

Минимизация набора сущностей для описания искусственных нейронных сетей и их свойств

В.Г. Абашин

Финансовый университет, Москва

Аннотация: В статье рассматриваются этапы выделения сущностей для построения языка описания искусственных нейронных сетей (ИНС), упрощающего обмен сетями между людьми. Разрабатываемый язык сводит задачу миграции ИНС между разными аппаратными и программными технологиями к задаче преобразования текстового описания на требуемую платформу, которая намного проще разработки интеллектуальных моделей. В результате минимизации набора сущностей получен язык, состоящий из полутора десятков тегов, позволяющих описывать наиболее популярные модели ИНС, используемые для задач классификации и распознавания образов.

Ключевые слова: искусственные нейронные сети, нейрон, синапс, многослойный персептрон, прямой проход, математическое моделирование.

Успехи в развитии многоядерных архитектур микропроцессоров привело к повышению популярности многопоточных вычислений. Одним из наиболее универсальных математических аппаратов, позволяющих достичь максимальной эффективности при использовании многопоточности является аппарат искусственных нейронных сетей (ИНС) [1-3].

Разработка ИНС требует значительных человеческих ресурсов. Часто требуется ручная подготовка данных для обучения, также для обучения необходим значительный объем машинного времени, а тестирование обученных моделей является отдельной задачей, решение которой невозможно без участия человека [4, 5]. С другой стороны значительная часть патентов на интеллектуальные модели в виде ИНС являются закрытыми [6, 7].

Серьёзная проблема требующая решения заключается в закрытости технологий, используемых для создания и эксплуатации ИНС. С одной стороны, многие технологии ИНС поставляются вместе с текстами программ,

их реализующими, а с другой, переносимость самих ИНС, между этими технологиями фактически отсутствует.

Описанные выше проблемы приводят к торможению развития интеллектуальных систем, основанных на ИНС, а их решение приведёт к появлению новых сценариев разработки ИНС, позволит внедрять этот механизм разработчикам, тестировщикам, интеграторам, экспертам, обучающимся, преподавателям.

Простота передачи ИНС между группами разработчиков упрощает интеграцию разработанных ИНС в различные аппаратные платформы. Этот эффект достигается за счёт того, что для разработки конвертора ИНС требуется меньшая квалификация, чем для её разработки.

Для унификации ИНС необходимо провести минимизацию сущностей, с помощью которых они будут описываться. Главным методом для исследования возможности минимизации сущностей является анализ математических моделей ИНС.

По мнению Хебба [8] обучение нейронных сетей должно приводить к изменению шаблона активности пространственно-распределённого «ансамбля нервных клеток». Таким образом, с самой первой модели, пространственная распределённость нейронной сети считается одной из важнейших характеристик. В связи с этим, пространственное расположение элементов ИНС должно быть включено в сущности ИНС и его элементов.

Синапсы Хебба обладают следующими свойствами [9-11], которые должны быть учтены при определении свойств сущности синапс:

- зависимы от времени;
 - локальны, т. к. нейрон реализует передачу данных между элементами находящимися в пространственно-временной близости;
 - интерактивны, т. е. влияют как на значение после синапса при прямом проходе, так и на значение до синапса при обратном проходе;
-

- реализуют коррелятивную связь значения до и после синапса.

На основании свойства локальности сделан вывод о главенствующей роли примитивов в ИНС, по сравнению с составными слоями современных моделей, реализуемыми с применением сложных алгоритмов.

В первую очередь разрабатываемый язык нужен для упрощения построения систем распознавания и классификации, которые строятся на многослойных персептронах, так как это наиболее популярная модель ИНС в современных информационных технологиях. Другой базовой сущностью являются синапсы, которые хранят в себе матрицу весов ИНС. Само обучение ИНС как раз и заключается в подборе значений этой матрицы весов в соответствии с правилом обучения.

В связи с вышесказанным далее рассмотрим изменения в алгоритмах работы ИНС и её математических моделях, с учётом необходимости минимизации количества сущностей для описания ИНС.

Разделение сущностей нейрона и синапса не позволяет использовать математическую формулу, обычно применяемую для описания суммирования входных сигналов нейрона:

$$\sum_{j=1}^N w_{ij} x_j + w_{i0} x_c, \text{ где}$$

w_0 — вес константного значения;

x_c — константное значение.

Более корректным является разделение представленной выше формулы на два этапа:

1. Вычисление весовых коэффициентов, включая значение константы $u_i = w_i x_i$, где w — вес синапса, x — входное значение синапса, u — результирующее значение синапса, i — номер синапса.

2. Суммирование полученных значений происходит после расчёта всех выходных значений синапсов:

$$u = \sum_{i=1}^N u_i, \text{ где } u \text{ — входной сигнал нейрона, } u_i \text{ — результирующее}$$

значение синапса, i — номер синапса, N — количество синапсов.

Часто константное значение и вес константного значения не разделяют, так как оба значения обычно не изменяются на протяжении работы ИНС. В этом случае используется одно значение называемое «порог». Здесь значение синапса константы меняется, т. к. это позволяет изменять значение сдвига суммы входных сигналов, при этом уменьшив количество сущностей.

Объединение математических моделей суммирования входных сигналов нейрона и функции активации не произведено, т. к. теоретически возможна замена как операции суммирования сигналов, например, умножением, так и функции активации.

Выделение синапса в отдельную сущность также сказывается на общей формуле многослойной ИНС.

$$y = f \left(\sum_{i=0}^n w_{2i} f \left(\sum_{i=0}^n w_{1i} x \right) \right)$$

Приведённая выше формула не позволяет увидеть различия между ИНС в виде полносвязного графа и неполносвязного.

В результате минимизации набора сущностей, был определён набор тегов, их свойств и свод правил их написания, названный ИНСМЛ (INSML) [12]. Сегодня актуальной является четвёртая версия языка ИНСМЛ.

С целью популяризации полученных результатов создан Интернет ресурс по адресу: abashin.ru/laboratory/xmlins/. Данный язык используется при обучении студентов. На текущий момент времени актуальной задачей для разработчиков языка является формирование первых аппаратно и программно независимых ИНС в виде описания на языке ИНСМЛ, для последующей эксплуатации в течение нескольких лет на разных платформах.

Ниже приведены два примера описания ИНС с помощью языка ИНСМЛ. В свойстве тега «сеть» указано что данная ИНС реализует функцию логическое И для диапазона входных значений [0;1].

```
<сеть ид=0 описание="Логическое И для диапазона [0;1]">
  <нейрон ид=12 константа=2 координаты=100;10;0/>
  <нейрон      ид=11      активация="1/(1+pow(e,-1*accum));0,5"
координаты=100;100;0/>
  <синапс ид=21 вес=1 приемник=11 передатчик=12/>
  <синапс ид=22 вес=-1,5 приемник=11 передатчик=""/>
  <синапс ид=23 вес=-1,5 приемник=11 передатчик=""/>
  <синапс ид=24 вес=1 приемник="" передатчик=11/>
</сеть>
```

Второй пример реализует функцию логическое И для диапазона [-1;1].

```
<сеть ид=1 описание="Логическое И для диапазона [-1;1]">
  <тип      ид=100      активация="if(accum<1)output=-
1;if(accum>=1)output=1"/>
  <нейрон ид=11 тип=100 координаты=100;100;0/>
  <синапс ид=21 вес=1 приемник=11 передатчик="">
  <синапс ид=22 вес=1 приемник="" передатчик=11
</сеть>
```

Описанный минимальный набор сущностей и их свойств позволяет представить прямой проход сигнала в большей части ИНС разработанных на сегодняшний день на языке ИНСМЛ. Таким образом решается проблема

передачи ИНС между группами разработчиков, тестеров, экспертов, обучающихся, преподавателей и других заинтересованных групп.

Упрощение передачи ИНС между заинтересованными группами, значительно упрощает интеграцию ИНС описанных на ИНСМЛ в контроллеры систем автоматизации и управления. В этом случае требуются усилия по переложению ИНСМЛ на языки программирования микроконтроллеров, а не повторная разработка и оптимизация ИНС для различных архитектур.

Литература

1. Царькова Е.Г. Математическая модель оптимального управления надежностью в охранных системах с нейросетевыми контроллерами // Инженерный вестник Дона, 2022, №4. URL: ivdon.ru/uploads/article/pdf/IVD_67__3_tsarkova.pdf_a231fe88d9.pdf.

2. Фирюлина М.А., Каширина И.Л. Описание процесса прогнозирования проблемных состояний с применением ансамблевых методов машинного обучения // Инженерный вестник Дона, 2022, №4. URL: ivdon.ru/uploads/article/pdf/IVD_104__3_Firyulina_Kashirina.pdf_f3b84b0bd0.pdf.

3. Евсина В.А., Широбокова С.Н., Жжонов В.А., Евсин В.А. Моделирование и реализация процесса распознавания знаков дорожного движения при определении ситуации на дороге с помощью искусственных нейронных сетей // Инженерный вестник Дона, 2022, №4. URL: ivdon.ru/uploads/article/pdf/IVD_95__3_Evsina.pdf_233a6e90a3.pdf.

4. Yejiang Yang, Tao Wang, Jefferson P.Woolard, Weiming Xiang. Guaranteed approximation error estimation of neural networks and model modification. / Neural Networks, Volume 151, July 2022, Pp. 61-69. doi.org/10.1016/j.neunet.2022.03.023.

5. Микеллучи У. Прикладное глубокое обучение. Подход к пониманию глубоких нейронных сетей на основе метода кейсов. СПб.: БХВПетербург, 2020. 368 с.
6. Song Yao. Method for optimizing an artificial neural network (ANN) (U.S. Patent No 10,621,486). U.S. Patent and Trademark Office. 2020. URL: patentimages.storage.googleapis.com/a3/a7/2a/cc5cbe75287ad1/US10621486.pdf
7. Lutich Andrey A. SRAF insertion with artificial neural network. (U.S. Patent No 10,198,550). U.S. Patent and Trademark Office. 2019. URL: patentimages.storage.googleapis.com/e9/a4/92/c31c85c13b47b0/US10198550.pdf
8. Hebb D.O. The Organization of Behavior: A Neuropsychological Theory, New York: Wiley, 1949, DOI: 10.1016/s0361-9230(99)00182-3.
9. Brown T.H., Kairiss E.W. and Keenan C.L. "Hebbian synapses: Biophysical mechanisms and algorithms", Annual Review of Neuroscience, 1990, vol. 13, p. 475-511.
10. Constantine-Paton M., Cline H.T. and Debski E. "Patterned activity, synaptic convergence and the NMDA receptor in developing visual pathways", Annual Review of Neuroscience, 1990, vol. 13, p. 129-154.
11. Fregnac Y. and Schulz D. "Models of synaptic plasticity and cellular analogs of learning in the developing and adult vertebrate visual cortex", Advances in Neural and Behavioral Development, 1994, vol. 4, p. 149-235.
12. Абашин В.Г. Язык разметки искусственных нейронных сетей INSML 0.3. // Монография. Орел: ОГУ имени И.С. Тургенева, 2019. 167 с.

References

1. Car'kova E.G. Inzhenernyj vestnik Dona, 2022, №4. URL: ivdon.ru/uploads/article/pdf/IVD_67__3_tsarkova.pdf_a231fe88d9.pdf.
-



2. Firyulina M.A., Kashirina I.L. Inzhenernyj vestnik Dona, 2022, №4. URL:ivdon.ru/uploads/article/pdf/IVD_104__3_Firyulina_Kashirina.pdf_f3b84b0bd0.pdf.
 3. Evsina V.A., Shirobokova S.N., Zhzhonov V.A., Evsin V.A. Inzhenernyj vestnik Dona, 2022, №4. URL:ivdon.ru/uploads/article/pdf/IVD_95__3_Evsina.pdf_233a6e90a3.pdf
 4. Yejiang Yang, Tao Wanga, Jefferson P.Woolard, Weiming Xiang. Neural Networks, Volume 151, July 2022, Pages 61-69 doi.org/10.1016/j.neunet.2022.03.023.
 5. Mikelluchi, U. Prikladnoe glubokoe obuchenie. Podhod k ponimaniyu glubokih nejronnyh setej na osnove metoda kejsov [Applied deep learning. An approach to understanding deep neural networks based on the case method]. Spb.: BHV-Peterburg, 2020. 368p.
 6. Song Yao. Method for optimizing an artificial neural network (ANN) (U.S. Patent No 10,621,486). U.S. Patent and Trademark Office. 2020. URL:patentimages.storage.googleapis.com/a3/a7/2a/cc5cbe75287ad1/US10621486.pdf
 7. Lutich Andrey A. SRAF insertion with artificial neural network. (U.S. Patent No 10,198,550). U.S. Patent and Trademark Office. 2019. URL:patentimages.storage.googleapis.com/e9/a4/92/c31c85c13b47b0/US10198550.pdf
 8. Hebb D.O. The Organization of Behavior: A Neuropsychological Theory, New York: Wiley, 1949, DOI: 10.1016/s0361-9230(99)00182-3.
 9. Brown T.H., Kairiss E.W. and Keenan C.L. Annual Review of Neuroscience, 1990, vol. 13, pp. 475-511.
 10. Constantine-Paton M., Cline H.T. and Debski E. Annual Review of Neuroscience, 1990, vol. 13, pp. 129-154.
 11. Fregnac Y. and Schulz D. Advances in Neural and Behavioral Development, 1994, vol. 4, pp. 149-235.
-



12. Abashin V.G. YAzyk razmetki iskusstvennyh nejronnyh setej INSML 0.3. [Artificial Neural Network Markup Language INSML 0.3]. Monografiya. Orel: OGU imeni I.S. Turgeneva, 2019. 167 p.