

# АВТОРИТЕТНОЕ МНЕНИЕ

## Нейросети в судебной экспертологии и экспертной практике: проблемы и перспективы

**Аннотация.** В статье с позиций теории цифровизации судебно-экспертной деятельности как частной теории судебной экспертологии рассматриваются перспективы внедрения нейронных сетей в судебную экспертизу и возникающие при этом актуальные проблемы. Автор отмечает изменения в методологии и технологиях разработки экспертных методик в связи с внедрением алгоритмов искусственного интеллекта — нейронных сетей. Обозначены сферы применения нейронных сетей для решения разнообразных задач науки и практической деятельности. На конкретных примерах продемонстрированы возможности использования алгоритмов контролируемого обучения нейросетей в судебно-экспертной практике. Дан подробный анализ причин, по которым использование нейронных сетей в судебной экспертизе может привести к ошибочным заключениям. Особое внимание уделено галлюцинациям нейронных сетей глубокого обучения на больших языковых моделях. Существует опасность, что эксперт, всецело полагаясь на нейросеть, может сделать неправильный вывод, поскольку самообучаемые генеративные нейросети не дают объяснения, почему приняли то или иное решение. Для разработки методик решения типичных экспертных задач на основе нейросетей предлагается создание баз данных (Dataset) для анализа и машинного обучения по различным объектам судебной экспертизы. Для хранения Dataset необходимо организовать репозитории, которые могут содержать наборы данных по родам (видам) судебных экспертиз. Dataset и репозитории обеспечат контроль качества данных и верификацию моделей. В статье обоснована необходимость новых компетенций специалиста Data Scientist, который формирует инструменты для решения судебно-экспертных задач при внедрении нейросетей и других алгоритмов искусственного интеллекта в судебную экспертологию, а также работающего с ним в контакте инженера по машинному обучению.

**Ключевые слова:** цифровизация, теория цифровизации судебно-экспертной деятельности, судебная экспертология, нейронная сеть, алгоритмы машинного обучения, контролируемое обучение, галлюцинации нейронных сетей, Dataset, Data Scientist.



### **Елена Рафаиловна РОССИНСКАЯ,**

заведующий кафедрой  
судебных экспертиз  
Университета имени  
О.Е. Кутафина (МГЮА),  
доктор юридических наук,  
профессор,  
заслуженный деятель  
науки РФ,  
почетный работник  
высшего  
профессионального  
образования РФ,  
академик Российской  
академии естественных  
наук,  
президент Палаты  
судебных экспертов  
имени Ю. Г. Корухова  
**elena.rossinskaya@  
gmail.com**  
125993, Россия, г. Москва,  
ул. Садовая-Кудринская, д. 9

DOI: 10.17803/2311-5998.2024.115.3.021-033

© Россинская Е. Р., 2024

**Elena R. ROSSINSKAYA,**

*Head of the Forensic Expertise's Department  
of the Kutafin Moscow State Law University (MSAL), Dr. Sci. (Law), Professor,  
Honored Scientist of the Russian Federation, Honorary Worker  
of Higher Professional Education of the Russian Federation,  
Active Member of the Russian Natural Sciences Academy, President  
of the Chamber of Forensic Experts named after Yu. G. Korukhov  
elena.rossinskaya@gmail.com  
9, ul. Sadovaya-Kudrinskaya, Moscow, Russia, 125993*

### **Neural Networks in Forensic Expertology and Expert Practice: Problems and Prospects**

**Abstract.** *The article, from the perspective of the theory the forensic activity digitalization as a particular theory of forensic expertise, examines the prospects for the introduction of neural networks in forensic examination and the current problems arising in this case. The author notes changes in the methodology and technologies for developing expert techniques in connection with the introduction of artificial intelligence algorithms — neural networks. The areas neural networks' application for solving various problems of science and practical activity are outlined. Specific examples demonstrate the possibilities of using supervised learning algorithms for neural networks in forensic practice. A detailed analysis of the reasons why the use of neural networks in forensic science can lead to erroneous conclusions is given. Particular attention is paid to hallucinations of deep learning neural networks on large language models. There is a danger that an expert, relying entirely on a neural network, may give the wrong conclusion, since self-learning generative neural networks do not provide an explanation for why they made a particular decision. To develop expert methods for solving typical expert problems based on neural networks, it is proposed to create databases (Dataset) for various forensic objects for analysis and machine learning. To store the Dataset, it is necessary to organize repositories that can contain data sets on types (kinds) of forensic examinations. Dataset and repositories will provide data quality control and model verification. The article substantiates the need for new competencies of a Data Scientist — a specialist who develops tools for solving forensic problems when introducing neural networks and other artificial intelligence algorithms into forensic science, as well as a machine learning engineer working in contact with him.*

**Keywords:** *digitalization, theory of digitalization for forensic expert activity, forensic expertology, neural network, machine learning algorithms, supervised learning, neural network hallucinations, Dataset, Data Scientist.*

Одними из основных закономерностей судебной экспертологии являются закономерности информационно-компьютерного обеспечения судебно-экспертной деятельности, информатизации и компьютеризации судебной экспертизы, включающие закономерности активного использования

информационных компьютерных технологий (процессов, методов поиска, сбора, хранения, обработки, представления, распространения информации), создания и совершенствования информационных систем судебно-экспертного назначения и информационно-телекоммуникационных сетей специального назначения<sup>1</sup>. Позиции, близкой к нашей, придерживалась и Т. В. Аверьянова, которая рассматривала использование компьютерных средств и баз данных, а также систем гибридного интеллекта, говоря об информационных процессах в судебно-экспертной деятельности<sup>2</sup>.

Глобальные процессы цифровизации во всех сферах жизни социума дали новый импульс развитию IT-технологий в судебно-экспертной деятельности вследствие замены традиционных аналоговых способов отображения объектов судебных экспертиз (следов, фотоснимков, видео- и звукозаписей и др.) качественно новыми объектами, представленными в цифровом виде — цифровыми следами. Это породило целый ряд теоретических, правовых и организационных проблем, решение которых обусловило формирование новой частной теории экспертологии — теории цифровизации судебно-экспертной деятельности<sup>3</sup>, которая включает два направления:

1) информационно-компьютерное обеспечение судебно-экспертной деятельности как методологическую и технологическую основу использования IT-технологий в экспертных исследованиях, включая разработку компьютерных систем анализа изображений; создание баз данных по объектам судебной экспертизы, программных комплексов для выполнения расчетов по известным алгоритмам; автоматизацию сбора и обработки экспериментальных данных; создание гибридных человеко-машинных систем автоматизированного решения экспертных задач<sup>4</sup>;

2) судебно-экспертное исследование цифровых следов, включая их природу, происхождение, механизм следообразования, свойства и признаки, информационную составляющую в зависимости от рода (вида) судебных экспертиз и формы представления: на отдельных носителях; в компьютерных системах и сетях, мобильных коммуникаторах, планшетах, на серверах, в облачных хранилищах; в виде электронных документов; образов, отображаемых в цифровом виде с помощью программных продуктов, образов цифровых следов, отображаемых на бумажных носителях<sup>5</sup>.

В XXI в. на смену экспертным системам, описывающим алгоритм действий по выбору решения в зависимости от конкретных условий, пришло машинное обучение — класс методов автоматического создания прогнозных моделей, представляющих собой инструменты для анализа данных и прогнозирования событий или

<sup>1</sup> Россинская Е. Р., Галяшина Е. И., Зинин А. М. Теория судебной экспертизы (Судебная экспертология) : учебник / под ред. Е. Р. Россинской. М. : Норма ; Инфра-М, 2020. С. 48—50.

<sup>2</sup> Аверьянова Т. В. Судебная экспертиза: курс общей теории. М. : Норма, 2006. С. 382—391.

<sup>3</sup> Россинская Е. Р. Концепция частной теории цифровизации судебно-экспертной деятельности // Вестник экономической безопасности. 2022. № 5. С. 173—176.

<sup>4</sup> Теория информационно-компьютерного обеспечения криминалистической деятельности : монография / под ред. Е. Р. Россинской. М. : Проспект, 2022. С. 165—168.

<sup>5</sup> Россинская Е. Р. Концепция частной теории цифровизации судебно-экспертной деятельности.



тенденций, которые дают возможность определить поведение системы и принимать обоснованные решения. Выбор алгоритма машинного обучения зависит от типа решаемой задачи, доступных вычислительных ресурсов и характера данных.

Машинное обучение может быть *контролируемым* — это подход к обучению алгоритмов на основе размеченных данных. Он предполагает, что имеется набор данных, для каждого элемента которого известно, к какому классу он относится. Затем алгоритм изучает признаки, характерные для каждого класса, и строит модель, которая позволяет определять класс нового объекта на основе его признаков. Эти алгоритмы уже давно реализованы в криминалистических учетах и судебно-экспертных исследованиях<sup>6</sup>, например в автоматизированной дактилоскопической идентификационной системе АДИС «Папиллон».

При *неконтролируемом* обучении машина учится без явного вмешательства человека. Этот метод часто используется для анализа больших объемов данных, когда невозможно вручную определить все зависимости и закономерности.

Третьим вариантом является обучение *с подкреплением*, когда используются алгоритмы, которые обучаются на результатах и определяют, какое действие следует предпринять. После каждого действия алгоритм получает отзыв, помогающий определить, является ли сделанный выбор правильным, нейтральным или неправильным. Этот метод удобно использовать в автоматизированных системах, которые должны принимать много мелких решений без вмешательства человека.

В связи с увеличением вычислительных возможностей программно-аппаратных комплексов, в том числе в результате использования графических процессоров и распределенных архитектур вычислительных систем, стало доступным широкое применение машинного обучения на базе множества вычислительных систем, что привело к значительному повышению качества разрабатываемых технологических решений<sup>7</sup>.

За последние годы в науке, технологиях и социальной сфере России в полтора раза расширилось использование решений в области искусственного интеллекта. Об это говорил Президент России В. В. Путин на пленарном заседании Международной конференции по искусственному интеллекту и машинному обучению Artificial Intelligence Journey 2023 на тему «Революция генеративного ИИ: новые возможности» 24 ноября 2023 г.<sup>8</sup>

В русле Национальной стратегии был утвержден паспорт национального проекта «Национальная программа “Цифровая экономика Российской Федерации”» со сроком реализации до конца 2024<sup>9</sup>. К сожалению, проблемы судебно-экспертной деятельности не нашли в нем своего отражения, хотя, как верно указывают авторы статьи «Искусственный интеллект в судебной экспертиологии», судебная экспертиза «является самостоятельной сферой деятельности федеральных

<sup>6</sup> Россинская Е. Р., Галяшина Е. И., Зинин А. М. Указ. соч. С. 323—359.

<sup>7</sup> Национальная стратегия развития искусственного интеллекта на период до 2030 года (утв. Указом Президента РФ от 10.10.2019 № 490). П. 7.

<sup>8</sup> URL: <http://kremlin.ru/events/president/news/72811> (дата обращения: 28.12.2023).

<sup>9</sup> Паспорт национального проекта «Национальная программа “Цифровая экономика Российской Федерации”» (утв. протоколом заседания президиума Совета при Президенте РФ по стратегическому развитию и национальным проектам от 04.06.2019 № 7).

органов исполнительной власти, осуществляющих функции по выработке и реализации государственной политики, нормативно-правовому регулированию, надзору и контролю в области внутренних дел в различных сферах»<sup>10</sup>, а судебная экспертология — это отрасль науки и технологии, весьма важная для теории и практики судопроизводства.

Среди алгоритмов машинного обучения и искусственного интеллекта можно выделить один из подходов — нейронные сети. Нейронная сеть — это математическая модель, ее программное и аппаратное воплощение, построенные по принципу организации и функционирования сетей нервных клеток живого организма.

Сферы применения нейронных сетей для решения задач весьма разнообразны, например:

- распознавание образов (изображений, текста, звуков и других типов данных);
- прогнозирование различных тенденций (например, экономических показателей);
- кластеризация данных, разделения их на группы на основе сходства;
- анализ больших объемов данных, прогнозирование и выявления закономерностей и тенденций;
- создание систем управления беспилотными аппаратами, роботами;
- обучение и оптимизация других алгоритмов, оптимизация процессов и систем.

Современные нейросети работают по следующей схеме. В нейросеть загружается некоторое количество конкретных, необходимых для эксперимента или исследования данных. Информация передается с помощью искусственных синапсов (соединений) от нейрона к нейрону, от слоя к слою, каждый нейрон может иметь несколько входящих синапсов с данными. Данные, полученные каждым нейроном, представляют собой сумму всех данных, умноженных на коэффициент веса каждого искусственного синапса. Полученные значения формируют выходные сигналы, которые передаются до тех пор, пока информация не достигнет конечного выхода.

В последние годы в России и за рубежом появились научные исследования, направленные на применение нейросетей в криминалистике и судебной экспертизе. При этом, как представляется, в частной теории цифровизации судебно-экспертной деятельности должны быть разработаны методологическая и технологическая основа использования нейронных сетей как нового инструментария экспертных исследований, так и как новых экспертных объектов.

Рассмотрим некоторые примеры применения нейросетей в судебной экспертизе. Одним из вариантов является использование алгоритмов контролируемого обучения нейросетей для автоматизации этапа работы, традиционно считавшегося «ручным» — просмотра экспертом рекомендательных списков, сформированных в результате автоматических поисков АДИС. Нейросеть, внедренная в ЭКЦ МВД по Республике Крым в 9-ю версию АДИС (AFIS) «Папиλλον» («Папилон-АДИС (AFIS)-9-НейроЭксперт»), позволяет: в десятки раз сократить трудозатраты специалистов на просмотр рекомендательных списков; уменьшить

<sup>10</sup> Чеснокова Е. В., Усов А. И., Омелянюк Г. Г., Никулина М. В. Искусственный интеллект в судебной экспертологии // Теория и практика судебной экспертизы. 2023. Т. 18. № 3. С. 60—77.



вероятность пропуска цели; идентифицировать следы по «родному» кандидату, находящемуся в недоступной для просмотра области результатов поисков, и за счет этого увеличить результативность. Внедрение данного программного обеспечения способствовало раскрытию преступлений прошлых лет на полуострове<sup>11</sup>.

Другим примером использования нейросетей для дактилоскопических исследований является сверточная нейронная сеть на основе глубокого обучения, которая используется для обнаружения отпечатков пальцев на изображениях. Она включает в себя два пути: доменно-адверсивный и состязательный. Первый путь помогает сети распознавать объекты независимо от их позиции, масштаба или ориентации. Однако внутренняя среда может измениться, и ранее созданный объект может оказаться недействительным для изменившейся среды. Чтобы адаптироваться к изменившейся среде, требуется собрать огромное количество размеченных выборок данных и заново провести обучение, что трудоемко и требует много времени.

Для преодоления этого недостатка предлагается второй путь — доменная состязательная нейронная сети (DANN), которой нужны только немаркированные образцы данных из измененной среды для обновления отпечатка пальца до адаптивирования к изменившейся среде<sup>12</sup>.

В судебной экспертизе могут найти применение исследования по использованию нейросетей для компьютерного зрения (computer vision) — области искусственного интеллекта и компьютерных наук, которая занимается разработкой алгоритмов и систем, способных распознавать и интерпретировать объекты на изображениях и видео для извлечения из них полезной информации. Нейросети при обучении на больших наборах данных, содержащих изображения с метками, учатся распознавать объекты на изображениях и определять их класс, что затем может использоваться для анализа новых изображений. Компьютерное зрение уже применяется в различных приложениях для распознавания объектов, обнаружения и отслеживания объектов, обработки медицинских изображений, изображений автономных транспортных средств и робототехники, систем видеонаблюдения и безопасности, анализа документов и др.

Благодаря развитию инструментов для редактирования видео, каждый может легко им манипулировать, не оставляя заметных визуальных следов. Повышение частоты кадров (FRUC), основанное на схемах интерполяции кадров, используется злоумышленниками для подделки видео, например для создания видео с фальшивой частотой кадров. Артефакты, которые остаются в интерполированных кадрах, часто трудно различимы. Нейросеть использует стек последовательных кадров в качестве входных данных и эффективно изучает артефакты интерполяции, используя сетевые блоки для изучения пространственно-временных особенностей<sup>13</sup>. Для проверки подлинности цифровых изображений и преодоления

<sup>11</sup> URL: <https://www.papillon.ru/about/blog/v-9-uju-versiju-adis-afis-papillon-vnedrena-nejroset/> (дата обращения: 29.12.2023).

<sup>12</sup> *Wei Liu, Zhiqiang Dun*. D-Fi: Domain adversarial neural network based CSI fingerprint indoor localization // *Journal of Information and Intelligence*. Vol. 1. Is. 2. July 2023. P. 104—114.

<sup>13</sup> *Minseok Yoon, Seung-Hun Nam, In-Jae Yu, Wonhyuk Ahn, Myung-Joon Kwon, Heung-Kyu Lee*. Framerate up-conversion detection based on convolutional neural network for learning spatiotemporal features // *Forensic Science International*. Vol. 340. November 2022. 111442.

сомнений в их редактировании<sup>14</sup>, дипфейков<sup>15</sup> также предлагаются многочисленные алгоритмы нейронных сетей. Нейронные сети могут быть использованы для распознавания лиц на фотографиях или видеозаписях, для определения возраста, пола и других характеристик человека.

В литературе описаны многочисленные исследования документов для выявления признаков подделок, распознавания текстов, добавления новых страниц, идентификации почерка и подписей<sup>16</sup>, исследования следов оружия на пулях и гильзах<sup>17</sup>.

Интересны разработки использования нейросетей в автотехнических исследованиях. Авторами разработан алгоритм машинного обучения, позволяющий решить обратную диагностическую задачу на основе остаточной поврежденной геометрической конфигурации автомобиля (его обломков) и восстановить пред-аварийное исходное состояние автомобиля<sup>18</sup>.

Нейронные сети могут служить эффективным инструментарием в компьютерно-технической экспертизе, например, сверточная нейросеть используется для распознавания сетевых атак путем анализа собранного сетевого трафика на уровне пакетов<sup>19</sup>. Авторы отмечают высокую эффективность и производительность предлагаемой ими структуры, когда она классифицирует отдельные классы атак с точностью 99,4, 99,0 и 90 % на каждый набор данных соответственно.

Этот краткий обзор показывает новые возможности в развитии технологий, которые за счет использования нейросетей и других алгоритмов искусственного

<sup>14</sup> См., например: *Yuxue Zhang, Yunfeng Yan, Guorui Feng*. Feature compensation network based on non-uniform quantization of channels for digital image global manipulation forensics // *Signal Processing: Image Communication*. Vol. 107. September 2022. 116795.

<sup>15</sup> *Samuel Henrique Silva, Mazal Bethany, Alexis Megan Votto, Ian Henry Scarff, Nicole Beebe, Peyman Najafirad*. Deepfake forensics analysis: An explainable hierarchical ensemble of weakly supervised models // *Forensic Science International: Synergy*. Vol. 4. 2022. 100217.

<sup>16</sup> См.: *Бахтеев Д. В.* Особенности распознавания подлога подписи человеком как первичные критерии для разработки системы искусственного интеллекта // *Сибирское юридическое обозрение*. 2020. № 17 (4). С. 514—522; *Wuyang Shan, Yaohua Yi, Ronggang Huang, Yong Xie*. Robust contrast enhancement forensics based on convolutional neural networks // *Signal Processing: Image Communication*. Vol. 71. February 2019. P. 138—146.

<sup>17</sup> *Федоренко В. А., Сорокина К. О., Гиверц П. В.* Классификация изображений следов бойков по экземплярам оружия с помощью полносвязной нейронной сети // *Известия Саратовского университета. Новая серия. Серия: Экономика. Управление. Право*. 2022. Т. 22. Вып. 2. С. 184—190; *Giverts P., Sorokina K., Fedorenko V.* Examination of the possibility to use Siamese networks for the comparison of firing marks // *Journal of Forensic Sciences*. 2022. 2022 Nov;67(6):2416-2424.

<sup>18</sup> *Yuxi Xie, C.T. Wu, Boyuan Li, Xuan Hu, Shaofan Li*. Computer Methods in Applied Mechanics and Engineering. A feed-forwarded neural network-based variational Bayesian learning approach for forensic analysis of traffic accident // *Computer Methods in Applied Mechanics and Engineering*. Vol. 397. 1 July 2022. 115148.

<sup>19</sup> *Sonam Bhardwaj, Mayank Dave*. Enhanced neural network-based attack investigation framework for network forensics: Identification, detection, and analysis of the attack // *Computers & Security*. Vol. 135. December 2023. 103521.



интеллекта открываются в судебной экспертизе. Заманчивы также перспективы быстрого получения новой доказательственной информации, расширения круга решаемых диагностических и идентификационных задач при минимизации затрат и ресурсов.

В XXI в. вычислительные мощности, оперирующие большими массивами данных (Big Data), позволяют создавать большие языковые модели глубокого обучения, которые предварительно обучены на огромных объемах данных<sup>20</sup>. Лежащий в их основе трансформер — это набор нейронных сетей, каждая из которых состоит из кодера и декодера с возможностью самонаблюдения. Кодер и декодер извлекают значения из последовательности текста и понимают отношения между имеющимися в ней словами и фразами.

Глубокие нейронные сети, или сети глубокого обучения, имеют несколько скрытых слоев с миллионами связанных друг с другом искусственных нейронов. Теоретически глубокие нейронные сети могут сопоставлять любой тип ввода с любым типом вывода. Они способны осуществлять самообучение, усваивать знания. Однако стоит учитывать, что им требуется гораздо более сложное обучение, чем другим методам машинного обучения. Таким узлам нужны миллионы примеров обучающих данных, а не сотни или тысячи, как в случае с простыми нейронными сетями.

Но в отличие от описанных нами выше рекуррентных нейронных сетей, которые последовательно обрабатывают входные данные, трансформеры обрабатывают целые последовательности параллельно, что значительно сокращает время обучения. Архитектура нейронной сети трансформера позволяет использовать очень большие модели, часто с сотнями миллиардов параметров. Такие сверхбольшие модели могут получать огромные объемы данных из Интернета и, например, из таких источников, как индекс Common Crawl (более 50 миллиардов веб-страниц) или Википедии (около 57 миллионов страниц). Большое количество исследований, проведенных за последнее время, показало, что предварительно обученные (pre-trained models — PTM) модели нейронных сетей обработки естественного языка позволяют избежать обучения новой модели с нуля<sup>21</sup>.

Последнее четвертое поколение моделей PTM, таких как GPT-4, имеют значительно большее количество параметров по сравнению с их предшественниками, что позволяет им улавливать более широкий спектр языковых моделей. Они проявляют лучшее понимание контекста, грамматики и семантики. Это приводит к более связному и контекстуально точному созданию текста, что позволяет с большой производительностью эффективно настраивать модели PTM для решения широкого спектра задач, таких как перевод, реферирование и ответы на вопросы из различных областей знания.

Внедрение нейросетей в различные сферы науки и практической деятельности для междисциплинарных исследований, общедоступность существующих алгоритмов и лавинообразное нарастание разработки новых затрагивают и сферу

<sup>20</sup> Бахтизин А. Р., Брагин А. В., Макаров В. Л. Большие языковые модели четвертого поколения как новый инструмент в научной работе // Искусственные общества. 2023. Т. 18. Вып. 1. URL: <https://artsoc.jes.su/s207751800025046-9-1> (дата обращения: 02.01.2024).

<sup>21</sup> Pre-trained models for natural language processing: A survey / Qiu X. [et al.] // Science China Technological Sciences. 2020. Vol. 63. № 10.

производства судебных экспертиз. Этому может способствовать соблазн упрощения поиска, быстроты получения результатов, возможности получения новых данных. Пока, как показывают приведенные выше примеры, используются нейросети с контролируемым обучением и с проверкой на традиционных признаках объектов, выявленных вручную.

Мы полагаем, что использование нейросетевых алгоритмов в судебно-экспертных технологиях должно базироваться на методологии теории цифровизации судебно-экспертной деятельности, поскольку при этом открываются новые возможности, но и предвидятся серьезные ограничения применения подобного инструментария.

Для разработки технологии экспертного исследования с применением нейросети необходимо определить следующее.

<i>Задача, которую надо решить</i>	
Модель, которая может решить эту задачу	Какую точность считать приемлемой для модели, если ее применять на практике
Данные, на которых можно обучиться (провести эксперименты)	Достаточно ли данных, чтобы при верификации модели сказать, что модель можно применять для решения задачи, результаты работы модели необходимо модерировать
<i>Данные, на которых эта модель верифицируется</i>	
Математический аппарат, который показывает статистическую значимость того или иного объема данных	Работа с моделями — это работа с вероятностями, результат — это всего лишь его очень высокая вероятность

Причины, по которым использование нейронных сетей в судебной экспертизе может привести к ошибочным заключениям:

- некорректность или недостаточная точность моделей, не подходящих для использования в судебной экспертизе;
- недостаток данных или их некорректная разметка. Нейронные сети требуют большого количества данных для обучения. Если данных недостаточно или они некачественные, то сеть может обучиться неправильным зависимостям и делать ошибки;
- предвзятость данных, поскольку нейросети обучаются на тех данных, которые им предоставляют. Если данные содержат предубеждения, то и нейросеть будет их воспроизводить. Например, если в обучающей выборке большинство дактилоскопических отпечатков были оставлены мужчинами, то нейросеть может начать ассоциировать преступность с мужским полом;
- неправильный алгоритм обучения нейросети, который может быть связан с выбором архитектуры сети, обучающих данных, с использованием неподходящего оптимизатора, выбором функции ошибки, неправильной настройкой гиперпараметров, отсутствием проверки на переобучение и т.д.;
- неправильная интерпретация результатов, поскольку нейросети не дают объяснения, почему они приняли то или иное решение. Эксперты, полагаясь всецело на нейросеть, не понимают, на чем основаны ее выводы;
- подмена понятий. Работа с моделями — это всегда работа с вероятностями, но эксперты забывают о том, что категорический вывод нейросети — это



всего лишь очень высокая вероятность результата. Истина на 90 % в сознании людей подменяется на безусловную истину.

Возможны и некоторые другие причины ошибок, например, галлюцинации нейросети, когда нейросеть «видит» или «слышит» то, чего на самом деле нет, выдает за факты придуманную ею же информацию, причем очень убедительно. Галлюцинации опасны тем, что искусственный интеллект упрямо настаивает на своем. Если уличить нейросеть в ложности ответа, она продолжит отстаивать свою правоту.

Галлюцинации нейросетей могут представлять собой:

- создание реалистичных изображений, которые кажутся настоящими, но на самом деле не существуют, например вставление в видео несуществующих людей и создание иллюзии, что они были там изначально;
- неправильное распознавание объектов или представление несуществующих объектов, например, беспилотный автомобиль может увидеть несуществующие препятствия или пешехода на пустой дороге; робот «думает», что он держит инструмент, когда в действительности инструмент отсутствует;
- восприятие прочитанных или произнесенных слов или фраз, когда на самом деле их не было, что может привести к неправильному пониманию или интерпретации текста;
- ошибочную интерпретацию результатов инструментальных исследований и анализов, например, симптомов для нейросетей, обученных на медицинских данных;
- выдача нерелевантных или нежелательных рекомендаций, вызванных некорректной обработкой входных данных.

Приведем несколько конкретных примеров. Во время одного эксперимента чат-бот ChatGPT написал текст на медицинскую тематику. В нем встретился неверный факт, и исследователи попросили чат-бот подтвердить информацию научными источниками. Модель выдала названия пяти работ, но ни одной из них не существовало в реальности. При этом нейросеть дала реальные ID исследований, но все они принадлежали статьям, которые не имели отношения к теме<sup>22</sup>.

Примером галлюцинации нейросети в области юриспруденции является первый иск в федеральный суд США из-за галлюцинаций нейросети к компании OpenAI, который подал радиоведущий из штата Джорджия (США) Марк Уолтерс из-за того, что ChatGPT обвинил его в мошенничестве и растрате средств некоммерческой организации<sup>23</sup>. Ответ ChatGPT на запрос, сгенерированный журналистом Фредом Рилом, гласил, что Уолтерс незаконно присвоил средства некоммерческой организации по защите прав на оружие на сумму более 5 млн долларов. Журналист решил перепроверить всю информацию и убедился в том, что никаких подобных дел против Уолтерса не существует.

Галлюцинации нейронных сетей зависят от их архитектуры и того, как они обучаются. Например на вопрос «Какое число больше: -999 934 или -32 323 232 332?» нейронная сеть отвечает, что -32 323 232 332 больше,

<sup>22</sup> URL: <https://vc.ru/future/918465-opasnye-fantazii-chto-takoe-gallyucinacii-ii-i-kak-ih-obnaruzhit> (дата обращения: 02.01.2024).

<sup>23</sup> URL: <https://www.pvsm.ru/news/385293> (дата обращения: 02.01.2024).

поскольку в обучающий алгоритм было заложено, что чем больше цифр в числе, тем оно больше<sup>24</sup>.

Резюмируя вышеизложенное, полагаем, что в настоящее время нейронные сети и другие алгоритмы искусственного интеллекта следует внедрять в судебно-экспертные технологии с большой осторожностью.

Для разработки экспертных методик на основе нейросетей необходимо создание баз данных (Dataset) для анализа и машинного обучения — совокупностей данных по различным объектам судебной экспертизы, для решения типичных экспертных задач. Это могут быть наборы верифицированных структурированных (например, таблицы), или неструктурированных (например, текстовые документы или изображения) данных. Для их хранения необходимо организовать репозитории (хранилища) наборов данных по родам (видам) судебных экспертиз — централизованные места для хранения с регламентированной системой доступа, которые можно использовать как основу для создания судебно-экспертных методик.

Dataset и репозитории для хранения и систематизации данных обеспечат возможность сравнения результатов различных исследований, контроль качества данных и верификацию (воспроизводимость) моделей. Если модель была обучена год назад, то факты, которые имели место за последний год, ей недоступны для оперирования, поэтому процесс обучения и верификации должен осуществляться на постоянной основе с разумной периодичностью.

Для создания судебно-экспертных репозиториях можно использовать уже существующие наиболее известные платформы<sup>25</sup>:

- Kaggle для соревнований по машинному обучению и анализу данных, где участники могут загружать свои наборы данных и делиться ими с другими пользователями;
- UCI Machine Learning Repository, где содержатся наборы данных для различных задач машинного обучения;
- OpenML — репозиторий для машинного обучения, который предоставляет доступ к наборам данных и моделям для научных исследований;
- Figshare — хранилище для научных данных, включая наборы данных, алгоритмы и программное обеспечение и др.

Создание Dataset и репозиториях по родам (видам) объектов возможно только после согласования методических подходов к исследованию объектов судебных экспертиз, разработанных различными ведомствами.

Изменения в методологии и технологиях судебно-экспертных исследований требуют новых экспертных компетенций. Внедрение нейросетей и других алгоритмов искусственного интеллекта в судебной экспертологии, как и в других отраслях науки, обуславливает появление специалиста Data Scientist, который формирует инструменты для решения судебно-экспертных задач, выбирая ранее созданные алгоритмы и анализируя данные, строит модели и тестирует их. Он работает на стыке статистики, машинного обучения, программирования и специальных знаний в данном роде (виде) судебных экспертиз. С ним в контакте работает инженер по машинному обучению, в обязанности которого входит

<sup>24</sup> URL: <https://dzen.ru/a/zfdu0hv67xcpoath> (дата обращения: 02.01.2024).

<sup>25</sup> URL: <https://paperswithcode.com/datasets> (дата обращения: 02.01.2024).



автоматизировать работу моделей, следить, чтобы они работали качественно, устранять ошибки. Если точность модели падает, инженер выявляет причины и переобучает алгоритм<sup>26</sup>. Все модели на чем-то учатся, а значит, отстают от актуального состояния ноосферы.

В заключение необходимо подчеркнуть, что целесообразными представляются постановка вопроса о создании системы искусственного интеллекта судебно-экспертной деятельности в рамках федеральной программы «Искусственный интеллект» либо принятие федеральной программы «Искусственный интеллект в судебной экспертизе», как это предлагают авторы из РФЦСЭ при Минюсте России<sup>27</sup>.

### БИБЛИОГРАФИЯ

1. *Аверьянова Т. В.* Судебная экспертиза : курс общей теории. — М. : Норма, 2006. — 480 с.
2. *Бахтеев Д. В.* Особенности распознавания подлога подписи человеком как первичные критерии для разработки системы искусственного интеллекта // Сибирское юридическое обозрение. — 2020. — № 17 (4). — С. 514—522.
3. *Бахтизин А. Р., Брагин А. В., Макаров В. Л.* Большие языковые модели четвертого поколения как новый инструмент в научной работе // Искусственные общества. — 2023. — Т. 18. — Вып. 1. — URL: <https://artsoc.jes.su/s207751800025046-9-1> (дата обращения: 02.01.2024).
4. *Россинская Е. Р.* Концепция частной теории цифровизации судебно-экспертной деятельности // Вестник экономической безопасности. — 2022. — № 5. — С. 173—176.
5. *Россинская Е. Р., Галяшина Е. И., Зинин А. М.* Теория судебной экспертизы (Судебная экспертология) : учебник / под ред. Е. Р. Россинской. — М. : Норма ; Инфра-М, 2020. — 268 с.
6. Теория информационно-компьютерного обеспечения криминалистической деятельности : монография / под ред. Е. Р. Россинской. — М. : Проспект, 2022. — 256 с.
7. *Федоренко В. А., Сорокина К. О., Гиверц П. В.* Классификация изображений следов бойков по экземплярам оружия с помощью полносвязной нейронной сети // Известия Саратовского университета. — Новая серия. — Серия : Экономика. Управление. Право. — 2022. — Т. 22. — Вып. 2. — С. 184—190.
8. *Чеснокова Е. В., Усов А. И., Омелянюк Г. Г., Никулина М. В.* Искусственный интеллект в судебной экспертологии // Теория и практика судебной экспертизы. — 2023. — Т. 18. — № 3. — С. 60—77.

<sup>26</sup> *Humberto Ferreira, Pedro Ruivo, Carolina Reisb.* How do data scientists and managers influence machine learning value creation // Procedia Computer Science. Vol. 181. 2021. Pages 757—764.

<sup>27</sup> *Чеснокова Е. В., Усов А. И., Омелянюк Г. Г., Никулина М. В.* Искусственный интеллект в судебной экспертологии // Теория и практика судебной экспертизы. 2023. Т. 18. № 3. С. 60—77.

9. *Giverts P., Sorokina K., Fedorenko V.* Examination of the possibility to use Siamese networks for the comparison of firing marks // *Journal of Forensic Sciences*. — 2022 Nov. — 67(6). — P. 2416-2424. — DOI: 10.1111/1556-4029.15143.
10. *Humberto Ferreira, Pedro Ruivo, Carolina Reisb.* How do data scientists and managers influence machine learning value creation // *Procedia Computer Science*. — 2021. — Vol. 181. — P. 757—764.
11. *Minseok Yoon, Seung-Hun Nam, In-Jae Yu, Wonhyuk Ahn, Myung-Joon Kwon, Heung-Kyu Lee.* Framerate up-conversion detection based on convolutional neural network for learning spatiotemporal features // *Forensic Science International*. — Vol. 340. — Nov. 2022. — 111442.
12. Pre-trained models for natural language processing: A survey / Qiu X. [et al.] // *Science China Technological Sciences*. — 2020. — Vol. 63. — № 10.
13. *Samuel Henrique Silva, Mazal Bethany, Alexis Megan Votto, Ian Henry Scarff, Nicole Beebe, Peyman Najafirad.* Deepfake forensics analysis: An explainable hierarchical ensemble of weakly supervised models // *Forensic Science International: Synergy*. — 2022. — Vol. 4. — 100217.
14. *Sonam Bhardwaj, Mayank Dave.* Enhanced neural network-based attack investigation framework for network forensics: Identification, detection, and analysis of the attack // *Computers & Security*. — Vol. 135. — December 2023. — 103521.
15. *Wei Liu, Zhiqiang Dun.* D-Fi: Domain adversarial neural network based CSI fingerprint indoor localization // *Journal of Information and Intelligence*. — Vol. 1. — Iss. 2. — July 2023. — P. 104—114.
16. *Wuyang Shan, Yaohua Yi, Ronggang Huang, Yong Xie.* Robust contrast enhancement forensics based on convolutional neural networks // *Signal Processing: Image Communication*. — Vol. 71. — February 2019. — P. 138—146.
17. *Yuxi Xie, C. T. Wu, Boyuan Li, Xuan Hu, Shaofan Li.* Computer Methods in Applied Mechanics and Engineering. A feed-forwarded neural network-based variational Bayesian learning approach for forensic analysis of traffic accident // *Computer Methods in Applied Mechanics and Engineering*. — Vol. 397. — 1 July 2022. — 115148.
18. *Yuxue Zhang, Yunfeng Yan, Guorui Feng.* Feature compensation network based on non-uniform quantization of channels for digital image global manipulation forensics // *Signal Processing: Image Communication*. — Vol. 107. — September 2022. — 116795.