Economic Crime Detection Using Support Vector Machine Classification

Andriy Krysovatyy, Hrystyna Lipyanina-Goncharenko, Svitlana Sachenko and Oksana Desyatnyuk

West Ukrainian National University, Lvivska Str., 11, Ternopil, 46000, Ukraine

Abstract.
Fictitious business – is the creation or acquisition of business entities in order to cover up illegal activities or activities that are prohibited. Investigation of economic crime takes a lot of time for law enforcement officers, so in this regard, the development of an algorithm for detecting a fictitious enterprise based on the classic method of machine learning, namely Support Vector Machine Classification, will develop a single software environment for rapid detection of economic crimes. To build the method, data from 1,100 companies operating in Ukraine were used. The data presented in in the set logical binary values are from 355 fictitious enterprises. Modeling of the Support Vector Machine was performed by 3 approaches: linear, polynomial and radial. The best results are obtained from classification by polynomial approach. The training sample showed evaluation results at 100%, and the test sample showed evaluation at 99.7%. Also, the confusion matrix showed quite good results.

Keywords 1
fictitious enterprises, business entities, classification, machine learning, Support Vector Machine Classification.

1. Introduction

A fictitious enterprise in Ukraine should be understood the following: a business entity that is registered in violation of the established procedure (legal norms) of registration with state bodies, the constituent documents of which do not comply with applicable law, or to carry out activities contrary to law or constituent documents, or violation of the procedure for tax accounting and deadlines for filing tax returns and financial statements, or violation of the deadlines for submission of information to government agencies about the change of name, organizational form, form of ownership and location [1].

The main reasons for the emergence and existence of economic crime and fictitious entrepreneurship are: imperfection of legislation governing economic activity, high levels of corruption, bondage of taxes, control of corrupt individuals in major industries, low professional level of law enforcement officers in detecting, documenting, investigating these crimes.

However, the investigation of economic crime often takes a lot of time for law enforcement officers, so in this regard, the development of an algorithm for detecting a fictitious enterprise based on the classical method of machine learning, namely Support Vector Machine Classification, will develop a single software environment that is one of the most promising areas for the rapid detection of economic crimes.

This article is devoted to this topic, the rest of which is distributed as follows. Section 2 discusses the analysis of related work; section 3 presents the algorithm for detecting a fictitious enterprise based

MoMLeT+DS 2021: 3rd International Workshop on Modern Machine Learning Technologies and Data Science, June 5, 2021, Lviv-Shatsk, Ukraine
EMAIL: rektor@wunu.edu.ua (A. Krysovatyy); xrustya.com@gmail.com (H. Lipyanina-Goncharenko); s_sachenko@yahoo.com (S. Sachenko); o.desyatnyuk@wunu.edu.ua (O. Desyatnyuk)
ORCID: 0000-0002-5850-8224 (A. Krysovatyy); 0000-0002-2441-6292 (H. Lipyanina-Goncharenko); 0000-0001-8225-1820 (S. Sachenko); 0000-0002-1384-4240 (O. Desyatnyuk)

on Support Vector Machine Classification; in section 4 the implementation of the algorithm itself. Section 5 presents the conclusions of the study.

2. Related work

Paper [3] forms the basis for the analysis of the interaction of legal and illegal entities: the definition of the terms “international crime” and “cross-border crime”; separates for analytical purposes “enterprise crime” from “political crime”. Organized crime focuses on the “model” of the enterprise, which is really [4] an approach to the study of organized crime, based on the idea that legal and illegal business are quite similar. Based on the paradigm of organized crime, paper [5] presents the concept of criminal entrepreneurship, the work developed a research model that links the degree of entrepreneurship with the stage of organized crime growth. The article [6] has the application of known theories of organized crime, such as agency theory, alliance theory, network theory, resource-based theory and other organizational theories to shed light on criminal organizations. Theoretical and applied aspects of the impact of criminal activity on the economic security of business structures have been studied [7]; the structural and dynamic tendency of the main threats to the economic security of enterprises on the basis of the application of integrated assessment depending on partial indicators is studied; the peculiarities of the multi-vector mechanism of counteracting crime with the help of economic security of enterprises are substantiated as the base of measures set choosing.

The study [8] considered the measure of corporate tax evasion, which reduces both financial and taxable income, which is called tax evasion. Simulation analyzes, transformations of US LIFO / FIFO inventory methods, and samples from private and public firms are used to confirm the results. The results indicate that the measure of compliance with tax avoidance successfully records transactions that comply with the accounting tax. As expected, it was found that the degree of participation of state-owned firms in compliance with tax evasion systematically changes depending on the pressure on the capital market.

It has been studied [2] as illegal mining, which is very common in Colombia, to overcome the problem of measuring illegal activity, a new data set has been built using machine learning methods on the satellite images functions. The study [9] listed opportunities in the virtual environment through ICT, such as FinTrack software, etc. to prevent financial crime. Face recognition plays a central role in many security programs [10], which are used to establish a huge system of social credit to control the entire population.

A generalized structure of a high-performance adaptive system for detecting cyberattacks based on neural networks and artificial immune systems has been developed [15, 28]. Various data from machine learning technologies applied to crime data to monitor the impact of the economic crisis on crime in India have been investigated in [16]. Possibilities of detailed study of data from huge repositories, analysis of various socio-economic factors related to crime incidents, detection of deviations, classification of patterns and development of effective computational models for crime prediction using data analysis and machine learning have been studied [17]. The proposed [18] system is tested for the problem of predicting crimes using data, and experimental results show that the proposed system provides better results and search for possible solutions and patterns of crime. The experiment [19-21] collected data on the crime scenario from the police, which were then simulated on the data set using machine learning algorithms to predict some attributes.

The researches [11, 23, 24] present a modern, structured and well-organized review of one-class support classifiers. The study [12] used the methods of vector support machines (SVM) and neural networks, where the SVM model received the highest efficiency among the classifiers for each data set. Research [13] provides an understanding of the use of stochastic gradient descent algorithms for large data applications, for example, to accelerate SVM or controlled regression on a large scale, or to increase the effectiveness of online learning or real-time forecasting (control). [22, 29] proposed an approach to machine learning for crime detection and crime location, through Twitter posts and vector-based filtering to eliminate noise.

It should be noted that the above-mentioned works do not describe the detection of fictitious enterprises with the help of information technology. Accordingly, the aim of this article is to develop a method of detecting a fictitious enterprise based on Support Vector Machine Classification, which will

further develop a software environment for public sector employees to prevent economic crimes and quickly track fictitious enterprises.

There are several close analogues to the purpose of the study [5, 6, 17], which analyze the applied area, but they do not investigate the use of information technology in this matter and further software development.

3. Materials and methods

The above-described factors in the study lead to mass shadowing of the Ukrainian economy [30]. Accordingly, there is an urgent need to develop special state programs to combat crime in the economic sphere (public procurement, use of state budget funds, taxation of economic entities, works, services at the expense of the state budget, etc.), certain areas of the national economy.

In order to properly organize the activities of law enforcement agencies to detect, promptly verify and investigate crimes committed using the capabilities of economic entities with signs of fictitiousness, it is necessary to determine their characteristics and features of the content of criminal activity and types of such entities. In such case the developed method may be used (Fig. 1.) to identify a fictitious enterprise on the basis of Support Vector Machine Classification, which will allow to quickly make decisions about the performance of an individual enterprise.

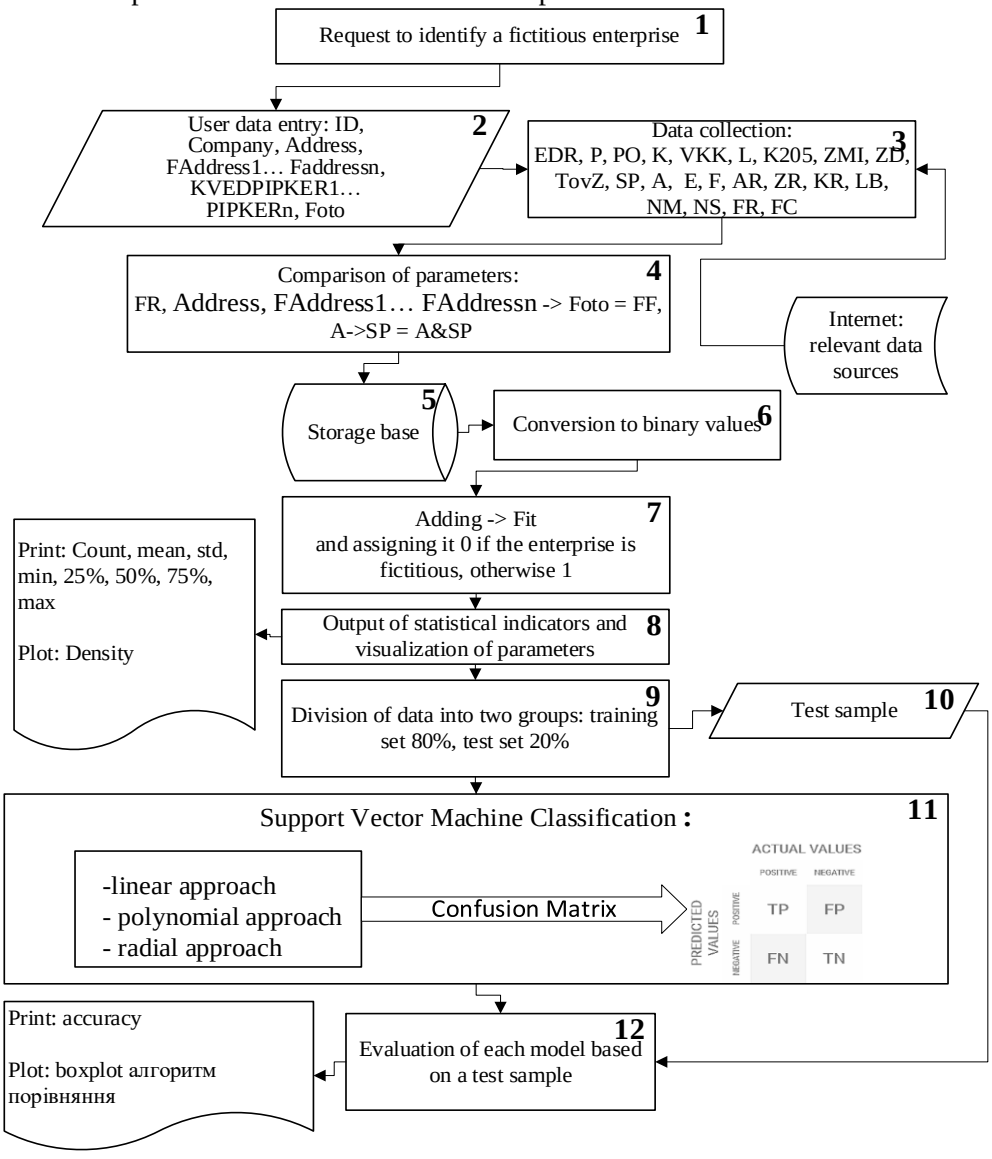

Figure 1: Schematic representation of the fictitious enterprise detection method based on Support Vector Machine Classification

At the first stage (block 1) the user needs to submit a request to identify a fictitious enterprise.

To build a method based on Support Vector Machine Classification, it is first needed to obtain a set of data on which the Support Vector Machine Classification algorithm will be formed. In future research it is planned to develop ready-made software. Accordingly, the required data to be entered (block 2) directly by the user into the system (Table 1): company code (generated automatically in the system), legal address, physical address, codes of economic activities, names of managers (maybe several), photos of equipment with geolocation. All these parameters can be supplemented with new ones and generalize the existing ones, the algorithm is easily adapted.

Based on the data entered by the user, the analysis of the following parameters (Table 2) (block 3) and search for appropriate values from sources of information created using the appropriate API and photo processing method of pattern recognition, which is the purpose of further research.

Table 1.

User-entered data

Parameter	Name	Data type
ID	Company code	int64
Company	The company name	object
Address	Legal address	object
FAddress1,… FAddressn	Physical address	object
KVED	Code of economic activity	object
PIPKER1,… PIPKERn	Surnames of managers	object
Foto	Photo of equipment with geolocation	object.jpg

Table 2.

Data collection on open data of Ukraine

Para-meter	Explanation	Data type	Data source
ID	Company code	int64	
fit	Determining the fictitiousness of the enterprise.	bool	
EDR	Availability of a register of legal entities and individuals in a single database	bool	https://usr.minjust.gov.ua/
P	Availability of VAT, SSC and a single tax in the database	bool	https://cabinet.sfs.gov.ua/cabinet/faces/public/reestr.jspx
PO	Carrying out timely payment of taxes	bool	https://cabinet.sfs.gov.ua/cabinet/faces/public/reestr.jspx
K	Availability of settlements with co-agents	bool	https://youcontrol.com.ua/landing_002/
VKK	Information on the presence of company executives in the state register of declarations	bool	https://public.nazk.gov.ua/
L	Availability of licenses in accordance with the NACE	bool	http://irc.gov.ua/ua/Poshuk-v-YeLR.html
K205	The presence of criminal cases under Art. 205 of the Criminal Code of Ukraine	bool	http://www.reyestr.court.gov.ua/
ZMI	Presence of mentions of company executives with keywords: criminal case, corruption, offshore accounts, etc.	bool	http://dzmi.informjust.ua/
ZD	Availability of land at the legal or physical address	bool	http://map.land.gov.ua/kadastrova-karta
TovZ	Availability of registered trademarks and services, database of industrial marks, database of inventions and other databases of the Institute of Industrial Property of Ukraine	bool	http://www.uipv.org/ua/bases2.html
SP	Availability of issued motor third party insurance policies, MTIBU policy check, motor third party database, search by state car number, check of the	bool	https://mail.mtibu.kiev.ua/Login.aspx?ReturnUrl=/Cbd/MTSBU_Pages/Tree.aspx

Para-meter	Explanation	Data type	Data source
	status of the Green Card policy for cars owned by the company		
A	The presence of cars and their owners issued to the company.	bool	https://igov.org.ua/service/1397/general
A&SP	Coincidence of registered cars with insurance policies	bool	
E	Availability in the database of exporters	bool	http://ukrexport.gov.ua/rus/ukr_export_exporters/?country=ukr
F	Availability in the stock market database	bool	http://www.nssmc.gov.ua/fund/registers
AR	The presence of cars and their owners registered for the company is wanted	bool	http://wanted.mvs.gov.ua/searchtransport/
ZR	The presence of weapons of the owners of the company is wanted	bool	http://wanted.mvs.gov.ua/searchorj/
KR	The presence of cultural values of the owners of the company is wanted	bool	http://wanted.mvs.gov.ua/searchart/
LB	Availability of building licenses in the company	bool	http://dabi.gov.ua/license/list.php
NM	The presence of real estate in the company	bool	https://kap.minjust.gov.ua/services?keywords=&product_id=1&usertype
NS	Availability of the company's website	bool	www.google.com
FR	Availability of equipment, recognition of equipment by the available photo and determination of compliance of geolocation with the production address	bool	Photo of equipment with geolocation
FC	Availability of the company's social networks and affiliates	bool	Facebook.com

Next, a comparison of data (block 4) with each other, namely: FR with Foto, to determine whether the geolocation coincides with one of the addresses of the company and whether the photo shows the relevant equipment; SP with A to check whether the registered car companies match the insurance policies.

Then all the data are transferred to the storage database (block 5) after which the data is converted into binary values (block 6).

Add the fit parameter (block 7), which shows the value relative to whether the company is fictitious or not. Of course, machine learning algorithms operate on numerical values, so we assign the corresponding discrete values 0 or 1.

It is needed always display the basic statistical characteristics of each numerical feature to verify that all data are displayed correctly. Accordingly, we display (block 8): the number of values, the average value, the minimal and maximal values. The std line shows the standard deviation (which measures how scattered the values are). 25%, 50% and 75% of the rows show the corresponding percentages of values in the corresponding parameter. With the correct visualization of the data, it is clear trends and patterns, the ratio of variables, which allows very well notice the trends in the figures. Therefore, graphs of data density are also displayed.

Before data investigation based on machine learning, one of the most important steps should be to distribute the data (block 9). The data are divided into two groups: training set 80%, test set 20%. The training kit should be used to build machine learning models. The test kit (block 10) should be used to see how well the model works on unfamiliar data.

Next, we classify the data using the method of the classical method of machine learning Support Vector Machine Classification (block 11) [25, 26]. For the purpose of fictitious enterprises selection, it is necessary to solve the problem of binary classification. Initially, the algorithm is trained on objects (block 10) from the training sample, for which the class labels are known in advance. Next, the already learned algorithm predicts the class label for each object from the deferred test sample. Class labels can take the values $Y = \{-1, +1\}$.Object – is a vector with N signs $x = (x_1, x_2, \dots, x_n)$ in the Rn space.

When learning, the algorithm must construct a simple function $F(x) = y$, where the argument x is an object from the Rn space and gives a label of class i. The main goal of Support Vector Machine (SVM) as a classifier is to find the equation that divides the hyperplane [14] $w_1x_1 + w_2x_2 + \cdots + w_nx_n + w_0 = 0$ in Rn space, which would divide the two classes optimally. General view of the transformation F of an object x into a label of class Y: $F(x) = sign(w^{Tx} - b)$. And denote $w = (w_1, w_2, \dots, w_n), b = -w_0$. After adjusting the weights of the algorithm w and b w and b, all objects that fall on one side of the constructed hyperplane will be defined as the first class, and objects that fall on the other side – the second class. Inside the function $sign()$ is a linear combination of object features with algorithm weights, that is why SVM refers to linear algorithms. The hyperplane partition can be constructed in different ways, but in SVM the weights w and b are adjusted so that the class objects lie as far as possible from the hyperplane distribution. In other words, the algorithm maximizes the gap (Margin) between the hyperplane and the objects of the classes that are closest to it (Fig. 2). Such objects are called reference vectors. If we consider the positive and negative sides of SVM, the positive thing is that: SVM works well with a space of features of large size, with data of small volume; SVM maximizes the separation of the hyperplane and objects, which reduces the number of classification errors; also, the algorithm is reduced to solving the problem of quadratic programming in the 3D domain, which makes it possible to divide the hyperplane with certain hyperparameters of the algorithm. The disadvantages of this method include the fact that: it takes a long time to learn, especially for large data sets; instability to noise.

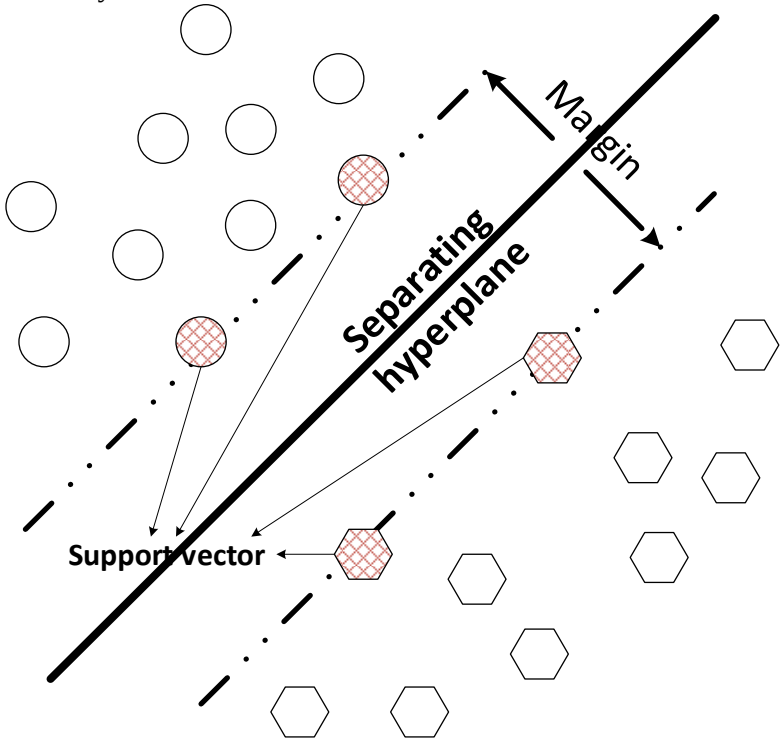

Figure 2: Classification hyperplane of SVM

Support Vector Machine should be carried out in three main approaches: linear, polynomial and radial approaches. Also, it is necessary to display the Confusion Matrix. The confusion matrix is deciphered as follows:

- The target variable has two values: Positive or Negative;
- These columns are the actual values of the target variable;
- These strings represent the predicted values of the target variable.

Next, we evaluate the model (block 12) based on the classical method of machine learning Support Vector Machine, which can be quite complex. Usually, the model is estimated on the basis of the error rate. However, this method is not very reliable, because the accuracy obtained for one test set can be very different from the accuracy obtained for another test set. The key to a fair comparison of machine learning algorithms is to ensure that each algorithm is evaluated equally on the same data.

4. Experimental results and discussion

Python was chosen as programming language, as this programming language works best with machine-based data analysis. The following libraries were used for analysis: pandas, numpy, train_test_split, SVC and cross_val_score.

To build the method, data from 1,100 companies operating in Ukraine were used. The data are presented in logical binary values (Fig. 3.). There are 355 fictitious enterprises in the set (Fig. 4).

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 1100 entries, 0 to 1099
Data columns (total 25 columns):
 #   Column  Non-Null Count  Dtype
---  ------  --------------  -----
 0   ID      1100 non-null   int64
 1   fit     1100 non-null   int64
 2   EDR     1100 non-null   int64
 3   P       1100 non-null   int64
 4   PO      1100 non-null   int64
 5   K       1100 non-null   int64
 6   VKK     1100 non-null   int64
 7   L       1100 non-null   int64
 8   K205    1100 non-null   int64
 9   ZMI     1100 non-null   int64
 10  ZD      1100 non-null   int64
 11  TovZ    1100 non-null   int64
 12  SP      1100 non-null   int64
 13  A       1100 non-null   int64
 14  A&SP    1100 non-null   int64
 15  E       1100 non-null   int64
 16  F       1100 non-null   int64
 17  AR      1100 non-null   int64
 18  ZR      1100 non-null   int64
 19  KR      1100 non-null   int64
 20  LB      1100 non-null   int64
 21  NM      1100 non-null   int64
 22  NS      1100 non-null   int64
 23  FF      1100 non-null   int64
 24  FC      1100 non-null   int64
dtypes: int64(25)
memory usage: 215.0 KB
```

Figure 3: Data structure

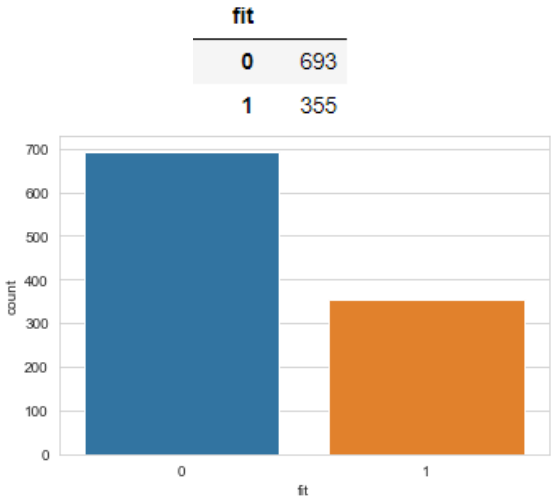

Figure 4: The number of fictitious enterprises in the set

The degree of dependence between the indicators with the definition of a fictitious enterprise is presented in the form of a correlation matrix (Fig. 5). The matrix shows that the fit indicator has a low correlation with other parameters, so it is difficult to clearly determine the dependence on individual parameters of whether the company is conducting economic crime or not. Therefore, it is advisable to use machine learning algorithms that will more accurately determine economic crime. To do this, modeling using the method of classical machine learning Support Vector Machine.

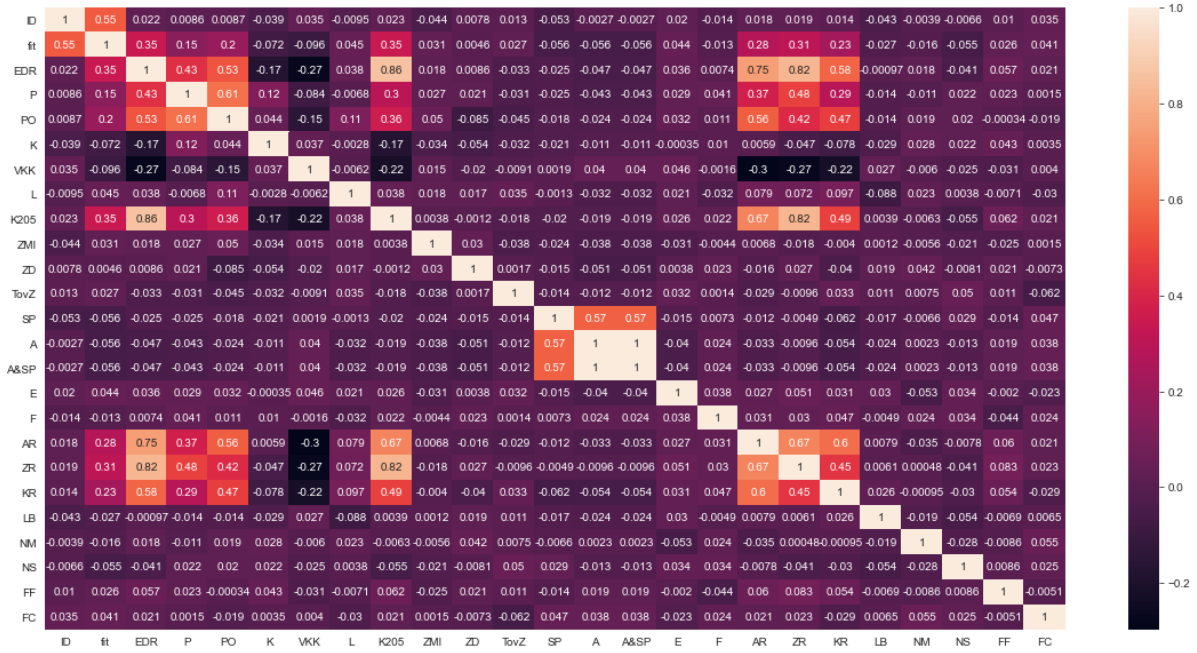

Figure 5: Correlation matrix

Modeling of the Support Vector Machine was performed by 3 approaches: linear (Fig. 6), polynomial (Fig. 7) and radial (Fig. 8). Other approaches also exist, but are used less frequent.

We will first consider the linear approach (Fig. 6), the evaluation of the simulation was performed on a test sample and training. The training sample showed evaluation results at 98.05%, and the test sample showed evaluation at 97.88%. The standard deviation is 98% for both samples.

An important result is the Confusion Matrix. The different values of the Confusion matrix will be as follows for the training sample:

- True positive (TP) = 143; this means that 143 indicators of positive class data are correctly classified by the model;
- True negative (TN) = 612; this means that 612 data points of negative class were correctly classified by the model;
- False positive (FP) = 15; this means that 15 indicators of negative class data were incorrectly classified as models belonging to the positive class;
- False negative (FN) = 0; this means that 0 data indicators of the positive class were incorrectly classified as models belonging to the negative class.

For the test calibration, the confusion matrix shows that 158 and 268 indicators of correctly defined positive and negative class, respectively, 7 indicators of incorrectly defined negative class and 0 indicators of data of positive class were incorrect.

```
Train Result:
================================================
Accuracy Score: 98.05%
_________________________________________________
CLASSIFICATION REPORT:
                    0           1  accuracy   macro avg  weighted avg
precision    1.000000    0.976077  0.980519    0.988038      0.980986
recall       0.905063    1.000000  0.980519    0.952532      0.980519
f1-score     0.950166    0.987893  0.980519    0.969030      0.980152
support    158.000000  612.000000  0.980519  770.000000    770.000000
_________________________________________________
Confusion Matrix:
 [[143  15]
 [  0 612]]

Test Result:
================================================
Accuracy Score: 97.88%
_________________________________________________
CLASSIFICATION REPORT:
                    0           1  accuracy   macro avg  weighted avg
precision    1.000000    0.974545  0.978788    0.987273      0.979328
recall       0.887097    1.000000  0.978788    0.943548      0.978788
f1-score     0.940171    0.987109  0.978788    0.963640      0.978290
support     62.000000  268.000000  0.978788  330.000000    330.000000
_________________________________________________
Confusion Matrix:
 [[ 55   7]
 [  0 268]]
```

Figure 6: Simulation Support Vector Machine linear approach

Next, consider the polynomial approach (Fig. 7). The training sample showed evaluation results at 100%, and the test sample showed evaluation at 99.7%. The standard deviation is close to 100% for both samples. For the training calibration, the confusion matrix shows that 158 and 612 indicators of correctly defined positive and negative class, respectively, and 0 indicators are incorrectly defined. In the test sample, the confusion matrix showed that 61 and 268 indicators were defined correctly in the positive and negative classes, respectively, and 1 indicator was determined incorrectly.

The last approach is radial (Fig. 8), which showed not very good results. The training sample showed evaluation results at 79.8%, and the test sample showed evaluation at 81.21%. The standard deviation is close to 80% for both samples. For the training calibration, the confusion matrix shows that 158 indicators are incorrectly defined, 612 indicators are correctly defined and 0 indicators are defined correctly and incorrectly. In the test sample, the confusion matrix showed that 268 indicators were defined correctly, but 62 indicators were determined incorrectly.

```
Train Result:
================================================
Accuracy Score: 100.00%
________________________________________________
CLASSIFICATION REPORT:
                0      1  accuracy  macro avg  weighted avg
precision     1.0    1.0       1.0        1.0           1.0
recall        1.0    1.0       1.0        1.0           1.0
f1-score      1.0    1.0       1.0        1.0           1.0
support     158.0  612.0       1.0      770.0         770.0
________________________________________________
Confusion Matrix:
 [[158   0]
 [  0 612]]

Test Result:
================================================
Accuracy Score: 99.70%
________________________________________________
CLASSIFICATION REPORT:
                   0           1  accuracy   macro avg  weighted avg
precision   1.000000    0.996283   0.99697    0.998141      0.996981
recall      0.983871    1.000000   0.99697    0.991935      0.996970
f1-score    0.991870    0.998138   0.99697    0.995004      0.996960
support    62.000000  268.000000   0.99697  330.000000    330.000000
________________________________________________
Confusion Matrix:
 [[ 61   1]
 [  0 268]]
```

Figure 7: Modeling Support Vector Machine by polynomial approach

```
Train Result:
================================================
Accuracy Score: 79.48%
________________________________________________
CLASSIFICATION REPORT:
                0           1  accuracy   macro avg  weighted avg
precision     0.0    0.794805  0.794805    0.397403      0.631715
recall        0.0    1.000000  0.794805    0.500000      0.794805
f1-score      0.0    0.885673  0.794805    0.442836      0.703937
support     158.0  612.000000  0.794805  770.000000    770.000000
________________________________________________
Confusion Matrix:
 [[  0 158]
 [  0 612]]

Test Result:
================================================
Accuracy Score: 81.21%
________________________________________________
CLASSIFICATION REPORT:
                0           1  accuracy   macro avg  weighted avg
precision     0.0    0.812121  0.812121    0.406061      0.659541
recall        0.0    1.000000  0.812121    0.500000      0.812121
f1-score      0.0    0.896321  0.812121    0.448161      0.727921
support      62.0  268.000000  0.812121  330.000000    330.000000
________________________________________________
Confusion Matrix:
 [[  0  62]
 [  0 268]]
```

Figure 8: Simulation of Support Vector Machine radial approach

When considering 3 approaches of the Support Vector Machine, the best results are obtained when classifying by the polynomial approach. The training sample showed evaluation results at 100%, and the test sample showed evaluation at 99.7%. Also, the confusion matrix showed quite good results.

Thus, returning to the analogues discussed above [5, 6, 17], the developed method makes it possible to develop a single software environment for public sector employees to prevent economic crimes and quickly track fictitious enterprises.

5. Conclusions

A method of detecting a fictitious enterprise based on the Support Vector Machine is proposed, which allows to quickly track fictitious enterprises, which is useful for public sector employees to prevent economic crimes.

The developed method is implemented on the basis of data of 1100 companies that conducted economic activity in Ukraine and 355 of which were defined as fictitious. The results of experimental studies showed that when considering 3 approaches of the Support Vector Machine, the best results were obtained when classifying the polynomial approach, where the training sample showed evaluation results at 100%, and the test sample showed evaluation at 99.7%. The confusion matrix for the training set shows that 158 and 612 indicators of correctly defined positive and negative class, respectively, and 0 indicators are incorrectly defined. In the test sample, the confusion matrix showed that 61 and 268 indicators were defined correctly in the positive and negative classes, respectively, and 1 indicator was determined incorrectly.

In further scientific research it is planned to conduct a more detailed analysis of the detection of fictitious enterprises based on the methods: Gaussian Naive Bayes, Logistic Regression.

6. References

[1] Yu. Piliukov, ІО. Fictitious entrepreneurship in Ukraine. The concept and connection with other economic crimes. Actual problems of jurisprudence, 1 (2019): 140-144. (in Ukrainian)

[2] S. Saavedra, & M. Romero, Local incentives and national tax evasion: The response of illegal mining to a tax reform in Colombia, SIEPR Discussion Paper No. 17-009, Stanford Institute for Economic Policy Research, (2017). https://siepr.stanford.edu/sites/default/files/publications/17-009_0.pdf.

[3] N. Passas, Cross-border crime and the interface between legal and illegal actors, Security Journal 16 (2003) 19–37. doi:10.1057/palgrave.sj.8340123.

[4] D. Liddick, The enterprise "model" of organized crime: Assessing theoretical propositions, Justice Quarterly 16 (1999) 403–430. doi:10.1080/07418829900094191.

[5] P. Gottschalk, Illegal entrepreneurialism as determinant of organised business crime maturity, International Journal of Business and Systems Research 3 (2009) 297. doi: 10.1504/ijbsr.2009.026185.

[6] P. Gottschalk, How criminal organisations work: Some theoretical perspectives, The Police Journal: Theory, Practice and Principles 81 (2008) 46–61. doi: 10.1350/pojo.2008.81.1.400.

[7] O. Vivchar, The influence of crime activity on economic security of enterprise structures in post-conflict conditions: identification of threats and mechanisms of antiaction, Actual Problems of Law, 1 (2019) 113-119. (in Ukrainian) doi: 10.35774/app2019.01.113.

[8] B. Badertscher, S. Katz, S. O. Rego, & R. J. Wilson, Conforming tax avoidance and capital market pressure, The Accounting Review 94 (2019) 1-30. doi: 10.2308/accr-52359.

[9] K. Olowu, & P. Gabasa, Financial crime, ICT & E-governance: Libraries role, Advances in Social Sciences Research Journal 7 (2020) 609-612. https://doi.org/10.14738/assrj.71.7476.

[10] R.T. Kreutzer, M. Sirrenberg, Fields of application of artificial intelligence – Security sector and military sector, in: Understanding Artificial Intelligence. Management for Professionals, Springer, Cham, 2020, pp. 225-233. https://doi.org/10.1007/978-3-030-25271-7_9.

[11] S. Alam, S. K. Sonbhadra, S. Agarwal, & P. Nagabhushan, One-class support vector classifiers: A survey, Knowledge-Based Systems, (2020) 105754. doi: 10.1016/j.knosys.2020.105754.

[12] N. L. Costa, L. A. G. Llobodanin, I. A. Castro, & R. Barbosa, Using support vector machines and neural networks to classify Merlot wines from South America, Information Processing in Agriculture 6 (2018) 265-278. doi: 10.1016/j.inpa.2018.10.003.

[13] W. He, & Y. Liu, To regularize or not: Revisiting SGD with simple algorithms and experimental studies, Expert Systems with Applications 112 (2018) 1–14. doi: 10.1016/j.eswa.2018.06.026.

[14] A. Blumer, A. Ehrenfeucht, D. Haussler, & M. K. Warmuth, Learnability and the Vapnik-Chervonenkis dimension, Journal of the ACM 36 (1989) 929–965. doi: 10.1145/76359.76371.

[15] Komar, M., Golovko, V., Sachenko, A., Bezobrazov, S. Intelligent system for detection of networking intrusion. Proceedings of the 6th IEEE International Conference on Intelligent Data Acquisition and Advanced Computing Systems: Technology and Applications, IDAACS'2011, pp. 374-377

[16] Komar, M., Kochan, V., Dubchak, L., Sachenko, A., Golovko, V., Bezobrazov, S., Romanets, I. High performance adaptive system for cyber attacks detection, in: Proceedings of the 2017 9th

IEEE International Conference on Intelligent Data Acquisition and Advanced Computing Systems: Technology and Applications (IDAACS) (2017) 853-858. doi: 10.1109/IDAACS.2017.8095208.
[17] M. Mittal, L.M. Goyal, J.K. Sethi, et al., Monitoring the impact of economic crisis on crime in India using machine learning, Comput Econ 53 (2019) 1467–1485. https://doi.org/10.1007/s10614-018-9821-x.
[18] P. Saravanan, J. Selvaprabu, L. Arun Raj, A. Abdul Azeez Khan, K. Javubar Sathick, Survey on crime analysis and prediction using data mining and machine learning techniques, in: N. Zhou, S. Hemamalini (Eds.), Advances in Smart Grid Technology, volume 688 of Lecture Notes in Electrical Engineering, Springer, Singapore, (2021) 435-448. doi: 10.1007/978-981-15-7241-8_31.
[19] R. Kumar, B. Nagpal, Analysis and prediction of crime patterns using big data, Int. J. Inf. Tecnol. 11 (2019) 799–805. https://doi.org/10.1007/s41870-018-0260-7.
[20] A. A. Biswas and S. Basak, Forecasting the trends and patterns of crime in Bangladesh using machine learning model, in: Proceedings of the 2019 2nd IEEE International Conference on Intelligent Communication and Computational Techniques (ICCT) (2019) 114-118. doi: 10.1109/ICCT46177.2019.8969031.
[21] L. G. Alves, H. V. Ribeiro, & F. A. Rodrigues, Crime prediction through urban metrics and statistical learning, Physica A: Statistical Mechanics and its Applications 505 (2018) 435-443. https://doi.org/10.1016/j.physa.2018.03.084
[22] S. Kim, P. Joshi, P. S. Kalsi and P. Taheri, Crime analysis through machine learning, in: Proceedings of the 2018 IEEE 9th Annual Information Technology, Electronics and Mobile Communication Conference (IEMCON) (2018) 415-420. doi: 10.1109/IEMCON.2018.8614828.
[23] S. P. C. W. Sandagiri, B. T. G. S. Kumara and B. Kuhaneswaran, Detecting crimes related Twitter posts using SVM based two stages filtering, in: Proceedings of the 2020 IEEE 15th International Conference on Industrial and Information Systems (ICIIS) (2020) 506-510. doi: 10.1109/ICIIS51140.2020.9342698.
[24] H. Lipyanina, V. Maksymovych, A. Sachenko, T. Lendyuk, A. Fomenko, I. Kit, Assessing the investment risk of virtual IT company based on machine learning, in: S. Babichev, D. Peleshko, O. Vynokurova (Eds.), Data Stream Mining & Processing, DSMP 2020, volume 1158 of Communications in Computer and Information Science, Springer, Cham, (2020) 167-187. doi: 10.1007/978-3-030-61656-4_11.
[25] R. Gramyak, H. Lipyanina-Goncharenko, A. Sachenko, T. Lendyuk, and D. Zahorodnia, Intelligent method of a competitive product choosing based on the emotional feedbacks coloring, in: Proceedings of the 2nd International Workshop on Intelligent Information Technologies & Systems of Information Security with CEUR-WS (IntelITSIS 2021), Khmelnytskyi, Ukraine, March 24–26 2853 (2021) 346-357. http://ceur-ws.org/Vol-2853/paper31.pdf
[26] Cherrat, E. M., Alaoui, R., & Bouzahir, H. SCORE FUSION OF FINGER VEIN AND FACE FOR HUMAN RECOGNITION BASED ON CONVOLUTIONAL NEURAL NETWORK MODEL. International Journal of Computing, 19(1) (2020) 11-19. https://doi.org/10.47839/ijc.19.1.1688
[27] Jonaitis, D., Raudonis, V., & Lipnickas, A. APPLICATION OF NUMERICAL INTELLIGENCE METHODS FOR THE AUTOMATIC QUALITY GRADING OF AN EMBRYO DEVELOPMENT. International Journal of Computing, 15(3) (2016) 177-183. https://doi.org/10.47839/ijc.15.3.850
[28] Golovko, V., Komar, M., Sachenko, A. Principles of neural network artificial immune system design to detect attacks on computers. Modern Problems of Radio Engineering, Telecommunications and Computer Science - Proceedings of the 10th International Conference, TCSET'2010, 237
Lipyanina, H., Sachenko, A., Lendyuk, T., Nadvynychny, S., Grodskyi, S. Decision tree based targeting model of customer interaction with business page. CEUR Workshop Proceedings 2608 (2020) 1001-1012
[29] Gozhyj, A., Kalinina, I., Vysotska, V., Sachenko, S., Kovalchuk, R. Qualitative and quantitative characteristics analysis for information security risk assessment in e-commerce systems. CEUR Workshop Proceedings, 2762 (2020) 177–190