



ins**ideBIGDATA**

InsideBIGDATA

Руководство по предиктивной аналитике

Daniel D. Gutierrez

TIBCO[™] Spotfire[®]

Определение предиктивной аналитики

Предиктивная аналитика, иногда называемая продвинутой аналитикой — это термин, описывающий ряд аналитических и статистических методов прогнозирования действий или поведения в будущем. В бизнесе предиктивная аналитика используется для принятия упреждающих решений и определения оптимальных действий.

Она основана на статистических моделях и позволяет находить закономерности в исторических и транзакционных данных и определять потенциальные риски и возможности.

Предиктивная аналитика включает в себя ряд действий, которые мы рассмотрим в этой статье: подключение к данным, анализ и визуализация данных исследований, развитие предположений и моделей данных, применение предиктивных моделей, оценка и/или прогнозирование будущих результатов.

История предиктивной аналитики

Корни современной предиктивной аналитики уходят в далекие 1940-е годы, когда правительства начали использовать первые вычислительные модели: метод Монте-Карло, вычислительные модели нейронных сетей и линейное программирование. Это применялось для расшифровки немецких сообщений во время Второй мировой войны, самонаведения орудий ПВО и прогнозного моделирования ядерных цепных реакций в проекте «Манхэттен». В 1960-х годах корпорации и исследовательские институты начали эпоху коммерциализации аналитики с помощью нелинейного программирования и решения эвристических задач на основе компьютеров. Это легло в основу первых моделей прогноза погоды, решения «задачи кратчайшей пути» для авиаперевозок и логистики, а также предиктивного моделирования для принятия решений о кредитном риске. Затем в 1970-х — 1990-х годах получила большее распространение в организациях. Технологические стартапы сделали реальными предписывающую аналитику (Prescriptive analytics) и анализ в режиме реального времени. Тем не менее, предиктивная аналитика была инструментом, в первую очередь для специалистов по статистике и доходила до бизнес-пользователей только в виде статичных отчетов.

Содержание

Определение предиктивной аналитики.....	2
Применение в бизнесе.....	3
Виды предиктивной аналитики.....	4
ПО для предиктивной аналитики.....	6
Язык R как инструмент для предиктивной аналитики.....	7
Подключение к данным для анализа	8
Исследовательский анализ данных.....	9
Предиктивное моделирование.....	10
Внедрение в рабочие процессы.....	11
Заключение.....	11

Сегодня предиктивная аналитика стала одним из важных направлений корпоративной аналитики, которая используется для решения широкого круга задач. Этот тренд обусловлен реалиями мировой экономики: организации постоянно ищут конкурентные преимущества и внедряют новые технологии.

Эти технологические инновации включают в себя более масштабируемые вычислительные мощности, реляционные базы данных, новые технологии на основе больших данных, как Hadoop и программное обеспечение самостоятельного анализа, которое обеспечивает свободный доступ к данным и предиктивным моделям для лиц, принимающих решения. Всё это позволяет организациям успешно конкурировать за счёт аналитических инноваций.

Сначала бизнес использовал аналитику для обнаружения данных, чтобы понять состояние компании и глубоко изучить данные, и понять причины различных событий. Но по мере изучения данных, появляются возможности, чтобы еще больше превзойти своих конкурентов с помощью передовой аналитики.

Использование для бизнеса

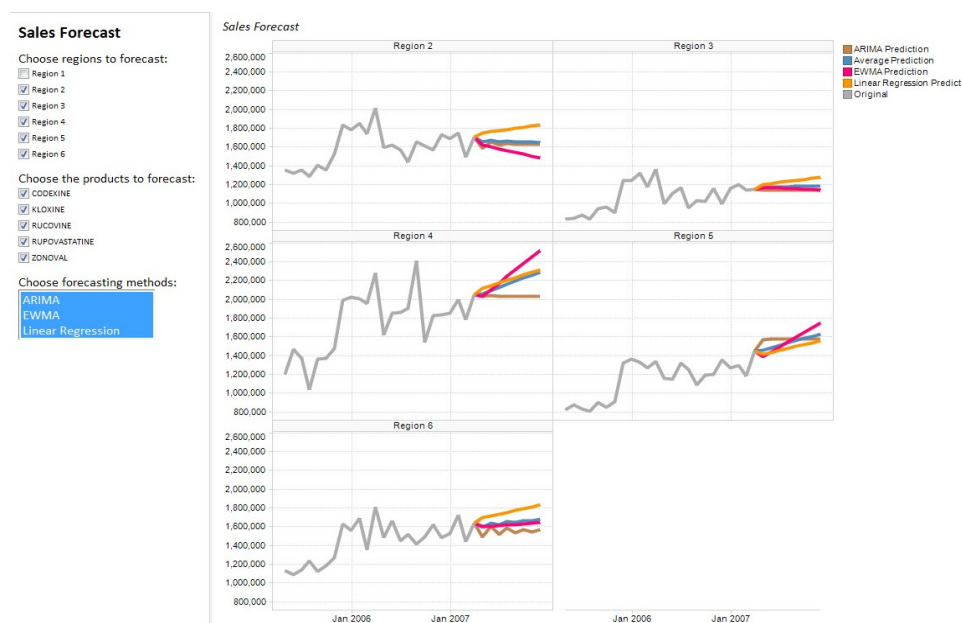
Польза предиктивной аналитики для бизнеса очевидна: она предоставляет более интеллектуальный анализ для принятия оптимальных решений, повышения конкурентоспособности, выявления возможностей и угроз, уменьшения неопределенности и управления рисками, проактивного планирования и действий, обнаружения значимых закономерностей, прогнозирования и своевременного реагирования на изменения.

Продвинутый количественный анализ продемонстрировал преимущества практически во всех сферах экономики.

С помощью предиктивной аналитики многие бизнес-задачи могут быть решены. Некоторые из них:

- **Прогноз продаж** — делайте прогнозы, чтобы знать, каких результатов вам следует ожидать в этом месяце или в квартале, с учетом ваших исторических коэффициентов конверсии. Например, применить процентное соотношение успешных продаж прошлых периодов к количеству текущих заявок, чтобы прогнозировать доходы организации.

- **Обнаружение мошенничества** — находите сомнительные кредитные заявки, мошеннические транзакции, совершаемые офлайн и онлайн, кражи личных данных, мошенничество со страховкой и т.д.
- **Оптимизация стимулирующих кампаний в ритейле** — позволяет маркетологам моделировать результаты кампаний на основе глубокого анализа поведения клиентов, предпочтений и данных профиля.
- **Маркетинговый и клиентский анализ** — собирайте данные цифрового маркетинга, социальных сетей, центров обработки вызовов, мобильных приложений и др. источников. Используйте информацию о прошлых действиях клиентов, чтобы оценить их действия в будущем.
- **HR-аналитика** — позволяет организациям анализировать прошлое и выявлять будущие тенденции ключевых факторов, связанных с добровольным увольнением, неявками и другими источниками риска, а также выявлять тенденции необходимого набора кадров с учетом текущих ресурсов.
- **Управление рисками** — прогнозируйте оптимальную структуру портфеля для максимизации прибыли и оценки соответствующего уровня вероятных рисков.



Предиктивная аналитика для прогноза продаж обеспечивает бизнес-пользователей таргетированной, релевантной аналитикой для улучшения процесса принятия решений. Изображение предоставлено TIBCO Spotfire.

Виды предиктивной аналитики

По своей сути, предиктивная аналитика основана на выявлении связей между историческими данными и прогнозировании будущих результатов на их основе. Для прогнозирования на заданном наборе данных используются одна или несколько переменных-предикторов. Проще говоря, предиктивная аналитика помогает в разработке прогнозов для принятия бизнес-решений.

Чтобы справиться с более высокими требованиями, применяются расширенные методы предиктивной аналитики для управления важными бизнес-процессами.

В этом разделе будут детально рассмотрены основные типы предиктивной аналитики: **контролируемое и неконтролируемое обучение**.

Регрессия — наиболее распространенная форма предиктивной аналитики. Во время регрессии используется количественная переменная ответа (то, что пытаются предсказать).

Контролируемое обучение делится на две большие категории: регрессия для количественных ответов (числовая ценность), например, расход топлива для конкретного автомобиля и классификация для ответов, которые могут иметь только несколько известных значений, таких как «правда» или «ложь».

Регрессия — наиболее распространенная форма предиктивной аналитики. Во время регрессии используется количественная переменная ответа (то, что пытаются предсказать). Например, цена продажи дома основана на ряде предикторных переменных: количество квадратных метров, количество спален и средняя цена в этом районе. Взаимосвязь между ценой и предикторами будет в основе предиктивной модели.

Существует много типов регрессии, включая многомерную линейную регрессию, полиномиальную регрессию и регрессионные деревья.

Классификация — еще один популярный вид предиктивной аналитики. Во время классификации используется категориальная вариация ответа, например, уровень дохода, разделённый на три класса или категории — высокий, средний и низкий доход. Классификатор исследует набор данных, в котором каждое наблюдение содержит информацию о переменной ответа, а также о предикторах.

Предположим, что аналитик хотел бы классифицировать уровень доходов лиц, не входящих в набор данных, исходя из таких связанных с человеком характеристик, как возраст, пол и профессия. Это — задача классификации, которая решается следующим образом: исследуется набор данных, содержащий как предикторные переменные, так и классифицированную переменную ответа — уровень дохода.

Так, алгоритм находит, какие комбинации переменных связаны с уровнем дохода. Этот набор данных называется обучающим набором. Затем алгоритм будет рассматривать новые наблюдения, для которых отсутствует информация о доходе.

Основываясь на классификации учебного набора данных, алгоритм присваивает классификацию новым наблюдениям. Например, 51-летняя женщина, директор по маркетингу, может быть отнесена к группе с высоким уровнем доходов.

Существуют множество методов классификации: логистическая регрессия, деревья решений, поддерживающие векторные машины, случайные леса, k-ближайшие соседи, наивный байесовский классификатор.

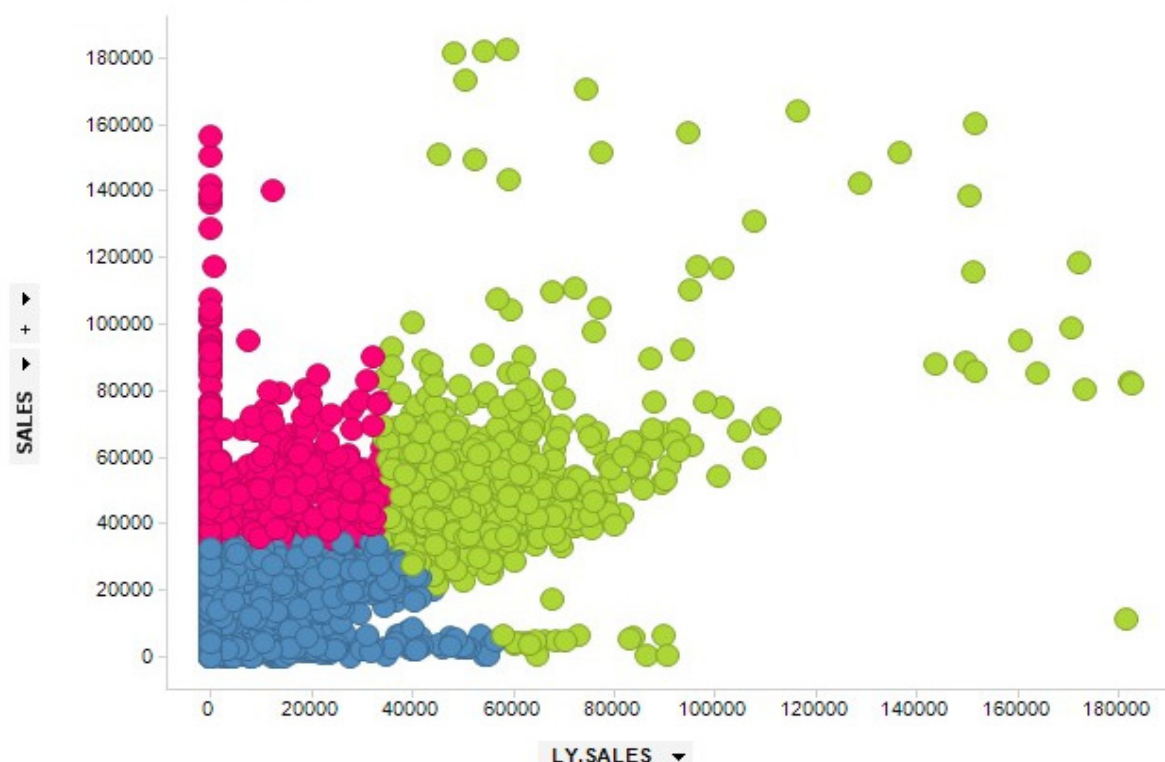
Неконтролируемое обучение применяется для получения выводов из наборов данных, состоящих из входных данных без указанных ответов. Наиболее распространенным методом неконтролируемого обучения является кластерный анализ, который используют для исследования и поиска скрытых закономерностей в данных.

Используя такие неконтролируемые методы, как **кластеризация**, мы можем понять взаимосвязи между переменными или между наблюдениями, определив, попадают ли наблюдения в отличающиеся друг от друга группы. Например, во время анализа сегментов клиентов мы можем наблюдать несколько переменных: пол, возраст, почтовый индекс, доход и др. Наша гипотеза может заключаться в том, что клиенты попадают в разные группы, например — частые и нечастые покупатели.

Классификационный анализ был бы возможен, если бы была доступна история покупок клиентов, но это не относится к неконтролируемому обучению — у нас нет переменных ответа, рассказывающих нам, является ли клиент частым покупателем или нет. Вместо этого мы можем попытаться сгруппировать клиентов на основе переменных, чтобы идентифицировать отдельные группы.

Существуют и другие типы неконтролируемого статистического обучения, в том числе кластеризация k-средних, иерархическая кластеризация, анализ основных компонентов и др.

SALES vs. LY.SALES



Number of Clusters



Кластеризация показывает взаимоотношения между переменными или наблюдениями, определяя, попадают ли они в относительно разные группы. Изображение предоставлено TIBCO Spotfire.

ПО для предиктивной аналитики

Существует множество инструментов интеллектуального анализа с разным функционалом. Программное обеспечение сильно отличается с точки зрения функционала и удобства использования — не все решения могут соответствовать различным потребностям в области расширенной аналитики.

Существуют разные типы пользователей аналитики: некоторые создают статистические модели, другие — применяют.

Продвинутые пользователи больше заинтересованы в таком инструменте, который бы позволил просто создавать аналитические модели и предоставлять их бизнес-пользователям, не обременяя последних необходимостью программирования.

Эффективные инструменты для предиктивной аналитики предоставляют широкий спектр алгоритмов и методов для всех видов данных и бизнес-задач, с которыми сталкиваются пользователи, а также возможность гибкого применения этих алгоритмов по мере необходимости.

Бизнес-пользователи понимают, как предсказательная аналитика может помочь им в работе, но они не всегда могут/понимают как разработать нужную модель.

Таким образом, оптимальные инструменты дают возможность использовать опыт и разработки статистиков при создании аналитических предиктивных моделей, которые бизнес-пользователи самостоятельно могут применять к своим данным и получать ценные сведения. Это позволяет совершенствовать предиктивную аналитику и сделать её доступной по всей организации.

При выборе подходящего инструмента необходимо убедиться, что он обладает полным спектром возможностей, от базового функционала для решения простейших задач до самых современных статистических возможностей.

Эффективные инструменты для предиктивной аналитики предоставляют широкий спектр

алгоритмов и методов для всех видов данных и бизнес-задач, с которыми сталкиваются пользователи, а также возможность гибкого применения этих алгоритмов по мере необходимости.

Масштабируемость системы, позволяющая легко интегрировать новые аналитические методы, по мере их появления, также имеет решающее значение для максимизации выгоды от аналитики. Важно, чтобы функционал соответствовал типу данных, которые вы исследуете, и действительно приносил пользу аналитикам и статистикам.

Правильный инструмент обычно сочетает в себе мощные возможности интеграции и преобразования данных, функционал для исследований, аналитические алгоритмы и обладает интуитивно понятным интерфейсом.

По сути, есть **три важных компонента**, которые являются рецептом успеха использования предиктивной аналитики:

1. статистики создают наиболее конкурентоспособные модели
2. автор аналитического приложения встраивает модели в аналитическое приложение
3. бизнес-пользователь применяет модели в своей работе на регулярной основе.

Ниже приведен **краткий список характеристик**, на которые следует обратить внимание **при оценке инструмента для предиктивного анализа**:

- Поддержка полного цикла предиктивной аналитики: структурирование данных, исследование данных, методы предиктивного моделирования (прогнозирование, кластеризация, скоринг), а также оценка эффективности модели.
- Интеграция знаний, полученных из аналитики в другие части бизнеса, и объединение данных с определенными процедурами, инструментами и рабочими процессами. Это поможет значительно упростить и оптимизировать весь процесс от подготовки данных до прогнозирования.

- Хороший инструмент должен легко интегрироваться с источниками данных, необходимыми для поиска ответов на важные бизнес-вопросы.
- Инструмент должен быть удобным для пользователей всех типов: бизнес-пользователей, аналитиков, статистиков, разработчиков приложений и системных администраторов.
- Рассмотрите инструменты, которые помогут свести к минимуму вмешательство технических специалистов при подключении к новым источникам данных.

Зачем нужен надёжный инструмент? Чтобы гарантированно получить все необходимые возможности для предсказательной аналитики: от простого функционала для прогнозирования и нахождения трендов до создания единой экосистемы моделей и алгоритмов.

Стандартные алгоритмы перестают быть конкурентным преимуществом, как только другие компании начинают тоже их использовать. И это происходит быстро. Поэтому важно обладать инструментом, который позволил создавать собственные модели на основе корпоративных данных, на постоянной основе.

Самые практичные системы — это решения, которые интегрируют предиктивную аналитику в структуру всего аналитического процесса принятия решений, позволяя самостоятельно использовать ее в рамках одного интерфейса. Такой подход обеспечивает доступ к расширенной аналитике для всех пользователей, предоставляя им инструменты, необходимые для выявления новых возможностей, управления рисками и быстрого реагирования на непредвиденные обстоятельства.

В результате профессионалы, управляющие важными департаментами и глобальными процессами, могут быстро формулировать нужные гипотезы и получать ответы из своих данных, предвидеть последующие события и предпринимать быстрые, аргументированные действия.

Язык R как инструмент для предиктивной аналитики

Несмотря на то, что существует множество способов для выполнения задач, связанных с аналитикой, моделированием данных и предиктивным анализом, язык R сегодня является основным фаворитом. Это связано с широким использованием R во время обучения специалистов для работы с коммерческими продуктами, как SAS и SPSS.

В настоящее время проходят оживленные дискуссии между сообществом пользователей R и сообществами SAS и Python относительно того, что является лучшим инструментом для статистики.

Присутствует ряд убедительных фактов, которые выступают за R: открытый исходный код R, широко используемая расширяемая аналитическая среда, более 5000 пакетов доступных на CRAN для расширения функциональности R и возможности визуализации с высоким рейтингом на основе ggplot2.

Кроме того, у пользователей R большое сообщество, в т.ч. местные группы Meetup, онлайн-курсы и специальные блоги (см. топ-блоги через консоль: r-bloggers.com).

Единственная проблема с движком с открытым исходным кодом R — это свойственное ему ограничение масштабируемости во время выполнения задач. R может работать в рамках ограниченной памяти его вычислительной среды.

Движок R является первым, что приходит на ум в качестве движка для предиктивной аналитики. Это объясняется тем, что базовый пакет R уже содержит ряд алгоритмов, которые могут быть увеличены за счёт дополнительных пакетов.

- Линейная регрессия с использованием `lm()`;
- Логистическая регрессия с использованием `glm()`;
- Регрессия с регуляризацией с использованием пакета `glmnet`;
- Нейронные сети, использующие `nnet()`;
- Поддержка векторных машин с использованием пакета `e1071`;
- Модели наивного байесовского классификатора с использованием пакета `e1071`;
- Классификация K-ближайших соседей с использованием функции `knn()` из пакета `класс`;
- Деревья принятия решений с использованием `деревя()`;
- Группы деревьев, использующие пакет случайные леса (`Random Forest`);
- Градиентный бустинг с использованием пакета `gbm`;
- Кластеризация с использованием метода k-средних (`kmeans()`), иерархической кластеризации (`hclust()`).

Единственная проблема с движком с открытым исходным кодом R — это свойственное ему ограничение масштабируемости во время выполнения задач. R может работать в рамках ограниченной памяти его вычислительной среды.

Единственный способ внедрения R для практического использования на предприятии — это использование коммерческой платформы корпоративного уровня для запуска языка R. Например, TERR (TIBCO Enterprise Runtime для R) дает возможность воспользоваться большими преимуществами R, избегая при этом проблем масштабируемости.

Доступ к данным для предиктивной аналитики

Наборы данных предприятия лежат в основе предиктивного аналитического процесса, и любой инструмент должен облегчить интеграцию со всеми источниками данных разных ти-

пов, необходимых для поиска ответов на важные вопросы бизнеса. Надежная предиктивная аналитика должна иметь доступ к аналитическим и реляционным базам данных, OLAP кубам, плоским файлам и корпоративным приложениям.

Для предиктивной аналитики необходимо интегрировать данные из следующих источников:

- Структурированные — БД SQL и хранилища данных, которые уже используются на предприятии.
- Неструктурированные — социальные сети, электронная почта и др.
- Внешние данные от поставщиков — Salesforce и т.д.

Инструменты для предиктивной аналитики должны обеспечить быструю и надёжную интеграцию с несколькими источниками, без необходимости масштабного вовлечения ИТ-специалистов и статистиков.

Пользователи должны обладать гибкостью для быстрого объединения собственных источников данных (таблиц Excel или баз данных Access), с корпоративными хранилищами данных — Hadoop или облачные приложения (например, Hadoop/Hive, Netezza, HANA, Teradata и многие другие). Поэтому широко используются прямые коннекторы для всех методов поддержки аналитики: внутри базы данных, внутри памяти и по требованию.

Внедрение аналитики на R требует меньших вложений, благодаря множеству пакетов, которые обеспечивают подключение к широкому спектру источников данных: ODBC, Excel, CSV, Twitter, Google Analytics и др.

Лучшие предиктивные аналитические решения обеспечивают доступ ко всем источникам данных, что позволяет объединять и сопоставлять данные любым способом, и за счёт этого получать полное представление о бизнесе. Причём без программирования или привлечения ИТ-специалистов. Это позволит пользователям получать ценные сведения и принимать аргументированные бизнес-решения в режиме реального времени.

Интеграция во время подготовки к предиктивному анализу подразумевает ознакомление с данными. Этот процесс называется *исследовательский анализ данных*.

Исследовательский анализ данных

Интеграция во время подготовки к предиктивному анализу подразумевает ознакомление с данными. Этот процесс называется *исследовательский анализ данных*.

Чёткое понимание данных обеспечивает основу для выбора модели — выбор соответствующего алгоритма интеллектуального анализа для решения бизнес-задачи. Пользователи могут применять различное ПО для первоначального исследования данных.

Одним из способов добиться глубокого изучения данных является *использование многочисленных функций статистической среды R*: числовые итоги, графики, агрегации, распределения, удельный вес. Анализируются все уровни факторных переменных и применяются общие статистические методы.

Для исследовательского анализа данных могут успешно применяться и другие средства: TIBCO Spotfire, SAS, SPSS, Statistica, Matlab и др.

Статистическое ПО для R позволяет гибко исследовать и визуализировать данные, но требует высокого уровня знаний в сфере программирования и написания скриптов.

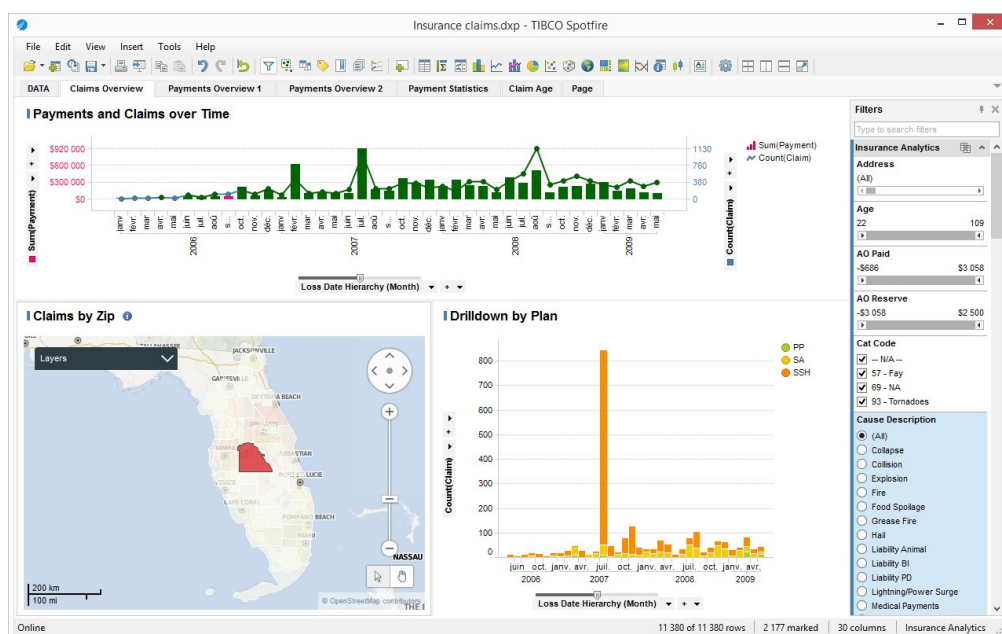
С помощью *детального исследовательского анализа данных вы можете получить ценные сведения и использовать эту информацию для точного прогнозирования*.

R поддерживает множество видов визуализаций для исследовательского анализа данных, включая гистограммы, диаграммы размаха (Boxplots), столбчатые диаграммы, диаграммы рассеяния, тепловые карты и многие другие, входящие в расширение ggplot2.

Применение этих инструментов позволяет глубоко понять данные, которые используются для интеллектуальной аналитики.

Другим методом исследовательского анализа данных является *программное обеспечение для обнаружения данных*.

Функционал обнаружения данных, которым обладает TIBCO Spotfire, помогает широкому кругу пользователей визуально исследовать и понимать свои данные, не обладая глубокими знаниями в области статистики.



Программное обеспечение для обнаружения данных обладает богатым интерактивным аналитическим интерфейсом для исследовательского анализа данных, включая доступ и управление данными, а также проведение анализа. Изображение предоставлено TIBCO Spotfire.

Бизнес-пользователи могут выполнять комплексные задачи исследовательского анализа данных, не обращая за помощью к ИТ-специалистам или статистикам.

Объединение возможностей обнаружения данных и средств предиктивной аналитики в единой платформе аналитики является передовой практикой, которая позволяет беспрепятственно переключаться между этими задачами и экономить на обслуживании.

Предиктивное моделирование

Использование интеллектуальной аналитики предполагает понимание и подготовку данных, определение модели прогнозирования и следование предиктивному процессу.

Предиктивные модели могут иметь разные формы и масштабы в зависимости от их сложности и приложения, для которого были разработаны. Первый шаг моделирования — понять, на какие вопросы бизнеса нужно найти ответ. Уровень детализации и сложности вопросов будет возрастать по мере расширения навыков работы с аналитическим процессом.

Наиболее важными шагами в процессе предиктивной аналитики являются:

- Определение результатов проекта, объема усилий, установки бизнес-целей и выбор используемых наборов данных.
- Сбор данных.
- Обработка данных: проверка, очистка и трансформация данных.
- Исследовательский анализ данных — графический анализ для получения полезной информации и вывода результатов. Для проверки допущений, гипотез и тестов достаточно стандартных статистических методов.
- Моделирование, для автоматического создания точных предиктивных моделей в будущем.

- Оценка выбранной модели, для проверки надежности и внесения корректив во время работы. Тестирование моделей существующих данных и применение прогнозов к новым данным.
- Выбор такого варианта развертывания, который позволит применять результаты аналитики для принятия решений каждый день.

Каждый из вышеперечисленных шагов может повториться и может быть пересмотрен по мере необходимости. Следует отметить, что этап сбора данных часто занимает много времени, в зависимости от уровня структурированности входящих данных и может составлять до 70% от общего времени проекта.

Характеристики данных часто помогают находить, какие методы интеллектуального моделирования могут наилучшим образом удовлетворить потребности аналитика.

Ниже представлены несколько подсказок, которые следует учитывать при определении того, какой метод лучше применять на основе ваших данных и задач, которые вы хотите решить.

- Когда данные группируются на основе наблюдений, такие инструменты, как кластерный анализ, ассоциативные правила и k-ближайшие соседи, как правило, обеспечивают наилучшие результаты.
- Используйте классификацию для разделения данных на классы на основе переменной ответа — как двоичных классов, таких как True или False, так и многоклассовых ситуаций.
- Используйте одиночную, множественную и полиномиальную регрессию во время прогноза, а не классификации.
- При низком качестве или ограниченных данных A/B является самой подходящей. Примерами A/B являются статистические эксперименты, которые помогут вам решить, действительно ли изменение оказывает существенное влияние на ваш продукт.

Внедрение в рабочие процессы

На заключительном этапе проекта по созданию системы предиктивной аналитики нужно определить, как лучше всего внедрить её в рабочие процессы. Первоочередной задачей является использование открытого источника R в больших наборах данных, где важна производительность. Пакет статистической обработки с открытым исходным кодом R не был создан для использования корпоративного уровня.

Развертывание R с открытым исходным кодом может быть проблематичным по следующим причинам:

1. Низкий уровень управления памятью.

R имеет невысокие показатели использования памяти. Следовательно, при высоких нагрузках появляются сбои во время работы вне памяти, а также производительность не растёт равномерно к увеличению нагрузки.

2. Риск развертывания R с открытым исходным кодом с лицензией GPL — поставщикам программного обеспечения запрещается встраивать или распространять открытый источник R в составе любого коммерческого программного обеспечения с закрытым исходным кодом.

Чтобы избежать этой проблемы, аналитики часто предпочитают преобразовывать свое рабочее решение, разработанное на R в другие среды программирования, такие как C++ или Python. Этот путь не самый лучший, поскольку требует значительных усилий перекодировки и повторного тестирования.

Лучшее решение — использовать коммерческие системы корпоративного уровня на R, такое как TIBCO Runtime for R (TERR) для устранения вышеуказанных ограничений и создания надежной рабочей среды. Поскольку у многих предприятий уже есть внедренные предиктивные модели, рекомендуется также, чтобы ваша платформа аналитики поддерживала модели TERR, R с открытым исходным кодом, S+, MATLAB и SAS, чтобы с легкостью использовать все экосистемы.

Заключение

Мы рассмотрели, как предиктивная аналитика помогает вашей организации с уверенностью прогнозировать будущие события, чтобы бизнес-пользователи могли принимать более аргументированные решения и улучшать бизнес-результаты. Важно выбрать такое решение для предиктивной аналитики, которое отвечает конкретным потребностям разных пользователей, с разными навыками — и новичков, и опытных аналитиков и статистиков.

С помощью программного обеспечения для предиктивной аналитики вы сможете:

- Преобразовать данных в точные прогнозы, что позволит принимать аргументированные решения и предпринимать нужные действия.
- Прогнозировать потребности для повышения прибыльности и лояльности клиентов.
- Максимизировать производительность сотрудников и процессов.
- Увеличить ценность ваших данных.
- Находить и избегать риски утечек и мошенничества, прежде чем они реализуются
- Выполнять статистический анализ, включая регрессионный анализ, классификацию и кластерный анализ.
- Измерять влияние социальных медиа на ваши продукты, услуги и маркетинговые кампании.

О TIBCO Spotfire

TIBCO Spotfire - это аналитическое решение от лидера в области инфраструктурной и бизнес-аналитики TIBCO Software. Благодаря интуитивно понятному программному обеспечению Spotfire от интерактивных панелей мониторинга и обнаружения данных до предиктивных приложений и аналитики в режиме реального времени обеспечивается удивительно быстрая и гибкая среда для визуализации и анализа ваших данных.

Так как компании нуждаются в расширении аналитических возможностей, вы можете с легкостью интегрировать решения от TIBCO Software и использовать расширенный аналитический функционал, чтобы принимать аргументированные решения и предпринимать необходимые действия.

TIBCO Spotfire имеет долгую и богатую историю в области предиктивной аналитики. Со Spotfire вы можете разрабатывать собственные модели и использовать свои разработки на R, S+, SAS, MATLAB, проводить анализ больших данных внутри базы данных, как Teradata Aster и другие источники. TERR является статистическим пакетом корпоративного уровня, который был разработан компанией TIBCO для полной совместимости с языком R, реализуя многолетний опыт компании в аналитической системе, связанный с S+. Это позволяет заказчикам продолжать разработку приложений и моделей не только используя открытый код R, но и интегрировать и развернуть свой код R на коммерческой надежной платформе без необходимости переписывать свой код.

TIBCO регулярно вносит свой вклад в сообщество R, включая обратную связь с командой R Core, и предлагает широкую совместимость с R и растущим количеством пакетов CRAN (в настоящее время 1800+).



**УЗНАЙТЕ БОЛЬШЕ
О ПЛАТФОРМЕ TIBCO**

www.spotfiretibco.ru
www.tern.ru

Москва
+7 (495) 234-9885
sales@tern.ru

Пермь
+7 (342) 219-7963
sales@tern.perm.ru