A Simple Credit Rating Prediction Model for FinTech Companies Using SMOTE and MRMR Techniques

Modelo sencillo para la predicción de la calificación crediticia para empresas fintech aplicando técnicas SMOTE y MRMR

Jesús Gopar Sánchez

EGADE Business School, Tecnológico de Monterrey, Mexico

Received: July 10, 2024.
Approved: November 13, 2024.

Abstract

FinTech companies have made the financial industry more efficient and have increased financial inclusion. However, it has also brought new risks to the financial system. Regulators, investors, and researchers are concerned that their financial difficulties could affect the financial system. Our study aims to delve deeper into the effectiveness of machine learning techniques in identifying early warnings of FinTech companies' credit risk impairment. Using commonly employed accounting and market measures in the literature, we created various classifiers to predict FinTech credit ratings. Classification algorithms face a challenge when the number of observations between classes is not equivalent, affecting their performance. Due to the limited size of publicly traded FinTech stocks with an issuer-level credit rating, our database has few observations and is highly imbalanced. The results of our study show that the SMOTE oversampling technique improves the predictive power of machine learning algorithms and that feature selection algorithms such as MRMR allow the generation of less complex and easier-to-understand models. Our results suggest that the KNN classification algorithm has higher accuracy in predicting FinTech's credit ratings.

Keywords: *FinTech, Credit Rating, Machine Learning, SMOTE, MRMR.*
JEL Classification: *C45, G17, G23, G24, G32.*

Resumen

Las empresas fintech han mejorado la eficiencia de la industria financiera y han aumentado la inclusión financiera. Sin embargo, también han incorporado nuevos riesgos al sistema financiero. Los reguladores, los inversionistas y los investigadores están preocupados de que sus dificultades financieras puedan afectar a todo el sistema financiero. Nuestro estudio tiene como objetivo profundizar en la eficacia de las técnicas de machine learning *(aprendizaje automático) para identificar alertas tempranas de deterioro del riesgo crediticio de las fintech. Valiéndonos de medidas contables y de mercado comúnmente empleadas en la literatura, creamos varios clasificadores para predecir las calificaciones crediticias de las fintech. Los algoritmos de clasificación enfrentan un desafío cuando el número de observaciones entre clases no es equivalente, lo que afecta su desempeño. Debido al tamaño limitado de las fintech que cotizan en la bolsa y que tienen una calificación crediticia a nivel de emisor, nuestra base de datos incluye pocas observaciones y está muy desequilibrada. Los resultados de nuestro estudio muestran que la técnica de sobremuestreo SMOTE mejora el poder predictivo de los algoritmos de aprendizaje automático y que los algoritmos de selección de características como MRMR permiten la generación de modelos más sencillos y fáciles de entender. Nuestros resultados sugieren que los algoritmos de clasificación basados en KNN tienen mayor precisión para predecir las calificaciones crediticias de las fintech.*

Palabras clave: *fintech, calificación crediticia, aprendizaje automático, SMOTE, MRMR.*
Clasificación JEL: *C45, G17, G23, G24, G32.*

1. Introduction

The number of FinTech companies (companies that use technology to deliver financial products and services) has been increasing. By 2024, the number of FinTech companies globally amounted to 29,955, representing a Compound Annual Growth Rate (CAGR) of 19.8% since January 2018 (Statista, 2024b). The number of users is expected to increase from 4.7 billion in 2023 to 6.8 billion in 2028 (Statista, 2024a).

FinTech has transformed the traditional banking model by digitally delivering financial products and services (Agarwal & Zhang, 2020; Anagnostopoulos, 2018; Chaudhry et al., 2022). The global FinTech industry transaction value is estimated to reach USD 25.22 trillion by 2028, with a CAGR of 15.5% since January 2018 (Statista, 2023). In particular, the Neobank segment will experience the fastest growth (CAGR of 22.7%) from 2023 to 2028 (Statista, 2023). While FinTech has improved efficiency and increased financial inclusion, its rise presents new risks and vulnerabilities for the financial system (Tello-Gamarra et al., 2022; Treu et al., 2021). Over the past decade, technological advancements, such as mobile money, have significantly impacted financial inclusion. Between 2011 and 2021, the percentage of adults with financial accounts increased by 30 points. Additionally, the proportion of adults engaging in digital payments surged to 57% in 2021, up from 35% in 2014 (Demirgüç-Kunt et al., 2021). On the other hand, recent events have shown that systemic risk can arise from institutions that are not individually critical to the financial system (Cevik, 2024). The possibility that financial difficulties faced by FinTech could spread to the financial system has attracted the attention of regulators, investors, and researchers (Milian et al., 2019; Agarwal & Zhang, 2020; Al-Shari & Lokhande, 2023; Junarsin et al., 2023).

The literature presents conflicting results regarding the impact of FinTech companies on the financial system, which is contingent on their business model (Cevik, 2024). In 2024, Cevik discovered that the overall impact of all FinTech on the financial system, as measured by the bank z-score (which indicates the probability of default of a country's banking system), is negative (Cevik, 2024). This is primarily due to FinTech companies' significant presence in digital lending. In advanced economies, the impact of FinTech companies engaged in fundraising is positive, while in developing economies, the effect is negative. Rapid growth and innovation in the FinTech sector significantly amplify systemic risk, potentially outstripping the ability of regulators to monitor and mitigate risks effectively (Anagnostopoulos, 2018).

Policymakers and regulators are actively working to oversee the FinTech sector and ensure financial stability, as emphasized by Anagnostopoulos (2018). Regulators, therefore, require early warning systems on the credit health of FinTech to take measures to prevent potential contagion to the financial system. In contrast, investors require tools that allow them to identify changes in their credit quality in a timely and reliable manner to make investment decisions. Unfortunately, credit analysis performed by credit rating agencies (CRAs) is a lengthy and costly process based on a combination of qualitative and quantitative methods. Among the quantitative models commonly found in the literature are the multiple discriminant analysis (Altman, 1968), the logistic model (Durand, 1941), the structural credit risk model (Merton, 1974), the multivariate adaptive regression splines (MARS) (Friedman, 1991), the J. P. Morgan's CreditMetrics™ model (Gupton et al., 1997), the McKinsey & Company's CreditPortfolioView™ model (Wilson, 1998), and the KMV model (Kealhofer et al., 1997).

In recent years, there has been a significant increase in the application of machine learning techniques in various fields. The financial market prediction has gained considerable attention in research (Henrique et al., 2019). These techniques have been widely applied in credit rating prediction, which is crucial for lenders and borrowers in the financial industry (Dastile et al., 2020). Unlike statistical models, machine learning-based credit rating forecasting models can capture non-linear relationships in financial variables and are not limited by statistical assumptions. Within this category, models based on artificial neural networks (ANN) and support vector machines (SVM) have demonstrated superior performance to other algorithms, with accuracy ranging from 36% to 88.44% and 60.1% to 89.76%, respectively (Golbayani et al., 2020). Galil et al. (2023) found that while SVM outperforms classification and regression trees (CART) in accuracy, the latter is superior in interpretability. Li et al. (2020) found that random forests (RF) showed the highest accuracy in predicting credit scores, followed by ANNs.

Our study aims to close the knowledge gap by exploring the effectiveness of machine learning techniques in identifying early warnings of credit risk impairment in FinTech companies. This research will help regulators prevent potential contagion within the financial system and enable investors to make more informed decisions. To the best of our knowledge, the use of these techniques to predict credit ratings in the FinTech sector has not been previously studied.

This study is structured as follows: the database and methodology are described in Section 2; Section 3 discusses the results of our research; and Section 4 presents conclusions and future lines of research.

The Anáhuac Journal, **Vol. 24, núm. 2, 2024.**
https://doi.org/10.36105/theanahuacjour.2024v24n2.2516

2. Database and Methodology

2.1 Data

Our sample includes all publicly traded FinTech companies globally with an issuer-level credit rating from Standard & Poor's (S&P Capital IQ, 2024). The analysis period covers the years 2021 to 2023. In our study, the target variable (dependent variable) is the credit rating, and the features (independent variables) are commonly employed accounting and market measures in the literature. All data were obtained from Capital IQ and Bloomberg (2024). Our original database includes 34 credit-rating observations with an imbalanced number of observations in each class. For example, there is only one observation in the database for the AA-, A+, and BBB+ ratings, while the BB- rating has six observations.

Also, our database does not contain observations across the entire S&P rating scale. To solve this problem, Doumpos et al. (2015) regrouped the companies in their sample into five classes: (1) AA- to AAA; (2) A-, A and A+; (3) BBB-, BBB, BBB+; (4) BB-, BB, BB+; and (5) D to B+. Regrouping credit ratings into classes is problematic because the distance between credit ratings is unknown. In our study, we used, as a criterion to solve this problem, the description of the S&P rating scale (S&P Global, 2024), which groups ratings into three categories based on the ability of issuers to meet their financial obligations. P 1 presents the S&P rating scale and regrouping based on the issuers' capacity in Table 1 (see Table 1).

Table 1. Distribution of Credit Ratings After Class Regrouping

Grade	S&P Rating Scale	Capacity	Credit Rating Classes		Number of Fintech Firms	As a % of Total
Investment Grade	AAA	Strong	Class 1			
	AA+	Strong	Class 1			
	AA	Strong	Class 1			
	AA-	Strong	Class 1			
	A+	Strong	Class 1			
	A	Strong	Class 1			
	A-	Strong	Class 1			
				Subtotal	5	14.7%
	BBB+	Adequate	Class 2			
	BBB	Adequate	Class 2			
	BBB-	Adequate	Class 2			
				Subtotal	11	26.5%

Grade	S&P Rating Scale	Capacity	Credit Rating Classes		Number of Fintech Firms	As a % of Total
Non-Investment Grade	BB+	Vulnerable	Class 3			
	BB	Vulnerable	Class 3			
	BB-	Vulnerable	Class 3			
	B+	Vulnerable	Class 3			
	B	Vulnerable	Class 3			
	B-	Vulnerable	Class 3			
	CCC+	Vulnerable	Class 3			
	CCC	Vulnerable	Class 3			
	CCC-	Vulnerable	Class 3			
	CC	Vulnerable	Class 3			
	C	Vulnerable	Class 3			
	D	Vulnerable	Class 3			
				Subtotal	18	58.8%
				Total	**34**	**100.0%**

Source: Prepared by the author.

Some FinTech companies in our database have a banking license, while others do not. Companies with a banking license must disclose information about their risk exposure. Various studies on bank credit risk have utilized the CAMELS methodology, which stands for Capital Adequacy, Asset Quality, Management, Earnings, Liquidity, and Sensitivity to Market Risk, as a guide for selecting feature variables. Since applying the CAMELS methodology to all FinTech companies in our database is not feasible, we selected the features by researching common financial ratios and financial variables reported in various CRAs methodologies and the literature.

Some studies use only accounting measures, while others use a combination of accounting and market metrics. For example, Galil et al. (2023) identified that market capitalization and accounting measures, such as interest coverage and dividends, are influential variables in determining credit ratings. Doumpos et al. (2015) used the distance-to-default obtained from a structural model and accounting data as explanatory variables in a classification model to predict credit ratings. Hajek and

Michalak (2013) found that firm size and market value ratios are the most critical parameters in the United States rating methodology. In contrast, the rating process for European firms relies heavily on profitability and leverage ratios. Jiang (2022) found that equity risk measures (i.e., beta, alpha, and idiosyncratic risk) account for an increasing share of the total rating variation. In addition to the features used in these studies and those commonly used in the literature, we included valuation measures such as the spread between the cost of equity (measured through the capital asset model, CAPM) and the return on equity (ROE) as well as measures of quality of earnings such as the M-Score (Beneish, 1999), the accrual ratio, and the cash flow from operations (CFO) to net income ratio.

In our study, we focused on variables that were not strongly correlated with one another. For instance, we used the DuPont identity to represent return on equity (ROE) to minimize redundancy. Table 2 displays the features we used to predict FinTech's credit ratings (see Table 2).

Table 2. Features Used to Predict Credit Ratings

Feature	Type	Symbology
Ebit/Interest	Coverage	ICR
CFO/Interest	Coverage	CICR
Current Assets/Current Liabilities	Liquidity	CR
Working Capital Cycle	Liquidity	WoCC
Growth in Revenue	Growth	Grwt_Rev
Growth in Net Income	Growth	Grwt_NI
Growth in CFO	Growth	Grwt_CFO
Ebit margin	Profitability	Ebit_Mgn
Net margin	Profitability	Net_Mgn
RoCE	Profitability	RoCE
Sales/Assets	Profitability	S_A
Debt/Equity	Solvency	D_E
Debt/Ebit	Solvency	D_Ebit

Feature	Type	Symbology
Assets/Equity	Solvency	A_E
Distance-to-default	Risk	DD
Beta	Risk	Beta
CFO/Net Income	Quality of Earnings	CFO_NI
Accruals Ratio	Quality of Earnings	Accruals
M-Score	Quality of Earnings	M-Score
ROE - Cost of Equity	Valuation	RR_Spread
Market Cap	Valuation	Mkt_Cap

Source: Prepared by the author.

CRAs use a rating-through-the-cycle methodology to capture long-term solvency (Kiff et al., 2013). To reflect FinTech's permanent economic attributes, we calculated the average of the independent variables used during the analysis period. This criterion is consistent with the one of Hajek and Michalak (2013). Table 3 shows the descriptive statistics of the features (see Table 3).

Table 3. Descriptive Statistics of Features*

Panel A: Global Descriptive Statistics											
	ICR	**CICR**	**CR**	**WoCC**	**S_A**	**Grwt_ Rev**	**Grwt_NI**	**Grwt_ CFO**	**Ebit_ Mgn**	**Net_ Mgn**	**RoCE**
Average	-0.10	-0.07	-0.04	0.00	-0.07	0.00	0.00	-0.02	-0.17	0.00	0.00
Median	-0.38	-0.28	-0.28	-0.27	-0.48	-0.21	-0.01	-0.22	-0.32	-0.07	-0.29
Std. Dev.	0.93	0.92	0.99	1.00	1.01	1.00	0.82	0.96	0.95	1.00	1.00
Max	4.29	4.95	5.22	4.35	3.42	4.03	2.43	4.93	2.46	1.76	3.57
Min	-0.79	-0.46	-0.75	-1.23	-1.24	-1.65	-2.06	-1.00	-2.50	-3.53	-1.79
	D_E	**D_Ebit**	**A_E**	**DD**	**Beta**	**CFO_NI**	**Accruals**	**M-Score**	**RR_ Spread**	**Mkt_ Cap**	
Average	-0.03	-0.04	0.00	0.00	0.00	0.00	0.00	0.00	0.00	-0.07	
Median	-0.22	-0.21	-0.18	0.15	0.08	0.19	-0.01	-0.20	-0.12	-0.43	
Std. Dev.	0.94	0.91	1.00	1.00	1.00	1.00	1.00	1.00	1.00	0.94	
Max	5.28	5.09	4.76	1.03	2.40	0.41	1.23	5.63	5.35	3.76	
Min	-0.25	-0.32	-2.25	-5.41	-1.48	-5.46	-4.96	-0.24	-1.73	-0.49	
Panel B: Descriptive Statistics by Credit Rating											
	ICR	**CICR**	**CR**	**WoCC**	**S_A**	**Grwt_ Rev**	**Grwt_NI**	**Grwt_ CFO**	**Ebit_ Mgn**	**Net_ Mgn**	**RoCE**
Class 1											
Average	1.42	1.09	-0.26	-0.13	-0.24	-0.01	0.19	-0.14	1.33	1.05	1.64
Median	0.93	0.11	-0.27	0.03	-0.22	0.00	0.32	-0.11	1.67	1.45	1.45
Std. Dev.	1.69	2.16	0.09	0.76	0.49	0.15	0.29	0.12	1.02	0.72	1.20
Max	4.29	4.95	-0.16	0.86	0.51	0.19	0.45	0.01	2.46	1.76	3.57
Min	0.10	0.00	-0.38	-1.23	-0.70	-0.22	-0.24	-0.29	-0.02	0.20	0.42

Panel B: Descriptive Statistics by Credit Rating											
	D_E	**D_Ebit**	**A_E**	**DD**	**Beta**	**CFO_NI**	**Accruals**	**M-Score**	**RR_ Spread**	**Mkt_ Cap**	
Class 1											
Average	-0.22	-0.22	-0.61	0.38	-0.50	0.19	0.16	-0.16	0.02	1.31	
Median	-0.23	-0.23	-0.26	0.19	-0.19	0.18	-0.06	-0.15	-0.05	0.71	
Std. Dev.	0.04	0.01	0.92	0.55	0.95	0.04	0.53	0.05	0.37	1.99	
Max	-0.16	-0.20	-0.01	1.03	0.66	0.25	1.04	-0.11	0.61	3.76	
Min	-0.25	-0.23	-2.25	-0.34	-1.48	0.15	-0.24	-0.22	-0.41	-0.49	
	ICR	**CICR**	**CR**	**WoCC**	**S_A**	**Grwt_ Rev**	**Grwt_NI**	**Grwt_ CFO**	**Ebit_ Mgn**	**Net_ Mgn**	**RoCE**
Class 2											
Average	-0.37	-0.26	-0.29	-0.48	-0.65	0.12	0.32	0.29	-0.39	-0.12	-0.41
Median	-0.34	-0.27	-0.30	-0.42	-0.64	-0.22	0.04	-0.15	-0.31	-0.03	-0.37
Std. Dev.	0.24	0.17	0.24	0.32	0.51	1.05	0.98	1.54	0.43	1.30	0.30
Max	0.18	0.09	0.22	-0.10	0.72	2.96	2.43	4.93	0.19	1.35	0.26
Min	-0.58	-0.44	-0.75	-1.21	-1.24	-0.76	-1.04	-0.34	-0.81	-3.53	-0.72
	D_E	**D_Ebit**	**A_E**	**DD**	**Beta**	**CFO_NI**	**Accruals**	**M-Score**	**RR_ Spread**	**Mkt_ Cap**	
Class 2											
Average	-0.21	-0.21	-0.06	0.18	0.01	-0.30	-0.30	-0.17	-0.11	-0.12	
Median	-0.22	-0.21	-0.15	0.21	0.18	0.20	-0.07	-0.19	-0.12	-0.17	
Std. Dev.	0.04	0.02	0.36	0.19	0.88	1.72	1.65	0.09	0.07	0.35	
Max	-0.12	-0.16	0.70	0.48	1.66	0.41	1.23	0.06	0.01	0.72	
Min	-0.25	-0.23	-0.37	-0.13	-1.48	-5.46	-4.96	-0.24	-0.19	-0.49	

Panel B: Descriptive Statistics by Credit Rating											
	ICR	**CICR**	**CR**	**WoCC**	**S_A**	**Grwt_ Rev**	**Grwt_NI**	**Grwt_ CFO**	**Ebit_ Mgn**	**Net_ Mgn**	**RoCE**
Class 3											
Average	-0.36	-0.28	0.17	0.33	0.33	-0.07	-0.25	-0.17	-0.45	-0.22	-0.20
Median	-0.50	-0.34	-0.28	-0.09	0.01	-0.22	-0.01	-0.24	-0.43	-0.26	-0.35
Std. Dev.	0.43	0.23	1.32	1.22	1.18	1.13	0.75	0.53	0.78	0.66	0.78
Max	1.18	0.61	5.22	4.35	3.42	4.03	0.97	1.46	1.32	0.79	1.85
Min	-0.79	-0.46	-0.75	-0.55	-1.24	-1.65	-2.06	-1.00	-2.50	-2.07	-1.79
	D_E	**D_Ebit**	**A_E**	**DD**	**Beta**	**CFO_NI**	**Accruals**	**M-Score**	**RR_ Spread**	**Mkt_ Cap**	
Class 3											
Average	0.13	0.11	0.21	-0.21	0.13	0.13	0.14	0.15	0.06	-0.42	
Median	-0.19	-0.21	-0.08	0.11	0.08	0.17	0.01	-0.20	-0.15	-0.47	
Std. Dev.	1.29	1.24	1.23	1.32	1.09	0.35	0.43	1.37	1.38	0.12	
Max	5.28	5.09	4.76	0.53	2.40	0.34	1.00	5.63	5.35	0.04	
Min	-0.25	-0.32	-0.92	-5.41	-1.48	-1.22	-0.63	-0.24	-1.73	-0.49	

*Note: Panel A shows descriptive statistics of features, and Panel B shows descriptive statistics of features by credit rating class.

Source: Prepared by the author.

Our sample's ratios and financial measures show a monotonic relationship with the credit ratings, as described by Metz and Cantor (2006).

2.2 Data Preprocessing

The assumptions of traditional statistical models do not bind machine learning algorithms, but it is assumed that the model's features contribute equally to the prediction of the target variable. The scale of the features influences their predictive importance, so it is necessary to transform them into a standard scale. To improve the efficiency and performance of the algorithms used, we standardized the data with mean 0 and variance 1.

2.3 Resampling

Because of the nature of the problem, credit rating forecasting models use imbalanced datasets, which poses a challenge for model training. In an imbalanced database, the number of observations between classes differs, leading to models with poor predictive performance. This issue arises because most classification algorithms were designed assuming an equal number of observations between classes (Brownlee, 2021). In the literature, two alternatives exist to solve this problem: improving the algorithm or balancing the database (Sundar & Punniyamoorthy, 2019). Ensemble models are generally used to improve the algorithm, combining multiple learning algorithms with low predictive power to improve their accuracy. Data balancing involves using undersampling (i.e., reducing the samples of the majority class) and oversampling techniques (i.e., creating new samples of the minority class).

Dastile et al. (2020) found that only 18% of the top-rated studies have balanced their databases, and that the most common technique is undersampling the majority class.

To solve the class imbalance problem, Chawla et al. (2002) proposed the Synthetic Minority Oversampling Technique (SMOTE), which creates new synthetic examples of the minority class by joining the nearest neighbors in the feature space. Dastile et al. (2020) argue that SMOTE is the recommended methodology for imbalanced databases.

As shown in Table 2, Class 3 is the majority class, and Class 1 is the minority class (see Table 2). Balancing the database is critical to avoiding bias towards the majority class. Due to the limited number of observations in our database, employing an undersampling technique may result in the loss of valuable information and subsequently lead to reduced classifier performance. To balance the database, we oversampled the minority class using the SMOTE family of algorithms for R Studio, developed by Wacharasak Siriseriwan (Siriseriwan, 2021). The SMOTE technique allows the creation of synthetic examples of minority classes that are required to improve the predictive ability of machine learning algorithms (Brownlee, 2021). Our study established an equal distribution with 20 observations in each class.

Following standard practice in the literature, we partitioned the database into three groups: 1) training (60%), 2) validation (20%), and 3) testing (20%), to reduce sampling bias. As our sample size is small, we used the resampling technique known as K-fold cross-validation and thus avoided significantly reducing the training set. This technique allows the model to be evaluated multiple times by creating random data combinations and grouping them into K folds with K-1 training samples and one validation sample. We set the parameter K to 10. Table 4 displays the distribution of observations for each credit rating class after synthetic data has been generated (see Table 4).

Table 4. Database Resampling

Panel A: Original Dataset						
Credit Rating Classes	**Number of Observations**	**As a % of Total**	**Trai-ning**	**Valida-tion**	**Testing**	**Total**
Class 1	5	8.3%	3	1	1	5
Class 2	11	18.3%	7	2	2	11
Class 3	18	30.0%	10	4	4	18
Total	34	56.7%	20	7	7	34
Panel B: SMOTE						
Credit Rating Classes	**Number of Observations**	**As a % of Total**	**Trai-ning**	**Valida-tion**	**Testing**	**Total**
Class 1	20	33.3%	12	4	4	20
Class 2	20	33.3%	12	4	4	20
Class 3	20	33.3%	12	4	4	20
Total	60	100.0%	36	12	12	60

Source: Prepared by the author.

2.4 Supervised Machine Learning Algorithms Used

a) Classification and Regression Trees (CART)

CART is a supervised machine learning algorithm based on logical conditions to classify or predict data. The technique creates binary trees composed of an initial root node, decision nodes, and terminal nodes representing a single feature (f) and a cutoff value I for that feature. This generates the most comprehensive separation of the labeled data while minimizing the classification error. This error is the decision tree's impurity function E(T), with the entropy function and the Gini diversity index being the most used. This recursive binary partitioning process continues by forming smaller and smaller subgroups at the decision nodes until the terminal nodes, where labels are assigned to the input data, are obtained. The key hyperparameter of the algorithm is the maximum depth of the tree.

b) K-Nearest Neighbors (KNN)

KNN is a supervised machine learning algorithm used primarily to solve classification problems. This technique classifies a new observation x_0 by finding the K_i instances with the most votes for the nearest neighbors between it and the N training samples $x_i,\ldots x_n$. Commonly, Euclidean distance is used to find the K_i instances in the training dataset that are most similar to the new observation. KNN classification is generally represented as:

$$p(c|x_0, d, k) = \frac{1}{K} \sum_{i \in N_k(x_0,d)} M(c_i)$$

$N_k(x_0,d)$ are the k closest neighbors of x_0 in terms of d, and M is a marker that takes the value of 1 if true and 0 if false.

c) Support Vector Machine (SVM)

SVM is a supervised machine learning algorithm for classification and regression problems. This technique maximizes the distance (margin) between the separating hyperplane and the training data closest to the hyperplane (i.e., the support vectors). Typically, data cannot be perfectly separated by the hyperplane, so the algorithm can be adapted using a soft-margin classification, which consists of introducing a tuning parameter (C) that allows for some errors in the classification while penalizing them. Alternatively, Kernels or non-linear separating boundaries can be used. The SVM

The Anáhuac Journal, **Vol. 24, núm. 2, 2024.**
https://doi.org/10.36105/theanahuacjour.2024v24n2.2516

algorithm was initially developed for binary classification (an alternative to binary logistic regression), but its application was extended to multiclass classification.

d) Artificial Neural Networks (ANNs)

ANNs are algorithms for classification and regression in supervised and unsupervised learning. Artificial neural networks (ANNs) consist of interconnected nodes known as neurons. These neurons transmit signals from one node to another using a sum operator and an activation function, similar to how nerve impulses are transmitted between neurons in a biological brain. The most common activation functions are the sigmoid and the rectified linear unit (ReLU), which transform the data non-linearly. Information flows through the model via an input layer that receives the data, hidden layers, where learning occurs during training, and an output layer, where the results are obtained.

2.5 Hyperparameter Tuning

Machine learning algorithms involve two errors that cannot be eliminated: bias (related to underfitting) and variance (related to overfitting). The goal is to minimize the total error (bias error plus variance error) by balancing underfitting and overfitting. In practical terms, there is not a standard method for estimating hyperparameters. Instead, heuristics and automated procedures such as grid search, random search, Bayesian optimization, gradient-based hyperparameter optimization, and evolutionary algorithms are among the most used methods for finding the optimal values of hyperparameters.

We employed Bayesian optimization to fine-tune the hyperparameters of the algorithms utilized in our analysis. Bayesian optimization helps identify hyperparameter values that minimize the loss function. Typically, this approach models the algorithm's capacity to accurately predict outcomes on unseen data using a Gaussian process sample (Snoek et al., 2012). The algorithm samples the input space iteratively, gaining insights from each evaluation until it achieves convergence. Table 5 displays the hyperparameters utilized for training the classification algorithms (see Table 5).

Table 5. Hyperparameters of the Models

Panel A: Original Dataset				
Hyperparameter	**CART**	**KNN**	**SVM**	**ANN**
Maximum number of splits	1			
Split criterion	Max. deviance reduction			
Number of neighbors		10		
Distance metric		Correlation		
Distance weight		Squared inverse		
Box constraint level			0.66	
Kernel function			Linear	
Number of fully connected layers				1
Activation				ReLU
Lambda				0.0626
First layer size				90
Panel B: SMOTE Dataset				
Hyperparameter	**CART**	**KNN**	**SVM**	**ANN**
Maximum number of splits	35			
Split criterion	Gini's Diversity Index			
Number of neighbors		1		
Distance metric		Correlation		
Distance weight		Inverse		
Box constraint level			43.8513	
Kernel function			Quadratic	
Number of fully connected layers				3
Activation				Tanh
Lambda				2.8607
First layer size				77
Second layer size				14
Third layer size				79

Source: Prepared by the author.

3. Results

We initially utilized the 21 features presented in Table 2 to construct the FinTech credit rating prediction models (see Table 2). In line with standard practice in the literature, we assessed the models' performance by comparing their accuracy (i.e., the percentage of correctly predicted classes). Table 6 shows the accuracy of the models in the validation and test databases (see Table 6). Table 7 shows the confusion matrix (see Table 7).

Table 6. Accuracy by Model

Panel A: Original Dataset				
Algorithm	**Validation**		**Test**	
	Accuracy	**Total Cost**	**Accuracy**	**Total Cost**
CART	63.0%	10	57.1%	3
KNN	81.5%	5	57.1%	3
SVM	63.0%	10	57.1%	3
ANN	63.0%	10	57.1%	3
Panel B: SMOTE Dataset				
Algorithm	**Validation**		**Test**	
	Accuracy	**Total Cost**	**Accuracy**	**Total Cost**
CART	93.8%	3	50.0%	6
KNN	91.7%	4	75.0%	3
SVM	93.8%	3	75.0%	3
ANN	89.6%	5	75.0%	3

Source: Prepared by the author.

Table 7. Confusion Matrix

Panel A: Original Dataset									
CART					**SVM**				
		Predicted					Predicted		
		Class 1	Class 2	Class 3			Class 1	Class 2	Class 3
Actual	Class 1	66.7%		33.3%	Actual	Class 1	66.7%	33.3%	
	Class 2			100.0%		Class 2		71.4%	28.6%
	Class 3			100.0%		Class 3		20.0%	80.0%
KNN					**ANN**				
		Predicted					Predicted		
		Class 1	Class 2	Class 3			Class 1	Class 2	Class 3
Actual	Class 1	66.7%		33.3%	Actual	Class 1	66.7%	33.3%	
	Class 2		100.0%			Class 2		57.1%	42.9%
	Class 3		20.0%	80.0%		Class 3		20.0%	80.0%
Panel B: SMOTE Dataset									
CART					**SVM**				
		Predicted					Predicted		
		Class 1	Class 2	Class 3			Class 1	Class 2	Class 3
Actual	Class 1	93.8%	6.2%		Actual	Class 1	100.0%		
	Class 2	6.2%	93.8%			Class 2		87.5%	12.5%
	Class 3		6.2%	93.8%		Class 3		6.2%	93.8%
KNN					**ANN**				
		Predicted					Predicted		
		Class 1	Class 2	Class 3			Class 1	Class 2	Class 3
Actual	Class 1	100.0%			Actual	Class 1	93.8%	6.2%	
	Class 2		87.5%	12.5%		Class 2		81.2%	18.8%
	Class 3		12.5%	87.5%		Class 3		6.2%	93.8%

Source: Prepared by the author.

From the tables above, we observe that the average accuracy of the machine learning algorithms used to predict FinTech credit ratings improves in both the validation and test samples. However, the overall cost of the CART algorithm increases when oversampling is used. Table 8 shows the predicted credit rating class based on the test database (see Table 8).

Table 8. Predictions by Model

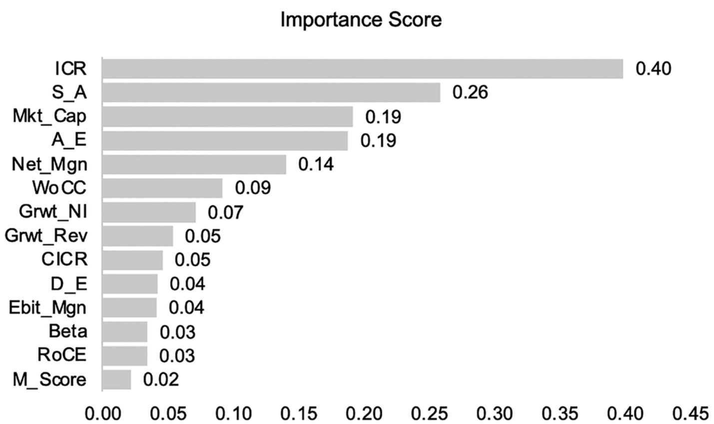
Panel A: Original Dataset					
		Predicted Classes			
Num.	**Actual Classes**	**CART**	**KNN**	**SVM**	**ANN**
1	Class 1	Class 3	Class 3	Class 3	Class 3
2	Class 2	Class 3	Class 3	Class 3	Class 3
3	Class 2	Class 1	Class 2	Class 3	Class 2
4	Class 3	Class 3	Class 3	Class 3	Class 3
5	Class 3	Class 3	Class 3	Class 3	Class 3
6	Class 3	Class 3	Class 2	Class 3	Class 2
7	Class 3	Class 3	Class 3	Class 3	Class 3
Panel B: SMOTE Dataset					
		Predicted Classes			
Num.	**Actual Classes**	**CART**	**KNN**	**SVM**	**ANN**
1	Class 1	Class 1	Class 1	Class 1	Class 1
2	Class 1	Class 3	Class 1	Class 1	Class 1
3	Class 1	Class 1	Class 1	Class 1	Class 1
4	Class 1	Class 1	Class 1	Class 2	Class 2
5	Class 2	Class 2	Class 2	Class 2	Class 2
6	Class 2	Class 3	Class 2	Class 2	Class 2
7	Class 2	Class 2	Class 2	Class 2	Class 2
8	Class 2	Class 3	Class 3	Class 3	Class 2
9	Class 3	Class 2	Class 3	Class 3	Class 3
10	Class 3	Class 3	Class 2	Class 2	Class 2
11	Class 3	Class 2	Class 3	Class 3	Class 3
12	Class 3	Class 1	Class 2	Class 3	Class 2

Source: Prepared by the author.

CRAs consider various factors when evaluating a company's solvency. However, these factors have different weights in their analysis. Focusing on the most critical features helps build simpler, easier-to-understand models (Bellotti & Crook, 2009). Using too many features can also cause overfitting and be computationally expensive.

We used the Minimum Redundancy Maximum Relevance (MRMR) feature selection algorithm proposed by Ding and Peng (2005). This algorithm eliminates redundant features, maintaining the minimum necessary to explain the target variable, unlike methods that select features based on the highest correlation with the target variable. Graph 1 shows the most relevant features to explain FinTech's credit ratings (see Graph 1).

Graph 1. Features with Minimum Redundancy and Maximum Relevance



Source: Prepared by the author.

When applying the MRMR algorithm to the balanced database through the SMOTE technique, it is observed that the interest coverage ratio (ICR) is the most critical variable for predicting the credit ratings of FinTech companies. After ICR, the asset turnover ratio (S_A), market capitalization (Mkt_Cap), leverage (A_E), and net margin are the features with the most significant statistical dependence on the target variable. Our result is consistent with the findings of Galil et al. (2023), who found that market capitalization is a crucial variable in determining credit ratings. Doumpos et al. (2015) found that incorporating distance-to-default alongside other accounting measures enhances the predictive power of credit rating models. However, they also noted that including market capitalization in the feature set reduces the predictive ability of distance-to-default, which appears to be consistent with our study's findings.

Our findings also support Hajek and Michalak (2013), who discovered that equity risk measures, such as the beta coefficient, help explain credit ratings. The cash conversion cycle (WoCC), the asset turnover ratio (S_A), and the net income growth (Grwt_NI) align with most studies on credit scoring and credit ratings. To the best of our knowledge, the existing literature has yet to examine the relationship between quality of earnings measures and credit ratings. A significant finding is that earnings manipulation, as measured by the M-Score, plays a crucial role in explaining credit ratings for FinTech companies. Table 9 presents the hyperparameters for the MRMR-based models (see Table 9), while Table 10 displays the accuracy of the algorithms with the most relevant features (see Table 10).

Table 9. Hyperparameters of MRMR-Based Models

SMOTE and MRMR				
Hyperparameter	**CART**	**KNN**	**SVM**	**ANN**
Maximum number of splits	8			
Split criterion	Max. deviance reduction			
Number of neighbors		1		
Distance metric		Cosine		
Distance weight		Inverse		
Box constraint level			872.1195	
Kernel function			Gaussian	
Kernel scale			3.8526	
Number of fully connected layers				3
Activation				Tanh
Lambda				0.0189
First layer size				17
Second layer size				19
Third layer size				25

Source: Prepared by the author.

Table 10. Accuracy of MRMR-Based Models

SMOTE and MRMR				
Algorithm	**Validation**		**Test**	
	Accuracy	**Total Cost**	**Accuracy**	**Total Cost**
CART	93.8%	3	50.0%	6
KNN	91.7%	4	83.3%	2
SVM	95.8%	2	75.0%	3
ANN	95.8%	2	75.0%	3

Source: Prepared by the author.

As seen in Table 9, the accuracy of the KNN algorithm improved when using features with minimum redundancy and maximum relevance (see Table 9). The in-sample accuracy of the CART, SVM, and ANN algorithms increased, but it remained unchanged in the test dataset. Table 11 shows the confusion matrix for the MRMR-based models (see Table 11), and Table 12 shows the predictions of these algorithms (see Table 12).

Table 11. Confusion Matrix of MRMR-Based Models

CART					**SVM**				
Predicted					Predicted				
Actual		Class 1	Class 2	Class 3	Actual		Class 1	Class 2	Class 3
	Class 1	93.8%	6.2%			Class 1	100.0%		
	Class 2	6.2%	93.8%			Class 2		93.8%	6.2%
	Class 3		6.2%	93.8%		Class 3		6.2%	93.8%
KNN					**ANN**				
Predicted					Predicted				
Actual		Class 1	Class 2	Class 3	Actual		Class 1	Class 2	Class 3
	Class 1	100.0%				Class 1	100.0%		
	Class 2		87.5%	12.5%		Class 2		100.0%	
	Class 3		12.5%	87.5%		Class 3		12.5%	87.5%

Source: Prepared by the author.

Table 12. Prediction by MRMR-Based Model

Predicted Classes					
Num.	**Actual Classes**	**CART**	**KNN**	**SVM**	**ANN**
1	Class 1	Class 1	Class 1	Class 1	Class 1
2	Class 1	Class 3	Class 1	Class 1	Class 1
3	Class 1	Class 1	Class 1	Class 1	Class 1
4	Class 1	Class 1	Class 1	Class 2	Class 2
5	Class 2	Class 2	Class 2	Class 2	Class 2
6	Class 2	Class 3	Class 2	Class 2	Class 2
7	Class 2	Class 2	Class 2	Class 2	Class 2
8	Class 2	Class 3	Class 2	Class 2	Class 2
9	Class 3	Class 2	Class 3	Class 3	Class 3
10	Class 3	Class 3	Class 2	Class 2	Class 2
11	Class 3	Class 2	Class 3	Class 3	Class 3
12	Class 3	Class 1	Class 2	Class 2	Class 2

Source: Prepared by the author.

4. Conclusion and Future Lines of Research

Our study used commonly employed accounting and market measures in the literature to predict FinTech credit ratings. The universe of FinTech companies required to report their financial data under securities market regulation is limited. Within this group, the number of FinTech companies with an issuer-level credit rating is even smaller. Due to the nature of the problem, our database is small in observations and highly imbalanced, which represents a challenge for classification algorithms since their performance improves when the number of observations between classes is equivalent. To solve this problem, oversampling the minority class has been proposed in the literature.

The results of our study suggest that the SMOTE technique, which is based on oversampling the minority class by creating synthetic observations, improves the performance of machine learning algorithms to predict FinTech's credit ratings. Our findings also show that feature selection algorithms such as Minimum Redundancy and Maximum Relevance allow the generation of less complex and easier-to-understand credit rating prediction models and improve the accuracy of KNN algorithms.

The interest coverage ratio is the key factor in determining credit ratings in this specific domain. Our results support the findings of Galil et al. (2023) by finding that company size (measured by market capitalization) is a crucial variable for estimating credit ratings. Credit risk measures such as distance-to-default appear not to have a statistical dependence on FinTech's credit rating when market capitalization is included in the features set. This finding aligns with the conclusions of Doumpos et al. (2015). Our findings suggest that equity risk measures, such as the Beta coefficient, influence credit rating prediction, as Hajek and Michalak (2013) reported. One key result is that the M-Score is essential in explaining FinTech firms' credit ratings. Our findings can help regulators and investors detect early changes in Fintech's credit health.

Incorporating market sentiment into the feature set and exploring alternative techniques to address imbalanced datasets, such as ensemble models, could be valuable directions for future research in FinTech credit rating prediction. Our model heavily depends on credit ratings to assess changes in FinTech's credit risk. However, in future research, exploring alternative credit health measures for publicly traded FinTech companies and expanding the dataset by obtaining accounting and credit risk metrics for privately held FinTech companies would be beneficial in increasing our knowledge about this phenomenon.

References

Agarwal, S., & Zhang, J. (2020). FinTech, Lending and Payment Innovation: A Review. *Asia-Pacific Journal of Financial Studies*, *49*(3), 353–367. https://doi.org/10.1111/ajfs.12294

Al-Shari, H. A., & Lokhande, M. A. (2023). The Relationship Between the Risks of Adopting FinTech in Banks and their Impact on the Performance. *Cogent Business and Management, 10*(1). https://doi.org/10.1080/23311975.2023.2174242

Altman, E. I. (1968). Financial Ratios, Discriminant Analysis and the Prediction of Corporate Bankruptcy. *The Journal of Finance, 23*(4), 589–609. https://doi.org/10.2307/2978933

Anagnostopoulos, I. (2018). Fintech and Regtech: Impact on Regulators and Banks. *Journal of Economics and Business, 100,* 7–25. https://doi.org/10.1016/j.jeconbus.2018.07.003

Bellotti, T., & Crook, J. (2009). Support vector machines for credit scoring and discovery of significant features. *Expert Systems with Applications, 36*(2), 3302–3308. https://doi.org/10.1016/j.eswa.2008.01.005

Beneish, M. D. (1999). The Detection of Earnings Manipulation. *Financial Analysts Journal, 55*(5), 24–36. http://www.jstor.org/stable/4480190

Bloomberg. (2024). Country Risk Premium. Retrieved on June 13, 2024, Bloomberg Professional (database).

Brownlee, J. (2021). *Imbalanced Classification with Python: Better Metrics, Balance Skewed Classes, and Apply Cost-Sensitive Learning.* Self-published.

Cevik, S. (2024) The Dark Side of the Moon? Fintech and Financial Stability. *International Review of Economics, 71,* 421–433. https://doi.org/10.1007/s12232-024-00449-8

Chaudhry, S. M., Ahmed, R., Huynh, T. L. D., & Benjasak, C. (2022). Tail Risk and Systemic Risk of Finance and Technology (FinTech) Firms. *Technological Forecasting and Social Change, 174,* 121191. https://doi.org/10.1016/j.techfore.2021.121191

Chawla, N. v, Bowyer, K. W., Hall, L. O., & Kegelmeyer, W. P. (2002). SMOTE: Synthetic Minority Over-Sampling Technique. *Journal of Artificial Intelligence Research, 16,* 321–357. https://doi.org/10.1613/jair.953

Dastile, X., Celik, T., & Potsane, M. (2020). Statistical and Machine Learning Models in Credit Scoring: A Systematic Literature Survey. *Applied Soft Computing Journal, 91,* 106263. https://doi.org/10.1016/j.asoc.2020.106263

Demirgüç-Kunt, A., Klapper,L., Singer, D., Ansar, S. 2022. The Global Findex Database 2021: Financial Inclusion, Digital Payments, and Resilience in the Age of COVID-19. World Bank. https://doi.org/10.48529/jq97-aj70

Ding, C., & Peng, H. (2005). Minimum Redundancy Feature Selection from Microarray Gene Expression Data. *Journal of Bioinformatics and Computational Biology, 3*(2), 185–205. https://doi.org/10.1142/s0219720005001004

Doumpos, M., Niklis, D., Zopounidis, C., & Andriosopoulos, K. (2015). Combining Accounting Data and a Structural Model for Predicting Credit Ratings: Empirical Evidence from European Listed Firms. *Journal of Banking and Finance, 50,* 599–607. https://doi.org/10.1016/j.jbankfin.2014.01.010

Durand, D. (1941). *Risk Elements in Consumer Installment Financing.* National Bureau of Economy Research (NBER). https://econpapers.repec.org/bookchap/nbrnberbk/dura41-1.htm

Friedman, J. H. (1991). Multivariate Adaptive Regression Splines. *The Annals of Statistics, 19*(1), 1–67. https://www.jstor.org/stable/2241837

Galil, K., Hauptman, A., & Rosenboim, R. L. (2023). Prediction of Corporate Credit Ratings with Machine Learning: Simple Interpretative Models. *Finance Research Letters, 58*, 104648. https://doi.org/10.1016/j.frl.2023.104648

Golbayani, P., Florescu, I., & Chatterjee, R. (2020). A Comparative Study of Forecasting Corporate Credit Ratings Using Neural Networks, Support Vector Machines, and Decision Trees. *North American Journal of Economics and Finance, 54,* 101251. https://doi.org/10.1016/j.najef.2020.101251

Gupton, G. M., Finger, C.C., & Bhatia, M. (1997). *CreditMetrics™: Technical Document.* J.P. Morgan & Company Incorporated. https://www.researchgate.net/publication/301776007_CreditMetrics_-_Technical_Document

Hajek, P., & Michalak, K. (2013). Feature Selection In Corporate Credit Rating Prediction. *Knowledge-Based Systems, 51,* 72–84. https://doi.org/10.1016/j.knosys.2013.07.008

Henrique, B. M., Sobreiro, V. A., & Kimura, H. (2019). Literature Review: Machine Learning Techniques Applied to Financial Market Prediction. *Expert Systems with Applications, 124,* 226–251. https://doi.org/10.1016/j.eswa.2019.01.012

Li, J. P., Mirza, N., Rahat, B., & Xiong, D. (2020). Machine Learning and Credit Ratings Prediction in the Age of Fourth Industrial Revolution. *Technological Forecasting and Social Change, 161*, 120309. https://doi.org/10.1016/j.techfore.2020.120309

Jiang, Y. (2022). Credit Ratings, Financial Ratios, and Equity Risk: A Decomposition Analysis Based on Moody's, Standard & Poor's and Fitch's ratings. *Finance Research Letters, 46,* 102512. https://doi.org/10.1016/j.frl.2021.102512

Junarsin, E., Pelawi, R. Y., Kristanto, J., Marcelin, I., & Pelawi, J. B. (2023). Does Fintech Lending Expansion Disturb Financial System Stability? Evidence from Indonesia. *Heliyon, 9*(9), e18384. https://doi.org/10.1016/j.heliyon.2023.e18384

Kealhofer, S., McQuown, J., & Vasicek, O. (1997). The KMV Model for Credit Portfolio Management. KMV Corporation.

Kiff, J., Kisser, M., & Schumacher, L. (2013). Rating Through-the-Cycle: What Does the Concept Imply for Rating Stability and Accuracy? IMF Working Paper *13*(64). https://doi.org/10.5089/9781475552119.001

Merton, R.C. (1974). On The Pricing Of Corporate Debt: The Risk Structure Of Interest Rates. *The Journal of Finance, 29*(2). 449-470. https://doi.org/10.1111/j.1540-6261.1974.tb03058.x

Metz, A. & Cantor, R. (2006). The Distribution of Common Financial Ratios by Rating and Industry for North American Non-Financial Corporations: July 2006, *Moody's Special Comment*. (Report Number: 98551). https://www.moodys.com/sites/products/defaultresearch/2005700000436062.pdf

Milian, E. Z., Spinola, M. de M., & Carvalho, M. M. de. (2019). Fintechs: A Literature Review and Research Agenda. *Electronic Commerce Research and Applications, 34*, 100833. https://doi.org/10.1016/j.elerap.2019.100833

Siriseriwan, W. (2021). Smotefamily: A Collection of Oversampling Techniques for Class Imbalance Problem Based on SMOTE. [R package smotefamily version 1.4.0]. https://CRAN.R-project.org/package=smotefamily

Snoek, J., Larochelle, H., & Adams, R P. (2012). Practical Bayesian Optimization of Machine Learning Algorithms. *arXiv (Cornell University)*. https://doi.org/10.48550/arxiv.1206.2944

Sundar, R., & Punniyamoorthy, M. (2019). Performance Enhanced Boosted SVM for Imbalanced Datasets. *Applied Soft Computing Journal, 83,* 105601. https://doi.org/10.1016/j.asoc.2019.105601

Statista. (2023). Transaction Value of Fintech Industry Worldwide from 2018 to 2023, with Forecasts from 2024 to 2028, by Segment (in Trillion U.S. dollars). *Statista Digital Market Insights.* Retrieved June 30, 2024. https://www.statista.com/statistics/1384088/estimated-global-fintech-transaction-value-by-segment/

Statista. (2024a). Number of Fintech Users Worldwide from 2023 to 2023, with Forecasts from 2024 to 2028, by Segment (in Billions). *Statista Digital Market Insights.* Retrieved June 27, 2024. https://www.statista.com/statistics/1384328/estimated-fintech-users-by-segment/

Statista. (2024b). Number of Fintechs Worldwide from 2018 to 2024, by Region. *Statista.* Retrieved June 27, 2024. https://www.statista.com/statistics/893954/number-fintech-startups-by-region/

S&P Global (2024). Understanding Credit Ratings. https://www.spglobal.com/ratings/en/about/intro-to-credit-ratings

S&P Capital IQ. (2024). Payment Processors; Payment Service Providers and Gateways; Mobile Wallets; Payments Fraud Management; Money Transfer and Remittance;

Payments Infrastructure: Public company profile. Retrieved on June 13, 2024 (database).

Tello-Gamarra, J., Campos-Teixeira, D., Longaray, A. A., Reis, J., & Hernani-Merino, M. (2022). Fintechs and Institutions: A Systematic Literature Review and Future Research Agenda. *Journal of Theoretical and Applied Electronic Commerce Research, 17*(2), 722-750. https://doi.org/10.3390/jtaer17020038

Treu, J., Elss, V. I., Buono, G., & Winkler, P. (2021). The Rising of Fintech–How the Tech Revolution in Financial Services Represents a Paradigm Shift. *Journal of International Business and Management, 4*(5), 1-8. https://doi.org/10.37227/JIBM-2021-03-718

Wilson, T. C. (1998). Portfolio Credit Risk. Economic Policy Review, Vol. 4, No. 3. http://dx.doi.org/10.2139/ssrn.1028756

https://doi.org/10.36105/theanahuacjour.2024v24n2.2516

About the author

Jesús Gopar Sánchez began his professional career in the financial sector in 1998. For ten years, he held various positions related to credit analysis and investment project evaluation in commercial and development banking. He served as the Director of Transactional Advisory Services at Salles Sainz Grant Thornton, S.C. Since 2015, he has been a partner responsible for the valuation, mergers, and acquisitions practice at Liberty Analysis, S.C. Jesús has provided advisory services to publicly traded and private companies across various industries, focusing on valuation, mergers and acquisitions, strategy, financing, recovery, reorganization, and operational optimization. He has conducted various seminars at the Mexican Institute for Finance Executives (IMEF). He has been a professor for the bachelor's degree in economics courses at Universidad Anáhuac and the Master's Degree in Finance at UNAM. He is a professor in the Department of Accounting and Finance at TEC de Monterrey. Jesús has a bachelor's degree in Economics and an MBA in Finance from TEC de Monterrey. He is currently a PhD in Financial Science candidate at TEC de Monterrey.

jesus.gopar@tec.mx
https://orcid.org/0000-0002-5616-3066