

The Effect of Digitalization and Human Intervention in Credit Decisions on the Risk of Late Payments in Multifinance Companies PT. XYZ

Ferlan Nopendi, Priyarsono, Heti Mulyati

Institut Pertanian Bogor, Indonesia

Email: ferlan.novendi@gmail.com, priyarsono@yahoo.com, heti@apps.ipb.ac.id

Keywords

digitalization, multifinance, first payment default, credit expert, logistics regression

ABSTRACT

This study analyzes the impact of digitalization and human intervention in credit decisions on the risk of late payments in multifinance companies PT. XYZ. The digital transformation in the financial industry has shifted the creditworthiness assessment process from a manual system to a fully automated one. The study used a quantitative method with binary logistic regression analysis on data from 533 credit contracts divided into two periods: January–April 2021 (196 contracts with human intervention) and June–September 2021 (337 contracts without human intervention). The variables analyzed included tenor, digital score, effective interest, monthly income, and credit expert involvement in First Payment Default (FPD). The results showed that digital scores had a significant influence on the probability of FPD, with an odds ratio of 1.852 ($p = 0.009$), meaning that every increase of one unit in digital score increased the probability of FPD by 85.2%. In contrast, the variables of tenor, interest, income, and credit expert intervention did not show a significant influence on FPD. The Chi-square test confirmed that there was no significant difference in FPD levels between credits with and without human intervention ($p = 0.148$). The model has a classification accuracy of 96.4%, with an AUC value of 0.738, indicating a fairly good discriminating ability. Strict implementation of Know Your Customer (KYC) plays an important role in risk mitigation. These findings provide managerial implications that the digital underwriting system can be relied upon as the main tool in the credit assessment process, supporting the decision of PT.

INTRODUCTION

Credit is one of the alternative solutions in meeting the needs for goods or services due to limited public income. Credit is the granting of the right to use money or goods to another party at a certain time, with or without collateral, with the provision of services or interest, or without interest (Rivai and Permata 2007). Credit is generally provided by Bank Financial Institutions or Non-Bank Financial Institutions, such as financing/multifinance companies.

According to the Financial Services Authority (OJK), multipurpose financing receivables at finance companies grew by 13.7% from January 2021 to May 2024 (Figure 1). The distribution of financing by financial institutions must be balanced with risk management to obtain maximum profitability. One of the factors affecting profitability is credit risk, which can be seen from the Non-Performing Financing (NPF) ratio. According to OJK Regulation No. 35 of 2018, the NPF ratio is the ratio of the quality of net non-performing financing receivables to total financing receivables.

The credit analysis process is carried out to ensure the debtor's eligibility. Credit analysis is a study conducted to determine the feasibility of a credit application using the 6C analysis technique, namely Character, Capacity, Capital, Collateral, Condition of Economic, and Constraints (Rivai and Permata 2007). The credit underwriting

The Effect of Digitalization and Human Intervention in Credit Decisions on The Risk of Late Payments in Multifinance Companies PT. XYZ

assessment process is a procedure that must be passed before a loan is granted and approved, to assess credit risk, creditor eligibility, and the borrower's ability to meet obligations (Pickert 2017).

The creditworthiness assessment process is essential for both lenders and borrowers to ensure that the loan corresponds with the financial circumstances of the prospective borrower and that the lender can offer loans with reasonable terms and interest rates tailored to the prospective borrower's needs. The creditworthiness assessment helps lenders manage the risk of default as part of credit risk by assessing the eligibility of the borrower.

The emergence of information technology and digitalization has brought new changes in credit provision. Consumer behavior in the digital era has undergone significant transformation. The use of smart devices and the internet has increased accessibility and convenience in financial transactions. The volume of digital-based financial transactions increased from 2015 to 2023 by 672% for SMS/Mobile Banking Transactions and 102% for Internet Banking Transactions. This must be balanced with the ability of financial service providers to adapt and serve the increasing number of users.

Digitalization is a mature stage in digital transformation, where the analog process switches entirely to digital. Moreover, digitalization enables the creation of new processing patterns and digital redefinitions that result in new value propositions in businesses, services, products, processes, and systems. The main goal of digitalization is to optimize business operations by using digital technology automatically and innovatively (Zimmermann et al. 2021).

Digitizing the creditworthiness assessment process can help financial institutions respond to increasing demand effectively and efficiently. Digitizing creditworthiness assessments can shorten the time required to evaluate creditworthiness, increasing the capacity to process more applications in less time.

POJK Number 44/POJK.05/2020 mentions the implementation of risk management for non-bank financial services institutions in Article 3, Paragraph 2, Point c, which states that risk management for non-bank financial services institutions at least includes the adequacy of risk identification, measurement, control, monitoring processes, as well as the risk management information system. The creditworthiness assessment is part of the risk assessment process, specifically measurement and control. Non-performing loans occur when there is a failure to meet the lender's requirements, and repayment cannot be made according to the terms of the credit that have reached maturity (Koç and SevgiLi 2020).

PT. XYZ is a multifinance company that has conducted semi-digitized creditworthiness assessments since its establishment in Indonesia in 2012, where human intervention was still present in the creditworthiness assessment process. On April 27, 2021, it was decided that the digitization process of the creditworthiness assessment would be fully implemented, eliminating human involvement from the entire creditworthiness assessment process. This forms the background for conducting research on multifinance PT. XYZ.

The study was conducted to examine the impact of digitalization of the creditworthiness assessment process on credit risk, particularly behavior regarding First Payment Default (FPD). This research involved collecting and analyzing data from finance companies using a digital creditworthiness assessment process and testing hypotheses about digitalization's impact on FPD. The findings of this study are expected to provide insights to optimize the digitization of the creditworthiness assessment process

The Effect of Digitalization and Human Intervention in Credit Decisions on The Risk of Late Payments in Multifinance Companies PT. XYZ

to better manage credit risk and improve the efficiency and effectiveness of the entire credit process.

RESEARCH METHOD

According to Sugiyono (2019), quantitative research methods refer to approaches based on the philosophy of positivism. This method is used to study a specific population or sample by employing research instruments to collect data. The data analysis is quantitative/statistical, with the primary goal of testing predetermined hypotheses. Sugiyono also emphasized that quantitative research employs statistical tools or quantitative analysis to process the collected data.

The population in this study consisted of all credit application data from PT. XYZ, divided into two periods: before and after full digitization. The period before full digitization covered January to April 2021, while the period after full digitization covered June to September 2021. For the pre-digitization period, only contracts involving credit experts were included, whereas the post-digitization period included all contracts as the process no longer involved credit experts.

The study used a saturated sampling technique, selecting the entire population as the sample. This included 196 contracts from January to April 2021 that involved human intervention in credit decisions, and 337 contracts from June to September 2021 where credit decisions were fully digital.

This research focused on multipurpose credit and used secondary data sourced from PT. XYZ's multifinance system. The data included all types of multipurpose loans related to ownership of goods such as mobile phones.

The study examined various relevant variables to evaluate the impact of digitalization on the creditworthiness assessment process and its outcomes. The following are the variables used in this research, along with their definitions, measurement methods, and observation details:

Table 1. Selected variables

No.	Column Name	Data Type	Information	Source
1	TENOR	Ordinal	Loan Tenor Length (Months) 1 = 6 months 2 = 7 months 3 = 9 months 4 = 12 months	(Zhu <i>et al.</i> 2024)
2	SCORE	Interval	Score From the digital scoring process 1 = 0 – 0.045 2 = 0.046 – 0.062 3 = 0.063 – 0.082 4 = >=0.083	(Pickert 2017)
3	FLOWER	Interval	Effective Interest 1 = <=0.25 2 = 0.26 – 0.50 3 = 0.51 – 0.75 4 = 0.76 – 0.10 5 = > 0.10	(Jansen <i>et al.</i> 2024 Apr 30)
4	MONTHLY INCOME	Ratio	Borrower's Monthly Income	(Jansen <i>et al.</i> 2024 Apr 30)
5	CREDIT EXPERT	Nominal	1 = "Yes" Using human judgment, 0 = "No"	
6	FPS	Nominal	1 = "Yes" Late loan	

The Effect of Digitalization and Human Intervention in Credit Decisions on The Risk of Late Payments in Multifinance Companies PT. XYZ

No.	Column Name	Data Type	Information	Source
			payments, 0 = "No"	

The data processing technique in this study involved dividing the data into two categories: creditworthiness assessments involving human decision-makers and assessments that were fully digitized without human intervention. The data were then tested to evaluate differences in risk levels related to the likelihood of first payment default (FPD) between these two groups. Data processing was conducted using SPSS version 25 to identify relevant patterns and support the analysis.

According to Sugiyono (2018:147), descriptive statistics describe the collected data without making generalizations. They provide an overview using measures such as mean, standard deviation, variance, maximum, minimum, sum, range, kurtosis, and skewness.

Binary logistic regression was used to test the hypothesis. This statistical method examines the relationship between a binary response variable and one or more independent variables. The purpose was to predict the probability of FPD based on the independent variables. Data processing was performed with SPSS. The logistic regression equation used in this study is as follows:

$$Y = \ln\left(\frac{p}{1-p}\right) = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 + \beta_4 X_4 + \beta_5 X_5 + \epsilon$$

$$p = \frac{\exp(\beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 + \beta_4 X_4 + \beta_5 X_5 + \epsilon)}{1 + \exp(\beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 + \beta_4 X_4 + \beta_5 X_5 + \epsilon)}$$

Where:

Y: variable dependent (FPD), 1=Yes (delay), 0=No (no delay).

β_0 : constant or intercept.

$\beta_1, \dots, \beta_5$: The regression coefficient for each independent variable

X_1, X_2, X_3, X_4, X_5 : Independent variables, namely:

X_1 : Tenor (length of loan tenor)

X_2 : Score (score from internal scorecard)

X_3 : Interest (Effective interest rate)

X_4 : Monthly Income (borrower's monthly income)

X_5 : Credit Expert (categorical human assessment)

ϵ : Error term, which reflects other variables that are not included in the model but contribute to Y.

To test the role of independent variables in the model simultaneously, the G test statistics (*omnibus test*) were used. The statistical formula of the G test is as follows:

$$G = -2 \ln \left[\frac{L_0}{L_1} \right]$$

Information:

L0: likelihood without independent variable

L1: likelihood with independent variable

With hypothesis:

H0: $\beta_1 = \beta_2 = \beta_3 = \dots = \beta_p = 0$

H1: there is at least one value $\beta_i \neq 0$, where $i = 1, 2, 3, \dots, p$

The G-test statistics follow the distribution χ^2 (*chi-square*) with the free degree p. The basis for decision-making is to subtract H0 when the statistical calculation value of $G > \chi^2(p)$ or the sig. value $< \alpha = 5\%$. Meanwhile, to test the partial logistic regression

The Effect of Digitalization and Human Intervention in Credit Decisions on The Risk of Late Payments in Multifinance Companies PT. XYZ

coefficient, Wald test statistics are used, with the following hypothesis:

H0: $\beta_i = 0$

H1: $\beta_i \neq 0$, where $i = 1, 2, 3, \dots, p$

The statistical formula of the Wald test is as follows:

$$W_i = \frac{\widehat{\beta}_i}{SE(\widehat{\beta}_i)}$$

$\widehat{\beta}_i$ is the β_i suspect and is the standard error suspect of β . The basis for decision-making is that the value of $\text{sig.} < \alpha = 5\%$ minus H0. $SE(\widehat{\beta}_i)$

Model fit testing on logistic regression using the *Hosmer-Lemeshow test*. The *Hosmer-Lemeshow statistical value* is distributed χ^2 (*chi-square*). The zero hypothesis in the *Hosmer-Lemeshow* test is that the empirical data has a match with the model. If the significance of Hosmer-Lemeshow is greater than 0.05, then the null hypothesis cannot be rejected, so it can be concluded that the model can be used because it has a match with its observations. The accuracy of the prediction in the logistic regression model can also be seen through the *correct classification table*. As with the linear regression model, the logistic regression model also has determination coefficient. The determination coefficient in the logistic regression model can be seen from the value of *Nagelkerke's R square*.

Here is a breakdown of the variables used in the model and the theoretical reasons for their selection:

1. TENOR (X1)

This variable measures the duration of the loan in months. Previous research (**Zhu et al., 2024**) showed that the length of a loan tenor has a significant influence on the risk of default, as longer tenors increase borrowers' risk exposure to changes in financial conditions.

2. SCORE (X2)

The score from the internal scorecard reflects the borrower's risk profile. Based on **Pickert (2017)**, this score is the main indicator of credit risk and can predict the likelihood of FPD from the digital calculation process.

3. FLOWERS (X3)

The effective interest rate is a major financial factor that affects a borrower's ability to pay installments. A study by **Jansen et al. (2024)** found that higher interest rates tend to increase the financial burden on borrowers, thereby increasing the risk of delays.

4. MONTHLY INCOME (X4)

Monthly income reflects the borrower's financial capacity. The same study by **Jansen et al. (2024)** showed that higher income is negatively related to the likelihood of default.

5. CREDIT EXPERT (X5)

These categorical variables reflect whether the manual assessment is carried out by credit experts. This assessment is important because it involves qualitative analysis that may not be covered by automated models.

6. FPD (Y)

As a dependent variable, FPD describes whether there is a delay in payment on the first payment. The model aims to predict the probability of delay based on independent variables.

RESULTS AND DISCUSSION

Logistic Regression Analysis Results

Assumption of Multicollinearity

The multicollinearity test ensures that there is no high correlation between independent variables, so that the estimation of logistic regression parameters is not biased or inefficient. The results of the multicollinearity test can be seen in Table 2.

Table 2 VIF		
Independent	Collinearity Statistics	
	Tolerance	VIVID
Tenor	0,963	1,039
Score	0,958	1,044
Flower	0,972	1,028
Income	0,976	1,025
Credit_expert	0,956	1,046

Source: SPSS Output

Based on the results in Table 2, all free variables had a VIF value that was well below the general threshold of 10 and a tolerance value close to 1, indicating the absence of symptoms of multicollinearity in the model.

Model Eligibility (Goodness of Fit)

The feasibility of the model can be seen from the Omnibus Test of Model Coefficients, the Hosmer-Lemeshow Test, and the Nagelkerke R². The respective values of the model feasibility indicator can be seen in Table 3.

Table 3 Model Eligibility		
Indicators	Value	Information
Omnibus Test	p-value step = 0.041 p-value block = 0.041 p-value model = 0.041	Models with predictors significant
Hosmer-Lemeshow Test	p-value = 0.392	The model matches the data
Nagelkerke R ²	0,081	Predictors are able to predict a response of 8.1%

Source: SPSS Output

The Omnibus Test of Model Coefficients gives a p-value of 0.041 at all levels (step, block, and model), which means that the addition of independent variables together makes a significant contribution to predicting dependent variables. In other words, the model as a whole is statistically meaningful. The Hosmer-Lemeshow test showed that there was no significant difference between the model's prediction results and the actual data, or that there was a correlation between the model and the data. The Nagelkerke R² shows a value of 0.081, which means that the model is only able to explain 8.1% of the variation in the dependent variable. This value is relatively low and indicates that the model's predictability is relatively weak. However, the value of Nagelkerke R² has a weakness, so it needs to be checked again with the AUC value. The AUC score results are as follows.

The Effect of Digitalization and Human Intervention in Credit Decisions on The Risk of Late Payments in Multifinance Companies PT. XYZ

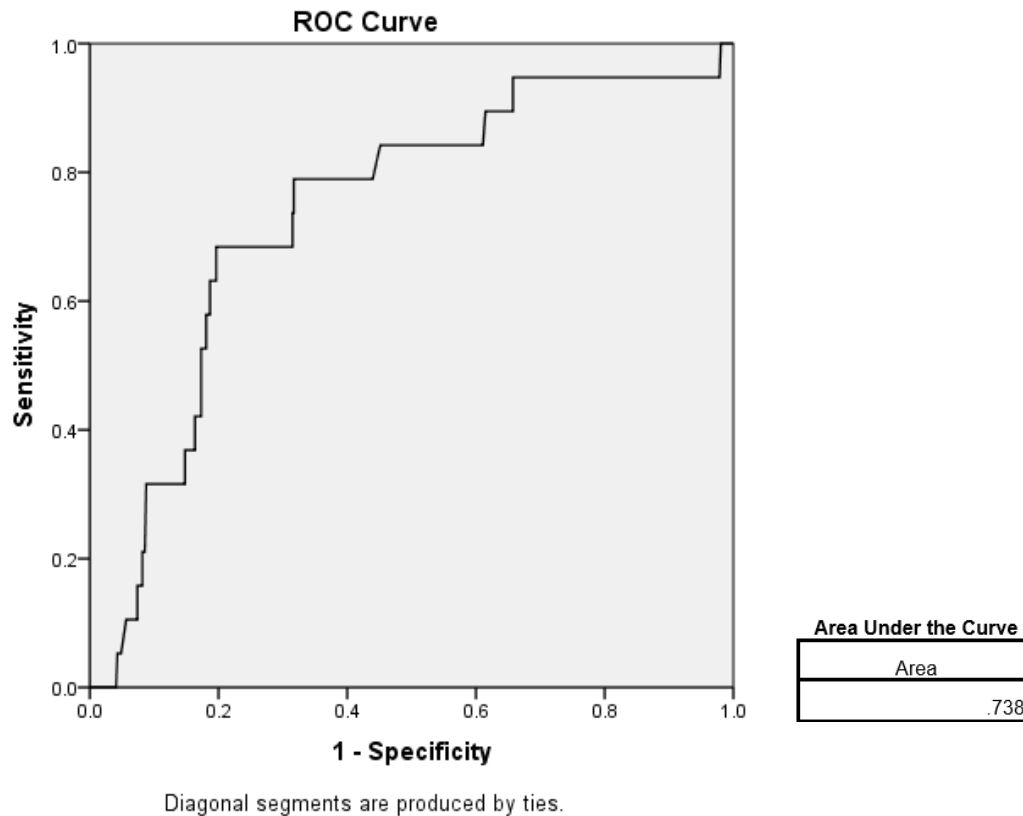


Figure 1. ROC Curve and AUC

Figure 1 shows the *Receiver Operating Characteristic* (ROC) curve of the logistics regression model used to predict *First Payment Default* events (FPD = 1). The *resulting Area Under the Curve* (AUC) value is 0.738, which indicates that the model has a fairly good discriminating ability in distinguishing between debtors who experience FPD (value 1) and those who do not experience FPD (value 0).

Based on guidelines from Hosmer and Lemeshow (2013), the AUC value of 0.738 is classified as quite good ($0.7 \leq \text{AUC} < 0.8$). This shows that the logistic regression model used has reasonable predictive capabilities, with an acceptable level of accuracy for the purposes of this classification.

In addition, AUCs in this range suggest that the model can distinguish well between debtors who are at risk of FPD and those who are not, although there is still room for improvement if a higher level of discrimination is desired (e.g., close to 0.9 which indicates excellent discriminating ability).

Test Coefficients and Odds Ratios

The coefficient test is partially using the Wald test. The results of the Wald test can be seen in Table 4.

Table 4 Wald Test Results

Independent	B	S.E.	Wald	Df	Sig.	Exp(B)
Tenor	-0,010	0,101	0,009	1	0,925	0,991
Score	0,616	0,236	6,798	1	0,009	1,852
Flower	0,394	0,311	1,609	1	0,205	1,484
Income	-0,274	0,496	0,304	1	0,581	0,761

The Effect of Digitalization and Human Intervention in Credit Decisions on The Risk of Late Payments in Multifinance Companies PT. XYZ

Independent	B	S.E.	Wald	Df	Sig.	Exp(B)
Credit_expert	-0,702	0,581	1,456	1	0,227	0,496
Constant	-1,43	7,701	0,035	1	0,853	0,239

Source: SPSS output

Based on the results of the coefficient test partially using the Wald test presented in Table 4, the analysis showed that of the five independent variables tested, only the Score variable had a significant influence on the probability of First Payment Default (FPD = 1), while the other variables showed no significant influence.

In more detail, the Score variable (digital scoring) has a coefficient of 0.616 with a p-value of 0.009, which is smaller than the significance level of 5% ($p < 0.05$). This indicates that digital scores have a significant effect on the likelihood of FPD. The odds ratio of 1.852 indicates that every one unit increase in the digital score will increase the probability of FPD occurring by 85.2%. Thus, a higher digital score is associated with an increased risk of late payments in the first month.

On the other hand, the variables Tenor, Interest, Income, and Credit Expert did not show a significant influence on the probability of FPD. The Tenor variable, with a coefficient of -0.010 and a p-value of 0.925, indicates that the change in the loan tenor does not significantly affect the likelihood of FPD. Although the coefficient is slightly negative, this relationship is not statistically significant, so it cannot be considered a factor determining the risk of late payment.

The Interest variable, with a coefficient of 0.394 and a p-value of 0.205, also did not provide significant evidence of FPD probability. Although the positive relationship between interest rates and the likelihood of FPD can be interpreted descriptively (each increase of one unit of interest rate increases the probability of FPD by 48.4%), but because of the p-value > 0.05 , this relationship is considered statistically insignificant and cannot be used as a policy basis.

The Income variable, with a coefficient of -0.274 and a p-value of 0.581, showed a negative relationship that was not significant with the probability of FPD. Although the revenue log shows that an increase in revenue can lower the probability of FPD by 33.9%, this influence is not strong enough to be considered relevant in the context of this analysis because the p-value is greater than 0.05.

The Credit Expert variable, which reflects human involvement in credit scoring, has a coefficient of -0.702 with a p-value of 0.227, which is also greater than the significance level of 5%. Although descriptively this variable shows that the involvement of human credit assessors can reduce the likelihood of FPD by 40.7%, this effect is not statistically significant and cannot be used as a basis for policymaking.

The model constant value was -1.43 with a p-value of 0.853 ($p > 0.05$). This constant value indicates the basic probability of FPD occurring when all independent variables are zero. In other words, if all the factors tested (tenor, digital score, interest, revenue, and expert credit involvement) are not affected, then the likelihood of FPD will be very low, which is reflected in the odds ratio of 0.239. However, since the p-value for this constant is very high (0.853), it can be concluded that the constant is insignificant and does not play an important role in the model.

Model Accuracy

The accuracy of the model can be seen from the accuracy value of the classification based on the model's prediction. The results of the classification accuracy can be seen in Table 5.

Table 5 Classification Accuracy

Observed	Predicted		Correct Percentage
	0	1	
FPS	0	514	100
	1	19	0
Overall Percentage			96,4

Source: SPSS output

Table 5 shows excellent performance in classifying non-FPD cases (FPD = 0), with a 100% accuracy rate for this category. A total of 514 cases classified as non-FPD were correctly classified as FPD = 0. This indicates that the model has an excellent ability to recognize individuals who did not experience payment delays in the first month.

However, the model shows significant limitations in classifying FPD = 1 cases. Of the 19 cases of FPD=1 that actually occurred, the model completely failed to predict those cases and classify them as FPD=0. Thus, the accuracy rate for the FPD = 1 category is 0%, which indicates that the model is not capable of detecting the risk of late payments in this group.

Although overall, the model has a classification accuracy rate of 96.4%, which seems to indicate good performance, but this figure is greatly influenced by the high accuracy in classifying non-FPDs. This model has an imbalance in predictions, where all FPD=1 predictions are wrong, leading to poor performance in identifying the risk of late payments.

FPD Comparison Test Results based on *Credit Expert*

A comparison of FPDs based on credit experts can be seen from the following table of cross-tabulation and Chi-square tests.

Table 6. Cross-Tabulation and Chi-square Test

Credit_expert	FPS		Total	p-value Chi-square
	0	1		
0	322	15	337	0,148
1	192	4	196	
Total	514	19	533	

Source: SPSS Output

Based on the results of cross-tabulation between the credit expert variable and the occurrence *First Payment Default* (FPD), an overview was obtained that out of a total of 337 observations that did not use credit experts (human assessment), as many as 322 (95.55%) did not experience FPD and 15 (4.45%) experienced FPD. Meanwhile, of the 196 observations using human credit experts, as many as 192 (97.96%) did not experience FPD and only 4 (2.04%) experienced FPD. Although descriptively it appears that the proportion of FPD occurrences is slightly lower in observations using credit experts, the results of the Chi-square test show that this difference is not statistically significant, with a p-value of 0.148 (greater than 0.05). This indicates that there is no significant relationship between the use of credit experts and the incidence of FPD. Thus, based on the available data, there is not enough evidence to conclude that the use of credit experts has an effect on the likelihood of FPD.

The implementation of the creditworthiness assessment process carried out by multifinance PT XYZ

The creditworthiness assessment process at PT XYZ is carried out automatically through a digital *underwriting* system, which utilizes a rating algorithm based on borrowers' historical data. This process generally includes three main stages: (1) data collection, (2) processing and scoring, and (3) automated decision-making. Credit can be approved immediately if the borrower has a certain minimum score based on the parameters that have been programmed. This process reflects full automation and reduces human interference in decision-making.

The results of the study showed that digital scores had a significant effect on FPD. Each increase of one unit of digital score increases the probability of FPD by 85.2%. Thus, this digital score is empirically proven to be a viable and effective indicator in the creditworthiness assessment process. This is consistent with the views of Horváth and Szabó (2019) and Bouncken et al. (2021) who state that digitalization not only automates administrative processes, but also revolutionizes the basic structure of business processes, including financing.

Digital transformation in the context of PT XYZ is reflected in the use of machine learning-based information systems, which are able to classify debtor risks more quickly, objectively, and consistently. This view is reinforced by Scardovi (2017) who emphasizes that the use of social data and financial behavior through multi-source integration can result in more accurate and inclusive risk assessments.

To fully evaluate the performance of the digital system, it can be seen from the influence between human intervention (period January – April 2021) and fully full system without human intervention (period June – September 2021). This human intervention is carried out to consider qualitative factors that cannot always be captured by the system, such as customer history in other institutions, job stability, or other non-formal information. Based on the results of the logistical regression that has been carried out, it is known that this variable is not statistically significant at a significance level of 5%, although the coefficient is negative, which indicates that human intervention tends to lower the risk of FPD.

These findings can be interpreted as an indication that although manual intervention policies are carried out with the aim of risk mitigation, their effectiveness has not been optimal or inconsistent in their implementation. These findings reinforce the results of previous research by Bisht et al. (2022) and Xiong (2023) which stated that digital systems are able to reduce cognitive bias in decision-making and are superior to manual approaches in terms of accuracy and processing speed. This can be caused by several factors, including:

1. Non-standardization of decision criteria by credit analysts.
2. Limited time and data available during the manual review process.
3. The use of intuition-based judgement that has not been supported by advanced analytics.

Thus, the implementation of the creditworthiness assessment system at PT XYZ can be entirely carried out by the system. PT XYZ can focus on improving the accuracy of the scoring system so that decisions can be made fully automatically without sacrificing the quality of the financing portfolio.

Other variables such as tenor, interest rates, and income also do not affect the likelihood of FPD occurring. This is because the company has strictly implemented *know your customers* (KYC) in the credit approval process. Through the implementation of

The Effect of Digitalization and Human Intervention in Credit Decisions on The Risk of Late Payments in Multifinance Companies PT. XYZ

KYC, companies conduct a thorough assessment of the customer's financial profile and credit history before giving approval. This ensures that approved credits have gone through careful selection and are only awarded to individuals who meet certain criteria, thereby reducing the risk of FPD occurring.

Thus, risk mitigation can be carried out through:

1. **Enhancements to the Digital Scoring Model**
The scoring model needs to be continuously calibrated by adding relevant non-traditional variables, such as transaction behavior, digital track record, and sectoral risk profile, as suggested by Kraus et al. (2022).
2. **Implementation of Early *Warning System***
Using real-time data to monitor anomalies in customer payment behavior can improve early detection of potential FPDs.
3. **Implementation of KYC is more intense**
Through the implementation of KYC, companies can conduct more in-depth verification of customers' financial conditions and payment behavior, allowing them to screen high-risk debtors before credit approval is granted. This reduces the likelihood of FPD occurring because only customers with a decent credit profile are accepted. KYC policies have played an important role in mitigating the risk of late payments and ensuring that credit is given only to individuals who meet the set standards.
4. **Use of the Hybrid Scoring Model**
Combining *machine learning* with human input for specific or borderline cases can be a strategic compromise. This is in line with the concept of augmented intelligence in risk management proposed by Ransbotham et al. (2017).
5. **Strengthening Governance and Periodic Model Validation**
Establish a sustainable scoring model validation framework to ensure that the model remains adaptive to new risk dynamics.

Comparison of late payment (FPD) between assessed credit applications and human intervention

Based on the results of the classification of the logistic regression model and descriptive statistics, it was found that the FPD rate in aggregate was 3.56% of all submissions. However, when separated based on human intervention (with variable `credit_expert`), it was observed that there was a difference in the proportion of FPD incidence between the two groups:

1. The group with human intervention (`credit_expert = 1`) had slightly lower levels of FPD than the fully automated group.
2. The non-intervention group (`credit_expert = 0`) showed a higher number of FPD cases in absolute terms, as the proportion was predominantly (the majority of submissions were processed without human intervention).

However, as shown in the results of logistic regression and Chi-square tests, human intervention was not statistically significant in lowering the risk of FPD. This means that even if there are descriptively differences, they are not statistically strong enough to be considered a determining factor. Possible causes of this phenomenon include:

1. Credits reviewed by humans are indeed included in the *borderline* category that naturally has higher risks, so even if interventions are made, the impact is only limited to offsetting the risk.
2. The quality of decisions made by humans is not always better than that of a system, especially if it is not supported by adequate data or tools in the manual evaluation

process.

Manual decision-making has not been systematically documented, so learning from previous decisions cannot be accumulated into future decision-making processes (there is no *feedback* loop in machine learning from human intervention).

These findings provide important implications for credit risk management, namely the need to evaluate the effectiveness of *credit expert teams*, including retraining, standardization of the assessment process, and integration of human analysis results into predictive models in order to produce more accurate hybrid decisions.

Managerial Implications

Important managerial implications for PT XYZ, especially in terms of credit risk management and optimization of the creditworthiness assessment process. First, the effectiveness of digital underwriting systems that have proven significant in predicting FPD indicates that companies can rely on this system as a key tool in the credit scoring process. This encourages management to continue to invest in the development and calibration of the machine learning models used, including by incorporating non-traditional variables such as digital behavioral data and sectoral indicators to improve the accuracy of risk predictions. In addition, the strict implementation of Know Your Customer (KYC) also plays an important role in the customer verification process, ensuring that credit is only given to individuals with a decent risk profile, which can reduce the likelihood of FPD occurring. Second, manual interventions that have not been proven to be statistically significant in lowering FPDs so that the Company's decision to eliminate the creditworthiness process with human intervention is appropriate. Management is advised to review processes in complex or borderline decision-making.

Third, it is important for management to implement an early warning system based on real-time data so that potential defaults can be detected early and interventions can be carried out in a timely manner. Fourth, management is advised to consider a hybrid scoring model approach by integrating advanced analytics and human intuition in a structured manner for cases with high uncertainty. This approach is in line with the principle of augmented intelligence, where technology and humans complement each other, not replace each other.

Fifth, strengthening governance through model validation that is carried out periodically is crucial so that the assessment system remains relevant and adaptive to changes in debtor behavior and market dynamics. Management also needs to ensure that there is a feedback loop from the results of previous decisions, both by the system and by humans, into the model learning process, so that the system can continue to learn and develop continuously. All of these implications underscore the importance of a data-driven, technological, and risk governance approach in supporting the success of a digital-based finance business.

Finally, the elimination of the creditworthiness process with human intervention has significant implications related to cost efficiency through the full implementation of digital underwriting by reducing the human cost component.

CONCLUSION

The study concludes that PT XYZ's implementation of a digital underwriting system, which automates credit eligibility assessments using borrower historical data and rating algorithms, effectively identifies credit risks, as evidenced by digital scores significantly increasing the likelihood of First Payment Default (FPD) by 85.2%. This demonstrates the system's objectivity and consistency in credit decisions while reducing

The Effect of Digitalization and Human Intervention in Credit Decisions on The Risk of Late Payments in Multifinance Companies PT. XYZ

dependence on subjective human judgment. However, no significant difference in FPD rates was found between credits assessed with or without human intervention, likely due to inconsistencies and limitations in manual processes. Additionally, traditional factors like tenor, interest rates, and income did not significantly impact FPD risk, possibly reflecting the company's strict Know Your Customer (KYC) policies focusing on borrower-specific characteristics. Future research could explore integrating more diverse data sources, including behavioral and external data, and applying advanced machine learning techniques to further enhance the predictive accuracy and robustness of the digital credit assessment system.

REFERENCES

- Bisht, D., Singh, R., Gehlot, A., Akram, S. V., Singh, A., & Montero, E. C. (2022). Imperative role of integrating digitalization in the finance firms: A technological perspective. *Electronics*, 11(19), 3252. <https://doi.org/10.3390/electronics11193252>
- Bouncken, R. B., Kraus, S., & Roig-Tierno, N. (2021). Knowledge- and innovation-based business models for future growth: Digitalized business models and portfolio considerations. *Review of Managerial Science*, 15(1), 1-14. <https://doi.org/10.1007/s11846-019-00366-z>
- Horváth, D., & Szabó, R. Z. (2019). Driving forces and barriers of Industry 4.0: Do multinational and small and medium-sized companies have equal opportunities? *Technological Forecasting and Social Change*, 146, 119-132. <https://doi.org/10.1016/j.techfore.2019.05.021>
- Jansen, M., Nguyen, H. Q., & Shams, A. (2024). Rise of the machines: The impact of automated underwriting. *Management Science*. Advance online publication. <https://doi.org/10.1287/mnsc.2024.4986>
- Koç, U., & Sevgili, T. (2020). Consumer loans' first payment default detection: A predictive model. *Turkish Journal of Electrical Engineering & Computer Sciences*, 28(1), 167-181. <https://doi.org/10.3906/elk-1809-190>
- Kraus, S., Durst, S., Ferreira, J. J., Veiga, P., Kailer, N., & Weinmann, A. (2022). Digital transformation in business and management research: An overview of the current status quo. *International Journal of Information Management*, 63, 102466.
- Otoritas Jasa Keuangan. (2018). *Peraturan Otoritas Jasa Keuangan Republik Indonesia Nomor 35/POJK.05/2018 tentang Penyelenggaraan Usaha Perusahaan Pembiayaan*.
- Otoritas Jasa Keuangan. (2020). *Peraturan Otoritas Jasa Keuangan Republik Indonesia Nomor 44/POJK.05/2020 tentang Penerapan Manajemen Risiko bagi Lembaga Jasa Keuangan Non-Bank*.
- Pickert, J. (2017). Risk assessment of unsecured loans - example of entering a new market. *CEREM*, 1(3). <https://doi.org/10.29015/cerem.449>
- Ransbotham, S., Kiron, D., Gerbert, P., & Reeves, M. (2017). Reshaping business with artificial intelligence: Closing the gap between ambition and action. *MIT Sloan Management Review*, 59(1), 1-17.
- Rivai, V., & Permata, A. (2007). *Credit management handbook: Theory, concepts, procedures, and applications guide for student practitioners, bankers and customers*. Raja Grafindo Persada.
- Scardovi, C. (2017). *Digital transformation in financial services*. Springer International Publishing. <https://doi.org/10.1007/978-3-319-66945-8>
- Sugiyono. (2018). *Metode penelitian kuantitatif, kualitatif, dan R&D*. Alfabeta.

The Effect of Digitalization and Human Intervention in Credit Decisions on The Risk of Late Payments in Multifinance Companies PT. XYZ

- Sugiyono. (2019). *Metode penelitian pendidikan pendekatan kuantitatif, kualitatif dan R&D*. Alfabeta.
- Xiong, Y. (2023). The impact of cloud computing and big data technology on digitalization in financial industry. *Financial Economics Research Management*, 6(1). <https://doi.org/10.23977/ferm.2023.060115>
- Zimmermann, A., Schmidt, R., & Jain, L. C. (Eds.). (2021). *Architecting the digital transformation: Digital business, technology, decision support, management* (Vol. 188). Springer International Publishing. <https://doi.org/10.1007/978-3-030-49640-1>
- Zhu, M., Shia, B.-C., Su, M., & Liu, J. (2024). Consumer default risk portrait: An intelligent management framework of online consumer credit default risk. *Mathematics*, 12(10), 1582. <https://doi.org/10.3390/math12101582>