

УДК 004.891.2

СОЗДАНИЕ ИНТЕЛЛЕКТУАЛЬНОГО ЧАТ-БОТА НА ОСНОВЕ РЕКУРРЕНТНОЙ НЕЙРОННОЙ СЕТИ С РАСПОЗНАВАНИЕМ РЕЧИ

Криливец Олег Викторович студент группы 345*

Национальный аэрокосмический университет им. Н.Е. Жуковского «ХАИ»

Существуют актуальные задачи обработки данных, при решении которых мы сталкиваемся не с отдельными объектами, но с их последовательностями, т.е. порядок следования объектов играет существенную роль в задаче. Например, это задача распознавания речи, где мы имеем дело с последовательностями звуков или некоторые задачи обработки текстов на естественном языке, где мы имеем дело с последовательностями слов.

Для решения такого рода задач можно применять рекуррентные нейронные сети, которые в процессе работы могут сохранять информацию о своих предыдущих состояниях.

В отличие от обычной сети прямого распространения, входной образ рекуррентной сети — это не один вектор, но последовательность векторов $\{x_1, x_2, \dots, x_n\}$ $\{x_1, x_2, \dots, x_n\}$ векторы входного образа в заданном порядке подаются на вход, при этом новое состояние скрытого слоя зависит от его предыдущих состояний.

$$\begin{aligned}h(t) &= f(V \cdot x(t) + U \cdot h(t-1) + bh) \\y(t) &= g(W \cdot h(t) + by),\end{aligned}\tag{1}$$

где $x(t)$ — входной вектор номер t , $h(t)$ — состояние скрытого слоя для входа $x(t)$ ($h(0)=0$), $y(t)$ — выход сети для входа $x(t)$, U — весовая матрица распределительного слоя, W — весовая (квадратная) матрица обратных связей скрытого слоя, bh — вектор сдвигов скрытого слоя, V — весовая матрица выходного слоя, by — вектор сдвигов выходного слоя, f — функция активации скрытого слоя, g — функция активации выходного слоя.

Входной вектор сигнала поступает на группу нейронов INPUT, на группе нейронов CONTEXT нулевой сигнал. Далее сигнал распространяется в группу нейронов скрытого слоя HIDDEN, а затем преобразуется ими и попадает на нейроны выходного слоя OUTPUT. На следующей итерации вместе с вектором сигнала INPUT, на контекстную группу нейронов поступает копии сигналов с выходного слоя OUTPUT прошлой итерации.

Структура рекуррентной нейронной сети для запоминания предложений имеет следующий вид:

Слой CONTEXT, INPUT и OUTPUT имеют по одному нейрону, значения сигнала на выходе которого ставится в соответствие индексу слова в наборе слов. Дополнительно вводится слово `__end__` соответствующее концу предложения. Объем слоев HIDDEN должен позволять запоминать весь набор предложений. Сеть последовательно обучается предложениям вида:

«Привет. Как дела? `__end__` Привет. Нормально. `__end__`»

В процессе обучения нейросети, использован подход обучения с учителем. Для этого каждый (дискретный) шаг времени на входные узлы подаются данные, а прочие узлы завершают свою активацию, и выходные данные готовятся для передачи нейронам следующего уровня. То есть в нашем случае на выходные узлы поступают распознанные слова. В качестве обучающей выборки был собран корпус часто задаваемых вопросов абитуриентов за последнюю вступительную кампанию в НАУ «ХАИ». Сеть обучается методом обратного распространения ошибки.

Методом простого подбора, наиболее удовлетворительный результат показала нейросеть с пятью HIDDEN слоями по 700 нейронов в каждом (+1 слой репрезентативных данных, тоже из 700 нейронов).

Таким образом была создана модель простейшего чат-бота на основе рекуррентной нейронной сети. Данный вид нейросети позволяет получать ответы на вопросы, которых не было в базе знаний, с помощью способности сети к обобщению.