

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»

УТВЕРЖДЕНА:

на заседании Совета института ИТиАД им. Е.И.Попова

Протокол №8 от 24 февраля 2025 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: очная

Документ подписан простой электронной подписью Составитель программы: Бучнев Олег Сергеевич Дата подписания: 04.06.2025

Документ подписан простой электронной подписью Утвердил: Говорков Алексей Сергеевич Дата подписания: 11.06.2025

Документ подписан простой электронной подписью Согласовал: Кононенко Роман Владимирович Дата подписания: 14.06.2025

Год набора – 2025

Иркутск, 2025 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-2.4

1.2 В результате освоения дисциплины у обучающихся должны быть сформированы

Код индикатора	Содержание индикатора	Результат обучения
ПКС-2.4	Способность 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», «Программирование», «Теория вероятностей, математическая статистика и случайные процессы», «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	Семестр № 5
Общая трудоемкость дисциплины	108	108
Аудиторные занятия, в том числе:	48	48
лекции	16	16
лабораторные работы	32	32

практические/семинарские занятия	0	0
Самостоятельная работа (в т.ч. курсовое проектирование)	60	60
Трудоемкость промежуточной аттестации	0	0
Вид промежуточной аттестации (итогового контроля по дисциплине)	Зачет	Зачет

4 Структура и содержание дисциплины

4.1 Сводные данные по содержанию дисциплины

Семестр № 5

№ 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	Сбор и подготовка данных; Разведочный анализ	1	2	1	8			1	12	Отчет по лаборатор ной работе
2	Исследование зависимостей	2, 3	4	2	8			1	16	Отчет по лаборатор ной работе
3	Задачи кластеризации и классификации	4, 5	4	3	6			1	10	Отчет по лаборатор ной работе
4	Методы понижения размерности и факторный анализ	6, 7	4	4	6			1	10	Отчет по лаборатор ной работе
5	Деревья решений для классификации	8	2	5	4			1	12	Отчет по лаборатор ной работе
	Промежуточная аттестация									Зачет
	Всего		16		32				60	

4.2 Краткое содержание разделов и тем занятий

Семестр № 5

№	Тема	Краткое содержание
1	Сбор и подготовка данных; Разведочный анализ	Сбор данных из различных источников. Типы шкал. Описательная статистика. Выявление выбросов
2	Исследование зависимостей	Корреляционный и регрессионный анализ. Частная, множественная корреляция, индекс корреляции. Нелинейная и множественная регрессия. Парадоксы регрессии.
3	Задачи кластеризации и классификации	Обучение с учителем и без учителя на примере кластерного и дискриминантного анализа.

		Иерархический кластерный анализ и метод k-средних. Дискриминантный анализ, квадратичный дискриминантный анализ. Дискриминантный анализ для многоклассовой классификации. Классификация методом kNN.
4	Методы понижения размерности и факторный анализ	Главные компоненты, главные факторы. Нахождение факторных нагрузок. Вращение факторов.
5	Деревья решений для классификации	Деревья решений CART и CTree. Критерии разбиения. Оценка качества классификации. Переобученные модели, способы борьбы с переобучением.

4.3 Перечень лабораторных работ

Семестр № 5

№	Наименование лабораторной работы	Кол-во академических часов
1	Сбор и подготовка данных для анализа	8
2	Корреляционный и регрессионный анализ	8
3	Решение задачи кластеризации и классификации	6
4	Метод главных компонент и факторный анализ	6
5	Построение деревьев решений для решения задачи классификации	4

4.4 Перечень практических занятий

Практических занятий не предусмотрено

4.5 Самостоятельная работа

Семестр № 5

№	Вид СРС	Кол-во академических часов
1	Оформление отчетов по лабораторным и практическим работам	60

В ходе проведения занятий по дисциплине используются следующие интерактивные методы обучения: публичная презентация, кейс-технология

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:

Бучнев О. С. Методы анализа данных : лабораторный практикум / О. С. Бучнев, 2022. - 114 с.

5.1.2 Методические указания для обучающихся по самостоятельной работе:

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 семестр 5 | 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.

Критерии оценивания.

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-2.4	Приводит примеры моделей, условий применимости моделей. Владеет инструментами и технологиями реализации методов и моделей для анализа данных. Работает со специализированными программными	Устное собеседование по теоретическим вопросам и защита лабораторных

	средствами. Способен осуществлять выбор методов и инструментов для анализа, осуществлять предобработку данных, интерпретировать полученные результаты.	работ.
--	--	--------

6.2.2 Типовые оценочные средства промежуточной аттестации

6.2.2.1 Семестр 5, Типовые оценочные средства для проведения зачета по дисциплине

6.2.2.1.1 Описание процедуры

Для того, чтобы получить зачет, необходимо защитить все лабораторные работы и пройти итоговое тестирование на странице курса "Методы анализа данных". Тест состоит из 10 вопросов, время тестирования ограничено 20 минутами. Порог успешного преодоления: 60% правильных ответов.

Пример задания:

Чем руководствуются при выборе метрики расстояния?

- a. количеством предполагаемых кластеров в данных;
- b. объемом выборки;
- c. способом нормировки;
- d. типами предполагаемых кластеров в данных;

6.2.2.1.2 Критерии оценивания

Зачтено	Не зачтено
Демонстрирует полученные знания. Умеет осуществлять постановку задачи, выбирать методы анализа. Работает со специализированными программными средствами и библиотеками. Знает ограничения применяемых методов. Владеет способами верификации моделей. Уверенно интерпретирует полученные результаты.	Непонимание назначения методов анализа данных и математических моделей, которые лежат в их основе. Не умеет осуществить постановку задачи.

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.
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.

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.

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