

МИНИСТЕРСТВО НАУКИ И ВЫСШЕГО ОБРАЗОВАНИЯ РОССИЙСКОЙ ФЕДЕРАЦИИ
федеральное государственное бюджетное образовательное учреждение
высшего образования
«Тольяттинский государственный университет»

Б1.В.05
(индекс дисциплины)

РАБОЧАЯ ПРОГРАММА ДИСЦИПЛИНЫ

Введение в интеллектуальный анализ данных

(наименование дисциплины)

по направлению подготовки
09.03.04 Программная инженерия

направленность (профиль)
Программная инженерия с применением ИИ-технологий

Форма обучения: очная

Год набора: 2026

Общая трудоемкость: 3 ЗЕ

Распределение часов дисциплины по семестрам

Семестр	3	Итого
Форма контроля	зачет	
Вид занятий		
Лекции	16	16
Лабораторные		
Практические	16	16
Руководство: курсовые работы (проекты) / РГР		
Промежуточная аттестация	0,25	0,25
Контактная работа	32,25	32,25
Самостоятельная работа	75,75	75,75
Контроль		
Итого	108	108

Рабочую программу составил(и):

доцент института цифровых технологий, канд. пед. наук, доцент Гущина О.М.

(должность, ученое звание, степень, Фамилия И.О.)

Рецензирование рабочей программы дисциплины:



Отсутствует



Рецензент

(должность, ученое звание, степень, Фамилия И.О.)

Рабочая программа дисциплины составлена на основании ФГОС ВО и учебного плана направления подготовки

09.03.04 Программная инженерия

Срок действия рабочей программы дисциплины до «31» августа 2030 г.

УТВЕРЖДЕНО

На заседании института цифровых технологий

(протокол заседания № 1 от «05» сентября 2025 г.).

1. Цель освоения дисциплины

Цель освоения дисциплины – формирование у обучающихся систематизированных теоретических знаний и практических навыков в области интеллектуального анализа данных (Data Mining), необходимых для решения прикладных задач извлечения знаний из данных.

2. Место дисциплины в структуре ОПОП ВО

Дисциплины и практики, на освоении которых базируется данная дисциплина: «Основы дискретной математики и логики», «Алгоритмы и программирование на основе Python».

Дисциплины и практики, для которых освоение данной дисциплины необходимо как предшествующее: «Машинное обучение и глубокий анализ данных».

3. Планируемые результаты обучения

Формируемые и контролируемые компетенции (код и наименование)	Индикаторы достижения компетенций (код и наименование)	Планируемые результаты обучения
ПК-10 Способен разрабатывать и применять классические модели машинного обучения для извлечения знаний из данных	ПК-10.1. Знает классические модели машинного обучения и основные методы их разработки и применения для извлечения знаний из данных	Знать: типы задач ML (классификация, регрессия, кластеризация); классические алгоритмы (линейная регрессия, логистическая регрессия, деревья решений, SVM, k-means, k-NN). Уметь: выбирать модель, адекватную типу задачи и данным. Владеть: навыками использования библиотек (Scikit-learn).
	ПК-10.2. Умеет применять классические модели машинного обучения для извлечения знаний из данных	Знать: процесс кросс-валидации, методы борьбы с переобучением. Уметь: проводить полный цикл ML-проекта: от сбора данных до интерпретации результатов модели. Владеть: навыками построения пайплайнов обработки данных и обучения моделей.
	ПК-10.3. Владеет навыками разработки классических моделей машинного обучения и применения их для извлечения знаний из данных	Знать: внутреннее устройство основных алгоритмов ML. Уметь: модифицировать и комбинировать классические алгоритмы для решения специфических задач. Владеть: навыками реализации алгоритмов "с нуля" и создания ансамблевых моделей.

4. Структура и содержание дисциплины

Модуль (раздел)	Вид учебной работы	Наименование тем занятий (учебной работы)	Семестр	Объем, ч.	Баллы	Интерактив, ч.	Формы текущего контроля (наименование оценочного средства)
Модуль 1. Введение в интеллектуальный анализ данных: основные понятия и задачи	Лек1	Лек1 Актуальность, базовая терминология и тенденции развития. Основные задачи, этапы и классификация методов анализа данных.	3	2	-	-	
	Лек2	Лек2 Предварительная обработка данных. Классификация.	3	2	-	-	
	Ср	Самостоятельное изучение методических рекомендаций при подготовке к практическим работам	3	20	-	-	
	Пр1	Пр1 Регрессия. Ассоциация, последовательная ассоциация, аномалии и визуализация.	3	2	15	-	Отчет по практической работе 1
	Пр2	Пр2 Высокопроизводительная обработка данных. Программные среды для интеллектуального анализа данных.	3	2	15	-	Отчет по практической работе 2
Модуль 2. Методы предобработки данных и классификации	Лек3	Лек3 Основные методы и предварительная обработка данных.	3	2	-	-	
	Лек4	Лек4 Оптимизация признакового пространства без трансформации пространства признаков.	3	2	-	-	
	Ср	Самостоятельное изучение методических рекомендаций при подготовке к практическим работам	3	20	-	-	
	Пр3	Пр3 Контролируемая непараметрическая нейросетевая классификация.	3	2	15	-	Отчет по практической работе 3
	Пр4	Пр4 Классификация по методу машины опорных векторов. Деревья решений.	3	2	15	-	Отчет по практической работе 4
Модуль 3. Регрессия, поиск закономерностей и анализ аномалий	Лек5	Лек5 Понятие регрессии. Основные этапы регрессионного анализа.	3	2	-	-	
	Лек6	Лек6 Описание алгоритма ассоциации.	3	2	-	-	
	Ср	Самостоятельное изучение методических	3	20	-	-	

Модуль (раздел)	Вид учебной работы	Наименование тем занятий (учебной работы)	Семестр	Объем, ч.	Баллы	Интерактив, ч.	Формы текущего контроля (наименование оценочного средства)
		рекомендаций при подготовке к практическим работам					
	Пр5	Пр5 Алгоритмы семейства «Априори». Алгоритм GSP.	3	2	10	-	Отчет по практической работе 5
	Пр6	Пр6 Обнаружение аномалий и методы визуализации.	3	2	10	-	Отчет по практической работе 6
Модуль 4. Технологии и инструменты для распределенной обработки данных	Лек7	Лек7 Принципы организации высокопроизводительных вычислений. SMP-системы.	3	2	-	-	
	Лек8	Лек8 Модели параллельных вычислений MPMD, SPMD.	3	2	-	-	
	Ср	Самостоятельное изучение методических рекомендаций при подготовке к практическим работам	3	15,75	-	-	
	Пр7	Пр7 Платформа программирования и выполнения распределённых вычислений. Нереляционные базы данных. Среды и языки программирования Python, R.	3	2	10	-	Отчет по практической работе 7
	ПА	Промежуточная аттестация	3	0,25	-	-	
	Псщ	Посещаемость	3	–	10		
	Пр8	Итоговое тестирование.	3	2	100		Тестовые задания 1-480
Итого:				108			

Схема расчета итогового балла: по накопительному рейтингу
Текущий рейтинг + Результат итогового теста и все делится на 2

5. Образовательные технологии

При изучении дисциплины используются следующие образовательные технологии:

- технологии традиционного обучения в форме лекций, практических работ и самостоятельной работы обучающихся;
- технология проектного обучения: реализация и защита отчетов по практическим работам.

Технологии традиционного обучения - организация учебного процесса в вузе, основанная на лекционных и практических формах обучения: объяснительно-иллюстративное обучение. Данная технология применяется во всех модулях курса.

Технология интерактивного обучения - организация учебного процесса, которая предполагает максимальную активность обучающихся в процессе формирования ключевых компетенций. На практическом занятии обучающиеся представляют результат выполнения заданной работы.

6. Методические указания по освоению дисциплины

6.1. Рекомендации по подготовке к практическим занятиям

Обучающимся следует:

- при подготовке к практическим занятиям следует обязательно использовать не только лекции, учебную литературу, но и другие источники;
- в начале занятий задать преподавателю вопросы по материалу, вызвавшему затруднения в его понимании и освоении при решении задач, заданных для самостоятельного решения;
- на занятии доводить каждую задачу до окончательного решения, демонстрировать понимание проведенных расчетов (анализов, ситуаций), в случае затруднений обращаться к преподавателю.

Для того чтобы практические занятия приносили максимальную пользу, необходимо помнить, что решение задач проводится по рассмотренному на лекциях материалу и связаны, как правило, с детальным разбором отдельных вопросов лекционного курса. Следует подчеркнуть, что только после усвоения лекционного материала с определенной точки зрения (а именно с той, с которой он излагается на лекциях) он будет закрепляться обучающимся на практических занятиях как в результате обсуждения и анализа лекционного материала, так и с помощью решения проблемных ситуаций, задач. При этих условиях обучающийся не только хорошо усвоит материал, но и научится применять его на практике, а также получит дополнительный стимул (и это очень важно) для активной проработки лекции.

При самостоятельном решении задач нужно обосновывать каждый этап решения, исходя из теоретических положений курса. Если обучающийся видит несколько путей решения проблемы (задачи), то нужно сравнить их и выбрать самый рациональный. Полезно до начала вычислений составить краткий план решения проблемы (задачи). Решение проблемных задач или примеров следует излагать подробно, вычисления располагать в строгом порядке, отделяя вспомогательные вычисления от основных. Решения при необходимости нужно сопровождать комментариями, схемами, чертежами и рисунками.

Следует помнить, что решение каждой учебной задачи должно доводиться до окончательного логического ответа, которого требует условие, и по возможности с выводом. Полученный ответ следует проверить способами, вытекающими из существа данной задачи. Полезно также (если возможно) решать несколькими способами и сравнить полученные результаты. Решение задач данного типа нужно продолжать до приобретения твердых навыков в их решении.

6.2. Рекомендации по подготовке к зачету

Подготовка к зачету способствует закреплению, углублению и обобщению знаний, получаемых, в процессе обучения, а также применению их к решению практических задач. Готовясь к зачету, обучающийся ликвидирует имеющиеся пробелы в знаниях, углубляет, систематизирует и упорядочивает свои знания. На зачете обучающийся демонстрирует то, что он приобрел в процессе обучения по конкретной учебной дисциплине.

На консультации перед экзаменом обучающиеся должны быть ознакомлены с основными требованиями и получить ответы на возникающие в процессе подготовки вопросы.

Необходимо ориентировать обучающихся на систематическую подготовку к занятиям в течение семестра, что позволит использовать время экзаменационной сессии для систематизации знаний.

7. Оценочные средства

7.1. Паспорт оценочных средств

Семестр	Код контролируемой компетенции (или ее части)	Наименование оценочного средства
3	ПК-10	Тестовые задания 1-480 Вопросы к зачету 1-64 Отчеты по практическим работам 1-7

7.2. Типовые задания или иные материалы, необходимые для текущего контроля

7.2.1. Типовые тестовые материалы

(наименование оценочного средства)

Типовые примеры заданий

Задание 1

Что является основной целью интеллектуального анализа данных (Data Mining)?

Выберите один из 4 вариантов ответа:

1. Хранение больших объемов информации.
2. + **Выявление скрытых, нетривиальных и практически полезных закономерностей в данных.**
3. Создание визуальных отчетов для руководства.
4. Перевод данных в удобный для восприятия формат.

Задание 2

Какая из перечисленных задач НЕ является типичной задачей интеллектуального анализа данных?

Выберите один из 4 вариантов ответа:

1. Классификация.
2. Кластеризация.
3. + **Нормализация данных.**
4. Поиск ассоциативных правил.

Задание 3

Какой этап процесса анализа данных (KDD) следует сразу после сбора данных?

Выберите один из 4 вариантов ответа:

1. Построение модели.
2. + **Предварительная обработка и очистка данных.**
3. Интерпретация результатов.
4. Развертывание решения.

Задание 4

Методы обучения с учителем (supervised learning) характеризуются тем, что:

Выберите один из 4 вариантов ответа:

1. Модель ищет внутреннюю структуру данных без заранее известных ответов.
2. **+ Для обучения модели используется набор данных с размеченными примерами (известными выходными значениями).**
3. Модель требует постоянного вмешательства человека-эксперта в процессе работы.
4. Алгоритм сам решает, какие данные ему необходимы для обучения.

Задание 5

Что из перечисленного является современным трендом в развитии Интеллектуального Анализа Данных (ИАД)?

Выберите один из 4 вариантов ответа:

1. Отказ от использования SQL.
2. **+ Интеграция с машинным обучением и искусственным интеллектом.**
3. Уменьшение объемов обрабатываемых данных.
4. Сосредоточение только на реляционных базах данных.

Задание 6

Понятие "Big Data" (Большие данные) традиционно характеризуется тремя "V". Что это за "V"?

Выберите один из 4 вариантов ответа:

1. Volume, Velocity, Veracity.
2. **+ Volume, Velocity, Variety.**
3. Volume, Value, Variety.
4. Velocity, Veracity, Value.

Задание 7

Какая из перечисленных задач относится к обучению без учителя (unsupervised learning)?

Выберите один из 4 вариантов ответа:

1. **+ Кластеризация.**
2. Прогнозирование временных рядов.
3. Классификация изображений.
4. Определение спама в email.

Задание 8

Что означает акроним KDD в контексте анализа данных?

Выберите один из 4 вариантов ответа:

1. Key Data Distribution.
2. Knowledge Driven Development.
3. **+ Knowledge Discovery in Databases.**
4. Kernel Data Description.

Задание 9

Какой тип задачи решается, когда нужно предсказать категорию объекта на основе его признаков?

Выберите один из 4 вариантов ответа:

1. Регрессия.
2. **+ Классификация.**
3. Кластеризация.
4. Снижение размерности.

Задание 10

Что такое "overfitting" (переобучение) в контексте машинного обучения?

Выберите один из 4 вариантов ответа:

1. Модель слишком проста для данных.
2. **+ Модель слишком точно соответствует обучающим данным и плохо обобщается на новые данные.**

3. Модель не может выучить закономерности в данных.
4. Модель требует слишком много вычислительных ресурсов.

Критерии оценки за пройденный тест:

100 баллов выставляется обучающемуся, если он ответил правильно на все вопросы рандомной выборки 30 тестовых заданий;

0-99 баллов выставляется обучающемуся в зависимости от количества верных ответов

7.2.2. Пример практической работы

Практическая работа 1. Регрессия. Ассоциация, последовательная ассоциация, аномалии и визуализация.

Цель работы: сформировать у обучающихся первичные практические навыки решения основных задач интеллектуального анализа данных (ИАД). Научиться применять простейшие алгоритмы регрессии, находить ассоциативные правила и последовательные шаблоны, выявлять аномалии и визуализировать данные для интерпретации результатов.

Порядок работы:

1. Подготовка данных:
 - Загрузите набор данных для задачи регрессии.
 - Проведите эксплоративный анализ: выведите основные статистики, постройте матрицу корреляций и гистограммы распределения признаков.
 - Разделите данные на обучающую и тестовую выборки в соотношении 70/30.
2. Построение и валидация модели линейной регрессии:
 - Импортируйте класс `LinearRegression` из библиотеки `sklearn.linear_model`.
 - Создайте экземпляр модели, обучите его на тренировочных данных и сделайте прогноз для тестовой выборки.
 - Рассчитайте метрики качества регрессии: Mean Absolute Error (MAE), Mean Squared Error (MSE), R2-score.
 - Визуализируйте результаты: постройте график, где по оси X - истинные значения целевой переменной, по оси Y - предсказанные.
3. Поиск ассоциативных правил:
 - Загрузите или создайте транзакционный набор данных (пример: рыночная корзина, где каждая строка - список товаров в одной покупке).
 - Используйте алгоритм `Apriori` из библиотеки `mlxtend.frequent_patterns`.
 - Найдите частые наборы элементов с заданными минимальными поддержкой (`min_support=0.05`) и уверенностью (`min_confidence=0.2`).
 - Сгенерируйте ассоциативные правила и отсортируйте их по "лифту".
4. Знакомство с последовательной ассоциацией:
 - На примере небольшого набора данных, имитирующего последовательности действий пользователя, сформулируйте, какие последовательные шаблоны можно искать.
 - Сделайте вывод о том, чем задача поиска последовательных шаблонов отличается от задачи поиска простых ассоциативных правил.
5. Обнаружение выбросов:
 - На том же наборе данных, что и для регрессии, примените метод `IsolationForest` из `sklearn.ensemble` для обнаружения аномальных наблюдений.
 - Визуализируйте аномалии на графике, используя первые два главных компонента (PCA) или два наиболее значимых признака.
6. Визуализация данных и результатов:

- Используя библиотеку `matplotlib` или `seaborn`, создайте комплексную визуализацию, включающую:
 - Диаграмму рассеяния с выделением аномальных точек.
 - Столбчатую диаграмму с метриками качества регрессионной модели.
 - Тепловую карту корреляций исходного набора данных.

Отчет должен включать:

1. Краткое теоретическое описание решаемых задач (регрессия, ассоциация, поиск аномалий).
2. Исходный код всех выполненных шагов с комментариями.
3. Результаты выполнения кода: рассчитанные метрики, найденные ассоциативные правила (топ-5 по лифту), список идентифицированных аномалий.
4. Все построенные графики и визуализации.
5. Выводы по работе: интерпретация полученных результатов, сравнение эффективности методов, возникшие трудности.

Требования к оформлению

Отчет должен содержать подробное описание (включая иллюстрации). Отчёт по практическому занятию выполняется на страницах формата А4 в электронном виде.

При оформлении отчёта используется сквозная нумерация страниц, считая титульный лист первой страницей. Номер страницы на титульном листе не ставится. Номера страницы ставятся по центру сверху.

При оформлении отчёта соблюдать следующие требования:

- Для заголовков: полужирный шрифт, 14 пт, центрированный.
- Для основного текста: нежирный шрифт, 14 пт, выравнивание по ширине.
- Во всех случаях тип шрифта – Times New Roman, отступ абзаца 1.25 см, полуторный междустрочный интервал.
- Поля: левое – 2 см, правое, верхнее и нижнее – 1 см.

Процедура оценивания

Оценка выполненной практической работы проводится по следующим критериям:

1. Наличие всей существенной информации по работе
2. Точность и полнота предоставляемых сведений
3. Непротиворечивость приводимой информации
4. Правильность интерпретаций и выводов, которые сделаны по результатам работы
5. Степень достижения обучающимся поставленной цели
6. Обоснованность применяемого решения
7. Грамотность (содержательная) используемых формулировок

Критерии оценки за отчеты по практическим работам:

Формы текущего контроля	Критерии и нормы оценки
Отчеты по практическим работам 1-4	<ul style="list-style-type: none"> • Работа выполнена в полном объеме с соблюдением необходимой последовательности действий; аккуратно, четко и без ошибок выполнил отчет, вывод исчерпывающий и доказательный. При защите отчета обучающийся ответил на все вопросы по теме; хорошо ориентируется в материале, умеет определить взаимосвязь факторов и их влияние на конечную цель, умеет графически отобразить важнейшие функциональные зависимости – 12-15 баллов

Формы текущего контроля	Критерии и нормы оценки
	<ul style="list-style-type: none"> • Работа выполнена в полном объеме с соблюдением необходимой последовательности действий; обучающийся без ошибок выполнил отчет, вывод исчерпывающий. При защите отчета хорошо разбирается в материале, но не уверен и неполно отвечает на вопросы. Способность к обобщению причинно-следственных связей важнейших факторов выражена недостаточно – 9-11 баллов; • Работа выполнена не полностью, но объем выполненной части таков, что позволяет получить правильные результаты и выводы; выполнен с несущественными замечаниями. Вывод по работе не раскрывает сути работы. Владение понятийным аппаратом темы недостаточны – 5-8 баллов; • Обучающийся выполнил работу не полностью или объем выполненной части работы не позволяет сделать правильных выводов. В ответах на вопросы есть грубые ошибки. Нет знания принципиальных теоретических положений темы –1-4 балла
Отчеты по практическим работам 5-7	<ul style="list-style-type: none"> • Работа выполнена в полном объеме с соблюдением необходимой последовательности действий; аккуратно, четко и без ошибок выполнил отчет, вывод исчерпывающий и доказательный. При защите отчета ответил на все вопросы по теме; хорошо ориентируется в материале, умеет определить взаимосвязь факторов и их влияние на конечную цель, умеет графически отобразить важнейшие функциональные зависимости – 9-10 баллов • Работа выполнена в полном объеме с соблюдением необходимой последовательности действий; обучающийся без ошибок выполнил отчет, вывод исчерпывающий. При защите отчета хорошо разбирается в материале, но не уверен и неполно отвечает на вопросы. Способность к обобщению причинно-следственных связей важнейших факторов выражена недостаточно – 6-8 баллов; • Работа выполнена не полностью, но объем выполненной части таков, что позволяет получить правильные результаты и выводы; выполнен с несущественными замечаниями. Вывод по работе не раскрывает сути работы. Владение понятийным аппаратом темы недостаточны 3-5 баллов; • Обучающийся выполнил работу не полностью или объем выполненной части работы не позволяет сделать правильных выводов. В ответах на вопросы есть грубые ошибки. Нет знания принципиальных теоретических положений темы. 1-2 балла

7.3. Оценочные средства для промежуточной аттестации по итогам освоения дисциплины

7.3.1. Вопросы к промежуточной аттестации

Семестр 3

№	Вопросы к зачету
1.	Что понимается под интеллектуальным анализом данных (Data Mining)?
2.	В чем заключается процесс Discovery in Databases (KDD)?
3.	Назовите основные тенденции развития интеллектуального анализа данных.
4.	Перечислите и охарактеризуйте основные задачи анализа данных.
5.	Дайте классификацию методов анализа данных по типу решаемых задач.
6.	Каковы основные этапы подготовки данных к анализу?
7.	Что такое «предварительная обработка данных» и почему она важна?
8.	Какие существуют методы обработки пропущенных значений в данных?
9.	Что такое «классификация» как задача интеллектуального анализа данных?
10.	Каковы цели этапа визуализации данных?
11.	Что такое «регрессия» и чем она отличается от классификации?
12.	Что понимается под «ассоциацией» и «последовательной ассоциацией»?
13.	Какие типы аномалий могут встречаться в данных?
14.	Что подразумевается под высокопроизводительной обработкой данных?
15.	Назовите основные программные среды, используемые для интеллектуального анализа данных.
16.	Каковы основные цели и задачи предварительной обработки данных?
17.	Что такое «нормализация» и «стандартизация» данных? В чем их различие?
18.	Какие методы кодирования категориальных признаков вам известны?
19.	Что такое «признаковое пространство»?
20.	Что понимается под оптимизацией признакового пространства?
21.	Какие методы отбора признаков без трансформации пространства вам известны?
22.	Что такое «контролируемое обучение» (supervised learning)?
23.	Дайте определение «непараметрическим методам» в машинном обучении.
24.	Опишите общий принцип работы искусственной нейронной сети для задачи классификации.
25.	Что такое метод опорных векторов (SVM) и в чем его основная идея?
26.	Как работает алгоритм построения дерева решений?
27.	Какие критерии используются для разделения данных в узлах дерева решений (например, энтропия, индекс Джини)?
28.	В чем заключаются преимущества и недостатки деревьев решений?
29.	Что такое «ядерный трюк» в методе опорных векторов и для чего он применяется?
30.	Какие параметры нейронной сети влияют на ее способность к обучению?
31.	Дайте определение задаче «регрессии» в машинном обучении.
32.	Каковы основные этапы регрессионного анализа?
33.	Что такое линейная регрессия и как находятся ее параметры?
34.	Какие метрики используются для оценки качества регрессионных моделей?
35.	Что такое «ассоциативные правила» и где они применяются?
36.	Опишите алгоритм Apriori для поиска ассоциативных правил.
37.	Какие основные показатели ассоциативных правил (support, confidence, lift)?
38.	Чем задача поиска последовательных шаблонов (sequential pattern mining)

№	Вопросы к зачету
	отличается от поиска ассоциативных правил?
39.	Опишите алгоритм GSP для поиска последовательных шаблонов.
40.	Что такое «аномалия» или «выброс» в данных?
41.	Какие основные подходы к обнаружению аномалий вам известны?
42.	Как визуализация помогает в анализе аномалий?
43.	Назовите и охарактеризуйте основные методы визуализации многомерных данных.
44.	Какова роль визуализации на этапе интерпретации результатов анализа?
45.	Какие библиотеки Python или R используются для визуализации данных и каковы их основные возможности?
46.	Как визуализация помогает в интерпретации результатов регрессионного анализа?
47.	Какие графики наиболее эффективны для визуального обнаружения аномалий в данных?
48.	В чем разница между описательной и прогнозной визуализацией?
49.	Какова роль визуализации на заключительном этапе процесса интеллектуального анализа данных (KDD)?
50.	Какие методы визуализации подходят для анализа многомерных данных?
51.	Что такое "интерактивная визуализация" и каковы ее преимущества?
52.	Как с помощью визуализации можно оценить качество кластеризации или классификации?
53.	Что подразумевается под высокопроизводительной обработкой данных (High-Performance Computing)?
54.	Опишите архитектуру SMP-систем. В чем их преимущества и недостатки?
55.	В чем разница между моделями параллельных вычислений MPMD и SPMD?
56.	Какую проблему решает платформа Hadoop в контексте больших данных?
57.	Опишите принцип работы алгоритма MapReduce.
58.	Каковы основные компоненты экосистемы Hadoop и их назначение (HDFS, MapReduce, YARN)?
59.	Что такое Apache Spark и каковы его преимущества перед Hadoop MapReduce?
60.	Что такое нереляционные (NoSQL) базы данных и когда их применение оправдано?
61.	Опишите основные особенности СУБД HBase и Cassandra.
62.	Каковы основные преимущества и области применения языка Python для интеллектуального анализа данных?
63.	Каковы основные преимущества и области применения языка R для интеллектуального анализа данных?
64.	Назовите ключевые библиотеки Python (например, Pandas, Scikit-learn) и R, используемые на разных этапах анализа данных.

7.3.2. Критерии и нормы оценки

Семестр	Форма проведения промежуточной аттестации	Критерии и нормы оценки	
3	Зачет (по накопительному рейтингу)	«зачтено»	55 и более баллов
		«не зачтено»	менее 55 баллов

8. Учебно-методическое и информационное обеспечение дисциплины

8.1. Обязательная литература

№ п/п	Авторы, составители	Заглавие (заголовок)	Тип (учебник, учебное пособие, учебно-методическое пособие, практикум, др.)	Год издания	Количество в научной библиотеке / Наименование ЭБС
1.	А. В. Замятин	Интеллектуальный анализ данных	Учебное пособие	2020	ЭБС «ZNANIUM.COM»
2	Д.М. Назаров [и др.]	Data Science и интеллектуальный анализ данных	Учебное пособие	2023	ЭБС “IPRbooks”

8.2. Дополнительная литература

№ п/п	Авторы, составители	Заглавие (заголовок)	Тип (учебник, учебное пособие, учебно-методическое пособие, практикум, др.)	Год издания	Количество в научной библиотеке / Наименование ЭБС
1.	Т. В. Афанасьева	Введение в проектирование систем интеллектуального анализа данных	Учебное пособие	2017	ЭБС «Лань»
2.	Е. П. Богданов	Интеллектуальный анализ данных	Практикум	2019	ЭБС «ZNANIUM.COM»
3.	А. В. Замятин	Введение в интеллектуальный анализ данных	Учебное пособие	2016	ЭБС «ZNANIUM.COM»

8.3. Перечень профессиональных баз данных и информационных справочных систем

№ п/п	Наименование	Ссылка
1	Springer Nature (Полнотекстовая коллекция журналов)	https://www.springernature.com/gp/products
2	Springer eBooks (Полнотекстовая коллекция электронных книг издательства Springer Nature)	https://link.springer.com/
3	«Кодекс»	https://kodeks.ru/
4	ELIBRARY.RU (электронная библиотека научных публикаций)	http://elibrary.ru
5	"Гарант"	https://www.garant.ru/
6	"КонсультантПлюс"	https://www.consultant.ru/
7	Техэксперт	https://cntd.ru/

8.4. Перечень программного обеспечения

№ п/п	Наименование ПО	Реквизиты договора (дата, номер, срок действия)
1	WinPro 10 RUS Upgrd OLP NL Acdmc	Договор № 757 от 04.07.2018, срок действия – бессрочно. Контракт № 1653 от 14.12.2018, срок действия – бессрочно.
2	Office Stdandard 2013 Russian OLP NL AcademicEdition	Контракт № 690 от 19.05.2015, срок действия - бессрочно
3	Python 3.11	Лицензия Python Software Foundation License (PSFL)
4	Pandas	Лицензия: BSD License (Open Source).
5	NumPy	Лицензия: BSD License (Open Source).
6	R Language	Лицензия: GNU GPL v2 (Open Source).

8.5. Описание материально-технической базы, необходимой для осуществления образовательного процесса по дисциплине

№ п/п	Наименование оборудованных учебных кабинетов, лабораторий, мастерских и др. объектов для проведения практических и лабораторных занятий, помещений для самостоятельной работы обучающихся (номер аудитории)	Перечень основного оборудования
1	Компьютерный класс. Учебная аудитория для проведения занятий лекционного типа. Учебная аудитория для проведения занятий семинарского типа. Учебная аудитория для проведения лабораторных работ. Учебная аудитория для курсового проектирования (выполнения курсовых работ). Учебная	Системные блоки (HP ProDesk), мониторы (Samsung), коммутатор (D-Link), столы ученические, столы компьютерные, стулья, доска аудиторная.

№ п/п	Наименование оборудованных учебных кабинетов, лабораторий, мастерских и др. объектов для проведения практических и лабораторных занятий, помещений для самостоятельной работы обучающихся (номер аудитории)	Перечень основного оборудования
	аудитория для проведения групповых и индивидуальных консультаций Учебная аудитория для проведения занятий текущего контроля и промежуточной аттестации. (УЛК-402)	
2	Помещение для самостоятельной работы обучающихся. (УЛК-105)	Столы, стулья, стеллажи (в т.ч. выставочные) с книгами, компьютеры, мобильные рабочие места.
3	Помещение для самостоятельной работы обучающихся. (УЛК-406)	Столы компьютерные, стулья, микрокомпьютеры raspberry pi 32 bit.