

Блог | 5 минут

20 сентября 2023

Почему специалисты по прогнозированию не используют машинное обучение?

Машинное обучение (ML) и ИИ сейчас очень популярно в сообществе аналитиков и специалистов по Data Science. Каждый уважающий себя отдел прогнозирования разрабатывает алгоритмы ML для предсказания того, кто кликнет, купит, соврет или умрет (если взять заголовок основополагающей работы Эрика Зигеля по этой теме).

Все конференции и публикации по аналитике наполнены модными терминами из мира ИИ. Однако, когда речь заходит о реальном использовании, большинство специалистов по прогнозированию осторожны в вопросах применения машинного обучения. Почему так происходит? Разве в их обязанности не входит предсказывать, что и есть суть машинного обучения? Давайте рассмотрим возможности и подводные камни применения машинного обучения в прогнозировании.



Специалисты по прогнозированию и специалисты по Data Science различно определяют понятие 'предсказание'

Есть тонкая разница в том, как прогнозирование и машинное обучение определяют

"предсказание". Когда специалисты по прогнозированию говорят о "предсказании", они имеют в виду прогнозирование будущего.

Традиционные методы прогнозирования включают в себя модели временных рядов, алгебраические уравнения и качественные оценки. В результате, традиционное прогнозирование относительно ручное и затратное по времени и может быть подвергнуто влиянию экспертного мнения. Однако результаты легко интерпретировать, и это гибкий процесс; специалист по прогнозированию понимает, откуда берутся числа, и может легко вносить коррективы по мере необходимости. Кроме того, традиционное прогнозирование можно проводить с ограниченными данными.

'Предсказание' в машинном обучении или статистической модели означает предсказание прошлого. Это может показаться контринтуитивным, но идея заключается в сравнении 'предсказания' модели с реальностью и измерении разницы или ошибки. Эти ошибки используются для дополнительной настройки модели для предсказания будущего. Следовательно, предсказания модели в значительной степени зависят от прошлой эффективности и почти невозможно их откорректировать. Также интерпретируемость моделей очень ограничена.

Другой фактор, который следует учитывать, заключается в том, что по своей конструкции ML требует большого объема данных. С другой стороны, машинное обучение быстрое, автоматизированное и объективное, поскольку свободно от человеческого суждения.

Машинное обучение создано для цифрового мира; специалисты по прогнозированию работают в реальном мире

Алгоритмы машинного обучения и ИИ созданы для цифрового мира с почти неограниченными данными о кликах, покупках и просмотрах пользователей. Как известно, эти алгоритмы отлично справляются с задачей стимулирования нас на повторные покупки, приобретение дополнительных товаров и регистрацию в программы лояльности. Потери от ошибок прогнозирования (нереализованные продажи) относительно невелики. Кроме того, каждая ошибка является возможностью для улучшения алгоритма машинного обучения.

Однако реальный рынок сильно отличается от цифрового. Данные здесь могут ограничиваться продажами через кассу, данными программ лояльности или данными по отправкам. Потери от ошибок прогнозирования могут быть довольно высокими, поскольку рестораны и розничные торговцы закупают товары оптом. Кроме того, прогнозы не могут сами себя улучшать, так как нет автоматической обратной связи. По этим причинам многие розничные торговцы и их поставщики по-прежнему полагаются на традиционные методы прогнозирования. Это не значит, что машинное обучение не может предложить возможности для улучшения прогнозирования, но есть ряд аспектов, которые необходимо учесть, прежде чем приступить к машинному обучению.

Машинное обучение требует гораздо больше данных, чем анализ временных рядов

Любой алгоритм машинного обучения требует большого количества данных. Под большим количеством данных я имею в виду не даты или переменные. Модели машинного обучения работают на определенных уровнях наблюдения — это может быть клиент, магазин и так далее. Для работы машинного обучения нужно как минимум несколько тысяч таких наблюдений. Если выборка ограничивается всего 10 магазинами, лучше воздержаться от машинного обучения и использовать методы временных рядов.

Еще одним фактором, который стоит учитывать, является стоимость поддержки данных. Данные доступны или их нужно вводить вручную? Требуется ли обработка данных? Это одноразовый процесс или непрерывная задача, требующая человеческих и вычислительных ресурсов? Каковы будут затраты на хранение данных на протяжении лет?



Интерпретировать работу алгоритмов машинного обучения гораздо сложнее, чем провести анализ временных рядов

По своей конструкции машинное обучение является "черным ящиком". Например, прогнозы могут быть сгенерированы на основе голосования тысяч деревьев принятия решений. Можно использовать красочные гистограммы для иллюстрации веса каждого фактора в модели. Эти диаграммы выглядят очень умно на слайды презентаций, но они далеки от интуитивного понимания. Если стоимость ошибки в прогнозе составляет миллионы долларов, компании могут предпочесть методы временных рядов и арифметику, которые они могут понять, вместо элегантного "черного ящика". Это особенно актуально для новых продуктов без данных о продажах или с ограниченными

тестовыми данными.

Существует несколько обходных путей для понимания машинного обучения. Эксперименты с параметрами могут быть хорошим индикатором надежности результатов. Если незначительное изменение входных данных или спецификаций модели приводит к существенным изменениям в прогнозах, это может быть красным флагом.

В конце концов, надежность модели может быть подтверждена тестированием на новых данных. Нам не обязательно нужно понимать все тонкости алгоритмов, если мы уверены в конечном результате. Убедительность этого аргумента может зависеть от вашей аудитории. Как правило, профессионалы в области аналитики комфортно чувствуют себя, используя прогнозы машинного обучения, если они проверены.

Руководители по управлению цепочками поставок могут быть более осторожными в принятии бизнес-решений на основе "черного ящика". Хорошей проверкой на здравый смысл является параллельное использование традиционных методов прогнозирования и машинного обучения. Если между результатами есть заметная разница, это может указывать на проблему с моделью или на то, что при создании традиционного прогноза был упущен важный аспект.

Стоимость и выгода от использования машинного обучения не всегда очевидны

Само собой разумеется, что правильно настроенное машинное обучение работает чрезвычайно эффективно. Все, что нужно, это предоставить входные данные и нажать кнопку. Сложность настройки может варьироваться от относительно простой до очень сложной, в зависимости от цели прогноза и доступных данных.

Продукты с богатой историей могут быть легко прогнозируемы даже с использованием стандартных пакетов машинного обучения, таких как SAS или Azure, при условии, что данные легко доступны.

Прогнозы для новых продуктов могут потребовать сложных алгоритмов-заменителей для работы с ограниченными данными. Это может потребовать разработки алгоритмов машинного обучения с нуля. Кроме того, может потребоваться интеграция данных из разных источников для «питания» алгоритма. Это, в свою очередь, может потребовать существенных инвестиций в найм подрядчиков, расширение команды аналитиков или дополнительные нагрузки на существующие ресурсы.

Прежде чем наращивать команду специалистов по данным, компаниям стоит подумать о частоте использования алгоритма, приросте эффективности и ресурсах, необходимых для проекта.

Влияние на общее планирование бизнеса

Прогнозирование является ключевым элементом планирования бизнеса. Любые изменения в процессе прогнозирования могут повлиять на другие сферы бизнеса, такие как финансы и управление цепочками поставок.

Обычно традиционные методы прогнозирования основаны на верхнеуровневом подходе. Прогноз создается в агрегированной форме, а затем детализируется по магазинам/временным периодам и так далее. Эти детализации могут быть использованы для финансовых целей или для планирования спроса на уровне магазина.

В отличие от этого, прогноз на основе машинного обучения использует нисходящий подход. Прогноз создается на уровне магазина/временного периода и затем агрегируется. При переходе от традиционного прогнозирования к машинному обучению компании должны обеспечить плавный переход на всех этапах планирования бизнеса. Если это сделано неправильно, это может привести к расхождениям между прогнозами на основе машинного обучения и финансовыми планами и планами поставок.

В заключение, машинное обучение — отличный инструмент для оптимизации прогнозирования. Как и любой инструмент, у него есть свои применения, преимущества, стоимости и риски.

При использовании машинного обучения для прогнозирования, компании должны учитывать свои данные, бизнес-потребности, культуру принятия решений и процесс планирования. Хорошим началом может быть тестирование машинного обучения на ваших данных с использованием онлайн-решений «из коробки», таких как Azure и SAS.

Большинство этих решений имеют пошаговые обучающие видео, которые помогут настроить алгоритм машинного обучения под ваши данные. Эксперименты с этими решениями могут помочь определить, подходит ли машинное обучение для прогнозирования в вашей компании, и достаточно ли решения «из коробки», или требуется разработка внутри компании. Даже если окажется, что по каким-то причинам машинное обучение не подходит для вашей компании, инвестиции не будут потеряны, а некоторые аналитические знания все равно будут приобретены.

Автор: Olga Gerasymchuk, специалист по прогнозированию

Оригинал статьи: <https://demand-planning.com/2023/08/28/why-arent-demand-planners-adopting-machine-learning/>

**Узнайте, чем программа Forecast NOW!
будет полезна именно вашей компании**

+7 (495) 929-71-05

info@forecastnow.ru

fnow.ru

© 2011-2024 Российский разработчик - Инжэниус Тим