широким частотным диапазоном. Эти преобразователи работают по принципу сейсмического эффекта, т.е. инерционного смещения массы чувствительного элемента под действием вибрационного нагружения [3].

1. Конструкция и проектирование жидкостных ракетных двигателей Жуковский А. Е., Кондрусев В. С., Окорочков В. В. 2-е изд., перераб. и доп. М.: Машиностроение, 1992. – 352 с.
2. Автоматизация физических исследований и эксперимента: компьютерные измерения и виртуальные приборы Бутырин П. А., Васьковская Т. А., Каратаев В. В. – М.: ДМК пресс, 2005. – 264 с.
3. Испытания жидкостных ракетных двигателей Жуковский, А.Е. - М.: Книга по Требованию, 2012. - 101 с.

Митрофанова А.С., Комлев Г.В.
Машинное обучение

*Сибирский государственный университет науки и технологий имени
академика М.Ф. Решетнева
(Россия, Красноярск)*

doi:10.18411/lj-11-2018-180

idsp: ljournal-11-2018-180

Аннотация

Нет никакой связи между наукой о данных и машинным обучением, но машинное обучение может быть использовано для того, чтобы заставить машины учиться (улучшать производительность по конкретным задачам) из данных. Это область компьютерных наук, которая дает возможность компьютерам учиться с использованием различных статистических методов. Из огромных наборов данных машинное обучение может применять знания для превосходного распознавания речи, распознавания лиц или многих других задач. Машинное обучение предоставляет системам возможность распознавания шаблонов и составления прогнозов.

Ключевые слова: машинное обучение

Основная задача науки и реальной жизни – получение правильных предсказаний о будущем поведении сложных систем на основании их прошлого поведения. Многие задачи, возникающие в практических приложениях, не могут быть решены заранее известными методами или алгоритмами. Это происходит по той причине, что нам заранее не известны механизмы порождения исходных данных или же известная нам информация недостаточна для построения модели источника, генерирующего поступающие к нам данные. Как говорят, мы получаем данные из «черного ящика». В этих условиях ничего не остается, как только изучать доступную нам последовательность исходных данных и пытаться строить предсказания, совершенствуя нашу схему в процессе предсказания. Подход, при котором прошлые данные или примеры используются для первоначального формирования и совершенствования схемы предсказания, называется методом машинного обучения [1].

Машинное обучение (machine learning) – это область научного знания, имеющая дело с алгоритмами, «способными обучаться». Необходимость использования методов машинного обучения объясняется тем, что для многих сложных – «интеллектуальных» – задач (например, распознавание рукописного текста, речи и т. п.) очень сложно (или даже невозможно) разработать «явный» алгоритм их решения, однако часто можно научить компьютер обучиться решению этих задач. Одним из первых, кто использовал термин «машинное обучение», был изобретатель первой самообучающейся компьютерной программы игры в шашки А. Л. Самуэль в 1959 г. [2].

Одних только данных недостаточно. Домингос писал: «...нет ничего удивительного в том, что для обучения нужны знания. Машинное обучение не может получить что-то из ничего, но может получить большее из меньшего. Обучение похоже на сельское хозяйство, где большую часть работы делает природа. Фермеры дают семенам питательные вещества, чтобы вырастить урожай. Так и здесь: чтобы создать программу, нужно совместить знания и данные».

Большое количество данных лучше детально продуманного алгоритма. Не пытайтесь изобретать велосипед и усложнять решения: выбирайте кратчайший путь, ведущий к цели. Домингос говорит: «Как правило, «глупый» алгоритм с большим количеством данных превосходит «умный» алгоритм с небольшим количеством данных. В машинном обучении главную роль всегда играют данные» [3].

Основная масса задач, решаемых при помощи методов машинного обучения, относится к двум разным видам: обучение с учителем (supervised learning) либо без него (unsupervised learning). Однако этим учителем вовсе не обязательно является сам программист, который стоит над компьютером и контролирует каждое действие в программе. «Учитель» в терминах машинного обучения – это само вмешательство человека в процесс обработки информации. В обоих видах обучения машине предоставляются исходные данные, которые ей предстоит проанализировать и найти закономерности. Различие лишь в том, что при обучении с учителем есть ряд гипотез, которые необходимо опровергнуть или подтвердить. Эту разницу легко понять на примерах.

Машинное обучение с учителем. Предположим, в нашем распоряжении оказались сведения о десяти тысячах московских квартир: площадь, этаж, район, наличие или отсутствие парковки у дома, расстояние от метро, цена квартиры и т. п. Нам необходимо создать модель, предсказывающую рыночную стоимость квартиры по её параметрам. Это идеальный пример машинного обучения с учителем: у нас есть исходные данные (количество квартир и их свойства, которые называются признаками) и готовый ответ по каждой из квартир – её стоимость. Программе предстоит решить задачу регрессии. Ещё пример из практики: подтвердить или опровергнуть наличие рака у пациента, зная все его медицинские показатели. Выяснить, является ли входящее письмо спамом, проанализировав его текст. Это всё задачи на классификацию.

Машинное обучение без учителя. В случае обучения без учителя, когда готовых «правильных ответов» системе не предоставлено, всё обстоит ещё интереснее. Например, у нас есть информация о весе и росте какого-то количества людей, и эти данные нужно распределить по трём группам, для каждой из которых предстоит пошить рубашки подходящих размеров. Это задача кластеризации. В этом случае предстоит разделить все данные на 3 кластера (но, как правило, такого строгого и единственного возможного деления нет). Если взять другую ситуацию, когда каждый из объектов в выборке обладает сотней различных признаков, то основной трудностью будет графическое отображение такой выборки. Поэтому количество признаков уменьшают до двух или трёх, и становится возможным визуализировать их на плоскости или в 3D. Это – задача уменьшения размерности [4].

1. Математические основы теории машинного обучения и прогнозирования. В.В.Вьюгин. - М.: 2013. – 390 с.
2. Введение в разработку мультимедийных приложений с использованием библиотек OpenCV и IPP. Лекция №2. Введение в машинное обучение [Электронный ресурс]. Режим доступа: <http://www.intuit.ru/studies/courses/19469/1105/lecture/17981>.
3. A Few Useful Things to Know about Machine Learning. Pedro Domingos. Department of Computer Science and Engineering. University of Washington Seattle, WA 98195-2350, U.S.A. [Электронный ресурс]. Режим доступа: <https://homes.cs.washington.edu/~pedrod/papers/cacm12.pdf>.
4. Машинное обучение: виды, алгоритмы, примеры. [Электронный ресурс]. Режим доступа: https://www.gd.ru/articles/9348-mashinnoe-obuchenie?from=PW_blurring&ustp=W.