DOI 10.34185/1562–9945–5–166–2026–04
UDC 681.5:004.89:004.942:622.24

O.D. Klymenko, S.M. Vovk

HYBRID NEURAL NETWORK MODEL FOR PROBABILITY DISTRIBUTION IDENTIFICATION IN THE ASSESSMENT OF THE STATE OF NATURAL AND TECHNICAL OBJECTS

Abstract. The aim of this study is to develop a hybrid neural network model for the automatic identification of the probability distribution through the combined use of the empirical distribution function and statistical characteristics of the sample. The research methodology integrates approaches from probability theory, mathematical statistics, deep learning, and computer simulation. The proposed architecture combines a one–dimensional convolutional neural network (1D CNN) for analyzing the shape of the empirical distribution with a multilayer perceptron (MLP) for processing a vector of 12 robust statistical features. A dataset of 60,000 synthetic samples with sizes ranging from 20 to 1000 observations was generated for model training and testing. The experimental results demonstrated a classification accuracy of 97.8% on an independent test set containing 18,000 samples, outperforming standalone CNN and MLP models as well as conventional statistical goodness–of–fit tests. The obtained results confirm the effectiveness of the proposed hybrid neural network approach for the automatic identification the probability distributions under conditions of limited data availability.
Keywords: hybrid neural network; probability distribution identification; empirical distribution function; statistical features; deep learning; mathematical statistics.

Introduction. Classical methods used for identification or recognition of probability distributions are one of the rather problematic issues in applied statistics tasks. For example, in decision support systems, in information systems, in data analysis in automated systems, in forecasting and modulation of complex processes, etc.

The reliability of statistical conclusions, as well as the adequacy of mathematical models and the efficiency and accuracy of algorithms in data analysis, directly affect the reliability of statistical conclusions, as well as the adequacy of mathematical models and the efficiency and accuracy of algorithms in data analysis. Based on this, it can be determined that the accuracy of probability distribution identification provides the opportunity for the correct selection of statistical methods for data analysis, and this in turn provides the opportunity to build reliable models and make informed decisions.

Classical approaches of applied statistics, including Kolmogorov–Smirnov, Anderson–Darling and Shapiro–Wilk tests, demonstrate low statistical power when applied, especially in conditions of small and medium sample sizes ($n < 50$), especially when statistical distributions have a similar shape (normal, logistic and Student distribution). Existing machine learn-

© Klymenko O.D., Vovk S.M., 2026
Copyright for this paper by its authors. Use permitted under License CC BY 4.0.

ISSN 1562-9945 (Print)
ISSN 2707-7977 (Online)

ing methods rely on only one type of data representation (measurement samples), or rely on geometric characteristics of measurement samples, or sets of statistical features of measurement samples.

All this limits the ability to automatically identify distribution laws in conditions when the measurement sample has noise or in conditions when there is a measurement limitation (sample length). Therefore, the development of automatic identification of distribution laws is relevant for many fields of knowledge (rocket and space, information technology, computer–integrated technologies, radio electronics, etc.).

Problem Statement. Conventionally, the problem of determining whether a sample follows a particular probability distribution is addressed using goodness–of–fit tests, including the Kolmogorov–Smirnov (KS) test, the Anderson–Darling test, the Shapiro–Wilk test (for normality), and the Pearson chi–square test [1, 2]. These methods evaluate candidate distributions individually by estimating their parameters and calculating the corresponding *p*–values.

Despite their widespread application, these approaches exhibit several important limitations:

- reduced statistical power for small sample sizes ($n < 50$), making it difficult to distinguish between distributions with similar shapes, such as the normal and logistic distributions or Student's *t*–distributions with different degrees of freedom;
- sensitivity to outliers and the need for preliminary data preprocessing;
- increased computational cost when numerous alternative hypotheses must be tested sequentially;
- inability to provide a unified probabilistic assessment of all candidate distributions within a single inference procedure.

Furthermore, probability distributions with similar characteristics, such as the normal, Student's *t*, and logistic distributions, often exhibit highly similar empirical distribution functions, making reliable visual and statistical discrimination challenging. These factors substantially limit the applicability of conventional statistical techniques in automated data analysis.

Recent advances in deep learning provide new opportunities for addressing this problem. In particular, one–dimensional convolutional neural networks (1D CNNs) are capable of automatically extracting structural features from empirical distribution functions, whereas multilayer perceptrons (MLPs) effectively process statistical descriptors of data. Combining these complementary sources of information enables the construction of hybrid multimodal models that integrate distribution shape analysis with robust statistical characteristics, resulting in improved classification accuracy, better generalization, and enhanced robustness under limited–data conditions.

Therefore, an important scientific challenge is the development of a hybrid multimodal neural network architecture (1D CNN + MLP) for the automatic identification of probability distributions that provides higher recognition accuracy, improved generalization capability, and seamless integration into modern intelligent data analysis and decision–support systems.

The purpose and methods of the research. The objective of this study is to improve the accuracy and robustness of automatic identification of parametric probability distributions under conditions of limited sample sizes by developing a hybrid neural network model that

combines the analysis of the empirical distribution function with the statistical characteristics of the data.

To achieve this objective, the following tasks are addressed:

- to develop a hybrid neural network architecture integrating a one–dimensional convolutional branch (1D CNN) for processing the empirical distribution function and a multilayer perceptron (MLP) for analyzing a vector of 12 robust statistical features;
- to investigate the contribution of each component (1D CNN and MLP) to the overall classification performance;
- to perform a comparative evaluation of the proposed approach against conventional goodness–of–fit tests (Kolmogorov–Smirnov and Anderson–Darling) as well as state–of–the–art machine learning methods;
- to analyze the dependence of classification accuracy on sample size, with particular emphasis on small samples ($n \leq 50$), and to assess the robustness of the proposed model under data–scarce conditions;
- to evaluate the computational efficiency of the proposed model and investigate its suitability for real–time applications and integration into intelligent information systems.

Analysis of Recent Research and Publications. An analysis of recent research demonstrates considerable interest in statistical data analysis, pattern identification, and the application of intelligent methods for automating these tasks. Conventional approaches, particularly statistical goodness–of–fit tests such as the Kolmogorov–Smirnov and Anderson–Darling tests, are widely used for distribution analysis and the detection of hidden patterns in data. However, their effectiveness decreases when applied to small sample sizes or structurally complex datasets [1]. In particular, the study by K. Tryfonova demonstrated the application of the Kolmogorov–Smirnov test to hidden information detection, confirming its practical utility while also emphasizing its limited sensitivity under certain conditions.

Recent advances in information systems have been driven by the widespread use of neural networks. Neural networks have demonstrated a remarkable ability to capture complex nonlinear relationships in large datasets and are extensively employed in forecasting, optimization, and classification problems [2]. In scheduling and optimization tasks, neural network models are used both independently and in combination with evolutionary algorithms, resulting in improved solution quality and adaptability to changing conditions. Nevertheless, their practical application remains constrained by their dependence on training data, the complexity of hyperparameter tuning, and limited model interpretability. Now, the hybrid neural network approaches that combine multiple architectures or incorporate statistical information into deep learning models have attracted increasing attention. Studies [3, 9] have shown that integrating convolutional, recurrent, and attention–based mechanisms improves model prediction accuracy and adaptability under complex dynamic conditions. Such hybrid architectures have been successfully applied to financial risk analysis, time–series forecasting, and cybersecurity, achieving 9–13% higher accuracy than individual neural network architectures [3]. The effectiveness of hybrid models has also been demonstrated in the analysis of critical infrastructures and cybersecurity applications. In [4], a multimodal framework for attack detection in

ISSN 1562-9945 (Print)
ISSN 2707-7977 (Online)

SCADA/ICS systems was proposed, combining multiple signal representations to improve robustness against incomplete and noisy data. Similarly, studies on IoT networks have shown that combining machine learning and deep learning techniques (CNN, RNN, and GAN) enables effective anomaly detection and enhances cybersecurity performance [5]. Another important research direction focuses on integrating deep learning models with statistical characteristics of data. Abbasimehr et al. [6] demonstrated that incorporating statistical features into neural network architectures significantly improves prediction accuracy and robustness to noise. These findings support the effectiveness of multimodal approaches that exploit complementary sources of information. Furthermore, convolutional neural networks (CNNs) have been widely employed for extracting structural patterns from data and time series, as demonstrated in [7]. Advanced architectures such as InceptionTime have shown high performance in classification and forecasting tasks. Moreover, combining CNNs with recurrent neural networks (LSTM and GRU) enables the simultaneous modeling of local and long–term dependencies, making these architectures highly effective in applications such as energy systems, finance, and the oil and gas industry [8].

However, most existing studies are limited to a relatively small set of probability distributions, typically focusing on normal distribution, and rely on a single type of input representation, either the distribution shape or statistical descriptors. This limitation reduces their generalization capability and robustness when distinguishing between statistically similar probability distributions. The analysis of the reviewed literature leads to the following conclusions:

- conventional statistical methods remain essential tools for probability distribution analysis but exhibit significant limitations under modern data processing conditions;
- neural network–based approaches provide high classification accuracy but require further improvements in robustness and generalization capability;
- hybrid models that combine different neural network architectures with multiple types of input features represent the most promising direction for future research;
- integration of statistical characteristics with structural analysis of empirical distributions can substantially enhance the performance of intelligent data analysis systems.

Presentation of the main research material. The proposed hybrid model is designed to overcome the limitations of conventional statistical approaches by integrating deep learning techniques with statistical data analysis. The use of convolutional neural networks enables the automatic extraction of structural features from the shape of the empirical distribution function, while the incorporation of robust statistical descriptors captures sample characteristics that are invariant to scaling and translation. By combining these complementary sources of information, the proposed architecture generates a unified hybrid representation of the input data, thereby improving classification accuracy, enhancing robustness under limited sample sizes, and increasing the capability to discriminate between probability distributions with similar shapes. The architecture of the proposed model is illustrated in Fig. 1.

To train, validate, and evaluate the proposed model, exclusively synthetic datasets were employed [10–13]. This approach enables complete control over the true underlying probability distribution and its parameters while eliminating the subjectivity associated with manual

class labeling. The datasets were generated from six parametric distribution families, as summarized in Table 1.

Table 1

Parameters Used for Synthetic Data Generation

Class	Distribution	Parameters
0	Normal	$\mu \in [-100, 100]$, $\sigma \in [0.1, 5.0]$
1	Student's t	$\nu \in \{1,2,3,\ldots,10\}$ (discrete choice)
2	Logistic	$\mu \in [-100, 100]$, $s \in [0.1, 5.0]$
3	Binomial	$n \in [10, 100]$ (integer), $p \in [0.1,0.9]$
4	Exponential	$\lambda \in [0.1, 10.0]$
5	Uniform	$a \in [-100,100]$, $b \in [1,100]$

For each probability distribution, the parameters were randomly sampled from uniform distributions within the specified ranges, ensuring substantial variability in distribution shapes within each class. The sample size (n) was also selected randomly from the interval [20, 1000], simulating realistic scenarios involving limited amounts of measurement data. A total of 13,000 samples were generated for each distribution class, of which 10,000 were used for model training and 3,000 for testing.

To provide the input for the convolutional branch of the network, each sample of arbitrary size n is transformed into a fixed–length representation of L=500 through the following preprocessing steps:

1. Sorting the sample values:

$$x_1 \leq x_2 \leq \cdots \leq x_n. \tag{1}$$

2. Z–score normalization:

$$z_i = \frac{x_i - \bar{x}}{s}, \tag{2}$$

where

$$\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i, \tag{3}$$

$$s = \sqrt{\frac{1}{n-1}\sum_{i=1}^{n}(x_i - \bar{x})^2}. \tag{4}$$

3. Linear interpolation to obtain a sequence of fixed length (L = 500). Let $t \in [0,1]$ denote the normalized position index. The interpolated value at position t is computed as

$$f(t) = z_k + \frac{t - t_k}{t_{k+1} - t_k}(z_{k+1} - z_k), \tag{5}$$

where

$$t_k = \frac{k-1}{n-1}, k = \lfloor t(n-1) \rfloor + 1. \tag{6}$$

ISSN 1562-9945 (Print)
ISSN 2707-7977 (Online)

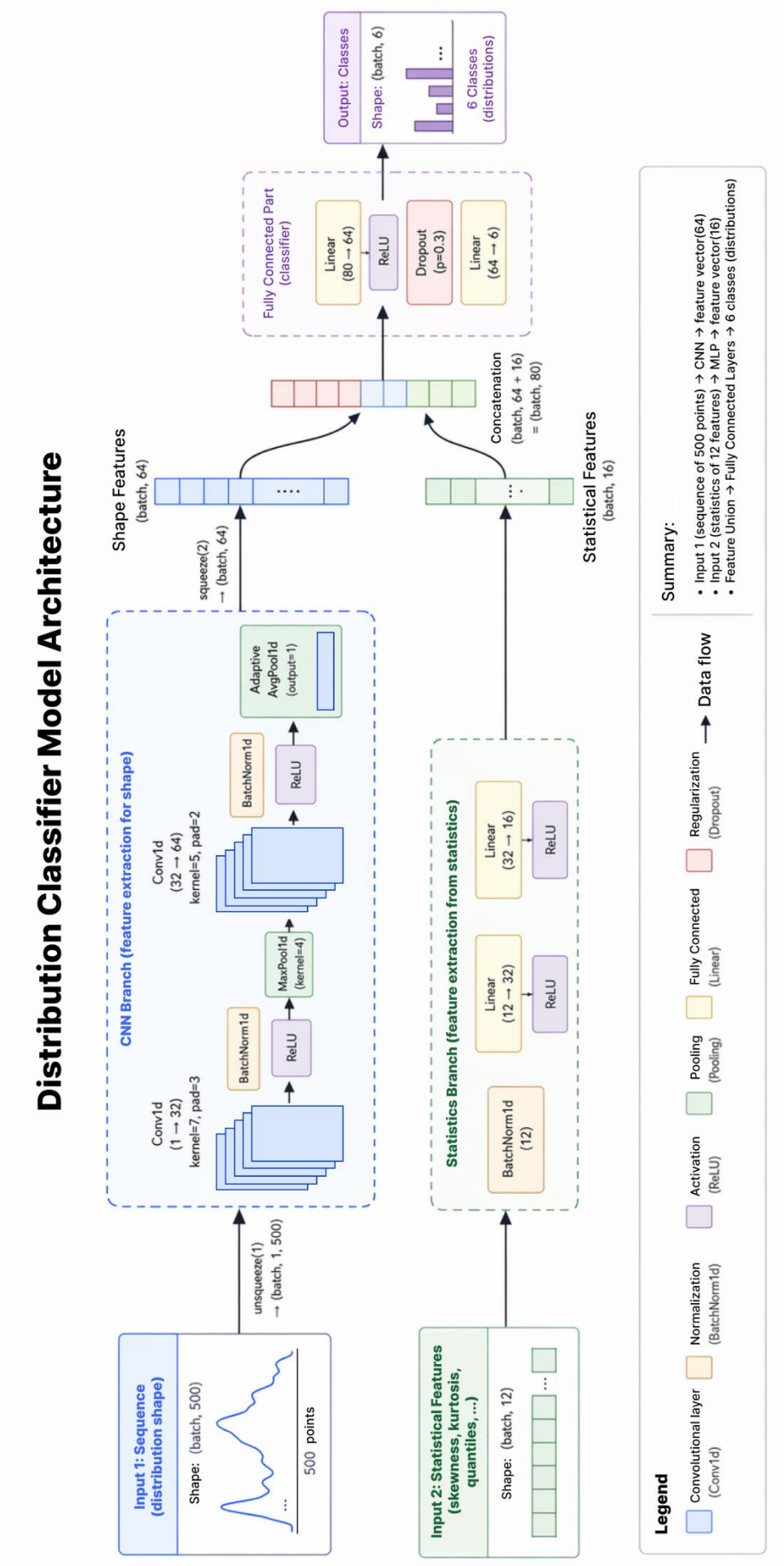


Figure 1 – Architecture of the proposed hybrid Distribution Classifier. The model integrates a one–dimensional convolutional neural network (1D CNN) for extracting structural features from the empirical distribution function with a multilayer perceptron (MLP) for processing a vector of robust statistical features.

The sorting and normalization procedures ensure invariance with respect to translation and scaling of the original data, whereas the resampling step enables samples of different lengths to be processed within a unified CNN architecture.

Simultaneously with the preparation of the input sequence for the convolutional branch, a feature vector s consisting of 12 statistical descriptors is computed. These nonparametric features characterize the shape of the underlying probability distribution and serve as input to the multilayer perceptron (MLP). The selected features were chosen based on their robustness, discriminative capability, and statistical informativeness, enabling reliable characterization of distribution properties even under limited sample sizes and noisy conditions. The complete set of statistical features is presented in Table 2, where IRQ=Q_3–Q_1 is the interquartile range and where Q_1 is the first quartile, Q_3 is the third quartile, Q_{99} is the 99th percentile.

Table 2

Statistical Features Used for the MLP Branch and Their Rationale

No.	Feature	Formula	Rationale
1	Skewness	$\text{Skewness} = \frac{1}{n}\sum_{i=1}^{n}(\frac{x_i - \bar{x}}{s})^3$	Characterizes the degree of asymmetry of the probability distribution.
2	Log–transformed kurtosis	$\tilde{k} = sign(k)\ln(1 + \lvert K\rvert)$	Characterizes tail heaviness and peakedness of the distribution while reducing the influence of extreme kurtosis values.
3	Median–to–mean ratio	$R_{med/mean} = \frac{median(x)}{\bar{x}}$	Reflects distribution asymmetry and the sensitivity of the sample mean to outliers.
4	Tail weight	$\text{T} = \frac{Q_{95} - Q_5}{IQR}$	Characterizes the heaviness of the distribution tails.
5	Quantile peakedness index	$P = \frac{IQR}{Q_{90} - Q_{10}}$	Characterizes the peakedness of the probability distribution.
6	Logarithm of sample size	$L_n = \log_{10}(n)$	Represents the sample size on a logarithmic

ISSN 1562-9945 (Print)
ISSN 2707-7977 (Online)

No.	Feature	Formula	Rationale
			scale.
7	Normalized first quartile (Q1)	$\frac{Q_1}{s}$	Characterizes the position of the first quartile relative to the scale of variation.
8	Normalized 5th percentile	$\frac{Q_5}{s}$	Characterizes the position of the 5th percentile relative to the scale of variation.
9	Normalized median (Q50)	$\frac{Q_{50}}{s}$	Characterizes the central tendency of the distribution in normalized form.
10	Normalized 95th percentile	$\frac{Q_{95}}{s}$	Characterizes the position of the 95th percentile relative to the scale of variation.
11	Normalized 99th percentile	$\frac{Q_{99}}{s}$	Characterizes the position of the 99th percentile relative to the scale of variation.
12	Normalized range	$\frac{\max(x) - \min(x)}{s}$	Characterizes the normalized spread (variation width) of the distribution.

The proposed Distribution Classifier model consists of two parallel branches whose outputs are concatenated and passed to the final classifier (Fig. 2). A detailed description of the network layers is provided in Table 3.

The model was trained using the cross–entropy loss function and the Adam optimizer with an initial learning rate of $\eta = 4 \times 10^{-4}$. An adaptive learning rate scheduling strategy (ReduceLROnPlateau) was employed with a reduction factor of 0.5 and patience = 5 epochs. The batch size was set to 64, and the maximum number of training epochs was 300.

The architecture of the proposed model is summarized in Table 3. The CNN branch is designed to analyze the shape of the probability distribution. A one–dimensional input signal of length 500, representing the normalized and sorted sample, is processed by a sequence of one–dimensional convolutional layers that automatically learn local structural patterns, such as slopes, inflection points, plateaus, and other characteristic features of empirical distribution functions corresponding to different probability distributions.

Table 3

Architecture of the proposed Distribution Classifier model

Block	Layer	Output Size	Parameters
CNN branch	Conv1d(1, 32, kernel = 7, padding = 3)	(batch, 32, 500)	256
	BatchNorm1d + ReLU	(batch, 32, 500)	64
	MaxPool1d(4)	(batch, 32, 125)	–
	Conv1d(32, 64, kernel = 5, padding = 2)	(batch, 64, 125)	10,304
	BatchNorm1d + ReLU	(batch, 64, 125)	128
	AdaptiveAvgPool1d(1)	(batch, 64, 1)	–
	Squeeze	(batch, 64)	–
Statistical branch	BatchNorm1d(12)	(batch, 12)	24
	Linear(12 → 32)	(batch, 32)	416
	ReLU	(batch, 32)	–
	Linear(32 → 16)	(batch, 16)	528
	ReLU	(batch, 16)	–
Classifier	Concatenate (CNN, Statistical Features)	(batch, 80)	–
	Linear(80 → 64)	(batch, 64)	5,184
	ReLU	(batch, 64)	–
	Dropout (p = 0.3)	(batch, 64)	–
	Linear(64 → 6)	(batch, 6)	390
Total	Trainable parameters		17,290

The statistical branch processes the numerical features. A vector of 12 statistical features (Table 2) is fed into a fully connected neural network. Unlike the CNN branch, it does not directly analyze the shape of the empirical distribution; instead, it operates on mathematically derived statistical descriptors that, under sufficiently large sample sizes, theoretically provide sufficient information to distinguish between different probability distributions.

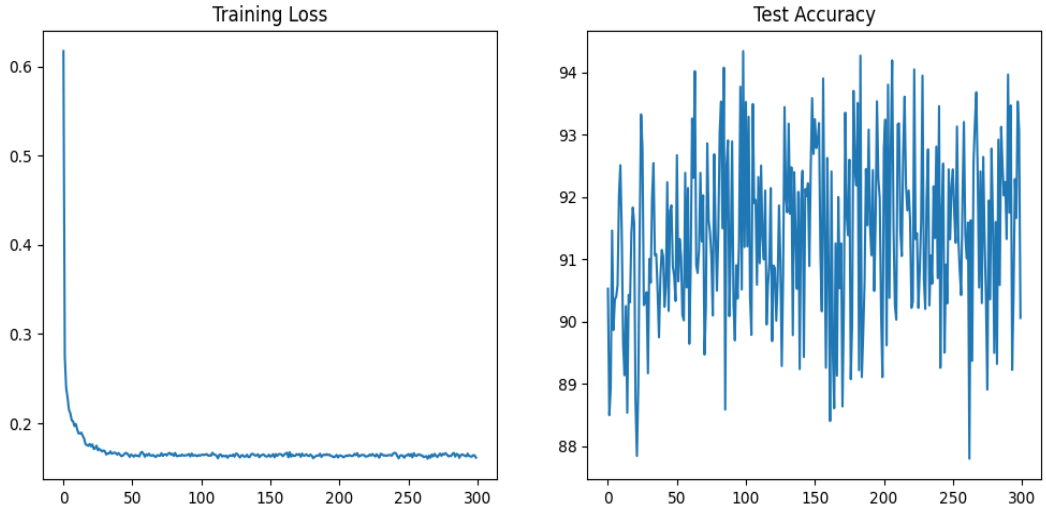
The model is trained by minimizing the multiclass cross–entropy loss function:

$$\mathcal{L} = -\sum_{c=1}^{6} y_c \log \hat{y}_c,$$

ISSN 1562-9945 (Print)
ISSN 2707-7977 (Online)

where $y_c \in \{0,1\}$ denotes the ground–truth class label, and $\hat{y}_c$is the predicted probability of class C.

The complete model contains 17,290 trainable parameters, making it computationally lightweight, easy to train, and suitable for fast inference on standard CPU–based hardware.
Fig. 2 illustrates the convergence behavior of the proposed model by presenting the evolution of the training loss on the training dataset throughout the learning process (the horizontal axis is the learning epoch, the vertical axis is the value of the cross–entropy loss function).



a b

Figure 2 – Evolution of the training loss on the training dataset as a function of the number of training epochs: a – Train Loss , b – Train Accuracy

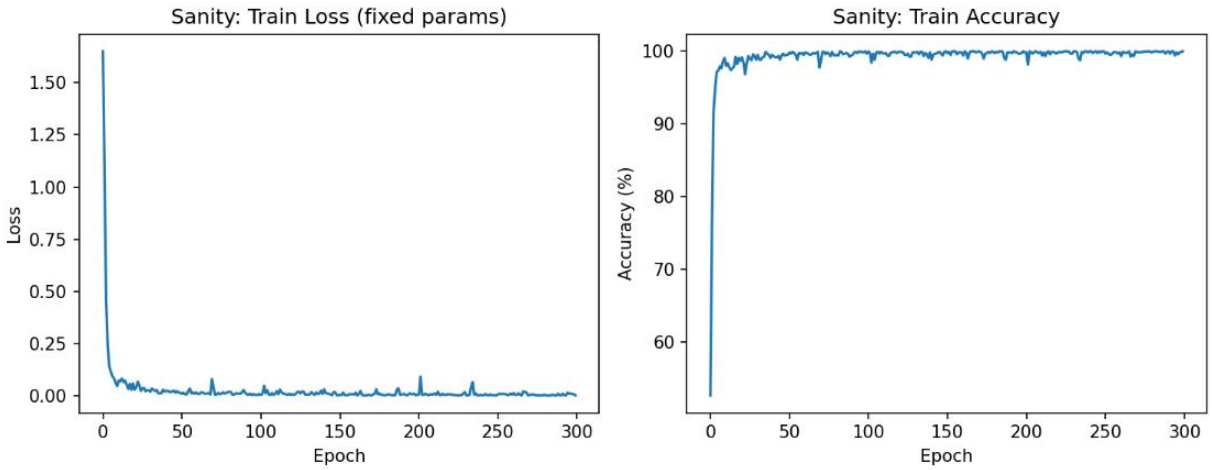
The progressive decrease in the training loss indicates that the proposed model successfully learns the underlying patterns in the training data and improves its approximation capability throughout the optimization process. The corresponding classification accuracy on the test dataset is shown on the right. The simultaneous decrease in the loss function and increase in classification accuracy demonstrate stable convergence of the learning process and confirm the effectiveness of the proposed architecture.

To further verify that the proposed architecture has sufficient representational capacity and can learn correctly, an additional sanity–check experiment was conducted. In this experiment, a simplified training dataset was generated in which the parameters of each of the six probability distributions were fixed. Specifically, the parameters were set as follows: Normal distribution ($\mu = 0$, $\sigma = 1$); Student's t distribution ($df = 3$); Logistic distribution ($\mu = 0$, $s = 1$); Binomial distribution ($n = 50$, $p = 0.5$); Exponential distribution ($\lambda = 1$, equivalent to scale = 1); and Uniform distribution ($a = -5$, $b = 5$). The dataset consisted of 200 samples per class, and the model was trained for 300 epochs without applying regularization.

As shown in Fig.3, under this simplified experimental setting the training loss decreases monotonically toward zero, while the training classification accuracy reaches 100% after approximately 50 epochs.

These results confirm that:

- the proposed model has sufficient representativeness for a small training data set with fixed distribution parameters, which confirms the obtained results (Fig. 2 and Fig. 3);
- the optimization process, including gradient–based learning and the multiclass cross–entropy loss function, operates correctly. Therefore, the non–zero loss and imperfect accuracy observed in the primary experiments rather reflect the high complexity of the target problem, where distribution parameters vary wide and there is a considerable overlap between classes.



a b

Figure 3 – Sanity check demonstrating the overfitting capability of the proposed model using fixed distribution parameters: a – Train Loss , b – Train Accuracy

The results of the sanity check therefore validate the correctness of the proposed implementation and confirm that the limitations observed in the main experiments are determined by the intrinsic difficulty of the probability distribution identification task rather than by insufficient model capacity or training instability.

Conclusions. This study is devoted to the development and experimental verification of a hybrid neural network architecture using for classifying parametric probability distributions. The proposed model demonstrated high generalization capability over a wide range of sample sizes (20–1000 observations). The experimental results showed an overall classification accuracy exceeding 97% for six classes of statistical distributions, confirming the effectiveness of the proposed approach for both limited and medium–sized datasets.

A structural analysis of the proposed architecture enabled a formal assessment of the contribution of its principal components to the overall classification performance. The 1D convolutional neural network (1D CNN) automatically extracts informative features from the shape of the empirical distribution function, capturing local structural patterns that characterize different probability distributions. In parallel, the multilayer perceptron (MLP) processes a vector of twelve robust statistical features, providing complementary quantitative information about the distribution. The experimental evaluation demonstrated that integrating these two components within a unified multimodal architecture yields a statistically significant improvement in classification accuracy compared to using either branch independently.

ISSN 1562-9945 (Print)
ISSN 2707-7977 (Online)

The obtained results confirm the effectiveness of combining deep learning techniques with methods of mathematical statistics for automatically identifying probability distributions. The proposed hybrid architecture exhibits increased robustness to variations in sample size and effectively distinguishes statistically similar distribution classes.

A comparative evaluation against conventional goodness–of–fit tests demonstrated the superiority of the proposed approach, particularly for small sample sizes, where classical statistical methods have reduced statistical power. The greatest performance improvement was observed under conditions of limited data size and partial overlapping between the statistical characteristics of different probability distributions.

The analysis of classification performance with respect to sample size confirmed the expected increase in accuracy as the number of observations grows. Nevertheless, even for small datasets, the proposed model provides stable and reliable results owing to the combined analysis of the empirical distribution shape and robust statistical descriptors. These findings indicate the model's robustness to noisy measurements, data variability, and the statistical similarity of different probability distributions.

The proposed architecture has moderate computational complexity and can be employed as an intelligent module for automatic statistical model selection in information–measuring systems, decision support systems, computer–integrated systems, and cyber–physical systems, where rapid identification of the underlying probability distribution is required without expert intervention.

ЛІТЕРАТУРА

1. Трифонова К. Виявлення прихованих даних в зображеннях на основі критерію Колмогорова–Смирнова // Інформаційні управляючі системи і технології (ІУСТ–Одеса–2025) : матеріали конференції. Одеса, 2025. С. 57–59. URL: https://icst–conf.com/2025.pdf (дата звернення: 11.07.2026).

2. Тулизик О., Рудніченко М. Аналіз методів складання розкладу за допомогою генетичних алгоритмів, алгоритму відпалу, ройового інтелекту та нейронних мереж // Інформаційні управляючі системи і технології (ІУСТ–Одеса–2025) : матеріали конференції. Одеса, 2025. С. 204–206. URL: https://icst–conf.com/2025.pdf (дата звернення: 11.07.2026).

3. Шведов Д., Рудніченко М., Петров І. Концепція гібридної моделі глибинного навчання для аналізу фінансових ризиків // Інформаційні управляючі системи і технології (ІУСТ–Одеса–2025) : матеріали конференції. Одеса, 2025. С. 209–211. URL: https://icst–conf.com/2025.pdf (дата звернення: 11.07.2026).

4. Судин А., Сцісло Л., Перекрест А., Онищенко Ю. Гібридна нейромережева модель оцінки ризику таргетованих атак на SCADA/ICS системи // Інформаційні технології та суспільство. 2025. № 4(19). С. 162–169. DOI: https://doi.org/10.32689/maup.it.2025.4.25

5. Іванюк І., Гнатюк Д. Дослідження вразливостей та захисту в IoT–мережах із використанням ШІ // Тези доповідей Чернівецького національного університету. 2025. С. 158–159. URL: https://www.chnu.edu.ua/media/leeld04z/2025–nniftkn_compressed.pdf (дата звернення: 11.07.2026)

6. Abbasimehr H., Paki R., Bahrini A. Improving the performance of deep learning models using statistical features // Mathematical Methods in the Applied Sciences. 2021. Vol. 49, No. 5. P. 3496–3510. DOI: https://doi.org/10.1002/mma.7500
7. Fawaz H. I., Lucas B., Forestier G. та ін. InceptionTime: Finding AlexNet for time series classification // Data Mining and Knowledge Discovery. 2020. Vol. 34. P. 1936–1962. DOI: https://doi.org/10.1007/s10618–020–00710–y
8. Ванотчак О. Розробка гібридної моделі CNN–LSTM для прогнозування ускладнень у бурінні // Вісник Хмельницького національного університету. 2026. № 363(2). С. 125–134. DOI: https://doi.org/10.31891/2307–5732–2026–363–16
9. Супруні О. М., Огірко О. І., Кримська А. О. Розроблення гібридних нейромережевих моделей для прогнозування // Вісник Херсонського національного технічного університету. 2025. № 1(92), ч. 2. С. 222–230. DOI: https://doi.org/10.35546/kntu2078–4481.2025.1.2.30
10. Klymenko O., Vovk S. Comparative analysis of the normality of statistical criteria for samples of contaminated data // Journal of Rocket–Space Technology. 2025. Vol. 34, No. 4. P. 97–104. DOI: https://doi.org/10.34185/1562–9945–5–160–2025–03
11. Klymenko O. Sensitivity analysis of statistical criteria to pseudorandom variable generation parameters // Journal of Rocket–Space Technology. 2025. Vol. 34, No. 2. P. 122–130. DOI: https://doi.org/10.15421/452521
12. Klymenko O., Malaichuk V. Methods of assessing the conformity of a pseudo–random sequence to a random distribution in the problems of automation, cybersecurity, computer–integrated technologies and non–destructive testing // Інформаційні технології в металургії та машинобудуванні : матеріали Міжнародної науково–технічної конференції. 2024. С. 327–334. DOI: https://doi.org/10.34185/1991–7848.itmm.2024.01.062
13. Клименко О., Астахов Д., Клименко С. Теоретичні дослідження вимірювань за допомогою функцій розподілу ймовірностей // Journal of Rocket–Space Technology. 2021. Vol. 28, No. 4. P. 152–158. DOI: https://doi.org/10.15421/452020
14. Моцний Ф. В. Статистичні розподіли хі–квадрат, Стьюдента, Фішера–Снедекора та їх застосування // Статистика України. 2018. № 1(80). С. 16–23. DOI: https://doi.org/10.31767/su.1(80).2018.01.02

REFERENCES

1. Tryfonova, K. (2025). Detection of hidden data in images based on the Kolmogorov–Smirnov criterion. In Information Control Systems and Technologies (ICST–Odesa–2025) (pp. 57–59). Odesa, Ukraine. Available at: https://icst–conf.com/2025.pdf
2. Tulyzyk, O., & Rudnichenko, M. (2025). Analysis of scheduling methods using genetic algorithms, simulated annealing, swarm intelligence, and neural networks. In Information Control Systems and Technologies (ICST–Odesa–2025) (pp. 204–206). Odesa, Ukraine. Available at: https://icst–conf.com/2025.pdf
3. Shvedov, D., Rudnichenko, M., & Petrov, I. (2025). Concept of a hybrid deep learning model for financial risk analysis. In Information Control Systems and Technologies

ISSN 1562-9945 (Print)
ISSN 2707-7977 (Online)

(ICST–Odesa–2025) (pp. 209–211). Odesa, Ukraine. Available at: https://icst–conf.com/2025.pdf
4. Sudyn, A., Stsislo, L., Perekrest, A., & Onyshchenko, Y. (2025). Hybrid neural network model for targeted attack risk assessment in SCADA/ICS systems. Information Technologies and Society, 4(19), 162–169. https://doi.org/10.32689/maup.it.2025.4.25
5. Ivaniuk, I., & Hnatiuk, D. (2025). Study of vulnerabilities and protection in IoT networks using artificial intelligence. In Proceedings of Chernivtsi National University (pp. 158–159). Available at: https://www.chnu.edu.ua/media/leeld04z/2025–nniftkn_compressed.pdf
6. Abbasimehr, H., Paki, R., & Bahrini, A. (2021). Improving the performance of deep learning models using statistical features. Mathematical Methods in the Applied Sciences, 49(5), 3496–3510. https://doi.org/10.1002/mma.7500
7. Fawaz, H. I., Lucas, B., Forestier, G., Pelletier, C., Schmidt, D. F., Weber, J., Webb, G. I., Idoumghar, L., Muller, P.–A., & Petitjean, F. (2020). InceptionTime: Finding AlexNet for time series classification. Data Mining and Knowledge Discovery, 34, 1936–1962. https://doi.org/10.1007/s10618–020–00710–y
8. Vanotchak, O. (2026). Development of a hybrid CNN–LSTM model for drilling complication prediction. Herald of Khmelnytskyi National University, 363(2), 125–134. https://doi.org/10.31891/2307–5732–2026–363–16
9. Supruni, O. M., Ohirko, O. I., & Krymska, A. O. (2025). Development of hybrid neural network models for forecasting. Bulletin of Kherson National Technical University, 1(92), Part 2, 222–230. https://doi.org/10.35546/kntu2078–4481.2025.1.2.30
10. Klymenko, O., & Vovk, S. (2025). Comparative analysis of the normality of statistical criteria for samples of contaminated data. Journal of Rocket–Space Technology, 34(4), 97–104. https://doi.org/10.34185/1562–9945–5–160–2025–03
11. Klymenko, O. (2025). Sensitivity analysis of statistical criteria to pseudorandom variable generation parameters. Journal of Rocket–Space Technology, 34(2), 122–130. https://doi.org/10.15421/452521
12. Klymenko, O., & Malaichuk, V. (2024). Methods of assessing the conformity of a pseudorandom sequence to a random distribution in the problems of automation, cybersecurity, computer–integrated technologies and non–destructive testing. In Proceedings of the International Scientific and Technical Conference "Information Technologies in Metallurgy and Mechanical Engineering" (pp. 327–334). https://doi.org/10.34185/1991–7848.itmm.2024.01.062
13. Klymenko, O., Astakhov, D., & Klymenko, S. (2021). Theoretical studies of measurements using probability distribution functions. Journal of Rocket–Space Technology, 28(4), 152–158. https://doi.org/10.15421/452020
14. Motsnyi, F. V. (2018). Chi–square, Student's t, and Fisher–Snedecor statistical distributions and their applications. Statistics of Ukraine, 80(1), 16–23. https://doi.org/10.31767/su.1(80).2018.01.02

Received 04.06.2026.
Accepted 06.08.2026.
Published 30.10.2026

Гібридна нейромережева модель ідентифікації ймовірнісних розподілів у задачах оцінки стану природних і технічних об'єктів

Анотація. У статті запропоновано методику автоматизованого визначення типу закону розподілу на основі гібридної нейромережевої архітектури, яка поєднує оброблення кількох інформаційних представлень даних. Гібридна нейромережева модель ідентифікації ймовірнісних розподілів у задачах оцінки стану природних і технічних об'єктів складається з двох модулів. Перший модуль архітектури реалізовано у вигляді одновимірної згорткової нейронної мережі, що здійснює аналіз емпіричної функції розподілу та форми вибірки. Другий модуль базується на багатошаровому перцептроні, який обробляє вектор із 12 робастних статистичних параметрів, включаючи моментні оцінки, коефіцієнти асиметрії та ексцесу, а також характеристики розсіювання. Інтеграція результатів цих двох модулів у єдиній архітектурі дозволяє підвищити точність класифікації та забезпечити стійкість до малих вибірок, шумів і часткового перекриття характеристик різних розподілів. Методика реалізована на основі поєднання методів теорії ймовірностей, математичної статистики, глибокого навчання та комп'ютерного моделювання. Для навчання та тестування моделі сформовано набір із 60 000 синтетичних вибірок обсягом від 20 до 1000 елементів із випадковими параметрами. Модель забезпечує класифікацію шести поширених параметричних розподілів: нормального, логістичного, Стьюдента, біноміального, експоненційного та рівномірного. Наукова новизна роботи полягає у розробленні гібридного підходу до автоматичної ідентифікації законів розподілу, що базується на одночасному використанні емпіричної функції розподілу та статистичних параметрів у межах єдиної нейромережевої архітектури. Запропонований підхід забезпечує підвищену точність класифікації для малих вибірок та складних випадків близькості розподілів, що відрізняє його від відомих методів. Практична значущість роботи полягає у створенні програмної моделі, придатної для інтеграції в системи аналізу даних, інтелектуальні інформаційні системи, системи підтримки прийняття рішень, а також у задачі автоматизації статистичного аналізу в комп'ютерно–інтегрованих і кіберфізичних системах. Запропоноване рішення дозволяє автоматизувати вибір статистичної моделі без залучення експерта. Експериментальні результати підтверджують ефективність запропонованого підходу: досягнуто точності класифікації 97,8% на тестовому наборі з 18 000 вибірок, що перевищує результати окремих моделей CNN та MLP, а також класичних статистичних критеріїв. Для малих вибірок (n = 20) точність становить 59,6%, що суттєво перевищує показники критерію Колмогорова–Смирнова (36,4%). Отримані результати підтверджують доцільність гібридного використання нейромережевих підходів для задач автоматичного визначення типів статистичних розподілів в умовах обмеженого обсягу даних.

Ключові слова: гібридна нейронна мережа; ідентифікація ймовірнісних розподілів; емпірична функція розподілу; статистичні ознаки; глибоке навчання; математична статистика.

ISSN 1562-9945 (Print)
ISSN 2707-7977 (Online)

Клименко Олексій Денисович – здобувач третього (освітньо–наукового) рівня вищої освіти кафедри кібербезпеки і комп'ютерно–інтегрованих технологій фізико–технічного факультету Дніпровського національного університету імені Олеся Гончара.
ORCID https://orcid.org/0009–0009–5164–1688
Вовк Сергій Михайлович – канд. фіз.–мат. наук, доцент, Дніпровський національний університет імені Олеся Гончара.
ORCID https://orcid.org/0000–0002–8116–3500.

Klymenko Oleksii Denysovych – postgraduate at the department of cybersecurity and computer–integrated technologies, faculty of physics and technology, Oles Honchar Dnipro national university.
ORCID https://orcid.org/0009–0009–5164–1688
Vovk Sergiy Mykhailovych – PhD in Physics and Mathematics, Associate Professor, Oles Honchar Dnipro National University.
ORCID https://orcid.org/0000–0002–8116–3500.