Financial Innovation

To supervise or to self-supervise: a machine learning based comparison on credit supervision

José Américo Pereira Antunes[*]

*Correspondence:
jose.antunes@bcb.gov.br
Central Bank of Brazil, Av
Presidente Vargas, 730, Rio de
Janeiro 20231-044, Brazil

Abstract

This study investigates the need for credit supervision as conducted by on-site banking supervisors. It builds on a real bank on-site credit examination to compare the performance of a hypothetical self-supervision approach, in which banks themselves assess their loan portfolios without external intervention, with the on-site banking supervision approach of the Central Bank of Brazil. The experiment develops two machine learning classification models: the first model is based on good and bad ratings informed by banks, and the second model is based on past on-site credit portfolio examinations conducted by banking supervision. The findings show that the overall performance of the on-site supervision approach is consistently higher than the performance of the self-supervision approach, justifying the need for on-site credit portfolio examination as conducted by the Central Bank.

Keywords: Bank supervision, Machine learning, Loan loss provisions, On-site credit supervision

JEL Classification: C45, G21, G28, M48

Introduction

The need for banking supervision is the result of a chain of events triggered by the market failure that results in the financial system itself. Informational asymmetry between economic units makes the allocation of resources inefficient without a hub to connect them. Hence, financial intermediary is the prerequisite for financial intermediation. Overcoming the first market failure gives rise to the second one: principal–agent problem. Intermediating means capturing someone else's deposits and directing it to a third party at the intermediary's will. From the depositor's perspective, the sounder the bank, the safer the deposits. However, this may not be the case from the perspective of the management, who can decide for a riskier, thus profitable, path. This environment can impede the alignment of interests between depositors and management, making financial intermediation inefficient without the presence of an independent external agent, namely, banking supervision, which asserts the solvency of intermediaries.

The purpose of banking supervision is to keep the financial system sound and safe, ensuring that the financial regulation, the set of rules governing the financial system, is followed (Masciandaro and Quintyn 2015).[1] The flagship of financial regulation is the Basel Accords, which are the policy directives prepared by the Basel Committee on Banking Supervision, a high-level committee of the Bank of International Settlements (BIS). Moreover, this financial regulation is adopted worldwide. The third Basel Accord, which emerged in the aftermath of the Great Financial Crisis (2008/2009), broadened the scope of prudential regulation and embraced liquidity and leverage as relevant microprudential issues. However, the solvency-based perspective remains the focus of prudential regulation, highlighting the capital adequacy ratio as its leading indicator.

From the solvency perspective, keeping the financial system sound and safe means asserting the worth of the bank's assets (Hellwig 2014). In particular, credit portfolio assessment, due to its relevance among assets, is an important task assigned to banking supervision. From the accounting standpoint, the credit portfolio is often measured by amortized cost deduced from the loan loss provision. The loan loss provision is a combination of incurred and expected losses, which is designed to adjust the credit portfolio to its fair value. The role of banking supervision is to assess loan portfolios and check whether banks comply with rules and regulatory requirements, especially the adequacy of loan loss provision to the loan portfolio risk profile.

Although credit portfolio assessment is a classic banking supervision predicate, the Great Financial Crisis slowed down a self-regulation process that gradually increased the reach of internal-based models. The internal ratings-based approach, an important innovation introduced in the Second Basel Accord, allows banks to replace regulatory standardized models with proprietary versions internally developed. Continuous innovation in the financial system, brought about by the technology revolution, may suggest the reignition of this process in the spirit of Stefanadis (2003). De Chiara et al. (2018) analyzed the effect of tighter regulation and powerful supervision in the financial sector and the consequent social costs. The authors argued that the optimal supervisory architecture combines a supervisory regime where direct assessment by a supervisor is always required (mandatory supervision) with a flexible supervision regime where banks self-select the regulatory contract designed for their level of risk.

In this sense, this study investigates the need for credit supervision as conducted by the Central Bank of Brazil (CBB). It builds on a real bank on-site credit examination to compare the performance of a hypothetical self-supervision approach, where banks themselves assess their loan portfolios without external intervention, with CBB's on-site banking supervision. To conduct this experiment, we train two machine learning classification models: (1) a model based on good and bad ratings informed by banks and (2) a model based on past on-site loan portfolio examinations conducted by CBB's banking supervision.

---

[1] Although complementary, banking regulations and banking supervision are separate activities, usually performed by different actors. The former concerns the rules governing the financial system, whereas the latter regards the enforcement of such rules (Masciandaro and Quintyn 2015). In the Brazilian financial system, the National Monetary Council is responsible for banking regulations and the Central Bank of Brazil (CBB) is responsible for banking supervision.

The findings show that CBB's on-site supervision consistently outperforms the self-supervision approach, which justifies the necessity of on-site credit portfolio examination as conducted by CBB.

The remainder of this paper is structured as follows. Section 2 discusses the related literature on financial supervision and loan loss provisioning. Section 3 highlights the Brazilian financial system, credit regulation, and supervision. Section 4 presents the empirical analysis comprising (1) the machine learning algorithm used to develop classification models based on on-site supervision and banks' experience; (2) on-site examination procedure that produced the ground truth against which both supervisory approaches are compared; and (3) the analysis of the results. Section 5 concludes the paper.

Banking supervision and loan loss provisioning regulation

The financial crisis casted doubts over policy certainties ranging from monetary policy to financial regulation and supervision. Barth et al. (2013) and Blanchard (2009) argued that the crisis was the result not only of incomplete regulation but also of ineffective supervision. Bernanke (2010) ascertained that stronger regulation and supervision aimed at problems with underwriting practices and lender's risk management would have been a more effective and surgical approach to constraining the housing bubble than a general increase in interest rates. Their assertion is based on evidence of declining lending standards during the boom.

Viñals et al. (2010) drew lessons from the financial crisis to answer why countries with similar financial systems, operating under the same set of global rules, were less affected than others. The authors argued that, besides the need for better regulations in areas such as capital, liquidity, provisioning, and others, financial supervision was not effective as it should have been. Moreover, they mentioned that to be effective, financial supervision must be intrusive, adaptive, skeptical, proactive, comprehensive, and conclusive. Therefore, a twofold approach was needed. On the one hand, the regulation was broadened and enhanced, including the explicit financial stability mandate, headed by financial stability committees. On the other hand, supervisory skills incorporated additional toolkits to face the forward-looking assessments of risks and the challenging macroprudential dimension the crisis added to supervision.

Masciandaro and Quintyn (2013) stated that financial supervision is the vital link between financial regulation and financial sector stability. Financial supervision acts as an essential complement to financial regulation in the authorities' pursuit of financial stability. The importance of financial supervision as an independent policy area motivated the development of strands in the literature to understand its role. Among the topics that gained attention are as follows: the relationship between the supervision and monetary policy (Goodhart and Schoenmaker 1995; Poloz 2015; Antunes, Moraes and Montes 2016); supervisory architecture (Taylor 1995, 1996); and supervisory governance (Kane 1989; Randall 1993).

Establishing clearly the roles assigned to financial regulation and supervision is a starting point to Freixas and Santomero's (2002) thorough review of the theoretical framework of banking regulation and supervision. On the one hand, financial intermediaries present the solution to market imperfections derived from asymmetric information

problems. On the other hand, regulation and supervision are the response to avoid excessive risk-taking or undesired monopolistic powers that can emerge as consequences of financial intermediaries' actions. Whenever a financial intermediary addresses a market failure, it works as a second-best solution, for it causes another market failure and requires financial regulation and supervision.

Among the market failures addressed by financial intermediaries, such as providing liquidity risk insurance, creating safe assets, screening of potential borrowers, and monitoring customers' actions and efforts (Freixas and Santomero 2002), we draw attention to the screening and monitoring activities as those directly linked to this study.

The quality of banks' assets, as well as the quality of their balance sheets, points to the quality of screening and monitoring activities. Gorton (1988) argued that a bank failure may signal both a weakness limited to the bank and fragility in the system as a whole. Thus, the systemic risk may emerge from microprudential failures that make depositors question the soundness of the financial system. Financial supervision is entitled to assess the quality of screening and monitoring practices. Moreover, for credit, it is responsible to spot bad credit clusters and properly resolve them before they turn into a going-concern problem. Such action avoids spillover effects and mitigates systemic risk.

An extensive number of studies have considered the influence of financial supervision on bank's risk-taking to be relevant. However, results are mixed when it comes to the effects of supervision on financial stability. For instance, Bhattacharya et al. (2002) concluded that intense supervision can improve the timeliness of supervisory intervention, whereas Delis and Staikouras (2009) showed that intense supervision can limit banks' risk-taking. White (2006) defended supervision and regulation as the best instruments to achieve financial stability, whereas Barth et al. (2004, 2008, 2013) argued that the efficiency of financial intermediation, hence financial performance, is reduced by financial supervision. Meanwhile, Brown and Dinç (2011) used a competing risk hazard model for bank survival to study bank failures in 21 emerging market countries in the 1990s and show that a government is less likely to take over or close a failing bank if the banking system is weak, hence establishing a Too-Many-to-Fail effect based on regulatory forbearance.

Brazilian financial system: credit regulation and supervision

In the Brazilian financial system, henceforth financial system, different types of financial institutions coexist, ranging from niche institutions, which explore specific types of activities, to universal banks, which gather many different activities in the same entity. The financial system is complex and well developed. In June/2019, it comprises 178 banks, mounting to 126% of GDP in assets, and 47% of GDP in credit,[2] which makes Brazil an interesting case study. National Monetary Council (NMC) is the financial regulator, and differently from other jurisdictions, CBB is responsible for all aspects regarding financial institutions oversight, from entry to the resolution, concordantly with Barth et al.'s (2004) public interest view.

---

[2] Data collected from the CBB website financial series repository: https://www3.bcb.gov.br/ifdata/.

Table 1 Ratings, provision and delinquency in Brazilian financial regulation

Ratings	Provison (% of the amount due)	Delinquency (days)
AA	$< 0.5\%$	–
A	$\geq 0.5\%; < 1\%$	–
B	$\geq 1\%; < 3\%$	$\geq 15; < 30$
C	$\geq 3\%; < 10\%$	$\geq 30; < 60$
D	$\geq 10\%; < 30\%$	$\geq 60; < 90$
E	$\geq 30\%; < 50\%$	$\geq 90; < 120$
F	$\geq 50\%; < 70\%$	$\geq 120; < 150$
G	$\geq 70\%; < 100\%$	$\geq 150; < 180$
H	100%	≥ 180

In Brazil, the supervisory process partially follows the Twin Peaks model (Group of Thirty 2008; FSI 2018), which recommends supervisory specialization by objectives: prudential monitoring of regulated institutions and oversight of business conduct. Although the Twin Peaks model expects two separate financial supervision authorities to tackle banking supervision, the Brazilian solution is a hybrid model in which an integrated supervisor, namely, CBB, holds both objectives inside the same authority.

The prudential regulation (henceforth banking supervision) is the focus of our analysis. The objective of banking supervision is to assess the soundness of financial institutions, mainly commercial banks, and to assert that regulation is complied. It consists of two cornerstones: examination, or on-site supervision, and monitoring, or off-site supervision. On-site supervision follows a supervision cycle and involves sending supervisory staff to banks to conduct specific examinations. Off-site supervision is a permanent process that analyzes bank's performance and compliance to regulation based on multiple sources of data, as well as the outcomes of on-site supervision.

Brazilian financial regulator, the NMC, still has not adopted IFRS 9 as loan loss provisioning regulation for the financial system. To date, NMC resolution 2682/99 (NMC 1999) defines a loan loss provisioning regulation. It combines expected loss and incurred loss approaches in the same framework. Accordingly, financial intermediaries are bound to assign an individual rating to each credit operation booked in the loan portfolio. As presented in Table 1, nine different ratings reflect the minimum and maximum provisions as percentage points of loan amount due. Whether the credit is due or past due defines the way ratings are assigned. For due credits, banks apply the expected loss approach, in which they assign ratings as they find best, as long as based on consistent credit risk assessment. The expected loss approach assigns ratings compatible to the loss banks expect to face in each credit operation along its lifetime. However, when credit is past due, incurred loss approach can be used, and banks are deemed to assign ratings compatible with the extension of delinquency, as determined by the regulation (see Table 1 for more details).

In June 2019, the amount of loan loss provisions (LLP) in the Brazilian financial system equaled to 6.22% of credit portfolios and 18.2% of equity, reflecting the relevance of

credit activity, hence loan portfolios, to the financial system.[3] The larger the loan portfolio, the more vulnerable banks are to an increase in loan default arising from deteriorating economic conditions (Laeven and Majnoni 2003). Therefore, monitoring and supervising LLP is a crucial microprudential surveillance tool that bank supervisors use to assess banks' loan portfolio quality (Ozili and Outa 2017).

On-site prudential credit supervision works out under two perspectives, namely, credit management and credit risk. Credit management inspections focus on credit processes and compliance of internal credit policies to credit regulation and good practices. As for credit risk inspections, the objective is to assess the quality of credit portfolios and sufficiency of LLP. On-site credit risk supervision focuses on incurred losses insufficiently provisioned and is centered on the borrower's financial performance, which involves intense cash flow analysis, following Antunes et al. (2017) and Antunes et al. (2019). Banks that fall short of incurred loss provisions are demanded to increase them to match their loan portfolios' risk.

From the loan portfolio information banks file monthly at CBB's credit bureau repository, it is possible to derive elementary cash flow variables, such as expected cash flows, received cash flows and disbursed cash flows. Those variables are calculated at loan-level monthly. The analysis is focused on borrowers; therefore, the loan-level cash flow variables are aggregated and turned into borrower-level cash flow variables. Then, the following step is to calculate borrowers' financial performance indices, such as borrower's cash performance (BCP) and borrower's liquidity performance (BLP). These indices are calculated considering a six-month period before the starting date of analysis. Equations (1) and (2) below present the calculation of indices.

$$BCP = \frac{\sum_{t_{-5}}^{t_0} Net\ Received\ Cash\ Flows}{\sum_{t_{-5}}^{t_0} Expected\ Cash\ Flows} \tag{1}$$

$$BLP = \frac{\sum_{t_{-5}}^{t_0} Net\ Received\ Cash\ Flows}{Maximum\ Borrower's Loan Portfolio} \tag{2}$$

Cash-flow-based analysis selects borrowers under cash flow indices (1) and (2). Once selected, on-site credit risk examination assesses borrowers to confirm bad credit risk suggested and possible LLP insufficiency. Borrowers presenting credits over 90-day past due are considered bad and the provision assigned by the bank is compared to regulation disposals.

This study investigates the need for credit supervision as conducted by CBB. It builds on a real bank on-site credit examination to compare the performance of CBB on-site supervision with a hypothetical self-supervision, where banks themselves assess their loan portfolios without external intervention. The experiment trains a machine learning algorithm to develop two classification models: the first one based on past on-site loan portfolio examinations conducted by CBB's banking supervision, and the second one based on good and bad ratings informed by banks. Then, both classification models

---

[3] Data collected from the CBB website financial series repository: https://www3.bcb.gov.br/ifdata/.

Fig. 1 Scheme of the analysis
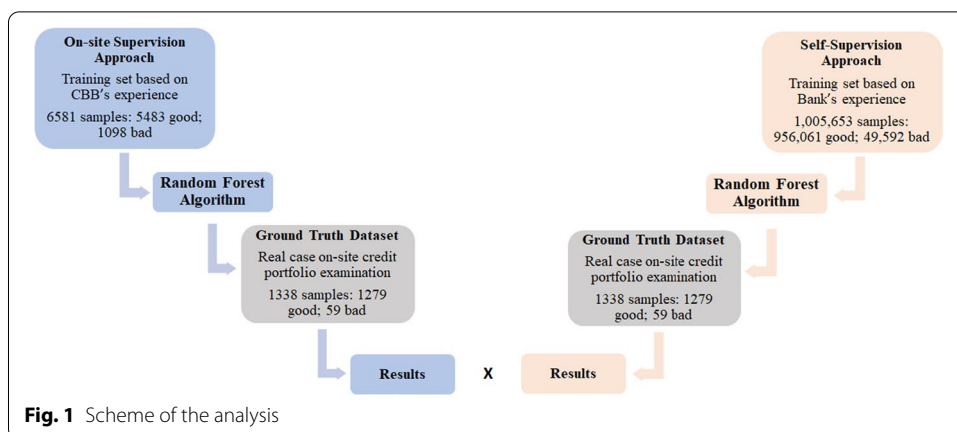
are applied to the real bank loan portfolio to compare performances. Figure 1 presents the procedure adopted to conduct the analysis.

Empirical analysis

Technology revolution reached the financial system. Although the extension and depth of its effects in the conduct of business is yet to be fully realized, the only certainty is that business will change. Alongside the emergence of new entrants in the financial system, supervisory policymakers and standard setters around the world draw attention to risks and opportunities for financial stability. Technology-enabled innovation in financial services (FinTech) develops rapidly and demands a continuous assessment of the adequacy of regulatory frameworks (FSB 2017).

Financial supervision gradually absorbs innovative technology approaches and the terms regtech and suptech were coined to capture a series of initiatives that use innovative technologies in the financial supervision domain (FSI 2018). Regtech accounts for innovative technologies used in support of compliance with the financial regulation, whereas suptech refers to the conduct of financial supervision underpinned by innovative technologies. Kou et al. (2019) surveyed the existing literature on assessment and measurement of financial systemic risk combined with machine learning technologies, including big data analysis, network analysis, and sentiment analysis. They introduced the current research on financial systemic risk with machine learning methods and proposed directions for future work.

Concerning financial supervision, artificial intelligence techniques, mainly those involving machine learning algorithms, are the most used (FSI 2018). Samuel (1959) defined machine learning as the *field of study that gives computers the ability to learn without being explicitly programmed*. Generally, machine learning deals with (automated) optimization, prediction, and categorization, not with causal inference. In other words, classifying whether the borrower is a good or bad credit risk is a machine learning task. However, determining what factors drive credit quality is unlikely to be a machine learning challenge (FSB 2017_A).

The different categories of machine learning algorithms relate to the extent of the human intervention required. In supervised learning, the algorithm receives a set of training data containing labels that classify the observations. Contrarily, unsupervised

Fig. 2 Statistical model × machine learning. Source: Bacham and Zhao (2017)
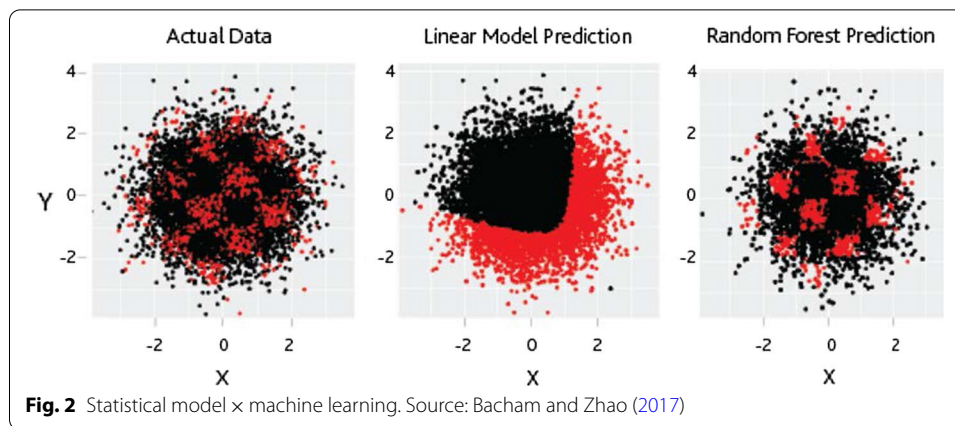
learning detects patterns in the data through similar underlying characteristics, making labels needless (Kou et al. 2014).

Methodology

According to Kirasich et al. (2018), selecting a learning algorithm to implement for a particular application remains an ad hoc process based on fundamental benchmarks such as the classifier's overall loss function and misclassification metrics.

Even so, models for predicting bankruptcy and default events have been the object of intensive research. Kou et al. (2021) proposed a bankruptcy prediction model for SMEs that use transactional data and payment-network-based variables instead of financial (accounting) data. Meanwhile, Wang et al. (2020) investigated credit default risk in P2P lending, arguing that standard binary classifiers are inappropriate in P2P lending because of multiple credit classes, in which misclassification costs vary largely across classes in P2P lending. Using publicly available data from Lending Club, the authors modeled credit rating in P2P lending as a cost-sensitive multi-class classification problem and showed that the cost-sensitive classifiers can significantly reduce the total cost. Meanwhile, Shen et al. (2020) proposed a novel three-stage reject inference learning framework using unsupervised transfer learning and three-way decision theory to infer the possible repayment behavior of rejected credit applicants. The framework was validated on Chinese credit data.

In particular, traditional statistics techniques were compared with artificial intelligence models. Barboza et al. (2017) tested different machine learning models, such as the random forest, to predict bankruptcy 1 year before the event, and compare their performance with results from statistics techniques, such as logistic regression. Using data from 1985 to 2013 on North American firms, the authors concluded that machine learning models show, on average, approximately 10% more accuracy than the traditional models. Comparing the best models, random forest led to 87% accuracy, whereas logistic regression analysis led to 69% accuracy in the testing sample. Addo et al. (2018) built binary classifiers based on machine and deep learning models and used real data to predict loan default probability. Their findings suggest that tree-based models are more stable than models based on multilayer artificial neural networks.
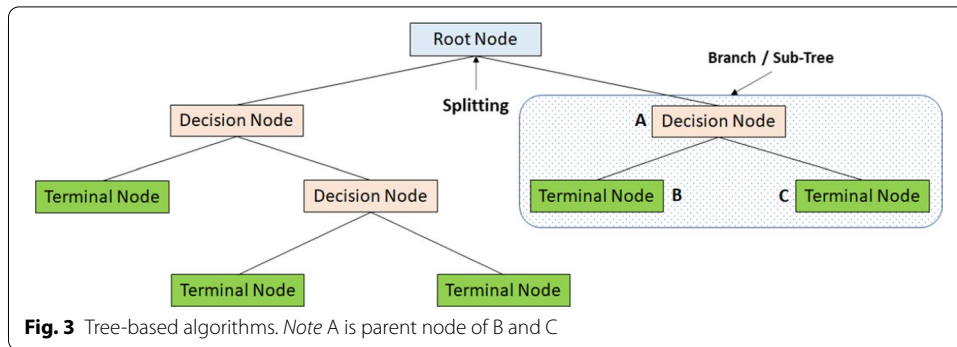
Fig. 3 Tree-based algorithms. *Note* A is parent node of B and C

Table 2 Building a machine learning based classification device

Steps	Description
1	Define the exogenous variables that compound the datasets, known as the matrix of features
2	Define the borrowers from which the matrices of features will be built and whose labels (good or bad borrowers, the endogenous variable) are known. In this study, we use two different sets of borrowers, belonging to the two supervisory approaches investigated
3	Build the two datasets that will be used to train the algorithms, according to the two supervisory approaches analyzed
4	Run (train) the algorithms on the datasets and evaluate their performance
5	Build the validation set from the real bank credit portfolio to be classified
6	Apply the trained algorithms to the validation set and compare the outcomes

A more traditional statistic technique, such as logistic regression, might lead to similar results concerning the comparison between the supervisory approaches. However, one must keep in mind that when dealing with more complex datasets, a linear or continuous statistical model cannot fit complex non-linear and non-monotonic behavior and may not be efficient to segment the class labels, thus leading to poor accuracies. More
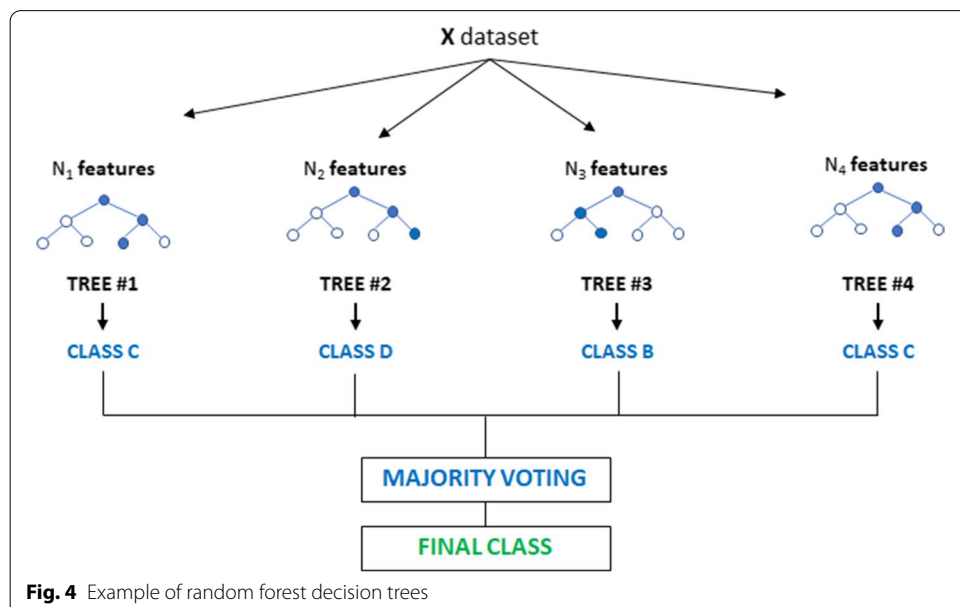

Fig. 4 Example of random forest decision trees

Table 3 Parameters and criteria adopted to classify borrowers according to CBB on-site credit supervision experience

Parameters	Description
Banks	12 previous on-site credit supervision performed by CBB
Extraction dates	From 2015 to 2018
Good borrowers	Considered as AA, A or B by CBB's on-site credit supervision
Bad borrowers	Considered as F, G or H (LLP equal or higher than 50%) by CBB's on-site credit supervision
Materiality	Loans in excess of R$ 200,000[a]

[a] Approximately US$ 40,000

Table 4 Parameters and criteria adopted to classify borrowers according to banks' experience

Parameters	Description
Banks	20 biggest loan portfolios in financial system at the starting date (June/2019)
Extraction dates	December/ 2016; December, 2017; December 2018
Good borrowers	Rated as AA, A or B (LLP inferior to 3%) in each one of the last seven months previously to the dates chosen for extractions
Bad borrowers	Rated as F, G or H (LLP equal or higher than 50%) in each one of the last three months previously to the dates chosen for extractions. A bad borrower was excluded from the dataset if, in the 6-month period after extraction dates: (1) the debt was paid; (2) the debt was reduced; (3) the rating assigned by the bank improved
Materiality	Loans in excess of R$ 200,000[a]

[a] Approximately US$ 40,000

sophisticated algorithms, such as random forest, may then be required because they can learn from a non-linear decision boundary and thus achieve higher accuracy scores, as presented in Fig. 2 (Bacham and Zhao 2017).

The statistical method used in this study is random forest (RF), introduced by Breiman (2001) as an extension of the decision tree method (Breiman et al. 1984).

Random forest consists of a large number of decision trees that operate as an ensemble. Each individual tree in the random forest is made of successive splits of the sample into two leaves, according to a single exogenous variable exceeding or not a threshold.

The quality of each split is measured at the node by an impurity function, such as entropy or information gain, as presented in Eq. (3).

$$Entropy = -p\,log_2(p) - q\,log_2(q) \tag{3}$$

where p and q are the probability of success and failure, respectively, in each node. Entropy measures the degree of disorganization in a system, hence the amount of information necessary to describe it. The node chosen to be split is the one that has the lowest entropy as compared with the parent node. By the end of the process, each tree defines a class prediction, which equals to one vote. The most voted class is the model prediction (Figs. 3 and 4).

However, before applying a machine learning algorithm, one must train it on a dataset with known outcomes, namely, a labeled training set. Therefore, to turn a machine learning algorithm into a classification device, the steps presented in Table 2 shall be followed.

Table 5 Parameters and criteria adopted to classify borrowers in the real bank examination

Parameters	Description
Bank	Medium-sized bank (R\$ 12 billion loan portfolio)[a]
Extraction date	June/2019
Good borrowers	Considered as under 90 day past-due by CBB's on-site credit supervision
Bad borrowers	Considered as over 90 day past-due by CBB's on-site credit supervision
Materiality	Loans in excess of R\$ 10,000

[a] Approximately US\$ 2.400 billion

The first step is to choose the exogenous variables to make up the datasets. Instead of adding every information about borrowers available in CBB's databases, which would lead us to a matrix of features with hundreds of variables and a computational consuming process, we opted by a parsimonious approach. We applied the experience of years of on-site supervision to choose which variables better inform about the risk quality of a borrower. In other words, we developed 26 proxies that reflect on-site experience in classifying good and bad borrowers. The proxies reflect ex-post information related to the borrower's past credit risk behavior because on-site supervision focuses on incurred loss, instead of expected loss (Table 9 in Appendix describes the matrix of features employed in the study and Table 10 in Appendix presents the descriptive statistics for the datasets used).

The next stage is to choose the labeled borrowers whose data will form the datasets. The labeled borrowers are the endogenous variable of the datasets. In particular, "1" is assigned to bad borrowers and "0" is assigned to good borrowers. The first set of borrowers comprises 6,581 samples of good and bad borrowers (5,483 good and 1,098 bad) derived from 12 previous on-site credit portfolio examinations conducted by CBB's banking supervision from 2015 to 2018. Table 3 presents the parameters used to select borrowers and the criteria used to label them as good or bad.

The second set of borrowers gathers 1,005,653 samples of good and bad borrowers (956,016 good and 49,592 bad) obtained from banks' experience and extracted from credit risk information banks file monthly in CBB's repositories.

As for the dataset built from banks' experience, Table 4 presents the parameters used to select borrowers and the criteria used to label them as good or bad.

After these preliminary steps, datasets are gathered through the selection of the variables that constitute the matrix of features for each labeled borrower. In other words, the datasets are the merging of the endogenous variable and the exogenous variables. These datasets are used to train the algorithm.

The training procedure is to apply the machine learning algorithm to the datasets. That allows the algorithm to combine the matrix of features (26 fields of information about each borrower) and the labels to learn the general rule of classification to predict labels in any other out-of-sample dataset. When running the training, the dataset is split into two subsets: the training set and the test set. Following a usual rule of thumb, we

assigned 70% of the dataset to form the training set and the remaining 30% to the test set.

We applied the random forest algorithm (RF) to the datasets (Table 11 in Appendix details the settings used to tune the algorithm). The algorithm is coded in Python and run on scikit-learn package. The algorithm was subject to regularization procedures and K-fold cross validation. The datasets are quite homogeneous, and thus, not much difference is detected between the training and test sets. Therefore, independently of the results obtained in the training phase, trained models are not guaranteed to perform properly in an out-of-sample dataset.

Having trained two models to classify good and bad borrowers, according to the supervisory approaches under comparison, the next step is to apply them to the real bank dataset and compare performances. The real bank dataset comprises 1,338 borrowers, with a minimum amount due of R\$ 10,000.[4] To establish a common ground truth against which the performance of both supervisory approaches can be assessed, the other front of analysis involves the mapping of good and bad borrowers in the real bank dataset through an on-site credit examination. After excluding all borrowers rated as "H" by the bank, that is, 100% provisioned, on-site examination concluded that 1,279 borrowers were considered good ("0") and 59 borrowers were considered bad ("1") (Table 5).

The assumption that the results of the on-site examination are the correct classification, that is, the ground truth, is central to the analysis. For that reason, one could consistently argue that this procedure biases the results toward on-site supervision approach. To oppose this argument, we posit that on-site supervision focuses on incurred losses, which renders less judgmental analysis, and follows objective regulatory rules, which make little room for discretionary decisions, thus mitigating this issue.

The criteria used by on-site credit supervision to identify bad loans are rather straightforward and clarify the frontier between objectivity and discretion. A credit that is over 90-day past due (rating "E," or worse) is considered a bad credit; hence, it is labeled as "1," and "0" otherwise. However, it is common to find evergreened credits, that is, credits artificially kept under the 90-day past-due threshold through successive rollovers. Another practice used to evergreen credits is to distribute the expected cash flow asymmetrically. In other words, small installments, smaller than the interest accrued, are concentrated at the beginning of the credit cash flow, while principal and the remaining interest are placed long in the future. That makes the credit easy to be paid, though artificially. In both cases, the effect of these practices is disregarded, and borrowers are considered bad and thus labeled as "1."

Having mapped the real bank credit portfolio, to evaluate the supervisory approaches is just a matter of matching results.

Results analysis

If the comparison proves that self-supervision approach built upon bank's experience outperforms CBB's on-site credit risk supervision, there would be a strong argument in favor of revising the scope assigned to on-site credit risk supervision. Therefore, the last

---

[4] Approximately U\$ 2,000.

Table 6 Confusion matrix—on-site supervision approach

	Actual values	
	Bad borrower Positive (1)	Good borrower Negative (0)
Predicted values		
Bad borrower Positive (1)	39	31
Good borrower Negative (0)	20	1248

Table 7 Confusion matrix—self-supervision approach

	Actual values	
	Bad borrower Positive (1)	Good borrower Negative (0)
Predicted values		
Bad borrower Positive (1)	17	6
Good borrower Negative (0)	42	1273

Table 8 Performance comparison—on-site supervision and self-supervision approaches

Efficiency metrics	On-site supervision approach	Self-supervision approach
True positive rate (recall)	0.661	0.288
False negative rate	0.339	0.712
False positive rate	0.024	0.005
Accuracy	0.962	0.964
Precision	0.557	0.739
F1 Score	0.605	0.415

step is to compare the performance of on-site banking supervision and the hypothetical self-supervision approaches against the ground truth provided by the real bank on-site examination results.

As discussed before, the role assigned to on-site credit supervision is to detect bad borrowers classified as good ones and to quantify the consequent amount of insufficient incurred LLP. Therefore, picking up bad borrowers is central to the analysis, which makes type-2 errors, that is, classifying bad borrowers as good ones, much worse than type-1 errors. Good borrowers, even those mistakenly classified as such, are not revised during an on-site credit examination. Thus, from the supervisory standpoint, minimizing type-2 errors is crucial, even if the cost is maximizing type-1 errors, because these cases are revised and dumped during the examination.

Another aspect to highlight is that the distribution of good and bad borrowers in credit portfolios is heavily unbalanced, because typically the number of good borrowers is much greater than bad borrowers. Accordingly, some measures used to assess machine learning algorithms performance may present the false sense of efficiency. The

real bank under analysis proves the point. Of the 1,338 borrowers portfolio, only 59 or 4.6% are classified as bad borrowers by on-site supervision staff, whereas 1,279 or 95.4% are deemed good borrowers. An algorithm that classifies the whole portfolio as good borrowers is 95.4% accurate, even failing to catch a single bad borrower.

Tables 6 and 7 present the confusion matrix for both supervisory approaches, while Table 8 presents the efficiency measures.

The confusion matrix is a performance measurement device for machine learning classification algorithms. It combines actual and predicted values to produce the elementary outcomes, which allow one to compute the efficiency metrics used to assess performance. This study consists of a binary classification problem; thus, four outcomes are derived from the confusion matrix, namely, True Positive (TP), True Negative (TN), False-Positive (FP, also type-1 error), and False-Negative (FN, also type-2 error). In the Appendix, Table 12 provides details on the confusion matrix specifics, whereas Table 13 describes the efficiency metrics.

From the confusion matrices, one can notice that the self-supervised approach classified less bad borrowers (23) than on-site supervision approach (70), which is positive from the efficiency standpoint, as it demands less work hours to examine the loan portfolio. However, the efficiency comes at a cost, because the narrower the sample, the harder it is to minimize type-2 error. On this matter, of the 59 bad borrowers in the loan portfolio, 39 bad borrowers were correctly classified by on-site supervision, resulting in a true positive rate of 0.66. As for the type-2 error, the approach failed to identify 20 of 59 bad borrowers, leading to a false-negative rate of 0.33. As for the self-supervised approach, only 17 bad borrowers are correctly classified, a true positive rate of 0.28 and a type-2 error of 0.71, which evidences the 42 bad borrowers it failed to identify.

For the sake of completeness, Table 8 also displays other performance measures. However, due to specificities of the borrowers' classification issue addressed in the study, the information they convey is minor. The number of bad borrowers is insignificant; thus, comments on the false-positive rate are irrelevant. Similarly, the heavily unbalanced distribution of good and bad borrowers in loan portfolios makes accuracy a fragile indicator. Regarding the precision measure, the main focus was on the correctly classified bad borrowers and does not consider false negatives, which, as commented before, is crucial for credit supervision. Therefore, although the self-supervision approach presents a higher precision, it has no significant meaning.

Fig. 5 ROC curves and AUC measure for both supervisory approaches
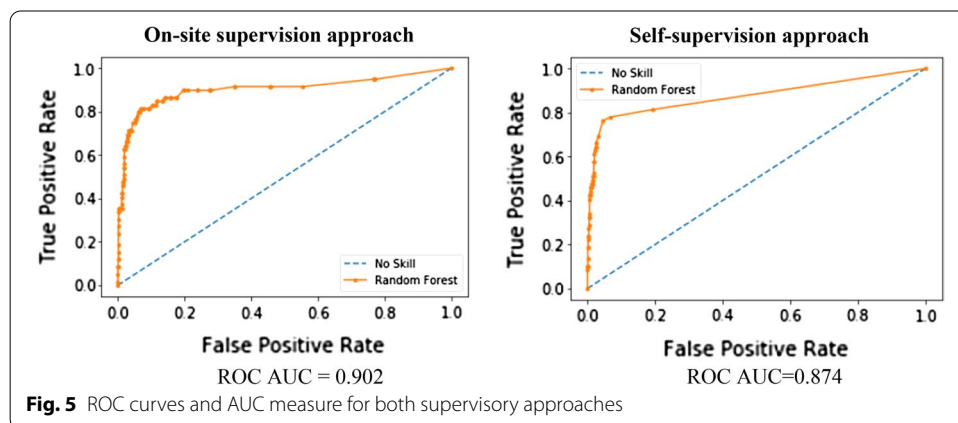
F1 score combines precision and recall (true positive rate) in the same measure. Hence, it informs how precise the classifier is, as well as how robust it is. The greater the F1 score, the better the performance. Although the on-site supervision approach is less precise because it produces more false positives, this approach presents a much better recall than the self-supervision approach, as the number of false negatives is smaller. Consequently, on-site supervision approach shows a better F1 score, that is, a better performance, than the self-supervision approach.

Figure 5 provides ROC curves and AUC measures for both approaches. Once more, the heavily unbalanced dataset compromises the explanatory power of this classical measure. Nonetheless, the performance of CBB's supervisory approach is clearly superior.

In brief, the overall efficiency of CBB's supervisory approach is higher, and the number of bad borrowers unidentified by the self-supervised approach, the type-2 error, is nearly twice as big as CBB's approach. Apart from moral hazard issues, which do not belong to the scope of this analysis, the results of the self-supervision approach could be worse in the absence of CBB's on-site supervision because ratings "F," "G," and "H" that banks assign to their credit portfolios are sometimes imposed by CBB's supervision.

Concluding remarks

This study investigates the need for credit supervision as conducted by the CBB. To the extent the revised literature informs, this study is the first to use a real bank on-site credit examination to compare the performance of supervisory approaches. In particular, two machine learning classification models are employed: the first one is based on past on-site loan portfolio examinations conducted by CBB's banking supervision, and the second one is based on good and bad ratings informed by banks.

The overall efficiency of CBB's supervisory approach is higher, and the number of bad borrowers unidentified by the self-supervised approach, the type-2 error, is nearly twice as big as CBB's approach. The on-site supervision approach is capable of identifying 39 of 59 bad borrowers, which correspond to a true positive rate of 0.66. Meanwhile, the self-supervision approach catches 17 of 59 bad borrowers, denoting a true positive rate of 0.28. From the type-2 error standpoint, on-site supervision approach failed to identify 20 bad borrowers, leading to a false-negative rate of 0.33, whereas self-supervised approach failed to identify 42 bad borrowers, leading to a type-2 error rate of 0.71.

The consistently higher performance of CBB's supervisory approach as compared with the self-supervision approach implies the necessity of on-site credit portfolio examination, as conducted by CBB. Alternatively, from the opposite perspective, the self-supervision approach fails to inform bad credit risk in accordance with the regulation.

The contribution of this study is threefold. First, it innovates by comparing the performance of on-site supervision with that of self-supervision against a common ground represented by a real bank on-site credit examination. Second, Regarding the methodology, this study uses recently available machine learning algorithms to develop classification models based on on-site credit supervision experience and banks' experience to establish the comparison. Third, it asserts the necessity of on-site credit supervision conducted by an independent external agent, such as the Central Bank.

Appendix

See Tables 9, 10, 11, 12 and 13.

Table 9 Variables descriptions

Variable	Description	Format
BCP_0	Borrower's cash performance index: realized cash flows to expected cash flows ratio. Reflects the amount paid by the borrower in relation to the contractual payment forecast. Considers a 6-month period previously to the starting date and does not net off rollovers from the amount realized	From 0 to 1, with 4 decimal places
BCP_1	Borrower's cash performance of first order index: <u>net</u> realized cash flows to expected cash flows ratio. Reflects the <u>net</u> amount paid by the borrower in relation to the contractual payment forecast. Considers a six-month period previously to the starting date and nets off rollovers from the amount realized	From 0 to 1, with 4 decimal places
Grace_period	Credit balance under grace period	Dummy (1 if yes; 0 otherwise)
BLP	Borrower's liquidity performance index: <u>net</u> realized cash flows to the amount due ratio. Reflects the <u>net</u> amount paid by the borrower in relation to the credit balance. Considers a six-month period previously to the starting date and nets off rollovers from the amount realized	From 0 to 1, with 4 decimal places
Revolving_credit_growth	Revolving credit balance rate of change between $MONTH_{-5}$ and $MONTH_0$. (starting date)	From -1 to $+2$, with 4 decimal places
Credit_balance_1	Credit balance rate of change between $MONTH_{-1}$ and $MONTH_0$ (starting date)	From -1 to $+2$, with 4 decimal places
Credit_balance_2	Credit balance rate of change between $MONTH_{-2}$ and $MONTH_{-1}$	From -1 to $+2$, with 4 decimal places
Credit_balance_3	Credit balance rate of change between $MONTH_{-3}$ and $MONTH_{-2}$	From -1 to $+2$, with 4 decimal places
Credit_balance_4	Credit balance rate of change between $MONTH_{-4}$ and $MONTH_{-3}$	From -1 to $+2$, with 4 decimal places
Credit_balance_5	Credit balance rate of change between $MONTH_{-5}$ and $MONTH_{-4}$	From -1 to $+2$, with 4 decimal places
Credit_balance_growth	Credit balance rate of change between $MONTH_{-5}$ and $MONTH_0$. (starting date)	From -1 to $+2$, with 4 decimal places
Provisions_Bank	Credit provisions to credit portfolio ratio in the bank	From 0 to 1, with 4 decimal places
Revolving_credit	Revolving credit to credit portfolio ratio	From 0 to 1, with 4 decimal places
Credit_past_due_FS	Amount of credit past due to the credit portfolio ratio in the financial system (except for the bank)	From 0 to 1, with 4 decimal places
Credit_past_due_FS_1	Amount of credit 90 days past due to the credit portfolio ratio in the financial system (except for the bank)	From 0 to 1, with 4 decimal places
Credit_past_due_FS_2	Amount of credit 180 days past due to the credit portfolio ratio in the financial system (except for the bank)	From 0 to 1, with 4 decimal places
Write-offs_FS	Write-offs to credit portfolio ratio in the financial system	From 0 to 1, with 4 decimal places
Provisions_FS	Credit provisions to credit portfolio ratio in the financial system (except for the bank)	From 0 to 1, with 4 decimal places
Credit_due_360	Amount of credit due up to 360 days to the credit balance ratio	From 0 to 1, with 4 decimal places

Table 9 (continued)

Variable	Description	Format
Credit_due_720	Amount of credit due up to 720 days to the credit balance ratio	From 0 to 1, with 4 decimal places
Credit_due_1440	Amount of credit due up to 1440 days to the credit balance ratio	From 0 to 1, with 4 decimal places
Credit_due_1800	Amount of credit due as of 1800 days from starting date to the credit balance ratio	From 0 to 1, with 4 decimal places
Credit_past_due_Bank	Amount of credit past due to the credit balance ratio in the bank	From 0 to 1, with 4 decimal places
Credit_past_due_Bank_1	Amount of credit 90 days past due to the credit balance ratio in the bank	From 0 to 1, with 4 decimal places
Credit_past_due_Bank_2	Amount of credit 180 days past due to the credit balance ratio in the bank	From 0 to 1, with 4 decimal places
Loan	Loans to credit portfolio ratio	From 0 to 1, with 4 decimal places

Table 10 Variables descriptive statistics

Variable	On-site supervision approach					Self-supervision approach					Ground truth dataset				
	Mean	Std	Min	Máx	Count	Mean	Std	Min	Máx	Count	Mean	Std	Min	Máx	Count
BCP_0	0.795	0.296	0.000	1.000	6581	0.899	0.206	0.000	1.000	1,005,653	0.874	0.217	0.000	1.000	1338
BCP_1	0.506	0.354	0.000	1.000	6581	0.757	0.303	0.000	1.000	1,005,653	0.690	0.331	0.000	1.000	1338
Grace_period	0.052	0.222	0.000	1.000	6581	0.043	0.202	0.000	1.000	1,005,653	0.206	0.404	0.000	1.000	1338
BLP	0.137	0.165	0.000	3.000	6581	0.142	0.158	0.000	3.000	1,005,653	0.202	0.273	0.000	3.000	1338
Revolving_credit_growth	0.580	0.887	0.000	4.000	6581	0.151	0.477	0.000	4.000	1,005,653	0.175	0.470	0.000	3.000	1338
Credit_balance_1	0.023	0.205	−0.970	2.000	6581	0.002	0.116	−1.000	2.000	1,005,653	0.022	0.332	−1.000	2.000	1338
Credit_balance_2	0.018	0.166	−1.000	2.000	6581	0.003	0.111	−1.000	2.000	1,005,653	0.014	0.286	−1.000	2.000	1338
Credit_balance_3	0.016	0.175	−1.000	2.000	6581	0.002	0.120	−1.000	2.000	1,005,653	0.002	0.321	−1.000	2.000	1338
Credit_balance_4	0.034	0.220	−1.000	2.000	6581	0.002	0.115	−1.000	2.000	1,005,653	0.019	0.285	−1.000	2.000	1338
Credit_balance_5	0.031	0.195	−0.995	2.000	6581	0.003	0.118	−1.000	2.000	1,005,653	0.007	0.287	−1.000	2.000	1338
Credit_balance_growth	0.141	0.569	−0.965	4.000	6581	0.013	0.282	−1.000	4.000	1,005,653	0.071	0.569	−1.000	2.000	1338
Provisions_Bank	0.097	0.191	0.000	1.000	6581	0.052	0.210	0.000	1.000	1,005,653	0.026	0.073	0.000	0.700	1338
Revolving_credit	0.078	0.177	0.000	1.000	6581	0.022	0.105	0.000	1.000	1,005,653	0.095	0.274	0.000	1.000	1338
Credit_past_due_FS	0.037	0.141	0.000	1.000	6581	0.016	0.102	0.000	1.000	1,005,653	0.008	0.067	0.000	0.980	1338
Credit_past_due_FS_1	0.023	0.112	0.000	1.000	6581	0.010	0.083	0.000	1.000	1,005,653	0.003	0.036	0.000	0.630	1338
Credit_past_due_FS_2	0.012	0.078	0.000	1.000	6581	0.006	0.066	0.000	1.000	1,005,653	0.002	0.029	0.000	0.630	1338
Write-offs_FS	0.064	0.377	0.000	3.000	6581	0.030	0.267	0.000	3.000	1,005,653	0.027	0.264	0.000	3.000	1338
Provisions_FS	0.088	0.215	0.000	3.000	6581	0.035	0.146	0.000	3.000	1,005,653	0.035	0.125	0.000	1.000	1338
Credit_due_360	0.568	0.372	0.000	1.000	6581	0.290	0.286	0.000	1.000	1,005,653	0.664	0.353	0.000	1.000	1338
Credit_due_720	0.722	0.325	0.000	1.000	6581	0.418	0.323	0.000	1.000	1,005,653	0.865	0.253	0.000	1.000	1338
Credit_due_1440	0.874	0.249	0.000	1.000	6581	0.577	0.336	0.000	1.000	1,005,653	0.957	0.172	0.000	1.000	1338
Credit_due_1800	0.904	0.230	0.000	1.000	6581	0.633	0.333	0.000	1.000	1,005,653	0.966	0.163	0.000	1.000	1338
Credit_past_due_Bank	0.045	0.173	0.000	1.000	6581	0.015	0.088	0.000	1.000	1,005,653	0.025	0.149	0.000	1.000	1338
Credit_past_due_Bank_1	0.023	0.126	0.000	1.000	6581	0.012	0.076	0.000	1.000	1,005,653	0.005	0.068	0.000	1.000	1338
Credit_past_due_Bank_2	0.010	0.085	0.000	1.000	6581	0.008	0.059	0.000	1.000	1,005,653	0.003	0.056	0.000	1.000	1338
Loan	0.314	0.431	0.000	1.000	6581	0.164	0.330	0.000	1.000	1,005,653	0.352	0.465	0.000	1.000	1338

Table 11 Machine learning algorithm settings

Machine learning algorithm	Settings
Random forest	Number of trees: 100; Information criteria: entropy; tenfold cross-validation; Further settings as default

Table 12 Confusion matrix

	Actual values	
	Positive (1)	Negative (0)
Predicted values		
Positive (1)	TP	FP
Negative (0)	FN	TN

where, True Positive (TP): Correctly predicted positive sample

True Negative (TN): Correctly predicted negative sample

False Positive (FP, also type-1 error): Negative sample mistakenly predicted as positive

False Negative (FN, also type-2 error): Positive sample mistakenly predicted as negative

Table 13 Efficiency metrics description

Efficiency metrics	Description	Formulae
Accuracy	Number of correct predictions to total number of samples ratio	$\frac{True\ Positives + True\ Negatives}{Total\ number\ of\ samples}$
Precision	Proportion of correct positive predictions, in relation to the total of positive predictions	$\frac{True\ Positives}{False\ Positives + True\ Positives}$
F1 Score	Harmonic mean between precision and recall	$2 * \frac{1}{\frac{1}{precision} + \frac{1}{recall}}$
True positive rate (recall or sensitivity)	Proportion of correct positive predictions, in relation to all relevant samples, i.e., all positive samples	$\frac{True\ Positives}{True\ Positives + False\ Negatives}$
False positive rate (specificity)	Proportion of negative samples mistakenly predicted as positive, in relation to all negative samples	$\frac{False\ Positives}{False\ Positives + True\ Negatives}$
False negative rate	Proportion of positive samples mistakenly predicted as negative, in relation to all relevant samples, i.e., all positive samples	$\frac{False\ Negative}{False\ Negative + True\ Positive}$

Declarations

Competing interests
The authors declare that they have no competing interests.

References

Addo PM, Guegan D, Hassani B (2018) Credit risk analysis using machine and deep learning models. Working Papers Department of Economics, University of Venice "Ca' Foscari"

Antunes JAP, De Moraes CO, Montes GC (2016) How does capital regulation react to monetary policy? New evidence on the risk-taking channel. Econ Model 56:177–186

Antunes JAP, De Moraes CO, Rodrigues A (2017) How financial intermediation impacts on financial stability? Appl Econ Lett 26:1–5

Antunes JAP, De Moraes CO, Rodrigues A (2019) Financial intermediation analysis from financial flows. J Econ Stud 46(3):727–747

Bacham D, Zhao J (2017) Machine learning: challenges, lessons, and opportunities in credit risk modeling. In: Moody's analytics risk perspectives/managing disruptions, vol IX, pp 1–5

Barboza F, Kimura H, Altman E (2017) Machine learning models and bankruptcy prediction. Expert Syst Appl 83:405–417

Barth J, Caprio G, Levine R (2004) Bank regulation and supervision: what works best? J Financ Intermed 13(2):205–248

Barth J R, Caprio G, Levine R (2008) Bank regulations are changing: for better or worse? World Bank Policy Research Working Paper 4646

Barth J, Lin C, Ma Y, Seade J, Song F (2013) Do bank regulation, supervision and monitoring enhance or impede bank efficiency? J Bank Finance 37(8):2879–2892

Basel Committee on Banking Supervision (2001) The internal ratings-based approach. Consultative Document

Bhattacharya S, Plank M, Strobl G, Zechner J (2002) Bank regulation with random audits. J Econ Dyn Control 26:1301–1321

Bernanke B (2010) Monetary policy and the housing bubble. Speech on January 3, 2010 at the Annual Meeting of the American Economic Association, Atlanta, GA

Blanchard O (2009) The state of macro. Annu Rev Econ Annu Rev 1(1):209–228

Breiman L (2001) Random forests. Mach Learn 45:5–32

Breiman L, Friedman JH, Olshen RA, Stone CJ (1984) Classification and regression trees. Chapman & Hall/CRC, Boca Raton

Brown C, Dinç S (2011) Too many to fail? Evidence of regulatory forbearance when the banking sector is weak. Rev Financ Stud 24(4):1378–1405

De Chiara A, Livio L, Ponce J (2018) Flexible and mandatory banking supervision. J Financ Stab 34:86–104

Delis M, Staikouras P (2009) Supervisory effectiveness and bank risk. Rev Finance. https://doi.org/10.1093/rof/rfq035

Freixas X, Santomero A (2002) An overall perspective on banking regulation. Working Papers 02-1, Federal Reserve Bank of Philadelphia

FSB (2017) Financial stability implications from FinTech: supervisory and regulatory issues that merit authorities' attention. Financial Stability Board

FSB (2017_A) Artificial intelligence and machine-learning in financial services. Market developments and financial stability implications. Financial Stability Board

FSI (2018) Innovative technology in financial supervision (suptech)—the experience of early users. Financial Stability Institute. FSI Insights on policy implementation No 9

Goodhart CAE, Schoenmaker D (1995) Should the functions of monetary policy and bank supervision be separated? Oxf Econ Pap 39:75–89

Gorton G (1988) Banking panics and business cycles. Oxf Econ Pap 40:751–778

Group of Thirty (2008) The structure of financial supervision approaches and challenges in a global marketplace. Washington, DC

Hellwig M (2014) Financial stability, monetary policy, banking supervision, and central banking. Max Planck Institute for Research on Collective Goods

Kane E (1989) Changing incentives facing financial services regulators. J Financ Serv Res 2:265–274

Kirasich K, Smith T, Sadler B (2018) Random forest vs logistic regression: binary classification for heterogeneous datasets. SMU Data Sci Rev 1(3):9

Kou G, Peng Y, Wang G (2014) Evaluation of clustering algorithms for financial risk analysis using MCDM methods. Inf Sci 27(2014):1–12. https://doi.org/10.1016/j.ins.2014.02.137

Kou G, Chao X, Peng Y, Alsaadi FE, Herrera-Viedma E (2019) Machine learning methods for systemic risk analysis in financial sectors. Technol Econ Dev Econ 25(5):716–742. https://doi.org/10.3846/tede.2019.8740

Kou G, Xu Y, Peng Y, Shen F, Chen Y, Chang K (2021) Ko S (2021) Bankruptcy prediction for SMEs using transactional data and two-stage multiobjective feature selection. Decis Support Syst. https://doi.org/10.1016/j.dss.2020.113429

Laeven L, Majnoni G (2003) Loan loss provisioning and economic slowdowns: too much, too late? J Financ Intermed 12(2):178–197

Masciandaro D, Quintyn M (2013) The evolution of financial supervision: the continuing search for the Holy Grail. In: SUERF 50th Anniversary Volume, chap 263–318

Masciandaro D, Quintyn M (2015) The Governance of financial supervision: recent developments. J Econ Surv 1–25

NMC (1999) National Monetary Council Resolution 2682. https://www.bcb.gov.br/pre/normativos/busca/downloadNormativo.asp?arquivo=/Lists/Normativos/Attachments/44961/Res_2682_v2_P.pdf

Ozili P, Outa E (2017) Bank loan loss provisions research: a review. Borsa Istanbul Rev 17(3):144–163

Poloz SS (2015) Integrating financial stability into monetary policy. Bus Econ 50(4):200–205

Randall R (1993) Lessons from New England bank failures, New England Economic Review, Federal Reserve Bank of Boston, May-June:13–38

Samuel A (1959) Some studies in machine-learning using the game of checkers. IBM J 3:211–229

Shen F, Zhao X, Kou G (2020) Three-stage reject inference learning framework in credit scoring via unsupervised transfer learning and three-way decision theory. Decis Support Syst. https://doi.org/10.1016/j.dss.2020.113366

Stefanadis C (2003) Self-regulation, innovation, and the financial industry. J Regul Econ 23(1):5–25

Taylor M (1995) Twin peaks: a regulatory structure for the new century. Centre for the Study of Financial Innovation, London

Taylor M (1996) Peak practice. Centre for the Study of Financial Innovation, London

Viñals J, Fiechter J, Narain A, Elliot J, Tower I, Bologna P, Hsu M (2010) The making of good supervision: learning to say "no". IMF Staff Position Note

Wang H, Kou G, Peng Y (2020) Multi-class misclassification cost matrix for credit ratings in peer-to-peer lending. J Oper Res Soc. https://doi.org/10.1080/01605682.2019.1705193

White WR (2006) Is price stability enough? BIS Working Paper 205

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.