# ТЕХНОЛОГИЧЕСКИЕ ОСНОВЫ ПОВЫШЕНИЯ НАДЕЖНОСТИ И КАЧЕСТВА ИЗДЕЛИЙ

# TECHNOLOGICAL BASIS FOR IMPROVING RELIABILITY AND PRODUCT QUALITY

**УДК 004.838.2** **DOI 10.21685/2307-4205-2020-3-9**

**А. И. Иванов, И. А. Кубасов, А. М. Самокутяев**

## ТЕСТИРОВАНИЕ БОЛЬШИХ НЕЙРОННЫХ СЕТЕЙ НА МАЛЫХ ВЫБОРКАХ

**A. I. Ivanov, I. A. Kubasov, A. M. Samokutyaev**

## TESTING LARGE NEURAL NETWORKS ON SMALL SAMPLES

***Аннотация.*** *Актуальность и цели.* Исследованы проблемы тестирования нейронных сетей, применяемых в интересах повышения надежности и качества сложных технических систем. Выявлены условия, при которых возможно оперативное и корректное тестирование качества принимаемых решений большими нейронными сетями на малых выборках. *Материалы и методы*. Выполнена оценка вероятности ошибок первого рода (ошибочного отказа от признания образа «Свой») на основе тестирования без сокращения тестовой выборки. Показано, что для нейросетевых решений в форме бинарного кода оценка вероятности ошибок второго рода (ошибочное принятие образа «Чужой») может быть осуществлена при значительном сокращении объема тестовой выборки. *Результаты и выводы.* Выявлено логарифмическое снижение объема тестовой выборки при переходе от статистического анализа обычных кодов к статистическому анализу расстояний Хэмминга между кодом образа «Свой» и кодами образов «Чужой». Представлена математическая модель вычисления вероятностей ошибок второго рода доверенного нейросетевого приложения на малых выборках. Обоснована необходимость дальнейшей стандартизации доверенных приложений искусственного интеллекта, позволяющих повысить надежность и качество сложных технических систем.

***Abstract.*** *Background.* The paper investigated the problems of testing neural networks used to improve the reliability and quality of complex technical systems. The conditions under which rapid and correct testing of the quality of decisions made by large neural networks on small samples is possible have been revealed. *Materials and methods.* The probability of errors of the first kind (erroneous rejection of recognition of the "Own" image) was estimated based on testing without reducing the test sample. It has been shown that for neural network solutions in the form of a binary code, the probability of errors of the second kind (erroneous adoption of the "Alien" image) can be estimated with a significant reduction in the volume of the test sample. *Results and conclusions.* A logarithmic decrease in the volume of the test sample was revealed when moving from statistical analysis of ordinary codes to statistical analysis of Hamming distances between the image code "Own" and the image codes "Alien." The mathematical model of calculation of probabilities of errors of the second kind of trusted neural network application on small samples is presented. The need for further standardization of trusted applications of artificial intelligence is justified, allowing to increase the reliability and quality of complex technical systems.

***Ключевые слова*:** искусственный интеллект, нейронные сети, доверенные нейросетевые приложения, объем тестовой выборки.

***Keywords***: artificial intelligence, neural networks, trusted neural network applications, test sample size.

## Введение

В соответствии с Указом Президента РФ В. В. Путина № 480 от 10 октября 2019 г. «О развитии искусственного интеллекта в Российской Федерации» в период с 2020 по 2030 г. выделяются значительные материальные ресурсы на развитие этой технологической ветви. При этом к искусственному интеллекту, применяемому в интересах повышения надежности и качества сложных технических систем, должны предъявляться особые требования. По аналогии с вычислениями, выполняемыми в доверенной вычислительной среде, следует выделять и доверенный искусственный интеллект. Если приложение искусственного интеллекта создано в интересах повышения надежности и качества ведомственных сложных технических систем, то обучать и тестировать достигнутое качество принимаемых решений должны сотрудники данного ведомства. Это коренное отличие ведомственных доверенных приложений искусственного интеллекта.

Еще важнейшими характеристиками являются продолжительность обучения доверенного искусственного интеллекта и продолжительность его тестирования. Скорость обучения, как правило, связана с объемом обучающей выборки. Так, современные алгоритмы глубокого обучения сверточных нейронных сетей требуют миллионы образов, размеченных «в ручную», и огромных затрат вычислительных ресурсов [1]. Дополнительно при тестировании сетей глубокого обучения также требуются огромные тестовые базы, сформированные уже без ручной разметки.

Читая зарубежную литературу по искусственному интеллекту, может создаться впечатление, что чем «умнее» и «многослойнее» сеть искусственных нейронов, тем сложнее ее обучение и тестирование. Это, действительно, было так для нейросетевых решений прошлого века. Глубокие нейронные сети и алгоритм их обучения придумал наш соотечественник А. И. Галушкин, а англичанин Джеффри Хинтон уже в 2006 г. слегка «доработал» обучение первых слов нейронных сетей.

Авторы данной статьи уверены в том, что нейросетевые решения прошлого века не совсем подходят для разработки нейросетевых доверенных вычислений. Для этой цели выгоднее применять архитектуру нейронных сетей, созданную в соответствии с требованиями пакета отечественных национальных стандартов ГОСТ Р 52633.xx-20xx. Причина такой уверенности проста: современные архитектуры нейронных сетей удается быстро и автоматически обучать, а также тестировать качество работы на малых выборках – всего 20 примеров. Следовательно, отпадает необходимость использования миллионов примеров, предварительно размеченных «вручную». Все инструменты управления нейросетевым искусственным интеллектом при использовании современных архитектур оказываются в руках владельца искусственного интеллекта. При этом формирование тестовой и обучающей выборки вполне по силам одному человеку-эксперту.

В данной статье исследованы условия, при которых возможно оперативное и корректное тестирование качества принимаемых решений доверенными нейронными сетями на малых выборках, и представлено математическое описание данной процедуры.

## Оценка уровней доверия к человеку-эксперту и нейросетевым решениям искусственного интеллекта

В настоящее время экспертиза надежности и качества сложных технических систем строится, в основном, на выводах людей-экспертов. Доверие к результатам экспертизы опирается на ряд факторов, таких как: уровень образования эксперта, наличие у эксперта аппаратно-программных средств, опыт эксперта, доверие к методике проведения той или иной экспертизы. Люди-эксперты в силу своей природы не могут выполнять большие объемы работ в очень короткие сроки. Все эти проблемы могут быть ослаблены, если создать приложения искусственного интеллекта под решение той или иной конкретной задачи.

Заметим также, что человек-эксперт не может дать достоверную вероятностную оценку ошибок принятого им решения. В этом отношении приложения искусственного интеллекта оказываются в более выгодном положении. Каждое нейросетевое решение, обученное алгоритмом

ГОСТ Р 52633.5 [2], удается быстро и автоматически тестировать как на больших, так и на малых тестовых выборках.

Для примера рассмотрим частный случай проведения почерковедческой экспертизы. Для тестирования уровня доверия к решениям конкретного эксперта формально можно сформировать, например, 10 000 тестовых заданий, заранее зная верный результат. Однако на практике столкнемся с проблемой низкой производительности «ручного» сложного труда эксперта. Проведение «ручной» почерковедческой экспертизы нескольких слов записки или одного автографа под документом может занимать десятки часов рабочего времени эксперта, т.е. на выполнение этих заданий может потребоваться несколько лет работы квалифицированного эксперта, что нерационально.

Рациональный путь решения проблемы найден. Для использования в учебном процессе российских учебных заведений создан свободно распространяемый программный продукт «БиоНейроАвтограф» [3], который позволяет видеть 416 биометрических параметров динамики воспроизведения конкретного рукописного слова [4].

### Классическое решение задач распознавания образов

Как показывают результаты исследования, стандартное отклонение большинства из 416 биометрических параметров оказывается примерно в три раза меньше, чем стандартное отклонение всех биометрических параметров всего словаря образов «Чужой». Эта ситуация отображена на рис. 1.

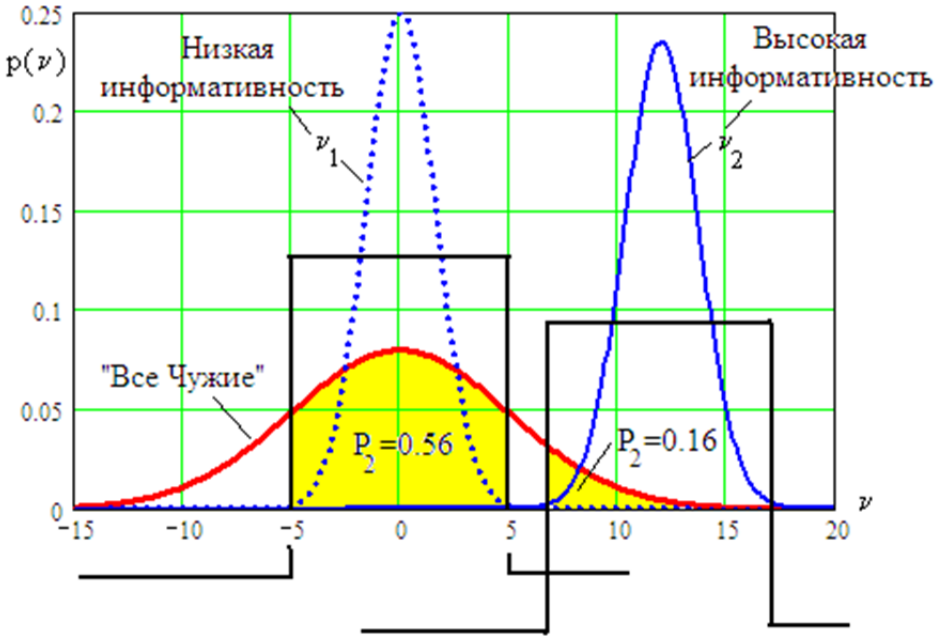


Рис. 1. Функции плотностей распределения значений двух биометрических параметров с низкой и высокой информативностью

Как видно из рис. 1, биометрические параметры могут быть выделены простым решающим правилом. Всегда можно вычислить математическое ожидание $i$-го биометрического параметра $E(\nu_i)$ и его стандартное отклонение $\sigma(\nu_i)$. Если гипотеза нормального распределения биометрических параметров верна, то интервал $\pm 3\sigma(\nu_i)$ вокруг математического ожидания должен накрывать состояния биометрического параметра с вероятностью 0.997. То есть попадание 416 параметров образа «Все Чужие» в 416 их допустимые интервалы $\{E(\nu_i) \pm 3\sigma(\nu_i)\}$ с высокой вероятностью свидетельствует об обнаружении образа «Все Чужие» (вероятность ошибок второго рода $P_2 \approx 0{,}003$ для каждого биометрического параметра). При этом вероятность ошибок второго рода $P_2$ будет зависеть от положения математического ожидания биометрического параметра. Самая высокая вероятность ошибок второго рода будет наблюдаться у биометрических параметров с математическим ожиданием в центре распределения данных «Все Чужие». Чем больше математическое ожидание биометрического параметра удалено от центра данных «Все Чужие», тем меньше оказывается вероятность ошибки

второго рода. Учитывая это, можно оценить информативность биометрических параметров, логарифмируя вероятности ошибок второго рода:

$$I(\nu_i) = -\log_2(P_2(\nu_i)), \quad (1)$$

где $I(\nu_i)$ – показатель информативности биометрического параметра $\nu_i$, вычисляемый как энтропия по Шеннону для алфавита из двух символов «0» и «1».

Из рис. 2 видно, что из 416 биометрических параметров 336 обладают низкой информативностью (менее одного бита), а только 80 биометрических параметров обладают приемлемой информативностью (более одного бита).

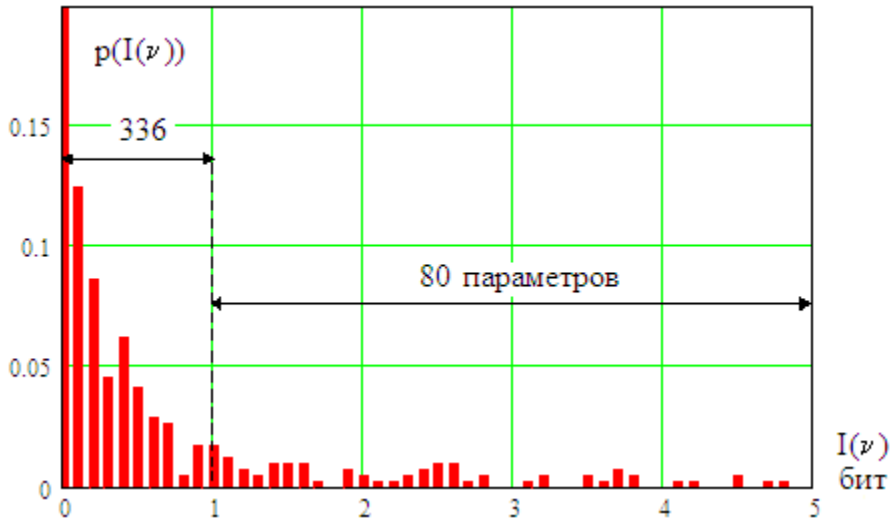


Рис. 2. Пример плотности распределения информативности 416 биометрических параметров динамики воспроизведения рукописного образа подписи

Если пользоваться парадигмой прошлого века, то следует отсечь бритвой Оккама[1] 336 биометрических параметров с низкой информативностью и работать только с 80 биометрическими параметрами приемлемой информативности. Эти 80 информативных биометрических параметров можно преобразовать в 80-битный криптографический ключ. Однако разряды такого ключа будут очень чувствительны к нестабильности примеров биометрического образа. По данным численного эксперимента, проведенного в среде моделирования «БиоНейроАвтограф» [3], каждый пример рукописного образа подписи будет давать примерно 16 % ошибок в разрядах 80-битного ключа.

Обнаружить и исправить 16 % ошибок можно за счет применения кодов с 20-кратной избыточностью [5]. В итоге получим криптографический ключ длиной 80/20 = 4 бита. Усечение криптографического ключа бритвой Оккама с 416 до 80 бит и последующее 20-кратное сжатие его длины до 4 бит представляет собой слабую технологию, так называемую «нечетких экстракторов» [6].

Следует заметить, что использование Россией технологии «нечетких экстракторов» будет означать неспособность защитить персональные биометрические данные своих граждан.

### Отказ «нечетких экстракторов» в пользу применения отечественных сетей, состоящих из большого числа искусственных нейронов

Отечественные криптографические стандарты (на шифрование, проверку и формирование цифровой подписи) используют ключи длинной 256 бит (а не 4 бита). В связи с этим в среде моделирования «БиоНейроАвтограф» реализована сеть из 256 искусственных нейронов, каждый из которых обучен преобразовывать 24 входных биометрических параметров образа «Свой» в один бит

[1] Уильям Оккам – средневековый английский монах-философ, сформулировавший принцип: «Не следует преумножать число сущностей сверх необходимого». Этот принцип в XX в. использовался как базовый многочисленными сторонниками поиска малого числа наиболее информативных параметров и «отсечения» всех остальных, менее информативных параметров.

криптографического ключа. Обучение нейросети выполняется автоматом, реализующим алгоритм ГОСТ Р 52633.5 [2]. Во время обучения вычисляются весовые коэффициенты сумматоров нейронов, осуществляющих обогащение «сырых» биометрических данных. Благодаря обучению на 20 примерах образа «Свой» на выходах сумматоров каждого нейрона информативность обогащенных данных значительно увеличивается.

Эффект нейросетевого «обогащения» биометрических данных позволяет получать верный 256-битный код криптографического ключа с вероятностью близкой к 0,95 ($P_1 \approx 0{,}05$) при обучении нейросети на 20 примерах образа «Свой». Чем больше примеров образа «Свой» использовано при обучении нейронной сети, тем ниже оказывается вероятность ошибок первого рода. В первом приближении можно считать связь вероятности ошибок первого рода обратно пропорциональной размеру обучающей выборки, если не выполнялось дополнительное тестирование на выборке, не участвовавшей в обучении.

Если есть возможность выполнить тестирование на ограниченной тестовой выборке примеров образа «Свой», то оценка вероятности ошибок первого рода вычисляется следующим образом:

$$\begin{cases} P_1 \approx \dfrac{n}{N} \quad \text{при} \quad n \neq 0; \\ P_1 \approx \dfrac{1}{N+1} \quad \text{при} \quad n = 0, \end{cases} \tag{2}$$

где $n$ – число обнаруженных ошибок; $N$ – объем тестовой выборки.

Вторая формула системы (2) построена, исходя из предположения, что следующий опыт малой выборки может обнаружить одну ошибку. Для малых выборок такая гипотеза достаточно хорошо работает.

### Оценка вероятности ошибок второго рода на больших тестовых выборках

Описанный выше отказ от процедур учета только малого числа наиболее информативных параметров и переход к нейросетевому статистическому анализу большинства контролируемых биометрических параметров позволяет получить выходной код длиной 265 бит.

Так, если создать базу из 100 миллионов уникальных образов подписей граждан РФ, обучить нейросеть на одном образе «Свой» и подавать на ее входы случайные образы, то будем получать случайные отклики нейросети в виде случайных выходных кодов. Обученная на образе «Свой» нейросеть будет выдавать один и тот же код, если подавать на ее входы примеры образа «Свой» из тестовой выборки, несмотря на наличие в них значительных вариаций. Нейросеть обучена учитывать естественные вариации биометрических параметров образа «Свой», представленные 20 примерами обучающей выборки. Далее при последующей эксплуатации обученная нейросеть выполняет функцию, обратную функции хэширования. Она почти устраняет влияние естественной нестабильности примеров рукописного образа «Свой». Для примеров образа «Чужой» нейросеть выполняет хэширование биометрических кодов, значительно усиливая естественную нестабильность непрерывных данных примеров образа «Чужой». Кроме того, выходной код нейросети всегда дискретен, т.е. должен иметь $2^{256}$ спектральных линии с очень малыми значениями амплитуды вероятности их появления.

Формально можно попытаться оценить вероятность ошибок второго рода обученной нейронной сети, подставляя поочередно на ее входы $10^8$ (100 миллионов) образов «Все Чужие». Далее для оценок следует применить формулу похожую на формулу (2):

$$\begin{cases} P_2 \approx \dfrac{n}{N} \quad \text{при } n \neq 0; \\ P_2 \approx \dfrac{1}{N+1} \quad \text{при} \quad n = 0. \end{cases} \tag{2a}$$

К сожалению, этот методически понятный подход полного перебора большой базы тестовых образов «Чужой» с достаточно высокой вероятностью дает значительную ошибку. Ее появление обусловлено тем, что число выходных состояний нейросети – $2^{256}$ много больше, чем размер тесто-

вой базы образов «Чужой» (создать тестовою базу из $2^{256}$ образов «Чужой» технически невозможно, да и в этом нет необходимости). При высокой стабильности и высокой уникальности автографов будем всегда сталкиваться с ситуацией отсутствия в большой тестовой базе данных, дающих точное повторение кода «Свой».

### Оценка вероятности ошибок второго рода на малых тестовых выборках

Решить проблему тестирования на малых выборках удается, воспользовавшись рекомендациями ГОСТ Р 52633.3 [7]. Этот национальный стандарт рекомендует отказаться от попыток статистического анализа длинных выходных кодов обученной нейросети. Для упрощения задачи необходимо перейти в пространство расстояний Хэмминга между кодом «Чужой» и кодом «Свой»:

$$h = \sum_{i=1}^{256} x_i \oplus c_i \,, \tag{3}$$

где $\oplus$ – логическая операция сложения по модулю два; $x_i$ – значение $i$-го разряда двоичного кода образа «Чужой»; $c_i$ – значение $i$-го разряда двоичного кода образа «Свой».

Блок-схема тестирования на малых выборках приводится на рис. 3.

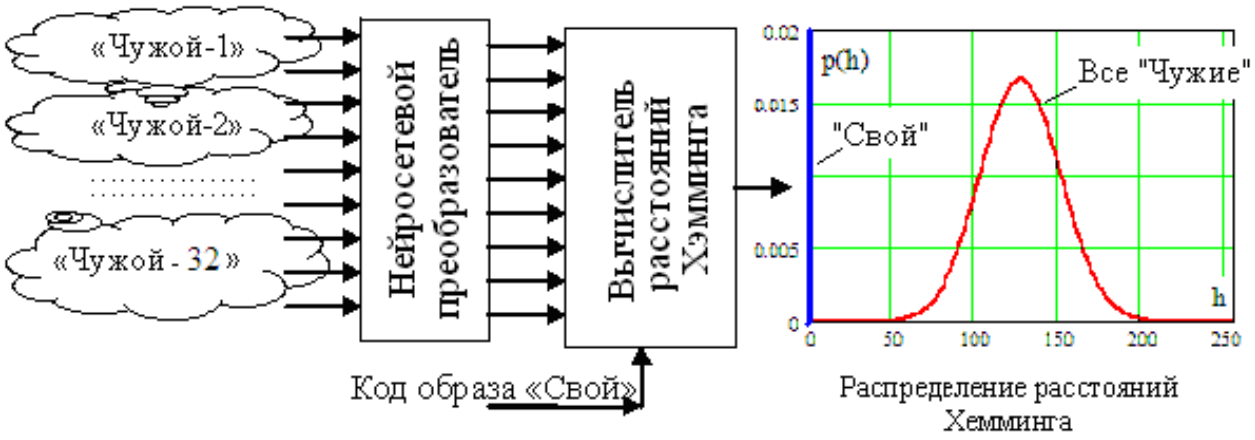


Рис. 3. Блок-схема тестирования на малых выборках в пространстве расстояний Хэмминга

Преимуществом перехода к расстояниям Хэмминга является то, что $2^{256}$ состояний выходных кодов сжимаются до 257 состояний: $h = \{0,1,\ldots,256\}$. Кроме того, распределение расстояний Хэмминга нормализуется по основной теореме статистики из-за 256-кратного суммирования случайных величин при вычислении расстояний Хэмминга (3).

Опираясь на эти обстоятельства, национальный стандарт [7] рекомендует на малой выборке вычислить математическое ожидание расстояний Хэмминга – $E(h)$ и его стандартное отклонение $\sigma(h)$. Далее в рамках гипотезы нормального закона распределения следует выполнить следующие вычисления вероятности ошибок второго рода:

$$P_2 \approx \frac{1}{\sigma(h)\sqrt{2\pi}} \int_0^1 \exp\left\{ \frac{-\left(E(h)-u\right)^2}{2\left(\sigma(h)\right)^2} \right\} \cdot du. \tag{4}$$

В итоге имеем корректную математическую модель вычисления вероятностей ошибок второго рода доверенного нейросетевого приложения на малых выборках.

### Заключение

Современные информационно-телекоммуникационные технологии развиваются стремительно. Уже сегодня возможна дистанционная надежная биометрическая аутентификация личности человека. Скорее всего в базе данных уполномоченного органа должны будут храниться образы владельцев паспорта (лицо [1], голос [8], автограф [3, 4], рисунки отпечатков пальцев [9], рисунки

радужной оболочки глаз [8]). Под каждую из биометрических технологий придется модифицировать и стандартизовать варианты сетей искусственных нейронов и разную предобработку данных для этих сетей. В этом отношении Россия сегодня является безусловным лидером, так как первой начала создавать национальные стандарты по автоматическому обучению [2] больших нейронных сетей и их автоматическому тестированию [7].

Эту работу по стандартизации необходимо активно продолжать в интересах развития доверенных приложений искусственного интеллекта, позволяющих повысить надежность и качество сложных технических систем. Наличие общедоступных отечественных стандартов должно породить конкуренцию на российском рынке, что в конечном итоге должно снизить стоимость продуктов искусственного интеллекта при качестве существенно выше мирового уровня.

## Библиографический список


1. Николенко, С. Глубокое обучение / С. Николенко, А. Кадурин, Е. Архангельская. – Санкт-Петербург : Питер, 2018. – 480 с.
2. ГОСТ Р 52633.5–2011 Защита информации. Техника защиты информации. Автоматическое обучение нейросетевых преобразователей биометрия-код доступа.
3. *Иванов, А. И.* Среда моделирования «БиоНейроАвтограф» / А. И. Иванов, О. С. Захаров. – URL: http://пниэи.рф/activity/science/noc/bioneuroautograph.zip
4. *Иванов, А. И.* Автоматическое обучение больших искусственных нейронных сетей в биометрических приложениях : учеб. пособие / А. И. Иванов. – Пенза, 2013. – 30 с. – URL: http://пниэи.рф/activity/science/noc/tm_IvanovAI.pdf
5. *Морелос-Сарагоса, Р.* Искусство помехоустойчивого кодирования / Р. Морелос-Сарагоса. – Москва : Техносфера, 2007. – 320 с.
6. *Juels, A.* A Fuzzy Commitment Scheme / A. Juels, M. Wattenberg // Proc. ACM Conf. Computer and Communications Security. – Singapore, 1999. – P. 28–36.
7. ГОСТ Р 52633.3–2011 Защита информации. Техника защиты информации. Тестирование стойкости средств высоконадежной биометрической защиты к атакам подбора.
8. Руководство по биометрии / Р. Болл, Дж. Х. Коннел, Ш. Панканти, Н. К. Ратха, Э. У. Сеньор. – Москва : Техносфера, 2007. – 368 с.
9. *Шапкин, А. В.* Развитие отечественного нейросетевого искусственного интеллекта в защищенном исполнении / А. В. Шапкин, И. А. Кубасов, А. И. Иванов // Вестник Воронежского института ФСИН России. – 2019. – № 4. – С. 132–144.


## References


1. Nikolenko S., Kadurin A., Arkhangel'skaya E. *Glubokoe obuchenie* [Deep learning]. Saint-Petersburg: Piter, 2018, 480 p. [In Russian]
2. GOST R 52633.5–2011 *Zashchita informatsii. Tekhnika zashchity informatsii. Avtomaticheskoe obuchenie neyrosetevykh preobrazovateley biometriya-kod dostupa* [GOST R 52633.5-2011 information Security. Information security techniques. Automatic training of neural network converters biometrics-access code]. [In Russian]
3. Ivanov A. I., Zakharov O. S. *Sreda modelirovaniya «BioNeyroAvtograf»* [Environment modeling "Bioneurological»]. Available at: http://pniei.rf/ activity/science/noc/bioneuroautograph.zip [In Russian]
4. Ivanov A. I. *Avtomaticheskoe obuchenie bol'shikh iskusstvennykh neyronnykh setey v biometricheskikh prilozheniyakh: ucheb. posobie* [Automatic training of large artificial neural networks in biometric applications: textbook]. Penza, 2013, 30 p. Available at: http://pniei.rf/activity/science/noc/tm_IvanovAI.pdf [In Russian]
5. Morelos-Saragosa R. *Iskusstvo pomekhoustoychivogo kodirovaniya* [The art of noise-tolerant coding]. Moscow: Tekhnosfera, 2007, 320 p. [In Russian]
6. Juels A. A, Wattenberg M. *Proc. ACM Conf. Computer and Communications Security*. Singapore, 1999, pp. 28–36.
7. GOST R 52633.3–2011 *Zashchita informatsii. Tekhnika zashchity informatsii. Testirovanie stoykosti sredstv vysokonadezhnoy biometricheskoy zashchity k atakam podbora* [GOST R 52633.3-2011 information Security. Information security techniques. Testing the resistance of highly reliable biometric security tools to matching attacks]. [In Russian]
8. Boll R., Konnel Dzh. Kh., Pankanti Sh., Ratkha N. K., Sen'or E. U. *Rukovodstvo po biometrii* [Guide to biometrics]. Moscow: Tekhnosfera, 2007, 368 p. [In Russian]
9. Shapkin A. V., Kubasov I. A., Ivanov A. I. *Vestnik Voronezhskogo instituta FSIN Rossii* [Bulletin of the Voronezh Institute of the Federal penitentiary service of Russia]. 2019, no. 4, pp. 132–144. [In Russian]

**Иванов Александр Иванович**
доктор технических наук, доцент,
ведущий научный сотрудник,
Пензенский научно-исследовательский
электротехнический институт
(Россия, г. Пенза, ул. Советская, 9)
E-mail: ivan@pniei.penza.ru

**Кубасов Игорь Анатольевич**
доктор технических наук, профессор,
кафедра информационных технологий,
Академия управления МВД России
(Россия, г. Москва, ул. Зои и Александра
Космодемьянских, 8)
E-mail: igorak@list.ru

**Самокутяев Александр Михайлович**
Герой России, летчик-космонавт,
заместитель командира отряда космонавтов
ФГБУ НИИ ЦПК им. Ю. А. Гагарина
(Россия, Московская область, Щелковский район,
Звездный городок)
E-mail: samo@gctc.ru

**Ivanov Alexander Ivanovich**
doctor of technical sciences, associate professor,
senior researcher,
Penza Scientific Research Electrotechnical Institute
(9 Sovetskaya street, Penza, Russia)

**Kubasov Igor Anatolyevich**
doctor of technical sciences, professor,
sub-department of information technologies,
Academy of Management of the Ministry
of Internal Affairs of the Russian Federation
(8 Zoy and Alexander Kosmodemianskih street,
Moscow, Russia)

**Samokutyaev Alexander Mikhailovich**
Hero of Russia, pilot-cosmonaut,
Deputy Commander of the Cosmonaut Detachment
of the Gagarin Research Institute of the CTC
(Star City, Shchelkovsky district, Moscow region,
Russia)

**Образец цитирования**:
Иванов, А. И. Тестирование больших нейронных сетей на малых выборках / А. И. Иванов, И. А. Кубасов, А. М. Самокутяев // Надежность и качество сложных систем. – 2020. – № 3 (31). – С. 72–79. – DOI 10.21685/2307-4205-2020-3-9.