

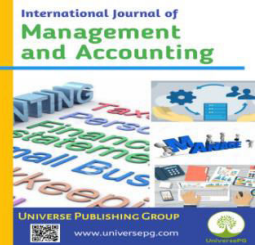


Publisher homepage: www.universepg.com, ISSN: 2707-4641 (Online) & 2707-4633 (Print)

<https://doi.org/10.34104/ijma.021.01140121>

International Journal of Management and Accounting

Journal homepage: www.universepg.com/journal/ijma



Machine Learning application for selecting efficient Loan Applicants in Private Banks of Bangladesh

Hasan M Sami^{1*}, Md. Rafatuzzaman¹, and Anirudha Bar¹

¹School of Business, Dept. of Accounting, Canadian University of Bangladesh, Pragati Sharani, Dhaka-1213, Bangladesh.

*Correspondence: samifrnds90@gmail.com (Hasan M Sami, Senior Lecturer, School of Business, Canadian University of Bangladesh, Pragati Sharani, Dhaka-1213, Bangladesh).

ABSTRACT

Machine Learning Applications have been well accepted for various financial processes throughout the world. Supervised Learning processes for objective classification by Naïve Bayes classifiers have been supporting many definitive segregation processes. Various banks in Bangladesh have found challenging moments to identify financially and ethically qualified loan applicants. In this research process, we have confirmed the safe applicant's list using definitive variable measures through identifiable questions. Our research process has successfully segregated the given applicants using Naïve Bayes classifier with the proof of lowering loan default rate from an average of 23.26% to 11.76% and development of financial ratios as performance indicators of these banks through various financial ratios as indicators of these banks.

Keywords: Applicant feature, Supervised learning, EPS, Efficient loan, PE ratio, and Naïve Bayes Classifier.

INTRODUCTION:

Private Banks in Bangladesh has been facing serious challenges since the very beginning in terms of identifying appropriate loan applicants. Various factors such as growing inflation, property price rise, nepotism and biasness, personal networking have made the application process highly inaccurate for proper identification (Ziaul, 2003) of loan applicants. The NPL (non-performing loans) within the growing banking sector of Bangladesh since the beginning of financial flexibility in Bangladesh has played as a positive inhibitor for financial independence (Rezina, 2020) but failed to effectively high-light correct loan provision policy. Although lots of surrounding political and environmental factor plays a crucial role for NPL growth but the majorities of growths are still inscribed within the human sentiments (Ali, 2012). In this research paper we collectively focus on factors that are associated with human faults rather than honest

mistakes in order to make the research process effective. Our primary analysis that was raised through Ziaul's question refers to determining variable criterions of the loan applicants and effective data analysis for suggestive solution formation. Vedala *et al.* have found that Naïve Bayes classifier is an effective method for identifying good and bad borrowers for better financial decisions (Vedala *et al.*, 2012). In order to understand the possible criterions to judge the loan applicant factors related to their current situation needs to be stipulated (Krichene, 2017). Although Rezina already mentioned that in Bangladesh despite of growth there are various identifiable factors those needs to be judged in order to make the loan provision process acceptable, but Bangladesh Bank has yet to make proper adjustments accordingly.

Banks overall financial development through this process needs to be an important indirect factor for loan approval. Sami *et al.* (2021) have shown that how

financial ratio becomes an important inhibitor for asset selection in case of financial investment (Sami *et al.*, 2021). It's also shown by HM Sami that in case of effective asset selection both financial ratios and classification method seems important (Sami, 2021). This research process will associate banks financial performances as an important background justifying classification process as an inhibitory action for selection of borrowers. In order effectively understand the criteria for selection various similar situation sorts needs to be analyzed and justified as important criteria for loan approval (Majeske, 2013).

In order to understand the financial implications of the financial viewpoint, Bangladesh Bank has governed several parameters to ideally identify a potential borrower (Weber *et al.*, 2015). Weber *et al.* have further evaluated the strengths of the given criteria as important standpoints for financial sustenance through loan provision method for banks. Hence in this research process we will use successful classification process to clearly allow only definitive criteria to check the financial perseverance of the borrowers for selected business plans. In general, our overall process is highly data supported and variable oriented hence we didn't propose any deep differentiation of loan types, fund requirements and business types associated with fund requirements as an important benchmark for financial analysis.

Literature Review

Singh *et al.* (2021) have figured out that through modernized methods of loan approvals based on machine learning parameters such as XGBoost and Random Forest process of decision tree have clearly considered credit score and other parameters as important decision-making parameter for loan approval. Similarly, it's also seen that through financial transactions credit scores could be determined which are important basis for financial decision makings (Lohokare *et al.*, 2017). It's seen in the research process of Ziaul's suggestion, although surrounding environmental factors plays a crucial role in loan approval decision both from borrower and lender perspective but this research would concentrate specifically on the financial aspect terms with association of observable factors as important borrowing determinants (Kemalbay *et al.*, 2014). Kemalbay *et al.* also critically

suggested that in case of minorities with low income and shorter employment history has higher loan default rate along with higher rates of disapproval (Kemalbay *et al.*, 2014). Calem has made an effective suggestion that shows the true relative analysis of income as benchmark for loan approval and promissory loan return (Calem, 1996). Squires & Kim, (1995) suggested in their primary overview, with effectivity in terms of gender, minority and income level loan approvals remain an effective classification parameter. It's also seen that with increasing development around the world and also with more ease and effectivity in reducing discriminating parameters more subjective issues became better determiner for loan approvals such as income level and employment history rather than discriminant factors (Hassan *et al.*, 2021; Harrison & Glover, 2008).

A major segregation or classification comes around the factor of financial condition of the applicant with reference to their asset condition and liquidity of collateral in reference to loan. There are various financial stages which generally classify a person as a better borrower than another (Carter *et al.*, 2007). With NPL rise in Bangladesh although Rezina clearly suggested a stricter approach towards approval but Naïve Bayesian methods will provide a definitive benchmark for applicant's eligibility through past data's (Rezina, 2020). In reference to information available for the public, financial ratios as performance benchmark plays an important role for judgment of the bank's present performance standards in terms of EPS & PE Ratio (Sami HM, 2021). It's also seen that in Bangladesh especially for private banks the performance measurement factors are highly reliant on financial ratios such as EPS and PE Ratio (Siddik *et al.*, 2017). Mayer & Gary has identified that in case of making loan approvals, most defaulters tend to set a smaller down payment boundary which seems more effective towards default rather a larger payment always makes the default risk more mitigated (Mayer & Gary, 1996). Loutskina & Strahan has made further contributions towards the field of available liquidity of borrower's assets as an important background loan approval (Loutskina & Strahan, 2009). Although most operating industries effectively applies for loans using project profile and previous work history but the effective results need to be

portrayed in financial ratios for performance development or evaluation (Sami, 2021). Referring to the length of the loan term associated with loan approval Heider & Inderst made the most appropriate suggestions. It's found that on the basis of length of the loan terms a higher interest rate seems more favorable for the loan because of increasing inflation (Heider & Inderst, 2012). Kathleen & Li demonstrated that personal debt and income ratio plays an important criterion for selection of borrowing individuals (Kathleen & Li, 2010). In relation to finding equality it seen that people with higher incomes generally tend to be better borrowers for banks in reference to same debt levels in terms of money (Kathleen & Li, 2010). But in general, lower debt to income ratio remains a favorable background for banks in order to identify a better borrower. Bian *et al.* have found that with smaller deposits on loan or less liquid assets as collaterals the amount of loan to the original property value for loan borrowing becomes highly close to 1 which remains a signal for bad borrower (Bian *et al.*, 2018). Blackwell & Winters have found out that if a borrower with more assets rather than their liabilities then its more favorable for banks to make lending's for them (Blackwell & Winters, 1997). Its also seen that the best possible loan default prohibitions happen only against greater than 1 asset liability ratio.

Theoretical Framework

Debt to Income Ratio: The debt-to-income ratio (DTI) is a lending ratio that represents a personal finance measure, comparing an individual's debt repayments to his or her gross income on a monthly basis. Gross income is simply a monthly paycheck before one pays off the costs, such as taxes, interest expense, etc.

$$\text{Debt-to-Income Ratio} = \frac{\text{Total Monthly Debt Payments}}{\text{Gross Monthly Income}}$$

In the research perspective, for the associated benefits Bangladesh bank have arranged a guideline to govern each banking activities as subordinating regulation by which their regular debt payments are supplemented by regular monthly payments against any loans. Under this perspective the newly formatted Debt to Income Ratio if greater than 1 then the banks are at risk and if lower then the payments are smooth. Hence the reformulated formula is:

$$\text{Debt to Income Ratio} = \frac{\text{Total Monthly Debt Balance}}{\text{Total Monthly Loan repayments}}$$

Similarly, the effectivity of Loan to Value Ratio plays an important role for loan approvals. It seems that

$$\text{Loan to Value} = \frac{\text{Amount of Loan}}{\text{Property Value}}$$

In reference to the approval of loan if the ratio is very close to 1 then the possibility of approval should be very limited whereas in case when the loan is very close to 0 then approval should be easier. The borrower's notion to current ratio will have to modified for loan approval view. In accordance to company-based scenario if an organization has more than 1 current ratio then the organization has more assets than its liabilities. Similarly, if a borrower has more assets rather than liabilities through asset collaterals or by financially liquidating sources. So, working capital ratio or current ratio for loan borrower is

$$\text{Current Ratio} = \frac{\text{Current Assets Owned}}{\text{Current Liabilities as loans}}$$

In this research we would arrange a framework to identify applicants with above mentioned ratios and use various statuses including credit status, income status, networking abilities, employment length, collateral liquidity, down payment size, number of loan terms and loan length plays vital role in identifying a qualified borrower. The three financial variables will indicate the financial conditions of the borrower whereas the other variable features from the perspective of previous performance of the borrowers would generate the bets possible notion of borrower's behavior in reference to lending.

Research Framework

In order to make this process effective, in this research we have initiated the primary process of selecting applicants from 5 different banks. Each of these banks have a NPL of more than 30% and weaker payment pattern of 56%. Most of these banks have extremely large capital but the loan approval process was highly non-conforming. Using K Means clustering we have been able to distinguish bad chunk of loan applicants pretty easily using the parameter setup through algorithmic boundaries. With the usability of high networking parameters, low debt to payment ratio, low loan to value ratio and more than 1 current ratio of the

borrowers, the primary borrower background should be selected. After making the selection of these borrowers, using these borrowers previous loan payment pattern as training data, a section of testing was arranged as a pilot process for major evaluation by naïve bayes classifier to judge the overall criteria that could be accepted from a group of selected applicants for loan approval.

METHODOLOGY:

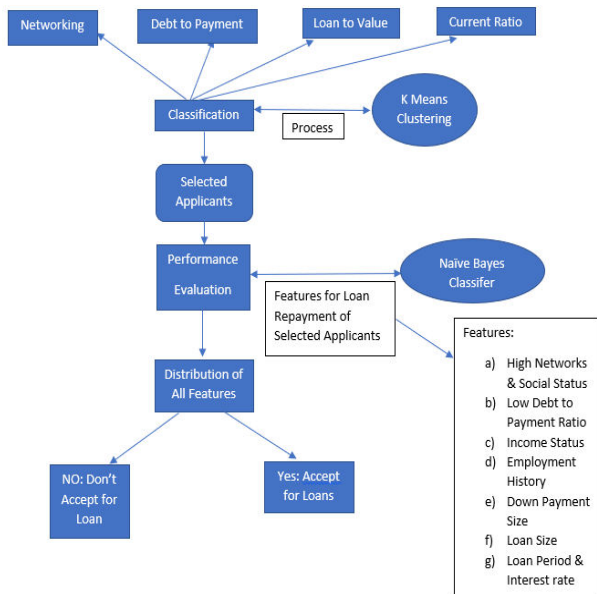


Fig 1: Outline of the Research plan.

Using K Means Clustering we would select applicants by following algorithmic features:

- Xi = Networking & Social status
- Yi = Loan to Payment Condition
- Zi = Loan to Value
- Ki = Current Ratio of Borrower

For loan applicants to be acceptable in accordance to the issue of bank approval for lending to a borrower of effective acceptability various deciding factors are followed –

- Xn = High Networking & Social Status
- Yn = High Loan to Payment Ratio
- Zn = Low Loan to Value Ratio
- Kn = More than 1 Current Ratio of the Borrower

Using these factors

Acceptable Terms are generated as:

$$\text{Acceptable Borrower} = \text{Min} (\text{Abs} (\text{Xn} - \text{Xi}) + \text{Abs} (\text{Yn} - \text{Yi}) + \text{Abs} (\text{Zn} - \text{Zi}) + \text{Abs} (\text{Kn} - \text{Ki}))$$

Only the Lowest possible values would then be accepted. After the selection of primarily selected borrowers with good loan paying status, we would then extract data of similar features from the previously applied loan applicants. It's been seen and observed that various loan paying individuals are capable of various positive features but has exhibited under payment or no payment despite of positive features. Through Naïve Bayes Classifier these data are the further evaluated for finding out effective loan applicants. While proceeding with Naive Bayesian classification process we need to administer various background to develop its importance.

$$P(y|x_1, \dots, x_n) = \frac{P(x_1|y)P(x_2|y)\dots P(x_n|y)P(y)}{P(x_1)P(x_2)\dots P(x_n)}$$

Where each of the criteria are separately calculated on the basis of given probability.

$$P(y|x_1, \dots, x_n) = \frac{P(y) \prod_{i=1}^n P(x_i|y)}{P(x_1)P(x_2)\dots P(x_n)}$$

Here all the criteria are being exhibited together.

$$P(y|x_1, \dots, x_n) \propto P(y) \prod_{i=1}^n P(x_i|y)$$

If all the probabilities of the events are justified are denoting a joint probability due to being mutually exclusive, hence the Probability of lender satisfaction being earned given the events are contained is directly proportional to the probability of satisfaction multiplied with probability of satisfaction earned when each event are satisfied.

$$y = \text{argmax}_y P(y) \prod_{i=1}^n P(x_i|y)$$

In the final step of argmax, it's clearly denoting if any specific criteria remain positively influencing towards a pathway of decision, then it would probably denote the definitive answer. Hence by this process the selected pool of applicants within common boundary criteria could be selected with positive and effective traits. All the variables of xi refer to the exhibiting criteria from feature set that would be defined under various strength and week parameter. Generally, from a lenders perspective the strength parameters make the suitable decision for loan approvals. Among the selected more than 3000 applicants of different status and nature from of pool of 10000 applicants for the term of 2021 January- March session, we were able to easily classify

best applicants using K Means Clustering. Among the 3000-applicant pool in various private banks there were quite a lot of them with good features but those features didn't make good examples previously. After extracting previous data of 12343 applicants of similar features it was found that nearly 4678 of them with similar features were running underpayment or didn't complete loan payments for the banks. Hence it was of paramount importance to understand definitive features by which a loan applicant will be able to make payments properly and who will not be able to do it.

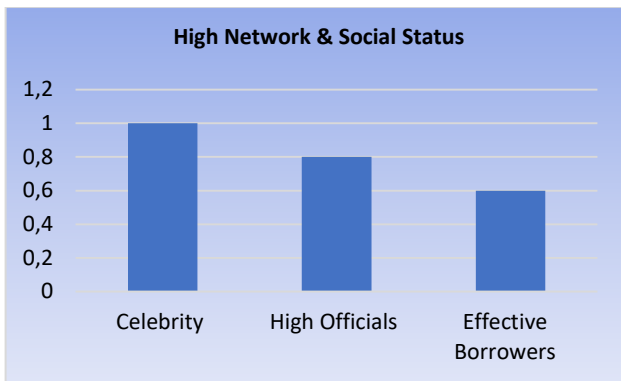


Fig 2: High Network & Social Status.

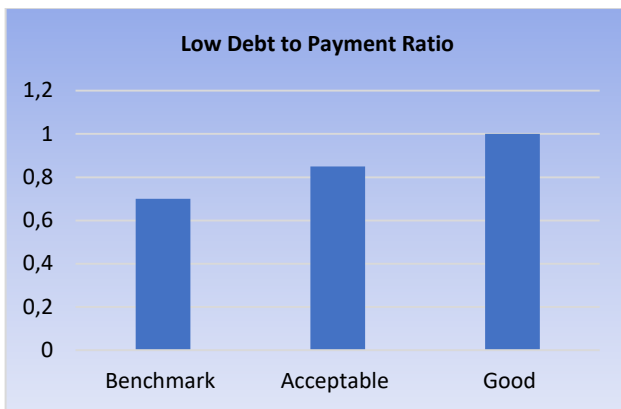


Fig 3: Low Debt to Payment Ratio.



Fig 4: Income Status.

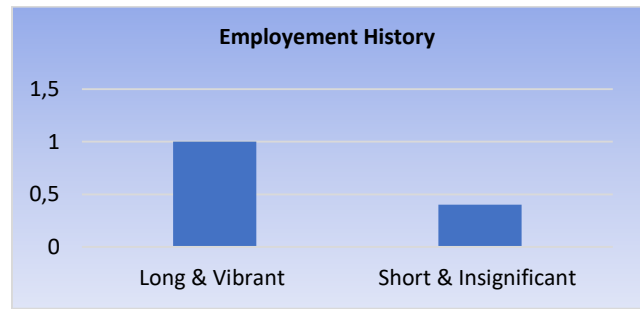


Fig 5: Employment History.

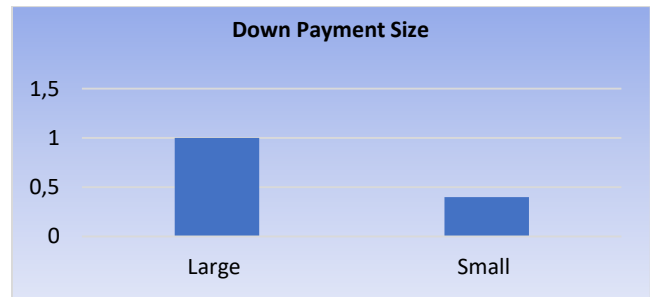


Fig 6: Down Payment Size.

In order to effectively segregate the selected applicant, pool it's been found that with the above average range of expected growth from the minimum till 1 there has been a credibility of return of more than 75% with proper repayments and no late or underpayments. Hence

$$[P(\text{High Network \& Social Status} > 0.6)] = P(X_a) = 1 - P(\sim X_a)$$

$$[P(\text{Low Debt to Payment Ratio} > 0.4)] = P(Y_a) = 1 - P(\sim Y_a)$$

$$[P(\text{Income Status} > 0.4)] = P(Z_a) = 1 - P(\sim Z_a)$$

$$[P(\text{Employment History} > 0.4)] = P(W_a) = 1 - P(\sim W_a)$$

$$[P(\text{Down Payment Size} > 0.4)] = P(k_a) = 1 - P(\sim K_a)$$

Algorithm

So, if

$X_a = []$
 $Y_a = []$
 $Z_a = []$
 $W_a = []$
 $K_a = []$
 For a in range:
 If $(P(X_a) \wedge P(Y_a) \wedge P(Z_a) \wedge P(W_a) \wedge P(K_a) > P(\sim X_a) + P(\sim Y_a) + P(\sim Z_a) + P(\sim W_a) + P(\sim K_a))$:
 Accept the Loan Applicant
 Else:
 Reject the Loan Applicant

Our results are demonstrated under two different parameters. Primarily we were capable of successfully segregating the borrower group of better promise and capability. Then among those 3000 capable loan applicants we were able to select the best 2133 applicants based on features which all exhibited positive notions of repayment. In this process for January till March, the Bank has found that around 98.3% of the selected applicants were making regular payments with only 4 people making no payments in 3 months interval.

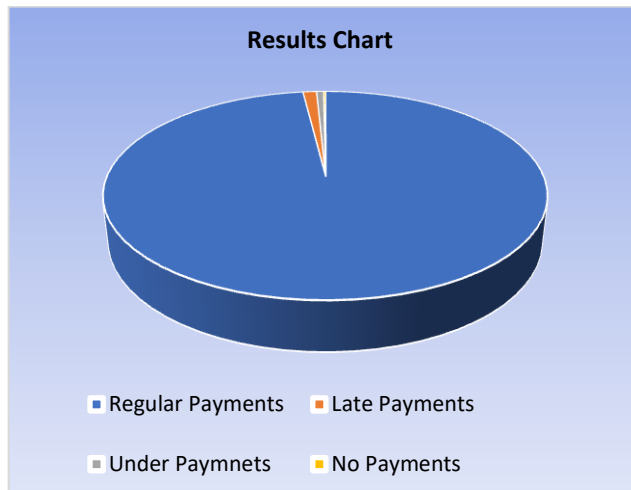


Fig 7: Results Analysis.

Scopes of Improvement & Limitations

This research is conducted strictly under the notion of statistical dependency to construct a definitive background for eligibility of loan candidates. But the overall process of Loan approval needs to be highlighted under the context of fundamental business development process along with various scopes of business growth and proposition, hence this research process needs to be evaluated under a fundamental perspective of understanding loan approval processes. Similarly, various benchmark features suggested by various other researches are considered as pioneering variables in this research process for determining loan approval decision but in reality, various other variables could have been evaluated and judged for better performance appraisals.

Finally, the process is a high degree reduction policy which does not invite more investment opportunity and is very suitable for economies with tightening policies as a whole but does not exactly exhibit the effective nature of finding out true qualifiers for Banks and furthermore economic benefit.

CONCLUSION:

For Bangladesh with growth of 5.5% inflation and around 7% GDP growth rate this process of stringent loan approval could be highly useful. Although it is highly statistical but such a process would eliminate applicants of low-quality return capability. Albeit it is observed that in our research process we lack the evidence of support for newer or potentially capable firms in many ways but this has reduced the loss of loan default to less than 1% which could be an extremely good method of support.

ACKNOWLEDGEMENT:

We do acknowledge the continuous support and help from various privatized banks who have helped us relentlessly through data provision and continuous support with managerial application changes for the application of such suggestive reforms. Due to market conflict the names of such private organizations are remained anonymous in this research process.

CONFLICTS OF INTEREST:

We declare no single potential conflict of interest.

REFERENCES:

- 1) Ali, Md. Yousuf, (2012). Bank Specific Determinants of Nonperforming Loan: Empirical Study in Case of State-Owned Commercial Banks in Bangladesh Available at SSRN: <http://dx.doi.org/10.2139/ssrn.3116433>
- 2) Bian, X., Lin, Z., & Liu, Y. (2018). House price, loan-to-value ratio and credit risk. *Journal of Banking & Finance*, **92**, 1–12. <https://doi.org/10.1016/j.jbankfin.2018.04.00>
- 3) Blackwell, D. W., & Winters, D. B. (1997). Banking Relationship and The Effect of Monitoring on Loan Pricing. *Journal of Financial Research*, **20**(2), 275–289. <https://doi.org/10.1111/j.1475-6803.1997.tb00249.x>
- 4) Calcagnini, G., Cole, R., Giombini, G., & Grandicelli, G. (2018). Hierarchy of bank loan approval and loan performance. *Economia Politica*. <https://doi.org/10.1007/s40888-018-0109-3>
- 5) Calem, P. (1996). Mortgage credit availability in low- and moderate-income minority neighborhoods: Are information externalities critical? *The Journal of Real Estate Finance and Economics*, **13**(1), 71–89. <https://doi.org/10.1007/bf00174552>

- 6) Carter, S., Shaw, E., Lam, W., & Wilson, F. (2007). Gender, Entrepreneurship, and Bank Lending: The Criteria and Processes Used by Bank Loan Officers in Assessing Applications. *Entrepreneurship Theory and Practice*, **31**(3), 427–444.
<https://doi.org/10.1111/j.1540-6520.2007.00181.x>
- 7) Ezeala-Harrison, F., & Glover, G. B. (2008). Determinants of Housing Loan Patterns Toward Minority Borrowers in Mississippi. *Journal of Economic Issues*, **42**(1), 75–96.
<https://doi.org/10.1080/00213624.2008.115071>
- 8) Hassan MK, Hassan MR, Ahmed MMT, Sabbir MSA, Ahmed MS, and Biswas M. (2021). A survey on an intelligent system for persons with visual disabilities. *Aust. J. Eng. Innov. Technol.*, **3**(6), 97-118.
<https://doi.org/10.34104/ajeit.021.0970118>
- 9) Heider, F., & Inderst, R. (2012). *Loan Prospecting*. *Review of Financial Studies*, **25**(8), 2381–2415. <https://doi.org/10.1093/rfs/hhs051>
- 10) Johnson, K. W., & LI, G. (2010). The Debt-Payment-to-Income Ratio as an Indicator of Borrowing Constraints: Evidence from Two Household Surveys. *Journal of Money, Credit and Banking*, **42**(7), 1373–1390.
<https://doi.org/10.1111/j.1538-4616.2010.00345.x>
- 11) Kemalbay, G., & Korkmazoğlu, Ö. B. (2014). Categorical Principal Component Logistic Regression: A Case Study for Housing Loan Approval. *Procedia - Social and Behavioral Sciences*, **109**, 730–736.
<https://doi.org/10.1016/j.sbspro.2013.12.537>
- 12) Krichene, A. (2017). *Using a naive Bayesian classifier methodology for loan risk assessment*. *Journal of Economics, Finance and Administrative Science*, **22**(42), 3–24.
<https://doi.org/10.1108/jefas-02-2017-0039>
- 13) Lohokare, J., Dani, R., & Sontakke, S. (2017). Automated data collection for credit score calculation based on financial transactions and social media. 2017 *International Conference on Emerging Trends & Innovation in ICT (ICEI)*.
<https://doi.org/10.1109/etiict.2017.7977024>
- 14) Loutskina, E., & STRAHAN, P. E. (2009). Securitization and the Declining Impact of Bank Finance on Loan Supply: Evidence from Mortgage Originations. *The Journal of Finance*, **64**(2), 861–889.
<https://doi.org/10.1111/j.1540-6261.2009.01451.x>
- 15) Majeske, K. D., & Lauer, T. W. (2013). The bank loan approval decision from multiple perspectives. *Expert Systems with Applications*, **40**(5), 1591–1598.
<https://doi.org/10.1016/j.eswa.2012.09.001>
- 16) Mayer, C. J., & Engelhardt, G. V. (1996). Gifts, Down Payments, and Housing Affordability. *Journal of Housing Research*, **7**(1), 59–77.
<http://www.jstor.org/stable/24832849>
- 17) Rezina, Sonia, Non-Performing Loan in Bangladesh: A Comparative Study on the Islamic Banks and Conventional Banks (April 6, 2020). *Indian Journal of Finance and Banking*; **4**(1), 2020.
<https://ssrn.com/abstract=3660812>
- 18) Sami HM. (2021). Optimizing organizational overall performance, the use of quantitative choice of HR in carrier quarter enterprise of Bangladesh, *Can. J. Bus. Inf. Stud.*, **3**(3), 49-59
- 19) Sami HM. (2021). Portfolio construction using financial ratio indicators & classification through machine learning, *Int. J. Manag. Account.* **3**(4), 83-90. <https://doi.org/10.34104/ijma.021.083090>
- 20) Sami HM, Fardous L, and Ruhit DS. (2021). Portfolio optimization in DSE using financial indicators, LSTM & PyportfolioOpt, *Int. J. Mat. Math. Sci.*, **3**(4), 74-84.
<https://doi.org/10.34104/ijmms.021.074084>
- 21) Siddik, M., Kabiraj, S., & Joghee, S. (2017). Impacts of Capital Structure on Performance of Banks in a Developing Economy: Evidence from Bangladesh. *International Journal of Financial Studies*, **5**(2), 13.
<https://doi.org/10.3390/ijfs5020013>
- 22) Singh, V., Yadav, A., Awasthi, R., & Partheeban, G. N. (2021). Prediction of Modernized Loan Approval System Based on Machine Learning Approach. 2021 *International Conference on Intelligent Technologies (CONIT)*.
<https://doi.org/10.1109/conit51480.2021.94984>
- 23) Squires, G. D., & Kim, S. (1995). Does Anybody Who Works Here Look Like Me: Mortgage Lending, Race, and Lender Employment. *Social Science Quarterly*, **76**(4), 823–838.
<http://www.jstor.org/stable/44072686>

- 24) Vedala, R., & Kumar, B. R. (2012). An application of Naive Bayes classification for credit scoring in e-lending platform. 2012 *International Conference on Data Science & Engineering (IC-DSE)*.
<https://doi.org/10.1109/icdse.2012.6282321>
- 25) Weber, O., Hoque, A., & Ayub Islam, M. (2015). Incorporating environmental criteria into credit risk management in Bangladeshi banks. *Journal of Sustainable Finance & Investment*, 5(1-2), 1–15.
<https://doi.org/10.1080/20430795.2015.1008736>
- 26) Ziaul Hoque, M. (2003). *Moral Hazard and Industrial Loan Default: Evidences from Bangladesh*. *Humanomics*, 19(3), 7–17.
<https://doi.org/10.1108/eb018885>

Citation: Sami HM, Rafatuzzaman M, and Bar A. (2021). Machine learning application for selecting efficient loan applicants in private banks of Bangladesh, *Int. J. Manag. Account.* 3(5), 114-121.

<https://doi.org/10.34104/ijma.021.01140121> 