

Практическая работа №1

Тема: Аналитическая платформа Deductor: начало работы с системой, создание проекта

Теоретические сведения

Deductor – это аналитическая платформа для создания законченных прикладных решений анализа данных. Реализованные в Deductor технологии позволяют на базе единой архитектуры пройти все этапы построения аналитической системы: от консолидации данных до построения моделей и визуализации полученных результатов.

Современным потребностям бизнеса в анализе данных отвечают новые парадигмы: хранилища данных, машинное обучение, Data Mining, Knowledge Discovery in Databases. Это позволяет использовать анализ данных для решения большого числа бизнес-задач, направленных на получение экономического эффекта.

Развитие информационных технологий анализа данных стали специализированные программные системы – аналитические платформы, которые полностью автоматизируют все этапы анализа от консолидации данных до эксплуатации моделей и интерпретации результатов, среди которых – Deductor, представитель системы анализа данных последнего поколения.

Общие сведения об аналитической платформе Deductor

Аналитическая платформа Deductor состоит из пяти частей:

- Warehouse – хранилище данных, консолидирующее информацию из разных источников;
- Studio – приложение, позволяющее пройти все этапы построения прикладного решения, рабочее место аналитика;
- Viewer – рабочее место конечного пользователя, одно из средств тиражирования знаний (т.е. когда построенные аналитиком модели используют пользователи, не владеющие технологиями анализа данных);
- Server – служба, обеспечивающая удаленную аналитическую обработку данных;
- Client – клиент доступа к Deductor Server. Обеспечивает доступ к серверу из сторонних приложений и управление его работой.

Платформа Deductor поставляется в трех вариантах:

- Enterprise;
- Professional;
- Academic.

В зависимости от типа поставки набор доступных компонентов может различаться.

Версия Enterprise предназначена для корпоративного использования. В ней присутствуют:

- Серверные компоненты Deductor Server и Deductor Client.
- Интерфейс доступа к Deductor через механизм OLE Automation.

- Традиционное хранилище данных Deductor Warehouse на трех СУБД: Firebird, MS SQL, Oracle.

- Виртуальное хранилище данных Deductor Virtual Warehouse.

Версия Professional предназначена для небольших компаний и однопользовательской работы. В ней отсутствуют серверные компоненты, поддержка OLE, виртуальное хранилище, а традиционное хранилище данных можно создавать только на СУБД FireBird. Автоматизация выполнения сценариев обработки данных осуществляется только через пакетный режим.

Версии Professional и Enterprise требуют установки драйверов Guardant для работы с лицензионным ключом.

Версия Academic предназначена для образовательных и обучающих целей. Ее функционал аналогичен версии Professional за исключением:

- функции пакетного запуска сценариев, т.е. работа в программе может вестись только в интерактивном режиме;

- функции импорта из промышленных источников данных: 1С, СУБД, файлы MS Excel, Deductor Data File;

- некоторых других возможностей.

Категории пользователей Deductor

В процессе развертывания и использования аналитической платформы с ней взаимодействуют различные категории пользователей. Можно выделить четыре основные категории:

- аналитик;
- пользователь;
- администратор;
- программист.

Функции аналитика:

- создание в Deductor Studio сценариев – последовательности шагов, которую необходимо провести для получения нужного результата;

- построение, оценка и интерпретация моделей;

- настройка панели отчетов для пользователей Deductor Viewer;

- настройка сценария на поточную обработку новых данных.

Функции пользователя – просмотр готовых отчетов в Deductor Viewer.

Функции администратора:

- установка компонентов Deductor на рабочих местах и сервера ключей Guardant при необходимости;

- развертывание традиционного хранилища данных на сервере;

- контроль процедур регулярного пополнения хранилища данных;

- конфигурирование сервера Deductor Server;

- настройка пакетной и/или серверной обработки сценариев Deductor;
- оптимизация доступа к источникам данных, в том числе к хранилищу данных.

Функции программиста:

- интеграция Deductor с источниками и приемниками данных;
- вызов Deductor из внешних программ различными способами, в том числе взаимодействие с Deductor Server.

Такая работа как проектирование и наполнение хранилище данных часто выполняется коллективно аналитиком, администратором и программистом. Аналитик проектирует семантический слой хранилища данных, то есть определяет, какие данные необходимо иметь в хранилище. Администратор создает хранилище данных и наполняет его данными. Программист при необходимости создает программные модули, выполняющие выгрузку информации из учетных систем в промежуточные источники (так называемые транспортные таблицы).

Установка Deductor

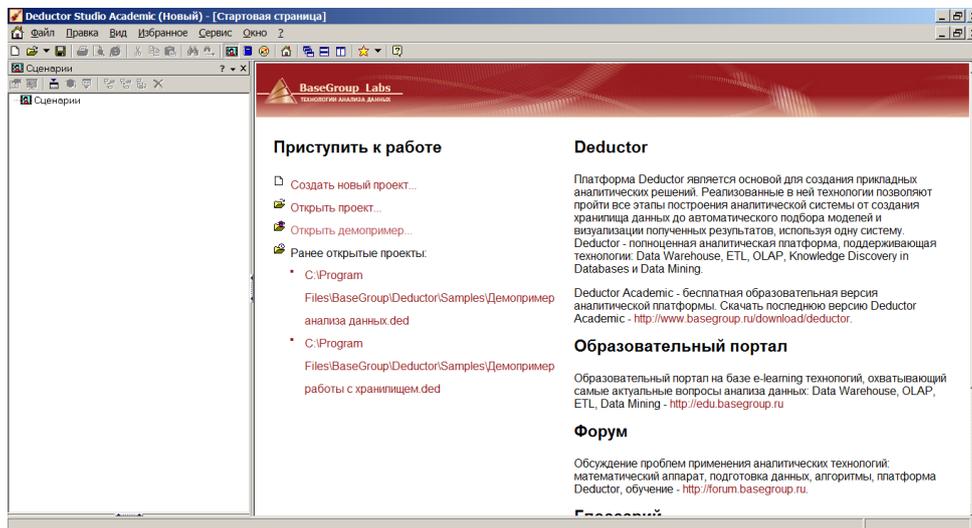
Установку Deductor рекомендуется проводить администратору системы, однако, при наличии прав администратора в Windows это может сделать и аналитик. Установка может быть произведена на компьютер с операционной системой MS Windows 2000 и выше. Системные требования к компьютеру изложены в справочной системе.

Для установки Deductor Professional/Academic запускается файл инсталлятора и следуют инструкциям по установке. На странице «Выбор компонентов» программы установки предоставляется выбор, какой набор компонентов пакета Deductor необходимо установить на компьютер. В выпадающем списке можно выбрать predetermined configurations установки платформы, и программа установки сама предложит нужный набор компонентов.

После установки программ серии Professional и Enterprise дополнительно потребуется настроить работу с электронным ключом защиты от копирования. Установку и подсоединение электронного ключа осуществляет администратор. Существуют два вида ключей – локальный и сетевой. Локальный ключ устанавливается на том же компьютере, что и Deductor, и работать с ним можно только с этой рабочей станции. Сетевой ключ устанавливается на сервере, и к нему могут подключаться несколько пользователей одновременно (количество пользователей ограничивается типом приобретаемой лицензии).

При каждом запуске Deductor пытается найти доступный электронный ключ. В случае если ключ не найден, могут появиться следующие сообщения об ошибке. При наличии ошибок следует обратиться к администратору.

После запуска главное окно Deductor Studio выглядит следующим образом.



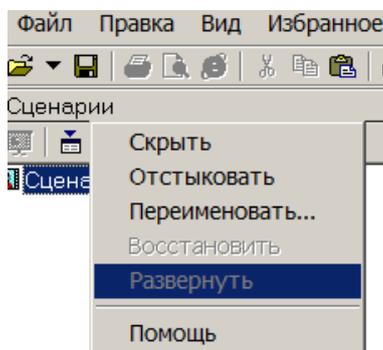
По умолчанию панель управления представлена одной вкладкой «Сценарии». Кроме того, доступны еще две вкладки: Отчеты и Подключения. Сделать их видимыми можно следующими способами:

- главное меню **Вид >Отчеты** и **Вид >Подключения**

- кнопки  и  на панели инструментов.

Можно производить «drag & drop» манипуляции с вкладками, меняя их расположение и порядок.

При нажатии правой кнопки мыши на любой вкладке появляется контекстное меню:



- Скрыть – делает вкладку невидимой;
- Переименовать – переименовывает название вкладки;
- Закладки – переключается на выбранную закладку;
- Верх/Низ – задает расположение названий вкладок: вверху либо внизу;
- Помощь – открывает раздел справки.

Справка по программе вызывается из главного окна системы следующими способами:

- главное меню ? >**Справка**, клавиша **F1**,
- кнопка на панели инструментов .

Помощь содержит подробное описание работы с Deductor Studio: системные требования, настройки узлов, способы осуществления действий с объектами системы.

Понятие проекта

В Deductor Studio ключевым понятием является проект. Это файл с расширением *.ded, по структуре соответствующий стандартному xml-файлу. Он хранит в себе:

- последовательности обработки данных (сценарии);
- настроенные визуализаторы;
- переменные проекта и служебную информацию.

Пример фрагмента файла *.ded:

```
<?xml version="1.0" encoding="UTF-8"?>
<Document>
<Version>
<Comments>Deductor Studio Academic</Comments>
<CompanyName>BaseGroup Labs</CompanyName>
<FileDescription>Deductor Studio Academic</FileDescription>
<FileVersion>5.2.0.89</FileVersion>
<InternalName>Deductor Studio Academic</InternalName>
<LegalCopyright>BaseGroup</LegalCopyright>
<LegalTrademarks>BaseGroup</LegalTrademarks>
<OriginalFilename>DStudio.exe</OriginalFilename>
<ProductName>Deductor Studio Academic</ProductName>
<ProductVersion>5.2</ProductVersion>
</Version>
<Properties>
<Author>USer</Author>
<DisplayName>p1</DisplayName>
</Properties>
<EnvironmentVariables>
<Version>1</Version>
<Vars/>
</EnvironmentVariables>
<Script>
<NodeCount>1</NodeCount>
<Root>
<DisplayName>Сценарии</DisplayName>
<NodeIndex>1</NodeIndex>
</Root>
</Script>
<ModelBookmarks/>
</Document>
```

Замечание. Каждый проект имеет авторские сведения: Название, Версия, Автор, Компания, Описание. Они заполняются в диалоговом окне Свойства проекта (меню **Файл>Свойства проекта...**).

Создать новый проект можно следующими способами:

- главное меню **Файл>Создать**;
- кнопка **Создать новый проект** на панели инструментов;
- клавиша **Ctrl+N**.

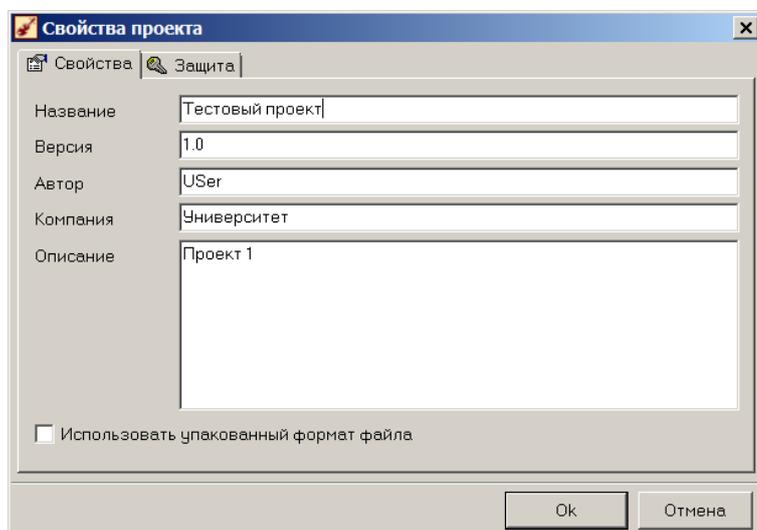
Открытие существующего проекта:

- главное меню **Файл>Открыть**;
- кнопка **Открыть проект** на панели инструментов;
- клавиша **Ctrl+O**.

Открыть проект можно еще одним способом – в главном меню **Файл>История найти имя проекта**. Способ работает в том случае, если недавно открывался этот проект, и он сохранился в менеджере истории проектов.

В одной запущенной копии Deductor Studio можно открыть только один проект.

По умолчанию файл проекта Deductor при сохранении запаковывается, что позволяет уменьшить его размер, поэтому просмотреть запакованный файл в виде xml невозможно. Для этого нужно снять опцию **Использовать упакованный формат файла** в диалоговом окне **Свойства проекта** (меню **Файл > Свойства проекта...**).



Создать новый проект можно следующими способами:

- главное меню **Файл>Создать**;
- кнопка **Создать новый проект** на панели инструментов;
- клавиша **Ctrl+N**.

Для сохранения проекта под текущим именем нужно выбрать главное меню **Файл>Сохранить**, нажать кнопку или комбинацию **Ctrl+S**.

Для сохранения текущего проекта под другим именем: главное меню **Файл>Сохранить как...**

В Deductor Studio вся работа ведется с использованием пяти мастеров:

- мастера импорта;
- мастера экспорта;
- мастера обработки;
- мастера визуализации;
- мастера подключений.

С помощью мастеров импорта, экспорта и обработки формируется сценарий. Сценарий состоит из узлов. Мастер подключений предназначен для создания настроек подключений к различным источникам и приемникам данных. Мастер визуализации настраивает визуализаторы для конкретного узла.

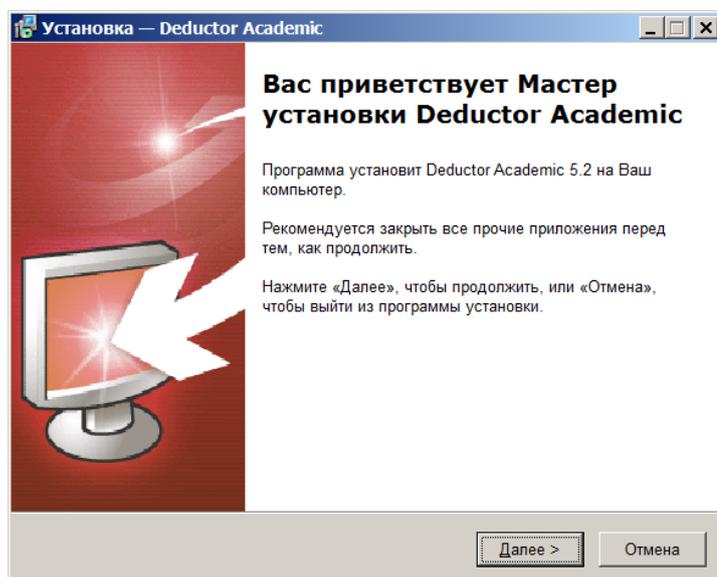
Визуализатором называется любое представление набора данных в каком-либо виде:

- табличном, графическом, описательном. Примеры визуализаторов: таблица, дерево,
- гистограмма, диаграмма, OLAP-куб и т.д.

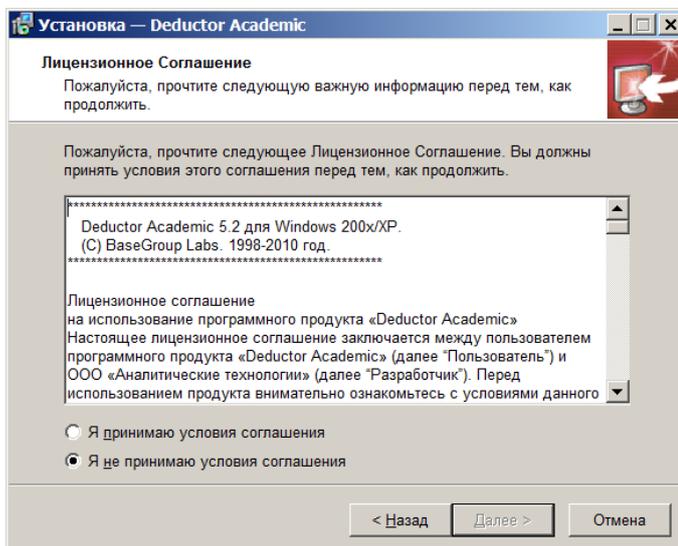
Ход работы

1. Скачайте инсталляционный пакет аналитической платформы Deductor. Свободно распространяемая платформа в виде инсталляционного пакета доступна на сайте разработчика <http://www.basegroup.ru/>., например, deductor5.2setupacademic.exe.

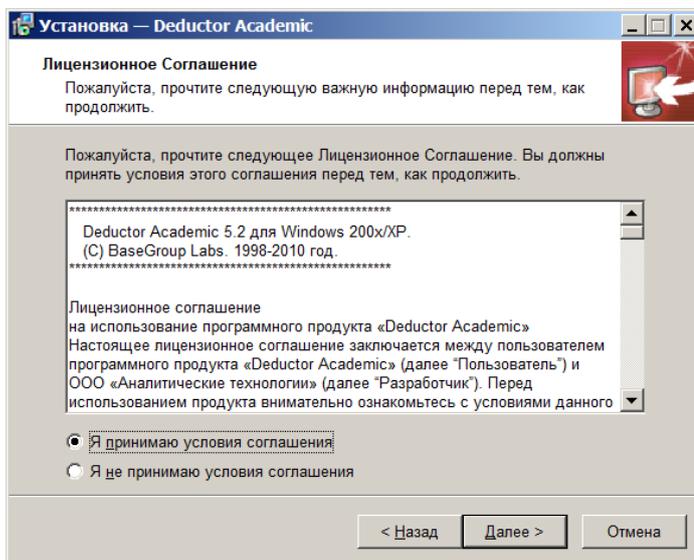
2. Запустите на выполнение файл deductor5.2setupacademic.exe. Появляется окно установки программы:



3. По клику кнопки «Далее» появляется окно лицензионного соглашения:

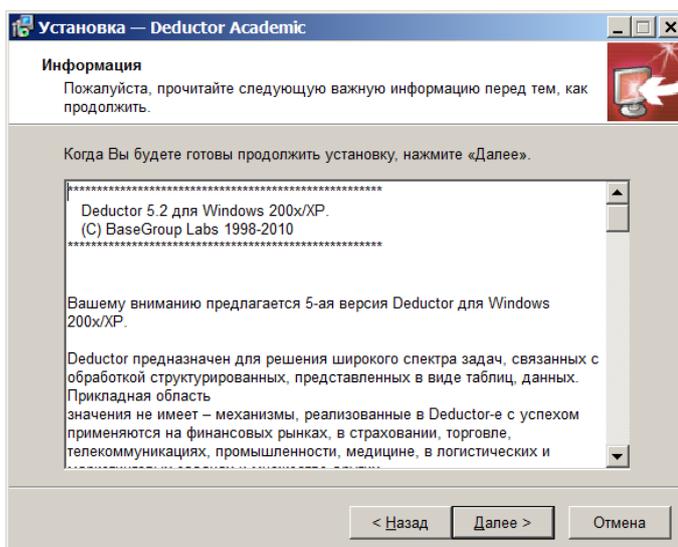


4. Установив флажок «Я принимаю условия соглашения»,



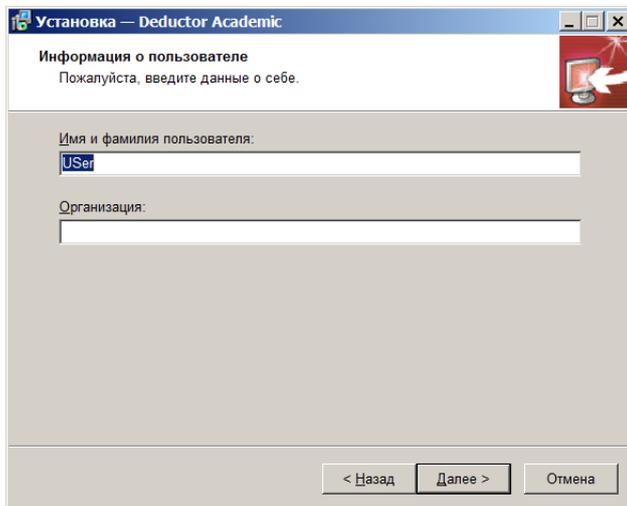
кликнуть кнопку «Далее».

5. После ознакомления с сообщением



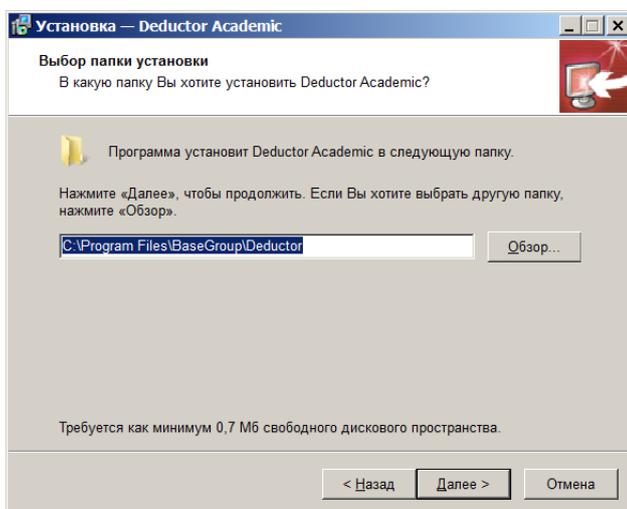
кликнуть кнопку «Далее».

6. Будет предложено ввести информацию о пользователе:



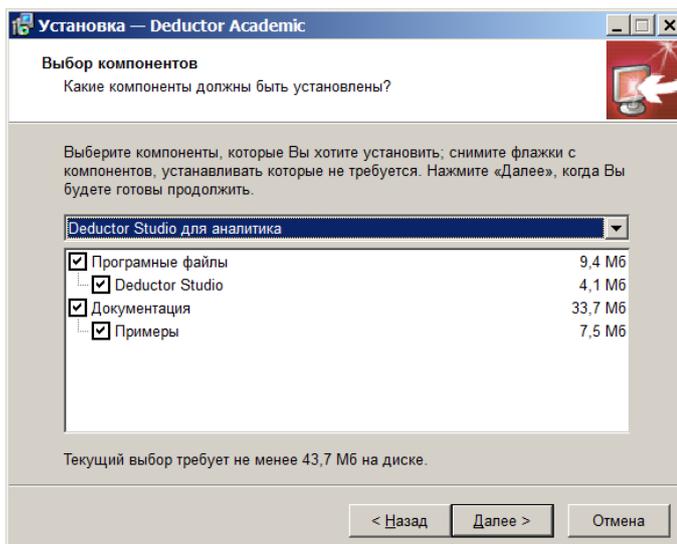
Заполнить имя пользователя. Кликнуть кнопку «Далее».

7. Выбрать папку установки:



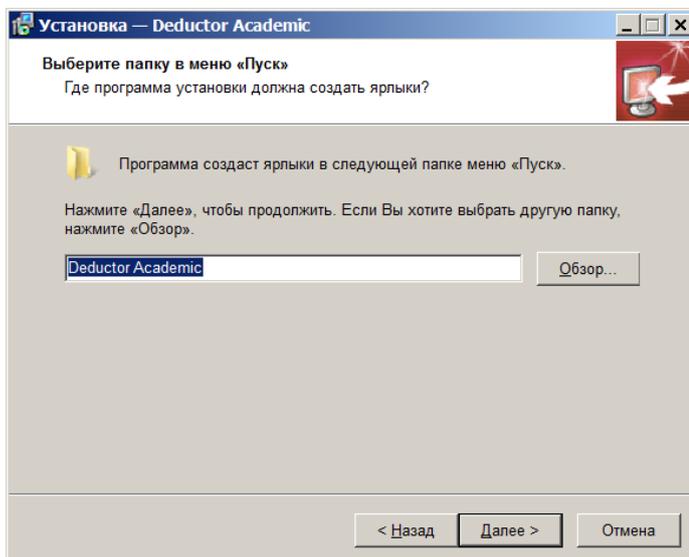
Кликнуть кнопку «Далее».

8. Выбрать компоненты программы:



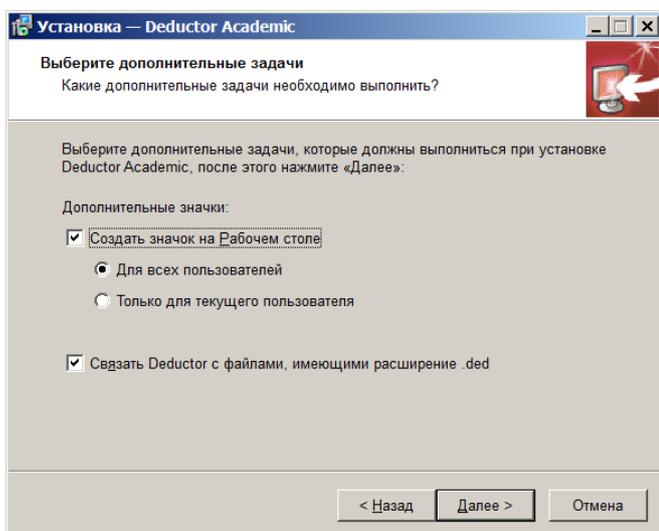
Кликнуть кнопку «Далее».

9. Выбрать папку в меню «Пуск»



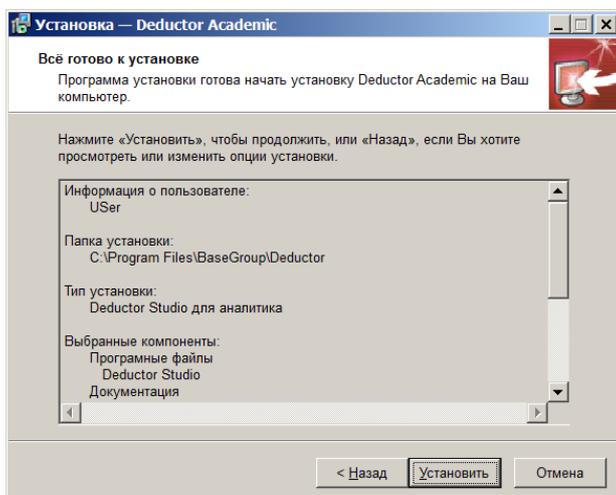
«Далее»

10. Выбрать дополнительные задачи:



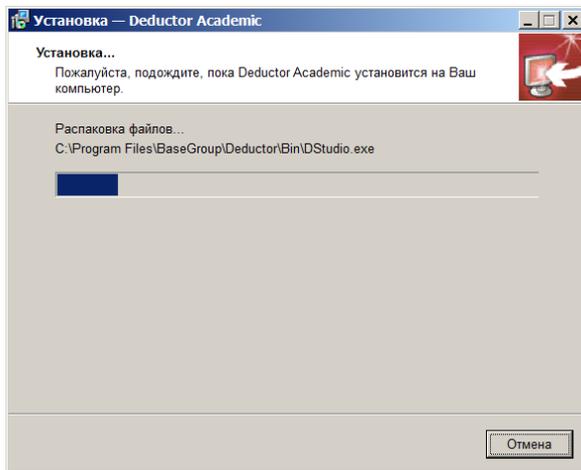
Кликнуть кнопку «Далее».

11. Приступить к установке,

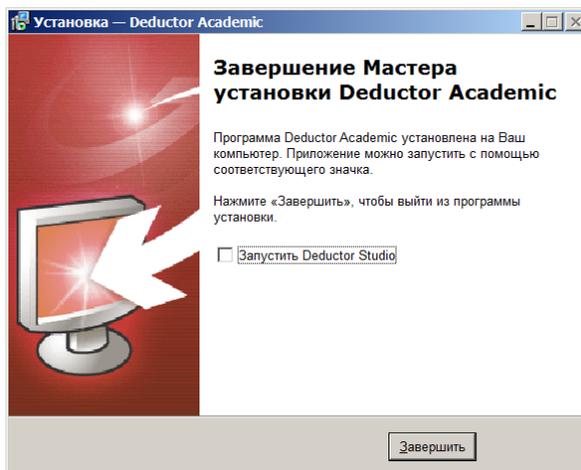


Нажатием кнопки «Установить».

12. Проходит процесс установки программы:

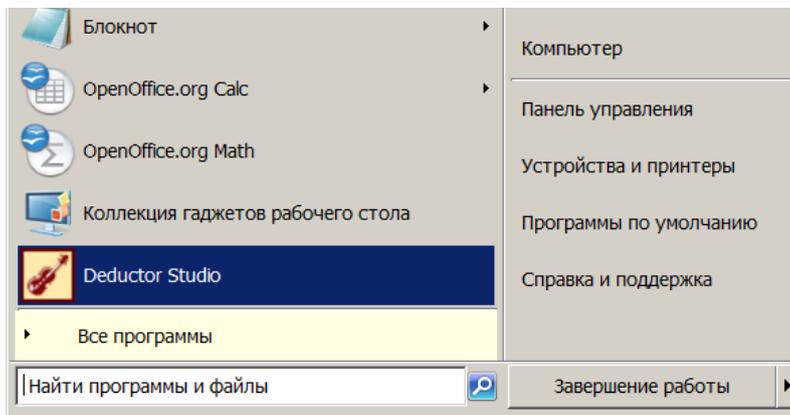


а после его завершения выдается сообщение:

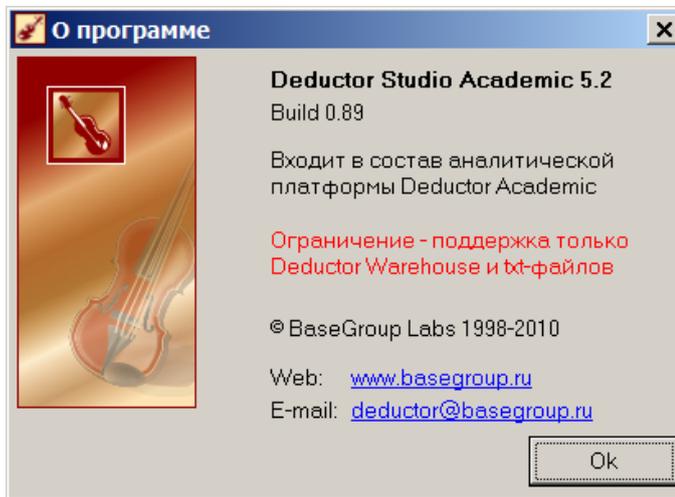


Следует кликнуть: «Завершить».

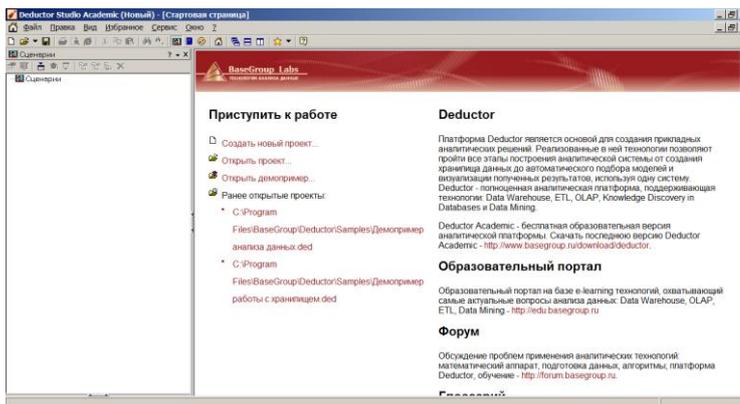
13. В пусковом меню появляется строка Deductor Studio:



14. Кликком Deductor Studio загружается программа и выводится сообщение о ней:



15. Клик кнопки «ОК» разворачивает окно программы.



16. Далее может выполняться аналитическая работа в Deductor.

17. Создайте новый проект и сохраните его под именем p1.ded. Не используйте упакованный формат файла.

18. Заполните свойства проекта.

19. Просмотрите файл проекта через любой текстовый редактор.

20. Сделайте видимой вкладку Подключения.

21. Поменяйте местами порядок вкладок Сценарии и Подключения.

22. Найдите в помощи раздел «Системные требования».