
ПРЕДИКТИВНАЯ АНАЛИТИКА: ИЗ ИСТОРИЧЕСКИХ ДАННЫХ К СТРАТЕГИЧЕСКОМУ ПРОГНОЗИРОВАНИЮ

Маштаков Никита Сергеевич

Студент бакалавриата 4 курс

МИРЭА-Российский технологический университет (РТУ МИРЭА)

Институт информационных технологий

Часов Павел Сергеевич

Студент бакалавриата 4 курс

МИРЭА-Российский технологический университет (РТУ МИРЭА)

Институт информационных технологий

Аннотация

В данной статье мы исследуется роль предиктивной аналитики в улучшении принятия решений в различных отраслях. Рассматривается использование статистических методов, предиктивного моделирования и машинного обучения для прогнозирования будущих событий на основе анализа исторических данных. Статья подчеркивает важность предиктивной аналитики как инструмента для достижения стратегических целей организации.

Ключевые слова: предиктивная аналитика, предписывающая аналитика, финансовые показатели, прогнозирование.

PREDICTIVE ANALYTICS: FROM HISTORICAL DATA TO STRATEGIC FORECASTING

Nikita S. Mashtakov

Bachelor's degree student 4nd year

MIREA-Russian Technological University (RTU MIREA)

Institute of Information Technology

Pavel S. Chasov

Bachelor's degree student 4nd year

MIREA-Russian Technological University (RTU MIREA)

Institute of Information Technology

ABSTRACT

In this article, we explore the role of predictive analytics in improving decision-making in various industries. The use of statistical methods, predictive modeling and machine learning to

predict future events based on the analysis of historical data is considered. The article emphasizes the importance of predictive analytics as a tool for achieving the strategic goals of the organization.

Keywords: predictive analytics, prescriptive analytics, financial indicators, forecasting.

Предиктивная аналитика – это механизм, необходимый организации для улучшения процесса принятия решений, независимо от функции или отрасли, в которой работает организация. Кроме того, внедрение прогностической аналитики приводит к конкурентному преимуществу, которое трудно найти в другом месте [2].

В этой статье мы обсудим, что такое предиктивная аналитика, рассмотрим несколько примеров ее использования и посмотрим, как она работает.

Что такое предиктивная аналитика?

Предиктивная аналитика – это обобщающий термин, который описывает различные статистические методы и методы анализа данных, включая интеллектуальный анализ данных, прогностическое моделирование и машинное обучение. Основная цель прогностической аналитики – делать прогнозы относительно результатов, тенденций или событий на основе шаблонов и выводов из исторических данных.

Прогнозирующая аналитика – это вторая из четырех стадий аналитических возможностей в организации. Четыре стадии аналитики по порядку таковы:

1. Описательная аналитика – выявление того, что происходило в прошлом
2. Диагностическая аналитика – понимание того, почему это произошло
3. Предиктивная аналитика – предсказывание того, что произойдет дальше
4. Предписывающая аналитика – оптимизация и эксперименты с тем, как наилучшим образом это реализовать

Организации должны проходить эти этапы аналитики в таком порядке, поскольку эффективно предсказывать будущее можно, только понимая прошлое. Таким образом, организации переходят от понимания того, что произошло и почему это произошло, к прогнозированию того, что произойдет дальше. Дополнительный заключительный этап аналитики включает в себя полностью оптимизированные автономные аналитические системы, которые постоянно обучаются с течением времени и, по сути, являются "интеллектуальными".

Прогностическое моделирование включает в себя два типа алгоритмов машинного обучения: контролируемые и неконтролируемые. Алгоритмы машинного обучения с контролем используются для прогнозирования целевого результата и являются основными инструментами прогностической аналитики.

Существует два основных типа контролируемых алгоритмов машинного обучения:

- Классификационные модели – используются для прогнозирования того, попадут ли наблюдения в определенную категорию или класс. Например, для прогнозирования того, будет ли отток клиентов или нет. Распространенные методы классификации включают деревья решений и модели логистической регрессии.
- Регрессионные модели – используются для прогнозирования значения. Например, прогнозирование количества кликов по рекламе.

Распространенными методами прогнозирования подобных значений являются модели линейной регрессии и полиномиальной регрессии.

С другой стороны, неконтролируемые алгоритмы машинного обучения не делают прогнозов, а скорее стремятся выявить закономерности в данных, которые затем могут быть использованы для обозначения или группировки сходных точек данных. Например, одним из самых популярных неконтролируемых алгоритмов является кластеризация k-средних, при которой похожие точки данных, такие как клиенты, группируются в кластеры.

Предиктивная аналитика может также включать другие методы статистики и интеллектуального анализа данных для выявления текущих тенденций, прогнозирования будущего и прогнозирования результатов. Далее в этой статье мы обсудим несколько конкретных примеров того, как организации могут использовать предиктивную аналитику.

Предписывающая и предиктивная аналитика

Предписывающая аналитика – это третий этап аналитических возможностей организации, который основывается на прогностических моделях, созданных на предыдущем этапе.

В то время как прогностическая аналитика говорит нам, почему что-то происходит и что может произойти дальше, основное внимание в предписывающей аналитике уделяется оптимизации и экспериментам с уже построенными моделями. Это отвечает на ваши вопросы "что, если" и позволяет вам действовать по наилучшему возможному сценарию, основанному на информации, которую вы получаете в результате проведения экспериментов и моделирования.

Джефф Безос, генеральный директор Amazon, как известно, сказал: "Наш успех в Amazon зависит от того, сколько экспериментов мы проводим в год, в месяц, в неделю, в день".

Проведение экспериментов по всем аспектам аналитических процессов и проектов является ключевым требованием для успешной директивной аналитики.

Примеры прогностической аналитики

Прогностическая аналитика преобразует точки данных в ценную информацию, которая может управлять многими аспектами организации.

Вот лишь несколько примеров того, как организации могут использовать прогностическую аналитику:

1. Прогнозирование финансовых показателей эффективности

Прогнозирование ключевых финансовых показателей, таких как выручка, расходы и запасы, приводит к более эффективному и обоснованному принятию решений на основе фактов и данных, а не только интуиции.

2. Выявление и сокращение масштабов мошенничества в банковской сфере

Одной из наиболее дорогостоящих и разрушительных ситуаций для банка является мошенническая деятельность. Предиктивная аналитика может помочь выявить отклонения и уязвимости, которые могут указывать на мошенничество, чтобы эти учреждения могли быстро принять меры.

3. Прогнозирование того, допустит ли клиент дефолт по кредиту

Предоставление займов по своей сути сопряжено с риском для страховых и финансовых учреждений. Использование прогностических моделей для прогнозирования вероятности дефолта клиента по кредиту – лучший способ для этих учреждений значительно снизить этот риск.

4. Прогнозирование выбытия сотрудников

Прогностическая аналитика может помочь улучшить управление человеческими ресурсами вашей организации, прогнозируя выбытие сотрудников. Это предполагает предвидение будущих требований к найму и поиск подходящего времени для стимулирования сотрудников.

5. Понимание покупательского поведения клиентов

Организации могут увеличить продажи и коэффициент конверсии, выявляя закономерности, стоящие за покупками клиентов, и исследуя причины их покупательского поведения. Вы можете использовать аналитику клиентов и A/B тестирование на Python, чтобы понять это поведение [1].

6. Нацеливание маркетинговых кампаний на нужных клиентов

Компании могут увеличить количество переходов по рекламе и конверсий маркетинговых кампаний в целом, ориентируясь на нужных клиентов в нужное время.

7. Сокращение производственных отходов

Предиктивная аналитика может помочь вашей организации понять факторы, связанные с производственными отходами, чтобы они могли принимать меры в нужных областях. Использование прогнозных моделей для понимания этих факторов может привести к существенной экономии средств в организациях [3].

Как работает предиктивная аналитика?

Доступно множество различных инструментов прогнозной аналитики. Выбранный вами инструмент зависит от цели вашего анализа – от инструментов бизнес-аналитики и визуализации, таких как Tableau и Power BI, до сложных языков программирования, таких как R Programming и Python.

Большинство проектов по предиктивной аналитике следуют аналогичным рабочим процессам. В этом разделе мы рассмотрим некоторые из основных этапов, наиболее часто встречающихся в проекте.

Цель

Каждый проект в области прогнозной аналитики должен начинаться с понимания цели, выявления проблемы и выбора наилучшего решения.

Проекты прогнозной аналитики направлены на то, чтобы помочь организации в достижении ее стратегических целей. Когда цель проекта связана с критически важной задачей для организации, более вероятно, что его поддержат на всех уровнях организации, к которым он имеет отношение. Это гарантирует, что проект не только ценен, но и успешен.

Как только цель проекта определена, это естественным образом приводит к четкой проблеме, которую необходимо решить проекту, и, следовательно, информирует о решении, необходимом для достижения этой цели.

Данные

Данные могут поступать из различных источников, таких как CSV-файлы, базы данных, хранилища и сторонние приложения. Если это еще не сделано, эти источники данных следует консолидировать и управлять ими в централизованном месте, прежде чем вы сможете использовать их в прогнозной аналитике и моделировании. Это гарантирует безопасность данных и то, что качество и управление являются главными приоритетами.

Как гласит старая поговорка в области анализа данных: "мусор на входе, мусор на выходе". Старайтесь избегать хранения всех ваших данных в электронных таблицах. Несмотря на гибкость, электронные таблицы легко редактируются и используются совместно без контроля качества и безопасности содержащихся в них данных.

Кроме того, учитывайте объем данных, которые у вас есть или которые вы хотели бы сгенерировать в будущем. Построение прогностических моделей может стать неоправданно трудоемким и неэффективным для задач, которые они должны решать, без надлежащих систем хранения и процессов для обработки больших объемов данных.

Преобразовать

Этап преобразования в проекте прогнозной аналитики включает очистку, изучение и подготовку данных для последующего анализа или модели.

При очистке данных следует искать отсутствующие данные и выбросы или подозрительные значения, которые не имеют смысла в контексте проблемы. Это идет рука об руку с исследованием данных, поскольку именно изучая данные, вы начинаете их понимать, и аномалии становятся для вас более заметными.

Наконец, данные должны быть подготовлены для следующего шага проекта. Точный процесс здесь зависит от алгоритма и типа анализа, который необходимо выполнить.

Например, предположим, что вы применяете модель линейной регрессии для прогнозирования количества переходов по ссылкам в рекламной кампании. В этом случае вы должны учитывать допущения линейной регрессии и то, удовлетворяют ли им данные. Также необходимо преобразовать любые категориальные переменные в ваших данных в фиктивные переменные, известные как одноуровневое кодирование.

Прежде чем переходить к построению прогностической модели, обязательно разделите свои данные на обучающие, тестирующие и валидационные наборы. Вы подгоняете модель к обучающему набору, и именно так она будет изучать закономерности в данных для составления прогнозов. Однако у вас также должен быть набор для проверки, чтобы выполнить итерацию и улучшить модель, а затем получить окончательную, непредвзятую оценку точности модели с использованием набора тестов.

Анализировать

Если вы строите прогностическую модель, используя относительно простой алгоритм машинного обучения с контролем, такой как логистическая регрессия, то на этом этапе вам потребуется подогнать модель и оценить результаты. Однако некоторые сложные алгоритмы, такие как нейронные сети, требуют тщательной настройки и корректировок для получения точных прогнозов.

Важно помнить, что многие прогностические модели требуют больших объемов данных для точного обобщения в реальном мире. Если вам все еще нужно получить объемы данных, необходимые для этих моделей, вы можете рассмотреть другие методы, которые позволяют прогнозировать результаты в меньшем масштабе. Это также зависит от вашей цели и бизнес-задачи, которую вы пытаетесь решить, и включает в себя такие методы, как интеллектуальный анализ данных и применение различных статистических методов.

Развертывать

Заключительным шагом в проекте прогнозной аналитики является развертывание. Это конечный результат проекта, который будет служить средой, с помощью которой модель повышает ценность вашей организации. В зависимости от проекта и решения, выбранного вами для решения вашей конкретной проблемы, этот шаг может включать в себя что угодно – от простого отчета или панели мониторинга до сложных развертываний на существующих платформах.

Подумайте, достаточно ли в вашей организации специалистов для более сложных внедрений, чтобы обеспечить бесперебойный и эффективный процесс. Например, инвестиции как минимум в 1 специалиста по обработке данных могут значительно способствовать успешному внедрению модели.

Вывод

Аналитика, и, в частности, предиктивная аналитика, предназначена не только для нескольких технологических гигантов и крупных корпораций или даже для немногих избранных внутри организации. Сегодня организации любого размера используют аналитику, и ее можно применять практически в любой отрасли. Кроме того, предиктивная аналитика теперь является функцией, которая распределяется и принадлежит командам по всей организации.

Предиктивная аналитика может предложить невероятное конкурентное преимущество практически каждой организации. Однако крайне важно учитывать и понимать элементы, которые входят в успешный проект предиктивной аналитики. В этой статье представлено руководство по предиктивной аналитике, о том, как ее можно применить к бизнес-задачам, и о процессе, лежащем в основе ее работы.

Список литературы:

1. «А/В-тестирование для продакт-менеджеров». [Электронный ресурс] / Режим доступа: <https://productlab.ru/tpost/614h697r11-ab-testirovanie-dlya-prodakt-menedzherov> (Дата обращения: 03.01.2024)
2. «Предиктивная аналитика: что это такое, методы и инструменты прогностического анализа». [Электронный ресурс] / Режим доступа: <https://www.cleverence.ru/articles/auto-busines/prediktivnaya-analitika-hto-eto-takoe-metody-i-instrumenty-prognosticheskogo-analiza/> (Дата обращения: 03.01.2024)
3. «Что реально происходит в предиктивной аналитике на производствах? » [Электронный ресурс] / Режим доступа: <https://habr.com/ru/companies/T1Holding/articles/656085/> (Дата обращения: 04.01.2024)

References:

1. "A/B testing for product managers." [Electronic resource] / Access mode: <https://productlab.ru/tpost/614h697r11-ab-testirovanie-dlya-prodakt-menedzherov> (Access date: 01/03/2024)
2. "Predictive analytics: what it is, methods and tools for predictive analysis." [Electronic resource] / Access mode: <https://www.cleverence.ru/articles/auto-busines/prediktivnaya-analitika-hto-eto-takoe-metody-i-instrumenty-prognosticheskogo-analiza/> (Date of access: 03.01. 2024)
3. "What is really happening in predictive analytics in manufacturing?" [Electronic resource] / Access mode: <https://habr.com/ru/companies/T1Holding/articles/656085/> (Access date: 01/04/2024)