

**Приложение 1 к РПД Методы и технологии анализа данных
09.03.02 Информационные системы и технологии
Направленность (профиль) – Программно-аппаратные комплексы
Форма обучения – очная
Год набора - 2018**

**МЕТОДИЧЕСКИЕ УКАЗАНИЯ ДЛЯ ОБУЧАЮЩИХСЯ ПО ОСВОЕНИЮ
ДИСЦИПЛИНЫ (МОДУЛЯ)**

1.	Кафедра	Информатики и вычислительной техники
2.	Направление подготовки	09.03.02 Информационные системы и технологии
3.	Направленность (профиль)	Программно-аппаратные комплексы
4.	Дисциплина (модуль)	Методы и технологии анализа данных
5.	Форма обучения	очная
6.	Год набора	2018

1. Методические рекомендации

Приступая к изучению дисциплины, обучающемуся необходимо внимательно ознакомиться с тематическим планом занятий, списком рекомендованной литературы. Следует уяснить последовательность выполнения индивидуальных учебных заданий. Самостоятельная работа обучающегося предполагает работу с научной и учебной литературой, умение создавать тексты. Уровень и глубина усвоения дисциплины зависят от активной и систематической работы на лекциях, изучения рекомендованной литературы, выполнения контрольных письменных заданий.

При изучении дисциплины обучающиеся выполняют следующие задания:

- изучают рекомендованную научно-практическую и учебную литературу;
- выполняют задания, предусмотренные для самостоятельной работы.

Основными видами аудиторной работы обучающихся являются лекции и практические/практические занятия.

1.1. Методические рекомендации по организации работы обучающихся во время проведения лекционных занятий

В ходе лекций преподаватель излагает и разъясняет основные, наиболее сложные понятия темы, а также связанные с ней теоретические и практические проблемы, дает рекомендации на семинарское занятие и указания на самостоятельную работу.

Знакомство с дисциплиной происходит уже на первой лекции, где от обучающегося требуется не просто внимание, но и самостоятельное оформление конспекта. При работе с конспектом лекций необходимо учитывать тот фактор, что одни лекции дают ответы на конкретные вопросы темы, другие – лишь выявляют взаимосвязи между явлениями, помогая обучающемуся понять глубинные процессы развития изучаемого предмета как в истории, так и в настоящее время.

Конспектирование лекций – сложный вид вузовской аудиторной работы, предполагающий интенсивную умственную деятельность обучающегося. Конспект является полезным тогда, когда записано самое существенное и сделано это самим обучающимся. Не надо стремиться записать дословно всю лекцию. Такое «конспектирование» приносит больше вреда, чем пользы. Целесообразно вначале понять основную мысль, излагаемую лектором, а затем записать ее. Желательно запись осуществлять на одной странице листа или оставляя поля, на которых позднее, при самостоятельной работе с конспектом, можно сделать дополнительные записи, отметить непонятные места.

Конспект лекции лучше подразделять на пункты, соблюдая красную строку. Этому в большой степени будут способствовать вопросы плана лекции, предложенные преподавателям. Следует обращать внимание на акценты, выводы, которые делает лектор, отмечая наиболее важные моменты в лекционном материале замечаниями «важно», «хорошо запомнить» и т.п. Можно делать это и с помощью разноцветных маркеров или ручек, подчеркивая термины и определения.

Целесообразно разработать собственную систему сокращений, аббревиатур и символов. Однако при дальнейшей работе с конспектом символы лучше заменить обычными словами для быстрого зрительного восприятия текста.

Работая над конспектом лекций, всегда необходимо использовать не только учебник, но и ту литературу, которую дополнительно рекомендовал лектор. Именно такая серьезная, кропотливая работа с лекционным материалом позволит глубоко овладеть теоретическим материалом.

1.2. Методические рекомендации по подготовке к практическим занятиям

Подготовку к каждому практическому занятию обучающийся должен начать с ознакомления с его планом, отражающим содержание предложенной темы. Продумывание и изучение вопросов плана основывается на проработке текущего материала лекции, и изучения рекомендованной обязательной и дополнительной литературы. Новые понятия по изучаемой теме необходимо проработать и внести в глоссарий.

Результат такой работы должен проявиться в способности обучающегося свободно ответить на теоретические вопросы практикума, его выступлении и участии в коллективном обсуждении вопросов изучаемой темы, правильном выполнении лабораторных заданий и контрольных работ.

В процессе подготовки к практическим занятиям, обучающимся необходимо обратить особое внимание на самостоятельное изучение рекомендованной литературы. При всей полноте конспектирования лекции в ней невозможно изложить весь материал из-за лимита аудиторных часов. Поэтому самостоятельная работа с учебниками, учебными пособиями, научной, справочной литературой, материалами периодических изданий и Интернета является наиболее эффективным методом получения дополнительных знаний, позволяет значительно активизировать процесс овладения информацией, способствует более глубокому усвоению изучаемого материала, формирует у обучающихся свое отношение к конкретной проблеме.

Качество учебной работы обучающихся преподаватель оценивает с использованием технологической карты дисциплины, размещенной на сайте филиала МАГУ.

1.3. Методические рекомендации по работе с литературой

Работу с литературой целесообразно начать с изучения общих работ по теме, а также учебников и учебных пособий. Далее рекомендуется перейти к анализу монографий и статей, рассматривающих отдельные аспекты проблем, изучаемых в рамках курса, а также официальных материалов и неопубликованных документов (научно-исследовательские работы, диссертации), в которых могут содержаться основные вопросы изучаемой проблемы.

Работу с источниками надо начинать с ознакомительного чтения, т.е. просмотреть текст, выделяя его структурные единицы. При ознакомительном чтении закладками отмечаются те страницы, которые требуют более внимательного изучения.

В зависимости от результатов ознакомительного чтения выбирается дальнейший способ работы с источником. Если для разрешения поставленной задачи требуется изучение некоторых фрагментов текста, то используется метод выборочного чтения. Если

в книге нет подробного оглавления, следует обратить внимание ученика на предметные и именные указатели.

Избранные фрагменты или весь текст (если он целиком имеет отношение к теме) требуют вдумчивого, неторопливого чтения с «мысленной проработкой» материала. Такое чтение предполагает выделение: 1) главного в тексте; 2) основных аргументов; 3) выводов. Особое внимание следует обратить на то, вытекает тезис из аргументов или нет.

Необходимо также проанализировать, какие из утверждений автора носят проблематичный, гипотетический характер и уловить скрытые вопросы.

Понятно, что умение таким образом работать с текстом приходит далеко не сразу. Наилучший способ научиться выделять главное в тексте, улавливать проблематичный характер утверждений, давать оценку авторской позиции – это сравнительное чтение, в ходе которого обучающийся знакомится с различными мнениями по одному и тому же вопросу, сравнивает весомость и доказательность аргументов сторон и делает вывод о наибольшей убедительности той или иной позиции.

Если в литературе встречаются разные точки зрения по тому или иному вопросу из-за сложности прошедших событий и правовых явлений, нельзя их отвергать, не разобравшись. При наличии расхождений между авторами необходимо найти рациональное зерно у каждого из них, что позволит глубже усвоить предмет изучения и более критично оценивать изучаемые вопросы. Знакомясь с особыми позициями авторов, нужно определять их схожие суждения, аргументы, выводы, а затем сравнивать их между собой и применять из них ту, которая более убедительна.

Следующим этапом работы с литературными источниками является создание конспектов, фиксирующих основные тезисы и аргументы. Можно делать записи на отдельных листах, которые потом легко систематизировать по отдельным темам изучаемого курса. Другой способ – это ведение тематических тетрадей-конспектов по одной какой-либо теме. Большие специальные работы монографического характера целесообразно конспектировать в отдельных тетрадях. Здесь важно вспомнить, что конспекты пишутся на одной стороне листа, с полями и достаточным для исправления и ремарок межстрочным расстоянием (эти правила соблюдаются для удобства редактирования). Если в конспектах приводятся цитаты, то непременно должно быть дано указание на источник (автор, название, выходные данные, № страницы). Впоследствии эта информация может быть использована при написании текста реферата или другого задания.

Таким образом, при работе с источниками и литературой важно уметь:

- сопоставлять, сравнивать, классифицировать, группировать, систематизировать информацию в соответствии с определенной учебной задачей;
- обобщать полученную информацию, оценивать прослушанное и прочитанное;
- фиксировать основное содержание сообщений; формулировать, устно и письменно, основную идею сообщения; составлять план, формулировать тезисы;
- готовить и презентовать развернутые сообщения типа доклада;
- работать в разных режимах (индивидуально, в паре, в группе), взаимодействуя друг с другом;
- пользоваться реферативными и справочными материалами;
- контролировать свои действия и действия своих товарищей, объективно оценивать свои действия;
- обращаться за помощью, дополнительными разъяснениями к преподавателю, другим обучающимся.
- пользоваться лингвистической или контекстуальной догадкой, словарями различного характера, различного рода подсказками, опорами в тексте (ключевые слова, структура текста, предваряющая информация и др.);

- использовать при говорении и письме перифраз, синонимичные средства, слова-описания общих понятий, разъяснения, примеры, толкования, «словотворчество»;
- повторять или перефразировать реплику собеседника в подтверждении понимания его высказывания или вопроса;
- обратиться за помощью к собеседнику (уточнить вопрос, переспросить и др.);
- использовать мимику, жесты (вообще и в тех случаях, когда языковых средств не хватает для выражения тех или иных коммуникативных намерений).

1.4. Методические рекомендации по подготовке к сдаче экзамена

Подготовка к экзамену способствует закреплению, углублению и обобщению знаний, получаемых, в процессе обучения, а также применению их к решению практических задач. Готовясь к экзамену, обучающийся ликвидирует имеющиеся пробелы в знаниях, углубляет, систематизирует и упорядочивает свои знания. На экзамене обучающийся демонстрирует то, что он приобрел в процессе изучения дисциплины.

В условиях применяемой в МАГУ балльно-рейтинговой системы подготовка к экзамену включает в себя самостоятельную и аудиторную работу обучающегося в течение всего периода изучения дисциплины и непосредственную подготовку в дни, предшествующие экзамену по разделам и темам дисциплины.

При подготовке к экзамену обучающимся целесообразно использовать не только материалы лекций, а и рекомендованные основную и дополнительную литературу.

При подготовке к промежуточной аттестации целесообразно:

- внимательно изучить перечень вопросов и определить, в каких источниках находятся сведения, необходимые для ответа на них;
- внимательно прочитать рекомендованную литературу;
- составить краткие конспекты ответов (планы ответов).

Качество учебной работы обучающихся преподаватель оценивает с использованием технологической карты дисциплины, размещенной на сайте филиала МАГУ.

1.5. Методические рекомендации по подготовке доклада

Алгоритм создания доклада:

- 1 этап – определение темы доклада
- 2 этап – определение цели доклада
- 3 этап – подробное раскрытие информации
- 4 этап – формулирование основных тезисов и выводов.

1.6. Методические рекомендации по составлению глоссария

1. Внимательно прочитайте и ознакомьтесь с текстом. Вы встретите в нем много различных терминов, которые имеются по данной теме.

2. После того, как вы определили наиболее часто встречающиеся термины, вы должны составить из них список. Слова в этом списке должны быть расположены в строго алфавитном порядке, так как глоссарий представляет собой не что иное, как словарь специализированных терминов.

3. После этого начинается работа по составлению статей глоссария. Статья глоссария - это определение термина. Она состоит из двух частей: 1. точная формулировка термина в именительном падеже; 2. содержательная часть, объемно раскрывающая смысл данного термина.

При составлении глоссария важно придерживаться следующих правил:

- стремитесь к максимальной точности и достоверности информации;

- старайтесь указывать корректные научные термины и избегать всякого рода жаргонизмов. В случае употребления такого, дайте ему краткое и понятное пояснение;
- излагая несколько точек зрения в статье по поводу спорного вопроса, не принимайте ни одну из указанных позиций. Глоссарий - это всего лишь констатация имеющихся фактов;
- также не забывайте приводить в пример контекст, в котором может употребляться данный термин;
- при желании в глоссарий можно включить не только отдельные слова и термины, но и целые фразы.

1.7. Методические рекомендации для занятий в интерактивной форме

В учебном процессе, помимо чтения лекций и аудиторных занятий, используются интерактивные формы. В сочетании с внеаудиторной работой это способствует формированию и развитию профессиональных навыков обучающихся.

Интерактивное обучение представляет собой способ познания, осуществляемый в формах совместной деятельности обучающихся, т.е. все участники образовательного процесса взаимодействуют друг с другом, совместно решают поставленные проблемы, моделируют ситуации, обмениваются информацией, оценивают действие коллег и свое собственное поведение, погружаются в реальную атмосферу делового сотрудничества по разрешению проблем.

В курсе изучаемой дисциплины «Методы и технологии анализа данных» в интерактивной форме часы используются в виде: групповой дискуссии, заслушивании и обсуждении подготовленных обучающимися докладов с презентациями по тематике дисциплины.

Тематика занятий с использованием интерактивных форм

№ п/п	Тема	Интерактивная форма	Часы, отводимые на интерактивные формы	
			Лекции	Практические занятия
1.	Работа в электронной таблице Calc.	Групповая дискуссия	-	2
2.	Информационный процесс накопления данных.	Групповая дискуссия	-	2
3.	Разработка баз данных в СУБД Access.	Групповая дискуссия	-	4
4.	Хранилища данных. Технология OLAP	Групповая дискуссия	-	2
5.	Классификация и регрессия	Групповая дискуссия	-	2
6.	Поиск ассоциативных правил	Групповая дискуссия	-	2
7.	Кластеризация	Групповая дискуссия, выступление с докладом	-	2
ИТОГО			16 часов	

2. План практических занятий

1 семестр

Практическая работа № 1. Обработка данных, построение таблиц и построение диаграмм в электронной таблице Calc.

План:

1. Принципы работы с электронной таблицей.
2. Ввод и форматирование текста.
3. Форматирование ячеек и таблиц.

4. Ввод формул. Вычислительные возможности электронных таблиц.
5. Работа с математическими формулами в электронной таблице.
6. Статистические функции Calc для обработки данных.
7. Принципы построения диаграмм в электронной таблице.
8. Абсолютные и относительные ссылки.
9. Алгоритм построения диаграмм.
10. Правила редактирования диаграмм.

Литература: [2, с. 2-150].

Вопросы для групповой дискуссии:

1. Для каких целей применяются в основном электронные таблицы?
2. Каким образом осуществляется набор текста, в электронной таблице?
3. Каким образом осуществляется набор формул, в электронной таблице?
4. Как выполняется построение диаграмм, в электронной таблице?
5. Типы диаграмм в Calc?
6. В чем отличие относительных и абсолютных ссылок?
7. Какие действия необходимо выполнить в Calc для построения графика заданной функции?
8. Что такое легенда диаграммы?
9. Каким образом можно выполнить редактирование осей диаграммы?

Задание для самостоятельной работы

1. Построить таблицу и диаграмму к ней согласно варианту задания.
2. Предусмотреть в таблице статические, изменяемые и вычисляемые ячейки.
3. Подготовиться к ответам на вопросы.

Практическая работа № 2. Построение баз данных в Microsoft Office Access 2007 на примере готовой базы данных «Деканат».

План:

1. Введение в реляционные базы данных. Основные понятия.
2. Системы управления базами данных.
3. Понятие схемы данных.
4. Понятие целостности данных.
5. Ключи: первичные и вторичные.

Литература: [4, с. 4-180].

Вопросы для групповой дискуссии:

1. Что такое сущность в реляционной базе данных?
2. Что такое атрибут в реляционной базе данных?
3. Что такое ключ в реляционной базе данных? Какие бывают ключи?
4. Как выполняется связь сущностей в реляционной базе данных?
5. Что такое сущность в реляционной базе данных?
6. Что такое сущность в реляционной базе данных?
7. Что такое система управления базами данных?
8. Какие таблицы используются в базе данных «Деканат»?
9. Какие запросы используются в базе данных «Деканат»?
10. Опишите схему данных базы данных «Деканат»?

Задание для самостоятельной работы

1. Изучить готовую базу данных «Деканат».
2. Подготовиться к ответам на вопросы.

Практическая работа № 3. Разработка базы данных согласно варианту задания в Microsoft Office Access 2007

План:

1. Разработка таблиц в Microsoft Office Access 2007.
2. Разработка схемы данных в Microsoft Office Access 2007.
3. Разработка запросов в Microsoft Office Access 2007.
4. Типы запросов в Microsoft Office Access 2007.
5. Разработка форм в Microsoft Office Access 2007.
6. Разработка отчетов в Microsoft Office Access 2007.

Литература: [4, с. 4-180].

Вопросы для групповой дискуссии:

1. Как выполняется построения таблиц в Microsoft Office Access 2007.
2. Как выполняется построения запросов в Microsoft Office Access 2007.
3. Как выполняется построения форм в Microsoft Office Access 2007.
4. Как выполняется построения отчетов в Microsoft Office Access 2007.

Задание для самостоятельной работы

1. Выполнить построение собственной базы данных, согласно варианта задания.
2. Подготовиться к ответам на вопросы.

2 семестр

Практическая работа № 1. Проектирование хранилищ данных в Deductor Warehouse 6.

План:

1. Эволюция корпоративных информационных систем.
2. Архитектура хранилища данных.
3. Технология OLAP. OLTP vs OLAP.
4. Объекты хранения хранилища данных Deductor Warehouse: измерение, атрибут, факт.
5. Многомерная модель данных.
6. Недостатки реляционной модели данных.
7. Проектирование структуры хранилища данных (ХД).
8. Схема снежинка. Пример схемы снежинка.
9. Создание нового хранилища данных.
10. Наполнение хранилища данных.
11. Извлечение информации из ХД.
12. Кубы в Deductor Warehouse.

Литература: [3, с. 16-30]; [3, с. 73-145];

Вопросы для групповой дискуссии:

1. В чем состоит отличие баз данных от хранилищ данных?
2. Что означает OLTP?
3. Что означает OLAP?
4. Какая схема реляционного ХД используется в Deductor Warehouse?
5. Перечислите объекты хранилища Deductor Warehouse и дайте их определения.
6. Чем отличается атрибут процесса от измерения?
7. Какова последовательность загрузки информации в хранилище?
8. Что такое куб в Deductor Warehouse 6?
9. Как создать куб?

Задание для самостоятельной работы

1. Разработать структуру хранилища данных (ХД) для следующего примера: имеется история продаж различных товаров по дням в нескольких торговых объектах. Товары объединены в группы. Все данные представлены в 4 таблицах: Товарные группы, Товары, Отделы, Продажи. Их фрагменты приведены ниже (таблицы 1-4).

Таблица 1 – Товарные группы

Код группы	Наименование группы
33	Иммуномодуляторы
48	Общетонизирующие средства и адаптогены
50	Местные анестетики
108	Микро- и макроэлементы
198	Витамины и витаминоподобные средства
223	Желчегонные средства и препараты желчи
247	Антисептики и дезинфицирующие средства

Таблица 2 – Товары (фрагмент)

Код товара	Наименование товара	Код группы
774	Альмагель	1
810	Иммунонорм	33
824	Ревит	198
898	Настойка пустырника	48
...

Таблица 3 – Отделы

Код отдела	Наименование отдела
1	Аптека 1
2	Аптека 2
3	Аптека 3

Таблица 4 – Продажи (фрагмент)

Дата	Код отдела	Код товара	Час покупки	Количество	Сумма
01.01.2006	1	31052	13	1	56,5
01.01.2006	1	36259	16	1	72,48
01.01.2006	1	40315	15	1	15,84
01.01.2006	3	810	14	1	163,50
01.01.2006	2	40315	15	3	47,52
...

2. Подготовиться к ответам на вопросы.

Практическая работа № 2. Ассоциативные правила в стимулировании розничных продаж.

План:

1. Методы и стадии Data Mining.
2. Задачи Data Mining.
3. Сферы применения Data Mining.
4. Методы поиска ассоциативных правил.
5. Решение задачи поиска ассоциативных правил в Deductor Studio.

Литература: [1, с. 194-208].

Вопросы для групповой дискуссии:

1. Что такое Data Mining?
2. Какова сфера применения методов Data Mining?
3. Перечислите основные задачи Data Mining.
4. В чем состоит основная идея метода ассоциативных правил?

5. Какой алгоритм используется в Deductor Studio в обработчике «Ассоциативные правила», для решения задач ассоциации.
6. Поясните алгоритм работы с обработчиком «Ассоциативные правила» в Deductor Studio.
7. Какой алгоритм реализован в Deductor Studio для решения задач поиска ассоциативных правил.

Задание для самостоятельной работы

1. Изучите особенности применения ассоциативных правил для решения бизнес-задачи стимулирования продаж в среде Deductor Academic, используя исходные данные поставляемые со средой Deductor Academic.
2. Подготовиться к ответам на вопросы.

Практическая работа № 3. Сегментация клиентов телекоммуникационной компании.

План:

1. Методы и стадии Data Mining.
2. Задачи Data Mining.
3. Сферы применения Data Mining.
4. Задача кластеризации.
5. Сети и карты Кохонена в Deductor Studio.
6. Визуализатор карт Кохонена в Deductor Studio.

Литература: [1, с. 155-193].

Вопросы для самоконтроля

1. Что такое Data Mining?
2. Какова сфера применения методов Data Mining?
3. Перечислите основные задачи Data Mining.

Задание для групповой дискуссии:

1. Изучите особенности применения ассоциативных правил для решения бизнес-задачи сегментации клиентов телекоммуникационной компании, используя исходные данные поставляемые со средой Deductor Academic.
2. Исходные данные: активные абоненты, которые регулярно пользовались услугами сотовой связи в течение последних нескольких месяцев. Данные находятся в файле mobile.txt.
3. Подготовиться к ответам на вопросы.

Практическая работа № 4. Скоринговая карта для оценки кредитоспособности заемщиков.

План:

1. Задача логистической регрессии.
2. Деревья решений.
3. Построение модели логистической регрессии в Deductor Studio.
4. Визуализатор
5. ROC – кривая.
6. Построение модели деревьев решений в Deductor Studio.

Литература: [1, с. 107-114]; [4, с. 115-126].

Вопросы для самоконтроля:

1. Как выполняется построение модели логистической регрессии в Deductor Studio?
2. Как выполняется построение модели деревьев решений в Deductor Studio?

3. Что такое ROC – кривая?

Задание для самостоятельной работы

1. Изучите особенности применения ассоциативных правил для решения бизнес-задачи оценки кредитоспособности клиента, с помощью логистической регрессии и деревьев решений, используя исходные данные, поставляемые со средой Deductor Academic.

2. Подготовиться к ответам на вопросы.