

УДК 622.32:004.85

АСПЕКТЫ МАШИННОГО ОБУЧЕНИЯ В КРУПНОЙ КОМПАНИИ НЕФТЕГАЗОВОЙ ОТРАСЛИ

К. И. Бушмелева, А. Н. Васильчук
Сургутский государственный университет
bkiya@yandex.ru, alex.hack89@gmail.com

В статье рассмотрены базовые теоретические понятия, а также актуальность и практическая значимость применения машинного обучения в нефтегазовой промышленности. Представлена информация о преимуществах использования машинного обучения.

Ключевые слова: искусственный интеллект, машинное обучение, автоматизация, нефтегазовая промышленность.

MACHINE LEARNING ASPECTS IN A MAJOR COMPANY OF OIL AND GAS INDUSTRY

K. I. Bushmeleva, A. N. Vasilchuk
Surgut State University
bkiya@yandex.ru, alex.hack89@gmail.com

The article discusses standard theoretical conceptions, relevance and practical significance of machine learning in the oil and gas industry. The article provides information on the benefits of using the machine learning.

Keywords: artificial intelligence, machine learning, automation, oil and gas industry.

Введение

Как известно, мы различаем механическое заучивание и интеллектуальное осмысление. Забуривание телефонных номеров или инструкций, без сомнения, тоже относится к процессу обучения. Но, как правило, под этим понятием мы подразумеваем кое-что другое.

Ребенок, играющий с друзьями, наблюдает реакцию других членов группы на свои действия. Этот опыт влияет на его будущее поведение в социуме. Он не вспоминает и не проигрывает заново свое прошлое, а опирается на определенные, легко опознаваемые характеристики прошлых взаимодействий: детская площадка, класс, мама, папа, сестры и братья, друзья, незнакомцы, взрослые, дети, в помещении или на улице. Оценка новой ситуации базируется на признаках, с которыми ему доводилось сталкиваться раньше. Обучение при этом является не просто сбором информации. Формируется то, что можно назвать аналитической оценкой.

Представьте, как вы по картинкам учите ребенка отличать собаку от кошки. Показанная картинка кладется в одну из двух стопок, в зависимости от правильности ответа. Чем дольше продолжается процесс, тем выше эффективность распознавания. Интересно, что нет необходимости специально учить ребенка отличать собаку от кошки. Человеческое сознание обладает встроенными механизмами классификации. Ему требуются только образцы. Научившись работать с картинками, ребенок сможет опознать практически любое изображение кошки или собаки, не говоря уже о реальных животных. Эта способность обобщать, применяя полученные в процессе тренировок знания к новым, ранее не встречавшимся образцам, является ключевой характеристикой как человеческого, так и машинного обучения.

Разумеется, процесс получения знаний человеком превосходит своей сложностью самые совершенные алгоритмы машинного обучения, но у компьютера есть преимущество в

виде большой емкости для запоминания, извлечения и обработки данных. Накапливаемый им опыт представлен в форме данных за длительный период времени, причем это представление позволяет получать и оптимизировать алгоритмы, реализующие если не аналитическую оценку, то хотя бы способность к обобщениям.

Аналогия между человеческим и машинным обучением закономерно заставляет вспомнить такое явление, как *искусственный интеллект* (англ. *Artificial Intelligence, AI*). При этом естественным образом возникает вопрос «Чем искусственный интеллект отличается от машинного обучения?» [1]. Для ответа на этот вопрос необходимо дать определение данным понятиям.

В настоящее время тематика искусственного интеллекта охватывает огромный перечень направлений, начиная с таких задач общего характера, как обучение и восприятие, и заканчивая такими специальными задачами, как игра в шахматы, доказательство математических теорем, сочинение поэтических произведений и диагностика заболеваний. В искусственном интеллекте систематизируются и автоматизируются интеллектуальные задачи, и поэтому эта область касается любой сферы интеллектуальной деятельности человека. В этом смысле искусственный интеллект является поистине универсальной научной областью [3].

Из сказанного выше, можно сделать вывод, что понятие «Искусственный интеллект» не имеет точного и четкого определения. Приведем несколько определений искусственного интеллекта:

«Наука и технология создания интеллектуальных машин, интеллектуальных компьютерных программ» [7].

«Изучение умственных способностей с помощью вычислительных моделей» [5].

«Автоматизация действий, которые мы ассоциируем с человеческим мышлением, т. е. таких действий, как принятие решений, решение задач, обучение» [4].

«Искусство создания машин, которые выполняют функции, требующие интеллектуальности при их выполнении людьми» [6].

«Наука о проектировании интеллектуальных агентов» [9].

«Наука, посвященная изучению интеллектуального поведения артефактов (артефакт – искусственный объект)» [8].

Машинное обучение (англ. *Machine Learning, ML*) является одной из форм искусственного интеллекта, которая заключается в извлечении знаний из данных. Это научная область, находящаяся на пересечении статистики, искусственного интеллекта и компьютерных наук, также известная как прогнозная аналитика или статистическое обучение [2].

Для обучения и обобщения люди используют различные данные из примеров, включая формы, цвета, текстуры, пропорции и другие характеристики. Машинное обучение также применяет множество стратегий в различных комбинациях в зависимости от поставленной задачи.

Эти стратегии нашли свое воплощение в наборе алгоритмов, разработанных в течение последних десятилетий, как учеными, так и практиками в самых разных дисциплинах – от статистики, компьютерной науки, робототехники и прикладной математики до поиска в Интернете, развлекательной сферы, цифровой рекламы и переводов с одного языка на другой. Алгоритмы крайне разнообразны и имеют как сильные, так и слабые стороны. Некоторые относят объекты к определенному классу, другие предсказывают числовые значения. Существуют и алгоритмы, определяющие сходства и различия допускающих сравнение сущностей. При этом все алгоритмы обучаются на примерах (опыте) и умеют применять полученные знания к новым, ранее не встречавшимся случаям, то есть способны к обобщению.

Актуальность

В настоящее время машинное обучение применяется к широкому кругу экономических задач – от обнаружения мошенничества до выбора целевой аудитории и рекомендации товара. В таблице представлены некоторые задачи машинного обучения.

Таблица 2

Задачи машинного обучения

Задача	Примеры
Классификации	Медицинская диагностика: по набору медицинских характеристик требуется поставить диагноз. Геологоразведка: по данным зондирования почв определить наличие полезных ископаемых. Кредитная оценка: по анкете заемщика принять решение о выдаче/отказе кредита
Восстановления регрессии	Оценка стоимости недвижимости: по характеристике района, экологической обстановке, транспортной связности оценить стоимость жилья. Кредитная оценка: по анкете заемщика оценить величину кредитного лимита
Кластеризации	Финансовая сфера: по сводкам банковских операций выявить группы «подозрительных», нетипичных банков, сгруппировать остальные по степени близости проводимой стратегии
Идентификации	Банковское дело: определить подлинность подписи на чеке. Медицинская диагностика: по набору медицинских характеристик требуется установить наличие/отсутствие конкретного заболевания
Прогнозирования	Биржевое дело: прогнозирование биржевых индексов и котировок. Экономика: прогноз цен на недвижимость
Извлечения знаний	Медицина: поиск взаимосвязей между различными показателями при фиксированной болезни. Научные исследования: получение новых знаний об исследуемом процессе. Биржевое дело: определение закономерностей между различными биржевыми показателями

Но какое применение машинное обучение может найти в нефтегазовой промышленности? В современной нефтегазовой отрасли при разработке месторождений принятие решений базируется на данных, объем которых растет со временем экспоненциально. Замерные эксплуатационные данные поступают со всех скважин, кроме этого формируются месячные данные по добыче, также имеется информация о проведенных исследованиях и физических характеристиках пласта. В то же время, качество этих данных не всегда позволяет провести полноценный анализ: может отсутствовать информация для определенных временных интервалов, некоторые измерения не всегда соответствуют физической модели или не согласовываются друг с другом. Присутствие в отчетах некорректных данных может быть вызвано как сбоем в работе замерного оборудования, так и «человеческим фактором».

Определить ошибку силами специалистов не всегда возможно, а некорректная информация повлечет за собой неверные выводы о текущем состоянии скважин и месторождении в целом, вследствие чего могут быть приняты неверные решения по проведению геолого-технических мероприятий.

Применение методов машинного обучения повысят скорость обработки и анализа больших объемов информации, которые поступают с месторождения. Кроме того, методы машинного обучения позволят интегрировать разнородные данные, проанализировать каждый мегабайт имеющейся информации, что приведет к появлению новых выводов способных вывести качество данных на новый уровень. Это, несомненно, повлияет на повышение эксплуатационных показателей.

В результате, применение новых алгоритмов может существенно повысить скорость и эффективность работы специалистов по разработке месторождений, снизить риски принятия неверных капиталоемких решений при разработке, вызванных «человеческим фактором», и уменьшить время простоя скважин.

Ниже представлены самые существенные преимущества систем с машинным обучением в сравнении с общепринятыми альтернативами, такими как ручной анализ, жестко запрограммированные бизнес-правила и простые статистические модели [1]:

Точность. Машинное обучение использует данные для создания принимающей решение программы, оптимизированной под поставленную задачу. По мере накопления данных автоматически возрастает точность прогнозов.

Автоматизация. По мере подтверждения и отбрасывания ответов ML-модель может автоматически обнаруживать новые шаблоны. Это позволяет встраивать машинное обучение непосредственно в автоматизированные рабочие процессы.

Скорость. Машинное обучение дает ответы за доли секунды после поступления новой информации, позволяя системам реагировать в реальном времени.

Возможность настройки. Многие задачи, управляемые данными, можно решить с помощью машинного обучения. Модели строятся на базе ваших собственных данных и допускают настройку под любую систему мер, принятую в вашем бизнесе.

Масштабируемость. При росте бизнеса ML-модель легко приспособливается к увеличению объема данных. Некоторые алгоритмы можно использовать для обработки множества данных на разных вычислительных машинах в облаке.

Вывод

Сегодня машинное обучение, несомненно, является одним из самых перспективных и актуальных направлений информационных технологий, которое применяется к широкому кругу задач.

Цифровые технологии меняют подходы нефтяных компаний к выбору вариантов разработки и эксплуатации месторождений. Новейшие способы работы с информацией позволяют повысить эффективность использования собираемых на месторождениях данных, принимать более взвешенные решения.

Литература

1. Бринк Х., Ричардс Дж., Феверолф М. Машинное обучение. СПб. : Питер, 2017. 336 с.
2. Мюллер А., Гвидо С. Введение в машинное обучение с помощью Python : пер. с англ. М. : Вильямс, 2017. 480 с.
3. Стюарт Р., Норвиг П. Искусственный интеллект: современный подход : пер. с англ. 2-е изд. М. : Вильямс, 2006. 1408 с.
4. Bellman R. E. An Introduction to Artificial Intelligence: Can Computers Think? San Francisco : Boyd & Fraser Publishing Company, 1978.
5. Charniak E., McDermott D. Introduction to Artificial Intelligence. Cambridge, Massachusetts : Addison-Wesley, Reading, 1985.
6. Kurzweil R. The Age of Intelligent Machines. Cambridge; Massachusetts : MIT Press, 1990.
7. McCarthy J. What is Artificial Intelligence? [Электронный ресурс]. URL: <http://www-formal.stanford.edu/jmc/whatisai/whatisai.html> (дата обращения: 03.09.2017).
8. Nilsson N. J. Artificial Intelligence: A New Synthesis. San Mateo, California : Morgan Kaufmann, 1998.
9. Poole D., Mackworth A. K., Goebel R. Computational intelligence: A logical approach. Oxford, UK : Oxford University Press, 1998.