

**Федеральное государственное бюджетное образовательное
учреждение высшего образования
«РОССИЙСКАЯ АКАДЕМИЯ НАРОДНОГО ХОЗЯЙСТВА
И ГОСУДАРСТВЕННОЙ СЛУЖБЫ
ПРИ ПРЕЗИДЕНТЕ РОССИЙСКОЙ ФЕДЕРАЦИИ»**

Экономический факультет, отделение национальной экономики

(наименование структурного подразделения (института/факультета/филиала))

Кафедра национальной экономики

(наименование кафедры)

УТВЕРЖДЕНА

решением кафедры _____

Протокол от «__» _____ 201__ г.

№ _____

РАБОЧАЯ ПРОГРАММА ДИСЦИПЛИНЫ

**Б1.В.02 ИНФОРМАЦИОННЫЕ ТЕХНОЛОГИИ
В ПРОГНОЗНО-АНАЛИТИЧЕСКОЙ ДЕЯТЕЛЬНОСТИ**

направление подготовки

38.04.01 Экономика

(код, наименование направления подготовки)

Экономика и право

(направленность (профиль))

магистр

(квалификация)

очная, очно-заочная, заочная

(формы обучения)

Год набора 2016г.

Москва, 2016

Автор–составитель:

Баркова Е.А., ст. преподаватель

(ученая степень и(или) ученое звание, должность)

Кафедра национальной экономики

(наименование кафедры)

(ученая степень и(или) ученое звание, должность)

(наименование кафедры)

(ученая степень и(или) ученое звание, должность)

(наименование кафедры)

(ученая степень и(или) ученое звание, должность)

(наименование кафедры)

Заведующий кафедрой

Национальной экономики

(наименование кафедры)

Казарян М.А., д.э.н., доцент

(ученая степень и(или) ученое звание)

СОДЕРЖАНИЕ

1. ПЕРЕЧЕНЬ ПЛАНИРУЕМЫХ РЕЗУЛЬТАТОВ ОБУЧЕНИЯ ПО ДИСЦИПЛИНЕ, СООТНЕСЕННЫХ С ПЛАНИРУЕМЫМИ РЕЗУЛЬТАТАМИ ОСВОЕНИЯ ПРОГРАММЫ	4
2. ОБЪЕМ И МЕСТО ДИСЦИПЛИНЫ В СТРУКТУРЕ ОП ВО.....	5
3. СОДЕРЖАНИЕ И СТРУКТУРА ДИСЦИПЛИНЫ (МОДУЛЯ)	7
4. ФОНД ОЦЕНОЧНЫХ СРЕДСТВ ПРОМЕЖУТОЧНОЙ АТТЕСТАЦИИ ПО ДИСЦИПЛИНЕ (МОДУЛЮ)	10
5. МЕТОДИЧЕСКИЕ УКАЗАНИЯ ДЛЯ ОБУЧАЮЩИХСЯ ПО ОСВОЕНИЮ ДИСЦИПЛИНЫ.....	43
6. УЧЕБНАЯ ЛИТЕРАТУРА И РЕСУРСЫ ИНФОРМАЦИОННО-ТЕЛЕКОММУНИКАЦИОННОЙ СЕТИ «ИНТЕРНЕТ», ВКЛЮЧАЯ ПЕРЕЧЕНЬ УЧЕБНО-МЕТОДИЧЕСКОГО ОБЕСПЕЧЕНИЯ ДЛЯ САМОСТОЯТЕЛЬНОЙ РАБОТЫ ОБУЧАЮЩИХСЯ ПО ДИСЦИПЛИНЕ (МОДУЛЮ)	45
6.1. ОСНОВНАЯ ЛИТЕРАТУРА	45
6.2. ДОПОЛНИТЕЛЬНАЯ ЛИТЕРАТУРА.....	46
6.3. УЧЕБНО-МЕТОДИЧЕСКОЕ ОБЕСПЕЧЕНИЕ САМОСТОЯТЕЛЬНОЙ РАБОТЫ	46
6.4. НОРМАТИВНЫЕ ПРАВОВЫЕ ДОКУМЕНТЫ	46
6.5. ИНТЕРНЕТ-РЕСУРСЫ	46
6.6. ИНЫЕ ИСТОЧНИКИ.....	46
7. МАТЕРИАЛЬНО-ТЕХНИЧЕСКАЯ БАЗА, ИНФОРМАЦИОННЫЕ ТЕХНОЛОГИИ, ПРОГРАММНОЕ ОБЕСПЕЧЕНИЕ И ИНФОРМАЦИОННЫЕ СПРАВОЧНЫЕ СИСТЕМЫ.....	46

1. ПЕРЕЧЕНЬ ПЛАНИРУЕМЫХ РЕЗУЛЬТАТОВ ОБУЧЕНИЯ ПО ДИСЦИПЛИНЕ, СООТНЕСЕННЫХ С ПЛАНИРУЕМЫМИ РЕЗУЛЬТАТАМИ ОСВОЕНИЯ ПРОГРАММЫ

1.1. Цели и задачи дисциплины

Дисциплина Б1.В.02 «Информационные технологии в прогнозно-аналитической деятельности» имеет своей целью реализацию требований к освоению соответствующих компонентов профессиональной компетенции ПК-10 на основе формирования у студентов системных и глубоких теоретических знаний, умений и практических навыков в области когнитивного анализа данных.

Изучение дисциплины осуществляется в соответствии с требованиями ФГОС ВО по направлению подготовки магистров 38.04.01 «Экономика» с учетом специфики направленности (профиля подготовки) – «Экономика и право».

Задачи освоения учебной дисциплины «Информационные технологии в прогнозно-аналитической деятельности»:

- получение обучающимися целостного представления о методах, моделях и средствах когнитивного анализа, используемых для поддержки бизнес - решений в процессе управления компанией;
- развитие способности оценить необходимость и возможность применения интеллектуального анализа данных при решении конкретных бизнес-задач;
- развитие способности формулировать задачи интеллектуального бизнес-анализа экономического объекта, осуществлять выбор методов анализа, подготовку исходных данных, визуализацию и интерпретацию полученных результатов;
- формирование навыков эффективной практической работы в среде аналитической платформы Deductor Academic.

Дисциплина «Информационные технологии в прогнозно-аналитической деятельности» обеспечивает овладение следующими компетенциями с учетом этапа:

Код компетенции	Наименование компетенции	Код этапа освоения компетенции	Наименование этапа освоения компетенции
ПК-10	Способность составлять прогноз основных социально-экономических показателей деятельности предприятия, отрасли, региона и экономики в целом	ПК-10.2.1 Этап 2	Способность составлять прогноз основных социально-экономических показателей деятельности предприятия

1.2. Планируемые результаты обучения по дисциплине «Информационные технологии в прогнозно-аналитической деятельности»

В результате освоения дисциплины у студентов должны быть сформированы:

Код этапа освоения компетенции	Результаты обучения
ПК-10.2.1	<p>Знания:</p> <ul style="list-style-type: none">– места интеллектуального анализа данных в системах поддержки принятия решений;– методов, моделей и средств когнитивного анализа, используемых для поддержки бизнес - решений в процессе управления компанией;– структуры и принципов работы современных аналитических платформ;– моделей, методов и алгоритмов интеллектуального анализа данных;– целей и задач визуализации данных на разных этапах аналитического процесса, визуализаторы общего назначения, OLAP-анализ, визуализаторы для оценки качества моделей, визуализаторы, применяемые для интерпретации результатов анализа;– способов оценки качества данных и методов подготовки данных к анализу. <p>Умения:</p> <ul style="list-style-type: none">– оценить необходимость и возможность применения интеллектуального анализа данных при решении конкретных бизнес-задач;– формулировать задачи интеллектуального бизнес-анализа экономического объекта,– осуществлять выбор методов анализа, подготовку исходных данных, визуализацию и интерпретацию полученных результатов;– решать задачи ассоциации, кластеризации, классификации, регрессии, задачи анализа и прогнозирования временных рядов;– строить аналитическую отчетность. <p>Навыки:</p> <ul style="list-style-type: none">– решения задач ассоциации, кластеризации, классификации, регрессии, задачи анализа и прогнозирования временных рядов;– построения аналитической отчетности;– эффективной аналитической работы в среде программы Deductor Academic.

2. ОБЪЕМ И МЕСТО ДИСЦИПЛИНЫ В СТРУКТУРЕ ОП ВО

2.1. Объем дисциплины

Общая трудоемкость дисциплины Б1.В.02 «Информационные технологии в прогнозно-аналитической деятельности» составляет 2 зачетные единицы (72 акаде-

мических часа).

Дисциплина реализуется в рамках следующих форм обучения:

- очной: 1 курс, 2 семестр;
контактная работа с преподавателем – 54 ч., из них:
18 ч. лекций, 18 ч. практических занятий, 18 ч. лабораторных работ;
самостоятельная работа студента – 18 ч.
- очно-
заочной: 1 курс, 2 семестр;
контактная работа с преподавателем – 54 ч., из них:
18 ч. лекций, 18 ч. практических занятий, 18 ч. лабораторных работ;
самостоятельная работа студента – 36 ч.
- заочной: 1 курс, 3 семестр;
контактная работа с преподавателем – 10 ч., из них:
4 ч. лекций, 4 ч. лабораторных работ;
самостоятельная работа студента – 28 ч.
- 2 курс, 1 семестр;
контактная работа с преподавателем – 16 ч., из них:
12 ч. лабораторных работ, 4 ч. практических занятий;
самостоятельная работа студента – 16 ч.

Форма промежуточной аттестации – зачет с оценкой.

2.2. Место дисциплины в структуре ОП ВО

Дисциплина Б1.В.02 «Информационные технологии в прогнозно-аналитической деятельности» относится к блоку дисциплин вариативной части учебного плана направления подготовки магистров 38.04.01 «Экономика» с профилем подготовки «Экономика и право».

В методологическом плане дисциплина «Информационные технологии в прогнозно-аналитической деятельности» опирается на минимально необходимый объем знаний, умений и навыков, приобретенных ранее при изучении дисциплин Б1.В.03 «Социально-экономическая статистика», Б1.Б.03 «Эконометрика (продвинутый уровень)».

Знания, умения и навыки, полученные в результате освоения дисциплины Б1.В.02 «Информационные технологии в прогнозно-аналитической деятельности», могут быть использованы при изучении дисциплин: Б1.В.06 «Финансовая диагностика и моделирование перспектив развития компании», Б1.В.01 «Корпоративные финансы (продвинутый уровень)», Б2.В.01(Н) «Научно-исследовательская работа и семинар», Б2.В.03(Пд) «Преддипломная практика», Б3.Б.01(Д) «Подготовка и защита выпускной квалификационной работы».

3. СОДЕРЖАНИЕ И СТРУКТУРА ДИСЦИПЛИНЫ (МОДУЛЯ)

3.1. Распределение объема дисциплины по разделам (темам), семестрам, видам учебной работы и формам контроля

№ п/п	Наименование тем (разделов),	Объем дисциплины (модуля), час.					СР	Форма текущего контроля успеваемости, промежуточной аттестации**
		Всего	Контактная работа обучающихся с преподавателем по видам учебных занятий					
			Л/ЭО, ДОТ*	ЛР/ЭО, ДОТ*	ПЗ/ЭО, ДОТ*	КСР		
Очная форма обучения								
Тема 1	Технологии анализа данных	6	2	-	2		2	О, Т
Тема 2	Визуализация данных	8	2	2	2		2	О, ПОЗ
Тема 3	Классификация и регрессия (статистические методы)	14	4	2	4		4	О, ПОЗ
Тема 4	Классификация и регрессия (Машинное обучение)	10	2	4	2		2	Т, ПОЗ
Тема 5	Кластеризация	8	2	2	2		2	О, ПОЗ
Тема 6	Поиск ассоциативных правил	8	2	2	2		2	О, ПОЗ
Тема 7	Анализ и прогнозирование временных рядов	10	2	4	2		2	О, ПОЗ
Тема 8	Ансамбли моделей	8	2	2	2		2	ПКР
Промежуточная аттестация								Зачет с оценкой
Всего:		72	18	18	18	-	18	
Очно-заочная форма обучения								
Тема 1	Технологии анализа данных	6	2	-	2		2	О, Т
Тема 2	Визуализация данных	8	2	2	2		2	О, ПОЗ
Тема 3	Классификация и регрессия (статистические методы)	14	4	2	4		4	О, ПОЗ
Тема 4	Классификация и	10	2	4	2		2	Т, ПОЗ

№ п/п	Наименование тем (разделов),	Объем дисциплины (модуля), час.						Форма текущего контроля успеваемости, промежуточной аттестации**
		Всего	Контактная работа обучающихся с преподавателем по видам учебных занятий				СР	
			Л/ЭО, ДОТ*	ЛР/ЭО, ДОТ*	ПЗ/ЭО, ДОТ*	КСР		
	регрессия (Машинное обучение)							
Тема 5	Кластеризация	8	2	2	2		2	О, ПОЗ
Тема 6	Поиск ассоциативных правил	8	2	2	2		2	О, ПОЗ
Тема 7	Анализ и прогнозирование временных рядов	10	2	4	2		2	О, ПОЗ
Тема 8	Ансамбли моделей	8	2	2	2		2	ПКР
Промежуточная аттестация								Зачет с оценкой
Всего:		72		18	18	-	18	
Заочная форма обучения								
1 курс 3 семестр								
Тема 1	Технологии анализа данных	8	2				6	О, Т
Тема 2	Визуализация данных	8		2			6	О, ПОЗ
Тема 3	Классификация и регрессия (статистические методы)	8	2				6	О, ПОЗ
Тема 4	Классификация и регрессия (Машинное обучение)	8		2			6	Т, ПОЗ
Тема 5	Кластеризация	4					4	О, ПОЗ
2 курс 1 семестр								
Тема 6	Поиск ассоциативных правил	12		4	2		6	О, ПОЗ
Тема 7	Анализ и прогнозирование временных рядов	12		4	2		6	О, ПОЗ
Тема 8	Ансамбли моделей	8		4			4	ПКР
Промежуточная аттестация		4						Зачет с оценкой
Всего:		72	4	16	4		44	

* *Примечание: формы текущего контроля успеваемости на семинарах и практических занятиях - опрос (О), тестирование (Т), практико-ориентированные задания (ПОЗ), практическая контрольная работа (ПКР).*

3.2. Наименование и содержание разделов (тем) дисциплины

№ темы	Наименование раздела (темы)	Содержание раздела (темы)
1	Технологии анализа данных	Введение в анализ данных. Принципы анализа данных. Виды и типы данных. Задача консолидации. Введение в хранилища данных. Основные концепции хранилищ данных. Подготовка данных к анализу. Технологии KDD и Data Mining. Задачи Data Mining. Модели Data Mining. Методы Data Mining. Практическое применение Data Mining. Аналитические платформы.
2	Визуализация данных	Введение в визуализацию. Визуализаторы общего назначения. OLAP-анализ. Визуализация данных в аналитической платформе Deductor. Компьютерный практикум «Разработка системы аналитической отчетности».
3	Классификация и регрессия (статистические методы)	Решение задачи классификации в аналитической платформе Deductor. Введение в классификацию и регрессию. Множественная линейная регрессия. Регрессия с категориальными входными переменными. Методы отбора переменных в регрессионные модели. Ограничения применимости регрессионных моделей. Нелинейные регрессионные модели. Основы логистической регрессии. Множественная логистическая регрессия. Оценка эффективности и сравнение моделей. ROC-анализ. Оценка ошибки модели. Издержки ошибочной классификации. LIFT- и PROFIT-кривые. Компьютерный практикум «Скоринговые модели для оценки кредитоспособности заемщиков».
4	Классификация и регрессия (Машинное обучение)	Введение в деревья решений. Алгоритмы построения деревьев решений. Упрощение деревьев решений. Введение в нейронные сети. Принципы построения нейронных сетей. Процесс обучения нейронной сети. Обучение в условиях несбалансированности классов. Компьютерный практикум «Повышение эффективности массовой рассылки».
5	Кластеризация	Введение в кластеризацию. Меры близости, используемые в алгоритмах кластеризации. Методы кластерного

		анализа. Иерархическая кластеризация. Алгоритм кластеризации k-means. Сети Кохонена. Карты Кохонена. Решение задачи кластеризации средствами аналитической платформы Deductor. Компьютерный практикум «Сегментация клиентов компании».
6	Поиск ассоциативных правил	Ассоциативные правила. Алгоритм Apriori. Иерархические ассоциативные правила. Последовательные шаблоны. Компьютерный практикум «Ассоциативные правила в стимулировании розничных продаж».
7	Анализ и прогнозирование временных рядов	Введение в прогнозирование. Временной ряд и его компоненты. Модели прогнозирования. Исследование наборов данных и выбор метода прогнозирования. Ошибки прогноза. Оценка адекватности выбранного метода прогнозирования. Методы сглаживания и скользящие средние. Декомпозиция временного ряда. Регрессионный анализ временных рядов. Анализ временных рядов средствами аналитической платформы Deductor. Компьютерный практикум «Прогнозирование продаж».
8	Ансамбли моделей	Введение в ансамбли моделей. Бэггинг. Бустинг. Альтернативные методы построения ансамблей.

4. ФОНД ОЦЕНОЧНЫХ СРЕДСТВ ПРОМЕЖУТОЧНОЙ АТТЕСТАЦИИ ПО ДИСЦИПЛИНЕ (МОДУЛЮ)

4.1. Формы и методы текущего контроля успеваемости и промежуточной аттестации

Текущий контроль является одним из составляющих оценки качества освоения образовательных программ, направленный на проверку знаний, умений и навыков обучающихся. Основными задачами текущего контроля успеваемости является повышение качества и прочности знаний студентов, приобретение и развитие навыков самостоятельной работы, а также повышение активности студентов на занятиях. Данный вид контроля стимулирует у студентов стремление к систематической самостоятельной работе по изучению дисциплины. Текущий контроль представляет собой регулярно осуществляемую проверку усвоения учебного материала. Данная оценка предполагает систематичность, непосредственно коррелирующая с требованием постоянного и непрерывного мониторинга качества обучения.

Оценка знаний, умений и навыков осуществляется на всех практических занятиях и лабораторных работах по всем формам обучения в соответствии с целями и задачами занятия.

Промежуточная аттестация представляет собой процедуру, проводимую с целью определения степени освоения обучающимися образовательной программы, в т. ч. отдельной части или всего объема учебного предмета, курса, дисциплины образовательной программы. Она проводится в формах, определенных учебным планом, и в порядке, установленном РАНХиГС.

4.1.1. В ходе реализации дисциплины «Информационные технологии в прогнозно-аналитической деятельности» **при проведении практических занятий и лабораторных работ** используются следующие методы текущего контроля успеваемости обучающихся:

устный и письменный контроль, предполагающий использование следующих форм контрольных средств:

– **устный опрос**;

– **тестирование** – учебная технология, позволяющая измерять знания, умения и навыки студентов, состоящая из тестовых заданий и формализованных процедур проведения, обработки и анализа результатов;

– **практико-ориентированное задание (компьютерный практикум)** – задание, в котором обучающемуся предлагают решить реальную профессионально-ориентированную аналитическую задачу на компьютере с использованием программы Deductor Academic.

Формы текущего контроля успеваемости по темам (разделам) дисциплины:

Тема и/или раздел	Формы текущего контроля успеваемости
Тема 1. Технологии анализа данных	Опрос Тестирование
Тема 2. Визуализация данных	Опрос Практико-ориентированное задание (компьютерный практикум)
Тема 3. Классификация и регрессия (статистические методы)	Опрос Практико-ориентированное задание (компьютерный практикум)
Тема 4. Классификация и регрессия (Машинное обучение)	Тестирование Практико-ориентированное задание (компьютерный практикум)
Тема 5. Кластеризация	Опрос Практико-ориентированное задание (компьютерный практикум)

Тема 6. Поиск ассоциативных правил	Опрос Практико-ориентированное задание (компьютерный практикум)
Тема 7. Анализ и прогнозирование временных рядов	Опрос Практико-ориентированное задание (компьютерный практикум)
Тема 8. Ансамбли моделей	Практическая контрольная работа

4.1.2. Промежуточная аттестация по дисциплине «Информационные технологии в прогнозно-аналитической деятельности» происходит в форме зачета с оценкой.

Зачет с оценкой проводится методом устного опроса студента (диалога преподавателя со студентом), цель которого заключается в выявлении индивидуальных достижений студента по освоению основных положений дисциплины в объеме требований учебной программы, и решения практической задачи в компьютерной программе Deductor Academic.

Процедура проведения зачета с оценкой предусматривает:

- получение студентом билета, включающего один вопрос и одно практическое задание для решения на компьютере в программе Deductor Academic;
- подготовку в течение 30 минут, в процессе которой студент может составить конспект ответа на вопрос билета и решить практическую задачу;
- ответ на вопрос, сформулированный в билете, и объяснение решения практической задачи, а также по мере необходимости на дополнительные вопросы.

Аттестация производится в соответствии со шкалой оценивания результатов обучения (пункт 4.3 настоящей рабочей программы).

4.2. Материалы текущего контроля успеваемости обучающихся

4.2.1. Средства оценивания результатов текущего контроля

Оценивание знаний, умений и навыков по учебной дисциплине «Информационные технологии в прогнозно-аналитической деятельности» осуществляется посредством использования следующих видов оценочных средств:

- устное собеседование (в рамках проведения опросов);
- тестирование;
- выполнение практико-ориентированных практических заданий на компьютере с использованием аналитической платформы Deductor Academic;
- контрольная работа.

Устное собеседование

Устные собеседования проводятся во время практических занятий и лабораторных работ.

Тематика обсуждаемых вопросов собеседования не должна выходить за рамки объявленной для данного занятия темы.

При оценке устного собеседования анализу подлежит точность формулировок, связность изложения материала, обоснованность суждений, опора на нормативные документы.

Тестирование

Тестирование — учебная технология, позволяющая измерять знания, умения и навыки студентов, состоящая из тестовых заданий и формализованных процедур проведения, обработки и анализа результатов; тесты предоставляются обучающимся в качестве раздаточного материала на бумажном или электронном носителе. При проведении тестирования обучающемуся запрещается пользоваться дополнительной литературой. На выполнение теста обучающемуся дается 30 минут.

Практические задания

Практические задания осуществляются в форме выполнения практико-ориентированных заданий на компьютере с использованием аналитической платформы Deductor Academic. В течение семестра студент под руководством преподавателя выполняет шесть практических работ. После выполнения практического задания преподаватель проводит опрос по теме занятия.

Контрольная работа

На последнем занятии студент выполняет практическую контрольную работу с использованием программы аналитической платформы Deductor Academic. Данный вид работы проверяет усвоение полученных в ходе обучения умений и навыков. Время выполнения контрольной работы составляет 2 часа. Успешное выполнение контрольной работы является обязательным для допуска к промежуточной аттестации.

4.2.2. Шкалы оценивания результатов текущего контроля

Шкала оценивания результатов устных собеседований на практических занятиях и лабораторных работах.

Обозначения		Формулировка требований к степени освоения дисциплины
Цифр.	Оценка	
2	Неудовлетворительно	Студент показывает недостаточные знания программного материала, не способен аргументировано и последовательно его излагать, допускаются грубые ошибки в ответах, неправильно отвечает на поставленный вопрос или затрудняется с ответом.

3	Удовлетворительно	Студент показывает достаточные, но не глубокие знания программного материала; при ответе не допускает грубых ошибок или противоречий, однако в формулировании ответа отсутствует должная связь между анализом, аргументацией и выводами. Для получения правильного ответа требуются уточняющие вопросы.
4	Хорошо	Студент показывает глубокие знания материала, грамотно его излагает, достаточно полно отвечает на поставленный вопрос и дополнительные вопросы, умело формулирует выводы. В то же время при ответе допускает несущественные погрешности.
5	Отлично	Студент показывает полные и глубокие знания программного материала, логично и аргументировано отвечает на поставленный вопрос, а также дополнительные вопросы, показывает высокий уровень теоретических знаний.

Шкала оценивания тестирования.

Обозначения		Формулировка требований к степени освоения дисциплины
Цифр.	Оценка	
2	Неудовлетворительно	50% и меньше правильных ответов.
3	Удовлетворительно	51–69% правильных ответов.
4	Хорошо	70–84% правильных ответов
5	Отлично	85–100% правильных ответов.

Шкала оценивания результатов практических заданий на практических занятиях.

Обозначения		Формулировка требований к степени освоения дисциплины
Цифр.	Оценка	
2	Неудовлетворительно	Студент неправильно решает практическую задачу, не делает выводов по ее результатам, не может объяснить технологию ее решения, показывает полное незнание теоретических аспектов, на дополнительные, уточняющие вопросы не отвечает.
3	Удовлетворительно	Студент допускает несколько незначительных ошибок в решении практической задачи, делает неполные выводы по ее результатам либо недостаточно аргументирует свое решение; отвечает на вопрос о технологии ее решения, но при ответе допускает неточности, что требует дополнительных вопросов.
4	Хорошо	Студент логично и правильно решает практическую зада-

		чу, делает грамотные выводы по ее результатам, отвечает на вопрос о технологии ее решения, достаточно аргументирует свое решение, но при ответе допускает погрешности.
5	Отлично	Студент логично и правильно решает практическую задачу, делает грамотные выводы по ее результатам, полно отвечает на вопрос о технологии ее решения, подробно аргументирует свое мнение со ссылками на норму закона, показывает хорошее знание теоретических аспектов.

Шкала оценивания выполнения контрольной работы.

Обозначения		Формулировка требований к степени освоения дисциплины
Цифр.	Оценка	
2	Неудовлетворительно	Выполнение менее 50 % объема задания контрольной работы, наличие существенных замечаний, студент делает неправильные выводы по ее результатам либо не аргументирует свое решение.
3	Удовлетворительно	Выполнение не менее 50 % объема задания контрольной работы (возможно наличие замечаний), студент делает неполные выводы по ее результатам либо недостаточно аргументирует свое решение.
4	Хорошо	Выполнение не менее 70 % объема задания контрольной работы (возможно наличие несущественных замечаний), наличие в работе аргументированных выводов на основе проведенного анализа
5	Отлично	Выполнение не менее 85% объема задания контрольной работы без замечаний, наличие в работе аргументированных выводов на основе проведенного анализа.

4.2.3. Типовые контрольные задания или иные оценочные материалы, необходимые для оценки знаний, умений, навыков и (или) опыта деятельности

Типовые оценочные материалы по теме 1 «Технологии анализа данных»:

Опрос

1. Аналитический и информационный подход к моделированию.
2. Основные этапы построения моделей.
3. Формы представления данных.
4. Типы и виды данных.
5. Что понимается под транзакцией?
6. KDD — методика извлечения знаний.

7. Классификация задач Data Mining.
8. Какие основные виды источников данных вы знаете?
9. В чем заключается процедура консолидации данных, каковы ее цели?
10. Какие задачи решаются при консолидации?
11. Какие причины мешают корректной аналитической обработке и требуют использования методов очистки данных?
12. В чем заключается цель процедуры обогащения данных?
13. Назначение систем OLTP и СППР (Системы Поддержки Принятия Решения)?
14. Отличия OLTP-систем и СППР?
15. В чем заключаются основные различия между ХД и обычными базами данных?
16. Какие функциональные требования предъявляются к ХД?
17. Зачем выполняется агрегирование данных?
18. Что такое метаданные и какова их роль в процессе функционирования ХД?
19. Какие архитектуры ХД вам известны?
20. Какова роль измерений и фактов в многомерной модели данных?
21. Какие действия применяются к измерениям для извлечения нужной информации из многомерного куба?
22. В какой форме хранятся данные в РХД?
23. Что такое витрина данных?

Тестирование

1. Проставьте соответствие между классами задач Data Mining и их определениями:

Классификация	выявление закономерностей между связанными событиями
Регрессия	установление зависимости непрерывной выходной переменной от входных переменных
Кластеризация	группировка объектов (наблюдений, событий) на основе данных, описывающих свойства объектов
Ассоциация	установление зависимости дискретной выходной переменной от входных переменных

2. К какой форме данных относится информация на веб-странице в Интернете?
 - a. структурированные
 - b. слабоструктурированные
 - c. неструктурированные
3. С какими формами данных, как правило, работают методы анализа данных?
 - a. структурированные

- b. неструктурированные и слабоструктурированные
- c. слабоструктурированные
- d. неструктурированные
- e. структурированные и слабоструктурированные

4. Проставьте соответствие между пропущенными словами и текстом:

[...] данные отражают отдельные факты предметной области.	Слабоструктурированные
[...] данные имеют некоторые правила и форматы их задания, но в самом общем виде.	Неструктурированные
[...] данные являются произвольными по форме.	Структурированные

5. Можно ли статистический пакет отнести к классу аналитических платформ?
- a. Да
 - b. Нет
6. Кто является основателем направления Knowledge Discovery in Databases и Data Mining?
- a. Л. Заде
 - b. Б. Инмон
 - c. Г. Пятецкий-Шапиро
 - d. Б. Коско
7. Каковы особенности данных, накапливаемых в компаниях? (Выберите три варианта)
- a. Данные редко накапливаются специально для решения задач анализа
 - b. Данные всегда представлены в структурированной форме
 - c. Как правило, данные содержат ошибки, аномалии и пропуски
 - d. Нередко имеют большой объем
 - e. Почти всегда носят неполный, фрагментарный характер
8. Могут ли числовые данные быть дискретного вида при решении задачи анализа?
- a. Нет
 - b. Да
9. Выберите неверный вариант:
- a. Эксперт выдвигает гипотезы и строит модели для проверки достоверности гипотез
 - b. Аналитик – это специалист в области анализа и моделирования

- с. Эксперт – это специалист предметной области, профессионал, который за годы обучения и практической деятельности научился эффективно решать задачи, относящиеся к конкретной предметной области
- д. Эксперт является связующим звеном между специалистами разных уровней и областей

10. Проставьте соответствие между пропущенными словами и текстом:

При [...] подходе модель не «подстраивается» под действительность, а исследователь пытается подогнать модель таким образом, чтобы она адекватно отражала реальность.	информационном
В [...] подходе отправной точкой являются данные, характеризующие исследуемый объект, и модель «подстраивается» под действительность.	аналитическом

11. Укажите основные виды моделей:

- а. эмпирические
- б. теоретические
- с. смешанные
- д. экспериментальные
- е. статистические

12. Как переводится Knowledge Discovery in Databases (KDD)?

- а. извлечение знаний из баз данных
- б. тиражирование знаний
- с. «раскопка» данных

13. Отметьте неверное рассуждение:

- а. Под моделированием понимается процесс исследования реальной системы, включающий построение модели, изучение ее свойств и перенос полученных сведений на моделируемую систему
- б. Построение моделей является единственным способом изучения окружающего мира
- с. Модель – это объект, который имеет сходство в некоторых отношениях с прототипом и служит средством описания и объяснения, прогнозирования поведения прототипа
- д. Модель – это объект или описание объекта, системы для замещения одной системы другой системой для лучшего изучения оригинала или воспроизведения каких-либо его свойств.

14. Какие существуют основные способы решения аналитических задач?

- а. Построение и использование моделей
- б. Верны оба варианта

с. Извлечение и визуализация данных

15. Какая самая распространенная модель хранения структурированных данных?
- а. текст
 - б. граф
 - с. дерево
 - д. матрица
 - е. таблица
16. На какие две группы подразделяются методы Data Mining по принципу работы с исходными обучающими данными?
- а. статистические методы
 - б. непосредственное использование данных или сохранение данных
 - с. кибернетические методы
 - д. выявление и использование формализованных закономерностей
17. Какие из перечисленных ниже пунктов являются названиями стадий Data Mining?
- а. индукция правил
 - б. свободный поиск
 - с. анализ исключений
 - д. прогностическое моделирование
18. Любые, неизвестные ранее сведения о каком-либо событии, сущности, процессе и т.п., являющиеся объектом некоторых операций, для которых существует содержательная интерпретация, являются...
- а. знаниями
 - б. информацией
 - с. данными
19. У основания так называемой информационной пирамиды находится категория ...
- а. информация
 - б. знания
 - с. данные
20. Задачи Data Mining, в зависимости от используемых моделей подразделяются на:
- а. обучение без учителя
 - б. прогнозирующие
 - с. дескриптивные
 - д. обучение с учителем

Типовые оценочные материалы по теме 2 «Визуализация данных»:

Опрос

1. Каковы цели и задачи визуализации данных в аналитических технологиях?
2. Зачем нужен визуальный контроль данных после их загрузки в аналитическое приложение?
3. Чем отличаются средства визуализации общего назначения от специализированных средств визуализации?
4. Какие средства визуализации данных относят к визуализаторам общего назначения и почему?
5. В чем заключается OLAP-анализ и каковы его цели?
6. Какова структура OLAP-куба? Какие манипуляции с измерениями можно производить, чтобы сделать представление куба более информативным?

Практико-ориентированное задание (компьютерный практикум) «Получение аналитической отчетности в Deductor Academic»

Легенда: аптечная сеть ООО «ФАРМА – Центр» имеет три торговых точки в разных городах России. Вы получили задание проанализировать работу фирмы за прошедший период и подготовить наглядный отчет.

Задача:

- Сформировать многомерный отчет, отражающий динамику сумм продаж по месяцам года в разрезе товарных групп и аптек
- Сформировать многомерный отчет и график загруженности торговых точек по времени суток.
- Сформировать многомерный отчет и график загруженности торговых точек по дням недели.

1. Подключить хранилище данных «Фармация».
2. Извлечение данных из хранилища.

С помощью Мастера импорта выбрать тип источника данных Deductor Warehouse, на следующем шаге — ХД Продажи аптечной сети, затем — процесс Продажи.

Определить, какие измерения и атрибуты из выбранного на предыдущем шаге процесса следует импортировать: Дата, Наименование отдела, Наименование товара, Товарная группа, Час.

Указать импортируемые факты и способы их агрегирования: Количество – Сумма, Сумма – Сумма.

Определить срезы для выбранных измерений: необходимы данные за последнее полугодие.

Запустить процесс извлечения данных.

3. Трансформация данных.

3.1 Обработчик «Дата и Время».

В вашем распоряжении имеется только измерение Дата, а для построения отчетов требуются измерения Месяц, День недели. Месяц года и День недели из даты можно получить, применив к узлу импорта Мастер обработки, а затем выбрав обработчик «Дата и время» из группы Трансформация данных.

3.2 Обработчик «Квантование» — преобразование непрерывных значений в дискретные.

В вашем распоряжении имеется измерение Час, а для построения отчета требуется измерение Время суток.

Применить к узлу «Импорт» обработчик «Квантование». Задать параметры квантования непрерывной переменной: Способ — По интервалам; Количество интервалов — 4, Используемое поле — Время суток. Задать границы интервалов: 8–12, 12–15, 15–19, 19–21.

4. Получение OLAP-отчета

Применить к узлу «Настройка набора данных» Мастер визуализации со способом отображения данных — куб. Указать измерения и факты: в данном случае измерения — это Месяц, День недели, Аптека и Товарная группа, а факты — Количество и Сумма проданных товаров (с агрегацией «Сумма»).

Задать размещение измерений по строкам/столбцам: Колонки — Аптека, Строки — Месяц

На последнем шаге определить, какие факты нужно отображать в кубе на пересечении измерений, способы их агрегации и варианты отображения.

Далее задания выполните самостоятельно.

Скопируйте последний узел сценария. Сформируйте многомерный отчет и график загруженности торговых точек по времени суток. На какие часы приходятся пики продаж?

Скопируйте последний узел сценария. Сформируйте многомерный отчет и график загруженности торговых точек по дням недели. На какие дни приходятся пики продаж?

Скопируйте последний узел сценария. Определите 10 товаров, дающих самый большой объем продаж в ноябре и декабре. Обратите внимание, что при последних настройках OLAP-куба поле Товар было информационным. К последнему узлу сценария нужно применить Мастер визуализации еще раз и назначить Товар измерением.

5. Аналитическая отчетность.

Для создания аналитической отчетности необходимо в меню Вид выбрать пункт Отчеты. Отчеты строятся в виде древовидного иерархического списка, каждым узлом которого является отдельный отчет или папка, содержащая несколько отчетов. Каждый узел дерева отчетности связан со своим узлом в дереве сценария.

Для каждого отчета настраивается свой способ отображения (таблица, гистограмма, куб, кросс-диаграмма и т.п.).

Задание: создать папку Динамика продаж, в которую поместить все созданные отчеты.

Типовые оценочные материалы к семинару по теме 3 «Классификация и регрессия (статистические методы)»:

Опрос

1. В чем заключается задача классификации?
2. Какие алгоритмы используются для решения задач классификации?
3. По какому принципу строится таблица сопряженности и каковы цели ее использования?
4. Как рассчитываются показатели чувствительности и специфичности модели, в чем состоит их смысл?
5. Как строится ROC-кривая?
6. Как можно определить оптимальный порог отсечения?
7. Как строится LIFT-кривая?
8. Какие достоинства и недостатки имеют модели логистической регрессии применительно к задаче кредитного скоринга?

Практико-ориентированное задание (компьютерный практикум) «Скоринговые модели для оценки кредитоспособности заемщиков»

Постановка задачи.

В коммерческом банке имеется продукт «Нецелевой потребительский кредит»: кредиты предоставляются на любые цели с принятием решения в течение нескольких часов. За это время проверяются минимальные сведения о клиенте, в основном, такие, как отсутствие криминального прошлого и кредитная история в других банках.

В банке накоплена статистическая информация о заемщиках и качестве обслуживания ими долга за несколько месяцев. Руководство банка, понимая, что отсутствие адекватных математических инструментов, позволяющих оптимизировать риски, не способствует расширению розничного бизнеса в области потребительского кредитования, поставило перед отделом розничных рисков задачу разработать скоринговые модели с различными стратегиями кредитования, которые позволили бы управлять рисками и минимизировать число «безнадежных» заемщиков.

Решение задачи.

Импортировать файл с кредитными историями Loans.txt.

В задаче представлено 2709 кредитов с известными исходами платежей. Набор данных уже разбит на два множества — обучающее (80%) и тестовое (20%) — та-

ким образом, чтобы в каждом множестве доля «плохих» кредитов была примерно одинаковой.

В структуре данных:

Поле	Описание
Код	Служебный код заявки
Дата	Дата выдачи кредита
О/Д, %	Коэффициент Обязательства/Доход в процентах (Вещественный)
Возраст	Возраст заемщика на момент принятия решения о выдаче кредита
Проживание	Основание для проживания: собственник, муниципальное жилье, аренда
Срок проживания в регионе	Менее 1 года, от 1 года до 5 лет, свыше 5 лет
Семейное положение	Холост/не замужем, женат/замужем, разведен/вдовство, другое
Образование	Среднее, среднее специальное, высшее
Стаж работы на последнем месте	Менее 1 года, от 1 года до 3 лет, свыше 3 лет
Уровень должности	Сотрудник, руководитель среднего звена, руководитель высшего звена
Кредитная история	Информация из бюро кредитных историй. Если есть информация о просрочках по прошлым кредитам, то клиенту присваивается категория «отрицательная» (тип данных Строковый);
Просрочки свыше 60 дней	0 – отсутствовали, 1 – имели место (тип данных Целый);
Тестовое множество	Служебный признак, TRUE соответствует тестовому множеству (тип данных Логический).

1. С помощью обработчика «Калькулятор» на основе поля Просрочки более 60 дн. (COL12) получите новое поле Класс заемщика.

2. Постройте модель Логистической регрессии. Выходная переменная — Класс заемщика.

3. Запишите уравнение логистической регрессии.

4. Используя визуализатор «Таблица сопряженности», определите Чувствительность и Специфичность полученной модели.

5. Оцените качество модели.

6. Чему равен оптимальный порог отсечения?

— Определить оптимальный порог отсечения, используя визуализатор ROC-кривая.

— Определить оптимальный порог отсечения, используя правило

$$P = \frac{1}{\left(1 + C_{FN}/C_{FP}\right)}$$

C_{FN} — издержки, связанные с ложноотрицательными наблюдениями;

C_{FP} — издержки, связанные с ложноположительными наблюдениями.

Отношение издержек C_{FN}/C_{FP} задать самостоятельно.

7. Как изменяется кредитный рейтинг заемщика с каждым дополнительным иждивенцем?

8. Используя визуализатор «Что–если», определите к какому классу будут отнесены заемщики

— Мужчина, 43-х лет, состоит в браке, 2 иждивенца, доход — 50000 руб., опыт работы — 20 лет, срок проживания на одном месте — 10 лет., недвижимость 0, месячный платеж 7500 руб.

— Женщина, 30-ти лет, не замужем, 1 иждивенец, доход — 40000 руб., опыт работы — 6 лет, срок проживания — 10 лет, недвижимость 0, месячный платеж 5000 руб.

Типовые оценочные материалы по теме 4 «Классификация и регрессия (Машинное обучение)»:

Тестирование

1. Классификация — это ...
 - a. отнесение объектов к одной из заранее неизвестных групп
 - b. процесс формирования групп и отнесения объектов к одному из них.
 - c. отнесение объектов к одному из заранее известных классов
2. Задачи классификации решаются следующими алгоритмами:
 - a. алгоритмом Apriori
 - b. нейронные сети
 - c. линейной регрессии
3. Деревья решений относятся к группе (-ам) ...
 - a. кибернетических методов
 - b. методов кросс-табуляции
 - c. статистических методов
 - d. логических методов
4. Внутренний узел дерева решений называют также ...
 - a. узлом проверки
 - b. конечным узлом
 - c. вершиной
 - d. листом

5. Конечный узел дерева решений называют также ...
- a. узлом проверки
 - b. узлом решения
 - c. листом
6. При помощи метода деревьев решений решаются задачи ...
- a. классификации и кластеризации
 - b. классификации и прогнозирования
 - c. кластеризации, классификации и прогнозирования
 - d. кластеризации и прогнозирования
7. Процесс отсечения ветвей или замена некоторых ветвей поддеревом ...
- a. происходит сверху вниз, т.е. является нисходящим
 - b. происходит снизу вверх, т.е. является восходящим
 - c. может быть, как нисходящим, так и восходящим
8. Алгоритм конструирования дерева решений ...
- a. не требует от пользователя выбора из набора входных атрибутов (независимых переменных), наиболее значимых
 - b. требует от пользователя выбора из набора входных атрибутов (независимых переменных), наиболее значимых
 - c. на вход алгоритма можно подавать все существующие атрибуты, алгоритм сам выберет наиболее значимые среди них, и только они будут использованы для построения дерева
9. Если зависимая переменная принимает непрерывные значения, то дерево решений решает задачу:
- a. классификации и численного прогнозирования
 - b. численного прогнозирования
 - c. классификации
10. Если зависимая переменная принимает дискретные значения, при помощи метода дерева решений решается задача:
- a. классификации и численного прогнозирования
 - b. классификации
 - c. численного прогнозирования
11. Метод, который делает заключения относительно данной ситуации по результатам поиска аналогий, хранящихся в базе прецедентов относится к категории ...
- a. обучение с учителем"
 - b. самообучающейся системы
 - c. "обучение без учителя"
12. Явление переобучения характеризуется ...
- a. чрезмерно точным соответствием нейронной сети конкретному набору обучающих примеров, при котором сеть теряет способность к обобщению

- b. возникновением, в случае слишком долгого обучения, недостаточного числа обучающих примеров или слишком сложной структуры нейронной сети
 - c. возникновением, в случае слишком долгого обучения, слишком большого числа обучающих примеров или слишком сложной структуры нейронной сети
13. Многослойный персептрон – это сеть ...
- a. обратного распространения сигнала (с обратными связями)
 - b. в которой входной сигнал преобразуется в выходной, проходя последовательно через несколько слоев
 - c. прямого распространения сигнала (без обратных связей)
14. В многослойном персептроне ...
- a. должен быть хотя бы один скрытый слой
 - b. присутствие нескольких скрытых слоев оправдано лишь в случае использования нелинейных функций активации
 - c. может быть какое угодно количество скрытых слоев, они также могут вообще отсутствовать
15. Заполните пропуски в формулировке: "... – выходная связь нейрона, с которой сигнал (возбуждения или торможения) поступает на... следующих нейронов"
- a. синапс, аксоны
 - b. аксон, точку ветвления
 - c. аксон, синапсы
16. Однонаправленные входные связи, соединенные с выходами других нейронов – это ...
- a. синапсы
 - b. слои сети
 - c. аксоны
17. Слой нейронной сети – это ...
- a. выходная связь данного нейрона, с которой сигнал (возбуждения или торможения) поступает на синапсы следующих нейронов
 - b. один или несколько нейронов, на входы которых подается один и тот же общий сигнал
 - c. однонаправленные входные связи, соединенные с выходами других нейронов
18. Ошибкой обучения нейронной сети называется ...
- a. переобучение нейронной сети
 - b. разность между желаемым и полученным на выходе сигналами
 - c. целевая функция, требующая минимизации в процессе управляемого обучения нейронной сети

19. Целевая функция, требующая минимизации в процессе управляемого обучения нейронной сети – это ...

- а. функция ошибок
- б. функция переобучения
- с. ошибка обучения

Практико-ориентированное задание (компьютерный практикум) «Прогнозирование ухода к другому оператору клиентов телекоммуникационной сети»

Предлагается решить задачу из области прогнозирования ухода к другому оператору клиентов телекоммуникационной сети.

В двух файлах telcojan.txt (обучающая выборка) и telcofeb.txt (тестовая выборка) имеются статистические данные клиентах телекоммуникационной компании по состоянию на январь и февраль некоторого года.

Структура файлов одинакова и содержит уже отобранные аналитиками наиболее значимые характеристики, влияющие на уход клиента (отказ от пользования услугами): демографическая информация и используемый тариф.

№ п/п	Имя поля	Описание	Тип
1	Срок клиента	Количество месяцев, прошедших с момента подключения	Целый
2	Возраст	Возраст клиента	Целый
3	Адрес	Срок проживания по последнему адресу	Целый
4	Стаж	Стаж работы общий	Целый
5	Семейное положение	Семейное положение клиента	Строковый: – не в браке – в браке
6	Образование	Уровень образования	Строковый: – начальное – среднее проф. – среднее спец. – высшее – уч. степень
7	Пенсионер	–	да/нет
8	Пол	–	муж/жен
9	Тариф	Текущий тариф клиента	Строковый: – «Базовый»

			– «Плюс» – «Эконом» – «Анлим»
10	Уход	Отказ от пользования услугами (соответствует значению 1)	0/1

1. Постройте модель «Дерево решений».
2. Используя визуализатор «Таблица сопряженности», определите Чувствительность и Специфичность полученной модели.
3. Оцените качество модели. Сравните ее с моделью, полученной на предыдущем этапе.
4. С помощью визуализатора «Значимость атрибутов» охарактеризуйте, степень нелинейной зависимости между каждым фактором, включенным в модель, и независимой переменной.
5. Постройте модель «Нейронная сеть». Используйте алгоритм BACKPROP.
7. Используя визуализатор Таблица сопряженности, определите Чувствительность и Специфичность полученной модели.
8. Оцените качество модели. Сравните ее с моделью, полученной на предыдущем этапе.
9. Переобучите нейронную сеть
 - алгоритмом RPOP;
 - изменив количество скрытых слоев и количество нейронов в каждом слое;
 - уменьшив количество входных переменных (в этом случае следует определить новое значение требуемого количества нейронов).
- Удалось ли значительно улучшить общую ошибку классификации?
10. Сравните качество моделей между собой и с моделью, построенной на шаге Дерево решений.
11. Сформулируйте рекомендации по использованию результатов моделирования в работе компании.

Типовые оценочные материалы по теме 5 «Кластеризация»:

Опрос

1. Дайте формальную постановку задачи кластеризации.
2. Каковы цели кластеризации?
3. Что такое метрика расстояния?
4. Каковы основные шаги алгоритма k-means?
5. Как в алгоритме k-means определяется принадлежность точки к тому или иному кластеру?
6. Что такое центроид (центр тяжести кластера) и какова его роль в алгоритме k-means? Что является условием остановки алгоритма k-means?

7. Каковы преимущества и недостатки алгоритма k-means?
8. Какова структура сети Кохонена? Какова роль входных и выходных нейронов карты?
9. В чем заключаются процедуры конкуренции, объединения и подстройки весов?
10. По какому правилу производится подстройка весов нейронов карты?
11. Что такое радиус обучения и какую роль он играет в процессе построения карты?
12. Чем определяется требуемое число выходных нейронов сети Кохонена?

Практико-ориентированное задание (компьютерный практикум) «Сегментация клиентов компании»

Легенда: руководство отдела маркетинга сети книжных магазинов по продаже книг различной направленности (от деловой литературы до художественной и развлекательной) поставило задачу сегментации своих постоянных покупателей, выявления их предпочтений. Предполагается, что это будет использоваться в будущем для оптимизации ассортимента книжных магазинов, проведения рекламных акций и для других целей.

В качестве исходных данных для сегментации выступают данные о постоянных покупателях книжной сети, у которых имеются дисконтные карты. Благодаря ним, о них известна следующая информация: пол, возраст и осуществленные покупки.

Задание. Вы выступаете в роли бизнес-аналитика. В вашем распоряжении имеются данные клиентах и их покупках за последний год (файл bookstore.txt).

Название поля	Описание
Код клиента	Уникальный код клиента
Возраст	Возраст клиента
Пол	Пол клиента
Посещений за год	Число посещений клиентом книжных магазинов, закончившихся покупкой товаров
Книги категории А	Число приобретенных за год клиентом книг из категории А
Книги категории В	Число приобретенных за год клиентом книг из категории В
Книги категории С	Число приобретенных за год клиентом книг из категории С
Книги категории D	Число приобретенных за год клиентом книг из категории D
Книги категории E	Число приобретенных за год клиентом книг из категории E

Среднее число книг в чеке	Среднее число книг в чеке (за год)
Средняя сумма чека	Средняя сумма чека (за год)

Для удобства проведения анализа все книги разделены на пять основных категорий.

Категория книжной продукции	Описание
Книги категории А	Деловая, научная, техническая и учебная литература
Книги категории В	Художественная литература
Книги категории С	Книги по литературе и искусству и культуре
Книги категории D	Народная медицина и здоровье
Книги категории Е	Детективы, любовные романы, фантастика

Перед вами стоит задача сегментации клиентской базы при помощи самоорганизующихся карт Кохонена.

1. Разбейте всех клиентов на сегменты. Перед построением аналитической модели рекомендуется ответить на следующие вопросы:

- Какие признаки включить в алгоритм кластеризации?
- Какие их признаков более значимые, какие – нет?
- Какой размер карты Кохонена выбрать?

2. Проанализировав карты, определите, какое количество кластеров дает самую наглядную картину?

3. Оцените емкость каждого сегмента (мощность кластера).

4. Охарактеризуйте каждый сегмент и дайте ему краткое название.

5. Постарайтесь описать социальный портрет покупателей каждого сегмента. Дайте название каждому сегменту. При интерпретации кластеров вы можете привлекать любую статистическую и (или) графическую информацию по кластеру: количество объектов, среднее, минимум, максимум и т.д.

Типовые оценочные материалы по теме 6 «Поиск ассоциативных правил»:

Опрос

1. Что такое ассоциация? В чем заключается основная задача анализа рыночной корзины?
2. Что представляет собой транзакция в теории ассоциативных правил?
3. Как определяются поддержка и достоверность и какова их роль в процессе поиска ассоциативных правил?
4. Чем определяется значимость и полезность ассоциативных правил и какими показателями она характеризуется?

5. Использование алгоритма `apriori` для поиска ассоциативных правил.
6. На какие группы можно разделить массив ассоциативных правил?

Практико-ориентированное задание (компьютерный практикум) «Ассоциативные правила в стимулировании розничных продаж»

Легенда: компания является дистрибьютером спортивных, горных и дорожных велосипедов и комплектующих к ним компании Adventure Work Cycle на территории России и стран СНГ. Офисы компании работают в шести городах России, на Украине и Казахстане. У фирмы есть склад и собственная сборочная база.

Отдел маркетинга заинтересован в увеличении продаж через интернет-магазин, размещенный на сайте компании. Для этого важно знать, какие товары покупатели смогут выбрать в дальнейшем в зависимости от того, что уже имеется в их корзинах. Такой прогноз позволит также оптимизировать структуру сайта — товары, часто покупаемые вместе, будут расположены по соседству на одной web — странице. Для решения поставленной задачи отдел маркетинга предоставил данные о 5 тыс. чеков от предыдущих покупателей.

Решение.

1. Импортировать данные из текстового файла `cycle_store.txt`.

В этом файле ID — код чека; ITEM — наименование товара.

2. К узлу импорта добавить обработчик Ассоциативные правила.

Столбец ID – идентификатор транзакции, а ITEM – ее элемент.

3. Настроить параметры построения ассоциативных правил (параметры алгоритма `apriori`): максимальная поддержка — 20%, минимальная достоверность ассоциативного правила 40%, максимальная достоверность — 90%

4. Выбрать все доступные специализированные визуализаторы (Правила, Популярные наборы, Дерево правил) и визуализатор Таблица.

5. Проанализировать сформированные ассоциативные правила. Сколько правил было выявлено?

6. Рассмотрите правило Шина HL Mountain → Велокамера Mountain.

- Если покупатель решил приобрести что-либо в данном магазине, то какова вероятность того, что это будет набор Шина HL Mountain + Велокамера Mountain?
- Если клиент положил в корзину товар Шина HL Mountain, то с какой вероятностью он купит еще и Велокамеру Mountain?
- Во сколько раз чаще клиент, купивший Шину HL Mountain, выбирает Велокамеру Mountain, чем какой-либо другой товар?

7. Проанализируйте Популярные наборы. Какие камеры продаются чаще всего? Какие велосипеды продаются чаще всего?

8. Какую достоверность имеют полученные правила?

Какие из полученных правил являются Полезными, Тривиальными, Непонятными?

9. Запустить заново алгоритм А PRIORI с интервалом допустимой достоверности 25—40%; не рассматривать правила, в следствиях и условиях которых содержатся велосипеды, шины и велокамеры (очевидные лидеры продаж).

Получены ли после второго прогона полезные правила?

10. Основываясь на полученных результатах,

- предсказать то, какие товары покупатели могут выбрать в зависимости от того, что уже есть в их корзинах;
- выявить наиболее популярные товарные наборы, состоящие из более чем 1 предмета;
- предложить рекламные акции типа «каждому купившему А и В товар С в подарок»;
- предложить комплекс мероприятий типа
 - а) разместить товар А и товар В рядом на витрине;
 - б) разместить товар А и товар В на большом расстоянии друг от друга;
 - в) сформировать подарочные наборы «Товар А + Товар В»;
 - г) сформируйте подарочные наборы «Товар А + Товар В + плохо продаваемый товар»;
 - д) поднять цену на одно, снизить на другое;
 - е) заказать комплекты товаров А и товаров В одного бренда и серии.

Типовые оценочные материалы по теме 7 «Анализ и прогнозирование временных рядов»:

Опрос

1. Детерминированная и случайная составляющие временного ряда.
2. Модели временных рядов.
3. Компоненты временного ряда.
4. Исследование временных рядов и автокорреляция.
5. В чем заключается принцип построения ретропрогноза?
6. Оценка адекватности выбранного метода прогнозирования.
7. Методы сглаживания.
8. Прогнозирование методом скользящего среднего.
9. Декомпозиция временного ряда.
10. Регрессионный анализ временных рядов.

Практико-ориентированное задание (компьютерный практикум) «Прогнозирование объема продаж»

Легенда: оптовая компания занимается сбытом строительных материалов, ассортимент насчитывает несколько тысяч торговых позиций, объединенных в группы (сухие смеси, плитка, напольные покрытия, грунтовка и т.п.)

Перед вами поставлена задача: автоматизировать ежемесячный расчет товарного прогноза на следующие три периода.

1. Подключить хранилище данных materials.gdb
 2. Для импорта выбрать только процесс Продажи
 3. Выбрать для импорта все измерения, атрибуты и факты процесса Продажи, для фактов агрегирующей функцией является Сумма.
 4. Определить срез для выбранных измерений. Типовой сценарий будет построен на примере товарной группы «Грунтовка». Установить фильтр «Кроме последнего периода», т.к. данные текущего, еще не завершившегося месяца, в построении модели участия не принимают.
 5. Для последующей группировки по месяцам выполнить Преобразование даты к виду «Дата (Год+Месяц)».
 6. Выполнить Группировку по измерению «Дата (Год+Месяц)» и факту «Сумма» и/или «Количество».
 7. Построить график продаж. Оценить, содержит ли Кривая продаж шумы и выбросы, которые необходимо удалить для получения более качественного прогноза. Для сглаживания данных воспользоваться «Парциальной обработкой».
- Парциальная обработка может включать в себя удаление аномалий и сглаживание (фильтрацию) данных (если требуется). При настройке парциальной обработки указать для поля Количество:
- Редактирование аномальных значений с о Степенью подавления (Малая),
 - Сглаживание данных с Полосой пропускания — 80.
- Отобразить результат преобразования данных на диаграмме. Оценить качество сглаживания.
8. С помощью обработчика «Автокорреляция» определить сезонность продаж данной группы товара. Какие значения АКФ говорят о наличии Тренда и наличии Сезонности?
 9. Для построения модели прогноза воспользоваться инструментом «Скользящее окно». Скользящее окно — преобразование временного ряда к таблице смежных значений.
 10. Построить несколько моделей временных рядов.
- В качестве моделей использовать:
- «Наивную» модель скользящего среднего за предыдущие три месяца.

- Линейную регрессию с периодами (1,2,3) и (1,2,12).
- Нейронную сеть периодами (1,2,12) — Рекомендуется перебрать 2-3 модели нейронной сети, изменяя число нейронов в скрытом слое.

В качестве визуализаторов выбрать Диаграмму рассеивания и Диаграмму, на которой отобразить Количество и Количество_OUT (ретропрогноз).

11. По диаграммам оценить качество моделирования.

12. Вычислить среднюю процентную ошибку для каждой модели. Оценить качество моделирования.

13. К узлу лучшей модели применить обработчик «Прогнозирование». Обработчик Прогнозирование — позволяет получать предсказание значений временного ряда на число отсчетов, соответствующее заданному горизонту прогнозирования.

14. К узлу Прогнозирование применить Фильтр: Шаг прогноза = Не пустой.

15. После получения прогнозных цифр по группе товаров произвести разгруппировку прогнозных значений для получения прогноза для каждой товарной позиции (обработчик «Разгруппировка»).

Типовые оценочные материалы по теме 8 «Ансамбли моделей»:

Типовой вариант практической контрольной работы:

Предлагается решить задачу из области коллекторского скоринга на тему «Оптимизация выбора стратегии взыскания просроченной задолженности».

В файле debtors.txt имеются статистические данные о заемщиках банка, которые стали должниками, то есть имеют просроченную задолженность. Структура полей файла приведена в следующей таблице.

Имя поля	Описание	Тип данных
Код	Код заемщика	Целый
Сумма кредита	Сумма кредита в рублях	Вещественный
Срок кредита	Срок кредита в мес. (6, 12, 18)	Целый
Ежемесячный платеж	Размер аннуитетного ежемесячного платежа в рублях	Вещественный
Возраст	Возраст, лет	Целый
Пол	Пол клиента	Строковый
Срок просрочки	Срок просрочки в днях	Целый
Число платежей до выхода на просрочку	Сколько сделал заемщик платежей до того, как была впервые зафиксирована просроченная задолженность	Целый
Исполнительный лист	Имеется ли по заемщику исполнительный лист (судебное решение)	0 / 1
Остаток долга (тело)	Сумма в рублях остатка «тела» кредита	Вещественный

Остаток долга (проценты)	Сумма в рублях остатка «процентов» по кредиту	Вещественный
Штрафы	Начисленные штрафы и пенни	Вещественный
Просрочка/Долг	Отношение, рассчитанное на основе трех предыдущих полей по формуле: $(12) / (11 + 10)$	Вещественный
Возобновление платежей	Начал ли заемщик снова оплачивать кредит после работы специалиста коллекторского подразделения: 0 – работа велась, и возобновления платежей не последовало, 1 – работа велась, и последовало возобновление платежей, «пусто» – работа не велась	Целый
Тестовое множество	Признак участия примера в тестовом множестве	Логический

Известна следующая дополнительная информация:

Ориентировочная стоимость плана мероприятий по взысканию просроченной задолженности с 1 заемщика оценивается в 1 000 рублей (зарплата персоналу, оплата звонков, выездов и т.д.)

Средняя ожидаемая сумма погашения от 1 должника 12 000 рублей.

Задание:

1. Построить модели, позволяющие классифицировать должников на 2 класса в зависимости от исхода мероприятий, проводимого специалистом коллекторского подразделения банка (логистическая регрессия, дерево решений, многослойный персептрон). Разделение на обучающее и тестовое множество уже проведено – используйте поле Тестовое множество. За положительный исход примите Возобновление платежей = 1.

2. Построить LIFT-кривые для моделей (там, где это возможно) и сравнить модели.

3. Используя дополнительную информацию, выбрать лучшую модель с двух позиций:

- Максимизация ожидаемых поступлений от должников;
- Минимизация рабочего времени, которое будет потрачено на мероприятия по взысканию.

4. Ответьте на вопрос: какой из алгоритмов оказался наименее чувствителен к несбалансированности классов? Комментарии впишите к любому узлу-обработчику.

5. Объедините лучшие модели в Ансамбль моделей.

6. «Прогоните» через ансамбль моделей тех должников, к которым еще не применялись мероприятия по взысканию и получите список клиентов, с которыми нужно работать, ранжированный по убыванию вероятности положительного исхода.

7. Все важные визуализаторы вывести на панель отчетов. Свои комментарии впишите к любому узлу-обработчику.

4.3. Оценочные средства для промежуточной аттестации

Зачет с оценкой по дисциплине «Информационные технологии в прогнозно-аналитической деятельности» служит для оценки работы студента в течение всего времени обучения по данной дисциплине.

Зачет с оценкой проводится в конце семестра и призван выявить уровень, прочность и систематичность полученных студентом теоретических и практических знаний, приобретения навыков самостоятельной работы, развития творческого мышления, умение синтезировать полученные знания и применять их в решении практических задач.

Зачет с оценкой осуществляется в форме устного ответа и решения практической задачи на компьютере. Задание билета включает в себя один теоретический вопрос и одно практическое задание для решения в программе Deductor Academic.

Пример типового билета к зачету с оценкой по дисциплине «Информационные технологии в прогнозно-аналитической деятельности»:

РОССИЙСКАЯ АКАДЕМИЯ НАРОДНОГО ХОЗЯЙСТВА И ГОСУДАРСТВЕННОЙ СЛУЖБЫ ПРИ ПРЕЗИДЕНТЕ РОССИЙСКОЙ ФЕДЕРАЦИИ

Кафедра Национальной экономики

Дисциплина

«Информационные технологии в прогнозно-аналитической деятельности»

Билет №1

1. Основные этапы построения моделей. KDD — методика извлечения знаний.
2. Задача.

Импортировать файл Loans.txt.

С помощью обработки «Нейронная сеть» построить модель, которая могла бы предсказывать наличие в будущем у заемщика просрочек по обслуживанию долга.

1. Оценить качество модели.
2. К какому классу будут отнесен заемщик

Мужчина, 45-х лет, состоит в браке, 2 иждивенца, доход – 50000 руб., опыт работы – 20 лет, срок проживания на одном месте – 10 лет, недвижимость – 0, месячный платеж 8000 руб.

Заведующий кафедрой

/Казарян М.А./

При оценке знаний студента преподаватель должен руководствоваться приведенной в рабочей программе шкалой оценивания результатов, а также следующими критериями:

- правильность ответа на вопросы билета;
- полнота ответа, знание терминологии;
- логика и аргументированность ответа;
- способность делать выводы.

Шкала оценивания результатов промежуточной аттестации (зачет с оценкой)

Обозначения		Формулировка требований к степени освоения дисциплины
Цифр.	Оценка	
2	Неудовлетворительно Не зачтено	Оценка «неудовлетворительно» выставляется студенту, который не знает значительной части программного материала, допускает существенные ошибки, неуверенно, с большими затруднениями выполняет практические работы. Как правило, оценка «неудовлетворительно» ставится студентам, которые не могут продолжить обучение без дополнительных занятий по соответствующей дисциплине. Учебные достижения в семестровый период и результаты рубежного контроля демонстрировали низкую степень овладения программным материалом.
3	Удовлетворительно зачтено	Оценка «удовлетворительно» выставляется студенту, если он имеет знания только основного материала, но не усвоил его деталей, допускает неточности, недостаточно правильные формулировки, нарушения логической последовательности в изложении программного материала, испытывает затруднения при выполнении практических работ. Учебные достижения в семестровый период и результаты рубежного контроля демонстрируют достаточную (удовлетворительную) степень овладения программным материалом.
4	Хорошо зачтено	Оценка «хорошо» выставляется студенту, если он твердо знает материал, грамотно и по существу излагает его, не допуская существенных неточностей в ответе на вопрос, правильно применяет теоретические положения при решении практических вопросов и задач, владеет необходимыми навыками и приемами их выполнения. Учебные достижения в семестровый период и результаты рубежного контроля демонстрируют хорошую степень овладения программным материалом.
5	Отлично	Оценка «отлично» выставляется студенту, если он глубоко

	зачтено	ко и прочно усвоил программный материал, исчерпывающе, последовательно, четко и логически стройно излагает его на зачете, умеет тесно увязывать теорию с практикой, свободно справляется с задачами, вопросами и другими видами применения знаний, причем не затрудняется с ответом при видоизменении заданий, использует в ответе материал монографической литературы, правильно обосновывает принятое решение. Учебные достижения в семестровый период и результаты рубежного контроля демонстрируют высокую степень овладения программным материалом.
--	---------	--

Перечень вопросов для проведения промежуточной аттестации (теоретическая часть):

1. Аналитический и информационный подход к моделированию.
2. Основные этапы построения моделей.
3. Формы представления, типы и виды данных.
4. Классификация задач Data Mining.
5. Виды источников данных.
6. Хранилища данных.
7. Многомерная модель данных.
8. Оценка качества данных.
9. Предобработка данных.
10. Трансформация данных.
11. Цели и задачи визуализации данных в аналитических технологиях.
12. Виды визуализаторов данных.
13. OLAP-анализ.
14. Задача поиска ассоциаций.
15. Значимость и полезность ассоциативных правил.
16. Алгоритм apriori.
17. Задача кластеризации.
18. Алгоритмы решения задачи кластеризации.
19. Алгоритм k-means.
20. Сети и карты Кохонена.
21. Задача классификации.
22. Статистические методы решения задачи классификации.
23. Деревья решений.
24. Нейронные сети.
25. Эффект переобучения.
26. Оценка качества классификационных моделей.

27. Модели временных рядов.
28. Методы сглаживания временных рядов.
29. Оценка адекватности выбранного метода прогнозирования.
30. Ансамбли моделей. Бэггинг.
31. Ансамбли моделей. Бустинг.

Примеры задач для решения в компьютерной программе Deductor Academic для промежуточной аттестации (практическая часть):

Вариант 1.

На основе данных хранилища данных «Банк» получить следующие отчеты и вынести их на Панель Отчетов.

1. Количество одобренных/отклоненных заявок по каждому филиалу и отделению на основе всех имеющихся данных.
2. 10 «наиболее злостных» неплательщиков по состоянию на 10.11.2015
3. Среднюю интенсивность поступления заявок по дням недели каждого из отделений банка на основании всех имеющихся данных. Отчет дополните соответствующим графиком.

Вариант 2.

Импортировать текстовый файл Loans_demo.txt.

С помощью обработки «Логистическая регрессия» построить модель, которая могла бы предсказывать наличие в будущем у заемщика просрочек по обслуживанию долга.

3. Оценить качество модели.
4. Чему равен оптимальный порог отсечения?
5. Как изменяется кредитный рейтинг заемщика с каждым дополнительным иждивенцем?
6. К какому классу будут отнесен заемщик
Мужчина, 43-х лет, состоит в браке, 3 иждивенца, 20500 доход, опыт работы 20, срок проживания 10, недвижимость 0, месячный платеж 6725.

Вариант 3.

Импортировать текстовый файл responses.txt.

С помощью обработок «Дерево решений» и «Нейронная» сеть построить модели, позволяющие предсказать отклик клиента на рассылку.

1. Сравнить качество моделей.
2. Какую модель для отбора респондентов из рассмотренных вы выберете, если известно, что бюджет рассылки составляет 95 000 руб.?
3. Скольким клиентам пришлось бы отправить письма, если бы мы проводили рассылку случайно? Какую экономию дает наша модель?

Вариант 4.

Импортировать текстовый файл mobile.txt.

1. Провести сегментацию абонентов оператора мобильной связи.
2. Построить профили абонентов путем выявления их схожего поведения в плане частоты, длительности и времени звонков, а также ежемесячных расходов;
3. Провести оценку наиболее и наименее доходных сегментов.

Вариант 5.

Исходные данные представлены в файле Чеки.txt.

Проанализировать данные и решить следующие задачи:

1. Предсказать то, какие товары покупатели могут выбрать в зависимости от того, что уже есть в их корзинах;
2. Выявить наиболее популярные товарные наборы, состоящие из более чем 1 предмета;
3. Предложить рекламные акции типа «каждому купившему А и В товар С в подарок».

Вариант 6.

На основе данных хранилища данных Materials

1. Построить модель, позволяющую, автоматизировать ежемесячный расчет потоварного прогноза на следующие три периода. Типовой сценарий построить на примере товарной группы «Грунтовка».
2. Оценить качество предложенной модели.

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

4.4.1. Показатели и критерии оценивания компетенций, используемые шкалы оценивания

Код этапа компетенции	Показатель оценивания	Критерий оценивания	Средства оценивания	Шкала оценивания
ПК-10.2.1	Знания: места интеллектуального анализа данных в системах поддержки принятия решений; методов, моделей и средств когнитивного анализа, используемых для поддержки бизнес - решений в процессе управления ком-	Правильность и полнота ответов, глубина понимания вопроса	<u>Текущий контроль</u> выполнение устных и письменных заданий <u>Промежуточная аттестация</u>	Шкала 1

	панией; структуры и принципов работы современных аналитических платформ; моделей, методов и алгоритмов интеллектуального анализа данных; целей и задач визуализации данных на разных этапах аналитического процесса, визуализаторы общего назначения, OLAP-анализ, визуализаторы для оценки качества моделей, визуализаторы, применяемые для интерпретации результатов анализа; способов оценки качества данных и методов подготовки данных к анализу.		Зачет с оценкой	
	Умения: оценить необходимость и возможность применения интеллектуального анализа данных при решении конкретных бизнес-задач; формулировать задачи интеллектуального бизнес-анализа экономического объекта, осуществлять выбор методов анализа, подготовку исходных данных, визуализацию и интерпретацию полученных результатов; решать задачи ассоциации, кластеризации, классификации, регрессии, задачи анализа и прогнозирования временных рядов; строить аналитическую отчетность.	Правильность выполнения учебных заданий, аргументированность выводов	<u>Текущий контроль</u> выполнение устных и письменных заданий <u>Промежуточная аттестация</u> Зачет с оценкой	Шкала 1

	Навыки: решения задач ассоциации, кластеризации, классификации, регрессии, задачи анализа и прогнозирования временных рядов; построения аналитической отчетности; эффективной аналитической работы в среде аналитической платформы Deductor Academic.	Обоснованность и аргументированность выполнения учебной деятельности	<u>Текущий контроль</u> выполнение практических заданий <u>Промежуточная аттестация</u> Зачет с оценкой	Шкала 2
--	--	--	--	---------

4.4.2. Шкалы оценивания результатов обучения

Уровень знаний определяется оценками «отлично», «хорошо», «удовлетворительно», «неудовлетворительно».

Шкала 1. Оценка сформированности отдельных элементов компетенций

Обозначения		Формулировка требований к степени сформированности компетенции		
Цифр.	Оценка	Знания	Умения	Навыки
2	Неудовлетворительно Не зачтено	Отсутствие знаний	Отсутствие умений	Отсутствие навыков
3	Удовлетворительно зачтено	Фрагментарные, не структурированные знания	Частично освоенное, не систематически осуществляемое умение	Фрагментарное, не систематическое применение
4	Хорошо зачтено	Сформированные, но содержащие отдельные пробелы знания	В целом успешное, но содержащие отдельные пробелы умение	В целом успешное, но содержащее отдельные пробелы применение навыков
5	Отлично зачтено	Сформированные систематические знания	Сформированное умение	Успешное и систематическое применение навыков

Шкала 2. Комплексная оценка сформированности знаний, умений и навыков

Обозначения		Формулировка требований к степени сформированности компетенции
Цифр.	Оценка	

2	Неудовлетворительно Не зачтено	Студент не имеет необходимых представлений о проверяемом материале.
3	Удовлетворительно зачтено	Знания не структурированы, на уровне ориентирования , общих представлений. Студент допускает неточности, приводит недостаточно правильные формулировки, нарушает логическую последовательность в изложении ответа на вопросы или в демонстрируемом действии.
4	Хорошо зачтено	Знания, умения, навыки на аналитическом уровне. Компетенции в целом сформированные, но содержащие отдельные пробелы. Студент твердо знает материал, грамотно и по существу его излагает, однако допускает несущественные погрешности при ответе на заданный вопрос или в демонстрируемом действии.
5	Отлично зачтено	Знания, умения, навыки на системном уровне. Студент глубоко и прочно усвоил программный материал, исчерпывающе, последовательно и четко его излагает, умеет тесно увязывать теорию с практикой, свободно справляется с задачами, вопросами и другими видами применения знаний, в том числе при видоизменении и решении нестандартных практических задач, правильно обосновывает принятое решение.

5. МЕТОДИЧЕСКИЕ УКАЗАНИЯ ДЛЯ ОБУЧАЮЩИХСЯ ПО ОСВОЕНИЮ ДИСЦИПЛИНЫ

Освоение дисциплины «Информационные технологии в прогнозно-аналитической деятельности» предполагает, как аудиторную, так и самостоятельную работу студентов.

Аудиторная работа проводится в форме практических занятий и лабораторных работ. Подготовка к занятиям должна носить систематический характер. Это позволит обучающемуся в полном объеме выполнить все требования преподавателя.

Самостоятельная работа является обязательным компонентом процесса подготовки магистров, она формирует самостоятельность, познавательную активность, вырабатывает практические навыки работы с научной литературой.

Общий объем аудиторной и самостоятельной работы определяется учебно-тематическим планом. Изучение дисциплины завершается зачетом с оценкой.

Успешное изучение дисциплины требует посещения и активной работы на практических занятиях и лабораторных работах, выполнения учебных заданий преподавателя, ознакомления с основной и дополнительной литературой.

5.2. Методические указания для обучающихся по подготовке к практическим и лабораторным занятиям.

Практические занятия служат для закрепления изученного материала, развития умений и навыков работы в компьютерном классе в среде аналитической платформы Deductor Academic, а также для контроля преподавателем степени подготовленности студентов по изучаемой дисциплине.

При подготовке к практическому занятию, обучающемуся необходимо:

- до очередного практического занятия по рекомендованным литературным источникам проработать теоретический материал, соответствующий теме занятия;
- в начале занятия задать преподавателю вопросы по материалу, вызвавшему затруднения в его понимании и освоении при решении задач, заданных для самостоятельного решения;
- в ходе занятия давать конкретные, четкие ответы по существу вопросов;
- на практическом занятии доводить каждую задачу до окончательного решения, демонстрировать понимание проведенных расчетов (анализов, ситуаций), в случае затруднений обращаться к преподавателю.

Студентам, пропустившим практическое занятие, либо не подготовившимся к нему, рекомендуется не позже чем в 2-недельный срок явиться на консультацию к преподавателю и отчитаться по теме занятия.

5.3. Методические указания по организации самостоятельной работы обучающихся

Основной целью самостоятельной работы студентов является улучшение профессиональной подготовки обучающихся, направленное на формирование действенной системы фундаментальных и профессиональных знаний, умений и навыков, которые они могли бы свободно и самостоятельно применять в практической деятельности.

Самостоятельная работа студентов по дисциплине «Информационные технологии в прогнозно-аналитической деятельности» способствует более глубокому усвоению изучаемого курса и проводится в следующих видах:

- подготовка к занятиям в соответствии с заданиями на самостоятельную работу с использованием материалов практических занятий и приведенных ниже источников литературы;
- выполнение домашних заданий;
- подготовка к промежуточной аттестации.

Приступая к изучению той или иной темы, выделяемой по предметно-систематизированному принципу, необходимо по отдельности и последовательно рассмотреть каждую из частей, из которых состоит тема. При изучении курса, обу-

чающиеся должны уметь пользоваться научной литературой для самостоятельной подготовки к занятиям.

5.4. Методические указания по подготовке обучающихся к зачету с оценкой

Подготовка к зачету с оценкой осуществляется студентом самостоятельно с использованием перечня вопросов к зачету с оценкой, конспекта лекций по дисциплине и рекомендованных литературных источников.

В период подготовки к зачету с оценкой студенты вновь обращаются к пройденному учебному материалу. При этом они не только укрепляют полученные знания, но и получают новые.

Подготовка студента к зачету с оценкой включает в себя два этапа:

- самостоятельная работа в течение семестра;
- непосредственная подготовка в дни, предшествующие зачету с оценкой по темам курса.

Основным источником подготовки к зачету с оценкой является конспект лекций, где учебный материал дается в систематизированном виде, ключевые его положения детализируются, подкрепляются современными фактами и информацией, которые в силу новизны не вошли в опубликованные печатные источники.

В ходе подготовки к зачету с оценкой студентам необходимо обращать внимание не только на уровень запоминания, но и на степень понимания излагаемых проблем.

6. УЧЕБНАЯ ЛИТЕРАТУРА И РЕСУРСЫ ИНФОРМАЦИОННО-ТЕЛЕКОММУНИКАЦИОННОЙ СЕТИ «ИНТЕРНЕТ», ВКЛЮЧАЯ ПЕРЕЧЕНЬ УЧЕБНО-МЕТОДИЧЕСКОГО ОБЕСПЕЧЕНИЯ ДЛЯ САМОСТОЯТЕЛЬНОЙ РАБОТЫ ОБУЧАЮЩИХСЯ ПО ДИСЦИПЛИНЕ (МОДУЛЮ)

6.1. Основная литература

1. Анализ данных: учебник для академического бакалавриата / В. С. Мхитарян [и др.]; под ред. В. С. Мхитаряна. — М.: Издательство Юрайт, 2017. — 490 с. — (Серия: Бакалавр. Академический курс). — ISBN 978-5-534-00616-2. — Режим доступа: www.biblio-online.ru/book/CC38E97A-CCE5-4470-90F1-3B6D35ACC0B4
2. Миркин, Б. Г. Введение в анализ данных: учебник и практикум / Б. Г. Миркин. — М.: Издательство Юрайт, 2017. — 174 с. — (Серия: Авторский учебник). — ISBN 978-5-534-03762-3. — Режим доступа: www.biblio-online.ru/book/46A41F93-BC46-401C-A30E-27C0FB60B9DE.

3. Подкорытова, О. А. Анализ временных рядов: учебное пособие для бакалавриата и магистратуры / О. А. Подкорытова, М. В. Соколов. — 2-е изд., перераб. и доп. — М.: Издательство Юрайт, 2017. — 267 с. — (Серия: Бакалавр и магистр. Модуль.). — ISBN 978-5-534-02556-9. — Режим доступа: www.biblio-online.ru/book/7132122F-D176-4118-AD03-D43A9FA2FF86

6.2. Дополнительная литература

1. Паклин Н.Б., Орешков В.И Бизнес-аналитика: от данных к знаниям. — СПб.: Питер, 2013. — 704 с.

6.3. Учебно-методическое обеспечение самостоятельной работы

1. Национальный Открытый Университет «ИНТУИТ». Курс «Data Mining» [Электронный ресурс]. — Режим доступа: <http://www.intuit.ru/studies/curriculums/6/courses/6/info>

6.4. Нормативные правовые документы

Не предусмотрено

6.5. Интернет-ресурсы

1. www.biblio-online.ru — Электронно-библиотечная система [ЭБС] Юрайт;
2. <http://www.iprbookshop.ru> — Электронно-библиотечная система [ЭБС] «Iprbooks»
3. <https://e.lanbook.com> — Электронно-библиотечная система [ЭБС] «Лань».
4. <https://dlib.eastview.com> — East View Information Services, Inc. Коллекции электронных научных и практических журналов.
5. <http://elibrary.ru/> — Научная электронная библиотека Elibrary.ru.

6.6. Иные источники

Статистические данные из различных открытых источников для проведения анализа.

7. МАТЕРИАЛЬНО-ТЕХНИЧЕСКАЯ БАЗА, ИНФОРМАЦИОННЫЕ ТЕХНОЛОГИИ, ПРОГРАММНОЕ ОБЕСПЕЧЕНИЕ И ИНФОРМАЦИОННЫЕ СПРАВОЧНЫЕ СИСТЕМЫ

Аудитории с мультимедийным оборудованием для визуализации материала во время проведения лекционных и практических занятий.

Компьютерные классы, оборудованные персональными компьютерами, объединенными в локальную сеть с выходом в Интернет.

Пакет приложений Microsoft Office.

Аналитическая платформа Deductor Academic.