

ТЕХНОЛОГИИ ИСКУССТВЕННОГО ИНТЕЛЛЕКТА В МЕНЕДЖМЕНТЕ

ОБЗОР НЕКОТОРЫХ СОВРЕМЕННЫХ ТЕНДЕНЦИЙ В ТЕХНОЛОГИИ МАШИННОГО ОБУЧЕНИЯ

Получено: 17.09.2018; одобрено: 24.09.2018; опубликовано: 26.11.2018

УДК 004.89 JEL C45 DOI 10.26425/2658-3445-2018-1-26-35

Коротеев Михаил Викторович

Канд. экон. наук, доцент, ФГБОУ ВО «Финансовый университет при Правительстве Российской Федерации», г. Москва, Россия
Старший научный сотрудник, ФГБУН «Институт проблем управления им. В.А. Трапезникова» Российской академии наук,
г. Москва, Россия

e-mail: mvkoroteev@fa.ru

АННОТАЦИЯ

Построение систем машинного обучения является на сегодняшний день одной из самых популярных, актуальных и современных областей человеческой деятельности на стыке информационных технологий, математического анализа и статистики. Машинное обучение все глубже проникает в нашу жизнь посредством пользовательских продуктов, созданных с помощью методов искусственного интеллекта. Очевидно, что данные технологии будут развиваться и дальше, постепенно становясь частью повседневной рутины во многих областях человеческой профессиональной деятельности. Однако со времен своего появления, машинное обучение успело обзавестись многочисленными проблемами, главная из которых – достаточно высокая трудоемкость. Построение систем машинного обучения требует огромного количества времени высокопрофессиональных специалистов как в сфере искусственного интеллекта, так и в той предметной области, к которой эта технология применяется.

В статье рассмотрены основные новации в области методологии машинного обучения, которые могут оказать значительное влияние на развитие данной отрасли. Выполнен анализ современной научной литературы, посвященной вопросам развития методологии и областей прикладного использования рассматриваемых тем. Сформулированы предположения о будущих тенденциях развития машинного обучения как сферы научно-прикладного знания и предложены наиболее перспективные направления исследований. Рассмотрены такие современные технологии в машинном обучении, как использование предобученных моделей, построение мультизадачных систем, нейроэволюция, проблема создания интерпретируемых моделей. Наиболее перспективной и актуальной в настоящее время технологией авторы полагают автоматизированное машинное обучение – комплекс инструментальных и методических средств, позволяющий значительно сократить долю человеческого участия в создании систем искусственного интеллекта, в том числе средствами автоматической валидации результатов моделирования.

КЛЮЧЕВЫЕ СЛОВА

Машинное обучение, интерпретируемость, мультизадачные модели, перенос обучения, нейроэволюция, автоматизированное обучение.



ARTIFICIAL INTELLIGENCE TECHNOLOGIES IN MANAGEMENT

REVIEW OF SOME CONTEMPORARY TRENDS IN MACHINE LEARNING TECHNOLOGY

Received: 17.09.2018; approved: 24.09.2018; published: 26.11.2018

JEL CLASSIFICATION C45 DOI 10.26425/2658-3445-2018-1-26-35

Koroteev Mihail

Candidate of Economic Sciences, Associate Professor, Financial University Under the Government of the Russian Federation, Moscow, Russia

Senior Researcher, Institute of Management Problems Named After V.A. Trapeznikov of the Russian Academy of Sciences, Moscow, Russia

e-mail: mvkoroteev@fa.ru

ABSTRACT

The construction of machine learning systems constitutes today one of the most popular, relevant and modern areas of human activity at the interface of information technology, mathematical analysis and statistics. Machine learning penetrates deeper into our lives through custom products created with the assistance of artificial intelligence methods. Obviously, that these technologies will develop further, gradually becoming a part of everyday routine in many areas of human professional activity. However, since its occurrence, machine learning has managed to acquire numerous problems, the main of which, according to authors, is a rather high labor intensity. The construction of machine learning systems requires a huge amount of time of highly professional specialists both in the field of artificial intelligence and in the subject area to which this technology is applied.

In this article we reviewed the main innovations in the field of machine learning methodology, which can influence significantly on the development of this industry. Also an analysis of modern scientific literature devoted to the development of methodology and areas of applied employment of the issues, we are considering, has been carried out. In addition, assumptions were formulated about future trends in the development of machine learning as a field of scientific and applied knowledge and suggested the most promising areas of research. Such modern technologies in machine learning as the use of pre-trained models, the construction of multitasking systems, neuroevolution, the problem of creating interpreted models were considered. The authors believe that the most promising and relevant technology at the moment is automated machine learning, a complex of instrumental and methodological tools that allows to significantly reduce the share of human participation in the creation of artificial intelligence systems, including the means for automatic validation of simulation results.

KEYWORDS

Machine learning, interpretability, multitask models learning, transfer learning, neuroevolution, automated machine learning.

CITATION

Koroteev M.V. (2018). Review of some contemporary trends in machine learning technology. *E-Management*, vol. 1, № 1, pp. 26–35. DOI: 10.26425/2658-3445-2018-1-26-35



В статье рассмотрены наиболее интересные тенденции в машинном обучении и искусственном интеллекте, сформировавшиеся на начало 2018 г. за пределами конкретных математических методов оптимизации, обработки и анализа данных. Все большее внимание исследователей занимает вопрос методологий, или метамоделей (от англ. *metamodel*): принципов использования, комбинирования и выбора конкретных моделей и методов машинного обучения. Многолетний прогресс в разработке методов машинного обучения породил не только разнообразные математические, программные и даже аппаратные решения, предназначенные для задач предиктивного и генеративного анализа данных в самых разных областях, но встретил на своем пути немало трудностей и препятствий.

Основная трудность, с которой сталкивается человек в процессе знакомства с областью машинного обучения, – огромное количество разрозненных методов, каждый из которых обладает своими особенностями, областью использования и преимуществами. Однако такое разнообразие подчас ставит в тупик и искушенных исследователей. С развитием математических и алгоритмических методов становится все труднее хорошо ориентироваться во всех нюансах применяемых алгоритмов. К сожалению, методологическая база значительно отстает от быстрого процесса разработки новых алгоритмов обучения, и процесс выбора обучаемой модели подчас сводится к простому перебору.

В индустрии машинного обучения давно назрела необходимость создания более простых в использовании подходов, которые можно применять в широком кругу неэкспертов. Часто процесс использования систем машинного обучения предполагает выполнение более или менее полного цикла прикладных исследовательских работ по обработке (препроцессингу) данных, выделения признаков, выбора вида модели, обучения параметров, гиперпараметров, и т. д.

Цель статьи – обзор актуальных приемов, технологий и методик, применяемых при решении прикладных задач машинного обучения, по материалам научных статей в высокорейтинговых журналах зарубежных исследователей, аналитических и обзорных заметок из открытых источников а также технической документации и пресс-релизов технических и программных решений.

ПРОБЛЕМА ИНТЕРПРЕТИРУЕМОСТИ

Несмотря на бурное развитие машинного обучения в последнее десятилетие, искусственный интеллект остается весьма расплывчатым понятием. Оно включает множество предметных областей: от предсказания временных рядов до генерации правдоподобных изображений по определенной теме. Методы машинного обучения, составляющие вычислительную основу технологий искусственного интеллекта, все еще остаются узкоспециализированными под каждую конкретную задачу.

В качестве математической и инструментальной основы машинного обучения сегодня центральное место занимают искусственные нейронные сети. Сейчас они стали своеобразным универсальным языком представления обучаемых моделей. И хотя искусственный интеллект и машинное обучение как области знаний гораздо шире и включают целые семейства других методов, нейросетевые модели в настоящий момент так или иначе фигурируют в 90 % научных публикаций в данных областях. «Руководитель направления «Поиск Mail.ru» в Mail.Ru Group Андрей Калинин отмечает, что нейронные сети способны решать такие же задачи, как и другие алгоритмы машинного обучения, разница заключается лишь в подходе к обучению» [Хохлова, 2016].

С разработкой все более сложных и глубоких архитектур нейронных сетей вместе с несомненными достоинствами все явственнее проявляются общие проблемы этого подхода. Наиболее существенной из них можно отметить принцип работы любой нейросети как «черного ящика» – даже должным образом обученная сеть не дает исследователям информации о внутренней структуре проблемы и выявленных зависимостях в данных. Обученная нейросеть является набором матриц весов, и смысловая интерпретация этих весов в общем случае не предполагается. С этой точки зрения нейросети представляют собой лишь инструмент решения конкретной задачи машинного обучения, но не дают экспертам аналитической информации для исследования проблемы.

Этот недостаток заставляет исследователей заниматься проблемами интерпретируемости нейросетевых моделей. В частности, довольно широкую известность приобрела опубликованная в 2017 г. работа о применении довольно старого метода интерпретации предсказаний к обученным

нейронным сетям через использование функций влияния [Koch, 2017]. В начале 2018 г. появилась книга, посвященная интерпретации моделей «черного ящика» [Molnar, 2018].

Интерпретируемость модели машинного обучения понимается как «степень, с которой человек может понять причины того или иного решения». Построение интерпретируемых моделей позволяет находить ответ на такие вопросы как: «Как алгоритм создает модель?», «Как обученная модель делает предсказания?», «Как составные части модели влияют на предсказание?», «Как модель принимает определенное решение для определенного объекта/группы объектов?».

В настоящее время выделяют класс так называемых «естественно интерпретируемых моделей» – моделей, человеческая интерпретация которых не представляет труда в силу самой архитектуры модели. К ним относят, в частности: линейные и логистические регрессии, деревья решений, наивные байесовские классификаторы, k ближайших соседей, модели правил вывода и другие [Molnar, 2018]. Объяснение внутреннего устройства других типов моделей, в том числе глубоких нейронных сетей, представляет собой открытую научную задачу.

Существуют следующие основные методы для построения моделенезависимых объяснительных механизмов (англ. model-agnostic methods):

- функции и графики частичных зависимостей, которые показывают предельное влияние признаков на результирующее значение;
- механизм индивидуального условного ожидания – расширение метода функций частичных зависимостей на локальный анализ отдельных предсказаний;
- исследование уверенности модели (англ. model class reliance) – моделенезависимый вариант алгоритма важности признаков, который позволяет оценить зависимость средней ошибки модели в зависимости от дисперсии отдельных признаков;
- метод LIME (от англ. local interpretable model-agnostic explanations) основан на обучении локальных интерпретируемых моделей по предсказаниям модели, исследуемой для объяснения конкретных предсказаний.

В целом, область исследований интерпретируемости моделей машинного обучения можно назвать очень актуальной. Несомненно, данная проблематика получит свое развитие в дальнейших исследованиях в ближайшее время.

ПРЕДОБУЧЕННЫЕ МОДЕЛИ

Использование нейронных сетей неразрывно связано с процессом их обучения. Даже самая продвинутая и сложная глубокая нейронная сеть, будучи необученной на правильно выбранном наборе данных, не может решить даже самых простых задач анализа данных. Эффективность работы обученных нейросетей напрямую зависит от объема и качества обучающей выборки и времени обучения. А время обучения, в свою очередь, зависит от вариативности модели. Чем более вариативная, глубокая и сложная модель используется, тем больше времени на обучение она требует и тем больший объем данных нужен для предотвращения переобучения.

Таким образом, перед исследователями в области машинного обучения стоит не только задача поиска качественной обучающей выборки и очистки данных, но и выбора архитектуры сети (процесс, больше напоминающей искусство, чем науку) и проведения довольно затратного в плане процессорного времени обучения сети. Но в результате обученная сеть представляет собой всего лишь набор матриц весов и описание архитектуры сети. Процессорное время является узким местом в процессе анализа данных. В противоположность этому объем оперативной памяти не так дорогостоящ. Это приводит к вполне естественному предположению о возможности повторного использования обученных нейросетей.

Широкое распространение предобученных моделей неразрывно связано с развитием инструментальных средств моделирования нейронных сетей. Использование обученных нейронных сетей получило в научной и технической литературе название «перенос обучения» (англ. transfer learning).

Существует несколько способов использования предобученных моделей в зависимости от размера доступной исследователям выборки и схожести выборок используемых для предобучения и непосредственно для анализа.

1. Автоматизированное создание признаков. Предобученные модели могут использоваться для автоматического извлечения значимых признаков из анализируемой выборки данных для последующего анализа «своими силами».

2. Использование только архитектуры. Предобученные модели, показавшие свою эффективность на определенном модельном наборе данных, могут послужить образцом архитектуры для построения новой модели и обучения ее на анализируемом наборе. В таком случае не происходит экономии процессорного времени, а сокращаются издержки проектирования неэффективных моделей.

3. Частичное дообучение. Также показал свою эффективность подход, при котором предобученные модели используют частично, а обучение происходит для части слоев такой сети. Обычно веса начальных слоев замораживают, а конечных обучают на непосредственно анализируемой выборке. Также количество дообучаемых слоев может настраиваться путем перекрестной проверки.

4. ANN-НММ (от англ. artificial neural networks – hidden Markov models) – использование глубоких сетей доверия для начальной инициализации весов нейронной сети с последующим дообучением методом обратного распространения ошибки.

Последний метод неоднократно протестирован и доказал свою эффективность в задачах распознавания речи.

Использование предобученных нейронных сетей, несомненно, является актуальным направлением развития технологии машинного обучения. Однако в настоящее время широкое распространение переноса обучения затруднено, в частности, отсутствием общепринятого формата хранения и распространения моделей нейронных сетей. Сейчас формат модели остается жестко привязанным к конкретному инструментальному средству реализации вычислений. Нам представляется, что в будущем разработка такого формата явится актуальным, естественным и неизбежным следствием развития и широкого распространения библиотек глубокого обучения.

МУЛЬТИЗАДАЧНЫЕ СЕТИ

Одной из характеристик классических систем машинного обучения является специфичность модели для определенной задачи. Если строится нейронная сеть для распознавания рукописных цифр, ее нельзя применять для распознавания лиц, синтеза музыки или предсказания трендов на фондовом рынке. Более того, различные архитектуры нейронных сетей имеют свою специфику в отношении предметной области: рекуррентные сети хороши для моделирования последовательностей, сверточные сети – для распознавания изображений, порождающие сети – для генерации правдоподобных изображений. Естественным является вопрос о построении универсальной нейронной сети, которую можно было бы с минимальным переобучением использовать для решения многих различных задач.

Исследования в этой области велись с конца 1990-х гг. Однако широкое распространение данная тема получила недавно. В литературе устоялось название «мультизадачное обучение» (от англ. multitask learning, MTL), хотя иногда употребляется термин «гетерогенное обучение» [Misra и др., 2016]. Исследования по данной теме начинались в университете Карнеги-Меллона [Caruana, 1998] и университете Монреаля [Ghoshn, Bengio, 1997]. В.R. Paredes [2012] предлагает разделение задач на основные и дополнительные, где обучение дополнительным задачам повышает эффективность решения основных. Показан значительный рост эффективности. Здесь же дается математическое доказательство того, что при использовании специальных методов регуляризации, проблема мультизадачной оптимизации становится выпуклой, что значительно облегчает обучение. S. Li [2014] с соавторами получили рекордные результаты в задачах распознавания поз людей по фотографиям через отдельное обучение моделей распознавания частей тела. J. Dai и K. He [2014] исследовали семантическую сегментацию изображений. Результаты в идентификации лиц также улучшаются в условиях многозадачности. М.-Т. Luong [2015] с соавторами также установили рекордные показатели эффективности в области машинного перевода с использованием разделяемых энкодеров и декодеров. D. Dong [2015] с соавторами исследовали преимущество в точности при одновременном обучении машинному переводу между многими разными парами языков. Показано, что при наличии достаточно обширного корпуса, эффективность перевода повышается за счет мультизадачности.

Примерно в это же время исследователи начали систематизировать накопленные знания в области мультизадачного обучения. Сейчас выделяют два принципиальных подхода к построению мультизадачных моделей: жесткое и мягкое разделение параметров.

В 2017 г. произошел прорыв в области многозадачного обучения. Команда Google Brain выпустила работу [Kaiser и др., 2017], в которой представила универсальную многокомпонентную архитектуру нейронной сети, названную MultiModel, которая достигает высоких результатов в одновременном обучении задачам из совершенно разных областей: обработка естественного языка, машинный перевод (несколько языков одновременно), распознавание лиц на изображении, распознавание речи. Важным методологическим результатом стало то, что такая сеть способна улучшать результаты одной задачи с обучением другой, совершенно, на первый взгляд, с ней не связанной.

По мнению автора, область мультизадачного обучения в ближайшие годы получит дальнейшее развитие. Представляется весьма интересным анализ мультизадачных сетей методами интерпретации черного ящика, развивающихся сегодня параллельно. Исследователи мультизадачного обучения Y. Zhang и Q. Yang [2017] выделяют следующие проблемы описанного выше подхода: преимущественный фокус на задачах обучения с учителем, необходимость анализа полезности различных задач в ансамбле, выявление задач-вредителей, необходимость более глубокого теоретического анализа полученных результатов. Несомненно, эти проблемы станут темами будущих исследований в самое ближайшее время.

НЕЙРОЭВОЛЮЦИЯ

Искусственные нейронные сети доказали свою эффективность в решении многих задач, однако эта эффективность в большой мере зависит от умения подбирать правильную архитектуру сети под каждую конкретную задачу анализа данных.

При проектировании архитектуры сети эксперт сталкивается с необходимостью принять множество решений, как количественных, так и качественных, от реализации которых напрямую зависит производительность результирующей модели. Среди них такие как: выбор количества слов, количества нейронов в каждом слое, выбор функции активации, использование рекуррентных и сверточных слоев. В процессе такого проектирования задача эксперта – балансирование между вариативностью модели и склонностью к переобучению. Существуют еще и внешние факторы, которые необходимо брать в расчет: располагаемая вычислительная мощность, временные рамки решения задачи и т. д.

Пространство возможных нейронных сетей, в котором ведется поиск, огромно. При проектировании нейронных сетей специалисты используют определенные эвристические правила и инструменты диагностики, однако полноценной методологией такой поиск назвать сложно, это скорее творческий процесс.

Естественным направлением исследований в области искусственного интеллекта является построение методов и инструментов сокращения человеческого участия в построении систем машинного обучения, автоматизация данного процесса. В настоящее время существует два принципиально отличающихся подхода: автоматизация машинного обучения (AutoML) и нейроэволюция (от англ. evolutionary artificial neural networks, EANN).

Фундаментальной работой в области нейроэволюции стало исследование Яо [Yao, 1999], в которой исследована связь между процессами обучения искусственных нейронных сетей и эволюционными алгоритмами. Работа Яо дала старт исследованиям применения эволюционного программирования на различных этапах проектирования и обучения нейронных сетей. Дальнейшие исследования привели к созданию метода NEAT (от англ. neuroevolution through augmenting topologies), более оптимизированного по вычислительным ресурсам. Совсем недавно этот метод был адаптирован для эволюционного поиска структур глубоких нейронных сетей.

На сегодняшний день эволюционное программирование дает возможность создавать сети, по производительности сравнимые с лучшими в своем классе при нулевом участии человека в процессе проектирования и обучения для задач классификации изображений и задач распознавания речи [Yao, 1999].

В настоящее время ведутся активные работы с использованием нейроэволюции в таких областях как: предсказание временных рядов ансамблевыми моделями, предсказание энергопотребления компьютерными кластерами, построение межпланетных траекторий, распознавание языка речи, предсказание цен на нефть.

Главным недостатком нейроэволюционного подхода являются весьма высокие требования к вычислительной мощности для поддержки данного процесса до сходимости. Развитие глубоких нейронных сетей и распространение больших данных еще выше поднимает эту планку. Поэтому, в настоящее время мы наблюдаем спад интереса к эволюционному программированию искусственных нейронных сетей: время программиста и специалиста в области анализа данных сейчас стоит дешевле, нежели требуемое процессорное время.

Однако большие корпорации, располагающие высокими вычислительными мощностями, могут позволить себе эксперименты с нейроэволюцией. По нашему мнению, с ростом производительности вычислительной техники интерес к эволюционным методам в машинном обучении будет возрастать (возможно итерационно), пусть не в ближайший год, но в обозримом будущем.

АВТОМАТИЗИРОВАННОЕ ОБУЧЕНИЕ

На протяжении последнего времени (примерно с 2013 г.) автоматизация машинного обучения стала широко обсуждаемой темой и одной из наиболее быстрорастущих областей теоретических и практических разработок. Р. Олсон, один из разработчиков библиотеки автоматизированного машинного обучения, выделяет три главных процесса, автоматизация которых возможна и высвобождает большое количество временных ресурсов экспертов:

- подбор гиперпараметров моделей;
- испытание большого количества разных моделей;
- использование разных признаков, выделенных из данных [Olson, 2016].

Таким образом, автоматизированное машинное обучение (от англ. automated machine learning, AutoML) можно охарактеризовать как набор технологий и методов алгоритмического выбора, оценки эффективности моделей машинного обучения и итеративного моделирования.

Сложность традиционного подхода к построению систем машинного обучения состоит в необходимости знания всех существующих алгоритмов искусственного интеллекта, умения их правильно применить и настроить.

Исторически первым механизмом автоматизации процесса машинного обучения является основанный на байесовской оптимизации метод AutoWEKA [Thornton C. и др., 2013].

Предлагаемые методы автоматизации машинного обучения связаны с существующими инструментальными средствами моделирования. Большинство из них ориентируется на популярную библиотеку scikit-learn языка Python. Например, довольно многообещающая система Auto-Sklearn, развивающая идеи байесовской оптимизации. Эта библиотека использует 15 классификаторов, 14 методов обработки признаков, 4 метода препроцессинга, позволяя вести автоматизированный поиск в пространстве моделей с более сотни гиперпараметров.

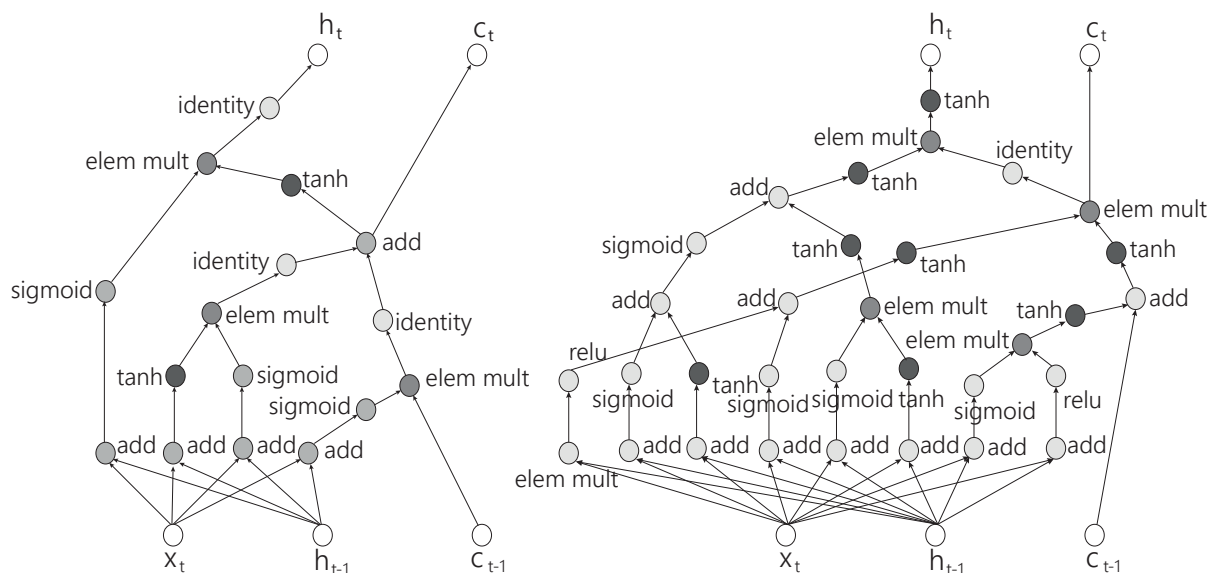
Однако байесовская оптимизация не единственная методологическая основа автоматизации машинного обучения. Существуют решения, основанные на генетическом подходе [Olson, Moore, 2016]. Такие системы способны строить сложные процессы машинного обучения без какого-либо вмешательства человека в процесс проектирования.

Современные рыночные системы AutoML, построенные в виде облачного сервиса могут строить модели, превосходящие созданные человеком в 30 % случаев. Однако экономия временных ресурсов колоссальна – до двух порядков.

Активным направлением исследований в области автоматизации машинного обучения является проработка вопроса использования сложных нелинейных конвейеров обработки данных. Преимуществом методом нахождения таких конвейеров является генетическое программирование.

В середине прошлого года исследовательский отдел Google представил архитектуру AutoML, основанную на обучении с подкреплением [Zoph, Le, 2016]. Эта система строит рекуррентные сети, схожие по своей архитектуре с построенными человеком, но более сложные.

В конце 2017 г. усовершенствованная система AutoML смогла построить сеть, превосходящую все существующие в задачах распознавания объектов на изображениях (см. рис).



Источник: [Le, Zoph, 2017]

Рис. Принципиальная схема одного элемента нейронной сети для распознавания изображений, созданная человеком (слева) и с помощью системы Google AutoML (справа)

Конечно, инструменты автоматизации машинного обучения не способны полностью вытеснить человека как участника процесса обработки данных. В конце концов, функцию целеполагания никакой инструмент обеспечить не в состоянии. Также одни из барьеров перед использованием систем AutoML – повышенные требования к вычислительным ресурсам. Однако с учетом экономии времени проектирования преимущество использования автоматизированного машинного обучения неоспоримо.

Таким образом, можно с уверенностью назвать автоматизацию машинного обучения актуальной и активно развивающейся областью исследований. Поиск новых методов выбора моделей, перекрестной проверки, эволюционного и аналитического подбора алгоритмов обучения представляет как научный, так и чисто практический интерес.

ЗАКЛЮЧЕНИЕ

Развитие технологий машинного обучения только ускорится в ближайшем будущем. В настоящее время мы наблюдаем прогресс в развитии методик автоматизированного поиска путей построения эффективных обучающихся моделей анализа данных, применимых ко многим практическим задачам интеллектуального анализа данных.

В ходе обзора современных тенденций в машинном обучении нами выделены следующие перспективные направления фундаментальных и прикладных исследований в данной области:

1. Теоретические исследования в области интерпретируемости моделей искусственного интеллекта в сочетании с анализом автоматически построенных моделей.
2. Практические исследования мультизадачных, генеративных (порождающих) моделей.
3. Более широкое распространение автоматизированных средств машинного обучения.
4. Развитие и унификация инструментальных средств, в том числе облачных средств и сервисов интеллектуального анализа данных.
5. Разработка новых интеллектуальных продуктов пользовательского уровня, основанных на вышеперечисленных достижениях методологии машинного обучения.

Кроме того, стоит отметить почти полное отсутствие научных публикаций на русском языке по анализируемым в обзоре темам. Существует ряд работ, посвященных методологии интеллектуального моделирования сложных систем, которые закладывают основы. Кроме того, в данном обзоре рассматривались публицистические и научно-популярные работы в данных областях машинного обучения. Их наличие является неоспоримым плюсом и дает надежду, что широкое распространение данных тем исследований в отечественных научных кругах еще впереди.

БИБЛИОГРАФИЧЕСКИЙ СПИСОК

- Хохлова Д.* (2016). Бум нейросетей: Кто делает нейронные сети, зачем они нужны и сколько денег могут приносить, 12.06.2016. Режим доступа: <https://vc.ru/16843-neural-networks> (дата обращения: 06.09.2018).
- Caruana R.* (1998). *Multitask learning* springer, Boston, MA, P. 95–133.
- Dai J., He K, J. Sun.* (2014). Instance-aware semantic segmentation via multi-task network cascades. Режим доступа: https://www.cv-foundation.org/openaccess/content_cvpr_2016/app/S14-02.pdf (дата обращения: 06.09.2018).
- Dong D., Wu H., He W., Yu D. and Wang H.* (2015). Multi-task learning for multiple language translation // Proceedings of the 53rd Annual Meeting of the Association for Computational Linguistics and the 7th International Joint Conference on Natural Language Processing, July 26–31. P. 1723–1732.
- Ghosh J., Bengio Y.* (1997). Multi-task learning for stock selection. Режим доступа: <http://papers.nips.cc/paper/1221-multi-task-learning-for-stock-selection.pdf> (дата обращения: 06.09.2018).
- Kaiser L.* [и др.]. (2017). One model to learn them all. ArXiv:1706.05137 [cs.LG]. Режим доступа: <https://arxiv.org/abs/1706.05137> (дата обращения: 06.09.2018).
- Koh P.W., Liang P.* (2017). Understanding black-box predictions via influence functions // Proceedings of the 34th International Conference on Machine Learning, PMLR. Vol. 70. P. 1885–1894.
- Le Q. & Zoph B.* (2017). Google using machine learning to explore neural network architecture // Research Blog, Wednesday, May 17, 2017. Режим доступа: <https://research.googleblog.com/2017/05/using-machine-learning-to-explore.html> (дата обращения: 06.09.2018).
- Li S., Liu Z.-Q., Chan A. B.* (2014). Heterogeneous multi-task learning for human pose estimation with deep convolutional neural network. Режим доступа: https://www.cv-foundation.org/openaccess/content_cvpr_workshops_2014/W15/papers/LI_Heterogeneous_Multi-task_Learning_2014_CVPR_paper.pdf (дата обращения: 06.09.2018).
- Luong M.-T.* [и др.]. (2015). Multi-task sequence to sequence learning. ICLR 2016. Режим доступа: <https://arxiv.org/abs/1511.06114> (дата обращения: 06.09.2018).
- Misra I., Shrivastava A., Gupta A., Hebert M.* (2016). Cross-stitch networks for multi-task learning. Режим доступа: https://www.cv-foundation.org/openaccess/content_cvpr_2016/papers/Misra_Cross-Stitch_Networks_for_CVPR_2016_paper.pdf (дата обращения: 06.09.2018).
- Molnar C.* (2018). Interpretable machine learning. Режим доступа: <https://christophm.github.io/interpretable-ml-book/> (дата обращения: 06.09.2018).
- Olson R.* (2016). TPOT: A Python tool for automating data science. Режим доступа: <https://www.kdnuggets.com/2016/05/tpot-python-automating-data-science.html/2> (дата обращения: 06.09.2018).
- Olson R.S., Moore J.H.* (2016). TPOT: A tree-based pipeline optimization tool for automating machine learning. P. 66–74.
- Paredes B.R.* [и др.]. (2012). Exploiting Unrelated tasks in multi-task learning // PMLR. Vol. 22. P. 951–959.
- Thornton C.* [и др.]. (2013). Auto-WEKA: combined selection and hyperparameter optimization of classification algorithms ACM, P. 847–855.
- Yao X.* (1999). Evolving artificial neural networks // Proceedings of the IEEE. 1999. № 9 (87). С. 1423–1447.
- Zhang Y., Yang Q.* (2017). A survey on multi-task learning. Режим доступа: <https://arxiv.org/abs/1707.08114> (дата обращения: 06.09.2018).
- Zoph B. & Le Q.V.* (2016). Neural architecture search with reinforcement learning. Режим доступа: <https://arxiv.org/abs/1611.01578> (дата обращения: 06.09.2018).

REFERENCES

- Caruana R. (1998), *Multitask learning springer*, Boston, MA, P. 95–133.
- Dai J., He K, Sun J. (2014), “Instance-aware semantic segmentation via multi-task network cascades”, available at: https://www.cv-foundation.org/openaccess/content_cvpr_2016/app/S14-02.pdf (accessed September 6, 2018).
- Dong D., Wu H., He W., Yu D. and Wang H. (2015), “Multi-task learning for multiple language translation”, *Proceedings of the 53rd annual meeting of the association for computational linguistics and the 7th international joint conference on natural language processing*, July 26–31, pp. 1723–1732.
- Ghosh J. & Bengio Y. (1997), “Multi-task learning for stock selection”, available at: <http://papers.nips.cc/paper/1221-multi-task-learning-for-stock-selection.pdf> (accessed September 6, 2018).
- Kaiser L. et al. (2017), “One model to learn them all”, available at: <https://arxiv.org/abs/1706.05137> (accessed September 6, 2018).
- Khokhlova D. (2016), “Neural networks boom: Who makes neural networks, why they are needed and how much money can bring”, June 2, [“Bum nejrosetej: Kto delaet nejronnye seti, zachem oni nuzhny i skol'ko deneg mogut prinosit”, 02.06.2016], available at: <https://vc.ru/16843-neural-networks> (accessed September 6, 2018) (accessed September 6, 2018).
- Koh P.W. & Liang P. (2017), “Understanding black-box predictions via influence functions”, *Proceedings of the 34th international conference on machine learning*, PMLR, vol. 70, pp. 1885–1894.
- Le Q. & Zoph B. (2017), “Google using machine learning to explore neural network architecture”, *Research blog*, Wednesday, May 17, available at: <https://research.googleblog.com/2017/05/using-machine-learning-to-explore.html> (accessed September 6, 2018).
- Li S., Liu Z.-Q. & Chan A. B. (2014), “Heterogeneous multi-task learning for human pose estimation with deep convolutional neural network”, available at: https://www.cv-foundation.org/openaccess/content_cvpr_workshops_2014/W15/papers/LI_Heterogeneous_Multi-task_Learning_2014_CVPR_paper.pdf (accessed September 6, 2018).
- Luong M.-T. et al. (2016), “Multi-task sequence to sequence learning”, *ICLR*, available at: <https://arxiv.org/abs/1511.06114> (accessed September 6, 2018).
- Misra I., Shrivastava A., Gupta A. & Hebert M. (2016), “Cross-stitch Networks for Multi-task Learning”, available at: https://www.cv-foundation.org/openaccess/content_cvpr_2016/papers/Misra_Cross-Stitch_Networks_for_CVPR_2016_paper.pdf (accessed September 6, 2018).
- Molnar C. (2018), “Interpretable machine learning”, available at: <https://christophm.github.io/interpretable-ml-book/> (accessed September 6, 2018).
- Olson R. (2016), “TPOT: A Python tool for automating data science”, available at: <https://www.kdnuggets.com/2016/05/tpot-python-automating-data-science.html/2> (accessed September 6, 2018).
- Olson R.S. & Moore J.H. (2016), “TPOT: A tree-based pipeline optimization tool for automating machine learning”, pp. 66–74.
- Paredes B.R. et al. (2012), “Exploiting unrelated tasks in multi-task learning”, PMLR, vol. 22, pp. 951–959.
- Thornton C. et al. (2013), “Auto-WEKA: combined selection and hyperparameter optimization of classification algorithms ACM”, pp. 847–855.
- Yao X. (1999). “Evolving artificial neural networks”, *Proceedings of the IEEE*, no 9 (87), pp. 1423–1447,
- Zhang Y. & Yang Q. (2017), “A survey on multi-task learning”, available at: <https://arxiv.org/abs/1707.08114> (accessed September 6, 2018).
- Zoph B. & Le Q.V. (2016), “Neural architecture search with reinforcement learning”, available at: <https://arxiv.org/abs/1611.01578> (accessed September 6, 2018).