

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

«АНАЛИЗ ДАННЫХ»

Направление: 09.04.02 Информационные системы и технологии

Цифровизация промышленных предприятий

Квалификация: Магистр

Форма обучения: очная

Документ подписан простой
электронной подписью
Составитель программы:
Кононенко Роман
Владимирович
Дата подписания: 18.06.2026

Документ подписан простой
электронной подписью
Утвердил: Говорков Алексей
Сергеевич
Дата подписания: 18.06.2026

Документ подписан простой
электронной подписью
Согласовал: Кононенко Роман
Владимирович
Дата подписания: 18.06.2026

Год набора – 2026

Иркутск, 2026 г.

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

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

Код, наименование компетенции	Код индикатора компетенции
ПК-4 Способен разрабатывать планы и контролировать работы по управлению ИТ-проектами по созданию информационных систем	ПК-4.1

1.2 В результате освоения дисциплины у обучающихся должны быть сформированы

Код индикатора	Содержание индикатора	Результат обучения
ПК-4.1	Способен описывать текущие бизнес-процессы с применением инструментов моделирования и производить сбор запросов и потребностей от заказчика для формирования требований к протоколам передачи данных в информационных системах	Знать Основные нотации и инструменты моделирования бизнес-процессов: BPMN 2.0, IDEF0, EPC, DFD (диаграммы потоков данных), их элементы и правила построения. Методологию сбора требований к информационным системам: техники интервьюирования, анкетирования, наблюдения, мозгового штурма, анализа документации. Классификацию протоколов передачи данных (прикладные, транспортные, сетевые) и их характеристики: REST, SOAP, GraphQL, MQTT, AMQP, WebSocket – области применения, преимущества и ограничения. Принципы формирования технического задания на разработку протоколов обмена данными: структура требований к формату сообщений, частоте, объёму, надёжности и безопасности. Основы управления требованиями: отслеживание изменений, приоритизация, проверка достижимости. Уметь Проводить обследование предметной области и выявлять текущие бизнес-процессы, фиксируя

		их в виде графических моделей с использованием CASE-средств (например, Bizagi, ARIS, Camunda Modeler, Draw.io). Формулировать и документировать функциональные и нефункциональные требования к протоколам передачи данных на основе анализа потребностей заказчика и стейкхолдеров. Сопоставлять модели бизнес-процессов с потоками данных, определяя критичные точки интеграции, сущности и атрибуты, передаваемые между компонентами системы. Выбирать подходящий протокол или набор протоколов в зависимости от требований к производительности, задержке, надёжности и топологии сети. Разрабатывать спецификации интерфейсов (API-контракты, схемы данных) в форматах OpenAPI (Swagger), JSON Schema или XML Schema для согласования с разработчиками. Владеть Навыками построения контекстных диаграмм, диаграмм декомпозиции и потоков данных для визуализации бизнес-процессов и информационных обменов. Техниками проведения интервью и рабочих сессий с заказчиком для выявления скрытых потребностей и формализации бизнес-правил. Приёмами составления матриц соответствия между бизнес-функциями, информационными объектами и используемыми протоколами (например, матрица CRUD). Методами валидации требований: проверка полноты,
--	--	---

		непротиворечивости, однозначности и тестируемости требований к протоколам. Инструментами прототипирования интерфейсов взаимодействия (Postman, Insomnia) для демонстрации форматов сообщений заказчику.
--	--	--

2 Место дисциплины в структуре ООП

Изучение дисциплины «Анализ данных» базируется на результатах освоения следующих дисциплин/практик: «Академическое письмо»

Дисциплина является предшествующей для дисциплин/практик:

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 № 3
Общая трудоемкость дисциплины	108	108
Аудиторные занятия, в том числе:	28	28
лекции	14	14
лабораторные работы	0	0
практические/семинарские занятия	14	14
Самостоятельная работа (в т.ч. курсовое проектирование)	44	44
Трудоемкость промежуточной аттестации	36	36
Вид промежуточной аттестации (итогового контроля по дисциплине)	Экзамен	Экзамен

4 Структура и содержание дисциплины

4.1 Сводные данные по содержанию дисциплины

Семестр № 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	Введение в анализ данных.									Тест
2	Разведочный анализ данных (EDA) и предобработка	2	2			1	4			Тест

3	Статистические методы и проверка гипотез	3	2							Тест
4	Регрессионные модели	4	2			2	4			Тест
5	Классификация и кластеризация	5	2			3	6			Тест
6	Анализ временных рядов	6	2							Тест
7	Визуализация данных и коммуникация результатов	7	2					1	44	Тест
	Промежуточная аттестация								36	Экзамен
	Всего		12				14		80	

4.2 Краткое содержание разделов и тем занятий

Семестр № 3

№	Тема	Краткое содержание
1	Введение в анализ данных.	Жизненный цикл проекта по данным Рассматриваются этапы работы аналитика: постановка задачи, сбор данных, их очистка и предобработка, исследовательский анализ, построение моделей и интерпретация результатов. Обсуждаются роли специалистов, основные инструменты (Python, R, SQL) и типичные вызовы, возникающие на каждом этапе.
2	Разведочный анализ данных (EDA) и предобработка	Изучаются методы первичного анализа структуры данных: описательная статистика, визуализация распределений, выявление выбросов и пропусков. Осваиваются приёмы заполнения пропусков, нормализации, стандартизации, кодирования категориальных признаков и обнаружения аномалий.
3	Статистические методы и проверка гипотез	Посвящена основам математической статистики: точечные и интервальные оценки, параметрические и непараметрические критерии (t-критерий Стьюдента, критерии Манна–Уитни, Уилкоксона, хи-квадрат). Рассматривается корреляционный анализ (Пирсон, Спирмен) и интерпретация p-значений.
4	Регрессионные модели	Подробно разбираются линейная и множественная регрессия, логистическая регрессия для бинарных исходов. Акцент делается на оценке качества моделей (R^2 , скорректированный R^2 , AIC/BIC, ROC-кривая), проверке остатков, проблеме мультиколлинеарности и методах регуляризации (ридж, лассо).
5	Классификация и кластеризация	Изучаются задачи обучения с учителем (k-ближайших соседей, деревья решений, метод опорных векторов) и без учителя (k-средних,

		иерархическая кластеризация, DBSCAN). Обсуждаются метрики оценки качества классификации (точность, полнота, F-мера) и кластеризации (индекс силуэта, коэффициент Дэвиса–Болдина).
6	Анализ временных рядов	Вводятся понятия тренда, сезонности, циклической компоненты и шума. Изучаются методы сглаживания (скользящее среднее, экспоненциальное сглаживание), классические модели ARIMA/SARIMA и современные подходы (Prophet, LSTM) для прогнозирования и обнаружения аномалий.
7	Визуализация данных и коммуникация результатов	Рассматриваются принципы эффективной визуализации (выбор типа графика, цветовые схемы, иерархия информации) и инструменты построения интерактивных отчётов (Plotly, Dash, Tableau). Особое внимание уделяется сторителлингу – способности связывать числовые выводы с бизнес-контекстом и ясно доносить их до заказчика.

4.3 Перечень лабораторных работ

Лабораторных работ не предусмотрено

4.4 Перечень практических занятий

Семестр № 3

№	Темы практических (семинарских) занятий	Кол-во академических часов
1	Разведочный анализ и предобработка данных	4
2	Построение регрессионных моделей для прогнозирования	4
3	Кластеризация клиентской базы и интерпретация сегментов	6

4.5 Самостоятельная работа

Семестр № 3

№	Вид СРС	Кол-во академических часов
1	Подготовка к практическим занятиям	44

В ходе проведения занятий по дисциплине используются следующие интерактивные методы обучения: кейс-технологии

5 Перечень учебно-методического обеспечения дисциплины

5.1 Методические указания для обучающихся по освоению дисциплины

5.1.1 Методические указания для обучающихся по практическим занятиям

5.1.2 Методические указания для обучающихся по самостоятельной работе:

Самостоятельная работа студентов является необходимым условием формирования профессиональных навыков.

Подготовка к лабораторным работам является важной частью самостоятельной работы.

Приступая к подготовке к занятиям по конкретной теме, на начальном этапе самостоятельной работы студент должен подробно изучить основные вопросы темы, их последовательность, список рекомендуемой литературы.

Следующий этап самостоятельной работы – изучение темы занятия по учебникам и учебным пособиям. Наряду с основной литературой при подготовке к лабораторной работе целесообразно использовать законодательные и нормативные акты и дополнительные источники: специальную научную, научно-популярную, справочную, а также материалы, размещенные в глобальной сети Интернет, статья из периодических изданий.

Это определяющий этап самостоятельной работы, он очень сложен и важен, так как самостоятельные суждения по изучаемой проблеме формируются именно здесь, в том числе и в умении студента работать с научной литературой.

Завершающий этап подготовки к лабораторным работам – ответы на контрольные вопросы и выполнение заданий для самостоятельной работы, которые помогут правильно осмыслить изученный материал и проверить приобретенные знания.

Самостоятельная работа реализуется непосредственно в процессе аудиторных занятий на лабораторных занятиях, а также в контакте с преподавателем вне рамок расписания (на консультациях по учебным вопросам, в ходе творческих контактов, при ликвидации задолженностей, при выполнении индивидуальных заданий и т.д.), в библиотеке, дома.

Самостоятельная работа обучающихся предполагает следующие виды отчетности: подготовку и защиту отчета, выполнение домашних заданий, поиск и отбор информации по отдельным разделам курса в сети Интернет, выполнение творческих заданий.

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 3 | Тест

Описание процедуры.

Метод контроля, при котором обучающиеся устно отвечают на вопросы преподавателя, излагая изученный материал. Это может быть как беседа, так и рассказ, объяснение или даже чтение текста. Устный опрос позволяет преподавателю оценить не только знания обучающегося, но и его умение связно излагать мысли, развивать речь и память.

Критерии оценивания.

Решает нестандартные профессиональные задачи с применением знаний о цифровой трансформации.

Умеет создавать блок-схемы алгоритмов функционирования программного обеспечения.

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
ПК-4.1	Пороговый (удовлетворительно) Студент воспроизводит основные нотации моделирования и общие принципы сбора требований, может построить простую диаграмму процесса по готовому описанию, но испытывает трудности при самостоятельном анализе предметной области. Формулирует требования фрагментарно, без учёта протокольных ограничений. Базовый (хорошо) Самостоятельно проводит интервью с заказчиком (в рамках учебного кейса), строит модели процессов средней сложности, выделяет ключевые потоки данных. Обосновывает выбор протокола передачи, составляет формализованное описание требований в виде таблиц или спецификаций. Учитывает нефункциональные характеристики (время отклика, надёжность). Повышенный (отлично) Предлагает оригинальную архитектуру взаимодействия компонентов с оптимизацией выбора протоколов под конкретные бизнес-сценарии. Разрабатывает полную документацию требований, включая диаграммы последовательности, схемы данных и прототипы запросов. Владеет техниками управления требованиями и умеет защищать принятые решения перед заказчиком-оппонентом, аргументируя каждый выбор.	Выполнение и защита практических работ

6.2.2 Типовые оценочные средства промежуточной аттестации

6.2.2.1 Семестр 3, Типовые оценочные средства для проведения экзамена по дисциплине

6.2.2.1.1 Описание процедуры

Ответ на контрольные вопросы

Пример задания:

1. Опишите типовой жизненный цикл проекта по анализу данных. Какие этапы выделяют и какова их последовательность?
2. В чём различие между первичными и вторичными данными? Приведите примеры для каждого типа.
3. Какие основные приёмы предобработки данных вы знаете? Опишите способы обработки пропущенных значений и обнаружения аномалий (выбросов).
4. Что такое разведочный анализ данных (Exploratory Data Analysis, EDA)? Назовите основные инструменты (графические и числовые), используемые на этом этапе.
5. Объясните разницу между параметрическими и непараметрическими статистическими тестами. Приведите по одному примеру для каждой группы и укажите условия их применения.
6. Дайте определение корреляции. Чем отличается коэффициент корреляции Пирсона от коэффициента ранговой корреляции Спирмена? В каких случаях предпочтителен каждый из них?
7. Сформулируйте предположения (условия Гаусса–Маркова) классической линейной регрессионной модели. Почему их важно проверять?
8. Что такое мультиколлинеарность и как она влияет на оценки параметров регрессии? Какими способами можно диагностировать мультиколлинеарность?
9. В чём сущность регуляризации L1 (Лассо) и L2 (Ридж)? Сравните их свойства и укажите, в каких ситуациях каждый из методов предпочтительнее.
10. Перечислите основные метрики для оценки качества регрессионных моделей (R^2 , скорректированный R^2 , MAE, RMSE, MAPE). Дайте интерпретацию каждой метрики.
11. Для каких задач применяется логистическая регрессия? Опишите её функцию потерь (логистическую) и способ оценки параметров (метод максимального правдоподобия).
12. Как оценивается качество модели классификации? Поясните элементы матрицы ошибок (confusion matrix), а также метрики: точность (precision), полнота (recall), F_1 -мера и ассурасу.
13. Что такое ROC-кривая и показатель AUC-ROC? Как они помогают при выборе оптимального порога классификации в бинарных задачах?
14. Опишите алгоритм k-ближайших соседей (k-NN) для классификации. Как выбор значения k влияет на качество модели и склонность к переобучению?
15. Поясните принцип работы деревьев решений. Какие критерии используются для выбора наилучшего разбиения (энтропия, индекс Джини, ошибка классификации)? В чём их различия?
16. Что представляет собой задача кластеризации? Опишите алгоритм k-средних (k-

means), его преимущества и ключевые ограничения.

17. Назовите методы оценки качества кластеризации без использования меток классов. Опишите коэффициент силуэта (silhouette score) и индекс Дэвиса–Болдина – что они показывают?

18. Из каких компонентов состоит классический временной ряд? Определите тренд, сезонность, циклическую составляющую и случайный шум. Какие методы декомпозиции рядов вы знаете?

19. Для чего применяется модель ARIMA? Расшифруйте параметры p , d , q . В чём отличие сезонной ARIMA (SARIMA) от обычной?

20. Назовите основные принципы эффективной визуализации данных. Какие типы графиков наиболее часто используются для отображения распределений, зависимостей и динамики? Приведите примеры.

6.2.2.1.2 Критерии оценивания

Отлично	Хорошо	Удовлетворительно	Неудовлетворительно
Демонстрирует глубокое и системное понимание всех разделов дисциплины: от этапов жизненного цикла данных до современных методов машинного обучения и анализа временных рядов. Свободно ориентируется в статистических критериях, знает условия их применимости и умеет интерпретировать p-значения с учётом контекста задачи. Понимает	Хорошо знает основные понятия и методы анализа данных, но допускает незначительные неточности в деталях (например, в условиях применения конкретных критериев или в формулах). Понимает разницу между основными типами моделей и может назвать их ключевые особенности, но не всегда глубоко осознаёт математический аппарат.	Воспроизводит основные определения и базовые методы (EDA, линейная регрессия, простые критерии), но знания носят фрагментарный характер. Знает названия ключевых алгоритмов, но путается в их математических формулировках и условиях применимости. Может перечислить этапы анализа данных, но не понимает их логической взаимосвязи.	Не владеет базовыми понятиями курса: не может дать определение регрессии, кластеризации, не знает основных статистических критериев. Не понимает различий между типами данных и не может назвать этапы анализа.

математические основы алгоритмов (градиентный спуск, регуляризация, декомпозиция рядов) и может объяснить их сильные и слабые стороны.			
--	--	--	--

7 Основная учебная литература

1. Тюрин Ю. Н. Анализ данных на компьютере : учебное пособие по направлениям "Математика" , "Математика . Прикладная математика" / Ю. Н. Тюрин , А. А. Макаров, 2011. - 366,[1] с.

[Сайт] – URL: нет

8 Дополнительная учебная литература и справочная

1. Александров В. В. Анализ данных на ЭВМ : (На примере системы СИТО) / В. В. Александров, А. И. Алексеев, Н. Д. Горский, 1990. - 191 с.

[Сайт] – URL: нет

9 Ресурсы сети Интернет

1. <http://library.istu.edu/>
2. <https://e.lanbook.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 Microsoft Windows (Подписка DreamSpark Premium Electronic Software Delivery (3 years). Сублицензионный договор №14527/МОС2957 от 18.08.16г.)

12 Материально-техническое обеспечение дисциплины

1. Сервер специал. HP Multiseat 6000 Q9500 500G 6G 31PC Intel Core2 Quad /DVD+RW
2. экран Projecta

3. Проектор TOSHIBA TLP-X3000

4. Доска магнитно-маркерная INDEX настенная ,размер 1x1.8 м