

ХАРАКТЕРИСТИКИ РАЗВИТИЯ НЕЙРОННЫХ СЕТЕЙ

Федорович В.Г.

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

Научный руководитель: Косарева Е.М. – ассистент кафедры ПИКС

Аннотация. Нейронные сети стремительно захватывают рынок информационных технологий, в статье рассматриваются характеристики нейронных сетей, динамика их изменения, отличия проектирования нейронных сетей в нынешнее время в сравнении с вариациями прошлых лет.

Ключевые слова: разработка искусственного интеллекта, нейронные сети, программирование, характеристика.

Введение. Нейронные сети – это математические модели, вдохновленные структурой и функционированием мозга.

Они состоят из множества взаимосвязанных элементов (нейронов), которые обрабатывают информацию и учатся на данных. Проще говоря, это алгоритмы, которые помогают компьютерам учиться и принимать решения, похожие на человеческие.

Основная часть. Проанализировав изменение в архитектуре нейронных сетей, можно выявить основные аспекты:

– нейронные сети состоят из множества слоев, каждый из которых содержит множество нейронов (нейронов). глубина сети (количество слоев) и количество нейронов в каждом слое играют важную роль в её способности обрабатывать сложные данные;

– важным аспектом является выбор активационных функций, которые помогают сети обрабатывать нелинейные зависимости;

– нейронные сети обучаются с помощью алгоритмов, таких как градиентный спуск и его вариации. эти алгоритмы позволяют сети адаптироваться к данным и улучшать свою точность;

– чтобы избежать переноса (*overfitting*), используются различные методы регуляризации, такие как *Dropout* и регуляризация;

– в компьютерном зрении часто используются конволюционные слои, которые позволяют сети обрабатывать изображения и выделять важные признаки.

Для примера, сравним по этим 5-ти параметрам примеры нейронных сетей 2017 года, 2020 года и 2024 года.

Архитектура нейронных сетей версии 2017 года:

1 Слои и глубина: В 2017 году архитектуры нейронных сетей, такие как *ResNet* и *Inception*, стали популярными благодаря увеличению глубины (более 100 слоев), что позволило моделям лучше распознавать сложные паттерны.

2 Активационные функции: Наиболее часто используемые функции активации были *ReLU (Rectified Linear Unit)*, которые помогали избежать проблем с затухающими градиентами.

3 Обучение: Алгоритмы обучения, такие как стохастический градиентный спуск (*SGD*) с *Momentum*, были основными методами оптимизации. Использовались большие объемы данных и продолжительное время тренировки.

4 Регуляризация: *Dropout* и *L2* регуляризация были стандартными методами для предотвращения переноса (*overfitting*).

5 Конволюционные слои: *CNN*, такие как *ResNet*, активно использовали конволюционные слои с остаточными соединениями, что улучшило производительность в задачах компьютерного зрения.

Архитектура нейронных сетей версии 2020 года:

1 Слои и глубина, в 2020 году архитектуры стали еще более глубокими и сложными. Например, модели трансформеров, такие как *BERT* и *GPT-2*, имели миллионы параметров и сотни слоев внимания.

2 Активационные функции – помимо *ReLU*, начали использоваться *GELU* (*Gaussian Error Linear Unit*), которые улучшали обучение моделей.

3 Обучение – алгоритмы адаптивного обучения, такие как *Adam* и *RMSprop*, стали стандартом, что ускорило процесс тренировки и улучшило сходимость моделей.

4 Регуляризация – помимо *Dropout*, начали использоваться техники, такие как *Batch Normalization* и *Layer Normalization*, которые стабилизируют процесс обучения.

5 Конволюционные слои – хотя *CNN* оставались популярными для компьютерного зрения, трансформеры начали находить применение и в этой области благодаря механизмам внимания.

Архитектура нейронных сетей версии 2024 года:

1 Архитектуры стали еще более сложными и масштабируемыми. Модели, такие как *GPT-4*, обладают миллиардами параметров и способны обучаться на огромных наборах данных.

2 Более продвинутые функции активации, такие как *Swish* и *Mish*, улучшили обучение и производительность моделей.

3 Обучение – использование обучения с подкреплением и смешанного обучения (*multi-task learning*) позволило моделям решать более сложные и разнообразные задачи. Также применялись более мощные вычислительные ресурсы, такие как специализированные аппаратные ускорители (*TPU*).

4 Регуляризации – новые методы регуляризации, такие как *Stochastic Depth* и *DropConnect*, были разработаны для улучшения обобщающих способностей моделей.

5 Модели компьютерного зрения продолжили развиваться с использованием трансформеров и новых методов конволюции, таких как *Vision Transformers (ViT)*.

Показательно сравнить эффективность некоторых отдельных элементов невозможно, в силу временного промежутка, недостатка информации и различия характера их работы.

Но мы можем уверенно сказать, что эффективность нейронных сетей 2024-го года превышает эффективность нейронных сетей 2017-го года в множество раз. Однако точное соотношение зависит от конкретных задач и моделей. Например, некоторые исследования показывают, что требуемое количество операций с плавающей запятой (*FLOPs*) для обучения моделей снизилось в несколько раз за этот период. Мы можем назвать лишь примерное количество проводимых операций в секунду, учитывая, что результат разнится в зависимости от архитектуры и от оборудования, в 2017 году количество операций в секунду для нейронной сети составляло $10^3 - 10^5$, а в 2024 году $10^6 - 10^9$, для сравнения в человеческом мозге происходит примерно 10^{16} операций в секунду. Даже используя один такой показатель можно сказать, что нейронные сети с каждым днем все ближе к тому, чтобы имитировать мозговую деятельность человека.

Заключение. Рост характеристик нейронных сетей, таких как слоистость и глубина, выбор активационных функций, совершенствование алгоритмов обучения, методы регуляризации и развитие конволюционных слоев, свидетельствует о значительном прогрессе в данной области.

Увеличение количества слоев и нейронов, улучшение обработки нелинейных зависимостей и внедрение передовых технологий позволяют сетям справляться с все более сложными задачами.

Эти улучшения способствуют повышению их точности, универсальности и эффективности, открывая новые горизонты для их применения в различных сферах, включая компьютерное зрение, анализ данных и искусственный интеллект.

Динамика роста данных характеристик обещает дальнейшее развитие и внедрение инноваций в научную и практическую деятельность.

Список литературы

1. Анисимова, Э. С. Введение в архитектуру нейронных сетей / Э. С. Анисимова // Экономика и социум. – 2015. – № 2–5 (15).
2. Гудфеллоу, Иэн. *Deep Learning* / Иэн Гудфеллоу, Йошуа Бенжю, Аарон Курвил. – MIT Press, 2016.
3. Заковоротный, А. Ю. Новые архитектуры и алгоритмы обучения нейронных сетей адаптивной резонансной теории / А. Ю. Заковоротный // Научный результат. Информационные технологии. – 2016. – № 1.
4. Кузнецова, И. О. Принцип работы и архитектура нейронных сетей / И. О. Кузнецова, Д. А. Малютов // Евразийская интеграция: современные тренды и перспективные направления. – 2024.
5. Малыгина, Ю. П. Нейронные сети: особенности, тенденции, перспективы развития / Ю. П. Малыгина // Молодой исследователь Дона. – 2018. – № 5 (14).

UDC 004.032.26

CHARACTERISTICS OF NEURAL NETWORK DEVELOPMENT

Fedorovich V.G.

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

Kosareva E.M. – assistant of the Department of ICSD

Annotation Neural networks are rapidly taking over the information technology market. This article explores the characteristics of neural networks, the dynamics of their development, and the differences in neural network design today compared to previous years.

Keywords: artificial intelligence development, neural networks, programming, characteristics.