Machine Learning and Possibility of its Application for Forensic Researches on Firearms

Pavel Giverts * [a], Oleksandr Kolomiitsev ** [b]

* * PhD in Law, Criminalistics and Forensic Science, Israel Police Ballistic Laboratory, Jerusalem, Israel, ORCID: https://orcid.org/0000-0002-4019-5055, e-mail: pavel.giverts@gmail.com
** PhD in Technology, NSC « Hon. Prof. M. S. Bokarius FSI», Kharkiv, Ukraine, ORCID: https://orcid.org/0000-0003-1932-1034, e-mail: sashagun@ukr.net

[a] Methodology, Concept, Writing and administration of original project.
[b] Formal analysis, Writing – original project, Resources, Supervision.

DOI: 10.32353/khrife.2.2024.03 UDC 004.81:159.953.5::343.983
Received: 22.04.2024 / Reviewed: 22.04.2024 / Accepted for print: 26.06.2024 / Available online: 30.06.2024

Role of machine learning in modern society cannot be overestimated, especially given its steadily increasing importance in recent years. Various machine learning models are applied not only in everyday life but also in solving practical, engineering, and scientific challenges. However, the methods and capabilities of machine learning remain relatively obscure to most professionals across various fields of knowledge, including specialists in forensic firearm examination. The article discusses basic tasks that machine learning methods can facilitate and provides examples of tasks faced by forensic experts during firearm examinations that can be solved utilizing these methods. The article purpose is to familiarize professionals in the field of forensic science with basic capabilities of machine learning methods. The research paper demonstrates which specific types of tasks in forensic firearms examination it is advisable to apply these methods to, and what is required for this purpose. In order to achieve this goal, general scientific methods (dialectical, comparative, analysis, synthesis, induction, deduction) and special scientific methods (formal-logical, systemic-structural) have been employed. Specific expertise

This article is translation of the original Ukrainian content, which source is available at the link: *https://khrife-journal.org/index.php/journal* (translated by Andriy Bublikov). The author acknowledges translation as corresponding to the original.

and skills in forensic identification methods (including ballistics, microscopic comparative analysis of traces on bullets and casings, detection and examination of gunshot residue to estimate shooting distance, etc.), principles of machine learning, and mathematical algorithms used to construct various models have been applied.

Keywords: *forensic firearm examination; shooting distance estimation; machine learning; neural network; classification; clustering; regression.*

Research Problem Formulation

During forensic examination of firearms, forensic expert usually faces a variety of identification and diagnostic issues. In order to answer some of them, forensic expert performs quantitative measurements (determines geometric dimensions of the object under research: weapon, bullet, or a cartridge; measures speed of the bullet; determines force that should be applied to the trigger to fire, etc.) certain objective actions, the results of which do not depend on the expert. While solving other problems, for example, finding out identity of the weapon and fired bullets (cartridge), forensic expert makes a decision based on personal experience and knowledge. Therefore, this kind of research is subjective and, theoretically, the obtained result depends precisely on the expert who performs the research. Subjectivity of some types of forensic research, including research on firearms is discussed in the *NRC* [1] and *PCAST* [2] reports, as well as in research papers of I. E. Dror and N. Scurich [3], J. W. Osterburg [4], M. Cuellar with co-authors [5]. In this regard, intensive development of tools has been underway in recent years, which use will make it possible to reduce the level of subjectivity during expert research and increase the reliability of the results obtained. Such tools can be considered systems of automatic comparison

1 Strengthening Forensic Science in the United States: A Path Forward. Washington, D. C., 2009. 348 p. DOI: 10.17226/12589 (date accessed: 02.04.2024).

2 Report to the President's. Forensic Science in Criminal Courts: Ensuring Scientific Validity of Feature-Comparison Methods. Sept 2016. 174 p. URL: https://obamawhitehouse.archives.gov/sites/default/files/microsites/ostp/PCAST/pcast_forensic_science_report_final.pdf (date accessed: 02.04.2024).

3 Dror I. E. The most consistent finding in forensic science is inconsistency. *Journal of Forensic Sciences*. 2023. Vol. 68. Is. 6. Pp. 1851—1855. DOI: 10.1111/1556-4029.15369 (date accessed: 02.04.2024) ; Dror I. E., Scurich N. (Mis) use of scientific measurements in forensic science. *Forensic Science International: Synergy*. 2020. Vol. 2. Pp. 333—338. DOI: 10.1016/j.fsisyn.2020.08.006 (date accessed: 02.04.2024).

4 Osterburg J. W. The Evaluation of Phisical Evidenve in Criminalistics: Subjective or Objective Process? *The Journal of Criminal Law, Criminalogy and Police Science*. 1969. Vol. 60. No. 1. Pp. 97—101. DOI: 10.2307/1141742 (date accessed: 02.04.2024).

5 Cuellar M., Gonzalez C., Dror I. E. Human and machine similarity judgments in forensic firearm comparisons. *Forensic Science International: Synergy*. 2022. Vol. 5. 9 p. DOI: 10.1016/j.fsisyn.2022.100283 (date accessed: 02.04.2024).

of bore traces on bullets, as well as details and parts of weapons on casings, which not only differentiate images of traces by the degree of their similarity (as with the help of Automated Ballistic Identification Systems, hereinafter referred to as ABIS), but provide statistical reasonable conclusions about [6].

In recent years, we have seen quite rapid progress in development and implementation of machine learning methods. Using these methods, they perform a significant number of tasks that until recently computers were not entrusted with. Instead, machine learning methods are not actually used to solve the tasks of forensic examination, in particular, due to the lack of information about possibilities of machine learning. In view of this, leading experts in various fields of forensic science, as well as experts in forensic weapon research do not consider machine learning as a tool for performing the tasks assigned to them by pre-trial investigation bodies.

Analysis of Essential Researches and Publications

Despite rather slow introduction of machine learning methods for solving forensic problems in Ukraine, such work is actively carried out abroad. For example, a group of statistical analysts led by A. Carriquiry published a brief overview of possibilities of applying machine learning methods in forensic science [7]. The article by Ch. Kingston [8], G. Margagliotti and T. Bollé is devoted to a similar topic [9]. Separately, the intensive use of machine learning methods for the forensic fingerprint should be noted [10]. The same meth-

6 Song J., Song H. Reporting likelihood ratio for casework in firearm evidence identification. *Journal of Forensic Sciences*. 2023. Vol. 68. Is. 2. Pp. 399—406. DOI: 10.1111/1556-4029.15186 (date accessed: 02.04.2024) ; Chen Zh., Song J., Soons J. A., Thompson R. M., Zhao X. Pilot study on deformed bullet correlation. *Forensic Science International*. 2020. Vol. 306. Art. 110098. DOI: 10.1016/j.forsciint.2019.110098 (date accessed: 02.04.2024) ; Chen Zh., Chu W., Soons J. A., Thompson R. M., Song J., Zhao X. Fired bullet signature correlation using the Congruent Matching Profile Segments (CMPS) method. *Ibid*. 2019. Vol. 305. Art. 109964. DOI: 10.1016/j.forsciint.2019.109964 (date accessed: 02.04.2024) ; Chu W., Tong M., Song J. Validation Tests for the Congruent Matching Cells (CMC) Method Using Cartridge Cases Fired with Consecutively Manufactured Pistol Slides. *Journal of the Association of Firearms and Toolmarks Examiners*. 2013. Vol. 45. Is. 4. Pp. 361—366. URL: https://www.nist.gov/publications/validation-tests-congruent-matching-cells-cmc-method-using-cartridge-cases-fired (date accessed: 02.04.2024) ; Tong M., Song J., Chu W., Thompson R. M. Fired Cartridge Case Identification Using Optical Images and the Congruent Matching Cells (CMC) Method. *Journal of Research of the National Institute of Standards and Technology*. 2014. Vol. 119. Art. 575. Pp. 575—582. DOI: 10.6028/jres.119.023 (date accessed: 02.04.2024).

7 Carriquiry A., Hofmann H., Tai X. H., VanderPlas S. Machine learning in forensic applications. *Significance*. 2019. Vol. 16. Is. 2. Pp. 29—35. DOI: 10.1111/j.1740-9713.2019.01252.x (date accessed: 02.04.2024).

8 Kingston Ch. Neural Networks in Forensic Science. *Journal of Forensic Sciences*. 1992. Vol. 37. Is. 1. Pp. 252—264. DOI: 10.1520/jfs13232j (date accessed: 02.04.2024).

9 Margagliotti G., Bollé T. Machine learning & forensic science. *Forensic Science International*. 2019. Vol. 298. Pp. 138—139. DOI: 10.1016/j.forsciint.2019.02.045 (date accessed: 02.04.2024).

10 Chugh T., Cao K., Zhou J., Tabassi E., Jain A. K. Latent Fingerprint Value Prediction: Crowd-Based Learning. *IEEE Transactions on Information Forensics and Security*. 2018.

ods are used for handwriting analysis [11]. Machine learning methods are widely used in forensic biological examinations in particular for DNA analysis [12].

Regarding forensic research on firearms, machine learning is most often perceived as a tool for analyzing and comparing images of traces on bullets and cartridges. Thus, researchers from New Zealand are considering the use of the *support vector method* for comparing striker traces [13]. P. Pisantanaroj and co-authors published research results s on the classification of traces on balls by the method of *deep learning* [14]. Machine learning methods are also used to compare the traces of

Vol. 13. Is. 1. Pp. 20—34. DOI: 10.1109/TIFS.2017.2721099 (date accessed: 02.04.2024) ; Stojanović B., Marques O., Nešković A., Puzović S. Fingerprint ROI segmentation based on deep learning. *24th Telecommunications Forum (TELFOR).* IEEE (Belgrade, Serbia,22—23 Nov 2016). 2016. Pp. 1—4. DOI: 10.1109/TELFOR.2016.7818799 (date accessed: 02.04.2024) ; Yih Ch.-H., Hung J.-L., Wu J.-A., Chen L.-M. Overlapped Fingerprint Separation Based on Deep Learning. *International Conference on Digital Medicine and Image Processing* (Okinawa, Japan, Nov 12—14, 2018). 2018. Pp. 14—18. DOI: 10.1145/3299852.3299857 (date accessed: 02.04.2024) ; Sharma A., Kaur M. Automatic segmentation for separation of overlapped latent fingerprints. *International Journal of Computer Sciences and Engineering*. 2018. Vol. 6. Is. 7. Pp. 484—490. DOI: 10.26438/ijcse/v6i7.484490 (date accessed: 02.04.2024) ; Wong W. J., Lai Sh.-H. Multi-task CNN for restoring corrupted fingerprint images. *Pattern Recognition.* 2020. Vol. 101. Art. 107203. DOI: 10.1016/j.patcog.2020.107203 (date accessed: 02.04.2024) ; Ezeobiejesi J. C. Deep Learning for the Analysis of Latent Fingerprint Images. University of California Riverside, 2019. 168 p.URL: https://escholarship.org/uc/item/564185wc (date accessed: 02.04.2024).

11 Lemos N., Shah K., Rade R., Shah D. Personality Prediction based on Handwriting using Machine Learning. *International Conference on Computational Techniques, Electronics and Mechanical Systems (CTEMS).* IEEE (Belgaum, India, 21—22 Dec 2018). 2018. Pp. 110—113. DOI: 10.1109/CTEMS.2018.8769221 (date accessed: 02.04.2024) ; Wang S., Jia Sh. Signature handwriting identification based on generative adversarial networks. *Journal of Physics: Conference Series.* 2019. Vol. 1187. Is. 4. Art. 042047. DOI: 10.1088/1742-6596/1187/4/042047 (date accessed: 02.04.2024) ; Durou A., Al-Maadeed S., Aref I., Bouridane Ah., Elbendak M. A Comparative Study of Machine Learning Approaches for Handwriter Identification. *12th International Conference on Global Security, Safety and Sustainability (ICGS3).* IEEE (London, UK, 16—18 Jan 2019). 2019. Pp. 206—212. DOI: 10.1109/ICGS3.2019.8688032 (date accessed: 02.04.2024).

12 Marciano M. A., Adelman J. D. PACE: Probabilistic Assessment for Contributor Estimation — A machine learning-based assessment of the number of contributors in DNA mixtures. *Forensic Science International: Genetics*. 2017. Vol. 27. Pp. 82—91. DOI: 10.1016/j.fsigen.2016.11.006 (date accessed: 02.04.2024) ; Barash M., McNevin D., Fedorenko V., Giverts P. Machine learning applications in forensic DNA profiling: A critical review. *Ibid.* 2024. Vol. 69.Art. 102994. DOI: 10.1016/j.fsigen.2023.102994 (date accessed: 02.04.2024).

13 Wen Zh., Curran J. M., Harbison S.-A., Wevers G. E. Classification of firing pin impressions using HOG-SVM. *Journal of Forensic Sciences*. 2023. Vol. 68. Is. 6. Pp. 1946—1957. DOI: 10.1111/1556-4029.15377 (date accessed: 02.04.2024).

14 Pisantanaroj P., Tanpisuth P., Sinchavanwat P., Phasuk S., Phienphanich Ph., Jangtawee P., Yakoompai K., Donphoongpi M., Ekgasit S., Tantibundhit Ch. Automated Firearm Classification From Bullet Markings Using Deep Learning. *IEEE Access*. 2020. Vol. 8. Art. 78236—78251. DOI: 10.1109/ACCESS.2020.2989673 (date accessed: 02.04.2024).

firearms on bullets and cartridges [15] and to analyze the sounds that mechanisms of the weapon make during firing and operation [16].

Article Purpose

Acquaint experts in the field of criminalistics with the main capabilities of machine learning methods. Demonstrate what types of weapons forensics problems it is appropriate to use these methods for, and what it takes.

Research Methods

In order to achieve this goal, general scientific (dialectical, comparison, analysis,

15 Kamaruddin S. B. Ah., Ghanib N. A., Liong Ch.-Y., Jemain A. A. Firearm Classification using Neural Networks on Ring of Firing Pin Impression Images. *ADCAIJ: Advances in Distributed Computing and Artificial Intelligence Journal.* 2012. Vol. 1. No. 3. Pp. 27—34. DOI: 10.14201/ADCAIJ20121312734 (date accessed: 02.04.2024) ; Giverts P., Sorokina K., Fedorenko V. Examination of the possibility to use Siamese networks for the comparison of firing pin marks. *Journal of Forensic Sciences.* 2022. Vol. 67. Is. 6. Pp. 2416—2424. DOI: 10.1111/1556-4029.15143 (date accessed: 02.04.2024) ; Giudice O., Guarnera L., Paratore A. B., Farinella G. M., Battiato S. Siamese Ballistics Neural Network. *International Conference on Image Processing (ICIP).* IEEE (Taipei, Taiwan, 22—25 Sept 2019). 2019. Pp. 4045—4049. DOI: 10.1109/ICIP.2019.8803619 (date accessed: 02.04.2024) ; Banno A. Estimation of Bullet Striation Similarity Using Neural Networks. *Journal of Forensic Sciences.* 2004. Vol. 49. Is. 3. Pp. 1—5. DOI: 10.1520/JFS2002361 (date accessed: 02.04.2024) ; Li D. A New Approach for Firearm Identification with Hierarchical Neural Networks Based on Cartridge Case Images. *5th IEEE International Conference on Cognitive Informatics.* IEEE (Beijing, China, 17—19 July 2006). 2006. Vol. 2. Pp. 923—928. DOI: 10.1109/COGINF.2006.365616 (date accessed: 02.04.2024) ; Petraco N. D. K, Chan H., De Forest P. R., Diaczuk P., Gambino C., Hamby J., Kammerman F. L., Kammrath B. W., Kubic T. A., Kuo L., et al. Application of Machine Learning to Toolmarks: Statistically Based Methods for Impression Pattern Comparisons. Final Report. U. S. Department of Justice. Dec 2011. 101 p. URL: https://www.ojp.gov/pdffiles1/nij/grants/239048.pdf (date accessed: 02.04.2024) ; Le Bouthillier M.-E., Hrynkiw L., Beauchamp A., Duong L., Ratté S. Automated detection of regions of interest in cartridge case images using deep learning. *Journal of Forensic Sciences.* 2023. Vol. 68. Is. 6. Pp. 1958—1971. DOI: 10.1111/1556-4029.15319 (date accessed: 02.04.2024).

16 Giverts P., Eckert J., Sofer S., Solewicz Y. Preliminary Investigation of Acoustic Identification of Firearms Mechanisms and its Application to Case Investigations. *AFTE Journal.* 2022. Vol. 53. Is. 4. Pp. 152—158. URL: https://www.researchgate.net/publication/358931773_Preliminary_Investigation_of_Acoustic_Identification_of_Firearms_Mechanisms_and_its_Application_to_Case_Investigation (date accessed: 02.04.2024) ; Varer B., Giverts P. The Comparison of Feature Engineering Methods Used for Acoustic Identification of Firearms. *International Conference on Information and Telecommunication Technologies and Radio Electronics (UkrMiCo).* IEEE (Odesa, Ukraine, 29 Nov 2021). 2021. Pp. 164—167. DOI: 10.1109/UkrMiCo52950.2021.9716587 (date accessed: 02.04.2024) ; Гиверц П., Грибер А. Применение машинного обучения при исследовании акустических сигналов огнестрельного оружия. *Актуальні питання судової експертизи і криміналістики* : зб. мат-лів міжнар. наук.-практ. конф.-полілогу (Харків, 15—16.04.2021) Харків, 2021. С. 133—134. URL: https://www.hniise.gov.ua/uploads/files/public-folder/2021_tezy_konference%20in%20print5.pdf (date accessed: 02.04.2024) ; Giverts P., Sofer S., Solewicz Y., Varer B. Firearms identification by the acoustic signals of their mechanisms. *Forensic Science International.* 2020. Vol. 306. Art. 110099. DOI: 10.1016/j.forsciint.2019.110099 (date accessed: 02.04.2024).

synthesis, induction, deduction) and special scientific methods (formal logical, system-structural) were used, as well as the method of morphological analysis to bring complex task of specific forensic research to a set of morphological units that can coincide for different types of studied processes and tasks.

Main Content Presentation

Machine learning methods and their varieties, known as *artificial intelligence* have become widespread in the modern world. Currently, it is difficult to find a person who did not seek the help of artificial intelligence to translate texts, improve photographic images, search for information, etc. In recent years, the most popular intelligent virtual assistants and systems like *ChatGPT* (*Chat Generative Pre-trained Transformer*) *and Bard* (*Building Automated Resilient Definition*). Many web-sites and phone apps are able to generate images based on the suggested description or write text on the recommended topic. At the same time, paradoxically, few people (except data scientists) know what machine learning is and how it can be involved in solving specialized problems.

Machine learning as a method of solving specialized problems

Let us explore what machine learning is and how it differs from other methods. According to the classical scientific approach to the study of some phenomenon or effect, the researcher studies all the connections between the elements of the studied system, analyzes them, and based on this creates a mathematical model that describes the entire system. It is in this way that the basic formulas in physics, mechanics, thermodynamics and other sciences were derived. At the same time, the results may differ from the ideal model, since a significant number of random factors arise, so empirical formulas (containing coefficients that take into account influence of various environmental factors) are developed so that the results of the calculations based on them are within the limits of permissible errors. Instead, such results can be closer to reality (compared to results obtained using analytical formulas for ideal models) and will cost much less to obtain. Dependence functions for empirical formulas can be obtained from the results of experiments and observations. Based on received data, build a function that will most fully reflect the relationship between these data. The form of obtained function, amount of analyzed data, the final values of the parameters are usually determined by the researcher, but their relationship is known. However, the question arises: what to do if the relationship between various data is unknown, and the amount of data itself is huge? In this case, statistical analysis, in particular, machine learning, will come in handy.

Machine learning is a statistical analysis system that allowing identify dependencies when analyzing a large amount of data. The first methods of machine learning began to be developed in the 1950s. Currently, for various data and types of problems to be solved, many different methods of machine learning have been developed: for example, linear regression, linear discriminant analysis, the method of *Decision tree, Naive Bayes*

Classifier, Random forests, various types of neural networks, etc.

Many different tasks that are solved using machine learning methods can be outlined in several main ways:

- *classification,* when objects (data) belonging to several groups are known and it is necessary to determine which group the object under research belongs to;
- *clustering* will facilitate distribution of all data into several groups;
- *regression construction* is analysis of obtained data and construction of functions that make possible to predict situation development (used to predict development of various processes);
- *dimensional reduction* allows to identify implicit dependencies between various parameters that describe a process or object and, as a result, reduce their number (used for process analysis, dependency detection, as well as for data compression and recovery).

In addition, these methods can be combined at different stages of the task.

Peculiarity of machine learning methods is that (unlike classical methods) the algorithm (model) analyzes data itself and chooses options to achieve the goal while assessing the error at each stage. Upon receipt of the minimum error value, research purpose is considered to be achieved. The process of choosing optimal parameter is training.

Let us outline several types of such training:

- *supervised learning* is algorithm receives sets of initial data, as well as known results that should be obtained with correct interpretation of available data;
- *unsupervised learning*: there are no known outcomes, only raw data sets and the desired outcome.

Let us take a closer look at each of the main tasks that are solved using machine learning. For this purpose, imagine many balls of different diameters, made of different materials. Each of these objects has its own characteristics: diameter, mass, color, electrical conductivity, jump height after falling from a certain height to a solid surface, deformation coefficient (change in geometric size) after two-sided compression with a given force, etc.

Classification. Suppose that our task is to automatically divide the balls into metallic and non-metallic ones. Solution for this problem is the development of a classification function. For this purpose, it is possible to apply the method of learning with a teacher, that is, build a function based on a training sample in which the parameters of objects and the group to which they belong are known. For simplicity, let us classify balls by only two characteristics: mass and diameter. **Fig. 1** clearly demonstrates that light and dark points representing metallic and non-metallic balls here, form two groups. Dividing line can be drawn between these groups. The function describing this distinction is classification function. The distance of the point to the line given by the classification function will indicate the degree of reliability with which the object belongs to a certain group.

Another method of machine learning used for classification is the *Decision Tree* method, according to which decisions are made by answering hierarchically

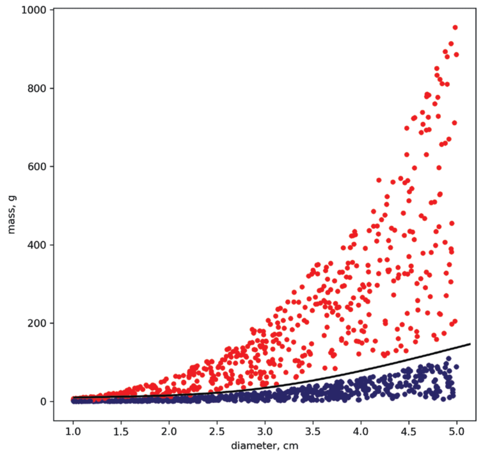

Fig. 1. Distribution of mass and diameter values for balls made of metal (light) and non-metallic materials (dark)

organized questions: each subsequent question depends on the answer to the previous one. According to our example, it is possible to build a *Decision Tree* (q.v. **Fig. 2**) with nodes based on the data of the training sample. Thus, for example, based on the results of this training sample, it can be stated that the diameter of the smallest metal ball is equal to $\mathbf{d_1}$, and its mass is greater than $\mathbf{m_1}$. Thus, if the ball diameter is equal to or less than $\mathbf{d_1}$ and at the same time its mass is greater than or equal to m1, then the ball is made of metal. On the other hand, the diameter of the largest non-metallic ball is $\mathbf{d_2}$ and its mass is $\mathbf{m_2}$, so a ball with a diameter greater than or equal to $\mathbf{d_2}$, and a mass less than or equal to $\mathbf{m_2}$ is a non-metallic ball. Continuing to consider such data for training sample and creating nodes (questions), you can build a *Decision Tree,* in which all branches will end with a *leaf* indicating that the object belongs to one of the groups.

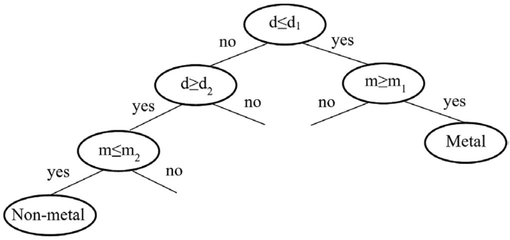

Fig. 2. *Decision tree* for classification of balls into metal and non-metal

With a small number of parameters and a small amount of data in the training group, it is possible to build a classification function or *Decision Tree* without involving technical means. However, as the amount of data increases, the number of computations increases rapidly. In this case, in fact, the only method is machine learning algorithms. Their use makes it quite easy to build a classification system that will contain more than two groups and many parameters. In addition, you can use error estimation functions that will allow to build a model that will fully correspond to studied data.

Clustering (cluster analysis): Clustering task is similar to the classification task, but here you need to distribute objects into groups. Since there are no examples of such groups, the tasks of clustering require training without a teacher. With clustering, the algorithm will try to find common characteristics for different groups of objects and form their cluster. Let us return to our example with balls. We know that all balls of the two types are metallic and non-metallic, but it is not known which of them belong to which group. Clustering algorithms will help to distribute objects according to relevant groups according to identified specifics.

Different methods can be used for both classification and clustering: for example, *k-nearest neighbors' algorithm* or the *hierarchical cluster analysis*. According to the *k-nearest neighbors' algorithm*, the number of points equal to the number of clusters is first selected from all the analyzed data, and each of these points is considered the centroid of the cluster. In the next iteration, for each centroid, *nearest neighbor* is searched, which is added to the cluster, after which the position of the centroid is calculated again, taking into account the new data, and the position of the centroid is defined as the minimum sum squared deviation of the cluster points from the centroid. Procedure is repeated until the coordinates of the centroids stop changing. Having identified clusters, it is possible to find out which of them the new object belongs to.

According to the hierarchical clustering method, at the first stage of data processing, each object is provided with its own cluster. In the next steps, the two closest clusters are merged into one, etc., until they receive one cluster.

Constructing regressions: Investigating data and processing the results of experiments usually involves determining the relationship between the parameters. Thus, the data of an experiment describe the process under research under specific conditions at specific points, and in order to apply the results obtained, it is necessary to find out the dependence of some parameters on others. In order to determine such a functional dependence, select the type of function that describes the process (linear, quadratic, power, etc.). After selecting the type of function, coefficients are selected so that resulting function describes all the data as accurately as possible. By constructing regression, it is possible calculate its value for source data, as well as between points determined experimentally, or for points that are outside the experiment range.

Let us consider regression use on the example with balls. Suppose that we know the mass of ten balls of different metals with a diameter from 1 to 10 cm. Relationship between the mass of these balls and their diameter can be represented using a graph. According to proposed data, dependence between mass and diameter can be constructed in the form of various functions (E.g., **Fig. 3**). In the example, linear regression, exponential (exponential) and polynomial function of the third order are constructed. Applying these functions, one can calculate the estimated mass of a metal ball with a diameter of 4.5 or 12 cm (light dots on the graph). According to the graph, forecast using linear regression will give an incorrect value, exponential regression predicts a fairly close weight value for a ball with a diameter of 4.5 cm, but already for a ball with a diameter of 12 cm, predicted value of this parameter is quite far from the control value. The third-order polynomial regression most closely describes the relationship between the diameter and the ball mass. This is quite expected, since relationship between volume and diameter is due to a third-order power relationship (volume depends on the diameter cube), and deviation of the points from regression function is due to the fact that the balls are made of metals of different densities.

In the above example, calculations are simple and the type of regression function is easy to determine, but what to do when there are thousands or hundreds of thousands of samples and dozens of parameters in the study sample, the relationship between which is not so obvious? Here, machine learning algorithms come to the rescue. These

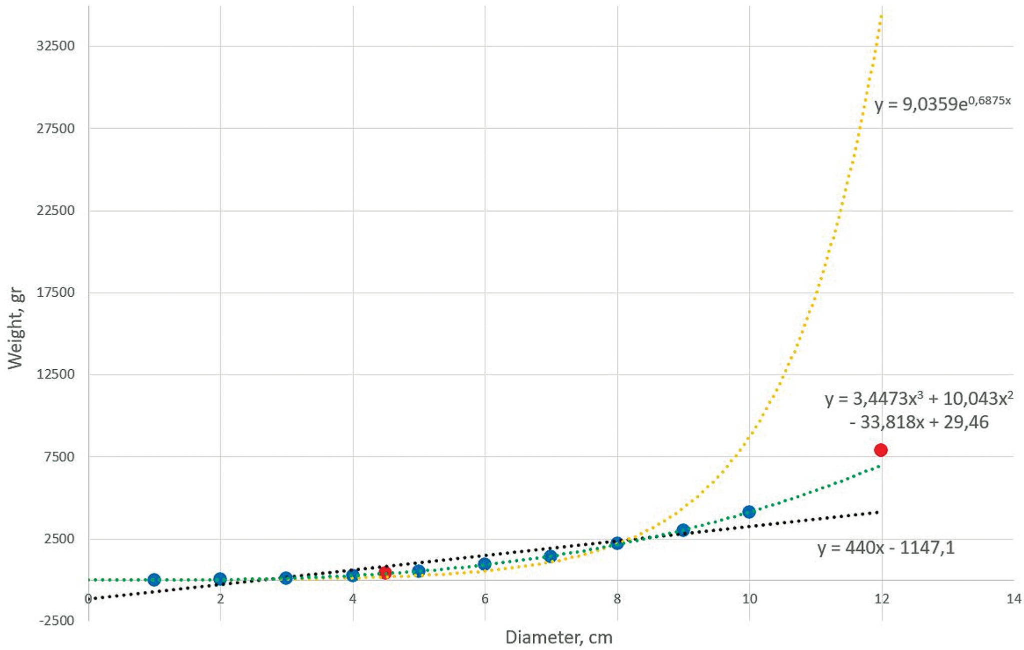

Fig. 3 Example of constructing regression function

algorithms make it possible to build a model that will more reliably describe the processed data. For this purpose, various methods of machine learning, various error estimation functions can be applied, which will allow to select the optimal regression, as well as *methods for detecting emissions:* erroneous data, which impact on the model should be minimized.

Dimension reduction. Suppose it is necessary to compare two different balls from our example, only partial data is known for each of them. Thus, for one ball, weight, coefficient of electrical conductivity, height of bouncing during a fall from a given height are known, and for the second: diameter, hardness and coefficient of deformation (changes in geometric size) after bilateral compression with a given force. In this case, it is easy to associate all the specified parameters with characteristics of the material which the balls are made from. Based on the set of parameters describing the first ball, it is possible to find out what material it is made of, therefore, to determine missing characteristics or, on the basis of known partial indicators, to obtain some other characteristics (for example, material density) that can be compared.

In availability of many data described by *N*-parameters, using machine learning methods, it is possible to build a model and determine the dependence of these parameters on other *M*-parameters (where *M* is less than *N*). This reduction in data dimension makes it easier to describe phenomena under research, compare objects and processes for which only partial parameters are known, etc.

Application of machine learning methods to solve the tasks of forensic weapons research

While examining cartridges recovered from the crime scene, the caliber is first determined. Most often, this task is not difficult for forensic expert. At the same time, software, that will determine the cartridge caliber based on photo (for example, taken by a smartphone), can become a useful investigator. This task can be considered a classification problem, for which solution it is advisable to use different variants of neural networks trained on a significant number of images of casings of different calibers. It is possible to use software that separates various geometric elements in the image (circles, arcs, ellipses, straight lines, angles, etc.): the scheme of the photographed sleeve obtained in this way is compared with the database of the geometric characteristics of the sleeves or processed by machine learning methods (for example, *Decision tree*).

The next issue to be considered by ballistic expert is to find out the group belonging to the cartridge, in other words, to determine from which type (model) of the weapon it was fired. For this, expert analyzes group specifics found on the cartridge surface and compares them with specifics described in professional readings, computer databases and collections of shell samples. At the same time, decision made is subjective and depends on experience of the specialist performing this research.

The above task can be considered as a classification task, when it is necessary to find out to which of the predefined classes the traces on the studied cartridge belong. Such a task is one of the most studied in machine learning. More effective is the use *Convolutional Neural Networks,* trained on a significant number of images of different cartridges fired from different samples of firearms. It is possible to apply another approach, in which different group characteristics are first distinguished, and then obtained result is compared with existing databases. In this case, it is possible to detect group features both by machine vision methods and by machine learning methods. Similarly, database searches can be performed both using classical methods and data comparison, and using machine learning methods (such as *Decision Trees* or *k-nearest neighbors*).

Next, forensic expert compares individual specifics available on examined cartridges or bullets in order to determine their identity. This task is the main one for forensic ballistics examination, although no objective, mathematical method of solving it has been developed to date. The result obtained by forensic expert is subjective: its reliability depends on the expert's experience. At the same time, this task can be limited to a simple classification into two categories: availability and lack of identity (for the sake of fairness and to more accurately reflect reality, intermediate, non-categorical answer should be introduced).

It would seem that if this is classification task, then machine learning methods can easily cope with it. However, difficulty lies in the fact that usually models are trained on a significant number of examples for each class, which allows machine learning algorithms to distinguish sets of features that characterize each class. In the case of comparing two cartridges or bullets, this method does not work: it is appropriate to apply an approach called *One Shot Learning* [17]. One implementation of this approach is the construction of a Siamese neural network: model containing two identical neural networks. In this case, the model learns to identify specifics that can theoretically characterize compared object, and then compares specifics of the two objects under research.

Bullets and cartridges examined by a ballistics expert are often damaged both by mechanical damage after firing and by corrosion. In addition, casings and bullets may contain marks related to the manufacturing process. Such "noise" damage can cause difficulties when using ABIS. The problem of cleaning images from "noise" is solved with the help of various intelligent filters. Cleaning the image from excess "noise" and damage is denoising it can be considered a regression construct. Thus, for example, if there is a secondary trace of the rifling field on a bullet, which surface is damaged by a shot, and there is a scratch crossing the informative traces, then such a scratch can be automatically detected and removed from the image, and the image of damaged traces can be restored. Similarly, it is possible to generate reconstructed images of deformed primary and secondary traces from the fields and the bottom of the cuts. Certainly, for such an application, existing machine learning models should be modified accordingly.

In order to determine the shot distance, shot traces on the victim body, on his clothes or on the surfaces of various obstacles are analyzed (in particular, the distribution of traces of burning gunpowder, traces of lead and copper). Based on this analysis, a decision is made about the possible range of shot distances. Like many others, this task belongs to the classification tasks, so analyzing the image will allow to determine what range of shot distances for this caliber, type of ammunition and weapon this shot belongs to. In some cases, when there is no data for specific cartridges or weapons in the relevant information bases, it is necessary to make an experimental shot at all distances, but most often cartridges used to commit the crime may not be enough to obtain samples of distribution nature of the shot products at all distances. In this case, you can build a machine learning model (regression), which, based on a smaller number of shots, will build a function that will describe the distribution of shot products and will make it possible to determine the shot distance even with a small amount of experimental data. Similarly, a model can be built and trained to take into account the influence of wind strength and direction, humidity and many other parameters that are often overlooked, which will increase accuracy and reliability of research results.

For solving tasks of external and wound ballistics, forensic expert should often determine:

- parameters of flight trajectory of bullets of various calibers and designs;
- bullet behavior and change in its trajectory after hitting an obstacle both in the case of an obstacle breakthrough and in the case of ricochet;

17 Jadon Sp., Garg A. Hands-On One-shot Learning with Python. Birmingham : Packt Publishing Ltd, 2020. 156 p.

- characteristics of the wound canal and injuries inflicted on the victim.

For this purpose, it is necessary to conduct dozens and hundreds of experimental shots and subsequently perform an approximation of collected data to identify an empirical dependence, that is, actually build a regression describing the process under research. In order to solve such a problem, it is more appropriate to use various methods of machine learning. In addition, the application of these methods will help to identify the main factors affecting the phenomenon under research (the task of reducing dimensions) and facilitate its description. In turn, isolating the main factors will significantly reduce the number of experiments to be conducted.

Conclusions

Machine learning systems and methods, which application has contributed to the solution of many applied problems in various fields of science and practice are becoming more common day by day, becoming a common feature of our daily life. At the same time, only isolated studies on the use of these methods to perform the tasks of forensic research on firearms have been published, since forensic experts are not aware of the possibilities of machine learning, and machine learning specialists do not know the tasks and needs of forensic science. Presented research takes the first steps towards solving this problem and demonstrates the main capabilities of machine learning and describes ways to apply them to perform some firearm research tasks.

Машинне навчання і можливість його застосування для криміналістичних досліджень вогнепальної зброї

Павел Ґіверц, Олександр Коломійцев

Роль машинного навчання для сучасного суспільства складно переоцінити, особливо зважаючи на те, що останніми роками ця роль постійно зростає. Різноманітні моделі машинного навчання застосовують як у повсякденному житті, так і для розв'язання прикладних, конструкторських і наукових проблем. Водночас для більшості фахівців у різноманітних галузях знань, зокрема для спеціалістів із криміналістичного дослідження зброї, методи й можливості машинного навчання до сьогодні залишаються маловідомими. У статті розглянуто базові завдання, виконанню яких сприятимуть методи машинного навчання, а також наведено приклади завдань, які постають перед експертами під час проведення криміналістичних досліджень вогнепальної зброї та які можна розв'язати за допомогою згаданих методів. Метою статті є ознайомити фахівців у галузі криміналістики з основними можливостями методів машинного навчання. Продемонстровано також, у розв'язанні яких саме типів завдань криміналістичного дослідження зброї ці методи доцільно застосовувати і що для цього потрібно. Для досягнення поставленої мети застосовано загальнонаукові (діалектичний, порівняння, аналіз, синтез, індукцію, дедукцію) і спеціальні наукові методи (формально-логічний, системно-структурний), а також спеціальні знання та навички з оперування методами криміналістичної ідентифікації (зокрема, балістики, мікроскопічних порівняльних досліджень слідів на кулях і гільзах, виявлення й дослідження

продуктів пострілу з метою визначити дистанцію пострілу тощо), принципами машинного навчання та математичними алгоритмами, використовуваними для побудови різних моделей.

Ключові слова: *криміналістичне дослідження вогнепальної зброї; балістична ідентифікація; оцінювання відстані пострілу; машинне навчання; нейронна мережа; класифікація; кластеризація; регресія.*

Financing

This research did not receive any specific grant from funding institutions in the public, commercial or non-commercial sectors.

Disclaimer

Founders had no role in the research design, data collection and analysis, decision to publish or manuscript preparation.

Participants

Authors have contributed solely to intellectual discussion underlying this document, case law research, writing and editing, and assumes responsibility for its content and interpretation.

Declaration of Competing Interest

Authors declare no conflict of interest.

References

Banno, A. (2004). Estimation of Bullet Striation Similarity Using Neural Networks. *Journal of Forensic Sciences*. Vol. 49. Is. 3. DOI: 10.1520/JFS2002361.

Barash, M., McNevin, D., Fedorenko, V., Giverts, P. (2024). Machine learning applications in forensic DNA profiling: A critical review. *Forensic Science International: Genetics*. Vol. 69. Art. 102994. DOI: 10.1016/j.fsigen.2023.102994.

Carriquiry A., Hofmann H., Tai X. H., VanderPlas S. (2019). Machine learning in forensic applications. *Significance*. Vol. 16. Is. 2. DOI: 10.1111/j.1740-9713.2019.01252.x.

Chen, Zh., Chu, W., Soons, J. A., Thompson, R. M., Song, J., Zhao, X. (2019). Fired bullet signature correlation using the Congruent Matching Profile Segments (CMPS) method. *Forensic Science International*. Vol. 305. Art. 109964. DOI: 10.1016/j.forsciint.2019.109964.

Chen, Zh., Song, J., Soons, J. A., Thompson, R. M., Zhao, X. (2020). Pilot study on deformed bullet correlation. *Forensic Science International*. Vol. 306. Art. 110098. DOI: 10.1016/j.forsciint.2019.110098.

Chu, W., Tong, M., Song, J. (2013). Validation Tests for the Congruent Matching Cells (CMC) Method Using Cartridge Cases Fired with Consecutively Manufactured Pistol Slides. *Journal of the Association of Firearms and Toolmarks Examiners*. Vol. 45. Is. 4. URL: https://www.nist.gov/publications/validation-tests-congruent-matching-cellscmc-method-using-cartridge-cases-fired.

Chugh, T., Cao, K., Zhou, J., Tabassi, E., Jain, A. K. (2018). Latent Fingerprint Value Prediction: Crowd-Based Learning. *IEEE Transactions on Information Forensics and Security*. Vol. 13. Is. 1. DOI: 10.1109/TIFS.2017.2721099.

Cuellar, M., Gonzalez, C., Dror, I. E. (2022). Human and machine similarity judgments in forensic firearm comparisons. *Forensic Science International: Synergy*. Vol. 5. DOI: 10.1016/j.fsisyn.2022.100283.

Dror, I. E. (2023). The most consistent finding in forensic science is inconsistency. *Journal of Forensic Sciences*. Vol. 68. Is. 6. DOI: 10.1111/1556-4029.15369.

Dror, I. E., Scurich, N. (2020). (Mis) use of scientific measurements in forensic science. *Forensic Science International: Synergy*. Vol. 2. DOI: 10.1016/j.fsisyn.2020.08.006.

Durou, A., Al-Maadeed, S., Aref, I., Bouridane, Ah., Elbendak, M. (2019). A Comparative Study of Machine Learning Approaches for Handwriter Identification. *12th International Conference on Global Security, Safety and Sustainability (ICGS3). IEEE* (London, UK, 16—18 Jan 2019). DOI: 10.1109/ICGS3.2019.8688032.

Ezeobiejesi, J. C. (2019). *Deep Learning for the Analysis of Latent Fingerprint Images*. University of California Riverside. URL: https://escholarship.org/uc/item/564185wc.

Giudice, O., Guarnera, L., Paratore, A. B., Farinella, G. M., Battiato, S. (2019). Siamese Ballistics Neural Network. *International Conference on Image Processing (ICIP). IEEE* (Taipei, Taiwan, 22—25 Sept 2019). DOI: 10.1109/ICIP.2019.8803619.

Giverc, P., Griber, A. (2021). Primenenie mashinnogo obucheniya pri issledovanii akusticheskih signalov ognestrelnogo oruzhiya [Application of machine learning in the study of acoustic signals of firearms]. *Aktualni pytannia sudovoi ekspertyzy i kryminalistyky*: zb. mat-liv mizhnar. nauk.-prakt. konf.-polilohu (Kharkiv, 15—16.04.2021) Kharkiv. URL: https://www.hniise.gov.ua/uploads/files/public-folder/2021_tezy_konference%20in%20print5.pdf [in Russian].

Giverts, P., Eckert, J., Sofer, S., Solewicz, Y. (2022). Preliminary Investigation of Acoustic Identification of Firearms Mechanisms and its Application to Case Investigations. *AFTE Journal*. Vol. 53. Is. 4. URL: https://www.researchgate.net/publication/358931773_Preliminary_Investigation_of_Acoustic_Identification_of_Firearms_Mechanisms_and_its_Application_to_Case_Investigation.

Giverts, P., Sofer, S., Solewicz, Y., Varer, B. (2020). Firearms identification by the acoustic signals of their mechanisms. *Forensic Science International*. Vol. 306. Art. 110099. DOI: 10.1016/j.forsciint.2019.110099.

Giverts, P., Sorokina, K., Fedorenko, V. (2022). Examination of the possibility to use Siamese networks for the comparison of firing pin marks. *Journal of Forensic Sciences*. Vol. 67. Is. 6. DOI: 10.1111/1556-4029.15143.

Jadon, Sp., Garg, A. (2020). *Hands-On One-shot Learning with Python*. Birmingham : Packt Publishing Ltd.

Kamaruddin, S. B. Ah., Ghanib, N. A., Liong, Ch.-Y., Jemain, A. A. (2012). Firearm Classification using Neural Networks on Ring of Firing Pin Impression Images. *ADCAIJ: Advances in Distributed Computing and Artificial Intelligence Journal*. Vol. 1. No. 3. DOI: 10.14201/ADCAIJ20121312734.

Kingston, Ch. (1992). Neural Networks in Forensic Science. *Journal of Forensic Sciences*. Vol. 37. Is. 1. DOI: 10.1520/jfs13232j.

Le Bouthillier, M.-E., Hrynkiw, L., Beauchamp, A., Duong, L., Ratté, S. (2023). Automated detection of regions of interest in cartridge case images using deep learning. *Journal of Forensic Sciences*. Vol. 68. Is. 6. DOI: 10.1111/1556-4029.15319.

Lemos, N., Shah, K., Rade, R., Shah, D. (2018). Personality Prediction based on Handwriting using Machine Learning. *International Conference on Computational Techniques, Electronics and Mechanical Systems (CTEMS). IEEE* (Belgaum, India, 21—22 Dec 2018). DOI: 10.1109/CTEMS.2018.8769221.

Li, D. (2006). A New Approach for Firearm Identification with Hierarchical Neural Networks Based on Cartridge Case Images. *5th IEEE International Conference on Cognitive Informatics. IEEE* (Beijing, China, 17—19 July 2006). Vol. 2. DOI: 10.1109/COGINF.2006.365616.

Marciano, M. A., Adelman, J. D. (2017). PACE: Probabilistic Assessment for Contributor Estimation — A machine learning-based assessment of the number of contributors in DNA mixtures. *Forensic Science International: Genetics*. Vol. 27. DOI: 10.1016/j.fsigen.2016.11.006.

Margagliotti, G., Bollé, T. (2019). Machine learning & forensic science. *Forensic Science International*. Vol. 298. DOI: 10.1016/j.forsciint.2019.02.045.

Osterburg, J. W. (1969). The Evaluation of Phisical *Evidenve* in *Criminalistics*: Subjective or Objective Process? *The Journal of Criminal Law, Criminalogy and Police Science*. Vol. 60. No. 1. DOI: 10.2307/1141742.

Petraco, N. D. K, Chan, H., De Forest, P. R., Diaczuk, P., Gambino, C., Hamby, J., Kammerman, F. L., Kammrath, B. W., Kubic, T. A., Kuo, L., et al. (2011). *Application of Machine Learning to Toolmarks: Statistically Based Methods for Impression Pattern Comparisons. Final Report*. U. S. Department of Justice. URL: https://www.ojp.gov/pdffiles1/nij/grants/239048.pdf.

Pisantanaroj, P., Tanpisuth, P., Sinchavanwat, P., Phasuk, S., Phienphanich, Ph., Jangtawee, P., Yakoompai, K., Donphoongpi, M., Ekgasit, S., Tantibundhit, Ch. (2020). Automated Firearm Classification From

Bullet Markings Using Deep Learning. *IEEE Access*. Vol. 8. Art. 78236—78251. DOI: 10.1109/ACCESS.2020.2989673.

Report to the President's. Forensic Science in Criminal Courts: Ensuring Scientific Validity of Feature-Comparison Methods. (2016). URL: https://obamawhitehouse.archives.gov/sites/default/files/microsites/ostp/PCAST/pcast_forensic_science_report_final.pdf.

Sharma, A., Kaur, M. (2018). Automatic segmentation for separation of overlapped latent fingerprints. *International Journal of Computer Sciences and Engineering*. Vol. 6. Is. 7. DOI: 10.26438/ijcse/v6i7.484490.

Song, J., Song, H. (2023). Reporting likelihood ratio for casework in firearm evidence identification. *Journal of Forensic Sciences*. Vol. 68. Is. 2. Pp. 399—406. DOI: 10.1111/1556-4029.15186.

Stojanović, B., Marques, O., Nešković, A., Puzović, S. (2016). Fingerprint ROI segmentation based on deep learning. *24th Telecommunications Forum (TELFOR). IEEE* (Belgrade, Serbia,22—23 Nov 2016). DOI: 10.1109/TELFOR.2016.7818799.

Strengthening Forensic Science in the United States: A Path Forward. (2009). Washington, D. C. DOI: 10.17226/12589.

Tong, M., Song, J., Chu, W., Thompson, R. M. (2014). Fired Cartridge Case Identification Using Optical Images and the Congruent Matching Cells (CMC) Method. *Journal of Research of the National Institute of Standards and Technology*. Vol. 119. Art. 575. DOI: 10.6028/jres.119.023.

Varer, B., Giverts, P. (2021). The Comparison of Feature Engineering Methods Used for Acoustic Identification of Firearms. *International Conference on Information and Telecommunication Technologies and Radio Electronics (UkrMiCo). IEEE* (Odesa, Ukraine, 29 Nov 2021). DOI: 10.1109/UkrMiCo52950.2021.9716587.

Wang, S., Jia, Sh. (2019). Signature handwriting identification based on generative adversarial networks. *Journal of Physics: Conference Series*. Vol. 1187. Is. 4. Art. 042047. DOI: 10.1088/1742-6596/1187/4/042047.

Wen, Zh., Curran, J. M., Harbison, S.-A., Wevers, G. E. (2023). Classification of firing pin impressions using HOG-SVM. *Journal of Forensic Sciences*. Vol. 68. Is. 6. DOI: 10.1111/1556-4029.15377.

Wong, W. J., Lai, Sh.-H. (2020). Multi-task CNN for restoring corrupted fingerprint images. *Pattern Recognition*. Vol. 101. Art. 107203. DOI: 10.1016/j.patcog.2020.107203.

Yih, Ch.-H., Hung, J.-L., Wu, J.-A., Chen, L.-M. (2018). Overlapped Fingerprint Separation Based on Deep Learning. *International Conference on Digital Medicine and Image Processing* (Okinawa, Japan, Nov 12—14, 2018). DOI: 10.1145/3299852.3299857.

Giverts, P., Kolomiitsev, O. (2024). Machine Learning and Possibility of its Application for Forensic Researches on Firearms. *Theory and Practice of Forensic Science and Criminalistics*. Issue 2 (35). P. 28—43. DOI: 10.32353/khrife.2.2024.03.