

ВОЗМОЖНОСТИ ПРИМЕНЕНИЯ МЕТОДОВ МАШИННОГО ОБУЧЕНИЯ В СФЕРЕ НЕДВИЖИМОСТИ

М. В. Нарчук, студентка 4 курса ГИУСТ БГУ

Научный руководитель:

кандидат технических наук, доцент **Т. В. Борздова**
(ГИУСТ БГУ)

Реалии таковы, что выживают только те идеи, которые проходят путь от красивого обещания обществу до решения, приносящего реальную пользу. Причем чем универсальнее формулировки и возможности применения, тем большую огласку получает идея, что и закономерно. Машинное обучение является ярким примером именно такой идеи и в данный момент находится на пике интереса у представителей разных направлений бизнеса.

Объяснение сути машинного обучения стоит начать с того, что процесс обработки данных подразумевает наличие трех элементов: входных данных, логики и выходных данных. Написание кода заключается в том, что логика строится в зависимости от входных данных, представляя собой пассивный элемент процесса – набор предписаний – и являясь актуальной только для набора данных конкретного типа. Анализ данных не типичного для разработанной логики формата приводит к необходимости написания другого кода, то есть выработки другой логики. Машинное обучение базируется на идее о существовании общего алгоритма, способного адаптироваться под входные данные, выстраивая собственную логику.

Понятие «машинное обучение» тесно связано с концепцией интеллектуального анализа данных, которая предполагает обработку огромных массивов данных и поиск среди них неочевидных взаимосвязей и закономерностей. Человек, разрабатывая ту или иную логику, физически не способен выделить и учесть большое количество факторов. Многофакторный анализ подразумевает обработку огромного количества разнотипных данных.

Проблема популяризованных идей зачастую заключается в том, что представители бизнеса сосредотачиваются на обещанных выгодах, не уделяя должного внимания переходу от идейной части к ее техническому воплощению. Машинное обучение в своем развитии проходит стадию «завышенных ожиданий», переходя к стадии поиска путей практической реализации идей.

Человек на протяжении долгих лет получает многомерные входные данные непосредственно из окружающей среды посредством органов чувств, то есть он способен эффективно распознавать данные, но не способен эффективно их хранить, обрабатывать и анализировать. Компьютер не способен рас-

познавать данные, но намного более эффективен, когда речь заходит о хранении, обработке и анализе. Человек способен воспринимать окружающий мир, а компьютер – данные, описывающие окружающий мир, из чего следует, что входные данные являются для компьютера ретранслируемыми. Данные, передаваемые человеком компьютеру как входные, зачастую являются неточными, противоречивыми, содержат пропуски и в некоторых случаях имеют субъективный характер, что приводит к возникновению «предубеждений» по поводу некоторых факторов и получению результатов, не соответствующих действительности. Идейная часть машинного обучения ограничивается не только техническими возможностями машин, но и человеческими возможностями. Как бы последние ни расширились, этого всегда будет недостаточно.

К настоящему моменту с учетом всех ограничений и возможностей машинное обучение позволяет эффективно решать три типа задач: классификация, регрессия и выявление аномалий. Первые две задачи реализуются через обучение с учителем. Последняя – через обучение без учителя. Существует принципиальная разница между этими подходами. Обучение с учителем реализуется в два этапа: на первом этапе входными данными служит выборка прецедентов «объект-ответ»; на выходе формируется алгоритм, предсказывающий ответ по любому объекту. На втором этапе по входящим новым данным предсказывается ответ на основе полученного на первом этапе «опыта». Обучение без учителя не предполагает наличия четко поставленной задачи и правильных ответов.

На данном этапе машинное обучение как практическое решение для бизнеса является достаточно сложно реализуемой, трудоемкой задачей в силу новизны темы. Отдельные достижения сосредоточены в разных сферах деятельности, целостная картина только начинает вырисовываться, отсутствует теоретический базис, делающий решение общедоступным. Как методы машинного обучения могут быть реализованы в сфере недвижимости? Какую же пользу машинное обучение может принести участникам рынка недвижимости?

Объекты недвижимости долговечны, и их стоимость имеет свойство меняться с течением времени под влиянием множества факторов. Многофакторный анализ – верный путь к определению наиболее точной и объективной стоимости объекта.

Опытный оценщик или риэлтер может «интуитивно» определить стоимость объекта исходя из тех или иных внешних признаков или сочетания определенных факторов. Обучение с учителем может помочь сохранить и приумножить накопленный опыт. Специалист задает параметры, на которых система учится оценивать новые данные. В полученную систему можно продолжить вносить изменения, обеспечивая применение такого количества подходов

к обработке и анализу данных, какое опять же на физическом уровне не в силах обеспечить ни один специалист.

Машинное обучение позволяет эффективно решать задачи классификации, начиная от группировки клиентов, тем самым обеспечивая риэлтеров возможностью делать предложения, наиболее точно «попадающие в цель», и заканчивая возможностью фильтровать объявления по фото, выбирая квартиры с фотографиями планировки, фасада здания или обустройства двора.

С помощью машинных методов решаются задачи поиска недооцененных или переоцененных квартир. Решение этой задачи специалистом без применения машинных методов потребовало бы реализации ряда точечных операций и не гарантировало бы точности результатов. Машинное обучение позволяет превратить решение задач данного типа в процесс [1].

Одним из основных рисков рынка недвижимости является риск мошенничества. А учитывая то, что недвижимость является дорогостоящим активом, цена ошибки действительно велика. Появление платформ для размещения объявлений об аренде или продаже объектов недвижимости было призвано значительно упростить процесс поиска. Однако распространенной проблемой стало размещение объявлений как злоумышленниками, желающими путем обмана быстро получить крупные суммы, так и недобросовестными риэлтерами, желающими привлечь больше клиентов, публикуя предложения настолько же привлекательные, насколько нереалистичные. В борьбе с этим негативным фактором значительную роль мо-

жет сыграть многофакторный анализ, реализуемый с применением методов машинного обучения. Анализ в этом случае подлежит весь спектр доступной информации: от описания объявления, контактных данных и цены до поиска по фотографии [2].

Также машинное обучение позволяет оптимизировать график работы сотрудников агентств недвижимости в соответствии с интенсивностью потока клиентов, прогнозировать конверсии сделок и многое другое.

Таким образом, возможности применения машинного обучения на данном этапе полностью зависят от креатива и знаний разработчиков-энтузиастов. Хотя потенциал технологии и огромен, существующие ограничения также значительны. Многие решения позволяют автоматизировать рутинные процессы, смещая фокус деятельности специалистов на решение созидательных задач. В сфере недвижимости, как и в прочих областях, до повсеместного применения технологии далеко, однако реализованные сегодня кейсы даже на промежуточных этапах показывают результаты, превосходящие все ожидания.

Литература

1. Machine Learning is Fun! [Электронный ресурс] // E-magazine Medium. – Режим доступа: <https://medium.com/@ageitgey/machine-learning-is-fun-80ea3ec3c471>. – Дата доступа: 01.03.2018.
2. Три кейса применения анализа данных в сфере недвижимости [Электронный ресурс] // Электронный журнал «Tmfeed». – Режим доступа: <https://habrahabr.ru/company/newprolab/blog/339116/>. – Дата доступа: 01.03.2018.