QUANTIFYING CREDIT RISK IN LENDING INDUSTRY: A MONTE CARLO SIMULATION APPROACH

Olalekan M. Durojaiye, Ramanjit K. Sahi

Department of Mathematics & Statistics, Austin Peay State University, TN, USA

ABSTRACT

The loan data simulated with Monte Carlo approach and analyzed in the research work provides valuableinsights into the borrowers' financial positions and loan performance. Debt-To-Income ratio (DTI) was calculated and we identified 141 *(57.8%) loans that were at high risk of default. We then adopted a risk-based pricing (RBP) to mitigate the risk of default by assigning higher interest rates to riskier loans by taking into consideration some parameters like credit score and risk premium. The analysis revealed that a higher DTI is associated with a higher risk of default, and a higher RBP is associated with a higher interest rate. Therefore, it is essential to use these metrics when assessing loan applications to ensure a healthy loan portfolio. This research can be used to inform loan officers, risk analysts, and other stakeholders involved in the lending process.*

KEYWORDS

Loan Simulation, Interest rate, Risk Mitigation, Debt-To-Income Ratio, Risk-Based Pricing

1. INTRODUCTION

Loan recovery and risk mitigation are crucial for lenders to minimize losses and ensure the repayment of loans. To undertake a risk analysis data project to mitigate risk, we need to follow a structured approach that involves several key steps: Identify the Risks, Quantify the Risks, Prioritize the Risks, Develop Mitigation Strategies, Implement and Monitor Mitigation Strategies, and Review and Update.

The 5 Cs of Credit (Character, Capacity, Capital, Collateral, Conditions) is a risk analysis system used by lenders, such as banks and institutional lenders, to determine the creditworthiness of borrowers.

Lenders carefully evaluate multiple factors when making decisions about loan applications. These factors influence not only whether the application is approved or denied, but also shape the specific terms of the financing product. Key elements such as the interest rate and repayment schedule are determined based on this comprehensive assessment. Ultimately, each factor plays a crucial role in crafting a loan structure that balances the lender's risk with the borrower's needs [1].

Numerous studieshave examined the role of risk analysis and mitigation in the lending process. One study by [2] found that banks that used more sophisticated risk management techniques tend

David C. Wyld et al. (Eds): NWCOM, ASOFT, ITCCMA, CSITY, BDIoT, MLNLP, SIGPRO, AIFZ– 2024
pp. 123 -133, 2024. - CS & IT - CSCP 2024 DOI: 10.5121/csit.2024.141911

to experience lower loan losses, i.e., banks with higher capital ratios and more diversified portfolios experienced lower loan losses. Similarly, a study by [3] found that banks that use credit scoring models to evaluate loan applications had lower default rates and lower loan losses. Another study referenced in explored the effect of collateral on loan recovery rates. The findings revealed that lenders who required collateral were more successful in recovering their loans in cases of default [4].

The impact ofeconomic conditions and market factors on lender loan losses has been studiedextensively in the literature. For example, a study by [5] found that economicfactors such as GDP growth and inflation were significant predictors of loan defaults and loan losses. Researchers have conducted various investigations into how external economic factors influence lenders' loan losses.

These studies have focused on key variables such as: Interest rate fluctuations, Exchange rate volatility and Commodity price movements. By analyzing these elements, scholars aim to understand their role in shaping the risk landscape for lenders and their potential impact on loan performance.

Finally, the impact of lender characteristics on loan losses has also been studied in the literature. A study by [6] found that larger banks were less likely to experience loan losses due to their greater diversification and risk management capabilities.

Overall, the literature indicates that lenders employ various techniques and strategies to analyze and mitigate risk, such as credit scoring models, collateral requirements, portfolio diversification, and adherence to regulatory standards. By thoroughly evaluating the risks associated with lending to borrowers and implementing measures to mitigate those risks, lenders can minimize the likelihood of loan defaults and safeguard their investments. This approach allows them to continue providing access to credit for individuals and businesses in need.

The aim here is to see if the lender understands the credit profile of the borrower while using quantitative data of the loan applicants. The two major objectives of this research work are:

- To quantify and mitigate the risk of loan defaults and
- Improve the chances of loan recovery with the help of some metrics such that the risk of incurring a capital loss is reduced.

2. METHODOLOGY

In this section, Monte Carlo simulation approach was applied to generate significant parameters that help in understanding loans in the loan industry. Moreover, a critical risk method was used to quantify and mitigate the risk.

2.1. Monte Carlo Simulations

To study loans and loan defaults and mitigation of risk, specific parameters that are used in the loan data by creditors in real life, were taken into consideration. These parameter values generated using Monte Carlo simulations are between a certain range. Monte Carlo simulation uses random sampling to simulate potential future outcomes. The loan data was simulated using one of the appropriate tools (Python) with various mathematical and financial functions to create a dataset. A sample size of 500 borrowers' information was generated with different loan parameters. With some basic assumptions, the loan data table was simulated by creating variables such as loan amount, interest rate, term, monthly payment and other relevant data.

First, we wrote functions to generate random loan amounts within a specified range. For example, to generate loan amounts between $1000 and $50000, that is, loan_amounts = *np.random.randint*(1000, 50001, *size* = 500). Then, we calculated randon interest rates within a range of 4.6% and 35.9%. The rates depend on the borrower's credit score and other eligibility criteria. Borrowers with a credit score between 610 and 640 are generally eligible, though some lenders may consider scores as low as 300 [7]. Then we used a function to generate random loan terms between 36 and 60 months. After that we used numpy finance to generate the ”PMT” function to calculate the monthly payment for each loan. That is, ”*PMT*(*rate, nper, pv,* [*fv*], [*type*])”, where ”rate” is the interest rate per period, ”nper” is the total number of periods, ”pv” is the present value of the loan, ”fv” is the future value (optional), and ”type” is the timing of payments (optional). Also, we added other relevant data to the loan table, such as loan purpose, collateral value, credit score, borrower income, loan status etc. Once we have finished generating the loan data, we will continue to use the same tool (python) to analyze the data.

2.2. Risk Analysis

Among every other approach like data preparation and exploratory data analysis, we would look into some key methods highlighted below to quantify the risk in the data and proffer better ways to mitigate them.

2.2.1. Debt-To-Income (DTI) Ratio

It is a relevant credit metric that compares the borrower’s monthly debt payments, i.e., The lender analyzes the borrower’s debt-to-income ratio to determine their ability to repay the loan. A borrower with a high debt-to-income ratio is more likely to default, increasing the risk of loan recovery. More so, the target DTI differsby lender, most prefer a DTI around 45% or less to approve the applicant.

$$DTI = \frac{Total\ Loan\ Amount}{Gross\ Monthly\ Income}$$

2.2.2. Risk- Based Pricing (RBI)

It is a method used by lenders to determine the interest rate and loan terms based on the risk of lending to a particular borrower. There isn’t a single formula for risk-based pricing as it involves a complex assessment of various factors that can affect the risk of default. However, a basic approach to risk-based pricing involves adding a risk premium to the base interest rate, which is determined by the lender’s cost of funds and market conditions. The risk premium reflects the lender’s assessment of the borrower’s creditworthiness and the likelihood of default. Find below a general formula for risk-based pricing:

RBI = *Base Interest Rate + Risk Premium*

The base interest rate is the interest rate that the lender would charge for a loan to a borrower with average creditworthiness and a low risk of default. The risk premium is the additional interest rate charged to reflect the increased risk of lending to a borrower with below-average creditworthiness or a higher risk of default.

The amount of the risk premium is determined by the lender’s internal risk assessment and can vary depending on the borrower’s credit score, credit history, income, loan purpose, loan-to-value ratio, and other factors. Generally, a higher credit score and lower risk factors will result in a lower risk premium, while a lower credit score and higher risk factors will result in a higher risk premium.

The lender will typically assign a risk grade to the borrower based on the information gathered and use that grade to determine the appropriate interest rate and loan terms. The higher the risk grade, the higher the interest rate and the more stringent the loan terms.

3. RESULT AND DISCUSSION

We start by showing the excerpt of the simulated data and explore the data further for some key information. Total Loan amount available for 500 borrowers was $13,468,495 with an average interest rate of 15.16% per borrowers. 20 of the 500 simulated data shown in Table 1 below:

Table 1. Excerpt from the Simulated Data

	loan amount	interest rate	term months	monthly payment	loan purpose	collateral value	borrower income	loan status
1	26967	0.1019	60	575.49	student loan	40452	7725	approved
2	42745	0.2239	55	1249.62	debt consolidation	43182	86746	rejected
3	5684	0.0566	41	152.8	student loan	66637	16658	approved
4	44834	0.1718	51	1244.76	car loan	46012	97705	approved
5	26632	0.2479	50	859.29	small business	14308	77269	approved
6	32441	0.0675	52	721.29	personal loan	67799	21498	pending
7	23219	0.2235	53	693.09	small business	55415	68076	rejected
8	19620	0.1469	51	519.49	small business	67333	88097	rejected
9	17759	0.2277	47	574.41	small business	60426	49597	approved
10	11954	0.2628	47	409.85	small business	12942	50472	approved
11	2991	0.1379	53	75.67	car loan	35752	69699	approved
12	4752	0.1878	55	129.49	personal loan	32549	96893	approved
13	12238	0.2017	49	368.55	house	69691	33921	pending
14	7911	0.1059	59	172.59	personal loan	62030	7635	approved
15	18401	0.0731	42	497.88	house	24968	21521	approved
16	47727	0.0614	48	1123.94	house	28992	50660	rejected
17	14714	0.06	48	345.56	debt consolidation	34930	19840	pending
18	35771	0.208	55	1014.14	debt consolidation	57873	49487	approved
19	9988	0.0891	52	232.24	student loan	49770	9620	rejected
20	29278	0.2282	46	960.61	small business	27500	14672	approved

3.1. Exploratory Data Analysis

The table 2 provides summary statistics for a dataset containing 500 loans, including various attributes such as loan amount, interest rate, term (in months), monthly payment, payment expected by term, collateral value, and borrower income. The loan amounts range from $1,074 to $49,983, with a median value of $27,676. The high standard deviation indicates significant variability in loan amounts. Interest rates vary from 4.00% to 26.99%, with a median of 15.28%. The rates are fairly spread out, as indicated by the standard deviation.

Table 2. Summary Statistics of Simulated Data

	loan amount	interest rate	term months	monthly payment	pmt expected by term	collateral value	borrower income
count	500	500	500	500	500	500	500
mean	26936.99	0.1516	47.58	770.66	36017.11	41279.31	50961.97
std	13745.60	0.0643	7.16	415.84	18784.82	16935.73	27045.36
min	1074.00	0.0400	36.00	27.04	1253.45	10174.00	5183.00
25%	14965.75	0.0971	42.00	429.71	20350.62	26706.00	27303.00
50%	27676.00	0.1528	47.00	765.73	36278.02	42279.00	50216.50
75%	38859.00	0.2053	53.25	1075.82	50960.89	55167.00	73818.25
max	49983.00	0.2699	60.00	1971.16	81388.51	69881.00	99597.00

3.2. Distribution of Loan Amount by Loan Purpose

The visualizations provide insights into the distribution and status of loans based on their purpose.The donut chart in Figure 1 illustrates the proportion of loans by their purposes, which shows afairly even distribution of loans across the different purposes, with slight variations in the percentages.

The bar chart (Figure 2) presents the count of loan applications based on their status (approved,rejected, pending) across various loan purposes. The Personal Loans, Small Business Loans, andDebt Consolidation Loans are more likely to be approved. Student Loans have a higher rejectionrate while Pending loan counts are generally lower, indicating fewer loans are in the pending statusacross all purposes.

These visualizations help identify trends in loan approval, rejection, and pending status, whichcan be useful for understanding loan distribution and decision-making processes.

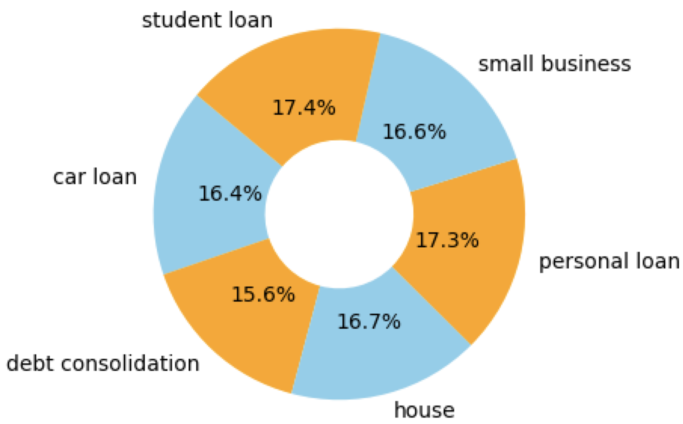


Figure 1: Distribution of loan amount by loan purpose

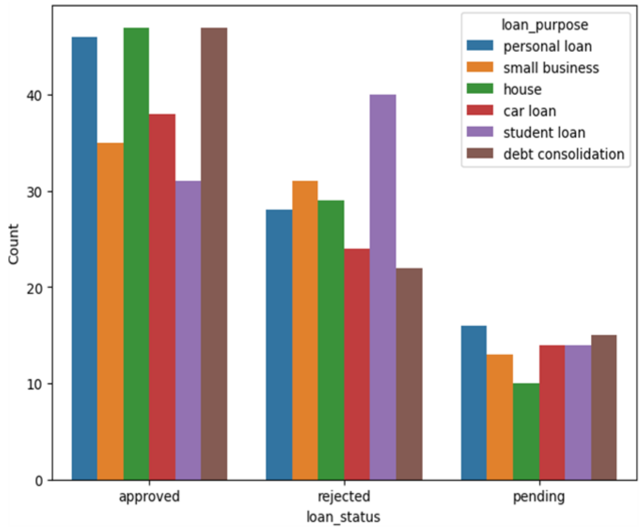


Figure 2: Number of loans by purpose and status

Table 3. Loan Purpose Statistics

loan purpose	**Avg loan amount**	**count**
car loan	26469.486842	76
debt consolidation	25143.738095	84
house	26983.825581	86
personal loan	27929.222222	90
small business	26809.151899	79
student loan	28147.976471	85

3.3. Loan Status Distribution

We were able to analyze the data to see the number of loan approved, rejected or pending. In this research work, our focus would be based on the approved loans. See below Table 4 showing the result.

Table 4. Distribution of loans

loan_status	No_of_loan	loan_amt_total
Approved	**244**	**$6,527,028**
Rejected	174	$2,262,299
Pending	82	$4,679,168
Total	500	$13,468,495

Therefore, we calculated from the data the expected payment by the borrowers to know the lender's expected profit, i.e.,

- **Total Loan Amount:** $ 6,527,028
- **Total Expected Payment:** $ 8,842,806.55
- **Total Expected Profit:** $ 2,315,778.55 (Not a bad business)

3.4. Debt-To-Income (DTI) Ratio

In quantifying the risk involved in all the loans approved, we analyzed the DTI for each of the borrowers. We found that out of the 244 (48.8%) approved loan data, 103 (42.2%) fulfilled their promises and 141 (57.8%) defaulted. See in the table and plot below:

Table 5. DTI Analysis

	loan amount	borrower income	loan status	DTI(%)	DTI target(<=45%)	Promised to Pay(P2P)
1	26967	7725	approved	349.09	high	defaulted
2	18401	21521	approved	85.50	high	defaulted
3	5684	16658	approved	34.12	low	fulfilled
4	44834	97705	approved	45.89	high	defaulted
5	26632	77269	approved	34.47	low	fulfilled
.	.	.	.	.	.	.
.	.	.	.	.	.	.
.	.	.	.	.	.	.
238	4627	14657	approved	55.10	high	defaulted
239	43824	84739	approved	48.19	high	defaulted
240	37326	67211	approved	31.41	low	fulfilled

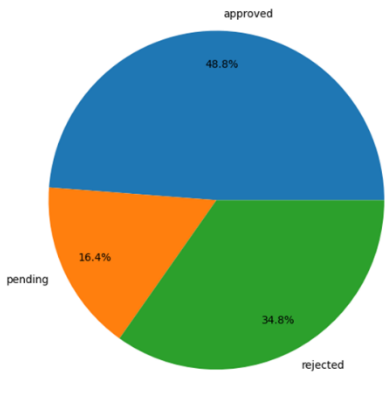


Figure 3: percent distribution of loan by status

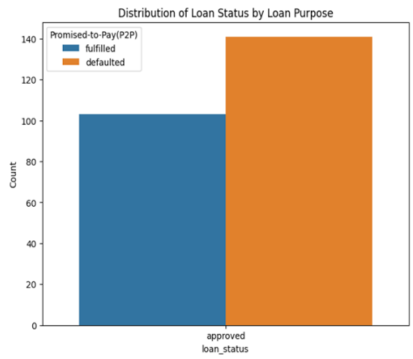


Figure 4: Approved loan by P2P

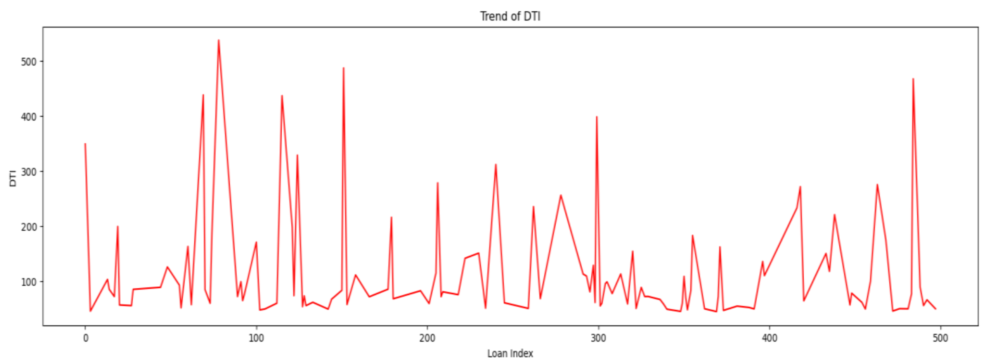


Figure 5: Trend of DTI for approved loans

From the figures (3-5) above, we see that 57.8% of the 244 loans are exposed to possible risk. Further analysis on the lender's expected profit shows the following:

- **Total Expected Payment** by the 244 loans was$ 8,842,806.55
- **Total Payment defaulted** by the 141 loanswas$ 5,476,354.05 . We can see that 62% is not recovered.
- **Total Loss** expected by the lender is -$3,359,267.04

3.5. Risk Mitigation

We were able to identify the risky loans from the 244 approved and adopted RBP to mitigate the risk. We generate random credit scores between 300 and 850 based on defaulted borrowers, which helps to determine the risky loan approved to borrowers and check expected profit or amount of loss that would accrue.

These risk-based pricing (RBP) is determined by adding a risk premium to a base rate. The base rate is assumed to be 26% and the risk premium is added based on the borrower's credit risk, that is 'excellent' = -2.0, 'good' = 0.5, 'fair' = 1.5, 'poor' = 3.0. These highlight the importance of creditworthiness and lender-specific criteria in determining the interest rate offered to borrowers. Lenders use these rates to balance the risk of default against the need to offer competitive terms to attract customers.

Table 6. RBP Analysis

	interest rate	**loan status**	**DTI(%)**	**DTI target(<=45%)**	**Promised to Pay(P2P)**	**credit score**	**credit score range**	**risk premium**	**RBP**
1	10.19	approved	349.09	high	defaulted	346	poor	3.0	29.0
2	17.18	approved	45.89	high	defaulted	560	poor	3.0	29.0
3	10.59	approved	103.61	high	defaulted	743	excellent	-2.0	24.0
4	7.31	approved	85.50	high	defaulted	559	poor	3.0	29.0
5	20.80	approved	72.28	high	defaulted	467	poor	3.0	29.0
6	22.82	approved	199.55	high	defaulted	470	poor	3.0	29.0
7	20.26	approved	57.16	high	defaulted	706	excellent	-2.0	24.0
8	22.90	approved	56.00	high	defaulted	662	good	0.5	26.5
9	18.65	approved	85.48	high	defaulted	428	poor	3.0	29.0
10	14.33	approved	89.24	high	defaulted	380	poor	3.0	29.0
.	.	.	.	.	.	.	.	.	.
.	.	.	.	.	.	.	.	.	.
.	.	.	.	.	.	.	.	.	.
120	21.10	approved	49.20	high	defaulted	612	good	0.5	26.5
121	19.01	approved	70.18	high	defaulted	400	poor	3.0	29.0
122	20.33	approved	77.24	high	defaulted	391	poor	3.0	29.0

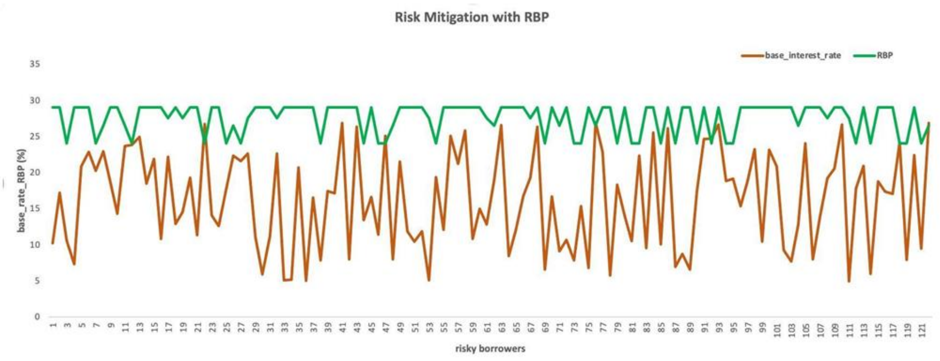


Figure 6. Relationship between Interest Rate and Risk-Based Pricing

4. CONCLUSION

Lenders employ a diverse array of risk analysis techniques to evaluate borrower creditworthiness and assess lending risks. Key factors in this research work include: Credit scores, Collateral value, Debt-to-income ratio (DTI), Payment history, Industry and Market Conditions.

The key findings from the analysis of the Monte Carlo simulation data revealed several important insights:

1. Debt-to-Income Ratio (DTI):

   - Used to gauge a borrower's loan repayment capacity
   - High DTI indicated an elevated default risk
   - Identified 141 high-risk loans, potentially resulting in losses of approximately $3.4 million

2. Risk-Based Pricing (RBP):

   - Implemented as a strategy to adjust interest rates based on perceived borrower risk
   - Potential correlation observed between DTI and RBP, with higher DTI potentially leading to higher RBP
   - Other factors like credit score, loan purpose, and amount may also influence RBP
   - Further analysis recommended for more definitive conclusions

Therefore for the risk mitigation strategies to minimize loan default risks and enhance recovery prospects, lenders may implement measures such as: Requiring collateral, Charging higher interest rates or fees, Mandating a cosigner etc.

Overall, lenders must meticulously evaluate the risks and benefits associated with each borrower, considering the multitude of factors that could impact loan repayment. This thorough approach enables lenders to make well-informed lending decisions, reduce the likelihood of loan defaults, and maintain a balance between risk management and credit accessibility.

By adhering to these principles, lenders can continue to fulfill their crucial role in providing financial support to individuals and businesses while safeguarding their own interests and maintaining a stable lending environment.

5. FUTURE RESEARCH

Future study may also want to establish a Loan-to-value (LTV) approach to quantify risk and ensure that loans are offered only to borrowers who have the ability to repay them. LTV is commonly used in the mortgage industry, where lenders use it to determine the amount of a mortgage they are willing to offer to a borrower. Generally, lenders prefer lower LTV ratios and may require borrowers to pay for mortgage insurance or other risk-mitigation measures if the LTV exceeds a certain threshold. The exact LTV requirements may vary depending on the lender's risk appetite and the type of loan being offered. In addition, lenders may use risk management tools, such as credit scoring models and loan portfolio diversification strategies, to minimize the risk of losses.

ACKNOWLEDGEMENTS

The authors would like to thank the Department of Mathematics & Statistics, College of STEM, and College of Graduate Studies at Austin Peay State Universityfor their support. Lastly, we would like to thank our families and friends for their encouragement and support.

REFERENCES

[1] Segal, T. (2023). 5 cs of credit: What they are, how they're used, and which is most important. investopedia: Personal Finance and Loan. https://www.investopedia.com/terms/f/five-c-credit.asp#toc-understanding-the-5-cs-of-credit.

[2] Demirgüç-Kunt, A., & Huizinga, H. (1999). Market discipline and financial safety net design (Vol. 2311). World Bank Publications.

[3] Berger, A. N., & DeYoung, R. (1997). Problem loans and cost efficiency in commercial banks. Journal of banking & finance, 21(6), 849–870.

[4] Yoon, Y., Li, Y., & Feng, Y. (2019). Factors affecting platform default risk in online peer-to-peer lending business: an empirical study using chinese online p2p platform data. Electronic Commerce Research, 19, 131–158.

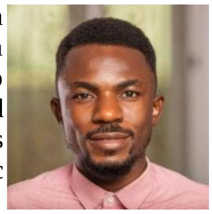
[5] Altman, E. I., & Sabato, G. (2007). Modeling credit risk for smes: Evidence from the US market. Abacus, 43(3), 332–357.

[6] Harhoff, D., &Körting, T. (1998). Lending relationships in Germany–empirical evidence from survey data. Journal of Banking & Finance, 22(10-11), 1317–1353.

[7] Kat Aoki& Holly Jennings (2023). Bankrate personal loans review: Compare multiple lenders in one place. https://www.finder.com/personal-loans/bankrate-personal-loans-review

AUTHORS

Olalekan M. Durojaiye is a dedicated and results-driven professional with a passion for leveraging data to drive informed business decisions. With a strong foundation in Data Science, Engineering, Machine Learning, Artificial Intelligence, and Deep Learning. He holds a Master of Science (MSc.) degree in Computer Science and Quantitative Methods (Mathematical Finance concentration). Olalekan delivers innovative solutions that optimize processes, enhance efficiency, and facilitate strategic planning.



Dr. Ramanjit K. Sahi is a Professor and Graduate Coordinator in the Dept. of Mathematics & Statistics, Austin Peay State University, TN, USA. She received her Ph.D. from University of Texas at Dallas and joined APSU in Fall 2007. Her research interests are in the field of Mathematical Finance, Mathematical Modeling, Risk Analysis, and Knot Theory.