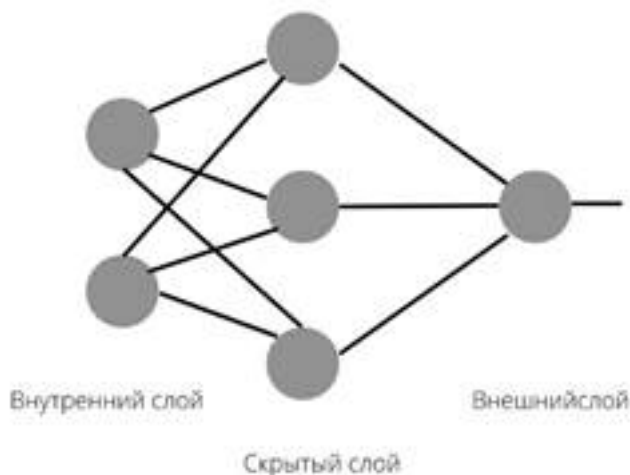


ИСПОЛЬЗОВАНИЕ LSTM-АРХИТЕКТУРЫ В ИНФОРМАЦИОННЫХ СИСТЕМАХ**Рахманкулов Ерман Даниярович**

магистрант, Евразийский национальный университет, РК, г. Нур-Султан

В целях распознавания речи, распознавания голоса, прогнозирования временных рядов и обработки естественного языка используется один из разновидностей нейронных сетей – рекуррентная нейронная сеть (РНС). Рекуррентные нейронные сети работают по принципу сохранения вывода определенного слоя и передачи его обратно на вход, чтобы предсказать вывод слоя, то есть РНС дают прогностические результаты в последовательных данных. Последовательные данные – это ряд упорядоченных данных, в которых связанные элементы следуют друг за другом. Примеры подобных данных можно наблюдать в финансовом секторе или в последовательности нуклеотидов молекулы ДНК. Одними из самых часто встречаемых типов последовательных данных являются данные, которые представляют собой примитивный ряд точек данных, перечисленных во временном порядке. Когда РНС принимает решение, сеть учитывает текущие входные данные, а также то, что она узнала из ранее полученных входных данных.

**Рисунок 1. Пример РНС-архитектуры**

Наблюдается два недостатка, с которыми сталкиваются при использовании архитектуры РНС. Данные недостатки связаны с градиентом. Градиент представляет собой частную производную по собственным входным данным. Градиент измеряет, насколько изменится вывод функции, если изменять входные данные. Чем выше градиент, тем быстрее может обучаться модель. Но если градиент равен нулю, модель прекращает обучение. Градиент просто измеряет изменение всех весов относительно изменения ошибки. Длинная цепь элементов краткосрочной памяти представляет собой особый вид рекуррентной нейронной сети, способной обрабатывать долгосрочные зависимости. Исчезающие градиенты возникают вследствие того, что значения градиента слишком малы, и в результате модель перестает

обучаться или занимает слишком много времени. Для решения проблемы исчезающих градиентов разработана концепция сети с долгосрочной краткосрочной памятью.

Сети с долгой краткосрочной памятью (LSTM) являются расширением рекуррентных нейронных сетей, которые в основном увеличивают память. LSTM имеет структуру, подобную цепочке, но повторяющийся модуль имеет другую структуру. Вместо одного слоя нейронной сети их четыре, взаимодействующих особым образом (Рисунок 2).

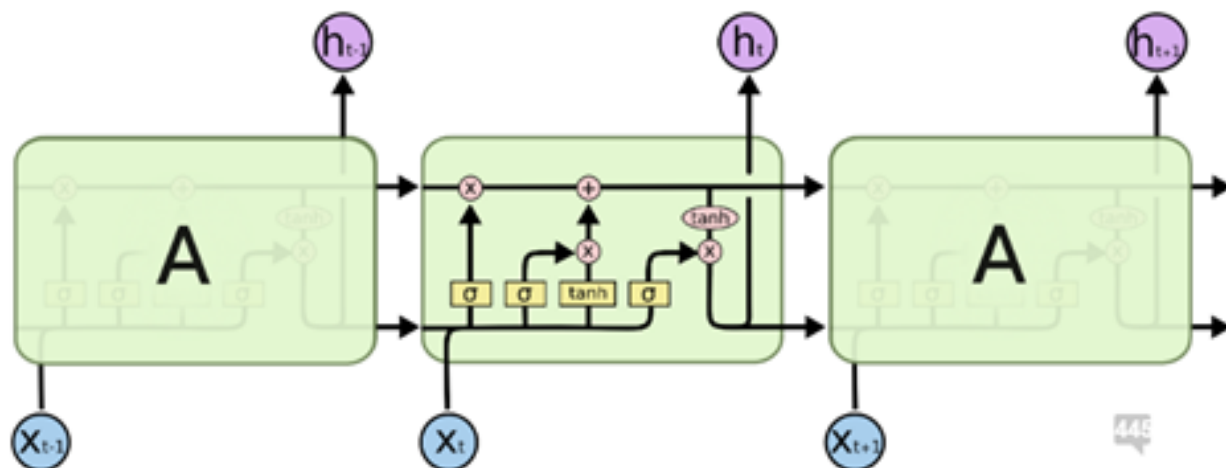


Рисунок 2. Архитектура LSTM сети

На приведенной выше диаграмме каждая линия несет целый вектор от выхода одного узла к входам других. Розовые кружки обозначают точечные операции, такие как сложение векторов, а желтые прямоугольники — слои обученной нейронной сети. Слияние строк означает конкатенацию, а разветвление строки означает, что ее содержимое копируется, а копии перемещаются в разные места.

Во-первых, на базовом уровне выходные данные LSTM в определенный момент времени зависят от трех вещей:

- Текущая долговременная память сети, известная как состояние ячейки
- Выходные данные в предыдущий момент времени — известное как предыдущее скрытое состояние
- Входные данные на текущем временном шаге

LSTM используют серию «ворот», которые контролируют, как информация в последовательности данных поступает, сохраняется и покидает сеть. В типичном LSTM есть три типа ворот; забытые ворота, входные ворота и выходные ворота. Эти ворота можно рассматривать как фильтры, и каждый из них представляет собой собственную нейронную сеть. Первым шагом в данном процессе являются забытые ворота. На данном этапе определяется какие биты состояния ячейки полезны, учитывая какое было предыдущее скрытое состояние, так и новые входные данные. Второй этап включает в себя новую сеть памяти и входные ворота. Цель этого этапа — определить, какую новую информацию следует добавить в долговременную память сети (состояние ячейки), но при этом учитывая предыдущее скрытое состояние и новые входные данные, которые уже были зафиксированы на предыдущем этапе. После обновления в долговременной памяти сети завершены, идет переход к последнему этапу, к выходным воротам, определяющим новое скрытое состояние. Для решения данной задачи необходимы:

- недавно обновленное состояние ячейки;
- предыдущее скрытое состояние;

- новые входные данные.

В зависимости от анализируемой информации описанный выше процесс может продолжаться в определенное количество итераций, пока LSTM не придет к результату.

Как и было сказано ранее, LSTM-архитектуры успешно применяются для прогнозирования временных рядов, в том числе распознавания звуков рядов, распознавания речи и ритму, и их обучению. Подобный архитектурный подход внедрен в информационные системы, предназначенные для перевода речи на язык жестов и наоборот. LSTM успешно внедрена в процессы управления автоматизированными работами. Помимо этого, технологии на основе LSTM применяются в области управления бизнес-процессами, анализе рынка труда, прогнозирования путей оказания медицинской помощи, в логистике аэропортов и т.д.

Список литературы:

1. A Gentle Introduction to Long Short-Term Memory Networks by the Experts - [Электронный ресурс] - Режим Доступа. -URL: <https://machinelearningmastery.com/gentle-introduction-long-short-term-memory-networks-experts/> (Дата обращения: 16.04.2022).
2. What is LSTM? Introduction to Long Short Term Memory - [Электронный ресурс] - Режим Доступа. -URL: <https://intellipaat.com/blog/what-is-lstm/> (Дата обращения: 16.04.2022).
3. LSTM Neural Network: The Basic Concept - [Электронный ресурс] - Режим Доступа. -URL: <https://towardsdatascience.com/lstm-neural-network-the-basic-concept-a9ba225616f7> (Дата обращения: 16.04.2022).
4. History of Neural Networks - [Электронный ресурс] - Режим Доступа. -URL: https://www.sas.com/ru_ru/insights/analytics/neural-networks.html (Дата обращения: 16.04.2022).
5. Neural Networks and Deep Learning - [Электронный ресурс] - Режим Доступа. -URL: <http://neuralnetworksanddeeplearning.com/> (Дата обращения: 16.04.2022).