




МИНОБРНАУКИ РОССИИ  
Федеральное государственное бюджетное образовательное учреждение  
высшего образования  
«Волгоградский государственный технический университет»

Факультет архитектуры и градостроительного развития

УТВЕРЖДЕНО

Факультет архитектуры и градостроительного  
развития

Деканом  Назарова Марина Петровна  
02.11.2023 г.

## Методы и технологии обработки больших данных

### рабочая программа дисциплины (модуля, практики)

Закреплена за кафедрой **Цифровые технологии в урбанистике, архитектуре и строительстве**

Учебный план 09.04.02 Информационные системы и технологии

Магистерская программа **Цифровые технологии в урбанистике, архитектуре и строительстве**

Квалификация **магистр**

Срок обучения **2 года**

Форма обучения **очная**

Виды контроля в семестрах: **зачет, экзамен**

| Семестр(Курс.Номер семестра на курсе) | 1(1.1) |       | 2(1.2) |       | Итого  |        |
|---------------------------------------|--------|-------|--------|-------|--------|--------|
|                                       | УП     | ПП    | УП     | ПП    | УП     | ПП     |
| Лекции                                | 8      | 8     | 12     | 12    | 20     | 20     |
| Лабораторные                          | 12     | 12    | 24     | 24    | 36     | 36     |
| Итого ауд.                            | 20     | 20    | 36     | 36    | 56     | 56     |
| Контактная работа                     | 20,25  | 20,25 | 36,35  | 36,35 | 56,6   | 56,6   |
| Сам. работа                           | 87,75  | 87,75 | 72     | 72    | 159,75 | 159,75 |
| Часы на контроль                      | 0      | 0     | 35,65  | 35,65 | 35,65  | 35,65  |
| Практическая подготовка               | 0      | 0     | 0      | 0     | 0      | 0      |
| Итого трудоемкость в часах            | 108    | 108   | 144    | 144   | 0      | 0      |

## ЛИСТ ОДОБРЕНИЯ И СОГЛАСОВАНИЯ РАБОЧЕЙ ПРОГРАММЫ

Разработчик(и) программы:

профессор Садовникова Наталья Петровна д.т.н.



Рецензент(ы):  
(при наличии)

Рабочая программа дисциплины (модуля, практики)

**Методы и технологии обработки больших данных**

Разработана в соответствии с ФГОС ВО:

Федеральный государственный образовательный стандарт высшего образования - магистратура по направлению подготовки 09.04.02 Информационные системы и технологии (приказ Минобрнауки России от 19.09.2017 г. № 917)

Составлена на основании учебного плана:

09.04.02 Информационные системы и технологии

Магистерская программа: Цифровые технологии в урбанистике, архитектуре и строительстве

Утвержденного учёным советом вуза от 31.01.2024 г. протокол №6.

Рабочая программа одобрена на заседании кафедры

**Цифровые технологии в урбанистике, архитектуре и строительстве**

Протокол №2 от 12.10.2023 г.

Зав. кафедрой Парыгин Данила Сергеевич

СОГЛАСОВАНО НМС ВОЛГГТУ:

Протокол заседания НМС от 20.12.2023 г. №4

| <b>1. ЦЕЛИ И ЗАДАЧИ ОСВОЕНИЯ ДИСЦИПЛИНЫ (МОДУЛЯ, ПРАКТИКИ). ВИД, ТИП ПРАКТИКИ, СПОСОБ И ФОРМА (ФОРМЫ) ЕЕ ПРОВЕДЕНИЯ.</b>  |  |
|---|--|
| Целью освоения дисциплины является формирование у студентов профессиональной компетенции в области разработки и использования систем обработки и анализа больших массивов данных, изучению теоретических основ построения и функционирования подобных систем. |  |
| Задачи:   |  |
| - Разработка методик автоматизации обработки больших данных.  |  |
| - Изучить теоретические основы построения систем обработки и анализа больших данных.  |  |
| - Изучить основные языковые и визуальные способы организации, отображения и манипулирования данными под управлением систем обработки и анализа больших данных.  |  |
| - Изучить теоретические основы построения и функционирования подобных систем  |  |

| <b>2. МЕСТО ДИСЦИПЛИНЫ (МОДУЛЯ, ПРАКТИКИ) В СТРУКТУРЕ ОБРАЗОВАТЕЛЬНОЙ ПРОГРАММЫ</b>                 |  |
|---|--|
| Цикл (раздел) ОП:   | Б1.О   |
| <b>2.1</b>  | <b>Требования к предварительной подготовке обучающегося:</b>   |
| 2.1.1   | Изучение дисциплины базируется на результатах обучения, полученных на предыдущем уровне (уровень бакалавриата, специалитета), в рамках таких дисциплин как: технологии обработки информации, анализ больших данных и др. |
| <b>2.2</b>  | <b>Дисциплины (модули) и практики, для которых освоение данной дисциплины (модуля) необходимо как предшествующее:</b>  |
| 2.2.1   | Прикладные информационные технологии в строительстве и архитектуре   |
| 2.2.2   | Применение AR и VR технологий в урбанистике, архитектуре и строительстве   |
| 2.2.3   | Разработка и проектирование интеллектуальных систем  |
| 2.2.4   | Управление проектами разработки систем искусственного интеллекта   |
| <b>3. КОМПЕТЕНЦИИ ОБУЧАЮЩЕГОСЯ, ФОРМИРУЕМЫЕ В РЕЗУЛЬТАТЕ ОСВОЕНИЯ ДИСЦИПЛИНЫ (МОДУЛЯ, ПРАКТИКИ)</b> |  |

**ОПК-6: Способен использовать методы и средства системной инженерии в области получения, передачи, хранения, переработки и представления информации посредством информационных технологий;**

*ОПК-6.1: Знать: основные положения системной инженерии в области получения, передачи, хранения, переработки и представления информации посредством информационных технологий*

Результаты обучения: знать принципы и методы обработки больших данных, алгоритмы и структуры данных, анализ данных, машинное обучение, возможности высокопроизводительных вычислительных систем, технологии распределенных вычислений, методы и модели Data Mining.

*ОПК-6.2: Уметь: применять методы и средства системной инженерии в области получения, передачи, хранения, переработки и представления информации посредством информационных технологий*

Результаты обучения: уметь применять методы обработки больших данных, анализа, проектировать и оптимизировать системы обработки больших данных, использовать оптимальное распределение ресурсов, обеспечивать реализацию масштабируемых архитектур, разрабатывать и анализировать концептуальные и теоретические модели прикладных задач анализа больших данных, использовать и применять углубленные знания в области обработки и анализа больших данных, оценивать время и необходимые аппаратные ресурсы для решения задач анализа и обработки данных, создавать алгоритмы анализа и обработки большого объема данных с применением моделей Data Mining.

*ОПК-6.3: Иметь навыки: применения методов и средств системной инженерии в области получения, передачи, хранения, переработки и представления информации посредством информационных технологий*

Результаты обучения: иметь навыки применения программных систем, предназначенных для анализа больших данных.

| <b>4. СТРУКТУРА И СОДЕРЖАНИЕ ДИСЦИПЛИНЫ (МОДУЛЯ, ПРАКТИКИ)</b> |  |                |       |                |
|--|--|----------------|-------|----------------|
| Код занятия  | Наименование разделов и тем /вид занятия/  | Семестр / Курс | Часов | Форма контроля |
| 1  | <b>Раздел 1. Обучение (1 семестр)</b>  |                |       |                |
| 1.1  | Математические основы анализа данных. /Тема/   | 1              | 0     |                |
| 1.1.1  | Применение различных математических методов для анализа сложности алгоритмов. /Лек/    | 1              | 1     | Ко, К, З       |
| 1.1.2  | Лабораторная №1. Изучение основ анализа данных /Лаб/                                   | 1              | 4     | Ко, К, З       |
| 1.1.3  | Подготовка к текущему контролю успеваемости (Лаб, Ко, К). /Ср/                         | 1              | 22    | Ко, К, З       |
| 1.2  | Технологии анализа данных. /Тема/  | 1              | 0     |                |
| 1.2.1  | Процесс анализа. Извлечение и визуализация данных. Технологии KDD и Data Mining. /Лек/ | 1              | 2     | Ко, К, З       |

|       |   |   |       |           |
|-------|---|---|-------|-----------|
| 1.2.2 | Подготовка к текущему контролю успеваемости (Лаб, Ко, К). /Ср/  | 1 | 12    | Ко, К, З  |
| 1.2.3 | Лабораторная работа № 2. Применение технологии и методы анализа   | 1 | 2     | Ко, К, З  |
| 1.3   | Обзор Big-Data. Особенности Big-Data. /Тема/  | 1 | 0     |           |
| 1.3.1 | Обзор Big-Data. Методы и средства. Используемые программы. Особенности Big-Data. /Лек/  | 1 | 2     | Ко, К, З  |
| 1.3.2 | Подготовка к текущему контролю успеваемости (Лаб, Ко, К). /Ср/  | 1 | 15    | Ко, К, З  |
| 1.3.3 | Лабораторная работа № 3. Реализация Big-Data /Лаб/  | 1 | 2     | Ко, К, З  |
| 1.4   | SAS Interprise Miner. /Тема/  | 1 | 0     |           |
| 1.4.1 | Принцип анализа данных SEMMA. Основные инструменты и узлы. Возможности построения моделей. /Лек/  | 1 | 2     | Ко, К, З  |
| 1.4.2 | Подготовка к текущему контролю успеваемости (Лаб, Ко, К). /Ср/  | 1 | 12    | Ко, К, З  |
| 1.4.3 | Лабораторная работа № 4. Применение SAS Interprise Miner /Лаб/  | 1 | 2     | Ко, К, З  |
| 1.5   | Определение источника данных. Исследование источника данных. /Тема/   | 1 | 0     |           |
| 1.5.1 | Библиотеки и диаграмм SAS. Настройки источника данных. /Лек/  | 1 | 1     | Ко, К, З  |
| 1.5.2 | Лабораторная работа № 5. Работа с источниками данных /Лаб/  | 1 | 2     | Ко, К, З  |
| 1.5.3 | Подготовка к текущему контролю успеваемости (Лаб, Ко, К). /Ср/  | 1 | 18    | Ко, К, З  |
| 1.6   | Подготовка к промежуточной аттестации (зачет) /Ср/  | 1 | 8,75  | З         |
| 2     | <b>Раздел 2. Промежуточная аттестация (1 семестр)</b>   |   |       |           |
| 2.1   | Зачет /Тема/  | 1 | 0     |           |
| 2.1.1 | Контактная работа с ППС /КоПа/  | 1 | 0,25  |           |
| 3     | <b>Раздел 3. Обучение (2 семестр)</b>   |   |       |           |
| 3.1   | Прогнозное моделирование. /Тема/  | 2 | 0     |           |
| 3.1.1 | Область прикладных задач с использованием прогнозного моделирования. /Лек/  | 2 | 4     | Ко, К, Эк |
| 3.1.2 | Подготовка к текущему контролю успеваемости (Лаб, Ко, К). /Ср/  | 2 | 10    | Ко, К, Эк |
| 3.1.3 | Лабораторная № 1. Использование прогнозного моделирования /Лаб/   | 2 | 4     | Ко, Эк    |
| 3.2   | Прогнозная модель, использующая дерево решений. /Тема/  | 2 | 0     |           |
| 3.2.1 | Создание дерева решений: структура. Алгоритм построения. /Лек/  | 2 | 2     | Ко, К, Эк |
| 3.2.2 | Лабораторная работа №2. Реализация прогнозной модели, использующей дерево решений /Лаб/   | 2 | 4     | Ко, К, Эк |
| 3.2.3 | Подготовка к текущему контролю успеваемости (Лаб, Ко, К). /Ср/  | 2 | 8     | Ко, К, Эк |
| 3.3   | Прогнозное моделирование: работа с регрессионными моделями. /Тема/  | 2 | 0     |           |
| 3.3.1 | Регрессия. Оценка параметров. Обработка пропущенных значений. Выбор входных переменных. Оптимизация сложности. Интерпретация регрессии. /Лек/ | 2 | 2     | Ко, К, Эк |
| 3.3.2 | Подготовка к текущему контролю успеваемости (Лаб, Ко, К). /Ср/  | 2 | 14    | Ко, К, Эк |
| 3.3.3 | Лабораторная работа № 3 Реализация прогнозной модели, использующей регрессию/Лаб/   | 2 | 4     | Ко, К, Эк |
| 3.4   | Прогнозное моделирование: нейронные сети. /Тема/  | 2 | 0     |           |
| 3.4.1 | Особенности нейронных сетей. Обучение нейронной сети. Инструмент AutoNeural. /Лек/  | 2 | 2     | Ко, К, Эк |
| 3.4.2 | Подготовка к текущему контролю успеваемости (Лаб, Ко, К). /Ср/  | 2 | 28    | Ко, К, Эк |
| 3.4.3 | Лабораторная работа № 4. Обучение нейронных сетей /Лаб/   | 2 | 8     | Ко, К, Эк |
| 3.5   | Ассоциативные правила. Кластеризация. /Тема/  | 2 | 0     |           |
| 3.5.1 | Ассоциативные правила. Поиск ассоциативных правил. Кластеризация. Алгоритм кластеризации k-means. /Лек/                                       | 2 | 2     | Ко, К, Эк |
| 3.5.2 | Лабораторная работа № 5. Реализация ассоциативных правил /Лаб/  | 2 | 4     | Ко, К, Эк |
| 3.5.3 | Подготовка к текущему контролю успеваемости (Лаб, Ко, К). /Ср/  | 2 | 12    | Ко, К, Эк |
| 3.6   | Подготовка к промежуточной аттестации /Ср/  | 2 | 35,65 | Эк        |
| 4     | <b>Раздел 4. Промежуточная аттестация (2 семестр)</b>   |   |       |           |
| 4.1   | Экзамен /Тема/  | 2 | 0     |           |
| 4.1.1 | Контактная работа с ППС /КоПа/  | 2 | 0,35  |           |

Примечание. Формы контроля: Эк – экзамен, К- контрольная работа, Ко- контрольный опрос, Сз- семестровое задание, З-зачет, ОП- отчет по практике.

## 5. ФОНД ОЦЕНОЧНЫХ СРЕДСТВ

Оценочные средства планируемых результатов обучения представлены в виде фондов оценочных средств (ФОС),

разработанных в соответствии с локальным нормативным актом университета. ФОС может быть представлен в Приложении к рабочей программе.

Перечень компетенций, формируемых в результате освоения дисциплины и показатели их оценивания:

ОПК-6: Способен использовать методы и средства системной инженерии в области получения, передачи, хранения, переработки и представления информации посредством информационных технологий

ОПК-6.1: Знать: основные положения системной инженерии в области получения, передачи, хранения, переработки и представления информации посредством информационных технологий. Раздел 1.1 -1.5, 3.1-3.5 содержания дисциплины.

Оценочные средства: контрольный опрос, контрольная работа, зачет, экзамен.

ОПК-6.2: Уметь: применять методы и средства системной инженерии в области получения, передачи, хранения, переработки и представления информации посредством информационных технологий. Раздел 1.1 -1.5, 3.1-3.5 содержания дисциплины.

Оценочные средства: контрольный опрос, контрольная работа, зачет, экзамен.

ОПК-6.3: Иметь навыки: применения методов и средств системной инженерии в области получения, передачи, хранения, переработки и представления информации посредством информационных технологий. Раздел 1.1-1.5, 3.1-3.5 содержания дисциплины. Оценочные средства: контрольный опрос, контрольная работа, зачет, экзамен.

Оценочными средствами по дисциплине являются: контрольная работа, контрольный опрос, зачет, экзамен.

Контрольный опрос – средство контроля, организованное как специальная беседа преподавателя с обучающимся на темы, связанные с изучаемой дисциплиной, и рассчитанное на выяснение объема знаний обучающегося по дисциплине.

Контрольная работа – средство проверки умений применять полученные знания для решения задач определенного типа по теме, разделу или дисциплине.

Зачет – средство контроля для оценки окончательных результатов обучения по дисциплине.

Экзамен – средство контроля для оценки окончательных результатов обучения по дисциплине.

Критерии шкала оценивания по оценочному средству «Контрольный опрос»

9-10 Дан полный, развернутый ответ на поставленный вопрос, показана совокупность осознанных знаний об объекте, проявляющаяся в свободном оперировании понятиями, умении выделить существенные и несущественные его признаки, причинно-следственные связи. Ответ формулируется в терминах науки, изложен литературным языком, логичен, доказателен, демонстрирует авторскую позицию студента.

7-8 Дан полный, развернутый ответ на поставленный вопрос, показана совокупность осознанных знаний об объекте, доказательно раскрыты основные положения темы; в ответе прослеживается четкая структура, логическая последовательность, отражающая сущность раскрываемых понятий, теорий, явлений. Ответ изложен литературным языком в терминах науки. Могут быть допущены недочеты в определении понятий, исправленные студентом самостоятельно в процессе ответа

5-6 Дан полный, но недостаточно последовательный ответ на поставленный вопрос, но при этом показано умение выделить существенные и несущественные признаки и причинно-следственные связи. Ответ логичен и изложен в терминах науки. Могут быть допущены 2-3 ошибки в определении основных понятий, которые студент затрудняется исправить самостоятельно

0-4 Дан неполный ответ, представляющий собой разрозненные знания по теме вопроса с существенными ошибками в определениях. Присутствуют фрагментарность, нелогичность изложения. Студент не осознает связь данного понятия, теории, явления с другими объектами дисциплины. Отсутствуют выводы, конкретизация и доказательность изложения. Речь неграмотная. Дополнительные и уточняющие вопросы преподавателя не приводят к коррекции ответа студента не только на поставленный вопрос, но и на другие вопросы дисциплины.

Критерии и шкала оценивания по оценочному средству «Контрольная работа»

8-10 Контрольная работа (семестровая) выполнена на высоком уровне (правильные ответы даны на 90-100% вопросов/задач)

6-7 Контрольная работа (семестровая) выполнена на среднем уровне (правильные ответы даны на 70-89% вопросов/задач)

4-5 Контрольная работа (семестровая) выполнена на низком уровне (правильные ответы даны на 50-69% вопросов/задач)

0-3 Контрольная работа (семестровая) выполнена на неудовлетворительном уровне (правильные ответы даны менее чем 50%) или не выполнена.

Критерии и шкала оценивания по оценочным средствам «Зачет» и «Экзамен»

35-40 Продемонстрировано уверенное владение материалом теоретического курса, студент разбирается в категориальном аппарате дисциплины, владеет научными терминами, может их использовать в ответе, показывает глубину знаний, критерий отражен полностью (ответы на 90-100% правильные).

27-34 Продемонстрировано хорошее владение материалом теоретического курса, студент разбирается в категориальном аппарате дисциплины, владеет научными терминами, может их использовать в ответе, показывает глубину знаний, критерий отражен частично (ответы на 70-89 % правильные).

20-26 Продемонстрировано удовлетворительное владение материалом теоретического курса, студент имеет общее представление о тематике дисциплины, умеет различать и пояснять основные понятия. В категориальном аппарате дисциплины разбирается слабо (ответы на 50 – 69 % правильные).

менее 20 Ответ студента демонстрирует слабые знания теоретического курса или полное их отсутствие (ответы правильные менее, чем на 50 %).

1-й семестр

Контрольный опрос проводится пять раз в течение семестра на лабораторных работах.

Примеры тем по оценочному средству «Контрольный опрос»

Вопросы для лабораторной работы №1:

1. Асимптотические обозначения и сравнение скоростей роста функций. (ОПК-6.1)

2. Стандартные функции и обозначения. (ОПК-6.1)

3. Суммирование и оценки сумм. (ОПК-6.1)

4. Применение различных математических методов для анализа сложности алгоритмов. (ОПК-6.1, ОПК-6.2)

5. Рекуррентные соотношения. (ОПК-6.1)

6. Метод подстановки. (ОПК-6.1)
7. Метод итераций. (ОПК-6.1)
8. Общая теорема о рекуррентных соотношениях и примеры её применения. (ОПК-6.1)

Вопросы для лабораторной работы №2:

1. Процесс анализа. (ОПК-6.1)
2. Общая схема анализа. (ОПК-6.1)
3. Извлечение и визуализация данных. (ОПК-6.1, ОПК-6.2)
4. Этапы моделирования. (ОПК-6.1)
5. Процесс построения моделей. (ОПК-6.1, ОПК-6.2)
6. Формы представления данных, типы и виды данных. (ОПК-6.1)
7. Представления наборов данных. (ОПК-6.1)
8. Технологии KDD и Data Mining. (ОПК-6.2, ОПК-6.3)
9. Подготовка данных к анализу. (ОПК-6.2)
10. Методика извлечения знаний. (ОПК-6.2)
11. Data Mining. (ОПК-6.2, ОПК-6.3)
12. Мультидисциплинарный характер Data Mining. (ОПК-6.1)
13. Причины распространения KDD и Data Mining. (ОПК-6.1)
14. Актуальность технологий Data Mining как средств обработки больших объемов информации. (ОПК-6.1)

Вопросы для лабораторной работы №3:

1. Обзор Big-Data. (ОПК-6.1)
2. Методы и средства. (ОПК-6.1)
3. Используемые программы. (ОПК-6.2)
4. Особенности (ОПК-6.1)
5. Большие данные (Big Data): современные подходы к обработке и хранению. (ОПК-6.1)
6. Проблема множественного сравнения данных. (ОПК-6.1)

Вопросы для лабораторной работы №4:

1. Стандарт SEMMA. (ОПК-6.1)
2. Принцип анализа данных SEMMA. (ОПК-6.1)
3. Основные инструменты и узлы. (ОПК-6.2)
4. Возможности построения моделей (ОПК-6.1)
5. Построение математической модели ИИ. (ОПК-6.2)
6. Методы оценки качества алгоритмов машинного обучения. (ОПК-6.2, ОПК-6.3)

Вопросы для лабораторной работы №5:

1. Создание проекта, библиотеки и диаграмм SAS. (ОПК-6.2, ОПК-6.3)
2. Настройки источника данных. (ОПК-6.2, ОПК-6.3)
3. Типы переменных. (ОПК-6.1)
4. Изменение размера выборки. (ОПК-6.2)
5. Создание диаграмм. (ОПК-6.2)
6. Исследование взаимосвязей между переменными. (ОПК-6.3)

Оценочное средство "Контрольная работа".

Контрольная работа реализуется в виде семестровой работы.

Список примерных тем:

1. Использование Big Data в городском планировании и развитии инфраструктуры.
2. Сравнительный анализ методов анонимизации данных для обеспечения конфиденциальности в городских Big Data.
3. Исследование влияния Big Data на повышение эффективности городского управления и принятия решений.
4. Разработка системы умного города на основе сбора и анализа метрик потребления энергии.
5. Применение машинного обучения для прогнозирования погоды в городе на основе крупных данных.

Структура отчета семестровой работы

1. Титульный лист.
2. Оглавление.
3. Введение
4. Основная часть работы.
5. Заключение.
6. Список использованных источников.

Если при реализации семестровой работы разрабатывался программный продукт, то прикладывается архив с исходным кодом программы.

Протокол семестровой работы и архив с исходным кодом (при наличии) загружаются в ЭИОС.

Оформление протокола семестровой работы

Семестровая работа должна быть представлена: на листах формата А4; редактор Word; межстрочный интервал – полуторный; основной шрифт – Times New Roman 14пт. Общй объем работы не должен превышать 25 страниц.

Текст семестровой работы следует располагать, соблюдая следующие размеры полей: левое - 30 мм; правое - 15 мм; верхнее - 20 мм; нижнее - 20 мм. Страницы семестровой работы нумеруются арабскими цифрами. Титульный лист включают в общую нумерацию работы, но номер на нем не ставится, на последующих страницах номер проставляется внизу по середине страницы без точек. Каждый раздел семестровой работы рекомендуется начинать с нового листа (страницы).

Подразделы нумеруются арабскими цифрами в пределах каждого раздела. Номер подраздела состоит из номера раздела и подраздела, разделенных точкой, например: "3.2" (второй подраздел третьего раздела).

Изучение дисциплины в первом семестре заканчивается сдачей студентом зачета. Зачет проводится устно в виде собеседования по вопросам, составленным на основе вопросов к разделам изучаемой дисциплины. Зачет по дисциплине может проводиться в одной из двух форм – очной или дистанционной. Зачет в дистанционной форме проводится с помощью электронной информационной образовательной среде (ЭИОС).

Независимо от формы проведения, зачет включает предварительную часть и окончательное собеседование: при проведении зачета студенту выдается 2 вопроса из приведенного ниже перечня. На протяжении 30 минут студент кратко (конспективно) излагает в письменной форме ответы на вопросы. После написания ответа состоится собеседование, в ходе которого преподаватель уточняет отдельные элементы ответа и делает вывод о степени сформированности компетенций студента.

Примерные вопросы, выносимые на зачет:

1. Асимптотические обозначения и сравнение скоростей роста функций. (ОПК-6.1)
2. Стандартные функции и обозначения. (ОПК-6.1)
3. Суммирование и оценки сумм. (ОПК-6.1)
4. Применение различных математических методов для анализа сложности алгоритмов. (ОПК-6.1, ОПК-6.2)
5. Рекуррентные соотношения. (ОПК-6.1)
6. Метод подстановки. (ОПК-6.1)
7. Метод итераций. (ОПК-6.1)
8. Общая теорема о рекуррентных соотношениях и примеры её применения. (ОПК-6.1)
9. Процесс анализа. (ОПК-6.1)
10. Общая схема анализа. (ОПК-6.1)
11. Извлечение и визуализация данных. (ОПК-6.1, ОПК-6.2)
12. Этапы моделирования. (ОПК-6.1)
13. Процесс построения моделей. (ОПК-6.1, ОПК-6.2)
14. Формы представления данных, типы и виды данных. (ОПК-6.1)
15. Представления наборов данных. (ОПК-6.1)
16. Технологии KDD и Data Mining. (ОПК-6.2, ОПК-6.3)
17. Подготовка данных к анализу. (ОПК-6.2)
18. Методика извлечения знаний. (ОПК-6.2)
19. Data Mining. (ОПК-6.2, ОПК-6.3)
20. Мультидисциплинарный характер Data Mining. (ОПК-6.1)
21. Причины распространения KDD и Data Mining. (ОПК-6.1)
22. Актуальность технологий Data Mining как средств обработки больших объемов информации (ОПК-6.1)
23. Обзор Big-Data. (ОПК-6.1)
24. Методы и средства. (ОПК-6.1)
25. Используемые программы. (ОПК-6.2)
26. Особенности используемых программ для работы с Big-Data (ОПК-6.1)
27. Большие данные (Big Data): современные подходы к обработке и хранению. (ОПК-6.1)
28. Проблема множественного сравнения данных. (ОПК-6.1)
29. Стандарт SEMMA. (ОПК-6.1)
30. Принцип анализа данных SEMMA. (ОПК-6.1)
31. Основные инструменты и узлы. (ОПК-6.2)
32. Возможности построения моделей (ОПК-6.1)
33. Построение математической модели ИИ (ОПК-6.2)
34. Методы оценки качества алгоритмов машинного обучения. (ОПК-6.2, ОПК-6.3)
35. Создание проекта, библиотеки и диаграмм SAS. (ОПК-6.2, ОПК-6.3)
36. Настройки источника данных. (ОПК-6.2, ОПК-6.3)
37. Типы переменных. (ОПК-6.1)
38. Изменение размера выборки. (ОПК-6.2)
39. Создание диаграмм. (ОПК-6.2)
40. Исследование взаимосвязей между переменными. (ОПК-6.3)

2-й семестр

Контрольный опрос проводится пять раз в течение семестра на лабораторных работах.

Примеры тем по оценочному средству «Контрольный опрос»

Вопросы для лабораторной работы №1:

1. Область прикладных задач с использованием прогнозного моделирования. (ОПК-6.1)
2. Проклятие размерности. (ОПК-6.1)
3. Выбор входных переменных. (ОПК-6.1, ОПК-6.3)
4. Избавление от бесполезных и избыточных входных переменных. (ОПК-6.1, ОПК-6.3)
5. Создание обучающих и проверочных данных. (ОПК-6.2, ОПК-6.3)

Вопросы для лабораторной работы №2:

1. Создание дерева решений: структура. (ОПК-6.1, ОПК-6.2)
2. Алгоритм построения. (ОПК-6.1)
3. Поиск разбиений. (ОПК-6.1)
4. Прогнозная модель использующая дерево решений: построение (ОПК-6.1, ОПК-6.2)

5. Прогнозная модель использующая дерево решений: создание правила разбиения. (ОПК-6.2, ОПК-6.3)

6. Оптимизация сложности деревьев решений. (ОПК-6.2, ОПК-6.3)

7. Оценка качества дерева решений (ОПК-6.3)

Вопросы для лабораторной работы №3:

1. Регрессия. (ОПК-6.1)

2. Логистическая регрессия. (ОПК-6.1)

3. Полиномиальные регрессии. (ОПК-6.1)

4. Оценка параметров. (ОПК-6.2)

5. Обработка пропущенных значений. (ОПК-6.2)

6. Оптимизация сложности. (ОПК-6.2)

7. Интерпретация регрессии. (ОПК-6.1)

8. Регрессии с преобразованными входными переменными. (ОПК-6.1)

9. Категориальные переменные в регрессионной модели. (ОПК-6.1)

Вопросы для лабораторной работы №4:

1. Основные понятия теории нейронных сетей. (ОПК-6.1)

2. Основные парадигмы нейронных сетей. (ОПК-6.1)

3. Особенности нейронных сетей. (ОПК-6.1)

4. Многослойный персептрон: класс решаемых задач, архитектура. (ОПК-6.1)

5. Определение дерева решений. (ОПК-6.1, ОПК-6.2)

6. Причины популярности и условия применимости. (ОПК-6.1, ОПК-6.3)

7. Структура дерева решений. (ОПК-6.1)

8. Выбор атрибута разбиения в узле. (ОПК-6.1, ОПК-6.3)

Вопросы для лабораторной работы №5:

1. Ассоциативные правила. (ОПК-6.1)

2. Аффинитивный анализ, предметный набор. (ОПК-6.1)

3. Поддержка и достоверность ассоциативного правила. (ОПК-6.1)

4. Значимость ассоциативных правил, лифт и левередж. (ОПК-6.1)

5. Поиск ассоциативных правил. (ОПК-6.2)

6. Частые предметные наборы и их обнаружение. (ОПК-6.2)

7. Алгоритм генерации ассоциативных правил. (ОПК-6.1)

8. Иерархические ассоциативные правила. (ОПК-6.1)

9. Методы поиска иерархических ассоциативных правил. (ОПК-6.1)

10. Определение кластеризации. (ОПК-6.1)

11. Постановка задачи кластеризации. (ОПК-6.2, ОПК-6.3)

12. Цели кластеризации в Data Mining. (ОПК-6.3)

13. Примеры кластеризации в различных областях. (ОПК-6.1)

14. Виды метрик. (ОПК-6.1)

15. Шаги алгоритма. (ОПК-6.1)

16. Меры расстояний. (ОПК-6.1)

17. Пример работы алгоритма k-means. (ОПК-6.1)

18. Проблемы алгоритмов кластеризации. (ОПК-6.3)

Оценочное средство "Контрольная работа".

Контрольная работа реализуется в виде семестровой работы.

Список примерных тем:

1. Анализ данных о транспортных потоках для оптимизации городского движения.

2. Оценка эффективности и влияния городских мобильных приложений на поведение жителей.

3. Анализ данных социальных медиа для мониторинга общественного мнения в городской среде.

4. Создание прототипа системы дистанционного мониторинга экологических показателей в городе с использованием IoT.

5. Применение глубокого обучения для обнаружения и анализа аномалий в городских сенсорных сетях.

Структура отчета семестровой работы

1. Титульный лист.

2. Оглавление.

3. Введение

4. Основная часть работы.

5. Заключение.

6. Список использованных источников.

Если при реализации семестровой работы разрабатывался программный продукт, то прикладывается архив с исходным кодом программы.

Протокол семестровой работы и архив с исходным кодом (при наличии) загружаются в ЭИОС.

Оформление протокола семестровой работы

Семестровая работа должна быть представлена: на листах формата А4; редактор Word; межстрочный интервал – полуторный; основной шрифт – Times New Roman 14пт. Общий объем работы не должен превышать 25 страниц.

Текст семестровой работы следует располагать, соблюдая следующие размеры полей: левое - 30 мм; правое - 15 мм; верхнее - 20 мм; нижнее - 20 мм. Страницы семестровой работы нумеруются арабскими цифрами. Титульный лист включают в общую нумерацию работы, но номер на нем не ставится, на последующих страницах номер проставляется внизу по середине страницы без точек.

Каждый раздел семестровой работы рекомендуется начинать с нового листа (страницы).

Подразделы нумеруются арабскими цифрами в пределах каждого раздела. Номер подраздела состоит из номера раздела и подраздела, разделенных точкой, например: "3.2" (второй подраздел третьего раздела).



Изучение дисциплины во втором семестре заканчивается сдачей студентом экзамена. Экзамен проводится устно в виде собеседования по вопросам, составленным на основе вопросов к разделам изучаемой дисциплины. Экзамен по дисциплине может проводиться в одной из двух форм – очной или дистанционной. Экзамен в дистанционной форме проводится с помощью электронной информационной образовательной среде (ЭИОС). Независимо от формы проведения, экзамен включает предварительную часть и окончательное собеседование: при проведении экзамена студенту выдаётся 2 вопроса из приведённого ниже перечня. На протяжении 30 минут студент кратко (конспективно) излагает в письменной форме ответы на вопросы. После написания ответа состоится собеседование, в ходе которого преподаватель уточняет отдельные элементы ответа и делает вывод о степени сформированности компетенций студента.

Примерный перечень вопросов для подготовки к экзамену:

1. Область прикладных задач с использованием прогнозного моделирования. (ОПК-6.1)
2. Проклятие размерности. (ОПК-6.1)
3. Выбор входных переменных. (ОПК-6.1, ОПК-6.3)
4. Избавление от бесполезных и избыточных входных переменных. (ОПК-6.1, ОПК-6.3)
5. Создание обучающих и проверочных данных. (ОПК-6.2, ОПК-6.3)
6. Создание дерева решений: структура. (ОПК-6.1, ОПК-6.2)
7. Алгоритм построения. (ОПК-6.1)
8. Поиск разбиений. (ОПК-6.1)
9. Прогнозная модель использующая дерево решений: построение (ОПК-6.1, ОПК-6.2)
10. Прогнозная модель использующая дерево решений: создание правила разбиения. (ОПК-6.2, ОПК-6.3)
11. Оптимизация сложности деревьев решений. (ОПК-6.2, ОПК-6.3)
12. Оценка качества дерева решений. (ОПК-6.3)
13. Регрессия. (ОПК-6.1)
14. Логистическая регрессия. (ОПК-6.1)
15. Полиномиальные регрессии. (ОПК-6.1)
16. Оценка параметров. (ОПК-6.2)
17. Обработка пропущенных значений. (ОПК-6.2)
18. Оптимизация сложности. (ОПК-6.2)
19. Интерпретация регрессии. (ОПК-6.1)
20. Регрессии с преобразованными входными переменными. (ОПК-6.1)
21. Категориальные переменные в регрессионной модели. (ОПК-6.1)
22. Основные понятия теории нейронных сетей. (ОПК-6.1)
23. Основные парадигмы нейронных сетей. (ОПК-6.1)
24. Особенности нейронных сетей. (ОПК-6.1)
25. Многослойный персептрон: класс решаемых задач, архитектура. (ОПК-6.1)
26. Определение дерева решений. (ОПК-6.1, ОПК-6.2)
27. Причины популярности и условия применимости. (ОПК-6.1, ОПК-6.3)
28. Структура дерева решений. (ОПК-6.1)
29. Выбор атрибута разбиения в узле. (ОПК-6.1, ОПК-6.3)
30. Ассоциативные правила. (ОПК-6.1)
31. Аффинитивный анализ, предметный набор. (ОПК-6.1)
32. Поддержка и достоверность ассоциативного правила. (ОПК-6.1)
33. Значимость ассоциативных правил, лифт и леввередж. (ОПК-6.1)
34. Поиск ассоциативных правил. (ОПК-6.2)
35. Частые предметные наборы и их обнаружение. (ОПК-6.2)
36. Алгоритм генерации ассоциативных правил. (ОПК-6.1)
37. Иерархические ассоциативные правила. (ОПК-6.1)
38. Методы поиска иерархических ассоциативных правил. (ОПК-6.1)
39. Определение кластеризации. (ОПК-6.1)
40. Постановка задачи кластеризации. (ОПК-6.2, ОПК-6.3)
41. Цели кластеризации в Data Mining. (ОПК-6.3)
42. Примеры кластеризации в различных областях. (ОПК-6.1)
43. Виды метрик кластеризации. (ОПК-6.1)
44. Шаги алгоритма кластеризации. (ОПК-6.1)
45. Меры расстояний кластеризации. (ОПК-6.1)
46. Пример работы алгоритма k-means. (ОПК-6.1)
47. Проблемы алгоритмов кластеризации. (ОПК-6.3)

## 6. УЧЕБНО-МЕТОДИЧЕСКОЕ И ИНФОРМАЦИОННОЕ ОБЕСПЕЧЕНИЕ ДИСЦИПЛИНЫ (МОДУЛЯ, ПРАКТИКИ)

### 6.1. Рекомендуемая литература

|     | Авторы, составители          | Заглавие                               | Издательство, год.       | Электронный адрес   |
|-----|------------------------------|--|--------------------------|---|
| Л.1 | Кузьмин В. И., Гадзаов А. Ф. | Методы анализа данных: Учебное пособие | Москва : РТУ МИРЭА, 2020 | <a href="https://e.lanbook.com/book/171433">https://e.lanbook.com/book/171433</a> |

|     | Авторы, составители                                  | Заглавие   | Издательство, год.             | Электронный адрес   |
|-----|--|--|--------------------------------|---|
| Л.2 | Садовникова Н. П.,<br>Щербаков М. В.                 | Технологии анализа данных: учеб.<br>пособие  | Волгоград:<br>ВолГТУ, 2015     |   |
| Л.3 | Кузнецов С. Ю.,<br>Костикова А. В.,<br>Сайкина Ю. А. | Интеллектуальный анализ данных: учеб.<br>пособие   | Волгоград:<br>ВолГТУ, 2019     |   |
| Л.4 | Ростовцев В. С.                                      | Искусственные нейронные сети:<br>учебник   | Санкт-Петербург:<br>Лань, 2021 | <a href="https://e.lanbook.com/reader/book/160142/#213">https://e.lanbook.com/reader/book/160142/#213</a>   |
| Л.5 | Садовникова Н.П.                                     | Методические указания по выполнению<br>лабораторных работ по дисциплине<br>«Методы и технологии обработки<br>больших данных»   | ВолГТУ, 2023                   | <a href="http://dump.vstu.ru/storage/Kafiedry/ctuas/09.04.02_magistratura_Informacionnye_sisitemy_i_tehnologii/Metodicheskie_material">http://dump.vstu.ru/storage/Kafiedry/ctuas/09.04.02_magistratura_Informacionnye_sisitemy_i_tehnologii/Metodicheskie_material</a> |
| Л.6 | Ерещенко Т.В.,<br>Ращевский Н.М.                     | Методические указания по выполнению<br>самостоятельной работы студентами по<br>направлению 09.04.02 Информационные<br>системы и технологии (программа<br>магистратуры: Цифровые технологии в<br>урбанистике, архитектуре и строительстве | ВолГТУ, 2023                   | <a href="http://dump.vstu.ru/storage/Kafiedry/ctuas/09.04.02_magistratura_Informacionnye_sisitemy_i_tehnologii/Metodicheskie_material">http://dump.vstu.ru/storage/Kafiedry/ctuas/09.04.02_magistratura_Informacionnye_sisitemy_i_tehnologii/Metodicheskie_material</a> |

## 6.2. Перечень ресурсов информационно-телекоммуникационной сети "Интернет"

Электронная информационно-образовательная среда ВолГТУ ([eos2.vstu.ru](http://eos2.vstu.ru))

Open Data Science, <https://opendatascience.com/>

## 6.3 Перечень программного обеспечения

|         |                         |
|---------|-------------------------|
| 6.3.1.1 | СДО "Moodle"            |
| 6.3.1.2 | Windows                 |
| 6.3.1.3 | Adobe Acrobat Reader DC |
| 6.3.1.4 | LibreOffice             |

## 6.4 Перечень информационных справочных систем и электронных библиотечных систем (ЭБС)

|         |  |
|---------|--|
| 6.3.2.1 | Научная электронная библиотека   |
| 6.3.2.2 | ЭБС "Book.ru"  |
| 6.3.2.3 | ЭБС "Лань"   |
| 6.3.2.4 | ИБЦ ВолГТУ <a href="http://library.vstu.ru/">http://library.vstu.ru/</a> |

## 7. МАТЕРИАЛЬНО-ТЕХНИЧЕСКОЕ ОБЕСПЕЧЕНИЕ ДИСЦИПЛИНЫ (МОДУЛЯ, ПРАКТИКИ) /ОБОРУДОВАНИЕ

|     |  |
|-----|--|
| 7.1 | Мультимедийная учебная аудитория для проведения занятий лекционного и семинарского типа, групповых и индивидуальных консультаций, текущего контроля и промежуточной аттестации. /Учебная доска, учебная мебель. Учебная мебель на 34 посадочных места, учебная доска – 1 шт.<br>Компьютерное и мультимедийное оборудование: интерактивная трибуна 21,5, Intel Core i3, конференц - микрофон, клавиатура беспроводная, мультисенсорная панель 3,5-дюйма, телевизор LG 55UQ75006LF.ARUB, 4K Ultra HD, черный – 2 шт. |
| 7.2 | Аудитория для самостоятельной работы обучающихся./Учебная мебель, компьютерная техника с возможностью подключения к сети "Интернет" и обеспечением доступа в электронную информационно-образовательную среду университета (читальный зал информационно-библиотечного центра)   |
| 7.3 | Лаборатория анализа больших данных<br>Учебная мебель на 26 посадочных мест, учебная доска – 2 шт.<br>Компьютерное и мультимедийное оборудование: рабочая станция Lenovo S40 40 All-In-One FS 21.5" -14 шт.;<br>Стационарный компьютер RAMEC (Процессор: i5 10400, ОЗУ: 16Gb, SSD: 500 Gb, монитор Philips 242V8A) -1 шт.;<br>экран + проектор SMART – 1 шт.  |

## 8. МЕТОДИЧЕСКИЕ УКАЗАНИЯ ДЛЯ ОБУЧАЮЩИХСЯ ПО ОСВОЕНИЮ ДИСЦИПЛИНЫ (МОДУЛЯ, ПРАКТИКИ)

Организация образовательного процесса по данной дисциплине регламентируется учебным планом и расписанием учебных занятий. При формировании своей индивидуальной образовательной траектории обучающийся имеет право на перезачет дисциплины (переаттестации ее части), если она была освоена в процессе предшествующего обучения. Перезачёт (переаттестации ее части) освобождает обучающегося от необходимости повторного освоения дисциплины (полностью или частично). Учебный процесс при преподавании курса основывается на использовании традиционных, инновационных и информационных образовательных технологий. Традиционные образовательные технологии представлены лекциями и лабораторными работами. Инновационные образовательные технологии используются в виде широкого применения активных и интерактивных форм проведения занятий. Информационные образовательные технологии реализуются путем активизации самостоятельной работы студентов в электронной информационной образовательной среде. Лекционный курс предполагает систематизированное изложение основных вопросов учебного плана.

На первой лекции лектор информирует студентов о рекомендуемой литературе и электронных источниках информации по дисциплине, с указанием, какой учебник (учебное пособие) является базовым.

Лабораторные работы предполагают выполнение и отчет заданий по темам, рассмотренным на лекционных и закрепленных на лабораторных работах. Каждому лабораторному занятию предшествует самостоятельная подготовка студента, включающая: ознакомление с содержанием лабораторной работы по методическим указаниям; проработку теоретической части по лекционному материалу и учебникам, рекомендованным в методических указаниях.

Самостоятельная работа студентов включает изучение законспектированного на лекционных занятиях материала, дополнение его с учетом рекомендованной по данной теме литературы, самостоятельную подготовку к лабораторным работам.

В течении семестра для студентов проводятся групповые текущие консультации по учебной дисциплине, а также консультация перед экзаменом.

Методические рекомендации по обучению лиц с ограниченными возможностями здоровья и инвалидов. Профессорско-педагогический состав знакомится с психолого-физиологическими особенностями обучающихся инвалидов и лиц с ограниченными возможностями здоровья (ОВЗ), индивидуальными программами реабилитации инвалидов (при наличии).

При необходимости осуществляется дополнительная поддержка преподавания тьюторами, психологами, социальными работниками, прошедшими подготовку ассистентами. В соответствии с методическими рекомендациями Минобрнауки РФ (утв. 8 апреля 2014 г. N АК-44/05вн), в курсе предполагается использовать социально-активные и рефлексивные методы обучения, технологии социокультурной реабилитации с целью оказания помощи в установлении полноценных межличностных отношений с другими студентами, создании комфортного психологического климата в студенческой группе.

Подбор и разработка учебных материалов производятся с учетом предоставления материала в различных формах: аудиальной, визуальной, с использованием специальных технических средств и информационных систем. Освоение дисциплины лицами с ОВЗ осуществляется с использованием средств обучения общего и специального назначения (персонального и коллективного использования). Материально-техническое обеспечение предусматривает приспособление аудиторий к нуждам лиц с ОВЗ (при необходимости). Форма проведения аттестации для студентов-инвалидов устанавливается с учетом индивидуальных психофизических особенностей. Для студентов с ОВЗ предусматривается доступная форма предоставления заданий оценочных средств. Студентам с инвалидностью увеличивается время на подготовку ответов на контрольные вопросы. Для таких студентов предусматривается доступная форма предоставления ответов на задания. При необходимости для обучающихся с инвалидностью процедура оценивания результатов обучения может проводиться в несколько этапов.