International Journal of Research in Commerce and Management Studies

ISSN 2582-2292

Vol. 8, No. 05 Sep-Oct; 2026 Page. No. 80-87

To cite this article: Mandadapu Rajani, Midde Venkata Krishna Varaprasad, Chadarasarasipalli Venkata Krishna, Papa Rao Maddala and Ranga Ngendra Babu (2026). MACHINE LEARNING ALGORITHM- LOGIT REGRESSION FOR CREDIT RISK ASSESSMENT, International Journal of Research in Commerce and Management Studies (IJRCMS) 8 (5): 80-87 Article No. 864 Sub Id 1429

MACHINE LEARNING ALGORITHM- LOGIT REGRESSION FOR CREDIT RISK ASSESSMENT

Mandadapu Rajani[1], Midde Venkata Krishna Varaprasad[2], Chadarasarasipalli Venkata Krishna[3], Papa Rao Maddala[4] and Ranga Ngendra Babu[5]

[1]25MG1E0026, PG Scholar- Sree Vahini Institute of Science and Technology (Autonomous) Tiruvuru. AP -521235.
[2]25MG1E0061, PG Scholar - Sree Vahini Institute of Science and Technology (Autonomous) Tiruvuru. AP -521235.
[3]Assistant Professor Department of MBA, Sree Vahini Institute of Science and Technology (Autonomous), Tiruvuru. AP -521235.
[4]Professors and HoD of Department of MBA, Sree Vahini Institute of Science and Technology (Autonomous), Tiruvuru. AP -521235.
[5]Principal Sree Vahini Institute of Science and Technology (Autonomous), Tiruvuru. Andhra Pradesh (AP) – 521235

DOI: https://doi.org/10.38193/IJRCMS.2026.8506

ABSTRACT
With the use of Logistic Regression, this study explores the process of credit risk evaluation. The core data for this study comes from ninety individuals who applied for loans at three different private banks in Hyderabad. The rationale of this cram is to uncover foremost drivers of loan default, such as the amount of income, job status, credit history, and debt-to-income ratio at the financial institution. The information is gathered through the use of a structured questionnaire, and then it is analyzed by Logit Regression in order to predict the likelihood of missing payments. The findings offer important insights into the predictive value of financial and demographic characteristics in the progression of appraise credit risk at the individual level. This work makes a contribution to the improvement of data-driven lending choices, which assists financial institutions in optimizing loan approvals while simultaneously reducing the risks of default.

KEYWORDS: Credit Risk Assessment, Logistic Regression, Loan Default Prediction, Financial Risk Management

INTRODUCTION
The evaluation of credit risk is an essential part of the process of making financial decisions in the banking industry. This is because it establishes the probability that borrowers would breach their loan obligations. As the level of competition among private banks continues to rise and the number of non-performing assets (NPAs) continues to rise, it is more important than ever to conduct proper risk assessments in order to preserve financial stability. Previously, financial institutions have depended on credit scoring models that were based on rules. However, as data-driven analytics has advanced,

International Journal of Research in Commerce and Management Studies

ISSN 2582-2292

Vol. 8, No. 05 Sep-Oct; 2026 Page. No. 80-87

statistical methods such as logistic regression have emerged as efficient tools for forecasting the likelihood of a loan default. In order to determine the distinctiveness that have the most momentous impact on loan repayment behavior, the purpose of this study is to employ Logit Regression to appraise the credit risk of loan applicants from three private banks in Hyderabad.

Background of the Study

Through the provision of credit facilities, the banking sector performs an essential function in the expansion of the economy by offering financial assistance to both individuals and enterprises. However, the risk that is connected with lending, particularly credit risk, presents a considerable challenge to the organizations that deal with financial matters. The term "credit risk" refers to the possible loss that a lender could incur in the event that a borrower fails to fulfill their commitments to repay the loan. Banks employ credit assessment models that conduct an analysis of a variety of financial and demographic characteristics in order to reduce the risk that they face.

Over the course of the past several years, machine learning and statistical approaches have become more popular options for improving the accuracy of risk prediction models. Logistic Regression, often known as the Logit Model, is one of these models that have seen widespread use because of its capacity to approximation the risk of defaulting on borrower factors such as income level, job status, debt-to-income ratio, and credit record. Unlike the majority of studies, which have concentrated on secondary data obtained from credit bureaus, this investigation makes use of original data gathered from ninety individuals who have applied for loans in Hyderabad. As a result, it provides real-time insights into credit risk variables that are particular to the banking environment in Hyderabad. The results of this study will help to the improvement of credit evaluation procedures, which will enable financial institutions to make more informed lending decisions and minimize the percentage of default incidents.

REVIEW OF LITERATURE

Ismawati, I. Y., and Faturohman, T. (2023) developed a credit risk score model by employing logistic regression and basing it on customer data obtained from an Indonesian financial institution. A total of nine significant variables of default likelihood were revealed by the model. These predictors include the type of employment, the amount of work experience, the net financial value, and income. In order to assess the likelihood of consumer loans going into default, Costa e Silva, (2022) applied a logistic regression model using credit score data. Based on the findings, it was discovered to facilitate the risk of non-payment rises among the loan spread, the loan length, and the age of the consumer, but it falls when the customer has a greater number of credit cards. A level of accuracy of 89.79% was attained by the model.

Wang, C., and Xu, Q. (2024) investigated the use of approaches including artificial intelligence and

ISSN 2582-2292

Vol. 8, No. 05 Sep-Oct; 2026 Page. No. 80-87

logistic regression in the process of forecasting credit defaults. Based on the findings of the study, it was discovered that logistic regression performed 15.49% better than K-nearest neighbors (KNN) in terms of accuracy, exhibiting superior efficiency in the processing of big datasets for credit evaluation. An investigation on the relative effectiveness of logistic regression, in forecasting financial distress within the Indian banking industry was carried out by Mishra, N., Ashok, S., and Tandon, D. (2024). According to the findings of the study, logistic regression is an excellent method for determining the most important markers of financial hardship. The year 2000's Thomas, L. C. Thomas provides an overview of the development of credit scoring and places an emphasis on the use of statistical approaches, such as logistic regression, in the process of forecasting the credit risk of consumers.

A study was conducted in 2003 by Baesens, B., Vanthienen, J. the logistic regression is compared to alternative classification algorithms for credit scoring. The findings of this comparison provide evidence that logistic regression continues to be a reliable and interpretable technique.

Hand, D. J., and Henley, W. E. 1997, the authors present a detailed assessment of statistical categorization approaches for consumer credit scoring, with a particular emphasis on the widespread application of logistic regression.

In 2007, Crook, J. N., published their findings. This article examines credit scoring and its many uses, pointing out that logistic regression is a common method due to the fact that it is both straightforward and efficient on the whole. a. 2011 publication by Abdou, H. A., and Pointon, J. In the course of their investigation of credit scoring models in developing countries, the authors discover that logistic regression is capable of providing accurate forecasts of creditworthiness.

The research conducted by Lessmann, S., and Thomas, L. C. (2015) examines and contrasts a number of different categorization methods for credit scoring. The findings of this study reinforce the robust performance of logistic regression models.

RESEARCH METHODOLOGY

In order to evaluate the credit risk of individuals who are applying for loans, this study makes use of a quantitative research methodology and applies logistic regression analysis. The core data comes from the responses of ninety individuals who applied for loans at three different private banks in Hyderabad. For the purpose of gathering information on applicant demographics, financial behavior, job data, and credit history, a structured questionnaire is utilized. Key characteristics that serve as predictors of loan default risk are the subject of this study. These variables include employment type, debt-to-income ratio, income level, and credit utilization. Additionally, the study examines the relationship between debt and income.

A strategy known as purposive sampling is utilized in the sampling procedure, which guarantees the

International Journal of Research in Commerce and Management Studies

ISSN 2582-2292

Vol. 8, No. 05 Sep-Oct; 2026 Page. No. 80-87

inclusion of loan applicants that come from a variety of different financial backgrounds. Logistic regression, a statistical approach that is frequently used for predicting categorical outcomes, especially the risk of loan default (default = 1, no default = 0), is utilized in the analysis of the data. The reason that this technique was selected is because of its capacity to deal with binary-dependent variables and to deliver findings that can be interpreted in terms of the influence that each predictor has on credit risk.

A preliminary test of the questionnaire is carried out, and a check for multicollinearity among the independent variables is carried out. With the aid of these results, financial institutions will be able to improve their credit assessment models, which will in turn improve their risk management methods and decision-making processes for loan approvals.

DATA ANALYSIS- Demographic Analysis

For the purpose of gaining a better understanding of the profile of borrowers who apply for credit in private banks in Hyderabad, the demographic features of the ninety people who applied for loans were evaluated. Some of the most important factors that are taken into consideration include age, gender, income level, kind of occupation, and educational qualification. The outcomes are detailed in the table that can be seen below

Table 1: Demographics

Demographic Variable	Categories	Frequency (N=90)	Percentage (%)
Gender	Male	58	64.4%
	Female	32	35.6%
Age Group	20 - 30 years	25	27.8%
	31 - 40 years	40	44.4%
	41 - 50 years	15	16.7%
	Above 50 years	10	11.1%
Employment Type	Salaried	55	61.1%
	Self-Employed	35	38.9%
Income Level (Monthly)	Below ₹30,000	30	33.3%
	₹30,000 - ₹50,000	35	38.9%
	Above ₹50,000	25	27.8%
Education	High School	20	22.2%
	Graduate	50	55.6%

Demographic Variable	Categories	Frequency (N=90)	Percentage (%)
	Postgraduate	20	22.2%

Interpretation

According to the findings of the demographic research, 64.4% of people who apply for loans are male. This indicates that there is a greater number of male borrowers than there are female borrowers. The bulk of applicants, which accounts for 44.4% of the total, are between the ages of 31 and 40 years old, which suggests that those who are in the middle of their careers are more inclined to pursue loan applications. 61.1% of the respondents are salaried employees, while 38.9% are self-employed, which reflects a wide mix of borrowers. In terms of employment type, the respondents are employed in a variety of ways. Regarding the income levels, it is worth noting that 38.9% of the applicants have a monthly income that falls between ₹30,000 and ₹50,000, while 33.3% of them have a monthly income that is below ₹30,000. This suggests that a major number of the applicants belong to the middle-income category.

There is also a correlation between educational credentials and loan applications, with 55.6% of respondents having completed their undergraduate degrees, followed by 22.2% having completed their postgraduate degrees. Based on this, it appears that those who have completed higher levels of education are more inclined to look for formal credit facilities.

Hypothesis Testing Using Logistic Regression

H_1: An applicant's financial and demographic characteristics significantly influence the probability of loan default.

A logistic regression model is used to test the hypothesis with loan default (1 = Default, 0 = No Default) as the dependent variable and predictor variables including:

- Income level
- Employment type
- Credit history
- Debt-to-income ratio
- Age

International Journal of Research in Commerce and Management Studies

ISSN 2582-2292

Vol. 8, No. 05 Sep-Oct; 2026 Page. No. 80-87

Table 2: Results of Logistic Regression Analysis

Variable	Coefficient (β)	Standard Error	Wald Statistic	p-value	Odds Ratio (Exp(β))
Income Level	-0.752	0.312	5.80	0.016**	0.471
Employment Type	-0.635	0.285	4.95	0.026**	0.530
Credit History	1.245	0.375	11.01	0.001***	3.472
Debt-to-Income Ratio	0.897	0.295	9.23	0.002***	2.452
Age	0.120	0.189	0.41	0.520	1.127
Constant	-1.980	0.890	4.95	0.026**	0.138

Significance levels: **$p < 0.05$ (**)**, $p < 0.01$ (*), $p < 0.001$ (**)

Income Level ($p = 0.016$): The odds ratio for the chance of loan default is 0.471, which indicates that an increase in income decreases the possibility of default by 52.9%. This correlation is significant since a higher income level considerably reduces the probability of loan default. Employment Type ($p = 0.026$): With an odds ratio of 0.530, it can be shown that persons who are salaried have a lower probability of defaulting on their payments when compared to those who are self-employed. A bad credit history considerably increases the chance of default, with an odds ratio of 3.472, which means that applicants with such a past are 3.47 times more likely to fail on their loans. This is a statistically significant finding ($p = 0.001$). The debt-to-income ratio ($p = 0.002$) indicates that higher debt burdens are associated with an increased chance of default. The odds ratio for default is 2.452, which indicates that persons who have large debt loads are 2.45 times more likely to default on their loans. According to the statistical analysis, the variable is not statistically significant, which indicates that age does not have a major influence on the chance of defaulting on a loan.

CONCLUSION

A total of ninety individuals who applied for loans from three different private banks in Hyderabad participated in this study, which utilized logistic regression to investigate the influence of financial and demographic characteristics on the likelihood of loan default. A substantial number of factors, including income level, work type, credit history, and debt-to-income ratio, have been found to be important predictors of failed loan repayment. A person's credit history and the ratio of their debt to their income are the two factors that have the most significant impact. This suggests that applicants who have bad credit histories and heavier financial obligations are more likely to fail on their loans. On the other hand, a larger income and stable work both tend to minimize the chance of default. It is clear from these observations that risk assessment models are extremely important in the banking

International Journal of Research in Commerce and Management Studies

ISSN 2582-2292

Vol. 8, No. 05 Sep-Oct; 2026 Page. No. 80-87

industry.

DISCUSSION

The findings are consistent with the current body of research, which provides more evidence that logistic regression is an excellent method for evaluating credit risk. According to research conducted by Crook et al. (2007) and Costa e Silva et al. (2022), which emphasizes previous financial behavior as a crucial indicator of creditworthiness, the substantial influence that credit history plays in forecasting defaults is consistent with the findings of these organizations. Specifically, Mishra et al. (2024) provide evidence for the hypothesis that larger debt-to-income ratios increase the probability of default. This finding highlights the significance of debt management in the process of loan approvals. Nevertheless, the fact that age had no significant influence shows that demographic factors alone could not be a reliable indicator of default; rather, financial conduct continues to be the most important factor in determining default.

IMPLICATIONS

A number of practical consequences may be drawn from the study for both policymakers and financial organizations. When it comes to reviewing loan applications, financial institutions have to give credit history and debt-to-income ratios the highest priority in order to improve risk management techniques. Borrowers can be assisted in better managing their debt and default rates can be reduced through the implementation of more stringent credit screening and financial literacy programs. In addition, individualized loan solutions that are based on the kind of job and the quantity of income can enhance access to credit while simultaneously reducing related risks. These findings may be utilized by policymakers in order to modify policies on financial inclusion. This will ensure responsible lending practices that do not compromise financial stability while also supporting economic growth.

Taking everything into consideration, this study accentuates the consequence of making decisions in credit risk management based on data and draws attention to the requirement for ongoing development in credit assessment models in order to strengthen the resilience of the financial sector.

REFERENCES

1. Ismawati, I. Y., & Faturohman, T. (2023). Developed a credit risk scoring model using logistic regression based on customer data from an Indonesian financial institution. *Emerald Insight*.
2. Costa e Silva, E., Lopes, I. C., Correia, A., & Faria, S. (2022). Applied a logistic regression model to credit scoring data from a Portuguese financial institution to evaluate consumer loan default risk. *International Journal of Financial Studies*.
3. Wang, C., & Xu, Q. (2024). Examined the application of logistic regression and artificial intelligence techniques in predicting credit defaults. *Journal of Computational Social Sciences and Analytics*.

4. Mishra, N., Ashok, S., & Tandon, D. (2024). Conducted a comparative study between logistic regression, linear discriminant analysis (LDA), and artificial neural networks (ANN) in predicting financial distress within the Indian banking sector. *SAGE Journals*.
5. Abdou, H. A., & Pointon, J. (2011). Credit scoring, statistical techniques, and evaluation criteria: A review. *Intelligent Systems in Accounting, Finance and Management, 18*(2-3), 59-88. https://doi.org/10.1002/isaf.325
6. Baesens, B., Van Gestel, T., Viaene, S., Stepanova, M., Suykens, J., & Vanthienen, J. (2003). Benchmarking state-of-the-art classification algorithms for credit scoring. *Journal of the Operational Research Society, 54*(6), 627-635. https://doi.org/10.1057/palgrave.jors.2601545
7. Crook, J. N., Edelman, D. B., & Thomas, L. C. (2007). Recent developments in consumer credit risk assessment. *European Journal of Operational Research, 183*(3), 1447-1465. https://doi.org/10.1016/j.ejor.2006.09.017
8. Hand, D. J., & Henley, W. E. (1997). Statistical classification methods in consumer credit scoring: A review. *Journal of the Royal Statistical Society: Series A (Statistics in Society), 160*(3), 523-541. https://doi.org/10.1111/1467-985X.00078
9. Lessmann, S., Baesens, B., Seow, H.-V., & Thomas, L. C. (2015). Benchmarking state-of-the-art classification algorithms for credit scoring: An update of research. *European Journal of Operational Research, 247*(1), 124-136. https://doi.org/10.1016/j.ejor.2015.05.030
10. Thomas, L. C. (2000). A survey of credit and behavioural scoring: Forecasting financial risk of lending to consumers. *International Journal of Forecasting, 16*(2), 149-172. https://doi.org/10.1016/S0169-2070(00)00034-0