

ВИЗУАЛИЗАЦИЯ РАБОТЫ АЛГОРИТМОВ ИСКУССТВЕННОГО ИНТЕЛЛЕКТА

Кухто В.С.

*Белорусский государственный университет информатики и радиоэлектроники,
г. Минск, Республика Беларусь*

Научный руководитель: Василькова А.Н. – ст. преподаватель кафедры ИПиЭ

Аннотация. Стремительное развитие искусственного интеллекта сопровождается проблемой непонимания их внутренней логики. Сложные нейросетевые модели остаются непрозрачными, что порождает риски в области безопасности, этики и доверия. Визуализация помогает исследовать и понять работу алгоритмов искусственного интеллекта.

Ключевые слова: искусственный интеллект, визуализация, алгоритм, нейронная сеть, иммерсивная визуализация, модель.

Введение. Стремительное развитие искусственного интеллекта (ИИ) стало цифровым двигателем современности. Технологии ИИ оказывают существенное влияние на все сферы человеческой деятельности. Вместе с тем темпы внедрения ИИ-решений существенно опережают возможности понимания механизмов их функционирования [1]. Современные нейронные сети зачастую характеризуются как «чёрные ящики»: их внутренняя логика принятия решений остаётся нераскрытой даже для разработчиков систем, что порождает риски в области безопасности, этики, доверия. Отсутствие прозрачности затрудняет прогнозирование поведения модели в нестандартных ситуациях. В результате возникают трудности внедрения ИИ-решений в различные сферы, так как невозможно гарантировать отсутствие скрытых ошибок в работе алгоритма. Данная проблема ставит серьёзный вопрос: «Как работает искусственный интеллект?».

Основная часть. Визуализация работы алгоритмов искусственного интеллекта – это фундамент понимания их внутреннего устройства. Правильно построенные структурные диаграммы, графы вычислений (рисунок 1), дендрограммы, тепловые карты позволяют визуализировать абстрактные математические концепции и сделать их доступными для широкого круга специалистов. Перечисленные двумерные методы визуализации, являясь плоскими и статичными, не передают топологию связей между нейронами и слоями, динамику распространения и масштаб параметров. Это означает, что разработчик видит архитектуру, но не видит процесса обучения: например, как отдельные нейроны активируются в ответ на различные входные данные.

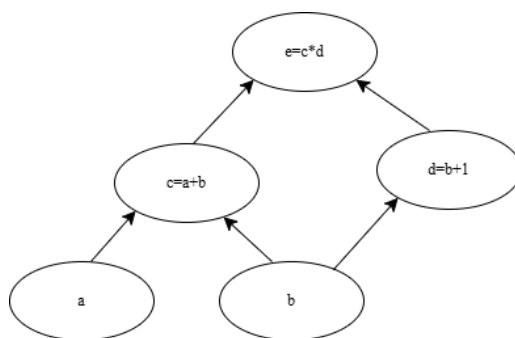


Рисунок 1 – Граф вычислений

Для решения проблем непонимания работы алгоритмов машинного обучения сегодня разработано множество инструментов и методов, выходящих за рамки статичных схем. Например, для преодоления описанных выше ограничений предлагается следующий подход:

иммерсивная научная визуализация внутренних пространств искусственного интеллекта в средах виртуальной (VR) и дополненной (AR) реальности. Преимуществом такой визуализации является интерактивность. Пользователь может «заглянуть» внутрь нейрона, получив информацию о типах входных данных, динамически изменять входные параметры, выявлять ложные корреляции [2].

Разнообразные способы визуализации признаков позволяют более точно определить, что именно активирует нейрон, проверять прогнозы относительно того, какие входные сигналы будут активировать нейрон.

Для визуализации алгоритмов нейронных сетей существует множество решений. Например, иммерсивная визуализация данных Google Experiments, где можно визуализировать данные в объемной форме [3].

На рисунке 2 можно увидеть, как нейронные сети формируют свое понимание изображений.

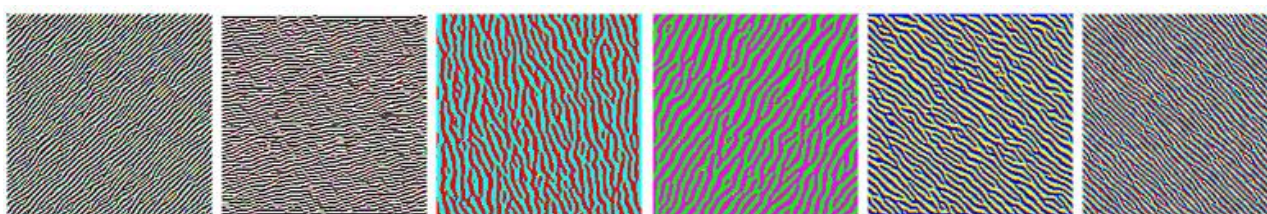


Рисунок 2 – Ребра (слой conv2d0)

Для образовательных целей может быть использован инструмент на базе TensorFlow Embedding Projector, позволяющий в интерактивном режиме исследовать многомерные пространства признаков [4]. Он дает возможность визуализировать 10 000 векторов word2vec, спроецированных в трехмерное пространство с использованием методов снижения размерности. Пользователю предоставляется возможность анализировать ближайших соседей, наблюдать семантические связи между словами, изучать кластеризацию данных в реальном времени, взаимодействуя с трехмерной моделью.

Такой же подход может быть применен к визуализации набора рукописных цифровых данных MNIST (рисунок 3), где каждое изображение проецируется в пространство объекта, а точки заменяются миниатюрами самих цифр, что позволяет визуально оценить, как нейронная сеть группирует визуально похожие образцы [5]. В результате разработчик может обнаружить закономерности в ошибках классификации: например, увидеть, как алгоритм путает графически схожие «5» и «6», или выявить аномальные экземпляры.

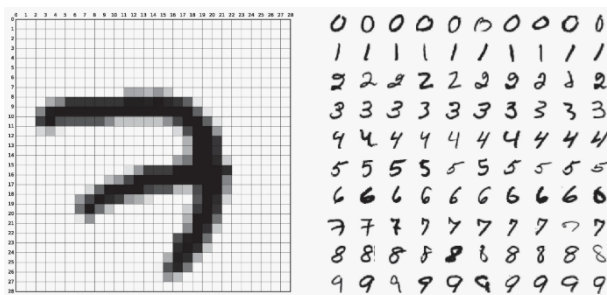


Рисунок 3 – Примеры рукописных цифр из обучающей выборки MNIST

Такая визуализация ценна для анализа качества обученной модели и выявления систематических ошибок, для получения представления о том, какие именно признаки усваивает сеть и где проходят границы принятия решений.

На практике ранее описанные методы визуализации находят прямое применение. В медицине врач-радиолог, используя установку с AR-приложением, может направить камеру

на КТ-снимок и увидеть тепловую карту, выделяющую подозрительную область. Дополнительное цветовое кодирование может отображать динамику изменений состояния здоровья пациента (например, при сравнении с предыдущими снимками пациента, где синим отмечены стабильные области, а красным – участки с негативной динамикой).

Студент, изучающий машинное обучение, может в интерактивном VR-симуляторе поместить любые данные на входную плоскость визуализированной сети и наблюдать, как данные проходят через фильтры и слои.

Описанные иммерсивные среды уже используются разработчиками и исследователями. Разработчик может визуально осмотреть трехмерный граф вычислений, приблизить интересующий участок и проверить, как проходят сигналы через отдельные нейроны. Это позволяет быстро находить неработающие элементы сети и исправлять ошибки еще до того, как начнется дорогостоящее обучение модели [6].

Закключение. Визуализация алгоритмов искусственного интеллекта – это не просто эффективный способ показать сложные вычисления. Это помогает исследовать и понять работу алгоритмов искусственного интеллекта, что является критически необходимым для понимания внутренней логики нейронных сетей. Без визуализации многие процессы остаются абстрактными математическими конструкциями, доступными лишь узким специалистам. Трехмерные графы вычислений, тепловые карты активности и иммерсивные среды превращают нейросети из «черного ящика» в наблюдаемую и изучаемую систему. По мере усложнения моделей роль визуализации будет только возрастать, становясь неотъемлемой частью рабочего процесса при создании и анализе различных систем искусственного интеллекта. Визуализация может стать ключевым инструментом в понимании работы алгоритмов, позволяя не просто наблюдать за результатом, но и анализировать причинно-следственные связи внутри модели.

Список литературы

1. Харбанс, Р. Грогаем алгоритмы искусственного интеллекта / Р. Харбанс ; пер. с англ. А. Киселёва. – СПб. : Питер, 2023. – 368 с.: ил. – (Серия «Библиотека программиста»). – ISBN 978-5-4461-2924-9.
2. Визуализация функций [Электронный ресурс]. Режим доступа: <https://distill.pub/2017/feature-visualization/>. Дата доступа: 10.02.2026.
3. Experiments with Google: Fractal [Электронный ресурс]. Режим доступа: <https://experiments.withgoogle.com/search?q=Fractal>. Дата доступа: 10.02.2026.
4. Встраивания: интерактивные упражнения [Электронный ресурс]. Режим доступа: <https://developers.google.com/machine-learning/crash-course/embeddings/interactive-exercises?hl=ru>. Дата доступа: 24.02.2026.
5. База данных MNIST с рукописными цифрами [Электронный ресурс]. Режим доступа: <https://learn.microsoft.com/ru-ru/azure/open-datasets/dataset-mnist>. Дата доступа: 24.02.2026.
6. Машинное обучение [Электронный ресурс]. Режим доступа: <https://education.yandex.ru/handbook/ml/article/mashinnoye-obucheniye>. Дата доступа: 24.02.2026.

UDC 004.923:004.8

VISUALIZATION OF ARTIFICIAL INTELLIGENCE ALGORITHMS

Kukhta V.S.

Belarusian State University of Informatics and Radioelectronics, Minsk, Republic of Belarus

Vasilkova A.N. – Senior Lecturer at the Department of EPE

Annotation. The rapid development of artificial intelligence is accompanied by a lack of understanding of its internal logic. Complex neural network models remain opaque, posing risks in terms of security, ethics, and trust. Visualization helps you explore and understand how artificial intelligence algorithms work.

Keywords: artificial intelligence, visualization, algorithm, neural network, immersive visualization, model.