ORIGINAL PAPER

DOI: 10.26794/2587-5671-2020-24-6-38-50
UDC 336.7(045)
JEL E58, G21, C53

Analysis of Possibilities to Automate Detection of Unscrupulous Microfinance Organizations Based on Machine Learning Methods

Yu.M. Beketnova
Financial University, Moscow, Russia
https://orcid.org/0000-0002-1005-6265

ABSTRACT

Microfinance is a way to fight poverty, and therefore is of high social significance. The microfinance sector in Russia is progressing. However, the engagement of microfinance organizations in illegal financial transactions associated with fraud, illegal creditors, money laundering, significantly limits their potential and has negative impact on their development. The **aim** of the paper is to study the possibilities to automate detection of unscrupulous microfinance organizations based on machine learning methods in order to promptly identify and suppress illegal activities by regulatory authorities. The author cites common fraudulent schemes involving microfinance organizations, including a scheme for cashing out maternity capital, a fraudulent lending scheme against real estate. The author carried out a comparative analysis of the results obtained by classification **methods** — the logistic regression method, decision trees (algorithms of two-class decision forest, Adaboost), support vector machine (algorithm of two-class support vector machine), neural network methods (algorithm of two-class neural network), Bayesian networks (algorithm of two-class Bayes network). The two-class support vector machine provided the most accurate results. The author analysed the **data** on microfinance institutions published by the Bank of Russia, the MFOs themselves, and *banki.ru*. The author **concludes** that the research results can be of further use by the Bank of Russia and Rosfinmonitoring to automate detection of unscrupulous microfinance organizations.

Keywords: microfinance organizations; financial monitoring; machine learning methods; classification algorithms

For citation: **Beketnova Yu.M. Analysis of possibilities to automate detection of unscrupulous microfinance organizations based on machine learning methods. *Finance: Theory and Practice.* 2020;24(6):38-50. (In Russ.). DOI: 10.26794/2587-5671-2020-24-6-38-50**

INTRODUCTION

Microfinance organizations are financial companies that offer small loans to unbanked or low income populations. MFOs provide microloans, insurance services, deposits and other services. These organizations are widespread in Russia, Asia, Europe, Africa and many other countries.

A. S. Sorokin and V. A. Shilov note that microfinance is a way to fight poverty, and therefore is of high social significance [1]. The microfinance sector in Russia is progressing, which is presented in the studies by E. B. Makarova [2], N. B. Balasheva [3], V.A. Tsvetkov [4] and other authors [5–7].

The article by Yu. S. Yevlakhova [8] proves that the engagement of microfinance organizations in illegal financial transactions associated with fraud, illegal creditors, money laundering, significantly limits their potential and has negative impact on their development.

Rosfinmonitoring Public Report “National Assessment of the Risks of Money Laundering. Main conclusions” also indicates the high risk of using MFOs in money laundering schemes.[1]

The MFOs vulnerability is partly due to the relatively easy registration of these organizations, as well as their specifics, in particular, the ability to legally raise funds from legal entities, redistributing them between individuals.

The Bank of Russia, together with law enforcement agencies, the General Prosecutor’s Office and the Federal Financial Monitoring Service, is consistently clearing the financial sector of unscrupulous organizations.

The high probability of client default is the main risk inherent in the work of microfinance organizations. The key to their success is to issue as much funds as possible with a minimum default and the lowest costs.

The two main money laundering schemes involving MFOs are as follows:

- an attempt on assets that belong to financial organizations and are included in their main activities;
- an attempt on assets that belong to investors and are attracted by financial organizations in order to steal them in the future.

The main vulnerabilities to money laundering involving MFOs are legal and financial ignorance of citizens, as well as an insufficient amount of resources necessary for effective fight against fraudsters among the majority of microfinance *market participants.*

The MFO owners and management may commit crimes related to attempts on assets that belong to investors and are attracted by financial organizations. Asset withdrawals can be disguised as standard procedures for issuing loans and raising funds.

The first stage of asset withdrawals is often falsified reporting, with either reduced the real attracted resources or increased microloans in order to formally comply with the established standards and not be excluded from the register.

ANALYSIS OF FRAUDULENT SCHEMES INVOLVING MFOS

S.E. Volkov and I.N. Loskutov [9] show some fraudulent schemes involving microfinance organizations:

Intermediary services. A future MFO client is imposed intermediary services that supposedly guarantee for borrowed funds. For this service, the intermediary company charges a commission of up to 50% of the loan amount. The intermediary company can conclude an agreement with the client, stating that the client is obliged to pay a one-time registration fee of 5% to 20% of the loan amount. Fraudsters may demand to

[1] Rosfinmonitoring. National assessment of the risks of money laundering. Main conclusions. 2017–2018. Rosfinmonitoring Public Report. 2018. URL: http://www.fedsfm.ru/content/files/documents/2018/%D0%BE%D1%86%D0%B5%D0%BD%D0%BA%D0%B0%20%D1%80%D0%B8%D1%81%D0%BA%D0%BE%D0%B2%20%D0%BE%D0%B4_5.pdf (accessed on 22.07.2020).

pay for checking the client's credit history, notary services, pay a membership fee in a credit cooperative, etc. This "intermediation" does not actually guarantee the client borrowed funds by a microfinance organization, and refunds are not provided.

Credit ladder. Supposedly to confirm his/her solvency, the client is given small amounts of borrowed funds at a high interest rate, after the next loan is repaid, the amount of the next one is increased, and the interest is reduced, and so on until the desired amount is reached. However, after the client pays the penultimate loan, the fraudster disappears. The victim ends up paying high interest rates without receiving the desired loan amount.

Activities of illegal lenders when providing microloans online. Applying for a microloan online on MFO websites, one must fill out a questionnaire that can be used by fraudsters to collect personal data. A potential borrower registers on this website and fills out a questionnaire with personal data. The victim sends an application and receives a message with a refusal of the loan and a proposal to reapply. The contact details of the registered person are used by fraudsters to apply for a loan in real microfinance companies that issue money online [9].

Now let us consider other common fraudulent and money laundering schemes involving microfinance organizations.

Cashing out maternity capital. Based on expert estimates, hundreds of companies may be involved in schemes for providing shadow services for cashing out maternity capital. They find troubled families that have received a birth certificate and provide them with services for cashing these funds.

Citizens are issued with a fake home purchase, usually unsuitable for living or from their own relatives (*Fig. 1*). Then, a loan for this transaction is issued in a microfinance organization and repaid with a certificate. Families receive part of their maternity capital in cash, and the intermediary lenders get the rest.

Fraudulent lending against real estate. An individual needs a loan, but banks refuse him/her or offer a loan at a high rate. S/he contacts a microfinance organization that is funded by private investors or banks. The borrower receives the money, and among other documents s/he has to sign a mortgage on his apartment.

If a client violates the terms of the loan agreement, the lender takes the apartment under the mortgage. The microfinance organization sells the apartment to a figurehead (holder). When the scandal ends, the MFO resells it to a bona fide buyer.

ANALYSIS OF THE DYNAMICS OF REGISTRATION AND LIQUIDATION OF MFOS

If the Bank of Russia detects unscrupulous microfinance market participants involved in dubious transactions, it excludes them from the register. As of July 27, 2020, 1,618 microfinance organizations are registered in Russia.[2]

Moscow is the leader by the number of MFOs registered in the territory (*Table 1*). Top ten also includes the Novosibirsk region, St. Petersburg, Irkutsk region, etc.

In terms of 1 million people living in the region (according to Rosstat data as of January 1, 2020[3]), the Altai Republic, the Arkhangelsk and Kostroma regions, and others are the leaders (*Table 2*). Moscow is on the 21st place.

To consider the dynamics of excluding MFOs from the register and the formation of new MFOs by region, we will correlate the number of MFOs excluded from the register for the entire time and the number of existing ones. The leaders of the rating are the Chechen Republic (more than 36 MFOs

[2] Bank of Russia. URL: https://www.cbr.ru/microfinance/registry/ (accessed on 27.07.2020).

[3] Rosstat. URL: http://www.statdata.ru/largest_regions_russia (accessed on 27.07.2020).

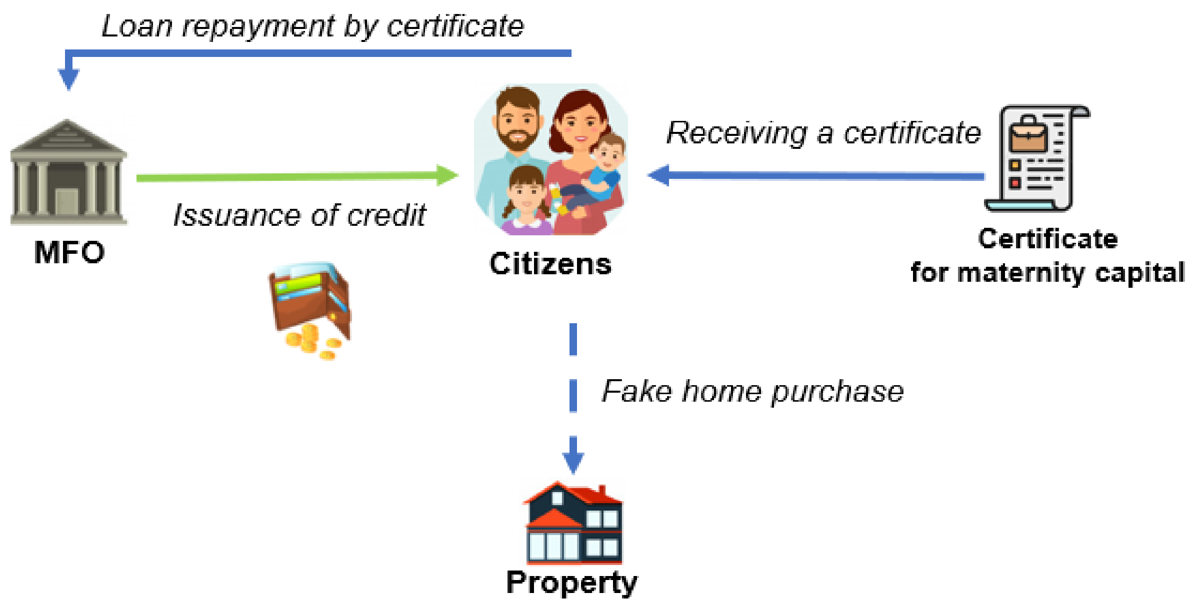

Fig. 1. **Maternity capital cashing scheme through MFOs**

Source: compiled by the author.

excluded from the register against one operating), the Yamalo-Nenets Autonomous District, Yaroslavl region, the Republic of Dagestan, etc. Moscow is on the 12th place (almost 8 excluded MFOs against one operating) (*Table 3*).

We will now consider the MFO average **operating time** before they are excluded from the Register (*Table 4*).

The MFO lowest operating time before exclusion from the register is in the Republic of Ingushetia — 14 months. In the Chechen Republic, this period is slightly longer — 15 months. There are more than 36 MFOs excluded from the register per one operating. Microfinance organizations in this region are very unstable.

The MFO greatest average operating time before exclusion from the register in the Kabardino-Balkarian Republic is 3.5 years. This period is about three years in the Sakhalin and Kaliningrad regions (*Table 5*).

Against the background of small values of the ratio of the number of MFOs excluded from the register to the number of operating organizations, these regions, as well as

Table 1

Top-10 regions by number of MFOs

Region	Number of MFOs
Moscow	226
Novosibirsk region	61
St. Petersburg	59
Irkutsk region	53
Arkhangelsk region	44
Krasnoyarsk region	44
Rostov region	44
Samara Region	43
Krasnodar region	42
Republic of Bashkortostan	42

Source: compiled by the author based on data from the Bank of Russia.

Table 2

Top-10 regions by number of MFOs per 1 million people

No.	Region	Number of MFOs	Population	Number of MFOs per 1 million people
1	Altai Republic	9	220 181	40.88
2	Arhangelsk region	44	1 136 535	38.71
3	Kostroma region	22	633 385	34.73
4	Republic of Sakha (Yakutia)	28	971 996	28.81
5	Republic of North Ossetia – Alania	19	696 837	27.27
6	Jewish Autonomous Region	4	158 305	25.27
7	Republic of Khakassia	13	534 262	24.33
8	Tomsk region	26	1 079 271	24.09
9	Amur region	19	790 044	24.05
10	Udmurtia	36	1 500 955	23.98
...				
21	Moscow	226	12 615 279	17.91

Source: compiled by the author based on data from the Bank of Russia.

Table 3

Top 10 regions by the number of MFOs excluded from the register to the existing ones

No.	Region	Number of MFOs excluded from the register	Number of MFOs	The ratio of the excluded MFOs to the existing MFOs
1	Chechen Republic	145	4	36.25
2	Yamalo-Nenets Autonomous District	16	1	16.00
3	Yaroslavl region	79	6	13.17
4	Republic of Dagestan	51	4	12.75
5	Pskov region	12	1	12.00
6	Saratov region	70	7	10.00
7	Volgograd region	99	10	9.90
8	Republic of Tatarstan	335	36	9.31
9	Omsk region	123	14	8.79
10	Kabardino-Balkar Republic	25	3	8.33
...				
12	Moscow	1734	226	7.67

Source: compiled by the author based on data from the Bank of Russia.

Table 4

Top 10 regions with the lowest average operating time

Region	Average work duration of an MFO (months)	Number of MFOs excluded from the register	Number of MFOs	The ratio of the excluded MFOs to the existing MFOs
Republic of Ingushetia	14	3	1	3.00
Chechen Republic	15	145	4	36.25
Pskov region	20	12	1	12.00
Altai Republic	21	28	9	3.11
Saratov region	21	70	7	10.00
Republic of Karelia	22	17	5	3.40
Kaluga region	22	20	8	2.50
Moscow	22	1734	226	7.67
Kurgan region	23	24	5	4.80
Tver region	23	29	5	5.80

Source: compiled by the author based on data from the Bank of Russia.

the Primorsky Krai and the Astrakhan region, preserve high stability of the microfinance sector.

DATA AND METHODS

The decision of the Bank of Russia to exclude an MFO from the register may indicate its involvement in shadow financial schemes. Automating the detection of unscrupulous microfinance organizations based on machine learning methods will allow regulatory authorities to promptly identify and suppress illegal activities, thereby contributing to the stability of the microfinance sector.

We will now examine the likelihood of excluding MFOs from the register. For this, we downloaded the data from the official website of the Bank of Russia.[4] We also collected MFO reports that must have been published on their websites. The information about organizations and the reports is supplemented by a rating by *banki.ru* based on the feedback from MFOs clients.

The following indicators were downloaded for each organization:

- company name;
- registration date;
- date of exclusion from the register;
- period of activity in months;
- region;
- city;
- activity profile;

[4] Bank of Russia. URL: https://www.cbr.ru/microfinance/registry/ (accessed on 27.07.2020).

Table 5

Top 10 regions with the greatest average operating time

Region	Average work duration of an MFO (months)	Number of MFOs excluded from the register	Number of MFOs	The ratio of the excluded MFOs to the existing MFOs
Kabardino-Balkar Republic	42	25	3	8.33
Sakhalin region	35	29	10	2.90
Kaliningrad region	35	44	11	4.00
Tula region	34	56	10	5.60
Karachay-Cherkess Republic	34	3	2	1.50
Leningrad region	33	28	7	4.00
Primorsky Krai	33	114	35	3.26
Republic of Kalmykia	33	6	3	2.00
Astrakhan region	33	55	11	5.00
Republic of North Ossetia – Alania	32	20	19	1.05

Source: compiled by the author based on data from the Bank of Russia.

- rating;
- authorized capital as of the end of the year;
- long-term financial obligations;
- book value of net assets as of the end of the year;
- book value of assets;
- net interest income;
- net profit;
- year of the submitted reports;
- how many years ago the last reporting was submitted;
- number of founders — individuals and legal entities;
- location of founders, managers and MFOs in different regions;
- information about the legal entities-founders in offshores.

Information on these indicators was collected by year for the period from 01.01.2015 to 09.05.2020. The final sample included 100 microfinance organizations, of which 50 were excluded from the MFO register for the specified period.

Logical variable "In the register" has been introduced as an MFO sustainability indicator. Its two states are defined as follows:

0 is the organization excluded from the register;

1 is the organization in the register.

Consequently, the final list of indicators consists of 100 organizations, 19 indicators (listed above) collected over 64 months, and the resulting column "In the register", which contains the information about the MFOs excluded from the register.

Identifying unscrupulous participants in the microfinance market can be viewed as a binary classification task. We will consider traditional methods and modern algorithms that solve classification problems in the field of financial monitoring.

Logistic regression

Logistic regression is a well-studied and widely used method in statistics. In modern studies highlighted in publications [10–12], logistic regression is used in combination with other methods or for comparison [13, 14].

Decision trees

Decision tress are advantageous for solving classification problems:

- efficient in computing and using computer memory, which makes them suitable for working with large amounts of data;
- the choice of functions is integrated into the learning and classification processes;
- are nonparametric models, which allows processing data with different distributions.

The disadvantages of these algorithms include the fact that the results may be variable and non-reproductive when the sample is changed.

To eliminate the disadvantages, the decision tree ensembles are used. The ensembles are based on a general principle that allows you to get the best results by combining several related models. Typically, ensemble models are more accurate than individual decision trees.

There are many different ways to ensemble decision trees. The *two-class decision forest* and the *Adaboost* algorithm are the most efficient for solving financial monitoring problems.

Using combined methods for solving practical problems is common. [15–17] prove this fact.

Vector Machine method

The algorithm was proposed in 1963 by Vladimir Vapnik and Aleksei Chervonenkis. The *Two-class Support Vector Machine* creates a binary classification model by a support vector machine. A two-class support vector machine is a supervised learning algorithm that trains on labeled data.

Neural networks

A lot of different models are based on neural networks. To solve financial monitoring

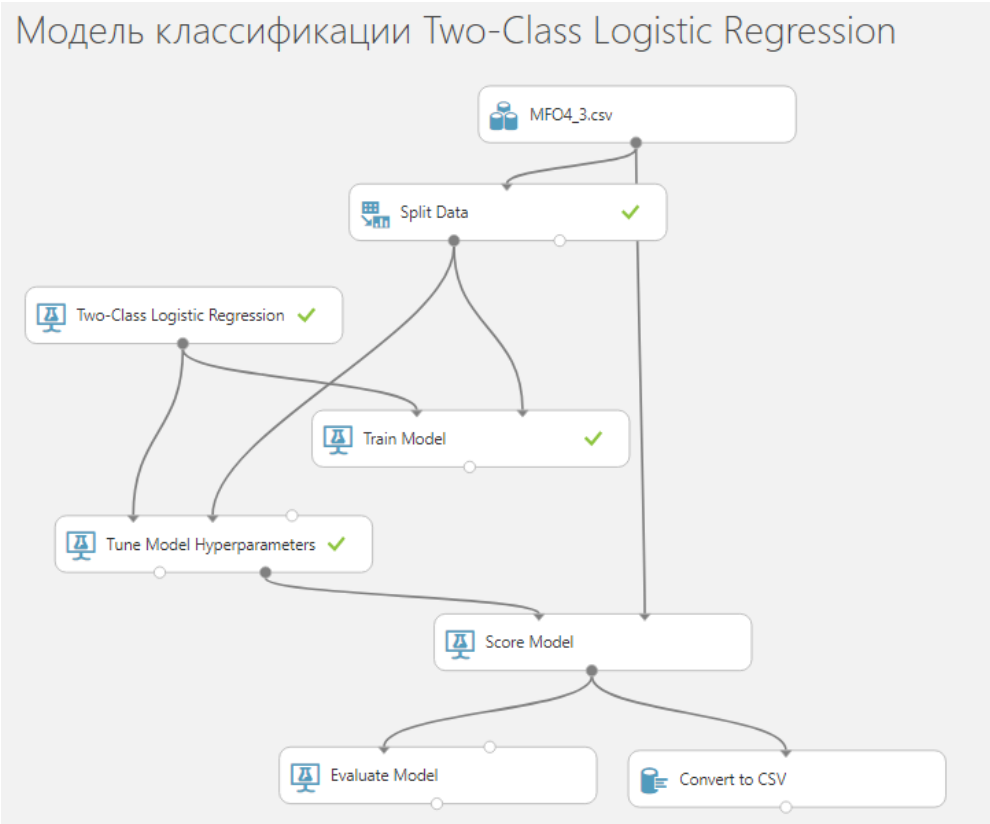

Fig. 2. **Two-Class Logistic Regression Classification Model**

Source: compiled by the author.

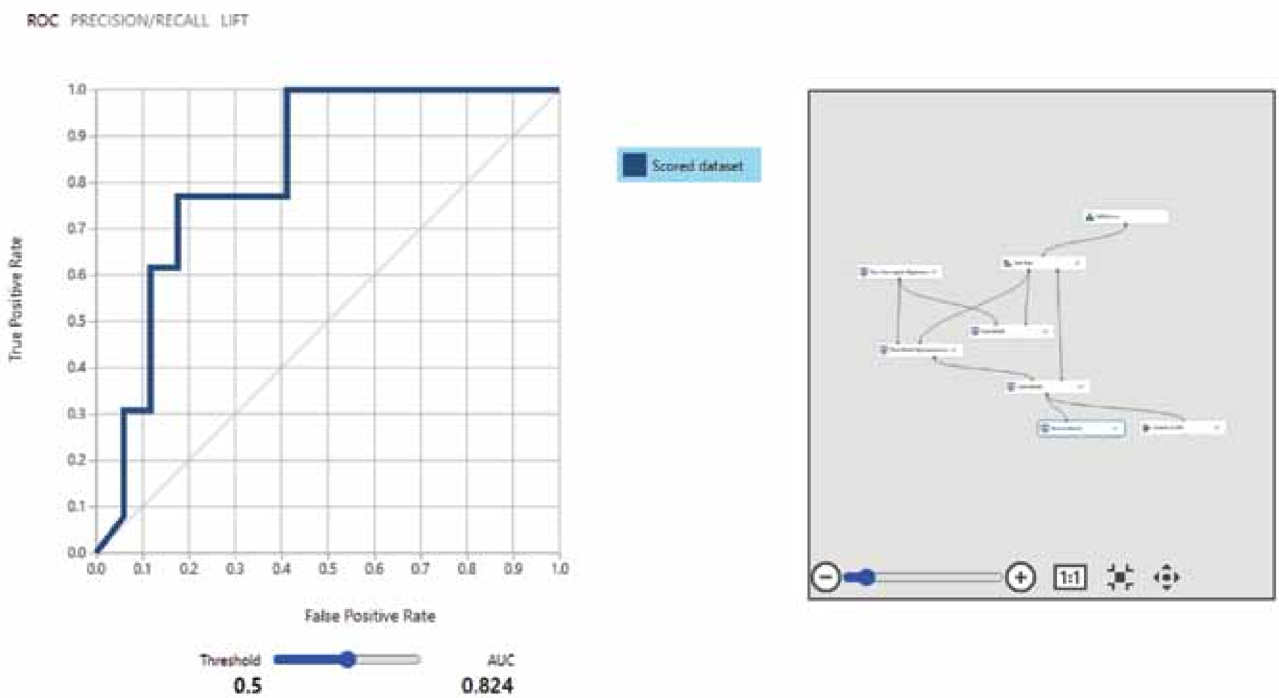

Fig. 3. **ROC Curve and Accuracy Ratings for the Loregression Model**

Source: compiled by the author.

Table 6

AUC indicators for classification algorithms

	Classification model					
	Two-Class Logistic Regression	Two-Class Decision Forest	Two-Class Boosted Decision Tree	Two-Class Neural Network	Two-Class Support Vector Machine	Two-Class Bayes Point Machine
AUC ratio	0.824	0.796	0.688	0.833	0.873	0.855

Source: compiled by the author.

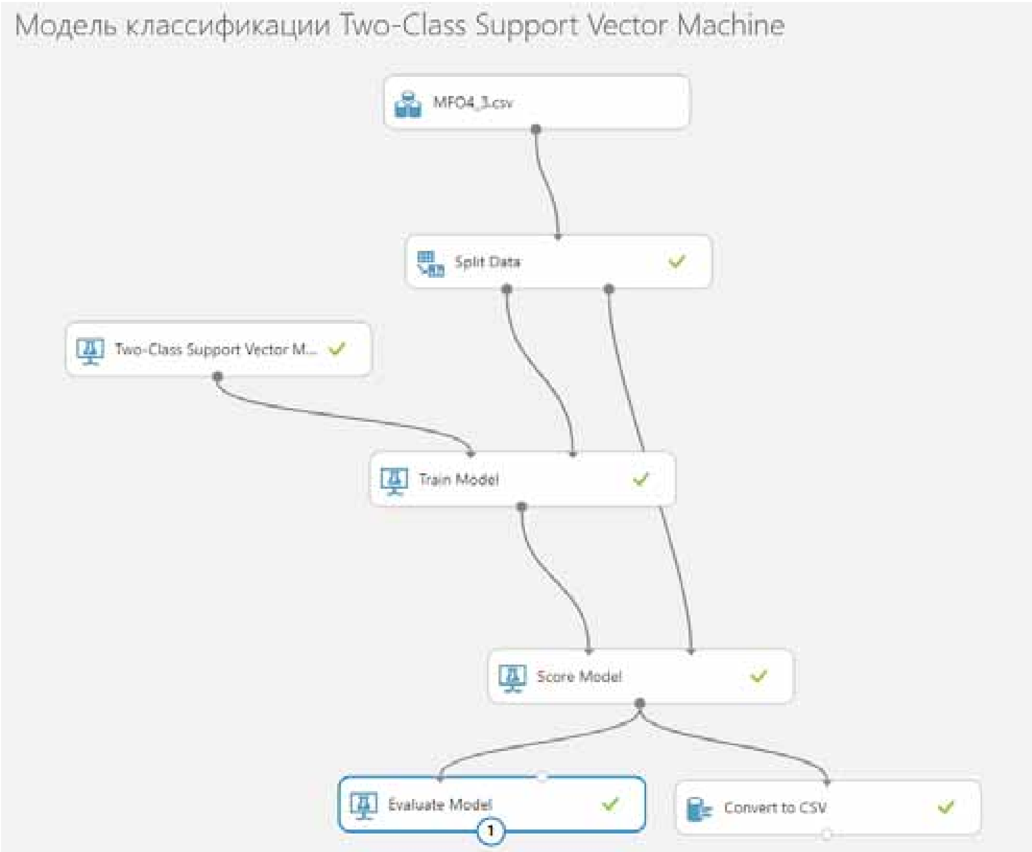

Fig. 4. **Two-Class Support Vector Machine Classification Model**

Source: compiled by the author.

problems, we will consider the two-class neural network algorithm.

Classification using neural networks is a supervised learning method and therefore requires a tagged dataset that includes a column of tags. The *Two-class neural network* algorithm is used to predict binary outcomes, such as whether a patient has a specific disease, whether a machine can fail within a certain period of time, or whether a particular financial monitoring object is deviant.

Bayes networks

Bayesian networks are mainly used for solving diagnostic problems. For example, they are often used in medicine, credit scoring [18–20], and other tasks requiring risk assessment.

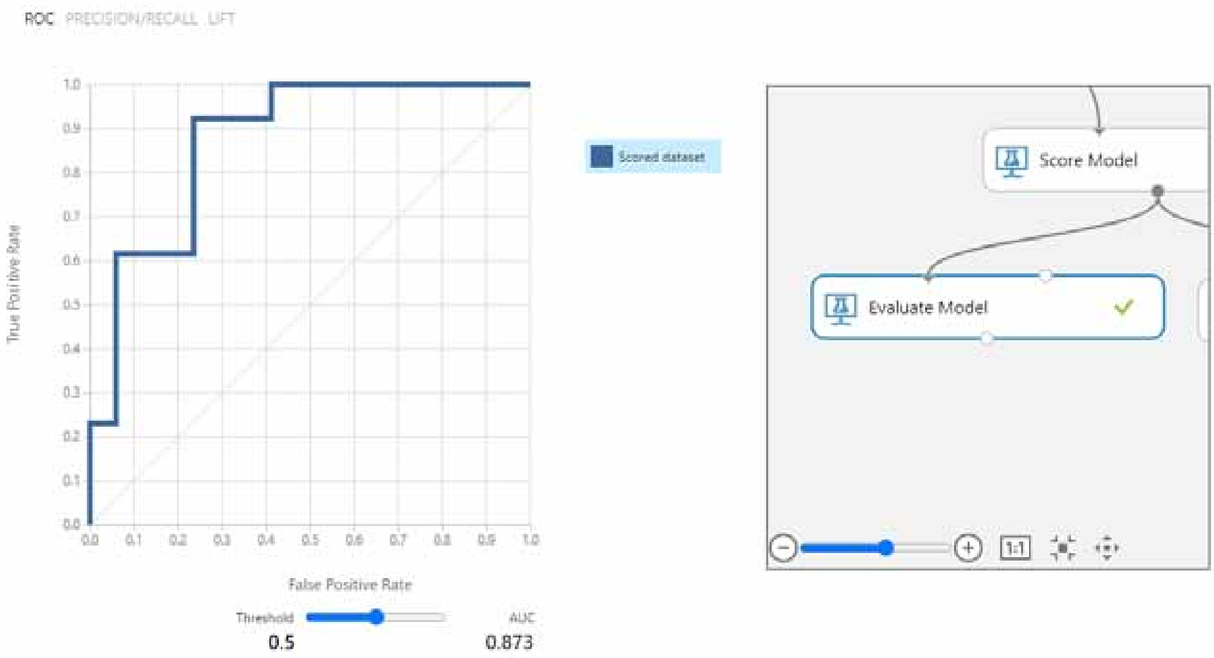

Fig. 5. **ROC curve and accuracy ratios for support vector machine**

Source: compiled by the author.

The *Two-class Bayes Point Machine* algorithm uses a Bayesian approach to linear classification, it effectively approximates the theoretically optimal Bayesian mean for linear classifiers (in terms of generalization efficiency) by choosing one "middle" classifier, the Bayes point. Being a Bayesian classification model, this algorithm cannot be over-trained.

CLASSIFICATION RESULTS

We will demonstrate the classification of MFOs based on prepared data using the *Two-Class Logistic Regression* algorithm. *Fig. 2* shows the machine learning diagram.

From the considered classification algorithms the most accurate results were shown by the *Two-Class Support Vector Machine* algorithm, the AUC ratio was 0.873 (*Fig. 4, 5*).

CONCLUSIONS

The microfinance sector in Russia is progressing and therefore is of high social significance. Short-term loans can be really helpful for citizens to quickly restore the financial balance.

However, there is a high risk of microfinance organizations to be involved in illegal financial transactions, fraud and money laundering.

Automating detection of unscrupulous microfinance organizations based on machine learning methods will allow for promptly identifying and suppressing illegal activities by regulatory authorities.

Comparative analysis of processing data on the MFOs activities by classification methods — logistic regression, decision trees (two-class decision forest algorithms, *Adaboost*), support vector machine (two-class support vector machine algorithm), neural network methods (two-class neural network algorithm), Bayesian networks (algorithm two-class Bayes network) — showed that the two-class support vector machine demonstrates the most accurate results. The research results may be of use to the Bank of Russia and Rosfinmonitoring to automate detecting unscrupulous microfinance organizations.

REFERENCES

1. Sorokin A. S., Shilov V. A. Multivariate statistical analysis of the structure of the microfinance market in Russia. *Internet-zhurnal Naukovedenie.* 2016;8(1):10. (In Russ.). DOI: 10.15862/10EVN 116
2. Makarova E. B. Features of microfinancing in Russia. *Vestnik Volgogradskogo gosudarstvennogo universiteta. Seriya 3: Ekonomika. Ekologiya = Science Journal of VolSU. Global Economic System*. 2017;19(4):80–86. (In Russ.). DOI: 10.15688/jvolsu3.2017.4.9
3. Balashev N. B., Barkinkhoeva M. Kh. Development trends of the microfinance market in the Russian Federation. *Ekonomika i biznes: teoriya i praktika = Economy and Business: Theory and Practice*. 2019;(10–1):27–31. (In Russ.). DOI: 10.24411/2411–0450–2019–11207
4. Tsvetkov V. A., Dudin M. N., Saifieva S. N. Problems and prospects for the development of microfinance organizations in the Russian Federation. *Finansy: teoriya i praktika = Finance: Theory and Practice*. 2019;23(3):96–111. (In Russ.). DOI: 10.26794/2587–5671–2019–23–3–96–111
5. Ordynskaya M. E., Silina T. A. Availability of microfinance services for small businesses (based on materials from the Republic of Adygea). *Vestnik Adygeiskogo gosudarstvennogo universiteta. Seriya 5: Ekonomika = Bulletin of the Adyghe State University. Series: Economics.* 2018;(3):213–224. (In Russ.).
6. Chernykh S. I. Microfinance organizations in the domestic financial and credit system: Problems of development. *Vestnik Instituta ekonomiki Rossiiskoi akademii nauk = Bulletin of the Institute of Economics of the Russian Academy of Sciences*. 2017;(2):139–146. (In Russ.).
7. Ershova I. V., Tarasenko O. A. Small and medium-sized enterprises: Transformation of the Russian crediting and microfinancing system. *Vestnik permskogo universiteta. Yuridicheskie nauki = Perm University Herald. Juridical Sciences.* 2018;(39):99–124. (In Russ.). DOI: 10.17072/1995–4190–2018–39–99–124
8. Evlakhova Yu. S. Russian microfinance organizations: Dynamics of development and the problem of involvement in illegal financial transactions. *Finansy i kredit = Finance and Credit*. 2018;24(7):1637–1648. (In Russ.). DOI: 10.24891/fc.24.7.1637
9. Volkov S. E., Loskutov I. N. Fraud in microfinance sphere. Moscow: Russian Academy of Natural Sciences; 2016. 20 p. (In Russ.).
10. Pavlidis N. G., Tasoulis D. K., Adams N. M., Hand D. J. Adaptive consumer credit classification. *Journal of the Operational Research Society*. 2012;63(12):1645–1654. DOI: 10.1057/jors.2012.15
11. Yap B. W., Ong S. H., Husain N. H.M. Using data mining to improve assessment of credit worthiness via credit scoring models. *Expert Systems with Applications*. 2011;38(10):13274–13283. DOI: 10.1016/j.eswa.2011.04.147
12. Khemais Z., Nesrine D., Mohamed M. Credit scoring and default risk prediction: A comparative study between discriminant analysis and logistic regression. *International Journal of Economics and Finance*. 2016;8(4):39. DOI: 10.5539/ijef.v8n4p39
13. Li Z., Tian Y., Li K., Yang W. Reject inference in credit scoring using support vector machines. *SSRN Electronic Journal*. 2016. DOI: 10.2139/ssrn.2740856
14. Louzada F., Anacleto-Junior O., Candolo C., Mazucheli J. Poly-bagging predictors for classification modelling for credit scoring. *Expert Systems with Applications*. 2011;38(10):12717–12720. DOI: 10.1016/j.eswa.2011.04.059
15. Vukovic S., Delibasic B., Uzelac A., Suknovic M. A case-based reasoning model that uses preference theory functions for credit scoring. *Expert Systems with Applications*. 2012;39(9):8389–8395. DOI: 10.1016/j.eswa.2012.01.181
16. Marqués A. I., García V., Sánchez J. S. Two-level classifier ensembles for credit risk assessment. *Expert Systems with Applications*. 2012;39(12):10916–10922. DOI: 10.1016/j.eswa.2012.03.033
17. Akkoç S. An empirical comparison of conventional techniques, neural networks and the three-stage hybrid Adaptive Neuro Fuzzy Inference System (ANFIS) model for credit scoring analysis: The case of Turkish credit card data. *European Journal of Operational Research*. 2012;222(1):168–178. DOI: 10.1016/j.ejor.2012.04.009

18. Wu W.-W. Improving classification accuracy and causal knowledge for better credit decisions. *International Journal of Neural Systems*. 2011;21(4):297–309. DOI: 10.1142/S 0129065711002845
19. Zhu H., Beling P. A., Overstreet G. A. A Bayesian framework for the combination of classifier outputs. *Journal of the Operational Research Society*. 2002;53(7):719–727. DOI: 10.1057/palgrave.jors.2601262
20. Eskindarov M. A., Solov'eva V.I., eds. Paradigms of the digital economy: Artificial intelligence technologies in finance and fintech. Moscow: Cogito-Center; 2019. 325 p. (In Russ.).

ABOUT THE AUTHOR

Yuliya M. Beketnova — Cand. Sci. (Eng.), Assoc. Prof., Faculty of Applied Mathematics and Information Technology, Financial University, Moscow, Russia
beketnova@mail.ru

The article was submitted on 10.08.2020; revised on 24.08.2020 and accepted for publication on 12.09.2020.
The author read and approved the final version of the manuscript.