

Министерство науки и высшего образования Российской Федерации
Федеральное государственное бюджетное образовательное учреждение высшего
образования
«ИРКУТСКИЙ НАЦИОНАЛЬНЫЙ ИССЛЕДОВАТЕЛЬСКИЙ ТЕХНИЧЕСКИЙ
УНИВЕРСИТЕТ»

Структурное подразделение «Институт информационных технологий и анализа данных
(121)»

УТВЕРЖДЕНА:

на заседании совета института ИТиАД им. Е.И. Попова
Протокол №8 от 16 февраля 2026 г.

Рабочая программа дисциплины

«МЕТОДЫ АНАЛИЗА ДАННЫХ»

Направление: 09.03.01 Информатика и вычислительная техника

Автоматизированные системы обработки информации и управления

Квалификация: Бакалавр

Форма обучения: заочная

Документ подписан простой электронной подписью Составитель программы: Бучнев Олег Сергеевич Дата подписания: 25.05.2026

Документ подписан простой электронной подписью Утвердил: Говорков Алексей Сергеевич Дата подписания: 10.06.2026

Документ подписан простой электронной подписью Согласовал: Кононенко Роман Владимирович Дата подписания: 31.05.2026

Год набора – 2026

Иркутск, 2026 г.

1 Перечень планируемых результатов обучения по дисциплине, соотнесённых с планируемыми результатами освоения образовательной программы

1.1 Дисциплина «Методы анализа данных» обеспечивает формирование следующих компетенций с учётом индикаторов их достижения

Код, наименование компетенции	Код индикатора компетенции
ПКС-2 Способность разрабатывать программное обеспечение и требования к программному обеспечению по обработке данных	ПКС-2.3

1.2 В результате освоения дисциплины у обучающихся должны быть сформированы

Код индикатора	Содержание индикатора	Результат обучения
ПКС-2.3	Способен обоснованно применять различные алгоритмы анализа данных; оценивать быстродействие различных методов анализа данных	Знать основные типы задач в современном анализе данных и методы их решения; возможности и ограничения применяемых методов и моделей для анализа. Уметь осуществлять постановку задачи и выбор методов для решения задачи с учетом имеющихся ограничений, осуществлять предобработку данных, осуществлять интерпретацию результатов анализа. Владеть современными средствами и инструментами для реализации этапов и методов анализа данных.

2 Место дисциплины в структуре ООП

Изучение дисциплины «Методы анализа данных» базируется на результатах освоения следующих дисциплин/практик: «Математика», «Теория вероятностей, математическая статистика и случайные процессы», «Программирование», «Базы данных»

Дисциплина является предшествующей для дисциплин/практик: «Системы искусственного интеллекта», «Нейронные сети и их приложения»

3 Объем дисциплины

Объем дисциплины составляет – 3 ЗЕТ

Вид учебной работы	Трудоемкость в академических часах (Один академический час соответствует 45 минутам астрономического часа)		
	Всего	Учебный год № 3	Учебный год № 4
Общая трудоемкость дисциплины	108	36	72

Аудиторные занятия, в том числе:	14	2	12
лекции	6	2	4
лабораторные работы	8	0	8
практические/семинарские занятия	0	0	0
Самостоятельная работа (в т.ч. курсовое проектирование)	90	34	56
Трудоемкость промежуточной аттестации	4	0	4
Вид промежуточной аттестации (итогового контроля по дисциплине)	, Зачет		Зачет

4 Структура и содержание дисциплины

4.1 Сводные данные по содержанию дисциплины

Учебный год № 3

№ п/п	Наименование раздела и темы дисциплины	Виды контактной работы						СРС		Форма текущего контроля
		Лекции		ЛР		ПЗ(СЕМ)				
		№	Кол. Час.	№	Кол. Час.	№	Кол. Час.	№	Кол. Час.	
1	2	3	4	5	6	7	8	9	10	11
1	Установочная лекция	1	2							Доклад
2	Сбор и подготовка данных; Разведочный анализ							1	14	Доклад
3	Исследование зависимостей							1	20	Доклад
	Промежуточная аттестация									
	Всего		2						34	

Учебный год № 4

№ п/п	Наименование раздела и темы дисциплины	Виды контактной работы						СРС		Форма текущего контроля
		Лекции		ЛР		ПЗ(СЕМ)				
		№	Кол. Час.	№	Кол. Час.	№	Кол. Час.	№	Кол. Час.	
1	2	3	4	5	6	7	8	9	10	11
1	Задачи кластеризации и классификации	1	1	1	3			1	18	Отчет по лабораторной работе
2	Методы понижения размерности и факторный анализ	2	2	2	3			1	24	Отчет по лабораторной работе
3	Деревья решений для классификации	3	1	3	2			1	14	Отчет по лабораторной работе
	Промежуточная								4	Зачет

	аттестация								
	Всего		4		8			60	

4.2 Краткое содержание разделов и тем занятий

Учебный год № 3

№	Тема	Краткое содержание
1	Установочная лекция	Знакомство с требованиями. Основные разделы курса. Источники. Язык Python, Jupyter Notebook.
2	Сбор и подготовка данных; Разведочный анализ	Сбор данных из различных источников. Типы шкал. Описательная статистика. Выявление выбросов.
3	Исследование зависимостей	Корреляционный и регрессионный анализ. Частная, множественная корреляция, индекс корреляции. Нелинейная и множественная регрессия. Парадоксы регрессии.

Учебный год № 4

№	Тема	Краткое содержание
1	Задачи кластеризации и классификации	Обучение с учителем и без учителя на примере кластерного и дискриминантного анализа. Иерархический кластерный анализ и метод средних. Дискриминантный анализ, квадратичный дискриминантный анализ. Дискриминантный анализ для многоклассовой классификации. Классификация методом kNN.
2	Методы понижения размерности и факторный анализ	Главные компоненты, главные факторы. Нахождение факторных нагрузок. Вращение факторов.
3	Деревья решений для классификации	Деревья решений CART и CTree. Оценка качества классификации.

4.3 Перечень лабораторных работ

Учебный год № 4

№	Наименование лабораторной работы	Кол-во академических часов
1	Решение задачи кластеризации и классификации	3
2	Метод главных компонент и факторный анализ	3
3	Построение деревьев решений для решения задачи классификации	2

4.4 Перечень практических занятий

Практических занятий не предусмотрено

4.5 Самостоятельная работа

Учебный год № 3

№	Вид СРС	Кол-во академических часов
1	Проработка разделов теоретического материала	34

Учебный год № 4

№	Вид СРС	Кол-во академических часов
1	Подготовка к сдаче и защите отчетов	56

В ходе проведения занятий по дисциплине используются следующие интерактивные методы обучения: компьютерные симуляции

5 Перечень учебно-методического обеспечения дисциплины

5.1 Методические указания для обучающихся по освоению дисциплины

5.1.1 Методические указания для обучающихся по лабораторным работам:

1. Бучнев О. С. Методы анализа данных : электронный курс / О. С. Бучнев
<https://el.istu.edu/course/view.php?id=957>
2. Бучнев, Олег Сергеевич. Методы анализа данных: лаб. Практикум. – Иркутск: Изд-во ИрННТУ, 2022. - 114 с.

5.1.2 Методические указания для обучающихся по самостоятельной работе:

1. Проработка разделов теоретического материала по темам «Сбор и подготовка данных; Разведочный анализ» и «Исследование зависимостей»

Цель работы:

Самоконтроль и получение теоретических знаний и, по возможности, практических навыков.

Содержание задания на СРС

Для успешной защиты конспекта студенту необходимо защитить подготовленный отчет о самостоятельной проработке. В ходе защиты отчета студенту необходимо дать краткое изложение основных результатов полученных в ходе написания конспекта, устно ответить на теоретические вопросы по изложенной теме. Защита отчета о самостоятельной проработке теоретического материала является необходимым условием для допуска до итоговой аттестации.

Список тем контрольных вопросов

Бучнев О. С. Методы анализа данных : электронный курс / О. С. Бучнев
<https://el.istu.edu/course/view.php?id=957>

Порядок выполнения: изучить темы контрольных вопросов, быть готовым ответить на вопросы по этим темам.

2. Подготовка к сдаче и защите отчетов

Цель работы:

Самоконтроль и закрепление полученных в ходе выполнения лабораторных работ теоретических знаний и практических навыков.

Содержание задания на СРС

Для успешной сдачи лабораторной работы студенту необходимо защитить подготовленный отчет. В ходе защиты отчета студенту необходимо дать краткое изложение основных результатов полученных в ходе выполнения лабораторной работы, устно ответить на теоретические вопросы по теме лабораторной работы, а также продемонстрировать умение ориентироваться в написанном программном коде. Защита

отчета является необходимым условием для сдачи лабораторной работы.

Список тем контрольных вопросов

Бучнев О. С. Методы анализа данных : электронный курс / О. С. Бучнев

<https://el.istu.edu/course/view.php?id=957>

Порядок выполнения: изучить темы контрольных вопросов, быть готовым ответить на вопросы по этим темам; повторно просмотреть написанный программный код и вставить комментарии в тех местах, где может потребоваться пояснение.

6 Фонд оценочных средств для контроля текущей успеваемости и проведения промежуточной аттестации по дисциплине

6.1 Оценочные средства для проведения текущего контроля

6.1.1 учебный год 3 | Доклад

Описание процедуры.

В семестре 3 необходимо, используя доступные источники, рекомендуемую литературу, подготовить конспект, в котором содержатся ответы на следующие вопросы:

- 1) Сбор и подготовка данных. Источники данных.
- 2) Разведочный анализ данных, основные этапы.
- 3) Структура данных, запись в таблице. Объекты и признаки.
- 4) Группирующие переменные и переменные отклика
- 5) Типы шкал, в которых могут быть измерены признаки. Особенности каждой шкалы.
- 6) Актуальность проблемы пропущенных значений. Способы решения, реализованные в библиотеке Pandas.
- 7) Выбросы, диаграмма размаха, диаграмма Кливленда.
- 8) Описательная статистика, основные показатели.
- 9) Диаграмма рассеяния.
- 10) Понятие корреляции, виды корреляции.
- 11) Парная линейная корреляция, анализ частных и множественных связей.
- 12) Проверка гипотезы о значимости коэффициента корреляции.
- 13) Регрессия, виды регрессии, задачи регрессионного анализа.
- 14) Парная линейная регрессия.
- 15) Метод наименьших квадратов.
- 16) Множественная линейная регрессия, полиномиальная регрессия.
- 17) Проверка гипотезы о значимости коэффициентов регрессионного уравнения.
- 18) Коэффициент детерминации, диаграмма остатков.

В ходе написания конспекта студент должен применить полученные теоретические знания, и, по возможности, получить практические навыки использования процедур и функций языка Python. Подготовка конспекта необходима для осмысления полученных знаний и навыков. Подготовленный доклад высылается ответом на задание в системе дистанционного обучения в курсе Бучнев О. С. Методы анализа данных : электронный курс / О. С. Бучнев <https://el.istu.edu/course/view.php?id=957>. Преподавателем осуществляется проверка содержания на соответствие структуре и содержанию.

Критерии оценивания.

Полнота изложения, наличие компьютерных симуляций.

6.1.2 учебный год 4 | Отчет по лабораторной работе

Описание процедуры.

Выполнение каждой лабораторной работы состоит из двух частей: практической части в ходе которой необходимо выполнить индивидуальное задание и подготовка и защита отчета по лабораторной работе. Реализация решения индивидуального задания осуществляется средствами языка программирования Python. В ходе выполнения лабораторной работы студент должен применить для выполнения задания теоретические знания, полученные в ходе изучения курса, а также приобрести практические навыки использования процедур и функций языка Python. Подготовка отчета по лабораторной работе необходима для закрепления и переосмысления полученных знаний и навыков. Для успешной сдачи лабораторной работы необходимо продемонстрировать работу программы, реализующую индивидуальное задание, устранить выявленные преподавателем недостатки, подготовить и защитить отчет. Требования к оформлению программного кода, а также содержанию отчета указаны в соответствующем разделе методических указаний для каждой лабораторной работы.

Вопросы для контроля:

В семестре №4:

К лабораторной работе №1:

- 1) Назначение кластерного анализа.
- 2) Метрики расстояния.
- 3) Иерархический кластерный анализ.
- 4) Алгоритм k-means.
- 5) График средних.
- 6) Оценка качества кластеризации.
- 7) Назначение дискриминантного анализа.
- 8) Дискриминантные функции.
- 9) Линейный дискриминантный анализ.
- 10) Квадратичный дискриминантный анализ.
- 11) Оценка качества модели дискриминантного анализа.

К лабораторной работе №2:

- 1) Проблема размерности, преимущества снижения размерности.
- 2) Главные компоненты, математическая модель.
- 3) Алгоритм метода главных компонент.
- 4) Назначение факторного анализа.
- 5) Метод главных факторов.
- 6) Метод максимального правдоподобия.
- 7) Факторные нагрузки, общности.
- 8) Вращение факторного пространства.

К лабораторной работе №3:

- 1) Решение задачи классификации с использованием деревьев решений.
- 2) Алгоритм построения деревьев решений CART.
- 3) Алгоритм построения деревьев решений CTree.
- 4) Ограничения на применение алгоритма.
- 5) Визуализация и интерпретация построенных деревьев.
- 6) Преимущества и недостатки каждого из методов построения деревьев решений.

Критерии оценивания.

Соответствие заданию, сложность набора данных и качество первичной обработки. В процессе защиты отчета по лабораторной работе, студенту задаются контрольные вопросы теоретического и практического характера, соответствующие теме работы. Для успешной защиты отчета студенту необходимо дать краткое изложение основных результатов полученных в ходе выполнения лабораторной работы, устно ответить на

теоретические вопросы по теме лабораторной работы, а также продемонстрировать умение ориентироваться в написанном программном коде. Успешная защита отчета является необходимым условием для выставления оценки «зачтено» по соответствующей лабораторной работе.

6.2 Оценочные средства для проведения промежуточной аттестации

6.2.1 Критерии и средства (методы) оценивания индикаторов достижения компетенции в рамках промежуточной аттестации

Индикатор достижения компетенции	Критерии оценивания	Средства (методы) оценивания промежуточной аттестации
ПКС-2.3	Приводит примеры моделей, условий применимости моделей. Владеет инструментами и технологиями реализации методов и моделей для анализа данных. Работает со специализированными программными средствами. Способен осуществлять выбор методов и инструментов для анализа, осуществлять предобработку данных, интерпретировать полученные результаты.	Устное собеседование по теоретическим вопросам и защита лабораторных работ.

6.2.2 Типовые оценочные средства промежуточной аттестации

6.2.2.1 Учебный год 4, Типовые оценочные средства для проведения зачета по дисциплине

6.2.2.1.1 Описание процедуры

До зачета допускаются обучающиеся, защитившие все лабораторные работы. Зачет проводится в устной или письменной форме. Оформляются 12 билетов, в каждом билете 2 вопроса. Зачет выставляется по результатам собеседования по вопросам в билете.

Список контрольных вопросов:

1. Современные методы анализа данных;
2. Сбор, предварительная обработка, оценивание исходных данных.
3. Парная линейная корреляция.
4. Множественная, нелинейная корреляция.
5. Частная корреляция, ранговые корреляции.
6. Парная линейная регрессия.
7. Множественная регрессия. Полиномиальная и нелинейная регрессия.
8. Дисперсионный анализ. Двух и трехфакторный дисперсионный анализ.
9. Кластерный анализ. Цели и задачи.
10. Иерархическая кластеризация.
11. Кластерный анализ методом k-средних.
12. Дискриминантный анализ. Цели и задачи.
13. Критерии применения параметрических методов дискриминантного анализа.

14. Непараметрический дискриминантный анализ.
15. Метод главных компонент. Собственные числа и собственные векторы.
16. Снижение размерности.
17. Факторный анализ. Интерпретация факторов.
18. Деревья решений CART.
19. Деревья решений CTree.

Пример задания:

Типовой билет к зачёту:

Вопрос 1:

Множественная, нелинейная корреляция.

Вопрос 2:

Метод главных компонент. Собственные числа и собственные векторы.

6.2.2.1.2 Критерии оценивания

Зачтено	Не зачтено
Демонстрирует знание предмета. На вопросы отвечает, приводит примеры. Работает со специализированными программными средствами. Знает ограничения применяемых методов, статистических критериев и способов верификации моделей. Умеет интерпретировать полученные результаты.	Непонимание назначения методов анализа данных и математических моделей, которые лежат в их основе.

7 Основная учебная литература

1. Татарникова, Т. М. Интеллектуальный анализ данных : учебное пособие / Т. М. Татарникова. — Вологда : Инфра-Инженерия, 2024. — 172 с. — ISBN 978-5-9729-1772-3. — Текст : электронный // Лань : электронно-библиотечная система. — URL: <https://e.lanbook.com/book/427826> (дата обращения: 03.06.2025). — Режим доступа: для авториз. пользователей.

[Сайт] – URL: <https://e.lanbook.com/book/427826>

2. Золкин, А. Л. Математическое моделирование и анализ данных : учебное пособие для вузов / А. Л. Золкин, М. В. Сартаков. — Санкт-Петербург : Лань, 2025. — 128 с. — ISBN 978-5-507-51616-2. — Текст : электронный // Лань : электронно-библиотечная система. — URL: <https://e.lanbook.com/book/455660> (дата обращения: 03.06.2025). — Режим доступа: для авториз. пользователей.

[Сайт] – URL: <https://e.lanbook.com/book/455660>

3. Маккинни, У. Python и анализ данных. Первичная обработка данных с применением pandas, NumPy и Jupiter : справочник / У. Маккинни ; перевод с английского А. А. Слинкина. — 3-е изд. — Москва : ДМК Пресс, 2023. — 536 с. — ISBN 978-5-93700-174-0. — Текст : электронный // Лань : электронно-библиотечная система. — URL:

<https://e.lanbook.com/book/348086> (дата обращения: 03.06.2025). — Режим доступа: для авториз. пользователей.

[Сайт] – URL: <https://e.lanbook.com/book/348086>

8 Дополнительная учебная литература и справочная

1. Миркин Б. Г. Введение в анализ данных : учебник и практикум для вузов / Б. Г. Миркин, 2022. - 174.
2. Брантон С. Л. Анализ данных в науке и технике. Машинное обучение, динамические системы и управление / С. Л. Брантон, Дж. Н. Куц, 2021. - 542.
3. Криволапов С. Я. Введение в анализ данных. Поиск структуры данных с применением языка Python : учебное пособие / С. Я. Криволапов, 2024. - 177.

9 Ресурсы сети Интернет

1. <http://library.istu.edu/>
2. <https://e.lanbook.com/>
3. <https://colab.research.google.com/>

10 Профессиональные базы данных

1. <http://new.fips.ru/>
2. <http://www1.fips.ru/>

11 Перечень информационных технологий, лицензионных и свободно распространяемых специализированных программных средств, информационных справочных систем

1. Свободно распространяемое программное обеспечение LibreOffice

12 Материально-техническое обеспечение дисциплины

1. Проектор Epson EB-W04LCD.WXGA 1280*800.3000:1.2800 ANSI Lumens
2. 313493 Персональный компьютер PC/AT-286