

РЕКУРРЕНТНЫЕ НЕЙРОННЫЕ СЕТИ

Сычѳв А.Ю., Стаселько И.Д., Аниховский М.А.

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

Стержанов М.В. – к.ф.-м.н., доцент

К нейронным сетям с обратными связями относят рекуррентные нейронные сети. Рекуррентные нейронные сети показывают хорошие результаты во многих задачах связанных с классификацией последовательностей, например, распознавание речи и подавление шума, более того сети имеют одни из лучших результатов на данный момент в некоторых задачах связанных с обработкой текста на естественном языке, например, в задаче классификации текста. Каждая цепочка вычисляет ошибку как суммарную девиацию по выходным сигналам сети. Если имеется набор образцов, ошибка вычисляется с учётом ошибок каждого отдельного образца.

Рекуррентные нейронные сети — подкласс нейронных сетей с обратными связями, которые используют предыдущие состояния сети для вычисления текущего. Сеть строится из узлов, каждый из которых соединѳн со всеми другими узлами. У каждого нейрона порог активации меняется со временем и является вещественным числом. Каждое соединение имеет переменный вещественный вес[1]. Узлы разделяются на входные, выходные и скрытые. Простейший пример такой сети проиллюстрирован на рисунке 1. Такая нейронная сеть имеет обратные связи. Данное отличительное свойство сетей позволяет провести аналогию с биологическим мозгом, так как он тоже способен сохранять свои состояния и реагировать на события, опираясь на предыдущий опыт.

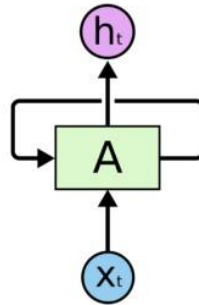


Рис. 1 - Архитектура рекуррентной нейронной сети с одним входным нейроном, одним скрытым нейроном и одним выходным нейроном

Основной класс задач, решаемый такими сетями - это обработка корпусов текстов, т.к. такие сети специально сделаны для того, чтобы обрабатывать последовательности данных неизвестной длины[2]. Каждая такая сеть имеет в себе особую структуру, похожую на звенья или модули, идущие один за другим. И каждый такой модуль представляет собой слой, с какой либо выходной функцией активации. Кроме этого, такая сеть может выполнять задачу предсказания слова в предложении, так как способна понимать контекст предложения. Рекуррентная нейронная может обрабатывать не только тексты, но и временные ряды, так как может обрабатывать их в двух направлениях. Способ обучения такой сети очень схож с обучением обычной нейронной сети. Разница лишь в том, что рекуррентная сеть требует использования метода обратного распространения ошибки во времени вместо привычного обратного распространения ошибки. Как следствие, на рекуррентные сети можно применять те же самые методы для решения проблемы переобучения, как и с обычными сетями.

Список использованных источников:

1. Elman J.L. Finding structure in time // Cognitive science, 1990. Vol. 14, no. 2. P. 179 – 211.
2. Lai S., Xu L., Liu K., Zhao J. Recurrent Convolutional Neural Networks for Text Classification. // AAAI, 2015. P. 2267 – 2273.