdavar et al., "Multimodal mental health analysis in social media," PLOS ONE, vol. 15, no. 4, art. e0226248, 2020.
T. Gui, L. Zhu, Q. Zhang, M. Peng, X. Zhou, K. Ding, and Z. Chen, "Cooperative multimodal approach to depression detection in Twitter," in Proc. AAAI Conf. Artif. Intell., vol. 33, no. 1, 2019, pp. 110–117.