Socio-Economic Planning Sciences 87 (2023) 101555

Contents lists available at ScienceDirect

Socio-Economic Planning Sciences

journal homepage: www.elsevier.com/locate/seps

A network based fintech inclusion platform ☆

Daniel Ahelegbey [a], Paolo Giudici [a,*], Valentino Pediroda [b]

[a] University of Pavia, 27100, Italy
[b] University of Trieste, 34100, Italy

ARTICLE INFO

Keywords:
COVID-19 pandemic
Credit rating models
Financial technologies

ABSTRACT

The paper evaluates, from a sustainable finance viewpoint, a machine learning model implemented in a fintech platform, whose aim is to assign credit ratings. The aim of the model is to learn from both micro economic data and macro economic trends the credit rating of companies that ask for credit. We show that the proposed model is able to reward the companies that have better financial performances with better ratings and, therefore, a higher probability/lower cost of obtaining credit. At the same time, the model correctly takes into account the overall evolution of the economy, favoring financial inclusion for the more penalized economic sectors, particularly during crisis times. The model, its application to credit rating, and its evaluation, are illustrated with reference to more than 100,000 European companies before and during the COVID-19 pandemic crisis. The results shows that, while the impact of the financial variables does not change over time, and particularly during the pandemic, the impact of sectors changes considerably, favoring financial inclusion and resilience.

1. Introduction

The assignment of a credit rating to a company determines the probability that, when the same company asks for a credit, the credit will be granted and, in addition, it determines its price (interest rate).

In the bank based credit rating, the traditional credit scoring models work well especially for firms with large dimension, credit access, cash, and/or collateral. Such models usually do not work for companies with no financial history or collateral even if they have payback capabilities. As a result, traditional credit rating and lending will not help the financial inclusion of a significant proportion of companies, and especially of the new ones. It is therefore crucial to develop novel credit scoring models, based on the application of machine learning, to allow companies without a traditional financial history or relationship, but with good financial conditions, and good networking capabilities, to have a credit rating that allows them to gain access to credit.

Recent advancements are gradually complementing the traditional bank lending system with platform based systems, known as "Financial technologies" (Fintechs). These systems present a paradigm shift from traditional intra-organizational systems to customer-oriented technological (digital) systems, and are gradually gaining ground in many economies across the world, both developed and in development. The emergence of business-to-customer (B2C), customer-to-customer (C2C), provider-oriented business-to-business (B2B) and peer-to-peer (P2P) platforms are typical examples of Fintech systems. Fintech companies offer solutions that differ from traditional institutions regarding the providers and the interaction types as well as regarding the banking and insurance processes they support [1,2].

Platform based systems aim at facilitating credit services by connecting individual lenders with individual borrowers without the interference of traditional banks as intermediaries. In this way, the platforms serve as a digital financial market and can significantly improve customer experience in terms of cost saving and speed of the services to both borrowers and lenders.

Despite their various advantages, digital platform systems also bear risks. Credit rating is purely based on the data available, and not on the financial history and banking relationship of the company. In addition, they are characterized by asymmetry of information and by a strong interconnectedness among their users [3–5] which may affect credit ratings.

There is, therefore, a need to evaluate the trustworthiness of the credit ratings assigned by fintech platforms and, in particular, their accuracy and robustness. The former measures how accurate the credit scores are in the estimation of the actual credit default; the latter measures how the same scores are stable across variations in the input data and/or in the surrounding economic scenarios.

In this paper we contribute to the literature with a fintech credit rating model, based on a multidimensional network model, whose aim is to improve financial inclusion. We also contribute to the literature

☆ Supported by European Commission Horizon2020 PERISCOPE project.
* Corresponding author.
E-mail address: giudici@unipv.it (P. Giudici).

https://doi.org/10.1016/j.seps.2023.101555
Received 19 June 2022; Received in revised form 29 October 2022; Accepted 23 February 2023
Available online 28 February 2023

with an assessment method which evaluates fintech ratings in terms of their accuracy and robustness, advancing the recent works of [6–8] Before describing the model and the evaluation framework, in the next Sections, we now briefly review the most important available rating models.

The use of statistical methods to evaluate the default risk of a company is a well-known problem in finance and in statistics. The relevant literature begins with the work of Beaver and Altman in the late 1960s. Beaver [9] proposed a univariate linear model, and found that a number of indicators could discriminate between matched samples of failed and non-failed firms for as long as five years prior to failure. One year later, Altman [10] proposed the Zscore, which, starting from the financial ratios of the companies, calculates a multivariate linear discriminant model to divide the companies in two likely categories: bankrupt and non-bankrupt. The advantage of the Zscore, which is still one of the most important reference models, is the simplicity and the explainability of its analytical form, with direct correlations between the input parameters (financial ratios) and the discriminant score. A common criticism of the method is that, in reality, the importance of financial variables for the score of a company depends on their economic environment, described for example by the country or economic sector of belonging. If the economic environment is not taken into account, the credit scores may suffer lack of robustness.

The Zscore methods of Altman has prompted the development of many credit scoring papers and real implementations based on statistical learning from the available data. All of them try to explain the probability of default of a company (response variable) by means of a set of explanatory financial variables. Different statistical learning methods have been attempted, ranging from the classical Logit and Probit models [11] and survival models [12,13] to more recent machine learning models such as neural networks [14] and support vector machines [15]. All methods minimize a distance function, based on the likelihood of the data or on the match between the predicted and the actual probabilities of default, in a cross-validation setting.

A peculiarity of these developments, and particularly of those based on machine learning models is that, differently from what occurs in the application of Altman's model, the discriminant coefficients are not defined ex-ante by an expert user, but they are automatically determined by the optimization algorithm. This is, one hand, an advantage, as it allows the wide applicability and reproducibility of scoring models. On the other hand, a possible disadvantage of scoring methods based on machine learning is the necessity to have an accurate and complete database of the companies, data that could be difficult to collect, especially for the companies which have experienced bankruptcy. And, in addition, a rating is a multidimensional opinion about the financial strengths or weaknesses of a company; hence it is important to integrate learning models with financial analysts' knowledge.

The need to combine machine learning with expert reasoning is fulfilled in an important newly emerging benchmark method, based on machine learning: the Maximum Expected Utility (MEU) model, developed at the Standard & Poor's Risk Solutions Group [16,17]. The model is not only an automatic algorithm, but is has a clear economic interpretation and it measures its performance in economic terms. The main idea is to seek a probability measure that maximizes the out-of-sample expected utility of an investor who chooses his investment strategy so as to maximize his expected utility under the model he believes to be efficient. The authors demonstrates how this new numerical method outperforms the Logit and Probit methodology, by adding multidimensional network effects, the interactions between the financial ratios of a company, to obtain a more accurate prediction of the real probability of default.

Another important emerging standard method is RiskCalcTM, employed by Moody's. The model is based on the Merton model [18] in which a firm's future asset value has a probability distribution characterized by its expected value and standard deviation. The method defines a "distance to default" which is the distance in standard deviations between the expected value of the assets and the value of the liabilities at time t. The greater the value of the firm and the smaller its volatility, the lower is the probability of default. The authors demonstrate how the performance of the Merton methodology outperforms that of the Zscores model using the well-known Area Under the ROC Curve predictive accuracy. As for the MEU model, RiskCalcTM achieves a high performance by combining learning methods with expert based input.

Following the previous developments, in this paper we introduce a different machine learning model: the Multi Objective Rating Evaluation (MORE) model, which combines a machine learning methodology, based on a multidimensional network model, with the utilization of expert opinions, notably in the choice of the explanatory variables.

The MORE rating model arises from an intensive data analysis on all available company data: to date, all yearly income statement/balance sheet data (plus register data) of more than 25 million companies worldwide have been used for its development. Company data are aggregated by sectors (defined by NACE codes) and by Country. Once aggregated, sector and country specific factors are inserted as explanatory factors in the MORE scoring model to provide more accurate ratings. Doing so, the macro economic evolution of each sector, within each country, is integrated into MORE. We point out that Modefinance, the owner of the MORE model, is a rating agency regulated by the ESMA (the European Securities and Markets Authority). This implies that the credit ratings obtained from the MORE model are to be validated every year, and that the validation results are provided to the ESMA authority for its external validation.

Differently from the MEU and RiskCalcTM models, which are based on proprietary data, all data behind the MORE model are accessible and, therefore, the performance of the model can be transparently evaluated in terms of accuracy and robustness. We will do so for the recent period of time, which includes the COVID-19 pandemic.

This will allow to evaluate the robustness of the model, that is, to investigate whether and how the MORE ratings have been affected during the COVID-19 pandemic, in terms of their (micro economic) financial drivers and their (macro economic) distribution across sectors.

To evaluate the robustness of the model we will take, as response variable, the variation in the rating classes over two consecutive time periods and evaluate its dependence on time, economic and financial covariates.

From a practical viewpoint, we will consider more than 100,000 European Small and Medium Enterprises, for which a MORE rating has been assigned, and for which balance sheet data is available. Without loss of generality, we will focus on the largest European countries: France, Germany, Italy, Spain, and on the financial periods between 2015 and 2020.

The paper is organized as follows. Section 2 presents the MORE fintech rating model; Section 3 presents the data which will be used to assess the model; Section 4 presents model assessment evaluation; Section 5 presents concluding remarks.

2. A fintech rating model

Modefinance is a FinTech company registered by the ESMA (European Securities and Markets Authority) as a Credit Rating Agency. Although the company does not operate as a peer-to-peer platform, it provides a scoring service for investors [19].

The data used by Modefinance is based on official financial information (balance sheets and income statements) which is publicly available through the network of European Chambers of Commerce. To make data consistent between different European countries, the data needs to be reclassified to minimize the differences between different fiscal legislations.

To determine credit ratings, Modefinance employs the Multi Objective Rating Evaluation (MORE) model. The MORE model is used to assess the level of distress of companies, using data included in financial statements. The basic idea of the model is to first select a set of financial ratios, which are used by expert financial analysts as relevant predictors of bankruptcy. The chosen financial ratios are then combined with the country and sectorial classification of each company, which summarizes the information on the economic environment. This with the purpose of creating a credit rating model that takes into account both micro economic and macro economic data. All financial, country and sectorial variables are then combined in a multidimensional network based machine learning algorithm, which produces a score for each company taking interactions between all variables into account and maximizing the overall predictive accuracy of the model.

Thanks to its interactive dimension, the MORE model can assign a rating to a company even in the presence of missing data, as "similar" companies can provide a complementary information. In particular, differently from what happens with most machine learning models, the MORE model provides an evaluation of the creditworthiness of a company even when the bankruptcy information is missing.

Thanks to a selection of variables which is based on financial knowledge, the MORE model is not solely a statistical learning tool, but is explainable in terms of the financial figures contained in the income statements of the companies and/or as a function of the economic evolution of countries and industrial sectors.

The core idea of the model is to assign credit ratings by observing every aspect of the economic and financial behavior of a company. The more a company is balanced in terms of key financial variables such as: profitability, liquidity, solvency, interest coverage and efficiency, the better the rating class will be.

The variables are chosen combining highly default correlated variables with financial analyst's opinion, separately for each economic sector and country. In this way the model can choose different ratios for different economic backgrounds. The choice of the explanatory variables is very important: they should be a limited number, to avoid multicollinearity and overfit of the model; and they should be representative of the financial and economical behavior of a company, to be explainable.

Once a set of financial ratio variables is selected, the MORE model "converts" each ratio value in a rating attribution, using fuzzy logic. This is based on the assumption that in each financial ratio there is the indication of a company's rating. The fuzzy analysis builds rating intervals from the observed distribution of the values of the financial ratios, using reference values which are typical benchmarks of financial analysts, following corporate finance theories.

More specifically, fuzzy logic is applied in the MORE model as follows. First, MORE assigns a rating class to the median value of the distribution of a specific ratio. The highest rating class and the lowest rating classes are then fixed from the same ratio distribution but conditional on the country and sector to which the company belongs. All rating classes are finally determined dividing the interval between the highest and lowest class in equally spaced intervals.

Foe example, the application of the MORE model may imply that Company X gets an AA rating value for its ratio because it performs relatively better than the other companies belonging to the same country. On the other hand, Company Y, with the same ratio value, may get a CCC rating class, because its ratio performs worse in comparison with the other companies belonging to the same country. It is important to underline that fuzzy intervals are not built solely on statistical models, but they are corrected by financial analysts. For example, if the leverage ratio is less than zero; the maximum rating class can only be equal to CC.

Once the rating value for each financial ratio is assigned, the MORE model agglomerates all the information to find one rating for each selected company. The MORE agglomeration is based on a multidimensional network approach for which the better the financial equilibrium of a company the better its final rating.

Table 1
Distribution of companies by economic Sectors.

No.	Sectors	Abbrev.	No. of. companies	Percentage
1	Automobiles & Components	Auto	3,160	2.90
2	Capital Goods	Capt	15,658	14.35
3	Commercial and Professional Services	Comm	9,509	8.71
4	Consumer Durables and Apparel	ConD	4,515	4.14
5	Consumer Services	ConS	3,193	2.93
6	Energy	Enrg	87	0.08
7	Food & Staples Retailing	FdStp	7,945	7.28
8	Food Beverage & Tobacco	FdBvg	6,191	5.67
9	Health Care Equipment & Services	HCare	2,631	2.41
10	Household & Personal Products	HsPd	386	0.35
11	Materials	Mtrls	11,281	10.34
12	Media & Entertainment	Media	2,634	2.41
13	Pharmaceuticals & Biotechnology	Pharm	578	0.53
14	Real Estate	REst	2,177	1.99
15	Retailing	Rtail	27,793	25.46
16	Software & Services	Sftw	1,869	1.71
17	Technology Hardware & Equipment	Tech	994	0.91
18	Telecommunication Services	Tcom	256	0.23
19	Transportation	Trnsp	6,809	6.24
20	Utilities	Util	1481	1.36

Table 2
Description of the data variables.

	Variables	Description
1	Turnover	Operating revenue (Turnover) in 1000 EUR
2	EBIT	Operating P/L [= EBIT] in 1000 EUR
3	PLTax	P/L after tax in 1000 EUR
4	Leverage	Leverage (ratio)
5	ROE	Returns on Equity (percentage)
6	TAsset	Total assets in 1000 EUR
7	ROA	Return on Assets (percentage)
8	MScore	MORE Rating Score

This can be understood by means of an example. Suppose there are three different companies, with different ratio behaviors: company ABC and company XYZ, which perform very well for one financial ratio (ratio 1 and ratio 2 respectively) but badly for others; and company UVW which has ratios with similar good values (although not very good). The model will assign a better rating to company UVW instead of company XYZ. For more details we refer to [19].

3. Data

To evaluate the MORE model, we consider annual balance sheet data on over 100,000 SMEs from 20 sectors across European countries, covering a period that ranges from 2015 to 2020. The total number of SMEs considered is 109,147, of which 31,249 (28.63%) are from France, 1437 (1.32%) from Germany, 53,121 (48.67%) from Italy, and 23,340 (21.38%) from Spain. Table 1 shows the distribution of the institutions according to their economic sector classification.

Table 1 shows that, consistently with the predominance of SMEs in Europe, companies are concentrated in the Retailing, Capital Goods, Materials sectors and Commercial and Professional services. Sectors characterized by large companies, such as Energy, Pharmaceuticals and Telecommunication services, are less populated.

For each company, seven financial ratios have been extracted, from their publicly deposited balance sheet. Doing so we approximate the functioning of the MORE algorithm, which is based on a limited selection of financial ratios, chosen with the help of expert financial analysts. Table 2 present a list and description of the considered financial ratios, which will be, along with the MORE rating, the data employed to evaluate the proposed model.

Among the eight variables in Table 2, the variable "MScore": the rating evaluation assigned by the MORE model, will be employed as a response target variable. The other seven variables will be taken as candidate predictor variables.

Table 3
Summary (Median and Median Absolute Deviation) of financials of companies by sectors for the year 2020.

		Turnover	EBIT	PLTax	TAsset	Leverage	ROE	ROA
Auto	Median	8472.50	421.50	301.50	8761.00	1.35	9.41	5.12
	MAD	6063.09	553.75	423.28	6570.14	1.22	10.46	5.86
Capt	Median	7304.00	317.00	217.00	7503.50	1.92	9.81	4.41
	MAD	4781.38	419.58	312.83	5931.14	1.86	12.21	5.37
Comm	Median	7582.00	311.00	213.00	6636.00	2.13	11.71	4.72
	MAD	5042.32	449.23	341.00	5295.85	2.13	15.11	6.35
ConD	Median	6733.00	225.00	128.00	7564.00	1.54	5.67	3.20
	MAD	4093.46	409.20	326.17	5589.40	1.45	10.56	5.31
ConS	Median	5534.00	56.00	18.00	6790.00	2.22	3.38	0.99
	MAD	2827.32	588.59	504.08	6378.15	2.52	21.17	7.93
Enrg	Median	9275.00	403.00	271.00	9253.00	1.37	8.35	5.16
	MAD	7411.52	429.95	326.17	8510.12	1.28	9.76	5.20
FdStp	Median	6699.00	194.00	135.00	3588.00	1.98	13.71	6.05
	MAD	4253.58	253.52	192.74	2868.83	1.96	16.23	7.16
FdBvg	Median	8491.00	229.00	132.00	8719.00	1.57	5.73	3.03
	MAD	6262.50	354.34	269.83	7564.23	1.60	8.09	4.05
HCare	Median	6850.00	373.00	273.00	6865.00	1.63	11.22	5.24
	MAD	4489.31	554.49	434.40	6401.87	1.65	16.72	8.41
HsPd	Median	6276.00	310.00	184.00	7403.00	1.56	10.42	4.53
	MAD	3785.82	710.17	564.87	6158.72	1.66	17.64	10.13
Mtrls	Median	7436.00	327.00	230.00	8316.00	1.36	7.56	4.09
	MAD	4956.33	444.78	341.00	6155.76	1.29	9.59	5.01
Media	Median	7350.50	245.00	195.50	7750.50	1.92	9.75	3.58
	MAD	4914.08	663.46	561.91	6246.94	1.96	17.53	7.73
Pharm	Median	11972.00	658.00	551.50	15466.50	1.07	11.04	5.35
	MAD	10054.99	1326.19	975.55	14194.41	1.12	13.53	7.99
REst	Median	7415.00	1356.00	737.00	48484.00	1.47	4.74	2.79
	MAD	5326.98	1974.82	1438.12	57293.59	1.75	6.94	3.38
Rtail	Median	8038.00	219.00	141.00	5297.00	1.81	8.89	4.40
	MAD	5805.86	281.69	203.12	4096.42	1.81	10.84	5.00
Sftw	Median	8079.00	448.00	356.00	6886.00	1.92	16.28	6.64
	MAD	5538.99	597.49	498.15	5307.71	1.79	18.92	8.52
Tech	Median	8216.00	425.00	320.50	9223.00	1.23	9.22	5.32
	MAD	5693.18	604.90	466.28	6801.43	1.24	11.88	6.71
Tcom	Median	8570.00	448.00	320.00	7987.50	1.73	11.88	6.11
	MAD	5893.34	640.48	475.91	7246.21	1.62	16.61	7.35
Trnsp	Median	7639.00	184.00	115.00	5397.00	2.36	9.46	3.42
	MAD	5141.66	312.83	247.59	4255.06	2.30	12.87	5.04
Util	Median	8694.00	1362.00	670.00	24035.00	2.15	9.14	5.45
	MAD	6616.84	1596.76	953.31	21482.87	2.48	12.42	5.46

The aim of the next Section will be to employ the above data to evaluate the accuracy and the robustness of the MORE model, particularly during the pandemic times.

Table 3 presents some summary statistics of the explanatory variables. For each sector, the median and median absolute deviation is calculated and reported. For the sake of clarity, we consider only the year 2020, remarking that the same statistics for the other years do not differ remarkably.

From Table 3 note that, as expected, the explanatory variables differ in their variability, being expressed in different scales. Some, like turnover and total assets, are expressed in absolute values; others, like EBIT and PL after Tax, although derived from other variables, are also in absolute values; finally, Leverage, ROE and ROA are ratios between variables.

Table 4 reports the summary statistics of the response variable: the rating of a company, along the considered years, that is, from 2015 to 2020. In the bottom part of the Table, the same ratings are aggregated in three rating groups, with all A's belonging to "High"; all B's to "Med" and the remaining ones to "Low".

From Table 4 note the stability of the rating distribution over time, with the exception of the year 2020, in which a larger proportion of companies falls into the "low" categories, with respect to the previous years.

Let $N_i(t_k)$, $i \in S$, be the number who are in state i at time t_k, and $N_{ij}(t_k, t_{k+1})$ be the number who move from state i at time t_k to state j at time t_{k+1}. The maximum likelihood estimator of the transition probability $P_{ij}(t_k, t_{k+1})$

$$\hat{P}_{ij}(t_k, t_{k+1}) = \frac{N_{ij}(t_k, t_{k+1})}{N_i(t_k)} \tag{1}$$

Table 4
Distributions of standard company annual ratings (top) and rating groups (bottom).

	2015	2016	2017	2018	2019	2020
Standard ratings						
AAA	0.78	0.89	1.00	1.10	1.22	0.99
AA	7.92	9.03	9.94	10.25	10.46	10.25
A	13.99	14.80	15.53	15.86	16.18	15.94
BBB	22.87	22.99	23.06	22.94	22.90	22.26
BB	24.91	24.54	24.12	23.69	23.24	21.90
B	19.05	18.32	17.47	17.04	16.53	15.36
CCC	6.44	5.71	5.27	5.28	5.43	7.33
CC	2.67	2.45	2.33	2.45	2.55	3.93
C	1.21	1.14	1.14	1.20	1.27	1.72
D	0.16	0.14	0.14	0.18	0.21	0.31
Rating groups						
High = AAA, AA, A	22.70	24.71	26.47	27.21	27.86	27.18
Med = BBB, BB, B	66.82	65.85	64.65	63.67	62.67	59.52
Low = CCC, CC, C, D	10.48	9.44	8.88	9.12	9.47	13.30

To better understand the phenomena, Table 5 shows the transition matrix of the probability of a change in the MORE rating between year t and $t+1$. Each row corresponds to the credit rating at time t and each column corresponds to the rating at time $t+1$.

Table 5 shows that the probability that a company rated AAA in 2019 will be rated AA in 2020 is 45.82%. This probability used to be around 37.27% between 2015–2016, 36.71% 2016–2017, 35.71% 2017–2018, and 35.78% in 2018–2019. This suggests that the chance of a AAA-rated company defaulting within one year continually decreased from 37% to 36% between 2015–2019, with 2019–2020 recording a rise in the transition probability. Thus, more AAA rated companies are likely to default in 2020 than in the preceding years. Similarly, the transition probability from a BBB-rated company to BB within one year ranged between 16%–18% between 2015–2020, with 2020 recording the highest transition probability over the sample period.

To summarize, the descriptive analysis done so far clearly indicates a deterioration of credit ratings of the considered European companies, during the COVID-19 outbreak year (2020). Worse ratings imply less financial inclusion and higher costs for companies.

However, an aggregate result of this kind is not sufficient. It is instead of extreme importance to understand whether and how the fintech MORE rating is robust; that is, whether it is able to maintain the importance weight of the financial performance variables, leaving differences in rating to depend only on sector specific characteristic. A positive answer would mean that the fintech rating model, while adjusting to the macro economic shocks caused by COVID-19, which create imbalance between sectors, steadily takes into account micro economic firm-specific characteristics, rewarding the best performing companies and encouraging movements from worsening sectors to more resilient ones, improving financial inclusion and sustainability.

4. Model assessment

A rating model can be assessed in different ways, which can be reflected in different statistical evaluation models. We follow the perspective of a supervisory authority like the European Security and Markets Authority (ESMA), which is the actual supervisor of the MORE model. In this perspective, to be validated from a statistical viewpoint, a rating model should be SAFE: Sustainable, Accurate, Fair and Explainable. "Sustainable" essentially means that the ratings are robust in time, allowing only for small variations in ratings, along all rating classes. "Accurate" essentially mean a high goodness of fit. "Fair" mean that they do not introduce biases among population groups, differing by

Table 5
Annual rating transition matrix, 2015–2020.

	AAA	AA	A	BBB	BB	B	CCC	CC	C	D
2015–2016										
AAA	53.97	37.27	6.19	1.52	0.47	0.35	0.12	0.12	0.00	0.00
AA	4.73	66.24	22.96	4.36	0.94	0.59	0.13	0.03	0.01	0.00
A	0.41	19.55	54.32	20.17	3.38	1.62	0.39	0.12	0.03	0.00
BBB	0.09	2.54	19.44	57.60	15.62	3.14	1.17	0.32	0.08	0.00
BB	0.02	0.40	2.38	21.20	57.84	14.35	2.76	0.86	0.17	0.01
B	0.02	0.26	1.01	5.05	25.48	57.97	7.60	2.12	0.47	0.01
CCC	0.00	0.30	1.21	4.58	13.21	31.95	36.26	9.77	2.60	0.13
CC	0.00	0.24	0.55	2.95	9.03	19.96	25.15	30.71	10.79	0.62
C	0.00	0.15	0.46	1.37	5.17	9.50	17.25	22.64	39.82	3.65
D	0.00	0.00	0.56	1.68	1.12	3.35	12.29	12.85	27.93	40.22
2016–2017										
AAA	54.40	36.71	5.58	1.24	0.72	0.93	0.41	0.00	0.00	0.00
AA	4.67	66.64	22.33	4.58	0.96	0.57	0.19	0.04	0.01	0.01
A	0.48	19.20	55.33	19.74	3.01	1.57	0.52	0.12	0.02	0.01
BBB	0.10	2.55	18.92	57.95	15.71	3.11	1.28	0.29	0.10	0.00
BB	0.03	0.38	2.48	21.03	59.16	13.43	2.54	0.77	0.17	0.01
B	0.00	0.27	1.28	4.70	24.11	59.34	7.53	2.15	0.61	0.03
CCC	0.00	0.37	1.14	4.90	14.05	29.95	35.61	10.85	3.02	0.11
CC	0.00	0.19	0.67	3.10	7.95	19.82	24.26	31.80	11.76	0.45
C	0.08	0.00	0.24	1.29	3.30	8.38	20.23	20.79	40.13	5.56
D	0.00	0.66	1.32	0.00	1.32	3.29	10.53	15.13	28.29	39.47
2017–2018										
AAA	54.61	35.71	6.85	2.01	0.46	0.37	0.00	0.00	0.00	0.00
AA	4.63	65.62	23.55	4.52	0.84	0.63	0.18	0.04	0.00	0.00
A	0.41	17.38	55.29	21.32	3.21	1.54	0.67	0.12	0.06	0.01
BBB	0.08	2.34	17.64	58.25	16.39	3.57	1.35	0.29	0.09	0.00
BB	0.02	0.34	2.16	19.37	59.63	14.25	3.01	0.99	0.20	0.02
B	0.01	0.19	1.13	4.11	22.60	60.19	8.29	2.77	0.65	0.05
CCC	0.00	0.26	1.13	4.51	13.67	27.66	38.09	11.11	3.32	0.24
CC	0.04	0.12	0.87	3.23	9.76	17.47	21.20	33.36	13.18	0.79
C	0.00	0.08	0.08	1.78	3.63	7.75	13.88	23.00	43.34	6.46
D	0.00	0.00	0.00	0.63	1.90	4.43	10.76	11.39	25.95	44.94
2018–2019										
AAA	55.63	35.78	5.25	2.25	0.50	0.50	0.08	0.00	0.00	0.00
AA	4.78	65.22	23.64	4.44	1.10	0.60	0.21	0.01	0.01	0.00
A	0.47	17.08	55.86	20.86	3.49	1.64	0.49	0.09	0.02	0.00
BBB	0.14	2.30	17.41	58.09	16.39	3.76	1.48	0.32	0.11	0.00
BB	0.01	0.36	2.27	19.56	59.14	14.10	3.22	1.14	0.19	0.00
B	0.03	0.24	1.26	4.62	22.32	58.95	8.73	3.06	0.72	0.06
CCC	0.07	0.29	1.35	4.96	13.60	26.54	37.98	11.75	3.19	0.26
CC	0.00	0.11	0.64	3.06	9.23	19.25	20.96	32.06	13.60	1.08
C	0.00	0.15	0.30	2.05	4.26	6.24	16.88	20.30	43.42	6.39
D	0.00	0.00	0.00	1.01	0.00	1.51	11.56	10.55	28.14	47.24
2019–2020										
AAA	39.43	45.82	9.41	2.63	0.98	1.28	0.30	0.15	0.00	0.00
AA	3.63	58.15	27.65	6.28	1.77	1.85	0.53	0.11	0.04	0.00
A	0.54	16.78	48.62	23.54	4.86	3.74	1.56	0.31	0.06	0.00
BBB	0.11	2.94	17.64	50.83	18.28	5.71	3.42	0.84	0.21	0.02
BB	0.04	0.54	2.92	20.42	51.03	15.26	6.81	2.47	0.48	0.02
B	0.05	0.38	1.46	5.58	23.13	47.57	13.28	6.89	1.61	0.04
CCC	0.05	0.42	1.43	6.22	13.36	24.35	32.51	16.34	4.79	0.52
CC	0.04	0.11	0.83	3.38	10.24	16.24	18.43	31.79	17.67	1.29
C	0.07	0.36	0.22	2.37	4.46	6.97	16.53	19.34	41.12	8.55
D	0.00	0.43	0.00	1.28	1.28	1.71	6.41	6.84	22.65	59.40

gender or country of nationality. "Explainable" means that the results of the models should be understood in terms of their drivers. According to this perspective, to evaluate the MORE model we consider, as a response variable, the variation in the rating class, once ordinal ratings are converted into numerical codes. Doing so, we measure the robustness of the ratings assuming linearity which implies, for example, that a two notches change in a high rating class has the same value as a two notches change in a low rating class. The obtained response variable will then be assumed to depend on a linear combination of company specific and of sector specific variables, allowing for a clear explainability of the results. At the same time, we aggregate results by countries, allowing for further analysis on model fairness which could compare results obtained separately for each country. Finally, as the aim of our validation exercise is essentially to explain the changes in ratings over time and, in particular, during COVID-19 times, we will measure accuracy in terms of goodness of fit, rather than in terms of predictive accuracy.

Table 6
Conversion of MORE ratings in numerical values.

MORE rating	D	C	CC	CCC	B	BB	BBB	A	AA	AAA
Numeric values	0	1	2	3	4	5	6	7	8	9

Table 7
ANCOVA parameter estimates for the model chosen by means of a stepwise model selection.

	Dependent variable: Δ MScore				
	2015–2016	2016–2017	2017–2018	2018–2019	2019–2020
Leverage	0.0045***	0.0041***	0.0041***	0.0038***	0.0049***
ROE	−0.00003**	−0.0002***	−0.0002***	−0.0001***	−0.0001***
ROA	−0.0150***	−0.0145***	−0.0138***	−0.0154***	−0.0172***
Auto	0.2649***	0.1775***	0.0834***	0.0804***	0.0775***
Capt	0.1453***	0.1410***	0.1038***	0.1337***	−0.0723***
Comm	0.1215***	0.1576***	0.0931***	0.0644***	−0.0601***
ConD	0.2093***	0.1706***	0.1151***	0.0867***	−0.2243***
ConS	0.1975***	0.1713***	0.0758***	0.0885***	−0.7566***
Enrg	0.0684	0.1707*	−0.0703	0.0769	0.1987*
FdBvg	0.1531***	0.1062***	0.0533***	0.0328***	0.0225
FdStp	0.1239***	0.1326***	0.0405***	0.1299***	0.1877***
HCare	0.1909***	0.2009***	0.0677***	0.0996***	−0.0177
HsPd	0.1704***	0.1065**	0.1933***	0.0157	−0.2525***
Media	0.0925***	0.1242***	0.0324*	0.0600***	−0.1563***
Mtrls	0.1999***	0.1433***	0.1013***	0.0864***	0.0055
Pharm	0.1818***	0.0887**	0.0117	0.1344***	0.1400***
REst	0.0936***	0.0982***	0.0245	0.0257	−0.0476**
Rtail	0.1459***	0.1305***	0.1027***	0.0847***	0.0042
Sftw	0.1856***	0.1937***	0.1893***	0.1398***	0.0326
Tcom	0.1312**	0.2046***	0.0434	0.2011***	−0.0051
Tech	0.1589***	0.2315***	0.1132***	0.1627***	−0.0912**
Trnsp	0.0962***	0.0633***	−0.0161	0.0560***	−0.0623***
Util	0.0287	0.2340***	0.1081***	0.1491***	0.0851***
Observations	109,147	109,147	109,147	109,147	109,147
R^2	0.0471	0.0404	0.0329	0.0344	0.0602
Adjusted R^2	0.0469	0.0402	0.0327	0.0342	0.0600

*$p < 0.1$
**$p < 0.05$
***$p < 0.01$

Let $S_{i,t}$ denote the numeric value of the MORE rating of institution-i at time t. For model assessment, we assign numeric values to each MORE rating as in Table 6 below.

4.1. ANCOVA model

We then model the relationship between changes in yearly ratings and financial performance variables as an analysis of covariance (ANCOVA) written as follows:

$$Y_{i,t} = \sum_{j=1}^{p} \beta_{j,t} X_{ij,t} + \sum_{j=1}^{q} \alpha_{j,t} D_{ji} + \varepsilon_{i,t} \tag{2}$$

where $Y_{i,t} = S_{i,t+1} - S_{i,t}$ is the dependent variable that measure changes in the yearly ratings of institutions-i between time t and $t+1$, $X_{ij,t}$ is the value of the quantitative covariate variable j for institution i at time t, D_{ij} is 1 if institution-i is in sector-j, and zero otherwise. In this application, the quantitative covariates are the financial performance variables (i.e., leverage, ROE, ROA). The error term $\varepsilon_{i,t}$ is typically assumed to be normal and independent.

The general application of ANCOVA models is to test whether the means of the fixed factors are equal, or at least one of the means is different from another. Furthermore, the models allows us to evaluate the effects of the fixed factors to identify which ones have weak, moderate or strong effects.

Table 8
Classification of the sector impacts on the 2019–2020 rating variations.

Mild positive impact	No impact	Mild negative impact	Moderate negative impact	High negative impact	Very high negative impact
Pharmaceuticals & Biotechnology	Utility	Software & Services	Real Estates	Consumer Durables	Consumer Services
Food & Staples Retailing	Automobiles & Components	Food Beverages & Tobacco	Commercial Services	Household Products	
Energy		Materials	Transportation		
		Retailing	Capital Goods		
		Telecommunication Services	Technology Hardware		
		Health Care Services	Media		

Table 7 presents the empirical findings derived from the application of the ANCOVA method to our data, leading to the estimated coefficients for the statistically selected explanatory variables and sectors, across different year-to-year variations.

From Table 7 note first that, among the seven candidate explanatory variables, three are selected by the stepwise model selection procedure and, therefore, reported in the Table: Leverage, ROE and ROA. These three variables are those that mostly affect the variations in credit ratings. Note that they balance different characteristics of a firm: its financial dependence (leverage); its profitability (ROA); its operational efficiency (ROA).

Second, Table 7 shows that the sign and magnitude of all the coefficients of the chosen three financial variables is stable over time, implying that the MORE rating maintains the importance weight of the micro economic characteristics, even during the pandemic.

Third, Table 7 shows that the importance weight of sectors, quite stable between 2015 and 2019, changes dramatically in 2020. While Food (Staples Retailing), Pharmaceutical and, to a lower extent, Energy, positively impact changes in credit ratings during the COVID-19 pandemic, all the others impact negatively. This indicate that the ratings have adjusted to take into account the higher demand of goods and services for certain sectors (food, pharmaceutical, energy) at the expense of all the others.

To further evaluate the results of the model, we have compared the 2019–2020 sector effects on the ratings by means of a set of pairwise statistical t-tests aimed at verifying the statistical significance of their difference. We have taken into account the dependence of all pairwise tests by the same model by means of the Bonferroni correction to the p-values. Fig. 1-panel (b) reports the resulting p-values, whereas Fig. 1-panel (a) reports the dendrogram of a hierarchical clustering algorithm on the same sectorial effects.

From Fig. 1, we observe six types of clusters for 2019–2020, whose description is reported in Table 8.

Table 8 classifies the changes in ratings during the pandemic by sector clusters. The leftmost cluster contains the sectors with a (mild) positive impact on the ratings: pharmaceuticals, food and staples retailing and energy. For these sectors, the pandemic improves the fintech rating and, therefore, when the fintech platform is employed, greater financial inclusion opportunities follow, for companies in the sector and for new entrants. The rightmost clusters contains sectors with a (high or very high) negative impact on the ratings: Consumer services and durables, Household products, for which the fintech platform implies lower financial opportunities for companies that remain in these sectors.

4.2. ANCOVA model with lags

This section considers more flexible models with lags on the dependent variable of the set-up in (2). This is to help us better understand the dynamics of these MORE rating. This new set-up is simply ANCOVA + ARDL (autoregressive distributed lag), given by:

$$Y_{i,t} = \sum_{l=1}^{q} \phi_l Y_{i,t-l} + \sum_{j=1}^{p} \beta_{j,t} X_{ij,t} + \sum_{j=1}^{q} \alpha_{j,t} D_{ji} + \varepsilon_{i,t} \quad (3)$$

Table 9
ANCOVA + ARDL parameter estimates for the model.

	Dependent variable: Δ MScore (t)			
	2016–2017	2017–2018	2018–2019	2019–2020
Δ MScore.t1	−0.2616***	−0.2586***	−0.2798***	−0.2681***
Leverage	0.0043***	0.0043***	0.0039***	0.0051***
ROE	−0.0001***	−0.0001***	−0.0001***	−0.0001***
ROA	−0.0073***	−0.0066***	−0.0073***	−0.0094***
Auto	0.1707***	0.0484***	0.0182	0.0191
Capt	0.1114***	0.0705***	0.0865***	−0.1148***
Comm	0.1121***	0.0544***	0.0051	−0.1276***
ConD	0.1601***	0.0896***	0.0427***	−0.2745***
ConS	0.1465***	0.0362**	0.0204	−0.8210***
Enrg	0.1254	−0.0911	−0.0021	0.1594
FdBvg	0.0917***	0.0255**	−0.0085	−0.0228
FdStp	0.0971***	0.0078	0.0726***	0.1532***
HCare	0.1535***	0.0226	0.0182	−0.0926***
HsPd	0.0703	0.1421***	−0.0190	−0.3332***
Media	0.0800***	−0.0064	−0.0032	−0.2091***
Mtrls	0.1286***	0.0684***	0.0396***	−0.0439***
Pharm	0.0718*	−0.0297	0.0728*	0.1105**
REst	0.0858***	0.0110	−0.0080	−0.0811***
Rtail	0.1016***	0.0672***	0.0381***	−0.0470***
Sftw	0.1450***	0.1395***	0.0782***	−0.0443*
Tcom	0.1589***	0.0134	0.1245**	−0.0386
Tech	0.1948***	0.0876***	0.0987***	−0.1409***
Trnsp	0.0276**	−0.0576***	−0.0034	−0.1002***
Util	0.1968***	0.1150***	0.1126***	0.0529*
Observations	109,147	109,147	109,147	109,147
R^2	0.1042	0.0946	0.1041	0.1057
Adjusted R^2	0.1040	0.0944	0.1039	0.1055

*p < 0.1
**p < 0.05
***p < 0.01

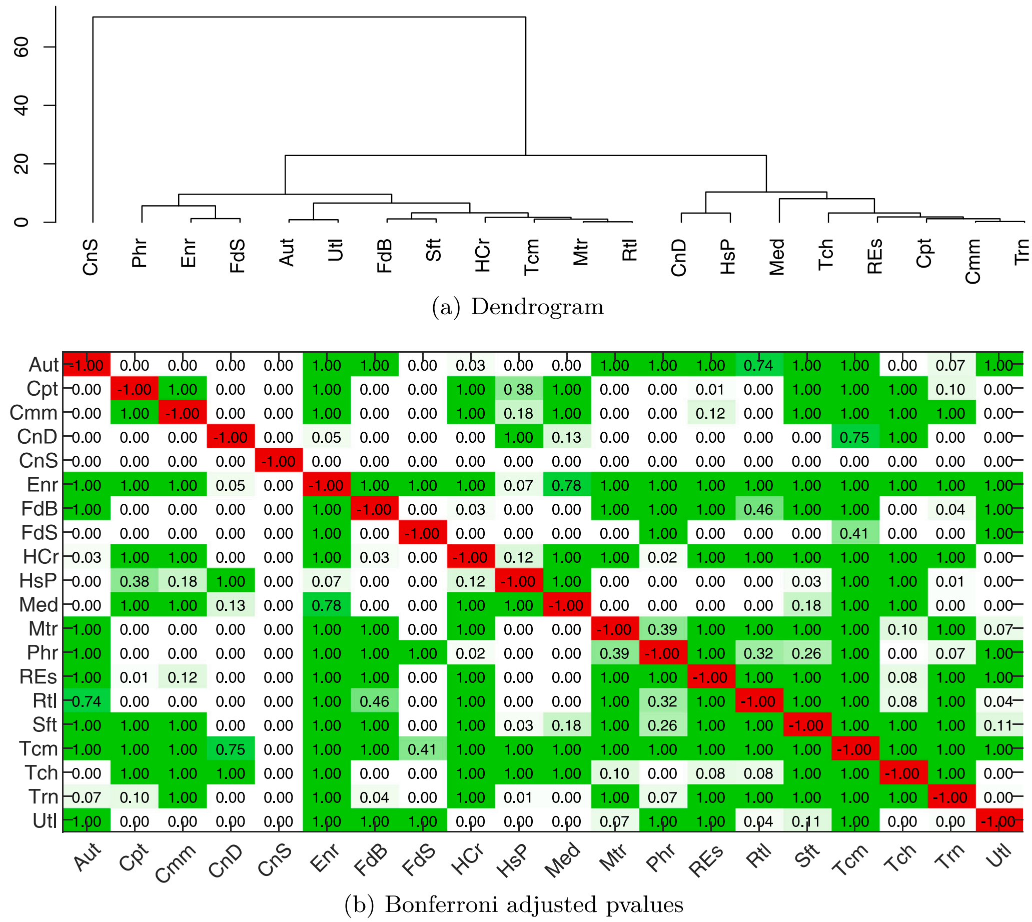
where q is the lag order of the dependent variable, and ϕ_l is the autoregressive coefficient at lag l. For simplicity, we consider a lag 1 set-up to model the dynamics of the MORE rating.

Table 9 reports the results of the parameter estimates with the lags of the MORE rating. Since there is no lag observation for 2015–2016, it is not included in the estimation. Fig. 2-panel (b) reports the associated p-values, whereas Fig. 2-panel (a) reports the dendrogram of a hierarchical clustering algorithm on the same sectorial effects.

The results from Table 9 and Fig. 2 confirm those in Table 7 and Fig. 1: the three explanatory variables that are selected as most significant by the model are still Leverage, ROE and ROA, measuring, respectively, the financial dependence, the profitability (ROA) and the operational efficiency of a company. The results that concern the impact of sectors remain unchanged.

5. Conclusions

The aim of the paper is to describe a machine learning fintech inclusion platform, and evaluate the sustainability of its credit ratings, in terms of their accuracy and robustness. This in particular during the COVID-19 pandemic, to understand how it favors financial inclusion.



(a) Dendrogram

(b) Bonferroni adjusted pvalues

Fig. 1. Top: A dendrogram of the cluster of sector effects of the change in MORE ratings between 2019–2020, using the ANCOVA model. Bottom: The matrix of Bonferroni adjusted p-value for pairwise t-test between means of the sectors, using the ANCOVA model. The off-diagonals are color-coded to describe the sign of the statistical relationships, with green signifying difference in means (i.e p-value > 0.05), and white for non-significant differences.

The machine learning model is based on financial data publicly available through the Chambers of Commerce, from which a set of explanatory variables are selected in terms of their combined financial and statistical importance. This implies a high degree of transparency and explainability of the obtained credit ratings.

To evaluate the accuracy and robustness of the proposed fintech rating model, we have analyzed data on a large collection of companies, and built an econometric model to explain the variation of the rating as a function of both micro economic factors (expressed by the company's financial ratios) and macro economic factors (expressed by the company's economic sector of activity).

The obtained empirical results indicate that the impact of company's financials on the ratings is highly robust: it does not change over time, even during the pandemic year. On the other hand, the ratings do change substantially across different sectors.

More precisely, during the COVID-19 pandemic, the ratings of the Food, Pharmaceutical and Energy sectors improve; whereas all the others, and especially Consumer services and household products, worsen.

In summary, our results indicate that fintech credit ratings based on the application of machine learning models, such as MORE, increase (decrease) the ratings of the companies that have the better (the worse) financials, and adjust their ratings to take into account variations in macro economic conditions, reflecting future market expectations. This implies that fintech credit ratings can improve financial inclusion, even during crisis times: maintaining the reward for the better performing companies, and incentivizing mobility of companies across sectors, to improve sustainability.

Future research may include comparison with Bayesian learning models for credit scoring (see e.g. [20], [21]).

CRediT authorship contribution statement

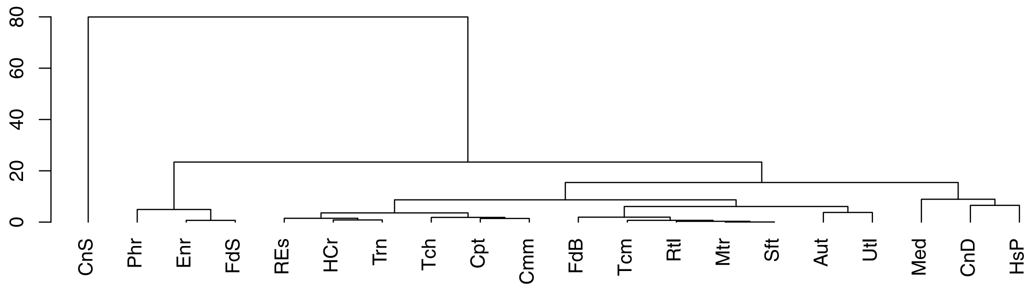
Daniel Ahelegbey: Conceptualization, Formal analysis, Investigation, Methodology, Software, Validation, Visualization, Writing – original draft, Writing – review & editing. **Paolo Giudici:** Conceptualization, Funding acquisition, Investigation, Methodology, Supervision, Writing – original draft, Writing – review & editing. **Valentino Pediroda:** Conceptualization, Funding acquisition, Investigation, Supervision, Writing – original draft, Writing – review & editing.

Data availability

Data will be made available on request.

Acknowledgments

The Authors acknowledge support from the Horizon2020 PERISCOPE (Pan-European Response to the Impacts of COVID-19 and future Pandemics and Epidemics), programme (contract number n. 101016233).



(a) Dendrogram

(b) Bonferroni adjusted pvalues

Fig. 2. Top: A dendrogram of the cluster of sector effects of the change in MORE ratings between 2019–2020, using the ANCOVA-ARDL model. Bottom: The matrix of Bonferroni adjusted p-value for pairwise t-test between means of the sectors, using the ANCOVA-ARDL model. The off-diagonals are color-coded to describe the sign of the statistical relationships with green signifying difference in means (i.e p-value > 0.05), and white for non-significant differences.

References

[1] Haddad C, Hornuf L. The emergence of the global fintech market: Economic and technological determinants. Small Bus Econ 2019;1(53):81–105.

[2] Puschmann T. Fintech. Bus Inf Syst Eng 2017;1(59):69–76.

[3] Ahelegbey D, Giudici P, Hadji-Misheva B. Latent factor models for credit scoring in P2P systems. Physica A 2019;1(522):112–21.

[4] Bussmann N, Giudici P, Marinelli D, Papenbrock J. Explainable machine learning in credit risk management. Comput Econ 2021;1(57):203–16.

[5] Giudici P, Hadji-Misheva B, Spelta A. Network based credit risk models. Qual Eng 2019;2(20):199–211.

[6] Emekter R, Yanbin T, Jirasakuldech B, Min L. Evaluating credit risk and loan performance in online peer-to-peer (P2P) lending. Appl Econ 2015;1(47):54–70.

[7] Serrano-Cinca C, Gutierrez-Nieto B. The use of profit scoring as an alternative to credit scoring systems in peer-to-peer lending. Decis Support Syst 2016;1(89):113–22.

[8] Giudici P, Raffinetti E. Shapley-lorenz explainable artificial intelligence. Expert Systems with applications 2021;167(114104).

[9] Beaver W. Financial ratios as predictors of failures. J Account Res 1967;1.

[10] Altman EI. Financial ratios, discriminant analysis and the prediction of corporate bankruptcy. J Finance 1968;4(23):589–609.

[11] Hastie T, Tibshirani R, Friedman J. The elements of statistical learning. Berlin: Springer; 2002.

[12] Andreeva G, Crook J, Ansell J. Modelling profitability using survival combination scores. European J Oper Res 2007;3(183):1537–49.

[13] Barrios LJS, Andreeva G, Ansell J. Monetary and relative scorecards to assess profits in consumer revolving credit. J Oper Res Soc 2014;3(65):443–53.

[14] Rojas R. Neural networks: A systematic introduction. Berlin: Springer; 1996.

[15] Cortez C, Vapnik V. Support vector networks. Mach Learn 1995;1(20):273–97.

[16] Friedman C, Sandow S. Model performance measures for expected utility maximizing investors. Int J Theor Appl Finance 2003;6(4):355–401.

[17] Friedman C, Sandow S. Learning probabilistic models: An expected utility maximization approach. J Mach Learn Res 2003;1(4):257–91.

[18] Merton RC. Theory of rational option pricing. J Econ Manage Sci 1973;1(4):141–83.

[19] Ciprian M, Di Stefano D, Kaucic M, Nogherotto G, Pediroda V. Multiattribute methodologies in financial decision aid. In: Handbook of research on nature inspired computing for economy and management. Hersey: Idea Group; 2016.

[20] Giudici P. Bayesian data mining, with application to credit scoring and benchmarking. Applied Stochastic Models in Business and Industry 2001;17:69–81.

[21] Giudici P, Mezzetti M, Muliere P. Mixtures of dirichlet process priors for variable selection in survival analysis. Journal of Statistical planning and inference 2003;111(1-2):101–15.

Paolo Giudici Professor of Statistics and of Financial Data Science. Supervisor of 200+ Master's students and of 21 Phd students, currently working as academic researchers, in the financial sector and in IT/consulting companies.

Author of several scientific publications, with an h-index of 41 (Google scholar), 29 (Scopus), 26 (Clarivate), which mainly concern statistical models to obtain sustainable, accurate , fair and explainable economic predictions and/or risk management models to improve financial inclusion, financial stability and financial sustainability.

The publications have appeared in top field scientific journals such as: Journal of the Royal Statistical Society, Journal of Business and Economics Statistics, Biometrika, Expert systems with applications, Machine Learning, Journal of Banking and Finance, Journal of Financial Stability, Finance research letters.

Coordinator of 12 funded scientific projects, among which the European Horizon2020 projects "PERISCOPE: Pan-European response to the impacts of COVID-19 and future pandemics and epidemics (2020-2023)", "FIN-TECH: Financial supervision and Technological compliance" (2019-2020).

Editor of ''Artificial Intelligence in Finance'', Frontiers; ''Risks'' and ''Fintech'', MDPI . Research fellow at the Bank for International Settlements and at the University College London center for Blockchain technologies. Expert advisor for the European Commission, the Bank of Italy, the European Insurance and Occupational Pensions Authority (EIOPA), the Italian Ministry of Development.

Member of the steering committee of the Italian Statistical Society (SIS) and honorary member of the Association of Italian Financial Risk Managers (AIFIRM). Member of the European Network for Business and Industrial Statistics (ENBIS), the European Big Data Value Association (BDVA) , the International Association for Trusted Blockchain Applications (INATBA), the Italian Econometric Society (SIDE).