

МИНИСТЕРСТВО НАУКИ И ВЫСШЕГО ОБРАЗОВАНИЯ  
РОССИЙСКОЙ ФЕДЕРАЦИИ  
ФЕДЕРАЛЬНОЕ ГОСУДАРСТВЕННОЕ АВТОНОМНОЕ  
ОБРАЗОВАТЕЛЬНОЕ УЧРЕЖДЕНИЕ ВЫСШЕГО ОБРАЗОВАНИЯ  
«СЕВЕРО-КАВКАЗСКИЙ ФЕДЕРАЛЬНЫЙ УНИВЕРСИТЕТ»

# МЕТОДИЧЕСКИЕ УКАЗАНИЯ

по выполнению лабораторных работ  
по дисциплине «Машинное обучение»  
для магистрантов направления 09.04.03 «Прикладная  
информатика» направленность (профиль) «Управление  
знаниями»

Ставрополь  
2026

## Введение

**OLAP** (англ. *online analytical processing*, аналитическая обработка в реальном времени) — технология обработки информации, включающая составление и динамическую публикацию отчётов и документов. Используется аналитиками для быстрой обработки сложных запросов к базе данных. Служит для подготовки бизнес-отчётов по продажам, маркетингу, в целях управления, т. н. **data mining** — добыча данных (способ анализа информации в базе данных с целью отыскания аномалий и трендов без выяснения смыслового значения записей).

**Data Mining** - современная технология анализа информации с целью нахождения в накопленных данных ранее неизвестных, нетривиальных и практически полезных знаний, необходимых для принятия оптимальных решений в различных областях человеческой деятельности.

**Microsoft SQL Server** предоставляет интегрированную среду для создания моделей **Data Mining** и работы с ними. В курсе описаны четыре сценария (адресная рассылка, прогнозирование, анализ покупательской корзины и анализ последовательности покупок) для демонстрации методов использования алгоритмов **Data Mining**, а также сопутствующего инструментария, входящего в состав **SQL Server Analysis Services**.

Средства анализа данных **OLAP** и **Data Mining** объединены в две среды разработки: **Business Intelligence Development Studio** и **SQL Server Management Studio**.

В **Business Intelligence Development Studio** можно создавать отсоединенные от сервера проекты. В этом случае готовый проект можно загрузить на сервер, но ничто не мешает работать с проектом, работающим с сервером напрямую.

Все средства для управления моделями **Data Mining** доступны в редакторе моделей. С его помощью можно создавать, просматривать, сравнивать разные модели, а также создавать на их базе прогнозы.

После создания модели можно провести ее анализ на предмет выявления интересных для нас шаблонов (паттернов) и правил. В зависимости от применяемого алгоритма в среде разработки используется различные представления просматриваемой модели.

Поскольку зачастую проект содержит несколько моделей **Data Mining**, предварительно созданных на основе наших предположений, то возникает задача определения наиболее адекватной из них. Для решения этой задачи в редакторе присутствует средство сравнения моделей **Mining Accuracy Chart**. С использованием этого инструмента можно предсказать точность модели и выбрать лучшую из них.

Для создания прогнозов используется язык **Data Mining Extensions (DMX)**, который является расширением **SQL** и содержит команды для создания, изменения и осуществления предсказаний на основании различных моделей. Создание таких прогнозов может быть сложной задачей, поэтому в

редакторе присутствует инструмент под названием Prediction Query Builder, который представляет собой визуальное средство создания DMX-запросов.

Кроме инструментария для работы с моделями, не менее значимыми являются и способы создания моделей. Ключевым моментом создания модели является выбор алгоритма обнаружения данных. SQL Server Analysis Services включает в себя следующие девять алгоритмов:

- Дерево решений (Microsoft Decision Trees)
- Кластеризация (Microsoft Clustering)
- "Наивный" Байес (Microsoft Native Bayes)
- Кластеризация последовательностей (Microsoft Sequence Clustering)
- Временные ряды (Microsoft Time Series)
- Ассоциативные правила (Microsoft Association)
- Нейронная сеть (Microsoft Neural Network)
- Линейная регрессия (Microsoft Linear Regression)
- Логистическая регрессия (Microsoft Logistic Regression)

Используя комбинацию этих алгоритмов можно создавать решения для большей части встречающихся задач по обнаружению скрытых закономерностей в больших объемах данных.

Наиболее важные шаги создания модели Data Mining состоят в получении, объединении, очистке и предобработке данных для дальнейшего использования алгоритмом. В состав SQL Server входят средства преобразования данных - SQL Server Integration Services (SSIS), которые позволяют объединять, очищать, проверять и проводить предварительную обработку данных.

Для демонстрации возможностей SQL Server мы будем использовать демо-проект AdventureWorksDW. Эта база включена в поставку SQL Server и спроектирована должным образом для работы с OLAP и моделями Data Mining. Чтобы эта база была доступна следует выбрать опцию установки этого проекта во время инсталляции SQL Server.

## Лабораторная работа 1. Реляционная база данных, хранимые процедуры, триггеры

**Тема** - Реляционная база данных, хранимые процедуры, триггеры.

**Цель работы** – приобретение практических навыков при работе с СУБД MSSQL в создании процедур и триггеров при работе с OLTP.

**Формируемые компетенции или их части** - ОПК-1; ОПК-2; ОПК-4; ПК-2

### Теоретическая часть.

Для проведения данной лабораторной работы должны быть установлены следующие компоненты:

- SQL Server Management Studio для работы с базами данных;
- SQL Server ;
- Образец базы данных **AdventureWorksDW**.

Можно загрузить образец базы данных со страницы SQL Server Samples and Sample Databases (образцы баз данных и другие образцы SQL Server) в центре загрузки Microsoft, либо использовать следующие указания для установки образца базы данных во время или после установки SQL Server.

Установка образца во время выполнения программы установки

Чтобы установить образец баз данных и сопутствующие примеры с помощью программы установки, выполните следующие действия:

- на странице **Компоненты для установки** выберите **Компоненты рабочей станции, электронная документация и средства разработки**;
- нажмите кнопку **Дополнительно** и раскройте пункт **Документация, примеры и образцы баз данных**;
- выберите **Образцы кода и приложений**;
- раскройте пункт **Образцы баз данных** и выберите образцы баз данных для установки;
- для завершения установки примеров после завершения установки выполните следующие шаги:
  - в меню **Пуск** последовательно выберите пункты **Программы, Microsoft SQL Server, Документация и учебные материалы, Примеры** и **Примеры Microsoft SQL Server**;
  - либо при помощи проводника Windows перейдите в папку <диск>:\Program Files\Microsoft SQL Server\90\Tools\Samples\ и дважды щелкните файл **SqlServerSamples.msi** для запуска программы установки.

Установка образца после завершения программы установки

Если образец базы данных не был установлен во время начальной установки SQL Server, его можно установить позже.

- в средстве **Установка и удаление программ** выберите **Microsoft SQL Server** и нажмите кнопку **Изменить**. Следуйте инструкциям мастера **Обслуживание Microsoft SQL Server**;

- в окне **Выбор компонентов** выберите **Компоненты рабочей станции** и нажмите кнопку **Далее**;
- на странице **Вас приветствует мастер установки SQL Server** нажмите кнопку **Далее**;
- на странице **Проверка конфигурации системы** нажмите кнопку **Далее**;
- в окне **Изменить или удалить экземпляр** выберите **Изменить установленные компоненты**;
- в окне **Выбор компонентов** раскройте узел **Документация, примеры и образцы баз данных**;
- выберите **Образцы кода и приложений**;
- раскройте пункт **Образцы баз данных** и выберите образец базы данных для установки. Нажмите кнопку **Далее**;
- чтобы установить и подключить образцы баз данных, на странице **Установка образцов баз данных** выберите **Установить и подключить образцы баз данных**, а затем нажмите кнопку **Далее**;

Файлы базы данных создаются и сохраняются в папке <диск>:\Program Files\Microsoft SQL Server\MSSQL.n\MSSQL\Data. Базы данных активизируются и готовы к использованию.

- чтобы установить образец базы данных без подключения, на странице **Установка образцов баз данных** выберите **Установить образцы баз данных** и нажмите кнопку **Далее**;

Файлы базы данных **AdventureWorksDW** создаются в папке <диск>:\Program Files\Microsoft SQL Server\90\Tools\Samples\AdventureWorks Data Warehouse. Перед использованием базу данных необходимо подключить.

- выберите экземпляр SQL Server, куда устанавливаются образцы базы данных и примеры;
- завершить шаги мастера;
- для завершения установки примеров после завершения установки выполните следующие шаги:
  - в меню **Пуск** последовательно выберите пункты **Программы, Microsoft SQL Server, Документация и учебные материалы, Примеры и Примеры Microsoft SQL Server**;
  - либо при помощи проводника Windows перейдите в папку <диск>:\Program Files\Microsoft SQL Server\90\Tools\Samples\ и дважды щелкните файл **SqlServerSamples.msi** для запуска программы установки.

**Хранимая процедура** - объект базы данных, представляющий собой набор SQL-инструкций, который компилируется один раз и хранится на сервере. Хранимые процедуры очень похожи на обыкновенные процедуры языков высокого уровня, у них могут быть входные и выходные параметры и локальные переменные, в них могут производиться числовые вычисления и операции над символьными данными, результаты которых могут

присваиваться переменным и параметрам. В хранимых процедурах могут выполняться стандартные операции с базами данных (как DDL – Data Definition Language, так и DML - Data Manipulation Language). Кроме того, в хранимых процедурах возможны циклы и ветвления, то есть в них могут использоваться инструкции управления потоком.

**Реализация хранимых процедур.** Хранимые процедуры обычно создаются с помощью языка SQL или конкретной его реализации в выбранной СУБД. Для этих целей в СУБД Microsoft SQL Server существует язык Transact-SQL.

Хранимая процедура создается оператором языка SQL CREATE PROCEDURE и имеет следующий синтаксис:

```
CREATE PROCEDURE name
[(param <datatype> [, param <datatype> ...])]
[RETURNS <datatype> [, param <datatype> ...]]
AS <procedure_body> [terminator]

<procedure_body> =
[<variable_declaration_list>]
<block>

<variable_declaration_list> =
DECLARE VARIABLE var <datatype>;

[DECLARE VARIABLE var <datatype>; ...]

<block> =
BEGIN
<compound_statement>
[<compound_statement> ...]
END

<compound_statement> = {<block> | statement;}

<datatype> = {
{SMALLINT | INTEGER | FLOAT | DOUBLE PRECISION}
| {DECIMAL | NUMERIC} [(precision [, scale])]
| DATE
| {CHAR | CHARACTER | CHARACTER VARYING |
VARCHAR} [(int)]
| [CHARACTER SET charname]
| {NCHAR | NATIONAL CHARACTER | NATIONAL
CHAR}
```

[VARYING] [(int)]}

**Триггер** - триггером называется именованный набор прекомпилированных команд SQL, хранящийся на сервере, который автоматически выполняется, когда происходит заданное событие. Например, триггер может выполняться при вставке, изменении или удалении строки в существующей таблице. Триггер сообщает СУБД, как нужно выполнять команды SQL INSERT, UPDATE или DELETE, чтобы выполнялись нормальные правила (business rules) организации.

**Триггер ссылочной целостности** - особый вид триггера, используемый для поддержания целостности между двумя таблицами, которые связаны между собой. Если строка в одной таблице вставляется, изменяется или удаляется, то триггер ссылочной целостности сообщает СУБД, что нужно делать с теми строками в других таблицах, у которых значение внешнего ключа совпадает со значением первичного ключа вставленной (измененной, удаленной) строки.

Триггер создается оператором языка SQL CREATE TRIGGER:

```
CREATE TRIGGER name FOR table
  [ACTIVE | INACTIVE]
  {BEFORE | AFTER}
  {DELETE | INSERT | UPDATE}
  [POSITION number]
  AS <trigger_body> terminator

<trigger_body> =
  [<variable_declaration_list>] <block>

<variable_declaration_list> =
  DECLARE VARIABLE variable <datatype>;
  [DECLARE VARIABLE variable <datatype>; ...]

<block> =
  BEGIN
    <compound_statement>
    [<compound_statement> ...]
  END

<compound_statement> = {<block> | statement;}

<datatype> = {
  {SMALLINT | INTEGER | FLOAT | DOUBLE PRECISION}
  | {DECIMAL | NUMERIC} [(precision [, scale))]
  | DATE
```

```

        | {CHAR | CHARACTER | CHARACTER VARYING |
VARCHAR}
        [(1...32767)] [CHARACTER SET char name]
        | {NCHAR | NATIONAL CHARACTER | NATIONAL
CHAR}
        [VARYING] [(1...32767)]}

```

## Оборудование и материалы

Для выполнения лабораторной работы рекомендуется использовать персональный компьютер со следующими характеристиками: 32-разрядный (x86) или 64-разрядный (x64) процессор с тактовой частотой 1 ГГц и выше, оперативная память – 1 Гб и выше, свободное дисковое пространство – не менее 1 Гб, графическое устройство DirectX 9. Программное обеспечение: операционная система WINDOWS 7 и выше, SQL Server.

## Указания по технике безопасности

Техника безопасности при выполнении лабораторной работы совпадает с общепринятой для пользователей персональных компьютеров. Самостоятельно не производить ремонт персонального компьютера, установку и удаление программного обеспечения; в случае неисправности персонального компьютера сообщить об этом обслуживающему персоналу лаборатории (оператору, администратору); соблюдать правила техники безопасности при работе с электрооборудованием; не касаться электрических розеток металлическими предметами; рабочее место пользователя персонального компьютера должно содержаться в чистоте; не разрешается возле персонального компьютера принимать пищу, напитки.

## Задания

- 1) Написать хранимую процедуру для получения списка сотрудников в формате (Фамилия Имя).
- 2) Написать хранимую процедуру для получения списка заказчиков, проживающих в городе.
- 3) Написать хранимую процедуру для получения списка всех наименований валют.
- 4) Написать хранимую процедуру для получения списка заказчиков – мужчин.
- 5) Написать хранимую процедуру для получения списка сотрудников – мужчин, имеющих день рождения в течение от 01 января 1979 г. по 01 января 1981 г.
- 6) Написать триггер, обеспечивающий целостность данных при удалении записи из таблицы DimProductSubcategory.
- 7) Написать триггер, обеспечивающий целостность данных при удалении записи из таблицы DimProduct.

## Содержание отчета и его форма



- краткий конспект теоретического обоснования;
- описание основных этапов выполнения работы.

### **Контрольные вопросы**

- 1) Каковы основные конструкции языка SQL?
- 2) Как создается хранимая процедура? Для чего они нужны?
- 3) Как создается триггер? Для чего они нужны?
- 4) Какие операторы языка SQL допустимы внутри триггера?
- 5) Как реализуется механизм событий?
- 6) Как указать действия, которые следует предпринять в случае ошибки?
- 7) Как удалить хранимую процедуру, триггер из базы данных?
- 8) С какими моделями транзакций связано исполнение триггеров?
- 9) Объясните структуру файла хранимой процедуры (триггера)? Из каких конструкций он состоит?

### **Список литературы, рекомендуемый к использованию по данной теме**

1. Лисьев, Г.А. Технологии поддержки принятия решений : учебное пособие / Г.А. Лисьев, И.В. Попова. - 2-е изд., стереотип. - М. : Флинта, 2011. - 133 с. - ISBN 978-5-9765-1300-6 ; То же [Электронный ресурс]. - URL: <http://biblioclub.ru/index.php?page=book&id=103806>

2. Соловьев, Н. Основы теории принятия решений для программистов : учебное пособие / Н. Соловьев, Е. Чернопрудова, Д.А. Лесовой ; Министерство образования и науки Российской Федерации, Федеральное государственное бюджетное образовательное учреждение высшего профессионального образования «Оренбургский государственный университет». - Оренбург : ОГУ, 2012. - 187 с. : табл., схем. ; То же [Электронный ресурс]. - URL: <http://biblioclub.ru/index.php?page=book&id=270301>

3. Балдин, К.В. Управленческие решения : учебник / К.В. Балдин, С.Н. Воробьев, В.Б. Уткин. - 7-е изд. - М. : Дашков и Ко, 2012. - 495 с. - ISBN 978-5-394-01458-1 ; То же [Электронный ресурс]. - URL: <http://biblioclub.ru/index.php?page=book&id=83585>

## Лабораторная работа 2. Подготовка базы данных SQL Server

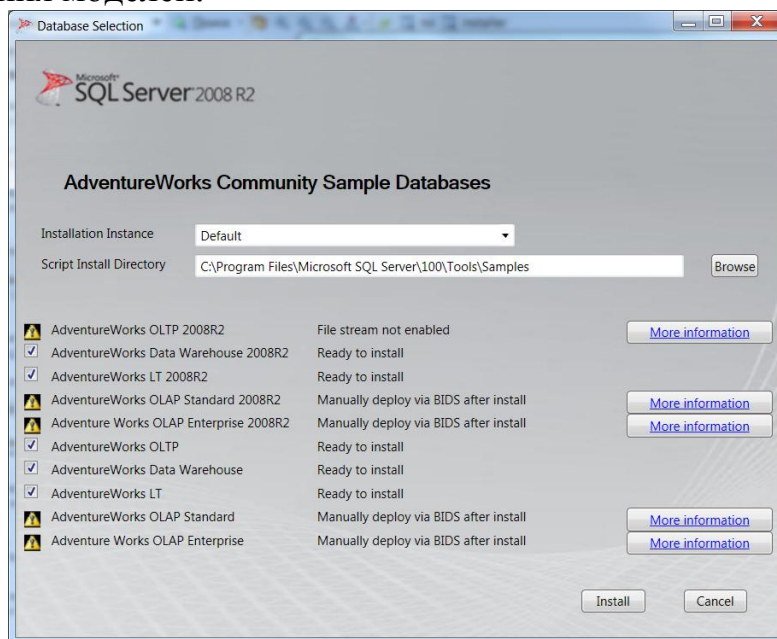
**Тема** - Подготовка базы данных SQL Server

**Цель работы** – познакомиться со средами разработки SQL Management Studio, BIDS.

**Формируемые компетенции или их части** - ОПК-1; ОПК-2; ОПК-4; ПК-2

**Теоретическая часть.**

База данных AdventureWorksDW, устанавливается вместе с SQL Server и уже содержит представления (views), которые будут использованы для создания моделей.



**Рисунок 1. Окно просмотра модели базы Adventure Works DW**

Разработка приложения бизнес-аналитики в SQL Server 2008 R2 начинается с создания в среде Business Intelligence Development Studio проекта служб SQL Server 2008 Analysis Services (SSAS) на основе шаблона проекта служб Analysis Services. После создания проекта служб Analysis Services необходимо задать один или несколько источников данных. Затем следует указать единое представление метаданных, называемое представлением источника данных, из выбранных таблиц и представлений источников данных.

На этом занятии предстоит создать проект служб Analysis Services, указать один источник данных и добавить подмножество таблиц в представление источника данных.

Для установки БД использования служб Analysis Services потребуется перенести файлы БД в рабочий каталог сервера и подключить их средствами VS.

Прежде чем вы начнете создавать и работать с моделями Data Mining, вы должны выполнить следующие действия:

1. Создать новый проект Analysis Services.
2. Создать новый источник данных (data source).

3. Создать новое представление данных (data source view).

### **Создание проекта Analysis Services**

Каждый проект Analysis Services определяет схему объектов для единственной базы данных Analysis Services, которая описывается моделями Data Mining, OLAP-кубами и дополнительными объектами.

1. Откройте Business Intelligence Development Studio.
2. Выберите **New** и **Project** из меню **File**.
3. Выберите проект Analysis Services в качестве типа нового проекта и назовите его **AdventureWorks**.
4. Нажмите **Ok**.

Новый проект откроется в Business Intelligence Development Studio.

Представление данных (Data Source View) обеспечивает уровень абстракции для источника данных, позволяя вам менять структуру исходных данных для удобства вашего проекта. Используя представление данных, вы можете выбрать только относящиеся к вашему проекту таблицы и представления, определить связи между таблицами, добавить вычисляемые поля, а также именованные запросы без необходимости вносить модификации в исходные данные.

1. В дереве решения щелкните правой кнопкой мыши на **Data Source View** и выберите **New Data Source View**.
2. На первой странице щелкните **Next**.
3. Выберите источник данных **Adventure Works DW**, который вы создали на прошлом шаге в окне **Relational data sources**. Нажмите **Next**.
4. Если вы хотите создать новый источник данных, нажмите **New Data Source**.
5. Выберите таблицы из списка нажмите на правую стрелку для включения их в представление данных:
  - ProspectiveBuyer
  - vAssocSeqLineItems
  - vAssocSeqOrders
  - vTargetMail
  - vTimeSeries
6. Нажмите **Next**.
7. По умолчанию представление данных будет названо Adventure Works DW. Нажмите **Finish**.

Далее можно производить изменения в представлении источника данных.

### **Оборудование и материалы**

Для выполнения лабораторной работы рекомендуется использовать персональный компьютер со следующими характеристиками: 32-разрядный (x86) или 64-разрядный (x64) процессор с тактовой частотой 1 ГГц и выше, оперативная память – 1 Гб и выше, свободное дисковое пространство – не менее 1 Гб, графическое устройство DirectX 9. Программное обеспечение: операционная система WINDOWS 7 и выше, SQL Server.

## **Указания по технике безопасности**

Техника безопасности при выполнении лабораторной работы совпадает с общепринятой для пользователей персональных компьютеров. Самостоятельно не производить ремонт персонального компьютера, установку и удаление программного обеспечения; в случае неисправности персонального компьютера сообщить об этом обслуживающему персоналу лаборатории (оператору, администратору); соблюдать правила техники безопасности при работе с электрооборудованием; не касаться электрических розеток металлическими предметами; рабочее место пользователя персонального компьютера должно содержаться в чистоте; не разрешается возле персонального компьютера принимать пищу, напитки.

### **Задание 1**

Установить примеры БД Adventure Works.

Создание проекта служб Analysis Services

### **Содержание отчета и его форма**

- краткий конспект теоретического обоснования;
- описание основных этапов выполнения работы.

### **Контрольные вопросы**

1. Какие сведения хранятся в БД AdventureWorksDW?
2. Выделите основные этапы подготовки данных для работы с моделями Data Mining.

### **Список литературы, рекомендуемый к использованию по данной теме**

1. Лисьев, Г.А. Технологии поддержки принятия решений : учебное пособие / Г.А. Лисьев, И.В. Попова. - 2-е изд., стереотип. - М. : Флинта, 2011. - 133 с. - ISBN 978-5-9765-1300-6 ; То же [Электронный ресурс]. - URL: <http://biblioclub.ru/index.php?page=book&id=103806>
2. Соловьев, Н. Основы теории принятия решений для программистов : учебное пособие / Н. Соловьев, Е. Чернопрудова, Д.А. Лесовой ; Министерство образования и науки Российской Федерации, Федеральное государственное бюджетное образовательное учреждение высшего профессионального образования «Оренбургский государственный университет». - Оренбург : ОГУ, 2012. - 187 с. : табл., схем. ; То же [Электронный ресурс]. - URL: <http://biblioclub.ru/index.php?page=book&id=270301>
3. Балдин, К.В. Управленческие решения : учебник / К.В. Балдин, С.Н. Воробьев, В.Б. Уткин. - 7-е изд. - М. : Дашков и Ко, 2012. - 495 с. - ISBN 978-5-394-01458-1 ; То же [Электронный ресурс]. - URL: <http://biblioclub.ru/index.php?page=book&id=83585>

### Лабораторная работа 3. Служба Microsoft SQL Server Analysis Services (SSAS)

**Тема** - Служба Microsoft SQL Server Analysis Services (SSAS)

**Цель работы** – Рассмотреть определение источников данных и их представлений, измерений и кубов в проекте служб Analysis Services в среде BI Development Studio.

**Формируемые компетенции или их части** - ОПК-1; ОПК-2; ОПК-4; ПК-2

#### **Теоретическая часть.**

Службы SSAS обеспечивают интерактивную аналитическую обработку (OLAP) и функции интеллектуального анализа данных для приложений бизнес-аналитики. Службы Analysis Services поддерживают OLAP, позволяя разрабатывать и создавать многомерные структуры, которые содержат данные, собранные из других источников, таких как реляционные базы данных, а также управлять этими структурами. Для приложений интеллектуального анализа данных службы Analysis Services позволяют разрабатывать, создавать и наглядно представлять модели интеллектуального анализа данных, построенные на основе других источников данных, используя с этой целью широкий спектр стандартных алгоритмов интеллектуального анализа данных.

Для разработки и развертывания проектов служб Analysis Services разработчики хранилищ данных используют среду Business Intelligence Development Studio (в более поздних версиях SQL Server используется SQL Server Data Tools), а для управления этими проектами — среду SQL Server Management Studio.

#### **а) Создание нового проекта служб Analysis Services**

Для создания нового проекта, необходимо открыть среду *SQL Server Business Intelligence Development Studio*, нажав меню *File->New->Project*, после чего появится новое окно с различными шаблонами проектов. В данном окне нужно выбрать шаблон *Analysis Services Project*, затем изменить имя проекта на «**ЛР по службам Analysis Services**», затем нажать кнопку «ОК» чтобы создать новый проект.

#### **б) Определение нового источника данных**

- В обозревателе решений щелкните правой кнопкой мыши элемент **Data sources** и выберите команду **New Data Source**, в результате будет открыт мастер источников данных;
- В новом окне перечисляются все созданные ранее **Data connections**, если в этом списке отсутствует соединение данных для работы с данной лабораторией, то нужно создать новое соединение данных (Data connection);
- Нажав на кнопку «New», новое окно Connection Manager появится. В данном окне нужно выбрать поставщик «Native OLE DB\Microsoft OLE DB Provider for SQL Server», затем нужно настроить сервер,

базу данных при этом необходимо обеспечить аутентификацию если требуется;

- Выбрав созданное соединение, нужно нажать на кнопку «Finish» для завершения этапа определения нового источника данных. После этого новый источник данных появится в папке **Data Sources** текущего проекта:

с) Определение представления источника данных

### **Оборудование и материалы**

Для выполнения лабораторной работы рекомендуется использовать персональный компьютер со следующими характеристиками: 32-разрядный (x86) или 64-разрядный (x64) процессор с тактовой частотой 1 ГГц и выше, оперативная память – 1 Гб и выше, свободное дисковое пространство – не менее 1 Гб, графическое устройство DirectX 9. Программное обеспечение: операционная система WINDOWS 7 и выше, PowerView.

### **Указания по технике безопасности**

Техника безопасности при выполнении лабораторной работы совпадает с общепринятой для пользователей персональных компьютеров. Самостоятельно не производить ремонт персонального компьютера, установку и удаление программного обеспечения; в случае неисправности персонального компьютера сообщить об этом обслуживающему персоналу лаборатории (оператору, администратору); соблюдать правила техники безопасности при работе с электрооборудованием; не касаться электрических розеток металлическими предметами; рабочее место пользователя персонального компьютера должно содержаться в чистоте; не разрешается возле персонального компьютера принимать пищу, напитки.

### **Задания**

На основе базы данных AdventureWorkDW постройте проект с использованием служб Analysis Services для получения результата «продажи через Интернет» по разным периодам и по регионам:

- 1) В феврале 2001 г. для города «Victoria» в Австралии.
- 2) В январе 2002 г. для города «Nord» во Франции.
- 3) В июне 2003 г. для всех городов Канады.
- 4) В периоде от января до декабря 2005 г. для города «Bayern» в Германии.
- 5) В периоде от 01 января 2002 г. по 01 января 2003 г. для всех городов Великобритании.
- 6) В периоде от 05 октября 2003 г. по 10 декабря 2003 г. для города «Florida» США.
- 7) В периоде от 2001 г. по 2009 г. в США.

### **Содержание отчета и его форма**

- краткий конспект теоретического обоснования;
- описание основных этапов выполнения работы.

### **Контрольные вопросы**

- 1) Что такие OLTP и OLAP системы?

- 2) Средства OLAP?
- 3) OLAP клиент с локальным кубом?
- 4) OLAP клиент и ROLAP сервер?
- 5) OLAP клиент с OLAP машиной?
- 6) Что такое измерения и для чего они нужны? Какие типы измерений Вы использовали в этой лабораторной работе?
- 7) Что такое источник данных, представление источника данных и куб?
- 8) Что такие пользовательские иерархии и для чего они нужны?

**Список литературы, рекомендуемый к использованию по данной теме**

1. Лисьев, Г.А. Технологии поддержки принятия решений : учебное пособие / Г.А. Лисьев, И.В. Попова. - 2-е изд., стереотип. - М. : Флинта, 2011. - 133 с. - ISBN 978-5-9765-1300-6 ; То же [Электронный ресурс]. - URL: <http://biblioclub.ru/index.php?page=book&id=103806>

2. Соловьев, Н. Основы теории принятия решений для программистов : учебное пособие / Н. Соловьев, Е. Чернопрудова, Д.А. Лесовой ; Министерство образования и науки Российской Федерации, Федеральное государственное бюджетное образовательное учреждение высшего профессионального образования «Оренбургский государственный университет». - Оренбург : ОГУ, 2012. - 187 с. : табл., схем. ; То же [Электронный ресурс]. - URL: <http://biblioclub.ru/index.php?page=book&id=270301>

3. Балдин, К.В. Управленческие решения : учебник / К.В. Балдин, С.Н. Воробьев, В.Б. Уткин. - 7-е изд. - М. : Дашков и Ко, 2012. - 495 с. - ISBN 978-5-394-01458-1 ; То же [Электронный ресурс]. - URL: <http://biblioclub.ru/index.php?page=book&id=83585>

## Лабораторная работа 4. Построение и работа с моделями Data Mining

Тема - Построение и работа с моделями Data Mining

Цель работы – знакомство с методами решения задач Data Mining.

Формируемые компетенции или их части - ОПК-1; ОПК-2; ОПК-4; ПК-2

### Теоретическая часть.

Редактор Data Mining содержит все инструменты и средства отображения для построения и работы с моделями Data Mining. Дополнительная информация по редактору доступна в разделе "Using the Data Mining Tools" в SQL Server Books Online.

Мы будете работать со следующими сценариями:

- Целевая отправка писем (Targeted mailing)
- Прогнозирование (Forecasting)
- Рыночная корзина (Market basket)
- Кластеризация последовательностей (Sequence clustering)

В сценарии «Целевая отправка писем» вы будете заниматься построением моделей, сравнением их ожидаемых возможностей (при помощи окна Mining Accuracy Chart), а также созданием прогнозов с использованием Prediction Query Builder.

В процессе решения этой задачи, отделу маркетинга потребуется выполнить следующие действия:

- Установить модели, способные выявить наиболее вероятных клиентов из списка потенциальных покупателей
- Кластеризовать своих существующих клиентов

Для выполнения этого сценария вам потребуется использовать Microsoft Native Bayes, Microsoft Decision Trees, и Microsoft Clustering алгоритмы. Сценарий содержит в себе 5 задач:

#### Создание проекта служб Analysis Services

- Откройте SQL Server Data Tools;
- Создайте новый проект типа Analysis Services Multidimensional and Data Mining Project;
- В строке название проекта укажите **BasicDataMining** и нажмите готово

#### Создание источника данных

- Правой кнопкой мыши щелкните по узлу **Data Source** в обозревателе решений вашего проекта и выберите **New Data Source**;
- На стартовой странице мастера создания источников данных нажмите **Next**;
- На странице **Select how to define the connection** нажмите **New**, чтобы добавить новое соединение к базе данных **AdventureWorksDW**;
- В появившемся окне Connection Manager в списке Provider выберите Native OLE DB\SQL Server Native Client 11.0;



- В поле для ввода имени сервера укажите Ваш сервер, на котором стоит база данных **AdventureWorksDW**;
- На странице Log onto the server выберите Use Windows Authentication;

На странице Select or enter a database name выберите имя AdventureWorksDW и нажмите Finish. Создание представления источника данных

- Правой кнопкой мыши щелкните по узлу **Data Source View** и выберите **New Data Source View**;
- На стартовой странице мастера установки представлений источников данных нажмите **Next**;
- На странице **Select a Data Source** выберите источник данных **AdventureWorksDW**, который Вы создали на предыдущем шаге;
- На странице Select Tables and Views выберите таблицы ProspectiveBuyer (dbo) и vTargetMail (dbo) и нажмите Next;
- На странице **Completing the Wizard** замените текущее имя представления на **Targeted Mailing** и нажмите **Finish**.

Создание структуры модели Data Mining

Первый шаг заключается в создании новой структуры Data mining модели при помощи Mining Model Wizard. При этом также создаётся начальная модель, основанная на алгоритме Microsoft Decision Trees.

Для того чтобы создать структуру модели:

1. В Solution Explorer кликните правой кнопкой мыши на **MiningStructure**, выберите **New MiningStructure**. Откроется Mining Model Wizard.
2. Нажмите **Next** на странице приветствия.
3. Выберите From existing relational database or data warehouse, затем Next.
4. В разделе **Which data mining technique do you want to use?** выберите **Microsoft Decision Trees**. Вы создадите несколько моделей, опирающихся на эту начальную структуру, в основе которой лежит алгоритм Microsoft Decision Trees.
5. Нажмите **Next**. По умолчанию в качестве источника данных здесь выбран Adventure Works DW.
6. На странице **Select Datasource View Page** в списке доступных представлений выберите Targeted Mailing. Нажмите **Next**.
7. Пометьте флагом **Case** таблицу **vTargetMail**, затем нажмите **Next**.
8. На странице Specify the Training Data необходимо определить по крайней мере один столбец predictable column, одну Key Column и одну input column; В качестве predictable column выберите **BikeBuyer**;
9. В качестве Key Column установите атрибут **CustomerKey**;
10. В роли Input Column задайте следующие атрибуты:
  - Age
  - CommuteDistance
  - EnglishEducation

- EnglishOccupation
  - Gender
  - GeographyKey
  - HouseOwnerFlag
  - MaritalStatus
  - NumberCarsOwned
  - NumberChildrenAtHome
  - Region
  - TotalChildren
  - YearlyIncome
11. В самой левой колонке отметьте следующие атрибуты:
- AddressLine1
  - AddressLine2
  - DateFirstPurchase
  - EmailAddress
  - FirstName
  - LastName
11. Нажмите Next.
12. На странице Create Testing Set для поля Percentage of data for testing оставьте значение по умолчанию равное 30. В строке Maximum number of cases in testing data set установите значение 1000. Нажмите Next;
13. На странице Completing the Wizard в строке **Mining structure name** введите Targeted Mailing;
14. В строке Mining model name введите TM\_Decision\_Tree;
15. Установите чекбокс рядом с надписью Allow drill through в состояние selected.
16. Нажмите Finish. При этом откроется редактор Data mining, отображающий структуру Targeted Mailing, которую вы только что создали.

### **Оборудование и материалы**

Для выполнения лабораторной работы рекомендуется использовать персональный компьютер со следующими характеристиками: 32-разрядный (x86) или 64-разрядный (x64) процессор с тактовой частотой 1 ГГц и выше, оперативная память – 1 Гб и выше, свободное дисковое пространство – не менее 1 Гб, графическое устройство DirectX 9. Программное обеспечение: операционная система WINDOWS 7 и выше, PowerView.

### **Указания по технике безопасности**

Техника безопасности при выполнении лабораторной работы совпадает с общепринятой для пользователей персональных компьютеров. Самостоятельно не производить ремонт персонального компьютера, установку и удаление программного обеспечения; в случае неисправности персонального компьютера сообщить об этом обслуживающему персоналу лаборатории (оператору, администратору); соблюдать правила техники безопасности при работе с электрооборудованием; не касаться электрических

розеток металлическими предметами; рабочее место пользователя персонального компьютера должно содержаться в чистоте; не разрешается возле персонального компьютера принимать пищу, напитки.

### **Задание**

1. Изучить различные алгоритмы для создания моделей Data Mining, позволяющих производить различные виды анализа данных, на основе базы данных AdventureWorksDW.
2. Построение классификатора для индивидуального набора данных при помощи следующих алгоритмов:
  - a. Наивный байесовский классификатор (Naive Bayes);
  - b. Деревья решений (Classification tree).
3. Тестирование эффективности алгоритмов для анализа индивидуального набора данных.

### **Содержание отчета и его форма**

- краткий конспект теоретического обоснования;
- описание основных этапов выполнения работы.

### **Контрольные вопросы**

1. Data Mining. Основные понятия и определения. Шкалы. Задачи анализа данных.
2. Классификация. Основные понятия и определения. Правила и деревья классификации. Их соответствие.
3. Классификация. Правила классификации. Алгоритм построения 1-правил. Сверхчувствительность.
4. Классификация. Правила классификации. Алгоритм Naive Bayes.
5. Классификация. Деревья классификации. Алгоритм ID3. Недостатки алгоритма ID3.
6. Классификация. Деревья классификации. Недостатки алгоритма ID3. Алгоритм C4.5.
7. Классификация. Правила и деревья классификации. Алгоритм покрытия.

### **Список литературы, рекомендуемый к использованию по данной теме**

1. Лисьев, Г.А. Технологии поддержки принятия решений : учебное пособие / Г.А. Лисьев, И.В. Попова. - 2-е изд., стереотип. - М. : Флинта, 2011. - 133 с. - ISBN 978-5-9765-1300-6 ; То же [Электронный ресурс]. - URL: <http://biblioclub.ru/index.php?page=book&id=103806>
2. Соловьев, Н. Основы теории принятия решений для программистов : учебное пособие / Н. Соловьев, Е. Чернопрудова, Д.А. Лесовой ; Министерство образования и науки Российской Федерации, Федеральное государственное бюджетное образовательное учреждение высшего профессионального образования «Оренбургский государственный университет». - Оренбург : ОГУ, 2012. - 187 с. : табл., схем. ; То же

[Электронный ресурс]. - URL:  
<http://biblioclub.ru/index.php?page=book&id=270301>

3. Балдин, К.В. Управленческие решения : учебник / К.В. Балдин, С.Н. Воробьев, В.Б. Уткин. - 7-е изд. - М. : Дашков и Ко, 2012. - 495 с. - ISBN 978-5-394-01458-1 ; То же [Электронный ресурс]. - URL:<http://biblioclub.ru/index.php?page=book&id=83585>

## Лабораторная работа 5. Исследование моделей

**Тема** - Исследование моделей.

**Цель работы** – рассмотреть возможности SSAS по предоставлению сведений о разработанной модели данных.

**Формируемые компетенции или их части** - ОПК-1; ОПК-2; ОПК-4; ПК-2  
**Теоретическая часть.**

После того как модели обработаны, вы можете просмотреть их, используя закладку Mining Model Viewer в редакторе Data Mining. При помощи выпадающего списка **Mining Model** в верхней части закладки можно исследовать модели, входящие в структуру.

Mining Model Viewer по умолчанию открывает модель Targeted Mailing первой в структуре. Tree viewer содержит две закладки, **Decision Tree** (дерево решений) и **Dependency Network** (сеть зависимостей).

На странице **Decision Tree** вы можете просмотреть все модели деревьев, составляющие модель Targeted\_Mailing. Для каждого прогнозируемого атрибута модели существует одно дерево, если только не задействовано feature selection. Так как ваша модель содержит всего один прогнозируемый атрибут, Bike Buyer, то и дерево тоже одно. Если бы их было несколько, можно было бы выбрать нужное при помощи Tree box.

Tree viewer по умолчанию показывает только первые три уровня дерева. Чтобы увидеть больше используйте Show Level slider или Default Expansion box. Для более подробной информации по настройке Tree viewer, обратитесь к разделу "Viewing with Tree Viewer" в SQL Server Books Online.

Чтобы создать дерево, установите Show Level 5 в списке Background нажмите 1. Здесь задаётся состояние прогнозируемого атрибута. Изменяя это свойство, вы можете быстро узнать число случаев попадания покупателя велосипеда в каждый узел. Чем темнее узел, тем больше случаев попадания он содержит.

Каждый узел в дереве решений несёт в себе следующую информацию:

- Условие, по которому попадаем в этот узел из предыдущего. Полный путь к узлу можно узнать из легенды или из подсказки.
- Гистограмма, описывающая распределение состояний для прогнозируемого столбца. Число состояний, отображаемых на гистограмме, регулируется при помощи Histogram control.
- Концентрацию случаев попадания, если состояние предсказываемого значения определено в контроле Background.

Если включен режим drillthrough ("проваливание" в исходные данные), вы можете просмотреть варианты для узла, кликнув по нему правой кнопкой мыши и выбрав Drillthrough.

На странице Dependency Network показаны отношения между атрибутами, влияющими на прогнозирующую способность модели.

Используйте combo box Mining Model для перехода к модели TM\_Clustering. Окно этой модели, Cluster viewer, содержит четыре закладки: Cluster Diagram, Cluster Profiles, Cluster Characteristics, и Cluster

Discrimination. По умолчанию при первом открытии отображается страница Cluster Diagram.

Для более подробной информации по использованию Cluster viewer, обратитесь к разделу " Viewing with Cluster Viewer " в SQL Server Books Online.

На странице Cluster Diagram вы можете изучить отношения между кластерами, обнаруженными алгоритмом. Длина линий, соединяющих кластеры, отображает "близость", а их интенсивность показывает, насколько они схожи. Цвет кластера отвечает за частоту появления переменной и состояния (выбранные в полях Shading Variable и State вверху экрана) в каждом кластере. Переменная по умолчанию population, но вы можете поменять её на любую другую из модели и найти нужные вам кластеры. Используя полосу прокрутки в левой части экрана, вы можете отфильтровать слабые связи и найти кластеры, расположенные наиболее близко к друг другу.

Страница Cluster Profiles предоставляет общий вид модели TM\_Clustering. Каждому кластеру модели на этой странице соответствует столбец с описывающими его данными. В первом столбце находится список атрибутов, с которыми связан по крайней мере один кластер. Все оставшиеся столбцы отображают распределение состояний этих атрибутов для каждого кластера. Распределение дискретной переменной показывается цветным столбцом, максимальное число отображаемых цветов-состояний определяется в поле Bars per histogram. Непрерывные значения отображаются diamond диаграммой, показывающей среднее и стандартное отклонение в каждом кластере.

### **Оборудование и материалы**

Для выполнения лабораторной работы рекомендуется использовать персональный компьютер со следующими характеристиками: 32-разрядный (x86) или 64-разрядный (x64) процессор с тактовой частотой 1 ГГц и выше, оперативная память – 1 Гб и выше, свободное дисковое пространство – не менее 1 Гб, графическое устройство DirectX 9. Программное обеспечение: операционная система WINDOWS 7 и выше, SQL Server.

### **Указания по технике безопасности**

Техника безопасности при выполнении лабораторной работы совпадает с общепринятой для пользователей персональных компьютеров. Самостоятельно не производить ремонт персонального компьютера, установку и удаление программного обеспечения; в случае неисправности персонального компьютера сообщить об этом обслуживающему персоналу лаборатории (оператору, администратору); соблюдать правила техники безопасности при работе с электрооборудованием; не касаться электрических розеток металлическими предметами; рабочее место пользователя персонального компьютера должно содержаться в чистоте; не разрешается возле персонального компьютера принимать пищу, напитки.

### **Контрольные вопросы**

1. Для чего предназначены средства по исследованию моделей?
2. Какие сведения они могут предоставить, как это может пригодиться в СППР?
3. Какие из рассмотренных алгоритмов предпочтительны для использования, в каких областях?

**Список литературы, рекомендуемый к использованию по данной теме**

1. Лисьев, Г.А. Технологии поддержки принятия решений : учебное пособие / Г.А. Лисьев, И.В. Попова. - 2-е изд., стереотип. - М. : Флинта, 2011. - 133 с. - ISBN 978-5-9765-1300-6 ; То же [Электронный ресурс]. - URL: <http://biblioclub.ru/index.php?page=book&id=103806>

2. Соловьев, Н. Основы теории принятия решений для программистов : учебное пособие / Н. Соловьев, Е. Чернопрудова, Д.А. Лесовой ; Министерство образования и науки Российской Федерации, Федеральное государственное бюджетное образовательное учреждение высшего профессионального образования «Оренбургский государственный университет». - Оренбург : ОГУ, 2012. - 187 с. : табл., схем. ; То же [Электронный ресурс]. - URL: <http://biblioclub.ru/index.php?page=book&id=270301>

3. Балдин, К.В. Управленческие решения : учебник / К.В. Балдин, С.Н. Воробьев, В.Б. Уткин. - 7-е изд. - М. : Дашков и Ко, 2012. - 495 с. - ISBN 978-5-394-01458-1 ; То же [Электронный ресурс]. - URL: <http://biblioclub.ru/index.php?page=book&id=83585>

## **Лабораторная работа 6. Тестирование точности Data Mining-моделей**

**Тема** - Тестирование точности Data Mining-моделей

**Цель работы** – выполнить проверку точности моделей. **Формируемые компетенции или их части** - ОПК-1; ОПК-2; ОПК-4; ПК-2

**Теоретическая часть.** Используя страницу Mining Accuracy Chart, вы можете вычислить, как

хорошо каждая из моделей предсказывает результат и сравнить их между собой. Этот метод сравнения также иногда называется диаграммой роста (lift chart). На странице The Mining Accuracy Chart использует данные, отделённые от первоисточника, что позволяет сравнивать прогнозы с известными результатами. Затем результаты сортируются и отображаются на графике вместе с идеальной моделью, показывая, насколько хорошо каждая модель делает прогноз. График идеальной модели соответствует теоретической модели, предсказывающей результат со 100% точностью.

Диаграмма роста является важным инструментом, поскольку она помогает различать модели практически идентичные по структуре, показывая, какая делает лучший прогноз. Подобным образом, из её рассмотрения можно сделать вывод о том, какие типы алгоритмов делают лучшие предсказания в данной ситуации. Для более подробной информации по использованию Mining Accuracy Chart, обратитесь к разделу "Comparing Data Mining Models with the Lift Chart" в SQL Server Books Online.

На первом шаге необходимо связать столбцы, используемые в модели со столбцами во входных данных. В том случае, если имена столбцов совпадают, связи создадутся автоматически.

Чтобы соотнести входные столбцы со структурой модели:

1. В поле Select Input Table(s) выберите Select case table. Откроется диалоговое окно выбора таблицы, где вы можете выбрать таблицу, содержащую данные, которые вы хотите использовать для прогнозирования и определения точности моделей. В этом руководстве используются те же данные, что и для обработки моделей.
2. В диалоговом окне Select table, раскройте базу AdventureWorks и выберите источник данных Adventure Works DW.
3. Нажмите vTargetMail. Столбцы data mining структуры автоматически будут соответствовать столбцам с теми же именами таблицы с исходными данными.

Запрос на прогноз генерируется для каждой модели Mining Structure на основании связанных столбцов. Вы можете удалить связь, выбрав линию, связывающую поля в Mining Structure и Select Input Table(s) и нажав DELETE. Также вы можете создать связь вручную, выбрав столбец в Select Input Table(s) и перетаскив его на соответствующий столбец в Mining Structure.

Для установки фильтров на входные данные вы можете использовать таблицу Filter the input data used to generate the lift chart. Вы можете



перетащить столбцы из Select Input Table(s) на таблицу с результатами или выбрать значения в выпадающих списках. Например, если необходимо учитывать только те строки в столбце Income, где его значение больше x, выберите Income в поле Field, а в поле Criteria/Argument введите >x.

Для выполнения данного упражнения вам не потребуется фильтровать данные.

Следующий шаг заключается в выборе моделей, которые вы хотите включить в рассмотрение, и прогнозируемого поля, по которому они будут сравниваться. По умолчанию, выбраны все модели.

Чтобы показать диаграмму lift chart:

1. Для каждой модели в поле **Predictable Column Name** выберите **Bike Buyer**.
2. Для каждой модели в поле **Predict Value** выберите 1.

Чтобы показать диаграмму точности моделей:

В поле Predictable Column Name, выберите Bike Buyer. Оставьте поле Predict Value пустым.

Если установлен флаг Synchronize Prediction Columns and Values, прогнозируемый столбец синхронизируется для каждой модели в структуре Data Mining.

### **Оборудование и материалы**

Для выполнения лабораторной работы рекомендуется использовать персональный компьютер со следующими характеристиками: 32-разрядный (x86) или 64-разрядный (x64) процессор с тактовой частотой 1 ГГц и выше, оперативная память – 1 Гб и выше, свободное дисковое пространство – не менее 1 Гб, графическое устройство DirectX 9. Программное обеспечение: операционная система WINDOWS 7 и выше, SQL Server.

### **Указания по технике безопасности**

Техника безопасности при выполнении лабораторной работы совпадает с общепринятой для пользователей персональных компьютеров. Самостоятельно не производить ремонт персонального компьютера, установку и удаление программного обеспечения; в случае неисправности персонального компьютера сообщить об этом обслуживающему персоналу лаборатории (оператору, администратору); соблюдать правила техники безопасности при работе с электрооборудованием; не касаться электрических розеток металлическими предметами; рабочее место пользователя персонального компьютера должно содержаться в чистоте; не разрешается возле персонального компьютера принимать пищу, напитки.

### **Задание**

1. Свяжите столбцы модели с соответствующими столбцами во входных данных.
2. Установите фильтр на входные данные.
3. Выберите сравниваемые модели, а также поля для которых делается прогнозирование и их значения.

### **Содержание отчета и его форма**

- краткий конспект теоретического обоснования;
- описание основных этапов выполнения работы.

### **Контрольные вопросы**

1. Прокомментируйте полученные результаты
2. Что такое идеальная модель
3. Какая из моделей наиболее близка к идеальной, как вы думаете почему?

### **Список литературы, рекомендуемый к использованию по данной теме**

1. Лисьев, Г.А. Технологии поддержки принятия решений : учебное пособие / Г.А. Лисьев, И.В. Попова. - 2-е изд., стереотип. - М. : Флинта, 2011. - 133 с. - ISBN 978-5-9765-1300-6 ; То же [Электронный ресурс]. - URL: <http://biblioclub.ru/index.php?page=book&id=103806>

2. Соловьев, Н. Основы теории принятия решений для программистов : учебное пособие / Н. Соловьев, Е. Чернопрудова, Д.А. Лесовой ; Министерство образования и науки Российской Федерации, Федеральное государственное бюджетное образовательное учреждение высшего профессионального образования «Оренбургский государственный университет». - Оренбург : ОГУ, 2012. - 187 с. : табл., схем. ; То же [Электронный ресурс]. - URL: <http://biblioclub.ru/index.php?page=book&id=270301>

3. Балдин, К.В. Управленческие решения : учебник / К.В. Балдин, С.Н. Воробьев, В.Б. Уткин. - 7-е изд. - М. : Дашков и Ко, 2012. - 495 с. - ISBN 978-5-394-01458-1 ; То же [Электронный ресурс]. - URL: <http://biblioclub.ru/index.php?page=book&id=83585>

## Лабораторная работа 7. Создание прогнозов

**Тема** - Создание прогнозов

**Цель работы** – Рассмотреть возможности пакета BI по решению задач прогнозирования

**Формируемые компетенции или их части** - ОПК-1; ОПК-2; ОПК-4; ПК-2  
**Теоретическая часть.**

Теперь, после того как вы закончили с моделями анализа, можете перейти к созданию DMX запросов, используя Prediction Query Builder. Prediction Query Builder аналогичен Access Query Builder, где вы можете использовать drag-and-drop для построения запросов. Инструмент содержит три окна:

- Design
- Query
- Result

Используя окна Design и Query, вы можете создавать и просматривать запросы. Результаты выполнения отображаются на странице Result.

Первым шагом при создании запроса является выбор модели анализа и таблицы с исходными данными.

Чтобы выбрать модель и таблицу с исходными данными

1. В окне Mining Model нажмите Select model. Откроется диалоговое окно Select Mining Model. По умолчанию выбрана первая модель в структуре.
2. Переместитесь по дереву к пункту Targeted Mailing и нажмите на него.
3. В поле Select Input Table(s) нажмите Select case table. Откроется диалоговое окно Select Table.
4. Спуститесь по дереву к таблице ProspectBuyer из data source AdventureWorksDW.

После того, как вы указали таблицу с исходными данными, Prediction Query Builder создаст связи по умолчанию между моделью и таблицей, основанные на совпадении имён столбцов.

Чтобы построить запрос на прогноз

1. В поле Source, нажмите на ячейку в первой пустой строке, затем кликните на ProspectBuyer table.
2. В поле Field, рядом с записью, созданной на первом шаге, кликните ProspectKey.

При этом добавится уникальный идентификатор к запросу, так что вы сможете определить кто склонен или наоборот, несклонен к покупке велосипеда.

3. Нажмите на следующей ячейке в столбце Source, затем кликните на Targeted Mailing.
4. В ячейке Field, выберите Bike Buyer. Таким образом, определяется, что для создания прогнозов будет использована модель Microsoft Clustering из структуры Targeted Mailing.

5. Нажмите на следующей ячейке в столбце Source, затем нажмите Prediction Function.
6. Рядом с Prediction Function, в поле Field, нажмите PredictProbability. Функции прогнозирования дают информацию о том, как делается прогноз. Функция PredictProbability предоставляет информацию о вероятности правильного прогноза. Вы можете определить параметры функции в столбце Criteria/Argument.
7. В поле Criteria/Argument, наберите [Targeted Mailing].[Bike Buyer]. Этим действием назначается целевой столбец для функции PredictProbability.

Для более подробной информации о функциях, обратитесь к разделу "DMX Function Reference" в SQL Server Books Online.

Нажав на иконку в левом верхнем углу, можно переключиться в Query view и просмотреть DMX код, созданный Prediction Query Builder. Кроме того, вы можете исполнить запрос, изменить и затем исполнить его, однако изменённый запрос не отобразится при переходе назад в Design View.

### **Оборудование и материалы**

Для выполнения лабораторной работы рекомендуется использовать персональный компьютер со следующими характеристиками: 32-разрядный (x86) или 64-разрядный (x64) процессор с тактовой частотой 1 ГГц и выше, оперативная память – 1 Гб и выше, свободное дисковое пространство – не менее 1 Гб, графическое устройство DirectX 9. Программное обеспечение: операционная система WINDOWS 7 и выше, SQL Server.

### **Указания по технике безопасности**

Техника безопасности при выполнении лабораторной работы совпадает с общепринятой для пользователей персональных компьютеров. Самостоятельно не производить ремонт персонального компьютера, установку и удаление программного обеспечения; в случае неисправности персонального компьютера сообщить об этом обслуживающему персоналу лаборатории (оператору, администратору); соблюдать правила техники безопасности при работе с электрооборудованием; не касаться электрических розеток металлическими предметами; рабочее место пользователя персонального компьютера должно содержаться в чистоте; не разрешается возле персонального компьютера принимать пищу, напитки.

### **Контрольные вопросы**

1. Для решения каких практических задач используется прогнозирование?
2. Какие инструменты для этого предлагается использовать в составе ПО от Microsoft?

### **Список литературы, рекомендуемый к использованию по данной теме**

1. Лисьев, Г.А. Технологии поддержки принятия решений : учебное пособие / Г.А. Лисьев, И.В. Попова. - 2-е изд., стереотип. - М. : Флинта, 2011.

- 133 с. - ISBN 978-5-9765-1300-6 ; То же [Электронный ресурс]. - URL: <http://biblioclub.ru/index.php?page=book&id=103806>

2. Соловьев, Н. Основы теории принятия решений для программистов : учебное пособие / Н. Соловьев, Е. Чернопрудова, Д.А. Лесовой ; Министерство образования и науки Российской Федерации, Федеральное государственное бюджетное образовательное учреждение высшего профессионального образования «Оренбургский государственный университет». - Оренбург : ОГУ, 2012. - 187 с. : табл., схем. ; То же [Электронный ресурс]. - URL: <http://biblioclub.ru/index.php?page=book&id=270301>

3. Балдин, К.В. Управленческие решения : учебник / К.В. Балдин, С.Н. Воробьев, В.Б. Уткин. - 7-е изд. - М. : Дашков и Ко, 2012. - 495 с. - ISBN 978-5-394-01458-1 ; То же [Электронный ресурс]. - URL: <http://biblioclub.ru/index.php?page=book&id=83585>

## Лабораторная работа 8. Анализ точности прогноза и использование модели интеллектуального анализа

**Тема** - Анализ точности прогноза и использование модели интеллектуального анализа.

**Цель работы** – освоить методы оценки точности прогноза и использование модели интеллектуального анализа

**Формируемые компетенции или их части** - ОПК-1; ОПК-2; ОПК-4; ПК-2

**Теоретическая часть.**

В предыдущей лабораторной работе мы создали модель для классификации клиентов магазина с целью определить, сделает ли данный клиент покупку или нет. Следующая задача - оценить *точность* построенной модели интеллектуального анализа. Для этого можно использовать инструменты из группы **Accuracy and Validation**.

*Диаграмма точности (AccuracyChart)* позволяет, применив модель на тестовой выборке данных, оценить результаты ее работы.

Запустим инструмент **AccuracyChart**. Первое окно мастера содержит краткое описание инструмента, в следующем - надо указать структуру или модель. Если для одной структуры определены несколько моделей, *по* диаграмме можно будет провести их сравнительный *анализ*. Следующее окно служит для выбора предсказываемого параметра и его значения. В нашем случае *параметр* - **BikeBuyer**, а оценивать будем *точность* предсказания значения "Yes". Далее требуется указать *источник данных* для тестирования. Это могут быть зарезервированные при создании модели данные, данные из таблицы или диапазона ячеек *Excel*, или из внешнего источника. Сейчас выберем данные из модели. В случае указания таблицы *Excel*, надо описать соответствие столбцов в модели и используемой для тестирования таблице. После этого будут сформированы и помещены на новый *лист Excel* *диаграмма точности* и *таблица* со значениями, представленными на диаграмме.

Данные на диаграмме и в таблице можно интерпретировать следующим образом. Пусть нам необходимо выбрать всех клиентов, которые сделают покупки. Формируемая идеальной моделью *выборка* объемом в 11% от числа исходных записей будет включать все 100% нужных записей (в тестовом множестве их видимо чуть меньше 11%). Случайная *выборка* объемом в 11% содержит 11% нужных записей, а *выборка* такого же объема, формируемая нашей моделью - 26,58%. В выборку в 25% от общего объема данных, наша модель поместит 52,7% "правильных" клиентов и т.д. Качество прогноза падает (горизонтальный участок *зеленого графика*) после обнаружения 76% интересующих случаев. При визуальном анализе - чем ближе *график* оцениваемой модели к идеальному, тем более точный прогноз она выдает.

**Задание 1.** Постройте диаграмму точности аналогичную той, что представлена выше (используемый файл - "Образцы данных Excel").

Дополнительно постройте диаграмму для BikeBuyer = "No". Объясните различие во внешнем виде графиков.

**Задание 2.** В предыдущем задании для целей тестирования использовались данные из модели. Модель формировалась на данных из таблицы **TrainingData**. В таблице **TestingData** находятся 30% данных из исходного набора **SourceData**. Проверьте точность модели на наборе **TestingData**.

**Задание 3.** Постройте матрицу классификации, проанализируйте полученный результат.

**Задание.** Проведите описанные в работе действия. Прокомментируйте результаты.

### **Оборудование и материалы**

Для выполнения лабораторной работы рекомендуется использовать персональный компьютер со следующими характеристиками: 32-разрядный (x86) или 64-разрядный (x64) процессор с тактовой частотой 1 ГГц и выше, оперативная память – 1 Гб и выше, свободное дисковое пространство – не менее 1 Гб, графическое устройство DirectX 9. Программное обеспечение: операционная система WINDOWS 7 и выше, SQL Server.

### **Указания по технике безопасности**

Техника безопасности при выполнении лабораторной работы совпадает с общепринятой для пользователей персональных компьютеров. Самостоятельно не производить ремонт персонального компьютера, установку и удаление программного обеспечения; в случае неисправности персонального компьютера сообщить об этом обслуживающему персоналу лаборатории (оператору, администратору); соблюдать правила техники безопасности при работе с электрооборудованием; не касаться электрических розеток металлическими предметами; рабочее место пользователя персонального компьютера должно содержаться в чистоте; не разрешается возле персонального компьютера принимать пищу, напитки.

**Задание.** Выполните запрос к модели интеллектуального анализа. Оцените полученные результаты.

### **Содержание отчета и его форма**

- краткий конспект теоретического обоснования;
- описание основных этапов выполнения работы.

### **Контрольные вопросы:**

1. Основные этапы использования конструктора запросов

### **Список литературы, рекомендуемый к использованию по данной теме**

1. Лисьев, Г.А. Технологии поддержки принятия решений : учебное пособие / Г.А. Лисьев, И.В. Попова. - 2-е изд., стереотип. - М. : Флинта, 2011.

- 133 с. - ISBN 978-5-9765-1300-6 ; То же [Электронный ресурс]. - URL: <http://biblioclub.ru/index.php?page=book&id=103806>

2. Соловьев, Н. Основы теории принятия решений для программистов : учебное пособие / Н. Соловьев, Е. Чернопрудова, Д.А. Лесовой ; Министерство образования и науки Российской Федерации, Федеральное государственное бюджетное образовательное учреждение высшего профессионального образования «Оренбургский государственный университет». - Оренбург : ОГУ, 2012. - 187 с. : табл., схем. ; То же [Электронный ресурс]. - URL: <http://biblioclub.ru/index.php?page=book&id=270301>

3. Балдин, К.В. Управленческие решения : учебник / К.В. Балдин, С.Н. Воробьев, В.Б. Уткин. - 7-е изд. - М. : Дашков и Ко, 2012. - 495 с. - ISBN 978-5-394-01458-1 ; То же [Электронный ресурс]. - URL: <http://biblioclub.ru/index.php?page=book&id=83585>



## Лабораторная работа 9. Построение модели кластеризации, трассировка и перекрестная проверка

**Тема** - Построение модели кластеризации, трассировка и перекрестная проверка

**Цель работы** – освоить методы построения модели кластеризации, трассировки и перекрестной проверки.

**Формируемые компетенции или их части** - ПК-2, ППК-1

**Теоретическая часть.**

Рассмотрим еще ряд возможностей, предоставляемых надстройками интеллектуального анализа данных.

Пусть необходимо провести сегментацию клиентов *Интернет* магазина, *список* которых находится в файле *Excel*. Если использовать **TableAnalysisTools**, для решения этой задачи надо применить инструмент **DetectCategories**. Также можно воспользоваться средствами **DataMiningClientforExcel**, где выбрать инструмент **Cluster**

Итак, откроем *файл* с образцами данных, идущий с надстройками интеллектуального анализа, перейдем на *лист* **TableAnalysisToolsSample** (или можно с первого листа с оглавлением перейти по ссылке "**Образцы данных для средств анализа таблиц**") и запустим инструмент **Cluster**.

**Задание 1.** По аналогии с рассмотренным примером создайте модель кластеризации. Изучите и проанализируйте полученные результаты. Откройте окно трассировки, проанализируйте отправляемые на сервер запросы.

**Задание 2.** Выполните перекрестную проверку для созданной модели интеллектуального анализа. Опишите и проанализируйте полученные результаты.

### Оборудование и материалы

Для выполнения лабораторной работы рекомендуется использовать персональный компьютер со следующими характеристиками: 32-разрядный (x86) или 64-разрядный (x64) процессор с тактовой частотой 1 ГГц и выше, оперативная память – 1 Гб и выше, свободное дисковое пространство – не менее 1 Гб, графическое устройство DirectX 9. Программное обеспечение: операционная система WINDOWS 7 и выше, SQL Server.

### Указания по технике безопасности

Техника безопасности при выполнении лабораторной работы совпадает с общепринятой для пользователей персональных компьютеров. Самостоятельно не производить ремонт персонального компьютера, установку и удаление программного обеспечения; в случае неисправности персонального компьютера сообщить об этом обслуживающему персоналу лаборатории (оператору, администратору); соблюдать правила техники безопасности при работе с электрооборудованием; не касаться электрических розеток металлическими предметами; рабочее место пользователя персонального компьютера должно содержаться в чистоте; не разрешается возле персонального компьютера принимать пищу, напитки.

**Список литературы, рекомендуемый к использованию по данной теме**

1. Лисьев, Г.А. Технологии поддержки принятия решений : учебное пособие / Г.А. Лисьев, И.В. Попова. - 2-е изд., стереотип. - М. : Флинта, 2011. - 133 с. - ISBN 978-5-9765-1300-6 ; То же [Электронный ресурс]. - URL: <http://biblioclub.ru/index.php?page=book&id=103806>

2. Соловьев, Н. Основы теории принятия решений для программистов : учебное пособие / Н. Соловьев, Е. Чернопрудова, Д.А. Лесовой ; Министерство образования и науки Российской Федерации, Федеральное государственное бюджетное образовательное учреждение высшего профессионального образования «Оренбургский государственный университет». - Оренбург : ОГУ, 2012. - 187 с. : табл., схем. ; То же [Электронный ресурс]. - URL: <http://biblioclub.ru/index.php?page=book&id=270301>

3. Балдин, К.В. Управленческие решения : учебник / К.В. Балдин, С.Н. Воробьев, В.Б. Уткин. - 7-е изд. - М. : Дашков и Ко, 2012. - 495 с. - ISBN 978-5-394-01458-1 ; То же [Электронный ресурс]. - URL: <http://biblioclub.ru/index.php?page=book&id=83585>

### **Список использованных источников**

1. Лисьев, Г.А. Технологии поддержки принятия решений : учебное пособие / Г.А. Лисьев, И.В. Попова. - 2-е изд., стереотип. - М. : Флинта, 2011. - 133 с. - ISBN 978-5-9765-1300-6 ; То же [Электронный ресурс]. - URL: <http://biblioclub.ru/index.php?page=book&id=103806>

2. Соловьев, Н. Основы теории принятия решений для программистов : учебное пособие / Н. Соловьев, Е. Чернопрудова, Д.А. Лесовой ; Министерство образования и науки Российской Федерации, Федеральное государственное бюджетное образовательное учреждение высшего профессионального образования «Оренбургский государственный университет». - Оренбург : ОГУ, 2012. - 187 с. : табл., схем. ; То же [Электронный ресурс]. - URL: <http://biblioclub.ru/index.php?page=book&id=270301>

3. Балдин, К.В. Управленческие решения : учебник / К.В. Балдин, С.Н. Воробьев, В.Б. Уткин. - 7-е изд. - М. : Дашков и Ко, 2012. - 495 с. - ISBN 978-5-394-01458-1 ; То же [Электронный ресурс]. - URL:<http://biblioclub.ru/index.php?page=book&id=83585>

МИНИСТЕРСТВО НАУКИ И ВЫСШЕГО ОБРАЗОВАНИЯ  
РОССИЙСКОЙ ФЕДЕРАЦИИ  
ФЕДЕРАЛЬНОЕ ГОСУДАРСТВЕННОЕ АВТОНОМНОЕ  
ОБРАЗОВАТЕЛЬНОЕ УЧРЕЖДЕНИЕ ВЫСШЕГО ОБРАЗОВАНИЯ  
«СЕВЕРО-КАВКАЗСКИЙ ФЕДЕРАЛЬНЫЙ УНИВЕРСИТЕТ»

# МЕТОДИЧЕСКИЕ УКАЗАНИЯ

по организации самостоятельной работы  
по дисциплине «Машинное обучение»  
для магистрантов направления 09.04.03 «Прикладная  
информатика» направленность (профиль) «Управление  
знаниями»

Ставрополь  
2022

## **Введение**

Дисциплина «Машинное обучение», формирует профессиональные навыки магистранта по направлению подготовки 09.04.03 «Прикладная информатика».

Изучение дисциплины «Машинное обучение» базируется на проведении лабораторных работ и индивидуальных занятий под руководством преподавателя, самостоятельной работе студентов

Задачами курса являются:

- формирование у магистра личностных и профессиональных качеств, позволяющих осуществлять профессиональную деятельность, связанную с анализом, разработкой и внедрением информационно-аналитических систем;
- изучение магистрами проблематики и областей использования методов автоматизации анализа информационной подготовки принятия управленческих решений с употреблением современных инструментальных средств широкого применения и специализированных пакетов прикладных программ;
- познание основ проблематики и областей использования искусственного интеллекта, экспертных и основанных на знаниях систем.

**1. Общая характеристика самостоятельной работы студента при изучении дисциплины** Машинное обучение – формирование набора профессиональных компетенций будущего магистра по направлению подготовки 09.04.03 «Прикладная информатика».

**2. Задачи освоения дисциплины:**

- формирование у магистра личностных и профессиональных качеств, позволяющих осуществлять профессиональную деятельность, связанную с анализом, разработкой и внедрением информационно-аналитических систем;

- изучение магистрами проблематики и областей использования методов автоматизации анализа информационной подготовки принятия управленческих решений с употреблением современных инструментальных средств широкого применения и специализированных пакетов прикладных программ;

- познание основ проблематики и областей использования искусственного интеллекта, экспертных и основанных на знаниях систем.

**3. Формируемые компетенции - ОПК-1; ОПК-2; ОПК-4; ПК-2**

**4. План-график выполнения самостоятельной работы**

| № п/п | Раздел дисциплины                                                                                                                  | Неделя семестра | СРС в часах | Формы текущего контроля успеваемости  |
|-------|------------------------------------------------------------------------------------------------------------------------------------|-----------------|-------------|---------------------------------------|
| 1     | Тема 1. Интеллектуальный анализ данных: базовые понятия                                                                            | 1.              | 7           | Собеседования по лабораторным работам |
| 2     | Тема 2. Интеллектуальный анализ данных в СУБД Microsoft SQL Server                                                                 | 2.              | 7           | Собеседования по лабораторным работам |
| 3     | Тема 3. Этапы проведения интеллектуального анализа данных                                                                          | 3.              | 7           | Собеседования по лабораторным работам |
| 4     | Тема 4. Краткий обзор алгоритмов интеллектуального анализа данных. Упрощенный алгоритм Байеса. Деревья решений. Линейная регрессия | 4.              | 7           | Собеседования по лабораторным работам |
| 5     | Тема 5. Краткий обзор алгоритмов интеллектуального анализа данных. Алгоритмы временных рядов и кластеризации                       | 5.              | 7           | Собеседования по лабораторным работам |
| 6     | Тема 6. Концепции языка DMX                                                                                                        | 6.              | 8           | Собеседования по лабораторным работам |
| 7     | Тема 7. DMX. Обработка, очистка, удаление и восстановление структур и моделей                                                      | 7.              | 8           | Собеседования по лабораторным работам |
| 8     | Тема 8. DMX. Запросы                                                                                                               | 8.              | 8           | Собеседования по лабораторным работам |

|              |                       |    |      |                                       |
|--------------|-----------------------|----|------|---------------------------------------|
| 9            | Тема 9. DMX. Прогнозы | 9. | 8,5  | Собеседования по лабораторным работам |
| <b>ИТОГО</b> |                       |    | 67,5 |                                       |

## **5. Методические рекомендации по изучению теоретического материала**

На первом этапе необходимо ознакомиться с рабочей программой дисциплины, в которой рассмотрено содержание тем дисциплины лекционного курса, взаимосвязь тем лекций с практическими занятиями, темы и виды самостоятельной работы. По каждому виду самостоятельной работы предусмотрены определённые формы отчетности.

### **Тема 1. Интеллектуальный анализ данных: базовые понятия**

1. Основные понятие и цель интеллектуального анализ данных
2. Задачи интеллектуального анализа данных

### **Список литературы, рекомендуемый к использованию по данной теме**

1. Лисьев, Г.А. Технологии поддержки принятия решений : учебное пособие / Г.А. Лисьев, И.В. Попова. - 2-е изд., стереотип. - М. : Флинта, 2011. - 133 с. - ISBN 978-5-9765-1300-6 ; То же [Электронный ресурс]. - URL: <http://biblioclub.ru/index.php?page=book&id=103806>

2. Соловьев, Н. Основы теории принятия решений для программистов : учебное пособие / Н. Соловьев, Е. Чернопрудова, Д.А. Лесовой ; Министерство образования и науки Российской Федерации, Федеральное государственное бюджетное образовательное учреждение высшего профессионального образования «Оренбургский государственный университет». - Оренбург : ОГУ, 2012. - 187 с. : табл., схем. ; То же [Электронный ресурс]. - URL: <http://biblioclub.ru/index.php?page=book&id=270301>

3. Балдин, К.В. Управленческие решения : учебник / К.В. Балдин, С.Н. Воробьев, В.Б. Уткин. - 7-е изд. - М. : Дашков и Ко, 2012. - 495 с. - ISBN 978-5-394-01458-1 ; То же [Электронный ресурс]. - URL: <http://biblioclub.ru/index.php?page=book&id=83585>

### **Тема 2. Интеллектуальный анализ данных в СУБД Microsoft SQL Server**

1. Структура подсистемы интеллектуального анализа данных
2. Алгоритмы интеллектуального анализа данных

### **Тема 3. Этапы проведения интеллектуального анализа данных**

1. Постановка задачи
2. Подготовка данных
3. Изучение данных
4. Построение моделей
5. Исследование и проверка моделей
6. Развертывание и обновление моделей

#### **Список литературы, рекомендуемый к использованию по данной теме**

1. Лисьев, Г.А. Технологии поддержки принятия решений : учебное пособие / Г.А. Лисьев, И.В. Попова. - 2-е изд., стереотип. - М. : Флинта, 2011. - 133 с. - ISBN 978-5-9765-1300-6 ; То же [Электронный ресурс]. - URL: <http://biblioclub.ru/index.php?page=book&id=103806>

2. Соловьев, Н. Основы теории принятия решений для программистов : учебное пособие / Н. Соловьев, Е. Чернопрудова, Д.А. Лесовой ; Министерство образования и науки Российской Федерации, Федеральное государственное бюджетное образовательное учреждение высшего профессионального образования «Оренбургский государственный университет». - Оренбург : ОГУ, 2012. - 187 с. : табл., схем. ; То же [Электронный ресурс]. - URL: <http://biblioclub.ru/index.php?page=book&id=270301>

3. Балдин, К.В. Управленческие решения : учебник / К.В. Балдин, С.Н. Воробьев, В.Б. Уткин. - 7-е изд. - М. : Дашков и Ко, 2012. - 495 с. - ISBN 978-5-394-01458-1 ; То же [Электронный ресурс]. - URL: <http://biblioclub.ru/index.php?page=book&id=83585>

#### **Тема 4. Краткий обзор алгоритмов интеллектуального анализа данных. Упрощенный алгоритм Байеса. Деревья решений. Линейная регрессия**

1. Упрощенный алгоритм Байеса (NaiveBayes)
2. Деревья решений (DecisionTrees)
3. Линейная регрессия

#### **Список литературы, рекомендуемый к использованию по данной теме**

1. Лисьев, Г.А. Технологии поддержки принятия решений : учебное пособие / Г.А. Лисьев, И.В. Попова. - 2-е изд., стереотип. - М. : Флинта, 2011. - 133 с. - ISBN 978-5-9765-1300-6 ; То же [Электронный ресурс]. - URL: <http://biblioclub.ru/index.php?page=book&id=103806>

2. Соловьев, Н. Основы теории принятия решений для программистов : учебное пособие / Н. Соловьев, Е. Чернопрудова, Д.А. Лесовой ; Министерство образования и науки Российской Федерации, Федеральное государственное бюджетное образовательное учреждение высшего профессионального образования «Оренбургский государственный университет». - Оренбург : ОГУ, 2012. - 187 с. : табл., схем. ; То же [Электронный ресурс]. - URL: <http://biblioclub.ru/index.php?page=book&id=270301>

3. Балдин, К.В. Управленческие решения : учебник / К.В. Балдин, С.Н. Воробьев, В.Б. Уткин. - 7-е изд. - М. : Дашков и Ко, 2012. - 495 с. - ISBN 978-5-394-01458-1 ; То же [Электронный ресурс]. - URL: <http://biblioclub.ru/index.php?page=book&id=83585>



## **Тема 5. Краткий обзор алгоритмов интеллектуального анализа данных. Алгоритмы временных рядов и кластеризации**

1. Анализ временных рядов
2. Кластеризация
3. Алгоритм взаимосвязей
4. Кластеризация последовательностей
5. Алгоритмы нейронных сетей и логистической регрессии

### **Список литературы, рекомендуемый к использованию по данной теме**

1. Лисьев, Г.А. Технологии поддержки принятия решений : учебное пособие / Г.А. Лисьев, И.В. Попова. - 2-е изд., стереотип. - М. : Флинта, 2011. - 133 с. - ISBN 978-5-9765-1300-6 ; То же [Электронный ресурс]. - URL: <http://biblioclub.ru/index.php?page=book&id=103806>

2. Соловьев, Н. Основы теории принятия решений для программистов : учебное пособие / Н. Соловьев, Е. Чернопрудова, Д.А. Лесовой ; Министерство образования и науки Российской Федерации, Федеральное государственное бюджетное образовательное учреждение высшего профессионального образования «Оренбургский государственный университет». - Оренбург : ОГУ, 2012. - 187 с. : табл., схем. ; То же [Электронный ресурс]. - URL: <http://biblioclub.ru/index.php?page=book&id=270301>

3. Балдин, К.В. Управленческие решения : учебник / К.В. Балдин, С.Н. Воробьев, В.Б. Уткин. - 7-е изд. - М. : Дашков и Ко, 2012. - 495 с. - ISBN 978-5-394-01458-1 ; То же [Электронный ресурс]. - URL: <http://biblioclub.ru/index.php?page=book&id=83585>

## **Тема 6. Концепции языка DMX**

1. Структура языка DMX
2. Создание структуры и модели

### **Список литературы, рекомендуемый к использованию по данной теме**

1. Лисьев, Г.А. Технологии поддержки принятия решений : учебное пособие / Г.А. Лисьев, И.В. Попова. - 2-е изд., стереотип. - М. : Флинта, 2011. - 133 с. - ISBN 978-5-9765-1300-6 ; То же [Электронный ресурс]. - URL: <http://biblioclub.ru/index.php?page=book&id=103806>

2. Соловьев, Н. Основы теории принятия решений для программистов : учебное пособие / Н. Соловьев, Е. Чернопрудова, Д.А. Лесовой ; Министерство образования и науки Российской Федерации, Федеральное государственное бюджетное образовательное учреждение высшего профессионального образования «Оренбургский государственный университет». - Оренбург : ОГУ, 2012. - 187 с. : табл., схем. ; То же [Электронный ресурс]. - URL: <http://biblioclub.ru/index.php?page=book&id=270301>

3. Балдин, К.В. Управленческие решения : учебник / К.В. Балдин, С.Н. Воробьев, В.Б. Уткин. - 7-е изд. - М. : Дашков и Ко, 2012. - 495 с. - ISBN 978-

5-394-01458-1 ; То же [Электронный ресурс]. - URL:<http://biblioclub.ru/index.php?page=book&id=83585>

## **Тема 7. DMX. Обработка, очистка, удаление и восстановление структур и моделей**

1. Запрос добавление
2. Создание структуры с двумя вложенными таблицами

### **Список литературы, рекомендуемый к использованию по данной теме**

1. Лисьев, Г.А. Технологии поддержки принятия решений : учебное пособие / Г.А. Лисьев, И.В. Попова. - 2-е изд., стереотип. - М. : Флинта, 2011. - 133 с. - ISBN 978-5-9765-1300-6 ; То же [Электронный ресурс]. - URL: <http://biblioclub.ru/index.php?page=book&id=103806>

2. Соловьев, Н. Основы теории принятия решений для программистов : учебное пособие / Н. Соловьев, Е. Чернопрудова, Д.А. Лесовой ; Министерство образования и науки Российской Федерации, Федеральное государственное бюджетное образовательное учреждение высшего профессионального образования «Оренбургский государственный университет». - Оренбург : ОГУ, 2012. - 187 с. : табл., схем. ; То же [Электронный ресурс]. - URL: <http://biblioclub.ru/index.php?page=book&id=270301>

3. Балдин, К.В. Управленческие решения : учебник / К.В. Балдин, С.Н. Воробьев, В.Б. Уткин. - 7-е изд. - М. : Дашков и Ко, 2012. - 495 с. - ISBN 978-5-394-01458-1 ; То же [Электронный ресурс]. - URL: <http://biblioclub.ru/index.php?page=book&id=83585>

## **Тема 8. DMX. Запросы**

1. Общая структура запроса
2. Детализация модели

### **Список литературы, рекомендуемый к использованию по данной теме**

1. Лисьев, Г.А. Технологии поддержки принятия решений : учебное пособие / Г.А. Лисьев, И.В. Попова. - 2-е изд., стереотип. - М. : Флинта, 2011. - 133 с. - ISBN 978-5-9765-1300-6 ; То же [Электронный ресурс]. - URL: <http://biblioclub.ru/index.php?page=book&id=103806>

2. Соловьев, Н. Основы теории принятия решений для программистов : учебное пособие / Н. Соловьев, Е. Чернопрудова, Д.А. Лесовой ; Министерство образования и науки Российской Федерации, Федеральное государственное бюджетное образовательное учреждение высшего профессионального образования «Оренбургский государственный университет». - Оренбург : ОГУ, 2012. - 187 с. : табл., схем. ; То же [Электронный ресурс]. - URL: <http://biblioclub.ru/index.php?page=book&id=270301>

3. Балдин, К.В. Управленческие решения : учебник / К.В. Балдин, С.Н. Воробьев, В.Б. Уткин. - 7-е изд. - М. : Дашков и Ко, 2012. - 495 с. - ISBN 978-

5-394-01458-1 ; То же [Электронный ресурс]. - URL:<http://biblioclub.ru/index.php?page=book&id=83585>

### **Тема 9. DMX. Прогнозы**

1. Одноэлементный запрос (прогноз в реальном времени)
2. Запрос с использованием функции OPENQUERY
3. Использование функции Predict

### **Список литературы, рекомендуемый к использованию по данной теме**

1. Лисьев, Г.А. Технологии поддержки принятия решений : учебное пособие / Г.А. Лисьев, И.В. Попова. - 2-е изд., стереотип. - М. : Флинта, 2011. - 133 с. - ISBN 978-5-9765-1300-6 ; То же [Электронный ресурс]. - URL: <http://biblioclub.ru/index.php?page=book&id=103806>

2. Соловьев, Н. Основы теории принятия решений для программистов : учебное пособие / Н. Соловьев, Е. Чернопрудова, Д.А. Лесовой ; Министерство образования и науки Российской Федерации, Федеральное государственное бюджетное образовательное учреждение высшего профессионального образования «Оренбургский государственный университет». - Оренбург : ОГУ, 2012. - 187 с. : табл., схем. ; То же [Электронный ресурс]. - URL: <http://biblioclub.ru/index.php?page=book&id=270301>

3. Балдин, К.В. Управленческие решения : учебник / К.В. Балдин, С.Н. Воробьев, В.Б. Уткин. - 7-е изд. - М. : Дашков и Ко, 2012. - 495 с. - ISBN 978-5-394-01458-1 ; То же [Электронный ресурс]. - URL:<http://biblioclub.ru/index.php?page=book&id=83585>

**6. Методические указания по подготовке к зачету**  
Процедура зачета как отдельное контрольное мероприятие не проводится, оценивание знаний обучающегося происходит по результатам текущего контроля.

### **7. Методические материалы, определяющие процедуры оценивания знаний, умений, навыков и (или) опыта деятельности, характеризующих этапы формирования компетенций**

Текущая аттестация студентов магистратуры проводится преподавателями, ведущими лабораторные работы по дисциплине, на основе результатов устного опроса на практических занятиях и выполнения заданий самостоятельной работы. К самостоятельному занятию магистр должен подготовить ответы на вопросы собеседования, выполнить дополнительные задания по теме занятия, проработать дополнительные научные источники, сделать конспект статьи (если это предусмотрено). Максимальную оценку студент получает, если он активно участвует в работе, владеет материалом, умеет логично и четко излагать мысли, творчески подходит к решению основных вопросов темы, показывает самостоятельность мышления.

Основанием для снижения оценки являются:

- слабое знание темы и основной терминологии;

- отсутствие умения применить теоретические знания для решения практических задач.