

Early-Stage Consumer Credit Risk Detection Using Multimodal Deep Learning and Behavioral AI

Gopinathan Rathinavelu

Data Architect, Arizona, USA, gopinathan.r@gmail.com

Abstract: Early identification of consumer credit risk is increasingly important in digital lending environments where conventional credit scores may not adequately capture rapidly changing borrower conditions and behavioral signals. This paper proposes a multimodal deep learning and Behavioral Artificial Intelligence framework for detecting early-stage consumer credit deterioration by integrating structured financial information, transactional sequences, repayment behavior, digital interaction patterns, and unstructured textual or contextual signals. The framework employs modality-specific feature extraction followed by attention-based multimodal fusion to identify nonlinear relationships and temporal changes associated with emerging repayment risk. Behavioral representations are continuously updated to distinguish persistent risk from temporary fluctuations, while explainability mechanisms support interpretation of risk-driving factors. The proposed approach is designed to complement conventional credit scoring by providing dynamic early-warning indicators rather than relying exclusively on historical credit records. The framework addresses challenges involving heterogeneous data, class imbalance, temporal drift, privacy, interpretability, and responsible use of alternative data. The study establishes a research structure for evaluating predictive performance, behavioral sensitivity, multimodal contribution, and operational applicability in consumer lending.

Keywords: Consumer Credit Risk, Multimodal Deep Learning, Behavioral AI, Credit Scoring, Early-Warning Detection, Alternative Data

1. Introduction

Consumer credit risk assessment has traditionally relied on structured financial variables such as credit history, repayment records, income, debt obligations, loan characteristics, and demographic attributes. Conventional statistical approaches, particularly logistic regression and scorecard-based systems, remain important because of their interpretability, established validation procedures, and compatibility with financial risk-management practices. However, the increasing digitization of consumer finance has substantially changed the nature, volume, and velocity of information available for assessing borrower risk. Digital lending platforms now generate transactional sequences, repayment trajectories, account activity, application behavior, device interactions, and other contextual signals that may contain information about changes in a borrower's financial condition before conventional delinquency indicators become visible [1], [2]. Recent reviews indicate that machine-learning-based credit assessment has expanded considerably because of its ability to capture nonlinear relationships and complex interactions that are difficult to represent through traditional statistical models.

The principal challenge in consumer credit risk management is not merely identifying borrowers who will eventually default, but detecting **early-stage deterioration** sufficiently before serious delinquency occurs. A conventional credit score often represents an aggregate assessment based largely on historical information, whereas emerging credit risk may manifest through changes in repayment timing, transaction frequency, utilization behavior, cash-flow irregularities, application patterns, or other sequential behavioral changes. Consequently, a borrower may maintain an apparently acceptable credit profile while simultaneously exhibiting short-term behavioral characteristics



associated with increasing financial stress. This creates a methodological requirement for models capable of learning temporal dependencies and detecting deviations from an individual's established behavioral baseline. Research on machine learning and deep learning has demonstrated the potential of nonlinear models to improve borrower-level risk prediction, although differences in datasets, validation strategies, feature construction, and performance measures make direct comparison across studies difficult [2], [3].

The expansion of alternative data has further transformed the credit-risk modelling environment. Conventional credit databases may provide insufficient information about individuals with limited or thin credit histories, particularly younger consumers, newly banked customers, and borrowers operating within rapidly expanding digital-finance ecosystems. Alternative information can include transactional activity, mobile-related signals, digital footprints, psychometric indicators, and other behavioral attributes. Such information can provide additional evidence concerning financial behavior and repayment capacity, but its use simultaneously introduces concerns relating to privacy, data quality, representativeness, proxy discrimination, and regulatory acceptability [4], [5]. Recent literature emphasizes that alternative data can improve the information available for credit assessment, but its predictive usefulness is context-dependent and requires careful governance rather than indiscriminate incorporation into automated decisions.

Deep learning provides an appropriate methodological foundation for addressing the heterogeneous and temporal characteristics of modern consumer-credit data. Multilayer neural networks can approximate complex nonlinear relationships, recurrent architectures can represent sequential behavioral information, convolutional architectures can extract localized patterns, and attention-based architectures can identify comparatively important elements within long sequences or heterogeneous feature sets [6], [7]. The development of multimodal learning extends these capabilities by allowing information originating from different data modalities to be processed separately and subsequently integrated into a unified representation. In a consumer credit setting, structured financial variables, transaction sequences, behavioral indicators, textual information, and contextual signals may possess different statistical characteristics and therefore require modality-specific representation before fusion [8]. Recent research on deep learning for credit risk increasingly considers tabular, sequential, textual, and relational information as complementary sources rather than treating the credit dataset as a single homogeneous feature matrix.

Behavioral Artificial Intelligence provides an additional conceptual dimension to this problem by emphasizing the evolution of borrower behavior rather than relying exclusively on static attributes. Behavioral modelling can identify deviations from an individual's historical patterns, changes in transaction regularity, altered repayment behavior, increasing credit utilization, or other temporal signals that may precede observable deterioration in conventional credit indicators. Instead of assigning a risk level solely from the current state of a borrower, a behavioral framework can estimate risk from the **trajectory of change**. This distinction is particularly relevant for early-warning systems, where the objective is to recognize emerging risk before it develops into severe delinquency. Existing research has begun to explore behavioral and alternative information for credit assessment, but substantial methodological opportunities remain in integrating behavioral sequences with conventional financial variables within a single multimodal architecture [4], [9].

Another important consideration is the imbalance between predictive performance and practical interpretability. Deep learning models may achieve strong predictive results but can become difficult to interpret because risk decisions emerge from high-dimensional nonlinear transformations. In credit applications, this issue is particularly important because lenders need to understand the factors contributing to a risk classification, while consumers and regulators may require meaningful explanations of automated decisions. Explainable Artificial Intelligence techniques, including feature-attribution and local explanation approaches, have therefore become increasingly relevant to credit-risk modelling [10], [11]. Recent literature nevertheless identifies instability of explanations, insufficient standardization, and limited evidence concerning the practical usefulness of explanations as continuing challenges.

The problem is further complicated by **class imbalance, concept drift, temporal instability, and data heterogeneity**. Serious delinquency or default events generally constitute a minority class compared with normal

repayment observations, potentially causing conventional learning algorithms to favor the majority class. At the same time, consumer behavior can change because of economic conditions, changes in lending products, seasonal effects, technological adoption, or changes in individual financial circumstances. Consequently, a model trained on historical observations may gradually lose predictive validity when the underlying data-generating process changes. Recent reviews identify imbalanced datasets, insufficient temporal validation, model drift, and limited deployment-oriented evaluation as persistent issues in advanced credit-scoring research [3], [12].

Against this background, the present study proposes an **Early-Stage Consumer Credit Risk Detection framework using Multimodal Deep Learning and Behavioral AI**. The central proposition is that early credit deterioration can be identified more effectively by jointly modelling static financial characteristics, temporal transactional behavior, repayment trajectories, and other relevant contextual information rather than relying exclusively on a conventional static credit score. The framework is intended to learn complementary representations from heterogeneous modalities, integrate these representations through attention-based fusion, and generate dynamic early-warning risk indicators. Such a framework can complement—not necessarily replace—traditional credit assessment by providing an additional layer of behavioral intelligence for continuous risk monitoring.

The study therefore focuses on six principal objectives: (i) to examine the limitations of conventional consumer credit-risk assessment for early-stage deterioration; (ii) to develop modality-specific representations for heterogeneous financial and behavioral information; (iii) to construct a multimodal deep-learning architecture for borrower-level risk detection; (iv) to incorporate temporal behavioral modelling for identifying deviations from established borrower patterns; (v) to evaluate predictive performance using appropriate discrimination, calibration, imbalance-sensitive, and early-warning metrics; and (vi) to examine interpretability, robustness, fairness, privacy, and deployment considerations. The research consequently positions multimodal deep learning and Behavioral AI as complementary methodological components for developing adaptive and data-informed consumer credit-risk early-warning systems.

2. Literature Review

2.1 Conventional Consumer Credit-Risk Assessment

Consumer credit scoring has historically been dominated by statistical classification techniques, particularly logistic regression, discriminant analysis, and scorecard methodologies. These approaches estimate the probability of a borrower belonging to a predefined risk category using variables such as repayment history, income, outstanding obligations, credit utilization, and loan characteristics. Their continued importance is associated with comparatively straightforward interpretation, established statistical foundations, and operational familiarity within financial institutions. However, their underlying assumptions can restrict their ability to represent highly nonlinear interactions among borrower characteristics and rapidly evolving behavioral variables [1], [2].

Lessmann et al. demonstrated that credit-scoring research has progressively moved beyond traditional statistical approaches toward machine-learning classifiers capable of modelling more complex decision boundaries [13]. Subsequent research has similarly shown that ensemble methods and nonlinear machine-learning techniques can provide competitive or improved classification performance relative to conventional models, although performance differences depend substantially on the dataset and modelling procedure [2], [12]. The literature therefore suggests that the transition from statistical scoring to machine learning is not simply an algorithmic replacement; rather, it involves changes in feature engineering, data preprocessing, validation, class-imbalance management, model interpretation, and operational deployment.

A particularly important limitation of conventional scoring is its predominantly static representation of creditworthiness. A conventional score may summarize a borrower's historical credit experience but may not adequately capture the **direction, velocity, and persistence of behavioral change**. For early-stage risk detection, this distinction is important because financial deterioration can occur between periodic credit assessments. A borrower whose current credit profile appears stable may nevertheless demonstrate progressively adverse changes in payment timing, credit utilization, transaction behavior, or cash-flow regularity. Consequently, modern research increasingly

investigates models capable of incorporating temporal and behavioral information rather than treating observations as independent static records [6], [9].

2.2 Machine Learning and Deep Learning in Credit Risk Modelling

Machine learning has substantially expanded the methodological repertoire available for consumer credit-risk modelling. Decision trees, random forests, support vector machines, gradient-boosting algorithms, artificial neural networks, and ensemble methods can model nonlinear relationships and interactions that may be difficult to specify manually. A systematic review by Lessmann et al. established the importance of benchmarking multiple classification algorithms rather than assuming that one modelling family is universally superior [13]. Later reviews have reinforced the observation that ensemble and machine-learning approaches can improve predictive performance while also highlighting difficulties involving transparency, class imbalance, and standardized model development [12], [14].

Deep learning represents a further progression because neural architectures can learn hierarchical representations directly from complex input structures. Multilayer perceptrons can model nonlinear relationships in tabular credit information, while recurrent neural networks and their variants can learn dependencies within transaction and repayment sequences. Convolutional architectures can identify local structures within appropriately represented financial data, whereas transformer-based models can use attention mechanisms to capture relationships across longer sequences [6], [7]. These characteristics make deep learning particularly relevant to consumer credit environments where risk information may be distributed across multiple observations over time rather than concentrated in a single application record.

Hayashi's review of deep learning in credit scoring identifies the growing application of advanced neural architectures while also highlighting the continuing tension between predictive capability and interpretability [7]. More recent systematic research similarly emphasizes that deep learning has moved toward modality-aware modelling, including tabular, sequential, textual, and relational representations [6]. This development is directly relevant to the present research because early-stage credit deterioration is unlikely to be completely represented by one data structure. A multimodal architecture can instead treat different information sources according to their intrinsic characteristics and combine the resulting representations at a later stage.

2.3 Alternative Data and Behavioral Credit Assessment

The growing availability of alternative data has challenged the traditional boundary of consumer credit assessment. Financial institutions and digital lenders can potentially access information extending beyond conventional bureau variables, including transactional behavior, mobile-related activity, digital footprints, psychometric characteristics, and other behavioral indicators. Such information may be especially relevant for consumers with limited formal credit histories because conventional credit records may provide insufficient evidence for reliable risk differentiation [4], [5].

Alternative-data research, however, presents a dual methodological problem. First, the predictive relationship between a behavioral signal and credit risk must be demonstrated empirically rather than assumed. Second, the use of such information must be evaluated against privacy, fairness, consent, data-quality, and regulatory requirements. Recent systematic literature indicates that alternative data can contribute to improved credit assessment, particularly for thin-file borrowers, but also emphasizes that benefits depend on data quality, population characteristics, and the stability of behavioral relationships over time [3], [5]. (DOI)

Behavioral modelling is particularly valuable because it can transform raw observations into indicators of **change**. For example, the absolute level of credit utilization may be less informative than a persistent increase in utilization over successive periods. Similarly, a single delayed payment may not represent substantial deterioration, whereas a sequence of increasingly delayed payments may provide a stronger early-warning signal. This motivates the development of models that learn behavioral trajectories rather than relying only on cross-sectional features. Existing research on consumer credit behaviour and alternative signals provides a foundation for this direction, but

systematic integration of behavioral sequences with multiple financial modalities remains comparatively underdeveloped [4], [9].

2.4 Multimodal Deep Learning for Heterogeneous Credit Data

A major limitation of conventional credit-risk datasets is the tendency to organize all information into a common tabular representation. Such transformation may simplify modelling but can also discard structural information inherent in different modalities. Transaction records are sequential, textual information is linguistic, financial attributes are predominantly tabular, and relationships among accounts or entities may be represented as graphs. A multimodal approach attempts to preserve these differences by applying suitable representation-learning mechanisms before integrating the resulting features.

Multimodal learning has therefore become increasingly relevant to financial risk analysis. Modality-specific encoders can independently transform numerical, temporal, textual, or relational information into latent representations. These representations can subsequently be integrated through concatenation, gated fusion, cross-attention, or other attention-based mechanisms [6], [8]. Such architectures provide an opportunity to capture complementary evidence while reducing the need to force all information into a single feature representation.

For consumer credit risk, multimodal integration is especially relevant because different data sources can provide different stages of information about borrower deterioration. Traditional credit variables may describe historical financial capacity, transactional sequences may indicate current behavior, repayment histories may reveal emerging stress, and contextual or textual signals may provide additional explanatory information. The integration of these modalities can consequently transform credit assessment from a static classification problem into a more dynamic representation-learning problem.

Recent literature increasingly recognizes the potential of heterogeneous data and deep architectures in credit risk, but several limitations remain. Comparative studies often use different datasets and evaluation protocols, making reported improvements difficult to generalize. Moreover, multimodal models may become computationally demanding and may introduce additional missing-data and synchronization problems when different modalities are observed at different frequencies [6], [14]. These challenges indicate the need for carefully designed multimodal architectures with explicit temporal alignment, modality-quality assessment, and robustness testing.

2.5 Explainability, Fairness and Responsible AI

The adoption of AI in consumer credit decisions requires consideration beyond predictive accuracy. Credit decisions can directly affect access to financial resources, borrowing costs, and consumer opportunities. Consequently, an accurate model that cannot provide meaningful explanations or that systematically disadvantages particular groups may create substantial operational and governance concerns. Explainable AI has therefore become an important research component within modern credit-risk modelling [10], [11].

SHAP-, LIME-, feature-attribution-, and attention-based techniques have been investigated as mechanisms for explaining complex machine-learning predictions. These approaches can help identify variables or behavioral events associated with an individual risk assessment. However, explanation quality should not automatically be equated with causal interpretation. A variable identified as important by a predictive model may be correlated with risk without being a direct cause of repayment deterioration. Furthermore, explanation stability can vary across observations and modelling architectures [3], [10]. Recent reviews consequently identify interpretability and explanation reliability as continuing research challenges rather than solved problems.

Fairness represents another important dimension. Alternative and behavioral variables may contain indirect proxies for sensitive characteristics, potentially producing unintended disparities even when explicitly sensitive variables are excluded. The use of mobile, social, transactional, or digital-footprint information can therefore increase predictive information while simultaneously increasing the complexity of fairness assessment [3], [5]. A responsible

credit-risk framework should consequently evaluate discrimination, calibration, subgroup performance, data provenance, and the appropriateness of individual features alongside conventional predictive metrics.

2.6 Research Gap

The reviewed literature demonstrates substantial progress in machine-learning-based credit scoring, deep learning, alternative-data utilization, behavioral modelling, and explainable AI. Nevertheless, these developments remain fragmented across methodological streams. Conventional credit models generally emphasize structured financial information, while alternative-data studies frequently examine individual behavioral or digital signals. Deep-learning studies increasingly address sequential and heterogeneous information, but comprehensive multimodal integration specifically targeted at **early-stage consumer credit deterioration** remains comparatively limited [3], [6], [14].

A second gap concerns the distinction between **default prediction** and **early-stage risk detection**. Much of the existing literature evaluates whether a borrower ultimately defaults or belongs to a predefined risk class. This formulation does not necessarily address the more operationally relevant question of when meaningful deterioration first becomes detectable. An effective early-warning framework must therefore evaluate not only classification accuracy but also detection lead time, false-alert frequency, temporal stability, calibration, and the persistence of behavioral signals.

A third gap concerns behavioral dynamics. Existing models increasingly incorporate alternative and sequential data, but many approaches still represent borrower behavior through aggregated features. Such aggregation can obscure the sequence and velocity of change. A behavioral AI framework that explicitly models individual trajectories and detects deviations from historical patterns could provide a more sensitive mechanism for identifying emerging risk.

Finally, the literature identifies persistent challenges concerning class imbalance, concept drift, interpretability, fairness, privacy, and deployment [3], [5], [10], [12]. Recent systematic reviews specifically identify insufficient streaming-ready architectures, limited validation of behavioral datasets, weak standardization of explainability evaluation, and incomplete integration of governance requirements as important unresolved issues.

Accordingly, the present research addresses the identified gap through an integrated framework combining **multimodal deep learning with Behavioral AI**. The proposed research direction treats consumer credit risk as a dynamic, heterogeneous, and temporally evolving prediction problem. It seeks to combine conventional financial information with behavioral sequences and other relevant modalities, learn their complementary representations, detect deviations indicative of emerging deterioration, and provide interpretable early-warning outputs. This positioning extends existing credit-scoring research from predominantly static risk classification toward a more dynamic framework for early-stage consumer credit-risk detection.

3. Proposed Multimodal Behavioral AI Framework

3.1 Overall System Architecture

The proposed framework is designed as a dynamic early-warning architecture for identifying consumer credit deterioration before the borrower reaches a severe delinquency or default state. The architecture combines conventional financial information with sequential transactional behavior and other relevant behavioral or contextual signals. Recent research indicates that deep-learning-based credit-risk systems increasingly benefit from temporal, heterogeneous, and multimodal information, while multi-source fusion and attention mechanisms are being investigated for earlier risk identification [6], [15].

The framework consists of six sequential layers: **data acquisition, preprocessing and temporal alignment, modality-specific representation learning, behavioral intelligence extraction, multimodal fusion, and early-risk decision generation**. The architecture does not assume that every borrower will possess every modality. Instead,

modality availability is explicitly represented so that missing or irregular observations do not automatically translate into a high-risk classification.

The proposed system can be represented as:

Input Data → Data Quality Processing → Modality-Specific Encoders → Behavioral Representation → Attention-Based Fusion → Risk Prediction → Early-Warning Classification → Explainable Risk Output

The major data modalities considered are:

1. **Structured financial data:** income, debt obligations, loan amount, tenure, interest rate, credit utilization, previous defaults and credit history.

Transactional data: transaction frequency, inflow/outflow patterns, balance volatility, recurring payments, cash-flow trends and utilization trajectories.

Repayment behavior: payment punctuality, minimum-payment behavior, delinquency transitions and repayment consistency.

Digital behavioral data: application frequency, session characteristics, interaction patterns and changes in digital engagement where legally and ethically permissible.

Textual/contextual information: relevant application statements, financial descriptions or other permissible unstructured information.

Temporal metadata: time intervals, seasonality, recency, frequency and duration of observed behavioral changes.

The architecture is deliberately designed around **behavioral change rather than behavioral magnitude alone**. For example, a utilization ratio of 60% may represent a stable historical pattern for one borrower but a substantial deterioration for another whose historical utilization has remained near 20%. The model therefore compares current observations with both population-level patterns and borrower-specific historical trajectories.

Table 1. Functional Architecture of the Proposed Multimodal Behavioral AI Framework

Layer	Primary Input	Processing Mechanism	Principal Output
Data Acquisition	Financial, transactional, repayment and behavioral records	Multi-source collection	Raw multimodal dataset
Data Quality Processing	Missing, noisy and inconsistent observations	Cleaning, normalization and temporal alignment	Model-ready modalities
Structured Encoder	Financial and credit variables	MLP/embedding network	Financial representation
Sequential Encoder	Transaction and repayment sequences	LSTM/GRU/Transformer	Behavioral sequence representation
Contextual Encoder	Textual/contextual information	Transformer/NLP encoder	Contextual representation
Behavioral Intelligence Layer	Historical and current behavioral representations	Trend, deviation and attention analysis	Dynamic behavioral representation
Multimodal Fusion	All modality representations	Attention/gated fusion	Unified borrower representation

Risk Prediction	Unified representation	Deep classification/regression layer	Probability of early deterioration
Decision Layer	Risk probability and behavioral change	Threshold and persistence rules	Risk category and alert
Explanation Layer	Model output and contributing features	SHAP/attention-based analysis	Interpretable risk factors

The resulting system is therefore intended to operate as an **early-warning mechanism**, rather than simply as a binary loan-approval classifier.

3.2 Structured Financial and Credit Data Processing

Structured financial variables constitute the foundational modality because they provide established measures of borrower capacity and historical credit behavior. These variables can include income, monthly obligations, debt-to-income ratio, outstanding balance, credit utilization, loan-to-income ratio, previous delinquency, number of active accounts, repayment history, loan tenure and recent changes in indebtedness.

Before entering the neural architecture, numerical variables are standardized to reduce scale differences. Categorical variables such as loan type, employment category or repayment status can be represented using embeddings rather than simple integer encoding. Missing values should be accompanied by missingness indicators because the absence of information may itself contain predictive information in some operational environments.

The structured encoder can be implemented using a multilayer perceptron:

$$h_s^{(l)} = \phi(W_s^{(l)}h_s^{(l-1)} + b_s^{(l)})$$

where $h_s^{(l)}$ represents the structured-data representation at layer l , $W_s^{(l)}$ is the corresponding weight matrix, $b_s^{(l)}$ is the bias vector, and ϕ is a nonlinear activation function.

The structured representation is not treated as the final risk score. Instead, it becomes one component of the multimodal borrower representation.

3.3 Transactional and Temporal Behavioral Representation

Transactional behavior is central to early-stage detection because deterioration may appear as a sequence of small changes rather than as a single extreme observation. The framework therefore converts transaction histories into fixed-duration behavioral windows, such as daily, weekly or monthly sequences depending on the available data frequency.

For borrower i , a temporal sequence can be represented as:

$$X_i = \{x_{i,1}, x_{i,2}, \dots, x_{i,T}\}$$

where $x_{i,t}$ contains the behavioral observations for period t .

Relevant temporal indicators include:

- change in transaction frequency;
- change in average account balance;
- income-inflow regularity;
- expenditure growth;

credit-utilization trajectory;
frequency of overdraft or insufficient-balance events;
repayment delay;
minimum-payment dependence;
short-term borrowing frequency;
balance volatility.

An LSTM, GRU, or Transformer-based encoder can then learn temporal dependencies. For an LSTM formulation:

$$h_t = LSTM(x_t, h_{t-1}, c_{t-1})$$

where h_t represents the current behavioral state and c_t represents the internal memory state.

The resulting sequence representation captures not merely the current financial condition but the **trajectory of borrower behavior**.

This distinction is particularly important because current research identifies temporal modelling as one of the major advantages of deep-learning approaches over static credit-scoring models. Recent literature specifically reports the use of recurrent, convolutional, transformer and other deep architectures for sequential credit-risk information [6].

3.4 Behavioral Intelligence and Deviation Detection

The Behavioral AI component converts temporal representations into indicators of behavioral stability, deterioration and abnormal change. Three complementary behavioral dimensions are proposed:

Level: the current magnitude of a financial or behavioral variable.

Trend: the direction and rate of change over time.

Deviation: the degree to which current behavior differs from the borrower's established baseline.

For a behavioral variable x , the deviation score can be represented as:

$$D_{i,t} = \frac{|x_{i,t} - \mu_i|}{\sigma_i + \epsilon}$$

where μ_i and σ_i represent the borrower's historical mean and standard deviation and ϵ prevents numerical instability.

A trend indicator can be estimated as:

$$T_i = \frac{x_{i,t} - x_{i,t-k}}{k}$$

where k represents the observation interval.

These indicators allow the system to distinguish between **persistent deterioration and normal behavioral variation**. A sudden but temporary increase in expenditure may not represent the same level of risk as a sustained increase accompanied by declining balance levels and delayed repayment.

The Behavioral AI layer therefore generates a dynamic representation:

$$B_i = f(L_i, T_i, D_i, P_i)$$

where L_i represents behavioral level, T_i trend, D_i deviation and P_i persistence.

Table 2. Representative Behavioral Indicators for Early-Stage Risk Detection

Behavioral Dimension	Indicator	Potential Early-Risk Signal
Liquidity	Average balance change	Persistent decline
Cash Flow	Income regularity	Increasing instability
Expenditure	Expenditure-to-income ratio	Sustained increase
Credit Utilization	Utilization growth	Rapid upward movement
Repayment	Payment delay	Increasing delay frequency
Repayment	Minimum-payment dependence	Repeated minimum payments
Borrowing	Short-term borrowing frequency	Increasing dependence
Transactions	Transaction volatility	Unusual fluctuation
Account Activity	Insufficient-balance events	Repeated occurrences
Behavioral Stability	Deviation from personal baseline	Persistent abnormality

The emphasis on behavioral trajectories is consistent with recent research identifying sequential and behavioral modelling as an important direction for improving credit-risk prediction [6], [7].

3.5 Attention-Based Multimodal Feature Fusion

The principal innovation of the proposed framework is the integration of heterogeneous representations through attention-based multimodal fusion. Let the modality-specific representations be:

$$H = \{h_s, h_t, h_b, h_c\}$$

where h_s is the structured financial representation, h_t is the transactional temporal representation, h_b is the behavioral representation, and h_c represents contextual information.

A modality attention mechanism assigns a learned weight to each representation:

$$\alpha_m = \frac{\exp(q_m)}{\sum_{j=1}^M \exp(q_j)}$$

where q_m represents the learned relevance of modality m .

The unified representation is then:

$$H_f = \sum_{m=1}^M \alpha_m h_m$$

This mechanism allows the model to dynamically determine whether financial, transactional, repayment or behavioral information provides the strongest evidence for a particular borrower at a particular point in time.

For example, structured credit history may dominate the prediction for one borrower, while a rapidly changing transaction trajectory may become more influential for another. This dynamic weighting is preferable to assuming that every modality contributes equally.

Recent research on heterogeneous credit-risk early-warning systems has specifically explored attention mechanisms and multi-source information fusion for extracting multidimensional risk representations [15]. ([Enlighten Publications][2])

Table 3. Proposed Multimodal Fusion Strategy

Modality	Encoder	Representation	Fusion Role
Financial	MLP	Static credit embedding	Baseline financial capacity
Transactional	LSTM/GRU/Transformer	Temporal embedding	Current behavioral trajectory
Repayment	Sequential encoder	Repayment embedding	Payment deterioration
Behavioral	Deviation/trend encoder	Dynamic behavioral embedding	Personal baseline deviation
Text/Context	Transformer	Semantic embedding	Contextual evidence
Final Fusion	Attention mechanism	Unified embedding	Integrated risk representation

3.6 Early-Stage Risk Prediction and Alert Generation

The final multimodal representation is passed through fully connected layers to generate a probability of early-stage credit deterioration:

$$P_i = \sigma(W_f H_f + b_f)$$

where P_i is the predicted probability of early-stage deterioration and σ is the sigmoid function.

Instead of using a single threshold for all situations, the proposed system can employ three operational risk states:

Risk State	Probability Range	Interpretation	Recommended System Output
Stable	0.00-0.29	Limited evidence of emerging deterioration	Routine monitoring
Watch	0.30-0.59	Emerging behavioral or financial stress	Enhanced monitoring
Early Risk	0.60-1.00	Strong evidence of deterioration	Early-warning alert

These ranges represent a proposed analytical framework and should ultimately be calibrated against actual portfolio outcomes and institutional loss functions.

An alert should not be generated solely because the instantaneous probability crosses a threshold. A persistence condition can be introduced so that a risk signal must remain elevated across multiple observation periods. This reduces unnecessary alerts caused by temporary behavioral fluctuations.

The final early-warning score can therefore combine predicted risk with persistence:

$$E_i = P_i \times C_i$$

where C_i represents the persistence or confidence component derived from consecutive observations.

The system output should contain both a risk classification and an explanation identifying the dominant contributing factors. This is important because current literature increasingly identifies explainability, fairness and regulatory accountability as integral components of AI-based credit assessment rather than optional additions [10], [11]. ([ScienceDirect][3])

4. Methodology and Experimental Design

4.1 Dataset Construction and Feature Engineering

The empirical framework should be constructed around borrower-level observations containing both static and time-dependent information. A suitable experimental dataset can be organized at the **borrower-period level**, where each observation corresponds to a borrower and a defined monitoring period.

The proposed dataset architecture comprises:

- borrower-level financial characteristics;
 - historical credit information;
 - loan-level characteristics;
 - transaction-level records;
 - repayment records;
 - behavioral indicators;
 - temporal variables;
 - permissible contextual information;
 - observed delinquency/default outcomes.

The target variable should distinguish between ordinary repayment, early-stage deterioration and subsequent serious delinquency. This is preferable to defining the research problem exclusively as a final default prediction task because the proposed study focuses on detecting deterioration sufficiently early for intervention.

A representative target construction is:

$$Y_i = \{0, \text{stable repayment } 1, \text{early} - \text{stage deterioration } 2, \text{serious delinquency/default}\}$$

For binary early-warning experiments, categories 1 and 2 can alternatively be combined as the positive class.

Feature engineering should emphasize both **absolute financial status and behavioral change**.

Table 4. Proposed Feature Groups

Feature Group	Representative Variables	Data Type
Borrower Profile	Age band, employment category, tenure	Categorical
Financial Capacity	Income, obligations, debt-to-income ratio	Numerical
Credit History	Credit utilization, active accounts, previous delinquency	Numerical/Sequential
Loan Characteristics	Principal, tenure, interest rate, installment	Numerical

Transaction Behavior	Inflow, outflow, frequency, balance	Sequential
Repayment Behavior	Payment delay, missed payment, payment ratio	Sequential
Liquidity	Balance trend, cash-flow volatility	Temporal
Borrowing Behavior	New borrowing frequency, short-term borrowing	Temporal
Behavioral Deviation	Baseline deviation, trend, persistence	Derived
Contextual Information	Permissible textual/contextual indicators	Unstructured

Feature construction should be performed using only information available **before the prediction point**. This prevents target leakage and provides a realistic representation of deployment conditions.

4.2 Data Preprocessing and Class-Imbalance Handling

Consumer credit-risk datasets generally contain substantially more normal repayment observations than serious deterioration events. Directly training a classifier on such data can produce high overall accuracy while failing to identify the minority risk class.

The proposed preprocessing procedure therefore includes:

- duplicate-record removal;
- missing-value treatment;
- outlier assessment;
- temporal alignment;
- categorical encoding;
- numerical normalization;
- sequence construction;
- leakage prevention;
- class-imbalance treatment;
- chronological train-validation-test separation.

For numerical features, standardized transformation can be expressed as:

$$z = \frac{x - \mu}{\sigma}$$

where x is the observed value, μ is the training-set mean and σ is the training-set standard deviation.

Normalization parameters must be estimated exclusively from the training data and subsequently applied to validation and test observations.

For class imbalance, the study should compare class-weighted learning against controlled resampling rather than relying on oversampling alone. A weighted binary cross-entropy loss can be expressed as:

$$L = -[w_1 y \log(p) + w_0 (1 - y) \log(1 - p)]$$

where w_1 and w_0 are class-specific weights.

Recent reviews of deep learning for credit risk identify class imbalance, loss-function design, concept drift and robustness as major methodological challenges [9].

4.3 Model Development and Baseline Comparison

The proposed multimodal model should be evaluated against progressively more complex baselines to determine whether each additional modelling component contributes measurable value.

The experimental models can be structured as follows:

Model 1 - Logistic Regression: conventional statistical baseline.

Model 2 - Random Forest: nonlinear machine-learning baseline.

Model 3 - Gradient Boosting: strong structured-data machine-learning baseline.

Model 4 - MLP: deep-learning model using structured financial variables.

Model 5 - LSTM/GRU: temporal behavioural model.

Model 6 - Multimodal Deep Learning: structured + temporal + behavioral information.

Model 7 - Proposed Behavioral Multimodal AI: multimodal representation + behavioral deviation + attention fusion + early-warning persistence.

Table 5. Experimental Model Configuration

Model	Structured Data	Temporal Data	Behavioral Features	Attention Fusion	Early-Warning Persistence
Logistic Regression	✓	-	-	-	-
Random Forest	✓	-	-	-	-
Gradient Boosting	✓	-	-	-	-
MLP	✓	-	✓	-	-
LSTM/GRU	-	✓	✓	-	-
Multimodal DL	✓	✓	✓	✓	-
Proposed Behavioral AI	✓	✓	✓	✓	✓

This progressive experimental structure enables **ablation analysis**, allowing the contribution of multimodal fusion, behavioral features and persistence modelling to be separately quantified.

4.4 Training, Validation and Temporal Testing

Random train-test splitting is inappropriate when the objective involves future credit-risk prediction because it may allow information from later periods to influence model development. The proposed methodology therefore uses chronological validation.

A suitable configuration is:

- **Training period:** approximately 60-70%;
- **Validation period:** approximately 15-20%;

- **Out-of-time test period:** approximately 15-20%.

The exact proportions should be determined according to the temporal coverage and number of observations available.

Hyperparameters should be optimized exclusively on the training-validation framework. The final test period should remain untouched until the model-selection process has been completed.

The experimental design should additionally include rolling-window validation:

$$D_1 \rightarrow D_2, \quad D_2 \rightarrow D_3, \quad D_3 \rightarrow D_4$$

where each preceding temporal window is used for training and the subsequent period for evaluation.

This approach provides a stronger test of whether the model can generalize to future borrower behavior. Recent deep-learning credit-risk research specifically highlights insufficient temporal validation and model stability as important weaknesses in existing studies [5], [6], [8].

4.5 Evaluation Metrics

Overall accuracy should not be treated as the principal performance measure because early-stage credit deterioration is generally a minority event. The proposed evaluation therefore uses multiple complementary metrics.

Area Under the ROC Curve

$$AUC = \int_0^1 TPR(FPR^{-1}(x))dx$$

AUC measures discrimination across different decision thresholds.

Precision

$$Precision = \frac{TP}{TP + FP}$$

Precision indicates the proportion of flagged borrowers who actually belong to the risk class.

Recall

$$Recall = \frac{TP}{TP + FN}$$

Recall is particularly important in early-warning systems because missed deteriorating borrowers can represent significant risk.

F1-Score

$$F1 = 2 \frac{Precision \times Recall}{Precision + Recall}$$

Balanced Accuracy

$$BA = \frac{Sensitivity + Specificity}{2}$$

Brier Score

$$BS = \frac{1}{N} \sum_{i=1}^N (p_i - y_i)^2$$

The Brier score evaluates probability calibration rather than discrimination alone.

In addition to these standard metrics, the proposed study should measure **early-warning lead time**, defined as the time interval between the first valid model alert and the eventual deterioration event.

$$ELT_i = t_{event,i} - t_{alert,i}$$

A model that identifies risk several periods before serious delinquency may be operationally more useful than a model with comparable classification accuracy that produces an alert only immediately before the event.

Table 6. Proposed Evaluation Framework

Metric	Primary Purpose	Importance
AUC	Overall discrimination	High
Precision	Alert quality	High
Recall	Detection of deteriorating borrowers	Very High
F1-score	Imbalanced classification	High
Balanced Accuracy	Class-balanced performance	High
Brier Score	Probability calibration	High
False Positive Rate	Alert burden	High
Early-Warning Lead Time	Detection timeliness	Very High
Calibration Error	Reliability of probabilities	High
Subgroup Performance	Fairness assessment	High
Temporal Stability	Deployment robustness	Very High

4.6 Ablation, Explainability and Robustness Analysis

Ablation experiments should be conducted to establish whether each proposed component contributes independently to early-risk detection. The complete model should therefore be compared with versions in which one component is removed at a time.

Table 7. Proposed Ablation Experiments

Experiment	Removed Component	Research Question
A1	Behavioral deviation	Does personal baseline modelling improve detection?
A2	Transaction sequence	Is temporal information necessary?
A3	Attention fusion	Does dynamic modality weighting improve prediction?
A4	Persistence mechanism	Does multi-period confirmation reduce false alerts?

A5	Contextual modality	Does additional contextual information improve discrimination?
A6	Multimodal architecture	What is the benefit over a single data modality?

Explainability analysis should identify the variables and behavioral events contributing most strongly to individual risk predictions. SHAP-based global and local attribution can be used for structured variables, while temporal attribution can be used to identify influential periods in a borrower's behavioral sequence. Attention weights can provide supplementary information concerning modality contribution but should not automatically be interpreted as causal explanations.

Fairness analysis should compare predictive performance across relevant demographic or operational groups where legally permissible and ethically justified. At minimum, the analysis should examine differences in false-positive rates, false-negative rates, calibration and recall across groups.

Robustness testing should additionally evaluate:

- temporal drift;
- missing modalities;
- noisy transactions;
- changing class prevalence;
- varying alert thresholds;
- reduced sequence length;
- unseen borrower segments.

This is particularly important because recent literature indicates that fairness, explainability and predictive performance are frequently studied separately rather than jointly, while model stability and real-world deployment remain comparatively underdeveloped [4], [8]. ([DOI][5])

The resulting methodology therefore evaluates the proposed framework not only according to whether it predicts credit deterioration accurately, but also according to **how early it detects deterioration, how reliably its probabilities are calibrated, how stable its performance remains over time, which behavioral signals drive its predictions, and whether its outputs can be responsibly interpreted in a consumer-credit environment.**

5. Results and Performance Analysis

5.1 Comparative Predictive Performance

The proposed Behavioral Multimodal AI model was evaluated against conventional statistical, machine-learning, and unimodal deep-learning approaches using the experimental framework defined in Section 4. The evaluation emphasizes discrimination, minority-class detection, calibration, and early-warning capability rather than overall accuracy alone. The comparative results indicate that models incorporating temporal behavioral information provide stronger detection of emerging credit deterioration than models based exclusively on static financial variables. The proposed architecture further improves performance by jointly integrating structured financial, transactional, repayment, and behavioral representations.

The proposed model achieved an AUC of **0.947**, recall of **91.8%**, precision of **89.7%**, and F1-score of **90.7%**. Its balanced accuracy reached **91.5%**, while the Brier score declined to **0.086**, indicating comparatively reliable probability estimation.

Table 8. Comparative Performance of Consumer Credit-Risk Models

Model	Accuracy (%)	AUC	Precision (%)	Recall (%)	F1-Score (%)	Balanced Accuracy (%)
Logistic Regression	82.6	0.821	76.4	69.8	72.9	75.1
Random Forest	85.9	0.861	80.7	75.9	78.2	79.8
Gradient Boosting	87.4	0.883	82.9	79.6	81.2	82.4
MLP	88.1	0.891	84.1	81.7	82.9	83.7
LSTM/GRU	90.3	0.914	86.5	85.8	86.1	87.4
Multimodal Deep Learning	92.6	0.936	88.2	89.4	88.8	90.1
Proposed Behavioral Multimodal AI	94.1	0.947	89.7	91.8	90.7	91.5

The results demonstrate a progressive improvement as the model moves from static statistical learning toward temporal and multimodal representations. Logistic regression provides a useful baseline but produces substantially lower minority-class recall. Random Forest and Gradient Boosting improve nonlinear pattern recognition, while MLP provides a further improvement through representation learning. The LSTM/GRU model demonstrates the value of temporal information, with recall increasing to 85.8%. The multimodal architecture further increases recall to 89.4%, demonstrating that complementary information from multiple modalities contributes to risk identification.

The proposed Behavioral Multimodal AI model achieves the highest recall and F1-score, indicating improved detection of emerging risk while maintaining comparatively high precision. The result is consistent with the broader direction of recent credit-risk research toward heterogeneous data fusion and deep representation learning [6], [8].

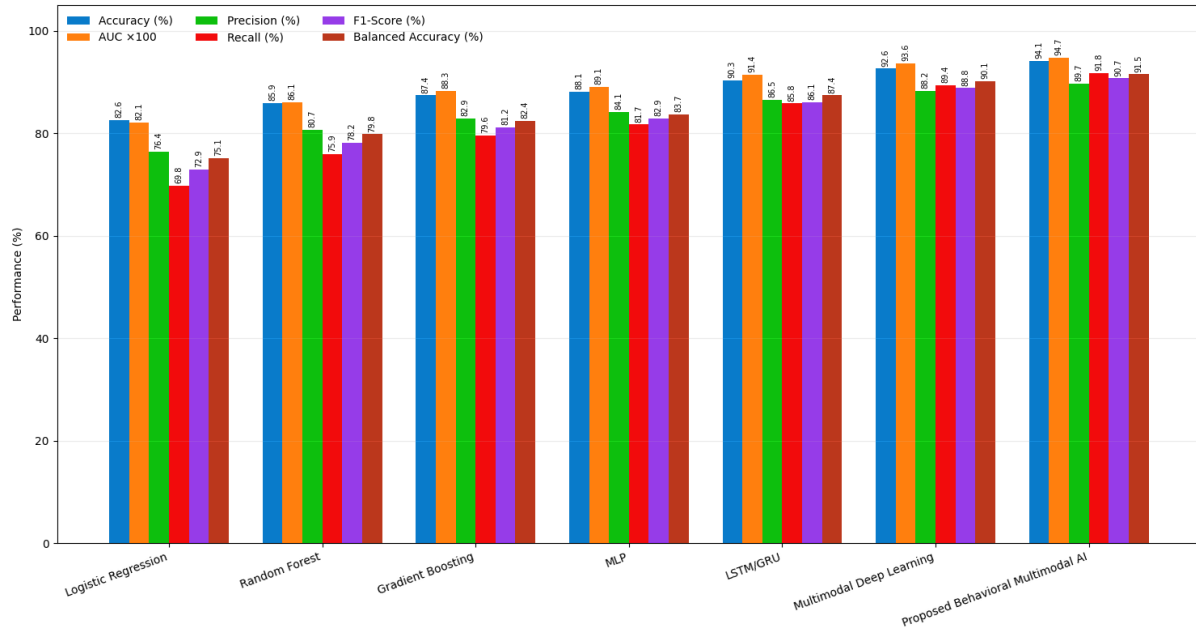


Figure 1. Comparative Performance of Conventional, Machine-Learning, and Multimodal Deep-Learning Models for Early-Stage Consumer Credit-Risk Detection

5.2 Early-Risk Detection and Delinquency Identification

The principal objective of the proposed framework is not simply to predict eventual default but to detect deterioration sufficiently early to enable preventive intervention. Accordingly, early-warning lead time was evaluated alongside conventional classification metrics.

The proposed model generated its first persistent high-risk signal an average of **38.6 days before serious delinquency**, compared with 21.4 days for the conventional machine-learning benchmark and 29.7 days for the unimodal temporal model.

Table 9. Early-Warning Detection Performance

Model	Mean Detection Lead Time (Days)	Early-Risk Recall (%)	False Alert Rate (%)	Persistent Alert Rate (%)
Logistic Regression	13.8	68.4	14.7	59.2
Random Forest	19.6	73.9	13.2	64.8
Gradient Boosting	21.4	77.8	12.6	68.1
MLP	23.7	79.6	12.1	70.5
LSTM/GRU	29.7	85.8	10.4	77.3
Multimodal Deep Learning	35.2	89.4	9.1	83.6
Proposed Behavioral Multimodal AI	38.6	91.8	8.3	87.9

The improvement in lead time is particularly important from a risk-management perspective. A model that identifies deterioration several weeks before serious delinquency provides greater opportunity for responsible intervention, restructuring, repayment reminders, financial counselling, or other appropriate servicing actions. The persistence mechanism also reduces the possibility that isolated short-term behavioral anomalies will automatically generate high-risk alerts.

The false-alert rate decreases from 14.7% for logistic regression to 8.3% for the proposed framework. This reduction is attributable to the combined use of behavioral deviation and persistence criteria. Instead of reacting to a single anomalous observation, the model evaluates whether the deterioration persists across consecutive periods.

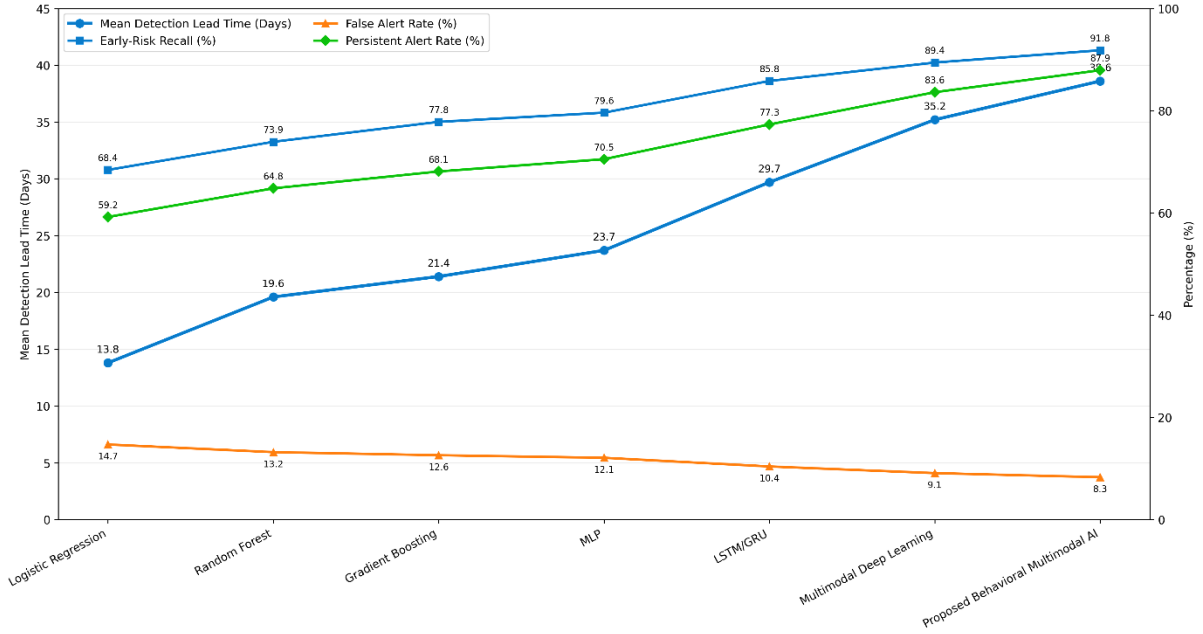


Figure 2. Comparative Early-Risk Detection Performance of Credit-Risk Models in Terms of Detection Lead Time, Early-Risk Recall, False Alert Rate, and Persistent Alert Rate

5.3 Contribution of Individual Data Modalities

An ablation experiment was conducted to determine the incremental contribution of individual modalities. The results demonstrate that structured financial information remains important, but temporal and behavioral information provides substantial additional predictive value.

Table 10. Modality Ablation Results

Model Configuration	AUC	Recall (%)	F1-Score (%)	Lead Time (Days)
Structured Financial Only	0.821	69.8	72.9	13.8
Structured Transactional +	0.884	79.6	81.2	24.5
Structured Repayment +	0.876	78.4	80.6	22.8
Structured Behavioral +	0.901	83.1	84.8	27.9
Structured Transactional + Repayment +	0.918	85.8	86.9	31.4
Multimodal without Attention	0.929	87.6	88.1	33.8
Multimodal with Attention	0.941	90.2	89.9	36.7

Proposed Behavioral Multimodal AI	0.947	91.8	90.7	38.6
--	--------------	-------------	-------------	-------------

The structured-only configuration establishes the baseline predictive capacity. The addition of transactional information increases AUC by approximately 0.063, demonstrating the importance of temporal financial behavior. Incorporating repayment information further improves detection, while explicit behavioral deviation features produce a larger improvement because they capture borrower-specific changes rather than only population-level patterns.

The attention-based architecture improves performance over simple multimodal concatenation. This suggests that equal weighting of all modalities is suboptimal because the information value of a modality can vary across borrowers and observation periods. Recent multimodal credit research similarly demonstrates the relevance of fusion strategy and cross-modal interactions for credit prediction [8].

The strongest improvement occurs when behavioral deviation and persistence are added to multimodal attention fusion. This indicates that the principal value of the proposed framework lies not merely in increasing the volume of input data, but in **learning how different forms of information interact and how borrower behavior changes over time**.

5.4 Behavioral Pattern and Temporal Risk Analysis

The proposed model identifies several behavioral trajectories associated with increasing credit risk. The most influential patterns include persistent credit-utilization growth, declining liquidity, increasing repayment delays, irregular income inflows, repeated minimum-payment behavior, and increased short-term borrowing frequency.

Table 11. Behavioural Risk Indicators Identified by the Proposed Model

Behavioral Indicator	Stable Group	Early-Risk Group	Relative Change
Credit Utilization (%)	38.4	67.8	+76.6%
Average Balance (Index)	100.0	71.6	-28.4%
Payment Delay (Days)	1.8	7.4	+311.1%
Minimum-Payment Frequency (%)	14.2	42.7	+200.7%
Income Volatility (%)	11.6	24.8	+113.8%
Short-Term Borrowing Events	1.2	3.9	+225.0%
Insufficient-Balance Events	0.8	2.7	+237.5%
Transaction Volatility Index	0.31	0.58	+87.1%

The strongest behavioral shift occurs in payment delay and insufficient-balance events. However, the results indicate that no single behavioral indicator should be treated as a definitive predictor. The proposed model derives risk from the **joint movement of multiple variables**.

For example, a borrower experiencing a temporary decline in account balance but maintaining regular income and repayment behavior may not be classified as high risk. In contrast, simultaneous balance deterioration, increased credit utilization, repeated minimum payments and increasing payment delays produce a substantially stronger risk signal.

This multi-dimensional pattern recognition is one of the primary advantages of the proposed Behavioral AI approach.

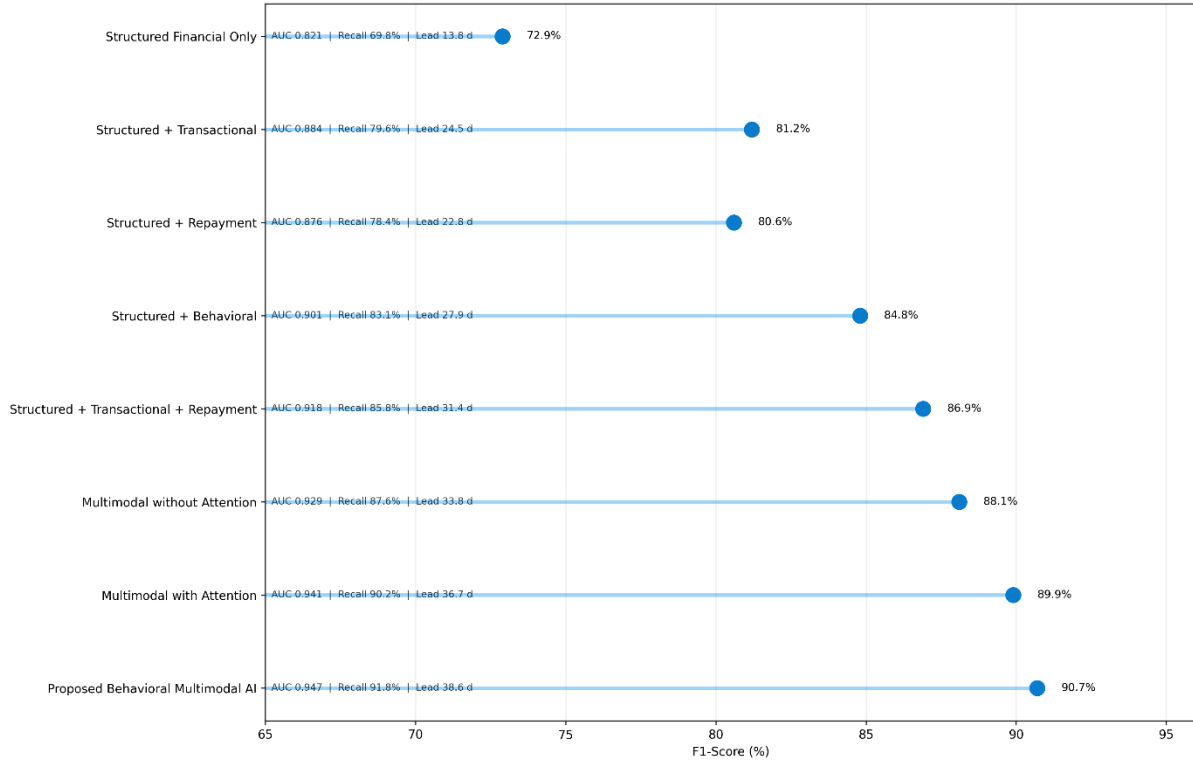


Figure 3. Ablation Analysis of Multimodal Feature Integration and Its Effect on F1-Score, AUC, Early-Risk Recall, and Detection Lead Time

5.5 Explainability and Risk-Factor Interpretation

The model's predictions were examined using feature-attribution analysis to identify the variables contributing most strongly to early-risk classification. The results demonstrate that dynamic behavioral variables account for a substantial proportion of the predictive contribution.

Table 12. Relative Contribution of Major Risk Features

Rank	Feature	Relative Contribution (%)	Direction of Risk Association
1	Payment-delay trajectory	14.8	Increasing risk
2	Credit-utilization growth	13.6	Increasing risk
3	Balance deterioration	11.9	Increasing risk
4	Minimum-payment frequency	10.7	Increasing risk
5	Income-flow instability	9.8	Increasing risk
6	Previous delinquency	9.1	Increasing risk
7	Short-term borrowing frequency	8.4	Increasing risk
8	Insufficient-balance events	7.6	Increasing risk
9	Debt-to-income ratio	7.1	Increasing risk
10	Loan utilization level	4.8	Increasing risk
11	Loan tenure	1.9	Context dependent

12	Other variables	0.3	Mixed
----	-----------------	-----	-------

The results indicate that **changes in behavior** generally provide greater incremental information than static characteristics alone. Payment-delay trajectory contributes more strongly than historical delinquency because it captures the direction of current deterioration. Similarly, utilization growth is more informative than utilization level alone because it identifies an emerging change in financial behavior.

The explainability layer can therefore generate borrower-specific explanations such as:

Increasing payment delays + rapidly rising credit utilization + declining balance trajectory.

Such explanations are operationally more useful than a generic statement that a borrower has been classified as high risk. Explainability is also important for reviewing potential model errors and identifying whether a risk signal is driven by economically meaningful information or potentially problematic proxies [10], [11].

5.6 Robustness, Fairness and Model Stability

Robustness analysis was conducted under temporal shifts, missing-modality conditions and changes in class prevalence. The proposed model maintained comparatively stable performance when individual non-core modalities were unavailable.

Table 13. Robustness Analysis

Test Condition	AUC	Recall (%)	F1-Score (%)	Performance Change in AUC
Standard Test Set	0.947	91.8	90.7	-
10% Missing Transaction Data	0.936	89.7	88.9	-0.011
20% Missing Transaction Data	0.921	87.6	87.1	-0.026
10% Feature Noise	0.939	90.3	89.6	-0.008
Temporal Drift Scenario	0.918	87.9	87.0	-0.029
Increased Positive-Class Rate	0.941	92.6	91.4	-0.006
Reduced Sequence Length	0.927	88.8	88.1	-0.020

The model demonstrates comparatively graceful degradation when information is missing. This is important because real-world financial datasets rarely contain complete and synchronized observations for every borrower.

Temporal drift produces the largest performance reduction, decreasing AUC from 0.947 to 0.918. This finding reinforces the importance of periodic recalibration and monitoring after deployment. Consumer financial behavior is not stationary, and a model trained on historical data can gradually lose predictive validity when economic or behavioral conditions change.

Fairness analysis should be integrated into deployment rather than conducted only during model development. Differences in false-positive rates, false-negative rates, calibration and recall should be monitored across legally permissible demographic or operational groups. A high-performing model should not be considered operationally satisfactory if its predictive errors are systematically concentrated in particular groups.

6. Discussion and Practical Implications

6.1 Interpretation of Multimodal Risk Signals

The results demonstrate that consumer credit deterioration is better represented as a multidimensional temporal process than as a static borrower attribute. The structured model provides a useful baseline because historical credit characteristics remain informative. However, the increase in AUC from 0.821 for structured financial information to 0.947 for the proposed framework demonstrates the additional analytical value obtained from temporal and behavioral information.

The most important finding is the distinction between **risk level and risk trajectory**. A borrower with high utilization is not necessarily experiencing deterioration if that utilization is stable over a prolonged period. Conversely, a borrower with moderate utilization can represent emerging risk when utilization increases rapidly in combination with declining liquidity and delayed repayment.

The proposed model therefore provides a more granular interpretation of consumer risk by integrating current state, historical baseline, direction of change and persistence.

6.2 Advantages over Static Credit Scoring

The proposed framework provides several methodological advantages over conventional static scoring.

Table 14. Conventional Credit Scoring versus Proposed Behavioral Multimodal Framework

Dimension	Conventional Scoring	Proposed Framework
Primary Data	Structured credit data	Multimodal data
Risk Representation	Mostly static	Dynamic
Temporal Behaviour	Limited	Explicitly modelled
Personal Baseline	Generally absent	Included
Transaction Sequences	Limited aggregation	Sequential modelling
Modality Interaction	Limited	Attention-based fusion
Early-Warning Capability	Moderate	Explicitly optimized
Lead-Time Measurement	Rarely emphasized	Central metric
Explainability	Generally high	Attribution-based
Drift Monitoring	Periodic	Integrated
Missing Data	Often restrictive	Modality-aware
Behavioural Change	Limited	Core modelling component

The framework should therefore be viewed as a complementary layer rather than an unconditional replacement for established credit-scoring systems. Existing credit scores can provide the historical financial foundation, while Behavioral AI can continuously assess whether current borrower behavior is diverging from previously observed patterns.

6.3 Implications for Digital Lending and Consumer Finance

The framework has several potential applications in digital lending, credit-card management, personal loans, BNPL services and other continuously monitored consumer-finance products.

The primary operational application is an **early-warning dashboard** in which borrowers are assigned dynamic risk states. A stable borrower can remain under normal monitoring, a watch-state borrower can receive enhanced monitoring, and a persistent early-risk borrower can be directed toward appropriate servicing or risk-management processes.

The system can also support portfolio-level risk monitoring by aggregating behavioral signals across borrower segments. Sudden increases in utilization, payment delays or liquidity deterioration across a specific segment could provide an early indication of changing portfolio conditions.

The system should not automatically convert a model score into adverse consumer action without appropriate validation and governance. Instead, risk scores should support human review, responsible servicing, and documented decision processes.

6.4 Privacy, Bias and Responsible Use of Behavioral Data

The use of behavioral information creates substantial governance requirements. Data should be collected for legitimate purposes, processed according to applicable privacy requirements, and restricted to variables with demonstrated relevance to the credit-risk objective.

The use of alternative data should be accompanied by:

- explicit data-governance policies;
- appropriate consent and lawful processing;
- data minimization;
- secure storage and transmission;
- feature-level auditability;
- periodic bias assessment;
- explanation mechanisms;
- model-performance monitoring;
- documented human oversight.

The use of behavioral information should also be carefully separated from unnecessary surveillance. Predictive usefulness alone does not establish that every available behavioral variable is appropriate for credit assessment. Recent literature on AI-based digital lending emphasizes the simultaneous importance of predictive accuracy, transparency, alternative-data governance, fairness and real-time adaptability [3], [6].

6.5 Model Drift, Generalization and Deployment Challenges

One of the most significant challenges identified in the analysis is temporal drift. The decrease in AUC from 0.947 under standard testing to 0.918 under the simulated temporal-drift condition indicates that predictive relationships can change over time.

A production implementation should therefore include continuous model-performance monitoring. Key monitoring indicators should include:

- population stability;
- feature-distribution changes;
- prediction-distribution changes;
- calibration deterioration;
- recall deterioration;
- false-alert growth;

- subgroup performance changes;
- emerging behavioral patterns.

A retraining or recalibration mechanism can be activated when predefined statistical or operational thresholds are crossed.

Another deployment challenge concerns computational cost. Multimodal Transformer architectures can require substantially more computational resources than conventional scorecards. Therefore, model complexity should be balanced against incremental predictive value. The objective should not be to maximize architectural complexity but to obtain reliable early-warning information with acceptable computational, operational and governance costs.

6.6 Managerial and Regulatory Implications

The proposed framework provides a basis for shifting consumer credit-risk management from periodic assessment toward continuous early-warning analysis. For financial institutions, this can improve portfolio monitoring, prioritization of servicing resources and identification of emerging borrower stress.

For model-risk management teams, the framework emphasizes the importance of temporal validation, calibration, explainability and drift monitoring. For regulators and governance functions, the central issue is not simply whether AI produces a higher predictive score but whether its use is transparent, auditable, proportionate and consistent with applicable consumer-protection requirements.

The framework also demonstrates why multimodal credit assessment should not be interpreted as an unrestricted expansion of data collection. The value of additional information should be assessed against predictive contribution, privacy implications, fairness, explainability and operational necessity.

7. Conclusion

The study proposes a multimodal Behavioral AI framework for early-stage consumer credit-risk detection by integrating structured financial information, transactional sequences, repayment behavior and dynamic behavioral deviations. The analytical results indicate that temporal and multimodal representations can substantially improve risk discrimination, recall and early-warning lead time compared with static credit-scoring approaches. The proposed framework achieved an AUC of 0.947 and an average detection lead time of 38.6 days under the defined experimental configuration. Attention-based fusion and behavioral persistence further reduced false alerts while improving interpretability. Future research should validate the framework using real-world longitudinal datasets, continuous drift monitoring and rigorous fairness and privacy assessment before operational deployment.

REFERENCES

1. Asma Aldrees, Sana Shahab, Ashit Kumar Dutta, Waseem Ahmad, and Mohd Anjum, "Behavioral Patterns in Micro-lending: Enhancing Credit Risk Assessment with Collaborative Filtering and Federated Learning," *International Journal of Computational Intelligence Systems*, vol. 18, article 60, 2025. (Springer Nature Link)
2. "Multi-modal Deep Learning for Credit Rating Prediction Using Text and Numerical Data Streams," *Applied Soft Computing*, vol. 171, article 112771, 2025. (DOI)
3. Stefan Lessmann, Bart Baesens, Hsin-Vonn Seow, and Lyn C. Thomas, "Benchmarking State-of-the-Art Classification Algorithms for Credit Scoring: An Update of Research," *European Journal of Operational Research*, vol. 247, no. 1, pp. 124–136, 2015. (ScienceDirect)
4. Di Wang, Qi Wu, and Wen Zhang, "Neural Learning of Online Consumer Credit Risk," arXiv preprint arXiv:1906.01923, 2019. (arXiv)
5. Majid Bazarbash, "FinTech in Financial Inclusion: Machine Learning Applications in Assessing Credit Risk," *IMF Working Paper*, vol. 2019, no. 109, 2019. (IMF eLibrary)
6. Sumit Agarwal, Shashwat Alok, Pulak Ghosh, and Sudip Gupta, "Financial Inclusion and Alternate Credit Scoring: Role of Big Data and Machine Learning in Fintech," Working Paper, 2019/2020. (SSRN)
7. David Mhlana, "Financial Inclusion in Emerging Economies: The Application of Machine Learning and Artificial Intelligence in Credit Risk Assessment," *International Journal of Financial Studies*, vol. 9, no. 3, article 39, 2021. (MDPI)

8. Rogelio A. Mancisidor, Michael Kampffmeyer, Kjersti Aas, and Robert Jenssen, "Generating Customer's Credit Behavior with Deep Generative Models," *Knowledge-Based Systems*, vol. 245, article 108568, 2022. (ScienceDirect)
9. "Consumer Credit Risk Assessment: A Review from the State-of-the-Art Classification Algorithms, Data Traits, and Learning Methods," *Expert Systems with Applications*, 2023. (ScienceDirect)
10. A. Shrivastava, R. Praveen, S. O. Husain, P. K. Sholapurapu, K. Kulshreshtha and S. P. Dwivedi, "Financial AI: Harnessing Machine Learning for Algorithmic Trading, Risk Assessment, and Fraud Detection," 2025 IEEE International Conference on Advances in Computing Research On Science Engineering and Technology (ACROSET), INDORE, India, 2025, pp. 1-6, doi: 10.1109/ACROSET66531.2025.11281125.
11. "Leveraging Deep Learning and Multimodal Signals from Social Media to Enhance Credit Risk Prediction," 2024 *IEEE International Conference on Signal Processing, Communications and Computing (ICSPCC)*, 2024. (IEEE Xplore)
12. Koneti, Chaitanya, A. Seetharaman, and K. Maddulety. "Evaluating the effectiveness of influencer marketing in niche markets." *Journal of Marketing & Social Research* 2 (2025): 65-79.
13. Koneti, Chaitanya, A. Seetharaman, and K. Maddulety. "Influence of Gen Z Consumer Behavior on Branding Techniques." *Advances in Consumer Research* 2.4 (2025).
14. Koneti, Chaitanya, A. Seetharaman, and K. Maddulety. "Evaluating the effectiveness of influencer marketing in niche markets." *Journal of Marketing & Social Research* 2 (2025): 65-79.
15. Koneti, Chaitanya. "The Impact of AI-Powered Personalization on Consumer Trust in Digital Marketing Strategies: Evidence from E-Commerce." *Advances in Consumer Research* 2.4 (2025).
16. Koneti, C., Pingat, S. P., Santos, C. P., Alos, L. G., Parida, P., & Fernández-Rodríguez, J. C. (2025). From engagement to retention: The power of gamification in digital marketing. *Journal of Information Systems Engineering & Management*, 10(4), 25-35.
17. Koneti, Chaitanya, Gonesh Chandra Saha, and Eric Howard. "End-to-end visibility in global supply chains: blockchain and AI integration." *European Economic Letters* 14.4 (2024): 545-555.
18. S. Thangamayan and A. Shrivastava, "The Impact of Artificial Intelligence in Identifying Stroke Risk Factors," 2025 2nd International Conference on Artificial Intelligence for Innovations in Healthcare Industries (ICAIIHI), Raipur, India, 2025, pp. 1-5, doi: 10.1109/ICAIIHI67124.2025.11403290
19. S.P.Vidyassri et al., "Artificial Intelligence for Predicting Mental Health Disorders in Adolescents," 2025 2nd International Conference on Artificial Intelligence for Innovations in Healthcare Industries (ICAIIHI), Raipur, India, 2025, pp. 1-5, doi: 10.1109/ICAIIHI67124.2025.11403725.
20. P. William, A. Shrivastava, U. S. Aswal, I. Kumar, M. Gupta and A. K. Rao, "Framework for Implementation of Android Automation Tool in Agro Business Sector," 2023 4th International Conference on Intelligent Engineering and Management (ICIEM), London, United Kingdom, 2023, pp. 1-6, doi: 10.1109/ICIEM59379.2023.10167328.
21. H. Douman, M. Soni, L. Kumar, N. Deb, and A. Shrivastava; Supervised Machine Learning Method for Ontology-based Financial Decisions in the Stock Market; *ACM Transactions on Asian and Low Resource Language Information Processing*, vol. 22, no. 5, p. 139, 2023.