Научная статья
УДК 004.8.032.26+519.7
ББК 32.813+32.818.1
Ч 54
DOI: 10.53598 / 2410-3225-2021-4-291-70-74

Обучение модульной нейронной сети для многозадачного искусственного интеллекта

(*Рецензирована*)

Петр Васильевич Четырбок[1], Марина Анатольевна Шостак[2]

[1, 2] *Гуманитарно-педагогическая академия (филиал) ФГАОУ ВО «Крымский федеральный университет им. В.И. Вернадского» в г. Ялте, Ялта, Россия*
[1] *petr58@mail.ru*
[2] *shostakma@inbox.ru*

Аннотация. *Рассматривается формализация алгоритмов для многозадачного искусственного интеллекта с использованием векторного критерия. Показаны проблемы, возникающие при использовании составления алгоритмов обучения для многозадачного искусственного интеллекта. Представлены методы разрешения проблем, включающие построения самообучающихся нейронных модулей, применения специальных типов нейронных сетей, основанных на векторном критерии и формализации алгоритмов для обработки больших данных с использованием векторного критерия.*

Ключевые слова: *информация, технология, алгоритм, интеллектуальная обработка больших данных, модульная нейронная сеть, векторный критерий, векторное пространство, категория*

Original Research Paper

Modular neural network training for multitasking artificial intelligence

Petr V. Chetyrbok[1], Marina A. Shostak[2]

[1, 2] *Humanities-Pedagogical Academy (Branch), V.I. Vernadsky Crimean Federal University in Yalta, Yalta, Russia*
[1] *petr58@mail.ru*
[2] *shostakma@inbox.ru*

Abstract. *The paper considers the formalization of algorithms for multitasking artificial intelligence using a vector criterion. The work shows problems that arise when compiling training algorithms for multitasking artificial intelligence. Methods of solving problems are presented, including constructing self-learning neural modules, using special types of neural networks based on a vector criterion and formalizing algorithms for processing large data using a vector criterion.*

Keywords: *information, technology, algorithm, intelligent big data processing, modular neural network, vector criterion, vector space, category*

Постановка проблемы. К области нейротехнологии относится направление, связанное с обработкой информации на основе принципов функционирования естественных нейронных систем. Математическое моделирование таких систем привело к созданию искусственных нейронных сетей. В биологических сетях принципиальное значение имеет структура связей между нервными клетками, называемыми нейронами.

Обособленные группы нейронов в нейробиологии называют нейронными модулями. Они взаимодействуют между собой только через внешние рецепторные и аксоновые поля. Через скрытые нейроны модулей идет обработка информации, которая поступает на рецепторы, и формируется реакция в выходном аксоновом поле модуля. Если предположить, что нейронные модули в сети не имеют пересечений рецепторных или аксоновых полей, то получим модель модульной нейронной сети с инъективными связями. Инъективность связей позволяет представить структуру нейронной сети в виде ориентированного графа. Погружение структурной модели в пространство рецепторных и аксоновых полей порождает топологическую модель нейронной сети, по которой можно реализовать адаптивный алгоритм обработки данных [1–3]. Особенно такая обработка актуальна для многозадачного искусственного интеллекта. Она позволяет обучить модули нейронной сети на различных обучающих выборках для решения различных задач, не меняя структуры модульной нейронной сети.

Цель статьи. Рассматриваются возможности модульной нейронной сети, топологическая модель которой построена с использованием не только пространств рецепторных и аксоновых полей, но и пространства ошибок, полученного с помощью векторного критерия, для многозадачного интеллекта.

Основная часть. Векторный критерий позволяет обучить модульную нейронную сеть для решения различных задач, включив в нее модули, обученные на различных обучающих выборках.

В конкретной реализации нейронной сети с каждым нейронным модулем связаны два линейных векторных пространства:

– пространство рецепторов;

– пространство аксонов.

Добавим к ним еще пространство ошибок. Пространство ошибок содержит векторы ошибок, полученные при распознавании объектов нейронным модулем.

Для нейронного модуля размерность пространства рецепторов равна числу рецепторов, а размерность пространства аксонов равна числу аксонов. То есть нейронный модуль является оператором, который преобразует вектор из пространства рецепторов в пространство аксонов. Оператор можно задать, вводя базисы в векторном пространстве.

Возьмем в качестве базиса – типичные представители классов объектов из обучающей выборки. Типичный представитель класса объектов выбирается из обучающей выборки среди эталонных образцов объектов этого класса по векторному критерию [4]. А именно, выбирается тот эталон, на котором векторный критерий максимален. Векторный критерий позволяет в обучающей выборке среди эталонных образцов выбрать опорные векторы для машины опорных векторов. Также сможем добавить еще один базис, а именно, векторы, полученные на ошибках распознавания опорных векторов в пространстве ошибок. Исходя из этого базиса, сможем вычислить векторный критерий в пространстве ошибок при распознавании нового объекта, и тем самым, используя обучающую выборку для данного модуля, обучить его решать конкретную задачу, поставленную перед этим модулем.

Современные нейронные сети (и другие алгоритмы машинного обучения) обычно решают одну проблему из одного примера – это может быть классификация, регрессия, структурированное прогнозирование, обнаружение аномалий или генерация объекта. Рассмотрим, как можно обучать модели с помощью многозадачного обучения.

Регуляризация L2 для функции потерь (функция потери нормы L2 также известна как ошибка наименьших квадратов (МНК)) [5]:

$$\min_{w} \boxed{\sum_{i=1}^{n} V(\hat{x}_i \cdot w, \hat{y}_i)} + \boxed{\lambda \|w\|_2^2},$$

Main optimization criteria

где V – функция потерь;

x_i – входные параметры;

y_i – выходные параметры;

w – вес;

λ – гиперпараметр.

L2-регуляризация содействует появлению малых весовых коэффициентов модели, но не способствует их точному равенству нулю. С помощью L2 решаем одновременно две задачи оптимизации: минимизация функции потерь и нормирование параметров.

Одним из эффектов регуляризации является решение обратных некорректных задач и добавление регуляризации по норме параметров, которая помогает добиться гладкости и хорошего кондиционирования задачи.

Если такая дополнительная потеря, как регуляризация параметров алгоритма, помогает повысить производительность, тогда алгоритмы будут выглядеть примерно так, как на рисунке 1: один вход, несколько слоев (в случае нейронных сетей) и множество выходов, которые могут играть роль регуляризатора с некоторым гиперпараметром *лямбда* или независимая цель, которая также должна быть оптимизированной и решенной [6].

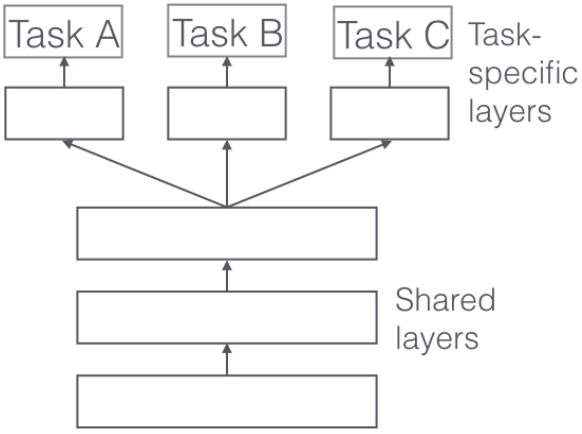

Рис. 1. Жесткий обмен параметрами для многозадачного обучения в глубоких нейронных сетях

Fig. 1. Hard exchange of multitasking settings in deep neural networks

Все эти «знания», «условия» включаются в алгоритм машинного обучения в качестве дополнительного входа. И если использовать модульную нейронную сеть, то алгоритм многозадачного обучения можно распараллелить. Правило разбиения на признаковые поля: если построенный на признаковом поле модуль нейронной сети на входных данных не изменяет векторный критерий распознавания образов, то это поле можно безболезненно, а также и модуль нейронной сети исключить из структуры модульной нейронной сети, тем самым информацию на данном признаковом поле необходимо отсеять из дальнейшего рассмотрения [7].

Для построения оптимальной модульной нейронной сети применим правило, описанное выше, и получим структуру сети, показанную на авторском рисунке 2.

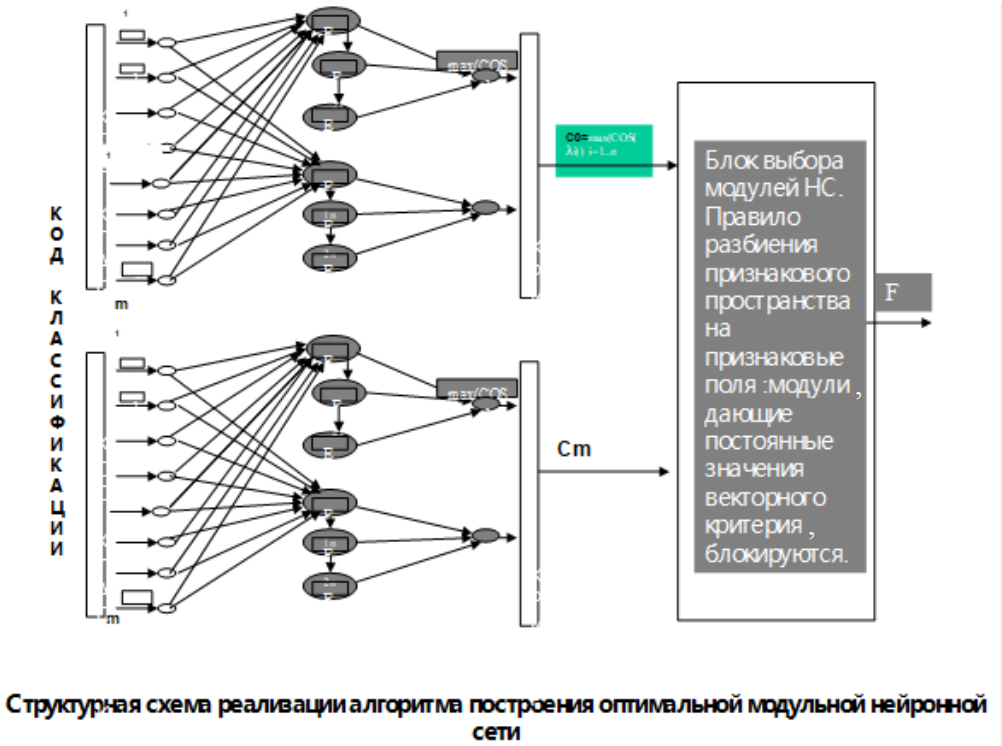

Рис. 2. Структура модульной нейронной сети
Fig. 2. Modular neural network structure

C_0=max(COS(λi)), i=1,…,n – векторный критерий модуля 0;
C_m – векторный критерий модуля m.

А именно модульная нейронная сеть с векторным критерием позволяет обучить модули нейронной сети на различных обучающих выборках для решения различных задач, не меняя структуры модульной нейронной сети.

Обучение нейронной сети для одновременного решения X задач вместо использования X отдельных нейронных сетей влияет на производительность, регуляризацию и обобщение.

Выводы. Модульная нейронная сеть с векторным критерием позволяет обучить модули нейронной сети на различных обучающих выборках для решения различных задач, не меняя структуры модульной нейронной сети. Модульная нейронная сеть сокращает время обучения, так как в традиционном варианте пришлось бы для решения каждой задачи строить архитектуру отдельной нейронной сети и обучать ее самостоятельно на обучающей выборке. После обучения модульной нейронной сети для решения нескольких задач возникает возможность ее самообучения для решения задач кластеризации.

Примечания

1. Дорогов А.Ю. Теория и проектирование быстрых перестраиваемых преобразований и слабосвязанных нейронных сетей. Санкт-Петербург: Политехника, 2014. 328 с.

2. Саймон Х. Нейронные сети: полный курс: пер. с англ. 2-е изд. Москва: Вильямс, 2006. 1104 с.

3. Четырбок П.В., Шостак М.А. Подготовка специалистов индустрии гостеприимства Южного региона России в условиях пандемии Covid-19 с использованием информационных технологий // Вестник Адыгейского государственного университета. Сер.: Естественно-

математические и технические науки. 2020. Вып. 3 (266). С. 87–91. URL: http://vestnik.adygnet.ru

4. Четырбок П.В. Построение решающего правила для классификации образов на основе векторов ошибок // Системні дослідження та інформаційні технології. Киев, 2013. № 2. С. 114–120.

5. Тихонов А.Н. О некорректных задачах линейной алгебры и устойчивом методе их решения // ДАН СССР. 1965. Т. 163, № 3. С. 591–594.

6. Рудер С. Обзор многозадачного обучения в глубоких нейронных сетях // arXiv: 1706.05098v1. 2017. 14 с. URL: https://arxiv.org/pdf/1706.05098.pdf

7. Chetyrbok P.V. Preliminary systemic decomposition of big data for their classification using vector criteria dynamic management model of innovations generations: 2017 20th IEEE International Conference on Soft Computing and Measurements (SCM). 2017. P. 762–764.

References

1. Dorogov A.Yu. Theory and design of fast tunable transformations and weakly connected neural networks. Saint Petersburg: Polytechnic, 2014. 328 p.

2. Simon H. Neural Networks: the complete course: transl. from English. 2nd ed. Moscow: Williams, 2006. 1104 p.

3. Chetyrbok P.V., Shostak M.A. Hospitality industry training in South Region (Russia) in conditions of Covid-19 pandemic using information technology // The Bulletin of the Adyghe State University. Ser.: Natural-Mathematical and Technical Sciences. 2020. Iss. 3 (266). P. 87–91. URL: http://vestnik.adygnet.ru

4. Chetyrbok P.V. Building a decision rule for the classification of patterns based on error vectors // Systematic advances and information technologies. Kiev, 2013. No. 2. P. 114–120.

5. Tikhonov A.N. On ill-posed problems in linear algebra and a stable method for their solution // DAN USSR. 1965. Vol. 163, No. 3. P. 591–594.

6. Ruder S. An Overview of Multi-Task Learning in Deep Neural Networks // arXiv:1706.05098v1. 2017. 14 p. URL: https://arxiv.org/pdf/1706.05098.pdf

7. Chetyrbok P.V. Preliminary systemic decomposition of big data for their classification using vector criteria dynamic management model of innovations generations: 2017 20th IEEE International Conference on Soft Computing and Measurements (SCM). 2017. P. 762–764.

Авторы заявляют об отсутствии конфликта интересов.

The authors declare no conflicts of interests.

Статья поступила в редакцию 5.11.2021; одобрена после рецензирования 2.12.2021; принята к публикации 3.12.2021.

The article was submitted 5.11.2021; approved after reviewing 2.12.2021; accepted for publication 3.12.2021.

© П.В. Четырбок, М.А. Шостак, 2021