

МИНИСТЕРСТВО НАУКИ И ВЫСШЕГО ОБРАЗОВАНИЯ РОССИЙСКОЙ ФЕДЕРАЦИИ  
Федеральное государственное бюджетное образовательное учреждение высшего образования  
«ТОМСКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ СИСТЕМ  
УПРАВЛЕНИЯ И РАДИОЭЛЕКТРОНИКИ»  
(ТУСУР)



УТВЕРЖДАЮ  
Проректор по УР

Документ подписан электронной подписью

Сертификат: a1119608-cdff-4455-b54e-5235117c185c

Владелец: Семенко Павел Васильевич

Действителен: с 17.09.2019 по 16.09.2024

РАБОЧАЯ ПРОГРАММА ДИСЦИПЛИНЫ

**АНАЛИЗ БОЛЬШИХ ДАННЫХ**

Уровень образования: **высшее образование - магистратура**

Направление подготовки / специальность: **38.04.05 Бизнес-информатика**

Направленность (профиль) / специализация: **Предпринимательство и организация бизнеса в сфере информационных технологий**

Форма обучения: **очная**

Факультет: **Факультет систем управления (ФСУ)**

Кафедра: **Кафедра автоматизации обработки информации (АОИ)**

Курс: **1**

Семестр: **2**

Учебный план набора 2023 года

Объем дисциплины и виды учебной деятельности

| Виды учебной деятельности          | 2 семестр | Всего | Единицы |
|------------------------------------|-----------|-------|---------|
| Лекционные занятия                 | 18        | 18    | часов   |
| Лабораторные занятия               | 36        | 36    | часов   |
| Самостоятельная работа             | 54        | 54    | часов   |
| Общая трудоемкость                 | 108       | 108   | часов   |
| (включая промежуточную аттестацию) | 3         | 3     | з.е.    |

| Формы промежуточной аттестация | Семестр |
|--------------------------------|---------|
| Зачет с оценкой                | 2       |

## 1. Общие положения

### 1.1. Цели дисциплины

1. Освоение теоретических основ и методов анализа данных, применяемых при решении прикладных задач с помощью языка программирования Python.

### 1.2. Задачи дисциплины

1. Формирование навыков обработки и анализа информации для получения общей оценки массива данных, и выявления тенденций, закономерностей в изучаемых наборах данных.

2. Освоение языка программирования Python, Interactive Python в виде Jupyter Notebook, библиотеки Pandas, применяемых при анализе данных.

## 2. Место дисциплины в структуре ОПОП

Блок дисциплин: Б1. Дисциплины (модули).

Часть блока дисциплин: Обязательная часть.

Модуль дисциплин: Модуль направления подготовки (hard skills - HS).

Индекс дисциплины: Б1.О.02.02.

Реализуется с применением электронного обучения, дистанционных образовательных технологий.

## 3. Перечень планируемых результатов обучения по дисциплине, соотнесенных с индикаторами достижения компетенций

Процесс изучения дисциплины направлен на формирование следующих компетенций в соответствии с ФГОС ВО и основной образовательной программой (таблица 3.1):

Таблица 3.1 – Компетенции и индикаторы их достижения

| Компетенция                             | Индикаторы достижения компетенции | Планируемые результаты обучения по дисциплине |
|---|-----------------------------------|---|
| <b>Универсальные компетенции</b>        |                                   |   |
| -                                       | -                                 | -   |
| <b>Общепрофессиональные компетенции</b> |                                   |   |

|   |   |   |
|---|---|---|
| ОПК-3. Способен принимать решения, осуществлять стратегическое планирование и прогнозирование в профессиональной деятельности с использованием современных методов и программного инструментария сбора, обработки и анализа данных, интеллектуального оборудования и систем искусственного интеллекта | ОПК-3.1. Знает теоретико-методологические основы принятия решений, осуществления стратегического планирования и прогнозирования в профессиональной деятельности   | Знает методы использования средств Python для обработки и анализа данных. Понимает структуру современного конвейера обработки данных и средства работы со сверхбольшими объемами данных |
|   | ОПК-3.2. Умеет разрабатывать, принимать и реализовывать стратегические решения в профессиональной деятельности с использованием современных методов и программного инструментария сбора, обработки и анализа данных, интеллектуального оборудования и систем искусственного интеллекта                            | Умеет выполнить очистку и объединение данных, представленных в разных форматах и из разных источников средствами Python   |
|   | ОПК-3.3. Владеет навыками использования современных методов и программного инструментария сбора, обработки и анализа данных, интеллектуального оборудования и систем искусственного интеллекта при принятии решений, осуществлении стратегического планирования и прогнозирования в профессиональной деятельности | Владеет навыками развертывания, настройки и применения кластера Hadoop для обработки сверхбольших массивов данных   |
| <b>Профессиональные компетенции</b>   |   |   |
| -   | -   | -   |

**4. Объем дисциплины в зачетных единицах с указанием количества академических часов, выделенных на контактную работу обучающихся с преподавателем и на самостоятельную работу обучающихся**

Общая трудоемкость дисциплины составляет 3 зачетных единиц, 108 академических часов.

Распределение трудоемкости дисциплины по видам учебной деятельности представлено в таблице 4.1.

Таблица 4.1 – Трудоемкость дисциплины по видам учебной деятельности

| Виды учебной деятельности   | Всего часов | Семестры  |
|---|-------------|-----------|
|   |             | 2 семестр |
| <b>Контактная аудиторная работа обучающихся с преподавателем, всего</b> | 54          | 54        |

|   |     |     |
|---|-----|-----|
| Лекционные занятия  | 18  | 18  |
| Лабораторные занятия  | 36  | 36  |
| <b>Самостоятельная работа обучающихся, в т.ч. контактная внеаудиторная работа обучающихся с преподавателем, всего</b> | 54  | 54  |
| Подготовка к зачету с оценкой   | 18  | 18  |
| Подготовка к тестированию   | 18  | 18  |
| Подготовка к лабораторной работе, написание отчета  | 18  | 18  |
| <b>Общая трудоемкость (в часах)</b>   | 108 | 108 |
| <b>Общая трудоемкость (в з.е.)</b>  | 3   | 3   |

## 5. Структура и содержание дисциплины

### 5.1. Разделы (темы) дисциплины и виды учебной деятельности

Структура дисциплины по разделам (темам) и видам учебной деятельности приведена в таблице 5.1.

Таблица 5.1 – Разделы (темы) дисциплины и виды учебной деятельности

| Названия разделов (тем) дисциплины               | Лек. зан., ч | Лаб. раб. | Сам. раб., ч | Всего часов (без экзамена) | Формируемые компетенции |
|--|--------------|-----------|--------------|----------------------------|-------------------------|
| <b>2 семестр</b>                                 |              |           |              |                            |                         |
| 1 Введение                                       | 2            | 4         | 6            | 12                         | ОПК-3                   |
| 2 Основные принципы и этапы анализа данных       | 2            | 4         | 6            | 12                         | ОПК-3                   |
| 3 Анализ данных может быть разным                | 2            | 4         | 6            | 12                         | ОПК-3                   |
| 4 Подготовка и очистка данных                    | 2            | 4         | 6            | 12                         | ОПК-3                   |
| 5 Выявление зависимостей и связей в данных       | 2            | 4         | 6            | 12                         | ОПК-3                   |
| 6 Методы машинного обучения для обработки данных | 2            | 12        | 10           | 24                         | ОПК-3                   |
| 7 Графическое представление данных               | 2            | 4         | 6            | 12                         | ОПК-3                   |
| 8 Большие данные. Экосистема Hadoop              | 2            | -         | 4            | 6                          | ОПК-3                   |
| 9 Распределенные системы анализа данных          | 2            | -         | 4            | 6                          | ОПК-3                   |
| Итого за семестр                                 | 18           | 36        | 54           | 108                        |                         |
| Итого  | 18           | 36        | 54           | 108                        |                         |

### 5.2. Содержание разделов (тем) дисциплины

Содержание разделов (тем) дисциплины (в т.ч. по лекциям) приведено в таблице 5.2.

Таблица 5.2 – Содержание разделов (тем) дисциплины (в т.ч. по лекциям)

| Названия разделов (тем) дисциплины | Содержание разделов (тем) дисциплины (в т.ч. по лекциям)                      | Трудоемкость (лекционные занятия), ч | Формируемые компетенции |
|------------------------------------|---|--------------------------------------|-------------------------|
| <b>2 семестр</b>                   |   |                                      |                         |
| 1 Введение                         | Причины появления термина "Большие данные". Роль данных в информационном мире | 2                                    | ОПК-3                   |
|                                    | Итого   | 2                                    |                         |

|  |   |    |       |
|--|---|----|-------|
| 2 Основные принципы и этапы анализа данных       | Методика CRISP-DM. Данные в покое и движении. Структурированные и неструктурированные данные  | 2  | ОПК-3 |
|  | Итого   | 2  |       |
| 3 Анализ данных может быть разным                | Описательный, прогнозный и рекомендательный анализ  | 2  | ОПК-3 |
|  | Итого   | 2  |       |
| 4 Подготовка и очистка данных                    | Методы преобразования форматов в Python. Методы статистической обработки данных   | 2  | ОПК-3 |
|  | Итого   | 2  |       |
| 5 Выявление зависимостей и связей в данных       | Выявление тренда, линейная и нелинейная регрессия. Методы кластеризации   | 2  | ОПК-3 |
|  | Итого   | 2  |       |
| 6 Методы машинного обучения для обработки данных | Общие принципы работы нейронных сетей, глубокое обучение, понятие "искусственный интеллект" и "машинное обучение", методы машинного обучения  | 2  | ОПК-3 |
|  | Итого   | 2  |       |
| 7 Графическое представление данных               | Библиотека Python Matplotlib, возможности, использование в Jupyter Notebook, особенности работы и настройки. Библиотеки PyPlot и Plotly   | 2  | ОПК-3 |
|  | Итого   | 2  |       |
| 8 Большие данные. Экосистема Hadoop              | Принцип Map-Reduce, основные характеристики кластера Hadoop. Файловая система HDFS, инструменты настройки YARN, управление доступом к данным Pig, Hive. Другие инструменты экосистемы | 2  | ОПК-3 |
|  | Итого   | 2  |       |
| 9 Распределенные системы анализа данных          | Процесс ETL. Традиционная архитектура обработки данных и архитектура Lambda, потоковая обработка данных в реальном времени. Архитектура Карра, проблемы распределенной обработки      | 2  | ОПК-3 |
|  | Итого   | 2  |       |
| Итого за семестр                                 |   | 18 |       |
| Итого  |   | 18 |       |

### 5.3. Практические занятия (семинары)

Не предусмотрено учебным планом

### 5.4. Лабораторные занятия

Наименование лабораторных работ приведено в таблице 5.4.

Таблица 5.4 – Наименование лабораторных работ

| Названия разделов (тем) дисциплины               | Наименование лабораторных работ   | Трудоемкость, ч | Формируемые компетенции |
|--|---|-----------------|-------------------------|
| <b>2 семестр</b>                                 |   |                 |                         |
| 1 Введение                                       | Анализ криминальной обстановки в Сан-Франциско                                    | 4               | ОПК-3                   |
|  | Итого   | 4               |                         |
| 2 Основные принципы и этапы анализа данных       | Измерение и анализ данных Интернет-соединения                                     | 4               | ОПК-3                   |
|  | Итого   | 4               |                         |
| 3 Анализ данных может быть разным                | Измерение и анализ Интернет-соединения с хранилищем данных в виде SQL базы данных | 4               | ОПК-3                   |
|  | Итого   | 4               |                         |
| 4 Подготовка и очистка данных                    | Описательная статистика в Python  | 4               | ОПК-3                   |
|  | Итого   | 4               |                         |
| 5 Выявление зависимостей и связей в данных       | Корреляционный анализ в Python  | 4               | ОПК-3                   |
|  | Итого   | 4               |                         |
| 6 Методы машинного обучения для обработки данных | Простая линейная регрессия в Python   | 4               | ОПК-3                   |
|  | Классификация с помощью дерева решений  | 4               | ОПК-3                   |
|  | Обработка ошибок при использовании метода линейной регрессии                      | 4               | ОПК-3                   |
|  | Итого   | 12              |                         |
| 7 Графическое представление данных               | Измерение и анализ Интернет-соединения, визуализация результатов                  | 4               | ОПК-3                   |
|  | Итого   | 4               |                         |
| Итого за семестр                                 |   | 36              |                         |
| Итого  |   | 36              |                         |

### 5.5. Курсовой проект / курсовая работа

Не предусмотрено учебным планом

### 5.6. Самостоятельная работа

Виды самостоятельной работы, трудоемкость и формируемые компетенции представлены в таблице 5.6.

Таблица 5.6. – Виды самостоятельной работы, трудоемкость и формируемые компетенции

| Названия разделов (тем) дисциплины | Виды самостоятельной работы | Трудоемкость, ч | Формируемые компетенции | Формы контроля |
|------------------------------------|-----------------------------|-----------------|-------------------------|----------------|
| <b>2 семестр</b>                   |                             |                 |                         |                |

|  |  |    |       |                     |
|--|--|----|-------|---------------------|
| 1 Введение                                       | Подготовка к зачету с оценкой                      | 2  | ОПК-3 | Зачёт с оценкой     |
|  | Подготовка к тестированию                          | 2  | ОПК-3 | Тестирование        |
|  | Подготовка к лабораторной работе, написание отчета | 2  | ОПК-3 | Лабораторная работа |
|  | Итого  | 6  |       |                     |
| 2 Основные принципы и этапы анализа данных       | Подготовка к зачету с оценкой                      | 2  | ОПК-3 | Зачёт с оценкой     |
|  | Подготовка к тестированию                          | 2  | ОПК-3 | Тестирование        |
|  | Подготовка к лабораторной работе, написание отчета | 2  | ОПК-3 | Лабораторная работа |
|  | Итого  | 6  |       |                     |
| 3 Анализ данных может быть разным                | Подготовка к зачету с оценкой                      | 2  | ОПК-3 | Зачёт с оценкой     |
|  | Подготовка к тестированию                          | 2  | ОПК-3 | Тестирование        |
|  | Подготовка к лабораторной работе, написание отчета | 2  | ОПК-3 | Лабораторная работа |
|  | Итого  | 6  |       |                     |
| 4 Подготовка и очистка данных                    | Подготовка к зачету с оценкой                      | 2  | ОПК-3 | Зачёт с оценкой     |
|  | Подготовка к тестированию                          | 2  | ОПК-3 | Тестирование        |
|  | Подготовка к лабораторной работе, написание отчета | 2  | ОПК-3 | Лабораторная работа |
|  | Итого  | 6  |       |                     |
| 5 Выявление зависимостей и связей в данных       | Подготовка к зачету с оценкой                      | 2  | ОПК-3 | Зачёт с оценкой     |
|  | Подготовка к тестированию                          | 2  | ОПК-3 | Тестирование        |
|  | Подготовка к лабораторной работе, написание отчета | 2  | ОПК-3 | Лабораторная работа |
|  | Итого  | 6  |       |                     |
| 6 Методы машинного обучения для обработки данных | Подготовка к зачету с оценкой                      | 2  | ОПК-3 | Зачёт с оценкой     |
|  | Подготовка к тестированию                          | 2  | ОПК-3 | Тестирование        |
|  | Подготовка к лабораторной работе, написание отчета | 6  | ОПК-3 | Лабораторная работа |
|  | Итого  | 10 |       |                     |

|   |  |    |       |                     |
|---|--|----|-------|---------------------|
| 7 Графическое представление данных      | Подготовка к зачету с оценкой                      | 2  | ОПК-3 | Зачёт с оценкой     |
|   | Подготовка к тестированию                          | 2  | ОПК-3 | Тестирование        |
|   | Подготовка к лабораторной работе, написание отчета | 2  | ОПК-3 | Лабораторная работа |
|   | Итого  | 6  |       |                     |
| 8 Большие данные. Экосистема Hadoop     | Подготовка к зачету с оценкой                      | 2  | ОПК-3 | Зачёт с оценкой     |
|   | Подготовка к тестированию                          | 2  | ОПК-3 | Тестирование        |
|   | Итого  | 4  |       |                     |
| 9 Распределенные системы анализа данных | Подготовка к зачету с оценкой                      | 2  | ОПК-3 | Зачёт с оценкой     |
|   | Подготовка к тестированию                          | 2  | ОПК-3 | Тестирование        |
|   | Итого  | 4  |       |                     |
| Итого за семестр                        |  | 54 |       |                     |
| Итого                                   |  | 54 |       |                     |

### 5.7. Соответствие компетенций, формируемых при изучении дисциплины, и видов учебной деятельности

Соответствие компетенций, формируемых при изучении дисциплины, и видов учебной деятельности представлено в таблице 5.7.

Таблица 5.7 – Соответствие компетенций, формируемых при изучении дисциплины, и видов занятий

| Формируемые компетенции | Виды учебной деятельности |           |           | Формы контроля                                     |
|-------------------------|---------------------------|-----------|-----------|--|
|                         | Лек. зан.                 | Лаб. раб. | Сам. раб. |  |
| ОПК-3                   | +                         | +         | +         | Зачёт с оценкой, Лабораторная работа, Тестирование |

## 6. Рейтинговая система для оценки успеваемости обучающихся

### 6.1. Балльные оценки для форм контроля

Балльные оценки для форм контроля представлены в таблице 6.1.

Таблица 6.1 – Балльные оценки

| Формы контроля           | Максимальный балл на 1-ую КТ с начала семестра | Максимальный балл за период между 1КТ и 2КТ | Максимальный балл за период между 2КТ и на конец семестра | Всего за семестр |
|--------------------------|--|---|---|------------------|
| <b>2 семестр</b>         |  |   |   |                  |
| Зачёт с оценкой          | 0  | 0   | 0   | 0                |
| Лабораторная работа      | 20   | 20  | 20  | 60               |
| Тестирование             | 10   | 10  | 20  | 40               |
| Итого максимум за период | 30   | 30  | 40  | 100              |
| Нарастающим итогом       | 30   | 60  | 100   | 100              |

### 6.2. Пересчет баллов в оценки за текущий контроль



Пересчет баллов в оценки за текущий контроль представлен в таблице 6.2.  
Таблица 6.2 – Пересчет баллов в оценки за текущий контроль

| Баллы на дату текущего контроля                       | Оценка |
|---|--------|
| ≥ 90% от максимальной суммы баллов на дату ТК         | 5      |
| От 70% до 89% от максимальной суммы баллов на дату ТК | 4      |
| От 60% до 69% от максимальной суммы баллов на дату ТК | 3      |
| < 60% от максимальной суммы баллов на дату ТК         | 2      |

### 6.3. Пересчет суммы баллов в традиционную и международную оценку

Пересчет суммы баллов в традиционную и международную оценку представлен в таблице 6.3.

Таблица 6.3 – Пересчет суммы баллов в традиционную и международную оценку

| Оценка                               | Итоговая сумма баллов, учитывает успешно сданный экзамен | Оценка (ECTS)           |
|--------------------------------------|--|-------------------------|
| 5 (отлично) (зачтено)                | 90 – 100   | A (отлично)             |
| 4 (хорошо) (зачтено)                 | 85 – 89  | B (очень хорошо)        |
|                                      | 75 – 84  | C (хорошо)              |
|                                      | 70 – 74  | D (удовлетворительно)   |
| 3 (удовлетворительно) (зачтено)      | 65 – 69  | E (посредственно)       |
|                                      | 60 – 64  |                         |
| 2 (неудовлетворительно) (не зачтено) | Ниже 60 баллов   | F (неудовлетворительно) |

## 7. Учебно-методическое и информационное обеспечение дисциплины

### 7.1. Основная литература

1. Маккинни, У. Python и анализ данных / У. Маккинни ; перевод с английского А. А. Слинкина. — 2-ое изд., испр. и доп. — Москва : ДМК Пресс, 2020. — 540 с. [Электронный ресурс]: — Режим доступа: <https://e.lanbook.com/book/131721>.

### 7.2. Дополнительная литература

1. Лебедев, А. С. Методы Big Data : учебно-методическое пособие / А. С. Лебедев, Ш. Г. Магомедов. — Москва : РТУ МИРЭА, 2021. — 91 с. [Электронный ресурс]: — Режим доступа: <https://e.lanbook.com/book/182452>.

2. Программные системы статистического анализа. Обнаружение закономерностей в данных с использованием системы R и языка Python : учебное пособие / В. М. Волкова, М. А. Семёнова, Е. С. Четвертакова, С. С. Вожов. — Новосибирск : НГТУ, 2017. — 74 с. [Электронный ресурс]: — Режим доступа: <https://e.lanbook.com/book/118287>.

### 7.3. Учебно-методические пособия

#### 7.3.1. Обязательные учебно-методические пособия

1. Анализ больших данных: Методические указания к лабораторным работам, и организации самостоятельной работы для студентов направлений «Бизнес-информатика» и «Программная инженерия» (уровень магистратуры) / Е. Ю. Агеев - 2021. 57 с. [Электронный ресурс]: — Режим доступа: <https://edu.tusur.ru/publications/10509>.

#### 7.3.2. Учебно-методические пособия для лиц с ограниченными возможностями здоровья и инвалидов

Учебно-методические материалы для самостоятельной и аудиторной работы обучающихся из числа лиц с ограниченными возможностями здоровья и инвалидов предоставляются в формах, адаптированных к ограничениям их здоровья и восприятия информации.

**Для лиц с нарушениями зрения:**

– в форме электронного документа;

– в печатной форме увеличенным шрифтом.

**Для лиц с нарушениями слуха:**

– в форме электронного документа;

– в печатной форме.

**Для лиц с нарушениями опорно-двигательного аппарата:**

– в форме электронного документа;

– в печатной форме.

#### **7.4. Современные профессиональные базы данных и информационные справочные системы**

1. При изучении дисциплины рекомендуется обращаться к современным базам данных, информационно-справочным и поисковым системам, к которым у ТУСУРа открыт доступ: <https://lib.tusur.ru/ru/resursy/bazy-dannyh>.

### **8. Материально-техническое и программное обеспечение дисциплины**

#### **8.1. Материально-техническое и программное обеспечение для лекционных занятий**

Для проведения занятий лекционного типа, групповых и индивидуальных консультаций, текущего контроля и промежуточной аттестации используется учебная аудитория с достаточным количеством посадочных мест для учебной группы, оборудованная доской и стандартной учебной мебелью. Имеются мультимедийное оборудование и учебно-наглядные пособия, обеспечивающие тематические иллюстрации по лекционным разделам дисциплины.

#### **8.2. Материально-техническое и программное обеспечение для лабораторных работ**

Компьютерный класс: учебная аудитория для проведения занятий практического типа, учебная аудитория для проведения занятий лабораторного типа, помещение для курсового проектирования (выполнения курсовых работ), помещение для проведения групповых и индивидуальных консультаций; 634034, Томская область, г. Томск, Вершинина улица, д. 74, 407 ауд.

Описание имеющегося оборудования:

- Интерактивный плоскпанельный дисплей SMART VIZION DC75-E4;
- Магнитно-маркерная доска;
- Комплект специализированной учебной мебели;
- Рабочее место преподавателя.

Программное обеспечение:

- Google Chrome, Open Source;
- Microsoft Office 2013 Standard;
- Python, свободно распространяемое ПО совместимое с GNU GPL;
- Windows 10 Professional, AZURE DEV TOOLS FOR TEACHING;

#### **8.3. Материально-техническое и программное обеспечение для самостоятельной работы**

Для самостоятельной работы используются учебные аудитории (компьютерные классы), расположенные по адресам:

- 634050, Томская область, г. Томск, Ленина проспект, д. 40, 233 ауд.;
- 634045, Томская область, г. Томск, ул. Красноармейская, д. 146, 209 ауд.;
- 634034, Томская область, г. Томск, Вершинина улица, д. 47, 126 ауд.;
- 634034, Томская область, г. Томск, Вершинина улица, д. 74, 207 ауд.

Описание имеющегося оборудования:

- учебная мебель;  
- компьютеры;  
- компьютеры подключены к сети «Интернет» и обеспечивают доступ в электронную информационно-образовательную среду ТУСУРа.

Перечень программного обеспечения:

- Microsoft Windows;
- OpenOffice;
- Kaspersky Endpoint Security 10 для Windows;

- 7-Zip;
- Google Chrome.

#### 8.4. Материально-техническое обеспечение дисциплины для лиц с ограниченными возможностями здоровья и инвалидов

Освоение дисциплины лицами с ограниченными возможностями здоровья и инвалидами осуществляется с использованием средств обучения общего и специального назначения.

При занятиях с обучающимися с **нарушениями слуха** предусмотрено использование звукоусиливающей аппаратуры, мультимедийных средств и других технических средств приема/передачи учебной информации в доступных формах, мобильной системы преподавания для обучающихся с инвалидностью, портативной индукционной системы. Учебная аудитория, в которой занимаются обучающиеся с нарушением слуха, оборудована компьютерной техникой, аудиотехникой, видеотехникой, электронной доской, мультимедийной системой.

При занятиях с обучающимися с **нарушениями зрения** предусмотрено использование в лекционных и учебных аудиториях возможности просмотра удаленных объектов (например, текста на доске или слайда на экране) при помощи видеоувеличителей для комфортного просмотра.

При занятиях с обучающимися с **нарушениями опорно-двигательного аппарата** используются альтернативные устройства ввода информации и другие технические средства приема/передачи учебной информации в доступных формах, мобильной системы обучения для людей с инвалидностью.

### 9. Оценочные материалы и методические рекомендации по организации изучения дисциплины

#### 9.1. Содержание оценочных материалов для текущего контроля и промежуточной аттестации

Для оценки степени сформированности и уровня освоения закрепленных за дисциплиной компетенций используются оценочные материалы, представленные в таблице 9.1.

Таблица 9.1 – Формы контроля и оценочные материалы

| Названия разделов (тем) дисциплины         | Формируемые компетенции | Формы контроля      | Оценочные материалы (ОМ)               |
|--|-------------------------|---------------------|--|
| 1 Введение                                 | ОПК-3                   | Зачёт с оценкой     | Перечень вопросов для зачета с оценкой |
|  |                         | Лабораторная работа | Темы лабораторных работ                |
|  |                         | Тестирование        | Примерный перечень тестовых заданий    |
| 2 Основные принципы и этапы анализа данных | ОПК-3                   | Зачёт с оценкой     | Перечень вопросов для зачета с оценкой |
|  |                         | Лабораторная работа | Темы лабораторных работ                |
|  |                         | Тестирование        | Примерный перечень тестовых заданий    |
| 3 Анализ данных может быть разным          | ОПК-3                   | Зачёт с оценкой     | Перечень вопросов для зачета с оценкой |
|  |                         | Лабораторная работа | Темы лабораторных работ                |
|  |                         | Тестирование        | Примерный перечень тестовых заданий    |

|  |       |                     |  |
|--|-------|---------------------|--|
| 4 Подготовка и очистка данных                    | ОПК-3 | Зачёт с оценкой     | Перечень вопросов для зачета с оценкой |
|  |       | Лабораторная работа | Темы лабораторных работ                |
|  |       | Тестирование        | Примерный перечень тестовых заданий    |
| 5 Выявление зависимостей и связей в данных       | ОПК-3 | Зачёт с оценкой     | Перечень вопросов для зачета с оценкой |
|  |       | Лабораторная работа | Темы лабораторных работ                |
|  |       | Тестирование        | Примерный перечень тестовых заданий    |
| 6 Методы машинного обучения для обработки данных | ОПК-3 | Зачёт с оценкой     | Перечень вопросов для зачета с оценкой |
|  |       | Лабораторная работа | Темы лабораторных работ                |
|  |       | Тестирование        | Примерный перечень тестовых заданий    |
| 7 Графическое представление данных               | ОПК-3 | Зачёт с оценкой     | Перечень вопросов для зачета с оценкой |
|  |       | Лабораторная работа | Темы лабораторных работ                |
|  |       | Тестирование        | Примерный перечень тестовых заданий    |
| 8 Большие данные. Экосистема Hadoop              | ОПК-3 | Зачёт с оценкой     | Перечень вопросов для зачета с оценкой |
|  |       | Тестирование        | Примерный перечень тестовых заданий    |
| 9 Распределенные системы анализа данных          | ОПК-3 | Зачёт с оценкой     | Перечень вопросов для зачета с оценкой |
|  |       | Тестирование        | Примерный перечень тестовых заданий    |

Шкала оценки сформированности отдельных планируемых результатов обучения по дисциплине приведена в таблице 9.2.

Таблица 9.2 – Шкала оценки сформированности планируемых результатов обучения по дисциплине

| Оценка                     | Баллы за ОМ                        | Формулировка требований к степени сформированности планируемых результатов обучения |   |   |
|----------------------------|------------------------------------|---|---|---|
|                            |                                    | знать   | уметь   | владеть   |
| 2<br>(неудовлетворительно) | < 60% от максимальной суммы баллов | отсутствие знаний или фрагментарные знания  | отсутствие умений или частично освоенное умение | отсутствие навыков или фрагментарные применение навыков |

|                          |  |   |   |  |
|--------------------------|--|---|---|--|
| 3<br>(удовлетворительно) | от 60% до 69% от максимальной суммы баллов | общие, но не структурированные знания                   | в целом успешно, но не систематически осуществляемое умение | в целом успешное, но не систематическое применение навыков           |
| 4 (хорошо)               | от 70% до 89% от максимальной суммы баллов | сформированные, но содержащие отдельные проблемы знания | в целом успешное, но содержащие отдельные пробелы умение    | в целом успешное, но содержащие отдельные пробелы применение навыков |
| 5 (отлично)              | ≥ 90% от максимальной суммы баллов         | сформированные систематические знания                   | сформированное умение                                       | успешное и систематическое применение навыков                        |

Шкала комплексной оценки сформированности компетенций приведена в таблице 9.3.

Таблица 9.3 – Шкала комплексной оценки сформированности компетенций

| Оценка                     | Формулировка требований к степени компетенции  |
|----------------------------|--|
| 2<br>(неудовлетворительно) | Не имеет необходимых представлений о проверяемом материале или<br>Знать на уровне <b>ориентирования</b> , представлений. Обучающийся знает основные признаки или термины изучаемого элемента содержания, их отнесенность к определенной науке, отрасли или объектам, узнает в текстах, изображениях или схемах и знает, к каким источникам нужно обращаться для более детального его усвоения. |
| 3<br>(удовлетворительно)   | Знать и уметь на <b>репродуктивном</b> уровне. Обучающихся знает изученный элемент содержания репродуктивно: произвольно воспроизводит свои знания устно, письменно или в демонстрируемых действиях.   |
| 4 (хорошо)                 | Знать, уметь, владеть на <b>аналитическом</b> уровне. Зная на репродуктивном уровне, указывать на особенности и взаимосвязи изученных объектов, на их достоинства, ограничения, историю и перспективы развития и особенности для разных объектов усвоения.   |
| 5 (отлично)                | Знать, уметь, владеть на <b>системном</b> уровне. Обучающийся знает изученный элемент содержания системно, произвольно и доказательно воспроизводит свои знания устно, письменно или в демонстрируемых действиях, учитывая и указывая связи и зависимости между этим элементом и другими элементами содержания дисциплины, его значимость в содержании дисциплины.                             |

### 9.1.1. Примерный перечень тестовых заданий

- Как с помощью NumPy найти ковариацию двух переменных?
  - Функция NumPy cov() вычисляет ковариационную матрицу.
  - Функция NumPy arange() вычисляет ковариационную матрицу.
  - Функция NumPy ravel() вычисляет ковариационную матрицу.
  - Функция NumPy where() вычисляет ковариационную матрицу.
- Какова была точность прогноза на обучающей и тестовой выборке данных при выполнении части 3 лабораторной работы "Классификация с помощью дерева решений"?
  - обучающая 90%, тестовая 82%
  - обучающая 80%, тестовая 71%

- в) обучающая 60%, тестовая 37%
- г) обучающая 100%, тестовая 102%
3. Линейную регрессию в Python можно выполнить разными способами, назовите несколько?
    - а) С помощью LinearRegression() из Scikit-learn - LinearRegression().fit(x, y)
    - б) С помощью функции ols() библиотеки Statsmodel
    - в) Можно написать функцию расчета в соответствии с известным алгоритмом
    - г) С помощью библиотеки SciPy и функции Optimize.curve\_fit()
  4. Является ли линейная регрессия методом машинного обучения?
    - а) Да
    - б) Нет
    - в) Возможно
    - г) Только если применяется к машинам.
  5. Какие библиотеки применяются для визуализации данных Python, назовите и кратко опишите пять таких библиотек.
    - а) Matplotlib, Seaborn, Folium, Bokeh, Plotly
    - б) Pandas, NumPy, Scikit-learn, Gradio, Keras
    - в) TensorFlow, Caffe, PyTorch, OpenCV, Theano
    - г) Pyevolve, NuPIC, Pylearn2, SciPy, SpaCy
  6. Каков точный формат задания подписей к осям графика, наименования графика, добавления легенды для выводимых на график величин?
    - а) plt.xlabel('Ось X'), plt.ylabel('Ось Y'), plt.title('Заголовок графика'), ax.legend();
    - б) axes.set\_xlabel('Ось X'), axes.set\_ylabel('ось Y'), axes.set\_title('Заголовок графика'), ax.legend();
    - в) axes.get\_xlabel('Ось X'), axes.get\_ylabel('ось Y'), axes.get\_title('Заголовок графика'), ax.legend();
    - г) axes.get\_xlabel('Ось X'), axes.get\_ylabel('ось Y'), axes.get\_title('Заголовок графика'), ax.legend();
  7. Если переменные сравниваемой пары данных представляют корреляцию, близкую к нулю, то что это означает?
    - а) Нет взаимной зависимости.
    - б) Есть полная взаимная зависимость.
    - в) Есть сильная корреляция.
    - г) Ничего не означает.
  8. Какие статистические значения для датафрейма позволяет получить метод describe() ?
    - а) Число записей, среднееквадратичное значение, стандартное отклонение, минимальное и максимальное значение, перцентиль для 25%, 50%, 75%.
    - б) Число, среднее геометрическое значение, нестандартное отклонение, минимальное и максимальное значение, перцентиль для 2%, 3%, 4%.
    - в) Число подписей, среднее аргументированное значение, стандартное приближение, минимальное и максимальное удаление, перцентиль для 5%, 10%, 20%.
    - г) Число описей, точное значение, приближенное отклонение, минимальное и максимальное увеличение, перцентиль для 125%, 500%, 1750%.
  9. Как при выводе значения переменной в текстовой строке оставить только 2 цифры после десятичной точки, если исходное значение содержит больше цифр?
    - а) С помощью оператора форматирования строки (%) - print("%.2f %")
    - б) С помощью строкового метода format - print('{:.2f}'.format(4/3))
    - в) С помощью форматированного строкового литерала - print(f'{a:.2f}')
    - г) Если значение переменной содержит больше цифр, то оставить в выводе только 2 невозможно.
  10. Как добавить в датафрейм новый столбец с данными?
    - а) Задать имя для нового столбца и определить его содержимое df["C"] = [10, 20, 30, 40]
    - б) С помощью функции вставки df.insert(1, "D", 5)
    - в) С помощью функции loc - df.loc[:, "E"] = list("abcd")
    - г) С помощью функции assign - df = df.assign(F = df.C \* 10)

### 9.1.2. Перечень вопросов для зачета с оценкой

1. У вас есть координаты расположения 10 АЗС в городской черте. Напишите программу отображения АЗС в сервисе OpenStreetMaps.
2. Вам дан датафрейм, содержащий записи в 10 колонках и 100 строках. Нужно удалить 7 и 9 колонку, и добавить две новые: одну со средним значением для строки, вторую с дисперсией в строке. Некоторые записи в исходном датафрейме содержат NaN. Как выполнить задачу с помощью Python?
3. Записи в датафрейме содержат URL веб-ресурсов. Как с помощью Python выделить для каждой записи только доменное имя первого уровня?
4. Мини-кластер Hadoop создан на одном PC. Какой выигрыш производительности при обработке задачи с хорошим параллелизмом можно получить, если запускать ее на таком кластере и применить алгоритм Map-Reduce?
5. Какие признаки должны иметь данные для отнесения их к категории "Большие данные"?

### **9.1.3. Темы лабораторных работ**

1. Анализ криминальной обстановки в Сан-Франциско
2. Измерение и анализ данных Интернет-соединения
3. Измерение и анализ Интернет-соединения с хранилищем данных в виде SQL базы данных
4. Описательная статистика в Python
5. Корреляционный анализ в Python
6. Простая линейная регрессия в Python
7. Классификация с помощью дерева решений
8. Обработка ошибок при использовании метода линейной регрессии
9. Измерение и анализ Интернет-соединения, визуализация результатов

### **9.2. Методические рекомендации**

Учебный материал излагается в форме, предполагающей самостоятельное мышление студентов, самообразование. При этом самостоятельная работа студентов играет решающую роль в ходе всего учебного процесса.

Начать изучение дисциплины необходимо со знакомства с рабочей программой, списком учебно-методического и программного обеспечения. Самостоятельная работа студента включает работу с учебными материалами, выполнение контрольных мероприятий, предусмотренных учебным планом.

В процессе изучения дисциплины для лучшего освоения материала необходимо регулярно обращаться к рекомендуемой литературе и источникам, указанным в учебных материалах; пользоваться через кабинет студента на сайте Университета образовательными ресурсами электронно-библиотечной системы, а также общедоступными интернет-порталами, содержащими научно-популярные и специализированные материалы, посвященные различным аспектам учебной дисциплины.

При самостоятельном изучении тем следуйте рекомендациям:

– чтение или просмотр материала осуществляйте со скоростью, достаточной для индивидуального понимания и освоения материала, выделяя основные идеи; на основании изученного составить тезисы. Освоив материал, попытаться соотнести теорию с примерами из практики;

– если в тексте встречаются незнакомые или малознакомые термины, следует выяснить их значение для понимания дальнейшего материала;

– осмысливайте прочитанное и изученное, отвечайте на предложенные вопросы.

Студенты могут получать индивидуальные консультации, в т.ч. с использованием средств телекоммуникации.

По дисциплине могут проводиться дополнительные занятия, в т.ч. в форме вебинаров. Расписание вебинаров и записи вебинаров публикуются в электронном курсе / электронном журнале по дисциплине.

### **9.3. Требования к оценочным материалам для лиц с ограниченными возможностями здоровья и инвалидов**

Для лиц с ограниченными возможностями здоровья и инвалидов предусмотрены дополнительные оценочные материалы, перечень которых указан в таблице 9.4.

Таблица 9.4 – Дополнительные материалы оценивания для лиц с ограниченными возможностями здоровья и инвалидов

| Категории обучающихся                         | Виды дополнительных оценочных материалов  | Формы контроля и оценки результатов обучения   |
|---|---|--|
| С нарушениями слуха                           | Тесты, письменные самостоятельные работы, вопросы к зачету, контрольные работы                        | Преимущественно письменная проверка  |
| С нарушениями зрения                          | Собеседование по вопросам к зачету, опрос по терминам   | Преимущественно устная проверка (индивидуально)  |
| С нарушениями опорно-двигательного аппарата   | Решение дистанционных тестов, контрольные работы, письменные самостоятельные работы, вопросы к зачету | Преимущественно дистанционными методами  |
| С ограничениями по общемедицинским показаниям | Тесты, письменные самостоятельные работы, вопросы к зачету, контрольные работы, устные ответы         | Преимущественно проверка методами, определяющимися исходя из состояния обучающегося на момент проверки |

#### 9.4. Методические рекомендации по оценочным материалам для лиц с ограниченными возможностями здоровья и инвалидов

Для лиц с ограниченными возможностями здоровья и инвалидов предусматривается доступная форма предоставления заданий оценочных средств, а именно:

- в печатной форме;
- в печатной форме с увеличенным шрифтом;
- в форме электронного документа;
- методом чтения ассистентом задания вслух;
- предоставление задания с использованием сурдоперевода.

Лицам с ограниченными возможностями здоровья и инвалидам увеличивается время на подготовку ответов на контрольные вопросы. Для таких обучающихся предусматривается доступная форма предоставления ответов на задания, а именно:

- письменно на бумаге;
- набор ответов на компьютере;
- набор ответов с использованием услуг ассистента;
- представление ответов устно.

Процедура оценивания результатов обучения лиц с ограниченными возможностями здоровья и инвалидов по дисциплине предусматривает предоставление информации в формах, адаптированных к ограничениям их здоровья и восприятия информации:

**Для лиц с нарушениями зрения:**

- в форме электронного документа;
- в печатной форме увеличенным шрифтом.

**Для лиц с нарушениями слуха:**

- в форме электронного документа;
- в печатной форме.

**Для лиц с нарушениями опорно-двигательного аппарата:**

- в форме электронного документа;
- в печатной форме.

При необходимости для лиц с ограниченными возможностями здоровья и инвалидов процедура оценивания результатов обучения может проводиться в несколько этапов.



## ЛИСТ СОГЛАСОВАНИЯ

Рассмотрена и одобрена на заседании кафедры АОИ  
протокол № 13 от «15» 12 2022 г.

### СОГЛАСОВАНО:

| Должность                           | Инициалы, фамилия | Подпись  |
|-------------------------------------|-------------------|--|
| Заведующий выпускающей каф. АОИ     | А.А. Сидоров      | Согласовано,<br>d4895b45-5d88-49f8-<br>b7c7-e8bf0196776a |
| Заведующий обеспечивающей каф. АОИ  | А.А. Сидоров      | Согласовано,<br>d4895b45-5d88-49f8-<br>b7c7-e8bf0196776a |
| И.О. начальника учебного управления | И.А. Лариошина    | Согласовано,<br>c3195437-a02f-4972-<br>a7c6-ab6ee1f21e73 |

### ЭКСПЕРТЫ:

|                               |              |  |
|-------------------------------|--------------|--|
| Доцент, каф. АОИ              | Н.Ю. Салмина | Согласовано,<br>ed28a52c-a209-461c-<br>b4ed-4e958affbfc7 |
| Заведующий кафедрой, каф. АОИ | А.А. Сидоров | Согласовано,<br>d4895b45-5d88-49f8-<br>b7c7-e8bf0196776a |

### РАЗРАБОТАНО:

|                                 |                |  |
|---------------------------------|----------------|--|
| Старший преподаватель, каф. АОИ | Л.И. Синчинова | Разработано,<br>90a7608e-274c-45a6-<br>b9cf-2c55c524e3f0 |
| Доцент, каф. АОИ                | Е.Ю. Агеев     | Разработано,<br>1380771b-dd3c-4ac1-<br>8e1d-30fb96b5fa40 |