

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 «Многомерный анализ данных /
Multidimensional data analysis»
для магистрантов направления 09.04.03 «Прикладная
информатика» направленность (профиль) «Управление
знаниями»

Ставрополь
2026

Введение

Информационные системы серьезного предприятия, как правило, содержат приложения, предназначенные для комплексного анализа данных, их динамики, тенденций и т.п. Соответственно, основными потребителями результатов анализа становится топ-менеджмент. Такой анализ, в конечном итоге, призван содействовать принятию решений. А чтобы принять любое управленческое решение необходимо обладать необходимой для этого информацией, обычно количественной. Для этого необходимо эти данные собрать из всех информационных систем предприятия, привести к общему формату и уже потом анализировать. Для этого создают хранилища данных (Data Warehouses).

Лабораторная работа 1. Архитектура OLAP-приложений

Цель работы – приобретение практических навыков анализа архитектуры OLAP-приложений.

Формируемые компетенции или их части - ОПК-2; ПК-2; ПК-3

Теоретическая часть.

Многомерность в OLAP-приложениях может быть разделена на три уровня:

Многомерное представление данных- средства конечного пользователя, обеспечивающие многомерную визуализацию и манипулирование данными; слой многомерного представления абстрагирован от физической структуры данных и воспринимает данные как многомерные.

Многомерная обработка- средство (язык) формулирования многомерных запросов (традиционный реляционный язык SQL здесь оказывается непригодным) и процессор, умеющий обработать и выполнить такой запрос.

Многомерное хранение- средства физической организации данных, обеспечивающие эффективное выполнение многомерных запросов.

Первые два уровня в обязательном порядке присутствуют во всех OLAP-средствах. Третий уровень, хотя и является широко распространенным, не обязателен, так как данные для многомерного представления могут извлекаться и из обычных реляционных структур; процессор многомерных запросов в этом случае транслирует многомерные запросы в SQL-запросы, которые выполняются реляционной СУБД.

Конкретные OLAP-продукты, как правило, представляют собой: средство многомерного представления данных, **OLAP-клиент**(например, Pivot Tables в Excel фирмы Microsoft или ProClarity фирмы Knosys),

многомерную серверную СУБД, **OLAP- сервер**(например, Oracle Express Server или Microsoft OLAP Services).

Слой многомерной обработки обычно бывает встроен в OLAP-клиент и/или в OLAP-сервер, но может быть выделен в чистом виде, как, например, компонент Pivot Table Service фирмы Microsoft.

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, MS Excel.

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. Установите надстройки интеллектуального анализа данных для MicrosoftOffice 2007. Выполните необходимую конфигурацию MS SQL Server для работы с надстройками. Создайте и протестируйте подключение.

2. Подготовленный набор данных отформатируйте как таблицу. Убедитесь, что вы можете получить доступ к вкладке с инструментами интеллектуального анализа таблиц.

Содержание отчета и его форма

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

Список литературы, рекомендуемый к использованию по данной теме:

1. Симчера, В.М. Методы многомерного анализа статистических данных : учебное пособие / В.М. Симчера. - М. : Финансы и статистика, 2008. - 398 с. - ISBN 978-5-279-03184-9 ; То же [Электронный ресурс]. - URL: <http://biblioclub.ru/index.php?page=book&id=59559>

2. Грекул, В.И. Проектирование информационных систем [Электронный ресурс] / В.И. Грекул, Г.Н. Денищенко, Н.Л. Коровкина. - М. : Интернет-Университет Информационных Технологий, 2005. - 304 с. - ISBN 5-9556-0033-7. - URL: <http://biblioclub.ru/index.php?page=book&id=233071>

3. Терещенко, П.В. Интерфейсы информационных систем: учебное пособие [Электронный ресурс] / П.В. Терещенко, В.А. Астапчук. - Новосибирск : НГТУ, 2012. - 67 с. - ISBN 978-5-7782-2036-2. - URL: <http://biblioclub.ru/index.php?page=book&id=228775>

4. <http://www.intuit.ru/studies/courses/2312/612/info>

Лабораторная работа 2. Использование инструментов "AnalyzeKeyInfluencers" и "DetectCategories"

Тема - Использование инструментов "AnalyzeKeyInfluencers" и "DetectCategories"

Цель работы – приобретение практических навыков использования инструментов "AnalyzeKeyInfluencers" и "DetectCategories"

Формируемые компетенции или их части - ОПК-2; ПК-2; ПК-3

Теоретическая часть.

Начнем непосредственное изучение инструментов интеллектуального анализа данных. В состав пакета надстроек для MS Office 2007 входит электронная *таблица* с образцами данных. Она может быть открыта из меню **Пуск->Надстройки интеллектуального анализа данных**. Microsoft SQL Server 2008.

Инструмент **AnalyzeKeyInfluencers** позволяет определить, как зависит интересующий нас *параметр* от других. При этом важно правильно определить, что и от чего может зависеть. Собственно в этом отчасти и заключается мастерство аналитика, основанное на его знании *предметной области* и используемых методов **DM**.

В связи с тем, что мы оцениваем степень взаимного влияния разных параметров друг на друга, стоит сразу убрать из рассмотрения полностью независимые и наоборот, полностью зависимые. Пусть, например, мы хотим оценить влияние различных факторов на уровень заработной платы человека. Если у нас есть *поле*, содержащее уникальный *идентификатор* (например, порядковый номер записи в таблицы или номер паспорта), его стоит убрать из рассмотрения, как не влияющий на *значение* исследуемого параметра. Другой пример, пусть у нас есть *значение* заработной платы за месяц и за год, рассчитываемое как *заработная плата* за месяц, умноженная на 12. Мы знаем, что эти значения всегда связаны, искать зависимость одного от другого средствами **DM** не имеет смысла, а имеющаяся сильная зависимость скроет влияние других факторов, которое мы как раз и хотим выявить.

После запуска процедуры анализа будет сформирован отчет о факторах влияния и предложено формирования дополнительного сравнительного отчета.

Если добавить сравнительный отчет для двух выбранных значений, можно увидеть, чем отличается выбор в пользу одного значения целевого столбца от выбора в пользу другого.

Задание 1. Проведите анализ в соответствии с рассмотренным примером.

Задание 2. На том же наборе данных проанализируйте зависимость уровня дохода от образования, семейного положения, типа работы, пола, возраста и региона проживания клиента. Опишите результаты.

Задание 3. Предложите свой вариант анализа данных, и пример использования полученных результатов.

Инструмент **Detect Categories** позволяет решить задачу кластеризации, т.е. разделения всего *множества* вариантов на "естественные" группы, члены которых наиболее близки *по* ряду признаков. Подобная задача также называется задачей сегментации.

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, MS Excel.

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. Переименуйте категорию **Category 3**.
2. Проведите анализ параметров, характеризующих оставшиеся категории, и дайте им осмысленные названия.

Содержание отчета и его форма

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

Список литературы, рекомендуемый к использованию по данной теме:

1. Симчера, В.М. Методы многомерного анализа статистических данных : учебное пособие / В.М. Симчера. - М. : Финансы и статистика, 2008. - 398 с. - ISBN 978-5-279-03184-9 ; То же [Электронный ресурс]. - URL: <http://biblioclub.ru/index.php?page=book&id=59559>
2. Грекул, В.И. Проектирование информационных систем [Электронный ресурс] / В.И. Грекул, Г.Н. Денищенко, Н.Л. Коровкина. - М. : Интернет-Университет Информационных Технологий, 2005. - 304 с. - ISBN 5-9556-0033-7. - URL: <http://biblioclub.ru/index.php?page=book&id=233071>
3. Терещенко, П.В. Интерфейсы информационных систем: учебное пособие [Электронный ресурс] / П.В. Терещенко, В.А. Астапчук. - Новосибирск : НГТУ, 2012. - 67 с. - ISBN 978-5-7782-2036-2. - URL: <http://biblioclub.ru/index.php?page=book&id=228775>
4. <http://www.intuit.ru/studies/courses/2312/612/info>

Лабораторная работа 3. Использование инструментов "FillFromExample" и "Forecast"

Тема - Использование инструментов "FillFromExample" и "Forecast"

Цель работы – приобретение практических навыков использования инструментов "FillFromExample" и "Forecast"

Формируемые компетенции или их части - ОПК-2; ПК-2; ПК-3

Теоретическая часть.

Оба рассматриваемых инструмента используются для решения задач прогнозирования неизвестных значений параметров. Поэтому в обоих случаях требуется обучающий набор данных, на базе которого строится модель, применяемая для предсказания.

Заполнение по примеру

Нужные данные находятся на листе "**Заполнение из примера**". Здесь описывается ряд клиентов магазина. Для некоторых из них отмечено, является ли данный клиент высокодоходным. Эти строки будут использоваться как *обучающая выборка*. Задачей анализа будет являться оценка остальных клиентов по этому параметру.

Для решения этой задачи используется *алгоритм* Microsoft Logistic Regression. Необходимо понимать, что для создания модели в обучающей выборке должны быть представлены варианты со всеми возможными значениями целевого столбца. Необходимое число примеров зависит от особенностей *предметной области*. Но во многих случаях справедливо, что чем больше характерных примеров в обучающей выборке, тем более качественно будет обучена модель.

Соответственно, данный инструмент непригоден для задачи предсказания значений параметра, который может принимать непрерывные числовые значения.

Еще одна особенность - *анализ* проводится по столбцам (т.е. предсказывается *значение* столбца). Если ряд, который необходимо заполнить, хранится в виде строки, перед началом анализа надо выполнить *транспонирование* (скопировать в *буфер*, выбрать в контекстном меню "**Специальная вставка**" и отметить флажок "**Транспонировать**").

Задание. Проведите анализ и опишите полученные результаты.

Инструмент **Forecast** позволяет построить прогноз значений числового ряда. Ряд должен быть представлен столбцом в таблице (если исследуемые значения организованы в виде строки, требуется, как и в случае инструмента "**FillFromExample**", выполнить *транспонирование*).

Инструмент ищет в анализируемой последовательности шаблоны следующих типов:

- **тренд** - тенденцию изменения значений. Тренд может быть восходящим (возрастание значений ряда) или нисходящим (уменьшение значений);
- **периодичность (сезонность)** - событие повторяется через определённые интервалы;
- **взаимная корреляция** - зависимость значений одного ряда от других

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, MS Excel.

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. С помощью инструмента постройте прогноз продаж на год (12 значений). Проанализируйте график. На ваш взгляд, какой тип периодичности обнаружил инструмент в исходных данных и использует для предсказания?

Содержание отчета и его форма

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

Список литературы, рекомендуемый к использованию по данной теме:

1. Симчера, В.М. Методы многомерного анализа статистических данных : учебное пособие / В.М. Симчера. - М. : Финансы и статистика, 2008. - 398 с. - ISBN 978-5-279-03184-9 ; То же [Электронный ресурс]. - URL: <http://biblioclub.ru/index.php?page=book&id=59559>

2. Грекул, В.И. Проектирование информационных систем [Электронный ресурс] / В.И. Грекул, Г.Н. Денищенко, Н.Л. Коровкина. - М. : Интернет-

Университет Информационных Технологий, 2005. - 304 с. - ISBN 5-9556-0033-7.
- URL:<http://biblioclub.ru/index.php?page=book&id=233071>

3. Терещенко, П.В. Интерфейсы информационных систем: учебное пособие [Электронный ресурс] / П.В. Терещенко, В.А. Астапчук. - Новосибирск : НГТУ, 2012. - 67 с. - ISBN 978-5-7782-2036-2. - URL: <http://biblioclub.ru/index.php?page=book&id=228775>

4. <http://www.intuit.ru/studies/courses/2312/612/info>

Лабораторная работа 4. Использование инструментов "HighlightExceptions" и "ScenarioAnalysis"

Тема - Использование инструментов "HighlightExceptions" и "ScenarioAnalysis"

Цель работы – приобретение практических навыков использования инструментов "HighlightExceptions" и "ScenarioAnalysis"

Формируемые компетенции или их части - ОПК-2; ПК-2; ПК-3

Теоретическая часть.

Как следует из названия, инструмент позволяет выявить данные, выделяющиеся среди имеющегося набора. Это может быть полезно в ряде случаев. Во-первых, это могут быть ошибочные данные (например, результаты ошибки оператора при вводе каких-то значений). Во-вторых, исключения могут представлять отдельный интерес (как, например, в случае обнаружения мошеннических действий с банковскими картами и т.п.). Кроме того, *анализ исключений* может рассматриваться как предварительная часть интеллектуального анализа данных с помощью других методов. В частности, это позволяет исключить попадание нетипичных примеров в обучающую выборку.

В ходе работы инструмент **HighlightExceptions** создает временную модель интеллектуального анализа с использованием алгоритма MicrosoftClustering. Для каждой анализируемой строки оценивается степень принадлежности выявленным кластерам. Значения, находящиеся далеко от всех кластеров, помечаются как исключения.

Задание 1. Проведите *анализ исключений* для набора данных "**Прогнозирование**". Предложите интерпретацию полученных результатов.

Инструмент **Scenario Analysis** позволяет моделировать влияние, оказываемое изменением одного из параметров (значений одного столбца) на другой, связанный с первым. В основе работы инструмента лежит использование алгоритма **Microsoft Logistic Regression**. Для формирования временной модели требуется обучающая *выборка*, содержащая не менее 50 записей.

Инструмент **Scenario Analysis** включает две составные части - "**Анализ сценария поиска решений**" (**GoalSeek**) и "**Анализ возможных вариантов**" ("**What-If**").

Использование инструмента **GoalSeek** позволяет оценить, сможем ли мы достичь желаемого значения в целевом столбце, меняя значения выбранного параметра. Инструмент позволяет провести анализ как для одной записи, так и для всей таблицы.

Используя этот инструмент надо быть готовым, что не для всех вариантов запроса может быть получен ответ. Это может быть связано с тем, что в исходных данных нет интересующих нас сочетаний. Также могут быть проблемы из-за типов данных.

Кроме того, нельзя забывать, что запрос нужно формировать с учетом знаний о предметной области.

Задание. Проведите анализ для отдельной строки и таблицы, аналогичный описанному выше. Прокомментируйте результаты.

Инструмент **What-If** позволяет решить обратную по отношению к **GoalSeek** задачу: оценить значение целевой переменной при определенном изменении заданного параметра.

Задание. Проведите анализ 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, MS Excel.

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. Симчера, В.М. Методы многомерного анализа статистических данных : учебное пособие / В.М. Симчера. - М. : Финансы и статистика, 2008. - 398 с. - ISBN 978-5-279-03184-9 ; То же [Электронный ресурс]. - URL: <http://biblioclub.ru/index.php?page=book&id=59559>

2. Грекул, В.И. Проектирование информационных систем [Электронный ресурс] / В.И. Грекул, Г.Н. Денищенко, Н.Л. Коровкина. - М. : Интернет-Университет Информационных Технологий, 2005. - 304 с. - ISBN 5-9556-0033-7. - URL:<http://biblioclub.ru/index.php?page=book&id=233071>

3. Терещенко, П.В. Интерфейсы информационных систем: учебное пособие [Электронный ресурс] / П.В. Терещенко, В.А. Астапчук. - Новосибирск : НГТУ, 2012. - 67 с. - ISBN 978-5-7782-2036-2. - URL: <http://biblioclub.ru/index.php?page=book&id=228775>

4. <http://www.intuit.ru/studies/courses/2312/612/info>

Лабораторная работа 5. Использование инструментов "Prediction Calculator" и "ShoppingbasketAnalysis"

Тема - Использование инструментов "Prediction Calculator" и "ShoppingbasketAnalysis"

Цель работы – приобретение практических навыков использования инструментов "Prediction Calculator" и "ShoppingbasketAnalysis"

Формируемые компетенции или их части - ОПК-2; ПК-2; ПК-3

Теоретическая часть.

Инструмент **Prediction Calculator** помогает сгенерировать и настроить "калькулятор", который позволяет оценить шансы на получение ожидаемого значения целевого параметра без подключения к аналитическим службам *SQL Server*. В частности, такая возможность может быть очень полезна для удаленных пользователей.

По итогам анализа формируется прогноз, который может быть отнесен к одной из четырех категорий:

- **истинный позитивный прогноз (TruePositive)** - верный прогноз. Например, клиент, для которого прогноз показал истину, на самом деле заинтересован в покупке велосипеда. Магазин получил прибыль;

- **истинный негативный прогноз (TrueNegative)** - верный негативный прогноз. Клиент, для которого прогноз показал незаинтересованность в покупке, на самом деле не собирается покупать велосипед. Магазин не получил прибыли, но и не понес затрат (на рассылку рекламных предложений и проч.);

- **ложный позитивный прогноз (FalsePositive; ошибка 1 рода)** - неверный прогноз, показывающий, что клиент хочет сделать покупку, хотя на самом деле это не так (может привести магазин к затратам на сопровождение клиента);

- **ложный негативный прогноз (FalseNegative; ошибка 2 рода)** - неверный прогноз, показывающий, что клиент не хочет сделать покупку, хотя на самом деле он в ней заинтересован (может привести к упущенной прибыли).

Задание. Проведите анализ для двух различных наборов значений прибыли от истинных прогнозов и убытков от ложных. Прокомментируйте результаты.

Инструмент **Shopping Basket Analysis**. Он позволяет, например, на основе данных о покупках выделить товары, чаще всего встречающиеся в одном заказе, и сформировать рекомендации относительно совместных продаж.

Задание 1. Проведите анализ аналогичный описанному выше.

Задание 2. Проанализируйте, какие товары (а не категории товаров, как было раньше), приобретаются вместе. Опишите 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, MS Excel.

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. Симчера, В.М. Методы многомерного анализа статистических данных : учебное пособие / В.М. Симчера. - М. : Финансы и статистика, 2008. - 398 с. - ISBN 978-5-279-03184-9 ; То же [Электронный ресурс]. - URL: <http://biblioclub.ru/index.php?page=book&id=59559>

2. Грекул, В.И. Проектирование информационных систем [Электронный ресурс] / В.И. Грекул, Г.Н. Денищенко, Н.Л. Коровкина. - М. : Интернет-Университет Информационных Технологий, 2005. - 304 с. - ISBN 5-9556-0033-7. - URL: <http://biblioclub.ru/index.php?page=book&id=233071>

3. Терещенко, П.В. Интерфейсы информационных систем: учебное пособие [Электронный ресурс] / П.В. Терещенко, В.А. Астапчук. - Новосибирск : НГТУ, 2012. - 67 с. - ISBN 978-5-7782-2036-2. - URL: <http://biblioclub.ru/index.php?page=book&id=228775>

4. <http://www.intuit.ru/studies/courses/2312/612/info>

Лабораторная работа 6. Использование инструментов Data Mining Client для Excel 2007 для подготовки данных

Тема - Использование инструментов Data Mining Client для Excel 2007 для подготовки данных

Цель работы – приобретение практических навыков использования инструментов Data Mining Client для Excel 2007 для подготовки данных

Формируемые компетенции или их части - ОПК-2; ПК-2; ПК-3

Теоретическая часть.

Рассмотренные в предыдущих лабораторных работах "Средства анализа таблиц для Excel" (**TableAnalysisTools**) позволяют быстро провести "стандартный" анализ имеющихся данных. В то же время, этот набор инструментов не предоставляет особых возможностей по подготовке данных к анализу, оценке результатов и т.д. Из Excel это можно сделать, используя клиент интеллектуального анализа данных (**DataMiningClient**), который также входит в набор надстроек интеллектуального анализа. В ходе "Настройки интеллектуального анализа данных для MicrosoftOffice", отмечалось, что желательно сделать полную установку надстроек, в которую входит и **DataMiningClient**.

Задание. Проведите описанную в лабораторной обработку выбранного набора 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, MS Excel.

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. Симчера, В.М. Методы многомерного анализа статистических данных : учебное пособие / В.М. Симчера. - М. : Финансы и статистика, 2008. - 398 с. - ISBN 978-5-279-03184-9 ; То же [Электронный ресурс]. - URL: <http://biblioclub.ru/index.php?page=book&id=59559>
2. Грекул, В.И. Проектирование информационных систем [Электронный ресурс] / В.И. Грекул, Г.Н. Денищенко, Н.Л. Коровкина. - М. : Интернет-Университет Информационных Технологий, 2005. - 304 с. - ISBN 5-9556-0033-7. - URL: <http://biblioclub.ru/index.php?page=book&id=233071>
3. Терещенко, П.В. Интерфейсы информационных систем: учебное пособие [Электронный ресурс] / П.В. Терещенко, В.А. Астапчук. - Новосибирск : НГТУ, 2012. - 67 с. - ISBN 978-5-7782-2036-2. - URL: <http://biblioclub.ru/index.php?page=book&id=228775>
4. <http://www.intuit.ru/studies/courses/2312/612/info>

Лабораторная работа 7. Использование инструментов Data Mining Client для Excel 2007 для создания модели интеллектуального анализа данных

Цель работы – приобретение практических навыков использования инструментов Data Mining Client для Excel 2007 для создания модели интеллектуального анализа данных

Формируемые компетенции или их части - ОПК-2; ПК-2; ПК-3

Теоретическая часть.

Инструменты **Data Modeling**, позволяющие создать модели интеллектуального анализа данных:

Классификация (Classify) создает модель классификации на основе существующих данных таблицы Excel, диапазона Excel или внешнего источника данных (**AnalysisServicesDataSource**). На основе обрабатываемых данных формируются шаблоны, которые при использовании позволяют отнести рассматриваемый пример к одному из возможных классов. По умолчанию используется алгоритм **DecisionTrees**, но также доступны **LogisticRegression, NaiveBayes, NeuralNetworks**.

Оценка (Estimate) позволяет создать модель оценки значения целевого параметра (он должен быть числовым) на основе данных из таблицы или диапазона ячеек Excel либо внешнего источника данных. По умолчанию используется алгоритм **Decision Trees**, также доступны **Linear Regression, Logistic Regression, Neural Networks**.

Кластер запускает мастер, позволяющий построить модель

(Cluster)	кластеризации на основе данных из таблицы или диапазона Excel, либо внешнего источника данных. Модель определяет группы строк со сходными характеристиками, для чего используется алгоритм MicrosoftClustering . Данная задача аналогична решаемой средством DetectCategories из набора TableAnalysisTools .
Поиск взаимосвязей (Associate)	помогает создать модель, описывающую взаимосвязь объектов (покупаемых товаров и т.д.), затрагиваемых одной транзакцией, для чего используется алгоритм AssociationRules . С подобной задачей мы сталкивались, используя инструмент ShoppingBasketAnalysis из TableAnalysisTools . Для построения модели анализа необходимо, чтобы исходные данные содержали столбец с идентификатором транзакций и были по нему отсортированы. В качестве источника данных может использоваться только таблица или диапазон ячеек Excel.
Прогноз (Forecast)	Данный мастер позволяет построить модель для прогнозирования новых значений в числовой последовательности, аналогично инструменту Forecast в TableAnalysisTools . Используется алгоритм TimeSeries , для работы которого требуется, чтобы столбец (или столбцы), в отношении которого будет выполняться прогноз, имели непрерывные числовые значения. Также может присутствовать столбец с отметкой времени (в этом случае, строки в таблице должны быть по нему отсортированы).
Дополнительно (Advanced)	позволяет создать структуру ¹ интеллектуального анализа данных или добавить в существующую структуру новую модель (например, для сравнения результатов, выдаваемых разными алгоритмами анализа).

Используем инструмент **Classify**. В поставляющемся с надстройками наборе данных (меню "**Пуск**"->"**Надстройки интеллектуального анализа данных**"->"**Образцы данных Excel**") выберем таблицу **TrainingData**, содержащую случайную выборку 70% данных из таблицы **SourceData**. Запустим мастер **Classify**, в первом окне которого будет комментарий *по* применению инструмента, а второе окно позволит указать *источник данных* для анализа (таблица **TrainingData**).

Задание 1. Создайте модель интеллектуального анализа, аналогичную описанной в лабораторной работе.

Задание 2. Воспользуйтесь набором данных в таблице **Associate** и одноименным мастером для создания модели, описывающей взаимосвязи между категориями товаров в одном заказе. При необходимости, воспользуйтесь

справочной системой по инструменту. Проанализируйте выявленные правила и диаграмму зависимостей.

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, MS Excel.

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. Симчера, В.М. Методы многомерного анализа статистических данных : учебное пособие / В.М. Симчера. - М. : Финансы и статистика, 2008. - 398 с. - ISBN 978-5-279-03184-9 ; То же [Электронный ресурс]. - URL: <http://biblioclub.ru/index.php?page=book&id=59559>
2. Грекул, В.И. Проектирование информационных систем [Электронный ресурс] / В.И. Грекул, Г.Н. Денищенко, Н.Л. Коровкина. - М. : Интернет-Университет Информационных Технологий, 2005. - 304 с. - ISBN 5-9556-0033-7. - URL: <http://biblioclub.ru/index.php?page=book&id=233071>
3. Терещенко, П.В. Интерфейсы информационных систем: учебное пособие [Электронный ресурс] / П.В. Терещенко, В.А. Астапчук. - Новосибирск : НГТУ, 2012. - 67 с. - ISBN 978-5-7782-2036-2. - URL: <http://biblioclub.ru/index.php?page=book&id=228775>
4. <http://www.intuit.ru/studies/courses/2312/612/info>

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

Цель работы – приобретение практических навыков анализа точности прогноза и использование модели интеллектуального анализа

Формируемые компетенции или их части - ОПК-2; ПК-2; ПК-3

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**.

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. Выполните запрос к модели интеллектуального анализа. Оцените 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, MS Excel.

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. Симчера, В.М. Методы многомерного анализа статистических данных : учебное пособие / В.М. Симчера. - М. : Финансы и статистика, 2008. - 398 с. - ISBN 978-5-279-03184-9 ; То же [Электронный ресурс]. - URL: <http://biblioclub.ru/index.php?page=book&id=59559>

2. Грекул, В.И. Проектирование информационных систем [Электронный ресурс] / В.И. Грекул, Г.Н. Денищенко, Н.Л. Коровкина. - М. : Интернет-Университет Информационных Технологий, 2005. - 304 с. - ISBN 5-9556-0033-7. - URL:<http://biblioclub.ru/index.php?page=book&id=233071>

3. Терещенко, П.В. Интерфейсы информационных систем: учебное пособие [Электронный ресурс] / П.В. Терещенко, В.А. Астапчук. - Новосибирск : НГТУ, 2012. - 67 с. - ISBN 978-5-7782-2036-2. - URL: <http://biblioclub.ru/index.php?page=book&id=228775>

4. <http://www.intuit.ru/studies/courses/2312/612/info>

Лабораторная работа 9. Построение модели кластеризации

Цель работы – приобретение практических навыков построения модели кластеризации

Формируемые компетенции или их части - ОПК-2; ПК-2; ПК-3

Теоретическая часть.

Существуют различные подходы к классификации кластеров, при этом используется достаточно широкий ряд признаков, по которым предприятий объединяются в однородные группы (кластеры). Наиболее часто кластеры предприятий классифицируют по таким параметрам, как: географическое положение, наличие капитала, близость поставщиков, состояние конкурентоспособности, доступ к специализированным услугам, развитие трудового потенциала, наличие специализированных учебных заведений и исследовательских организаций, отраслевая принадлежность и т.д.

В научной литературе существуют различные подходы к принципам формирования кластеров. В настоящее время описаны семь основных подходов, встречающихся на практике и представляющих собой комбинацию различных факторов, на которых, в свою очередь, базируется выбор той или иной кластерной стратегии. Кратко перечислим эти модели:

1. Географическая модель предполагает построение различных по масштабам пространственных кластеров экономической активности от сугубо местных (например, садоводство в Нидерландах), до подлинно глобальных (аэрокосмический кластер).
2. Горизонтальная модель применяется для описания тех случаев, когда несколько отраслей входят в более крупный кластер.
3. Вертикальная модель: в кластерах могут присутствовать предприятия смежных этапов производственного процесса. При этом важное значение имеет, кто именно из участников сети является инициатором и конечным исполнителем инноваций в рамках кластера.
4. Литеральная модель: в кластер объединяются разные секторы, которые могут обеспечить экономию за счет эффекта масштаба, что приводит к новым комбинациям (например, мультимедийный кластер).

5. Технологическая модель: описывает совокупность предприятий различных отраслей, пользующихся одной и той же технологией (например, биотехнологический кластер).
6. Фокусная модель: кластер фирм, сосредоточенных вокруг одного центра – предприятия, НИИ или учебного центра и являющихся потребителями знаний и инноваций [3].
7. Качественная модель: характеризует в качественных категориях наличие и способы взаимодействия фирм в рамках кластера. Для нее существенным является не только то, действительно ли фирмы сотрудничают, но и то каким образом они это делают [2].

Кластеры могут варьироваться по своим размерам, широте охвата, уровню развития, в зависимости от отраслей. Природа отраслей позволяет изменять границы кластеров вследствие появления новых отраслей или фирм, а также в результате развития, либо изменения условий осуществления бизнеса. Следует отметить, что кластеры встречаются в рамках, как развитых стран, так и развивающихся. Отличием будет являться наличие в более развитых промышленных кластерах специализированных баз поставщиков, наиболее разветвленный массив связанных отраслей, более глубокие связи с потребителями [1].

В настоящее время при формировании промышленных кластеров следует опираться на комплекс методов. Основные из них:

1. Аналитический метод: исследование данных о территории, ее природных, трудовых и финансовых ресурсах производительных сил.
2. Метод индикативного планирования конкурентоспособности: количественный и качественный анализы, существующей и определение перспективной конкурентоспособности.
3. Метод моделирования: исследование объектов кластера с помощью построения и изучения моделей реально существующих организаций, процессов или явлений.
4. Метод «затраты-выпуск»: исследование циркуляции товаров и услуг между отраслями промышленности и другими секторами экономики.

На практике формирования промышленных кластеров данные методы используются в совокупности, так как взятые по отдельности они не отображают объективной экономической ситуации. Аналитический метод отображает ситуацию на рынке в конкретный момент времени (данные о предприятиях, их производительности, численности занятых и др.). Метод индикативного планирования заключается в определении экономических приоритетов. Индикативный план имеет рекомендательный характер. Главные достоинства индикативного планирования – снижение неопределенности при принятии решений. Метод моделирования – это метод, позволяющий строить модели, описывающие процессы так, как они проходили бы в действительности. Данный метод позволяет избежать каких-либо ошибок в будущем. Метод «затраты-выпуск» исследует данные о циркуляции товаров на основе комплекса

экономических показателей (производственная мощность, рост производительности труда, объем выпуска продукции и др.).

Также практикой выделяется несколько этапов создания и управления кластерами.

1-й этап: идентификация кластеров (методика их отбора). Можно выделить множество парных категорий для характеристики процесса и техник идентификации кластеров, однако большинство из них исходят из двух основных подходов. В первом, который можно назвать «снизу», кластеры идентифицируются на конкретно выбранной территории, исходя из присутствия заранее известных предприятий и отраслей-лидеров. Второй подход использует методику, условно называемую «сверху», где ищутся пространственные локализации производства, ориентированные на специфические виды экономической деятельности. Подходы к идентификации кластеров «сверху», учитывая две неизменные характеристики кластеров – функциональную связанность и географическую близость, - традиционно подразделяются на 2 типа:

- функциональные, ориентированные на выявление промышленных кластеров;
- пространственные, ориентированные на выявление географических кластеров.

2-й этап: Управление существующими кластерами.

3-й этап: Оценка экономического эффекта от создания кластера.

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, MS Excel.

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. Симчера, В.М. Методы многомерного анализа статистических данных : учебное пособие / В.М. Симчера. - М. : Финансы и статистика, 2008. - 398 с. - ISBN 978-5-279-03184-9 ; То же [Электронный ресурс]. - URL: <http://biblioclub.ru/index.php?page=book&id=59559>

2. Грекул, В.И. Проектирование информационных систем [Электронный ресурс] / В.И. Грекул, Г.Н. Денищенко, Н.Л. Коровкина. - М. : Интернет-Университет Информационных Технологий, 2005. - 304 с. - ISBN 5-9556-0033-7. - URL: <http://biblioclub.ru/index.php?page=book&id=233071>

3. Терещенко, П.В. Интерфейсы информационных систем: учебное пособие [Электронный ресурс] / П.В. Терещенко, В.А. Астапчук. - Новосибирск : НГТУ, 2012. - 67 с. - ISBN 978-5-7782-2036-2. - URL: <http://biblioclub.ru/index.php?page=book&id=228775>

4. <http://www.intuit.ru/studies/courses/2312/612/info>

Лабораторная работа 10. Создание представления источника данных

Цель работы – приобретение практических навыков создания представления источника данных

Формируемые компетенции или их части - ОПК-2; ПК-2; ПК-3

Теоретическая часть.

В представлении источников данных содержится логическая модель схемы, используемой объектами многомерной базы данных служб Службы Analysis Services, а именно кубы, измерения и структуры интеллектуального анализа данных. Представление источника данных — это определение метаданных, хранящееся в формате XML, для элементов схемы, используемых в унифицированной многомерной модели (UDM) и структурах интеллектуального анализа данных. Представление источника данных:

- содержит метаданные, представляющие выбранные объекты из одного или нескольких базовых источников данных, или метаданные, которые будут использоваться для формирования базового реляционного хранилища данных при использовании нисходящего подхода в построении схемы.
- Может быть построено на основе одного или нескольких источников данных, что позволяет определять многомерные объекты и объекты интеллектуального анализа данных, объединяющие данные из нескольких источников.
- может содержать связи, первичные ключи, имена объектов, вычисляемые столбцы и запросы, отсутствующие в базовом источнике данных и существующие отдельно от него;
- не видно, но доступно для запросов со стороны клиентских приложений.

Представление источника данных — необходимый компонент многомерной модели. Большинство разработчиков Analysis Services создают

представление источника данных на ранних стадиях проектирования модели, обеспечивая хотя бы одно представление источника данных, основанное на внешней реляционной базе данных, предоставляющей базовые данные. Однако можно также создать представление источника данных на более поздней фазе, создав схему и базовые структуры базы данных после создания измерений и кубов. Этот второй подход иногда называют нисходящим и часто используют для создания прототипов и аналитического моделирования. В этом случае воспользуйтесь мастером создания схем для создания базового представления источников данных и объектов источника данных на основе объектов OLAP, определенных в проекте Analysis Services или базе данных. Независимо от того, как и когда создается представление источника данных, оно должно быть в каждой модели, прежде чем та сможет быть обработана.

Представление источника данных содержит следующие элементы.

- Имя и описание.
- Определение любого подмножества схемы, получаемого из одного или нескольких источников данных, вплоть до полной схемы и включая ее, в том числе:
 - имена таблиц;
 - имена столбцов;
 - типы данных;
 - допустимость значений NULL;
 - длины столбцов;
 - Первичные ключи.
 - связи «первичный-внешний ключ».
- Заметки к схеме из базовых источников данных, включая следующее:
 - понятные имена таблиц, представлений и столбцов;
 - именованные запросы, возвращающие столбцы из одного или нескольких источников данных (отображающихся в схеме как таблицы);
 - именованные вычисления, возвращающие столбцы из источника данных (отображающиеся как столбцы таблиц или представлений);
 - логические первичные ключи (необходимые, если первичный ключ в базовой таблице не определен или не включен в представление или именованный запрос);
 - связи «логический первичный ключ-внешний ключ» между таблицами, представлениями и именованными запросами.

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, MS Excel.

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. Симчера, В.М. Методы многомерного анализа статистических данных : учебное пособие / В.М. Симчера. - М. : Финансы и статистика, 2008. - 398 с. - ISBN 978-5-279-03184-9 ; То же [Электронный ресурс]. - URL: <http://biblioclub.ru/index.php?page=book&id=59559>

2. Грекул, В.И. Проектирование информационных систем [Электронный ресурс] / В.И. Грекул, Г.Н. Денищенко, Н.Л. Коровкина. - М. : Интернет-Университет Информационных Технологий, 2005. - 304 с. - ISBN 5-9556-0033-7. - URL: <http://biblioclub.ru/index.php?page=book&id=233071>

3. Терещенко, П.В. Интерфейсы информационных систем: учебное пособие [Электронный ресурс] / П.В. Терещенко, В.А. Астапчук. - Новосибирск : НГТУ, 2012. - 67 с. - ISBN 978-5-7782-2036-2. - URL: <http://biblioclub.ru/index.php?page=book&id=228775>

Лабораторная работа 11. Создание структуры и модели интеллектуального анализа. Задача кластеризации

Цель работы – приобретение практических навыков создания структуры и модели интеллектуального анализа

Формируемые компетенции или их части - ОПК-2; ПК-2; ПК-3

Теоретическая часть.

Задача кластерного анализа заключается в том, чтобы на основании данных, содержащихся во множестве X , разбить множество объектов G на m (m – целое) кластеров (подмножеств) $Q_1, Q_2, \dots, Q_m$, так, чтобы каждый объект G_j принадлежал одному и только одному подмножеству разбиения и чтобы объекты, принадлежащие одному и тому же кластеру, были сходными, в то время как объекты, принадлежащие разным кластерам были разнородными.

Первым шагом в создании модели интеллектуального анализа данных кластеризации последовательностей является использование мастера интеллектуального анализа данных для создания новой структуры интеллектуального анализа данных и модели интеллектуального анализа данных, основанной на алгоритме кластеризации последовательностей.

Для этой цели будет использоваться представление источника данных, которое использовалось для анализа покупательского поведения, к которому добавляется столбец, содержащий идентификатор **sequence**. В данном сценарии «sequence» означает порядок, в котором покупатели добавляют элементы в свою корзину во время покупок.

Также будут добавлены несколько столбцов, которые используются в одной из моделей для группировки покупателей по демографическому признаку.

Создание структуры кластеризации последовательностей и модели

1. В обозревателе решений в среде SQL Server Data Tools (SSDT) щелкните правой кнопкой мыши элемент **Структуры интеллектуального анализа данных** и выберите команду **Создать структуру интеллектуального анализа**.
2. На странице **Вас приветствует мастер интеллектуального анализа данных** нажмите кнопку **Далее**.
3. Убедитесь, что на странице **Выбор метода определения** выбран параметр **На основе существующей реляционной базы данных или хранилища данных**, а затем нажмите кнопку **Далее**.
4. На странице **Создание структуры интеллектуального анализа данных** проверьте, что был выбран параметр **Создать структуру и модель интеллектуального анализа данных**. Затем щелкните раскрывающийся список **Какой метод интеллектуального анализа данных использовать?** и выберите пункт **Алгоритм кластеризации последовательностей (Майкрософт)**. Нажмите кнопку **Далее**.

Будет открыта страница **Выбор представления источника данных**. В группе **Доступные представления источников данных** выберите **Заказы**.

Orders — это то представление источника данных, которое использовалось для сценария потребительской корзины.

5. Нажмите кнопку **Далее**.
6. На странице **Определение типов таблиц** установите флажок **Вариант** рядом с таблицей **vAssocSeqOrders** и флажок **Вложенный** рядом с таблицей **vAssocSeqLineItems**. Нажмите кнопку **Далее**.
7. На странице **Определение обучающих данных** выберите столбцы, которые будут использоваться в модели, путем установки флажка следующим образом.

- **IncomeGroup** Установите флажок **Ввод**.

Данный столбец содержит полезную информацию о покупателях, которую можно использовать для кластеризации. Такая информация будет использоваться в первой модели и не будет учитываться во второй.

- **OrderNumber** Установите флажок **Ключ**.

Данное поле будет использоваться в качестве идентификатора таблицы вариантов либо значения **Key**. В целом ключевое поле таблицы вариантов никогда не используется в качестве входных данных, поскольку ключ содержит уникальные значения, не представляющие интереса для кластеризации.

- **Region** Установите флажок **Ввод**.

Данный столбец содержит полезную информацию о покупателях, которую можно использовать для кластеризации. Такая информация будет использоваться в первой модели и не будет учитываться во второй.

- **LineNumber** Установите флажки **Ключ** и **Ввод**.

Поле **LineNumber** будет использоваться в качестве идентификатора для вложенной таблицы или значения **Sequence Key**. Ключ для вложенной таблицы всегда должен использоваться для ввода.

- **Model** Установите флажки **Ввод** и **Прогнозируемый**.

Убедитесь, что все флажки установлены правильно, после чего нажмите кнопку **Далее**.

8. На странице **Определение содержимого и типа данных столбцов** убедитесь, что в сетке содержатся столбцы, типы содержимого и типы данных, приведенные в следующей таблице, после чего нажмите кнопку **Далее**.

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, MS Excel.

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. Симчера, В.М. Методы многомерного анализа статистических данных : учебное пособие / В.М. Симчера. - М. : Финансы и статистика, 2008. - 398 с. - ISBN 978-5-279-03184-9 ; То же [Электронный ресурс]. - URL: <http://biblioclub.ru/index.php?page=book&id=59559>

2. Грекул, В.И. Проектирование информационных систем [Электронный ресурс] / В.И. Грекул, Г.Н. Денищенко, Н.Л. Коровкина. - М. : Интернет-Университет Информационных Технологий, 2005. - 304 с. - ISBN 5-9556-0033-7. - URL: <http://biblioclub.ru/index.php?page=book&id=233071>

3. Терещенко, П.В. Интерфейсы информационных систем: учебное пособие [Электронный ресурс] / П.В. Терещенко, В.А. Астапчук. - Новосибирск : НГТУ, 2012. - 67 с. - ISBN 978-5-7782-2036-2. - URL: <http://biblioclub.ru/index.php?page=book&id=228775>

Список использованной литературы

1. Грекул, В.И. Проектирование информационных систем [Электронный ресурс] / В.И. Грекул, Г.Н. Денищенко, Н.Л. Коровкина. - М. : Интернет-Университет Информационных Технологий, 2005. - 304 с. - ISBN 5-9556-0033-7. - URL:<http://biblioclub.ru/index.php?page=book&id=233071>
2. Терещенко, П.В. Интерфейсы информационных систем: учебное пособие [Электронный ресурс] / П.В. Терещенко, В.А. Астапчук. - Новосибирск : НГТУ, 2012. - 67 с. - ISBN 978-5-7782-2036-2. - URL: <http://biblioclub.ru/index.php?page=book&id=228775>
3. <http://www.intuit.ru/studies/courses/2312/612/info>
4. Малышев, С.Л. Основы интернет-экономики : учебное пособие [Электронный ресурс] / С.Л. Малышев. - М. : Евразийский открытый институт, 2011. - 120 с. - ISBN 978-5-374-00556-1. - URL: <http://biblioclub.ru/index.php?page=book&id=90789>
5. Гаврилов, Л.П. Основы электронной коммерции и бизнеса [Электронный ресурс] / Л.П. Гаврилов. - М. : СОЛОН-ПРЕСС, 2009. - 592 с. - ISBN 978-5-91359-065-7. - URL:<http://biblioclub.ru/index.php?page=book&id=118188>

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 «Многомерный анализ данных /
Multidimensional data analysis»
для магистрантов направления 09.04.03 «Прикладная
информатика» направленность (профиль) «Управление
знаниями»

Ставрополь
2022

Введение

Дисциплина «Многомерный анализ данных», формирует профессиональные навыки магистранта по направлению подготовки 09.04.03 «Прикладная информатика».

Изучение дисциплины «Многомерный анализ данных» базируется на проведении лабораторных работ и индивидуальных занятий под руководством преподавателя, самостоятельной работе студентов

Задачами курса являются:

1. формирование у магистра личностных и профессиональных качеств, позволяющих осуществлять профессиональную деятельность, связанную с анализом, разработкой и внедрением информационно-аналитических систем;
2. 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;
3. познание основ проблематики и областей использования искусственного интеллекта, экспертных и основанных на знаниях систем.

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

2. Задачи освоения дисциплины:

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

2. 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;

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

3. Формируемые компетенции - ОПК-2; ПК-2; ПК-3

4. План-график выполнения самостоятельной работы

№ п/п	Раздел дисциплины	Неделя семестра	СРС в часах	Формы текущего контроля успеваемости
1	Лабораторная работа №1. Архитектура OLAP-приложений	1.	6	Собеседования по лабораторным работам
2	Лабораторная работа №2. Использование инструментов "AnalyzeKeyInfluencers" и "DetectCategories"	2.	6	Собеседования по лабораторным работам
3	Лабораторная работа №3. Использование инструментов "FillFromExample" и "Forecast"	3.	6	Собеседования по лабораторным работам
4	Лабораторная работа №4. Использование инструментов "HighlightExceptions" и "ScenarioAnalysis"	4.	6	Собеседования по лабораторным работам
5	Лабораторная работа №5. Использование инструментов "Prediction Calculator" и "ShoppingbasketAnalysis"	5.	6	Собеседования по лабораторным работам
6	Лабораторная работа №6. Использование инструментов Data Mining Client для Excel 2007 для подготовки данных	6.	6	Собеседования по лабораторным работам
7	Лабораторная работа №7. Использование инструментов Data Mining Client для Excel 2007 для создания модели анализа данных	7.	7,5	Собеседования по лабораторным работам

8	Лабораторная работа №8. Анализ точности прогноза и использование модели анализа	8.	6	Собеседования по лабораторным работам
9	Лабораторная работа №9. Построение модели кластеризации	9.	6	Собеседования по лабораторным работам
10	Лабораторная работа №10. Создание представления источника данных	10.	6	Собеседования по лабораторным работам
11	Лабораторная работа №11. Создание структуры и модели анализа. Задача кластеризации	11.	6	Собеседования по лабораторным работам
ИТОГО			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.

Вопросы для защиты заданий по лабораторным занятиям

Лабораторная работа №1. Настройки интеллектуального анализа данных для MicrosoftOffice

1. Основные настройки интеллектуального анализа данных для MS Office 2007

2. Средство настройки интеллектуального анализа данных для MSSQLServer

Список литературы, рекомендуемый к использованию по данной теме:

1. Симчера, В.М. Методы многомерного анализа статистических данных : учебное пособие / В.М. Симчера. - М. : Финансы и статистика, 2008. - 398 с. - ISBN 978-5-279-03184-9 ; То же [Электронный ресурс]. - URL: <http://biblioclub.ru/index.php?page=book&id=59559>

2. Грекул, В.И. Проектирование информационных систем [Электронный ресурс] / В.И. Грекул, Г.Н. Денищенко, Н.Л. Коровкина. - М. : Интернет-Университет Информационных Технологий, 2005. - 304 с. - ISBN 5-9556-0033-7. - URL: <http://biblioclub.ru/index.php?page=book&id=233071>

3. Терещенко, П.В. Интерфейсы информационных систем: учебное пособие [Электронный ресурс] / П.В. Терещенко, В.А. Астапчук. - Новосибирск : НГТУ, 2012. - 67 с. - ISBN 978-5-7782-2036-2. - URL: <http://biblioclub.ru/index.php?page=book&id=228775>

Лабораторная работа №2. Использование инструментов "AnalyzeKeyInfluencers" и "DetectCategories"

1. Настройка интеллектуального анализа данных
2. Анализ ключевых факторов влияния
3. Обнаружение категорий

Список литературы, рекомендуемый к использованию по данной теме:

1. Симчера, В.М. Методы многомерного анализа статистических данных : учебное пособие / В.М. Симчера. - М. : Финансы и статистика, 2008. - 398 с. - ISBN 978-5-279-03184-9 ; То же [Электронный ресурс]. - URL: <http://biblioclub.ru/index.php?page=book&id=59559>
2. Грекул, В.И. Проектирование информационных систем [Электронный ресурс] / В.И. Грекул, Г.Н. Денищенко, Н.Л. Коровкина. - М. : Интернет-Университет Информационных Технологий, 2005. - 304 с. - ISBN 5-9556-0033-7. - URL: <http://biblioclub.ru/index.php?page=book&id=233071>
3. Терещенко, П.В. Интерфейсы информационных систем: учебное пособие [Электронный ресурс] / П.В. Терещенко, В.А. Астапчук. - Новосибирск : НГТУ, 2012. - 67 с. - ISBN 978-5-7782-2036-2. - URL: <http://biblioclub.ru/index.php?page=book&id=228775>

Лабораторная работа №3. Использование инструментов "FillFromExample" и "Forecast"

1. Использование последовательности шаблона типа тренд - тенденцию изменения значений. Тренд может быть восходящим (возрастание значений ряда) или нисходящим (уменьшение значений)
2. Использование последовательности шаблона типа периодичность (сезонность) - событие повторяется через определённые интервалы;
3. Использование последовательности шаблона типа взаимная корреляция - зависимость значений одного ряда от других.

Список литературы, рекомендуемый к использованию по данной теме:

1. Симчера, В.М. Методы многомерного анализа статистических данных : учебное пособие / В.М. Симчера. - М. : Финансы и статистика, 2008. - 398 с. - ISBN 978-5-279-03184-9 ; То же [Электронный ресурс]. - URL: <http://biblioclub.ru/index.php?page=book&id=59559>
2. Грекул, В.И. Проектирование информационных систем [Электронный ресурс] / В.И. Грекул, Г.Н. Денищенко, Н.Л. Коровкина. - М. : Интернет-Университет Информационных Технологий, 2005. - 304 с. - ISBN 5-9556-0033-7. - URL: <http://biblioclub.ru/index.php?page=book&id=233071>
3. Терещенко, П.В. Интерфейсы информационных систем: учебное пособие [Электронный ресурс] / П.В. Терещенко, В.А. Астапчук. - Новосибирск :

НГТУ, 2012. - 67 с. - ISBN 978-5-7782-2036-2. - URL:
<http://biblioclub.ru/index.php?page=book&id=228775>

**Лабораторная работа №4.Использование инструментов
"HighlightExceptions" и "ScenarioAnalysis"**

1. Выделение исключений
2. Анализ сценариев
3. Анализ возможных вариантов

Список литературы, рекомендуемый к использованию по данной теме:

1. Симчера, В.М. Методы многомерного анализа статистических данных : учебное пособие / В.М. Симчера. - М. : Финансы и статистика, 2008. - 398 с. - ISBN 978-5-279-03184-9 ; То же [Электронный ресурс]. - URL: <http://biblioclub.ru/index.php?page=book&id=59559>

2. Грекул, В.И. Проектирование информационных систем [Электронный ресурс] / В.И. Грекул, Г.Н. Денищенко, Н.Л. Коровкина. - М. : Интернет-Университет Информационных Технологий, 2005. - 304 с. - ISBN 5-9556-0033-7. - URL:<http://biblioclub.ru/index.php?page=book&id=233071>

3. Терещенко, П.В. Интерфейсы информационных систем: учебное пособие [Электронный ресурс] / П.В. Терещенко, В.А. Астапчук. - Новосибирск : НГТУ, 2012. - 67 с. - ISBN 978-5-7782-2036-2. - URL: <http://biblioclub.ru/index.php?page=book&id=228775>

**Лабораторная работа №5.Использование инструментов
"PredictionCalculator" и "ShoppingbasketAnalysis"**

1. Расчет прогноза
2. Категории прогноза
3. Анализ покупательской корзины

Список литературы, рекомендуемый к использованию по данной теме:

1. Симчера, В.М. Методы многомерного анализа статистических данных : учебное пособие / В.М. Симчера. - М. : Финансы и статистика, 2008. - 398 с. - ISBN 978-5-279-03184-9 ; То же [Электронный ресурс]. - URL: <http://biblioclub.ru/index.php?page=book&id=59559>

2. Грекул, В.И. Проектирование информационных систем [Электронный ресурс] / В.И. Грекул, Г.Н. Денищенко, Н.Л. Коровкина. - М. : Интернет-Университет Информационных Технологий, 2005. - 304 с. - ISBN 5-9556-0033-7. - URL:<http://biblioclub.ru/index.php?page=book&id=233071>

3. Терещенко, П.В. Интерфейсы информационных систем: учебное пособие [Электронный ресурс] / П.В. Терещенко, В.А. Астапчук. - Новосибирск :

НГТУ, 2012. - 67 с. - ISBN 978-5-7782-2036-2. - URL:
<http://biblioclub.ru/index.php?page=book&id=228775>

Лабораторная работа №6. Использование инструментов DataMiningClient для Excel 2007 для подготовки данных

1. Инструмент CleanData

Список литературы, рекомендуемый к использованию по данной теме:

1. Симчера, В.М. Методы многомерного анализа статистических данных : учебное пособие / В.М. Симчера. - М. : Финансы и статистика, 2008. - 398 с. - ISBN 978-5-279-03184-9 ; То же [Электронный ресурс]. - URL: <http://biblioclub.ru/index.php?page=book&id=59559>

2. Грекул, В.И. Проектирование информационных систем [Электронный ресурс] / В.И. Грекул, Г.Н. Денищенко, Н.Л. Коровкина. - М. : Интернет-Университет Информационных Технологий, 2005. - 304 с. - ISBN 5-9556-0033-7. - URL: <http://biblioclub.ru/index.php?page=book&id=233071>

3. Терещенко, П.В. Интерфейсы информационных систем: учебное пособие [Электронный ресурс] / П.В. Терещенко, В.А. Астапчук. - Новосибирск : НГТУ, 2012. - 67 с. - ISBN 978-5-7782-2036-2. - URL: <http://biblioclub.ru/index.php?page=book&id=228775>

Лабораторная работа №7. Использование инструментов DataMiningClient для Excel 2007 для создания модели интеллектуального анализа данных

1. Создать модели интеллектуального анализа данных.

Список литературы, рекомендуемый к использованию по данной теме:

1. Симчера, В.М. Методы многомерного анализа статистических данных : учебное пособие / В.М. Симчера. - М. : Финансы и статистика, 2008. - 398 с. - ISBN 978-5-279-03184-9 ; То же [Электронный ресурс]. - URL: <http://biblioclub.ru/index.php?page=book&id=59559>

2. Грекул, В.И. Проектирование информационных систем [Электронный ресурс] / В.И. Грекул, Г.Н. Денищенко, Н.Л. Коровкина. - М. : Интернет-Университет Информационных Технологий, 2005. - 304 с. - ISBN 5-9556-0033-7. - URL: <http://biblioclub.ru/index.php?page=book&id=233071>

3. Терещенко, П.В. Интерфейсы информационных систем: учебное пособие [Электронный ресурс] / П.В. Терещенко, В.А. Астапчук. - Новосибирск : НГТУ, 2012. - 67 с. - ISBN 978-5-7782-2036-2. - URL: <http://biblioclub.ru/index.php?page=book&id=228775>

Лабораторная работа №8. Анализ точности прогноза и использование модели интеллектуального анализа

1. Диаграмма точности.
2. Матрица классификации
3. Дерево решений
4. Запросы к модели DM

Список литературы, рекомендуемый к использованию по данной теме:

1. Симчера, В.М. Методы многомерного анализа статистических данных : учебное пособие / В.М. Симчера. - М. : Финансы и статистика, 2008. - 398 с. - ISBN 978-5-279-03184-9 ; То же [Электронный ресурс]. - URL: <http://biblioclub.ru/index.php?page=book&id=59559>
2. Грекул, В.И. Проектирование информационных систем [Электронный ресурс] / В.И. Грекул, Г.Н. Денищенко, Н.Л. Коровкина. - М. : Интернет-Университет Информационных Технологий, 2005. - 304 с. - ISBN 5-9556-0033-7. - URL: <http://biblioclub.ru/index.php?page=book&id=233071>
3. Терещенко, П.В. Интерфейсы информационных систем: учебное пособие [Электронный ресурс] / П.В. Терещенко, В.А. Астапчук. - Новосибирск : НГТУ, 2012. - 67 с. - ISBN 978-5-7782-2036-2. - URL: <http://biblioclub.ru/index.php?page=book&id=228775>

Лабораторная работа №9. Построение модели кластеризации, трассировка и перекрестная проверка

1. DMX-запрос на обработку структуры

Список литературы, рекомендуемый к использованию по данной теме:

1. Симчера, В.М. Методы многомерного анализа статистических данных : учебное пособие / В.М. Симчера. - М. : Финансы и статистика, 2008. - 398 с. - ISBN 978-5-279-03184-9 ; То же [Электронный ресурс]. - URL: <http://biblioclub.ru/index.php?page=book&id=59559>
2. Грекул, В.И. Проектирование информационных систем [Электронный ресурс] / В.И. Грекул, Г.Н. Денищенко, Н.Л. Коровкина. - М. : Интернет-Университет Информационных Технологий, 2005. - 304 с. - ISBN 5-9556-0033-7. - URL: <http://biblioclub.ru/index.php?page=book&id=233071>
3. Терещенко, П.В. Интерфейсы информационных систем: учебное пособие [Электронный ресурс] / П.В. Терещенко, В.А. Астапчук. - Новосибирск : НГТУ, 2012. - 67 с. - ISBN 978-5-7782-2036-2. - URL: <http://biblioclub.ru/index.php?page=book&id=228775>

Лабораторная работа №10. Начало работы в BIDevStudio

1. Работа с источником данных

2. Режимы развёртывания приложений

Список литературы, рекомендуемый к использованию по данной теме:

1. Симчера, В.М. Методы многомерного анализа статистических данных : учебное пособие / В.М. Симчера. - М. : Финансы и статистика, 2008. - 398 с. - ISBN 978-5-279-03184-9 ; То же [Электронный ресурс]. - URL: <http://biblioclub.ru/index.php?page=book&id=59559>

2. Грекул, В.И. Проектирование информационных систем [Электронный ресурс] / В.И. Грекул, Г.Н. Денищенко, Н.Л. Коровкина. - М. : Интернет-Университет Информационных Технологий, 2005. - 304 с. - ISBN 5-9556-0033-7. - URL: <http://biblioclub.ru/index.php?page=book&id=233071>

3. Терещенко, П.В. Интерфейсы информационных систем: учебное пособие [Электронный ресурс] / П.В. Терещенко, В.А. Астапчук. - Новосибирск : НГТУ, 2012. - 67 с. - ISBN 978-5-7782-2036-2. - URL: <http://biblioclub.ru/index.php?page=book&id=228775>

Лабораторная работа №11.Создание представления источника данных

1. Понятие представления
2. Создание представления
3. Создание именованных вычислений и запросов

Список литературы, рекомендуемый к использованию по данной теме:

1. Симчера, В.М. Методы многомерного анализа статистических данных : учебное пособие / В.М. Симчера. - М. : Финансы и статистика, 2008. - 398 с. - ISBN 978-5-279-03184-9 ; То же [Электронный ресурс]. - URL: <http://biblioclub.ru/index.php?page=book&id=59559>

2. Грекул, В.И. Проектирование информационных систем [Электронный ресурс] / В.И. Грекул, Г.Н. Денищенко, Н.Л. Коровкина. - М. : Интернет-Университет Информационных Технологий, 2005. - 304 с. - ISBN 5-9556-0033-7. - URL: <http://biblioclub.ru/index.php?page=book&id=233071>

3. Терещенко, П.В. Интерфейсы информационных систем: учебное пособие [Электронный ресурс] / П.В. Терещенко, В.А. Астапчук. - Новосибирск : НГТУ, 2012. - 67 с. - ISBN 978-5-7782-2036-2. - URL: <http://biblioclub.ru/index.php?page=book&id=228775>

Лабораторная работа №12.Создание структуры и модели интеллектуального анализа. Задача кластеризации

1. Тип обработки объекта
2. Задача кластеризации
3. Создание и выделение кластера

Список литературы, рекомендуемый к использованию по данной теме:

1. Симчера, В.М. Методы многомерного анализа статистических данных : учебное пособие / В.М. Симчера. - М. : Финансы и статистика, 2008. - 398 с. - ISBN 978-5-279-03184-9 ; То же [Электронный ресурс]. - URL: <http://biblioclub.ru/index.php?page=book&id=59559>

2. Грекул, В.И. Проектирование информационных систем [Электронный ресурс] / В.И. Грекул, Г.Н. Денищенко, Н.Л. Коровкина. - М. : Интернет-Университет Информационных Технологий, 2005. - 304 с. - ISBN 5-9556-0033-7. - URL: <http://biblioclub.ru/index.php?page=book&id=233071>

3. Терещенко, П.В. Интерфейсы информационных систем: учебное пособие [Электронный ресурс] / П.В. Терещенко, В.А. Астапчук. - Новосибирск : НГТУ, 2012. - 67 с. - ISBN 978-5-7782-2036-2. - URL: <http://biblioclub.ru/index.php?page=book&id=228775>

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 практические занятия по дисциплине, на основе результатов устного опроса на практических занятиях и выполнения заданий самостоятельной работы.

К самостоятельному занятию магистрант 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.