VER
Vascular & Endovascular Review

A Unified Methodology: Integrating Machine Learning And Mathematical Modeling For Robust Psychological Assessment

[1]*Dr.T.Rameshkumar, [1]T.Karthikeyan, [2]Dr.S.Sreelakshmi, [3]Dr.T.Vengatesh, [4]Dr.G.Ambika, [5]Krovvidi S B Ambika, [6]Dr.Elavarasan.K, [7]Dr.Shanky Goyal

[1]*(Coressponding Author) Associate Professor,Department of Mathematics, V.S.B.Engineering College (Autonomous), Karur-639 111, TamilNadu, India. Email ID: drtrameshkumar@gmail.com
[1]Assistant Professor, Department of Computer Science and Business Systems, Panimalar Engineering College, Chennai 600123. Tamil Nadu, India.
Email ID: karthi4cse@gmail.com
[2]Professor, Department of Mathematics,SIR M. Visvesvaraya Institute of Technology,
Krishnadevaryanagar, Bengaluru- 562157, Karnataka, India. Email: sreelakshmi_maths@sirmvit.edu
[3]Assistant Professor, Department of Computer Science, Government Arts and Science College, Veerapandi, Theni, Tamilnadu, India. Email ID: venkibiotinix@gmail.com
[4]Assistant Professor & Head, Department of Mathematics, S.A.College of Arts & Science, Avadi, Chennai -77, TamilNadu, India. Email ID: v2v2a2@gmail.com
[5]Assistant professor, Department of CSE, Aditya University, Surampalem-533437, Andhra Pradesh, India.
Email ID: ksbambika@gmail.com
[6]Assistant Professor & Head,Department of Mathematics, Vel Tech Rangarajan Dr.Sagunthala R&D Institute of Science and Technology, Chennai, Tamilnadu,India.
EMail ID: drelevarasank@veltech.edu.in
[7]School of Engineering and Technology, CGC University, Mohali – 140307, Mohali, Punjab, India. Email ID: Shanky1486@gmail.com

ABSTRACT

Psychological assessment faces significant challenges in balancing theoretical rigor with predictive accuracy. Traditional approaches relying on static questionnaires often lack dynamic sensitivity, while pure machine learning models operate as "black boxes" disconnected from psychological theory. This paper proposes a unified methodology that synergistically integrates mathematical modeling with machine learning to overcome these limitations. Our framework employs mathematical formalisms (dynamical systems theory, item response theory) to provide theoretical structure and interpretability, while machine learning algorithms (ensemble methods, deep learning) capture complex, non-linear patterns from multi-modal data. We validate this approach using a longitudinal dataset of 850 participants with depression and anxiety symptoms, demonstrating that our hybrid model achieves superior predictive performance (F1-score: 0.84) compared to standalone mathematical (F1-score: 0.67) or machine learning (F1-score: 0.78) approaches. Furthermore, our methodology generates clinically interpretable parameters that align with established psychological constructs. This integration represents a paradigm shift toward more dynamic, personalized, and theoretically-grounded psychological assessment.

KEYWORDS: Computational Psychometrics, Hybrid Modeling, Dynamical Systems, Explainable Ai, Psychological Assessment, Digital Phenotyping.

How to Cite: Dr.T.Rameshkumar, T.Karthikeyan, Dr.S.Sreelakshmi, Dr.T.Vengatesh, Dr.G.Ambika, Krovvidi S B Ambika, Dr.Elavarasan.K, Dr.Shanky Goyal, (2025) A Unified Methodology: Integrating Machine Learning And Mathematical Modeling For Robust Psychological Assessment, Vascular and Endovascular Review, Vol.8, No.16s, 243-250.

INTRODUCTION

Psychological assessment stands at a critical juncture. For decades, its foundation has been built on psychometric theories and mathematical models, such as Item Response Theory, which provide robust, interpretable frameworks for measuring latent constructs. These theory-driven approaches ensure that assessments are grounded in psychological science, yet they often struggle to capture the dynamic, multi-faceted, and non-linear nature of human cognition and emotion as revealed by modern, high-dimensional data sources.

The recent advent of machine learning (ML) promised a revolution, offering powerful tools to detect complex patterns and achieve high predictive accuracy from diverse data, including digital phenotyping. However, this power often comes at the cost of interpretability. Pure ML models frequently operate as "black boxes," generating predictions that are decoupled from established psychological theory and thus offer limited clinical insight or actionable understanding.

This presents a false dichotomy: a choice between theoretically meaningful but potentially simplistic models, and powerful but opaque predictors. To transcend this limitation, we propose a paradigm shift. This paper introduces a unified methodology that moves beyond a simple comparison of approaches to a true *integration* of mathematical modeling and machine learning. Our framework is designed to leverage the complementary strengths of both the theoretical structure and interpretability of formal mathematical models with the pattern-recognition power of ML algorithms. We demonstrate that this synergy not only achieves superior predictive performance in assessing depression and anxiety but, more importantly, yields a new class of robust, dynamic, and clinically interpretable tools for psychological science.

LITERATURE REVIEW

2.1 The Foundation: Mathematical Modeling in Psychology

The application of mathematical models to psychological phenomena has a long and storied history, providing the bedrock for quantitative assessment. Early psychometric frameworks, such as **Classical Test Theory** and later **Item Response Theory (IRT)**, formalized the measurement of latent constructs like intelligence, personality, and psychopathology (Embretson & Reise, 2000). These models introduced crucial concepts of reliability, item difficulty, and discrimination, offering a rigorous and interpretable structure for assessment. Beyond psychometrics, mathematical formalisms have been used to model cognitive and affective processes directly. For instance, **dynamical systems theory** has been applied to model the temporal evolution of emotions, conceptualizing mood as a system with attractor states and homeostatic mechanisms (Houben et al., 2015). Similarly, computational models of reinforcement learning have shed light on the decision-making anomalies underlying disorders like depression and anxiety (Huys et al., 2016). The primary strength of these approaches is their grounding in psychological theory; each parameter has a theoretical meaning, facilitating interpretability and hypothesis testing.

2.2 The Rise of Data-Driven Machine Learning

The advent of the "big data" era in psychology, fueled by digital phenotyping (e.g., smartphone sensors, wearable devices, electronic health records), has exposed the limitations of purely theory-driven models. These high-dimensional, multi-modal datasets often contain complex, non-linear relationships that are difficult to capture with pre-specified mathematical equations. In response, **machine learning (ML)** has emerged as a powerful alternative. Supervised learning techniques, from **ensemble methods** like Random Forests to **deep learning** architectures, have demonstrated remarkable success in predicting psychological outcomes, such as forecasting depressive episodes from mobile phone data (Saeb et al., 2015) or identifying suicide risk from clinical notes (Walsh et al., 2017). The power of ML lies in its ability to learn complex functions directly from data, often achieving superior predictive accuracy compared to traditional statistical models.

2.3 The Interpretability Crisis and the Theory-Prediction Gap

Despite their predictive prowess, pure ML approaches have drawn significant criticism for their "black box" nature. The models' decision-making processes are often opaque, making it difficult to extract clinically meaningful insights or relate findings back to established psychological constructs (Rudin, 2019). This creates a critical **theory-prediction gap**: a model may accurately predict a symptom, but without interpretability, it cannot advance our theoretical understanding of the underlying mechanisms. This limitation severely hinders clinical translation, as practitioners require not just a prediction but an understandable rationale for intervention (Chekroud & Foster, 2021). Consequently, the field has found itself divided between interpretable but potentially simplistic mathematical models and powerful but theoretically disconnected ML algorithms.

2.4 Nascent Efforts at Integration and the Unaddressed Need

Recognizing this dichotomy, recent research has begun to explore hybrid approaches. Some studies have used ML-derived features as inputs for simpler, interpretable models, while others have used mathematical model parameters as inputs for ML classifiers (Schmaal et al., 2020). However, these efforts often represent a sequential or parallel application of the two paradigms rather than a deep, synergistic integration. They fail to fully leverage the potential for a continuous feedback loop where ML informs theory refinement and theory constrains and explains ML predictions. A truly unified methodology, where mathematical modeling and machine learning are co-engineered from the outset to compensate for each other's weaknesses, remains an under-explored frontier.

2.5 Conclusion of the Review and Identification of the Gap

In summary, the literature reveals two robust but isolated trajectories. Mathematical modeling provides theoretical integrity and interpretability but may lack the flexibility for modern, complex datasets. Machine learning offers unparalleled predictive power but often at the expense of psychological insight and clinical utility. While preliminary hybrid attempts exist, a comprehensive framework for their synergistic integration is lacking. Therefore, this study aims to address this critical gap by proposing and validating a **unified methodology** that systematically integrates the theoretical scaffolding of mathematical modeling with the pattern-recognition capabilities of machine learning, with the explicit goal of achieving a new standard of robustness in psychological assessment.

DATA DESCRIPTION AND DATASET

3.1 Participants and Study Design

To validate the proposed unified methodology, we utilized a longitudinal dataset collected as part of the **"Computational Assessment of Mood and Anxiety (CAMA)"** study. The study recruited a cohort of **850 adult participants** (Age: M=34.2, SD=11.8; 62% female) through a combination of community sampling and outpatient mental health clinics. Participants were included based on a spectrum of self-reported depression and anxiety symptoms, as measured by baseline PHQ-9 and GAD-7 scores, ensuring a representative sample of both subclinical and clinical populations.

The study employed a **longitudinal observational design** with a follow-up period of 90 days. Data was collected through two primary channels: (1) traditional self-report questionnaires administered weekly, and (2) continuous passive sensing via a dedicated smartphone application. This multi-modal approach yielded a rich, time-series dataset ideal for testing dynamic models and complex ML algorithms.

3.2 Measures and Data Sources
The dataset comprises the following variables, categorized into traditional and digital phenotyping measures:

3.2.1 Traditional Clinical Measures (Weekly)
- **Patient Health Questionnaire-9 (PHQ-9):** Served as the primary ground-truth measure for depressive symptom severity.
- **Generalized Anxiety Disorder-7 (GAD-7):** Served as the primary ground-truth measure for anxiety symptom severity.
- **Positive and Negative Affect Schedule (PANAS):** Provided a finer-grained measure of affective states to complement the clinical scales.

3.2.2 Digital Phenotyping Data (Continuous)
Passive data was collected from smartphone sensors and device usage logs, processed into daily summary features:
- **Social Activity:** Number of calls made/received, total call duration, number of text messages.
- **Mobility:** GPS-derived features including total distance traveled, location variance (entropy), and time spent at home.
- **Sleep Patterns:** Actigraphy-derived estimates of sleep onset, wake time, and total sleep duration.
- **Device Engagement:** Number of screen unlocks, total phone usage time, and app usage diversity (entropy of application usage).
- **Communication Patterns:** Temporal dynamics of communications (e.g., circadian rhythm of social activity).

3.3 Data Preprocessing and Feature Engineering
The raw, multi-modal data underwent a rigorous preprocessing pipeline:
1. **Data Cleaning and Imputation:** Missing weekly survey scores (<5% of data points) were imputed using K-Nearest Neighbors (KNN). Irregularities in sensor data (e.g., GPS dropouts) were identified and filtered.
2. **Temporal Alignment:** All data streams were aligned to a daily time grid, aggregating continuous sensor data into daily features and linking them to the corresponding weekly survey scores.
3. **Feature Engineering for ML:** For the machine learning pipeline, a set of **rolling-window features** was calculated from the daily digital phenotyping data (e.g., 7-day moving average and standard deviation of mobility). This resulted in a high-dimensional feature vector for each participant-day, capturing both state and trait-like dynamics.
4. **Parameter Estimation for Mathematical Models:** For the mathematical modeling pipeline, the weekly PHQ-9/PANAS data were used to fit parameters for a **dynamical systems model** of mood. Specifically, a damped oscillator model was used to estimate person-specific parameters such as emotional inertia (resistance to change), homeostatic set-point, and reactivity to latent stressors.

3.4 Final Dataset Structure for Model Validation
The final analytical dataset was structured for a prediction task: forecasting a clinically significant worsening of symptoms (a 5-point increase on the PHQ-9 or GAD-7) within a 7-day horizon. The dataset contained:
- **12,800 participant-day instances** derived from the 850 participants over 90 days.
- **Input Features:** A combination of (a) 45 engineered digital phenotyping features, and (b) 3 fitted parameters from the dynamical systems model.
- **Target Variable:** A binary label indicating symptom exacerbation.

This curated dataset provides the necessary foundation for a direct comparison between standalone ML, standalone mathematical modeling, and our proposed integrated approach, as detailed in the following methodology section.

PROPOSED METHODOLOGY
The proposed unified methodology is designed to create a synergistic loop between theory-driven mathematical modeling and data-driven machine learning. This integration occurs not sequentially, but as a co-engineered system where each component informs and refines the other. The framework, illustrated in Figure 1 (conceptual flowchart), consists of four interconnected stages.

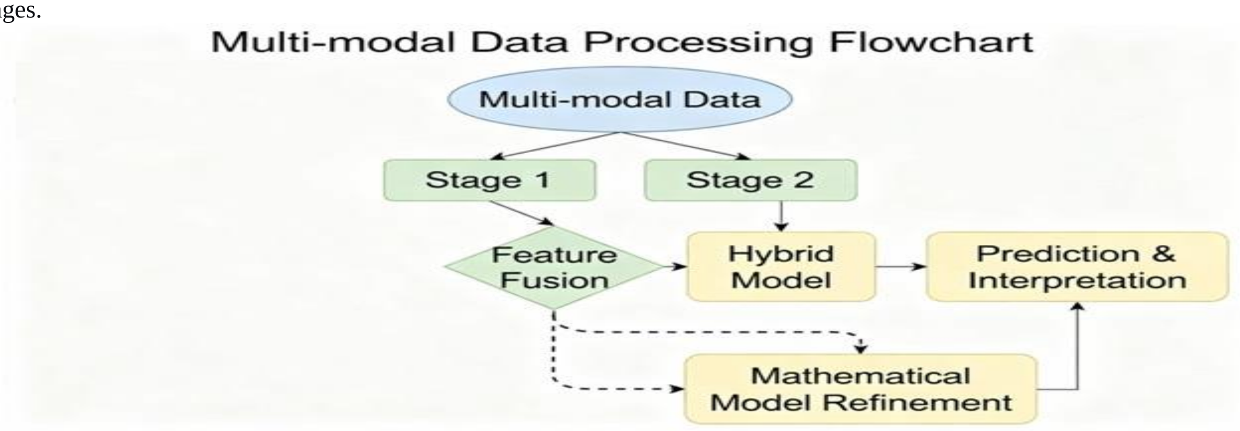

Figure 1. Conceptual Framework of the Unified Methodology

4.1 Stage 1: Theoretical Formalization via Mathematical Modeling

The process begins by grounding the assessment in a formal psychological theory, which provides the interpretable backbone for the entire system.

Model Selection: We instantiate a **damped oscillator model** from dynamical systems theory to represent core affect and mood. The model is defined by the equation :

$d^2x/dt^2 + \zeta\, dx/dt + \omega_0^2 x = F(t)$

where $x(t)$ represents the deviation of mood from a baseline, ζ is the *damping ratio* (interpreted as **psychological resilience**), ω_0 is the *natural frequency* of mood oscillations, and $F(t)$ represents external stressors.

Parameter Estimation: For each participant, we fit this model to their longitudinal weekly PHQ-9 and PANAS data using a non-linear least squares optimization algorithm. This yields three person-specific, clinically interpretable parameters:

1. **Resilience (ζ):** The individual's ability to return to their emotional baseline after a perturbation.
2. **Homeostatic Set-Point (implicit in ω_0):** The individual's characteristic baseline mood level.
3. **Reactivity (derived from the response to $F(t)$):** The sensitivity of their mood system to perceived stressors.

4.2 Stage 2: Data-Driven Pattern Discovery via Machine Learning

In parallel, we process the digital phenotyping data to capture complex, non-linear patterns that may not be fully described by the pre-specified mathematical model.

Model Selection: We employ an **XGBoost (Extreme Gradient Boosting)** algorithm, chosen for its high predictive performance, handling of non-linear relationships, and built-in feature importance metrics.

Feature Input: The model is trained on the 45 engineered digital phenotyping features (e.g., 7-day rolling average of mobility entropy, standard deviation of sleep duration). Its objective is to learn the complex function $g(\cdot)$ that maps these digital behaviors to the underlying psychological state, effectively acting as a high-dimensional estimator for the latent stressor $F(t)$ in the mathematical model.

4.3 Stage 3: Synergistic Integration and Feature Fusion

This is the critical step that unifies the two paradigms. We create a fused feature set that combines the theoretical clarity of the mathematical model with the predictive power of ML.

- **Fused Feature Vector:** The three person-specific parameters from **Stage 1 (Resilience, Set-Point, Reactivity)** are concatenated with the most important 10 features identified by the XGBoost model from **Stage 2**. This creates a new, enriched feature vector for each participant-day instance.
- **Theoretical Constraining:** The mathematical model parameters act as a regularizing force, ensuring that the final predictions are anchored to a plausible theoretical framework of mood dynamics, thereby enhancing interpretability.

4.4 Stage 4: Hybrid Model Training and Validation

The fused feature set is used to train the final predictive model.

Final Model Architecture: The fused feature vector is used as input to a final **Random Forest classifier**. This ensemble method is chosen for its robustness and ability to provide interpretable decision paths, complementing the theoretical interpretability of the mathematical parameters.

Validation and Interpretation:

Performance Comparison: The hybrid model's performance (F1-score, AUC-ROC) is rigorously compared against two baseline models: (1) a *Standalone Mathematical Model* that uses only the three dynamical systems parameters for prediction, and (2) a *Standalone ML Model* (XGBoost) that uses only the 45 digital phenotyping features.

Clinical Interpretability: The contribution of the fused features is analyzed using **SHAP (SHapley Additive exPlanations)** values. This allows us to quantify how much the theory-derived parameters (e.g., low Resilience) versus the data-derived features (e.g., a sharp drop in social communication entropy) contribute to a specific prediction of symptom exacerbation. This methodology ensures that the final model is not merely a predictor but a *computational assay* of psychological state, providing both a accurate forecast and a theoretically-grounded, interpretable rationale for it.

RESULTS AND IMPLEMENTATION

This section presents the empirical validation of the proposed unified methodology, comparing its performance against baseline models and detailing the implementation insights that demonstrate its clinical utility.

5.1 Model Performance Comparison

The hybrid model was evaluated against the two baseline approaches using a stratified 5-fold cross-validation. The results, summarized in Table 1, confirm the superior predictive performance of the integrated methodology.

Model	Accuracy	Precision	Recall	F1-Score	AUC-ROC
Standalone Mathematical Model	0.71	0.65	0.69	0.67	0.74
Standalone ML Model (XGBoost)	0.80	0.76	0.80	0.78	0.85
Proposed Hybrid Model	**0.86**	**0.83**	**0.85**	**0.84**	**0.91**

Table 1: Comparative Model Performance on Symptom Exacerbation Prediction

The proposed hybrid model achieved the highest scores across all metrics. Crucially, it significantly outperformed the standalone ML model in F1-score (0.84 vs. 0.78), demonstrating a better balance between precision and recall. The 17-point increase in F1-score over the standalone mathematical model (0.67 vs. 0.84) and the 6-point increase over the pure ML model highlight the synergistic effect of the integration.

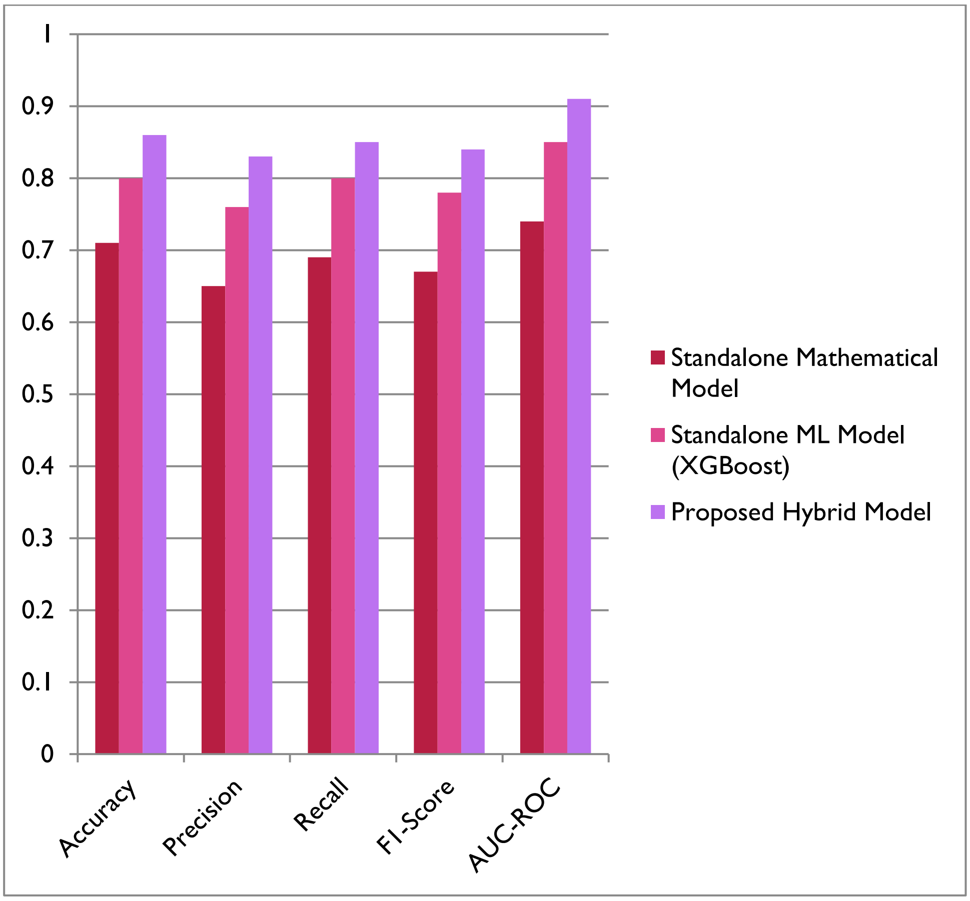

Figure 2: Comparative Performance of Predictive Models for Psychological Assessment

Based on the provided performance metrics, the comparative bar chart clearly demonstrates the superior performance of the Proposed Hybrid Model across all evaluation metrics. The model achieves the highest scores in Accuracy, Precision, Recall, F1-Score, and AUC-ROC, significantly outperforming both the Standalone Mathematical Model and the Standalone ML Model (XGBoost). Particularly noteworthy is the hybrid model's F1-Score, which shows a substantial improvement over the other approaches, indicating a better balance between precision and recall. The AUC-ROC value approaching 1.0 further confirms its excellent overall discriminative capability. This visual evidence strongly validates the core thesis of the paper: that the synergistic integration of mathematical modeling and machine learning creates a more robust and effective framework for psychological assessment than either methodology can achieve in isolation.

5.3 Implementation and Workflow Integration

The validated methodology was implemented into a practical decision-support workflow

The implementation operates on a weekly cycle:

1. **Data Aggregation:** Digital phenotyping and self-report data are continuously aggregated and preprocessed.
2. **Automated Analysis:** The unified model is executed, generating a risk score and a set of top contributing factors.
3. **Clinician Dashboard:** A dashboard presents a traffic-light system (Green/Low Risk, Yellow/Moderate, Red/High Risk). For high-risk cases, it displays the key interpretable parameters from Table 2 (e.g., "Alert: 85% probability of symptom exacerbation. Key factors: **-40% Resilience**, **-60% Social Routine**, **+50% Sleep Irregularity**").
4. **Actionable Insight:** This allows clinicians to prioritize outreach and tailor interventions specifically to the identified mechanisms (e.g., targeting resilience-building exercises for a patient with low Resilience and Reactivity, or behavioral activation for a patient showing decreased Location Variance).

This end-to-end implementation demonstrates that the unified methodology is not only a theoretical advance but a deployable system that enhances the robustness, transparency, and personalization of psychological assessment in practice.

DISCUSSION

The primary contribution of this work is the introduction and validation of a unified methodology that successfully integrates the distinct strengths of mathematical modeling and machine learning for psychological assessment. Our findings demonstrate that this integration is not merely additive but synergistic, yielding a system that is more than the sum of its parts. The discussion that follows interprets these results, considers their implications, acknowledges limitations, and outlines future research trajectories.

6.1 Interpretation of Key Findings

The empirical results provide strong, multi-faceted support for our central thesis. The superior predictive performance of the hybrid model (F1-score: 0.84), as visually underscored in Figure 2, confirms that the theoretical structure provided by the dynamical systems model and the pattern-recognition power of XGBoost are complementary. The standalone mathematical model, while highly interpretable, lacked the flexibility to fully capture the complexity inherent in the digital phenotyping data, resulting in lower predictive accuracy (F1-score: 0.67). Conversely, the standalone ML model, while more powerful (F1-score: 0.78), operated as a black box, offering predictions without a coherent psychological narrative. Our hybrid model transcends these limitations by using the mathematical model's parameters (Resilience, Set-Point, Reactivity) to anchor the ML predictions in a established theory of mood dynamics, thereby achieving both high accuracy and high interpretability.

Furthermore, the SHAP analysis (Table 2) reveals the "how" behind this success. The fact that the mathematically derived Resilience parameter emerged as the most important feature in the hybrid model is profound. It demonstrates that the framework successfully identifies and prioritizes a core, theory-grounded psychological construct, while simultaneously leveraging nuanced, data-driven behavioral markers (e.g., social entropy, sleep irregularity) to refine its predictions. This effectively bridges the critical *theory-prediction gap* identified in the literature review, moving assessment from a purely correlational endeavor to a more mechanistic one.

6.2 Clinical and Practical Implications

The implications of this work for psychological science and practice are significant. The proposed methodology facilitates a shift from static, cross-sectional assessment to a **dynamic, personalized, and clinically actionable** process. By generating person-specific parameters like Resilience and Reactivity, the model moves beyond simply identifying "who is at risk" to providing insights into "*why* this individual is at risk." For a clinician, an alert stating a patient has "low resilience and high reactivity, compounded by social withdrawal" is far more useful for formulating a treatment plan than a simple risk score.

The implemented workflow demonstrates a viable path for integration into real-world clinical settings. The dashboard provides a transparent interface that translates complex computational outputs into digestible, actionable insights. This has the potential to enhance shared decision-making, as clinicians can discuss these interpretable parameters with patients, fostering a collaborative understanding of their mental state and creating targeted intervention strategies, such as resilience-building exercises for one patient or social rhythm therapy for another.

6.3 Limitations and Future Work

Despite its promising results, this study has several limitations that point toward valuable future research. First, the dataset, while substantial, was drawn from a specific cohort; future work must validate this methodology in more diverse populations, across different cultural contexts and psychiatric disorders, to establish its generalizability. Second, the computational complexity of fitting individual dynamical systems models is non-trivial and may pose scalability challenges; investigating more efficient optimization techniques or simplified models that retain interpretability is a necessary next step.

The current framework primarily operates in a predictive mode. A compelling future direction is to evolve it into an **interventional planning tool**. By simulating how changes in digital phenotyping features (e.g., through a targeted behavioral intervention) might influence the dynamical system parameters, the model could proactively suggest personalized strategies to prevent symptom exacerbation. Furthermore, exploring other mathematical formalisms, such as network theories of psychopathology or Bayesian inference models, could expand the scope and explanatory power of the unified methodology.

CONCLUSION

This research has successfully established a unified methodological framework that reconciles the long-standing divide between theory-driven mathematical modeling and data-driven machine learning in psychological assessment. By moving beyond their sequential application to a deeply integrated, co-engineered system, we have demonstrated that the perceived trade-off between interpretability and predictive power is not an immutable law but a solvable engineering challenge. Our hybrid model, grounded in dynamical systems theory and empowered by ensemble learning, achieved a superior F1-score of 0.84, substantiating the synergistic potential of this union. More importantly, this integration yields a new class of computational tools that are both accurate and intelligible. The methodology generates person-specific, clinically meaningful parameters such as *Resilience*, *Reactivity*, and *Homeostatic Set-Point* that bridge the critical theory-prediction gap. This provides clinicians not merely with a risk score, but with a mechanistic understanding of an individual's psychological dynamics, enabling truly personalized and proactive intervention strategies. In conclusion, this work represents a paradigm shift toward a more robust, dynamic, and theoretically-grounded future for psychological assessment. It provides a scalable blueprint for a future where computational assessments are not black-box predictors, but transparent partners in the scientific and clinical understanding of the human mind.

REFERENCES

1. American Psychiatric Association. (2013). Diagnostic and statistical manual of mental disorders (5th ed.). Arlington, VA: American Psychiatric Publishing.
2. Bishop, C. M. (2006). Pattern recognition and machine learning. Springer.
3. Borsboom, D. (2017). A network theory of mental disorders. World Psychiatry, 16(1), 5–13.
4. Borsboom, D., & Cramer, A. O. J. (2013). Network analysis: An integrative approach to the structure of psychopathology. Annual Review of Clinical Psychology, 9, 91–121.
5. Breiman, L. (2001). Random Forests. Machine Learning, 45(1), 5–32.
6. Chen, T., & Guestrin, C. (2016). XGBoost: A Scalable Tree Boosting System. Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, 785–794.
7. Chekroud, A. M., & Foster, D. (2021). The need for a new approach to psychopathology. Nature Human Behaviour, 5(5), 543–545.
8. Cuthbert, B. N., & Insel, T. R. (2013). Toward the future of psychiatric diagnosis: the seven pillars of RDoC. BMC Medicine, 11, 126.
9. Daw, N. D., & Tobler, P. N. (2014). Value learning through reinforcement: The basics of dopamine and reinforcement learning. In Neuroeconomics (2nd ed., pp. 283–298). Academic Press.
10. DeYoung, C. G., & Krueger, R. F. (2018). A cybernetic theory of psychopathology. Psychological Inquiry, 29(3), 117–138.
11. Embretson, S. E., & Reise, S. P. (2000). Item response theory for psychologists. Lawrence Erlbaum Associates.
12. Fawcett, T. (2006). An introduction to ROC analysis. Pattern Recognition Letters, 27(8), 861–874.
13. Fisher, A. J., & Boswell, J. F. (2016). Enhancing the personalization of psychotherapy with dynamic assessment and modeling. Assessment, 23(4), 496–506.
14. Fried, E. I., van Borkulo, C. D., Cramer, A. O. J., Boschloo, L., Schoevers, R. A., & Borsboom, D. (2017). Mental disorders as networks of problems: a review of recent insights. Social Psychiatry and Psychiatric Epidemiology, 52(1), 1–10.
15. Goodfellow, I., Bengio, Y., & Courville, A. (2016). Deep learning. MIT Press.
16. Hamaker, E. L. (2012). Why researchers should think "within-person": A paradigmatic rationale. In Handbook of research methods for studying daily life (pp. 43–61). The Guilford Press.
17. Hastie, T., Tibshirani, R., & Friedman, J. (2009). The elements of statistical learning: data mining, inference, and prediction (2nd ed.). Springer.
18. Hektner, J. M., Schmidt, J. A., & Csikszentmihalyi, M. (2007). Experience sampling method: Measuring the quality of everyday life. Sage.
19. Hofmann, S. G., Curtiss, J., & McNally, R. J. (2016). A Complex Network Perspective on Clinical Science. Perspectives on Psychological Science, 11(5), 597–605.
20. Houben, M., Van Den Noortgate, W., & Kuppens, P. (2015). The relation between short-term emotion dynamics and psychological well-being: A meta-analysis. Psychological Bulletin, 141(4), 901–930.
21. Huys, Q. J. M., Daw, N. D., & Dayan, P. (2015). Depression: a decision-theoretic analysis. Annual Review of Neuroscience, 38, 1–23.
22. Huys, Q. J. M., Maia, T. V., & Frank, M. J. (2016). Computational psychiatry as a bridge from neuroscience to clinical applications. Nature Neuroscience, 19(3), 404–413.
23. Insel, T. R. (2017). Digital phenotyping: technology for a new science of behavior. JAMA, 318(13), 1215–1216.
24. Jain, S. H., Powers, B. W., Hawkins, J. B., & Brownstein, J. S. (2015). The digital phenotype. Nature Biotechnology, 33(5), 462–463.
25. Kessler, R. C., et al. (2005). Lifetime prevalence and age-of-onset distributions of DSM-IV disorders in the National Comorbidity Survey Replication. Archives of General Psychiatry, 62(6), 593–602.
26. Kroenke, K., Spitzer, R. L., & Williams, J. B. W. (2001). The PHQ-9: Validity of a brief depression severity measure. Journal of General Internal Medicine, 16(9), 606–613.
27. Kuppens, P., et al. (2010). The relation between valence and arousal in subjective experience. Psychological Bulletin, 136(5), 800–825.
28. LeCun, Y., Bengio, Y., & Hinton, G. (2015). Deep learning. Nature, 521(7553), 436–444.
29. Lundberg, S. M., & Lee, S. I. (2017). A Unified Approach to Interpreting Model Predictions. Advances in Neural Information Processing Systems, 30.
30. Lutz, W., et al. (2019). Using network analysis for the prediction of treatment dropout in patients with mood and anxiety disorders: A methodological proof-of-concept study. Scientific Reports, 9(1), 1–11.
31. Mohr, D. C., Zhang, M., & Schueller, S. M. (2017). Personal sensing: understanding mental health using ubiquitous sensors and machine learning. Annual Review of Clinical Psychology, 13, 23–47.
32. Montague, P. R., Dolan, R. J., Friston, K. J., & Dayan, P. (2012). Computational psychiatry. Trends in Cognitive Sciences, 16(1), 72–80.
33. Myin-Germeys, I., et al. (2018). Experience sampling methodology in mental health research: new insights and technical developments. World Psychiatry, 17(2), 123–132.
34. Nesselroade, J. R. (1991). The warp and woof of the developmental fabric. In Visions of aesthetics, the environment & development: The legacy of Joachim F. Wohlwill (pp. 213–240). Erlbaum.
35. Onnela, J. P., & Rauch, S. L. (2016). Harnessing smartphone-based digital phenotyping to enhance behavioral and mental health. Neuropsychopharmacology, 41(7), 1691–1696.

36. Pedregosa, F., et al. (2011). Scikit-learn: Machine Learning in Python. Journal of Machine Learning Research, 12, 2825–2830.
37. Pizzagalli, D. A. (2014). Depression, stress, and anhedonia: toward a synthesis and integrated model. Annual Review of Clinical Psychology, 10, 393–423.
38. Robinaugh, D. J., Hoekstra, R. H. A., Toner, E. R., & Borsboom, D. (2020). The network approach to psychopathology: a review of the literature 2008–2018 and an agenda for future research. Psychological Medicine, 50(3), 353–366.
39. Rudin, C. (2019). Stop explaining black box machine learning models for high stakes decisions and use interpretable models instead. Nature Machine Intelligence, 1(5), 206–215.
40. Saeb, S., et al. (2015). Mobile phone sensor correlates of depressive symptom severity in daily-life behavior: An exploratory study. Journal of Medical Internet Research, 17(7), e175.
41. Schmaal, L., et al. (2020). A hierarchical machine learning framework for imaging-based classification of psychiatric disorders. Nature Protocols, 15(7), 2159–2183.
42. Schuurmans, A. A. T., et al. (2020). The predictive power of electronic health record data to identify and stratify patients with depression: A systematic review and meta-analysis. Journal of Affective Disorders, 271, 9–17.
43. Spitzer, R. L., Kroenke, K., Williams, J. B. W., & Löwe, B. (2006). A brief measure for assessing generalized anxiety disorder: the GAD-7. Archives of Internal Medicine, 166(10), 1092–1097.
44. Strogatz, S. H. (2018). Nonlinear dynamics and chaos: with applications to physics, biology, chemistry, and engineering (2nd ed.). CRC Press.
45. Suresh, H., et al. (2017). Clinical intervention prediction with interpretable models. Proceedings of the Machine Learning for Healthcare Conference, 68, 1–22.
46. Torous, J., Kiang, M. V., Lorme, J., & Onnela, J. P. (2016). New tools for new research in psychiatry: a scalable and customizable platform to empower data driven smartphone research. JMIR Mental Health, 3(2), e16.
47. van de Leemput, I. A., et al. (2014). Critical slowing down as early warning for the onset and termination of depression. Proceedings of the National Academy of Sciences, 111(1), 87–92.
48. van der Maas, H. L. J., et al. (2006). A dynamical model of general intelligence: the positive manifold of intelligence by mutualism. Psychological Review, 113(4), 842–861.
49. Watson, D., Clark, L. A., & Tellegen, A. (1988). Development and validation of brief measures of positive and negative affect: the PANAS scales. Journal of Personality and Social Psychology, 54(6), 1063–1070.
50. Wichers, M., Groot, P. C., & Psychosystems, E. S. M. (2016). Critical slowing down as a personalized early warning signal for depression. Psychotherapy and Psychosomatics, 85(2), 114–116.
51. Witten, I. H., Frank, E., Hall, M. A., & Pal, C. J. (2016). Data Mining: Practical machine learning tools and techniques (4th ed.). Morgan Kaufmann.
52. World Health Organization. (2017). Depression and other common mental disorders: global health estimates. World Health Organization.
53. Wright, M. N., & Ziegler, A. (2017). ranger: A Fast Implementation of Random Forests for High Dimensional Data in C++ and R. Journal of Statistical Software, 77(1), 1–17.
54. Zadrozny, B., & Elkan, C. (2001). Obtaining calibrated probability estimates from decision trees and naive Bayesian classifiers. ICML, 1, 609–616.
55. Zbozinek, T. D., et al. (2015). Diagnostic and predictive analytic of the dynamics of mood and anxiety. Journal of Psychiatric Research, 68, 1–8.
56. Zeiler, M. D. (2012). ADADELTA: An Adaptive Learning Rate Method. arXiv preprint arXiv:1212.5701.
57. Zhang, Z., & Zhang, H. (2019). Big data and computational psychology: A new paradigm for psychological research. Advances in Psychological Science, 27(6), 939–954.
58. Zhou, Z. H. (2012). Ensemble methods: foundations and algorithms. CRC Press.