

ГЛУБОКОЕ ОБУЧЕНИЕ И ГЕНЕРАТИВНО-СОСТЯЗАТЕЛЬНЫЕ СЕТИ

DEEP LEARNING AND GENERATIVE ADVERSARIAL NETWORKS

V. Magomadov

Summary. this article explores the field of artificial intelligence known as generative-adversarial networks, which is one of the models of machine learning. This area, like machine learning in General, is a very active research topic that has the potential to make a major breakthrough in the field of artificial intelligence. This article explains what generative-adversarial networks are made of and what their purpose is.

Keywords: generative-adversarial networks, artificial intelligence, machine learning, deep learning, neural networks, discriminative models, adversarial networks, algorithms, generator, discriminator.

Магоматов Висхан Салманович

Старший преподаватель, ФГБОУ ВО Чеченский
Государственный университет, г. Грозный, Россия
vmagomadov@gmail.com

Аннотация. в этой статье исследуется область искусственного интеллекта, известная как генеративно-состязательные сети, являющихся одной из моделей машинного обучения. Эта область, как и машинное обучение в целом, является очень активной темой исследований, которая имеет потенциал сделать серьезный прорыв в сфере искусственного интеллекта. Данная статья объясняет, из чего состоят генеративно-состязательные сети, и каково их предназначение.

Ключевые слова: генеративно-состязательные сети, искусственный интеллект, машинное обучение, глубокое обучение, нейронные сети, дискриминативные модели, состязательные сети, алгоритмы, генератор, дискриминатор.

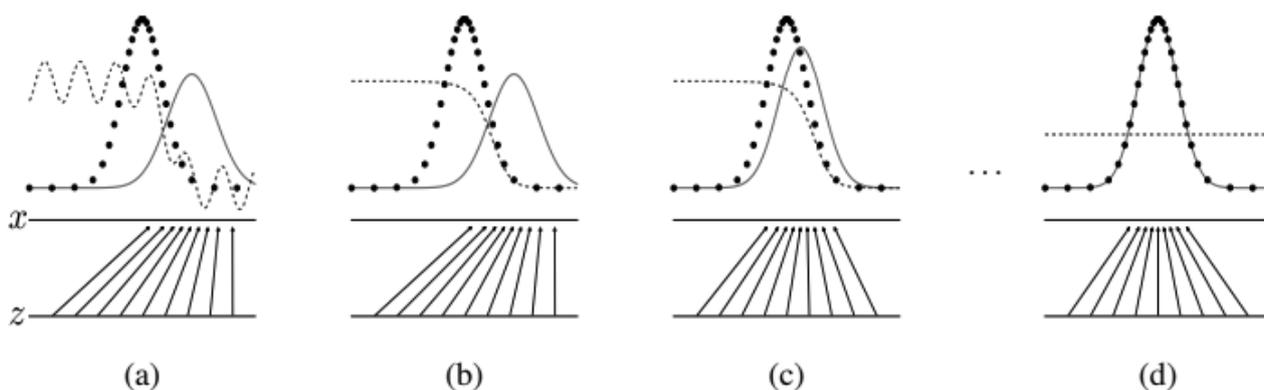
Перспектива глубокого обучения заключается в открытии богатых, иерархических моделей [1], представляющих вероятностные распределения по видам данных, встречающихся в приложениях искусственного интеллекта, таких как естественные изображения, звуковые волны, содержащие речь, и символы в корпусах естественного языка. До сих пор самые поразительные достижения в глубоком обучении были связаны с дискриминативными моделями, обычно с теми, которые наносят высокоразмерный, богатый сенсорный ввод на метку класса [2]. Эти достижения в первую очередь были основаны на алгоритмах обратного распространения и отсева, используя кусочно-линейные единицы [3], имеющие исключительно хороший градиент. Глубокие генеративные модели оказали меньшее влияние из-за сложности аппроксимации многих трудно-разрешимых вероятностных вычислений, возникающих при оценке максимального правдоподобия и связанных стратегий, а также из-за сложности использования преимуществ кусочно-линейных единиц в генеративном контексте. В этой статье, предлагается новая процедура оценки генеративной модели, позволяющей обойти эти трудности.

В системе состязательных сетей генеративная модель противостоит дискриминативной модели, которая учится определять, является ли образец из распределения модели или распределения данных. Генеративную модель можно рассматривать как нечто схожее с командой фальшивомонетчиков, пытающихся производить фальшивую валюту и использовать ее без обнаружения, в то время как дискриминативная модель аналогична полиции, пытающейся обнаружить фальшивую валюту.

Конкуренция в этой игре заставляет обе команды совершенствовать свои методы, пока подделки будут неотличимы от подлинных предметов.

Эта система может дать конкретные алгоритмы обучения для многих видов моделей и алгоритмов оптимизации. В этой статье исследуется особенный случай, когда генеративная модель генерирует образцы, передавая случайный шум через многослойный перцептрон. Этот особый случай называется состязательными сетями. В этом случае мы можем обучить обе модели, используя только очень успешные алгоритмы обратного захвата и отсева [4] и образцы из генеративной модели, используя только прямое распространение. Нет необходимости в приближительном выводе и цепях Маркова.

До недавнего времени большинство трудов по глубоким генеративным моделям были сосредоточены на моделях, обеспечивающих параметрическую спецификацию функции распределения вероятностей. Затем модель может быть обучена путем максимизации логарифмического правдоподобия. В этом семействе моделей, пожалуй, наиболее успешной является глубокая машина Больцмана [5]. Такие модели обычно имеют трудно-разрешимые функции правдоподобия и поэтому требуют многочисленных аппроксимаций к градиенту правдоподобия. Эти трудности стали причиной разработки «генеративных машин» — модели, которые явно не представляют вероятность, но способны генерировать образцы из желаемого распределения. Генеративные стохастические сети [6] являются примером генеративной машины, которая может быть обучена с точным обратным распространением, а не многочисленными



аппроксимациями, необходимыми для машин Больцмана. Эта работа расширяет идею генеративной машины, устраняя цепи Маркова, используемые в генеративных стохастических сетях.

Генеративно-состязательные сети — это мощный класс нейронных сетей, используемых для неконтролируемого обучения. Они были разработаны и представлены Йоном Дж. Гудфеллоу в 2014 году. Генеративно-состязательные сети состоят из системы двух нейросетевых моделей, конкурирующих друг с другом и способны анализировать, фиксировать и копировать вариации в наборе данных [7].

Было замечено, что большинство мейнстримных нейронных сетей можно легко обмануть неправильной классификацией вещей, добавив только небольшой объем шума в исходные данные. Удивительно то, что модель, после добавления шума, имеет больше уверенности в неправильном предсказании, чем, когда она предсказывала правильно. Причиной такого противника заключается в том, что большинство моделей машинного обучения учатся на ограниченном количестве данных, что является огромным недостатком, поскольку оно склонно к переоснащению. Кроме того, отображение между вводом и выводом является почти линейным. Может показаться, что границы разделения между различными классами являются линейными, но на самом деле, они состоят из линейностей и даже небольшое изменение точки в пространстве признаков может привести к неправильной классификации данных [8].

Состязательные сети

Структура состязательного моделирования наиболее проста в применении, когда модели являются многослойными перцептронами. Чтобы узнать распределения генератора p_g по данным x , определяются переменные входного шума $p_z(z)$, затем представляется отображение на пространство данных как $G(z; \theta_g)$, где G — дифференцируемая функция, представленная

многослойным перцептроном с параметрами θ_g . Также определяется второй многослойный перцептрон $D(x; \theta_d)$, который выводит один скаляр. $D(x)$ представляет вероятность того, что x пришел из данных, а не p_g . D обучается максимизировать вероятность присвоения правильной метки как учебным примерам, так и образцам из G . Мы одновременно обучаем G минимизировать $\log(1 - D(G(z)))$.

В следующем разделе, мы представляем теоретический анализ состязательных сетей, по существу, показывая, что критерий обучения позволяет восстановить распределение данных, поскольку G и D имеют достаточную емкость, т.е. в непараметрическом пределе. На практике, мы должны реализовать игру, используя итерационный, численный подход. Оптимизация D до завершения во внутреннем цикле обучения является вычислительно запретительной, а на конечных наборах данных приведет к переобучению. Вместо этого, мы чередуем между шагами k оптимизирования D и одним шагом G . Это приводит к тому, что D поддерживается вблизи своего оптимального решения, пока G изменяется достаточно медленно.

Теоретические результаты

Генератор G неявно определяет распределение вероятностей p_g как распределение образцов $G(z)$, полученных при $z \sim p_z$. Результаты этого раздела сделаны в непараметрической установке, например, мы представляем модель с бесконечным объемом путей изучения сходимости в пространстве функций плотности вероятности (см. рисунок).

Генеративные состязательные сети обучаются путем одновременного обновления дискриминативного распределения (D , синяя, пунктирная линия) таким образом, чтобы он дискриминировал между образцами генерирующего данные распределения (черная, пунктирная линия) p_x и образцами генеративного распределения p_g (G) (зеленая, сплошная линия). Нижняя горизонтальная

линия — это область, из которой берется образец от Z , в этом случае равномерно. Горизонтальная линия выше является частью области x . Стрелки вверх показывают, как отображение $x = G(z)$ накладывает неравномерное распределение p_g на преобразованные образцы. G сжимается в областях высокой плотности и расширяется в областях низкой плотности p_g . (a) Рассмотрим со-
 стязательную пару, близкой к конвергенции: p_g похож на p_{data} , а D является частично точным классификатором. (b) Во внутреннем цикле алгоритма, D обучается дискриминировать образцы из данных, приближающихся к

$$D * (x) = \frac{P_{data}(x)}{P_{data}(x) + P_g(x)}$$

(c) После обновления до G , градиент D направил $G(z)$ в регионы, которые с большей вероятностью будут классифицированы как данные. (d) После нескольких этапов обучения, если G и D имеют достаточно емкости, они достигнут точки, в которой они оба не могут больше улучшиться, поскольку $P_g = P_{data}$. Дискриминатор не в состоянии отличить два дискриминатора друг от друга, т.е. $D(x) = 1/2$.

Эксперименты

Мы оцениваем вероятность данных под P_g , приспособив Гауссово окно Парцена к образцам, сгенерированным с G , и сообщая правдоподобие при этом распределении. Параметр σ Гауссов был получен путем перекрестной валидации на наборе валидаций. Эта процедура используется для различных генеративных моделей, для которых точное правдоподобие не поддается обработке [9, 10]. Этот метод оценки вероятности имеет высокую дисперсию и плохо работает в многомерных пространствах, но пока что это лучший доступный метод. Достижения в области генеративных моделей, которые

могут отбирать образцы, но не оценивают вероятность, являются мотивацией для дальнейшего исследования того, как оценивать такие модели.

Преимущества и недостатки

Этот новый фреймворк имеет определенные преимущества и недостатки. Недостатки заключаются прежде всего в том, что нет явного представления $p_g(x)$, а также в том, что D должен быть хорошо синхронизирован с G во время обучения (в частности, G не должен обучаться слишком много без обновления D , чтобы избежать сценария, в котором G обваливает слишком много значений z до того же значения x , чтобы иметь достаточное разнообразие для моделирования p_{data}), так же, как и отрицательные цепочки машины Больцмана должны обновляться между шагами обучения. Преимущества в том, что марковские цепи не нужны, только метод обратного распространения ошибки используется для получения градиентов, никакого вывода не требуется во время обучения, и широкий спектр функций может быть включен в модель.

Вышеупомянутые преимущества в первую очередь являются вычислительными. Состязательные модели могут также получить некоторое статистическое преимущество, поскольку генераторная сеть не обновляется непосредственно примерами данных, а только градиентами, проходящими через дискриминатор. Это означает, что компоненты входного сигнала не копируются непосредственно в параметры генератора. Еще одним преимуществом состязательных сетей является то, что они могут представлять очень резкие, даже вырожденные распределения, в то время как методы, основанные на марковских цепях, требуют, чтобы распределение было несколько размытым, чтобы цепи могли смешиваться между режимами.

ЛИТЕРАТУРА

1. Домингос П. Верховный алгоритм: как машинное обучение изменит наш мир. Москва: Манн, Иванов и Фербер, 2016. 61 с.
2. Флах П. Машинное обучение: наука и искусство построения алгоритмов, которые извлекают знания из данных. Москва: ДМК, 2015. 42 с.
3. Бенгфорт Б. Прикладной анализ текстовых данных на Python: машинное обучение и создание приложений обработки естественного языка. Санкт-Петербург: Питер, 2019. С. 124–126.
4. Грифт Н. Прагматичный ИИ: машинное обучение и облачные технологии. Санкт-Петербург: Питер, 2019. С. 171–172.
5. Гудфеллоу Я., Бенджио И, Курвилль А. Глубокое обучение. Москва: ДМК, 2018. С. 118–121.
6. Вьюгин В. В. Математические основы машинного обучения и прогнозирования. Москва: МЦНМО, 2014. С. 113–116.
7. Lakhmi, C.J, Balas V. E., Johri, P. Data and Communication Networks: Proceeding of GUCON2018. Sydney: Springer, 2019. pp. 51–52.
8. Langr, J., Bok, V. GANs in Action: Deep Learning with Generative Adversarial Networks. New York: Manning Publications Company, 2019. pp. 66–67.
9. Николенко С., Кадурин А., Архангельская Е. Глубокое обучение. Погружение в мир нейронных сетей. Санкт-Петербург: Питер, 2017. С. 101–102.
10. Шолле Ф. Глубокое обучение на Python. Санкт-Петербург, 2018. С. 56–57.