

image not found or type unknown



Сегодня мы являемся свидетелями активного развития технологии интеллектуального анализа данных (ИАД или data mining), появление которой связано, в первую очередь, с необходимостью аналитической обработки сверхбольших объемов информации, накапливаемой в современных хранилищах данных. Возможность использования хорошо известных методов математической статистики и машинного обучения для решения задач подобного рода открыло новые возможности перед аналитиками, исследователями, а также теми, кто принимает решения - менеджерами и руководителями компаний. Сложность и разнообразие методов ИАД требуют создания специализированных средств конечного пользователя для решения типовых задач анализа информации в конкретных областях. Поскольку эти средства используются в составе сложных многофункциональных систем поддержки принятия решений, они должны легко интегрироваться в подобные системы. Одним из наиболее важных и перспективных направлений применения ИАД являются бизнес-приложения, поэтому опыт канадско-американской фирмы Cognos по реализации методов ИАД в составе интегрированных интеллектуальных систем поддержки принятия решений представляет интерес как для разработчиков, так и для пользователей.

Системы ИАД применяются в научных исследованиях и образовании, в работе правоохранительных органов, производстве, здравоохранении и многих других областях. Особенно широко технология ИАД используется в деловых приложениях. В таблице 1 приведено несколько примеров приложений в области финансов и маркетингового анализа.

Таблица 1. Примеры использования методов интеллектуального анализа данных в финансовых приложениях и маркетинговом анализе

Приложение (организация)	Описание
-------------------------------------	-----------------

FALCON (HNC Software, Inc.)	Инструментальное средство для оперативного выявления злоупотреблений с кредитными карточками; более 100 организаций-пользователей отмечают сокращение числа нарушений на 20-30%.
Классификатор дебиторских счетов (Internal Revenue Service)	Выявление счетов потенциально платежеспособных дебиторов на основе анализа больших объемов архивных данных по уплате налогов.
Повышение качества архивной финансовой информации (Lockheed)	Выявление закономерностей (в виде правил вывода) в архивных финансовых данных для использования в моделях прогнозирования, системах поддержки принятия решений по инвестированию и т.д.
Верификация данных по курсам валют (Reuters)	Система выявления ошибок в оперативно поступающих данных по курсам валют. С помощью нейронных сетей и индуктивного вывода правил строятся приблизительные прогнозы, которые сравниваются с поступающими данными. Большие отклонения рассматриваются как возможные ошибки.
Прогнозирование невыплат в сделках с недвижимостью (Leeds)	Анализ архивных данных по сделкам с недвижимостью и выявление паттернов, соответствующих проблемным сделкам, заканчивающимся невыплатами. Выявленные закономерности используются для оценки риска при заключении новых сделок.
Маркетинговые исследования (Dickinson Direct)	Определение характеристик типичных покупателей продукции компании для выявления новых потенциальных клиентов (профилирование клиентов).

Маркетинговые исследования (Reader's Digest Canada)	Выявление основных сегментов рынка и наиболее благоприятных подмножеств, а также исследование зависимостей между основными показателями и характеристиками сегментов.
Установка лотерейных автоматов (Automated Wagering, Inc.)	Объединение методов ИАД с географическим анализом для определения наилучших мест для установки лотерейных автоматов в штате Флорида.
Выявление потенциальных покупателей автомобильных стерео систем (Washington Auto Audio, Inc.)	Анализ демографической базы данных, содержащей информацию о 14000 реальных и потенциальных клиентов, позволил за 90 секунд получить 3 довольно надежных индикатора для прогноза спроса на продукцию и услуги компании. Аналогичные результаты были получены в результате традиционного исследования, выполненного одной из консалтинговых компаний, причем это исследование обошлось фирме на порядок дороже, чем автоматизированная система интеллектуального анализа данных.

Как видно из таблицы 1, результаты ИАД представляют большую ценность для руководителей и аналитиков в их повседневной деятельности. Поэтому перед разработчиками встает задача внедрения технологии ИАД в автоматизированные системы анализа и поддержки принятия решений. В отличие от исследовательских приложений, подобные системы должны ориентироваться на непрограммирующего пользователя и на решение конкретных проблем, которые оказываются довольно разнообразными и нередко требуют применения различных подходов. Например, от автоматизированной системы анализа финансового состояния коммерческого банка может потребоваться решение таких задач как:

- получение отчетности банка и проверка ее полноты и корректности;
- проведение группировки статей баланса и расчет экономических нормативов и аналитических коэффициентов (например, по методу CAMEL);
- проверка экономических нормативов по методике ЦБ РФ;
- оценка состояния банка по системе аналитических коэффициентов;

- определение рейтинга банка;
- анализ динамики основных показателей, выявление тенденций и прогнозирование состояния банка;
- анализ степени влияния тех или иных факторов на состояние банка;
- выработка рекомендаций по оптимизации банковского баланса.

Эта система должна обеспечивать решение нескольких задач, от сбора и проверки корректности информации, поступающей в базу данных, до традиционного и/или интеллектуального анализа данных (степень влияния различных факторов на состояние банка, прогнозирование) и оптимизационного анализа, причем от пользователя не требуется специальных знаний в области баз данных, факторного анализа или методов оптимизации.

К сожалению, универсальные средства ИАД довольно сложны и дороги, поэтому они не могут широко применяться в рамках интегрированных систем, ориентированных на конечного пользователя. Действительно, в основу технологии ИАД положен не один, а несколько принципиально различных подходов (таблица 2), причем использование некоторых из них невозможно без специальной подготовки. Выбор подхода нередко требует привлечения специалиста по ИАД.

Таблица 2. Основные технологии интеллектуального анализа данных

Технология	Достоинства	Недостатки
Правила вывода	Удобны в тех случаях, когда данные связаны отношениями, представимыми в виде правил "если то "	При большом количестве правил теряется наглядность; не всегда удастся выделить отношения "если то "
Нейронные сети	Удобны при работе с нелинейными зависимостями, зашумленными и неполными данными	"Черный ящик": модель не может объяснить выявленные знания; данные обязательно должны быть преобразованы к числовому виду

Нечеткая логика	Ранжируют данные по степени близости к желаемым результатам; нечеткий поиск в базах данных	Технология новая, поэтому сейчас известно ограниченное число специализированных приложений
Визуализация	Многомерное графическое представление данных, по которому пользователь сам выявляет закономерности - образцы и отношения между данными	Модели не исполняются, и их интерпретация полностью зависит от аналитика
Статистика	Есть большое число алгоритмов и опыт их применения в научных и инженерных приложениях	Больше подходят для проверки гипотез, чем для выявления новых закономерностей в данных
К-ближайший сосед	Выявление кластеров, обработка целостных источников данных	Большие затраты памяти, проблемы с чувствительностью
Интегрированные технологии	Можно выбирать подходы, адекватные задачам, или сравнивать результаты применения разных подходов.	Сложные средства поддержки; высокая стоимость; для каждой отдельно взятой технологии не всегда реализуется наилучшее решение

Однако вряд ли есть смысл использовать универсальные средства ИАД во всех случаях, поскольку накопленный опыт работы с методами интеллектуального анализа уже позволил выделить типовые задачи и определить наиболее эффективные методы их решения. Существующие системы ИАД можно

подразделить на исследовательские, ориентированные на специалистов и предназначенные для работы с новыми типами проблем, и прикладные, рассчитанные на непрограммирующих пользователей (аналитиков, менеджеров, технологов и т.д.) и решающие типовые задачи. Если в исследовательских системах ИАД важно разнообразие доступных методов обработки данных и гибкость используемых средств, то в прикладных системах такие широкие возможности оказываются излишними. В подобных системах целесообразно реализовывать не методы, а типовые виды рассуждений (анализа), характерные для проблемной области. Создание соответствующих инструментальных средств с дружественным интерфейсом обеспечивает доступность ИАД для широкого круга пользователей.

Средства создания интеллектуальных приложений

Среди инструментальных средств создания интеллектуальных приложений для бизнеса определенным интересом представляет семейство программных продуктов Business Intelligence (BI) компании Cognos. Системы Impromptu, PowerPlay, Scenario и 4Thought представляют собой взаимосвязанные и дополняющие друг друга инструментальные средства, поддерживающие наиболее эффективные технологии обработки данных и обеспечивающие решение широкого круга задач в бизнес-приложениях, от доступа к информации в распределенных базах данных до вычислительной обработки и интеллектуального анализа

Отличительной особенностью рассматриваемых средств является сочетание эффективности реализуемых в них методов с дружественным интерфейсом, что делает их доступными для непрограммирующих пользователей, а предлагаемая программа обучения позволяет осваивать каждое средство не более чем за 2 дня.

По оценкам Gartner Group, META Group, Lehman Bros., Seybold, Forrester и Hurwitz, продукция Cognos пользуется сегодня повышенным спросом - среди покупателей программных продуктов семейства BI, установивших у себя свыше 2000 лицензий, такие компании как Procter & Gamble, Chrysler, Dutch Telecom, Swedish Telecom и др. Программные продукты семейства BI обеспечивают создание корпоративных приложений в архитектуре "клиент-сервер". К достоинствам этих продуктов можно отнести объектно-ориентированную архитектуру средств администрирования, возможности контроля доступа к информации, поддержку технологии OLE в среде Windows. Дополнительные возможности по интеграции систем предоставляются за счет использования встроенного макроязыка, позволяющего описывать взаимодействие между различными приложениями и обмен информацией между

ними.

Система **Impromptu** обеспечивает доступ к базам данных и формирование отчетов. Она ориентирована на непрограммирующих пользователей и имеет графический интерфейс создания сложных отчетов и задания последовательностей вычислительной обработки, причем не требует знания SQL. Простота использования достигается за счет объектно-ориентированного подхода к построению отчетов, а также наличия библиотеки шаблонов стандартных отчетов и утилиты Report Wizard, направляющей пользователя в ходе работы. Отчеты могут быть табличными, графическими и гипертекстовыми (формат HTML). Impromptu поддерживает работу с распространенными СУБД - Oracle, MS SQL Server, Sybase SQL Server, Sybase NetGateway, OmniSQL Gateway, MDI DB2 Gateway, Informix, CA-Ingres, Gupta SQLBase, а также обеспечивает доступ через ODBC. Отчеты Impromptu могут использоваться в качестве источников данных для остальных систем семейства BI, выступая при этом в роли информационных витрин. По итогам опроса InfoWord, проведенного в 1995 г., система Impromptu была признана лучшей в своем классе.

Система **PowerPlay** является средством для проведения OLAP-анализа. Технология OLAP представляет собой развитие методов работы с реляционными базами данных, что позволило существенно повысить эффективность их использования в аналитической работе. Ключевые особенности баз данных OLAP следующие: многомерное представление данных в виде гиперкубов, проведение сложных вычислений над массивами данных и привязка информации ко времени, обеспечивающая возможность анализа динамики данных. Благодаря простоте извлечения информации из баз данных OLAP конечные пользователи стали меньше зависеть от разработчиков: с помощью средств OLAP многие задачи, ранее требовавшие специального программирования, они сами решают. Разработчики, в свою очередь, получили удобное средство для быстрого создания приложений. Выигрывает и компания в целом, поскольку хранение в гиперкубах агрегированной аналитической информации позволяет снижать число запросов к базе данных и уменьшает трафик в сети.

PowerPlay обеспечивает многомерный просмотр данных с нисходящим и уровневый анализом, работу с различными видами дисплеев (таблицами, двумерными и трехмерными графиками, диаграммами). Система позволяет выявлять и отмечать в просматриваемых отчетах исключения и особые случаи, ранжировать данные и проводить вычисления над массивами данных. Гиперкубы строятся на основе информации, поступающей из различных источников (баз данных и отчетов

системы Impromptu), причем возможна обработка более 10 млн. записей. Число размерностей в гиперкубах не ограничивается. По сравнению с известными средствами OLAP, PowerPlay обеспечивает самую высокую степень сжатия данных в гиперкубах. Гиперкубы могут создаваться и размещаться как на серверах, так и на клиентских компьютерах. Новая версия PowerPlay обеспечивает возможность работы с гиперкубами через Web-браузеры.

На основе обобщения опыта применения технологии OLAP была разработана методика ее использования в различных сферах бизнеса. Методика включает анализ 24 типовых приложений в таких областях как финансовая деятельность, работа с кадрами, продажи, маркетинг, закупки, производство, распределение продукции и обслуживание клиентов. Например, анализ объемов продаж и прибылей удобно проводить на основе гиперкубов с такими размерностями как даты, регионы, виды продукции, каналы сбыта, клиенты и др. Аналитик может самостоятельно получать ответы на вопросы вида: "В каких городах во II и III квартале прошлого года продажи витаминных препаратов обеспечили не менее 60% прибыли?" и др. В производстве OLAP-анализ может применяться для анализа причин появления брака и оценки использования ресурсов. Бухгалтерские балансы одного или нескольких предприятий, с привязкой ко времени, также могут быть представлены в виде гиперкуба. Действительно, баланс, по своей сути, является многомерным документом, и анализ десятков подобных документов за период в 24 месяца, представленных в традиционном табличном виде, произвести очень сложно. Агрегирование информации с возможностью нисходящего анализа в различных разрезах существенно упрощает работу аналитика.

Так, в системе анализа отчетности коммерческих банков совокупность балансов за несколько отчетных периодов может быть представлена в виде гиперкуба системы PowerPlay. Подобное представление обеспечивает простую и естественную реализацию метода группировок и нормативного метода. Аналитик, имеющий навыки работы с PowerPlay, может самостоятельно получать оценки состояния банков по системе аналитических коэффициентов (CAMEL), задавать порядок вычисления рейтингов банков, а также выводить динамику основных показателей в графическом виде.

Средства извлечения новой информации

Если первые две системы семейства BI предназначены для обеспечения доступа к информации, уже имеющейся в базе данных, то следующие программные продукты представляют собой собственно средства извлечения новой информации - ИАД и

позволяют получать знания, явно не содержащиеся в исходных данных [1,2]. Оба средства доступны на платформе Windows 95/NT и рассчитаны на использование непосредственно менеджерами и аналитиками как крупных организаций, так и относительно небольших компаний. Поскольку системы основаны на различных методах ИАД (простейшие статистические методы - в Scenario, нейронные сети - в 4Thought), в распоряжение пользователей предоставляется набор средств, для каждого из которых определена область эффективного применения.

Система **Scenario** предназначена для выявления взаимосвязей в данных. Анализируемая информация представляется в виде таблиц, столбцам которых соответствуют факторы (переменные), описывающие предметную область, причем один из этих столбцов отмечается пользователем как целевой показатель. Имея точное представление о том, в какой степени отдельные факторы (тип клиента, канал распространения продукции, затраты на рекламу и др.) влияют на целевой показатель (объем продаж), менеджеры могут оптимизировать решения по размещению ресурсов для обеспечения максимальной прибыли. Scenario фокусирует внимание аналитиков на измеряемых показателях и факторах, что позволяет им переходить от качественных рассуждений к построению более точных количественных моделей.

В Scenario реализовано автоматическое выявление зависимостей по критерию хи-квадрат (метод CHAID) для нахождения однородных групп (сегментов) данных с "аналогичным" поведением относительно целевого показателя. Система использует методы деревьев решений и строит правила классификации, которые можно применять к новым данным для построения прогнозов относительно возможных значений целевого показателя. Хотя ядром системы являются методы статистики, от пользователя не требуется специальных знаний в этой области. Основная статистическая информация отображается средствами интуитивно понятной графики, а более детальные сведения запрашиваются пользователем в случае необходимости. Сами методы ИАД являются лишь основой для реализации типовых видов анализа, используемых аналитиками и лицами, принимающими решения в бизнесе. К этим видам анализа относятся:

1. **Ранжирование** - упорядочивание факторов по степени их влияния на целевой показатель. С каждым фактором связывается весовой коэффициент, дающий количественную оценку степени влияния.
2. **Сегментация** - разделение области значений фактора на сегменты для проведения дальнейшего нисходящего анализа. Например, успешные показатели по продажам в целом могут скрыть от исследователя неудачи в

некоторых регионах. Проведение сегментации по географическим районам позволит выявить, что успешные продажи имели место только в двух регионах, А и В. Последующий нисходящий анализ может показать, что объем продаж в регионе А зависит от затрат на рекламу, а в регионе В на него влияют главным образом сезонные колебания.

3. **Профилирование наилучших достижений** - выявление основных характеристик наиболее успешных регионов, филиалов, клиентов и т.д. для планирования дальнейшей деятельности компании. Данный вид анализа также используется в банках и страховых компаниях для оценки риска при выдаче кредитов.
4. **Выявление ассоциаций** - поиск ассоциированных групп значений факторов, т.е. значений, почти всегда появляющихся вместе (например, продажа ПК в большинстве случаев сопровождается продажей источника бесперебойного питания). Этот вид анализа в первую очередь необходим при планировании продаж, усовершенствовании работы складов, проведении рекламных компаний.
5. **Выявление исключений** - поиск элементов, выпадающих из общей картины. Появление подобных элементов может быть вызвано как ошибками в данных, которые следует исправить, так и необычными ситуациями в работе компании, требующими немедленного вмешательства со стороны руководства.

Итак, с помощью Scenario можно получать ответы на многие вопросы, необходимые для принятия обоснованных решений.

- Кто из клиентов компании останется в следующем году, кто откажется от ее услуг ?
- Какие новые тенденции в спросе определяют деятельность компании?
- Что отличает лучших служащих компании (наиболее выгодные продукты, самые успешные филиалы) от остальных служащих (продуктов, филиалов) ?
- Каковы признаки/индикаторы неудач (злоупотреблений) в деятельности филиалов ?
- Как определить наилучший район для продвижения нового товара или услуги ?

Поскольку процесс ИАД почти всегда требует значительных затрат времени и вычислительных ресурсов, в Scenario реализованы три стратегии анализа.

1. Для проведения быстрого начального анализа задачи, отсева незначительных факторов и фокусировки внимания на ключевых параметрах рекомендуется

- использовать режим **исследования**. Этот режим позволяет выигрывать во времени за счет некоторого снижения точности результатов. Поскольку в дальнейшем предполагается проведение более детального анализа, требования к надежности не выдвигаются. Размер выборки не ограничивается.
2. Более точный анализ может быть проведен в режиме **тестирования**. Использование этого режима обеспечивает очень высокую точность и достаточную надежность результатов. Для его проведения требуется не менее 300 записей в базе данных.
 3. Наилучшие по точности и надежности результаты можно получить в режиме **верификации**. Этот режим, требующий не менее 1000 записей, рекомендуется использовать для проверки и уточнения результатов, полученных в ходе тестирования.

Результаты исследований могут представляться в виде графиков и деревьев решений.

График позволяет просматривать информацию о зависимости целевого показателя от того или иного фактора. На график выводятся описания сегментов данных, полученных путем разбиения значений исследуемого фактора, и вычисляются средние для целевого показателя по каждому сегменту. Пользователь может выбирать любой фактор для просмотра зависимости, группировать или разгруппировывать любые факторы, а кроме того проводить нисходящий анализ для любого сегмента данных. Может оказаться, что наибольшее влияние на продажи в этом магазине оказывают сезонные колебания, в то время как объем продаж в магазине А в значительной мере определяется ценой продукции.

Деревья решений разбивают исходную совокупность данных на группы (сегменты) с точки зрения значений целевого показателя. В приведенном примере все данные, попавшие в левый сегмент второго уровня, обеспечивают относительно небольшую прибыль. Выявленная закономерность состоит в том, что величина прибыли в значительной мере определяется затратами на рекламу: чем они выше, тем больше прибыль. Деревья решений генерируют иерархии правил вида "если ... то", позволяющих классифицировать данные в соответствии с их влиянием на целевой показатель.

Для просмотра больших деревьев удобно использовать фильтры, позволяющие получать подробную информацию только по тем сегментам данных, которые интересуют пользователя. Например, в дереве решений подробная информация дана по сегментам, включающим такие каналы сбыта как Магазин А, Магазин В и

Спортивные клубы. С помощью фильтров пользователь может также ранжировать сегменты по степени их влияния на целевой показатель и выделять подсветкой N наиболее значимых сегментов.

Ознакомление с результатами анализа обычно требует просмотра большого количества сегментов данных. Чтобы пользователь мог собирать интересующую его информацию в одном месте, в Scenario реализован механизм закладок. Статистическая информация по каждому из сегментов, на котором пользователь оставляет закладку, выводится системой в окно Закладки.

Выбор данных для анализа осуществляется с помощью специальной подсистемы Import Wizard из файлов в формате dBase, электронных таблиц Excel, текстовых файлов, а также отчетов Impromptu.

Scenario дополняет и расширяет возможности PowerPlay. Во-первых, система позволяет фокусировать внимание пользователя, осуществляющего OLAP-анализ, на наиболее значимых факторах, предварительно выявляемых средствами ИАД. Во-вторых, результаты ИАД могут помочь в формировании гиперкубов, поскольку они определяют, какие размерности представляют интерес для исследования и как сегментируются данные.

Как и PowerPlay, Scenario удобно использовать для анализа продаж, проведения маркетинговых исследований, а также проведения ИАД в области финансов, при работе с кадрами и в производстве. По мнению аналитиков, благодаря Scenario можно получать более содержательные результаты. Отмечаются также возможности средств визуализации результатов исследований, встроенные виды анализа и удачная документация - все это делает Scenario удобным инструментом непрограммирующего пользователя. Недостатком текущей версии системы является невозможность использования нечисловых полей в качестве целевых показателей.

В тех случаях, когда представление извлекаемых знаний в виде набора дискретных правил "если ... то" не удовлетворяет пользователя, он может применить методы нейронных сетей, реализованные в системе **4Thought** и позволяющие:

- моделировать сложные нелинейные зависимости между факторами и целевыми показателями;
- выявлять тенденции в данных (при наличии временных рядов) для построения прогнозов;
- работать с неполными и зашумленными данными;

- получать содержательные результаты при относительно небольшом объеме исходной информации с возможностью последующего усовершенствования модели по мере поступления новых данных.
- Типовыми видами анализа в системе 4Thought являются:
- выявление значимых факторов и получение оценок степени влияния каждого фактора на целевой показатель;
- проверка гипотез "что если" - каково будет значение целевого показателя при некотором гипотетическом наборе значений факторов, каков вклад каждого фактора в это значение;
- прогнозирование значений целевого показателя.

Процесс моделирования средствами 4Thought достаточно прозрачен, а источники данных для анализа такие же, как и в системе Scenario. Анализируемые данные представляются в виде электронной таблицы, столбцам которой соответствуют атрибуты из таблиц базы данных, а строкам - записи. При подготовке данных для анализа пользователь может редактировать таблицы, а также включать в них вычисляемые столбцы. Как и в Scenario, для постановки задачи моделирования должны быть указаны целевой (моделируемый) показатель и переменные-факторы.

Данные в таблице рассматриваются системой как примеры для обучения нейронной сети. Совокупность всех данных разбивается на два подмножества - модельное и тестовое. На каждом цикле обучения сначала сеть обучается на модельном подмножестве, а затем корректность ее работы проверяется на тестовых данных. Обучение продолжается до тех пор, пока не будет достигнуто предельно допустимое число итераций (5000) или пользователь не остановит процесс. Очевидно, что обучение сети целесообразно проводить до тех пор, пока точность результатов сети на модельном и на тестовом подмножествах растет. Убывание точности на тестовом подмножестве означает переобучение сети и снижение качества ее работы, даже при продолжающемся росте точности на модельных данных. Система запоминает состояние сети на момент последней итерации с максимальным значением точности на тестовом подмножестве, и использует именно его в рабочей модели (отметим, что здесь возможны локальные максимумы). Как правило, для обучения оказывается достаточно от 10 до 1000 итераций. При работе с небольшими объемами данных, когда выбор тестового подмножества существенно влияет на результаты обучения, пользователь может провести обучение в несколько циклов, выбирая для каждого цикла новое разбиение на тестовые и модельные данные. В процессе обучения пользователю

выдается табло, на которое выводятся данные о текущем состоянии сети (точность на модельном и на тестовом подмножестве, количество итераций, скорость обучения, итерация с наилучшим показателем точности на тестовых данных) в числовом и графическом виде.

По результатам моделирования могут быть получены следующие отчеты.

1. Основная статистика с комментариями системы. Указываются: точность модельных и тестовых данных; среднеквадратичная ошибка, стандартное отклонение, границы 95% доверительного интервала, среднее абсолютной ошибки, количество записей, количество переменных, оценки влияния факторов на целевой атрибут (аналоги t-статистик) и заключение системы относительно качества полученной модели.
2. Количественные зависимости целевого показателя от каждого фактора в графическом виде и приближенное описание модели с помощью уравнения линейной регрессии.
3. Коэффициенты корреляции для любого выбранного фактора относительно других факторов.
4. Результаты использования модели на любых задаваемых пользователем значениях факторов.
5. Объяснение для любого значения моделируемой переменной, вычисленной системой - как она была получена и какой вклад внес каждый фактор в это значение.