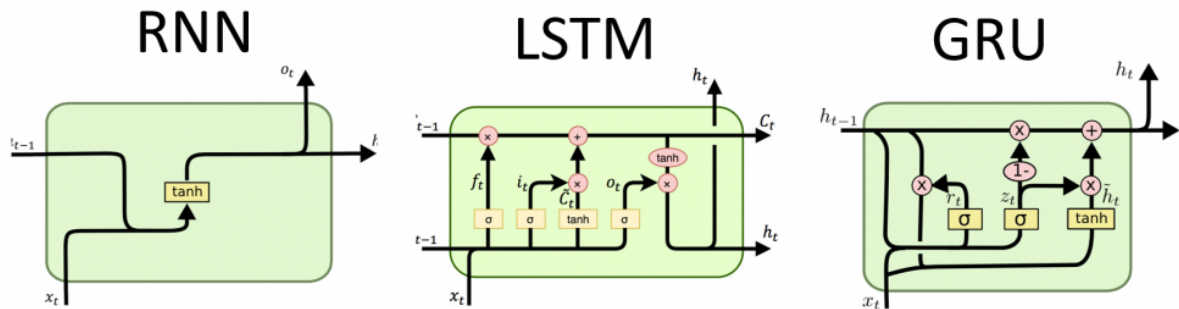


Архитектура RNN

Рекуррентные нейронные сети (RNN) — это класс нейронных сетей, которые хороши для моделирования последовательных данных, таких как временные ряды или естественный язык.



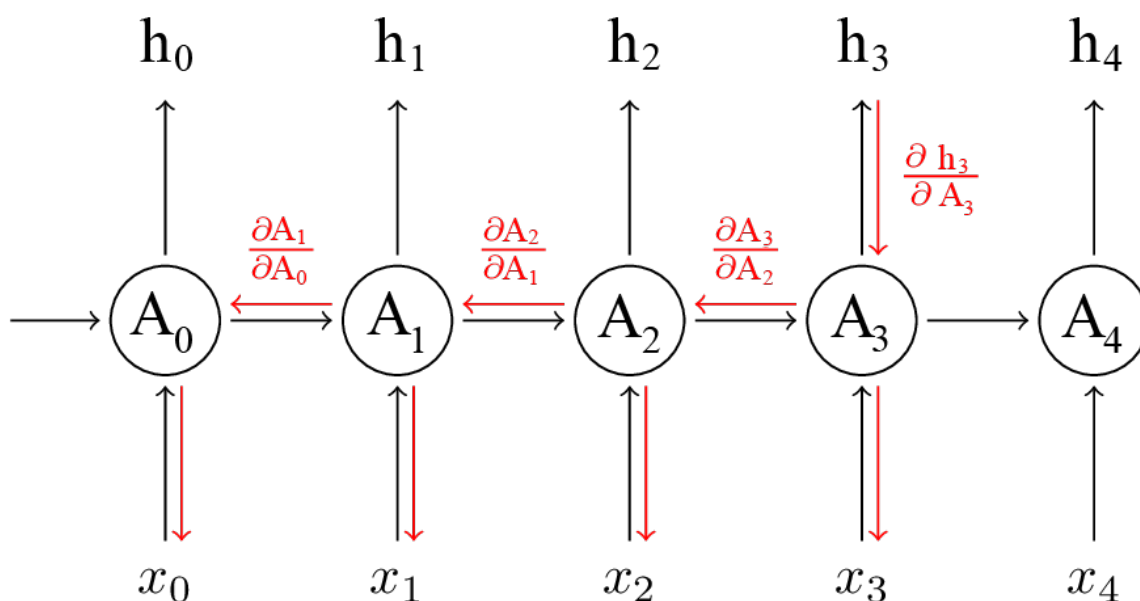
Схематично слой RNN использует цикл `for` для итерации по упорядоченной по времени последовательности, храня при этом во внутреннем состоянии, закодированную информацию о шагах, которые он уже видел.

Keras RNN API разработан с фокусом на:

- **Простоту использования:** встроенные слои `tf.keras.layers.RNN`, `tf.keras.layers.LSTM`, `tf.keras.layers.GRU` позволяют быстро построить рекуррентные модели без необходимости делать сложные конфигурационные настройки.
- **Простота кастомизации:** есть возможность задать собственный слой ячеек RNN (внутреннюю часть цикла `for`) с кастомным поведением и использовать его с общим слоем `tf.keras.layers.RNN` (сам цикл `for`). Это позволит быстро прототипировать различные исследовательские идеи в гибкой манере, с минимумом кода.

Обучение RNN аналогично обучению обычной нейронной сети. Здесь также используется алгоритм обратного распространения ошибки (Backpropagation), но с небольшим изменением. Поскольку одни и те же параметры используются на всех временных этапах в сети, градиент на каждом выходе зависит не только от расчетов текущего шага, но и от предыдущих временных шагов. Например, чтобы вычислить градиент для четвертого элемента последовательности, нам нужно было бы

«распространить ошибку» на 3 шага и суммировать градиенты. Этот алгоритм называется «алгоритмом обратного распространения ошибки сквозь время»:



В процессе работы RNN передают входные последовательности данных на скрытые слои. Также они имеют самозацикленный или повторяющийся рабочий процесс: скрытый слой может запоминать предыдущие входные данные в компоненте краткосрочной памяти и использовать их для будущих прогнозов. Модель использует текущее входное значение и сохраненное в памяти значение для прогнозирования следующей последовательности.

Для примера рассмотрим последовательность Apple is red. Вы хотите, чтобы RNN предсказывала слово red при получении входной последовательности Apple is. Когда скрытый слой обрабатывает слово Apple, он сохраняет его копию в своей памяти. Затем, увидев слово is, модель извлекает слово Apple из памяти и понимает контекст всей последовательности: Apple is. Это позволяет ей с высокой точностью прогнозировать слово red. Благодаря этому RNN полезны для распознавания речи, машинного перевода и других задач языкового моделирования.

RNN часто имеют архитектуру «один к одному», где одна входная последовательность строго сопоставляется с одним выходом. Но вы можете гибко настраивать их в разных конфигурациях с учетом конкретных целей. Ниже приведены несколько распространенных типов RNN.

Один ко многим

Такой тип RNN направляет один вход на несколько выходов. Он позволяет создать лингвистические приложения, например для создания субтитров к изображениям, которые генерируют предложение по одному ключевому слову.

Многие ко многим

Такая модель использует несколько входов для прогнозирования нескольких выходов. Например, вы можете создать с помощью RNN программу перевода, которая анализирует предложение на одном языке и правильно выстраивает слова на другом.

Многие к одному

Здесь несколько входов сопоставляются с одним выходом. Это полезно для таких приложений, как анализ настроений, в которых модель прогнозирует категорию «положительный, отрицательный или нейтральный» для настроения клиента, исходя из текста его отзыва.

Область применения.

Используются, когда важно соблюдать последовательность, когда важен порядок поступающих объектов.

- **Обработка естественного языка:**
 - Анализ текста - RNN используются для анализа тональности отзывов и комментариев. Например, модели RNN могут анализировать отзывы о продуктах на Amazon, чтобы классифицировать их как положительные, отрицательные или нейтральные. Успешные применения включают системы рекомендаций, основанные на анализе тональности
Актуальность: RNN по-прежнему применяются, но их постепенно заменяют модели на основе трансформеров, такие как BERT и GPT, которые демонстрируют лучшую производительность при обработке текста, поскольку способны захватывать более сложные контексты
 - Автоматический перевод - RNN используются в системах машинного перевода, таких как Google Translate, для перевода предложений с одного языка на другой, сохраняя контекст и грамматическую структуру
Актуальность: В современных системах перевода RNN заменены архитектурами на основе трансформеров, такими как Transformer и его производные (например, T5). Эти модели работают быстрее и обеспечивают более качественные переводы, особенно для длинных предложений
- **Обработка аудио:**
 - Автоматическое распознавание речи - RNN используются в системах распознавания речи, например, в программном обеспечении Siri от Apple или Google Assistant. Модели RNN способны преобразовывать аудиосигналы в текст, учитывая последовательность звуков и интонацию, что позволяет улучшить качество распознавание
Актуальность: современные подходы, основанные на моделях трансформеров, таких как Wav2Vec, дают лучшие результаты в области распознавания речи. Они обеспечивают большую точность и устойчивость к шумам
- **Обработка видео:**
 - Прогнозирование следующего кадра на основе предыдущих

Актуальность: RNN всё ещё актуальны для задач, где требуется предсказание на основе временных последовательностей, однако появляются новые архитектуры, такие как Temporal Convolutional Networks (TCN)

- Распознавание эмоций - для распознавания эмоций в видео RNN могут использоваться для анализа временных последовательностей изображений. Модели, такие как ConvLSTM, объединяют сверточные и рекуррентные слои для оценки эмоционального состояния человека на основе его лицевой мимики и жестов.

Актуальность: архитектуры на основе трансформеров, такие как Vision Transformers (ViT), постепенно вытесняют RNN

- Обработка изображений:

- Прогнозирование следующего пикселя на основе окружения - технологии, используемые в Adobe Photoshop, применяют RNN для заполнения пробелов в изображениях, создавая реалистичные текстуры и цвета

Актуальность: RNN теряют актуальность в области обработки изображений, так как сверточные нейронные сети (CNN) и модели на основе GAN (Generative Adversarial Networks) обеспечивают лучшее качество и скорость при решении таких задач

- Генерация описания изображений - модели, такие как Show and Tell от Google, используют RNN для создания описаний на основе содержимого изображения. Они анализируют изображение с помощью сверточных нейронных сетей (CNN), а затем генерируют последовательность слов с помощью RNN

Актуальность: были заменены трансформерами, которые более эффективно обрабатывают как визуальные, так и текстовые данные. Модели, такие как CLIP и DALL-E, способны генерировать более качественные и точные описания изображений, чем классические RNN