

A thick black L-shaped frame surrounds the text. It starts at the top left, goes right, then down, then right again, and finally down to the bottom right corner.

# АРХИТЕКТУРЫ НЕЙРОННЫХ СЕТЕЙ

ВВЕДЕНИЕ

2023

# ОГЛАВЛЕНИЕ

1. Введение
  1. Нейрон
  2. Нейронные сети
2. Краткая история
3. Архитектуры нейронных сетей
  1. Сверточные сети
  2. Рекуррентные сети
  3. Глубокие многослойные перцептроны
  4. Трансформеры
4. Заключение
5. Источники

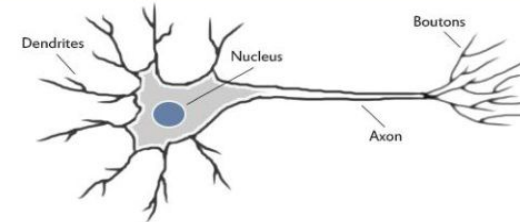
# Введение. Нейрон

- **Нейрон** в терминологии компьютерных нейронных сетей – это базовая единица нейронной сети. Он состоит из следующих компонентов:
  1. *Входы - это данные, которые поступают в нейрон от других нейронов.*
  2. *Веса - это коэффициенты, которые определяют, как сильно влияют входные данные на выход нейрона.*
  3. *Активационная функция - это функция, которая преобразует входные данные в выходные данные.*
  4. *Выходы - это данные, которые нейрон передает другим нейронам.*

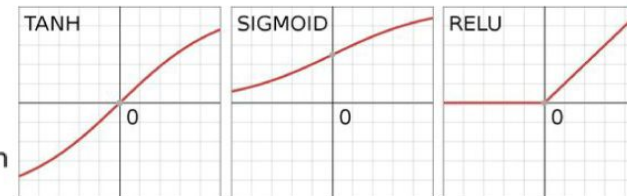
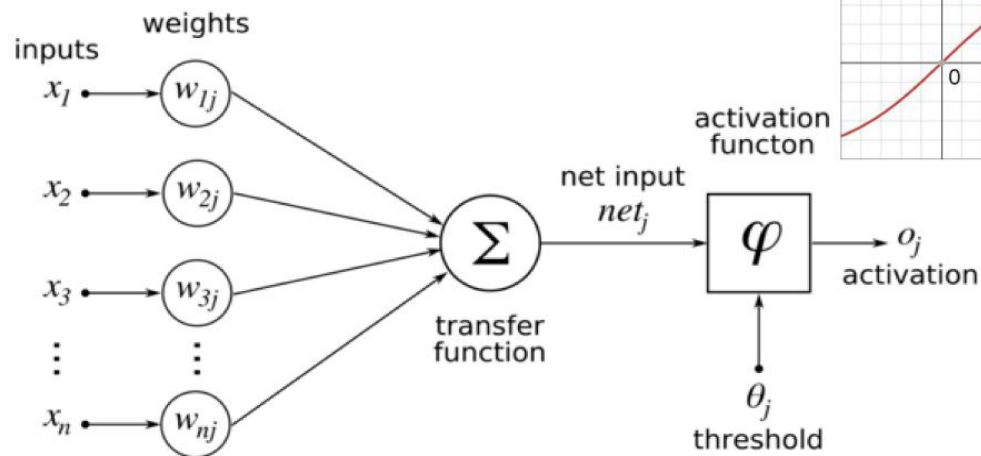
# Введение. Нейрон

В сущности нейрон представляет из себя функцию со входами. Входы умножаются на некие веса, далее веса суммируются, прогоняются через функцию и на выход передаётся результат функции. Соответственно нейросеть состоит из множества связанных нейронов.

## Искусственный нейрон



Искусственный нейрон — отдалённое подобие биологического. Базовый элемент искусственной нейронной сети



# Введение. Нейронные сети

- **Нейронные сети** - это искусственные интеллектуальные системы, которые имитируют работу человеческого мозга. Они состоят из множества нейронов, связанных друг с другом. Нейронные сети используются для решения широкого круга задач, включая распознавание изображений, распознавание речи, перевод языков и компьютерное зрение.

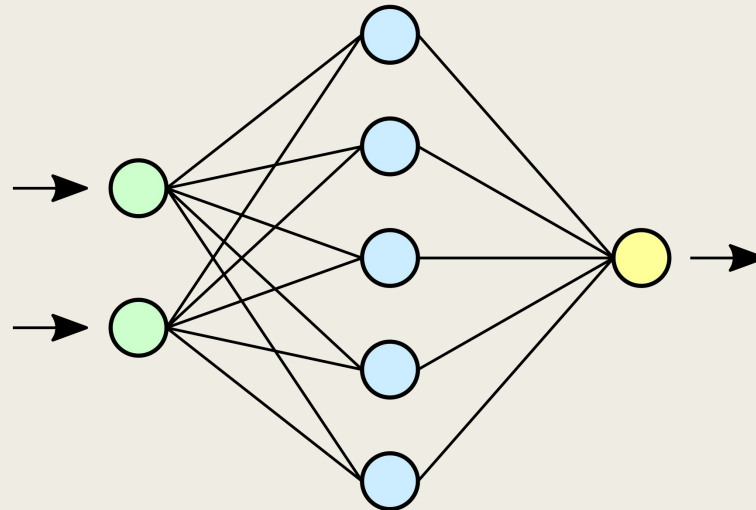


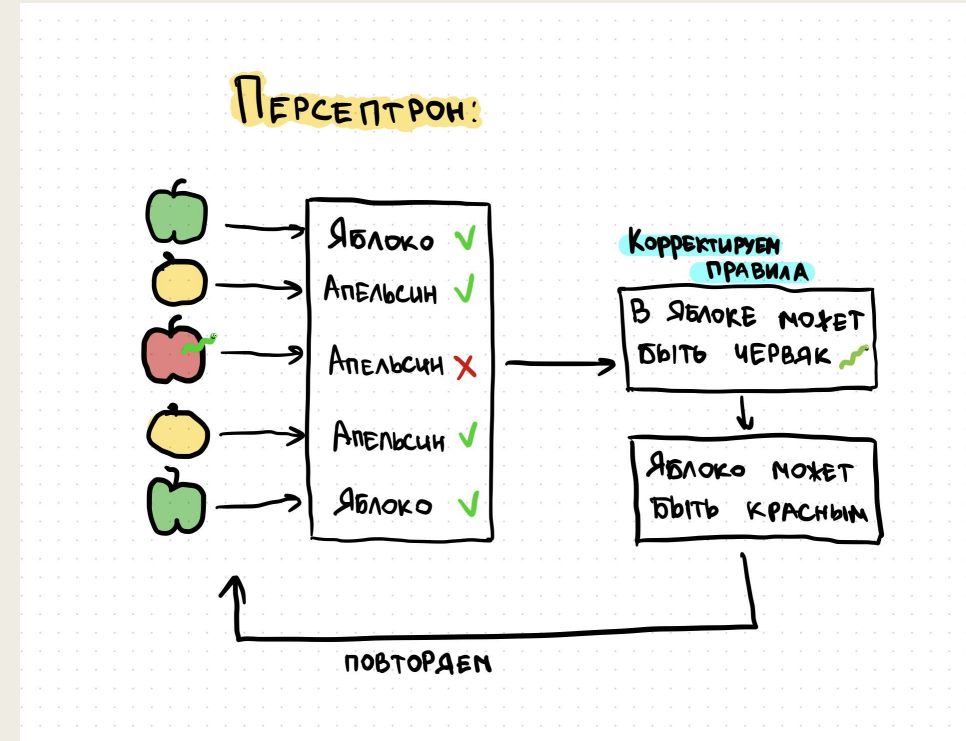
Схема простой нейросети. Зелёным цветом обозначены входные нейроны, голубым — скрытые нейроны, жёлтым — выходной нейрон

# Краткая история нейронных сетей (1940-1960)

- До нейросетей были первые наработки в области искусственного интеллекта и машинного обучения. Сама концепция нейросети начала формироваться в 40-50-х годах прошлого века. Тогда исследователи пытались создать машину, которая могла бы имитировать функции человеческого мозга.
- В 1943-м году Уоррен Мак-Каллок и Уолтер Питтс предложили математическую модель нейрона, а уже в конце 50-х Фрэнк Розенблатт представил персептрон (или же перцептрон). Именно персептрон можно назвать первой практической реализацией нейросети.

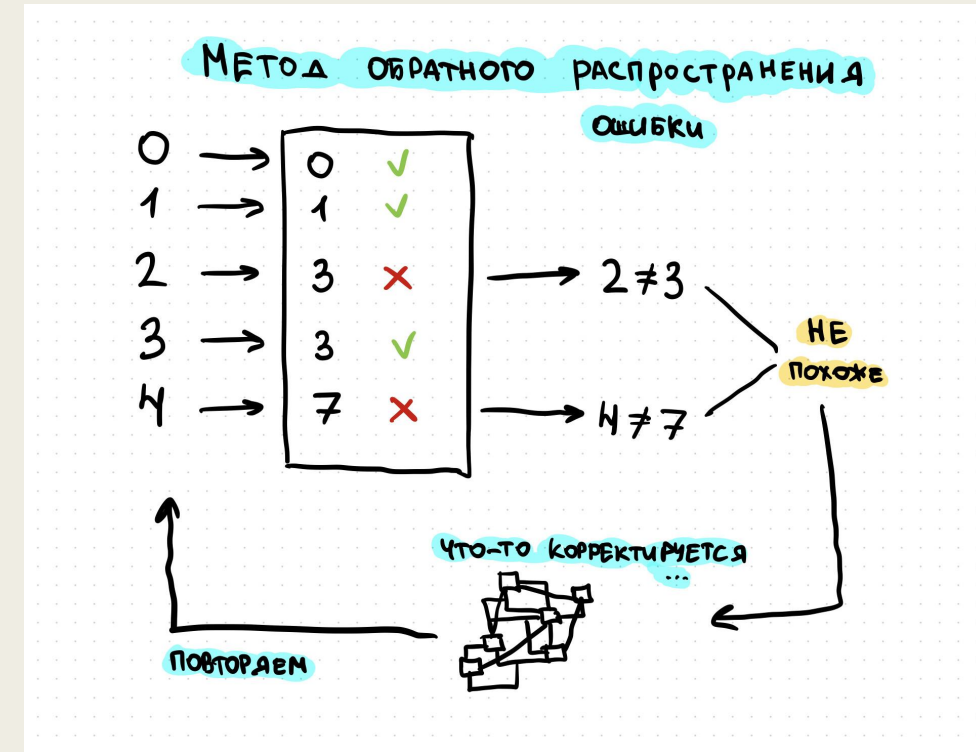
# Краткая история нейронных сетей

- **Персептрон** — это простая модель машинного обучения, которую создали для помощи компьютерам в обучении на разных данных. Допустим, мы хотим обучить персептрон классифицировать изображения на "яблоки" и "апельсины". Вот как мы будем это делать:
  1. Берём несколько изображений яблок и апельсинов.
  2. Даём персептрону по одной картинке. Он её обрабатывает и соотносит к яблоку или апельсину.
  3. Если ответ правильный, мы ничего не меняем. Если ответ неправильный, то мы корректируем правила внутри персептрона, чтобы в следующий раз он не ошибся.
  4. Повторяем процесс для всех изображений, пока он не перестает ошибаться.
  5. Далее мы тестируем персептрон уже с другими изображениями яблок и апельсинов.
- Это очень простой пример того, как можно обучить и протестировать персептрон. В реальности процесс более сложный и включает в себя настройку гораздо большего количества переменных и правил, но основная идея та же.
- Персептрон — это очень базовая модель, но она проложила путь к разработке более совершенных алгоритмов машинного обучения, которые используются сегодня.



# Краткая история нейронных сетей (1980-2000)

- Несмотря на крутость персептрона, тогдашние исследования и вычислительные возможности не позволяли сильно развить идею нейросетей. В какой-то степени, от неё даже на время отказались. Уже ближе к концу 20-го века исследователи добились большого прогресса. Был разработан «Метод обратного распространения ошибки»
- Допустим, нейросеть пытается распознать рукописные цифры от 0 до 9. Ей сначала дают примеры для обучения, а потом она их использует, чтобы выдавать собственные предположения. Нейросеть выдает предположение о конкретной цифре на изображении, а затем сравнивает это предположение с реальным значением и вычисляет разницу между ними (ошибку). Затем ошибка используется для корректировки весов нейронов, которые влияют на выходные данные нейросети. Процесс повторяется много раз, пока сеть не начинает распознавать цифры с максимально высокой точностью. До появления Метода обучать нейросети было сложно, потому что было трудно обновлять веса сети для оптимизации работы.
- В дополнение к Методу исследователи придумали нелинейные функции активации. Они позволяют сети моделировать сложные взаимосвязи между входами и выходами. Это заложило основу для недавнего "возрождения" нейронных сетей уже в лице Deep Learning (Глубокое обучение).
- Люди смогли обучать гораздо более крупные и сложные сети, что привело к прорывам в областях применения. Нейросети научились лучше распознавать изображения, речь и обрабатывать естественные языки.



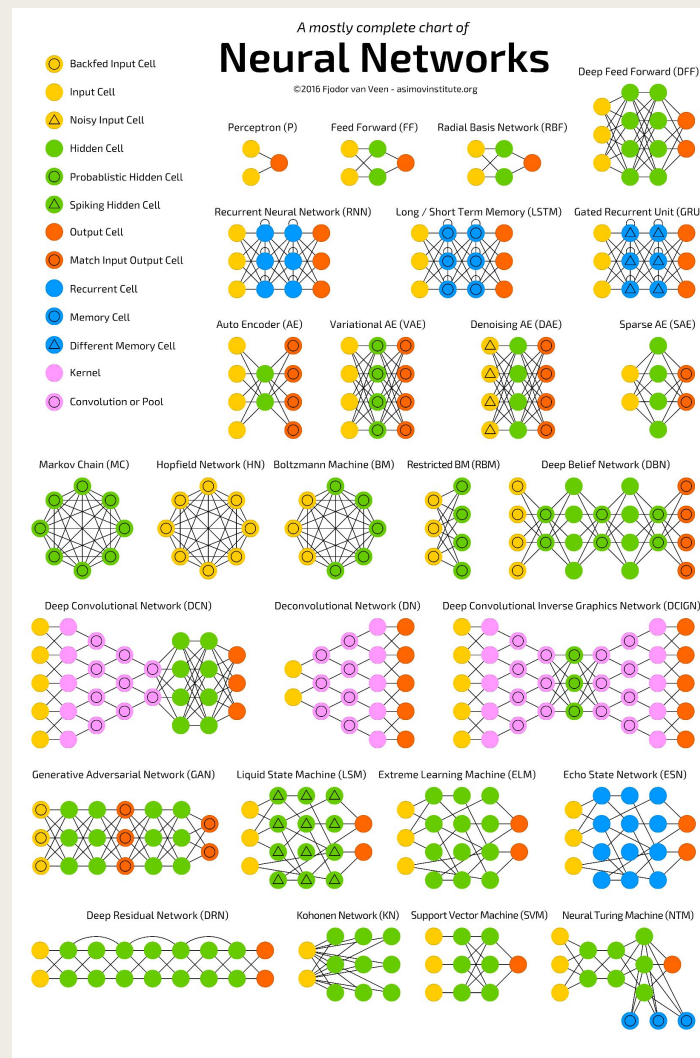


# Краткая история нейронных сетей (2000-2020)

- В "нулевых" появились мощные графические процессоры и стали доступны большие объёмы данных, что привело к разработке алгоритмов Deep Learning.
- Термин Deep Learning обрёл настоящую популярность только в 2010-х годах. Результаты исследований начали демонстрировать высокие результаты в обучении и составлении прогнозов на основе больших объёмов данных. Интерес к нейросетям начал сильно расти, а вместе с большим интересом пришли большие деньги.
- Развитие Deep Learning привело к разработке новых архитектур, алгоритмов и инструментов для построения и обучения нейросетей, а это привело к началу распространения практического применения во многих отраслях.

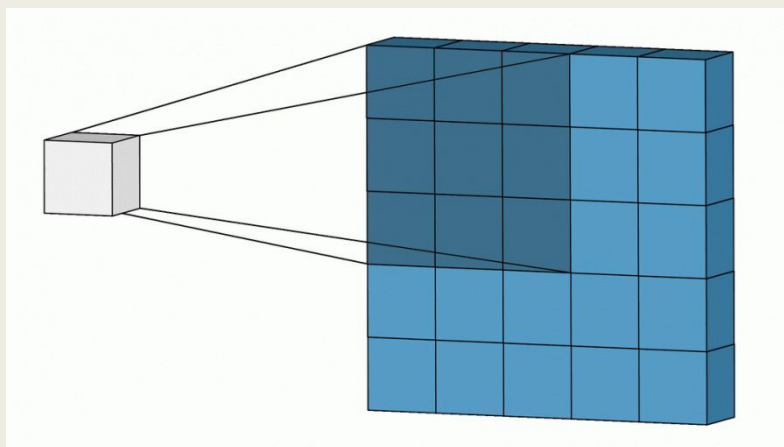
# Архитектуры нейронных сетей

- Новые виды архитектуры нейронных сетей появляются постоянно, и существует огромное их разнообразие. К самым распространённым архитектурам нейронных сетей относятся:
  1. Сверточные нейронные сети (CNN) - это тип нейронных сетей, который используется для обработки изображений и видео. Они хорошо подходят для задач распознавания образов, классификации изображений и сегментации изображений.
  2. Рекуррентные нейронные сети (RNN) - это тип нейронных сетей, который обрабатывает данные последовательно. Они хорошо подходят для задач обработки естественного языка, таких как перевод, распознавание речи и генерация текста.
  3. Глубокие многослойные перцептроны (MLP) - это тип нейронных сетей, который состоит из нескольких слоев полносвязных нейронов. Они хорошо подходят для широкого круга задач, включая классификации, регрессию и кластеризацию.
  4. Трансформеры - это тип нейронных сетей, который используется для обработки естественного языка. Они хорошо подходят для задач машинного перевода, обработки вопросов и ответов и генерации текста.



# Сверточные нейросети

- Сверточная нейронная сеть (CNN) — это тип нейронной сети, которая используется для обработки данных, имеющих сетчатую структуру, например, изображений. CNN состоят из нескольких слоёв, каждый из которых выполняет определённое преобразование над входными данными.
- Сверточные нейросети решают 3 основные задачи:
  1. Классификация. Вы подаете картинку, и нейросеть просто говорит — у вас картинка про собаку, про лошадь, еще про что-то, и выдает класс.
  2. Детекция – это более продвинутая задачка, когда нейросеть не просто говорит, что на картинке есть собака или лошадь, но находит еще Bounding box — где это находится на картинке.
  3. Сегментация. По сути, это попиксельная классификация. Здесь мы говорим про каждый пиксель изображения: этот пиксель относится к собаке, этот — к лошади, а этот еще к чему-то. На самом деле, если вы умеете решать задачу сегментации, то предыдущие 2 задачи уже автоматически даны.



# Сверточные нейросети

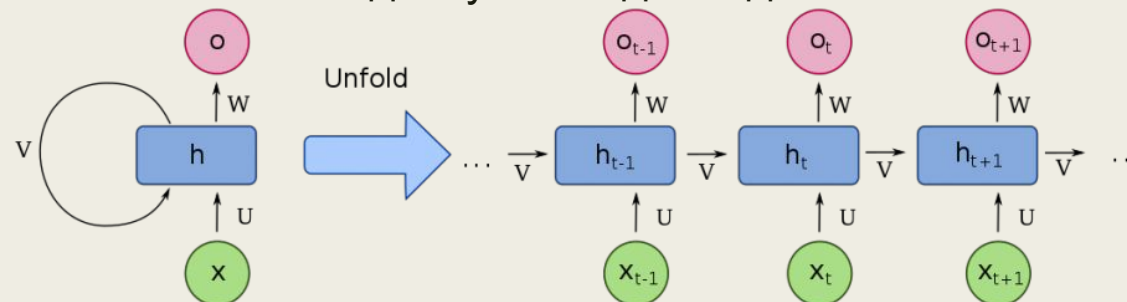
- Сверточные нейронные сети работают следующим образом:
  1. Входные данные, например, изображение, подаются на вход CNN.
  2. Сверточный слой выполняет операцию свертки, которая представляет собой умножение элементов входного массива с элементами весового фильтра.
  3. Функция активации применяется к выходу сверточного слоя. Функция активации позволяет нейронной сети учиться на входных данных.
  4. Результат сверточного слоя подаётся на следующий слой CNN.
  5. Эти шаги повторяются до тех пор, пока CNN не достигнет последнего слоя. Последний слой обычно является полносвязным слоем, который выполняет классификацию или регрессию.

# Сверточные нейросети

- Плюсы сверточных нейронных сетей:
  1. Высокая точность — CNN могут достигать высокой точности на задачах обработки изображений.
  2. Эффективность — CNN могут обрабатывать данные очень быстро.
  3. Универсальность — CNN могут использоваться для решения широкого круга задач обработки изображений.
- Минусы сверточных нейронных сетей:
  1. Требуются большие объёмы данных для обучения — CNN требуют больших объёмов данных для обучения.
  2. Требуется много вычислительных ресурсов для обучения — обучение CNN требует много вычислительных ресурсов.

# Рекуррентные нейросети

- Рекуррентные нейронные сети (RNN) - это тип нейронных сетей, который обрабатывает данные последовательно. Они хорошо подходят для задач обработки естественного языка, таких как перевод, распознавание речи и генерация текста. В отличие от обычных нейронных сетей, которые обрабатывают данные в виде независимых точек данных, RNN могут обрабатывать данные как последовательные цепочки. Это позволяет им учитывать контекст предыдущих данных при обработке текущих данных.
- Структура RNN состоит из нескольких слоев нейронов, соединенных друг с другом петлями. Каждый нейрон в слое получает входные данные от нейронов предыдущего слоя, а также от себя самого в предыдущем состоянии.
- RNN могут быть однослойными или многослойными. Однослойные RNN имеют только один слой нейронов, а многослойные RNN имеют несколько слоев нейронов, соединенных друг с другом.
- RNN могут быть как скрытые, так и видимые. Скрытые RNN не имеют доступа к входным данным, а видимые RNN имеют доступ к входным данным.



# Рекуррентные нейросети

- RNN используются для решения широкого круга задач обработки естественного языка, включая:
  1. Перевод. RNN могут использоваться для перевода текстов из одного языка на другой.
  2. Распознавание речи. RNN могут использоваться для распознавания речи и преобразования ее в текст.
  3. Генерация текста. RNN могут использоваться для генерации текста, например, для написания статей, создания сценариев или создания музыкальных произведений.
- Они могут обрабатывать данные последовательно. Это позволяет им учитывать контекст предыдущих данных при обработке текущих данных. Они могут быть использованы для решения широкого круга задач. RNN используются для решения задач обработки естественного языка, обработки сигналов и обучения машин. Они продолжают развиваться и улучшаться. RNN становятся все более мощными и эффективными.

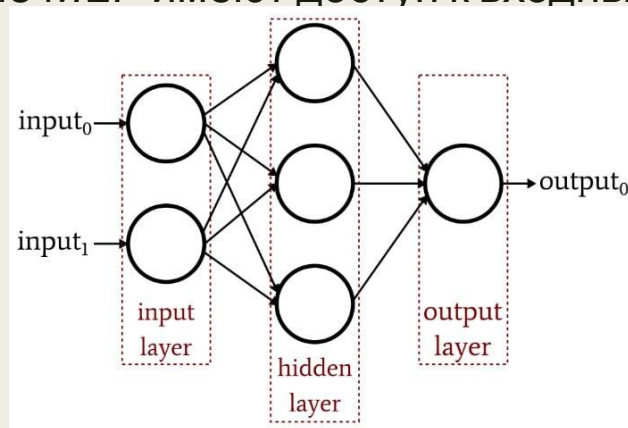
# Рекуррентные нейросети

- Однако рекуррентные сети не лишены недостатков:
  1. Они могут быть подвержены проблемам с заикливанием. Если RNN не будет правильно обучена, она может заиклиться на определенном состоянии.
  2. Они могут быть вычислительно дорогостоящими. Обучение RNN может быть сложной и вычислительно дорогостоящей задачей.
- Примерами рекуррентных нейросетей могут служить Google Translate, Siri, Alexa и GPT-3.



# Глубокие многослойные перцептроны

- Глубокие многослойные перцептроны (MLP) - это тип нейронных сетей, который состоит из нескольких слоев полносвязных нейронов. Они хорошо подходят для широкого круга задач, включая классификации, регрессию и кластеризацию. MLP работают, путем обработки входных данных через несколько слоев нейронов. Каждый слой нейронов принимает входные данные от предыдущего слоя и вычисляет выходной сигнал. Выходной сигнал последнего слоя используется для классификации или прогнозирования выходных данных.
- Такие нейросети могут быть однослойными или многослойными. Однослойные MLP имеют только один слой нейронов, а многослойные MLP имеют несколько слоев нейронов, соединенных друг с другом.
- Также они могут быть как скрытые, так и видимые. Скрытые MLP не имеют доступа к входным данным, а видимые MLP имеют доступ к входным данным.



# Глубокие многослойные перцептроны

- MLP используются для решения широкого круга задач, включая:
  1. Классификация. MLP могут использоваться для классификации объектов по категориям, например, для классификации изображений собак и кошек.
  2. Регрессия. MLP могут использоваться для прогнозирования значений непрерывных переменных, например, для прогнозирования цены акции или погоды.
  3. Кластеризация. MLP могут использоваться для группировки объектов, которые похожи друг на друга, например, для группировки клиентов по их покупательскому поведению.

# Глубокие многослойные перцептроны

## ■ Плюсы:

1. Они универсальны. MLP могут использоваться для решения широкого круга задач.
2. Они хорошо изучены. Существует большое количество литературы по MLP, что делает их относительно простыми для понимания и использования.
3. Они доступны. Существует множество библиотек и фреймворков, которые упрощают использование MLP.

## ■ Минусы:

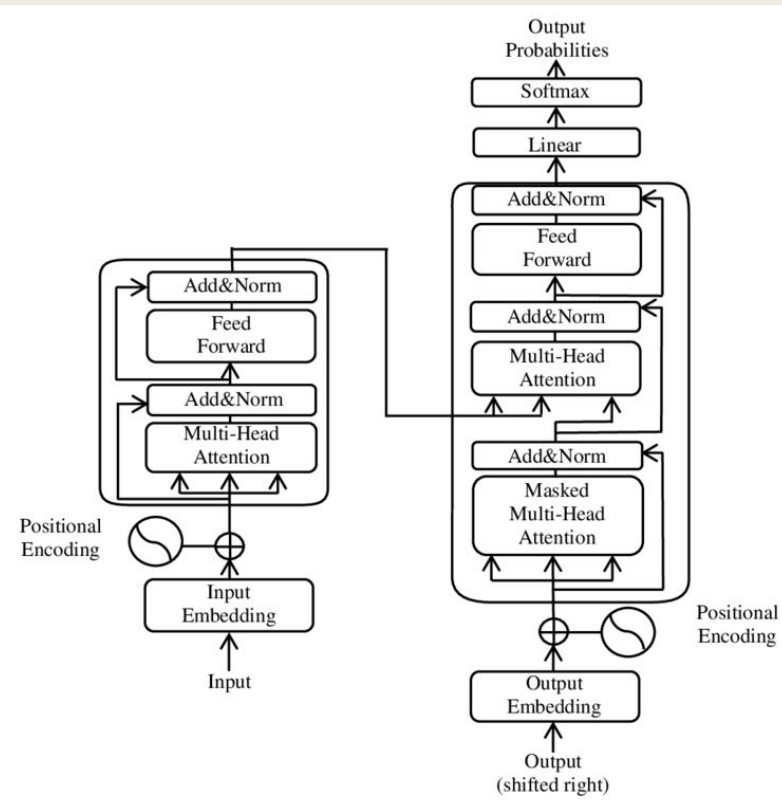
1. Они могут быть сложны в обучении. Обучение MLP может быть сложной и вычислительно дорогостоящей задачей.
2. Они могут быть подвержены переобучению. Если MLP будет обучена на слишком большом количестве данных, она может начать распознавать несуществующие паттерны.

# Трансформеры

- Трансформеры - это тип нейронных сетей, которые используются для обработки естественного языка. Они хорошо подходят для задач машинного перевода, обработки вопросов и ответов и генерации текста.
- Трансформеры работают, путем обработки входных данных через несколько слоев внимания. Внимание - это механизм, который позволяет сети учитывать контекст предыдущих данных при обработке текущих данных.
- Трансформеры состоят из нескольких слоев внимания. Каждый слой внимания принимает входные данные от предыдущего слоя и вычисляет выходной сигнал. Выходной сигнал последнего слоя используется для классификации или прогнозирования выходных данных.

# Трансформеры

- Трансформеры используются для решения широкого круга задач обработки естественного языка, включая:
  1. Машинный перевод.
  2. Обработка вопросов и ответов.
  3. Генерация текста.



# Трансформеры

- Плюсы:

1. Они более устойчивы к шуму и ошибкам в данных.
2. Они могут быть масштабированы до больших размеров, что позволяет им решать более сложные задачи.

- Минусы:

1. Они могут быть сложны в обучении. Обучение трансформеров может быть сложной и вычислительно дорогостоящей задачей.
2. Они могут быть подвержены переобучению. Если трансформер будет обучен на слишком большом количестве данных, он может начать распознавать несуществующие паттерны.

# ЗАКЛЮЧЕНИЕ

Новые виды архитектуры нейронных сетей появляются постоянно, поэтому в данной презентации были рассмотрены лишь самые популярные их классы, а именно сверточные сети, рекуррентные сети, глубокие многослойные перцептроны и трансформеры. Каждая из этих архитектур решает свой конкретный набор задач.

# ИСТОЧНИКИ

- Нейронные сети для начинающих. Habr. [Электронный ресурс] - <https://habr.com/ru/articles/312450/>
- Введение в архитектуры нейронных сетей. Олег Бунин. Habr [Электронный ресурс] - <https://habr.com/ru/companies/oleg-bunin/articles/340184/>
- Очень краткая история нейросетей: от разработок 20-го века до ChatGPT. Habr [Электронный ресурс] - <https://vc.ru/future/606777-ochen-kratkaya-istoriya-neyrosetey-ot-razrabotok-20-go-veka-do-chatgpt>
- Шпаргалка по разновидностям нейронных сетей. TProger [Электронный ресурс] - <https://tproger.ru/translations/neural-network-zoo-1>
- Что такое нейронная сеть? AWS [Электронный ресурс] - <https://aws.amazon.com/ru/what-is/neural-network/>