

Практическая работа 4

Построение нейронной сети в Deductor Studio

Бизнес-приложения интеллектуального анализа данных (BI-платформы) сегодня имеют мало общего с некогда популярными системами генерации отчетов, порождавшими "горы" распечаток, в которых было практически нереально отыскать нужную информацию.

Появление хранилищ данных открыло путь к извлечению и интеграции сведений, выдаваемых системами поддержки текущих операций, к их очистке и фильтрации с учетом бизнес-задач. К тому же накопление данных за определенные периоды позволило проводить их ретроспективный анализ для выявления трендов и скрытых закономерностей в длинных временных рядах.

В настоящее время организации стремятся приобретать интеллектуальное ПО для корпоративной информационной среды, которое функционирует поверх общей BI-платформы, легко настраивается на решение различных бизнес-задач и, главное, обеспечивает детальный анализ данных с учетом специфики конкретной отрасли.

Последнее требование является особенно важным и выгодно отличает специализированные аналитические системы от решений общего назначения. На примере рынка услуг связи этот тезис можно проиллюстрировать особенно наглядно.

Как известно, специфика телекоммуникационной отрасли заключается в обслуживании многочисленной абонентской базы, однако из-за постоянно обостряющейся конкуренции это не исключает необходимости бороться буквально за каждого высокоприбыльного клиента. Телекоммуникационный рынок характеризуется чрезвычайно быстрыми темпами развития технологий, вследствие чего для модернизации или расширения инфраструктуры, а также для создания и продвижения новых услуг от оператора требуются крупные инвестиции. Одновременно значительно возрастает риск потери вложенных средств, поскольку, как показывает практика последних лет, оценки

грядущего потребительского спроса частенько оказываются непомерно завышенными. Кроме того, практически все операторы страдают от значительной "текучности" абонентов, которых привлекают либо новые услуги, либо более выгодные тарифные планы конкурентов.

Современная бизнес-среда уже не позволяет телекоммуникационным компаниям надеяться на значительный рост прибыли за счет революционных технологических достижений или резкого всплеска активности пользователей. Его можно достичь лишь при внедрении эффективных схем удержания существующих и привлечения новых абонентов, извлечения максимальной выгоды из отношений с каждым клиентом, разработки динамичных бизнес-процессов, проведения высокорезультативных маркетинговых кампаний.

Один из путей достижения этих целей заключается в извлечении всей полезной информации из имеющихся у телекоммуникационных фирм данных об абонентах. Не секрет, что по объему и степени детализации сведений о клиентах операторы связи уступают лишь банкам и страховым компаниям. Проблема, однако, заключается в том, что обычно такая информация поступает из различных источников и распределяется по множеству хранилищ и баз данных. В этой ситуации какие-то данные дублируют друг друга, оказываются устаревшими и т. д., поэтому анализу подобных сведений должны предшествовать процедуры их извлечения из разных источников, очистки, фильтрации и преобразования к общему формату для записи в единое информационное хранилище. Кроме того, необходимо, чтобы результаты анализа накопленной оператором информации были доступны для сотрудников компании в нужное время и по разным коммуникационным каналам (через электронную почту, корпоративный портал, интрасеть и т. п.).

В решении всего комплекса перечисленных задач на помощь приходят системы бизнес-аналитики.

В качестве BI-решения для телекоммуникационных компаний используется аналитическая платформа Deductor, которая состоит из

нескольких интегрированных программных компонентов, поддерживающих разработку хранилищ данных и В 1-систем, а также управление ими.

В Deductor Studio реализованы механизмы трансформации, такие, как преобразование к скользящему окну, квантование, группировка, сортировка, приведение типов.

В случае прогнозирования необходимо решать задачу регрессии. В Deductor Studio имеются механизмы построения как линейных моделей (линейная регрессия), так и нелинейных (нейронные сети).

Из общеизвестных преимуществ нейросетевого подхода следует выделить одно, самое привлекательное в нем - отсутствие необходимости в строгой математической спецификации модели, что особенно ценно при анализе плохо формализуемых процессов. А большинство бизнес-задач плохо формализуется. Это означает, что, при наличии достаточно развитых и удобных инструментальных программных средств, пользователь может при построении модели прогнозируемого процесса руководствоваться такими понятиями, как опыт и интуиция. После получения прогностической модели можно получить, собственно, сам прогноз. Для этого можно, например, воспользоваться визуализатором «что-если».

В случае построения авторегрессионной модели, когда значения временного ряда зависят от предыдущих периодов, можно воспользоваться специальным обработчиком «прогнозирование» и, построив модель, прогнозирующую на один шаг, получить результат на сколько угодно шагов вперед. Результаты прогнозирования можно экспортировать на сторону для использования в программах планирования, учетных системах и прочих.

Deductor был разработан с учетом требований, предъявляемых к аналитическим системам. Среди них - возможности сокращения издержек и управления рисками, уменьшение оттока и удержание наиболее ценных клиентов, увеличение размера доходов, получаемых с одного клиента, повышение эффективности использования сетевых ресурсов, выявление

случаев мошенничества и управление отношениями с поставщиками дополнительных услуг.

Аналитические модели данных и процессов используются Deductor Studio для ранжирования данных, предсказания поведения клиентов и их сегментации. Они отражают процедуры аналитической обработки, специфические для конкретной отрасли, и охватывают такие области, как оценка кредитоспособности существующих и новых клиентов, предсказание оттока абонентов фиксированной и мобильной связи, сегментация клиентов и анализ рыночной корзины (например, типов звонков для организации "связанных" продаж и целевого маркетинга). Результаты применения аналитических моделей могут быть помещены в хранилище детальных данных либо использоваться системами поддержки текущих операций.

Гибридные OLAP-модели данных представляют собой многомерные серверные базы данных (иногда называемые кубами), в которые помещаются итоговые данные. Тесная интеграция этих БД с другими компонентами решения Deductor Studio позволяет пользователю оперативно перейти к работе с информацией, содержащейся в хранилище детальных данных либо представленной в моделях-схемах.

Приложение Deductor Studio осуществляет сбор данных из различных источников, имеющихся у оператора услуг связи, а затем преобразует их в информацию, которая необходима для оперативного реагирования на изменяющиеся условия бизнес-среды, для принятия стратегических решений и доведения их до сотрудников. Присутствующие в нем средства доступа к данным из любого источника, управления ими и их систематизации повышают ценность унаследованных систем, а значит, обеспечивают сохранение ранее сделанных инвестиций. Данный инструментарий позволяет гарантировать, что информация, используемая для принятия стратегических решений, является достоверной, адекватна бизнес-задачам компании и своевременно попадет к тем руководителям и сотрудникам, которым она действительно необходима.

Если верить оценкам аналитиков, ежегодный отток клиентов в телекоммуникационном секторе может достигать до 30%. Миграция абонентов от одной компании к другой самым негативным образом отражается на финансовых показателях операторов связи. К тому же не стоит забывать, что затраты на привлечение новых клиентов обычно в несколько раз превышают расходы на удержание уже существующих.

Большинство телекоммуникационных компаний кровно заинтересованы в сохранении существующих абонентов, в увеличении дохода от каждого из них и в повышении рентабельности своих отношений с абонентами. Для достижения данных целей операторы должны увеличить эффективность своих маркетинговых кампаний, сделать их более адресными и добиться того, чтобы предложение новых продуктов было дифференцированным как с точки зрения категорий клиентов, так и во времени. Первый шаг к реализации этого сценария заключается в проведении точной сегментации клиентов, а также в установлении многогранных и гибких отношений с клиентами с использованием различных коммуникационных каналов.

Deductor Studio интегрирует мощные функции анализа данных о клиентах и предварительно созданные аналитические модели со средствами анализа маркетинговой деятельности. Такая интеграция дает возможность быстро определить целевую аудиторию для конкретных маркетинговых мероприятий и сформулировать правила, на основе которых следует проводить сегментацию клиентов. У оператора появляется инструментарий для динамической обработки результатов рекламных кампаний или иных маркетинговых акций, позволяющий автоматически обновлять данные о контактах с клиентами и отслеживать их реакцию на маркетинговые мероприятия, будь то решения о приобретении тех или иных продуктов либо трудно выявляемые иными способами изменения в поведении.

Описываемое ПО позволяет использовать предварительно заданные планы (шаблоны) рекламных кампаний, учитывающие типичные

маркетинговые задачи операторов связи, пользоваться функциями детального управления рекламными кампаниями, планировать разносторонние, многоканальные и многоэтапные кампании и выполнять их приоритизацию, сводить воедино и фильтровать списки потенциальных клиентов, полученные из различных источников, оптимизировать и постоянно координировать связи с сотнями тысяч и даже миллионами клиентов по различным коммуникационным каналам.

Deductor Studio предназначен для сегментации клиентов, выявления трендов в их поведении и прогнозирования оттока, а также для выработки мер, призванных уменьшить или даже предотвратить миграцию наиболее ценных пользователей.

Переход абонентов от одного оператора к другому является неизбежным следствием предложения новых типов услуг и ценовых войн, время от времени разгорающихся на рынке услуг связи. Поскольку ослабления данного процесса ожидать не приходится, а уровень рентабельности телекоммуникационных компаний (ARPU) продолжает снижаться, операторам необходимо четко определить для себя, с какими клиентами можно расстаться без особого ущерба для бизнеса, а каких надо сохранить. В этой второй категории особенно важно своевременно выявить потенциальных кандидатов на уход к конкурентам. Принципиальным моментом здесь является определение причин такого ухода, поскольку лишь знание мотивации позволяет своевременно разработать эффективную стратегию удержания наиболее ценных клиентов.

На решение этой задачи как раз и ориентирована АП Deductor Studio. Используя аналитические процедуры и модели процессов, с его помощью можно точно определить потенциальных "беглецов", способных разорвать отношения с оператором в ближайшее время. Данные для анализа собираются из различных источников. Это сведения об оплаченных и просроченных счетах, о подписке на те или иные виды услуг, о характере сделанных звонков, о процедурах обслуживания клиентов и т. д. Основой описываемого

компонента являются модели и процедуры классификации клиентов и механизмы их реализации, готовые к внедрению и допускающие гибкую настройку под особенности бизнеса конкретной компании.

Deductor Studio. Его отличительной особенностью является способность классифицировать клиентов не только на основании простых демографических параметров или базовых показателей использования ими сетевых сервисов (как это делает подавляющее большинство операторов), но и с точки зрения динамики их поведения и смены продуктовых предпочтений, изменения уровня жизни, параметров звонков либо иных используемых сервисов, кредитных рисков, лояльности к данному оператору, текущей и прогнозируемой прибыльности. Только опора на множество разнообразных показателей позволяет построить детальный профиль клиента, а значит, повысить эффективность взаимодействия с ним и отдачу от маркетинговых кампаний, персонализировать предложения новых телекоммуникационных услуг и, в конечном счете, увеличить приносимую клиентом прибыль.

В Deductor Studio, для классификации клиентов может быть использован весь объем данных, накопленных оператором и обычно распределенных по разным источникам.

Эффективная организация продаж принципиально новых или дополнительных услуг - одно из необходимых условий выживания компании на фоне обостряющейся конкурентной борьбы. И опять же она возможна только при наличии надежных моделей детальной сегментации клиентов. Попытка предложить новые или более дорогие услуги абонентам, традиционно выбирающим самые экономичные тарифы, заранее обречена на провал, а средства, затраченные на продвижение услуг среди этой категории клиентов, окажутся выброшенными на ветер.

Поэтому первый шаг к организации "связанных" продаж должен состоять в выявлении пользователей, традиционно подписывающихся на несколько различных сервисов, стремящихся приобретать самое новое абонентское оборудование, активно пользующихся базовыми услугами,

имеющих подходящий демографический и социальный профиль. После того как эта группа пользователей четко определена, можно попытаться выявить другие категории клиентов, которые потенциально готовы к новым приобретениям, будь то услуги или абонентские устройства. Очевидно, что продвижение новых продуктов среди пользователей таких категорий должно вестись более экономичными способами, чем в первой категории.

На основе предварительно разработанных аналитических моделей Deductor Studio дает возможность спрогнозировать, какие продукты заинтересуют каждого из клиентов в будущем. Для этого, в частности, используются прогностическая модель рыночной корзины, средства анализа и генерации отчетов о лояльности существующих клиентов и возможностях ее повышения, об эффективности использования различных каналов распространения информации о новых продуктах.

Deductor Studio обеспечивает повышение конкурентоспособности телекоммуникационной компании несколько с иной стороны. Он позволяет точно оценивать кредитоспособность существующих и потенциальных клиентов и на основании этого анализа снижать бизнес-риски и издержки на списание безнадежных долгов. Его применение дает возможность установить, какие предложения услуг и методы оплаты наиболее приемлемы для той или иной категории клиентов, какие из потенциальных клиентов являются самыми перспективными, а какие должны быть отнесены к группе наивысшего риска. При этом модели, применяемые к существующим клиентам, опираются на скрупулезный анализ их поведения, сведения об оплаченных и неоплаченных счетах и предпочитаемых клиентами методах расчетов, данные кредитной истории и т. д. При построении предварительной классификации потенциальных клиентов используются модели прогнозирования кредитоспособности.

Таким образом, Deductor Studio содержит полный набор механизмов импорта, обработки, визуализации и экспорта данных для быстрого и эффективного анализа информации. Реализованные в Deductor Studio

механизмы позволяют в рамках одного приложения пройти весь цикл анализа данных - получить информацию из произвольного источника, провести необходимую обработку (очистку, трансформацию данных, построение моделей), отобразить полученные результаты наиболее удобным образом (OLAP, таблицы, диаграммы, деревья решений и др.) и экспортировать результаты.

Архитектура системы построена таким образом, что вся работа по анализу данных в Deductor Studio базируется на выполнении следующих действий :

- импорт данных;
- обработка данных;
- визуализация;
- экспорт данных.

Любой набор данных можно визуализировать каким-либо доступным способом или несколькими способами, поскольку визуализация помогает интерпретировать построенные модели.

В Deductor предусмотрены различные способы визуализации данных:

- OLAP. Многомерное представление данных. Любые данные, используемые в программе, можно посмотреть в виде кросс-таблицы и кросс-диаграммы.

- Таблица. Стандартное табличное представление с возможностью фильтрации данных.

- Диаграмма. График изменения любого показателя.

- Диаграмма рассеяния. График отклонения прогнозируемых при помощи модели значений от реальных. Может быть построена только для непрерывных величин и только после использования механизмов построения модели, например, нейросети или линейной регрессии. Используется для визуальной оценки качества построенной модели.

- «Что-если». Таблица и диаграмма. Позволяют «прогонять» через построенную модель любые интересующие пользователя данные и оценить влияние того или иного фактора на результат.

- Обучающая выборка. Набор данных, используемый для построения модели.

- Диаграмма прогноза. Применяется после использования метода обработки - прогнозирование. Прогнозные значения выделяются цветом.

- Граф нейросети. Визуальное отображение обученной нейросети. Отображается структура нейронной сети и значения весов;

- Дерево решений. Отображение дерева решений, полученного при помощи соответствующего алгоритма.

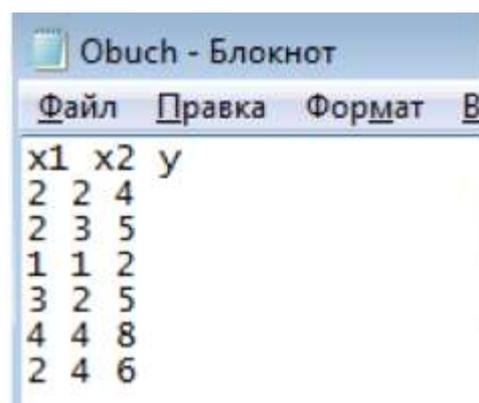
- Правила. Отображает в текстовом виде правила, полученные при помощи алгоритма построения деревьев решений или поиска ассоциаций.

- Карта Кохонена. Отображение карт, построенных при помощи соответствующего алгоритма.

- и другие.

Построение НС в системе Deductor Studio происходит в несколько этапов (в примере реализуется НС, суммирующая числа обучающего набора).

Статистическая информация, предназначенная для последующего обучения сети, представляется в формате файла .txt (рис.1). Причем первая строка - наименования столбцов (через пробел), далее обучающее множество (через пробел).



x1	x2	y
2	2	4
2	3	5
1	1	2
3	2	5
4	4	8
2	4	6

Рисунок 1 - Представление обучающего набора данных

Готовый файл с набором обучающих данных загружается в Deductor Studio (Сценарии —> Мастер импорта —> Text —> Далее (Рис. 2).

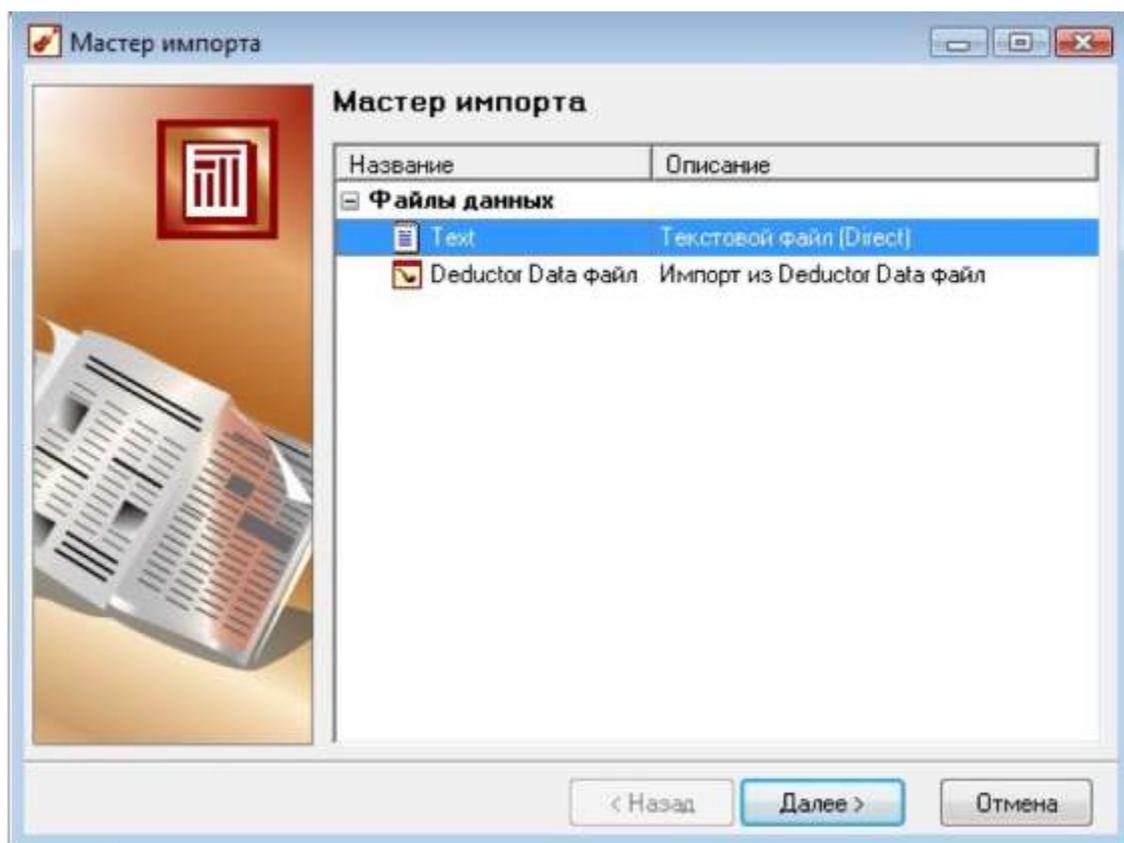


Рисунок 2 - Импорт обучающего набора данных

Выбирается база данных (файл *.txt с содержащейся в нем статистикой), необходимо поставить галочку в пункте «Первая строка является заголовком» (рис. 3).

Далее происходит настройка форматов импорта из файла (рис. 4) - с разделителями.

Затем необходимо настроить параметры импорта файла с разделителями (рис. 5), выбрав пробел - символом- разделителем.

Следующим этапом является задание назначения полей (рис. 6). Слева в окне выбирается поле, справа в поле назначения вводим: для x1 и x2 - входное, для y - выходное. Выбираем вид данных (дискретный или непрерывный) и тип данных (**логический** - данные в поле могут принимать только два значения: Истина или Ложь; **дата/время** - поле содержит данные типа дата/время; **вещественный** - данные в поле представляют собой числа с

плавающей точкой; **целый** - данные в поле представляют собой целые числа; **строковый** - данные в столбце представляют собой строки символов) в зависимости от поставленной задачи.

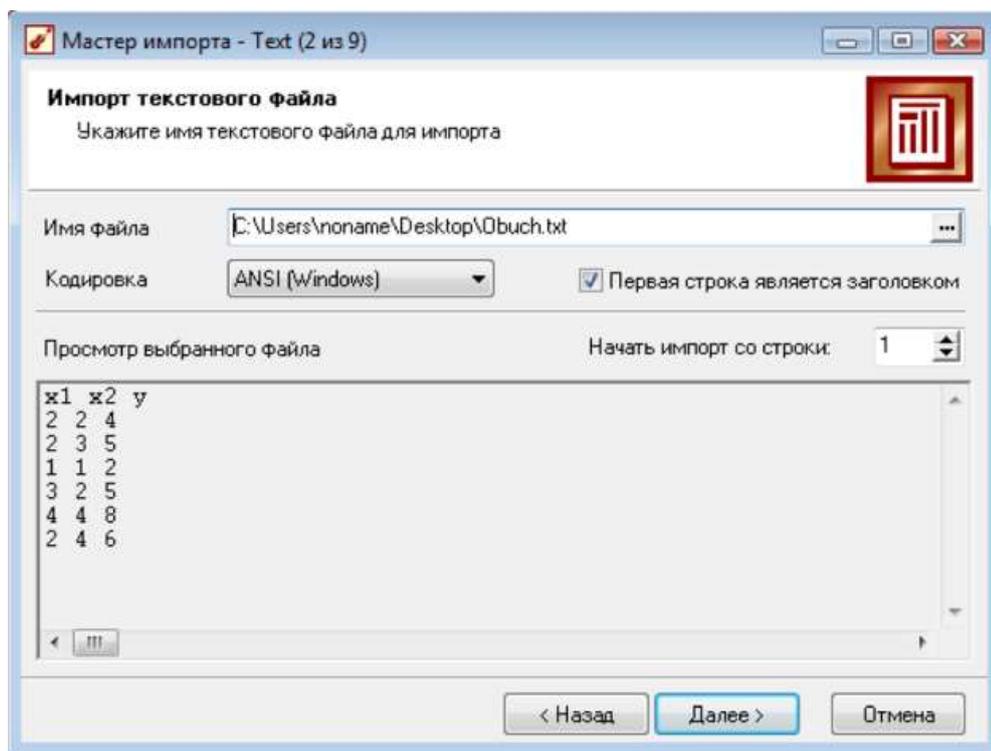


Рисунок 3 - Выбор обучающего набора данных

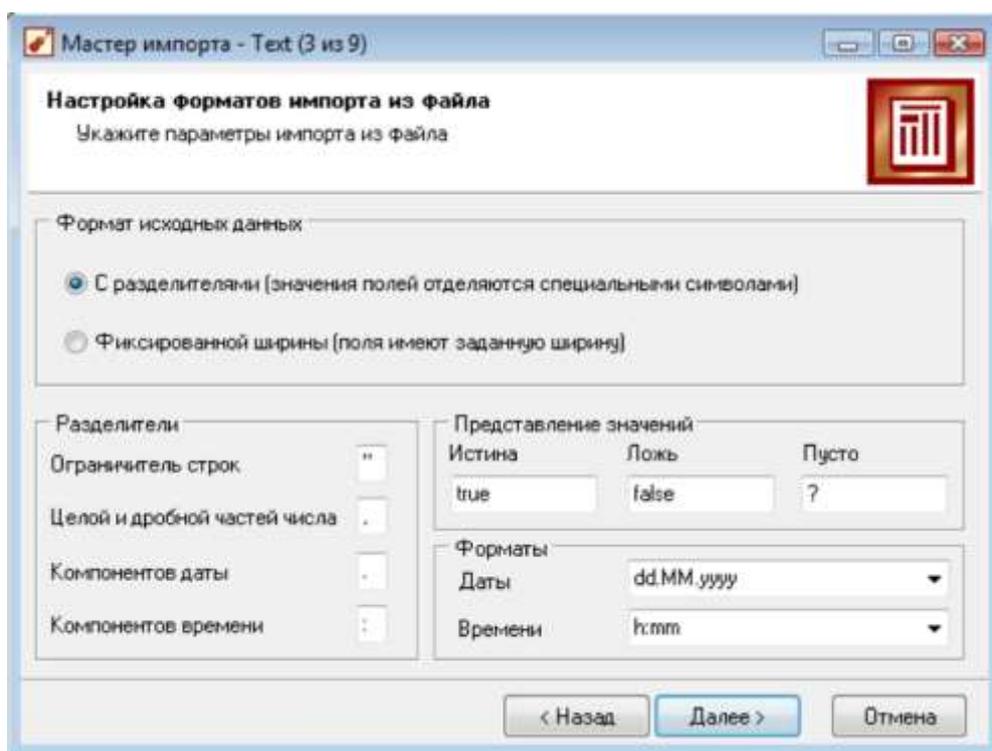


Рисунок 4 - Настройка форматов импорта из файла

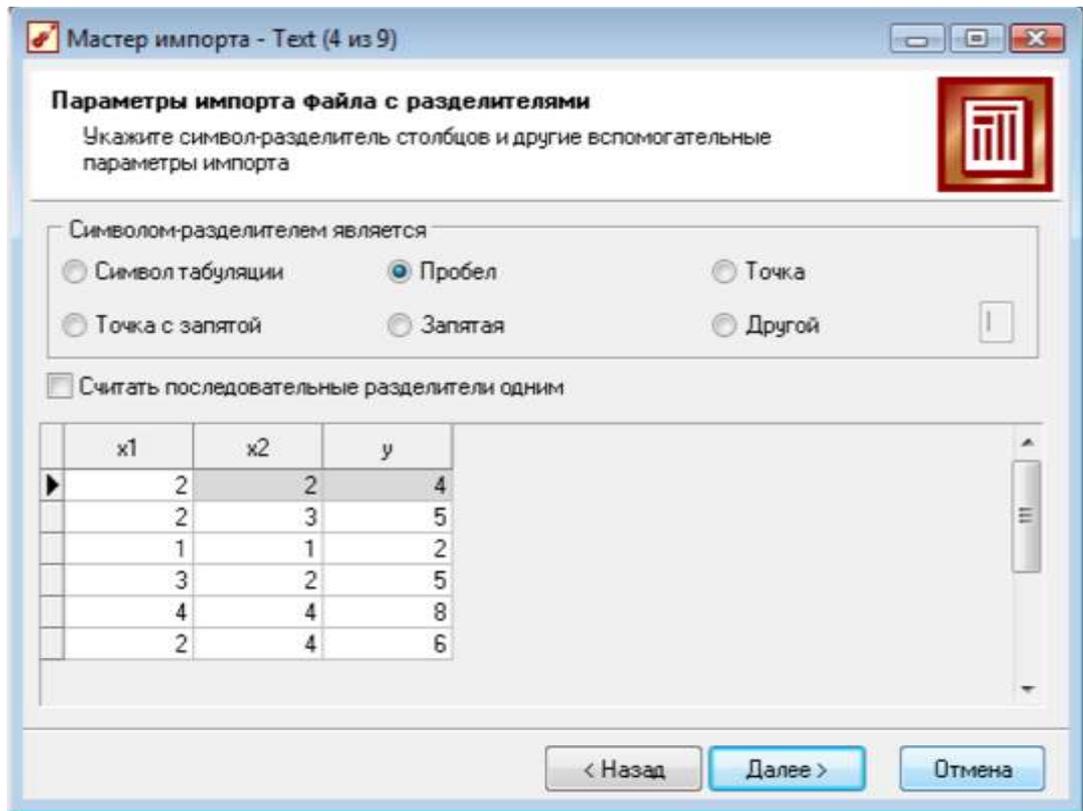


Рисунок 5 - Настройка параметров импорта файла

Запускаем процесс импорта (кнопка Пуск) и определим способ отображения как табличный (рис.7).

После импорта обучающего множества окно программы выглядит следующим образом (рис. 8).

Запускаем мастер обработки и выбираем пункт «Нейросеть» (рис. 9).

Далее происходит настройка назначения столбцов (рис. 10), здесь также можно при необходимости провести нормализацию значений.

Следующий шаг - разбиение исходного набора данных на обучающее и тестовое множества (рис. 11).

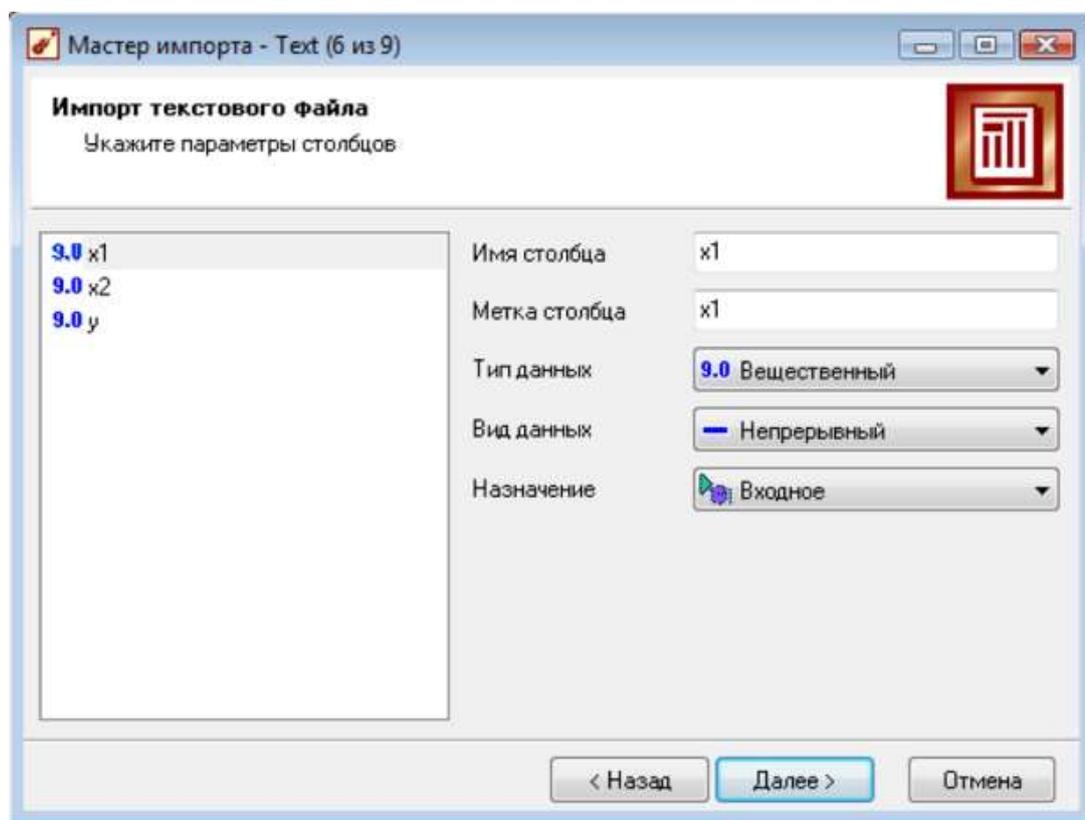


Рисунок 6 - Задание назначения полей

Обучающее множество включает записи (примеры), которые будут использоваться в качестве входных данных, а также соответствующие желаемые выходные значения.

Тестовое множество также включает записи, содержащие входные и желаемые выходные значения, но используемые не для обучения модели, а для проверки его результатов.

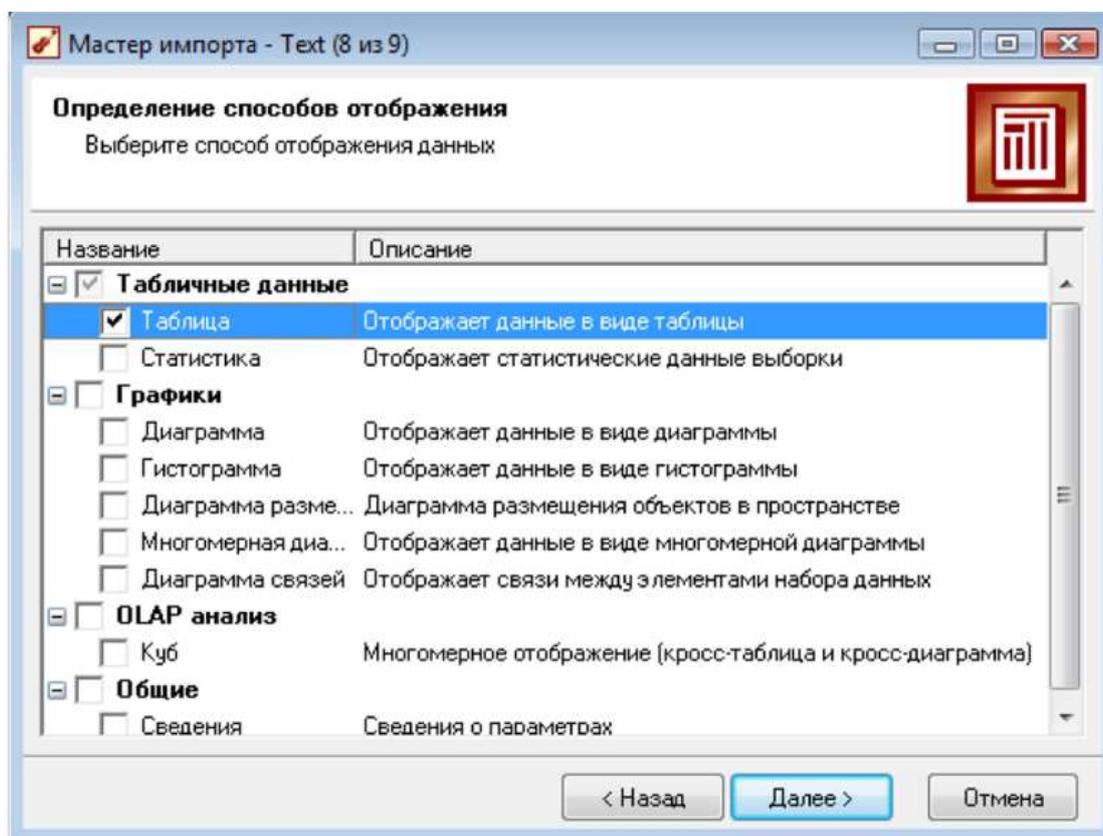


Рисунок 7 - Определение способов отображения

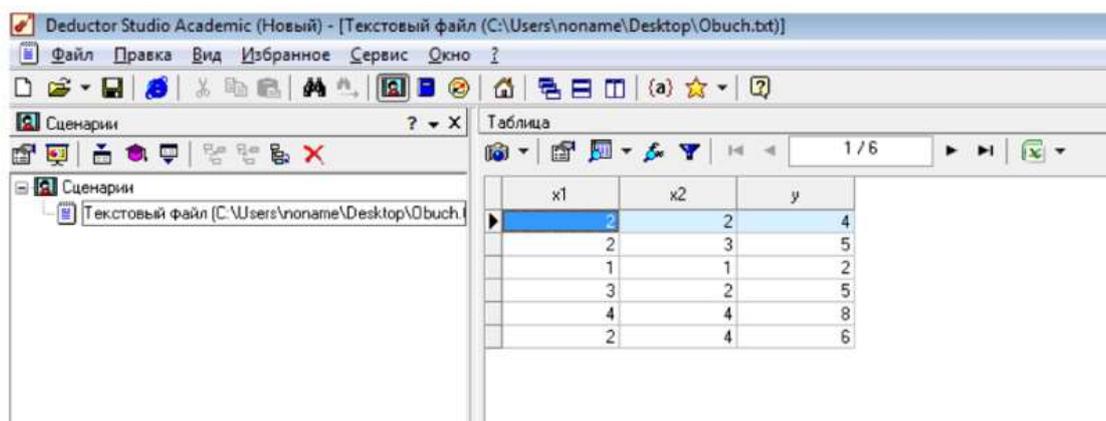


Рисунок 8 - Главное окно программы после импорта

Для разбиения исходного множества на обучающее и тестовое необходимо настроить несколько параметров: 1. Из списка "Способ разделения исходного множества" выбирается порядок отбора записей во все три множества: по порядку (порядок следования записей при их разделении не меняется); случайно (отбор записей происходит случайным образом); по столбцу (при выборе данного способа необходимо указать столбец для разделения исходного множества).

Затем необходимо указать, какие множества будут применяться. Для того чтобы множество было сформировано, нужно установить флажок слева от его названия. Если флажок сброшен, то множество использовано не будет. Обучающее множество используется всегда, поэтому сбросить флажок для него нельзя.

Для каждого из используемых множеств необходимо задать его размер. Размер может быть указан непосредственно количеством записей или в процентах от объема исходной выборки.

В поле «Количество строк (всего)» отображается общее количество записей в исходной выборке данных, которое может быть задействовано для формирования множеств.

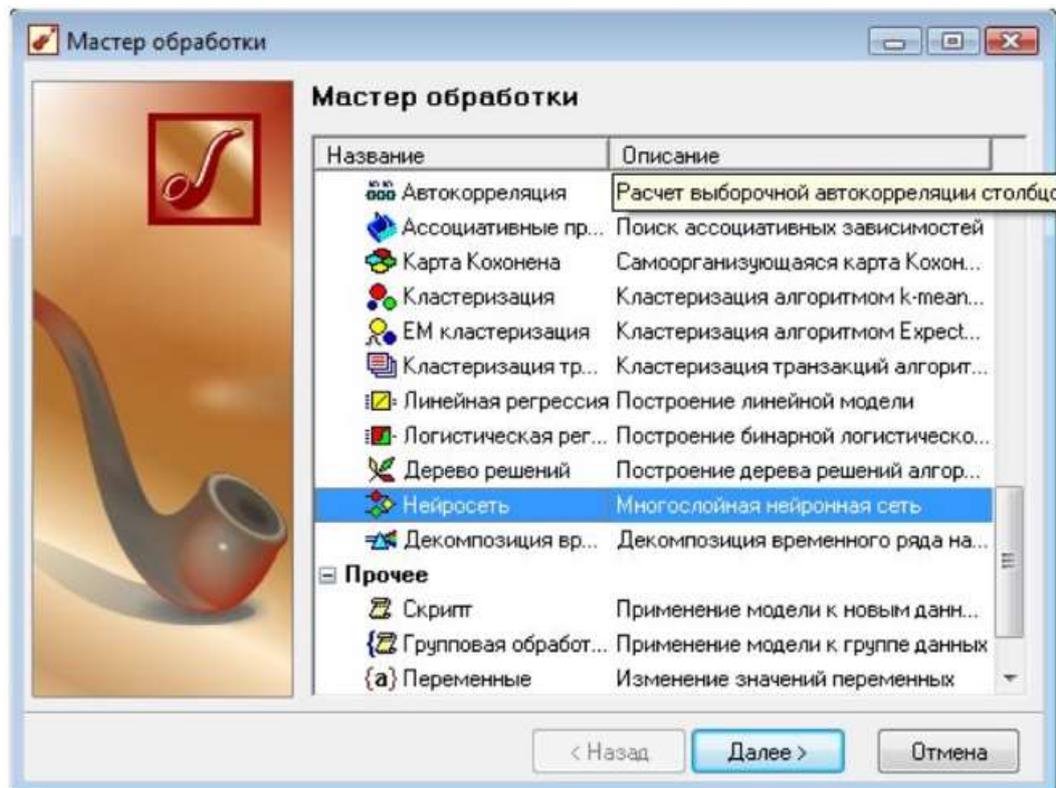


Рисунок 9 - Мастер обработки

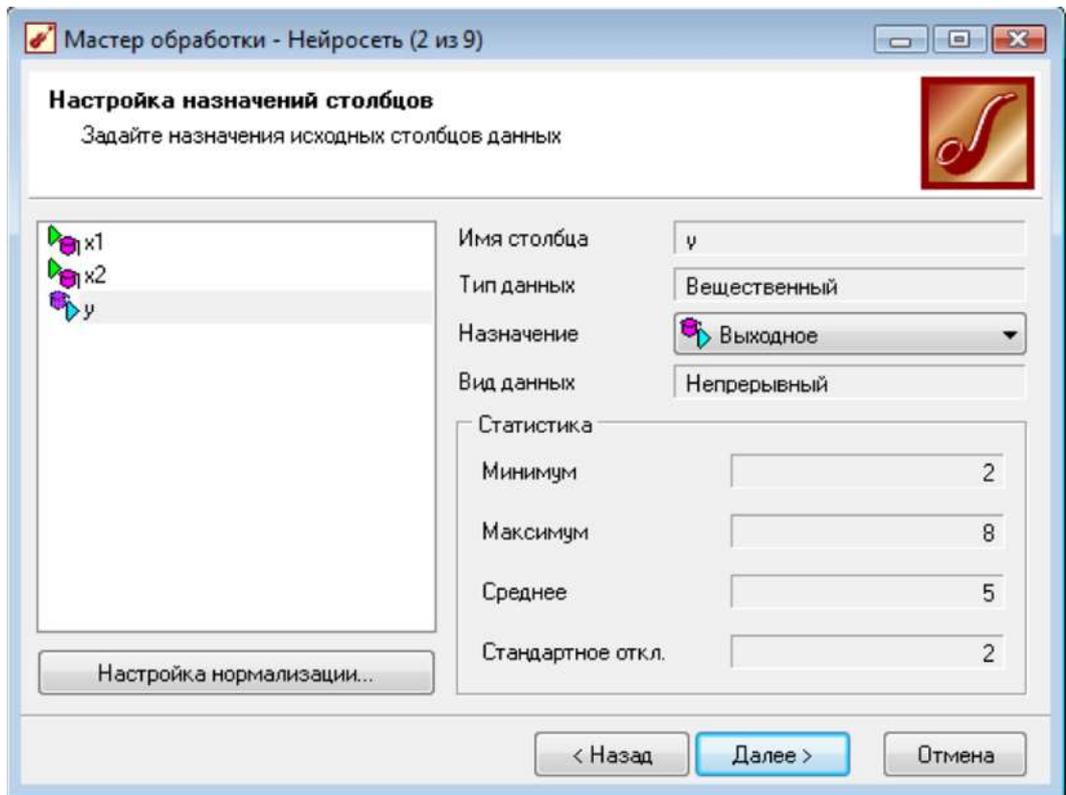


Рисунок 10 - Настройка назначений столбцов

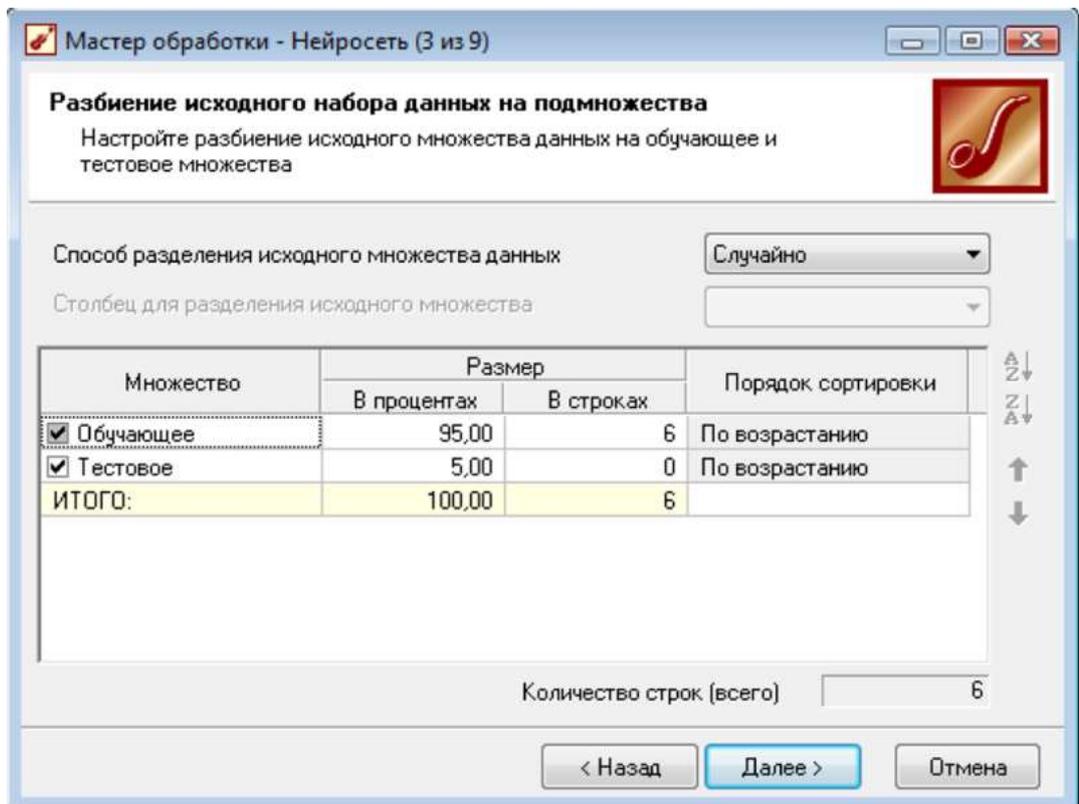


Рисунок 11 - Разбиение исходного набора данных

На рис. 12 показан этап настройки параметров, определяющих структуру нейронной сети, - количество скрытых слоев и нейронов в них, а также активационная функция нейронов.

В секции «Нейроны в слоях» необходимо указать количество скрытых слоев, т. е. слоев нейронной сети, расположенных между входным и выходным слоями. Число нейронов во входном и выходном слоях автоматически устанавливается в соответствии с числом входных и выходных полей обучающей выборки, и здесь изменить его нельзя.

Считается, что задачу любой сложности можно решить при помощи двухслойной нейросети, поэтому конфигурация с количеством скрытых слоев, превышающих 2, вряд ли оправдана. Для решения многих задач вполне подойдет однослойная нейронная сеть.

При выборе численности нейронов следует руководствоваться следующим правилом: «количество связей между нейронами должно быть примерно на порядок меньше количества примеров в обучающем множестве».

Число связей рассчитывается как связь каждого нейрона со всеми нейронами соседних слоев, включая связи на входном и выходном слоях. Слишком большая численность нейронов может привести к так называемому "переобучению" сети, когда она выдает хорошие результаты на примерах, входящих в обучающую выборку, но практически не работает на других данных.

В расположенной ниже таблице для каждого скрытого слоя необходимо указать число находящихся в нем нейронов.

В секции «Активационная функция» следует определить тип функции активации нейронов и ее крутизну. Для этого в списке «Тип функции» нужно выбрать требуемую функцию активации, а в поле «Крутизна» - задать ее крутизну (также крутизну можно задать с помощью ползунка, расположенного ниже). В нижней части окна отображается график выбранной функции в соответствии с установленной крутизной.

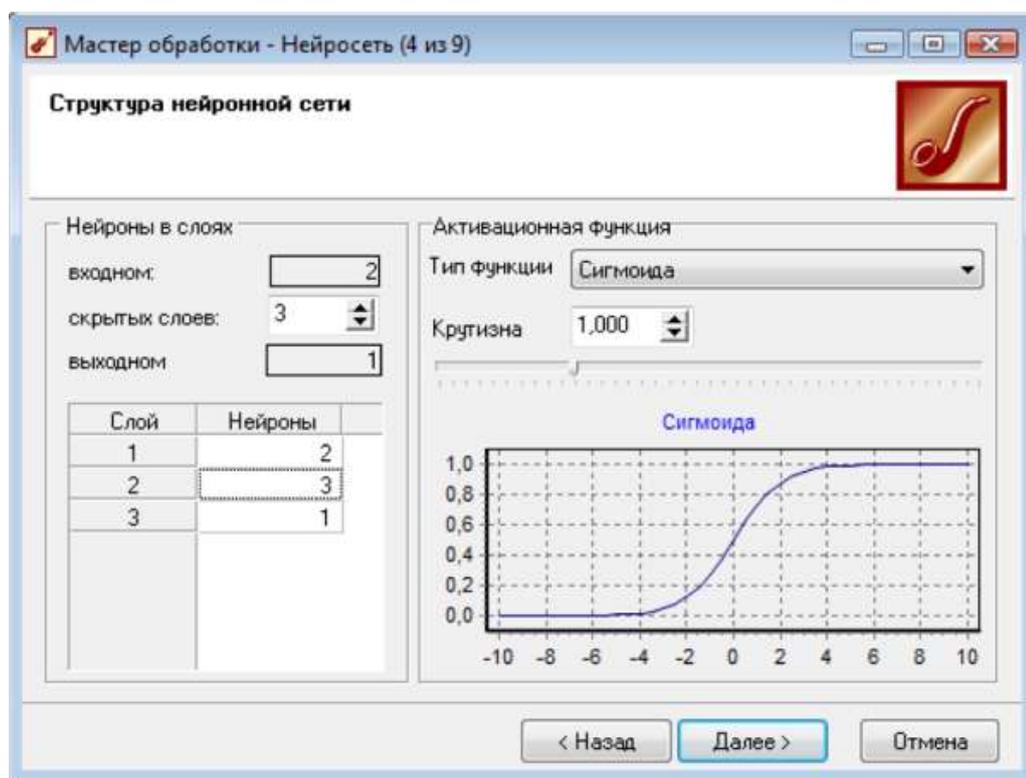


Рисунок 12 - Определение структуры нейронной сети

На рис. 13 представлена настройка процесса обучения нейронной сети. На данном шаге можно выбрать алгоритм обучения нейронной сети и задать его параметры.

Для алгоритма обратного распространения ошибки (Back- Propagation) задаются два параметра: скорость обучения - определяет величину шага при итерационной коррекции весов в нейронной сети (рекомендуется задавать в интервале 0...1); момент - учитывает величину последнего изменения веса при коррекции весов (задается в интервале 0... 1).

На шаге «Настройка параметров остановки обучения» (рис. 14) необходимо задать условия, при выполнении которых обучение будет прекращено: считать пример распознанным, если ошибка меньше (критерием останова в данном случае является условие, что рассогласование между эталонным и реальным выходом сети становится меньше заданного значения); по достижении эпохи (установка данного флажка позволяет задать число эпох (циклов обучения), по достижении которого обучение останавливается независимо от величины ошибки).

При выборе нескольких условий останова процесса обучения происходит по достижении хотя бы одного из них.

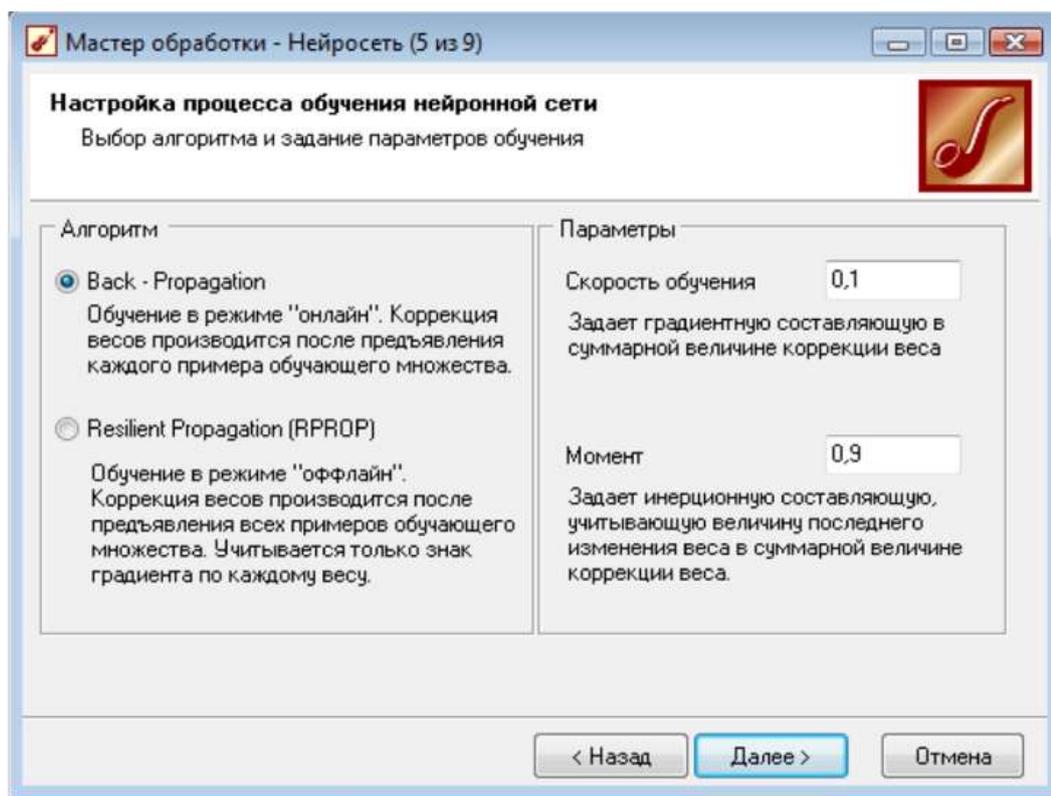


Рисунок 13 - Настройка процесса обучения нейросети

На следующем шаге производится собственно процесс обучения нейронной сети (рис. 15). В зависимости от объема обрабатываемых данных и быстродействия компьютера, он может занять определенное время.

Остановить процесс обучения можно с помощью кнопки «Стоп». После этого появляется возможность вернуться на предыдущие шаги мастера обработки для проверки и изменения параметров. Далее процесс обучения может быть запущен заново, продолжен или отменен.

В процессе обучения в окне отображаются следующие графики для обучающего (синяя линия) и тестового (красная линия) множеств: максимальная квадратичная ошибка на обучающем и тестовом множествах (сплошная линия); средняя квадратичная ошибка на обучающем и тестовом множествах (пунктирная линия).

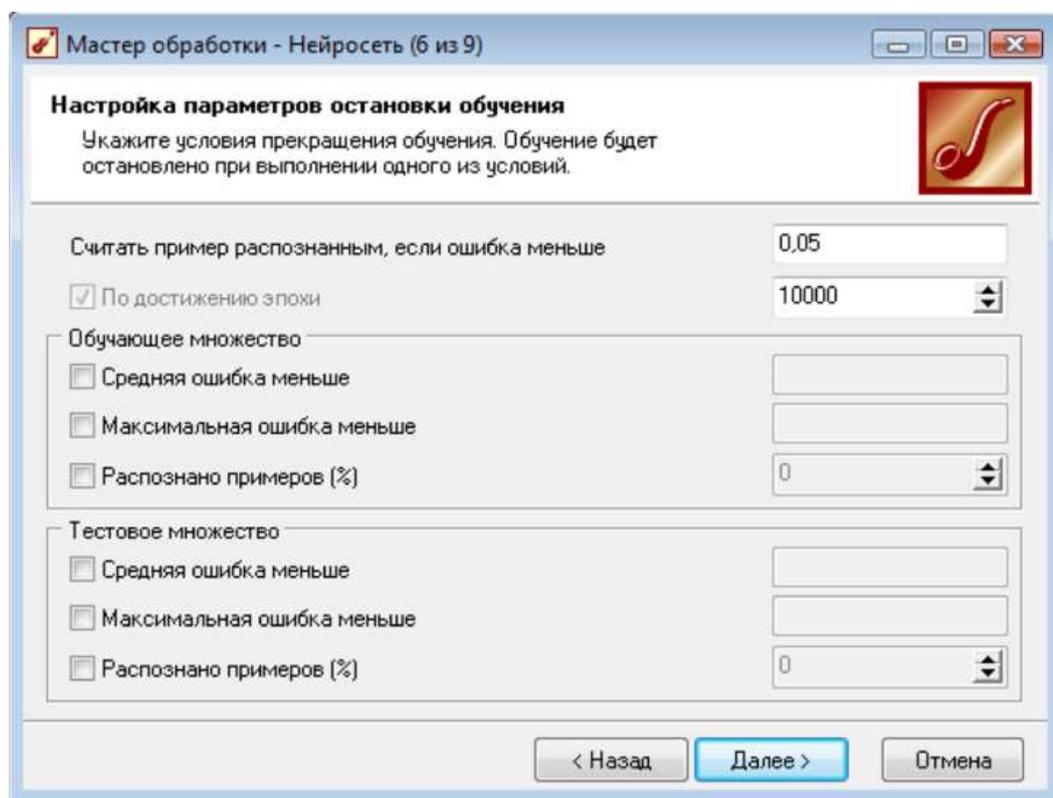


Рисунок 14 - Настройка параметров остановки обучения

В правой части окна постоянно отображаются номер текущей эпохи и время, прошедшее с начала обучения.

В поле «Темп обновления» можно задать число эпох обучения сети, через которое будет происходить обновление графика.

Флажок «Рестарт» позволяет включить режим инициализации начальных весов сети случайными значениями. Если он сброшен, то при повторном запуске обучения после остановки будет иметь место так называемое "дообучение сети", когда оно будет начато с текущими весами.

Далее необходимо определить способ отображения (рис. 16). Выберем пункты «Граф нейросети» и «Что-если».

Визуализатор «Граф нейросети» (рис. 17) будет доступен только для тех ветвей сценария, которые были созданы с помощью нейронной сети, которую он позволяет представить графически со всеми ее нейронами и синоптическими связями. При этом пользователь может увидеть не только ее структуру, но и значения весов, которые принимают те или иные нейроны. В

зависимости от веса нейрона он отображается определенным цветом, а соответствующее значение можно определить по цветовой шкале, расположенной внизу окна.

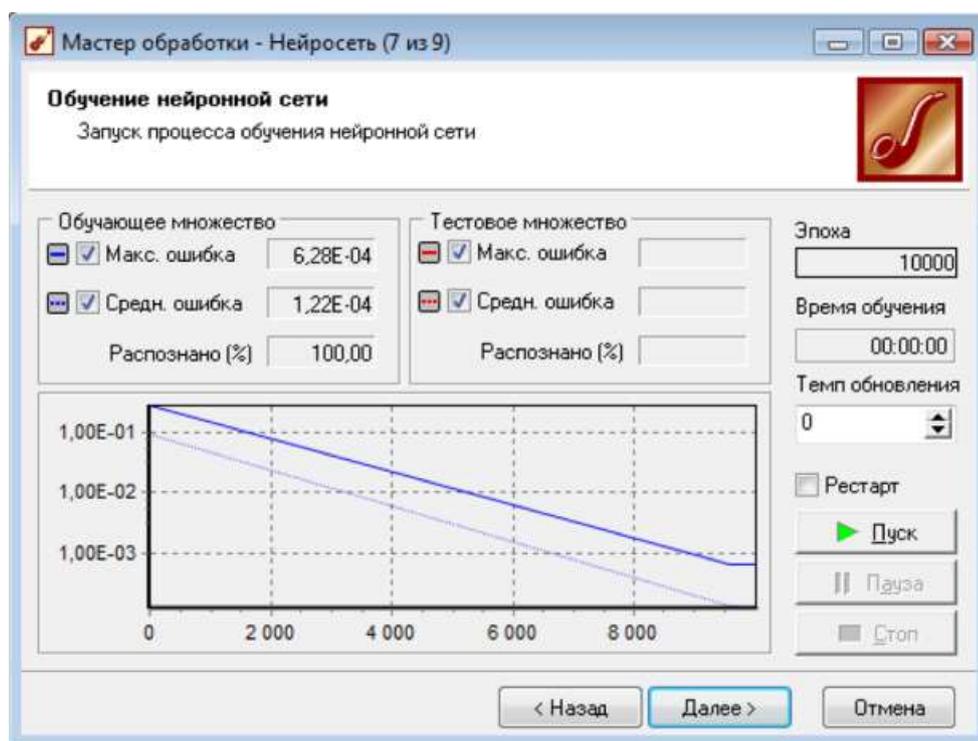


Рисунок 15 - Обучение нейросети

Анализ по методу «Что-если» позволяет исследовать, как будет вести себя построенная система обработки при подаче на ее вход тех или иных данных. То есть, проводится эксперимент, в котором, изменяя значения входных полей обучающей или рабочей выборки нейронной сети или дерева решений, пользователь наблюдает за изменением значений на выходе.

Возможность анализа по принципу "Что-если" особенно ценна, поскольку позволяет исследовать правильность работы системы, достоверность полученных результатов, а также ее устойчивость. Под устойчивостью понимается то, насколько снижается достоверность полученных результатов при попадании на вход системы нетипичных данных - выбросов, пропусков данных и т. д. Такой анализ дает возможность определить, какую предварительную обработку данных нужно провести перед подачей на вход системы.

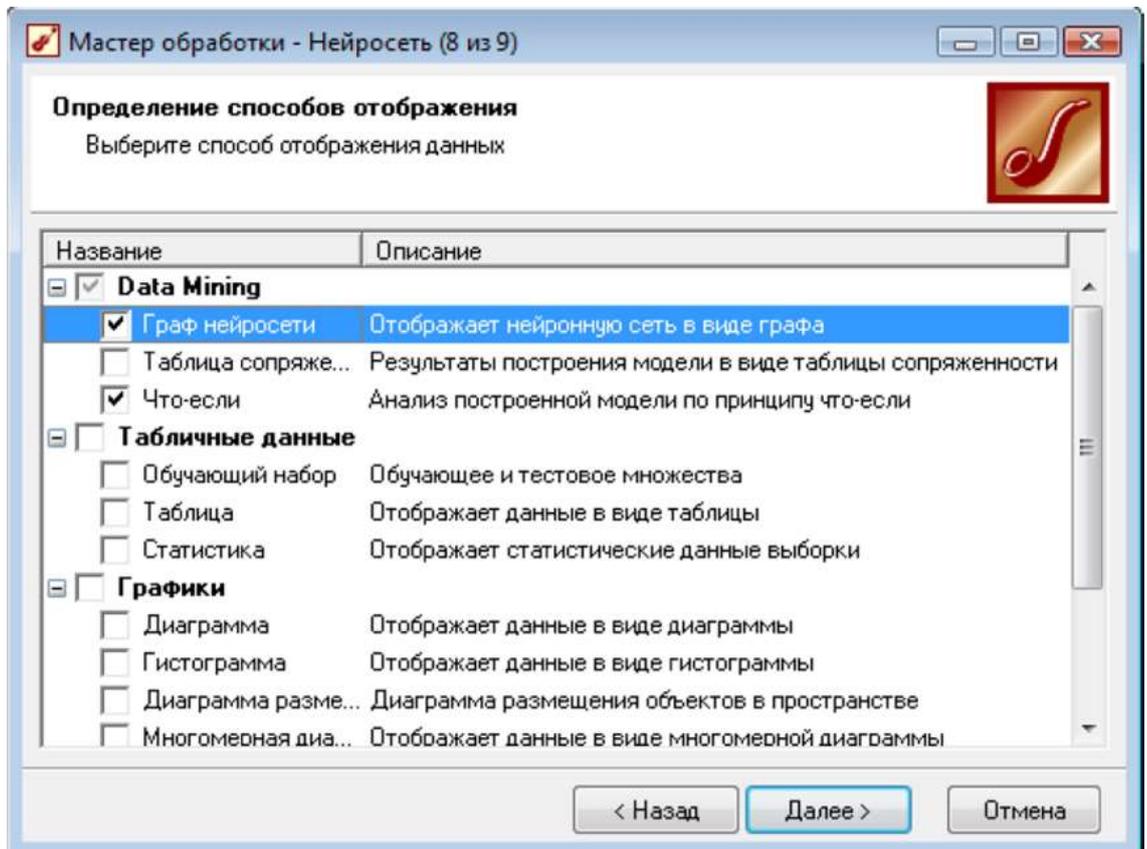


Рисунок 16 - Выбор способа отображения

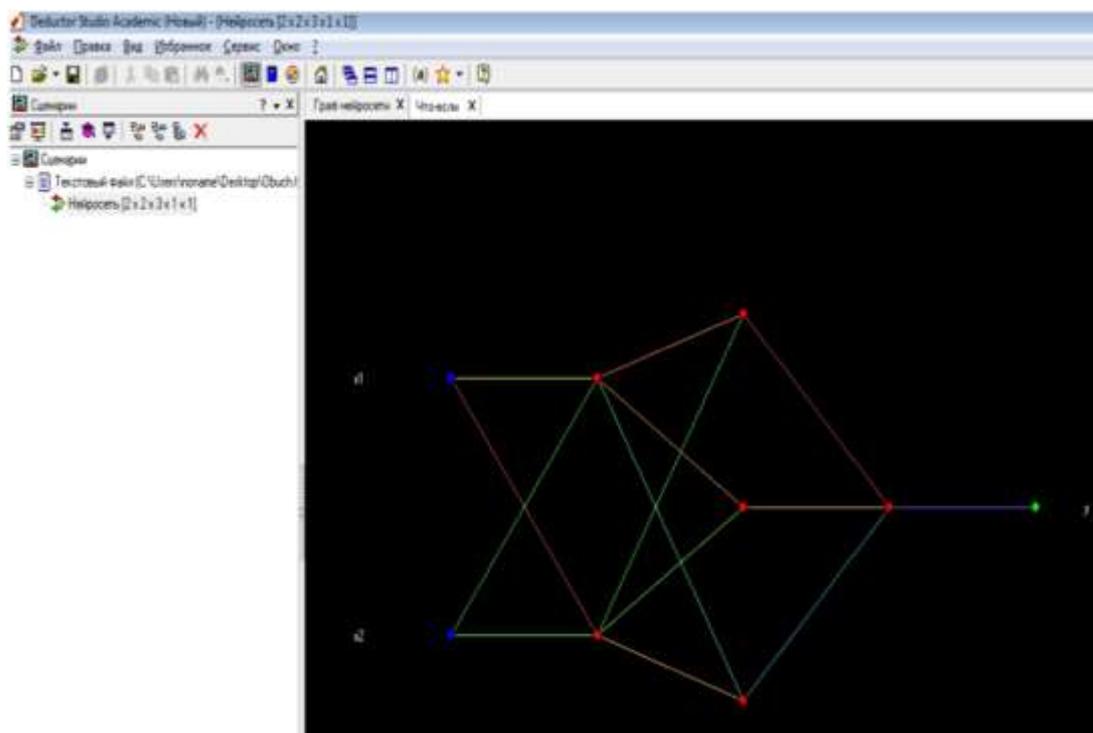


Рисунок 17 - Визуализатор «Граф нейросети»

Далее, на вкладке «Что-если», путем задания случайных входных значений проверяется работа нейросети (рис. 18).

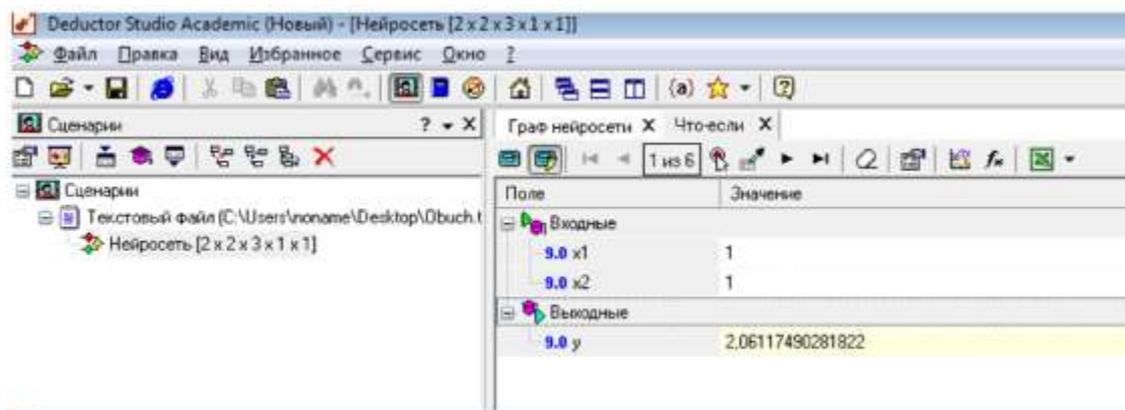


Рисунок 18 - Проверка работы нейросети

Выводы

В Deductor Studio приводятся специализированные визуализаторы для узлов, реализующих алгоритмы регрессии и классификации, такие как нейросеть. В этом режиме мастер обработки позволяет задать структуру нейронной сети, определить ее параметры и обучить с помощью одного из доступных в системе алгоритмов. В результате будет получен эмулятор нейронной сети, который может быть использован для решения задач прогнозирования, классификации, поиска скрытых закономерностей, сжатия данных и многих других приложений.

Вопросы для самоконтроля

1. Какие способы визуализации предусмотрены в Deductor Studio?
2. Что такое обучающая выборка и как импортировать ее из файла?
3. Как запустить мастер обработки?
4. Что такое разбиение исходного набора данных на подмножества?
5. Правила определения количества слоев и нейронов в скрытых слоях?

Задание:

Создать матрицу размером 3 на 20 и построить нейронную сеть в Deductor Studio.