

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

Факультет Высшая школа финансов и менеджмента

Кафедра финансового менеджмента, управленческого учета и международных
стандартов финансовой деятельности

УТВЕРЖДЕНА

решением кафедры

Протокол №6 от «07» сентября 2017 г.

РАБОЧАЯ ПРОГРАММА ДИСЦИПЛИНЫ

**Б1.В.ДВ.03.01 «Информационные технологии в прогнозно-аналитической дея-
тельности»**

направление подготовки
38.03.02 Менеджмент

направленность (профиль)
Финансовая математика

квалификация: бакалавр

формы обучения: очная

Год набора – 2017

Москва, 2017 г.

Автор(ы)–составитель(и):

Преподаватель кафедры финансового менеджмента, управленческого учета и международных стандартов финансовой деятельности Е.А. Баркова

Заведующий кафедрой финансового менеджмента, управленческого учета и международных стандартов финансовой деятельности д.э.н., профессор Е.Н. Лобанова

СОДЕРЖАНИЕ

1.Перечень планируемых результатов обучения по дисциплине, соотнесенных с

4.Материалы текущего контроля успеваемости обучающихся и фонд оценочных

5.Методические указания для обучающихся по освоению дисциплины **Ошибка!**

Закладка не определена.

6.Учебная литература и ресурсы информационно-телекоммуникационной сети Интернет», включая перечень учебно-методического обеспечения для самостоятельной работы обучающихся по дисциплине **Ошибка! Закладка не определена.**

6.1 Основная литература **Ошибка! Закладка не определена.**

6.2 Дополнительная литература **Ошибка! Закладка не определена.**

6.3 Интернет-ресурсы..... **Ошибка! Закладка не определена.**

7.Материально-техническая база, информационные технологии, программное

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

1.1. Дисциплина Б1.В.ДВ.03.01 «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
ОПК-7	Способность решать стандартные задачи профессиональной деятельности на основе информационной и библиографической культуры с применением информационно-коммуникационных технологий и с учётом основных требований информационной безопасности	ОПК-7.3	Способность решить стандартные задачи профессиональной деятельности на основе информационной культуры с применением информационно-коммуникационных технологий и с учетом основных требований информационной безопасности

1.2. В результате освоения дисциплины Б1.В.ДВ.03.01 «Информационные технологии в прогнозно-аналитической деятельности» у студентов должны быть:

ОТФ/ТФ (при наличии профстандарта)/ профессиональные действия	Код этапа освоения компетенции	Результаты обучения
	ОПК-7.3	Способен оценить необходимость и возможность применения интеллектуального анализа данных при решении конкретных бизнес-задач; Способен осуществлять выбор методов анализа, подготовку исходных данных, визуализацию и интерпретацию полученных результатов

2. ОБЪЕМ И МЕСТО ДИСЦИПЛИНЫ В СТРУКТУРЕ ОП ВО

Дисциплина Б1.В.ДВ.03.01 «Информационные технологии в прогнозно-аналитической деятельности» относится к блоку дисциплин вариативной части учебного плана направления подготовки бакалавров 38.03.02 «Менеджмент» с профилем подготовки

Финансовый менеджмент». Дисциплина Б1.В.ДВ.03.01 «Информационные технологии в прогнозно-аналитической деятельности» является дисциплиной по выбору и изучается в 5 семестре.

Общая трудоемкость дисциплины составляет 72/54 часов (2 ЗЕ)

Из них:

32 /24 часа — контактная работа с преподавателем (компьютерный практикум);

40 часов — самостоятельная работа обучающихся.

Формой промежуточной аттестации в соответствии с учебным планом является зачет.

В методологическом плане дисциплина Б1.В.ДВ.03.01 «Информационные технологии в прогнозно-аналитической деятельности» опирается на минимально необходимый объем знаний, умений и навыков, приобретенных ранее при изучении дисциплин Б1.В.06 «Информатика» и Б1.Б.08 «Информационные технологии в менеджменте».

Знания, умения и навыки, полученные в результате освоения дисциплины Б1.В.ДВ.03.01 «Информационные технологии в прогнозно-аналитической деятельности», могут быть использованы при изучении дисциплины Б1.В.ДВ.01.01 «Информационные системы в экономике».

3.СОДЕРЖАНИЕ И СТРУКТУРА ДИСЦИПЛИНЫ

Таблица 1

№ п/п	Наименование тем	Объем дисциплины (модуля), академический/ астрономический час.			Форма текущего контроля успеваемости, промежуточной аттестации*
		Всего	Контактная работа обучающихся с преподавателем по видам учебных занятий	СР	
			ПЗ		
Тема 1.	Аналитическая отчетность.	/6	/3	/3	О, ПЗ
Тема 2.	Классификация и регрессия. Статистические методы.	/9	4,5	4,5	О, ПЗ
Тема 3.	Классификация и регрессия. Машинное обучение.	9		,5	О, ПЗ

№ п/п	Наименование тем	Объем дисциплины (модуля), академический/ астрономический час.			Форма текущего контроля успеваемо сти, промежут очной аттестаци и*
		Всего	Контактная работа обучающихся с преподавателем по видам учебных занятий	СР	
			ПЗ		
Тема 4.	Кластеризация.	7,5		,5	О, ПЗ
Тема 5.	Поиск ассоциативных правил.	7,5	/3	,5	О, ПЗ
Тема 6.	Анализ и прогнозирование временных рядов.	7,5	/3	,5	О, ПЗ
Тема 7.	Ансамбли моделей.	7,5	/3	,5	О, КР
Промежуточная аттестация					Зачет
Всего:		54	/24	/30	

Примечание:

** – формы текущего контроля успеваемости: опрос (О), результаты выполнения практического задания на компьютере (ПЗ), контрольная работа (КР), тестирование (Т).*

3.1 Содержание дисциплины (модуля)

№ п/п	Наименование тем (разделов)	Содержание тем (разделов)
Тема 1	Аналитическая отчетность.	Введение в визуализацию. Визуализаторы общего назначения. OLAP-анализ. Визуализация данных в аналитической платформе Deductor. Компьютерный практикум №1. «Разработка системы аналитической отчетности».
Тема 2	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

		<p>модели.</p> <p>Основы логистической регрессии. Множественная логистическая регрессия.</p> <p>Оценка эффективности и сравнение моделей. ROC-анализ. Оценка ошибки модели. Издержки ошибочной классификации. Lift- и Profit-кривые.</p> <p>Компьютерный практикум №2. «Скоринговые модели для оценки кредитоспособности заемщиков».</p>
Тема 3	Классификация и регрессия. Машинное обучение	<p>Введение в деревья решений. Алгоритмы построения деревьев решений. Упрощение деревьев решений.</p> <p>Введение в нейронные сети. Принципы построения нейронных сетей. Процесс обучения нейронной сети. Обучение в условиях несбалансированности классов.</p> <p>Компьютерный практикум №3. «Повышение эффективности массовой рассылки».</p>
Тема 4	Кластеризация	<p>Введение в кластеризацию. Меры близости, используемые в алгоритмах кластеризации. Методы кластерного анализа.</p> <p>Иерархическая кластеризация. Алгоритм кластеризации k-means. Сети Кохонена. Карты Кохонена.</p> <p>Решение задачи кластеризации средствами аналитической платформы</p> <p>Компьютерный практикум №4. «Сегментация клиентов компании».</p>
Тема 5	Поиск ассоциативных правил	<p>Ассоциативные правила. Алгоритм Apriori. Иерархические ассоциативные правила. Последовательные шаблоны.</p> <p>Компьютерный практикум №5. «Ассоциативные правила в стимулировании розничных продаж».</p>
Тема 6	Анализ и прогнозирование временных рядов	<p>Введение в прогнозирование. Временной ряд и его компоненты.</p> <p>Модели прогнозирования. Исследование наборов данных и выбор метода прогнозирования. Ошибки прогноза. Оценка адекватности выбранного метода прогнозирования.</p> <p>Методы сглаживания и скользящие средние. Декомпозиция временного ряда. Регрессионный анализ временных рядов. Анализ временных рядов средствами аналитической платформы Deductor.</p> <p>Компьютерный практикум №6. «Прогнозирование продаж».</p>
Тема 7	Ансамбли моделей	<p>Введение в ансамбли моделей. Бэггинг. Бустинг. Альтернативные методы построения ансамблей</p>

4. МАТЕРИАЛЫ ТЕКУЩЕГО КОНТРОЛЯ УСПЕВАЕМОСТИ ОБУЧАЮЩИХСЯ И ФОНД ОЦЕНОЧНЫХ СРЕДСТВ ПРОМЕЖУТОЧНОЙ АТТЕСТАЦИИ ПО ДИСЦИПЛИНЕ

4.1. Формы и методы текущего контроля успеваемости и промежуточной аттестации

Формы текущего контроля успеваемости

В ходе реализации дисциплины «Информационные технологии в прогнозно-аналитической деятельности» используются следующие формы текущего контроля успеваемости обучающихся: практическое задание (компьютерный практикум), устный опрос, контрольная работа.

Практическое задание (компьютерный практикум).

Практические задания осуществляются в форме выполнения практико-ориентированных заданий на компьютере с использованием аналитической платформы. 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 аналитической платформы «Deductor Academic». Данный вид работы проверяет усвоение полученных в ходе обучения умений и навыков. Время выполнения контрольной работы составляет 2 часа. При выполнении контрольной работы обучающемуся запрещается пользоваться дополнительной литературой. Успешное выполнение контрольной работы является обязательным для допуска к промежуточной аттестации.

Формы текущего контроля успеваемости по темам (разделам) дисциплины:

Тема и/или раздел	Формы текущего контроля успеваемости
Тема 1. Аналитическая отчетность.	Опрос Практическое задание
Тема 3. Классификация и регрессия. Статистические методы	Опрос Практическое задание
Тема 4. Классификация и регрессия. Машинное обучение	Опрос Практическое задание
Тема 5. Кластеризация	Опрос Практическое задание
Тема 6. Поиск ассоциативных правил	Опрос Практическое задание
Тема 7. Анализ и прогнозирование временных рядов	Опрос Практическое задание
Тема 8. Ансамбли моделей	Контрольная работа

Форма промежуточной аттестации

Промежуточная аттестация по дисциплине «Информационные технологии в прогнозно-аналитической деятельности» проводится в форме зачета.

Зачет проводится методом устного опроса студента и решения практической задачи с использованием аналитической платформы «Deductor Academic».

4.2 Материалы текущего контроля успеваемости обучающихся

Типовые оценочные материалы по теме 1 «Аналитическая отчетность»:

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?

Практическое 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 последние полгода.

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.

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 2 «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?

Практическое 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%) — таким 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.) 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

$$P = \frac{1}{1 + C_{FN}/C_{FP}}$$

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 3 «Классификация и регрессия. Машинное обучение»:

Опрос

1. Какие бизнес-задачи можно отнести к задачам классификации?
2. Какие алгоритмы используются для решения задач классификации?
3. Деревья решений
4. Чем характеризуется явление переобучения дерева решений?
5. Что такое нейронная сеть?
6. Нейронная сеть как метод решения задачи классификации.
7. Что такое ошибка обучения нейронной сети?
8. Чем характеризуется явление переобучения нейронной сети?

Практическое 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
	Срок клиента	Количество месяцев, прошедших с момента подключения	Целый
	Возраст	Возраст клиента	Целый
	Адрес	Срок проживания по последнему адресу	Целый
	Стаж	Стаж работы общий	Целый
	Семейное положение	Семейное положение клиента	Строковый: – не в браке – в браке
	Образование	Уровень образования	Строковый: – начальное – среднее проф. – среднее спец. – высшее – уч. степень
	Пенсионер	–	да/нет
	Пол	–	муж/жен
	Тариф	Текущий тариф клиента	Строковый: – «Базовый» – «Плюс» – «Эконом» – «Анлим»
	Уход	Отказ от пользования услугами	

		(соответствует значению 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 RPROP;
 - 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 4 «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?

Практическое 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 E	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 5 «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?

Практическое 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 A 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 6 «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.

Практическое 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 потоварного 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 7 «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)	
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

4.3.1. Перечень компетенций с указанием этапов их формирования в процессе освоения образовательной программы. Показатели и критерии оценивания компетенций с учетом этапа их формирования

Код компетенции	Наименование компетенции	Код этапа освоения компетенции	Наименование этапа освоения компетенции
-----------------	--------------------------	--------------------------------	-----------------------------------------

ОПК-7	Способность решать стандартные задачи профессиональной деятельности на основе информационной и библиографической культуры с применением информационно-коммуникационных технологий и с учётом основных требований информационной безопасности	ОПК-7.3	Способность решить стандартные задачи профессиональной деятельности на основе информационной культуры с применением информационно-коммуникационных технологий и с учетом основных требований информационной безопасности
-------	----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------	---------	--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------

Этап освоения компетенции	Показатель оценивания	Критерий оценивания
ОПК-7.3 Способность решить стандартные задачи профессиональной деятельности на основе информационной культуры с применением информационно-коммуникационных технологий и с учетом основных требований информационной безопасности	Способен оценить необходимость и возможность применения интеллектуального анализа данных при решении конкретных бизнес-задач; Способен осуществлять выбор методов анализа, подготовку исходных данных, визуализацию и интерпретацию полученных результатов; Способен решать задачи ассоциации, кластеризации, классификации, регрессии, задачи анализа и прогнозирования с помощью временных рядов, строить аналитическую отчетность.	Проведена оценка возможности применения интеллектуального анализа данных при решении конкретных бизнес-задач; Сделан выбор методов анализа, подготовлены исходные данные, проведена интерпретация полученных результатов; Решена задача анализа и прогнозирования с помощью временных рядов, подготовлена аналитическая отчетность.

Типовые оценочные средства

Вопросы для подготовки к зачету

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 практических заданий для выполнения на зачете

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 10.11.2009
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.

1. Оценить качество модели.
2. Чему равен оптимальный порог отсечения?
3. Как изменяется кредитный рейтинг заемщика с каждым дополнительным иждивенцем?

4. 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 так же ежемесячных расходов;
3. Провести оценку наиболее и наименее доходных сегментов.

Вариант 5

Исходные данные представлены в файле Чеки.

Стоят 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 товарного прогноза на следующие три периода. Типовой сценарий построить на примере товарной группы «Грунтовка».
2. Оценить качество предложенной модели.

Шкала оценивания

Критерий оценивания	Шкала оценивания
---------------------	------------------

Проведена оценка возможности применения интеллектуального анализа данных при решении конкретных бизнес-задач; Сделан выбор методов анализа, подготовлены исходные данные, проведена интерпретация полученных результатов; Решена задача анализа и прогнозирования с помощью временных рядов, подготовлена аналитическая отчетность.	Зачтено (60-100 баллов)
Не проведена оценка возможности применения интеллектуального анализа данных при решении конкретных бизнес-задач; Не сделан выбор методов анализа, не подготовлены исходные данные, не проведена интерпретация полученных результатов; Не решена задача анализа и прогнозирования с помощью временных рядов, не подготовлена аналитическая отчетность.	Не зачтено (0-59 баллов)

4.4. Методические материалы

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

Зачет 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 осуществляется в форме устного ответа и решения практической задачи на компьютере по билетам. Задание билета включает в себя один теоретический вопрос и одно практическое задание для выполнения на компьютере с использованием аналитической платформы «Deductor Academic».

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

РОССИЙСКАЯ АКАДЕМИЯ НАРОДНОГО ХОЗЯЙСТВА И
ГОСУДАРСТВЕННОЙ СЛУЖБЫ ПРИ ПРЕЗИДЕНТЕ РОССИЙСКОЙ ФЕДЕРАЦИИ

Кафедра финансового менеджмента, управленческого учета и международных
стандартов финансовой деятельности
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.

5. Оценить качество модели.

6. 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

д.э.н., проф. Е.Н. Лобанова

Процедура проведения зачета предусматривает

- получение студентом билета, включающего один вопрос и одно практическое задание для решения на компьютере с использованием аналитической платформы
- 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 ответ на дополнительные вопросы.

Аттестация производится в соответствии со шкалой оценивания результатов обучения .

5.МЕТОДИЧЕСКИЕ УКАЗАНИЯ ДЛЯ ОБУЧАЮЩИХСЯ ПО ОСВОЕНИЮ ДИСЦИПЛИНЫ (МОДУЛЯ)

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

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

Самостоятельная работа является обязательным компонентом процесса подготовки бакалавров, 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.

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

. Методические указания для обучающихся по подготовке к практическим занятиям

Практические занятия служат для закрепления изученного материала, развития умений и навыков работы в компьютерном классе в среде программного комплекса «Информационные технологии в прогнозно-аналитической деятельности», 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 действий, 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.

. 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 компьютерного практикума с использованием материалов практических занятий и приведенных ниже источников литературы;
- 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, обучающиеся должны уметь пользоваться научной литературой для самостоятельной подготовки к занятиям.

. Методические указания по подготовке обучающихся к зачету

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

В период подготовки к зачету 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

6.1 Основная литература

1. Миркин Б.Г. Введение в анализ данных. — [М.: Издательство Юрайт](#), 2016. — 174 с.
2. Анализ данных/ В. С. Мхитарян [и др.]; под ред. В. С. Мхитаряна. — М.: Издательство Юрайт, 2017. — 490 с.
3. Подкорытова О.А., Соколов М.В. Анализ временных рядов. — [М.: Издательство Юрайт](#), 2016. — 266 с.

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>

Нормативные правовые документы

Не предусмотрено

6.5. Интернет-ресурсы

w

— Электронно-библиотечная система [ЭБС] «Iprbooks»

— Электронно-библиотечная система [ЭБС] «Лань».

4. <https://dlib.eastview.com> — East View Information Services, Inc. Коллекции электронных научных и практических журналов.
— Научная электронная библиотека Elibrary.ru.

l

6.6. Иные источники

i

Статистические данные из различных открытых источников для проведения анализа.

o

o

n

l

i

n

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

Практические занятия проводятся в компьютерном классе, оснащенном персональными компьютерами, объединенными в сеть с выходом в Интернет, а также оборудованием для показа презентаций и слайдов (компьютер, проектор, экран).

Используется следующее программное обеспечение:

Количество	2607
Правообладатель	Microsoft Corporation
Дата покупки / продления	06.12.2016
К	
о	Продавец ООО «ЛАНИТ-Интеграция»
н	Покупатель РАНХиГС
т	Дата окончания 31.12.2017
р	Срок подписки 1 год / 3 года
а	
к	Microsoft Office Professional 2016
т	Количество 2607
	Правообладатель Microsoft Corporation
	Дата покупки / продления 06.12.2016
	К
о	Продавец ООО «ЛАНИТ-Интеграция»
н	Покупатель РАНХиГС
т	Дата окончания 31.12.2017
р	Срок подписки 1 год / 3 года
а	
к	Acrobat Professional AcademicEdition License Russian
т	Multiple Platforms (Adobe, 65258631AE01A00)
	Количество 50
	Правообладатель Adobe
	Дата покупки / продления 03.04.2017
	Контракт #15/08-17

Продавец SoftLine

Покупатель РАНХиГС

Дата окончания 03.04.2018

Для выполнения практических заданий по дисциплине используется свободно распространяемое программное обеспечение Deductor Academic.