

Identifying Predictors of Treatment Dropout Using ML Analysis of Motivational Interviewing Processes and Attachment-Related Avoidance

Moriel. Davila Tenorio^{1*}, Jaime. Coello-Bravo², Micaela. Reyes Zevallos¹

¹ School of Higher Studies (F.E.S.) Zaragoza, National Autonomous University of Mexico, Mexico City CP 09230, Mexico

² Department of Psychology, Virginia Commonwealth University, United States

* Corresponding author email address: moridaviteno@gmail.com

Article Info

Article type:

Original Research

How to cite this article:

Davila Tenorio, M., Coello-Bravo, J., & Reyes Zevallos, M. (2026). Identifying Predictors of Treatment Dropout Using ML Analysis of Motivational Interviewing Processes and Attachment-Related Avoidance. *Journal of Assessment and Research in Applied Counseling*, 8(1), 1-12. <http://dx.doi.org/10.61838/kman.jarac.5194>



© 2026 the authors. Published by KMAN Publication Inc. (KMANPUB), Ontario, Canada. This is an open access article under the terms of the Creative Commons Attribution-NonCommercial 4.0 International (CC BY-NC 4.0) License.

ABSTRACT

Objective: This study aimed to utilize advanced machine learning algorithms to evaluate whether baseline attachment-related avoidance and specific linguistic frequencies of client change talk and sustain talk during early Motivational Interviewing sessions can reliably predict subsequent outpatient psychotherapeutic treatment dropout.

Methods and Materials: The study employed a prospective, longitudinal observational design with a sample of $N = 485$ adults from Mexico City. Dropout was defined as missing three consecutive sessions or mutually agreeing to terminate prematurely. Data collection included baseline demographics, the Experiences in Close Relationships-Revised questionnaire to assess attachment-related avoidance, and Motivational Interviewing Skill Code ratings derived from the audio-recorded and transcribed first two therapy sessions. Machine learning analysis incorporated k -nearest neighbors imputation, Synthetic Minority Over-sampling Technique (SMOTE) for class imbalance, and evaluated Random Forest, Support Vector Machines (SVM), and Extreme Gradient Boosting (XGBoost) classifiers using stratified 10-fold cross-validation. SHapley Additive exPlanations (SHAP) values were utilized for model interpretability.

Findings: Results indicated that 157 out of 485 participants (32.4%) dropped out of treatment. Bivariate analyses revealed that dropouts had significantly higher attachment-related avoidance ($M = 4.25$) compared to completers ($M = 3.15$). During early sessions, the dropout cohort exhibited significantly more Sustain Talk ($M = 24.6$ vs. $M = 15.2$) and less Change Talk ($M = 19.8$ vs. $M = 31.4$), alongside receiving fewer complex reflections and lower empathy ratings from therapists. The XGBoost model demonstrated superior predictive performance ($AUC = 0.87$), followed by Random Forest ($AUC = 0.84$) and SVM ($AUC = 0.78$). SHAP analyses identified attachment avoidance and patient sustain talk as the strongest overall predictors, highlighting a critical interaction effect where

highly avoidant patients receiving fewer therapist complex reflections exhibited a compounding probability of dropout.

Conclusion: Machine learning algorithms can accurately predict treatment dropout from early clinical interactions, emphasizing that patient retention fundamentally depends on the complex interplay between severe attachment-related avoidance and the therapist's technical proficiency in managing resistance.

Keywords: *Treatment Dropout; Machine Learning; Motivational Interviewing; Attachment Avoidance; Therapeutic Alliance*

1. Introduction

The premature termination of psychological and psychiatric treatment, colloquially and clinically referred to as treatment dropout, remains one of the most pervasive, costly, and therapeutically challenging impediments to the effective delivery of modern mental healthcare (Banham & Schweitzer, 2016). Dropout is generally characterized by a patient's unilateral decision to cease attending scheduled therapeutic sessions prior to the successful completion of a mutually agreed-upon clinical protocol, or before achieving any clinically significant or sustained symptomatic relief. Extensive systematic reviews and sophisticated meta-analytical investigations have repeatedly demonstrated that substantial proportions of treatment-seeking individuals disengage prematurely across a vast array of psychological disorders and diverse treatment modalities. For instance, the specific nature, timing, and symptomatic trajectories of treatment dropout in transdiagnostic and single-diagnosis cognitive-behavioral therapy (CBT) for severe anxiety disorders exhibit highly complex patterns that frequently result in early, abrupt clinical termination (Bentley et al., 2021). Similar attrition challenges have been thoroughly documented on a global scale, with critically high dropout rates significantly impairing the established efficacy of individual psychotherapy for generalized anxiety disorder (Gersh et al., 2017). Furthermore, this phenomenon negatively impacts the long-term patient adherence to structured CBT protocols that are specifically designed for obsessive-compulsive disorder (Leeuwerik et al., 2019). The clinical, ethical, and economic ramifications of such therapeutic attrition are profound, predictably leading to unresolved psychological distress for the patient, severely exacerbated psychiatric symptomatology over time, the highly inefficient allocation of finite community clinical resources, and an exponentially increased financial burden on broader social and national healthcare infrastructures. Even in rigorously controlled experimental environments that are explicitly designed to maximize patient adherence and engagement—such as the large-scale deployment of supported mindfulness-based

cognitive therapy self-help compared directly to traditional supported cognitive behavioral therapy self-help for depressed adults—retention metrics remain a paramount indicator of treatment acceptability and overall clinical effectiveness (Strauss et al., 2023). Consequently, identifying and deeply understanding the multidimensional etiology of treatment dropout is an absolute clinical prerequisite for optimizing individual patient outcomes and systematically enhancing therapeutic engagement across all mental health disciplines.

The specific etiology of premature treatment termination is rarely singular or linear; rather, it is a highly multifactorial phenomenon that is heavily influenced by broader systemic contexts, highly specific socio-demographic variables, and the unique structural nuances of varying therapeutic modalities. At a macroscopic, systemic level, premature disengagement from structured, rule-bound institutions is frequently driven by profound socio-economic dysfunctions and deeply ingrained prejudices, a destructive dynamic clearly visible in adjacent academic fields, such as the severe school dropout rates observed among marginalized demographics like Roma children due to pervasive anti-gypsyism and systemic educational failures (Rotaru, 2019). In specialized adult clinical settings, complex contextual pressures, such as navigating court-mandated treatment status combined synergistically with active antisocial personality traits, create highly volatile, non-linear interactive effects that strongly precipitate early and hostile dropout in community substance abuse rehabilitation programs (Daughters et al., 2024). Different psychotherapeutic paradigms also inherently face unique, protocol-specific retention challenges based on their distinct mechanisms of action. For instance, while solution-focused brief therapy is structurally designed to yield rapid, highly visible positive outcomes within a constrained timeframe, maintaining the necessary long-term adherence to its core principles necessitates exceptionally high initial client engagement and buy-in (Cortes et al., 2016). Acceptance and Commitment Therapy (ACT), a third-wave behavioral intervention that has proven highly effective for diverse and complex clinical challenges ranging from increasing vital

distress tolerance in mothers of children diagnosed with autism (Ahmadi & Raeesi, 2018) to managing the profound, existential psychological stress associated with medical infertility (Peterson & Eifert, 2011), similarly requires patients to actively confront, rather than experientially avoid, deep psychological discomfort. This intensive process of acceptance can easily trigger early emotional withdrawal and treatment refusal. Furthermore, highly structured, time-limited interventions like cognitive analytic therapy (CAT) have been subjected to rigorous, continuous benchmarking regarding both treatment refusal and clinical dropout rates, emphasizing the absolute necessity to empirically evaluate ongoing treatment acceptability (Simmonds-Buckley et al., 2022). This necessity is particularly acute and mathematically pronounced when treating chronic, severe psychopathology, where the very core nature of the psychological disorder inherently drives interpersonal instability and reflexive treatment rejection.

Nowhere is the clinical challenge of treatment retention more acute or statistically evident than in the psychiatric management of severe personality pathology, most notably Borderline Personality Disorder (BPD). Patients presenting with clinical BPD frequently exhibit intense affective instability, chronic and active suicidality, and highly chaotic, polarized interpersonal relations, making them highly susceptible to premature therapeutic dropout at rates frequently exceeding 40% to 50% in standard outpatient settings (Linehan et al., 2015). Specialized, highly structured evidence-based treatments have been painstakingly developed to specifically target these severe behavioral manifestations, yet successfully retaining these vulnerable patients in these demanding, multi-modal protocols remains an arduous clinical undertaking. Comprehensive, longitudinal randomized trials directly comparing intensive outpatient therapies like Schema-Focused Therapy and Transference-Focused Psychotherapy for BPD specifically highlight the critical, life-saving necessity of meticulously managing the therapeutic alliance to proactively prevent attrition (Giesen-Bloo et al., 2006). Dialectical Behavior Therapy (DBT), heavily recognized as a frontline psychological intervention for severe emotional dysregulation, has been heavily scrutinized in academic literature regarding its overall clinical effectiveness, nuanced patient response trajectories, and concerning dropout rates in highly controlled inpatient settings (Kröger et al., 2013), as well as in standard outpatient skills groups specifically tailored for mitigating severe suicidal behaviors (Stratton et al., 2020). Advanced statistical approaches, such

as mixed-effects modeling utilized in comprehensive meta-analyses of DBT, consistently underscore that while the therapeutic modality is highly efficacious in reducing self-harm, maintaining active patient engagement through the inherent emotional distress of early skills acquisition is a paramount clinical hurdle (Kliem et al., 2010). Adapting core DBT principles for vulnerable adolescents requiring prolonged exposure protocols (Lang et al., 2018) or rigorously comparing its effectiveness directly against CBT in treating complex, heavily stigmatized comorbidities like binge-eating disorder (Lammers et al., 2022) consistently reveals that the nuanced, real-time interpersonal dynamic occurring between patient and provider ultimately dictates whether the patient remains in care. To specifically address the deep-seated, rigid cognitive structures driving this interpersonal dysfunction, combined group and individual schema therapy models have been piloted with varying degrees of success regarding clinical adherence (Dickhaut & Arntz, 2014). Furthermore, intensive inpatient schema therapy for highly treatment-resistant patients relies heavily on systematically altering symptomatic distress, maladaptive coping styles, and deeply ingrained early maladaptive schemas to fundamentally foster long-term mental well-being and definitively prevent clinical elopement (Schaap et al., 2016).

Underpinning much of the interpersonal friction, resistance, and rupture that directly precipitates psychotherapeutic treatment dropout is the foundational psychological framework of adult attachment theory. Attachment theory powerfully posits that early relational experiences with primary caregivers permanently solidify into complex, internalized working models that unconsciously govern emotional regulation, epistemic trust, and rigid interpersonal expectations throughout the entirety of an individual's lifespan. These internal working models profoundly influence macroscopic relational dynamics, successfully bridging the theoretical gap between specific, targeted therapeutic interventions, such as schema therapy, and broader, highly complex psychosocial outcomes like marital intimacy (Mohammadi et al., 2025). In the intimate environment of the clinical setting, a patient's dominant attachment style largely dictates their inherent capacity to form a functional, collaborative, and trusting therapeutic alliance. Maladaptive adult attachment, specifically attachment-related avoidance, is clinically characterized by a compulsive, rigid self-reliance, an intense, visceral discomfort with emotional intimacy, and a predictable tendency to defensively deactivate the interpersonal

attachment system when faced with psychological stress or perceived emotional vulnerability. This avoidant pathology is intrinsically linked to profound emotional dysregulation, directly mediating the complex relationship between adult attachment styles and severe obsessive-compulsive symptomatology (Nielsen et al., 2025), and fundamentally altering critical cognitive emotion regulation strategies and subsequent psychological attitudes toward deep, binding interpersonal commitments, such as marriage (Rezapour et al., 2025). Furthermore, highly maladaptive attachment pairings within the context of intimate relationships significantly and demonstrably exacerbate negative emotions during interpersonal conflict, frequently escalating beyond verbal disputes into documented instances of intimate partner violence (Pudelko et al., 2025). In the direct context of severe psychopathology, highly avoidant attachment traits operate as critical psychological mediators in the ongoing development of BPD features, driven largely by severe clinical hypomentalization and profound, pervasive epistemic mistrust of caregivers and professionals alike (Kurt, 2025). When individuals burdened with high levels of attachment avoidance enter the psychotherapeutic environment, the fundamental psychological requirement of the therapeutic process—which demands vulnerable emotional disclosure, sustained interpersonal reliance, and the relinquishing of defensive autonomy—directly and acutely threatens their established psychological architecture. Consequently, these highly avoidant individuals are theoretically and empirically at a significantly elevated, compounding risk for abruptly terminating therapy as a primary self-protective mechanism utilized to rapidly re-establish perceived emotional distance and psychological safety.

To actively circumvent the rigid defensive barriers erected by profound attachment avoidance and to systematically foster intrinsic, enduring treatment engagement, mental health clinicians frequently utilize Motivational Interviewing (MI) as a crucial preliminary or highly integrated therapeutic framework. MI has undergone significant, well-documented conceptual evolution over recent decades, successfully transitioning from a highly specific, niche intervention for severe substance abuse into a ubiquitous, transdiagnostic therapeutic communication style explicitly designed to resolve deep psychological ambivalence and organically elicit sustainable behavioral change (Miller & Rollnick, 2023). The core theoretical architecture of MI hinges entirely on a highly collaborative, evocative, and profoundly empathetic relational stance that

prioritizing the client’s autonomy, amplifying their intrinsic motivations, and validating their complex emotional realities (Miller, 2023). Because of its inherently non-confrontational, non-judgmental, and highly supportive clinical nature, MI has consistently demonstrated remarkable therapeutic versatility and measurable efficacy across highly diverse, resistant, and clinically challenging populations. For instance, structured MI protocols have been successfully and safely employed to proactively mitigate severe physical aggression and measurably reduce high-risk, impulsive behaviors among vulnerable adolescent girls (Mansouri & Khodabakhshi-Koolae, 2024). It has also shown significant, quantifiable promise in dramatically enhancing treatment participation and cultivating vital emotional skills among adult populations continuously navigating the lifelong complexities of learning disabilities (Nejatifar & Abedi, 2023). Furthermore, recent pilot feasibility studies strongly indicate that the core principles of MI can be effectively and ethically adapted for young adults diagnosed with Autism Spectrum Disorder to foster initial clinical engagement and build therapeutic rapport (Pagan, 2024). Other parallel investigations highlight its statistical efficacy in substantially ameliorating chronic academic procrastination and reducing debilitating social anxiety among highly stressed university student populations (Parsafar, 2024). The vast clinical utility of MI is even clearly evident in multifaceted, highly complex medical contexts, where it has been successfully combined with third-wave interventions like ACT and specialized compassion-focused training to measurably improve psychological distress tolerance and significantly reduce health-related anxiety in patients suffering from the chronic, unpredictable progression of multiple sclerosis (Rezaei et al., 2023). A highly critical, quantifiable behavioral metric within the established MI theoretical framework is the strict linguistic categorization of client speech during the clinical hour into discrete categories: “Change Talk” (verbal statements favoring positive behavioral adjustment and deeper therapeutic engagement) and “Sustain Talk” (verbal statements actively favoring the pathological status quo and signaling rigid therapeutic resistance). The dynamic, real-time interplay between a trained therapist’s strategic reflective skills and the patient’s measurable emission of change versus sustain talk serves as a highly accurate, observable proxy for therapeutic engagement. This linguistic exchange provides a highly granular, empirical data source for mathematically predicting subsequent clinical behavior,

most notably the precise statistical likelihood of premature treatment dropout.

Despite the highly robust theoretical connections explicitly linking baseline attachment-related avoidance, the highly detailed micro-processes of Motivational Interviewing, and the overarching, costly phenomenon of psychotherapeutic treatment dropout, comprehensive empirical investigations successfully integrating these multidimensional constructs remain notably sparse within the academic literature. Traditional, linear statistical methodologies, such as standard logistic regression or basic survival analysis techniques, often mathematically fail to adequately capture the highly complex, synergistic, and profoundly non-linear interactions that occur organically within the therapeutic dyad. Human behavior within the unpredictable clinical setting is rarely, if ever, dictated by entirely isolated variables; rather, it emerges organically from a highly intricate, interconnected matrix of deeply rooted psychological traits, real-time linguistic exchanges, and shifting contextual clinical factors. To successfully address these profound analytical limitations, the current clinical research paradigm increasingly necessitates the direct application of advanced computational techniques and algorithmic processing. Machine Learning (ML) algorithms offer a highly sophisticated, data-driven mathematical approach to predictive modeling, fully capable of simultaneously processing vast, multidimensional arrays of clinical, demographic, and linguistic features to reliably identify subtle behavioral patterns and complex, multi-variable interactions that consistently elude standard, linear statistical inference. By deploying advanced supervised ML classification algorithms, such as Random Forests or Extreme Gradient Boosting architectures, clinical researchers can construct highly accurate, broadly generalizable mathematical models. These models are uniquely capable of calculating a highly individualized, precise probability of early treatment termination based entirely upon baseline psychometric data and the initial, quantifiable linguistic exchanges observed during the very first therapeutic encounter. Such advanced predictive capabilities hold immense, immediate translational clinical value, allowing overwhelmed mental health systems to proactively and accurately identify high-risk individuals from the very first clinical session. This foresight enables practitioners to rapidly deploy targeted, highly personalized treatment retention strategies, thereby successfully optimizing long-term therapeutic outcomes, significantly

minimizing critical resource attrition, and directly improving the standard of global mental healthcare.

Therefore, the primary aim of the present study is to utilize advanced machine learning algorithms to evaluate whether baseline attachment-related avoidance and the specific linguistic frequencies of client change talk and sustain talk during early Motivational Interviewing sessions can reliably predict subsequent outpatient treatment dropout.

2. Methods and Materials

2.1. Study Design and Participants

This research utilized a prospective, longitudinal observational design to investigate the predictors of early treatment termination among individuals receiving outpatient psychotherapy. The sample was drawn from multiple public and private community mental health and substance use treatment centers located in Mexico City, Mexico. Participants were consecutively recruited during their intake process over a span of twenty-four months to ensure a representative cross-section of treatment-seeking individuals. The final, exact sample size consisted of $N = 485$ adult patients who agreed to participate and provided written informed consent. Inclusion criteria mandated that participants must be at least 18 years of age, fluent in the Spanish language, and initiating a new episode of psychological treatment utilizing a Motivational Interviewing clinical framework. Individuals were specifically excluded from the study if they exhibited severe cognitive impairments, acute and active psychotic symptoms, or required immediate inpatient stabilization, as these acute clinical factors could profoundly confound the outpatient therapeutic process and the naturalistic manifestation of interpersonal attachment behaviors. Treatment dropout, serving as the primary outcome variable for the predictive algorithms, was operationally defined as the premature cessation of therapy. This was specifically characterized by a patient missing three consecutively scheduled therapy sessions without prior notification, or failing to formally and mutually agree with the attending therapist to terminate the treatment prior to the successful completion of the prescribed clinical protocol.

2.2. Measures

To systematically capture the diverse variables of interest required for the computational models, a comprehensive battery of standardized psychometric instruments and

observational coding systems was deployed. Attachment-related avoidance was quantified utilizing a culturally validated Mexican-Spanish adaptation of the Experiences in Close Relationships-Revised questionnaire. This self-report instrument requires participants to rate their level of agreement with various statements regarding their subjective comfort with emotional intimacy and their reliance on romantic partners or close others. The tool yields a continuous, standardized score for attachment avoidance, where mathematically higher values represent a greater defensive deactivation of the interpersonal attachment system and a profound psychological discomfort with interpersonal closeness. To empirically evaluate the Motivational Interviewing processes, the first two clinical sessions for each participant were audio-recorded, professionally transcribed into text, and subsequently subjected to rigorous behavioral coding using the Motivational Interviewing Skill Code system. Highly trained independent raters, completely blinded to the ultimate clinical outcomes and the baseline attachment scores of the patients, utilized this established system to code both therapist behaviors and client utterances. The raters tracked therapist metrics such as the frequency of complex reflections and the demonstration of empathy, while meticulously categorizing patient speech into distinct constructs including change talk, sustain talk, and neutral statements. The qualitative linguistic data and the sequential therapeutic interactions were thereby transformed into quantifiable frequencies and global therapeutic ratings suitable for algorithmic processing. Furthermore, baseline sociodemographic characteristics and clinical symptom severity profiles were gathered using standard clinic intake assessments to serve as foundational control variables and additional demographic features within the predictive modeling framework.

2.3. Data analysis

The predictive modeling was executed utilizing a comprehensive machine learning pipeline specifically designed to handle multidimensional psychological, demographic, and linguistic data. Prior to model training, rigorous data preprocessing was conducted, wherein missing psychometric values were computationally imputed using a k -nearest neighbors imputation algorithm. All continuous features, including the attachment avoidance scores and the Motivational Interviewing behavioral counts, were standardized to a statistical mean of $\mu = 0$ and a standard

deviation of $\sigma = 1$ to ensure algorithmic scale invariance. Given the inherent class imbalance typical in psychotherapeutic dropout research, where the minority class of dropouts is frequently eclipsed by the majority class of treatment completers, the Synthetic Minority Over-sampling Technique was applied exclusively to the training data partitions. This technique synthetically generated representative numerical instances of the dropout class in the feature space, thereby preventing the machine learning algorithms from developing a predictive bias toward the majority class. Three distinct supervised machine learning classifiers were deployed and evaluated: Random Forest, Support Vector Machines, and Extreme Gradient Boosting. The predictive performance and generalizability of these mathematical models were rigorously validated using a stratified 10-fold cross-validation methodology to guarantee that each distinct training and testing fold contained a proportional representation of the actual dropout instances. Model efficacy was evaluated utilizing a multidimensional statistical metric approach, calculating the Area Under the Receiver Operating Characteristic Curve, mathematical precision, recall sensitivity, and the F_1 -score to provide a holistic view of predictive accuracy across both the positive and negative classes. Finally, to ensure the clinical interpretability of the predictive models and to explicitly identify the most robust individual predictors of treatment dropout, SHapley Additive exPlanations values were computed. This advanced feature attribution technique allowed for the precise mathematical quantification of the directional impact of individual variables, such as specific elevations in attachment-related avoidance and the exact frequency of patient sustain talk during early Motivational Interviewing processes, on the overall algorithmic probability of a patient prematurely terminating their treatment.

3. Findings and Results

The comprehensive analysis of the $N = 485$ participants revealed significant insights into the behavioral and psychological predictors of early treatment termination. Out of the initial cohort, exactly $n = 157$ patients (32.4%) met the operational definition for treatment dropout, failing to complete the therapeutic protocol. The remaining $n = 328$ patients (67.6%) constituted the treatment-completer group. Baseline demographic and clinical characteristics were examined to contextualize the sample and identify potential confounding variables. The mean age of the overall

sample was $M = 34.5$ years ($SD = 10.2$), with a predominantly female composition (61.2%). Preliminary independent samples t -tests and Chi-square analyses indicated that there were no statistically significant differences between the dropout and completer groups regarding age, biological sex, or baseline depressive

symptom severity, suggesting that these standard demographic factors did not primarily drive the premature termination of therapy in this specific Mexican sample. A detailed breakdown of these baseline characteristics is presented in Table 1.

Table 1

Baseline Demographic and Clinical Characteristics of the Sample by Dropout Status

Characteristic	Total Sample ($N = 485$)	Completers ($n = 328$)	Dropouts ($n = 157$)	p -value
Age in years, $M(SD)$	34.5(10.2)	34.8(10.5)	33.9(9.6)	.374
Female gender, $n(\%)$	297(61.2%)	205(62.5%)	92(58.6%)	.412
Years of education, $M(SD)$	12.4(3.1)	12.6(3.0)	12.0(3.3)	.051
Baseline Depression, $M(SD)$	18.3(5.4)	18.1(5.2)	18.7(5.8)	.268
Baseline Anxiety, $M(SD)$	16.7(4.9)	16.5(4.8)	17.1(5.1)	.213

Bivariate analyses were subsequently conducted on the primary variables of interest: attachment-related avoidance and the coded Motivational Interviewing (MI) processes derived from the initial two therapy sessions. Patients who subsequently dropped out of treatment exhibited significantly higher baseline levels of attachment-related avoidance ($M = 4.25$, $SD = 1.10$) compared to those who completed treatment ($M = 3.15$, $SD = 0.95$), yielding a statistically significant difference of $t(483) = 11.24$, $p < .001$. Furthermore, the behavioral coding of the therapeutic dialogue revealed distinct interactional patterns. The dropout

cohort produced a significantly higher frequency of Sustain Talk ($M = 24.6$, $SD = 8.4$) per session compared to completers ($M = 15.2$, $SD = 6.1$). Conversely, treatment completers expressed significantly more Change Talk ($M = 31.4$, $SD = 9.2$) than those who dropped out ($M = 19.8$, $SD = 7.5$). Therapist behaviors also showed variance; lower observed therapist empathy ratings and fewer complex reflections were moderately associated with the dropout group, highlighting the dyadic nature of treatment retention. These comprehensive bivariate comparisons are summarized in Table 2.

Table 2

Bivariate Analysis of MI Process Variables and Attachment Avoidance by Dropout Status

Variable	Completers ($n = 328$)	Dropouts ($n = 157$)	t -statistic	p -value
Attachment Avoidance	3.15(0.95)	4.25(1.10)	11.24	<.001
Patient Change Talk	31.4(9.2)	19.8(7.5)	13.75	<.001
Patient Sustain Talk	15.2(6.1)	24.6(8.4)	-13.88	<.001
Therapist Complex Reflections	14.3(4.5)	11.2(5.1)	6.82	<.001
Global Therapist Empathy	4.1(0.7)	3.4(0.9)	9.35	<.001

To rigorously predict treatment dropout utilizing these multidimensional data sets, three supervised machine learning classifiers were evaluated: Random Forest, Support Vector Machines (SVM), and Extreme Gradient Boosting (XGBoost). Following the application of the Synthetic Minority Over-sampling Technique (SMOTE) to the training data to rectify class imbalances, the models were validated using a stratified 10-fold cross-validation approach. The XGBoost algorithm demonstrated the superior predictive capacity across all evaluated metrics. Specifically, the XGBoost model achieved an Area Under the Receiver Operating Characteristic Curve (AUC) of 0.87,

indicating excellent discriminative ability between dropouts and completers. It also generated a precision score of 0.79 and a recall of 0.82 for the minority class (dropouts), culminating in an F_1 -score of 0.80. The Random Forest model performed marginally lower but still exhibited robust predictive power ($AUC = 0.84$), whereas the SVM algorithm demonstrated the lowest comparative performance ($AUC = 0.78$), struggling slightly with the non-linear interactions among the linguistic and psychological variables. The full suite of performance metrics for all three machine learning models is detailed in Table 3.

Table 3

Predictive Performance Metrics of Machine Learning Models for Treatment Dropout

Machine Learning Model	AUC	Precision	Recall	F ₁ -score	Accuracy
Extreme Gradient Boosting (XGBoost)	0.87	0.79	0.82	0.80	0.85
Random Forest	0.84	0.75	0.78	0.76	0.82
Support Vector Machine (SVM)	0.78	0.68	0.71	0.69	0.76

Following the identification of XGBoost as the optimal predictive model, SHapley Additive exPlanations (SHAP) values were extracted to interpret the algorithmic decision-making process and quantify individual feature importance. The SHAP analysis definitively established that high baseline Attachment Avoidance was the single most influential predictor of treatment dropout, followed closely by a high frequency of patient Sustain Talk during the initial sessions. Furthermore, the SHAP dependence plots revealed critical non-linear interaction effects between patient pathology and therapist technique. Specifically, the algorithms identified that patients with highly avoidant attachment styles who were met with low frequencies of therapist complex reflections during the first session possessed a compound, exponentially higher algorithmic probability of terminating therapy prematurely. Conversely, high rates of patient Change Talk and elevated global ratings of Therapist Empathy served as the strongest protective features, significantly lowering the calculated probability of dropout even among individuals with moderate levels of attachment avoidance.

4. Discussion

The primary objective of this investigation was to employ advanced machine learning algorithms to determine whether baseline attachment-related avoidance and specific linguistic frequencies observed during early Motivational Interviewing sessions could reliably predict outpatient psychotherapeutic treatment dropout. The findings derived from the multidimensional analysis of *N* = 485 treatment-seeking adults unequivocally confirm that treatment attrition is heavily influenced by the complex, non-linear interplay between patient interpersonal pathology and therapeutic micro-processes. The overall dropout rate observed in this sample, calculated at 32.4%, is remarkably consistent with the broader psychiatric literature, which repeatedly highlights the systemic vulnerability of mental healthcare delivery to premature therapeutic termination across diverse clinical settings and treatment modalities (Bentley et al.,

2021; Gersh et al., 2017; Leeuwerik et al., 2019). The algorithmic modeling demonstrated that the Extreme Gradient Boosting (XGBoost) classifier possessed superior predictive capacity, achieving an excellent Area Under the Receiver Operating Characteristic Curve of *AUC* = 0.87. The SHapley Additive exPlanations analysis subsequently revealed that elevated baseline attachment-related avoidance was the single most dominant individual predictor of treatment dropout, functioning as a massive barrier to the successful establishment of a viable therapeutic alliance. This aligns profoundly with recent empirical models indicating that severe attachment avoidance intrinsically generates a pervasive epistemic mistrust and profound hypomentalization, fundamentally impairing the individual’s ability to rely on healthcare professionals or tolerate emotional vulnerability (Kurt, 2025).

The algorithmic identification of attachment avoidance as the primary driver of therapeutic elopement is theoretically supported by the foundational principles of adult attachment theory, which posits that highly avoidant individuals compulsively deactivate their emotional attachment systems when confronted with interpersonal stress. In the context of psychological treatment, the inherent demand for emotional disclosure operates as an acute psychological threat. Consequently, just as maladaptive attachment styles severely disrupt emotional regulation in obsessive-compulsive models (Nielsen et al., 2025) and significantly diminish cognitive flexibility regarding deep interpersonal commitments like marriage (Rezapour et al., 2025), this same avoidant pathology dictates a reflexive withdrawal from the clinical dyad. The intense discomfort with intimacy characteristic of these patients mimics the relational instability heavily documented in specialized interventions for personality pathology. For instance, the high risk of premature termination observed in this study mirrors the severe attrition challenges frequently documented in the prolonged treatment of Borderline Personality Disorder across intensive dialectical behavior therapy (Kliem et al., 2010; Kröger et al., 2013; Lang et al., 2018; Linehan et al., 2015; Stratton et al., 2020) and highly structured schema

therapy protocols (Dickhaut & Arntz, 2014; Giesen-Bloo et al., 2006; Schaap et al., 2016). Furthermore, as highly avoidant pairings frequently escalate into severe emotional distress and even intimate partner violence within romantic relationships (Mohammadi et al., 2025; Pudenko et al., 2025), the therapeutic relationship similarly fractures when the therapist attempts to foster a closeness that the patient fundamentally perceives as dangerous, leading to unilateral and abrupt dropout.

Beyond the static trait of attachment avoidance, the dynamic linguistic processes inherent to Motivational Interviewing provided critical, real-time predictive value. The machine learning models identified that a high frequency of patient Sustain Talk during the initial two clinical sessions was the second most powerful predictor of subsequent dropout, whereas high frequencies of Change Talk served as a robust protective factor. This finding provides profound empirical validation for the core mechanisms of Motivational Interviewing (Miller, 2023; Miller & Rollnick, 2023). When patients actively verbalize arguments against change and reinforce the pathological status quo (Sustain Talk), they are not merely expressing ambivalence; they are actively fortifying their behavioral resistance and signaling impending clinical disengagement. Conversely, the protective nature of Change Talk observed in this study is conceptually aligned with the successful application of Motivational Interviewing across highly diverse and resistant populations, such as its efficacy in reducing physical aggression in adolescents (Mansouri & Khodabakhshi-Koolaei, 2024), mitigating severe social anxiety (Parsafar, 2024), and improving emotional participation in learning-disabled adults (Nejatifar & Abedi, 2023). Even in uniquely challenging contexts, such as engaging young adults with Autism Spectrum Disorder (Pagan, 2024) or managing the chronic distress of multiple sclerosis alongside Acceptance and Commitment Therapy (Rezaei et al., 2023), the elicitation of intrinsic motivation is paramount. The current findings suggest that when this motivational elicitation fails, and Sustain Talk dominates the clinical hour, the statistical probability of the patient abandoning therapy increases exponentially.

Crucially, the advanced algorithmic modeling exposed critical non-linear interactions between the patient's pathological presentation and the therapist's technical execution. The predictive models demonstrated that patients harboring high levels of attachment avoidance were extraordinarily sensitive to deficits in therapist technique, specifically a low frequency of complex reflections and

lower global ratings of empathy. When a highly avoidant patient was paired with a therapist who failed to adequately utilize these advanced, validating communication strategies, the probability of dropout compounded dramatically. This complex, dyadic interaction effect highlights the immense responsibility placed upon the clinician to adapt their methodology to the patient's specific relational deficits (Banham & Schweitzer, 2016). In treatments that naturally provoke high distress, such as addressing infertility trauma (Peterson & Eifert, 2011), treating severe substance abuse under court mandate (Daughters et al., 2024), or managing complex co-morbidities like binge-eating disorder (Lammers et al., 2022), the therapist's ability to maintain an unbroken empathetic connection is the primary safeguard against therapeutic elopement. If therapeutic frameworks lack this robust relational scaffolding, they become structurally vulnerable to unacceptable rates of treatment refusal and attrition, a phenomenon heavily scrutinized in comparative efficacy trials (Ahmadi & Raeisi, 2018; Cortes et al., 2016; Rotaru, 2019; Simmonds-Buckley et al., 2022; Strauss et al., 2023). Ultimately, the findings strongly assert that treatment retention is not simply a function of patient compliance, but rather the measurable outcome of a highly sensitive, bi-directional therapeutic interaction where advanced clinical skills must proactively neutralize deep-seated interpersonal avoidance.

5. Conclusion

The application of advanced machine learning algorithms in this study provides compelling empirical evidence that outpatient psychotherapeutic treatment dropout can be accurately predicted from the earliest stages of clinical engagement. By integrating static psychological traits with dynamic, real-time communication metrics, the predictive models successfully mapped the complex etiology of premature therapeutic termination. Specifically, the findings established that severe baseline attachment-related avoidance acts as the primary systemic barrier to treatment retention, profoundly compounding a patient's natural resistance to therapeutic vulnerability. Furthermore, the precise quantification of Motivational Interviewing processes revealed that the dominance of patient sustain talk, heavily exacerbated by a distinct lack of advanced therapist complex reflections, serves as the observable behavioral mechanism through which this underlying avoidance clinically manifests into overt dropout. Consequently, this study definitively bridges the critical gap between abstract

attachment theory and observable clinical micro-processes, demonstrating that the therapeutic alliance is a highly quantifiable and statistically predictable phenomenon.

6. Limitations & Suggestions

Despite the robust predictive accuracy of the utilized machine learning models, several methodological limitations must be acknowledged to appropriately contextualize these findings. First, the clinical sample was drawn entirely from community mental health centers within a specific metropolitan area in Mexico, which inherently restricts the broad socio-cultural generalizability of the findings, as expressions of attachment avoidance and therapeutic communication styles may vary significantly across diverse global populations. Second, the reliance on the manual transcription and subsequent human behavioral coding of the initial therapy sessions using the Motivational Interviewing Skill Code system, while methodologically rigorous, is highly resource-intensive and practically precludes the real-time, instantaneous clinical application of these predictive models in fast-paced, underfunded community settings. Finally, while the algorithms effectively captured linguistic patterns and self-reported attachment metrics, the models did not incorporate potential macroscopic, systemic confounders, such as sudden shifts in the patient's socio-economic stability, acute transportation barriers, or sudden changes in health insurance coverage, all of which are known external precipitants of treatment dropout that operate entirely independently of the internal therapeutic dyad.

Future empirical investigations should prioritize the direct integration of automated natural language processing and advanced speech recognition technologies to bypass the severe time constraints associated with manual therapeutic coding. By training deep learning models to automatically analyze both the semantic content and the acoustic prosody of clinical sessions in real-time, researchers could develop instantaneous predictive dashboards for working clinicians, fundamentally transforming retrospective data analysis into proactive clinical intervention. Additionally, longitudinal studies must be conducted across vastly different cultural and linguistic contexts to rigorously determine whether the profound predictive dominance of attachment avoidance and sustain talk remains algorithmically stable, or if different socio-cultural paradigms generate entirely distinct mathematical pathways to treatment dropout. Finally, large-scale, randomized controlled trials are urgently needed to

empirically test whether providing therapists with algorithmically generated, session-by-session feedback regarding their patient's specific dropout risk mathematically improves actual clinical retention rates compared to standard, unguided psychiatric care.

Clinical administrators and frontline mental health practitioners should heavily prioritize the systematic, formal assessment of adult attachment styles immediately during the intake process, recognizing that highly avoidant scores represent a critical, latent emergency regarding treatment retention. When a patient is identified as possessing severe attachment-related avoidance, the clinical timeline must be fundamentally adjusted, prioritizing the meticulous, slow development of absolute epistemic trust over the rapid implementation of demanding behavioral change protocols. Furthermore, clinical supervisors must ensure that ongoing therapist training specifically targets the mastery of advanced Motivational Interviewing techniques, focusing heavily on the continuous generation of complex reflections and the intentional suppression of confrontational stances when managing high levels of patient sustain talk. Ultimately, therapists must recognize that early patient resistance is not a clinical failure, but rather a direct manifestation of the patient's core interpersonal pathology, requiring the strategic, compassionate application of advanced communication skills to successfully anchor the patient in the lifesaving process of therapy.

Acknowledgments

We would like to express our appreciation and gratitude to all those who cooperated in carrying out this study.

Declaration of Interest

The authors of this article declared no conflict of interest.

Ethical Considerations

The study protocol adhered to the principles outlined in the Helsinki Declaration, which provides guidelines for ethical research involving human participants.

Transparency of Data

In accordance with the principles of transparency and open research, we declare that all data and materials used in this study are available upon request.

Funding

This research was carried out independently with personal funding and without the financial support of any governmental or private institution or organization.

Authors' Contributions

All authors equally contributed in this article.

Declaration

In order to correct and improve the academic writing of our paper, we have used the language model ChatGPT.

References

- Ahmadi, A., & Raeisi, Z. (2018). The Effect of Acceptance and Commitment Therapy on Distress Tolerance in Mothers of Children with Autism [Research]. *Quarterly Journal of Child Mental Health*, 5(3), 69-79. <http://childmentalhealth.ir/article-1-349-en.html>
- Banham, J. A., & Schweitzer, R. D. (2016). Trainee-therapists are not all equal: Examination of therapeutic efficiency, effectiveness and early client dropout after 12 months of clinical training. *Psychology and Psychotherapy*, 89(2), 148-162. <https://doi.org/10.1111/papt.12071>
- Bentley, K. H., Cohen, Z. D., Kim, T., Bullis, J. R., Nauphal, M., Cassiello-Robbins, C., Sauer-Zavala, S., Sbi, S., Gallagher, M. W., Farchione, T. J., & Barlow, D. H. (2021). The Nature, Timing, and Symptom Trajectories of Dropout From Transdiagnostic and Single-Diagnosis Cognitive-Behavioral Therapy for Anxiety Disorders. *Behavior therapy*, 52(6), 1364-1376. <https://doi.org/10.1016/j.beth.2021.03.007>
- Cortes, B., Ballesteros, A., Collantes, J., & Aguilar, M. L. (2016). What makes for good outcomes in solution-focused brief therapy? A follow-up study. *European Psychiatry*, 33(S1), s232-s233. <https://doi.org/10.1016/j.eurpsy.2016.01.580>
- Daughters, S. B., Stipelman, B. A., Sargeant, M. N., Schuster, R., Bornoalova, M. A., & Lejuez, C. W. (2024). The interactive effects of antisocial personality disorder and court-mandated status on substance abuse treatment dropout. *Journal of Substance Abuse Treatment*, 34(2), 157-164. <https://doi.org/10.1016/j.jsat.2007.02.007>
- Dickhaut, V., & Arntz, A. (2014). Combined group and individual schema therapy for borderline personality disorder: A pilot study. *Journal of Behavior Therapy and Experimental Psychiatry*, 45(2), 242-251. <https://doi.org/10.1016/j.jbtep.2013.11.004>
- Gersh, E., Hallford, D. J., Rice, S. M., Kazantzis, N., Gersh, H., Gersh, B., & McCarty, C. A. (2017). Systematic review and meta-analysis of dropout rates in individual psychotherapy for generalized anxiety disorder. *Journal of anxiety disorders*, 52, 25-33. <https://doi.org/10.1016/j.janxdis.2017.10.001>
- Giesen-Bloo, J., van Dyck, R., Spinhoven, P., van Tilburg, W., Dirksen, C., van Asselt, T., Kremers, I., Nadort, M., & Arntz, A. (2006). Outpatient Psychotherapy for Borderline Personality Disorder: Randomized Trial of Schema-Focused Therapy vs Transference-Focused Psychotherapy. *Archives of General Psychiatry*, 63(6), 649-658. <https://doi.org/10.1001/archpsyc.63.6.649>
- Kliem, S., Kröger, C., & Kosfelder, J. (2010). Dialectical behavior therapy for borderline personality disorder: A meta-analysis using mixed-effects modeling. *Journal of consulting and clinical psychology*, 78(6), 936-951. <https://doi.org/10.1037/a0021015>
- Kröger, C., Harbeck, S., Armbrust, M., & Kliem, S. (2013). Effectiveness, response, and dropout of dialectical behavior therapy for borderline personality disorder in an inpatient setting. *Behaviour Research and Therapy*, 51(8), 411-416. <https://www.sciencedirect.com/science/article/pii/S0005796713000752>
- Kurt, Y. (2025). Attachment and Borderline Personality Features: The Mediating Roles of Hypomentalizing and Epistemic Mistrust. *Clinical Psychology & Psychotherapy*, 32(6). <https://doi.org/10.1002/cpp.70185>
- Lammers, M. W., Vroling, M. S., Crosby, R. D., & van Strien, T. (2022). Dialectical behavior therapy compared to cognitive behavior therapy in binge-eating disorder: An effectiveness study with 6-month follow-up. *International Journal of Eating Disorders*, 55(7), 902-913. <https://doi.org/10.1002/eat.23750>
- Lang, C. M., Edwards, A. J., Mittler, M. A., & Bonavitacola, L. (2018). Dialectical Behavior Therapy With Prolonged Exposure for Adolescents: Rationale and Review of the Research. *Cognitive and Behavioral Practice*, 25(3), 416-426. <https://doi.org/10.1016/j.cbpra.2017.12.005>
- Leeuwerik, T., Cavanagh, K., & Strauss, C. (2019). Patient adherence to cognitive behavioural therapy for obsessive-compulsive disorder: A systematic review and meta-analysis. *Journal of anxiety disorders*, 68, 102135. <https://doi.org/10.1016/j.janxdis.2019.102135>
- Linehan, M. M., Korslund, K. E., Harned, M. S., Gallop, R. J., Lungu, A., Neacsiu, A. D., McDavid, J., Comtois, K. A., & Murray-Gregory, A. M. (2015). Dialectical Behavior Therapy for High Suicide Risk in Individuals With Borderline Personality Disorder: A Randomized Clinical Trial and Component Analysis. *JAMA Psychiatry*, 72(5), 475-482. <https://doi.org/10.1001/jamapsychiatry.2014.3039>
- Mansouri, K., & Khodabakhshi-Koolae, A. (2024). The Effectiveness of Motivational Interview (MI) on Aggression and High-risk Behaviors of Adolescent Girls. *JPEN*, 10(4), 41-49. <https://civilica.com/doc/2029606/>
- Miller, W. R. (2023). The Evolution of Motivational Interviewing. *Behavioural and Cognitive Psychotherapy*, 51(6), 616-632. <https://doi.org/10.1017/s1352465822000431>
- Miller, W. R., & Rollnick, S. (2023). *Motivational interviewing: Helping people change*. Guilford Press. <https://psycnet.apa.org/record/2012-17300-000>
- Mohammadi, M., Shirafkan-Kopekan, A., Shafiabadi, A., & Farhangi, A. (2025). The effectiveness of schema therapy on attachment styles and marital intimacy in couples in Tehran. *Women Fam Stud Q*, 17(64), 173-195. <https://srb.sanad.iau.ir/en/Article/954007>
- Nejatifar, S., & Abedi, A. (2023). Effectiveness of Motivational Interviewing on Participation and Emotional Skills in Learning Disabled Adults. *Psychological Research in Individuals with Exceptional Needs*, 1(1), 32-40. <https://doi.org/10.61838/kman.prien.1.1.5>
- Nielsen, S. K. K., Stuart, A. C., Winding, C., Pedersen, M. Ø., Franksdatter Daniel, S. I., Vangkilde, S., Rosenberg, N., & Hageman, I. (2025). Adult attachment style, emotion regulation and obsessive-compulsive disorder-A preliminary cross-sectional mediational investigation of an attachment-based model. *Clinical Psychology & Psychotherapy*, 32(1), e70031. <https://doi.org/10.1002/cpp.70031>

- Pagan, A. F. (2024). Motivational Interviewing for Young Adults With Autism Spectrum Disorder: A Pilot Feasibility Study. <https://doi.org/10.31234/osf.io/mt247>
- Parsafar, A. (2024). Investigating the Effectiveness of Motivational Interviewing on Social Anxiety and Academic Procrastination Among Students. *Jayps*, 5(3), 55-64. <https://doi.org/10.61838/kman.jayps.5.3.6>
- Peterson, B. D., & Eifert, G. H. (2011). Using Acceptance and Commitment Therapy to Treat Infertility Stress. *Cognitive and Behavioral Practice*, 18(4), 577-587. <https://doi.org/10.1016/j.cbpra.2010.03.004>
- Pudelko, A. H., Ramos, B., Emond, M., Péloquin, K., & Daspe, M. É. (2025). Is Our Attachment Hurting Us? Unraveling the Associations Between Partners' Attachment Pairings, Negative Emotions During Conflict, and Intimate Partner Violence. *Journal of marital and family therapy*, 51(2). <https://doi.org/10.1111/jmft.70009>
- Rezaei, Z., Abolghasemi, S., Khalatbari, J., & Zarbakhsh, R. (2023). The effectiveness of a training package based on motivational interviewing, therapy based on acceptance and commitment, and therapy focused on compassion on tolerance of failure and health anxiety in patients with multiple sclerosis. *Journal of Adolescent and Youth Psychological Studies (JAYPS)*, 4(2), 86-96. <https://doi.org/10.61838/kman.jayps.4.2.10>
- Rezapour, Y., Rabaniayn, H., & Aminjafari, F. (2025). The Mediating Role of Cognitive Emotion Regulation in the Relationship between Attachment Styles and Attitudes Towards Marriage. *Journal of Culture of Counseling and Psychotherapy*, 16(61), 33-66. https://qccpc.atu.ac.ir/article_18877_en.html
- Rotaru, I. C. (2019). The school dropout of Roma children. Between anti-gypsism and the socio - economic dysfunction of Romanian educational system. *International Conference "Supervision in Psychotherapy"*. https://www.researchgate.net/publication/335293496_The_school_dropout_of_Roma_children_Between_anti-gypsism_and_the_socio-economic_dysfunction_of_Romanian_educational_system
- Schaap, G. M., Chakhssi, F., & Westerhof, G. J. (2016). Inpatient schema therapy for nonresponsive patients with personality pathology: Changes in symptomatic distress, schemas, schema modes, coping styles, experienced parenting styles, and mental well-being. *Psychotherapy*, 53(4), 402-412. <https://doi.org/10.1037/pst0000056>
- Simmonds-Buckley, M., Osivwemu, E.-O., Kellett, S., & Taylor, C. (2022). The acceptability of cognitive analytic therapy (CAT): Meta-analysis and benchmarking of treatment refusal and treatment dropout rates. *Clinical psychology review*, 96(1), 102187-102197. <https://doi.org/10.1016/j.cpr.2022.102187>
- Stratton, N., Mendoza Alvarez, M., Labrish, C., Barnhart, R., & McMain, S. (2020). Predictors of dropout from a 20-week dialectical behavior therapy skills group for suicidal behaviors and borderline personality disorder. *Journal of personality disorders*, 34(2), 216-230. https://doi.org/10.1521/pedi_2018_32_391
- Strauss, C., Bibby-Jones, A.-M., Jones, F., Byford, S., Heslin, M., Parry, G., Barkham, M., Lea, L., Crane, R., de Visser, R., Arbon, A., Rosten, C., & Cavanagh, K. (2023). Clinical Effectiveness and Cost-Effectiveness of Supported Mindfulness-Based Cognitive Therapy Self-help Compared With Supported Cognitive Behavioral Therapy Self-help for Adults Experiencing Depression: The Low-Intensity Guided Help Through Mindfulness (LIGHTMind) Randomized Clinical Trial. *JAMA Psychiatry*, 80(5), 415-424. <https://doi.org/10.1001/jamapsychiatry.2023.0222>