RISK-RETURN MODELLING IN THE P2P LENDING MARKET: TRENDS, GAPS, RECOMMENDATIONS AND FUTURE DIRECTIONS

MILLER-JANNY ARIZA-GARZÓN[1,2], MARÍA-DEL-MAR CAMACHO-MIÑANO[3], MARÍA-JESÚS SEGOVIA-VARGAS[4] and JAVIER ARROYO[1,5]

[1]Software Engineering and Artificial Intelligence Department, Complutense University of Madrid, 28040 Madrid, Spain

[2]Faculty of Statistical Studies, Complutense University of Madrid, 28040 Madrid, Spain

[3]Accounting and Finance Department. Faculty of Business Administration and Economics. Complutense University of Madrid, Campus de Somosaguas, 28223, Pozuelo de Alarcón, Madrid, Spain.

[4]Department of Financial and Actuarial Economics and Statistics, Complutense University of Madrid, 28040 Madrid, Spain

[5]Institute of Knowledge Technology, Complutense University of Madrid, 28040 Madrid, Spain

Funding: This work was supported in part by the European Union's H2020 Coordination and Support Actions [Grant 825215], by COST (European Cooperation in Science and Technology) [COST Action 19130] and by the Santander-UCM Research Project [Grant PR87/19-22586].

Corresponding author.
E-mail addresses: millerar@ucm.es (M.J. Ariza-Garzón),
Telephone number: 91394 2564, Fax number: 91 394 23 88

Abstract

The proposal for new financial products has been accompanied by new tools for risk management and profit in the Peer-to-Peer (P2P) lending market, one market in evolution, as an alternative for traditional investment and financing. For understanding this development, a systematic literature review and a bibliometric analysis of 104 papers published in the Web of Science database in the last decade are carried out using Scimat software. Our aim is to identify methodological elements, modelling components, analysis of variables and business aspects that generate opportunities for deepening its development and application. Developments of algorithms of artificial intelligence (AI) and machine learning (ML) support most of new proposals. Regulators, supervisors and users tend to increasingly seek these new alternatives in a natural project of financial digitalization demanded by technological advances, innovation and market development. Based on this study, recommendations in future research directions are provided for researchers.

Keywords
Risk management modelling, credit risk modelling, profit and invest modelling, P2P lending, machine learning, literature review, bibliometric analysis.

1. Introduction

The P2P lending market is based on technological platforms that charge a fee for the service of connecting lenders and borrowers together. Borrowers can obtain credit directly from lenders at lower rates and with greater accessibility than conventional credit alternatives, while lenders can obtain higher returns than other financial products (Emekter et al., 2015). The peer-to-peer (P2P) lending market emerges as an alternative for investment and financing that breaks down some of the intermediation barriers offered by traditional banking. P2P markets can be considered as a collaborative economy agreement (Serrano-Cinca et al., 2015) that even allows the financial inclusion of people excluded from traditional systems (Claessens et al., 2018). It is considered a complementary and non-competitive market to conventional banking (Milne and Parboteeah, 2016), an approach that different governments and regulatory bodies around the world are working on (ROFIEG, 2019).

The rise of this new market is also associated with factors such as the reduction of intermediation costs, the incipient regulation customized to this business and the evolution of big data and artificial intelligence tools that complement traditional modelling, allowing the experience to be personalized and the service to be improved (Financial Stability Board, 2017; Giudici et al., 2020). Claessens et al. (2018) differentiate leverage factors by country, by market and by the regulatory framework, and details some challenges of the new credit markets. Some of the most important risks are information asymmetries (Serrano-Cinca and Gutierrez-Nieto, 2016), adverse selection and moral hazard (Cummins et al., 2019), and in most platforms, that the lenders are the ones that bear the credit risk (Serrano-Cinca and Gutierrez-Nieto, 2016).

However, the P2P lending market has also its inherent risks. For example, in China, the lack of regulation of the P2P market caused the surge of fraud and illegal activity. As a result, there was a regression in this market in China. As mentioned in Gao et al. (2020), even though the P2P market is based on technological innovation, it remains a financial problem that demands proper credit and financial risk management. The Chinese collapse of the P2P lending market contributed to making visible the need of efficient regulatory policies, customised to the challenges of the new markets, and based on competent technicians, aspects also highlighted in ROFIEG (2019). As it is a different market, there is a need for customised risk management and regulation, but unified for the different countries involved, which truly protects investors and the entire market in a globalised context.

Quantitative modeling is considered an important activity for risk management. Still, it is also vital to maximize investors' profits and provide quality of service in a technological and globalized environment. Its use can be improved by recognizing factors such as appropriate knowledge of the business and, therefore, the associated risk factors, understanding the risk stages, and understanding of different techniques customized to its financial management's specific problems. It is also essential a modeling structure that ensures unbiasedness and efficiency, and concern not only for the accuracy of the other different metrics but also for what they mean in a business context (Xia et al., 2017). It is equally important to have sufficient data, to take advantage of various sources, with validation and appropriate treatment (Giudici et al., 2020; ROFIEG, 2019), to know a range of powerful prediction techniques but with explainability (Ariza-Garzón et al., 2020; Bussmann et al., 2020; ROFIEG, 2019), among other aspects. These provide elements to guarantee its use, control, and trust by regulators and users.

Bearing all these ideas in mind, we aim to find main research trends in the literature associated with the different quantitative methodologies used in the management of the pair risk-return in the P2P lending market, which includes several of the aspects described above. In other words, our goal is to know what topics have been already studied and addressed in the P2P lending market and what are the biggest challenges and limitations that need further study. To address these questions, we carry out a bibliometric analysis and a traditional systematic analysis over a sample of 104 papers from the Web of Science (WoS) bibliographic database, which corresponds to most of the papers published between 2010 and the first quarter of 2020 (15 April 2020).

The bibliometric analysis will show the main references, trends and the different developments that took place in the last decade. Our paper has a different approach as prior papers published as (Bachmann et al., 2011) which is focused mainly on the description of existing platforms, on the identification of the determinants associated with the proper management of financing sources and the obtaining of returns on investment, for the second half of the first decade of this century. The traditional systematic analysis complements the bibliometric analysis, by analysing in-depth several aspects of risk and return management in this market. The criteria used are such as the business focus of the proposals, the methodological contribution, the definition made of the target variables, the type of modelling paradigm used, the quantitative contribution, and the description of the modelling stages.

To achieve the purposes described, the rest of the paper is organized as follows. Section 2 presents the research methodology, including some general statistics of the P2P market in the world, publications by region, indicators of cooperation among authors, citation and co-occurrence analysis through graphs and cluster analysis of concepts and methodological elements, among other aspects. In section 3, we carry out a systematic analysis, including descriptive segmentation criteria such as the datasets and software used, characterize and classify by some components of the P2P business that are addressed in the papers, methodological aspects that are highlighted in each research, segment by the learning paradigm used and draw attention to some elements of modelling required by market regulators and supervisors. Finally, we present the main conclusions and draw attention to some identified gaps, which we consider are demanded to be investigated and deepened for an adequate risk and profit management and healthy development of the P2P lending market in our countries.

2. Research Methodology

We will carry out a bibliometric analysis based on connectivity networks and natural language processing techniques. This methodology will provide an overview of the research area, identifying the most relevant documents, topics and trends associated with the management of both risk and profitability in the P2P market. Natural language processing algorithms will allow us to extract the most relevant terms from titles and abstracts. We will explore citation and co-ocurrence with network analysis. In the first case, the network represents citations (links) between documents (nodes). While in the second one, nodes represent topics and there is a link between topics when they appear in the same document, more precisely, in their titles, abstracts, or keywords. For this analysis, we rely on the VOSviewer software tool to analyse and visualize bibliographic data, which is freely accessible through www.vosviewer.com (van Eck and Waltman, 2010). We will also use bibliometrix, a library on R developed by (Aria and Cuccurullo, 2017).

In particular, we will use three types of representations from VOSviewer:

- Clusters by color on a network built using measures of similarity associated with the connection or relationship between the observed units, which can be papers or terms. The links represent the connection or the relation between two units.

- A temporal trend on a network. The same network described in the previous representation but with a color degradation from green to yellow that represents how recent is the unit's appearance from older to more recent.
- Density maps that use the nodes (terms) in a similar way as in the network visualization use a degradation from blue (lowest density) to red (highest density) to represent the nodes that are less and more frequent, respectively.

It is worth mentioning that node proximity in a network diagram implies a stronger relationship between the items, either by citation or because there are more documents where the terms appear together. Moreover, the nodes are thicker when they are more relevant according to a measure of centrality (greater number of links). The graphs are undirected; however, the direction of the citation can be generally inferred from the year of publication of both nodes. A detailed explanation of the methods used can be found in Van Eck and Waltman (2010, 2014, 2017). For the definition of the networks that generate the visualization of co-occurrence maps, we previously unified the concepts and terms, extracting those that we considered evident in the context of analysis, such as p2p lending, credit, among others, to identify the real occurrence of the most relevant concepts.

2.1. Definition of research questions

The first stage of our study is the definition of the research question. The goal of this study is to provide an overview of the current research on P2P lending market. Therefore, we define 3 main research questions:

1) Research question 1 (RQ1): How has the literature on P2P lending market evolved, specifically, the most influential papers?

2) Research question 2 (RQ2): What are the main contents developed in P2P lending market in the last decade? Dataset? Methodology? Business problem? Learning paradigm used? and Main current research trends?

3) Research question 3 (RQ3): What are the main research gaps and future research directions in the P2P lending market?

2.2. Search and screening the relevant papers

The second stage consists in searching all the relevant scientific papers on P2P lending market. We define a search protocol to reduce researcher bias. We use the main worldwide scientific journal database (Birkle et al., 2020), the Web of Science (Wos), to gather all the papers relevant for our research topic. The terms used were chosen using widely known keywords in the area, after the mutual agreement between the four authors. The keywords used are the following:

TOPIC: ((("peer to peer market" OR "P2P Platforms" OR "Online Peer-to-Peer credit") OR ("peer-to-peer credit") OR ("P2P credit") OR ("peer to peer credit") OR ("Online Peer-to-Peer lending") OR ("peer-to-peer lending") OR ("P2P lending") OR ("peer to peer lending") OR ("social lending"))) AND TOPIC: ((Statistical Analysis OR Survival OR learning schemes OR supervised models OR classification models OR learning techniques OR model OR machine learning OR logistic OR cox OR method OR algorithm OR prediction)) AND TOPIC: ((credit risk OR risk assessment OR risk OR Risk assessment OR financial risk)) AND TOPIC: ((loan evaluation OR loan default OR probability of Default OR Lenders' profitability OR profit OR creditworthiness OR creditworthiness rates OR scoring OR pay loan OR pay back the loan OR default OR status OR borrower status))

We also refine by years of publication from between 2010 and the first quarter of 2020 (15 April 2020), which corresponds to most of the papers published in this topic.

For the bibliometric analysis, we considered all the documents retrieved by the query, 104 papers. This analysis used authors, citations, titles, abstract, and keywords. For the second part, we excluded grey literature and concentrated on peer-reviewed high-quality papers. We excluded the following: publications without full-text availability and non-English papers. We obtained the final sample of 98 papers, which we included in this study as primary papers.

3. Results

3.1. Bibliometric analysis

3.1.1. Publication year, authors, source and geographic distribution

Figure 1 shows the distribution of the selected papers by publication year. Since most papers were published in the last five years, we can conclude that the P2P lending market is a very recent research area that is gaining attention (almost one-third of the papers were published in 2019). Since 2020 has not complete information, it is difficult to assess whether it does follow or not the trend from the previous years.

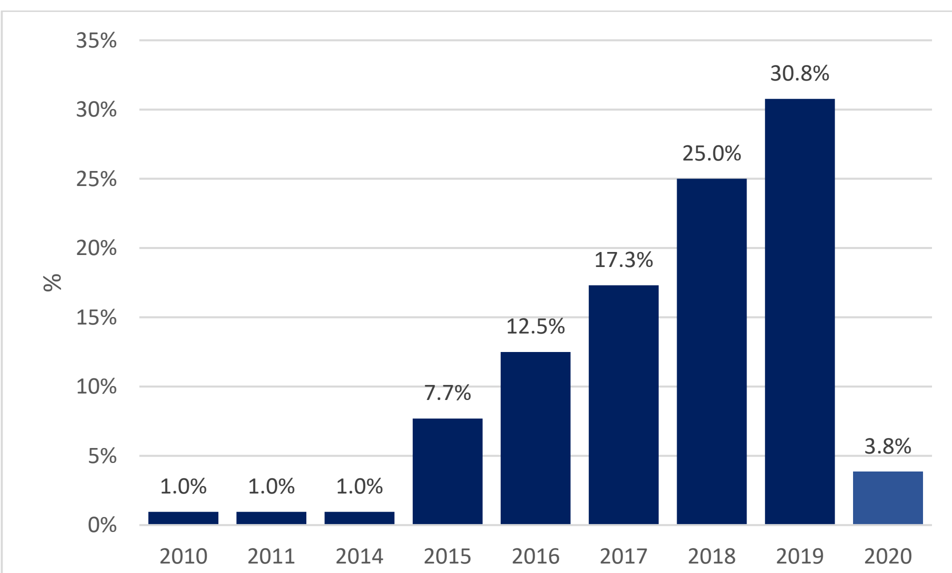

Figure 1. Distribution of primary papers selected 2010-2019 and 2020 until April (Source: Prepared by the authors using Web of Science data)

For Table 1 of authoring statistics, we will use the bibliometrix library of R developed by Aria and Cuccurullo (2017). There are only 12 papers with only one author. According to the data, each author wrote 37% of a document (documents/authors), which is equivalent to 2.7 authors per document (authors/documents), with 3.25 real authors on average per document (authors appearance/documents). The difference lies in that the first is calculated with the total of authors compared to the total of documents, and the second calculates the average number of authors per document. The collaboration index is almost 3, which means 3 authors per group work (total authors of multi-authored articles/total multi-authored articles), according to the indicator presented in Koseoglu (2016).

Table 1. Measures of contribution per author

Index	Value
Single-authored documents	12
Documents per author	0.367
Authors per Document	2.72
Co-Authors per Documents	3.25
Collaboration Index	2.96

(Source: Prepared by the authors using bibliometrix and Web of Science data)

Figure 2 shows the research areas of our papers according to the non-mutually exclusive categorization established by WoS. We can see that the publications mainly concentrate on Science Technology, Social Sciences and Technology, mainly in journals from Business Economics, Computer Science, Mathematics and Engineering.

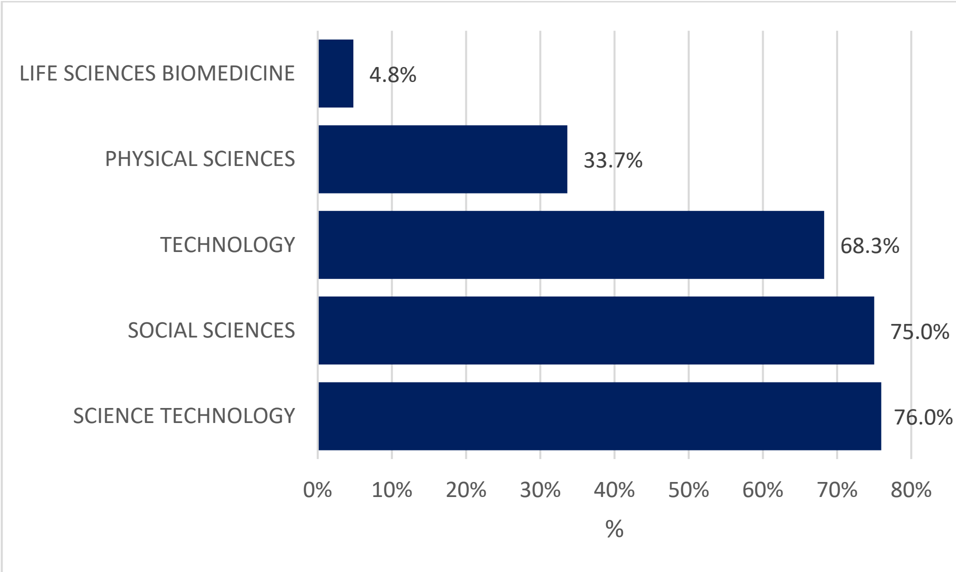

Figure 2. Research areas according to the WoS categories
(Source: Prepared by the authors using Web of Science data)

The geographical participation of the authors of the selected papers is shown in Figure 3. Most of the papers have had the participation of Chinese authors, approximately 94.2%, followed by the USA with 14.4%, then Italy and other Eastern countries.

Figure 3. Participation in papers of the final sample by country.
(Source: Prepared by the authors using Web of Science data)

It is also noteworthy that the most relevant inter-country collaborations by the number of published articles are between the USA and China, followed by the UK and China, with some inter-country contributions between authors from other Eastern countries, but in a smaller proportion (see Figure 4).

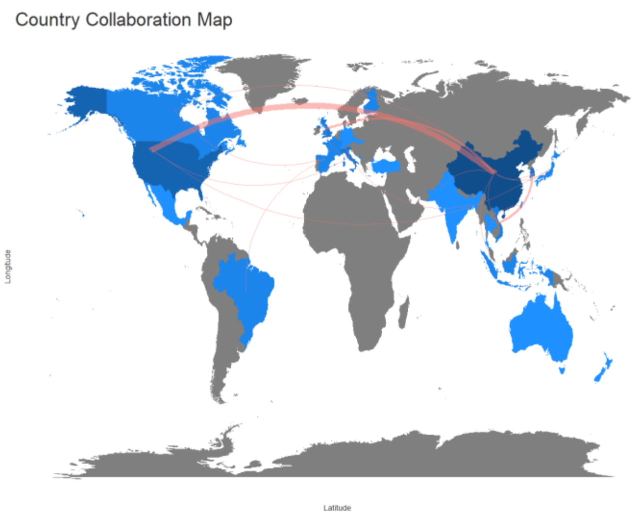

Figure 4. Map of the contributions between countries according to co-authorship
(Source: Prepared by the authors using bibliometrix and Web of Science data)

This result is unsurprising since P2P had tremendous growth in China in the last decade. However, at the time of this analysis, the market of this country has deteriorated due to the lack of an adequate regulatory framework. This market presented a significant increase in illegality and fraud, few risk managements, overshadowing the entities shown to have a robust offer and efficient risk management, forcing intervention and closure of many platforms. Figure 5 shows the number of P2P platforms across the world, but the figure was not available for China. The 10 countries in their order with the highest number of platforms are the USA, the UK, Indonesia, Germany, Mexico, Switzerland, Latvia, Estonia, Spain and South Korea.

Figure 5. Distribution of the number of platforms of P2P market lending by country.
(Source: Compiled by the authors based on data from https://P2Pmarketdata.com/ , accessed April 17, 2020)

3.1.2. Identification of the relevant papers using citation analysis

In this section, we use citation analysis to identify the papers with the greatest number of links, it means the greatest number of citations. Figure 6 shows the citation network, each circle represents a paper. The bigger circles represent papers that have more citations. Colors are associated with each article. This map allows us to visualize the citation structure and identify the most relevant scientific community documents. We identify the most relevant papers to answer RQ1.

Consequently, according to Figure 6, the main articles are: Greiner and Wang (2010), Herzenstein et al. (2011), Emekter et al. (2015), Serrano-Cinca et al. (2015), Malekipirbazari and Aksakalli (2015), Serrano-Cinca and Gutierrez-Nieto (2016) and Xia et al. (2017), being Emekter et al. (2015) the most cited publication. However, the relevance of the papers has to be leveraged by the time they have been published, because recent papers have fewer opportunities to be cited. Thus, we believe that papers such as Serrano-Cinca and Gutierrez-Nieto (2016) and Xia et al. (2017) will have a greater impact in the coming years. Both papers deal with risk management models with a business focus, which we believe make them very suitable for industry.

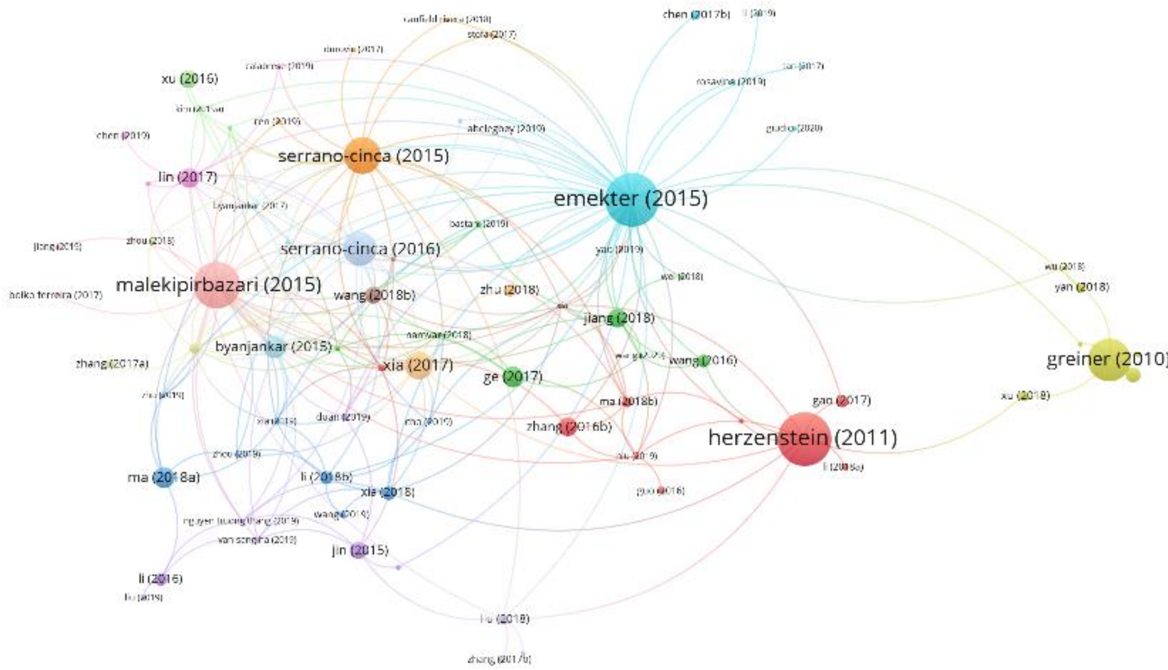

Figure 6. Main authors’ citation graph
(Source: Prepared by the authors using VOSviewer based on the Web of Science data)

The most cited papers are mostly devoted to identifying the key factors in the credit risk models in P2P lending. They use mainly statistical and econometric models. They propose cross-sectional models and some survival models. These models try to predict endogenous variables such as the approval of the credit contract, the prediction of default, the performance, or the profit on a credit obligation. However, works such as Malekipirbazari and Aksakalli (2015) and Xia et al. (2017) try to improve credit risk management prediction and obtain profits incorporating machine learning techniques, mainly decision tree methods.

After the citation analysis, we have identified the main topics, concepts, or research lines associated with risk and profit management in P2P lending through clusters. As previously mentioned, the clusters are based on identifying communities on the networks of occurrence and co-occurrence of terms (nodes) from English-language textual data found in the abstracts and the titles of publications or keywords prepared by the authors. Co-occurrence determines the links in the network. It is associated with sharing terms from titles and abstracts, which is informative about the relatedness of publications. The advantage of using VOSviewer for this analysis lies in combining clustering techniques to find communities in networks of terms obtained from text mining, the optimization of modularity functions, and mapping and visualization using multivariate analysis techniques. The resulting plots show as many labels and nodes as possible, prioritizing those most relevant nodes over the less critical ones.

For cluster analysis, we will use the descriptive summary in Table 2 and Figures 7, 8, and 9. In these figures, the nodes (terms) of the network are presented in two dimensions through nodes and labels so that nearby nodes represent related nodes. In Figures 7 and 8, the nodes' size and label indicate the number of publications in which each term appears. The color of the nodes in Figure 7 corresponds to the cluster to which each term has been assigned. For the definition of clusters, the variable of time of publication is used only as an illustrative variable. Figure 8 shows the terms associated with more recent publications in light colors and the older in dark colors. Finally, on the structure of the co-occurrence network's visualization map, Figure 9 displays a map where the terms' density is displayed through colors ranging from blue with low density to red with high density. The color depends on the density or frequency of each term of the density of the neighboring terms, allowing to complement the previous analyses by identifying the most relevant topics in the analyzed co-occurrence structure.

According to our cluster analysis, we found five clusters with 12, 10, 10, 8 and 7 elements respectively. We have named them as: new sources of information, evaluation and interpretability; investment and profit models; methodological innovations; business and regulation; and standard modelling, respectively.

Table 2 shows a summary of the five topics found. The first column (*Name of Cluster*) represents the name we assigned to each cluster according to our characterization. The second column symbolizes the labels of the terms grouped in each cluster. The third column (*Occur.*) represents the number of papers in which each term appears. The fourth one (*Avg. Occur*) shows the average number of documents in which each term appears. The fifth (*Avg. pub. Year*) indicates the average of the year of publication of the term considering the articles that appear in. The sixth column (Avg. Avg. pub. Year), the average year of publication for all the terms of the cluster. The seventh represents the number of links of the node (the links represent co-ocurrence). And finally, the eighth (*Avg. Links*) shows the average number of links for each term within a cluster.

Table 2. Descriptive summary of clusters (communities) on co-occurrence networks based on abstracts and titles.

Name of Cluster	Label	Occur.	Avg. Occur.	Avg. pub. Year	Avg. Avg. pub. Year	Links	Avg. Links
New sources of information, evaluation (red)	performance_accuracy	51	10.9	2017.6	2017.4	44	22.8
	ann	17		2017.4		33	
	machine learning	15		2018.1		33	
	evaluation	8		2017.0		21	
	cross-validation	7		2018.0		19	
	metrics	7		2017.1		25	
	big data	6		2017.7		14	
	opportunity	5		2017.2		23	
	network connectivity	5		2018.2		16	
	fraud	4		2017.3		17	
	transaction-features	3		2017.3		14	
	social data	3		2016.3		15	
Investment and profit models (green)	investment	25	10.6	2018.0	2017.9	34	22.8
	prediction	24		2018.2		38	
	challenge	12		2018.3		27	
	profit	11		2017.8		22	
	deep learning	10		2018.6		27	
	need	8		2018.0		18	
	efficiency	7		2018.0		26	
	traditional method	3		2017.0		10	
	effectiveness	3		2017.7		12	
	feature selection	3		2017.7		14	
Methodological developments (blue)	classifier	25	8.1	2018.1	2018.4	40	22.1
	decision trees	11		2017.8		29	
	regression	8		2018.4		22	
	gradient boosting	8		2018.5		26	
	natural language processing	6		2018.7		23	
	imbalance	6		2018.5		22	
	ensemble	5		2017.8		19	
	soft information	5		2019.2		16	
	hard information	4		2018.5		15	
	fuzzy	3		2019.0		9	

Business and regulation (yellow)	technology	10	7.5	2018.4	2017.4	26	17.6
	business knowledge	10		2017.4		22	
	regulatory authority	9		2017.9		21	
	information asymmetry	9		2016.4		19	
	operation	8		2017.4		20	
	regulation	8		2017.4		16	
	collateral	3		2016.3		9	
	government	3		2017.7		8	
Standard modelling and interpretability (purple)	scoring	33	14.0	2017.6	2017.7	41	25.4
	logistic regression	20		2018.2		32	
	probability	18		2017.9		31	
	trust	9		2016.6		19	
	random-forest	7		2017.1		21	
	interpretability	6		2018.7		20	
	survival analysis	5		2017.6		14	

(Source: Prepared by the authors using VOSviewer based on the Web of Science data)

We will characterize each of the clusters in the following way. The biggest and most diverse area is "*new sources of information and evaluation*" (red colour). The terms related to the evaluation, accuracy and performance measurement through different metrics stand out, together with the cross-validation process mainly in the machine learning proposals. According to Figure 9, the term of performance accuracy is one of the most frequent topics in the titles and abstracts of our sample. In addition, in the same cluster, we find that the use of new sources of data which include social data and information derived from connectivity networks and they are typically incorporated to traditional models mainly as alternatives to artificial neural networks (ANN) and machine learning methods in general. These alternatives are also shown for fraud models. This cluster, in particular, does not include topics heavily studied in the last two years, being the connectivity networks the most recent at the beginning of 2018 (see Figure 8).

In the topic of "*investment and profit models*" (green), there are terms related to alternatives to the probability models, as the profit models, the investment models, among other models related to business. In this kind of other options, obtaining the right prediction is a challenge, so realizing a fair feature selection process is crucial for an efficient modeling process.

For this segment, deep learning algorithms also stand out as alternatives in the most recent topics, see Figure 8.

The "*methodological developments*" (blue) cluster has the largest number of current terms. It includes terms related to new strategies used in credit risk modelling. For example, the term of class imbalance which is one of the main challenges in classifiers with low default rates. The term "ensemble" also appears, mainly, for the tree techniques based on boosting gradient. At the same time, the fuzzy and natural language processing (NLP) stand out with high occurrence, as new ways to represent or complement data (see Figures 7, 8 and 9 below).

The "*business and regulation*" (yellow) cluster is closely related to the business structure, regulation and elements of credit risk theory in the P2P market. These issues are more relevant here than in the traditional market due to the lack of central intermediaries. The terms are related to information asymmetry, the lack of collateral and the issue of governance (see Figure 7). The terms of this section are the least recent, (see Figure 8), with a time average associated with the first semester of 2017 (see Table 2). According to their frequency, the most relevant terms are technology, business and regulatory framework, being the first one studied in more recent publications (see Figure 8).

In the cluster of "*standard modelling*" (purple), there are elements and terms of the classic risk management, such as the scoring model, probability of default, for example, as one of its relevant inputs. Logistic regression is typically used as a benchmark model in many publications (see Figure 7). Random forests also appear as it is one of the most popular machine learning methods used in this market. In turn, the survival models, a classic alternative to default probability models, are widely used in expected loss estimation models, thanks to the longitudinal evaluation and their great interpretability. It is precisely the topic of interpretability, which is a relevant and outstanding aspect in the papers of recent years, one of the gaps found in several machine learning proposals (see Figure 8).

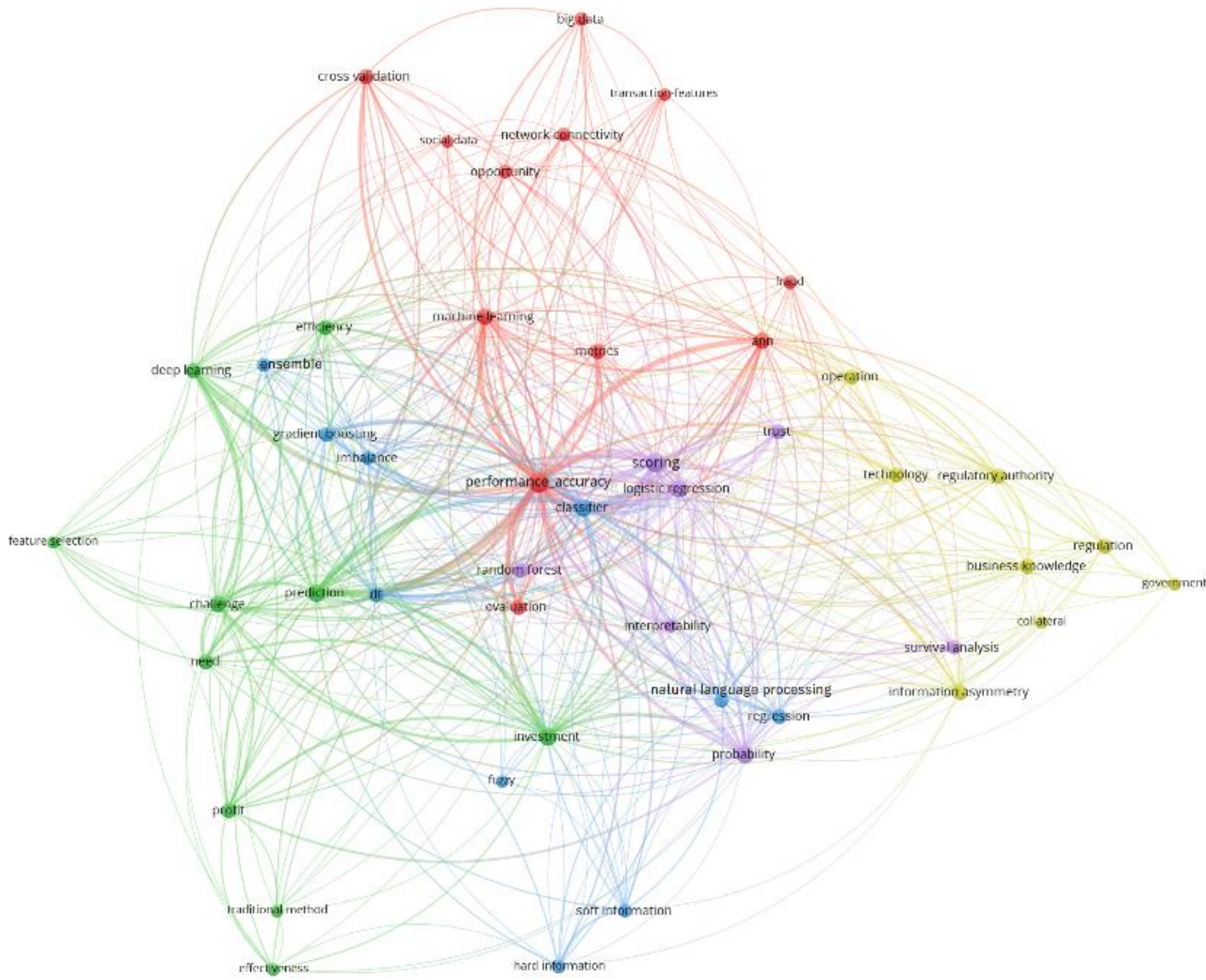

Figure 7. Graph of the co-occurrence and clusters for the concepts present in titles and abstracts. (Source: Prepared by the authors using VOSviewer based on the Web of Science data)

As mentioned before, Figure 8displays the evolution of the terms. According to the figure, a recent trend that consists of using new methodologies in the P2P lending market such as deep learning, gradient boosting, fuzzy hard and soft information.

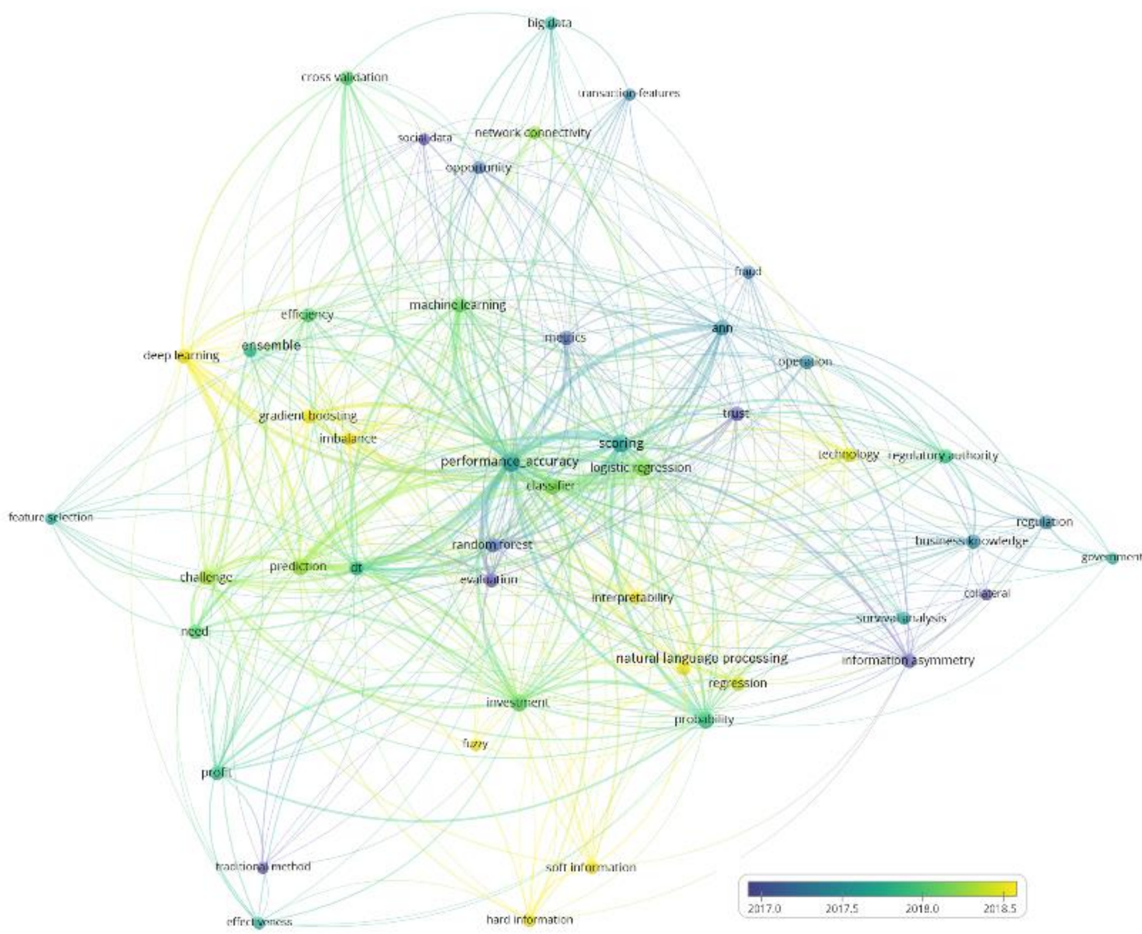

Figure 8. Graph of the co-occurrence and time trend for the concepts present in titles and abstracts.
(Source: Prepared by the authors using VOSviewer based on the Web of Science data)

According to Figure 9, the term of performance accuracy is one of the most frequent topics in the titles and abstracts of our sample, representing by hot color as red. In addition, in the same cluster, we find that the use of new sources of data which include social data and information derived from connectivity networks and are typically incorporated to traditional models, mainly, as alternatives to artificial neural networks (ANN) and machine learning methods in general. These alternatives are also shown for fraud models. This cluster, in particular, does not include topics heavily studied in the last two years, being the connectivity networks the most recent at the beginning of 2018 (see Figure 8).

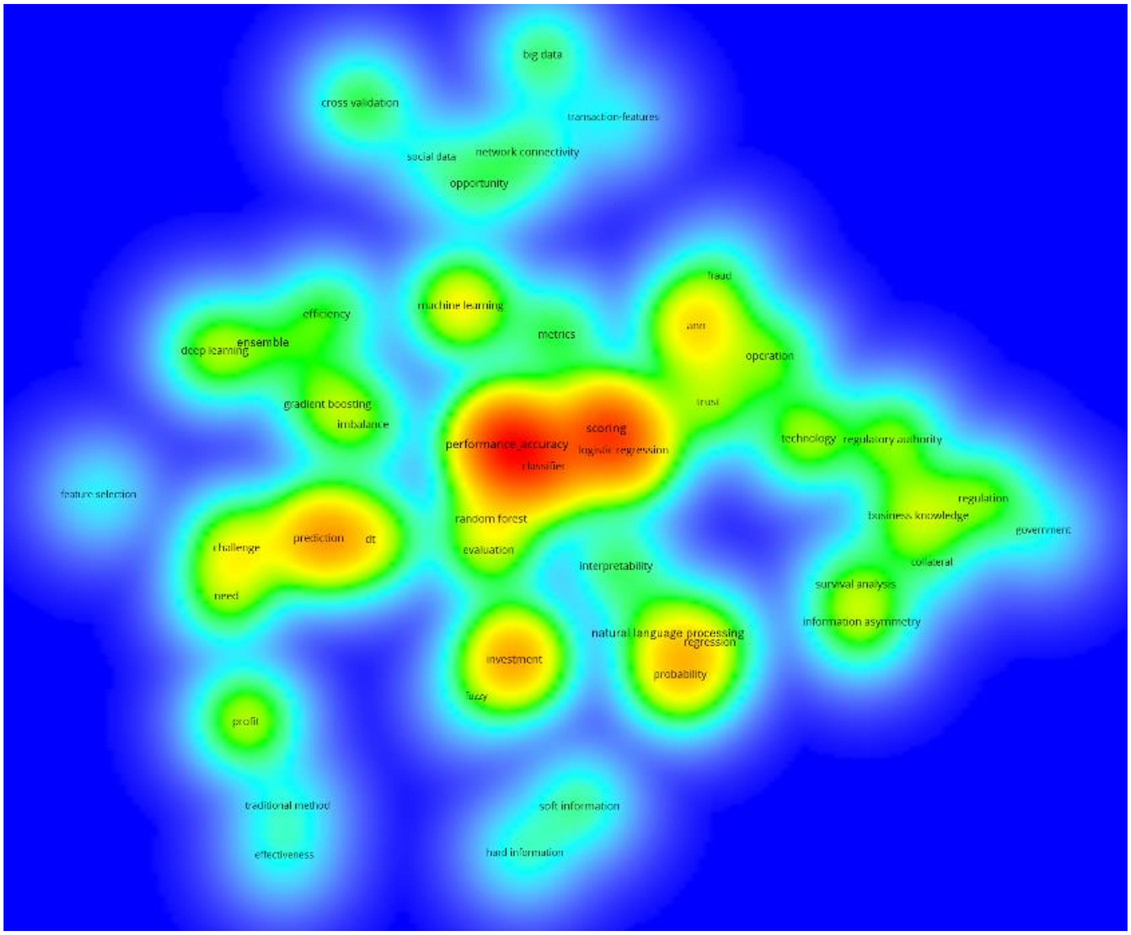

Figure 9. Heat map of the co-occurrence for the concepts present in titles and abstracts. (Source: Prepared by the authors using VOSviewer based on the Web of Science data)

From Figures 10 and 11 many aspects are similar to co-ocurrence terms analysis with abstracts and titles, but also new terms such as microfinance and microcredit. This indicates the association of the P2P market with the supply of alternative financial services for low-income people and small businesses, leveraging financial inclusion.

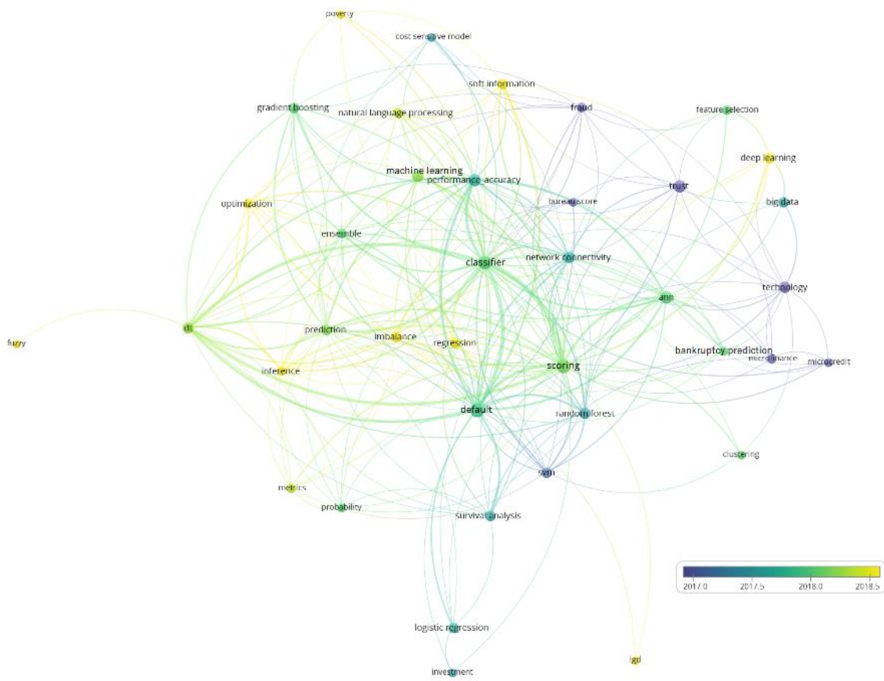

Figure 10. Graph of co-occurrence and time trend for concepts present in keywords.
(Source: Prepared by the authors using VOSviewer based on the Web of Science data)

We can find also that the use cost-sensitive models stand out in publications in mid-2017, as an alternative that includes business elements in credit risk management. Less frequently, the term of one of the relevant components of the expected loss calculation, the loss given default (LGD) appears during 2018, showing the potential that the research on this topic may have for this market. The machine learning methods used through time are support vector machine and random forest, around 2017, later the use of gradient boosting, and more recently the use of deep learning. Again logistic regression appears as a constant, as it is usually the method of reference. Finally, the issue of inference also appears in the last two years as one of the important elements to be incorporated into the new modelling alternatives to support their credibility (see Figure 10).

In Figure 11, the most co-occurrence for the keywords of the research papers from our sample is shown. Default, scoring, classifier, machine learning and performance accuracy are the most frequent topics, evidencing the problem, the tools proposed, and the purpose pursued.

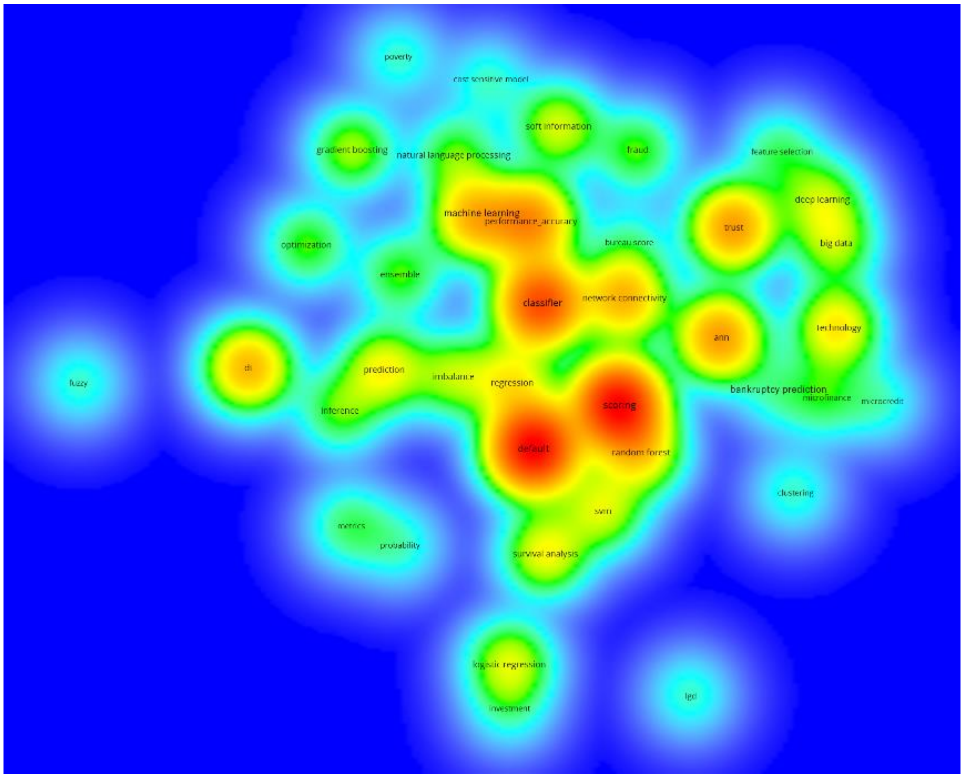

Figure 11. Heat map by co-occurrence for concepts present in keywords.
(Source: Prepared by the authors using VOSviewer based on the Web of Science data)

3.2. Selection and classification of the relevant papers using systematic review

In addition to the bibliometric analysis that shows the main trends and topics, in this section, we will carry out a systematic review of the literature on P2P lending. We will analyze several aspects of the modelling proposals, focusing on risk and profit management, in order to answer the RQ2. We analyzed a sample of 98 papers according to the filters defined in section 2.2.

3.2.1. General aspects. Datasets and software.

We will start our systematic review analysing the data set used to assess the strategies for risk and profit management in P2P lending. We will identify the dataset by the name of the company, but also by its country. Table 3 shows the frequency of the data sets in the literature. The Lending Club (LC) data set is the most popular. Lending Club is one of the platforms with the highest volume of users in the United States. The data has been used in 35 research papers analysed, about 35% of the papers analysed. Probably this is due to the quality, variety and quantity of the data. The fact of being available through Kaggle, the popular data science website, also makes it possible to work with it for many people from different countries. As a result, it is a data set studied from different perspectives and with both standard and novel methodological strategies.

Table 3. Datasets

Data sets (Observation Unit)	Country/Region	Number	Authors/year
Three undefined institutions (Individuals)	Australia, Germany and Brazil	1	Ding et al. (2017)
Two undefined institutions and Lending Club (Individuals)	Australia, Germany and the United States	2	Nguyen Truong et al. (2019); Van-Sang et al. (2019)
Eloan (Individuals)	China	1	Jiang et al. (2018)
HNL-Home of Network Loan (p2p Lending Intermediaries)	China	1	Li et al. (2019)
Jinan Hengxin Micro-Investment Advisory Co., Ltd. (Individuals)	China	2	Zhang, Wang, Chen, Shang, et al. (2017); Zhang, Wang, Chen, Zhao, et al. (2017)
LendingMarket (Individuals)	China	1	Xu and Chau (2018)
MyLending (Individuals)	China	1	Xu et al. (2016)
Paipai (Individuals)	China	1	Chen (2017)
Paipaidai (Individuals)	China	2	Xinmin et al. (2019); Zhu (2018)
Ppdai (Individuals)	China	6	Chen et al. (2019); Xu et al. (2015); Zhang, Wang, et al. (2016); Zhang, Jia, et al. (2016); Zhang, Geng, et al. (2017); Zhao (2015)
Renrendai (Individuals)	China	5	Gao et al. (2017); Li, Hao, et al. (2018); Liu et al. (2018); Tao et al. (2017); Yao et al. (2019)
Undefined platforms (Individuals and enterprises)	China	16	Chen et al. (2016); Guo et al. (2016); Jiang et al. (2019); Li, Ding, et al. (2018); Li et al. (2020); Niu et al. (2019); Ma, Zhao, et al. (2018); Wang, Jiang, et al. (2018); Wang et al. (2019); Xia and Li (2016); Xu and Zhang (2017); Yan et al. (2016); Yan et al. (2017); Yang et al. (2019); Yuan et al. (2018); Zhou et al. (2019)
Wangdaizhijia P2P online loan industry portal (Platforms)	China	2	Fu et al. (2019); Ge et al. (2017)
Yooli (Individuals)	China	1	Lin et al. (2017)
Lending Club, We.com (Individuals)	China and United States	3	Xia et al. (2017); Xia et al. (2018); Xia (2019)
European External Credit Assessment Institution-ECAI (SMEs)	Europe	1	Ahelegbey et al. (2019)
Modefinance a ECAI (SMEs)	Europe	2	Giudici et al. (2020); Hadji-Misheva et al. (2018)
Bondora (Individuals)	Europe. Estonia, Finland and Spain	2	Byanjankar et al. (2015); Byanjankar (2017)
Bandung (SMEs)	Indonesia	1	Rosavina et al. (2019)
Undefined platforms (Individuals)	Indonesia	1	Amalia et al. (2019)
Undefined platform (Individuals)	Mexico	1	Canfield (2018)
Funding Circle (Individuals)	United Kingdom	1	Pierrakis (2019)

KIVA non-profit organization (Individuals)	United States	1	Uddin et al. (2018)
Lending Club (Individuals)	United States	30	Bastani et al. (2019); Boiko Ferreira et al. (2017); Cai and Zhang (2020); Calabrese et al. (2019); Cho et al. (2019); Duan (2019); Durovic (2017); Emekter et al. (2015); Gourieroux and Lu (2019); Jin and Zhu (2015); A. Kim and Cho (2019); Kim and Cho (2019a); Kim and Cho (2019b); Kumar et al. (2016); Ma, Sha, et al. (2018); Malekipirbazari and Aksakalli (2015); Namvar et al. (2018); Rodrigues et al. (2018); Serrano-Cinca and Gutierrez-Nieto (2016); Serrano-Cinca et al. (2015); Stofa (2017); Wang, Kou, et al. (2018); Wan et al. (2019); Wang et al. (2020); Wei et al. (2018); Xia et al. (2019); Ye et al. (2018); Zang et al. (2015); Zhou et al. (2018); Zhu et al. (2019)
Prosper (Individuals)	United States	5	Greiner & Wang (2010); Herzenstein et al. (2011); Ren et al. (2019); Tan et al. (2017); Wang et al. (2016)
Data simulated	-	1	Lee et al. (2017)
No Information	-	2	Ji et al. (2020); Li et al. (2016)
No use data sets	-	5	Liu and Yan (2016); Pokorna and Sponer (2016); Pur et al. (2014); Wang (2018); Xiong (2018)

Source: Compiled by the authors

Apart from the Lending Club case, there is a strong tendency to use data from China. This may be caused due to the growth of the market there, the availability of the information, and the concern of an important group of researchers who seek to provide answers to the regulatory problems of this market. In the literature, the most used datasets are ppdai.com, renrendai.com, and we.com.

By country, the data use is closely related to the development of P2P markets globally, Figure 5, with data from China and the United States mostly, followed by Europe's data, including the United Kingdom, Mexico, and even data Southeast Asia in lesser proportion.

It is important to mention that most of the studies have been carried out on loans to individuals and, in a few cases, the P2P business lending market has been selected as the object of study, particularly SME's. Especially, the papers by Giudici et al. (2020), Ahelegbey et al. (2019) and Hadji-Misheva et al. (2018) use information from some European External Credit Assessment

Institution (ECAI), such as modeFinance, an agency specialized in companies and banks, a Fintech credit rating evaluation in Europe.

Table 4 focuses on the software used. *R* and *Python* are the most popular programming environments used. In Python, *Scikit-learn* (machine learning library) and *Keras* (neural networks library) stand out for their frequency of use. Other software such as *Tensorflow* for research involving deep learning, usually combined with *Python*, is also used. Software packages such as *SPSS*, *STATA* and *WEKA* are also employed, but to a lesser extent (the first two mainly for traditional statistical and econometric models). Interestingly, only 47 papers report the software used.

Table 4. Software used

Software	Number	Authors/Year
LISREL	1	Amalia et al. (2019)
MATLAB	2	Yang et al. (2019); Zang et al. (2015)
Python	10	Boiko Ferreira et al. (2017); Cho et al. (2019); Ding et al. (2017); Li et al. (2019); Li et al. (2020); Li, Ding, et al. (2018); Namvar et al. (2018); Rodrigues et al. (2018); Xia (2019); Xia et al. (2017)
Python, TensorFlow	5	Bastani et al. (2019); Fu et al. (2019); Nguyen Truong et al. (2019); Van-Sang et al. (2019); Wang et al. (2019)
Python, SPSS, TensorFlow	1	Duan (2019)
Python, STATA	1	Niu et al. (2019)
R	9	Byanjankar (2017); Byanjankar et al. (2015); Calabrese et al. (2019); Chen et al. (2019); Giudici et al. (2020); Jiang et al. (2019); Wan et al. (2019); Wang, Jiang, et al. (2018); Xia et al. (2019)
R, Weka	1	Malekipirbazari and Aksakalli (2015)
R, Stanford CoreNLP, SentiStrength	1	Wang et al. (2020)
R, MATLAB, Portfolio Safeguard (PSG)	1	Wei et al. (2018)
R, Bazhuayu web-crawler software	1	Yao et al. (2019)
SPSS	7	Emekter et al. (2015); Greiner and Wang (2010); Jin and Zhu (2015); Serrano-Cinca and Gutierrez-Nieto (2016); Yan et al. (2016); Yan et al. (2017); Xu and Chau (2018)
STATA	3	Canfield (2018); Chen et al. (2016); Chen (2017)
Weka	3	Cai and Zhang (2020); Wang et al. (2016); Xu et al. (2016)
Weka, Liblinear, LibSVM	1	Guo et al. (2016)

Source: Compiled by the authors

3.2.2. The business problem addressed in risk and profit management in the P2P market

In this section, we propose several categories associated with different problems of risk management, profit and market knowledge. It is worth mentioning that a paper can be found in several categories, because they can analyse and propose solutions, made up of various models dealing with different aspects of the market. The categories defined are listed in Table 5.

Table 5. Business problem

Business problem	Number	Authors/Year
Default classification /Default probability	)	Ahelegbey et al. (2019); Bastani et al. (2019); Boiko Ferreira et al. (2017); Byanjankar et al. (2015); Byanjankar (2017); Cai and Zhang (2020); Calabrese et al. (2019); Canfield (2018); Chen (2017); Chen et al. (2019); Chen et al. (2016); Cho et al. (2019); Ding et al. (2017); Duan (2019); Durovic (2017); Emekter et al. (2015); Gao et al. (2017); Ge et al. (2017); Giudici et al. (2020); Gourieroux and Lu (2019); Guo et al. (2016); Herzenstein et al. (2011); Jiang et al. (2018); Jiang et al. (2019); Jin and Zhu (2015); A. Kim and Cho (2019); Kim and Cho (2019a); Kim and Cho (2019b); Kumar et al. (2016); Li et al. (2016); Li, Ding, et al. (2018); Li et al. (2020); Lin et al. (2017); Liu et al. (2018); Ma, Sha, et al. (2018); Ma, Zhao, et al. (2018); Malekipirbazari and Aksakalli (2015); Hadji-Misheva et al. (2018); Namvar et al. (2018); Nguyen Truong et al. (2019); Niu et al. (2019); Rodrigues et al. (2018); Serrano-Cinca et al. (2015); Serrano-Cinca and Gutierrez-Nieto (2016); Stofa (2017); Tan et al. (2017); Tao et al. (2017); Uddin et al. (2018); Van-Sang et al. (2019); Wang et al. (2016); Wang et al. (2019); Wang et al. (2020); Wang, Jiang, et al. (2018); Wang, Kou, et al. (2018); Xia (2019); Xia et al. (2017); Xia et al. (2018); Xia et al. (2019); Xinmin et al. (2019); Xu and Chau (2018); Yang et al. (2019); Ye et al. (2018); Yuan et al. (2018); Zang et al. (2015); Zhang, Jia, et al. (2016); Zhang, Wang, Chen, Shang, et al. (2017); Zhang, Wang, Chen, Zhao et al., (2017); Zhou et al. (2019); Zhu (2018); Zhu et al. (2019)
Fraud	3	Li et al. (2019); Xu et al. (2015); Xu et al. (2016)
LGD	3	Gourieroux and Lu (2019); Xia et al. (2017); Zhou et al. (2018)
Market knowledge: Risk and regulation (Descriptive, theoretical and qualitative analysis)	5	Liu and Yan (2016); Pokorna and Sponer (2016); Pur et al. (2014), Wang (2018); Xiong (2018)
Prepayment	1	Wan et al. (2019)
Profit Model/profit scoring model	6	Bastani et al. (2019); Cho et al. (2019); Serrano-Cinca and Gutierrez-Nieto (2016); Tan et al. (2017); Xia and Li (2016); Ye et al. (2018)
Risk and investment/Lending decision/Asset allocation	25	Amalia et al. (2019); Chen et al. (2016); Fu et al. (2019)； Gao et al. (2017); Greiner and Wang (2010); Herzenstein et al. (2011); (Ji et al., 2020); Lee et al. (2017); Li, Hao, et al. (2018); Pierrakis (2019); Ren et al. (2019); Rosavina et al. (2019); Tan et al. (2017); Tao et al. (2017); Wei et al. (2018); Xinmin et al. (2019); Xu and Chau (2018); Xu and Zhang (2017); Yan et al. (2016), Yan et al. (2017); Yao et al. (2019); Zhang, Geng, et al. (2017); Zhang, Wang, et al. (2016); Zhao (2015); Zhu (2018)

Source: Compiled by the authors

The models that try to estimate the default and have an approximation to the probability of this event, "*Default classification/Default probability"* category, are the most frequent with 70 papers, more than two-thirds of the studies analysed. Many of them propose new methodological approaches, new variables or combine various types of information. They typically exhibit remarkable results in classification and performance indicators (see Annex at the end of the paper for more details about each proposal). However, we draw attention to some limitations that we have found in many of the studies. First, although many of them describe alternatives for managing credit risk and are even intended to be used in the lending process, few ones draw attention to the selection of variables for this purpose. More precisely, many proposals include input variables that are indeed the output of the risk assessment, e.g. grade or levels of risk estimated by the platforms or interest rates, among others. As a result, these proposals could over or underestimate the performance indicators, and they cannot be used because they do not suit a real-life lending assessment process.

Second, few papers specify the kind of models they propose for risk management (application model, behavioural model, collection model, among others). We believe that it is essential to make such a distinction since this determines the moment of application in risk management, the potential variables to be used, and the modelling methodology.

Third, there are many ways in which the default event is defined, but some of the studies do not even detail this aspect when it is crucial for interpreting the results. For example, it is different to have a model that flags default with 30 days past due in one year, to a model that defines default as an unpaid obligation during the life of the credit. This kind of distinction must be considered when designing proposals, and, in turn, it must be accompanied by a good understanding of the business and regulatory needs. As an example, some papers in which default is defined as in charged-off transactions, analyzed against fully paid transactions at the end of the credit life: Serrano-Cinca and Gutierrez-Nieto (2016), Kim and Cho (2019), Ye et al. (2018), Cho et al. (2019), Rodrigues et al. (2018) or Stofa (2017). The other definitions can be found in the Annex.

All these facts lead to the limitation of the application due to lack of uniformity and consistency. These two characteristics are objectives sought by the regulatory bodies, especially, for transnational growing markets such as P2P lending. From the joint work between academics, those who propose models for risk management, regulators, and the platforms, tailored solutions can be presented, which generate greater confidence and probability of application. This synergy would ensure a better understanding of the business's nature and structure, which is reflected in better modelling, risk management, and service, enhancing the sustainability and development of the P2P lending market.

We also want to highlight that most of the papers are focused on classification, which is valuable for lending processes. However, the robust estimation of the probability of default (PD) needs to be studied in more depth. This aspect is crucial in the application and lending processes, it is a component in the classification process, and assists in the rate allocation and provisioning processes to better manage risk, investment, and regulatory processes. The above implies calibration methods, incorporation of forward-looking macroeconomic factors, and the search for other different factors, which are recognized that they have been entering into the modeling processes, as shown in the list of independent variables in Annex.

Another field of work is the "*Fraud*" models, necessary for the sustainability and efficiency of the market. Financial institutions have made use of fraud early warning systems more frequently in recent years, and it has become one of the relevant tasks in risk management. In the case of the P2P market, we find the relevant works of Li et al. (2019), Xu et al. (2015) and Xu et al. (2016).

Besides the category of models for the estimation of default, other models recognise the different components of credit risk management, such as those associated with the estimation of the "*LGD*", the loss in case of default (see this category in Table 5), enabling investors to evaluate the interest rate according to the expected loss and make better decisions. This is an essential component for lenders to assess the profitability of their investment. In this line, we can find the paper of Zhou et al. (2018), that try to study the distribution and to identify some of their determinants. In addition, Gourieroux and Lu (2019) propose a stress testing tool through a Least Impulse Response Estimator. This tool evaluates the impact on the proportion of expected loss according to loan variables and others, recognizing the relationship between PD and LGD through an indicator called total loss ratio. On the other hand, the paper of Xia et al. (2017) incorporates the information of this LGD component in the models of the default estimation through a cost-sensitive proposal. There are also papers such as Lee et al. (2017) which, through network models and a theoretical financial model, evaluate and project the positive impact that the presence of collateral has on the P2P lending market.

We describe another category "*Market knowledge: Risk and Regulation"*, which present descriptive, theoretical, contextual and qualitative elements of the P2P markets, including benefits, actors, risks and failures market. These works without established datasets study how these markets can ensure their development, continuity and stability as an alternative for investment and financing of financial market agents. They suggest several elements of regulation and supervision. This group includes works such as Liu and Yan (2016), Wang (2018), Pokorna and Sponer (2016), Pur et al. (2014) and Xiong (2018). It is worth to mention a particular paper, Wan et al. (2019), in the "*Prepayment"* category (see Table 5). It provides further elements of analysis for investor lenders and platforms based on survival models, analysing prepayment risk.

Another line of research seeks to model the rate of return of investors, identify the determinants or offer sets of better investment alternatives through "*Profit Model/Profit Scoring" model* alternatives (see Table 5). Within this category, there are studies, such as Serrano-Cinca and Gutierrez-Nieto (2016), Xia and Li (2016) or Ye et al. (2018), that propose a profit score combining Random Forest (RF) and genetic algorithms (AG) (see the Annex for details of the objective, independent variables and the methodologies used). Other papers in this category that also propose default prediction models, incorporate these results in profit scoring models. These are the cases of Bastani et al. (2019) that propose a two-stage model using deep learning, or Tan et al. (2017) that establishes a recommendation model, maximizing the total profit of investors, incorporating a novel variable, the Time Value of Money Prediction (TVM). Also, Cho et al. (2019), with their proposal of an instance-based entropy fuzzy support vector machine (IEFSVM) and several methods for the treatment of imbalance, use the return on credit operations of the transactions with less risk of default as a final criterion for investment decisions.

Another important field of research is related to the investment, recommendation or decision models, which we have grouped under the category of "*Risk and investment/Lending decision/Asset allocation"* (see Table 5). Here it can be found papers such as those by Zhang, Wang, et al. (2016) and Zhang, Geng, et al. (2017), which use different collaborative filtering algorithms. Ren et al. (2019) raises the problem of investment decisions under the Markowitz theory, taking into account variables such as the number of days required for an application to be fully funded. Xu and Zhang (2017) propose an investment index per platform based on the Analytic Hierarchy Process. Ji et al. (2020) with their proposal for risk ranking through fuzzy methodologies and the interactive multi-criteria decision-making method (TODIM). See the Annex for further details. We include Lee et al. (2017), who evaluates the impact of collateral on P2P lending as a tool to generate trust and a more efficient operation based on risk and investment. Tan et al. (2017) also propose a recommender system, allowing investors to choose the most profitable and least risky borrowers. In this category, we have also included the papers by Herzenstein et al. (2011), Greiner and Wang (2010) and Chen et al. (2016). They study the determinants of variables, such as the percentage of the requested amount that is funded by investors, along with the final interest rate, or the variation between the maximum rate payable versus the assigned rate (see also the Annex for details). In this way, they shed light on the factors associated with applicants that are used by investors to make reliable investment decisions. For this same purpose, Herzenstein et al. (2011) study credit performance, including narrative variables, which are combined with demographic variables and loan characteristics. Chen et al. (2016) proposes to mix regression and survival models to investigate the possible gender discrimination in the P2P credit loan market in China. They find empirical evidence that supports their hypothesis.

In the same category, the papers by Tao et al. (2017), Zhu (2018), Xu and Chau (2018) and Xinmin et al. (2019) that model both the prediction of default and whether the credit requested can be funded under traditional statistical methodologies. They also propose new variables related to information disclosure, guarantees and the safety promise offered by the platforms. Concretely, Xu and Chau (2018) includes soft information, among other inputs, while Xinmin et al. (2019) focuses on the impact of the variable successful borrowing times. Yao et al. (2019) evaluates with text mining techniques the influence of the description of the loan purpose. Finally, Wei et al. (2018) uses alternative transformations on the variables and an alternative objective function (buffered AUC), to understand the loan approval processes. See the Annex for more details on each proposal.

In this broad category, we can find studies that try to characterize and explain the market. For example, Yan et al. (2016) and Yan et al. (2017) study the determinants of the number of investors per platform in China and identify variables that impact positively such as social capital, risk management, and operating duration. Gao et al. (2017), in addition to studying the determinants of the 30 days past due default, studies the determinants of the percentage of bad debt in the total amount of a loan, including variables from two segments forward-looking and backward-looking, the latter to identify variables from previous experience with the platforms. On the other hand, Li, Hao, et al. (2018), using connectivity networks and regression models, analyzes the systemic risk from topological variables derived from the network generated by the P2P market. Zhao (2015) simulates, using neural network models, expert risk assessments on credit granted.

Pierrakis (2019) analyses a survey to lenders from a platform in the United Kingdom, using Principal Component Analysis (PCA), to study their investment criteria and their motivation to invest. Rosavina et al. (2019) uses qualitative methods to explore the determinants of borrowing for SMEs in Indonesia. Also, they evaluate some aspects including collateral, requirements, process duration, interest rate, costs, and profits, among others.

We also include works such as Fu et al. (2019), which help us understand the dynamics of the P2P market for efficient decision making. They characterize the feelings of the investors' comments using Convolutional Neural Networks for Sentence Classification (TextCNN) and a Long and Short Time Memory neural network (LSTM). In addition, they predict an index of the daily trading volume of the Chinese platforms from a temporal series of changes that describes the dynamics of this variable together with other variables.

3.2.3. Methodological aspects.

In this section, we focus on the methodological aspects of the modelling process. We have identified ten non-exclusive categories that are shown in Table 6.

Table 6. Methodological aspects (93 papers, the 5 papers that did not use datasets were excluded)

Methodological aspects	Number	Authors/Year
Class imbalance treatment	9	Bastani et al. (2019); Boiko Ferreira et al. (2017); Cho et al. (2019); Ding et al. (2017); Malekipirbazari and Aksakalli (2015); Wang, Kou, et al. (2018); Xia et al. (2017); Ye et al. (2018); Zhu et al. (2019)
Descriptive analysis	3	Durovic (2017); Pierrakis (2019); Rosavina et al. (2019)
Ensemble methodology	7	Guo et al. (2016); Ding et al. (2017); A. Kim & Cho (2019); Li, Ding, et al. (2018); Li et al. (2020); Namvar et al. (2018); Zhou et al. (2019)
Explanation	37	Amalia et al. (2019); Byanjankar (2017); Calabrese et al. (2019); Canfield (2018); Chen (2017); Chen et al. (2019); Chen et al. (2016); Durovic (2017); Emekter et al. (2015); Gao et al. (2017); Ge et al. (2017); Giudici et al. (2020); Gourieroux and Lu (2019); Greiner and Wang (2010); Herzenstein et al. (2011); Jiang et al. (2018); Lee et al. (2017); Li, Hao, et al. (2018); Lin et al. (2017); Liu et al. (2018); Hadji-Misheva et al. (2018); Niu et al. (2019); Serrano-Cinca et al. (2015); Serrano-Cinca and Gutierrez-Nieto (2016); Stofa (2017); Tao

		et al. (2017); Wan et al. (2019); Xia and Li (2016); Xinmin et al. (2019); Xu and Chau (2018); Yan et al. (2016); Yan et al. (2017); Yang et al. (2019); Yao et al. (2019); Zhang, Jia, et al. (2016); Zhou et al. (2018); Zhu (2018)
Feature selection	9	Cai and Zhang (2020); Jiang et al. (2019); Jin and Zhu (2015); Kim and Cho (2019a); Kim and Cho (2019b); Nguyen Truong et al. (2019); Van-Sang et al. (2019); Xu et al. (2015); Zhu et al. (2019)
Inclusion of new variables	23	Ahelegbey et al. (2019); Fu et al. (2019); Gao et al. (2017); Ge et al. (2017); Giudici et al. (2020); Guo et al. (2016); Herzenstein et al. (2011); Jiang et al. (2018); Li et al. (2019); Li, Hao, et al. (2018); Ma, Zhao, et al. (2018); Hadji-Misheva et al. (2018); Niu et al. (2019); Wang et al. (2016); Wang et al. (2019); Wang et al. (2020); Wei et al. (2018); Xia et al. (2019); Xu et al. (2015); Xu and Chau (2018); Yao et al. (2019); Zhang, Jia, et al. (2016); Zhu (2018)
Outliers treatment	1	Li et al. (2016)
Performance/evaluation	42	Bastani et al. (2019); Boiko Ferreira et al. (2017); Byanjankar et al. (2015); Byanjankar (2017); Cai and Zhang (2020); Calabrese et al. (2019); Chen (2017); Chen et al. (2019); Cho et al. (2019); Duan (2019); Jiang et al. (2019); Jin and Zhu (2015); A. Kim and Cho (2019); Kim and Cho (2019a); Kim and Cho (2019b); Kumar et al. (2016); Li, Ding, et al. (2018); Li et al. (2019); Li et al. (2020); Liu et al. (2018); Ma, Sha, et al. (2018); Malekipirbazari and Aksakalli (2015); Nguyen Truong et al. (2019); Rodrigues et al. (2018); Uddin et al. (2018); Van-Sang et al. (2019); Wang et al. (2019); Wang, Jiang, et al. (2018); Wang, Kou, et al. (2018); Xia et al. (2017); Xia et al. (2018); Xia (2019); Xia et al. (2019); Xu et al. (2015); Yang et al. (2019); Ye et al. (2018); Yuan et al. (2018); Zang et al. (2015); Zhang, Wang, Chen, Shang, et al. (2017); Zhang, Wang, Chen, Zhao, et al. (2017); Zhao (2015); Zhu et al. (2019)
Rank by risk or profitability	6	Ji et al. (2020); Ren et al. (2019); Tan et al. (2017); Xu and Zhang (2017); Zhang, Wang, et al. (2016); Zhang, Geng, et al. (2017)
Reject inference	2	Xia et al. (2018); Xia (2019)

Source: Compiled by the authors

The largest category is "*Performance/evaluation*" because 42 publications address such aspect. Those articles focus on default classification and use performance measures as the main evaluation criterion. The most frequent measures are accuracy (ACC), even if the data is typically imbalanced, the area under the ROC curve (AUC), precision or recall. On the other hand, measures such as F1 score, a harmonic mean between recall and precision, or the H-measure, present a lower frequency. It is important to mention that just a few articles use inferential evaluation to establish which modelling alternative is better.

In this sense, we have paid attention to another aspect, "*Reject inference*". The articles in this category propose alternative reject inference techniques to correct the sampling bias generated by modeling only with the set of accepted loans, even though the models' use is established on the whole set of loan applications. This bias can affect the predictability of the models. The input features' distributions and the target variables' label proportions could be very different if researchers included the rejected applications' data set. In particular, Xia et al. (2018) and Xia (2019) combine semi-supervised methods and outlier detection methods with classifiers within the Gradient Boosting Decision Tree models (GBDT), respectively (see Annex for details). They use inferential tests (Friedman's or Shapiro-Wilk's) to evaluate their solutions' performance and strengthen their proposals. On the other hand, Niu et al. (2019) adds information from social networks, use the DeLong test for AUCs as a performance comparison criterion for the proposed models. This paper identifies the GBDT algorithm as the best alternative.

Another category includes the works that tackle the problem of unbalanced classes, present in P2P lending, "*Class imbalance treatment*" (Table 6). Typically, the percentage of default, fraud, or accepted applications, is the minority class and this is problematic since the algorithms mainly learn from the majority class. Additionally, the performance evaluation tends to give the same weights or costs to different classification errors, which is far from the market needs (Xia et al., 2017). Different strategies can alleviate this problem such as: resampling methods to balance classes, the adjustments of misclassification costs or the definition of cost functions to be minimized (either directly in the learning processes or after it), the calibration of the cut-off value of the models employing cost-sensitive models.

Bastani et al. (2019) and Zhu et al. (2019) use resampling techniques such as undersampling, oversampling or Synthetic Minority Over-sampling TEchnique (SMOTE). Regarding cost-sensitive models, Xia et al. (2017) provides an excellent description of this line of research and propose to adjusts the estimation process and learning of the extreme gradient boosting algorithm (XGBoost), taking into consideration the costs of misclassification. Similarly, Wang, Kou, et al. (2018), Ye et al. (2018), Boiko Ferreira et al. (2017) and Cho et al. (2019) use and compare cost-sensitive strategies with resampling strategies. However, Cho et al. (2019) additionally proposes and contrasts an instance-based entropy fuzzy support vector machine model (IEFSVM) for this problem, with which they select the least risky registers to compose a portfolio that seeks high expected returns based on an investment model. Ding et al. (2017) proposes an ensemble using a novel technique of undersampling, supported by clustering techniques. Likewise, Malekipirbazari and Aksakalli (2015) on their proposal of random forest (RF) incorporates a cost-weighted matrix that increases the cost of erroneous classification associated with borrowers with poor behaviour in default, through a cost-sensitive meta-algorithm incorporated in the WEKA software (see the Annex for more details).

On the other hand, another group of papers are focused on identifying determinants, validating relationship hypotheses, understanding the type of relationship of different potential variables with the target variable, for example, default, fraud, approval, interest rate decrease, number of investors per platform, among others (see Annex for the list of target variables used) . We have classified these studies in Table 6 under the "*Explanation*" category. This aspect is demanded by regulators, supervisors and users of machine learning methods. It is strongly associated with transparency and confidence in risk and profit management in the credit markets, particularly in P2P lending. We found that 31 out of 37 use classic statistical and econometric methods, such as linear regression models, logit, probit or survival models. Such models include the explanatory component through the interpretation of coefficients and testing of inferential hypotheses about them. The few proposals that present elements of explanation using machine learning alternatives only concentrate on feature importance. Therefore, they ignore relevant aspects such as monotonicity in general or by segments, the study of structural changes, identification of non-linearity, among others, which include theoretical contextual support. Thus, lacking a critical aspect demanded in the reliable use of methodological alternatives for managing risk and profit and strengthening the P2P lending market as a Fin-Tech alternative for financing and investment. We highlight the statistical and explainable methods' duality and note that few explainable machine learning applications, even though methodologically its progress has been substantial.

Among the proposals that try to emphasize the explanation component, we highlight Jiang et al. (2018) and Yao et al. (2019) for their methodological innovations. The authors, using text mining tools such as Latent Dirichlet Allocation (LDA), among other techniques, seek to incorporate and validate new information soft features through logistic regression (LR). Logistic regression (LR) is the most used model in the papers that have as the purpose of the prediction of default through classification models (see Annex for more details). Another important paper is Ge et al. (2017) that analyses the effect of social media information on the prediction of default. In addition, Chen et al. (2019) makes a proposal with logistic quantile regression (LQR) for the same objective variable and with explanatory elements. Amalia et al. (2019) uses the Structural Equation Model (SEM) to understand causally the aspects that generate confidence in the P2P market. Li, Hao, et al. (2018) and Hadji-Misheva et al. (2018) validate and incorporate determinants of credit risk associated with the topology of the p2p market agents' connectivity network into regression models. In line with them, Calabrese et al. (2019) proposes a flexible bivariate regression model, advisable for class imbalance events and recognizing the dependence in borrowers' behavior in the platforms and credit bureau background (see the Annex for details).

The category of “*Feature selection*” has 9 papers that highlight this component of the modelling (see Table 6). These studies are characterised by methodologies that select the best set of features, seeking to improve the indicators of evaluation of the performance of classification. Nevertheless, the explanation in the causal relationship that could exist between variables included is typically ignored in these studies. This aspect of machine learning and data science is still the subject of research and development today.

In the category of "*Inclusion of new variables*" we can find works that incorporate new sources and types of information that prove to be decisive in risk and profit management models in the P2P market. The main papers are written by Herzenstein et al. (2011), Wang et al. (2016), Jiang et al. (2018) and Yao et al. (2019), who use text mining algorithms within natural language processing (NLP) technologies. They include narrative features derived from aspects such as economic hardship, hardworking, successful, moral, and religious. Also, they add variables which are derived from aspects related to deception, subjectivity, sentiment, readability, personality, etc., or soft features related to asset, income, work, family, agriculture, length, capital turnover, investment and entrepreneurship, business expansion, the ambiguity of purpose, among others (see the Annex). These variables are mainly taken from loan titles, textual data generated from statements describing the purpose of the loan, and all descriptive text paragraphs related to each application.

Wang et al. (2020) proposes a soft factor mining method in terms of the distribution of the kinds of semantics expressed in the descriptive loan text. They also evaluate the inclusion of these features by performance, compared to linguistic and stylistic soft factors. In turn, Xia et al. (2019) complements demographic and financial information with narrative data and include soft information related to loan description, the character of borrowers, and variables obtained from a clustering procedure on soft information. Li et al. (2019), with advanced natural language processing (NLP) techniques, evaluates the risk of fraud associated with platforms in China. They use variables derived from management team members' working experience, educational background, and its composition. Text mining techniques are used on new information sources, and the explanation and prediction components in risk models are improved, benefiting management in the P2P market.

We can also find works that deal with internet and telecommunication information from users. Wang et al. (2019) uses information of borrower's online operation behaviour on P2P lending website, with variables derived from online operations of the users, such as registration records, login records, click records, browse records, authentication records, etc. Guo et al. (2016) focuses on the evaluation of Capacity, Character and Conditions, within the five Cs of credit risk management, excluding Capital and Collateral. They also include information obtained from the internet, heterogeneous social data, mainly demographic information derived from social networks, content generation, and the structure of each user's social network, through the demographic, tweet and network features components. Ge et al. (2017) and Zhang, Jia, et al. (2016) evaluate the role of social media information in the prediction of default. In the same line, Niu et al. (2019) and Xu et al. (2016) study the role of social network information. This last paper also evaluates information derived from herding manipulation to predict fraud. They find that herding behavior increases the likelihood that lenders will invest in listings that have already received bids from others. In addition, Ma, Zhao, et al. (2018) include patterns derived from mobile phone use (phone calls, text messages, and data traffic), demographics, mobility patterns, telecommunication patterns, App usage patterns and telecommunication records, to predict loan default.

Another trend consists of the use of network topology information, a field that has gained attention in recent years. Li, Hao, et al. (2018), Hadji-Misheva et al. (2018) and Giudici et al. (2020) include variables derived from the market connectivity network topology, for systematic risk assessment in the first paper, and credit risk in the latter two. Similarly, Ahelegbey et al. (2019), with inference derived from a latent factor model on financial indicators, distinguishes between connected and unconnected business communities. This distinction improves the predictive performance of scoring models in P2P lending for SME's.

Some other works study the impact of particular features of the P2P lending platforms. Zhu (2018) studies how the promise of security as guarantee support offered by the platform to back a loan, presents a level of negative relevance in the models, bringing problems of moral risk and adverse selection and affecting the probability of loan financing and default. They find that loans with security promises have a higher quoted amount, a lower interest rate, higher ratings, more investors, a higher funding success rate, but a higher default rate than unsecured loans. Xu and Chau (2018) examines the impact of communication between lenders and borrowers on financing outcomes and loan performance, collecting variables related to information disclosure, social influence, the quantity and quality of information exchanged, among other characteristics (see the Annex). Finally, Fu et al. (2019) characterises the sentiments of investors' comments, and from a time series of changes describing the dynamics of this variable, they predict an index of the daily trading volume of Chinese platforms. They take the comments from the first authorized P2P information platform in China and one of the largest portals in the P2P industry.

On the other hand, Gao et al. (2017) empirically validate the use of forwards-looking and backwards-looking mechanisms. The first mechanism with credit indicators, and the second with variables describe the experience of borrowers on the platform. These variables (Annex) are complemented with information on loans and the titles (e.g., length). On the other hand, Wei et al. (2018), trying to capture the non-linearity of characteristics, uses a cubic spline regression transformation, combining their proposal with a classification optimisation procedure through buffered AUC (bAUC). However, we consider that these performance-enhancing transformations must also be evaluated in the light of the explainability of the model. We include this article in the "*Inclusion of new variables"* segment due to the transformation proposed to test the variables differently from the traditional way.

3.2.4. Learning paradigm used

In Table 7, we segment the methodological paradigms that underlie the quantitative tools used, bearing in mind again, that a paper can use more than one paradigm depending on the objective of the research. Thus, it can appear in several categories.

Table 7. Paradigm used (93 papers, the 5 papers that did not use datasets were excluded)

Main paradigm used	Number	Authors/Year
Artificial Neural Networks	7	Byanjankar et al. (2015); Jin and Zhu (2015); Xu et al. (2015); Zang et al. (2015); Yuan et al. (2018); Zhang, Wang, Chen, Zhao, et al. (2017); Zhao (2015)
Deep Learning	11	Bastani et al. (2019); Duan (2019); Fu et al. (2019); Kim and Cho (2019a); Kim and Cho (2019b); Li, Ding, et al. (2018); Li et al. (2020); Nguyen Truong et al. (2019); Van-Sang et al. (2019); Wang et al. (2019); Zhang, Wang, Chen, Shang, et al. (2017)
Tree-based models	37	Bastani et al. (2019); Boiko Ferreira et al. (2017); Cai and Zhang (2020); Ding et al. (2017); Guo et al. (2016); Jin and Zhu, (2015); Jiang et al. (2019); A. Kim and Cho (2019); Kim and Cho (2019a); Kim and Cho (2019b); Kumar et al. (2016); Li, Ding, et al. (2018); Li et al. (2020); Ma, Zhao, et al. (2018); Malekipirbazari and Aksakalli (2015); Ma, Sha, et al. (2018); Hadji-Misheva et al. (2018); Namvar et al. (2018); Nguyen Truong et al. (2019); Niu et al. (2019); Rodrigues et al. (2018); Tan et al. (2017); Van-Sang et al. (2019); Wang et al. (2016); Wang et al. (2020); Wang, Kou, et al. (2018); Wang, Jiang, et al. (2018); Xia et al. (2017); Xia et al. (2018); Xia et al. (2019); Xia (2019); Xu et al. (2015); Xu et al. (2016); Ye et al. (2018); Zhang, Jia, et al. (2016); Zhou et al. (2019); Zhu et al. (2019)
Bayesian-probabilistic models	6	Boiko Ferreira et al. (2017); Guo et al. (2016); Jiang et al. (2018); Rodrigues et al. (2018); Wang et al. (2016); Wang, Kou, et al. (2018)
Fuzzy methods	2	Cho et al. (2019); Ji et al. (2020)
Qualitative models	2	Rosavina et al. (2019); Uddin et al. (2018)
Recommendation algorithms	4	Ren et al. (2019); Tan et al. (2017); Zhang, Wang, et al. (2016); Zhang, Geng, et al. (2017)
Statistics and econometrics models	47	Ahelegbey et al. (2019); Amalia et al. (2019); Boiko Ferreira et al. (2017); Byanjankar (2017); Cai and Zhang (2020); Calabrese et al. (2019); Canfield (2018); Chen et al. (2016); Chen (2017); Chen et al. (2019); Cho et al. (2019); Ding et al. (2017); Durovic (2017); Emekter et al. (2015); Gao et al. (2017); Ge et al. (2017); Gourieroux and Lu (2019); Greiner and Wang (2010); Herzenstein et al. (2011); Jiang et al. (2018); Jiang et al. (2019); Li et al. (2016); Li et al. (2020); Li, Ding, et al. (2018); Li, Hao, et al. (2018); Lin et al. (2017); Liu et al. (2018); Hadji-Misheva et al. (2018); Namvar et al. (2018); Pierrakis (2019); Serrano-Cinca et al. (2015); Serrano-Cinca and Gutierrez-Nieto (2016); Stofa (2017); Tan et al. (2017); Tao et al. (2017); Wan et al. (2019); Wang et al. (2016); Wei et al. (2018); Xia and Li (2016);

		Xinmin et al. (2019); Xu and Chau (2018); Xu and Zhang (2017); Yan et al. (2016); Yan et al. (2017); Yao et al. (2019); Zhou et al. (2018); Zhu (2018)
Support Vector Machine	25	Bastani et al. (2019); Cho et al. (2019); Ding et al. (2017); Duan (2019); Fu et al. (2019); Guo et al. (2016); Jiang et al. (2018); Jin and Zhu (2015); A. Kim and Cho (2019); Kim and Cho (2019a); Kim and Cho (2019b); Li et al. (2020); Malekipirbazari and Aksakalli (2015); Nguyen Truong et al. (2019); Rodrigues et al. (2018); Van-Sang et al. (2019); Wang, Kou, et al. (2018); Xia et al. (2018); Xia (2019); Xu et al. (2015); Xu et al. (2016); Yang et al. (2019); Ye et al. (2018); Zhou et al. (2019); Zhu et al. (2019)
Use ensemble models	22	Bastani et al. (2019); Boiko Ferreira et al. (2017); Cho et al. (2019); Jiang et al. (2019); Kim and Cho (2019a); Kim and Cho (2019b); Ma, Sha, et al. (2018); Ma, Zhao, et al. (2018); Malekipirbazari and Aksakalli (2015); Nguyen Truong et al. (2019); Niu et al. (2019); Rodrigues et al. (2018); Tan et al. (2017); Van-Sang et al. (2019); Wang, Jiang, et al. (2018); Wang et al. (2020); Xia et al. (2017); Xia et al. (2018); Xia et al. (2019); Xia (2019); Ye et al. (2018); Zhu et al. (2019)
Propose Heterogenous ensemble	6	Guo et al. (2016); Zhou et al. (2019); A. Kim & Cho (2019); Li, Ding, et al. (2018); Li, Ding, et al. (2018); Namvar et al. (2018)
Propose Homogenous ensemble	1	Ding et al. (2017)

Source: Compiled by the authors

The largest category is "*Statistics and econometrics models*" with 47 papers, almost all of the paper analysed. These papers typically use explanatory models to identify the type of relationship and measure the impact of the determinants of the dependent variable (default, fraud, investment, return, etc). We believe that it is the most widely used precisely because of its capacity for interpretation, the ease of comparison with business theory and knowledge, the ease of estimation and implementation, and the greater generalized understanding of researchers, users, and regulators.

The next category is "*Tree-based models*" with 37 papers. Decision trees have been used throughout the last decades as an alternative for managing credit risk, mainly because of their easy interpretation and application. However, in recent years their application has increased due to the advances in tree-based ensemble methods which obtain better classification performance tan standard decision trees, although sacrificing explanaibility. The main advances have come from homogeneous ensemble techniques, such as bagging and boosting; Random Forest (RF) and Extreme Gradient Boosting (XGBoost) being the most used (see the Annex). The developments of the boosting technique have also brought the application of algorithms such as Adaptive Boosting

(AdaBoost), Light Gradient Boosting Machine (LightGBM), among other alternatives. In the tree-based category, 22 papers use such tree-based ensemble methods.

In the category of "*Propose heterogeneous ensemble*", there are 6 papers that propose their own heterogeneous ensemble. As in homogenous ensemble techniques, their aim is to increase the learning ability of algorithms, reduce the variability of classifiers or predictive models, generalise performance, and guarantee scalability. They are also presented as an alternative to the problem of class imbalance. In other words, these methods are offered as an alternative to reduce the over-adjustment of individual options, to produce superior predictive results and a better capacity of generalization, which are visible in data sets with the class imbalance problem. We briefly describe some of these applications. Li, Ding, et al. (2018) and Li et al. (2020) make use of linear weighting method as a fusion strategy for individual classifiers, from models such as Extreme gradient boosting (XGBoost), Deep neural network (DNN) and logistics regression (LR). The default prediction of the ensemble model is better than that from the other alternatives. Namvar et al. (2018) presents an ensemble method based on Choquet fuzzy integral to improve the accuracy of multiple classifiers in predicting the credit risk of borrowers in P2P lending, specifically on the Adaptive Boosting (AdaBoost), Gradient Boosting (GB) and logistic regression (LR) methods. Zhou et al. (2019) uses Gradient boosting decision trees (GBDT), XGBoost and LightGBM as individual classifiers to create a heterogeneous ensemble learning-based default prediction model. Guo et al. (2016) proposes a two-tier learning ensemble with social data for credit scoring, making use of both stacking and boosting techniques for learning ensemble. Kim and Cho (2019) suggests a semi-supervised ensemble learning method through simple averaging by combining Label propagation and a transductive support vector machine (TSVM) with Dempster-Shafer probability theory.

Only one paper, Ding et al. (2017), "*propose a homogenous ensemble*". They propose a homogeneous ensemble based on clustering processes by weighting the distance function by the weight derived from the F-score for numerical attributes, and the Weight of Evidence (WOE) transformation and the Information Value (IV) for the nominal attributes. They also add an undersampling technique to equate clusters of positive instances with negative cases clusters. They generate the prediction and subsequent weighting, using as an underlying model to ensemble the logistic regression.

Another method widely used is "*Support Vector Machine (SVM)*", with 25 papers, although most papers present SVM as an alternative model for comparison (see Annex for more details of the models used in each paper).

In the case of "*Artificial Neural Networks (ANN)",* seven papers use them, mainly backpropagation algorithms. However, the more sophisticated "Deep Learning" methods are being more used (11). These accounts for the great development and applicability of these methods. The most popular techniques are Long Short-Term Memory (LSTM) techniques, wide deep learning (WDP), Deep Neural Network based on multilayer perceptron, Convolutional Neural Networks (CNN), among other options.

Finally, smaller categories include methods such as "*Bayesian-probabilistic models*", with (6) papers, mainly with the Naive Bayes (NB) algorithm, and "*Recommendation algorithms*" papers (4), focusing on the choice of profitable investment alternatives, as a tool for platforms and borrowing investors. "*Fuzzy-methods*" (2) and "*qualitative models*" (2) are less frequent. Some examples in the latter category are Rosavina et al. (2019), which based on semi-structured interviews, seeks to describe the factors which determine loan provision and Uddin et al. (2018), which proposes an expert-based risk rating to improve microcredit initiatives.

In summary, machine learning is being intensively used in credit risk and profit management, especially, ensemble methods, artificial neural networks and SVM. Interestingly, while these models usually have very good performance, they are black-box models difficult to interpret. Thus, the interpretability and explicability components are still a challenge and a necessity for machine learning proposals.

3.2.5. **Modeling elements**

We believe that modelling element is very important to ensure the replicability of the results, not only because replicable research affords greater confidence in the results, but also because it provides transparency and confidence in the methods. Furthermore, these requirements are especially crucial for machine learning alternatives, which are often more complex than statistical models. Transparency creates confidence and helps in the application, governance, and monitoring of the models and algorithms. Regulators and supervisors must have a standard and robust understanding of risk and profit management modeling proposals and their potential application to regulate and supervise effectively (ROFIEG, 2019). In this line, Table 8 presents some of the elements that we consider should be detailed and considered in the modeling process. We analyzed several modelling elements for 74 out of 98 papers (75% of the total) that propose classification models. It is worth mentioning that 49 publications include machine learning techniques.

Table 8. Modeling Elements

Component	Number	Authors/Year
Statistical comparison of the performance of classifiers	10	Cho et al. (2019); Jiang et al. (2019); Kim and Cho (2019a); Ma, Sha, et al., (2018); Niu et al. (2019); Wang, Jiang, et al. (2018); Wang et al. (2020); Xia et al. (2017); Xia et al. (2018); Xia (2019)
Explainability	38	Ahelegbey et al. (2019); Byanjankar et al. (2015); Byanjankar (2017); Calabrese et al. (2019); Canfield (2018); Chen (2017); Chen et al. (2019); Durovic (2017); Emekter et al. (2015); Gao et al. (2017); Ge et al. (2017); Giudici et al. (2020); Gourieroux and Lu (2019); Guo et al. (2016); Herzenstein et al. (2011); Jin and Zhu (2015); Jiang et al. (2018); Li, Ding, et al. (2018); Lin et al. (2017); Li et al. (2020); Liu et al. (2018); Ma, Sha, et al. (2018); Ma, Zhao, et al. (2018); Hadji-Misheva et al. (2018); Niu et al. (2019); Serrano-Cinca and Gutierrez-Nieto (2016); Serrano-Cinca et al. (2015); Stofa (2017); Tao et al. (2017); Uddin et al. (2018); Wan et al.

		(2019); Xinmin et al. (2019); Xu and Chau (2018); Xia et al. (2017); Yang et al. (2019); Yao et al. (2019); Zhu (2018); Zhang, Jia, et al. (2016)
Hyperparameters tuning	20	Boiko Ferreira et al. (2017); Cho et al. (2019); Jiang et al. (2019); Kim and Cho (2019a); Kim and Cho (2019b); Li, Ding, et al. (2018); Li et al. (2020); Ma, Zhao, et al. (2018); Malekipirbazari and Aksakalli (2015); Niu et al. (2019); Rodrigues et al. (2018); Wang et al. (2019); Wang et al. (2020); Xia et al. (2017); Xia et al. (2018); Xia et al. (2019); Xia (2019); Yang et al. (2019); Ye et al. (2018); Zhou et al. (2019)
Cross-validation	36	Bastani et al. (2019); Boiko Ferreira et al. (2017); Byanjankar (2017); Cai and Zhang (2020); Cho et al. (2019); Guo et al. (2016); Ding et al. (2017); Jiang et al. (2019); Jiang et al. (2018); Jin and Zhu (2015); Kim and Cho (2019a); Kim and Cho (2019b); Kumar et al. (2016); Li, Ding et al., (2018); Li et al. (2019); Li et al. (2020); Malekipirbazari and Aksakalli (2015), Namvar et al. (2018); Nguyen Truong et al. (2019); Niu et al. (2019); Rodrigues et al. (2018); Van-Sang et al. (2019); Wang et al. (2016); Wang, Jiang et al., (2018); Wang et al. (2020); Wei et al. (2018); Xia et al. (2019); Xia (2019); Xu et al. (2016); Yang et al. (2019); Ye et al. (2018); Zhang, Jia, et al. (2016); Zhang, Wang, Chen, Shang, et al. (2017); Zhang, Wang, Chen, Zhao, et al. (2017); Zhou et al. (2019); Zhu (2018)
Use of logistic regression	49	Ahelegbey et al. (2019); Boiko Ferreira et al. (2017); Byanjankar et al. (2015); Byanjankar (2017); Cai and Zhang (2020); Calabrese et al. (2019); Canfield (2018); Chen (2017); Chen et al (2019); Ding et al. (2017); Duan (2019); Jiang et al. (2019); Jiang et al. (2018); Ge et al. (2017); Giudici et al. (2020); Gourieroux and Lu (2019); Guo et al. (2016); Li, Ding, et al. (2018); Li et al. (2020); Lin et al. (2017); Liu et al. (2018); Ma, Zhao, et al. (2018); Malekipirbazari and Aksakalli (2015); Hadji-Misheva et al. (2018); Namvar et al. (2018); Nguyen Truong et al. (2019); Niu et al. (2019); Rodrigues et al. (2018); Wang et al. (2016); Serrano-Cinca and Gutierrez-Nieto (2016); Serrano-Cinca et al. (2015); Stofa (2017); Tan et al. (2017); Tao et al. (2017); Van-Sang et al. (2019); Wang, Kou, et al. (2018); Wei et al. (2018); Xinmin et al. (2019); Xu and Chau (2018); Xia et al. (2017); Xia et al. (2018); Xia et al. (2019); Xia (2019); Yao et al. (2019); Ye et al. (2018); Wang, Jiang, et al. (2018); Wang et al. (2020); Zhang, Jia, et al. (2016); Zhu et al. (2019)

Source: Compiled by the authors

49 papers "*Use logistic regression*" as the main model or as the benchmark methodology in the P2P lending market. However, only 10 papers use a statistical significance test to compare the performance of the competing alternatives. The tests used are paired-t, Friedman, DeLong or Shapiro-Wilk.

As it has been already mentioned, machine learning models typically present problems of explainability, making it difficult to understand for regulators and users. Although the modelling proposals present good performance indicators, entities, platforms, users and main regulators tend to prefer explanatory models that allow them to understand the logic of decisions. This fact avoids the stakeholders to take advantage of these alternatives. However, only 38 papers consider the "*Explainability*" component, and most of them do it through traditional models coming from statistical or econometric methodologies, through coefficient estimations and inferential tests. The few papers that include explanatory elements in machine learning only study feature importance, ignoring aspects such as monotonicity, non-linearity of relationships, structural changes, which could be taken into account through approaches such as those offered in recent years with techniques of local and global explicability and interpretability.

Regarding, model adjustment, only 20 papers mention "*Hyperparameter tuning*" processes. Machine learning methods have different parameters that control the learning process and it is essential to ensure their performance and stability. Among the most used ones are searching on grids, Bayesian optimization methods and genetic algorithms. Related to this is the use of "C*ross-validation*", a method for primarily assessing the predictive performance of a model, evaluating how the results can be generalized to an independent data set, and pointing out problems of overfitting. Cross-validation technique is used in 36 papers, mainly with 5 or 10 folds, on training and validation sets, some of these by replicating the process several numbers of times.

4. Discussions and conclusions

P2P loans and other new financial products have been growing hand in hand with further technological developments, and risk and profit management have also become more sophisticated. Although this has brought business alternatives as intermediaries for the platforms and has opportunities for investment and financing for lenders and borrowers, it has also generated difficulties in regulation and supervision. Governing bodies must adapt quickly to new market demands for risk and profit management to ensure participants' financial health and the sustainability of the market. For this purpose, they must cover not only financial elements but also technological and methodological ones, where machine learning and data science tools are strategical if used efficiently.

Current technologies, in particular machine learning, offer the potential for broader, deeper, and faster analysis of large data sets, including unique and diverse sources of information that may be relevant for risk assessment and profitability management of companies and individuals, much more so in a developing market such as P2P lending with new financing and investment alternatives (ROFIEG, 2019). The development of machine learning and new data science technologies helps provide more suitable and adapted products. For lenders, these solutions can help protect against credit risk and fraud, reduce the cost of the credit evaluation process, encourage product development, distribution, and monitoring, and enable better consumer-customer interaction (ROFIEG, 2019).

Therefore, given the importance of this emerging market, P2P lending, their rapid expansion, and the need for efficient risk and profit assessment, we considered it necessary to have carried out a bibliometric and systematic review of the papers associated mainly with models to risk and profit management in the P2P lending market in the last decade. A study that contributes to elements that help protect the participants and guarantee the market's continuity and development as a financing and investment alternative. It becomes relevant to identify the research focuses and the aspects of the business that have been mostly treated and supported with different modeling techniques, particularly with machine learning and artificial intelligence. With this, we have also identified research and development opportunities that contribute to the P2P lending market's healthy development. In turn, aspects demanded by regulators and users on the methodological use of artificial intelligence and machine learning in risk and profit management are described. We draw attention to some limitations that we have found in many of the studies.

Related to RQ1, we have found among the most relevant works by citation are Greiner and Wang (2010), Herzenstein et al. (2011), Emekter et al. (2015), Serrano-Cinca et al. (2015), Malekipirbazari and Aksakalli (2015) , Serrano-Cinca and Gutierrez-Nieto (2016) and Xia et al. (2017), being Emekter et al. (2015), the most cited publication. These papers are mostly devoted to identifying the key factors in the credit risk models in P2P lending. They use mainly statistical and econometric models. And we also highlight papers such as those by Serrano-Cinca and Gutierrez-Nieto (2016) and Xia et al. (2017) that have recently had a great impact and present proposals for risk management models with a business focus, which we believe make them very suitable for the industry.

In turn, we identify that China and the United States are the countries with the most outstanding experience in the P2P market. However, the eastern market shows deterioration and reduction of the market due to fraud and regulatory issues. However, the most significant number of publications and research are linked to eastern countries. The most frequent cooperation between researchers is between the two countries mentioned. Consequently, the countries with the largest market have the most significant number of publications or researchers participating in the publications, sensing an association between market development and research. However, we cannot establish a causal relationship.

Concerning to RQ2, we emphasize that the most used data are those of the Lending Club, significantly associated with the availability of the information, a comprehensive dataset with various variables of both the application and the behavior of the obligations, in addition to a complete description of these. We can see a strong tendency to use data from China that may be caused due to the growth of the market and the concern of a broad group of researchers seeking to answer regulatory problems. It highlights that most of the studies concentrate on models for managing loans to individuals. By country, data use is closely related to the development of P2P markets globally, mostly from China and the United States. With few datasets from Europe, including the United Kingdom, Mexico, and Southeast Asia.

Related to the software used is mainly free open source such as Python and R. Because of the opportunity, they generate modeling alternatives and use developments from other researchers, some at the edge of knowledge, with few restrictions.

At the business level, we have identified that many publications focus on credit risk classification models, bringing new algorithms and models derived from the advances of machine learning and artificial intelligence. These models typically improve the performance of traditional alternatives, which eventually favour investors, platforms, and consumers. Exist significant proportion of the publications have tended to use statistical and econometric models because of their capacity for interpretation, the ease of comparison with business theory and knowledge, the ease of estimation and implementation, and the greater generalized understanding of researchers, users, and regulators. It is worth mentioning that a considerable percentage of the works that propose classification models for credit risk management use logistic regression as their benchmark model for performance comparison or as their base model. Nevertheless, there is a tendency to use machine learning alternatives to take advantage of its predictive power. We highlight the statistical and explainable methods' duality and note that few explainable machine learning applications, even though methodologically its progress has been substantial.

In recent papers, we identified outstanding developments with machine learning, such as deep learning, the new alternatives of the ensemble in decision trees (such as random forest and gradient boosting algorithms), and heterogeneous ensembles. We also found that it has been proposed to complement traditional information with variables from various information sources, soft information, and natural language processing (NLP) methods, both for credit risk and investment models and profits. Specifically, many papers explore the use of new input variables derived from new information sources such as social networks (connections, consultations, photographs, messages, and videos), connectivity networks, the use of telecommunication devices, georeferencing, the textual analysis of the documents that accompany the applications, etc. Such variables help to improve the performance indicators and evaluation of the proposed models. Text mining techniques have also been used on new information sources, and the explanation and prediction components in risk models have improved in some cases, benefiting development in the P2P market.

Regarding RQ3, it is interesting to note that very few studies focus on P2P loans for businesses and companies, most of which have focused on loans for individuals. On the other hand, we consider that some aspects are understudied such as fraud, debt collection, credit provision -LGD, third party-collateral warranties, etc. All of them are necessary for the healthy development of the market. Surprisingly, we have found difficulties in the specification of the problem in many articles. For example, few papers clarify whether it is for granting or behavior when proposing a scoring model. This lack prevents the use of these models by regulators and end-users. We have also found problems in the default definition; even in some papers, there is no mention of how this event is defined, aspects that could cause an underestimation of the risk or an overestimation of the profit. Perhaps this could be due to the lack of a theory about business failure.

In turn, there should be an adequate selection of the independent variables according to the model's purpose. For example, for granting loans, it makes no sense to use decision variables such as interest rate or the level of risk estimated by the companies since they are closely linked to the event in analysis, generating biases in the results and limiting their use. The variables included must consider the operative of the business. These facts lead to the limitation of the application due to lack of uniformity and consistency, which are objectives sought by the regulatory bodies, especially for transnational growing markets such as P2P lending. From the joint work between academics, those who propose models for risk management, regulators, and the platforms, tailored solutions can be presented, which generate greater confidence and probability of application. This synergy would ensure a better understanding of the business's nature and structure, which is reflected in better modeling, risk management, and service, enhancing the sustainability and development of the P2P lending market.

On the appropriate choice of variables for the models, a feature selection process is still an object of research and development today. We emphasize that with a more precise business sense in the P2P loan have found proposals for modeling risk management through cost-sensitive techniques, useful for several years in credit risk and other areas. These techniques mitigate class imbalance problems and weigh the classification errors with more significant intuition and business sense, allowing optimization and learning to recognize these elements, providing useful results for platforms and P2P investors market risk managers. From our perspective, in P2P lending, cost-sensitive models are a research line that deserves to be further deepened.

Given the extended use of machine learning methods and the great need for model explainability and transparency required by regulators, supervisors, financial institutions, and governments, we believe that more research is needed. Most machine learning proposals typically focus on performance and neglect interpretability. The models with explainable purposes are still supported in traditional methodologies, econometrics, and statistics tools, being the logistic regression the most used one. The papers that use machine learning models rarely include this aspect, and if they do, they only include an analysis of feature importance. Thus, the interpretability components are still a challenge and a necessity for machine learning proposals. Interpretability is a necessary component to generate confidence and transparency in the use of machine learning models. The lack of transparency creates specific research challenges for supervisors regarding why and how digital models and algorithms work. As future research lines, there is a need to generate frameworks that ensure the ethical application of new solutions, particularly in credit, whose central elements rely on the explainability of decisions (ROFIEG, 2019). In P2P lending in the last year, some advances have been seen in the line of explicability of machine learning models for credit risk, as shown in Ariza-Garzón et al. (2020) and Bussmann et al. (2020). At a broader level, there is great interest in reflecting on the component of explainability in modeling and in offering alternative explanations for new technologies, as shown in Carvalho et al. (2019), Rudin (2019), and Barredo Arrieta et al. (2020).

Finally, we highlight the works that help to understand the market, the deficiencies in regulation from an economic, political and financial component, many of them qualitative, with theoretical and contextual elements that seek to identify the shortcomings and opportunities, to guarantee their development, continuity, and stability as an alternative for investment, inclusion, and financing of financial market agents, a transcendental focus that should be further deepened.

Regarding the limitations of this study, we point out that although the review of articles in the main journals has been extensive, the design established could have made some studies on the subject in other publications to be overlooked. For example, due to the limitation of the database selected. Assertions on findings are, therefore, only based on the journals analyzed. Therefore, further investigation will have to be conducted with new empirical studies to fill in the gaps and weaknesses discovered in this literature review. More articles focusing on risk management, fraud, ethical issues and the P2P loan markets. In sum, P2P lending is a flourishing market, and its research is also buoyant and active. Still, it has essential shortcomings and opportunities that must be addressed and studied to ensure its service, development, and sustainability. We hope that this study will serve to draw attention to some of these.

References

Ahelegbey, D. F., Giudici, P., and Hadji-Misheva, B. (2019). Latent factor models for credit scoring in P2P systems. *Physica A-Statistical Mechanics and Its Applications*, *522*, 112–121. https://doi.org/10.1016/j.physa.2019.01.130

Amalia, N., Dalimunthe, Z., and Triono, R. A. (2019). The Effect of Lender's Protection on Online Peer-to-Peer Lending in Indonesia. In K. S. Soliman (Ed.), *Education Excellence and Innovation Management through Vision 2020*.

Aria, M. and Cuccurullo, C. (2017). Bibliometrix: An R-tool for comprehensive science mapping analysis. *Journal of Informetrics. Elsevier.*, *11(4)*, 959–975.

Ariza-Garzón, M. J., Arroyo, J., Caparrini, A., and Segovia-Vargas, M. (2020). Explainability of a Machine Learning Granting Scoring Model in Peer-to-Peer Lending. *IEEE Access*, *8*, 64873–64890. https://doi.org/10.1109/ACCESS.2020.2984412

Bachmann, A., Becker, A., Buerckner, D., Hilker, M., Kock, F., Lehmann, M., Tiburtius, P., and Funk, B. (2011). Online peer-to-peer lending - A literature review. In *Journal of Internet Banking and Commerce*.

Bae, J. K. (2018). A Study on the Determinant Factors of P2P Loans and Activation Factors of P2P Lending Market - P2p. *Logos Management Review*, *16*(2), 21–36.

Barredo Arrieta, A., Díaz-Rodríguez, N., Del Ser, J., Bennetot, A., Tabik, S., Barbado, A., Garcia, S., Gil-Lopez, S., Molina, D., Benjamins, R., Chatila, R., & Herrera, F. (2020). Explainable Artificial Intelligence (XAI): Concepts, taxonomies, opportunities and challenges toward responsible AI. *Information Fusion*, *58*, 82–115. https://doi.org/https://doi.org/10.1016/j.inffus.2019.12.012

Bastani, K., Asgari, E., and Namavari, H. (2019). Wide and deep learning for peer-to-peer lending. *Expert Systems with Applications*, *134*, 209–224. https://doi.org/10.1016/j.eswa.2019.05.042

Birkle, C., Pendlebury, D. A., Schnell, J., and Adams, J. (2020). Web of Science as a data source for research on scientific and scholarly activity. *Quantitative Science Studies*, *1*(1), 363–376. https://doi.org/10.1162/qss_a_00018

Boiko Ferreira, L. E., Barddal, J. P., Enembreck, F., and Gomes, H. M. (2017). Improving Credit Risk Prediction in Online Peer-to-Peer {(P2P)} Lending Using Imbalanced Learning Techniques. In *2017 IEEE 29th International Conference on Tools with Artificial Intelligence* (pp. 175–181). https://doi.org/10.1109/ictai.2017.00037

Bussmann, N., Giudici, P., Marinelli, D., and Papenbrock, J. (2020). Explainable AI in Credit Risk

Management. *Frontiers in Artifical Intelligence. Artifical Intelligence in Finance*. https://doi.org/10.3389/frai.2020.00026

Byanjankar, A. (2017). Predicting Credit Risk in Peer-to-Peer Lending with Survival Analysis. In *2017 Ieee Symposium Series on Computational Intelligence*.

Byanjankar, A., Heikkila, M., and Mezei, J. (2015). Predicting Credit Risk in Peer-to-Peer Lending: A Neural Network Approach. In *2015 Ieee Symposium Series on Computational Intelligence*. https://doi.org/10.1109/ssci.2015.109

Cai, S., and Zhang, J. (2020). Exploration of credit risk of P2P platform based on data mining technology. *Journal of Computational and Applied Mathematics*, *372*. https://doi.org/10.1016/j.cam.2020.112718

Calabrese, R., Osmetti, S. A., and Zanin, L. (2019). A joint scoring model for peer-to-peer and traditional lending: a bivariate model with copula dependence. *Journal of the Royal Statistical Society Series A-Statistics in Society*, *182*(4), 1163–1188. https://doi.org/10.1111/rssa.12523

Canfield R, C. E. (2018). Determinants of Default in P2P Lending: The Mexican Case. *Independent Journal of Management and Production*, *9*(1), 1–24. https://doi.org/10.14807/ijmp.v9i1.537

Carvalho, D. V, Pereira, E. M., & Cardoso, J. S. (2019). Machine Learning Interpretability: A Survey on Methods and Metrics. *Electronics*, *8*(8). https://doi.org/10.3390/electronics8080832

Chen, C., Dong, M. C., Liu, N., and Sriboonchitta, S. (2019). Inferences of default risk and borrower characteristics on P2P lending. *North American Journal of Economics and Finance*, *50*. https://doi.org/10.1016/j.najef.2019.101013

Chen, D., Li, X., and Lai, F. (2016). Gender discrimination in online peer-to-peer credit lending: evidence from a lending platform in China. *Electronic Commerce Research*, *17*(4), 553–583. https://doi.org/10.1007/s10660-016-9247-2

Chen, Y. (2017). Research on the Credit Risk Assessment of Chinese Online Peer-to-peer Lending Borrower on Logistic Regression Model. In *3rd Asian Pacific Conference on Energy, Environment and Sustainable Development* (pp. 216–221).

Cho, P., Chang, W., and Song, J. W. (2019). Application of Instance-Based Entropy Fuzzy Support Vector Machine in Peer-To-Peer Lending Investment Decision. *IEEE Access*, *7*, 16925–16939. https://doi.org/10.1109/access.2019.2896474

Claessens, S., Frost, J., Turner, G., and Zhu, F. (2018). Fintech credit markets around the world: size, drivers and policy issues. *BIS Quarterly Review*. https://www.bis.org/publ/qtrpdf/r_qt1809e.pdf

Cummins, M., Lynn, T., Mac an Bhaird, C., and Rosati, P. (2019). Addressing Information Asymmetries in Online Peer-to-Peer Lending. In *Disrupting Finance*. https://doi.org/10.1007/978-3-030-02330-0_2

Ding, H., Zhang, P., Lu, T., Gu, H., and Gu, N. (2017). Credit Scoring Using Ensemble Classification Based on Variable Weighting Clustering. In W. Shen, P. Antunes, N. H. Thuan, J. P. Barthes, J. Luo, and J. Yong (Eds.), *2017 Ieee 21st International Conference on Computer Supported Cooperative Work in Design*.

Duan, J. (2019). Financial system modeling using deep neural networks (DNNs) for effective risk assessment and prediction. *Journal of the Franklin Institute-Engineering and Applied Mathematics*, *356*(8), 4716–4731. https://doi.org/10.1016/j.jfranklin.2019.01.046

Durovic, A. (2017). Estimating Probability of Default on Peer to Peer Market - Survival Analysis Approach. *Journal of Central Banking Theory and Practice*, *6*(2), 149–167. https://doi.org/10.1515/jcbtp-2017-0017

Elango, B., and Rajendran, P. (2012). Authorship trends and collaboration pattern in the marine sciences literature: a scientometric study. *International Journal of Information Dissemination and Technology*.

Emekter, R., Tu, Y., Jirasakuldech, B., and Lu, M. (2015). Evaluating credit risk and loan performance in online Peer-to-Peer (P2P) lending. *Applied Economics*, *47*(1), 54–70. https://doi.org/10.1080/00036846.2014.962222

Financial Stability Board. (2017). Artificial intelligence and machine learning in financial services. Market developments and financial stability implications. In *FSB. Financial Stability Board*. internal-pdf://0180047707/P011117.pdf

Fu, X., Zhang, S., Chen, J., Ouyang, T., and Wu, J. (2019). A Sentiment-Aware Trading Volume Prediction Model for P2P Market Using LSTM. *Ieee Access*, *7*, 81934–81944. https://doi.org/10.1109/access.2019.2923637

Gao, Y., Sun, J., and Zhou, Q. (2017). Forward looking vs backward looking An empirical study on the effectiveness of credit evaluation system in China’s online P2P lending market. *China Finance Review International*, *7*(2), 228–248. https://doi.org/10.1108/cfri-07-2016-0089

Gao, Y., Yu, S., Chen, M., and Shiue, Y. (2020). A 2020 perspective on “The performance of the P2P finance industry in China.” *Electronic Commerce Research and Applications*. https://doi.org/10.1016/j.elerap.2020.100940

Ge, R., Feng, J., Gu, B., and Zhang, P. (2017). Predicting and Deterring Default with Social Media Information in Peer-to-Peer Lending. *Journal of Management Information Systems*, *34*(2), 401–424. https://doi.org/10.1080/07421222.2017.1334472

Giudici, P., Hadji-Misheva, B., and Spelta, A. (2020). Network based credit risk models. *Quality Engineering*. https://doi.org/10.1080/08982112.2019.1655159

Gourieroux, C., and Lu, Y. (2019). Least impulse response estimator for stress test exercises. *Journal of Banking and Finance*, *103*, 62–77. https://doi.org/10.1016/j.jbankfin.2019.03.021

Greiner, M. E., and Wang, H. (2010). Building Consumer-to-Consumer Trust in E-Finance Marketplaces: An Empirical Analysis. *International Journal of Electronic Commerce*, *15*(2), 105–136. https://doi.org/10.2753/jec1086-4415150204

Guo, G., Zhu, F., Chen, E., Liu, Q., Wu, L., and Guan, C. (2016). From Footprint to Evidence: An Exploratory Study of Mining Social Data for Credit Scoring. *Acm Transactions on the Web*, *10*(4). https://doi.org/10.1145/2996465

Hadji-Misheva, B. H., Giudici, P., Pediroda, V., and Ieee. (2018). Network-based models to improve credit scoring accuracy. In *2018 Ieee 5th International Conference on Data Science and*

Advanced Analytics (pp. 623–630). https://doi.org/10.1109/dsaa.2018.00080

Herzenstein, M., Sonenshein, S., and Dholakia, U. M. (2011). Tell Me a Good Story and I May Lend You Money: The Role of Narratives in Peer-to-Peer Lending Decisions. *Journal of Marketing Research*, *48*, S138–S149. https://doi.org/10.1509/jmkr.48.SPL.S138

Hodge, D. R., and Lacasse, J. R. (2011). Ranking disciplinary journals with the Google Scholar h-index: A new tool for constructing cases for tenure, promotion, and other professional decisions. *Journal of Social Work Education*, *47*(3), 579–596. https://doi.org/10.5175/jswe.2011.201000024

Ji, X., Yu, L., and Fu, J. (2020). Evaluating Personal Default Risk in P2P Lending Platform: Based on Dual Hesitant Pythagorean Fuzzy TODIM Approach. *Mathematics*, *8*(1). https://doi.org/10.3390/math8010008

Jiang, C. Q., Wang, Z., Wang, R. Y., and Ding, Y. (2018). Loan default prediction by combining soft information extracted from descriptive text in online peer-to-peer lending. *Annals of Operations Research*, *266*(1–2), 511–529. https://doi.org/10.1007/s10479-017-2668-z

Jiang, C., Wang, Z., and Zhao, H. (2019). A prediction-driven mixture cure model and its application in credit scoring. *European Journal of Operational Research*, *277*(1), 20–31. https://doi.org/10.1016/j.ejor.2019.01.072

Jin, Y., and Zhu, Y. (2015). A data-driven approach to predict default risk of loan for online Peer-to-Peer {(P2P)} lending. In G. Tomar (Ed.), *2015 Fifth International Conference on Communication Systems and Network Technologies* (pp. 609–613). https://doi.org/10.1109/csnt.2015.25

Kim, A., and Cho, S.-B. (2019). An ensemble semi-supervised learning method for predicting defaults in social lending. *Engineering Applications of Artificial Intelligence*, *81*, 193–199. https://doi.org/10.1016/j.engappai.2019.02.014

Kim, J.-Y., and Cho, S.-B. (2019a). Predicting repayment of borrows in peer-to-peer social lending with deep dense convolutional network. *Expert Systems*, *36*(4). https://doi.org/10.1111/exsy.12403

Kim, J.-Y., and Cho, S.-B. (2019b). Towards Repayment Prediction in Peer-to-Peer Social Lending Using Deep Learning. *Mathematics*, *7*(11). https://doi.org/10.3390/math7111041

Koseoglu, M. A. (2016). Growth and structure of authorship and co-authorship network in the strategic management realm: Evidence from the Strategic Management Journal. *BRQ Business Research Quarterly*, *19 (3)*, 153–170. https://doi.org/10.1016/j.brq.2016.02.001

Kumar, V. L., Natarajan, S., Keerthana, S., Chinmayi, K. M., Lakshmi, N., and Ieee. (2016). Credit Risk Analysis in Peer-to-Peer Lending System. In *2016 Ieee International Conference on Knowledge Engineering and Applications*.

Lee, Y.-W., Chen, S., Yu, T., and Ieee. (2017). Analysis of the Impact of Collateral on Peer-to-Peer Lending. In *2017 Ieee/Sice International Symposium on System Integration* (pp. 77–82).

Li, H., Zhang, Y., Zhang, N., and Jia, H. (2016). Detecting the abnormal lenders from P2P lending data. In H. Lee, Y. Shi, J. Lee, F. Cordova, I. Dzitac, G. Kou, and J. Li (Eds.), *Promoting Business*

Analytics and Quantitative Management of Technology: 4th International Conference on Information Technology and Quantitative Management (Vol. 91, pp. 357–361). https://doi.org/10.1016/j.procs.2016.07.095

Li, L., Feng, Y., Lv, Y., Cong, X., Fu, X., and Qi, J. (2019). Automatically Detecting Peer-to-Peer Lending Intermediary Risk-Top Management Team Profile Textual Features Perspective. *Ieee Access*, *7*, 72551–72560. https://doi.org/10.1109/access.2019.2919727

Li, W., Ding, S., Chen, Y., and Yang, S. (2018). Heterogeneous Ensemble for Default Prediction of Peer-to-Peer Lending in China. *Ieee Access*, *6*, 54396–54406. https://doi.org/10.1109/access.2018.2810864

Li, W., Ding, S., Wang, H., Chen, Y., and Yang, S. (2020). Heterogeneous ensemble learning with feature engineering for default prediction in peer-to-peer lending in China. *World Wide Web-Internet and Web Information Systems*, *23*(1), 23–45. https://doi.org/10.1007/s11280-019-00676-y

Li, Y., Hao, A., Zhang, X., and Xiong, X. (2018). Network topology and systemic risk in Peer-to-Peer lending market. *Physica A-Statistical Mechanics and Its Applications*, *508*, 118–130. https://doi.org/10.1016/j.physa.2018.05.083

Lin, X. C., Li, X. L., and Zheng, Z. (2017). Evaluating borrower's default risk in peer-to-peer lending: evidence from a lending platform in China. *Applied Economics*, *49*(35), 3538–3545. https://doi.org/10.1080/00036846.2016.1262526

Liu, C., and Yan, J. (2016). Researches on Risks and Precautions of Chinese P2P Lending. In M. Kuek and R. Zhao (Eds.), *Proceedings of the 23rd International Business Annual Conference*.

Liu, H., Zhou, S., and Yang, W. (2019). Research on Intelligent Inter net Financial Investment Model. In R. Su (Ed.), *2019 International Conference on Image and Video Processing, and Artificial Intelligence* (Vol. 11321). https://doi.org/10.1117/12.2539006

Liu, Y., Zhou, Q., Zhao, X., and Wang, Y. (2018). Can Listing Information Indicate Borrower Credit Risk in Online Peer-to-Peer Lending? *Emerging Markets Finance and Trade*, *54*(13), 2982–2994. https://doi.org/10.1080/1540496x.2018.1427061

Ma, L., Zhao, X., Zhou, Z., and Liu, Y. (2018). A new aspect on P2P online lending default prediction using meta-level phone usage data in China. *Decision Support Systems*, *111*, 60–71. https://doi.org/10.1016/j.dss.2018.05.001

Ma, X., Sha, J., Wang, D., Yu, Y., Yang, Q., and Niu, X. (2018). Study on a prediction of P2P network loan default based on the machine learning LightGBM and XGboost algorithms according to different high dimensional data cleaning. *Electronic Commerce Research and Applications*, *31*, 24–39. https://doi.org/10.1016/j.elerap.2018.08.002

Malekipirbazari, M., and Aksakalli, V. (2015). Risk assessment in social lending via random forests. *Expert Systems with Applications*, *42*(10), 4621–4631. https://doi.org/10.1016/j.eswa.2015.02.001

Milne, A., and Parboteeah, P. (2016). The Business Models and Economics of Peer-to-Peer Lending. *Centre for European Policy Studies*, *17*, 36. internal-pdf://131.209.49.252/ECRI RR17 P2P Lending.pdf NV - Technical Report

Namvar, A., Naderpour, M., and Ieee. (2018). Handling Uncertainty in Social Lending Credit Risk Prediction with a Choquet Fuzzy Integral Model. In *2018 Ieee International Conference on Fuzzy Systems*.

Nguyen Truong, T., Khuat Thanh, S., Ngo Thi Thu, T., Nguyen Ha, N., and Tran Manh, D. (2019). Improve Risk Prediction in Online Lending (P2P) Using Feature Selection and Deep Learning. *International Journal of Computer Science and Network Security*, *19*(11), 216–222.

Niu, B., Ren, J., and Li, X. (2019). Credit Scoring Using Machine Learning by Combing Social Network Information: Evidence from Peer-to-Peer Lending. *Information*, *10*(12). https://doi.org/10.3390/info10120397

Park, S., and Daeseon, C. (2019). A Study on P2P Lending Deadline Prediction Model based on Machine Learning. *Journal of KIISE*, *46*(2), 174–183.

Pierrakis, Y. (2019). Peer-to-peer lending to businesses: Investors' characteristics, investment criteria and motivation. *International Journal of Entrepreneurship and Innovation*, *20*(4), 239–251. https://doi.org/10.1177/1465750319842528

Pokorna, M., and Sponer, M. (2016). Social lending and its risks. In S. Kapounek and V. Krutilova (Eds.), *19th International Conference Enterprise and Competitive Environment 2016* (Vol. 220, pp. 330–337). https://doi.org/10.1016/j.sbspro.2016.05.506

Pur, S., Huesig, S., Mann, H.-G., and Schmidhammer, C. (2014). How to Analyze the Disruptive Potential of Business Model Innovation in Two-Sided Markets?: The Case of Peer to Peer Lending Marketplaces in Germany. In D. F. Kocaoglu, T. R. Anderson, T. U. Daim, D. C. Kozanoglu, K. Niwa, and G. Perman (Eds.), *2014 Portland International Conference on Management of Engineering and Technology* (pp. 693–709).

Ren, K., Malik, A., and Acm. (2019). Investment Recommendation System for Low-Liquidity Online Peer to Peer Lending (P2PL) Marketplaces. In *Proceedings of the Twelfth Acm International Conference on Web Search and Data Mining*. https://doi.org/10.1145/3289600.3290959

Rodrigues, D. S., Brasil, A. R. A., Costa, M. B., Komati, K. S., Pinto, L. A., and Acm. (2018). A Comparative Analysis of Loan Requests Classification Algorithms in a Peer-to-Peer Lending Platform. In *Proceedings of the 14th Brazilian Symposium on Information Systems*. https://doi.org/10.1145/3229345.3229390

ROFIEG, Expert Group on Regulatory Obstacles to Financial Innovation .(2019). *Thirty recommendations on regulation, innovation and finance* (Issue December). Final Report to the European Commission. https://ec.europa.eu/info/files/191113-report-expert-group-regulatory-obstacles-financial-innovation_en

Rosavina, M., Rahadi, R. A., Kitri, M. L., Nuraeni, S., and Mayangsari, L. (2019). P2P lending adoption by SMEs in Indonesia. *Qualitative Research in Financial Markets*, *11*(2), 260–279. https://doi.org/10.1108/qrfm-09-2018-0103

Rudin, C. (2019). Stop explaining black box machine learning models for high stakes decisions and use interpretable models instead. *Nature Machine Intelligence*, *1*(5), 206–215. https://doi.org/10.1038/s42256-019-0048-x

Serrano-Cinca, C., and Gutierrez-Nieto, B. (2016). The use of profit scoring as an alternative to credit scoring systems in peer-to-peer (P2P) lending. *Decision Support Systems*, *89*, 113–122. https://doi.org/10.1016/j.dss.2016.06.014

Serrano-Cinca, C., Gutierrez-Nieto, B., and Lopez-Palacios, L. (2015). Determinants of Default in P2P Lending. *Plos One*, *10*(10). https://doi.org/10.1371/journal.pone.0139427

Soo H., A. (2016). FinTech supporting Government's Policy, its Implementing Measures and Legal Institution in UK- focused on the Payment Service Industry. *Kangwon Law Review*, *49*, 179–219. https://doi.org/10.18215/kwlr.2016.49..179

Stofa, T. (2017). Analysis of Repayment Failures in P2P Lending. In B. Gavurova and M. Soltes (Eds.), *Central European Conference in Finance and Economics*.

Sungbok, L. (2018). Study on the Financial Intermediary Role of P2P Lending Platform - P2p. *Journal of Money and Finance*, *32*(2), 21–62. https://doi.org/10.21023/jmf.32.2.2

Tan, Y., Zheng, X., Zhu, M., Wang, C., Zhu, Z., and Yu, L. (2017). Investment Recommendation with Total Capital Value Maximization in Online P2P Lending. In O. Hussain, L. Jiang, X. Fei, C. W. Lan, and K. M. Chao (Eds.), *2017 Ieee 14th International Conference on E-Business Engineering*. https://doi.org/10.1109/icebe.2017.32

Tao, Q., Dong, Y., and Lin, Z. (2017). Who can get money? Evidence from the Chinese peer-to-peer lending platform. *Information Systems Frontiers*, *19*(3), 425–441. https://doi.org/10.1007/s10796-017-9751-5

Uddin, M. J., Vizzari, G., Bandini, S., and Imam, M. O. (2018). A case-based reasoning approach to rate microcredit borrower risk in online Kiva P2P lending model. *Data Technologies and Applications*, *52*(1), 58–83. https://doi.org/10.1108/dta-02-2017-0009

Van-Sang, H., Dang-Nhac, L., Choi, G. S., Ha-Nam, N., and Yoon, B. (2019). Improving Credit Risk Prediction in Online Peer-to-Peer {(P2P)} Lending Using Feature selection with Deep learning. *2019 21st International Conference on Advanced Communication Technology*, *6*(1), 20–31. https://doi.org/10.23919/ICACT.2019.8701943

van Eck, N. J., and Waltman, L. (2010). Software survey: VOSviewer, a computer program for bibliometric mapping. *Scientometrics*, *84 (2)*, 523–538. https://doi.org/10.1007/s11192-009-0146-3

van Eck, N. J., and Waltman, L. (2014). Visualizing Bibliometric Networks. In Y. Ding, R. Rousseau, and D. Wolfram (Eds.), *Measuring Scholarly Impact: Methods and Practice* (pp. 285–320). Springer International Publishing. https://doi.org/10.1007/978-3-319-10377-8_13

van Eck, N. J., & Waltman, L. (2017). Citation-based clustering of publications using CitNetExplorer and VOSviewer. *Scientometrics*, *111*, 1053–1070. https://doi.org/10.1007/s11192-017-2300-7

Wan, J., Zhang, H., Zhu, X., Sun, X., and Li, G. (2019). Research on Influencing Factors of P2P Network Loan Prepayment Risk Based on Cox Proportional Hazards. In E. HerreraViedma, Y. Shi, D. Berg, J. Tien, F. J. Cabrerizo, and J. Li (Eds.), *7th International Conference on Information Technology and Quantitative Management* (Vol. 162, pp. 842–848). https://doi.org/10.1016/j.procs.2019.12.058

Wang, C., Han, D., Liu, Q., and Luo, S. (2019). A Deep Learning Approach for Credit Scoring of Peer-to-Peer Lending Using Attention Mechanism LSTM. *IEEE Access*, *7*, 2161–2168. https://doi.org/10.1109/access.2018.2887138

Wang, H., Kou, G., and Peng, Y. (2018). Cost-sensitive Classifiers in Credit Rating A Comparative Study on P2P Lending. In I. Dzitac, F. G. Filip, M. J. Manolescu, S. Dzitac, H. Oros, and D. Dzitac (Eds.), *2018 7th International Conference on Computers Communications and Control*.

Wang, L. (2018). Supervision of Peer-to-Peer Lending in China. In J. Liu and K. L. Teves (Eds.), *Proceedings of the 2018 2nd International Conference on Education, Economics and Management Research* (Vol. 182, pp. 291–293).

Wang, S., Qi, Y., Fu, B., and Liu, H. (2016). Credit Risk Evaluation Based on Text Analysis. *International Journal of Cognitive Informatics and Natural Intelligence*, *10*(1), 1–11. https://doi.org/10.4018/ijcini.2016010101

Wang, Z., Jiang, C., Ding, Y., Lyu, X., and Liu, Y. (2018). A Novel behavioral scoring model for estimating probability of default over time in peer-to-peer lending. *Electronic Commerce Research and Applications*, *27*, 74–82. https://doi.org/10.1016/j.elerap.2017.12.006

Wang, Z., Jiang, C., Zhao, H., and Ding, Y. (2020). Mining Semantic Soft Factors for Credit Risk Evaluation in Peer-to-Peer Lending. *Journal of Management Information Systems*, *37*(1), 282–308. https://doi.org/10.1080/07421222.2019.1705513

Wei, X., Gotoh, J., and Uryasev, S. (2018). Peer-To-Peer Lending: Classification in the Loan Application Process. *Risks*, *6*(4). https://doi.org/10.3390/risks6040129

Wu, C., Zhang, D., and Wang, Y. (2018). Evaluating the risk performance of online peer-to-peer lending platforms in China. *Journal of Risk Model Validation*, *12*(2), 63–87. https://doi.org/10.21314/jrmv.2018.187

Xia, L., and Li, J. (2016). Analysis on Credit Risk Assessment of P2P. In E. Qi, J. Shen, and R. Dou (Eds.), *Proceedings of the 22nd International Conference on Industrial Engineering and Engineering Management: Core Theory and Applications of Industrial Engineering*. https://doi.org/10.2991/978-94-6239-180-2_86

Xia, Y. (2019). A Novel Reject Inference Model Using Outlier Detection and Gradient Boosting Technique in Peer-to-Peer Lending. *IEEE Access*, *7*, 92893–92907. https://doi.org/10.1109/access.2019.2927602

Xia, Y., He, L., Li, Y., Liu, N., and Ding, Y. (2019). Predicting loan default in peer-to-peer lending using narrative data. *Journal of Forecasting*. https://doi.org/10.1002/for.2625

Xia, Y., Liu, C., and Liu, N. (2017). Cost-sensitive boosted tree for loan evaluation in peer-to-peer lending. *Electronic Commerce Research and Applications*, *24*, 30–49. https://doi.org/10.1016/j.elerap.2017.06.004

Xia, Y., Yang, X., and Zhang, Y. (2018). A rejection inference technique based on contrastive pessimistic likelihood estimation for P2P lending. *Electronic Commerce Research and Applications*, *30*, 111–124. https://doi.org/10.1016/j.elerap.2018.05.011

Xinmin, W., Peng, H., Akram, U., Yan, M., and Attiq, S. (2019). The effect of successful borrowing times on behavior of investors: An empirical investigation of the P2P online lending market.

Human Systems Management, *38*(4), 385–393. https://doi.org/10.3233/hsm-190517

Xiong, J. (2018). Risk Identification and Monitoring Model of Online P2P Lending. In J. Liu and K. L. Teves (Eds.), *Proceedings of the 2018 2nd International Conference on Education, Economics and Management Research* (Vol. 182, pp. 360–363).

Xu, J., and Chau, M. (2018). Cheap Talk? The Impact of Lender-Borrower Communication on Peer-to-Peer Lending Outcomes. *Journal of Management Information Systems*, *35*(1), 53–85. https://doi.org/10.1080/07421222.2018.1440776

Xu, J., Chen, D., and Chau, M. (2016). Identifying Features for Detecting Fraudulent Loan Requests on P2P Platforms. In L. Zhou, L. Kaati, W. Mao, and G. A. Wang (Eds.), *Ieee International Conference on Intelligence and Security Informatics: Cybersecurity and Big Data*.

Xu, J. J., Lu, Y., and Chau, M. (2015). P2P Lending Fraud Detection: A Big Data Approach. In M. Chau, G. A. Wang, and H. Chen (Eds.), *Intelligence and Security Informatics, Paisi 2015* (Vol. 9074, pp. 71–81). https://doi.org/10.1007/978-3-319-18455-5_5

Xu, L., and Zhang, Y. (2017). A Credit Rating Model for Online P2P Lending Based on Analytic Hierarchy Process. In J. Xu, A. Hajiyev, S. Nickel, and M. Gen (Eds.), *Proceedings of the Tenth International Conference on Management Science and Engineering Management* (Vol. 502, pp. 537–549). https://doi.org/10.1007/978-981-10-1837-4_46

Yan, Y., Lv, Z., and Hu, B. (2017). Building investor trust in the P2P lending platform with a focus on Chinese P2P lending platforms. *Electronic Commerce Research*, *18*(2), 203–224. https://doi.org/10.1007/s10660-017-9255-x

Yan, Y., Lv, Z., and Hu, B. (2016). Building Investor Trust in the P2P Lending Platform with a Focus on Chinese P2P Lending Platforms. *2016 International Conference on Identification, Information and Knowledge in the Internet of Things*, 470–474. https://doi.org/10.1109/iiki.2016.15

Yang, X., Fan, W., Wang, L., Yang, S., and Wang, W. (2019). Risk Control of Online {P2P} Lending in {China} Based on Health Investment. *Ekoloji*, *28*(107), 2013–2022.

Yao, J., Chen, J., Wei, J., Chen, Y., and Yang, S. (2019). The relationship between soft information in loan titles and online peer-to-peer lending: evidence from RenRenDai platform. *Electronic Commerce Research*, *19*(1), 111–129. https://doi.org/10.1007/s10660-018-9293-z

Ye, X., Dong, L.-A., and Ma, D. (2018). Loan evaluation in P2P lending based on Random Forest optimized by genetic algorithm with profit score. *Electronic Commerce Research and Applications*, *32*, 23–36. https://doi.org/10.1016/j.elerap.2018.10.004

Yuan, Z. N., Wang, Z. H., and Xu, H. (2018). Credit Risk Assessment of Peer-to-Peer Lending Borrower Utilizing {BP} Neural Network. In L. Barolli, M. Zhang, and X. A. Wang (Eds.), *Advances in Internetworking, Data and Web Technologies, Eidwt-2017* (Vol. 6, pp. 22–33). https://doi.org/10.1007/978-3-319-59463-7_3

Zang, D., Qi, M., and Fu, Y. (2015). The credit risk assessment of P2P lending based on BP neural network. In G. Lee (Ed.), *Industrial Engineering and Management Science*.

Zhang, Y., Geng, X., and Jia, H. (2017). The Scoring Matrix Generation Method and

Recommendation algorithm in P2P Lending. In R. Bahsoon and Z. Chen (Eds.), *2017 13th Ieee World Congress on Services* (pp. 86–89). https://doi.org/10.1109/services.2017.22

Zhang, Y., Jia, H., Diao, Y., Hai, M., and Li, H. (2016). Research on Credit Scoring by fusing social media information in Online Peer-to-Peer Lending. In H. Lee, Y. Shi, J. Lee, F. Cordova, I. Dzitac, G. Kou, and J. Li (Eds.), *Promoting Business Analytics and Quantitative Management of Technology: 4th International Conference on Information Technology and Quantitative Management* (Vol. 91, pp. 168–174). https://doi.org/10.1016/j.procs.2016.07.055

Zhang, Y., Wang, D., Chen, Y., Shang, H., and Tian, Q. (2017). Credit Risk Assessment Based on Long Short-Term Memory Model. In D. S. Huang, K. H. Jo, and J. C. FigueroaGarcia (Eds.), *Intelligent Computing Theories and Application, Icic 2017, Pt Ii* (Vol. 10362, pp. 700–712). https://doi.org/10.1007/978-3-319-63312-1_62

Zhang, Y., Wang, D., Chen, Y., Zhao, Y., Shao, P., and Meng, Q. (2017). Credit Risk Assessment Based on Flexible Neural Tree Model. In F. Cong, A. Leung, and Q. Wei (Eds.), *Advances in Neural Networks, Pt I* (Vol. 10261, pp. 215–222). https://doi.org/10.1007/978-3-319-59072-1_26

Zhang, Y., Wang, X., Qian, Y., and Jia, H. (2016). *The Research of Recommendation Algorithms in {P2P} Lending*.

Zhao, J. (2015). Research on Mathematical Model P2P Online Credit Risk Evaluation Based on Data Processing. In J. Wang and Y. Qin (Eds.), *Proceedings of the 2015 International Conference on Education Technology, Management and Humanities Science* (Vol. 27, pp. 897–900).

Zhou, G., Zhang, Y., and Luo, S. (2018). P2P Network Lending, Loss Given Default and Credit Risks. *Sustainability*, *10*(4). https://doi.org/10.3390/su10041010

Zhou, J., Li, W., Wang, J., Ding, S., and Xia, C. (2019). Default prediction in P2P lending from high-dimensional data based on machine learning. *Physica A-Statistical Mechanics and Its Applications*, *534*. https://doi.org/10.1016/j.physa.2019.122370

Zhu, L., Qiu, D., Ergu, D., Ying, C., and Liu, K. (2019). A study on predicting loan default based on the random forest algorithm. In E. HerreraViedma, Y. Shi, D. Berg, J. Tien, F. J. Cabrerizo, and J. Li (Eds.), *7th International Conference on Information Technology and Quantitative Management* (Vol. 162, pp. 503–513). https://doi.org/10.1016/j.procs.2019.12.017

Zhu, Z. (2018). Safety promise, moral hazard and financial supervision: Evidence from peer-to-peer lending. *Finance Research Letters*, *27*, 1–5. https://doi.org/10.1016/j.frl.2018.07.002

Annex. Variables and specific model of the primary papers of our sample (93 papers, the 5 papers that did not use datasets were excluded)

Year	Dependent variables	Independent variables	Specific Model	Authors/Year
2010	(1) % funded; (2) % reduction interest rate	financial, loan, soft (social capital, listing quality:description-image)	MLR	Greiner and Wang (2010)
2011	(1) % funded; (2) % reduction interest rate; (3) performance (4 categories:paid ahead of schedule and in full, paid as scheduled, payments between	demographic, loan, narrative variables (trustworthy, economic hardship, hardworking, successful, moral, and religious)	MR, MLR	Herzenstein et al. (2011)

	one and four months late, Default)			
2015	Default	demographic, loan, financial, employment, credit	ANN, LR	Byanjankar et al. (2015)
			S-CR, LR	Serrano-Cinca et al. (2015)
		loan, financial, credit	S-CR, LR	Emekter et al. (2015)
	Default; fully paid	demographic, loan, financial, employment, credit	RF, KNN, SVM, LR, FICO, Lcgrade	Malekipirbazari and Aksakalli (2015)
		loan, financial, credit	SVM, RBF, MLP, DT-CART, DT-CHAID	Jin and Zhu (2015)
	Default: with outstanding repayment records	loan, credit, employment, financial	BP-ANN	Zang et al. (2015)
	Expert assessment of the risk	demographic, loan behavior, credit	BP-ANN	Zhao (2015)
	Fraud	loan application, e-commerce, text description, social network	DT, SVM, ANN	Xu et al. (2015)
2016	(1) % funded; (2) final interest rate; (3) Default: 60+ days past due	demographic, loan, credit, experience lending website, interest rate offered, duration (list-expire or loan), role (borrower or borrower and lender)	HMR, LR, S-CR	Chen et al. (2016)
	(1) IRR; (2) Default: Charged off, fully paid	demographic, loan, financial, employment, credit	MLR, LR, DT	Serrano-Cinca and Gutierrez-Nieto (2016)
	# investors	financial, credit, social capital, risk management (mortgage collateral or third-party guarantee), listing information, operation duration	ELM-MLR	Yan et al. (2016)
	Default	heterogeneous social data: demographic, tweet, network, high-level (features derived from different classifiers)	GBDT, RF, BAG, NB, L1-LR, SVM; with LDA	Guo et al. (2016)
		loan, credit, soft information (deception, subjectivity, sentiment, readability, personality and mode of thought)	LR, DT-C4.5, NB, MLP, RF	Wang et al. (2016)
	Default; fully paid	demographic, loan, credit historic, social media	DT, LR, ANN	Zhang, Jia, et al. (2016)
	Default: 16+ days past due, charged off; Current, Fully Paid, In grace period	loan	RF, DT, BAG	Kumar et al. (2016)
	Default: overdue	demographic, credit, loan, platform variables (successful loan number, failed loan number, membership score, prestige, forum currency, contribution, group)	OC	Li et al. (2016)
	Fraud	loan, learning, past performance, social networking, herding manipulation	RF, SVM	Xu et al. (2016)
	repayment rate	loan application, loan, demographic, platform	MLR	Xia and Li (2016)
	top-N recommendation	Invest or not, proportion express bid amount user invest in a project, categorical proportion amount variable	UCF, ICF, LDA	Zhang, Wang, et al. (2016)
2017	(1) Funded; (2) interest rate; (3) Default	demographic, loan, credit, financial, third-party credit guarantee, offline authentication	LR, TOBIT, PROBIT	Tao et al. (2017)
	# investors	financial, credit, social capital, risk management (mortgage collateral or third-party guarantee), listing information, operation duration	ELM-MLR	Yan et al. (2016)
	Default	demographic, financial, loan, credit	LR	Chen (2017)
		NI	ECSC-L2-LR, ECSC-DT, ECSC-SVM, ECVWC-L2-	Ding et al. (2017)

			LR, ECVWC-DT, ECVWC-SVM	
		demographic, loan, financial, employment, credit	cs-XGBoost, LR, RF, cs-LR-T, cs-RF-T, cs-LR-sm, cs-RF-sm	Xia et al. (2017)
	Default: 120+ days past due	demographic, loan, campaign, image, social media	LR. PSM, IV, Dif-Dif	Ge et al. (2017)
	Default: 120+ days past due-charged off	credit, loan, financial, verification, number of derogatory public records	LR	Stofa (2017)
	Default: 150+ days past due-charged off	NI	cs-GNB, cs-DT, cs-LR, ru-GNB, sm-GNB, sm-DT, ru-DT, ru-LR, sm-LR, AdaBoost, BAG, RF	Boiko Ferreira et al. (2017)
	Default: 30+ days past due; Bad debt rate	Forward-looking credit indicators (demographic, financial, employment), Backward-looking credit indicators (Successful borrowing times, borrowing amount), loan, borrowing description and titles (lengthed)	PROBIT, TOBIT. Three-stage dynamic game to analyze the default behavior of borrowers	Gao et al. (2017)
	Default: three consecutive payments past due	demographic, loan, verification, employment, financial	S-DA(KM), S-CR, S-ANN, S-LR	Byanjankar (2017)
	Default: charged off	loan	S-DA(KM)	Durovic (2017)
	Default:overdue	demographic, employment, financial, credit	FNT, IBPNN, RBF, ANN	Zhang, Wang, Chen, Zhao, et al. (2017)
			LSTM, IBPNN, RBF, ANN	Zhang, Wang, Chen, Shang, et al. (2017)
		demographic, loan, financial, employment, credit	LR	Lin et al. (2017)
	Degree by node	profit (lenders and borrowers), collateral	BBN	Lee et al. (2017)
	Rank by risk or profitability	financial, management, operation, guarantee	AHP-Index	Xu and Zhang (2017)
	top-N recommendation	Default, time value of money, profits	TCVM (profit:term, r, default(RF, GBDT, KNN, NB LR, DT); TVM)	Tan et al. (2017)
		Invest or not, proportion express bid amount user's invest in a project, categorical proportion amount variable	UCF, ICF, LDA	Zhang, Geng, et al. (2017)
2018	(1) Funded; (2) Default	demographic, loan, credit historic, online seller, student, safety promise, number of bidders	L1-LR	Zhu (2018)
	(1) Funded; (2) interest rate; (3) % Funded; (4) # Bids; (5) Default	demographic, loan, Soft Information (communication, interaction, Lender Comment-Social Influence, Borrower Response-Info Quality)	MLR, LR, S-CR	Xu and Chau (2018)
	5 credit grades	demographic, loan, financial	BP-ANN	Yuan et al. (2018)
	7 credit grades	loan, financial, credit	csDT, csNB, csLR, csSVM with 3 cost matrices	Wang, Kou, et al. (2018)
	Approved; declined	financial, loan, credit historic, employment	LR	Wei et al. (2018)
	Default	demographic, loan, credit, network behavior, third-party, social network	E(XGBoost, DNN, LR), XGBoost, DNN, LR	Li, Ding, et al. (2018)
		demographic, financial, loan	LR, S-DA	Canfield (2018)
		demographic, loan, credit, financial	LR, PROBIT	Liu et al. (2018)
		demographic, loan, financial, credit	CPLE-LightGBM, CPLE-RF, CPLE-LR, CPLE-SVM, E-SVM, E-RF, E-	Xia et al. (2018)

			LightGBM, S3VM, Ex, Ag, LR, SVM, RF	
		demographic, loan,reputational risk, Macro (sectorial risk, country risk)	QM-EBRR	Uddin et al. (2018)
		demographics, mobility patterns, phone usage data: (phone calls, text messages, and data traffic), telecommunication patterns, App usage patterns and telecommunication records	AdaBoost, RF, LR	Ma, Zhao, et al. (2018)
		financial ratios+ degree and closeness centrality measures	CN-LR, CN-DT, LR, DT	Hadji-Misheva et al. (2018)
		loan, financial, loan behavior	LightGBM, XGBoost	Ma, Sha, et al. (2018)
	Default; fully paid	demographic, loan, financial, soft features (asset, income, work, family, agriculture, and length)	RF, LR, NB, SVM; with LDA	Jiang et al. (2018)
	Default: 120+ days past due-charged off	demographic, loan, financial, credit	RFoGAPS, RF, SVM, DT, KNN, LN, Actual profit, LR	Ye et al. (2018)
	Default: 150+ days past due-charged off	demographic, loan, financial	E-FIC(GBDT, AdaBoost, LR), E-MV(GBDT, AdaBoost, LR), E-OWAo(GBDT, AdaBoost, LR), E-OWAp(GBDT, AdaBoost, LR), GBDT, AdaBoost, LR	Namvar et al. (2018)
	Default: 90+ days past due	demographic, loan, financial, credit	S-EMRF, MCM, S-CR, LR	Wang, Jiang, et al. (2018)
	Default: charged off, fully paid	loan, financial	GBDT, SVM, LR, KNN, NB, RF	Rodrigues et al. (2018)
	LGD	loan, credit, financial	MLR	Zhou et al. (2018)
	network systemic risk: DVtf	network variables	CN, MLR	Li, Hao, et al. (2018)
2019	Grade≥A: 1, Grade<A: -1	demographic, employment, financial, loan, credit, institutional guarantee	GA-SVM	Yang et al. (2019)
	(1,3) Default; (2) Rejected	demographic, loan, employment, credit	LD, LR, SVM, ANN, KNN, RF with RBM-FS	Nguyen Truong et al. (2019)
	(1) % negative sentimen comments-(2) trading volume index-(3) Weekday	% negative sentimen comments, TVI (tradingvolumeindex-ND), Weekday	Text-CNN-LSTM, Text-CNN-VAR, Text-CNN-DNN, Text-CNN-MLR, Text-CNN-RF, Text-CNN-SVM	Fu et al. (2019)
	(1) default-charged off-(2) 1+ public record bankruptcies	Thau-kendall (copula), loan, financial, credit, spatial variables	BivGEV, BivProbit, LR	Calabrese et al. (2019)
	(1) Default; (2) IRR	demographic, loan, financial, credit	DP, WL, WDP, RF, GB, SVM with IHT, BC, ru, so, sm	Bastani et al. (2019)
	(1) Default; (2) Rejected	demographic, loan, employment, credit	LD, LR, SVM, ANN, KNN, RF with RBM-FS	Van-Sang et al. (2019)
	(1) Default; (2) Succesful	demographic, loan, verification and proof variables, successful borrowing times,	LR, Probit	Xinmin et al. (2019)
	(1) Potential defaulted; (2) Default	loan, credit, financial, employment, region information	LR, RF, SVM; S3VM, OD-LightGBM	Xia (2019)
	Default	demographic, financial, soft (loan description, character, clustering result based soft information)	CatBoost, LR, DT, BNN, RF, GBDT, XGBoost	Xia et al. (2019)
		demographic, registration, loan, social network	LightGBM, RF, AdaBoost and LR	Niu et al. (2019)
		financial ratios	CN(LF)-LR	Ahelegbey et al. (2019)

	NI	E(GBDT, XGBoost, LightGBM), NN, LR, RF, SVM, KNN, AdaBoost	Zhou et al. (2019)
	soft information (online operation behavior: data of borrower's online operation behavior on P2P lending website), credit	AM-LSTM, BOA-XGBoost, LSTM, BLSTM, BLSTM-Meanpool.	Wang et al. (2019)
Default: 16+ days past due, charged off, In Grace period; Current, Fully Paid, Issued	loan, verification, application type	RF, DT, SVM, LR with sm	Zhu et al. (2019)
Default: 90+ days past due	demographic, loan, employment, financial, credit	Latency and Incidence with MCM: RF-TDH, RF-Cox, LR-Cox, LR-TDH and RF, LR, DT, B-LR, B-DT, BAG-LR, BAG-DT	Jiang et al. (2019)
Default: charged off, fully paid	demographic, loan, credit	E(LP, TSVM), LP, TSVM, DT	A. Kim and Cho (2019)
	demographic, loan, employment, financial, credit	IEFSVM, cs-AdaBoost,cs-RF, EE ,ru-B, w-ELM, cs-XGBoost, EFSVM	Cho et al. (2019)
Default:120+ days past due, 120-days past due, 0 safe loans	demographic, loan, financial	DNN, LR, LD, DT, SVM, RBF, MLP, AdaBoost	Duan (2019)
Default: charged off, fully paid	demographic, loan, employment, financial, credit	CNN: (CNN, Inception, ResNet, DenseNet, Inception-ResNet); MLP, SVM, KNN, DT, RF	Kim and Cho (2019a)
		DP-CNN, CNN: (CNN, Inception, ResNet, DenseNet, Inception-ResNet); MLP, SVM, KNN, DT, RF	Kim and Cho (2019b)
Default:with outstanding repayment records	demographic, loan, platform authentication, regulation change of the governmen	LQR	Chen et al. (2019)
Hazard exposed (absconded with ill-gotten gains, difficult withdrawing, out of business and investigated by Economic Crime Investigation Police), Normal.	Text information: Management team members' working experience, educational background, and composition	MUN-LETCLA, Doc2vec, LDA, Dependency, Syntactic, Keyword	Li et al. (2019)
fully repay the loan in advance	demographic, loan, verification, financial, employment, credit	S-CR	Wan et al. (2019)
Funded	demographic, loan, credit, financial, employment, platform audition, soft information (capital turnover, Investment and entrepreneurship, expanding business, ambiguous purpose, house decoration, household expenses, Daily consumption, Car purchase)	LDA(VSM)-LR	Yao et al. (2019)
Number of loans by grade. top-N recommendation	return-risk trade-off (probability of appearance of loans in the future and maximum number of days the investor is willing to wait to invest their funds)	MO	Ren et al. (2019)
Total loss ratio (LGD, balance at default)	demographic, loan, employment, financial	SPM-LIR	Gourieroux and Lu (2019)
Lending Intention, Platform Trust	Protection Policies, Platform Trust	SEM	Amalia et al. (2019)
Survey: demographic, investment background, investment behaviour, investment criteria, motivation to lend variables		DA, PCA	Pierrakis (2019)
Semi-structured interviews (non-probability sampling): Collateral, requirement, procedure, online process, fast process, interest rate, credit scoring, profits, costs variables		QM, DA	Rosavina et al. (2019)

2020	Default	demographic, loan, financial, employment, credit	DT-J48, LR	Cai and Zhang (2020)
		demographic, network behavior, third party information, social network, loan transaction time, loan	E(XGBoost, DNN, LR), XGBoost, DNN, LR, AdaBoost, GBDT, RF, DT, KNN, SVM	Li et al. (2020)
		financial ratios, degree(number of neighbors of a node), strength(average distance of a node for its neighbors) and PageRank (importance of a node in a network by assigning relative scores to all nodes in the network)	LR	Giudici et al. (2020)
	Default: 120+ days past due	demographic, loan, credit, soft information (semantic vs. linguistic and stylistic factors)	LR, L1-LR, RF, XGBoost with Semantic Soft Factor Mining Method, LDA	Wang et al. (2020)
	Rank by risk or profitability	demographic, loan, employment, financial, credit	DHPF-TODIM	Ji et al. (2020)

Legend: AdaBoost: Adaptive Boosting, Ag: Augmentation, AHP: Analytic Hierarchy Process, AM: Attention Mechanism, ANN: Artificial Neural Network, Ap: Actual profit, B: Boosting, BAG: Bagging method, BBN: Bianconi-Barabasi network model, BC: Balance Cascade, BivGEV: Bivariate Generalized Extreme Value regression, BivProbit: Bivariate Probit regression, BLSTM: Bidirectional LSTM network, BNN: Bagging Neural Network, BOA: Bayesian hyper-parameter optimization, BP: Back Propagation, C4.5: C4.5 decision tree algorithm , CatBoost: Unbiased boosting with categorical features, CN: Connective Network model, CNN: Convolutional Neural Network, CPLE: Contrastive Pessimistic Likelihood Estimation, CR: Cox Regression, cs: cost sensitive, DA: Descriptive Analysis, DenseNet: Dense Convolutional Network, DHPF-TODIM: Dual Hesitant Pythagorean Fuzzy number + TODIM approach ((an acronym in Portuguese of interactive and multi-criteria decision making), Dif-Dif: Difference-in-Differences, DNN: Deep Neural Network , DP: Deep Learning, DST: Dempster–Shafer theory, DT: Decision Tree, E: Ensemble, EBRR: Expert Based Risk Rating, ECSC: Ensemble Classification based on Supervised Clustering , ECVWC: Ensemble Classification based on Variable Weighting Clustering , EE: Easy Ensemble, EFSVM: Entropy Fuzzy Support Vector Machine, ELM: Elaboration Likelihood Model, EMRF: survival Ensemble Mixture RF, Ex: Extrapolation, FIC: Fuzzy Integral Combination, FICO: Fair Isaac Corporation Score, FNT: Flexible Neural Tree, FS: Feature selection, GA: Genetic Algorithm, GB: Gradient Boosting regression, GBDT: Gradient Boosting Decision Tree, GNB: Gaussian Naive Bayes, HMR: Hierarchical Multiple Regression, IBPNN: Improved BP Neutral Network, ICF: Item based Collaborative Filtering, IEFSVM: Instance-based Entropy Fuzzy Support Vector Machine, IHT: Instance Hardness Threshold, IV: Instrumental Variables, KM: Kaplan-Meier method, KNN: K-Nearest Neighbor, L1: Lasso (L1 norm) Regularization, L2: Ridge (L2 norm) Regularization, LC: Lending Club, LD: Linear Discriminant Analysis, LDA: Latent Dirichlet Allocation, LF: Latent factor model, LightGBM : Light Gradient Boosting Machine, LIR: semi-parametric estimator Least Impulse Response, LP: Label Propagation, LQR: Logistic Quantile Regression, LR: Logistic Regression, LSTM: Long Short-Term Memory, MCM: Mixture Cure Model, MLP: Multilayer Perceptron , MLR: Multiple Linear Regression, MO: Markowitz optimization model, MR: Multinomial Regression, MUN-LETCLA: Multiple NLP (Natural Language Processing) Integrated Learning Text Classifier Model, MV: Majority voting, NB: Naive Bayes, NR: Nuclear Regression model, OC: Outliers Cluster-based method, OD: Outlier Detection, OWAo: Optimistic OWA (ordered weighting averaging), OWAp: Pessimistic OWA (ordered weighting averaging), PCA: Principle Component Analysis, PSM: Propensity Score Matching, QM: Qualitative Model, RBF: Radial Basis Function , RBM: Restricted Boltzmann Machines, ResNet: Residual Neural network, RF: Random Forest, RFoGAPS: Random Forest optimized using a genetic algorithm with profit score, ru: random undersampling, S: Survival, S3VM: semi-supervised variants of SVM, SEM: Structural Equation Model, sm: smote (Synthetic Minority Over-sampling Technique), so: random oversampling, SPM: Semi-Parametric Transformation model, SVM: Support Vector Machine , T: Threshold, TCVM: Total Capital Value Maximization, TDH: Time-Dependent Hazards, TSVM: Transductive Support Vector Machine, TVM: Time Value of Money, UCF: User based Collaborative Filtering, VAR: Vector Autoregressive Model, VSM: Vector Space Model, WDP: Wide and Deep Learning, WL: Wide Learning, XGBoost: Extreme Gradient Boosting. NI: No Information.

Source: Compiled by the authors