Comparative Analysis of Bayesian and Classical Logistic Regression for Credit Scoring in Kenya Using Non-Informative Priors

Ngoa E. Chiro[1], Aggrey Adem[1], Martin H. Kai[1], Josephat Mutembei[1] & Leonard K. Alii[2]

Technical University of Mombasa, Kenya

Pwani University, Kenya

**Corresponding author: evancechiro@gmail.com*

https://doi.org/10.62049/jkncu.v6i2.560

Abstract

Credit scoring is essential for financial institutions, particularly in Kenya, where limited or nonexistent credit histories complicate the assessment of new clients. Classical logistic regression models, though widely used, can underperform in such data-scarce environments. This study compared classical logistic regression with Bayesian logistic regression models employing non-informative priors, based on several classification criteria, to assess whether non-informative priors can be successfully used in developing credit scoring models without compromising predictive accuracy. Data was collected retrospectively from one of the tier-one commercial banks in Kenya. The data included borrower demographic information, financial attributes, credit history, and loan default status. The performance of the models was compared mainly using AUC. The results showed that the Uniform prior model achieved the highest AUC (0.8642), Normal prior model (0.8639), Jeffrey's prior model (0.8636), and Classical model (0.8635). The study provided a comprehensive overview of other metrics, such as residual deviance, precision, accuracy, recall, and specificity, in order to give room for the financier to choose their preferred model depending on the specific objectives and risk appetite. In addition, each model's strengths and limitations were highlighted in order to give room for further exploration of the models for better predictive performance.

Keywords: Credit Scoring, Bayesian Logistic Regression, Non-Informative Priors, Classical Logistic Regression, Credit Risk Assessment, Area Under Curve (AUC)

Introduction

In the highly competitive financial sector, credit scoring has become a fundamental tool for evaluating borrowers' creditworthiness. Financial institutions, including commercial banks, microfinance institutions, and cooperative societies, rely heavily on credit scoring models to assess potential borrowers before granting loans. These models not only guide lending decisions but also play a critical role in risk management, ensuring that institutions minimize loan defaults while maintaining profitability.

Classical logistic regression has been widely used in credit scoring. (Webster, 2011) proves the effectiveness of the Bayesian logistic regression models using informative priors in credit scoring compared to the classical models. The performance of Bayesian logistic regression models employing non-informative priors in the Kenyan banking sector remains insufficiently explored. This study aimed to determine how the predictive performance of Bayesian logistic regression models developed under Normal, Uniform, and Jeffrey's non-informative priors compared with the classical logistic regression model in credit scoring.

The main objective of this study was to compare the performance of the classical logistic regression and non-informative Bayesian models based on model fit, discriminatory ability, and classification performance as measured mainly by their area under the operating characteristic curve (AUC). This will improve predictive accuracy and support sound credit decision-making in Kenyan financial institutions.

Literature

Credit Scoring Models

Credit scoring is the process of using statistical or machine learning models to predict the probability of loan default (Zhang et al., 2025). Traditional models, particularly logistic regression, have dominated the field due to their interpretability, robustness, and suitability for binary classification problems (Hand & Henley, 1997). Logistic regression uses borrower characteristics such as income, age, employment status, and credit history to estimate the probability of default. Despite its popularity, logistic regression is limited in contexts where sample sizes are small or borrower histories are incomplete.

Recent studies have explored alternative methods, including decision trees, neural networks, and ensemble techniques (Ayari et al., 2026). However, these models often lack interpretability, which is a crucial requirement for financial regulators and lenders. Bayesian approaches, on the other hand, offer both interpretability and flexibility in handling data limitations, making them an attractive alternative for credit scoring.

Logistic Regression Model

Logistic regression is widely applied in credit scoring due to its ability to model a binary dependent variable, such as loan repayment status (default or non-default) (AJPAS, 2025).

The logistic regression model is given by: $p(Y_i = 1|X_i) = \frac{e^{x_i^{\mathrm{T}}\beta}}{\left(1+e^{x_i^{\mathrm{T}}\beta}\right)}$ ……………………….. (1)

where Y_i is the binary outcome, X_i is the vector of borrower characteristics, and β is the vector of regression coefficients. The model assumes that the log-odds of default are a linear function of the predictor $logit(P(Y_i = 1|X_i) = X^T \beta$s). Logistic regression coefficients are typically estimated using maximum likelihood estimation (MLE) (Barasa et al., 2025). However, in small sample settings or when predictors exhibit Multicollinearity, MLE can yield unstable results (Hosmer, Lemeshow, & Sturdivant, 2013). These limitations motivate the use of Bayesian methods.

Bayesian Logistic Regression

Bayesian logistic regression extends the classical approach by incorporating prior beliefs about parameters into the estimation process. Unlike MLE, which produces point estimates, Bayesian methods generate full posterior distributions for parameters. The posterior distribution is derived as:

$P(\beta|Data) \alpha L(\beta|Data) * \pi(\beta)$ where $L(\beta|Data)$ is the likelihood function, and $\pi(\beta)$ is the prior distribution. Posterior inference is often performed using Markov Chain Monte Carlo (MCMC) methods (He et al., 2021a). Bayesian methods are particularly advantageous when historical data is scarce or incomplete. By introducing priors, these models can stabilize parameter estimates and improve predictive performance (Kosmidis & Firth, 2021).

Prior Distributions in Bayesian Logistic Regression

Normal Prior

The choice of prior distributions in Bayesian logistic regression carries substantial implications for model regularization and inference, particularly in high-dimensional or sparse data contexts. Traditionally, large-variance normal priors such as $N(0, 100^6)$ have been employed to represent weakly informative or non-informative beliefs due to their computational convenience. However, (Gelman et al., 2008). Highlight key limitations of this approach, noting that such priors may lead to unstable estimates or poor predictive performance, especially in the presence of complete separation. To address these challenges, they proposed Cauchy priors centered at zero with a scale parameter of 2.5 as a more robust alternative. These heavier-tailed priors improve model stability while mitigating the risks of overfitting.

Uniform Prior

In this study, uniform priors were adopted. (Löffler et al., 2011) examined Bayesian methods for improving credit scoring models, proposing a Bayesian framework to enhance default probability estimates for banks with small datasets. Their methodology imposed prior information on parameter estimates by utilizing uniform prior as a Type-1 weight vector, meaning that each prior coefficient had an equal proportional influence on the posterior estimates. They further assumed that the prior covariance matrix was equal to the sample covariance matrix scaled by a constant factor, thereby representing prior weights. Based on this setup, (Löffler et al., 2011) derived an empirical Bayes estimator using the marginal distribution of the estimator in the current sample, conditional on the prior.

To evaluate accuracy and the gains from Bayesian estimation, they conducted simulations on a dataset of non-financial firms. Model coefficients were estimated using Maximum Likelihood Estimation (MLE) and compared with Bayesian, empirical Bayes, and Stein-rule estimators. Performance was assessed using the accuracy ratio and prior score. Their findings indicated that all three Bayesian estimators significantly outperformed the standard logistic regression model. However, a key limitation of the study was that it relied on coefficients derived from non-financial firms, meaning that the credit scoring model did not exclusively incorporate financial data.

A significant study by He et al. (2021a) explored Bayesian inference under small-sample conditions using various non-informative priors. The authors proposed a methodology for constructing the Bayesian posterior when data availability is limited. Their research examined a range of prior distributions, including classical non-informative priors and asymptotically locally invariant priors, to enhance inference reliability. They utilized normal-inverse gamma conjugate priors to evaluate the posterior distribution and introduced a $1/q - type$ prior formulation. A key finding of their study was that the conventional Jeffrey's prior, proportional to $\frac{1}{\theta}$ yielded estimates that closely aligned with maximum likelihood estimates (MLE). Building on this foundation, this study adopted a Bayesian logistic regression model that incorporates both uniform and Jeffrey's' priors. The primary objective is to compare the predictive performance of these priors in a low-data environment, where traditional methods may struggle due to data scarcity. By systematically analyzing the model under different prior specifications, this study seeks to assess the robustness, efficiency, and accuracy of Bayesian inference in credit scoring when conventional, data-rich methodologies are impractical.

Bayesian Approaches in Credit Scoring

Several empirical studies have demonstrated the usefulness of Bayesian methods in credit scoring. For example, Hand & Henley (1997) highlighted the flexibility of Bayesian models in incorporating prior knowledge. More recently, He et al. (2021b) proposed two-stage Bayesian credit scoring models, showing improvements in prediction when integrating demographic and behavioral factors. Similarly, Zhou et al. (2022) applied Jeffrey's priors in high-dimensional logistic regression and demonstrated their superiority in managing multicollinearity and improving classification accuracy. These findings suggest that Bayesian approaches can significantly enhance credit scoring models, particularly in contexts where data is limited or noisy.

Model Predictive Power Comparisons

To evaluate the performance of the proposed credit scoring models, this study employed an established predictive metric, AUC, commonly used in related research. Area Under the Receiver Operating Characteristic Curve (AUC) (AUC) measures the discriminatory ability to distinguish between good and bad borrowers. Other metrics, like residual deviance, were used to assess the goodness-of-fit of each model. These metrics provided a robust framework for comparing models developed under different prior specifications, particularly in small sample environments where data limitations challenge traditional scoring methods.

Methodology

Study Design

The study utilizes a model that factors a binary output. Bayesian logistic regression was applied. The logistic regression model considers a parameter as a random variable, the data are considered fixed, and the parameters have their prior distribution. The study design selected individuals who are servicing a 36-month loan term who meet the criteria that the i^{th} individual has not defaulted in the first six months, then, by using the same set of data, was observed for the next preceding six months, and filtered the defaulters to enable the variable selection process.

Study Population, Sample, and Sampling Technique

The study employed a retrospective method of data collection where stratified random sampling method was used to select loan applicants from the bank's dataset of 4,926 individuals for loan applicants in 2021 just after the COVID-19 Pandemic in Coastal region branches. The sample size was arrived by taking 60% of loan type per the branch's loan book sufficiently representing the total loan book per branch. Borrowers were grouped into strata based on credit history, repayment habits, employment status and income levels. From each stratum, a proportional random sample of non-defaulters was selected to ensure a balanced dataset with defaulters. This technique was chosen to improve representativeness, minimize sampling bias, and capture variations in borrower characteristics across different loan schemes. Stratified sampling enhanced the precision of parameter estimates and supported reliable evaluation of credit scoring models using Bayesian and classical logistic regression approaches.

Data Collection Methods

Data was collected retrospectively from the bank's historical loan application and repayment records. The dataset included borrower demographic information, financial attributes, credit history, and loan default status. Records were extracted electronically, cleaned, and anonymized to ensure confidentiality. Each borrower's information was organized to include predictor variables and a binary outcome variable indicating default (1) or non-default (0). The data collection focused on completeness and accuracy, ensuring all relevant variables for logistic and Bayesian logistic regression modeling were captured.

Ethical Considerations and Data Protection

This study adhered to ethical principles in handling sensitive financial and personal data. An ethical permit was obtained from the National Commission for Science, Technology, and Innovation (NACOSTI), License No. NACOSTI/P/25/4180050. Approval was sought from the Scientific and Ethical Review Committee of the Technical University of Mombasa before accessing the loan records. All borrower information was anonymized to protect privacy, and identifiers such as names, account numbers, or contact details were removed. Data were stored securely on password-protected systems and accessed only by the research team. The study ensured confidentiality, avoided misuse of personal or financial information, and complied with institutional and national guidelines, guaranteeing that findings were reported objectively without compromising borrower rights or privacy.

Measurements of Variables

This study focused on variables like: Owner's capital (OC), Collateral value (CV), Credit amount (CA), Owner's age (OA), and Marital status (MS). In this study, the logit model had five independent variables and only one dependent variable in the form of a dummy (the results of credit analysis, i.e., approved = 1 and not approved = 0)

Logistic Regression

Classical logistic regression was used as the baseline model for credit scoring. The model estimates the probability that a borrower defaults on a loan based on a set of demographics, financial, and credit-history characteristics. Parameter estimation was performed using Maximum Likelihood Estimation (MLE). The fitted model was used to identify significant predictors of loan default and served as a benchmark for comparison with the Bayesian logistic regression models.

Bayesian Logistic Regression

Bayesian logistic regression extends the classical logistic regression framework by assigning prior distributions to the regression coefficients and updating these beliefs using observed data through Bayes' theorem: Posterior ∝ Likelihood × Prior, where the response variable follows a Bernoulli distribution. Regression coefficients were estimated using Hamiltonian Monte Carlo (HMC) sampling implemented through the brms package in R.

Bayesian Logistic Regression Using Normally Distributed Prior

To provide mild regularization while allowing the data to drive inference, regression coefficients were assigned weakly informative Normal priors:

$\beta_j \sim N(0, \sigma^2)$ where σ^2 represents a large variance. This specification assumes that coefficient values are centered around zero while permitting substantial variability. The approach improves computational stability and reduces the likelihood of extreme parameter estimates.

Uniform and Jeffrey's Priors

Two non-informative priors were considered to facilitate objective Bayesian inference.

For the Uniform prior:

$$\beta_i = p(\beta) = \begin{cases} \dfrac{1}{b-a} & a \le \beta_i \le b \\ 0 & otherwise \end{cases} \qquad \text{………………………………………………………………...……..(2)}$$

where a and b represent sufficiently wide bounds, ensuring minimal prior influence on posterior estimates. Jeffrey's prior is defined as:

$$\rho\left(y_i = 1/x_i, \beta\right) \sim Bernoulli = \frac{1}{1+e^{\left(-x_i^T \beta\right)}} \qquad \text{……………………………………………………………(3)}$$

Jeffrey's prior is invariant under reparameterization and is widely used as an objective prior because it reduces estimation bias and improves parameter stability. Uniform and Jeffrey's priors were evaluated alongside the weakly informative Normal before assessing their suitability for Bayesian credit-scoring models.

Simulation of the Posterior Distribution

To obtain the posterior estimate, MCMC in brm engine was used, where the brm translated the formula and priors into stan code Bürkner, (2017). The stan uses Hamilton Monte Carlo (HMC), which is a highly efficient MCMC algorithm. The sampler then explores the posterior distribution of the parameters β by simulating trajectories in the posterior landscape. After running the MCMC the posterior means of the Bayesian estimate $\hat{E}\left[\beta_j\right]$, Est. Error, l-95% CI, u-95% CI, and the trace plots and diagnostics Rhat, Bulk ESS were obtained to ascertain convergence of the Markov chains.

Results

Descriptive Statistics

The dataset used in this study comprised the loan performance records of 4,926 loan applicants. The dependent variable was loan default, which was recorded as a binary outcome indicating whether or not a borrower defaulted during the loan repayment period. Specifically, a value of 1 denoted a default event, while a value of 0 indicated non-default. In addition to the dependent variable, the dataset included sixteen independent variables capturing borrower characteristics. These were drawn from four broad categories: demographic, financial, credit history, and loan characteristics. Demographic variables consisted of the borrower's gender, age, and number of dependents. Financial factors included monthly income, monthly expenses, the amount of the loan applied for, and the credit bureau (CRB) rating. Credit history was represented by the number of major derogatory reports, the number of trade lines, the number of delinquent trade lines, the age of the oldest trade line expressed in months, and the number of recent credit enquiries. Finally, loan-specific characteristics included the purpose of the loan classified as house construction, medical purposes, or school fees and the scheme under which the borrower fell, which was categorized as check-off TSC, check-off corporate, or non-check-off.

Table 1: Summary of Statistics for the Numerical Input Variables

Variable	Mean	Median	SD	Min	Max	Q1	Q3
Age	33	32	8	21	49	27	39
Dependents	5	5	1	3	10	4	6
Expenses	30408	31456	759	10000	69880	21890	36988
Income	45628	44334	13724	5567	116780	33456	54334
Loan amount	460,095	479000	199111	69800	1450000	323900	576334
CRB rating	447	440	143	200	878	339	540
Major derogatory reports	0	0	1	0	6	0	0
Number of trade lines	21	21	8	1	56	15	23
Number of delinquent trade lines	1	0	2	0	9	0	2
Oldest trade lines in months	3	2	1	1	6	2	4
Number of recent credit enquiries	0	0	1	0	5	0	1
Loan amount Z- score	4.53768		1			-0.684	0.58379

Table 1 precents the descriptive statistics for the numeric variables. Regarding measures of central tendency, most variables show close values between the mean and median, suggesting a relatively symmetric distribution. However, for Loan Amount and Income, a notable gap between the mean and median indicates positive skewness in the data. Additionally, variables such as the number of major derogatory reports, recent credit enquiries, and the number of delinquent trade lines are highly concentrated around zero, reflecting sparse distributions. These variables may therefore require transformation to improve model stability and predictive accuracy.

For the credit dataset, exploratory data analysis (EDA) was conducted to provide an initial understanding of the structure of the data, reveal underlying patterns, and detect potential issues such as skewness, outliers, or class imbalance before model development. This step was essential to ensure that subsequent modeling decisions were based on a well-understood dataset. Out of the 4,926 customers whose loans were booked during the study period, 1,578 were classified as bad loans, representing a default rate of 32.03%. This indicates that nearly one-third of the loan portfolio fell into default, underscoring the importance of an effective credit scoring model.

Distribution of Loan Amount, Income, and Expenses by Scheme

Deeper data exploration was done to examine the data structure of loan amount, expenses, and income segmented by scheme, as shown in Figure 1.

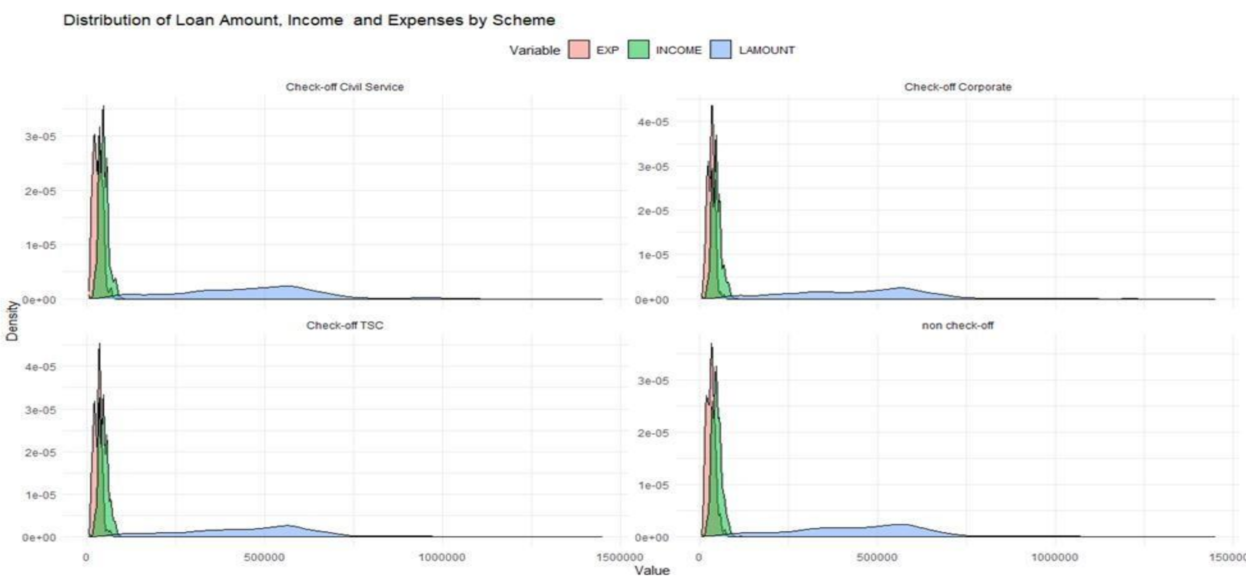

Figure 1: Distribution of loan amount, income, and expenses by scheme

Figure 1 precents the distribution of loan amounts, monthly income, and monthly expenses across the different scheme types. This visualization illustrates how scheme type is associated with the financial behavior of applicants. Loan amounts, represented in blue, display a positively skewed distribution across all schemes, with most applicants requesting relatively small amounts while a few outliers applied for significantly higher loans.

Monthly income, represented in green, also shows a positively skewed distribution, indicating that the majority of applicants fall within the lower income brackets, with only a small number earning substantially higher incomes. For applicants under the check-off civil service scheme, both income and expenses are concentrated at relatively low values, while loan amounts exhibit a wider spread, with some applicants requesting considerably higher amounts. By contrast, applicants under the check-off corporate scheme display a broader distribution of both income and expenses, though loan amounts remain skewed toward lower values, with occasional large loan requests.

Default Status by Variables

A box plot was used to graphically display the default status by variables, as given by Figure 2.

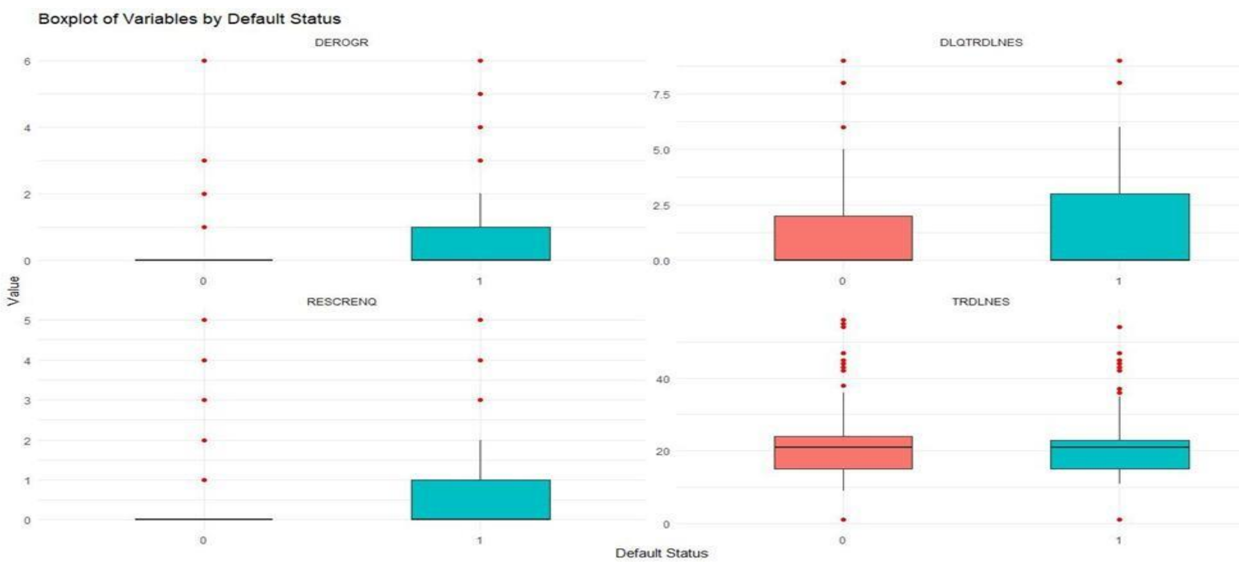

Figure 2: Box plot of variables by Default Status

The boxplot illustrates the distribution of the variables number of derogatory trade lines, number of delinquent trade lines, number of recent credit enquiries, and number of trade lines, grouped by loan default status. For derogatory trade lines, defaulters (coded as 1) exhibit greater variability compared to non-defaulters (coded as 0). While most non-defaulters have no derogatory records, defaulters are more likely to precent values above zero.

In the case of delinquent trade lines, defaulters show a wider range and higher median values relative to non-defaulters, who generally record fewer delinquencies. For recent credit enquiries, defaulters display higher variability and a greater number of outliers compared to non-defaulters, who report minimal credit enquiries. Similarly, the number of trade lines tends to be slightly higher among defaulters, although both groups precent some outliers. To evaluate the adequacy of the model, diagnostic checks were conducted focusing on multicollinearity among independent variables, as well as the presence of outliers and influential observations. The correlation matrix of the numerical independent variables is presented in Figure 2.

The correlation matrix of the numerical independent variables is given in Figure 3

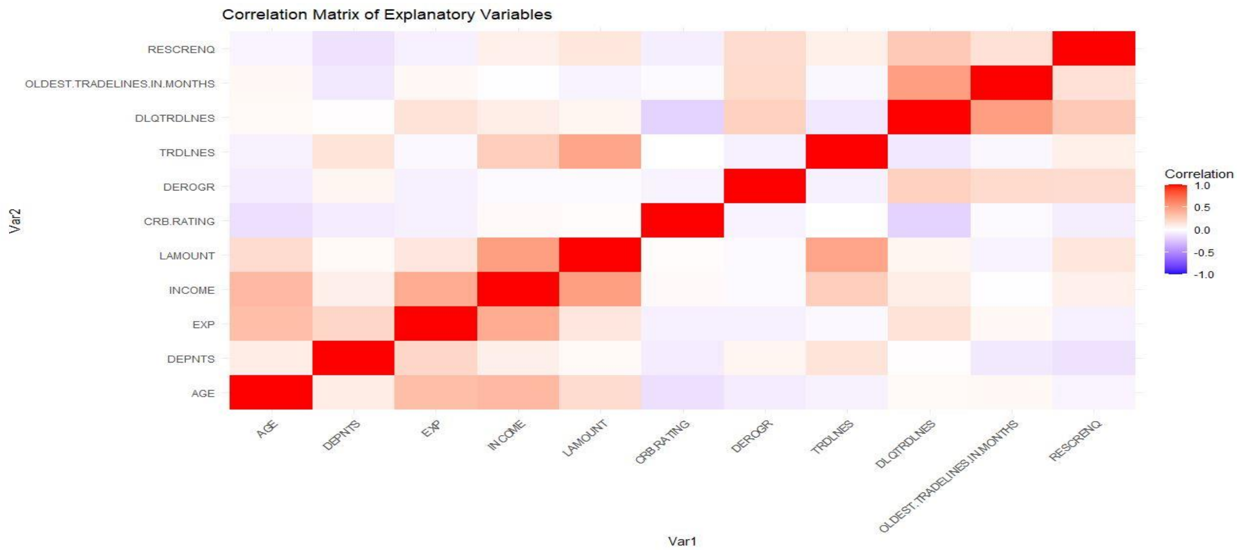

Figure 3: Heat map correlation matrix of numeric variables

The diagonal elements of the correlation matrix represent self-correlation, which is always equal to 1 and is shown in red. The correlation between loan amount and monthly income is relatively low at 0.165, suggesting that higher income does not necessarily translate to proportionately higher loan requests. Similarly, the age of the applicant and the number of dependents show a modest positive correlation, which can be attributed to the natural family size pattern that tends to increase with age. In contrast, variables such as derogatory reports exhibit little correlation with most other predictors. Importantly, the heatmap reveals no evidence of strong negative correlations, as indicated by the absence of dark blue regions. This suggests that none of the variables are highly inversely related.

To further assess potential collinearity among predictors, the Variance Inflation Factor (VIF) was computed. VIF quantifies the extent to which the variance of an estimated regression coefficient is inflated due to multicollinearity with other predictors, with values typically above 5 or 10 indicating potential issues. For categorical variables with more than two levels, the Generalized Variance Inflation Factor (GVIF) was applied. The GVIF results are presented in Table 3.

Table 2: Generalized Variance Inflation Factor (GVIF) of the Predictor Variables

Predictor V	GVIF	DF	GVIF1/(2*Df)
Age	1.0122	1	1.0061
Male	1.0184	1	1.0092
Number of dependents	1.0581	1	1.0286
Monthly expenses	1.0632	1	1.0369
Monthly income	1.0632	1	1.0311
Loan amount	1.054	1	1.0267
CRB rating	1.0172	1	1.0086
Derogatory lines	1.0271	1	1.0135
Number of trade lines	1.06	1	1.0296

Delinquent trade lines	1.0192	1	1.0095
Oldest trade lines	1.0026	1	1.0013
Recent credit enquiries	1.0052	1	1.0026
Loan purpose	1.0778	3	1.0125
Marital status	1.1085	3	1.0173

The results indicate that most predictors have GVIF values close to 1, and the adjusted GVIF values are also very close to 1. This clearly demonstrates that the predictors exhibit very low levels of multicollinearity. Therefore, based on the GVIF results, multicollinearity is not a concern for the model.

Logistic Regression Model

The CREDIT dataset was partitioned into training and testing subsets to facilitate model building and validation. The training subset was employed to estimate the model parameters, allowing the model to learn the underlying patterns and relationships in the data. The testing subset was reserved for evaluating the model's predictive performance on unseen data, thereby ensuring generalizability and reducing the risk of overfitting. The training subset comprised 3,449 observations and 16 predictor variables, while the testing subset consisted of 1,477 observations and the same 16 predictor variables. The estimated coefficients of the logistic regression model are presented in Table 3.

Table 3: Classical Logistic Regression Model Coefficient Estimates

Variable	Estimate	Std. Error	z-value	p-value	Significance
Intercept	-1.1991	0.211	-5.684	<0.001	Yes
Male	-0.1682	0.0922	-1.824	0.068	No
Age	-0.0476	0.0436	-1.094	0.274	No
Number of dependents	-0.4892	0.0467	-10.48	<0.001	Yes
Monthly expenses	1.0649	0.0524	20.332	<0.001	Yes
Monthly income	-0.0379	0.0433	-0.875	0.382	No
Loan amount	0.0248	0.0446	0.556	0.578	No
CRB rating	-1.1202	0.0521	-21.51	<0.001	Yes
Derogatory Credit Records	0.7835	0.0473	16.547	<0.001	Yes
Trade lines	-0.0888	0.0455	-1.953	0.051	No
Delinquent Trade Lines	0.114	0.042	2.712	0.007	Yes
Oldest trade lines in months	-0.0012	0.0425	-0.028	0.978	No
Recent Credit Enquiries	0.0673	0.0424	1.587	0.113	No
House Construction	-0.1138	0.1369	-0.831	0.406	No
Medical purposes	0.4057	0.149	2.722	0.006	Yes
School Fees	0.1376	0.15	0.917	0.359	No
Married	-0.1748	0.1306	-1.338	0.181	No
Single	0.2474	0.1337	1.85	0.064	No
Check-off Corporate	0.0696	0.1263	0.551	0.582	No
Check-off TSC	0.1297	0.1274	1.017	0.309	No
Non-check off	0.7932	0.155	5.117	<0.001	Yes

The intercept represents the baseline log-odds of default when all predictors are held constant. In this model, the estimated intercept is -1.1991, indicating a relatively low baseline probability of default when all explanatory variables are zero. Predictors with p-values less than 0.05 were considered statistically significant contributors to default risk. The number of dependents had a negative and statistically significant coefficient of – 0.4892 ($p < 0.001$), implying that applicants with more dependents were less likely to default. These results may be attributed to the increased financial responsibility associated with supporting more individuals, which potentially fosters greater financial discipline.

Monthly expenses emerged as a highly significant predictor with a positive coefficient of 1.0649 ($p < 0.001$). This suggests that higher monthly spending increases the probability of default, holding other factors constant. Elevated expenses may impose financial strain, thereby heightening the risk of repayment difficulties. The CRB rating exhibited a significant negative association with default, with a coefficient of – 1.1202 ($p < 0.001$). This implies that applicants with higher credit bureau ratings were less likely to default, which aligns with expectations, since strong credit ratings generally reflect better financial discipline and repayment history. Derogatory trade lines were also positively and significantly associated with default, with a coefficient of 0.7835 ($p < 0.001$). This indicates that applicants with more derogatory credit records had a higher likelihood of defaulting, reinforcing the importance of historical credit behavior in predicting repayment risk. On the other hand, variables such as loan amount, income, and age were not statistically significant in this model, suggesting that their individual predictive power was limited after accounting for the effects of other variables.

Bayesian Logistic Regression Models with Non-Informative Priors

To evaluate the robustness of the Bayesian approach, the logistic regression model was also fitted using non-informative priors, specifically the Uniform prior. Non-informative priors are designed to exert minimal influence on the posterior distribution, allowing the observed data to play the dominant role in shaping parameter estimates. For the Uniform prior, a distribution over the range $[-10,10]$ was specified for each regression coefficient. Within this interval, all values are assigned equal probability, while values outside are considered impossible. The choice of this range ensures that the true parameter values are unlikely to lie near the boundaries, thereby preventing undue constraint of the posterior distribution. Uniform prior allows the model to rely primarily on the data for inference, thereby offering an objective benchmark against the Bayesian model fitted with the Normal prior.

Table 4: Bayesian Logistic Regression Model Using Uniform Prior

Variable	Estimate	Est. Error	L-95% CI	U-95% CI	Rhat	Bulk_ESS	Tail_ESS	Significance
Intercept	-1.2	0.21	-1.62	-0.79	1	1781	2005	Yes
Male	-0.17	0.09	-0.35	0.01	1	4699	2826	No
Age	-0.05	0.04	-0.13	0.04	1	6907	2987	No
Number of dependents	-0.49	0.05	-0.58	-0.4	1	5250	3575	No
Monthly expenses	1.07	0.05	0.97	1.18	1	4728	2839	Yes
Monthly income	-0.04	0.04	-0.12	0.05	1	5968	3165	No
Loan amount	0.02	0.05	-0.07	0.11	1	6110	3116	No

CRB rating	-1.13	0.05	-1.23	-1.03	1	6466	3348	Yes
Derogatory Credit Records	0.79	0.05	0.7	0.88	1	6958	3274	Yes
Trade lines	-0.09	0.05	-0.18	0	1	5681	2680	Yes
Delinquent Trade Lines	0.11	0.04	0.04	0.2	1	5929	3033	Yes
Oldest trade lines in months	0	0.04	-0.08	0.08	1	6289	2721	No
Recent Credit Enquiries	0.07	0.04	-0.02	0.15	1	5681	2718	No
House Construction	-0.11	0.13	-0.38	0.15	1	1750	2525	No
Medical purposes	0.41	0.15	0.12	0.69	1	1621	2114	Yes
School Fees	0.14	0.14	-0.15	0.42	1	1760	2139	No
Married	-0.18	0.13	-0.43	0.08	1	2074	2177	No
Single	0.25	0.13	0	0.52	1	2100	2449	Yes
Check-off Corporate	0.07	0.12	-0.18	0.31	1	2162	2516	No
Check-off TSC	0.13	0.13	-0.11	0.39	1	2269	2427	No
Non-check-off	0.8	0.15	0.5	1.1	1	2337	2402	Yes

In the Bayesian logistic regression model fitted with a Uniform prior, statistical significance was assessed by examining whether the 95% credible intervals excluded zero. Several predictors were found to significantly influence the probability of default. The number of dependents had an estimate of –0.49 (95% CI: –0.58, –0.40), indicating that applicants with more dependents were less likely to default, possibly due to greater financial responsibility. Monthly expenses were strongly associated with higher default risk, with an estimate of 1.07 (95% CI: 0.97, 1.18), suggesting that heavier financial obligations increase vulnerability to default. Similarly, Credit Bureau Rating had a negative and significant effect (estimate –1.13, 95% CI: –1.23, –1.03), confirming that applicants with stronger credit histories are less likely to default.

Credit history variables also showed strong effects. Derogatory trade lines were positively associated with default (estimate 0.79, 95% CI: 0.70, 0.88), emphasizing the predictive importance of past negative credit records. Conversely, the number of trade lines was negatively associated with default (estimate –0.09, 95% CI: –0.18, –0.00), suggesting that broader credit experience reduces default risk. Delinquent trade lines were positively associated with default (estimate 0.11, 95% CI: 0.04, 0.20), reflecting the impact of recent repayment difficulties. Among categorical predictors, loan purpose for medical needs was significant (estimate 0.41, 95% CI: 0.12, 0.69), implying that applicants borrowing for medical reasons carry a higher risk of default compared to other purposes. Likewise, applicants under the non-check-off scheme exhibited greater default risk (estimate 0.80, 95% CI: 0.50, 1.10), consistent with the reduced repayment assurance in non-payroll-deducted loans. Overall, the results under the Uniform prior closely align with findings from the Normal prior model, reinforcing the robustness of key predictors such as expenses, credit rating, and derogatory records in explaining loan default behavior.

Comparative Performance of the Credit Scoring Models

The performance of the Bayesian logistic regression models was evaluated using the Area Under the Receiver Operating Characteristic Curve (AUC), confusion matrices, accuracy, sensitivity, specificity, and balanced accuracy. The AUC is particularly important in credit scoring because it measures the model's ability to correctly distinguish between defaulters and non-defaulters across all possible classification thresholds. An AUC value of 0.5 indicates no discriminatory ability, values between 0.7 and 0.8 indicate acceptable discrimination, values between 0.8 and 0.9 indicate excellent discrimination, while values above 0.9 indicate outstanding predictive performance.

In this study, all Bayesian models produced AUC values greater than 0.86, indicating excellent discriminatory ability. This means that for any randomly selected pair of borrowers consisting of one defaulter and one non-defaulter, the models have approximately an 86% probability of assigning a higher risk score to the defaulter than to the non-defaulter. The similarity of the AUC values further suggests that all three Bayesian models are capable of effectively separating high-risk borrowers from low-risk borrowers and can therefore be used as reliable credit scoring tools.

Table 5: Model Performances

Model	Residual Deviance	AUC
Classical Logistic	3415.048	0.8635
Bayesian Normal Prior	3457.805	0.8639
Bayesian Uniform Prior	3457.450	0.8642
Bayesian Jeffrey's Prior	3458.603	0.8636

Bayesian Logistic Regression under the Normal Prior

All three Bayesian priors yielded an AUC value above 0.86, which reinforces the conclusion that Bayesian models under non-informative priors offer reliable classification performance. The AUC value demonstrates excellent discrimination between defaulters and non-defaulters, indicating that the model can reliably rank borrowers according to their level of credit risk. The balanced accuracy further suggests that the model performs consistently in identifying both good and bad borrowers.

The Normal prior introduces moderate regularization by shrinking coefficient estimates toward zero, resulting in stable parameter estimates and reducing the risk of overfitting. Consequently, this model is suitable for banks seeking a balanced credit scoring framework that supports both portfolio growth and effective risk management. Institutions aiming to maintain credit expansion while preserving portfolio quality may therefore prefer the Normal prior model.

Bayesian Logistic Regression under the Uniform Prior

The Uniform prior model recorded the highest AUC of approximately 0.8642 among all Bayesian models. Although the difference is small, this result suggests that the Uniform prior model has the strongest ability to distinguish between borrowers who will default and those who will not. A higher AUC is particularly valuable in credit scoring because it improves borrower ranking and risk segmentation.

The confusion matrix results further indicate that the model maintains strong classification performance while relying primarily on information contained in the observed data. Therefore, this model may be most suitable for banks that prioritize risk ranking, credit scorecard development, risk-based pricing, and customer segmentation. Such institutions may benefit from the model's slightly superior discriminatory power when making lending decisions.

Bayesian Logistic Regression under Jeffrey's Prior

The Jeffrey's prior model achieved an AUC of approximately 0.8636 and a balanced accuracy of 0.8426. Although its AUC is marginally lower than that of the Uniform prior model, it remains within excellent range and demonstrates strong discriminatory performance. The model's balanced accuracy was among the highest, indicating an effective balance between correctly identifying defaulters and correctly identifying non-defaulters.

Jeffrey's prior is theoretically appealing because it is objective and invariant under parameter transformation. The confusion matrix results suggest that the model provides a stable balance between sensitivity and specificity, making it appropriate for institutions seeking cautious and objective credit risk assessment. Conservative lenders, microfinance institutions, and banks operating in uncertain environments may therefore find Jeffrey's prior model particularly useful.

Confusion Matrices

The confusion matrix provides additional insight into model suitability for credit scoring because it reveals the trade-off between correctly identifying defaulters (sensitivity) and correctly identifying non-defaulters (specificity).

Table 6: Performance metric for the Bayesian models

Metric	Normal Prior	Jeffrey's Prior	Uniform Prior
Accuracy	84.26%	84.37%	84.26%
Precision (PPV)	78.01%	78.09%	78.21%
Recall (Sensitivity)	90.77%	90.77%	90.92%
F1-Score	83.83%	83.89%	84.01%
Specificity	70.29%	70.61%	69.97%
False Positive Rate (FPR)	29.71%	29.39%	30.03%
False Negative Rate (FNR)	9.23%	9.23%	9.08%

Precision reflects the proportion of the predicted defaulters who actually defaulted and is highest in the Uniform prior model at 78.21%, followed by Jeffrey's and Normal prior models at 78.09% and 78.01%, respectively. F1- Score balances precision and recall, is highest in Uniform at 84.01%, followed by Jeffrey's and Normal priors at 83.89% and 83.83%, respectively. The FPR of non-defaulters classified as defaulters is lowest in Jefferey's at 29.39%.

Discussion of Results

Results showed that variables such as gender, number of dependents, monthly expenses, CRB rating, number of derogatory reports, loan purpose, marital status, and scheme check-off have a significant influence on the borrower's ability to repay a loan. These findings align with the research conducted by Kenyatta et al. (2012), who demonstrated that gender, age, marital status, occupation, and loan term significantly influence loan repayment ability in Kenya. Similarly, the results corroborate the study by

Daniel K et al. (2019), which found that marital status and number of dependents significantly affect a borrower's likelihood of repaying a loan.

Under the uniform prior, significant variables include gender, age, number of dependents, monthly expenses, CRB rating, number of derogatory reports, delinquent trade lines, loan purpose, marital status, scheme check-off, and oldest trade lines in months. These findings are consistent with those reported by Platur Gashi in their master's Thesis on explainable loan default prediction using machine learning, which identified credit rating (including derogatory and delinquent trade lines), loan rate, loan term, and loan amount as significant factors influencing loan repayment. The performance of the models was evaluated using the Area Under the Receiver Operating Characteristic Curve (AUC) (AUC). The Bayesian logistic regression model under a uniform prior demonstrated strong performance, achieving the highest AUC of 0.8642.

Conclusion

The findings demonstrate that all Bayesian logistic regression models possess excellent predictive performance, with AUC values exceeding 0.86 and balanced accuracy values above 0.80. The Uniform prior model provides the strongest discriminatory ability as reflected by its slightly higher AUC, making it attractive for borrower ranking and risk segmentation. The Normal prior model offers stable and balanced predictions suitable for institutions pursuing both growth and risk control. The Jeffrey's prior model provides objective inference and strong, balanced classification performance, making it particularly useful for conservative credit risk management strategies.

Therefore, all three Bayesian models are suitable for credit scoring applications, with the preferred model depending on the specific strategic objectives and risk appetite of the lending institution.

The findings may also inform credit risk management policies by demonstrating that Bayesian logistic regression models based on non-informative priors can provide robust and interpretable predictions without relying heavily on subjective assumptions. Such approaches may be valuable for institutions seeking to expand lending while maintaining prudent risk management practices.

References

Ayari, H., Guetari, P. R., & Kraïem, P. N. (2026). Machine learning powered financial credit scoring: A systematic literature review. *Artificial Intelligence Review, 59*(1). https://doi.org/10.1007/s10462-025-11416-2

Barasa, I. N., Wanyonyi, S. W., & Kololi, M. M. (2025). Application of logistic regression in enhancing digital credit risk management in commercial banks. *Asian Journal of Probability and Statistics, 27*(2), 13–26. https://doi.org/10.9734/AJPAS/2025/v27i2710

Bernardo, J. M., & Smith, A. F. M. (2000). *Bayesian theory*. John Wiley & Sons.

Bürkner, P.-C. (2017). brms: An R package for Bayesian multilevel models using Stan. *Journal of Statistical Software, 80*(1), 1–28. https://doi.org/10.18637/jss.v080.i01

Daniel K., B., F., O., J., N., & Samuel F., P. (2019). Performance evaluation criteria of credit scoring models for commercial lenders. *International Journal of Mathematics Trends and Technology, 65*(7), 23–30. https://doi.org/10.14445/22315373/IJMTT-V65I7P505

Del Signore, S., & Betancourt, M. (2022). Improving Bayesian credit risk modeling through hierarchical priors. *Journal of Applied Econometrics, 37*(5), 854–873. https://doi.org/10.1002/jae.2941

Gashi, P. (2023). *Loan default prediction model* [Master's thesis, Riinvest College]. Scribd. https://www.scribd.com/document/953349408/PlaturGashi-MasterThesis

Gelman, A., Jakulin, A., Pittau, M. G., & Su, Y.-S. (2008). A weakly informative default prior distribution for logistic and other regression models. *The Annals of Applied Statistics, 2*(4), 1360–1383. https://doi.org/10.1214/08-AOAS191

He, J., Wang, W., Huang, M., Wang, S., & Guan, X. (2021). Bayesian inference under small sample sizes using general noninformative priors. *Mathematics, 9*(21), Article 2810. https://doi.org/10.3390/math9212810

Hitchcock, C. (2024). Bayesian logistic regression in applied statistics: A modern pedagogical approach. *Statistics Education Review, 12*(2), 98–112.

Kenyatta, J., Adem, O., Gichuhi, A. W., & Otieno, R. O. (2023). Parametric modeling of the probability of bank loan default in Kenya. *Journal of Banking and Finance Studies, 14*(1), 45–58.

Kosmidis, I., & Firth, D. (2021). Jeffrey's prior penalty, finiteness, and shrinkage in binomial-response generalized linear models. *Biometrika, 108*(1), 71–82. https://doi.org/10.1093/biomet/asaa052

Kyeong, S., & Shin, J. (2022). Two-stage credit scoring using a Bayesian approach. *Journal of Big Data, 9*(1), 1–15. https://doi.org/10.1186/s40537-022-00665-5

Löffler, G., Posch, P. N., & Schone, C. (2011). Bayesian methods for improving credit scoring models. *SSRN Electronic Journal*. https://doi.org/10.2139/ssrn.742469

Ndiritu, S., & Waithaka, S. (2022). Bayesian credit scoring models for low-default portfolios: A Kenyan banking perspective. *African Journal of Financial Risk Analysis, 5*(2), 101–117.

Ntzoufras, I. (2011). *Bayesian modeling using WinBUGS*. John Wiley & Sons.

Okesola, O. J., Okokpujie, K. O., Adewale, A. A., John, S. N., & Omoruyi, O. (2018). An improved bank credit scoring model: A naïve Bayesian approach. In *Proceedings of the 2017 International Conference on Computational Science and Computational Intelligence (CSCI 2017)* (pp. 228–233). https://doi.org/10.1109/CSCI.2017.36

Oliveira, G., Bassinello, P. Z., & Rice, E. (2017). Comparing priors in Bayesian logistic regression for sensorial classification of rice. *Brazilian Journal of Applied Statistics, 12*(2), 1–10.

Olugbo, E., Obini, N. N., Olugbo, E. J., Adenomon, M. O., & Nweze, N. O. (2024). Comparison of Bayesian and classical approaches of logistic regression in modeling risk of preterm birth in Nasarawa

State, Nigeria. *BIMA Journal of Science and Technology, 8*(1B), 2536–6041. https://doi.org/10.56892/bima.v8i1B.648

Omondi, T., & Muthama, P. (2024). Advancing credit scoring accuracy through Bayesian hierarchical logistic regression in developing economies. *African Journal of Data Science and Analytics, 4*(1), 34–52.

Scricciolo, C. (1999). Probability matching priors: A review. *Journal of the Italian Statistical Society, 8*(1), 83–100. https://doi.org/10.1007/BF03178943

Szepannek, G. (n.d.). *Scorecard modelling using R*. Retrieved from https://cran.r-project.org/web/packages/scorecardModelR

Tahir, M., Syversveen, A. R., Abid, L., Zaghdene, S., Masmoudi, A., & Ghorbel, S. Z. (2017). Bayesian network modeling: A case study of credit scoring analysis of consumer loan default payment. *Asian Economic and Financial Review, 7*(12), 846–857. https://doi.org/10.18488/journal.aefr.2017.79.846.857

Tasche, D. (2011). Bayesian estimation of probabilities of default for low default portfolios. *Journal of Risk Management in Financial Institutions, 6*(3), 302–317.

Tierney, L. (1994). Markov chains for exploring posterior distributions. *The Annals of Statistics, 22*(4), 1701–1728. https://doi.org/10.1214/aos/117632575

Turkkan, N., & Pham-Gia, N. (1993). Computation of the highest posterior density interval in Bayesian analysis. *Journal of Statistical Computation and Simulation, 44*(3–4), 243–250. https://doi.org/10.1080/00949659308811461

Vidal, M. F., & Barbon, F. (2019). *Credit scoring in financial inclusion* (CGAP Research Paper). CGAP. https://www.cgap.org/research/publication/credit-scoring-financial-inclusion

Webster, G. (2011). *Bayesian logistic regression models for credit scoring* [Doctoral dissertation, University of Nottingham].

Zhang, X., Zhang, T., Hou, L., Liu, X., Guo, Z., Tian, Y., & Liu, Y. (2025). Data-driven loan default prediction: A machine learning approach for enhancing business process management. *Systems, 13*(7). https://doi.org/10.3390/systems13070581

Zhou, J., & Li, X. (2023). Bayesian regularization in credit scoring using hierarchical non-informative priors. *Journal of Quantitative Financial Analytics, 5*(2), 99–113. https://doi.org/10.2139/ssrn.4592219