

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

Институт экономики, математики и информационных технологий

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

Кафедра национальной экономики

(наименование кафедры)

УТВЕРЖДЕНА

решением Ученого Совета

Института ЭМИТ РАНХиГС

Протокол от «09» сентября 2020 г.

№ 1-20/21

РАБОЧАЯ ПРОГРАММА ДИСЦИПЛИНЫ

**Б1.В.ДВ.01.02 АНАЛИЗ ДАННЫХ В ПРОГНОЗНО-
АНАЛИТИЧЕСКОЙ ДЕЯТЕЛЬНОСТИ**

(индекс, наименование дисциплины (модуля), в соответствии с учебным планом)

направление подготовки (специальность)

38.04.01 Экономика

(код, наименование направления подготовки (специальности))

«Экономика медиаиндустрии»

(направленность(и) (профиль (и)/специализация(ии))

магистр

(квалификация)

очная

(форма(ы) обучения)

Год набора – 2021 г.

Москва, 2020

Автор(ы)–составитель(и):

Баркова Е.А.

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

Кафедра национальной экономики

(наименование кафедры)

Заведующий кафедрой

Национальной экономики

(наименование кафедры)

д.э.н., доцент Казарян М.А.

(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 образовательной программы.....	5
3. Содержание и структура дисциплины.....	5
4. Материалы текущего контроля успеваемости обучающихся и фонд оценочных средств промежуточной аттестации по дисциплине.....	8
5. Методические указания для обучающихся по освоению дисциплины.....	36
6. Учебная литература и ресурсы информационно-телекоммуникационной сети «Интернет», включая учебно-методическое обеспечение самостоятельной работы обучающихся по дисциплине.....	39
6.1. Основная литература.....	39
6.2. Дополнительная литература.....	39
6.3. Учебно-методическое обеспечение самостоятельной работы.....	40
6.4. Нормативные правовые документы.....	40
6.5. Интернет-ресурсы.....	40
7. Материально-техническая база, информационные технологии, программное обеспечение и информационные справочные системы	41

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

1.1. Дисциплина Б1.В.ДВ.01.02 «Анализ данных в прогнозно-аналитической деятельности» обеспечивает овладение следующими компетенциями с учетом этапа:

Код компетенции	Наименование компетенции	Код этапа освоения компетенции	Наименование этапа освоения компетенции
ПКс ОС III-1	Способен анализировать и использовать различные источники информации для проведения экономических расчетов, в том числе, с использованием информационных технологий	ПКс ОС III-1.1	Способен анализировать различные источники информации для проведения экономических расчетов

1.2. В результате освоения дисциплины у студентов должны быть сформированы:

Код этапа освоения компетенции	Результаты обучения
ПКс ОС III-1.1	<p>Знания:</p> <p>знание структуры и принципов работы современных аналитических платформ;</p> <p>знание моделей, методов и алгоритмов анализа данных;</p> <p>знание способов визуализации данных на разных этапах аналитического процесса;</p> <p>знание методов оценки качества данных и подготовки данных к анализу.</p> <p>Умения:</p> <p>умение осуществлять выбор методов анализа;</p> <p>умение оценить качество данных и подготовить исходные данные для анализа;</p> <p>умение визуализировать данные и интерпретировать полученные результаты;</p> <p>умение решать задачи кластеризации, классификации, регрессии, задачи анализа и прогнозирования временных рядов;</p> <p>умение строить аналитическую отчетность.</p> <p>Навыки:</p> <p>навыки решения задач кластеризации, классификации, регрессии, задачи анализа и прогнозирования временных рядов; навыки построения аналитической отчетности; навыки эффективной аналитической работы в среде компьютерных программ Loginom Community и Power BI</p>

2. ОБЪЕМ И МЕСТО ДИСЦИПЛИНЫ В СТРУКТУРЕ ОП ВО

Объем дисциплины

Общая трудоемкость Б1.В.ДВ.01.02 «Анализ данных в прогнозно-аналитической деятельности» составляет 3 зачётных единицы – 108 ак. ч.

Количество академических часов, выделенных на контактную работу с преподавателем для очной формы обучения, составляет 24 ак.ч.: лабораторные занятия – 16 ак.ч., практические занятия – 16 ак. ч. Самостоятельная работа составляет 76 ак. ч.; для очно-заочной формы обучения: лабораторные занятия – 8 ак.ч., практические занятия – 8 ак. ч. Самостоятельная работа составляет 92 ак. ч.

Форма промежуточной аттестации в соответствии с учебным планом – зачет с оценкой.

Место дисциплины в структуре ОП ВО

Дисциплина Б1.В.ДВ.01.02 «Анализ данных в прогнозно-аналитической деятельности» относится к блоку дисциплин вариативной части учебного плана направления подготовки магистров 38.04.01 «Экономика» с направленностью подготовки «Корпоративная экономика» и изучается во 2 семестре на 1 курсе – очная и в 3 семестре на 2 курсе очно-заочная форма обучения.

Дисциплина реализуется после изучения дисциплины Б1.В.01 «Социально-экономическая статистика».

Достижение планируемых результатов обучения служит основой для Б1.В.08 «Корпоративные финансы (продвинутый уровень)», Б1.В.09 «Управление проектами», Б1.В.10 «Инвестиционное планирование».

3. СОДЕРЖАНИЕ И СТРУКТУРА ДИСЦИПЛИНЫ

Очная форма обучения

№ п/п	Наименование тем (разделов),	Объем дисциплины, ак. час.					СР	Форма текущего контроля успеваемости, промежуточно й аттестации*
		Всего	Контактная работа обучающихся с преподавателем по видам учебных занятий					
			Л	ЛР	ПЗ	КСР		
Тема 1	Подготовка данных к анализу	17			2		15	О, ПОЗ
Тема 2	Визуализация данных	23		4	4		15	ПОЗ, ПКР
Тема 3	Задачи классификации и регрессии	22		4	2		16	О, Т, ПОЗ
Тема 4	Задача кластеризации	23		4	4		15	О, ПОЗ
Тема 5	Анализ и прогнозирование временных рядов	23		4	4		15	О, ПОЗ
Промежуточная аттестация								ЗаО
Всего:		108		16	16		76	

* формы текущего контроля успеваемости: опрос (О), тестирование (Т), онлайн тест (ОТ),

практико-ориентированные задания (ПОЗ), практическая контрольная работа (ПКР).

*** форма промежуточной аттестации: зачет с оценкой (ЗаО)*

Содержание дисциплины

Тема 1. Подготовка данных к анализу

Очистка и предобработка данных.

Оценка качества данных. Фильтрация данных. Обработка дубликатов и противоречий. Выявление аномальных значений. Восстановление пропущенных значений. Сокращение размерности.

Трансформация данных. Группировка данных. Слияние данных. Квантование. Нормализация и кодирование.

Тема 2. Визуализация данных

Введение в визуализацию. Визуализаторы общего назначения. OLAP-анализ. Визуализаторы для оценки качества моделей. Визуализаторы, применяемые для интерпретации результатов анализа.

Визуализация данных средствами MS Power BI.

Семейство MS Power BI. Конструктор кубов и аналитических представлений. Совместная работа с представлениями – серверные сервисы Power BI. Решение задач объединения данных из разных источников, вычисляемые поля и Меры, элементы языка DAX, построение различных визуальных представлений по данным из учетных систем и CRM–систем.

Тема 3. Задачи классификация и регрессии

Статистические методы

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. Оценка ошибки модели.

Машинное обучение

Введение в нейронные сети. Принципы построения нейронных сетей. Процесс обучения нейронной сети. Обучение в условиях несбалансированности классов.

Тема 4. Задача кластеризации

Введение в кластеризацию. Меры близости, используемые в алгоритмах кластеризации. Методы кластерного анализа. Иерархическая кластеризация. Алгоритм кластеризации k-means. Сети Кохонена. Карты Кохонена.

Тема 5. 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.

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

4.1. Формы и методы текущего контроля успеваемости и промежуточной аттестации

4.1.1. В ходе реализации дисциплины Б1.В.ДВ.01.02 «Анализ данных в прогнозно-аналитической деятельности» используются следующие методы текущего контроля успеваемости обучающихся:

Тема и/или раздел	Методы текущего контроля успеваемости
Подготовка данных к анализу	<i>О, ПОЗ</i>
Визуализация данных	<i>ПОЗ, ПКР</i>
Классификация и регрессия (Статистические методы)	<i>О, ПОЗ</i>
Классификация и регрессия (Машинное обучение)	<i>Т, ПОЗ, ПКР</i>
Кластеризация	<i>О, ПОЗ</i>
Анализ и прогнозирование временных рядов	<i>О, ПОЗ</i>
	<i>ЗаО</i>

4.1.2. Зачет с оценкой осуществляется в форме устного ответа и решения практической задачи на компьютере.

4.2. Материалы текущего контроля успеваемости обучающихся.

Типовые оценочные материалы по теме 1 Опрос

1. Для чего нужны очистка и предобработка данных.
2. Оценка качества данных.
3. Фильтрация данных.
4. Обработка дубликатов и противоречий.
5. Выявление аномальных значений.
6. Восстановление пропущенных значений.
7. Сокращение размерности.
8. Трансформация данных.

9. Группировка данных.
10. Слияние данных.
11. Квантование.
12. Нормализация и кодирование.

**Практико-ориентированное задание (компьютерный практикум)
«Решение задачи мониторинга корректности консолидированных данных
истории продаж сети предприятий розничной торговли»**

Для принятия управленческих решений о частоте проведения сезонных распродаж, которые должны увеличить объем выручки по итогам года, требуется анализ деятельности магазинов с возможностью прогнозирования выручки на следующий год. Для этого предоставляются данные о продажах 45 магазинов, расположенных в разных регионах - в каждом магазине есть несколько отделов. Компания также проводит несколько рекламных мероприятий по уценке в течение года, которые предшествуют знаменательным праздникам.

Табличные данные представлены в текстовых файлах в формате .csv:

1. stores.csv (магазины);
2. train.csv (продажи);
3. features.csv (признаки).

Магазины (stores.csv) — представлена анонимная информация о 45 магазинах (Store), с указанием типа (Type) и размера торговых площадей в sf (Size) магазина.

Продажи (train.csv) — представлены реальные данные о продажах в сети Walmart, которые охватывают период с 2010 по 2012 года.

Признаки (features.csv) — представлены дополнительные данные, относящиеся к деятельности магазина, отдела и региона на указанные даты.

Выявление аномалий (также обнаружение выбросов) — это опознавание редких данных, событий или наблюдений, которые вызывают подозрения ввиду существенного отличия от большей части данных. Аномалии в данных могут возникать по разным причинам от

неправильного ввода до ошибок оцифровки.

Анализ данных Stores

1. Преобразуйте категориальный признак Type в порядковый, так как в данном признаке есть скрытый порядок;
2. Преобразуйте признак Size в кв.м.;
3. Выявите аномалии (выбросы) в признаке Type и определите их в подходящую по значению к одному из типов магазинов, тем самым ослабив эффект выброса;
4. Сохраните полученные данные в файл: new _ stores . csv ;
5. Оцените качество очистки данных по каждому типу магазинов по признаку Size:

соотнесите межквартильный размах до очистки данных к межквартильному размаху после очистки.

Внешнее обогащение предполагает привлечение дополнительной информации из

внешних источников, что позволит повысить ценность и значимость данных с точки

зрения их анализа. Под повышением значимости данных подразумевается, что на

основе их анализа можно будет принимать управленческие решения принципиально нового уровня.

Анализ данных Train.

1. Выявите аномалии (выбросы) в признаке Weekly_Sales и преобразуйте их в подходящее значение, тем самым ослабив эффект выброса;

2. Сгруппируйте данные таким образом, чтобы получить новые данные следующего формата: сумма выручки по каждому магазину за определенный день;

3. Сформировать новую таблицу со следующими признаками: «Store», «Value», «Day», «Month», «Year», «IsHoliday»;

4. Сохраните полученные данные в файл: new_train.csv;

Зачастую данные, которые нам нужно использовать, имеют сложный вид; они не всегда бывают цельными, то есть разрозненный набор данных, который в процессе решения поставленной задачи требуется слить в один набор данных.

5. Соедините данные new_stores.csv и new_train.csv;

6. Сохраните полученные данные: value.csv;

7. Рассчитайте долю (%) магазинов сети Walmart каждого типа.

Внутреннее обогащение не предполагает привлечения какой-либо внешней информации.

Оно обычно связано с получением и включением в набор данных полезной информации, которая отсутствует в явном виде, но может быть тем или иным способом получена с помощью манипуляций с имеющимися данными.

Анализ данных Value.

Эффективность использования торговых площадей может оцениваться как по отделам внутри магазина, так и по магазину в целом.

1. Создайте новый признак Effic_space и рассчитайте по каждой позиции в данных эффективность использования одного квадратного метра торговой площади по каждому магазину (сколько выручки приходится на один квадратный метр торговой площади);

2. Сохраните полученные данные: new_value.csv;

3. Определите какой магазин эффективней всех использует торговую площадь, а какой неэффективно.

Шкалирование – это общепринятая задача предобработки данных в машинном обучении. Многие алгоритмы машинного обучения – градиентные методы, которые сильно чувствительны к шкалированию. Они отлично работают на признаках, которые находятся на одной шкале, как правило, от 0 до 1 или от -1 до 1.

Анализ данных Features

Распространенный метод шкалирования данных является стандартизация, при котором значение признаков должны быть приближенно стандартно распределены.

1. Соединить данные: new_value.csv и features.csv;
2. Сделайте преобразования данных (аномалии, выбросы, категориальные признаки и т.п.);
3. Удалить признаки, которые дублируют друг друга или которые информативно включают в себя другие признаки.

После преобразования должны остаться следующие признаки: «Store», «Value», «IsHoliday», «Size», «Effic _ space», «Temperature», «Fuel _ Price», «MarkDown 2», «MarkDown 3», «MarkDown 4», «MarkDown 5», «CPI», «Unemployment» ;

4. Преобразуйте признаки, чтобы они имели среднее значение 0 и стандартное отклонение 1;
5. Определите, какой признак не требуется шкалировать;
6. Определите, какой признак имеет максимальное значение после шкалирования, а какой минимальный.

Типовые оценочные материалы по теме 2

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?
7. Семейство MS Power BI. Конструктор кубов и аналитических представлений.
8. Совместная работа с представлениями – серверные сервисы Power BI.
9. Решение задач объединения данных из разных источников.
10. Вычисляемые поля и Меры.
11. Элементы языка DAX

Практико-ориентированное задание (компьютерный практикум)

«Получение аналитической отчетности в MS Power BI»

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.
- Сформировать многомерный отчет, отражающий динамику сумм продаж по месяцам года в разрезе аптек.
- Сформировать многомерный отчет и график продаж в регионах.

1. Откройте Power BI Desktop. Загрузите в Power BI данные для построения отчета. Тип данных — файлы Excel. Данные содержатся в файлах Продажи, Товары, Курс, География.

2. Объедините таблицы в единую модель данных.

3. Перейдите в представление Отчет. Используйте визуализатор «Карточка» для получения общей суммы продаж за период.

4. Постройте визуализатор «Гистограмма с группировкой», отметив поля Дата (таблица Курс) и Сумма (таблица Продажи). Получите представление по кварталам.

5. На панели Визуализации выберите Линейчатую диаграмму с группировкой. Отметьте поля Товарная группа (таблица Товары) и Сумма (таблица Продажи). Отсортируйте товарные группы по сумме выручки.

6. На панели Визуализации выберите Круговую диаграмму. Отметьте поля Область (таблица География), Отдел и Сумма (таблица Продажи).

7. На панели Визуализации выберите диаграмму Карта. Отметьте поля Город (таблица География) и Сумма (таблица Продажи). В области форматирования удалите название диаграммы. Параметр «Пузырьковая» установите — 10%. Измените цвета данных и стили карт.

8. Добавьте в отчет новую страницу. Новую страницу переименуйте в Тренд. На странице Тренд постройте визуализатор График. Отметьте поля Дата (таблица Курс) и Сумма (таблица Продажи). Добавьте Линию тренда.

9. На панели Визуализации выберите Срез: поле Товарная группа (таблица Товары).

10. Добавьте срез по городам.

Типовой вариант практической контрольной работы

1. Из нескольких открытых источников подобрать данные, которые позволят
- дать одного всестороннюю характеристику из регионов РФ;

- проанализировать динамику основных показателей социально-экономического развития региона.
- 2. Загрузить данные в MS Power BI, построить модель данных, провести необходимые преобразования данных, вычислить меры.
- 3. Подобрать средства визуализации данных. Построить отчет, который позволит проанализировать демографическое, социальное, экономическое положение региона РФ и динамику основных показателей.
- 4. Подготовить презентацию с использованием MS Sway для иллюстрации полученных результатов.

Типовые оценочные материалы по теме 3

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.

Руководство банка 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.

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 «Задачи классификации и регрессии. Машинное обучение»:

Тестирование

1. Классификация — это ...

- а. отнесение объектов к одной из заранее неизвестных групп
- б. процесс формирования групп и отнесения объектов к одному из них.
- с. отнесение объектов к одному из заранее известных классов

2. Задачи классификации решаются следующими алгоритмами:

- a. алгоритмом Apriori
 - b. нейронные сети
 - c. линейной регрессии
3. 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"
4. 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
5. 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)
6. 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
7. 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
8. Однонаправленные входные связи, соединенные с выходами других нейронов – это ...
- a. синапсы
 - b. слои сети
 - c. аксоны

9. Слой нейронной сети – это ...

- а. выходная связь данного нейрона, с которой сигнал (возбуждения или торможения) поступает на синапсы следующих нейронов
- б. один или несколько нейронов, на входы которых подается один и тот же общий сигнал
- с. однонаправленные входные связи, соединенные с выходами других нейронов

10. Ошибкой обучения нейронной сети называется ...

- а. переобучение нейронной сети
- б. разность между желаемым и полученным на выходе сигналами
- с. целевая функция, требующая минимизации в процессе управляемого обучения нейронной сети

11. 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.

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.

1. Постройте модель «Нейронная сеть».

2. Используя визуализатор Таблица сопряженности, определите Чувствительность и Специфичность полученной модели.

3. Оцените качество модели. Сравните ее с моделью, полученной на предыдущем этапе.

4. Переобучите нейронную сеть

– 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?

5. Сравните качество моделей.

6. Сформулируйте рекомендации по использованию результатов моделирования в работе компании.

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.

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.

Типовые оценочные материалы по теме 4

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 останова 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).

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

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

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 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»).

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
ПКс ОС III-1	Способен анализировать и использовать различные источники информации для проведения экономических расчетов, в том числе, с использованием информационных технологий	ПКс ОС III-1.1	Способен анализировать различные источники информации для проведения экономических расчетов

Этап освоения компетенции	Показатель оценивания	Критерий оценивания
ПКс ОС III-1.1	Анализирует различные источники информации для проведения экономических расчетов	Уровень усвоения учебного материала, глубина понимания вопроса, правильность и полнота ответов, четкость и логичность изложения его на промежуточной аттестации, аргументированность выводов, умение тесно увязывать теорию с практикой

4.3.2 Типовые оценочные средства

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

Зачет осуществляется в форме устного ответа и решения практической задачи на компьютере.

Список вопросов для подготовки к зачету с оценкой

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. Фильтрация данных
11. Обработка дубликатов и противоречий

12. Выявление аномальных значений
13. Восстановление пропущенных значений
14. Сокращение размерности
15. Трансформация данных
16. Группировка данных
17. Слияние данных
18. Квантование
19. Нормализация и кодирование
20. Цели и задачи визуализации данных в аналитических технологиях
21. Виды визуализаторов данных
22. OLAP-анализ
23. Задача кластеризации
24. Алгоритмы решения задачи кластеризации
25. Алгоритм k-means
26. Сети и карты Кохонена
27. Задача классификации
28. Статистические методы решения задачи классификации
29. Нейронные сети
30. Эффект переобучения
31. Оценка качества классификационных моделей
32. Модели временных рядов
33. Методы сглаживания временных рядов
34. Выбор метода прогнозирования
35. Оценка адекватности выбранного метода прогнозирования
36. Ансамбли моделей

Примеры задач для решения на компьютере

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 10.11.2019
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 «Логистическая регрессия» 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.

Шкала оценивания результатов промежуточной аттестации (зачет с оценкой)

Обозначения		Формулировка требований к степени освоения дисциплины
Цифр.	Оценка	
2	Неуд.	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	Удовл.	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	Хор.	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	Отл.	Оценка «отлично» выставляется студенту, если он глубоко и прочно усвоил программный материал, исчерпывающе, последовательно, четко и логически стройно излагает его на экзамене, 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.

4.4. Методические материалы

Шкалы оценивания результатов текущего контроля

Шкала оценивания результатов устных ответов на семинарских занятиях

Обозначения		Формулировка требований к степени освоения дисциплины
Цифр.	Оценка	
2	Неуд.	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	Удовл.	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	Хор.	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	Отл.	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 результатов практических заданий на практических занятиях

Обозначения		Формулировка требований к степени освоения дисциплины
Цифр.	Оценка	
2	Неуд.	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	Удовл.	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	Хор.	Студент логично и правильно решает практическую задачу, делает грамотные выводы по ее результатам, отвечает на вопрос о технологии ее решения, достаточно аргументирует свое решение, но при ответе допускает погрешности.
5	Отл.	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 результатов теста на практических занятиях

Обозначения		Формулировка требований к степени освоения дисциплины
Цифр.	Оценка	
2	Неуд.	Менее 50% правильных ответов

3	Удовл.	Более 50%, но не более 65% правильных ответов
4	Хор.	Более 65%, но не более 80% правильных ответов
5	Отл.	Более 80% правильных ответов

5. МЕТОДИЧЕСКИЕ УКАЗАНИЯ ДЛЯ ОБУЧАЮЩИХСЯ ПО ОСВОЕНИЮ ДИСЦИПЛИНЫ

Освоение дисциплины предполагает как аудиторную, так и самостоятельную работу студентов.

Аудиторная работа проводится в форме лекционных и 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 экзаменом.

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

5.1. Методические указания для обучающихся по подготовке к лекционным занятиям

Занятия лекционного вида дают систематизированные знания о наиболее сложных и актуальных проблемах изучаемой дисциплины.

Осуществляя учебные действия на лекционных занятиях, обучающиеся должны внимательно воспринимать материал, подготовленный преподавателем, мыслить, добиваться понимания изучаемого предмета, аккуратно вести конспект. В случае недопонимания какой-либо части предмета следует в установленном порядке задать вопрос преподавателю. В процессе работы на лекции необходимо также выполнять в конспектах модели изучаемого предмета (рисунки, схемы, чертежи и т.д.), которые использует преподаватель.

Самостоятельная подготовка обучающихся к занятиям лекционного вида включает в себя:

- доработку конспекта лекции, которую желательно осуществлять в тот же день, пока материал еще легко воспроизводим в памяти. Необходимо прочитать записи, расшифровать сокращения, доработать схемы, рисунки, таблицы;
- повторение изученного на предыдущем занятии материала. При затруднениях в восприятии материала следует обратиться к основным литературным источникам. Если разобраться в материале не удалось, то следует обратиться к преподавателю на занятиях или по графику его индивидуальных консультаций.

5.2. Методические указания для обучающихся по подготовке к практическим занятиям

Практические занятия завершают изучение наиболее важных тем учебной дисциплины. Они служат для закрепления изученного материала, развития умений и навыков подготовки докладов и эссе, приобретения опыта устных публичных выступлений, ведения дискуссии, аргументации и защиты выдвигаемых положений, а также для контроля преподавателем степени подготовленности студентов по изучаемой дисциплине.

При подготовке к практическому занятию обучающемуся необходимо:

- до очередного практического занятия по конспекту лекций и 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 семинара 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 способствует более глубокому усвоению изучаемого курса и проводится в следующих видах:

- подготовка к занятиям в соответствии с заданиями на самостоятельную работу с использованием конспекта лекций, 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, обучающиеся должны уметь пользоваться научной литературой для самостоятельной подготовки к занятиям.

В ходе самостоятельной работы студент может:

- освоить теоретический материал по изучаемой дисциплине (отдельные темы, вопросы тем, отдельные положения и т.д.);
- закрепить знания теоретического материала, используя необходимый инструментарий, практическим путем (решение практико-ориентированных задач, написание тестов для самопроверки);
- применить полученные знания и практические навыки для анализа ситуации и выработки правильного решения (подготовка к групповой дискуссии, участие в деловой игре, устный и письменный анализ конкретной ситуации);
- использовать полученные знания и умения для формирования собственной позиции, теории, модели (подготовка докладов, написание эссе).

Моделирование самостоятельной работы обучающихся:

1. Повторение пройденного теоретического материала.
2. Установление главных вопросов темы.
3. Определение глубины и содержания знаний по теме, составление тезисов по теме.
4. Упражнения, выполнение практико-ориентированных заданий.
5. Анализ выполняемой деятельности и ее самооценка.
6. Приобретение умений и навыков.
7. Составление вопросов по содержанию лекции.

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

6.1. Основная литература

1. Анализ данных : учебник для академического бакалавриата / В. С. Мхитарян [и др.] ; под редакцией В. С. Мхитаряна. — Москва : Издательство Юрайт, 2019. — 490 с. — (Бакалавр. Академический курс). — ISBN 978-5-534-00616-2. — Текст : электронный // ЭБС Юрайт [сайт]. — URL: <https://biblio-online.ru/bcode/432178>
2. Билл, Фрэнкс Революция в аналитике : как в эпоху Big Data улучшить ваш бизнес с помощью операционной аналитики / Фрэнкс Билл ; перевод И. Евстигнеева ; под редакцией В. Мылов. — Москва : Альпина Паблишер, 2017. — 320 с. — ISBN 978-5-9614-5302-7. — Текст : электронный // Электронно-библиотечная система IPR BOOKS : [сайт]. — URL: <http://www.iprbookshop.ru/58563.html>

6.2. Дополнительная литература

1. Чубукова, И. А. Data Mining : учебное пособие / И. А. Чубукова. — 3-е изд. — Москва, Саратов : Интернет-Университет Информационных Технологий (ИНТУИТ), Ай Пи Ар Медиа, 2020. — 469 с. — ISBN 978-5-4497-0289-0. — Текст : электронный // Электронно-библиотечная система IPR BOOKS : [сайт]. — URL: <http://www.iprbookshop.ru/89404.html>
2. Миркин, Б. Г. Введение в анализ данных : учебник и практикум / Б. Г. Миркин. — Москва : Издательство Юрайт, 2019. — 174 с. — (Авторский учебник). — ISBN 978-5-9916-5009-0. — Текст : электронный // ЭБС Юрайт [сайт]. — URL: <https://www.biblio-online.ru/bcode/432851>
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. — Москва : Издательство Юрайт, 2019. — 267 с. — (Бакалавр и магистр. Модуль). — ISBN 978-5-534-02556-9. — Текст : электронный // ЭБС Юрайт [сайт]. — URL: <https://biblio-online.ru/bcode/433180>
4. Билл, Фрэнкс Укрощение больших данных : как извлекать знания из массивов информации с помощью глубокой аналитики / Фрэнкс Билл ; перевод А. Баранов. — Москва : Манн, Иванов и Фербер, 2014. — 340 с. — ISBN 978-5-00057-146-0. — Текст : электронный // Электронно-библиотечная система IPR BOOKS : [сайт]. — URL: <http://www.iprbookshop.ru/39433.html>

6.3. Учебно-методическое обеспечение самостоятельной работы

Не предусмотрено.

6.4. Нормативные правовые документы

1. Федеральный закон Российской Федерации от 27 июля 2006 г. N 149-ФЗ Об информации, информационных технологиях и о защите информации, Собрание Законодательства РФ. Официальное издание. 2006 г.
2. Федеральный закон "Об информации, информатизации и защите информации", от 20 февраля 1995 Г. N 24-ФЗ (с изменениями от 10 января 2003 г.)

6.5. Интернет-ресурсы

3. www.biblio-online.ru — Электронно-библиотечная система [ЭБС] Юрайт;
4. <http://www.iprbookshop.ru> — Электронно-библиотечная система [ЭБС] «Iprbooks»
5. <https://e.lanbook.com> — Электронно-библиотечная система [ЭБС] «Лань».
6. <https://dlib.eastview.com> — East View Information Services, Inc. Коллекции электронных научных и практических журналов.
7. <http://elibrary.ru/> — Научная электронная библиотека Elibrary.ru.
8. <http://www.intuit.ru> — Национальный Открытый Университет «ИНТУИТ»;
9. <http://www.planetaexcel.ru> — сайт о возможностях MS Excel (приемы, видео-уроки, книги);
10. http://www.clouderp.ru/tags/BUSINESS_INTELLIGENCE/ Что такое Business Intelligence? Обзор BI систем
11. <http://netpeak.net/ru/blog/kak-rabotat-s-microsoft-power-bi-podrobnoe-rukovodstvo/> Как работать с Microsoft Power BI — подробное руководство
12. <http://community.powerbi.com/> Форум (англ. яз.) по Microsoft Power BI. Примеры применения. Ответы на вопросы. Использование на мобильных устройствах и др.
13. <https://www.microsoft.com/ru-RU/download/details.aspx?id=45331> – ссылка для скачивания Microsoft Power BI;
14. <https://docs.microsoft.com/ru-ru/power-bi/desktop-getting-started> – техническая документация Microsoft, документация Power BI;
15. <https://support.office.com/ru-ru/article/Вход-в-sway-db30c63b-6548-456d-b0e29fcd68bfa7a0?ui=ru-RU&rs=ru-RU&ad=RU> – руководство по созданию презентации в MS Sway;
16. <http://www.gks.ru> – официальный сайт Федеральной службы государственной статистики;
17. <http://moscow.gks.ru> – Территориальный орган Федеральной службы государственной статистики по городу Москве;
18. <https://data.mos.ru> – портал открытых данных Правительства Москвы

6.6. Иные источники

1. Паклин Н.Б., Орешков В.И Бизнес-аналитика: от данных к знаниям.: Учебное пособие — СПб.: Питер, 2013. — 704 с.
2. Виктор Майер-Шенбергер Большие данные. Революция, которая изменит то, как мы живем, работаем и мыслим [Электронный ресурс]/ Виктор Майер-Шенбергер, Кеннет Кукьер— Электрон. текстовые данные.— М.: Манн, Иванов и Фербер, 2014.— 220 с.— Режим доступа: <http://www.iprbookshop.ru/39150>.— ЭБС «IPRbooks»
3. Барский А.Б. Введение в нейронные сети [Электронный ресурс]/ А.Б. Барский— Электрон. текстовые данные. — М.: Интернет-Университет Информационных Технологий (ИНТУИТ), 2016. — 358 с. — Режим доступа: <http://www.iprbookshop.ru/52144.html>. — ЭБС «IPRbooks»
4. Ханк Д.Э., Уичерн Д.У., Райтс А.Дж. Бизнес-прогнозирование. — М.: Вильямс, 2013. — 656 с.

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

Для проведения занятий по дисциплине необходимо материально-техническое обеспечение учебных аудиторий (наглядными материалами, экраном, мультимедийным проектором с ноутбуками (ПК) для презентации учебного материала, выходом в сеть Интернет, программными продуктами Microsoft Office (Excel, Word, PowerPoint)) в зависимости от типа занятий: семинарского и лекционного типов, групповых и индивидуальных консультаций, текущего контроля и промежуточной аттестации.

Для самостоятельной работы обучающимся необходим доступ в читальные залы библиотеки и/или помещение, оснащенное компьютерной техникой с возможностью подключения к сети «Интернет», доступ в электронную информационно-образовательную среду организации и ЭБС.

Учебная аудитория для проведения занятий лекционного типа и промежуточной аттестации.

Оборудование:

Рабочие места студентов: парты, стулья;

Рабочее место преподавателя: стол, стул;

Доска для рисования маркерами;

Мультимедийный проектор.

Учебная аудитория для проведения практических занятий.

Оборудование:

Рабочие места студентов: столы, стулья;

Рабочее место преподавателя: стол, стул;

Доска для рисования маркерами,

Доска интерактивная;

Мультимедийный проектор;

Персональные компьютеры: Core i7 / 8Gb / 2000Gb -15 шт.

Программное обеспечение:

Microsoft Windows 10 Corporate 1909 (контракт с продавцом АО «Ланит» от 18.10.2019 №117/08-19);

Microsoft Office 2019 (контракт с продавцом АО «Ланит» от 18.10.2019 №117/08-19);

Google Chrome 76.0.3809.100 (свободная лицензия);

Deductor Academic 5.3.0.88 (свободная лицензия);

Loginom Community (свободная лицензия)

MS Power BI Desktop (свободная лицензия)

Библиотека (абонемент, читальный и компьютерный залы)

Учебная аудитория для самостоятельной работы студента.

Оборудование:

Рабочие места студентов: столы, стулья; Персональные компьютеры.

Программное обеспечение:

Microsoft Windows 10 Corporate 1909 (контракт с продавцом АО «Ланит» от 18.10.2019 №117/08-19);

Microsoft Office 2019 (контракт с продавцом АО «Ланит» от 18.10.2019 №117/08-19);

Deductor Academic 5.3.0.88 (свободная лицензия);

Loginom Community (свободная лицензия)

MS Power BI Desktop (свободная лицензия)

Базы данных, информационно-справочные и поисковые системы

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. <http://elibrary.ru/> - Научная электронная библиотека Elibrary.ru.