

ПРЕДСТАВЛЕНИЕ РЕКУРРЕНТНЫХ НЕЙРОННЫХ СЕТЕЙ В ВИДЕ НЕЛИНЕЙНЫХ ДИНАМИЧЕСКИХ СИСТЕМ

Чижов Пётр Сергеевич, Филиппов Андрей Александрович

Студент, студент

Факультет ВМК МГУ имени М. В. Ломоносова, Москва, Россия

E-mail: p.s.chizhov@gmail.com, andrey.filippov00@gmail.com

Научный руководитель — Ревизников Дмитрий Леонидович

В настоящее время рекуррентные нейронные сети (RNN) и остаточные нейронные сети (ResNet) широко применяются для решения самых различных задач в сфере искусственного интеллекта. В работе исследуется альтернативное представление указанных искусственных нейронных сетей в виде нелинейной динамической системы (Нейронные ОДУ/ Neural ODE) [1].

Указанные выше нейронные сети описывают последовательность преобразований скрытого состояния, которые можно представить в виде:

$$z_{t+1} = z_t + f(z_t, \theta_t), \quad (1)$$

в пределе, при увеличении глубины сети и переходе к непрерывному времени t , эти итеративные обновления z можно рассматривать как решение обыкновенного дифференциального уравнения (2), полученного методом Эйлера [2]:

$$\frac{dz(t)}{dt} = f(z(t), t, \theta), \quad (2)$$

Таким образом выход сети можно вычислять с помощью решателя задач Коши (ODE solver). Для обучения можно масштабировать алгоритм обратного распространение ошибки [3] через любой ODE solver без доступа к его внутренним операциям, с помощью решения сопряженной системы ОДУ [1]. Это позволяет проводить сквозное обучение нейронной ОДУ в рамках более крупных моделей.

В работе, с помощью пакета PyTorch, реализована описанная модель и соответствующая классическая рекуррентная нейронная сеть, проведено сравнение скорости и результатов работы.

На фиксированном наборе данных модель нейронных ОДУ продемонстрировала более стабильную и быструю сходимость, чем модель классической рекуррентной нейронной сети, что является хорошей предпосылкой для дальнейшего изучения представления ней-

ронных сетей в виде нелинейных динамических систем для более масштабных задач.

В заключение авторы выражают благодарность ведущему научному сотруднику ФИЦ ИУ РАН, д.ф.-м.н. Дмитрию Леонидовичу Ревизникову за научное руководство и помощь в проведении исследования.

Литература

1. Chen R. T. Q. et al. Neural ordinary differential equations //Advances in neural information processing systems. – 2018. – Т. 31..
2. Chang B. et al. Multi-level residual networks from dynamical systems view //arXiv preprint arXiv:1710.10348. – 2017.
3. Rumelhart D. E., Hinton G. E., Williams R. J. Learning representations by back-propagating errors //nature. – 1986. – Т. 323. – №. 6088. – С. 533-536.