

УДК 004.896:165.243.2:316.346.3

17. ПРОБЛЕМЫ И ПЕРСПЕКТИВЫ РАЗВИТИЯ НЕЙРОННЫХ СЕТЕЙ

Волкогон М.Р., студент гр.378108, Ходарёнок Н.А., магистрант гр.376741, Петрович Ю.Ю. магистрант 376701

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

Ефремов А.А. – канд. экон. наук, доцент каф. ЭИ

Аннотация. Эта работа посвящена нейронным сетям - их принципу работы, перспективам развития и влиянию на нашу жизнь. Мы обсудим сложности обучения, этические вопросы и будущее технологии. А также рассмотрим варианты решений ключевых проблем и перспективные их применения. Нейронные сети играют важную роль в развитии искусственного интеллекта, и это исследование поможет лучше понять их важность.

Ключевые слова. Нейронная сеть, веса, нейрон, синапс, этика, мозг.

Введение: В эпоху стремительного технологического развития, нейронные сети становятся ключевым элементом прогресса в области искусственного интеллекта. Они лежат в основе многих инновационных технологий, от автоматического перевода до автономных систем вождения. Несмотря на значительные успехи, мы часто сталкиваемся с различными проблемами - от вычислительных требований до этических вопросов. В то же время, перспективы развития нейронных сетей обещают революционные изменения в различных сферах жизни.

Понятие нейронных сетей и принцип их работы : Нейронные сети — это компьютерные системы, вдохновленные структурой мозга, состоящие из элементов, называемых нейронами. Эти нейроны могут быть активными или пассивными и взаимодействуют друг с другом, формируя сложные сетевые отношения. Малейшие изменения в одном нейроне могут вызвать значительные изменения во всей сети, подобно эффекту бабочки.

В искусственных нейронных сетях (ИНС) и неврологии различное понимание нейрона. В ИНС нейрон — это узел, который хранит данные, например, в формате float, и связан с другими узлами "синапсами", веса которых определяют активацию нейронов. Это отличается от биологических нейронов, где процессы более сложные и многообразные.

Так как же работает нейросеть? Нейросеть работает, имитируя структуру и функционирование человеческого мозга. Она состоит из слоев искусственных нейронов, которые обрабатывают входные данные и передают их через синапсы. В процессе обучения нейросеть корректирует веса синапсов, чтобы улучшить свою способность к распознаванию образов или выполнению задач. Это обучение происходит путем многократного предъявления примеров и корректировки весов с помощью алгоритмов обратного распространения ошибки.

На рисунке 1 представлена схема простейшей нейросети.

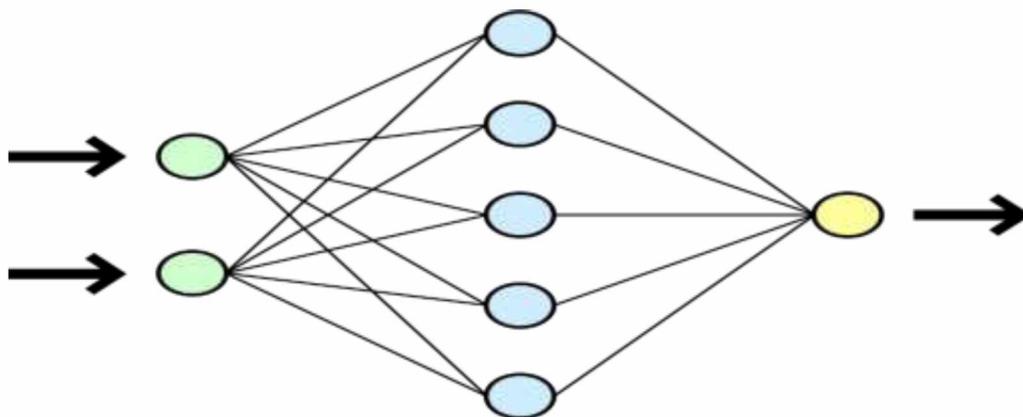


Рисунок 1 – Пример простой нейросети

Первый слой в нейронной сети известен как входной. Нейроны этого слоя представляют компоненты данных, которые нейронная сеть использует для изучения и прогнозирования.

Последний слой в нейронной сети известен как выходной. С ним связан процесс прогнозирования. Если учесть, что нейронная сеть классифицирует данные, в выходном слое будет по одному нейрону для каждой возможной категории входных данных. К примеру, если мы обучаем модель классификации изображений автомобилей и грузовиков, нам понадобятся два нейрона в выходном слое, чтобы представить две возможные категории, под которые могут подпадать наши данные.

Остальные слои, расположенные между входным и выходным слоями, называются скрытыми. Количество скрытых слоев, а также количество нейронов в каждом из них определяется произвольно. Это означает, что программист может экспериментировать и настраивать различные комбинации, пытаясь определить, что лучше всего подходит для конкретной модели.

Скрытые слои позволяют модели извлекать из входных данных функции, необходимые для распознавания шаблонов, чтобы с их помощью классифицировать различные входные данные.[1]

Это самое простое объяснение, но для понимания работы нейросетей этого вполне достаточно.

Проблемы развития нейронных сетей: Несмотря на значительные достижения в области искусственного интеллекта, нейронные сети по-прежнему сталкиваются с рядом проблем, которые ограничивают их эффективность и применимость.

Одна из таких проблем – серьезные вычислительные требования. Нейронные сети, несомненно, являются одним из самых мощных инструментов в области искусственного интеллекта, но их эффективность сопряжена с определенными требованиями к вычислительным ресурсам. Для удобства информация объединена в таблице 1.

Обучение нейронных сетей	Для обучения нейронных сетей требуются значительные вычислительные мощности, особенно для больших моделей и сложных задач. Это связано с необходимостью обработки огромных объемов данных и выполнения множества операций, таких как обратное распространение ошибки и оптимизация весов. Обучение может занимать от нескольких часов до нескольких недель, в зависимости от сложности модели и доступных ресурсов
Работа в реальном времени	Нейронные сети, используемые для задач, требующих немедленного реагирования, например, в автономных транспортных средствах или системах распознавания речи, должны работать в режиме реального времени. Однако высокая вычислительная нагрузка может привести к задержкам, что делает их использование непрактичным в критически важных приложениях, где каждая миллисекунда имеет значение
Устройства с ограниченными ресурсами	Многие устройства, такие как смартфоны, планшеты и некоторые IoT-устройства, имеют ограниченные вычислительные возможности и энергопотребление. Развертывание сложных нейронных сетей на таких устройствах может быть затруднено из-за ограничений по памяти, процессорному времени и энергии.

Таблица 1 – Проблема с вычислительными требованиями

Для преодоления этих ограничений исследователи и инженеры работают над созданием более эффективных алгоритмов и архитектур нейронных сетей. Это включает в себя разработку методов сжатия моделей, квантования весов и прунинга (обрезки) нейронов, которые не вносят значительного вклада в работу сети. Также исследуются новые подходы к обучению, которые могут уменьшить необходимость в больших вычислительных ресурсах.

Следующая проблема – недостаток данных для обучения нейросетей. Для эффективного обучения нейронных сетей требуется большой объем данных, что может быть проблематично в некоторых областях таких, как медицина и научные исследования. В этих областях данные часто ограничены из-за строгих требований к конфиденциальности и защите личной информации пациентов или участников исследований. Кроме того, некоторые события или состояния встречаются редко, что делает сбор достаточного объема данных для обучения нейросетей еще более затруднительным.

В медицине, например, существуют законы, такие как HIPAA в США, которые регулируют использование и распространение медицинской информации. Это означает, что исследователи должны найти способы обеспечения анонимности данных перед их использованием в обучении нейросетей. В науке же, особенно в областях с редкими явлениями, такими как некоторые астрономические события или редкие заболевания, данные могут быть не только ограниченными, но и

трудно воспроизводимыми, что требует разработки специализированных методов обучения, способных работать с малыми объемами данных.[2]

Для преодоления этих проблем используются различные подходы, включая техники аугментации данных, передачи обучения, где нейросеть предварительно обучается на одном наборе данных, а затем адаптируется к другому, и разработку синтетических данных, которые могут имитировать реальные данные без нарушения конфиденциальности. Такие методы помогают расширить возможности применения нейронных сетей, даже когда реальные данные ограничены или недоступны.

И последняя ключевая проблема в развитии, которую мы разберём – сложность интерпретации. Из-за их “черного ящика” сложно понять, как именно модель приходит к определенному решению. В критически важных сферах, таких как медицина, где каждое решение может иметь серьезные последствия, и в финансах, где требуется точность и возможность аудита, необходимо не только получать точные результаты, но и понимать логику их получения.

Интерпретируемость важна для доверия пользователей к моделям ИИ и для соответствия нормативным требованиям. Она также способствует более эффективному устранению ошибок и улучшению моделей, поскольку разработчики могут понять, почему модель ведет себя определенным образом. В ответ на эту проблему исследуются новые подходы и инструменты, которые могут повысить прозрачность и интерпретируемость нейронных сетей.

Перспективы развития нейронных сетей: Нейронные сети – одна из самых многообещающих технологий современности. Они преобразуют каждую отрасль, от медицины до искусства, обещая революционные изменения в способах нашего взаимодействия с миром и решения сложнейших задач. Огромное число инженеров и IT-специалистов ежедневно работают над улучшением различных аспектов работы нейросетей. Рассмотрим некоторые направления в улучшении нейросетей:

1. Улучшение алгоритмов обучения: Исследования в этой области направлены на создание более эффективных алгоритмов, которые могут обучаться быстрее и с меньшим количеством данных. Одним из направлений является разработка новых методов оптимизации, таких как адаптивные методы, которые могут автоматически корректировать скорость обучения в зависимости от сложности задачи. Также важно улучшение точности моделей, чтобы они могли лучше обобщать данные и быть применимыми в реальных условиях.

2. Развитие аппаратных решений: Прогресс в области аппаратных средств, включая GPU и TPU, позволяет нейронным сетям работать быстрее и эффективнее. Эти устройства специально разработаны для выполнения параллельных вычислений, которые необходимы для обучения и работы нейронных сетей. Увеличение производительности и снижение энергопотребления этих устройств делает искусственный интеллект более доступным и экологичным.

3. Интеграция с другими технологиями: Комбинирование нейронных сетей с квантовыми вычислениями может привести к созданию новых типов искусственного интеллекта, которые будут способны решать задачи, недоступные для классических компьютеров. Квантовые компьютеры могут обрабатывать огромные объемы данных за доли секунды, что может радикально изменить подходы к обучению и использованию нейронных сетей. Это открывает возможности для создания более мощных и интеллектуальных систем, способных выполнять сложные задачи, такие как моделирование биологических процессов или оптимизация больших систем.

В таблице 2 приведены перспективные сферы применения нейросетей.

Сфера применения	Область применения
Медицина	Диагностика заболеваний, анализ медицинских изображений, предсказание патологий.
Автономные транспортные средства	Разработка систем управления для беспилотных автомобилей.
Финансовый сектор	Прогнозирование цен на акции, оценка кредитоспособности, выявление мошенничества.
Робототехника	Управление роботами для выполнения сложных задач.
Распознавание изображений	От систем безопасности до фильтров в социальных сетях.
Обработка естественного языка	Перевод текстов, голосовые помощники, анализ тональности текста.
Рекомендательные системы	Подбор фильмов, музыки, товаров в интернет-магазинах.

Таблица 2 – Перспективные сферы и области применения нейросетей

Этические проблемы: Этические проблемы использования нейронных сетей охватывают широкий спектр вопросов, начиная от предвзятости в данных и заканчивая проблемами прозрачности и объяснимости их работы. Особую озабоченность вызывает использование нейронных сетей в военном деле, где они могут быть применены для разработки автономного оружия, что ставит под вопрос контроля за такими системами. Важно также учитывать вопросы конфиденциальности и защиты данных, поскольку нейронные сети требуют больших объемов информации для обучения. Социальные и экономические последствия, такие как влияние на рынок труда и структуру занятости, также являются значимыми аспектами. Все эти факторы требуют тщательного рассмотрения и разработки соответствующих этических принципов и регулирования для минимизации потенциальных рисков и негативных последствий.

Будущее нейросетей: Будущее нейросетей представляется как эра значительных технологических достижений, которые могут радикально изменить нашу жизнь. С развитием новых архитектур и алгоритмов, нейросети станут более мощными и способными к обучению с меньшим количеством данных, что позволит им проникнуть во все большее количество сфер нашей деятельности.

“Мы на пороге революции в области нейросетей, которая изменит наш подход к обработке данных и искусственному интеллекту.” - Дмитрий Коробченко, инженер NVIDIA”[3]. Нейросети уже сегодня играют ключевую роль в медицине, финансах, автономных транспортных средствах и искусстве, и их влияние будет только расти.

Одним из наиболее важных направлений развития является создание самообучающихся систем, которые могут непрерывно улучшать свои навыки и знания, работая с реальными данными. Это открывает путь к самозволюции и самопознанию искусственного интеллекта, что является увлекательным направлением исследований.

В то же время, необходимо учитывать потенциальные риски и этические дилеммы, связанные с использованием нейросетей. В частности, их применение в военном деле вызывает серьезные опасения, так как они могут быть использованы для создания автономного оружия, что ставит под вопрос возможность контроля за такими системами и соответствие этическим нормам.

“С каждым годом нейросети становятся все более мощными и автономными, что открывает новые горизонты для их применения в самых разных сферах.”- Источник из Skillbox Media. Это подчеркивает необходимость разработки строгих этических регламентов и безопасностных мер для обеспечения позитивного вклада нейросетей в общество и минимизации возможных негативных последствий их использования.

Заключение: Нейронные сети играют ключевую роль в развитии искусственного интеллекта и имеют огромный потенциал для трансформации различных сфер жизни, от медицины и образования до промышленности и финансов. Однако для реализации этого потенциала необходимо решить ряд проблем, связанных с вычислительными требованиями, недостатком данных и интерпретируемостью моделей. При этом продолжение исследований и инноваций в области нейронных сетей позволит преодолеть эти препятствия и раскрыть всю их силу в будущем.

Список использованных источников:

1. *medium.com* [Электронный ресурс]. Режим доступа: <https://medium.com/nuances-of-programming/o-нейронных-сетях-в-двух-словах-dc44de042ee5>. Дата доступа: 05.04.2024.
2. *bing.com* [Электронный ресурс]. Режим доступа: <https://www.bing.com>. Дата доступа: 05.04.2024.
3. *skiiibox.ru* [Электронный ресурс]. Режим доступа: <https://skiiibox.ru/media/business/inzhener-nvidia-o-budushchem-neyrosetey>. Дата доступа 05.04.2024.